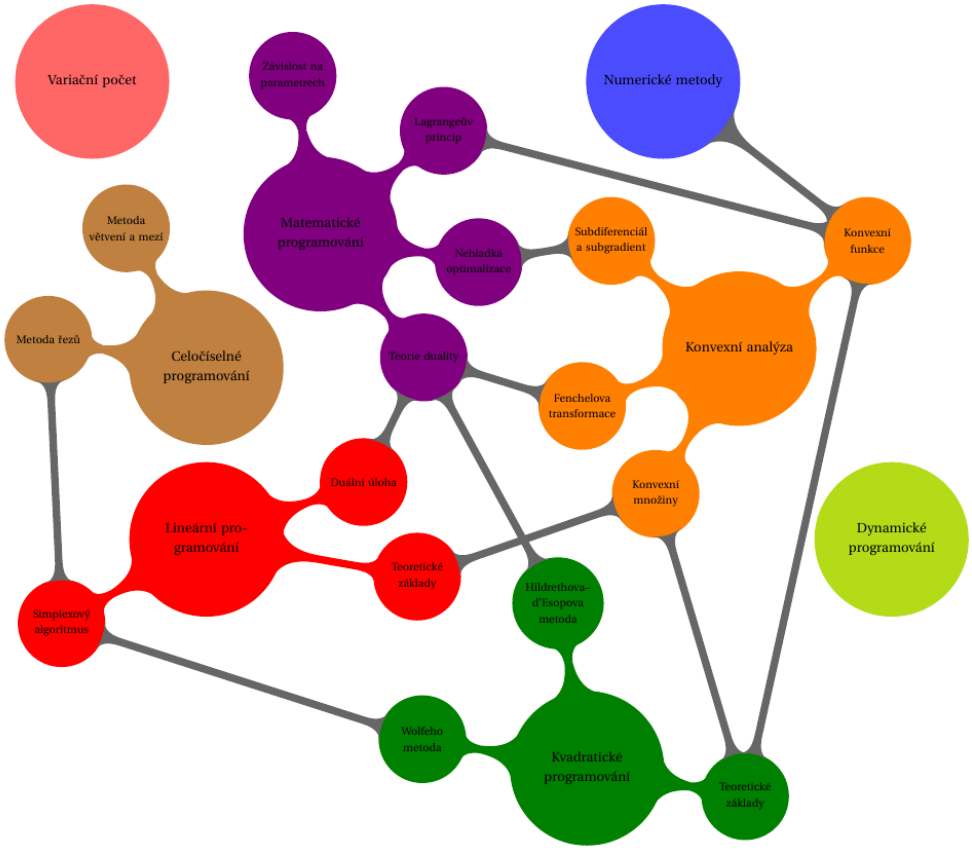
PETR ZEMÁNEK

MUNI SCI Ústav matematiky a statistiky

OPTIMALIZACE

ANEB KDYŽ MÉNĚ JE VÍCE

PRACOVNÍ VERZE: 4. června 2025



MUNI PŘÍRODOVĚDECKÁ FAKULTA

MASARYKOVA UNIVERZITA

Tato publikace vznikla v rámci projektu Fondu rozvoje Masarykovy univerzity (MUNI/FR/0901/2020) realizovaného v období 01/2020-12/2020.

Copyright © 2025 Petr Zemánek

PUBLIKOVÁNO MASARYKOVOU UNIVERZITOU

VYTVOŘENO S VYUŽITÍM BALÍČKU TUFTE LaTeX, VIZ HTTPS://GITHUB.COM/TUFTE-LATEX/TUFTE-LATEX

Toto je pracovní verze a *nesmí* být dále šířena. Určeno pro studenty Masarykovy univerzity v předmětu M5170 během podzimního semestru 2023.

První vydání, červen 2025

STELLINCE,
AUTOROVU JEDINÉMU GLOBÁLNÍMU ŘEŠENÍ

Obsah

Předmluva **9**

1 Úvod a motivace **15**

1.1 Pár (5202) slov na úvod, k zamyšlení a třeba i k diskuzi 16

1.2 O co nám zde vlastně půjde a co už známe? 20

1.3 První (ne)optimalizo(tel)ný historický exkurz 47

2 Lineární programování **61**

2.1 Trochu více historie (snad) nikoho nezabije 62

2.2 Seznamte se! Úlohy lineárního programování 80

2.3 Lineární programování všude kolem nás 91

2.4 Základem všeho je slušný náčrtek 105

2.5 Teoretické základy aneb být, či nebýt BPB 116

2.6 Dualita aneb peníze až na prvním místě 126

2.7 Svatý grál lineárního programování – simplexová metoda 139

2.8 Teď už si to jen správně zorganizovat 144

2.9 Aby začátky nebyly tak těžké 167

2.10 Simplexová metoda zprava dolů 175

2.11 Dopravní problém 187

2.12 Některé aplikace nyní podrobněji 208

2.13 Kleeho–Mintyho kostka je vržena 237

2.14 Lineární lomené programování 251

3 Celočíselné programování 255

3.1 Když reálná čísla nejsou příliš reálná 256

3.2 Celočíselnost až na prvním místě 258

3.3 Celočíselnost i bez celočíselnosti 274

3.4 Metoda větvení a mezí 280

3.5 Metoda řezů 287

3.6 Chcete být milionářem? 305

4 Kvadratické programování 319

4.1 Úvod a motivace 320

4.2 Když stačí tužka a pravítko... 328

4.3 Nepříliš zevrubný teoretický vhled 332

4.4 Hildrethova–d'Esopova metoda 335

4.5 Wolfeho metoda v krátkém tvaru 340

4.6 Wolfeho metoda v dlouhém tvaru 346

4.7 Než opustíme kvadratické programování... 359

5 Dynamické programování 367

5.1 Motivace 368

5.2 Konečně krokový deterministický proces 370

5.3 Stochastický proces 386

5.4 Nekonečně krokový deterministický proces 390

6 Základy konvexní analýzy 429

6.1 Konvexní množiny 430

6.2 Konvexní funkce 452

6.3 Oddělování konvexních množin a jeho důsledky 477

6.4 Vlastnosti konvexních funkcí 492

6.5 Subgradient a subdiferenciál 504

6.6 Fenchelova transformace 512

7 Numerické metody řešení optimalizačních úloh 533

7.1 Není konvergence jako konvergence... 535

7.2 Numerické metody nepodmíněné optimalizace v $\mathbb{R}$ 540

7.2.1 Metoda prostého dělení intervalu 542

7.2.2 Metoda půlení intervalu 544

7.2.3 Metoda zlatého řezu 548

7.2.4 Fibonacciho metoda 551

7.2.5 Metoda půlení intervalu II 560

7.2.6 Interpolační metody 562

7.3 Numerické metody nepodmíněné optimalizace v $\mathbb{R}^n$ 569

7.3.1 Metoda největšího spádu 572

7.3.2 Newtonova metoda 587

7.3.3 Metoda sdružených gradientů 593

8 Matematické programování 611

8.1 Obecná optimalizační úloha 619

8.2 Nutné a postačující podmínky optimality v matematickém programování 626

8.3 Teorie (Lagrangeovy) duality 651

8.4 Analýza citlivosti 674

9 Variační počet 699

9.1 Není to jen o nejrychlejší klouzačce 700

9.2 Variace bez kombinace 700

9.3 Základní úloha variačního počtu 705

9.4 Úloha o brachistochroně 710

Literatura 715

Zdroje převzatých obrázků 745

Seznam videozáznamů

1 Úvod a motivace

2 Lineární programování

Nedegenerovaná úloha v kanonickém tvaru . . 147

Nedegenerovaná úloha 147

Neohraničená úloha 148

Alternativní řešení #1 149

Alternativní řešení #2 149

Degenerovaná úloha 156

Dvoufázová metoda #1 170

Dvoufázová metoda #2 170

Dvoufázová metoda #3 170

Dvoufázová metoda #4 170

3 Celočíselné programování

Metoda větvení a mezí #1 282

Metoda větvení a mezí #2 282

Metoda větvení a mezí #3 287

Metoda řezů 296

4 Kvadratické programování

Grafické řešení #1 332

Grafické řešení #2 332

Grafické řešení #3 332

Hildrethova–D'Esopova metoda 338

Wolfeho metoda #1 345

Wolfeho metoda #2 346

5 Dynamické programování

Problém dostavníku 375

Problém batohu 375

Alokační problém 375

Spojitý rozhodovací proces 377

Stochastický rozhodovací proces 387

6 Základy konvexní analýzy

7 Numerické metody řešení optimalizačních úloh

8 Matematické programování

Primární úloha MP #1 640

Primární úloha MP #2 640

Primární úloha MP #3 640

Primární úloha MP #4 642

Duální úloha MP #1 668

Duální úloha MP #2 669

Parametrická úloha MP #1 687

Parametrická úloha MP #2 688

9 Variační počet

Předmluva

DĚLÁŠ TU NEJZBYTEČNĚJŠÍ PRÁCI NA SVĚTĚ.

AUTOROVA EXMANŽELKA
mimo jiné i v souvislosti se vznikem této knihy,
díky čemuž se autor nechal rozvést

Milé čtenářky, milí čtenáři,

k Vašim očím (a případně i do rukou) se dostává kniha, jejíž hlavní naplní je matematické hledání toho *nejlepšího* (ať už to znamená cokoli). Díky tomu si budeme moci lépe naplánovat naši dovolenou, investiční portfolio, vítěznou strategii do kasina či pomoci se zlepšením výrobního procesu v továrně. Také nám to pomůže porozumět mnoha přírodním jevům z našeho okolí, neboť příroda, ekosystém, vesmír (říkejme tomu jakkoli) je tím nejefektivnějším a nejúspornějším systémem[1]. Samozřejmě ne vždy se matematika dokáže popasovat se všemi detaily reálného života, takže si budeme muset některé situace zjednodušit (např. vyčíslit náš užitek z 10 párů ponožek zabalených na dvoudenní výlet či zanedbat odpor vzduchu při šestiminutové cestě tunelem mezi Prahou a Brnem), abychom se s pomocí námi budovaných nástrojů dopracovali k řešení. Z téhož důvodu zde však budeme raději sestavovat energeticky dostačující jídelníček s minimální cenou či rozvrhovat expresní rozvoz roušek do domovů seniorů místo toho, abychom se snažili optimalizovat náš spánek, mezilidské vztahy nebo náš politický postoj k Číně. Jako velmi názorná ukázka tohoto rozporu mezi matematikou a realitou může posloužit samotný autor, který navzdory svému nezměrnému úsilí o maximalizaci souladu mezi pracovním a rodinným životem zjistil, že dosáhl zcela opačného extrému. Ale Vy to jistě zvládnete mnohem lépe.

Pro koho je tato kniha určena? Pro všechny, kteří by chtěli objevit jednu z nejvyužívanějších a nejrychleji se rozvíjející oblastí matematiky. Např. v databázi MathSciNet[2] je toto obsahem kategorií *49*, *65K* a *90*, ve kterých najdeme od roku 2010 celkem 145347 záznamů, zatímco od roku 1901 to je 189180 záznamů. Také v databázi ArXiV najdeme třeba jen za leden 2020 více než 150 nových záznamů v kategorii *Optimization and Control*.

Jaké znalosti jsou potřebné k porozumění našemu výkladu? Kupecké počty, trocha zdravého rozumu, špetka maticové algebry a něco málo ze základního kurzu matematické analýzy. Ale ze všeho nejdůležitější je chuť objevovat a učit se nové věci.

Jak tuto knihu číst? Jakkoli triviální se tato otázka zdá, odpovědět na ni není úplně snadné. A stejně nesnadné vlastně bylo i zvolit vhodné uspořádání jednotlivých kapitol. Ono to totiž záleží na tom, zda nás zajímá především praktická stránka tohoto matematického aparátu nebo chceme raději celému výkladu porozumět do hloubky, aniž bychom příliš řešili „reálné“ úlohy. V prvním případě se stačí držet běžné konvence, která nám velí číst knihy od první do (snad) poslední stránky. Tím vlastně budeme částečně sledovat historický vývoj celé disciplíny a začneme od těch nejjednodušších úloh (lineární programování), přes ty složitější (celočí-

[1] Toto můžeme chápat i jako projev toho, co český geograf a popularizátor vědy, dr. Václav Cílek, označuje za *inteligenci přírody*.

[2] *MatSciNet* je interaktivní digitální mezinárodní matematická databáze vytvořená Americkou matematickou společností (AMS) v roce 1996, viz https://mathscinet.ams.org. Zahrnuje kompletní obsah časopisu *Mathematical Reviews* založeného v roce 1940. Obsahuje přes 3,6 miliónů recenzí, abstraktů a bibliografických údajů o matematické literatuře z celého světa a více než 2,3 milionů z těchto položek zahrnuje i odkaz na původní článek. Bibliografická data retrodigitalizovaných článků sahají až do roku 1864. Ročně přibývá v databázi více než 90 tisíc nových položek, které jsou klasifikovány podle kódů MSC (Mathematics Subject Classification). Spolu s evropskou databází *Zentralblatt MATH* (viz https://zbmath.org) jde o nejdůležitější referenční zdroj v oblasti matematiky.

selné a kvadratické programování) až se skrze konvexní analýzu dostaneme k těm nejobecnějším úlohám matematického programování, abychom se nakonec vrátili cca o 300 let zpět k variačnímu počtu. Pak ale nejméně v první polovině této cesty budeme narážet na různá tvrzení, jejichž důkazy nám v tu chvíli zůstanou skryty. To je pochopitelně z matematického pohledu zcela nepřijatelné a mělo by být tvrdě potrestáno, avšak my pro to máme jeden dobrý důvod: tato tvrzení jsou totiž *de facto* jen přímým důsledkem výsledků obecné teorie, takže bychom zde některé důkazy jen zbytečně duplikovali (i když občas s trochu jednoduššími argumenty). Pokud by nám ale ty důkazy výrazně chyběly, vydejme se více bourbakistickou[3] cestou a čtěme knihu od prostředku. Takto si nejdříve vybudujeme základy konvexní analýzy a na nich pak postavíme teorii matematického programování – zcela abstraktní a velmi obecnou se všemi podrobnostmi a důkazy. Poté se můžeme vrátit na začátek a věnovat se některým speciálním případům podrobněji i s různými alternativními možnostmi jejich řešení. Vydáme-li se touto druhou cestou, pak jen prosíme o shovívavost, neboť v takovém případě nám výklad v úvodních kapitolách může občas připadat zbytečně pomalý. Některé kapitoly lze také číst zcela či téměř zcela samostatně. Tyto singularity a vzájemné propojení některých témat jsme se pokusili schématicky znázornit na Obrázku 1 na konci předmluvy.

Kde se to všechno vzalo? Kniha vznikla na základě přednášek ke kurzům M5170 Matematického programování a M0160 (Teorie) Optimalizace na Přírodovědecké fakultě MU vedených autorem od roku 2015. První řádky tohoto textu spatřily světlo světa na jaře 2020 v průběhu pandemie způsobené koronavirem SARS-CoV-2. Základní inspirací pro tento text byla skripta prof. Došlého [143] z roku 2005 a jejich ruská „předloha“ [466], na jejichž obsah zde volně navazujeme a výrazně jej rozšiřujeme. Dalšími velmi užitečnými zdroji byly zejména [51, 57–59, 77, 79, 117, 250, 337, 360, 374, 427, 452, 467] a další články či monografie, které jsou zmíněny napříč celým textem. Některé pasáže v Kapitole 2 jsou převzaty (se svolením autorky; jde zejména o Podkapitoly 2.11–2.12) z diplomové práce Jany Zuzaňákové [528], která vznikla pod autorovým vedením. Podtitul této knihy je inspirován známým českým úslovím a také názvem skvělé knihy [391].

Kdo ovládá minulost, ovládá budoucnost(!?)[4] V souladu s tímto orwellovským výrokem nalezneme napříč textem řadu různých historických poznámek a faktografických údajů, které by měly napomoci k lepšímu pochopení některých souvislostí a detailů. Pro lepší autentičnost je obvykle uvádíme v původním jazyce, pokud se nám to podařilo najít. Pochopitelně nepředpokládáme u čtenářů znalost latiny, francouzštiny či němčiny, proto je doplňujeme i překladem do češtiny a většinou i do angličtiny (což vlastně je i náš primární zdroj), abychom minimalizovali riziko „ztrát v překladu“. Medailonky matematiků jsou převážně z `WikipediE` a z webové databáze [548].

Aby se ani Fermat[5] *nemohl vymlouvat.* Celý text je také doplněn řadou dalších poznámek doplňujících samotný výklad. Jde např. o důkazy některých vlastností, které jsou v textu zmíněny, ale nejsou formulovány ani jako věty. Pro pochopení našeho výkladu není nutné je číst, ale zvídavý čtenář v nich může najít mnoho zajímavých detailů. Zároveň s tím poskytujeme na levé části některých stran určitý prostor na poznámky, abychom alespoň částečně předešli problémům, do kterých nás kdysi dostal Fermat — jejichž řešení si vyžádalo více než 3 století a sto stran velmi komplikované matematiky. Jen kdyby na okraji bylo více místa…

Je to úplné? Samozřejmě, že ne. Ke každé z kapitol je možné dohledat nespočet velmi podrobných monografií věnovaných pouze a jenom jednomu tématu. Jenže výklad v takovém rozsahu by dalece překročil náš hlavní cíl[6], kterým je nabídnout čtenářům a čtenářkám prvotní vhled do různých oblastí optimalizace. Neodpustíme si sice zmínit některé nuance, ale v jednotlivých kapitolách nalezneme především základní výsledky, které však nepostrádají potřebný matematický aparát (to se týká zejména Kapitoly 6).

Věříte, že špetka praxe je cennější než tuna teorie? Proto je nedílnou součástí této učebnice více než 700 příkladů, které pocházejí (nejenom) ze cvičení k výše zmíněným předmětů a měly

[3] *Nicolas Bourbaki* byl pseudonym, pod kterým od roku 1935 publikovala svoje práce skupina (s oficiálním názvem *Association des collaborateurs de Nicolas Bourbaki*) tvořená převážně francouzskými matematiky, kteří byli absolventy École Normale Supérieure v Paříži (zakládajícími členy byli *Henri Cartan*, *André Weil*, *Claude Chevalley*, *Jean Delsarte* a *Jean Dieudonné*, ke kterým se později přidali např. *Jean Coulomb*, *Charles Ehresmann*, *René de Possel*, *Szolem Mandelbrojt*, *Laurent Schwartz*, *Jean-Pierre Serre*, *Alexander Grothendieck*, *Samuel Eilenberg*, *Serge Lang* či *Roger Godement*). Jejich původním záměrem bylo napsat novou učebnici matematické analýzy. V průběhu času však jejich cíl nabral ambiciózněji rozměrů a výsledkem byla série učebnic publikovaných pod Bourbakiho jménem s moderním výkladem čisté matematiky (*Éléments de mathématique*), čímž měla být matematika plně axiomatizována a postavena na pevný základ díky zavedení *matematických struktur*. Jednotlivé díly pojednávají o topologii (1940), algebře (1942), diferenciálním a integrálním počtu (1949 a 1952), teorii množin (1954) a jiných tématech. Skupina vznikla v reakci na První světovou válku, v níž zahynula řada francouzských matematiků, a tak nastupující generace byla nucena používat k výuce na univerzitách zastaralé texty. Díky jejich práci vznikla tzv. *Nová matematika*, což byl nový trend ve způsobu výuky matematiky v 60. letech 20. století s cílem posílit přírodovědné vzdělání a matematické dovednosti studentů tak, aby dokázali čelit technologické hrozbě sovětů, kteří byli skvělými matematiky. Šlo o reakci na tzv. sputnikovou krizi z roku 1957, kdy SSSR úspěšně vypustil první umělou družici Země, čímž potvrdil svoji schopnost odpalovat jaderné zbraně na mezikontinentální vzdálenost.

[4] *Who controls the past controls the future.* George Orwell: 1984 (str. 25)

[5] Viz str. 23 a další poznámky o Fermatovi v Kapitole 1.

[6] Asi bychom tím i čtenáře/čtenářky vystrašili a překročili jakoukoli únosnou míru z jejich strany. Také by ty bylo nad rámec autorových sil a časových možností.

by dostatečně posloužit k propojení teoretických poznatků s početními dovednostmi. Úplná řešení těchto příkladů z Kapitol 6–8 lze nalézt ve sbírce, za jejímž vznikem stojí především Markéta Zoubková a Jana Zuzaňáková. Navíc ke Kapitolám 7–8 jsou k dispozici naprogramované Maplety, které by měly poskytnout velmi podrobné řešení „libovolného“ příkladu z těchto oblastí odpovídající našemu současnému výkladu. Ke Kapitole 7 jsou dostupné také kódy v programu R, které vznikly v rámci diplomové práce Martina Habrovce [245] pod autorovým vedením a které nabízejí širší vhled do praktických výpočtů v představených metodách.

Psát, nebo mluvit? Součástí jednotlivých kapitol jsou i řešené ilustrativní příklady, z nichž většina je k dispozici ve formě videozáznamu (obvykle je uvedeno pouze zadání a odkaz na záznam). Jsme totiž přesvědčeni, že u takového druhu příkladů je forma „živého“ videozáznamu mnohem příjemnější, než pouhé čtení autorova řešení, v němž jsme navíc mnohdy nuceni se dopustit jistého zjednodušení.

Jste perfekcionisté? My také, ale současně jsme jen lidmi. Proto v duchu latinského rčení *Errare humanum est, in errore perseverare stultum*[7] už nyní, prosím, od nás přijměte omluvu za to, že se v textu jistě vyskytují drobné nejasnosti, překlepy, chyby… Vězte, že jsme je tentokrát neudělali záměrně, ale navzdory maximální snaze o dokonalost se nám je nepodařilo odstranit. Budeme velmi rádi za jakékoli upozornění na tyto nedostatky či náměty na vylepšení – využít můžete např. autorovu emailovou adresu zemanekp@math.muni.cz nebo se můžete zastavit za autorem osobně. Odměnou Vám bude autorův nekonečný vděk a věčné zapsání Vašeho jména do seznamu níže.

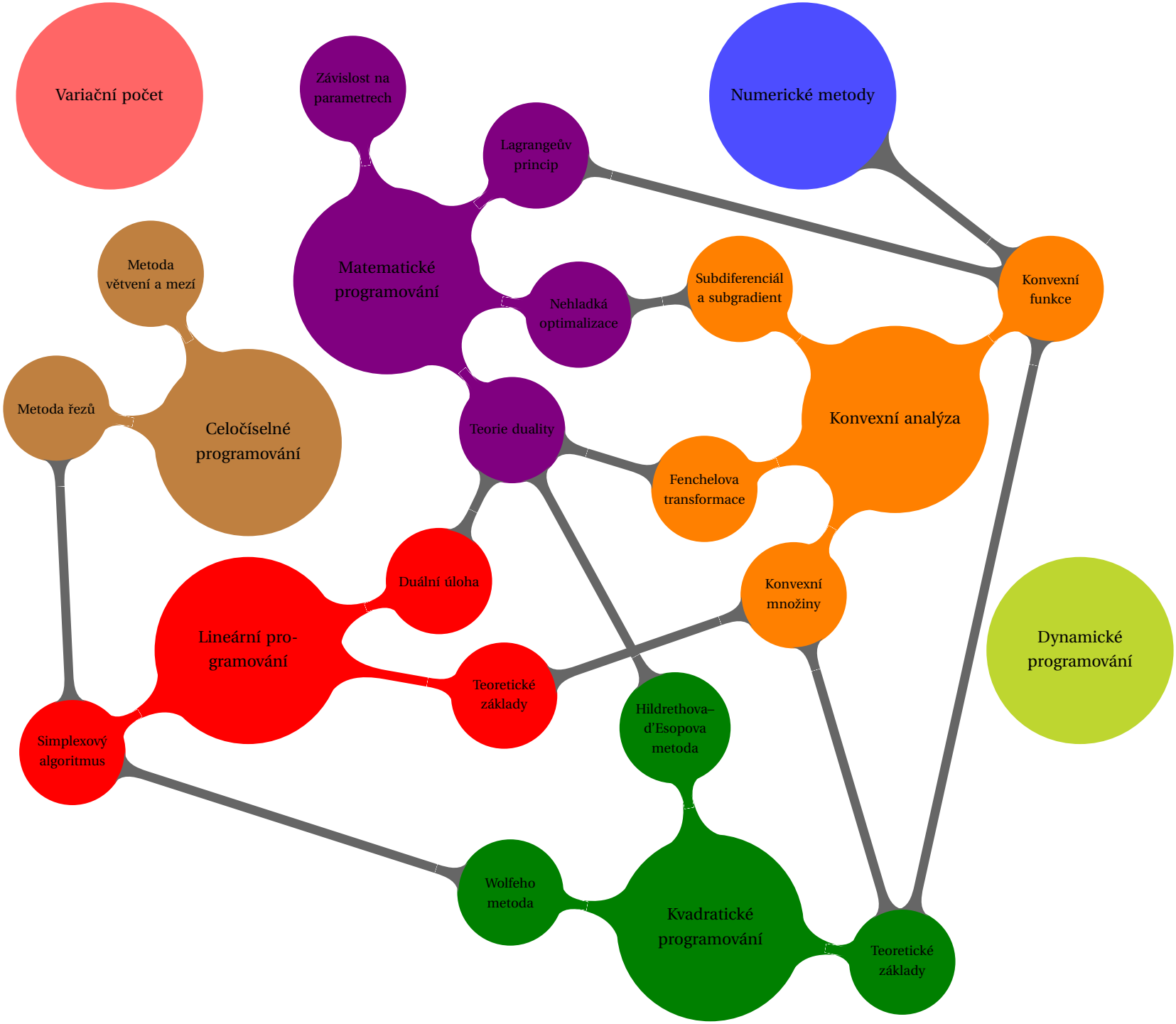
[7] *Chybovat je lidské, trvat na omylu hloupé* (či dábelské). Tento výrok je často připisován římskému filosofovi zejména Senecovi, ale objevují se v této souvislosti i jiná jména.

Vypadá to jako bod, hovoří se o tom jako o bodu, ale počítá se s tím ve skalárním součinu. Co to je? Je to bod a vektor současně. Jedné zdánlivé nepřesnosti jsme se totiž dopustili zcela úmyslně. V celém textu nerozlišujeme mezi body a vektory. Tedy lépe řečeno, někdy budeme s x pracovat jako s bodem (a psát jeho souřadnice v hranatých závorkách) a jindy o tomtéž x budeme hovořit jako o vektoru (psát jej jako sloupec v kulatých závorkách a počítat s ním např. ve skalárním součinu) či naopak. To vše činíme pouze z jediného důvodu: pro jednoduchost, abychom se vyhnuli např. zavádění duálního značení x a $\vec{x}$, o němž jsme přesvědčeni, že by celý výklad pouze zbytečně zkomplikovalo.

Kdo nám pomáhal? Samotný text byl připraven sázecím systémem TEX ve formátu LATEX 2ε s využitím modifikovaného stylového balíku TUFTE–BOOK. Některé obrázky byly generovány pomocí programu GEOGEBRA a jiné pomocí programu MAPLE. Obrázky 6.65–6.67 jsou interaktivní a umožňují rotování či přibližování, k čemuž je ale potřeba použít prohlížeč ADOBE ACROBAT READER. Za přečtení a výraznou korekturu první poloviny kapitol autor děkuje zejména Michaelu Novotnému. Vděk ale patří i velkému počtu dalších studentů předmětů M5170 a M0160 za mnoho jednotlivých připomínek a podnětných návrhů, které společně pomohly zvýšit (alespoň jazykovou) úroveň tohoto textu. Také bychom rádi alespoň touto cestou poděkovali Masarykově univerzitě za podporu prostřednictvím jejího fondu rozvoje, díky níž mohla vzniknout tato učebnice a především sbírka [527].

Toto je pracovní verze, i když Kapitoly 1–8 již lze považovat za kompletní.

Brno, červen 2025 Autor



Obrázek 1: Schématické propojení jednotlivých témat.

…Na počátku bylo *slovo*!

⋮

Pojem na počátku byl.

⋮

Na počátku byla síla!

⋮

Byl na počátku čin!

Johan Wolfgang Geothe: Faust. Prvý díl tragédie (str. 47)

Kapitola 1

Úvod a motivace

1.1 *Pár (5202) slov na úvod, k zamyšlení a třeba i k diskuzi* 16
1.2 *O co nám zde vlastně půjde a co už známe?* 20
1.3 *První (ne)optimalizo(tel)ný historický exkurz* 47

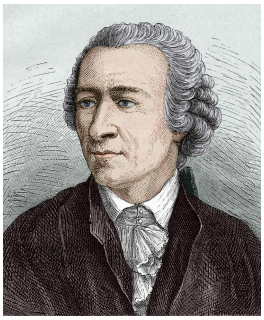
CUM ENIM MUNDI UNIVERSI FABRICA SIT PERFECTISSIMA ATQUE A CREATORE SAPIENTISSIMO ABSOLUTA, NIHIL OMNINO IN MUNDO CONTINGIT, IN QUO NON MAXIMI MINIMIVE RATIO QUAEPIAM ELUCEAT.[1]

LEONHARD EULER[2]

Viz Additamentum 1 (str. 245) ve spisu [159], jehož nádherná (a barevná) fotokopie je k dispozici na adrese https://ln.math.muni.cz/pz/opt/euler1.

[1] *Jelikož uspořádání vesmíru je na nejvýše dokonalé a ve skutečnosti dílem nejmoudřejšího Stvořitele, neexistuje ve vesmíru zhola nic, z čeho by nevyzařoval nějaký princip maxima nebo minima.*
(For since the fabric of the universe is most perfect and the work of a most wise Creator, nothing at all takes place in the universe in which some rule of maximum or minimum does not appear.)

[2]



Leonhard Euler (15. dubna 1707 – 18. září 1783) byl švýcarský matematik, fyzik, astronom a geograf. Byl matematickým pokračovatelem slavného rodu Bernoulliů, který zcela zásadním způsobem přispěl k rozvoji matematiky na přelomu 17. a 18. století. Euler výrazně ovlivnil mnohé (ne-li všechny tehdejší) oblasti matematiky (zejména infinitesimální počet a teorii grafů) a byl také průkopníkem několika nových směrů (zejména variačního počtu a analytické teorie čísel). Je autorem mnoha matematických pojmů a symbolů (obzvláště v matematické analýze). Euler počítal lehce, tak *jako člověk dýchá nebo orel létá* a jeho následovníci o něm např. řekli, že *Studium Eulerova díla zůstane nejlepší školou pro nejrůznější oblasti matematiky a nemůže je nic nahradit* (Carl Friedrich Gauss) a *Čtěte Eulera, čtěte Eulera, je to uči-*

1.1 Pár (5202) slov na úvod, k zamyšlení a třeba i k diskuzi

Matematika má nespočet podob, které však mají jedno společné – touhu po vědění čí poznání. Vždyť také základ slova *matematika* tvoří řecké *mathēma* (μαθημα) znamenající *vědění* či *poznání*[3] a *mathēmata* označovala to, co se lze naučit, a tudíž současně i to co může být vyučováno – jde vlastně o jakési složení slov *manthanein* znamenající poznání, vědění či naučení se a *mathēsis* znamenající učení. Heidegger[4] k tomu v jednom ze svých pojednání zaměřených na „moderní matematické přírodní vědy“ napsal následující slova.

> *Die μαθηματα das sind die Dinge, sofern wir sie in die Kenntnis nehmen, als das in die Kenntnis nehmen, als was wir sie eigentlich im voraus schon kennen, den Körper als das Körperhafte, an der Pflanze das Pflanzliche, am Tier das Tierische, am Ding die Dingheit usw. Dieses eigentliche Lernen ist somit ein höchst merkwürdiges Nehmen, ein Nehmen, wobei der Nehmende nur solches nimmt, was er im Grunde schon hat. Diesem Lernen entspricht auch das Lehren. Lehren ist ein Geben, Darbieten; aber dargeboten wird im Lehren nicht das Lernbare, sondern gegeben wird nur die Anweisung an den Schüler, sich selbst das zu nehmen, was er schon hat. Wenn der Schüler nur etwas Dargebotenes übernimmt, lernt er nicht. Er kommt erst zum Lernen, wenn er das, was er nimmt, als das erfährt, was er selbst eigentlich schon hat. Erst dort ist wahrhaftes Lernen, wo das Nehmen dessen, was man schon hat, ein Sichselbstgeben ist und als ein solches erfahren wird. Lehren heißt daher nichts Anderes, als die Anderen lernen lassen, d. h. sich gegenseitig zum Lernen bringen. Lernen ist schwerer als Lehren; denn nur wer wahrhaft lernen kann – und nur solange er es kann – der allein kann wahrhaft lehren. Der wahrhafte Lehrer unterscheidet sich vom Schüler nur dadurch, daß er besser lernen kann und eigentlicher lernen will. Bei allem Lehren lernt am meisten der Lehrer.*[5]

Matamatično je tedy to „na“ věcech, co vlastně už známe, co nezískáváme teprve z věcí, nýbrž jistým způsobem si již sami přinášíme s sebou. A ještě jedno zamyšlení od Heideggera, která by nám mohlo pomoci pochopit, proč kupř. číslo je něco matematického.

> *Wir sehen drei Stühle und sagen: Es sind drei. Was „drei“ ist, das sagen uns nicht die drei Stühle, auch nicht drei Äpfel oder drei Katzen oder sonst irgend drei Dinge. Vielmehr können wir Dinge nur als drei zählen, wenn wir schon die „drei“ kennen. Indem wir also die Dreizahl als solche fassen, nehmen wir nur etwas ausdrücklich zur Kenntnis, was wir irgendwie schon haben. Dieses Zur-Kenntnis-nehmen ist das eigentliche Lernen. Die Zahl ist etwas im eigentlichen Sinne Lernbares, ein mathēma, d. h. etwas Mathematisches. Um die Drei als solche, d. h. die Dreiheit, zu fassen, dazu helfen uns die Dinge nichts. Die Drei – was ist das eigentlich? Die Zahl, die in der natürlichen Zahlenreihe an dritter Stelle steht. An „dritter“! Die dritte Zahl ist es doch nur, weil es die Drei ist. Und „Stelle“ – woher Stellen? Die Drei ist nicht die dritte Zahl, sondern die erste Zahl, nicht etwa die Eins. Wir haben z. B. vor uns einen Laib Brot und ein Messer, dieses eine und dazu das andere. Wenn wir sie zusammen nehmen, sagen wir: diese beiden, das eine und das andere, aber nicht: diese zwei, nicht* $1+1$. *Erst wenn zu Brot und Messer z. B . ein Becher kommt und wir das Gegebene zusammennehmen, sagen wir: alle; jetzt nehmen wir sie als Summe, d. h. als ein Zusammen und soundso Viele. Erst vom Dritten her wird das vormalige Eine das erste und das vormalige Andere das zweite, wird eins und zwei, wird aus dem „und“ das „plus“, wird die Möglichkeit der Stellen und der Reihe. Was wir jetzt zur Kenntnis nehmen, schöpfen wir nicht*

tel nás všech (Pierre-Simon Laplace). Během svého života publikoval cca 500 knih a článků, což představuje za každý rok asi 800 stran textu. Po jeho smrti vyšlo ještě téměř 400 další prací. V roce 1907 (při příležitosti dvoustého výročí Eulerova narození) iniciovala Švýcarská akademie přírodních věd vydání Eulerova souhrnného díla. Tato kolekce momentálně čítá 72 svazků s desítkami tisíc stran, viz ln.math.muni.cz/pz/opt/euler6. Eulerovy práce jsou také dostupné v „jeho“ archivu ln.math.muni.cz/pz/opt/euler3. Velmi zajímavý a podrobný průřez Eulerovým dílem můžete nalézt i v krátkých komentářích *How Euler Did It* na adrese ln.math.muni.cz/pz/opt/euler4. Pro trochu kratší čtení doporučujeme článek dostupný na adrese ln.math.muni.cz/pz/opt/euler5. Emil Fellmann přesně ukázal, že Eulerova práce spočívá z 58 % v matematice a z 41 % v mechanice, astronomii, architektuře, dělostřelectvu, strojích a lodních pohonech. Na stránkách Eulerova archivu se také dočteme o odhadu amerického matematika a historika Clifforda Truesdella, v němž tvrdí, že *v* (potenciálním) *seznamu všech matematických, fyzických, mechanických, astronomických a navigačních prací vytvořených během 18. století by autorství 25 % z nich patřilo Leonhardu Eulerovi.*

[3] Angl. *knowledge, study, learning.*

[4] Martin Heidegger (26. září 1889– 26. května 1976) byl německý fenomenologický filosof, žák Edmunda Husserla a je považován za jednoho z nejvýznačnějších filosofů 20. století. Ovlivnil mnoho myslitelů, mimo jiné např. Jean-Paula Sartra. Jeho největší přínos tkví v tom, že obnovil otázku po bytí čili ontologii, kterou podle něho moderní filosofie zanedbala. Převzetí moci Adolfem Hitlerem roku 1933 chápal Heidegger jako příležitost k reformě německých univerzit. V dubnu 1933 byl zvolen rektorem univerzity ve Freiburgu a 1. května 1933 vstoupil do NSDAP. Během svého rektorství pomohl jedné ze svých židovských studentek do emigrace, zabránil pálení knih a vylepování protižidovských plakátů na freibugské univerzitě. Současně se ale snažil prosadit své reformy s pomocí nacistických aktivistů. Když se mu to nedařilo, rok po svém zvolení rezignoval. Zůstal však členem NSDAP a přestože byl v nemilosti, dále přednášel. Roku 1944

byl však jako „zcela postradatelný“ povolán na zákopové práce. Po roce 1945 měl v rámci denacifikace zákaz přednášet.

[5] *Ta mathēmata jsou věci, pokud je vezmeme na vědomí, které vezmeme na vědomí jako to, co již o nich dopředu víme, tělo jako tělesné, rostlinné v rostlině, zvířecí ve zvířeti, věcné ve věcech a tak dále. Toto skutečné učení je tedy nejpozoruhodnější braní, braní, ve kterém ten kdo, bere, bere jen to, co v podstatě již má. Výuka odpovídá tomuto učení. Učení je dávání, nabízení; ale to, co je nabízeno ve výuce, není to, co se lze naučit, protože student je pouze instruován, aby si vzal sám to, co již má. Pokud student pouze přebírá něco, co je nabízeno, neučí se. Učí se jen tehdy, když prožívá to, co bere, jako to, co už ve skutečnosti má. Skutečné učení je pouze tam, kde brát to, co již někdo má, je dávat sám sobě a jako takové to prožívat. Vyučování tedy neznamená nic jiného než nechat ostatní učit se, tj. navzájem se učit. Učení je obtížnější než vyučování: protože pouze ten, kdo se skutečně může učit – a jen tak dlouho, jak jen to dokáže – může skutečně učit. Skutečný učitel se liší od žáka pouze v tom, že se může lépe učit a že se skutečně chce učit. Při celé výuce se nejvíce učí učitel.*
(The mathēmata are the things insofar as we take cognizance of them as what we already know them to be in advance, the body as the bodily, the plant-like of the plant, the animal-like of the animal, the thingness of the thing, and so on. This genuine learning is therefore an extremely peculiar taking, a taking where he who takes only takes what he actually already has. Teaching corresponds to this learning. Teaching is a giving, an offering; but what is offered in teaching is not the learnable, for the student is merely instructed to take for himself what he already has. If the student only takes over something which is offered he does not learn. He comes to learn only when he experiences what he takes as something he himself already has. True learning only occurs where the taking of what one already has is a self-giving and is experienced as such. Teaching, therefore, does not mean anything else than to let the others learn, i.e., to bring one another to learning. Learning is more difficult than teaching: for only he who can truly learn – and only as long as he can do it – can truly teach. The genuine teacher differs from the pupil only in that he can learn better and that he more genuinely wants to learn. In

> *aus irgendwelchen Dingen. Wir nehmen, was wir irgendwie schon selbst haben. Es handelt sich um solches Lernbare, was als Mathematisches begriffen werden muß.*[6]

Matematikové neboli *mathēmatikoi* jsou ti, kteří *milují poznání*[7]. U pythagorejců se takto označovali *studenti* a jejich tehdejšími „konkurenty“ byli *posluchači* – *akusmatikové* neboli *akousmatikoi*, kteří se zaměřovali spíše na náboženské a rituální aspekty pythagorejského učení[8]. Při takovýchto úvahách autorovi vždy vytanou na mysli slova tří velkých matematiků 19. a 20. století, které vystihují podstatu matematiky. První z nich

> *Wahrlich es ist nicht das Wissen, sondern das Lernen, nicht das Besitzen sondern das Erwerben, nicht das Da-Seyn, sondern das Hinkommen, was den grössten Genuss gewährt.*[9]

je od Gausse[10]. Ten druhý

> *It can be of no practical use to know that π is irrational, but if we can know, it surely would be intolerable not to know.*[11]

je od Titchmarshe[12]. A konečně třetí citát

> *Wir müssen wissen. Wir werden wissen.*[13]

pochází od Hilberta[14] a zdobí jeho náhrobek v Göttingenu. Tato slova pronesl 8. září 1930 v závěru svého projevu ke Společnosti německých vědců a lékařů při odchodu do důchodu, který lze slyšet i na záznamu https://ln.math.muni.cz/pz/opt/hilbert1. Obšírnější verzi Hilbertova epitafu nalezneme již v jeho slavné přednášce [264] z roku 1900.

> *Diese Überzeugung von der Lösbarkeit eines jeden mathematischen Problems ist uns ein kräftiger Ansporn während der Arbeit; wir hören in uns den steten Zuruf: Da ist das Problem, suche die Lösung. Du kannst sie durch reines Denken finden; denn in der Mathematik gibt es kein Ignorabimus.*[15]

Toto byla reakce na tezi zformulovanou du Bois-Reymondem[16] a zpopularizovanou především v jeho knize *Über die Grenzen des Naturerkennens* (O mezích poznání přírody) vydané v roce 1872

> *Ignoramus et ignorabimus.*[17]

kterou Hilbert ve svém projevu z roku 1930 označil za pošetilou či bláhovou. Tato domněnka vyjadřuje postoj, že *člověk má ohraničené a nepřekročitelné hranice možností poznání přírody*, což bývá považováno za projev extrémního *agnosticismu*, tj. názoru, že pravdivost některých tvrzení, zejména těch, která se týkají existence či neexistence jakéhokoliv boha, se nedá prokázat ani vyvrátit a že totéž platí i pro další, zejména náboženská a metafyzická tvrzení. Jak už tomu bývá v takovýchto ostrých názorových střetech, pravda bude nejspíše někde uprostřed, neboť v roce 1931 publikoval Gödel[18] dvojici *Vět o neúplnosti*. První z nich totiž ukazuje, že *v žádné rozumné teorii hovořící o přirozených číslech není dokazatelné vše*, zatímco ta druhá udává konkrétní příklad takového nedokazatelného tvrzení.

Touha po poznání je tedy hnacím motorem matematiky. Ten jsme spustili ve chvíli, kdy jsme začali počítat své prsty na rukou, a od té doby už běží jako (téměř) dokonalé perpetuum mobile, neboť každé nové poznání obvykle přináší další a další otázky. Navíc existují také různé

all teaching, the teacher learns the most.) Viz [256, str. 73–74] a [257, str. 73].

[6] *Vidíme tři židle a říkáme: jsou tři. Co je „tři“, to nám ty tři židle neříkají, dokonce ani tři jablka nebo tři kočky ani žádné další tři věci. Víceméně můžeme věci počítat jako tři pouze tehdy, když už známe „tři“. Uchopením čísla tři jako takového tedy pouze explicitně poznáváme něco, co už nějakým způsobem máme. Toto braní-na-vědomí je skutečné učení. Číslo je něco, co se dá skutečně naučit neboli mathēma, tzn. něco matematického. K těm třem jako takovým, tj. trojitosti, k jejich pochopení nám žádné věci nepomohou. Ty tři – co to vlastně je? Číslo, které v řadě přirozených čísel je na třetím místě. Na „třetím“! Je to teprve třetí číslo, protože je to tři. A „místo“ – odkud pocházejí místa? Trojka není třetí číslo, ale první číslo, ne nějaká jednička. Máme např. před námi bochník chleba a nůž, jedna věc a navíc ještě jiná věc. Když je vezmeme dohromady, řekneme: obě dvě, jedna a druhá, ale ne: tyto dvě, ne* $1+1$. *Pouze když ke chlebu a noži např. přidáme hrnek a posbíráme, co máme, řekneme: vše; nyní je bereme jako součet, tj. jako celek a tak mnoho. Jenom když to vnímáme z pohledu té třetí se ten předcházející stane prvním a další předcházející druhým druhým, takže se objeví jednička a dvojka, „a“ se stává „plusem“, objeví se možná místa a řazení. To, čeho si nyní všímáme, nevychází z žádné věci. Bereme to, co už tak nějak máme. To, co musí být chápáno jako matematické, je to, co se můžeme naučit tímto způsobem.*
(We see three chairs and say that there are three. What „three“ is the three chairs do not tell us, nor three apples, three cats nor any other three things. Moreover, we can count three things only if we already know „three“. In thus grasping the number three as such, we only expressly recognize something which, in some way, we already have. This recognition is genuine learning. The number is something in the proper sense learnable, a mathēma, i.e., something mathematical. Things do not help us to grasp „three“ as such, i.e., threeness. „Three“ – what exactly is it? It is the number in the natural series of numbers that stands in third place. In „third“? It is only the third number because it is the three. And „place“ – where do places come from? „Three“ is not the third number, but the first number. „One“ isn't really the first number. For instance, we have before us one loaf of bread and one knife, this one and, in addition,

faktory, které mohou výrazně zvýšit výkon tohoto motoru. Např. jako ve sportu, kdy samozřejmě chcete být lepší, rychlejší a překonávat výkony svých soupeřů. V historii matematiky nalezneme celou řadu problémů, jejichž řešení se podařilo najít jen díky tomu, že se o nich někdo významný zmínil (jako v případě Hilbertova seznamu) nebo dokonce i oznámil znalost takového řešení, ale nechal si jej pro sebe (pokud jej tedy vůbec měl) – třeba když Targalia[19] nalezl vzorce pro řešení kubických rovnic (dnes známé spíše jako *Cardanovy*[20] vzorce) jen kvůli matematickému souboji s Fiorem[21] v roce 1535, protože si o něm myslel, že je zná (avšak ve skutečnosti neznal, ovládal pouze řešení jednoho typu tzv. redukovaných kubických rovnic). Druhý příklad stojící za zmínku je slavná *Velká Fermatova věta*[22], kterou si Fermat[23] v roce 1637 napsal na okraj slavné knihy *Aritmetika* od Diofanta z Alexandrie[24] a zároveň dodal

> *Cubum autem in duos cubos, aut quadrato-quadratum in duos quadrato-quadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exigitas non caperet.*[25]

Dnes už se asi nedozvíme, zda Fermat skutečně něco tak geniálního objevil nebo se mýlil. Jediný doposud známý důkaz této věty čítá více než 100 stran velmi složité matematiky z různých oblastí a pochází od Wilese[26], kterému se jej podařilo dokončit v roce 1994.

Druhým možným „urychlovačem“ může být *krása* matematického poznání, jak se o tom zmiňuje např. již dříve jednou citovaný Titchmarsh

> *Perhaps the most surprising thing about mathematics is that it is so surprising...*[27]

nebo na půl v žertu Weyl[28]

> *My work has always tried to unite the true with the beautiful and when I had to choose one or the other, I usually chose the beautiful.*[29]

nebo zcela vážně a poeticky Russell[30]

> *Mathematics, rightly viewed, possesses not only truth, but supreme beauty – a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show. The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as poetry.*[31]

K tomu však potřebujeme mít v sobě alespoň trochu matematického ducha, neboť asi ne úplně každý plně docení výsledky založené např. na překvapivém propojení dvou původně zcela nesouvisejících oblastí matematiky ilustrující dokonalost matematického soukolí, o kterém se zmiňuje Whitehead[32].

> *The science of pure mathematics, in its modern developments, may claim to be the most original creation of the human spirit.*[33]

Nebo můžeme obdivovat různé „singularity“ typu

$$10^2+11^2+12^2=13^2+14^2 \qquad \operatorname{tg}10^\circ=\operatorname{tg}20^\circ\times\operatorname{tg}30^\circ\times\operatorname{tg}40^\circ$$

another one. When we take both together we say, „both of these", the one and the other, but we do not say, „these two", or $1+1$. Only when we add a cup to the bread and the knife do we say „all". Now we take them as a sum, i.e., as a whole and so and so many. Only when we perceive it from the third is the former one the first, the former other the second, so that one and two arise, and „and" becomes „plus", and there arises the possibility of places and of a series. What we now take cognizance of is not created from any of the things. We take what we ourselves some how already have. What must be understood as mathematical is what we can learn in this way.) Viz [256, str. 74–75] a [257, str. 74–75].

[7] Angl. *fond of learning.*

[8] V tomto kontextu je pak leckdy sousloví „učit se matematiku" spíše oxymóronem než čímkoli jiným.

[9] *Nejsou to znalosti, ale akt učení, ne vlastnictví pravdy, ale cesta k ní, co přináší člověku největší potěšení.* (It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment.) Viz [211, str. 94].

[10]

Johann Carl Friedrich Gauss (30. dubna 1777 – 23. února 1855) byl jeden z největších (nejen) německých matematiků všech dob. Výrazně ovlivnil zejména rozvoj geometrie, matematické analýzy, teorie čísel, astronomie, elektrostatiky, geodézie a optiky. Nicméně v oblasti našeho současného zájmu tak významnou roli nesehrál. Koluje také řada různě důvěryhodných historek o jeho genialitě již v brzkém věku, viz např. [148].

[11] *Vědět, že číslo π je iracionální, může postrádat jakékoli praktické využití, ovšem pokud to můžeme vědět, bylo by jistě nepřijatelné to nevědět.* Viz [435, str. 113].

$$2=\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}}} \qquad e^{i\pi}+1=0 \qquad e=2+\cfrac{2}{3+\cfrac{3}{4+\cfrac{4}{5+\cfrac{5}{6+\ddots}}}}$$

$$\frac{\pi^2}{6}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\cdots \qquad \int_0^1 \frac{\mathrm{d}x}{x^x}=\sum_{k=1}^{\infty}\frac{1}{k^k}.$$

Obvyklá reakce ale bývá *No a dál?* (což bývá ještě ta přijatelnější verze). Zatímco projevem obdivu k dílu Shakespeara či Mozarta s doplněním nějakého střípku z jejich života mnohdy vzbudíme dojem vzdělaného a kulturního člověka, zmínkou o kráse matematiky podpořené třeba o některou z identit výše vyvoláme nejspíše dojem úplně jiný. Dokonce jsou situace, ve kterých bychom důrazně nedoporučovali se o něčem takovém zmiňovat. Na párty to totiž z nás asi hvězdu večera neudělá, a to ani když si k tomu připravíme výběr těch nejlepších matematických vtipů nebo (nedej bože) přidáte detailní rozbor „situace s elektronickou tužkou". Z autorovy vlastní zkušenosti můžeme i potvrdit, že se s tím nedá ani příliš uspět na první, druhé či třetí (tohle bylo maximum, další pokus už autor raději nezkoušel) schůzce s partnerem/partnerkou, pokud tedy zrovna nepatříte do téhož „$\varepsilon-\delta$ klubu".

„Odhalovat krásu matematiky" je také velmi oblíbené klišé v dnešních debatách o výuce matematiky. Jenže zkusme si do těch vět místo matematiky dosadit jiný středoškolský předmět. Český jazyk, biologii, chemii, fyziku… Copak se něco z toho učí kvůli tomu, abychom tím byli okouzleni? Ani náhodou. Tím třetím a zároveň nejdůležitějším palivem našeho matematického motoru je totiž *užitečnost*. Klasická „školská"[34] matematika by nás především měla naučit logickému, analytickému a abstraktnímu myšlení, což jsou dovednosti mnohdy zcela nezbytné pro běžný život. Zde se nám jistě vybaví všelijaké kvízy, hádanky či dokonce olympiády s obvykle velmi nepraktickými úlohami. Mozek je jako sval a potřeba jej procvičovat, jinak „zakrní" (a moderní technologie nám to příliš neulehčují). Bez ohledu na to, co slýcháme od různých „osobností" ze všech stran, novodobý člověk je především *Homo Mathematicus*. Kde bychom dnes asi byli bez matematiky? Zkuste si odpovědět sami…Velmi výstižně to napsal americký spisovatel Robert Heinlein:

> *If it can't be expressed in figures, it is not a science; it is opinion.*[35]

Matematika je dnes jazykem snad už skoro všech věd a zároveň je úžasným nástrojem pro modelování reálného světa, čehož si byl velmi dobře vědom už i slovutný Galileo Galilei[36]:

> *La filosofia [della natura] è scritta in questo grandissimo libro che continuamente ci sta aperto dinanzi a gli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscere i caratteri ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto.*[37]

Oblast, do které se společně v následujících kapitolách vydáme, patří v tomto ohledu k těm nejvyužívanějším – optimalizace. Jen si vzpomeňme na slova Leonharda Eulera z úvodu této kapitoly, která ideálně doplňují předchozí Galileův výrok. A co vy, čtenáři? Kdy jste naposledy zvažovali různé varianty a hledali *nejlepší rozhodnutí, ideální kompromis* či *největší užitek* (ať už se týkalo čehokoli)? Samozřejmě matematika není všemocná, takže asi nedokážeme vyřešit každý rozhodovací problém, na který v životě můžeme narazit – jako obvykle totiž platí, že čím realističtější je problém, tím složitější je jeho matematizace a samotné řešení. Proto vám zde asi nepomůžeme optimalizovat vaše studium, spánek, partnerský vztah či milostný život (ačkoli ten naštěstí trápí matematiky/matematičky ze všeho nejméně). Nicméně uvidíme, že s celou řadou mnohem praktičtějších úloh si už celkem snadno poradíme (a díky některým

[12]

Edward Charles Titchmarsh (1. června 1899 – 18. ledna 1963) byl významným anglickým matematikem 20. století. Je znám především díky jeho práci v analytické teorii čísel a Fourierově analýze.

[13] *Musíme vědět. Budeme vědět.* (We must know. We will know.)

[14]

David Hilbert (23. ledna 1862 – 14. února 1943) byl německý matematik, který velmi výrazně ovlivnil matematiku na přelomu 19. a 20. století. Jeho jméno je mimo jiné spojeno s tzv. *Hilbertovými problémy*. Ty spatřily světlo světa 8. srpna 1900, kdy Hilbert vystoupil během Druhého mezinárodního kongresu matematiků pořádaného v Paříži s přednáškou obsahující deset nevyřešených problémů. Tyto problémy spolu s dalšími třinácti pak publikoval v [264], díky čemuž vznikl seznam přitahující pozornost mnoha matematiků. Devět problémů je již zcela vyřešeno, sedm z nich má „částečné" řešení, formulace dvou problému je příliš vágní, aby se o jejich řešení dalo rozhodnout, a zbývají 4 problémy, které zůstávají nevyřešené dodnes (včetně Riemannovy hypotézy). Na počátku 21. století se vědecká rada Clayova matematického ústavu (ve složení Alain Connes, Arthur Jaffe, Edward Witten a Andrew Wiles) vyhlásila tzv. *Problémy tisíciletí* (*Millennium Prize Problems*) se sedmi nejdůležitějšími otevřenými problémy soudobé matematiky (Riemannova hypotéza, Problém P versus NP, Birchova a Swinnertonova-Dyerova domněnka, Hodgeova domněnka, Yangova–Millsova teorie a hypotéza hmotnostních rozdílů, Navierova–Stokesova rovnice

z nich lze dokonce i získat Nobelovu cenu), a sem tam odkryjeme i nějaké tajemství přírody. A pokud navíc v předloženém výkladu občas spatříte i něco krásného, bude naše putování zcela *optimální*, neboť jak řekl Godfrey Hardy[38]:

> *The mathematician's patterns, like the painter's or the poet's, must be beautiful; the ideas, like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics.*[39]

Ovšem dopředu upozorňujeme, že s jednou (byť velmi výraznou) výjimkou se zde rozhodně nebudeme zabývat nějakou hlubokou matematickou historií. Přestože se nám to může z dnešního pohledu zdát snad i nepochopitelné, vždyť všelijaké optimalizační úlohy a rozhodovací procesy jsou jistě staré jako lidstvo samo, jejich systematické studium začalo až ve druhé polovině 20. století a výzkum v těchto oblastech je (a jistě i nadále bude) velmi aktivní.

1.2 O co nám zde vlastně půjde a co už známe?

Optimalizační úlohou můžeme v tom nejobecnějším pojetí rozumět problém nalezení takového prvku dané *přípustné množiny* $\mathbb{X}$, že odpovídající *„účelové zobrazení"*[40] $\mathbb{F} : \mathbb{X} \to \mathbb{R}$ nabývá v jistém smyslu *optima* (typicky minima či maxima), tj. úlohu

$$\mathbb{F}(x) \to \operatorname{opt}, \quad x \in \mathbb{X}. \tag{1.2.1}$$

Řešením této úlohy pak rozumíme prvek $x^* \in \mathbb{X}$ splňující[41] podmínku $\mathbb{F}(x) \bowtie \mathbb{F}(y)$ pro všechna $y \in \mathbb{X} \setminus \{x^*\}$. Takový prvek ale rozhodně nemusí být jediný, takže řešením dané úlohy je spíše množina bodů

$$\mathbb{X}^* = \arg\operatorname{opt}_{x \in \mathbb{X}} \mathbb{F}(x) := \{x \in \mathbb{X} \mid \mathbb{F}(x) \bowtie \mathbb{F}(y) \text{ pro všechna } y \in \mathbb{X}\}.$$

V některých případech však nemusí být nutné (nebo dokonce možné) nějaký optimální bod $x^* \in \mathbb{X}^*$ najít a místo toho se můžeme spokojit „pouze" s určením *optimální hodnoty účelového zobrazení*, tj. čísla

$$\mathbb{F}^* := \operatorname{opt}\{\mathbb{F}(x) \mid x \in \mathbb{X}\},$$

kde symbol opt znamená buď hledání minima (v obecnějším případě infima) nebo maxima (případně suprema). Vlastně ani není příliš podstatné, zda je naším úkolem minimalizovat nebo maximalizovat „zobrazení" $\mathbb{F}$, neboť úlohy

$$\mathbb{F}(x) \to \min, \ x \in \mathbb{X} \quad \text{a} \quad -\mathbb{F}(x) \to \max, \ x \in \mathbb{X}$$

jsou *ekvivalentní*, tj. odpovídající množiny řešení $\mathbb{X}^*$ jsou totožné (obě úlohy buď mají tatáž řešení nebo žádná z nich nemá řešení) a v případě $\mathbb{X}^* \neq \emptyset$ zjevně platí

$$\mathbb{F}^* = \min\{\mathbb{F}(x) \mid x \in \mathbb{X}\} = -\max\{-\mathbb{F}(x) \mid x \in \mathbb{X}\}.$$

My se v našich úvahách omezíme především na minimalizaci, což je typické pro úlohy s fyzikálním podtextem (minimalizace času, práce, „akce"), zatímco maximalizační úlohy mívají obvykle ekonomický podtext (maximalizace zisku, užitku)[42].

Takto formulovanou obecnou optimalizační úlohu (1.2.1) můžeme v závislosti na charakteru vstupních veličin z množiny $\mathbb{X}$ a „zobrazení" $\mathbb{F}$ rozdělit do dvou základních skupin:

(i) *parametrická optimalizace*, což pro nás bude zejména *matematické programování*[43];

(ii) *funkční optimalizace*, což pro nás bude především *variační počet* a *optimální řízení*.

a Poincarého domněnka). Za řešení každého z nich nabízí ústav odměnu jeden milión dolarů a zatím jich zbývá vyřešit šest, neboť v roce 2010 se podařilo ruskému matematikovi Grigoriji Perelmanovi dokázat Poincarého domněnku. Odměnu si ale nepřevzal, neboť jak sám říká *„Vím, jak ovládat vesmír. Tak mi řekněte, proč bych se měl hnát za milionem?!“*

[15] *Toto přesvědčení o řešitelnosti jakéhokoli matematického problému je pro nás silnou motivací v práci; slyšíme v nás neustálé volání: je tu problém, hledejte řešení. Můžete jej najít čistým myšlením; neboť v matematice neexistuje žádný Ignorabimus.*
(This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no ignorabimus.)

[16]

Emil Heinrich du Bois-Reymond (7. listopadu 1818 – 26. prosince 1896) byl německý lékař a fyziolog, objevitel akčního potenciálu a zakladatel experimentální elektrofyziologie. Jedním z jeho mladších bratrů byl matematik Paul du Bois-Reymond, na kterého ještě později rozhodně narazíme.

[17] *Nevíme a nebudeme vědět.*
(We do not know and will not know.)

V parametrické optimalizaci máme přípustnou množinu $\mathbb{X} := X \subseteq \mathbb{R}^n$ a tato množina obvykle bývá vymezena podmínkami danými soustavou nějakých algebraických nebo transcendentních rovnic či nerovnic. Účelovým „zobrazením“ $\mathbb{F}$ je reálná funkce $f : \mathbb{R} \to \mathbb{R}$ definovaná (alespoň) na přípustné množině. V závislosti na kontextu nalezneme pro funkci f různá označení jako *účelová, cílová, ztrátová, nákladová, nepřímá užitková, užitková, energetická* atd. Úloha (1.2.1) pak znamená najít bod $x^* \in X$, ve kterém daná reálná funkce f nabývá své nejmenší (příp. největší) hodnoty mezi všemi body $x \in X$, tj. máme řešit úlohu

$$f(x) \to \min, \quad x \in X, \tag{1.2.2}$$

přičemž je velmi výhodné množinu X vyjádřit jako (je-li to možné)

$$X = \{x \in \mathbb{R}^n \mid g_1(x) \le 0, \dots, g_L(x) \le 0 \ \& \ h_1(x) = 0, \dots, h_K(x) = 0\}$$

pro $K, L \in \mathbb{N} \cup \{0\}$ a dané funkce $g_1, \dots, g_L, h_1, \dots, h_K$, tj. hledáme bod

$$x^* \in \operatorname*{arg\,min}_{x \in X} f(x) = \{x \in X \mid f(x) \le f(y) \text{ pro všechna } y \in X\}.$$

Úlohu (1.2.2) pak čteme jako:

> *Najdi bod x^*, v němž reálná funkce f dosahuje nejmenší hodnoty (tj. minima) mezi všemi body $x \in D(f) \subseteq \mathbb{R}^n$, které vyhovují soustavě L nerovností*
>
> $$g_1(x) \le 0 \quad \& \quad g_2(x) \le 0 \quad \& \quad \cdots \quad \& \quad g_L(x) \le 0$$
>
> *a soustavě K rovností*
>
> $$h_1(x) = 0 \quad \& \quad h_2(x) = 0 \quad \& \quad \cdots \quad \& \quad h_k(x) = 0.$$

Bod x^* pak v této úloze představuje (*optimální*) *řešení* a zejména v případě $X \subsetneq \mathbb{R}^n$ se hovoří o *optimální strategii* nebo *návrhovém vektoru* daného matematického programu. Místo úlohy (1.2.2) můžeme případně hledat řešení „pouze“ úlohy

$$f^* = \inf_{x \in X} f(x) \tag{1.2.3}$$

Pro úlohu (1.2.2) v jejím obecném pojetí samozřejmě nemáme zaručenou existenci jejího řešení, zatímco v případě úlohy (1.2.3) jej díky vlastnostem reálných čísel najdeme vždy – buď to bude nějaké reálné číslo $x^* \in X$ nebo $x = \pm\infty$, přičemž hodnota $x^* = +\infty$ odpovídá[44] situaci $X = \emptyset$, kdežto $x^* = -\infty$ znamená, že funkce f není zdola ohraničená na X. Pokud ale existuje konečné řešení úlohy (1.2.2), pak to není nic jiného než hledání vrstevnice funkce f na úrovni f^*. Úlohu (1.2.2) můžeme navíc rozdělit ještě do dvou subkategorií:

- úlohy bez omezujících podmínek (tzv. *volné extrémy*), tj. $X = \mathbb{R}^n$, např. $x^2 + 1 \to \min$;
- úlohy s omezujícími podmínkami (tzv. *vázané extrémy*), tj. $X \subsetneq \mathbb{R}^n$, např. $x \cos y \to \min$ pro $[x, y] \in [-5, 5] \times \mathbb{R}$.

S oběma typy úloh jsme se již setkali v základním kurzu matematické analýzy a my si nyní některé pojmy připomeneme.

Definice 1.2.1

Bod $x^* \in X$ nazveme *globálním minimem* funkce f na množině $X \subseteq \mathrm{D}(f)$, tj. (*globálním*) *řešením* úlohy (1.2.2), pokud $f(x^*) \le f(x)$ pro všechny $x \in X$. Bod $x^* \in X$ nazveme *lokálním minimem* funkce f na X, tj. *lokálním řešením* úlohy (1.2.2), jestliže existuje takové ε-okolí bodu x^* neboli množina

$$\mathcal{O}_\varepsilon(x^*) := \{x \in \mathbb{R}^n \mid \|x - x^*\| < \varepsilon\},$$

že platí $f(x^*) \le f(x)$ pro každé $x \in X \cap \mathcal{O}_\varepsilon(x^*)$. V případě ostrých nerovností hovoříme

DEFINICE 1.2.1

o *ostrých* globálních/lokálních minimech. Analogicky definujeme globální/lokální (ostrá) *maxima* funkce f na množině X.

Jak je ale taková řešení (ať již lokální či globální) najít? V případě diferencovatelné funkce f nám s hledáním lokálních řešení může pomoci následující tvrzení.

VĚTA 1.2.2

Nechť pro funkci $f:\mathbb{R}^n\to\mathbb{R}$ existují všechny parciální derivace prvního řádu v bodě x^*, ve kterém má funkce f lokální extrém na $\mathbb{R}^n$. Potom x^* je *stacionárním bodem* funkce f, tj. platí

$$\operatorname{grad} f(x^*) := \begin{pmatrix} f_{x_1}(x^*) \\ f_{x_2}(x^*) \\ \vdots \\ f_{x_n}(x^*) \end{pmatrix} = 0.$$

Důkaz. Nejdříve uvažme situaci $n=1$, tj. $f:\mathbb{R}\to\mathbb{R}$ a $\operatorname{grad} f(x) = f'(x)$. Je-li $f'(x^*)>0$, pak existuje $\mathcal{O}(x^*)$ takové, že $\frac{f(x)-f(x^*)}{x-x^*}>0$ pro všechna $x\in\mathcal{O}(x^*)$. V takovém případě je ale f rostoucí na $\mathcal{O}(x^*)$, a tudíž i v x^*. Analogicky bychom mohli ukázat, že v případě $f'(x^*)<0$ je funkce f klesající v x^*. Proto nutně $f'(x^*)=0$.

Nechť nyní $n\in\mathbb{N}\setminus\{1\}$ a připusťme, že některá z parciálních derivací funkce f v bodě x^* je nenulová, tj. $f_{x_i}\neq 0$ pro nějaké $i\in\{1,\dots,n\}$. V takovém případě je však funkce $g(t) := f(x^*+te_i) = f(x_1^*,\cdots,x_{i-1}^*,x_i^*+t,x_{i+1}^*,\cdots,x_n^*)$ v okolí bodu x^* a ve směru e_i rostoucí, takže nemůže mít v x^* lokální extrém ↯. ■

18

Kurt Gödel (28. dubna 1906 – 14. ledna 1978) byl matematik česko-rakouského původu (narozený v Brně), který se stal jedním z nejvýznamnějších logiků všech dob. Významné jsou i jeho příspěvky ve fyzice a ve filosofii matematiky.

19

Niccolò Fontana (1499/1500 – 13. prosince 1557) byl italský matematik a konstruktér, který se jako první věnoval studiu dráhy letu dělových koulí. Je znám spíše pod přezdívkou Tartaglia znamenající „koktavec“, což odkazuje na jeho vadu řeči způsobenou úderem šavle francouzského vojáka do čelisti při dobývání jeho rodné Brescie v roce 1512.

20

Gerolamo Cardano či Hieronymus

Poznámka 1.2.3 (historická, hodně dlouhá a hodně důležitá). Předchozí tvrzení je známo jako *Fermatova věta*[23], neboť Fermat byl asi prvním, kdo využil charakteristického chování algebraických výrazů poblíž (lokálních) extrémů jakožto kritéria pro jejich nalezení. To úzce souvisí s tzv. *Fermatovou metodou* pro hledání maxim a minim, kterou Fermat využíval k hledání těžiště či řešení problémů z teorie čísel. Tuto metodu „objevil“ Fermat kolem roku 1629 a v [174] ji popisuje takto[45,46,47]:

> *Omnis de inventione maximæ et minimæ doctrina duabus positionibus in notis innititur et hac unica præceptione:*
>
> *Statuatur quilibet quæstionis terminus esse A (sive planum, sive solidum aut longitudo, prout proposito satisfieri par est) et, inventa maxima aut minima in terminis sub A, gradu <aut gradibus>, ut libet, involutis, ponatur rursus idem qui prius terminus esse A+E, iterumque inveniatur maxima aut minima in terminis sub A et E gradibus, ut libet, cœfficientibus. Adæquentur, ut loquitur Diophantus, duo homogenea maximæ aut minimæ aqualia et, demptis communibus (quo peracto, homogenea omnia ex parte alterutra ab E vel ipsius gradibus afficiuntur), applicentur omnia ad E vel ad elatiorem ipsius gradum, donec aliquod ex homogeneis, ex parte utravis, affectione sub E omnino liberetur. Elidantur deinde utrimque homogenea sub E aut <sub> ipsius gradibus quomodolibet involuta, et reliqua equentur, aut, si ex una parte nihil superest, æquentur sane, quod eodem recidit, negata affirmatis. Resolutio ultimæ istius equalitatis dabit valorem A, quai cognita, maxima aut minima ex repetitis prioris resolutionis vestigiis innotescet.*[48]

Vyložený postup následně Fermat ilustruje na příkladu, ve kterém chce rozdělit úsečku spojující body A a C na dvě části tak, aby součin jejich délek byl maximální. Je-li délka původní úsečky AC rovna číslu B, pak jejím rozdělením získáme úsečky o délkách A a $B-A$, takže hledáme maximum výrazu $A(B-A)=AB-A^2$. Nahraďme nyní hodnotu A výrazem $A+E$, takže

Cardanus (24. září 1501 – 20. září 1576), byl italský polyhistor, jehož zájmy a dovednosti sahaly od matematiky přes lékařství, biologii, fyziku, astrologii a astronomii až po gamblerství. Je jedním z nejvýznamnějších představitelů rozvoje přírodních věd období renesance. Byl také klíčovou osobností na počátku studia pravděpodobnosti a snad i prvním, kdo „na západě“ využíval binomická čísla a binomickou větu.

[21] Antonio Maria del Fiore (přelom 15. a 16. století) byl italský renesanční matematik, který se od svého učitele Scipioneho del Ferra naučil řešit redukovanou kubickou rovnici $x^3 + px = q$ pro libovolná čísla $p, q > 0$.

[22] *Neexistují čísla* $x, y, z \in \mathbb{N}$ *a* $n \in \mathbb{N} \setminus \{1, 2\}$ *splňující* $x^n + y^n = z^n$.

[23]

Pierre de Fermat (někdy mezi 31. říjnem a 6. prosincem 1607 – 12. ledna 1665) byl francouzský právník a matematik, který se zasloužil o rozvoj teorie čísel, pravděpodobnosti a kterému také patří uznání za rozpracování metody hledání extrému křivky jakožto předchůdce infinitesimálního počtu, viz později. Je také znám díky tzv. *Fermatovu principu* z optiky. Více viz později.

[24]

Diofantos z Alexandrie byl starověký řecký matematik působící ve 3. století n. l. v egyptské Alexandrii. Je prvním matematikem, který řešil ryze algebraické problémy a zavedl pro ně důmyslný systém značek. Ve svém nejdůležitějším díle *Aritmetika* popisuje ve 13 knihách mj. všechna známá řešení lineárních a kvadratických rovnic. Ačkoli se nám dodnes z původních 13 knih dochovala pouze polovina, je zcela po zásluze označován za „otce algebry“.

[25] *Je nemožné, aby třetí mocnina byla součtem dvou třetích mocnin, čtvrtá mocnina součtem dvou čtvrtých mocnin nebo obecně aby libo-*

součin obou délek je $(A+E)(B-A-E) = AB - A^2 + BE - 2AE - E^2$. Tato hodnota musí „být rovna“ původní délce $AB - A^2$, což

po odstranění stejných výrazů na obou stranách rovnosti dává: $BE \sim 2AE + E^2$,

kde symbol ~ odpovídá Fermatovu označení „adæquentur“ (srovnání). Následným

podělením obou stran veličinou E dostaneme: $B \sim 2A + E$,

odkud

odstraněním E dospějeme k výsledku: $B = 2A$,

tj. maximum součinu získáme rozdělením úsečky na dvě stejné části[49]. Fermat tento příklad zakončuje slovy

> *… nec potest generalior dari methodus.*[50]

Ale měl skutečně pravdu? Z dnešního pohledu totiž Fermatova metoda není nic jiného než hledání stacionárního bodu, ačkoli zde explicitně nevystupují limity či derivace, které samozřejmě byly ve Fermatově době zcela neznámé. V moderním pojetí hledáme x takové, že

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 0.$$

Jenže Fermatův návod týkající se nahrazení A výrazem $A + E$ a následné odstranění stejných výrazů je mnohem sofistikovanější, neboť vede k získání rozvoje výrazu $f(x+h) - f(x)$ v mocninách h. Podělení E pak odpovídá zlomku $\frac{f(x+h)-f(x)}{h}$ a poslední krok týkající se „vynulování“ E je vlastně způsobem, jak získat lineární člen v rozvoji $f(x+h) - f(x)$. Jenže tento člen odpovídá první derivaci funkce f, což velmi dobře víme díky Taylorovu rozvoji, podle kterého při splnění patřičných předpokladů (což ve Fermatově případě jistě není problém) platí

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{f''(x)}{2!}h + \frac{f'''(x)}{3!}h^2 + \ldots, \tag{1.2.4}$$

takže „vynulováním“ h na pravé straně získáme $f'(x)$, což vzhledem k Fermatově předpokladu „srovnatelnosti“ $f(x+h)$ a $f(x)$ vede k rovnici $f'(x) = 0$. Fermat se také zmiňuje o tom, že v případě potřeby, je možné obě strany rovnice podělit i vyšší mocninnou E tak, aby alespoň na jedné straně zcela zmizelo E. Sám ale žádný takový příklad nenabízí – ani nemůže, protože z rozvoje (1.2.4) můžeme vidět, že by to bylo možné pouze v případě $f'(x) \equiv 0$, tj. je-li daná funkce konstantní, což je pochopitelně pro optimalizování zcela nezajímavý případ[51]. Proč to tedy Fermat uváděl? Nejjednodušší vysvětlení, které se zde nabízí, je to, že si nebyl ohledně této situace příliš jistý. Proto se raději ukryl za slovní spojení „v případě potřeby“, které už od dávných věků poskytuje matematikům velmi nedobytnou tvrz.

Neměli bychom ale opomenout ještě jeden velmi důležitý fakt. Fermatova metoda skutečně funguje pouze pro nalezení stacionárních bodů – nikoli extrémů! Aplikováním této metody na klasický příklad $f(x) = x^3$ nás totiž dovede k vyjádření

$$3A^2 + 3AE + E^2 \sim 0,$$

což po odstranění E dává $A = 0$. My ale velmi dobře víme, že toto rozhodně není extrém dané funkce. V tomto ohledu jsou všechny Fermatem nabízené příklady trochu zavádějící, neboť v nich vždy je stacionární bod zároveň hledaným extrémem.

Fermatovo pojednání, jehož součástí byl i druhý příklad zaměřený na hledání tečny k libovolné křivce, bylo publikováno až posmrtně v roce 1679 díky jeho synovi Samuelovi, a to v souhrnné podobě zahrnující několik jeho prací. Za Fermatova života nebyly vědecké práce publikovány tak, jak to známe z dnešní doby. Tehdejší učenci si své poznatky a objevy sdělovaly především

prostřednictvím vzájemné korespondence. V tomto ohledu zde sehrál klíčovou roli františkánský mnich Marin Mersenne[52], jehož prostřednictvím se dostalo Fermatovo pojednání k rukám (a očím) René Descartesovi[53]. U něj vyvolala Fermatova metoda značné pochybnosti, což ale jistě souviselo s Fermatovou předchozí kritikou Descartesova pojednání *La dioptrique*, které bylo Descartesovým vrcholným příspěvkem k rozvoji optiky (mj. je v něm poprvé popsán zákon lomu světelného paprsku, který je dnes znám spíše jako *Snellův zákon*). Avšak zde je nutné přiznat Mersennovi vzácný talent „činit špatné věci ještě horšími“, když opomněl v dopise Descratesovi dovysvětlit některé souvislosti vyplývající z jeho vzájemné komunikace s Fermatem. O Descartesově znalosti Fermatova pojednání víme z jeho odpovědi Mersennovi dopisem z 18. ledna 1638, ve kterém píše (aniž by byl ochoten napsat Fermatovo jméno) následující.

> *Byl bych rád, kdybych věděl, co řekne jak o připojeném dopise, kde reaguji na jeho pojednání o maximech a minimech, tak o předchozím, kde jsem odpověděl na jeho vystoupení proti mé La dioptrique. Jeden i druhý jsem napsal proto, aby je viděl, pokud chcete; ani jsem ho nechtěl jmenovat, aby se mohl cítit méně zahanbený chybami, které jsem tam našel, protože mým úmyslem není nikoho urážet, ale pouze bránit sám sebe. A protože mám pocit, že nezklame v překonávání mých předsudků vůči jeho mnoha spisům, považuji za vhodné, aby lidé viděli i moji obranu. Proto vás žádám, abyste mu je neposílal, aniž byste si neponechal jejich kopie. A pokud i poté bude hovořit o tom, že vám chce poslat další texty, prosím vás, abyste ho požádal, aby je promyslel pečlivěji než ty předchozí; jinak vás žádám, abyste nepřijal mé zmocnění k jejich přeposlání ke mně. Neboť, jen tak mezi námi, pokud mi chce prokázat tu čest předložením námitky a nechce si dělat více problémů, než tomu bylo poprvé, byl bych zahanben, kdyby bylo nutné, abych si dal tu práci odpovědět na takovou malichernost, ačkoli bych se tomu nemohl čestně vyhnout, kdyby věděl, že jste mi to poslal.*[54]

Jeden z přiložených dopisů pak je věnován právě Fermatově metodě a Descartes v něm pokračuje ve své kritice.

> *Raději bych neříkal nic o příspěvku, který jste mi zaslal, protože bych nemohl říci nic, co by bylo ve prospěch člověka, který jej napsal. Ale protože si uvědomuji, že je to stejný člověk, který se dříve pokoušel vyvrátit moji La dioptrique, a protože jste mě informoval, že vám zaslal svoji esej poté, co si přečetl moji La Géométrie, kde byl překvapen, že jsem nenašel to stejné – což chápu jako úmysl vyvolat spor a zároveň ukázat, že ví o tématu více než já – a poté také proto, že se z Vašich dopisů dozvídám, že má pověst velmi učeného v matematice, cítím povinnost mu odpovědět.*[55]

Nicméně spor byl celkem rychle zažehnán, když si oba matematici vzájemně vyjasnili své pochybnosti. Descartes pak v dopise z 27. července 1638 určeném Fermatovi píše:

> *… když vidím poslední metodu, kterou používáte pro hledání tečen křivek, nemohu na to odpovědět jiným způsobem, než říci, že je velmi dobrá a že kdybyste to takto vysvětlil hned na začátku, nebyli bychom vůbec v žádném rozporu.*[56]

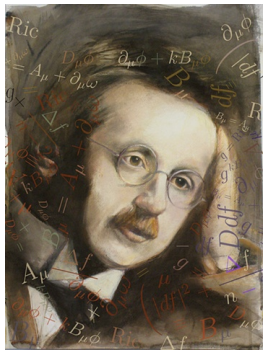
Každopádně to samozřejmě motivovalo Fermata k další práci na metodě hledání maxim a minim, výsledkem čehož byla série dalších několika pojednání v nichž přichází s dalšími příklady, na kterých ilustruje její využití např. při hledání těžiště paraboloidu; rozdělení úsečky tak, aby hodnota $A^2(B-A)$ byla maximální; nalezení válce s maximálním povrchem, jenž je vepsán

volná mocnina stupně vyššího než dva nějakého čísla byla součtem dvou mocnin téhož stupně. Objevil jsem vskutku pozoruhodný důkaz tohoto tvrzení, pro který je ale tento okraj příliš malý, aby se do něj vešel. (It is impossible for a cube to be the sum of two cubes, a fourth power to be the sum of two fourth powers, or in general for any number that is a power greater than the second to be the sum of two like powers. I have discovered a truly marvelous demonstration of this proposition that this margin is too narrow to contain.)

[26] Sir Andrew John Wiles (⋆11. dubna 1953) je anglický matematik, profesor na Oxfordské univerzitě a (mimo jiné) držitel Abelovy medaile z roku 2016.

[27] *Tím nejpřekvapivějším na matematice je asi to, že je tak překvapující…* Viz [435, str. 75].

[28]



Hermann Klaus Hugo Weyl (9. listopadu 1885 – 8. prosince 1955) byl německý matematik, teoretický fyzik a filosof, který patřil k nejvlivnějším matematikům 20. století. Ačkoli většina jeho profesního života je spojena se Spolkovou vysokou technickou školou v Curychu (ETH Zurich) a univerzitou v Princetonu (centrem Institute for Advanced Study), je brán především jako další pokračovatel matematické školy Univerzity v Göttingenu, kde v roce 1908 získal doktorát pod Hilbertovým vedením.

[29] *Ve své práci jsem se vždy snažil spojit pravdu s krásou, a když jsem si musel vybrat jedno nebo druhé, obvykle jsem si vybral tu krásu.* Viz [149].

[30] Bertrand Arthur William Russell (18. května 1872 – 2. února 1970), třetí hrabě Russell, byl britský polyhistor. Jeho práce měly značný vliv na matematiku, logiku, teorii množin (zde je jeho jméno spojováno zejména s paradoxem známým např. v podobě holiče holící všechny muže, kteří se neholí sami – tento paradox ukazuje vnitřní

spornost naivní teorie množiny), lingvistiku, umělou inteligenci, kognitivní vědy, informatiku a různé oblasti analytické filosofie (zejména filosofie matematiky, jazyku, epistemologie a metafyziky). Byl veřejným intelektuálem, historikem, sociálním kritikem, politickým aktivistou a nositelem Nobelovy ceny za literaturu z roku 1950. Narodil se do jedné z nejvýznamnějších aristokratických rodin ve Spojeném království.

[31] *Matematika se popravdě řečeno nepyšní pouze pravdivostí, ale také svrchovanou krásou – krásou chladnou a strohou, jakou nalézáme u soch; krásou, jež se neodvolává k žádné slabé stránce naší přirozenosti; krásou bez zbytečného pozlátka, které je běžné v hudbě či malířství – a přesto vznešeně čistou a stroze dokonalou, takovou, již se pyšní jen největší umělecká díla. Skutečného ducha slasti, povznesení, pocitu být více než Člověk, který je prubířským kamenem nejvyšší dokonalosti, lze nalézt v matematice stejně jako v poezii.* Viz [437, str. 60].

[32] Alfred North Whitehead (15. února 1861 – 30. prosince 1947) byl anglický matematik a filosof. Je považován za jednoho z nejvýznamnějších anglo-amerických filosofů 20. století. Společně se svým bývalým studentem Russellem napsal třísvazkové dílo *Principia Mathematica*, které patří k nejdůležitějším dílům matematické logiky 20. století.

[33] *Čistou matematiku v její moderní podobě můžeme prohlásit za nejoriginálnější výtvor lidského ducha.* Viz [509, str. 20].

[34] Tj. v rámci primárního a sekundárního vzdělávání na základních a středních školách.

[35] *Pokud to nelze vyjádřit v číslech, není to věda; je to pouhý názor.* Viz [258, str. 240(???)].

[36]

Galileo di Vincenzo Bonaulti de Galilei (15. února 1564 – 8. ledna 1642) byl italský astronom, fyzik a konstruktér. Bývá označován za „otce pozorovací astronomie“ a „otce moderní fyziky“ či dokonce „otce moderní vědy“.

do koule; a v dalších geometrických úlohách zahrnujících i hledání různých tečen. V pojednání [175], o jehož správné dataci panují jisté pochybnosti (1644 či dříve?), konečně nalezneme i Fermatovo zdůvodnění správnosti výsledku příkladu, který jsme si na začátku ukázali, tj. $A = B/2$ s maximálním součinem $B^2/4$:

At, si recta eadem B proponatur secanda eâ conditione ut rectangulum sub ejus segmentis sit æquale Z plano (quod supponendum minus quadrante B quadrati), tunc duo puncta proposito satisfacient, quæ quidem a puncto maximi rectanguli intercipiuntur.

Sit enim alicujus recte B segmentum A, fiet

$$B \text{ in } A - A \text{ quad.} \quad \text{æquale} \quad Z \text{ plano,}$$

quat æquatio est anceps et rectam A de duobus lateribus explicari posse indicat. Sit igitur sequatio correlata

$$B \text{ in } E - E \text{ quad.} \quad \text{æquale} \quad Z \text{ plano;}$$

ex methodo Vietæa comparentur hæ duæ æquationes:

$$B \text{ in } A - B \text{ in } E \quad \text{æquabitur} \quad A \text{ quad.} - E \text{ quad.,}$$

et, omnibus per $A - E$ divisis, fiet

$$B \quad \text{æqualis} \quad A + E,$$

ipsæque A et E erunt inæquales.

Si sumatur aliud planum, loco Z plani, quod sit majus quam Z planum, sed minus quadrante B quadrati, tune recte A et E minus inter se different quam superiores, quum puncta divisionis magis accedent ad punctum rectanguli maximi constitutivum, semperque, auctis divisionum rectangulis, ipsarum A et E. differentia minuetur, donec per ultimam maximi rectanguli divisionem evanescat, quo casu μοναχή vel unica continget solutio, quum duæ æquales <fient> quantitates, hoc est, A æquabitur E.

Quum igitur, in duabus superioribus æquationibus correlatis, per methodum Vietæam, B æquabitur $A + E$, si E æquetur ipsi A (quod contingere semper in puncto maximæ vel minimæ constitutivo apparet), ergo, in casu proposito,

$$B \quad \text{æquabitur} \quad A \text{ bis:}$$

hoc est, si recta B bifariam secetur, rectangulum sub ipsius segmentis erit maximum.[57]

Fermat dále nabízí podobné zdůvodnění i pro několik dalších dříve vyřešených příkladů a toto své pojednání zakončuje komentářem (zřejmě) narážejícím na Descartesovu kritiku[58] a jedním zajímavým příkladem k samostatnému řešení.

Si qui adhuc supersunt qui methodum hanc nostram debitam sorti pronuntiant,

Hos cupiam similes tentando excudere sortes.

Qui hanc methodum non probaverit, ei proponitur:

Datis tribus punctis, quartum reperire, a quo si ducantur tres rectae ad data puncta, summa trium harum rectarum sit minima quantitas.[59]

Pro více podrobností odkazujeme zvídavé čtenáře na překlad Fermatových prací [174]– [175] dostupný v anglickém jazyce společně i s dalšími texty obsahujícími výše zmíněné příklady, na Kapitolu IV v knize [361, str. 143–213] a také na články [319, 465]. ▲

Ale my velmi dobře víme (a vlastně jsme to zmínili i v předchozí poznámce), že Věta 1.2.2 udává pouze nutnou podmínku pro existenci lokálního extrému. Pro ověření, zda ve stacionárním bodě skutečně nastává extrém, potřebujeme nějaký další nástroj – například podmínku druhého řádu, která je založena na definitnosti *Hessovy matice* funkce f, tj. $n \times n$ matice

$$\nabla^2 f(x^*) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) \right)_{i,j=1,\dots,n}$$

která v případě funkce f se spojitými parciálními derivaci druhého řádu bude jistě symetrická. Proto si nyní nejdříve připomeneme pár pojmů z lineární algebry.

DEFINICE 1.2.4

Matice $M \in \mathbb{R}^{n\times n}$ se nazývá *pozitivně semidefinitní*, pokud

$$v^\top A\, v \geq 0 \quad \text{pro každé } v \in \mathbb{R}^n,$$

což zapisujeme jako $M \geq 0$. Jestliže dokonce platí $v^\top M\, v > 0$ pro všechny vektory $v \in \mathbb{R}^n \setminus \{0\}$, nazývá se matice M *pozitivně definitní*, což zapisujeme jako $M > 0$. Pokud platí opačné nerovnosti, nazývá se matice M *negativně definitní* či *negativně semidefinitní*. Konečně v případě, že existují vektory $v \in \mathbb{R}^n \setminus \{0\}$ a $u \in \mathbb{R}^n \setminus \{0\}$ takové, že $v^\top M\, v > 0$ a $u^\top M\, u < 0$, nazývá se matice M *indefinitní*.

Předchozí definice explicitně nevyžaduje, aby matice A byla symetrická. Avšak my budeme definitnost testovat právě na symetrických maticích. V takovém případě pak výraz $v^\top A\, v$ udává *kvadratickou formu* v proměnných $v_1, \dots, v_n$. Asi není příliš těžké ověřit podmínky z Definice 1.2.4, abychom zjistili, že platí např.

$$\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \\ 0 & \dots & 1 \end{pmatrix} > 0, \quad \begin{pmatrix} -1 & \dots & 0 \\ \vdots & \ddots & \\ 0 & \dots & -1 \end{pmatrix} < 0,$$

$$\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} < 0, \quad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} > 0, \quad \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \leq 0, \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \geq 0.$$

Avšak tento způsob určování definitnosti není příliš efektivní (a o to nám tu přeci jde především). První možná ekvivalentní charakterizace je založena na vlastních číslech.

VĚTA 1.2.5

Je-li M symetrická $n \times n$ matice s vlastními čísly $\lambda_1, \dots, \lambda_n$, pak matice M je

(i) pozitivně definitní právě tehdy, když $\lambda_i > 0$ pro $i = 1, 2, \dots, n$;

(ii) negativně definitní právě tehdy, když $\lambda_i < 0$ pro $i = 1, 2, \dots, n$;

(iii) pozitivně semidefinitní právě tehdy, když $\lambda_i \geq 0$ pro $i = 1, 2, \dots, n$;

(iv) negativně semidefinitní právě tehdy, když $\lambda_i \leq 0$ pro $i = 1, 2, \dots, n$.

Díky tomuto kritériu můžeme snadno ukázat, např. že

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} > 0 \quad \text{a} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \geq 0.$$

[37] *Filozofie [přírody] je psána v této úžasné knize (mluvím o vesmíru), kterou máme neustále otevřenou před očima. Ovšem nelze ji porozumět dříve, než se naučíme jazyk a znaky, ve kterých je napsána. Je psána jazykem matematiky a jeho znaky jsou trojúhelníky, kruhy a další geometrické útvary, bez nichž je lidsky nemožné porozumět jedinému slovu; bez nich se člověk marně potuluje temným labyrintem.*
(Philosophy [of nature] is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles and others geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth.) Viz [198, str. 25].

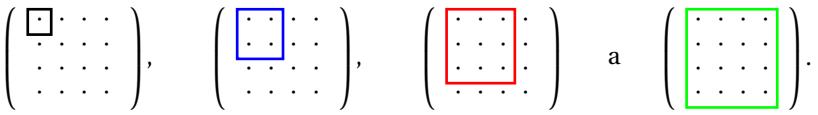
[38] Godfrey Harold Hardy (7. února 1877 – 1. prosince 1947) byl anglický matematik, který se věnoval především teorii čísle a matematické analýze. Jeho jméno je úzce spojeno s tzv. Hardyho-Weinbergovým principem v populační genetice. Mimo matematiku je Hardy znám díky své eseji [252], která bývá považována za jeden z nejlepších vhledů do matematikovy mysli určeným pro laiky. Když se slavný maďarský matematik Pál Erdös zeptal Hardyho, co bylo jeho největším příspěvkem na poli matematiky, Hardy bez rozmýšlení odpověděl, že to bylo objevení indického matematika Šrínivási Rámanudžana, jehož učitelem byl od roku 1914.

[39] *Matematikovy vzorce, stejně jako ty malířovy či básníkovy, musí být krásné; myšlenky, stejně jako barvy nebo slova, musí být v harmonickém souladu. Krása je první zkouškou: na světě není trvalé místo pro ošklivou matematiku.* Viz [252, str. 85].

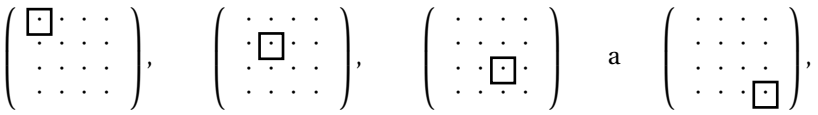
[40] Zde jsou uvozovky zcela na místě, neboť skutečné „zobrazení" by pro naše úvahy mohlo být příliš obšírné.

[41] Symbol $\bowtie$ nahrazuje některou z nerovností $\leq$, $\geq$, $<$ nebo $>$, jejichž konkrétní volba je odvislá od dané úlohy.

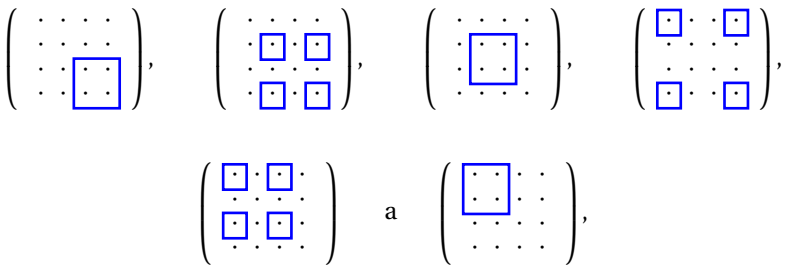
Avšak je to použitelné i pro matice větších rozměrů? Počítání vlastních čísel matice M vlastně znamená nalézt kořeny polynomu stupně n, což je snadné v případě $n = 2$, pro $n = 3$ už je to složitější ale stále řešitelné s využitím Cardanových vzorců, s jejichž pomocí lze díky Lodovicu Ferrarimu[60] vyřešit i případ $n = 4$, zatímco pro $n \geq 5$ už žádný funkční vzorec nemáme (a ani mít nemůžeme, viz tzv. *Abelovu–Ruffiniho větu*). Proto bychom se měli raději počítání vlastních čísel vyhnout a využít raději charakterizaci definitnosti matice založenou na počítání determinantů tzv. *vedoucích hlavních submatic* a *hlavních submatic*. Ty vzniknou tak, že v matici M vynecháváme stejné skupiny řádků a sloupců. Přesněji, vedoucí hlavní submatici řádu m získáme tak, že v matici M vynecháme posledních $n - m$ řádků a sloupců, takže pro $m = 1$, $m = 2$, $m = 3$ a $m = 4$ máme postupně submatice



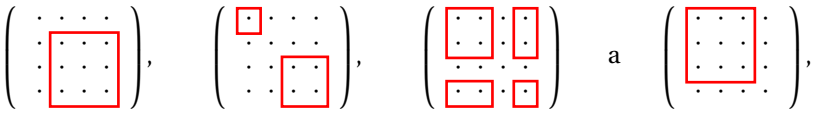
Takovýchto submatic budeme mít celkem n. Hlavní submatice řádu m získáme tak, že vynecháme skupinu $n - m$ řádků a sloupců s týmž indexem. Proto hlavní submatice řádu $m = 1$ jsou (vynecháme vždy $n - 1$ řádků a sloupců, což můžeme udělat celkem n způsoby)



hlavní submatice řádu $m = 2$ jsou tvořeny vyznačenými prvky (vynecháme vždy $n - 2$ řádků a sloupců, což můžeme udělat celkem $\binom{n}{2}$ způsoby)



hlavní submatice řádu $m = 3$ jsou tvořeny vyznačenými prvky (vynecháme vždy $n - 3$ řádků a sloupců, což můžeme udělat celkem $\binom{n}{3}$ způsoby)



atd. Takovýchto submatic budeme celkem mít

$$\binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n} = 2^n - 1.$$

Např. pro matici 3×3 máme 3 vedoucí hlavní submatice a 7 hlavních submatic. Pro definitnost matice M jsou klíčová znaménka determinantů těchto submatic, o kterých se hovoří jako o *vedoucích hlavních minorech* a *hlavních minorech*. Pro důkaz následujícího tvrzení viz např. [276, Theorem 7.2.5].

[42] Toto rozdělení mezi fyziku a ekonomii samozřejmě není nikterak striktní, můžeme mít např. maximalizaci účinnosti v elektrickém obvodu nebo minimalizaci nákladů. Navíc toto rozhodně nejsou jediné dvě oblasti, kde se můžeme s optimalizačními problémy setkat. Nicméně náš teoretický výklad je (především z historických důvodů) zaměřen výhradně na minimalizační úlohy.

[43] Z dnešního pohledu se může zdát označení „programování“ trochu nešťastné a zavádějící. Jedná se o pozůstatek vojensky motivované terminologie, která odkazuje na *plánování* či *rozvrhování* tréninků, logistiky či rozmístění mužstva. Více se o tom dozvíte v další sekci a na začátku následující kapitoly.

[44] Jelikož infimem rozumíme největší dolní závoru dané množiny, takže zcela „přirozeně“ může brát $\inf \emptyset = +\infty$.

[45] Fermat se zde drží Viètova způsobu zápisu rovnic, kdy používá samohlásky pro označení neznámých veličin (proměnných) – zejména A a E. Pro konstantní veličiny pak používá souhlásky počínaje B. Fermat také podobně jako Viète nikdy nezpřetrhal propojení mezi algebrou a geometrií, tj. uvažované hodnoty proměnných byly vždy limitovány požadavkem na jejich geometrický význam.

[46] François Viète (1540 – 23. února 1603) byl francouzský matematik, jehož práce na „nové algebře“ byla velmi důležitým krokem k dnešní moderní algebře díky jeho inovativnímu zápisu rovnic, v němž jsou jednotlivé parametry rovnic (proměnné) označovány písmeny. Byl to tedy on, kdo připravil půdu pro objevy Descartese, Fermata a Newtona. Povoláním byl advokát a sloužil i jako poradce Jindřichu III. a Jindřichu IV. Pro oba tyto králetaké rozluštil několik šifer. Objevil mimo jiné historicky první nekonečný součin pro vyjádření π:

$$\pi = 2 \cdot \frac{2}{\sqrt{2+\sqrt{2}}} \cdot \frac{2}{\sqrt{2+\sqrt{2+\sqrt{2}}}} \cdots$$

[47] Také bychom zde měli zdůraznit, že Fermat pracuje výhradně s polynomy (případně s odmocninami, které ale lze umocněním „odstranit“). S čím také jiným v tehdejší době, že?

VĚTA 1.2.6 SYLVESTROVO KRITÉRIUM

Nechť $M \in \mathbb{R}^{n\times n}$ je symetrická matice. Potom

(i) matice $M > 0$ právě tehdy, když její *její vedoucí hlavní minory* jsou kladné;

(ii) matice $M < 0$ právě tehdy, když její *vedoucí hlavní minory* střídají znaménko počínaje záporným;

(iii) matice $M \geq 0$ právě tehdy, když její *hlavní minory* jsou nezáporné;

(iv) $M \leq 0$ právě tehdy, když její *hlavní minory* lichých řádů jsou nekladné a HM sudých řádů jsou nezáporné.

Díky předchozímu tvrzení[61] už můžeme celkem snadno ověřit splnění ať už nutné, či postačující podmínky pro lokální extrém, jejichž důkaz je založen na Taylorově rozvoji[62].

VĚTA 1.2.7

Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ má v bodě $x^* \in \mathbb{R}^n$ a nějakém jeho okolí spojité parciální derivace druhého řádu. Nastává-li v x^* lokální minimum funkce f, pak nutně $\nabla^2 f(x^*) \geq 0$.

Důkaz. Nechť funkce f má v x^* lokálním minimum. Pro libovolná $\alpha > 0$ a $y \in \mathbb{R}^n \setminus \{0\}$ dostaneme z Taylorova rozvoje druhého řádu rovnost

$$f(x^* + \alpha y) = f(x^*) + \alpha \langle \operatorname{grad} f(x^*), y\rangle + \frac{\alpha^2}{2} y^\top \nabla^2 f(x^*)\, y + \omega(\alpha y).$$

která se vzhledem ke skutečnosti, že $\operatorname{grad} f(x^*) = 0$, redukuje na

$$\frac{f(x^* + \alpha y) - f(x^*)}{\alpha^2} = \frac{1}{2} y^\top \nabla^2 f(x^*)\, y + \frac{\omega(\alpha y)}{\alpha^2}.$$

Jelikož $f(x^* + \alpha y) - f(x^*) \geq 0$ pro všechna dostatečně malá α, tj. pro všechna $\alpha \in (0, \alpha_0)$ s vhodným $\alpha_0 > 0$, plyne odtud

$$0 \leq \frac{f(x^* + \alpha y) - f(x^*)}{\alpha^2} = \frac{1}{2} y^\top \nabla^2 f(x^*)\, y + \frac{\omega(\alpha y)}{\alpha^2},$$

z čehož limitním přechodem pro $\alpha \to 0^+$ získáme díky libovolnosti y, že nutně $\nabla^2 f(x^*) \geq 0$, neboť $\lim_{\alpha\to 0^+} \frac{\omega(\alpha y)}{\alpha^2} = \lim_{\alpha\to 0^+} \frac{\omega(\alpha y)}{\|\alpha y\|^2} \cdot \|y\|^2 = 0$. ■

VĚTA 1.2.8

Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ má v bodě $x^* \in D(f)$ a nějakém jeho okolí spojité parciální derivace druhého řádu. Pokud navíc $\operatorname{grad} f(x^*) = 0$ a $\nabla^2 f(x^*) > 0$, pak v x^* nastává ostré lokální minimum funkce f. Je-li $\nabla^2 f(x^*) < 0$, pak v x^* nastává ostré lokální maximum f, zatímco v případě indefinitní Hessovy matice $\nabla^2 f(x^*)$ extrém nenastává (x^* je sedlový bod).

Důkaz. Nechť $\operatorname{grad} f(x^*) = 0$ a $\nabla^2 f(x^*) > 0$, přičemž $\lambda_{\min} > 0$ značí nejmenší vlastní číslo této matice. Pak $\nabla^2 f(x^*) - \lambda_{\min} I \geq 0$, takže $y^\top \nabla^2 f(x^*)\, y \geq \lambda_{\min} \|y\|^2$ pro libovolné $y \in \mathbb{R}^n$. Potom opět s využitím Taylorova rozvoje druhého řádu získáme

$$f(x^* + y) - f(x^*) = \langle \operatorname{grad} f(x^*), y\rangle + \frac{1}{2} y^\top \nabla^2 f(x^*)\, y + \omega(y) = \frac{1}{2} y^\top \nabla^2 f(x^*)\, y + \omega(y) \geq$$
$$\geq \|y\|^2 \left(\frac{\lambda_{\min}}{2} + \frac{\omega(y)}{\|y\|^2} \right).$$

Protože $\lim_{\|y\|\to 0} \frac{\omega(y)}{\|y\|^2} = 0$, existuje číslo $\delta > 0$ takové, že

$$\left| \frac{\omega(y)}{\|y\|^2} \right| < \frac{\lambda_{\min}}{2} \quad \text{pro všechna } y \in \mathbb{R}^n \text{ splňující } \|y\| < \delta.$$

Proto $\frac{\lambda_{\min}}{2} + \frac{\omega(y)}{\|y\|^2} > 0$ pro všechna taková y, takže $f(x^* + y) - f(x^*) > 0$ neboli $f(x) > f(x^*)$ na dostatečně malém okolí $\mathcal{O}(x^*)$, tj. x^* je ostré lokální minimum f. Druhou část tvrzení můžeme dokázat analogicky s využitím největšího vlastního čísla $\lambda_{\max} < 0$ matice $\nabla^2 f(x^*) < 0$ a poslední část pak získáme zkombinováním obou těchto úvah. ■

[48] *Celá teorie studia maxim a minim vyžaduje zastoupení dvou veličin a takovéto jediné pravidlo:*
Nechť A je libovolně zvolená proměnná, kterou chceme určit (ať už je jednodimenzionální, dvoudimenzionální nebo třídimenzionální, jak vyplývá z tvrzení). Maximalizovanou nebo minimalizovanou veličinu vyjádříme pomocí A s využitím výrazů jakéhokoli stupně. Potom dosadíme $A+E$ za původní neznámou A a maximalizovanou nebo minimalizovanou veličinu vyjádříme s pomocí výrazů obsahujících A a E v jakémkoli stupni. Budeme, řečeno slovy Diofantos, „ad-equate" (srovnávat) ona dvě vyjádření pro maximalizovanou nebo minimalizovanou veličinu a odstraníme z nich výrazy společné pro obě strany. Po této úpravě zjistíme, že výrazy na obou stranách obsahují pouze E nebo mocninu E. Všechny výrazy podělíme E nebo vyšší mocninou E tak, že alespoň na jedné ze stran zcela zmizí E. Potom eliminujeme všechny výrazy, ve kterých ještě stále zůstalo E (nebo nějaká jeho mocnina), a všechny ostatní výrazy považujeme za rovnocenné nebo, pokud na jedné ze stran nic nezbylo, ztotožníme přičítané výrazy a odčítané výrazy, což je totéž. Řešením této poslední rovnice dostaneme hodnotu E, která nás přivede k maximu nebo minimu v původním výrazu.
(The whole theory of the study of maxima and minima assumes the position of two unknowns and this sole rule:
Let A be an arbitrarily chosen unknown of the question (whether it has one, two, or three dimensions, as follows from the statement). We will express the maximum or minimum quantity in terms of A, by means of terms of any degree. We will then substitute $A+E$ for the primitive unknown A, and express the maximum or minimum quantity in terms containing A and E to any degree. We will *ad-equate*, to speak like Diophantus, the two expressions of the maximum and minimum quantity, and we will remove from them the terms common to both sides. Having done this, it will be found that on both sides, all the terms will involve E or a power of E. We will divide all the terms by E, or

by a higher power of E, such that on at least one of the sides, E will disappear entirely. We will then eliminate all the terms where E (or one of its powers) still exists, and we will consider the others equal, or if nothing remains on one of the sides, we will equate the added terms with the subtracted terms, which comes to be the same. Solving this last equation will give the value of A, which will lead to the maximum or the minimum, in the original expression.)

[49] Toto je vlastně analogie klasického příkladu o nalezení rozměrů obdélníku daného odvodu majícího maximální obsah.

[50] *... je nemožné poskytnout obecnější metodou.*
(... it is impossible to give a more general method.)

[51] *Největší životní umění je neoptimalizovat konstantní funkce.* Viz předmluvu k [143].

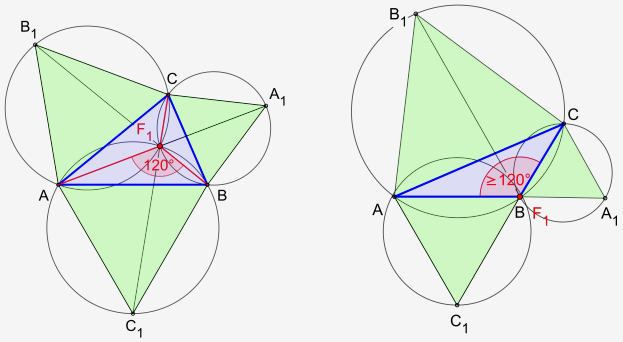
[52]

Marin Mersenne (8. září 1588 – 1. září 1648) byl francouzský kněz, teolog, filosof, matematik, fyzik a muzikolog. V matematice je jeho jméno spojeno především s tzv. *Mersennovými prvočísly*, což jsou prvočísla ve tvaru $2^n - 1$ pro $n \in \mathbb{N}$. V knize *Cogitata Physico–Mathematica* z roku 1644 totiž zveřejnil domněnku, že čísla ve tvaru $2^n - 1$ jsou prvočísla pouze pro $n \in \{2,3,5,7,13,17,19,31,67,127,257\}$, zatímco pro ostatní hodnoty $n \leq 257$ se jedná o čísla složená. Vzhledem k velikosti těchto čísel ale nebylo v Mersennových silách ověřit s právnost této domněnky. To se podařilo až o cca 300 let později, kdy se např. pomocí Lucasova–Lehmerova testu ukázalo, že v seznamu chybí 61, 89 a 107, zatímco hodnoty 67 a 257 do něj nepatří. Rozšiřování Mersennova seznamu o další hodnoty n je aktivní i v dnešní době a je to skvělý zdroj obřích prvočísel, které jsou užitečná např. v kryptografii. Aktuálně největší známé prvočíslo odpovídá $n = 82\,589\,933$ a bylo objeveno 7. prosince 2018. Avšak skutečný Mersennův přínos pro

Pochopitelně klasické příklady s funkcemi $f(x) = x^3$ a $f(x) = x^4$, ukazují, že samotná podmínka $\nabla^2 f(x) \geq 0$ nestačí k zaručení lokálního minima ve stacionárním bodě a naopak podmínka $\nabla^2 f(x) > 0$ nemusí být v lokálním minimu splněna. Tím je otázka existence lokálních extrémů víceméně uzavřena a my bychom tak mohli s pomocí předchozích výsledků vyřešit např. následující dva slavné problémy. Později uvidíme (Kapitola 6), že v těchto úlohách hraje nezanedbatelnou roli konvexnost účelových funkcí, která je činí „snadnými“, tj. stačí najít „pouze“ stacionární bod.

Příklad 1.2.9

Už jsme se zmínili o Fermatem nabízené úloze[59] pro zbývající pochybovače. Jejím řešením je tzv. Fermatův bod či Fermatův–Torricelliho bod, neboť Evangelista Torricelli[63] v roce 1646 nalezl tři základní řešení tohoto problému, která zveřejnil jeho žák Vincenzo Viviani v první „optimalizační“ knize *De maximis et minimis geometrica divinatio in quintum conicorum Apollonii Pergaei adhuc desideratum*[64] z roku 1659 společně s jedním svým vlastním, viz Obrázek 1.1.

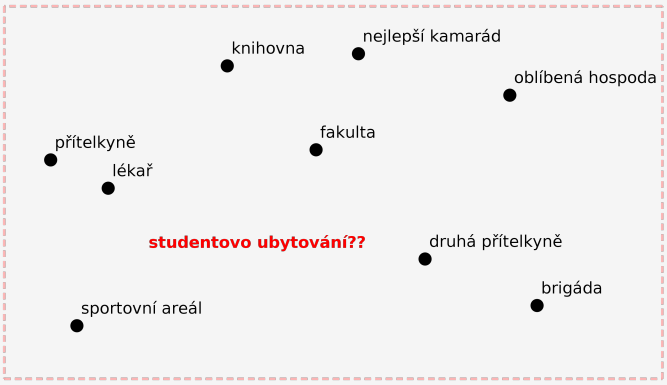


Obrázek 1.1: Dva možné Fermatovy–Torricelliho body v závislosti na velikosti vnitřních úhlů daného trojúhelníku.

Zobecněním této úlohy se dostaneme k jednomu z nejznámějších problémů z teorie lokalizace. Jedná se o tzv. *Weberův*[65,66,67] *problém*, ve kterém jde o nalezení bodu v rovině, který minimalizuje přepravní náklady z tohoto bodu do n dalších lokalit. Toto si můžeme ilustrovat na jednoduché úloze, ve které bychom chtěli najít optimální místo k bydlení pro běžného (pochopitelně matematiky znalého) studenta s obvyklými zájmy a povinnostmi tak, aby součet vzdáleností z tohoto místa do školy, k přítelkyni, na brigádu atd. byl minimální, viz Obrázek 1.2. Označíme-li souřadnice jednotlivých lokalit jako $y_1, \ldots, y_n \in \mathbb{R}^2$, pak jde o řešení úlohy nepodmíněné minimalizace

$$\sum_{i=1}^{n} \| x - y_i \| \to \min, \quad x \in \mathbb{R}^2,$$

kde $\|\cdot\|$ značí standardní eukl(e)idovskou normu v $\mathbb{R}^2$, tj. $\| x \| := \sqrt{\langle x, x \rangle} = \sqrt{\sum x_i^2}$.



Obrázek 1.2: „Jednoduchý“ Weberův problém běžného studenta.

matematiku, filosofii a vlastně celou vědu 17. století tkví v jeho „dopisovatelské“ činnosti. Stal se středobodem rozsáhlé korespondenční sítě pro výměnu vědeckých informací v 17. století, kdy ještě neexistovaly vědecké časopisy, jak je známe dnes, a hlavní formou šíření informací byla vzájemná korespondence mezi vědci. Dopisoval si s mnoha učenci a vzdělanci své doby, např. Descartem, Pascalem, Fermatem, Torricellim, Hobbesem, Gassendim, Leibnizem, Huygensem, Komenským aj., kopíroval jejich dopisy a rozesílal je dále, čímž vlastně koordinoval vědecké aktivity a předávání nových poznatků v celé Evropě. Také přeložil a připravil k vydání díla Eukl(e)ida, Apollonia i Archiméda.

[53]

René Descartes (31. března 1596 – 11. února 1650) byl francouzský filosof, matematik a fyzik. Významně se podílel na vytvoření číselné reprezentace geometrických objektů známé jako kartézská soustava souřadnic, na vzniku analytické geometrie a na matematizaci optiky. Základním Descartovým filosofickým východiskem je metodická skepse, soustavná pochybnost zejména o datech smyslového poznání. Tímto odvážným myšlenkovým krokem dokázal podstatně změnit evropské myšlení. Avšak Descartes odmítá běžnou skepsi (tj. „pochybování pro pochybování“) a pochybnost užívá jen proto, aby se dobral pravdy. Jeho nejznámějším filosofickým výrokem je *cogito, ergo sum* (*myslím, tedy jsem*) objevující se v jeho díle *Discours de la méthode pour bien conduire sa raison, et chercher la verité dans les sciences* (*Rozprava o metodě, jak správně vést svůj rozum a hledat pravdu ve vědách*) z roku 1637.

[54] I would be happy to know what he will say, both about the letter attached to this one, where I respond to his paper on maxima and minima, and about the one preceding, where I replied to his demonstration against my Dioptrics. For I have written the one and the other for him to see, if you please;

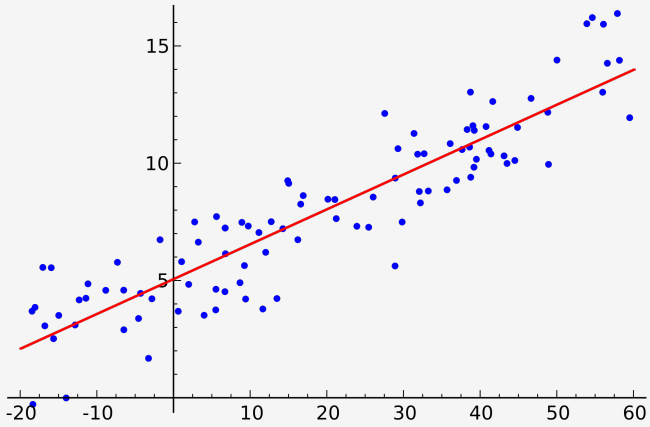
V tomto pojetí jsou Weberem zmíněné přepravní náklady shodné se vzdálenostmi mezi jednotlivými lokalitami. Pochopitelně je možné, že dopravní spojení do některých lokalit je náročnější nebo že některé lokality jsou významnější než jiné (všichni jistě víme, že dobře dostupná knihovna je pro studenta důležitější než nějaké povyražení). V takových případech bychom mohli pro jednotlivé lokality zavést váhy $w_1,\dots,w_n$ a řešit minimalizační úlohy s účelovou funkcí

$$\sum_{i=1}^{n} w_i \|x-y_i\| \quad \text{nebo} \quad \sum_{i=1}^{n} \|x-w_i y_i\|.$$

Námi uvedený příklad se studentem hledajícím ubytování sice nemá s realitou příliš společného, ale jistě si všichni dokážeme představit mnohem reálnější situaci, ve které znalost takového řešení může být velmi užitečná.

Příklad 1.2.10

Druhým velmi známým příkladem vedoucím na problém nepodmíněné minimalizace je tzv. *metoda nejmenších čtverců*. V její dvourozměrné variantě chceme najít přímku, která nejlépe (tj. v normě) aproximuje daná data, viz Obrázek 1.3.



Obrázek 1.3: Naměřená data a hledaná regresní přímka.

Uvažme lineární regresní model

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k,$$

kde $\beta_0,\dots,\beta_k$ jsou *regresní koeficienty*. Označme y_i jako výsledek i-tého pozorování závislé proměnné při hodnotách x_{ij} nezávislých proměnných, tj.

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + e_i, \quad i = 1,\dots,n,$$

kde e_i je *chyba*, která popisuje odchylku pozorování od skutečné regrese (tj. hledané podmíněné střední hodnoty). K tomu, abychom docílili nejlepšího možného modelu, je potřeba odhadnout $\beta_0,\dots,\beta_k$ tak, aby byla chyba nejmenší možná. Toho bychom mohli docílit pouze pomocí absolutní hodnoty (chceme stejně „vážit“ kladné i záporné odchylky, proto s absolutní hodnotou – bez ní by se nám odchylky s opačnými znaménky anulovaly), tj. řešením úlohy

$$\sum_{i=1}^{n} |e_i| \to \min.$$

Jenže naše nástroje jsou založené na derivování, což ale v případě absolutní hodnoty narazí na její možnou nediferencovatelnost. Proto je určení regresních koeficientů za-

loženo na minimalizaci sumy druhých mocnin jednotlivých chyb, tj.

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \ldots - \beta_k x_{ik})^2 \to \min, \quad \beta_0, \ldots, \beta_k \in \mathbb{R}.$$

S využitím vektorů $\beta \in \mathbb{R}^{k+1}$, $Y \in \mathbb{R}^n$ a matice $X \in \mathbb{R}^{n \times (k+1)}$ daných jako

$$\beta := \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X := \begin{pmatrix} 1 & x_{11} & \ldots & x_{1k} \\ 1 & x_{21} & \ldots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \ldots & x_{nk} \end{pmatrix}$$

lze tuto úlohu zapsat ve formě minimalizace euklidovské normy

$$\| Y - X\beta \|^2 \to \min, \quad \beta \in \mathbb{R}^{k+1}.$$

Je-li $k = 1$, tj. máme-li při měření pouze jednu nezávislou proměnnou, pak výsledkem budou koeficienty určující *regresní přímku* $y = \beta_1 x + \beta_0$. Lze se také setkat se situacemi, kdy mají být naměřená data aproximována jinými funkcemi, např. polynomem $p(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3 x^3$, což vede na úlohu

$$\sum_{i=1}^{n} \left[y_i - p(x_i) \right]^2 \to \min, \quad \gamma_0, \gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}.$$

Nalezení regresní přímky můžeme využít např. pro konstrukci jednoduché lineární předpovědi ceny akcií společnosti *Halíř a spol.* Označíme-li jako $c(t)$ cenu akcií v i-tém dni a chceme-li odhadnout cenu na základě předchozích čtyřdenních hodnot, dostaneme funkci

$$c(t) = \alpha_1 c(t-1) + \alpha_2 c(t-2) + \alpha_3 c(t-3) + \alpha_4 c(t-4),$$

pro kterou musíme určit koeficienty $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$. Máme-li k dispozici ceny za posledních n dní (např. 4 měsíce, tj. cca 120 dní), pak tyto koeficienty nalezneme řešením úlohy

$$\| Y - C\alpha \|^2 \to \min, \quad \alpha \in \mathbb{R}^4,$$

kde

$$Y = \begin{pmatrix} c(5) \\ c(6) \\ \vdots \\ c(n) \end{pmatrix}, \quad C = \begin{pmatrix} c(4) & c(3) & c(2) & c(1) \\ c(5) & c(4) & c(3) & c(2) \\ \vdots & \vdots & \vdots & \vdots \\ c(n-1) & c(n-2) & c(n-3) & c(n-4) \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}.$$

A co globální extrémy? *Viva la Weierstrass*[68]!

Věta 1.2.11 Weiestrass

Nechť množina $X \subseteq \mathbb{R}^n$ je kompaktní[69] a funkce $f : \mathbb{R}^n \to \mathbb{R}$ je spojitá na $X \subseteq D(f)$. Pak f nabývá své největší a nejmenší hodnoty na X, tj. existuje globální minimum i maximum f na X.

Důkaz. Ze spojitosti funkce f a kompaktnosti množiny X vyplývá kompaktnost obrazu X skrze f, tj. množiny $f(X)$, takže existuje

$$N := \sup_{x \in X} f(x)$$

a zbývá dokázat existenci $x^* \in X$ takového, že $f(x^*) = N$. V souladu s definicí suprema víme, že pro libovolné $k \in \mathbb{N}$ existuje bod $x_k \in X$ takový, že $f(x_k) > N - 1/k$. Posloupnost takovýchto

I did not even want to name him, so that he will feel less shame at the errors that I have found there and because my intention is not to insult anyone but merely to defend myself. And, because I feel that he will not have failed to vaunt himself to my prejudice in many of his writings, I think it is appropriate that many people also see my defense. That is why I ask you not to send them to him without retaining copies of them. And if, even after this, he speaks of wanting to send you still more papers, I beg of you to ask him to think them out more carefully than those preceding; otherwise I ask you not to accept the commission of forwarding them to me. For, between you and me, if when he wants to do me the honor of proposing objections, he does not want to take more trouble than he did the first time, I should be ashamed if it were necessary for me to take the trouble to reply to such a small thing, though I could not honestly avoid it if he knew that you had sent them to me. Viz [361, str. 177].

[55] I would prefer to say nothing about the paper you have sent me, because I could not say anything that would be to the advantage of the person who wrote it. But, because I recognize that it is same person who earlier tried to refute my Dioptrics and because you inform me that he sent the paper after having read my Geometry and having been surprised that I had not found the same thing – that is, as I interpret things, with the intention of entering into competition and of showing that he knows more about the subject than I – and then also because I learn by your letters that he has the reputation of being very learned in mathematics, I feel obligated to reply to him. Viz [361, str. 178].

[56] . . . seeing the last method that you use for finding tangents to curved lines, I can reply to it in no other way than to say that it is very good and that, if you had explained it in this manner at the outset, I would not have contradicted at all. Viz [361, str. 192].

[57] *Avšak pokud někdo navrhuje rozdělit úsečku B tak, že součin jejích části bude roven Z (o němž předpokládáme, že je menší než $B^2/4$), pak budeme mít dvě řešení daného problému, z nichž každé bude umístěno na jiné straně od bodu odpovídající-*

bodů $\{x_k\}_{k=1}^{\infty}$ je ohraničená, tudíž v ní existuje podposloupnost $\{x_{k_\ell}\}_{\ell=1}^{\infty}$ konvergující k nějakému x^*, tj. $x_{k_\ell} \to x^*$ pro $\ell \to \infty$. Navíc díky uzavřenosti X máme $x^* \in X$ a díky spojitosti f dokonce $\{f(x_{k_\ell})\}_{\ell=1}^{\infty} \to f(x^*)$. Protože $f(x_{k_\ell}) > N - 1/k_\ell$ pro všechna $\ell \in \mathbb{N}$, dostaneme limitním přechodem pro $\ell \to \infty$, že $f(x^*) \geq N$. Avšak vzhledem k tomu, že $f(x) \leq N$ pro všechna $x \in X$, musí být $f(x^*) = N$, a tedy x^* je skutečně oním hledaným bodem. Důkaz existence globální minima se provede analogicky s využitím infima hodnot $f(x)$ pro $x \in X$. ■

Nás ovšem budou zajímat především lokální/globální minima, a pro ta globální platí následující důsledek Weierstrassovy věty.

DŮSLEDEK 1.2.12

Nechť $X \subseteq \mathbb{R}^n$ je uzavřená množina, funkce $f : \mathbb{R}^n \to \mathbb{R}$ spojitá na $X \subseteq D(f)$ a nechť existuje $x^* \in X$ takové, že („dolní obrysová“) množina

$$N(x^*) := \{x \in X \mid f(x) \leq f(x^*)\}$$

je neprázdná a ohraničená. Pak funkce f nabývá své nejmenší hodnoty na X, tj. existuje globální minimum f na X.

Důkaz. Množina $N(x^*)$ je zjevně uzavřená, takže díky požadavku její ohraničenosti dostaneme z Weierstrassovy věty existenci globálního minima f na $N(x^*)$. To je ale nutně shodné s globálním minimem f na X, neboť případným zvětšováním hodnoty $f(x^*)$ budeme do množiny $N(x^*)$ přidávat pouze body s větší funkční hodnotou, které minimum f na $N(x^*)$ nikterak neovlivní. Ani zmenšováním $f(x^*)$ nedosáhneme žádné změny nejmenší hodnoty, pokud zůstane $N(x^*)$ neprázdné. ■

Zkombinováním Fermatovy a Weierstrassovy věty pak dokážeme extrémy „lokalizovat“.

DŮSLEDEK 1.2.13

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní množina a funkce $f : \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná na $X \subseteq D(f)$. Pak f nabývá svého maxima a minima na X buď ve stacionárním bodě ležícím uvnitř X nebo v některém hraničním bodě množiny X.

Všimněme si, že předpoklady předchozího tvrzení zaručují existenci globálního minima i maxima. Při jejich oslabení již některý z extrémů existovat nemusí a zároveň mohou přibýt/ubýt případní „kandidáti“, v nich se extrém může realizovat. Např. je-li množina X otevřená, pak možnými kandidáty jsou pouze stacionární body. Naopak pokud funkce f není na celém X diferencovatelná, pak je potřeba problematické body přidat mezi kandidáty na extrém. Jak moc je ale toto tvrzení praktické? To si můžeme vyzkoušet na jednoduchém příkladě. Další řešené úlohy lze najít v [255, Kapitola III.5].

Příklad 1.2.14

Určeme největší a nejmenší hodnotu funkce $f(x,y) = x^2 + y^2 + xy - 3x + 1$ na množině X, kterou je trojúhelník s vrcholy $[0,0]$, $[9,0]$, $[0,9]$.

Řešení. Množina X je zjevně kompaktní a funkce f je na této množině jistě spojitá, takže Věta 1.2.12 zaručuje existenci obou hledaných hodnot. V souladu s Důsledkem 1.2.13 se nejdříve podíváme na možné lokální extrémy uvnitř trojúhelníku. Parciální derivace prvního řádu funkce f jsou

$$f_x(x,y) = 2x + y - 3 \quad \text{a} \quad f_y(x,y) = 2y + x,$$

takže řešením soustavy $f_x(x,y) = 0$ a $f_y(x,y) = 0$ nalezneme jediný stacionární bod $[2,-1]$ s funkční hodnotou $f(2,-1) = -2$. Avšak tento bod leží mimo stanovenou přípustnou množinu, takže v něm jistě hledaný globální extrém nenastane. Nicméně v této souvislost ještě dodejme, že i kdyby tento bod ležel uvnitř přípustné množiny, nebylo by nutné zjišťovat, zda se skutečně jedná o lokální extrém a jakého druhu. Hle-

ho maximálnímu součinu.
Vskutku, nechť A je jedna z částí úsečky B. Pak budeme mít rovnici $BA - A^2 = Z$ s nejednoznačným řešením, neboť A může být jakýkoli ze dvou kořenů. Mějme tedy druhou rovnici $BE - E^2 = Z$. Pak následujícím Viètovu metodu dostaneme porovnáním obou hodnot $BA - BE = A^2 - E^2$. Poděleným obou stran výrazem $A - E$ získáme $B = A + E$, přičemž obě délky A a E budou různé.
Vezmeme-li součin větší než Z a současně menší než $B^2/4$, bude rozdíl mezi délkami A a E menší než předtím a body určující rozdělení úsečky budou blíže k bodu dávajícímu maximální součin. Čím více vzroste součin délek obou částí, tím naopak více se zmenší rozdíl mezi A a E, dokud nezmizí úplně při rozdělení odpovídajícímu maximálnímu součinu. V takovém případě pak máme pouze jedno jediné a singulární řešení: obě veličiny A a E si jsou rovny.
Nynější aplikování Viètovy metody na dvě přidružená řešení výše nás přivedlo k rovnosti $B = A + E$, a tudíž je-li $A = E$ (což se stane vždy v bodě, který udává maximum nebo minimum), dostáváme v daném případě $B = 2A$. To znamená, že vezmeme-li střed úsečky B, pak bude součin délek obou částí maximální.
(But if one proposed dividing the same line B such that the product of its segments be equal to Z (this area being assumed to be less than $B^2/4$), we will have two points satisfying the question, and they will be situated on either side of the point corresponding to the maximum product.
Indeed, let A be one of the segments of the line b. We will then have $BA - A^2 = Z$, an ambiguous equation, since for straight line A the value can be either of the two roots. Let the correlative equation therefore be $BE - E^2 = Z$. Let us compare these two methods following the method of Viète: $BA - BE = A^2 - E^2$. Dividing both sides by $A - E$, there results $B = A + E$; further, the lengths A and E will be unequal.
If we take an area larger than Z, while still less than $B^2/4$, the difference between lines A and E will be less than earlier, the points of division approaching closer to the point constituting the maximum product. The more the product of the segments increases, the more, on the contrary, the difference between A and E will shrink, until it vanishes completely at the division corresponding to the maxi-

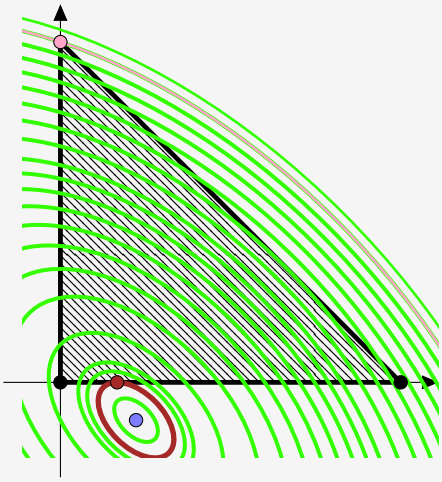
mum product. In this case, there is only one unique and singular solution: the two quantities A and E become equal.
Now the method of Viète, applied to the two correlative equations above, has led us to the equality $B = A+E$; therefore, if $E = A$ (which will constantly happen at the point constituting the maximum or minimum), we will have, in the proposed case, $B = 2A$. That is to say that if we take the center of the line B, the product of the two segments will be maximum.)

[58] Fermatova výzva pochybovačům také úzce souvisí s jeho jedinečným osobním stylem (nezapomínejme, že byl matematik–amatér a původní profesí právník). Měl pověst tajnůstkáře – někoho, kdo nechává výsledky obestřené tajemstvím a podává je matematické komunitě bez jakéhokoli náznaku toho, jak by na to někdo mohl přijít (rozumějme „bez důkazu“). Je to téměř jako hádanka, kterou by svět měl vyřešit! Historik matematiky Victor Katz o Fermatovi píše: „V mnoha případech není známo, zda vlastně Fermat vůbec nějaký důkaz skutečně podal, a také ne vždy nalezneme v jeho pracích vše systematicky vysvětlené. Fermat často mučil své dopisovatele s náznaky svých nových metod k řešení určitých problémů. Někdy poskytl hrubý náznak těchto metod, ale jeho sliby, že doplní chybějící výklad „když čas dovolí“ často zůstaly nenaplněny.“ Viz [318, str. 474].

[59] *Najde-li se stále někdo, kdo považuje úspěch této metody pouze za šťastnou náhodu, klidně může zkusit přijít se svojí podobnou metodou. Pokud jde o ty, kteří ji neuznávají, předkládám jim takovýto problém: Pro danou trojici bodů nalézt čtvrtý tak, že součet jednotlivých vzdáleností tohoto bodu od daných bodů je minimální.*
(If there is still someone who considers the success of this method due to serendipitous chance, he is welcome to try to come across a similar one.
As for those who do not approve of it, I pose to them this problem: Given three points, find a fourth point such that the sum of its distances to the three given points be a minimum.)

[60] Lodovico de Ferrari (2. února 1522 – 5. října 1565) byl italský matematik a Cardanův žák. Je autorem řešení kvartické rovnice, které bylo publikováno v Cardanově knize Ars Magna.

dáme totiž globální extrémy, a k tomu nám postačí pouze z nalezených kandidátů vybrat ty s největší a nejmenší funkční hodnotou.



Obrázek 1.5: Vrstevnice funkce f (zelená barva) se zvýrazněním vrstevnice na nejnižší (červená) a nejvyšší (fialová) úrovni mající neprázdný průnik s množinou X. Stacionární bod $[2,-1]$ je vyznačen modrou barvou, globální minimum $[3/2,0]$ červenou a globální maximum $[0,9]$ fialovou.

Nyní se podíváme na hranici množiny X. Ta je tvořena třemi úsečkami, které můžeme vyjádřit jako

$$\text{(i)}\ \ y=0\ \ \&\ \ x\in[0,9];\quad \text{(ii)}\ \ x=0\ \ \&\ \ y\in[0,9];\quad \text{(iii)}\ \ y=9-x\ \ \&\ \ x\in[0,9].$$

V prvním případě se funkce f redukuje na $g(x) := f(x,0) = x^2-3x+1$ a chceme určit její nejmenší a největší hodnotu pro $x\in[0,9]$, tj. máme vlastně stejný problém jako na začátku, ale tentokrát již jen pouze pro funkci jedné proměnné. Proto začneme se stacionárními body, tj. řešením rovnice $g'(x) = 2x-3 = 0$, což nás přivede k bodu $[3/2,0]$ s hodnotou $f(3/2,0) = -5/4$. V tomto případě je hranice přípustné množiny tvořena pouze krajní body intervalu $[0,9]$, tj. $x=0$ a $x=9$ s funkčními hodnotami $f(0,0)=1$ a $f(9,0)=55$. Podobně ve druhém případě hledáme nejmenší a největší hodnotu funkce $h(y) = f(0,y) = y^2+1$ pro $y\in[0,9]$. Stacionární bod splňuje $h'(y) = 2y = 0$, což odpovídá hodnotě $y=0$, která neleží uvnitř stanoveného intervalu. Jedná se ale o krajní bod úsečky, takže v něm beztak musíme určit funkční hodnotu, tj. opět $f(0,0)=1$. Ve druhém krajním bodě této úsečky máme $f(0,9)=82$. Zbývá poslední úsečka, na které dostaneme $\ell(x) := f(x,9-x) = x^2-12x+82$ pro $x\in[0,9]$ se stacionárním bodem splňujícím $\ell'(x) = 2x-12 = 0$, tj. $x=6$ a $y=3$ s funkční hodnotou $f(6,3)=46$. Krajní body odpovídají vrcholům $[9,0]$ a $[0,9]$, v nichž jsme již funkční hodnoty počítali.
Teď už jen stačí z vypočtených funkčních hodnot vybrat tu největší a tu nejmenší. Proto globální minimum nastává v bodě $[3/2,0]$ s $f_{\min} = -5/4$ a globální maximum v bodě $[0,9]$ s hodnotou $f_{\max} = 82$, viz také Obrázek 1.5.

Jenže takový postup může být poměrně zdlouhavý a komplikovaný, neboť je potřeba postupně analyzovat chování účelové funkce f na jednotlivých částech hranice přípustné množiny X. Nebylo by možné použít nějakou efektivnější metodu? Samozřejmě ano, a vlastně o tom jsou všechny následující kapitoly. Obecný princip pro libovolnou účelovou funkci na množině vymezené systémem algebraických rovnic a/nebo nerovnic si ukážeme v Kapitole 8. Nyní si ale připomeneme tvrzení pro případ, kdy je přípustná množina vymezena pouze rovnostmi, tj. budeme hledat řešení úlohy

$$f(x)\to\min,\quad x\in X := \textstyle\bigcap_{i=1}^{m}\{x\in\mathbb{R}^n \mid g_i(x)=0\}, \tag{1.2.5}$$

přičemž občas bude užitečné množinu X vyjádřit spíše jako $X = \{x \in \mathbb{R}^n \mid G(x) = 0\}$ s pomocí vektorové funkce $G : \mathbb{R}^n \to \mathbb{R}^m$ se složkami

$$G(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}.$$

Připomeňme, že množina

$$\{x \in \mathbb{R}^n \mid g_1(x) = 0\}$$

obvykle představuje nějakou *nadplochu*, tj. „útvar“ dimenze $n-1$ v $\mathbb{R}^n$. Přidáním další nadplochy dostaneme

$$\{x \in \mathbb{R}^n \mid g_1(x) = 0\} \bigcap \{x \in \mathbb{R}^n \mid g_2(x) = 0\},$$

což je v ne příliš patologických případech množina dimenze $n-2$ atd., viz např. Obrázek 1.6. Tedy X bývá velmi často $(n-m)$-dimenzionální množina v $\mathbb{R}^n$. V takovémto případě bude pro $m = n$ množina X jednobodová a pro $m > n$ bude dokonce prázdná, tj. úloha (1.2.5) bude nepřípustná (to se ale pochopitelně může stát i pokud $m \leq n$).

Poznámka 1.2.15 (opět historická, ne tak dlouhá, ale zase důležitá). Počátky následující (nám již dobře známé) metody pro řešení úlohy (1.2.5) jsou spojeny především s Josephem Lagrangem[70], a proto na jeho počest hovoříme o *Lagrangeově funkci*

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x) = f(x) + \lambda^\top G(x),$$

obsahující m-tici *Lagrangeových multiplikátorů* $\lambda = (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}^m$. Lagrange tuto metodu vyvinul nikoli v souvislosti s hledáním vázaných extrémů, ale při řešení úlohy týkající se rovnovážného stavu systému částic – konkrétně šlo o řešení soutěžní úlohy zveřejněné v roce 1762 Francouzskou akademií věd[71] týkající se librace[72] Měsíce. Jeho zpráva s řešením z roku 1764 byla zveřejněna v souhrnném vydání [341][73] a o patnáct let později Lagrange tuto metodu v podrobnější a obecnější podobě publikoval v [342]. Pochopitelně na ni nezapomněl ani ve svém nejslavnějším díle [343] z roku 1788. Na začátku této monografie se Lagrange věnuje statice a vyjmenovává tři její základní principy: (i) princip páky, (ii) princip skládání sil a (iii) princip virtuálních rychlostí. První dva principy jsou pro nás v tuto chvíli nedůležité, což už neplatí o tom třetím.

> *On doit onlendrc par vitesse virtuelle, celle qu'un corps en équilibre est disposé à recevoir, en cas que l'équilibre vienne à être rompu, c'est-à-dire, la vitesse que ce corps prendrait réellement dans le premier instant de son mouvement; et le principe dont il s'agit consiste en ce que des puissances sont en équilibre quand elles sont en raison inverse de leurs vîtesses virtuelles, estimées suivant les directions de ces puissances.*[74]

Tento princip lze v případě působení dvou sil reprezentovat tím, že těleso zůstane v rovnovážné poloze, pokud

$$P \frac{\mathrm{d}p}{\mathrm{d}t} + Q \frac{\mathrm{d}q}{\mathrm{d}t} = 0,$$

kde P, Q jsou ony dvě síly, p, q značí směry jejich působení a veličiny $\mathrm{d}p/\mathrm{d}t$, $\mathrm{d}q/\mathrm{d}t$ představují virtuální rychlosti způsobené působením sil P a Q. Veličiny $P\frac{\mathrm{d}p}{\mathrm{d}t}$ a $Q\frac{\mathrm{d}q}{\mathrm{d}t}$ tak vyjadřují momenty jednotlivých sil. Následně Lagrange místo podílu $\mathrm{d}p/\mathrm{d}t$ píše pouze $\mathrm{d}p$ apod., tj. obecná rovnice pro rovnovážnou polohu je

$$P\,\mathrm{d}p + Q\,\mathrm{d}q = 0.$$

Na začátku Odstavce 3 ve druhé sekci [343] píše následující.

[61] Je zcela zásadní dodržovat rozlišování mezi vedoucími hlavními minory a hlavními minory, jak je popsáno ve Větě 1.2.6, tj. pomocí vedoucích hlavních minorů zjistíme pozitivní/negativní definitnost, zatímco pro semidefinitnost musíme využít hlavní minory. Toto můžeme jednoduše ilustrovat na příkladu matice $\left(\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, která má nezáporné (nulové) oba vedoucí hlavní minory, ale rozhodně není pozitivně semidefinitní. Ve skutečnosti je negativně semidefinitní, což je v souladu s tím, že její hlavní minory jsou nekladné.

[62] Připomeňme, že pro diferencovatelnou funkci $f : \mathbb{R}^n \to \mathbb{R}$ na množině D a libovolné $x, x^* \in D$ platí

$$f(x) = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle + \\ + \omega(x - x^*)$$

pro nějakou funkci $\omega : \mathbb{R}^n \to \mathbb{R}$ splňující $\omega \in o(\|x - x^*\|)$, tj.

$$\lim_{x \to x^*} \frac{\omega(x - x^*)}{\|x - x^*\|} = 0.$$

Podobně pro dvakrát spojitě diferencovatelnou funkci $f : \mathbb{R}^n \to \mathbb{R}$ na množině D a libovolné $x, x^* \in D$ platí *Taylorův rozvoj druhého řádu*

$$f(x) = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle + \\ + \frac{1}{2} (x - x^*)^\top \nabla^2 f(x^*) (x - x^*) + \\ + \omega(x - x^*)$$

pro nějakou funkci $\omega : \mathbb{R}^n \to \mathbb{R}$ splňující $\omega \in o(\|x - x^*\|^2)$, tj.

$$\lim_{x \to x^*} \frac{\omega(x - x^*)}{\|x - x^*\|^2} = 0.$$

Obecněji pro $f \in \mathrm{C}^{m+1}$ platí

$$f(x^* + h) = f(x^*) + \mathrm{d}f(x^*)(h) + \cdots + \\ + \frac{1}{m!} \mathrm{d}^m f(x^*)(h) + R_m(x^*, h),$$

kde

$$\mathrm{d}^k f(x^*)(h) = \sum_{j_1 + \cdots + j_n = k} \frac{k!}{j_1! \times \cdots \times j_n!} \\ \frac{\partial^k f(x^*)}{\partial x_1^{j_1} \ldots \partial x_n^{j_n}} h_1^{j_1} \ldots h_n^{j_n}$$

a

$$R_m(x^*, h) = \frac{\mathrm{d}^{m+1}}{(m+1)!} f(x^* + \lambda h)(h)$$

pro $h = (h_1, \ldots, h_n)^\top \in \mathbb{R}^n$ a nějaké $\lambda \in (0, 1)$, přičemž platí

$$\lim_{h \to 0} \frac{\|R_m(x^*, h)\|}{\|h\|^m} = 0.$$

> *L'objet de cet Ouvrage étant de réduire la Mécanique à des opérations purement analytiques, la formule que nous venons de trouver est très-propre à le remplir. Il ne s'agit que d'exprimer analytiquement, et de la manière la plus générale, les valeurs des lignes p, q, r, etc., prises dans les directions des forces P, Q, R, elc, et l'on aura, par lu simple différentiation, les valeurs des vîtesses virtuelles $\mathrm{d}p, \mathrm{d}q, \mathrm{d}r$, etc.*[75]

Jinými slovy, jde o oproštění se od geometrických úvah, jakkoli mohou být intuitivní a přirozené. Proto na začátku čtvrté sekce v [343] přichází s *metodou multiplikátorů* (*Méthode des Multiplicateurs*), díky níž je schopen řešit problém ohledně rovnovážné polohy při posouvání či rotaci uvažovaný ve třetí sekci jednodušším a obecnějším způsobem. Ve čtvrté sekci se totiž věnuje rovnovážné poloze v situaci, kdy na jednotlivé body působí vnější omezení. Jsou-li tato omezení dána rovnicemi $L = 0$, $M = 0$, … a jedná-li se o funkce několika proměnných, pak musí být $\mathrm{d}L = 0$, $\mathrm{d}M = 0$ atd. Proto jsou-li $\lambda, \mu, \ldots$ libovolná reálná čísla (Lagrange je nazývá *neurčitými koeficienty* čili *coefficiens indéterminés*), dostaneme také $\lambda\,\mathrm{d}L = 0$, $\mu\,\mathrm{d}M = 0$ atd., což Lagrange následně vysvětluje takto:

> *De là résulte donc cette règle extrêmement simple pour trouver les conditions de l'équilibre d'un système quelconque proposé.*
> *On prendra la somme des momens de toutes les puissances qui doivent être en équilibre, et on y ajoutera les différentes fonctions différentielles qui doivent être nulles par les conditions du problème, après avoir multiplié chacune de ces fonctions par un coefficient indéterminé; on égalera le tout à zéro, et l'on aura ainsi une équation différentielle qu'on traitera comme une équation ordinaire de maximis et minimis, et d'où l'on tirera autant d'équations particulières finies qu'il y aura de variables; ces équations étnt ensuite débarrassées, par l'élimination, des coefficiens indéterminés, donneront toutes les conditions nécessaires pour l'équilibre.*
> *L'éqaution différentielle dont il s'agit sera donc de cette forme,*
>
> $$P\,\mathrm{d}p + Q\,\mathrm{d}q + \cdots + \lambda\mathrm{d}L + \mu\,\mathrm{d}M + \cdots = 0,$$
>
> *dans laquelle λ, μ, etc. sont des quantités indéterminées; nous la nommerons dans la suite, équation générale de l'équilibre*[76].

viz Odstavec 3 ve čtvrté sekci [343]. Pro každou jednotlivou proměnnou vztažné ortogonální soustavy daného systému dostaneme jednu rovnici, např.

$$P\frac{\mathrm{d}p}{\mathrm{d}x} + Q\frac{\mathrm{d}q}{\mathrm{d}x} + \cdots + \lambda\frac{\mathrm{d}L}{\mathrm{d}x} + \mu\frac{\mathrm{d}M}{\mathrm{d}x} + \cdots = 0.$$

Lagrange v navrženém postupu zmiňuje potřebu odstranění neurčitých koeficientů, což vlastně může být jediná skutečná komplikace při hledání řešení odvozené soustavy rovnic, jak sám uvádí v následujícím odstavci. Avšak ihned tyto pochybnosti rozhání, neboť stačí přidat jednotlivá omezení, čímž dostaneme soustavu tolika rovnic, kolik je dohromady neznámých a neurčitých koeficientů, tj. např. pro dvě proměnné a jedno omezení získáme soustavu

$$P\frac{\mathrm{d}p}{\mathrm{d}x} + Q\frac{\mathrm{d}q}{\mathrm{d}x} + \lambda\frac{\mathrm{d}L}{\mathrm{d}x} = 0 \quad \& \quad P\frac{\mathrm{d}p}{\mathrm{d}y} + Q\frac{\mathrm{d}q}{\mathrm{d}y} + \lambda\frac{\mathrm{d}L}{\mathrm{d}y} = 0 \quad \& \quad L = 0.$$

Pro více podrobností viz např. [86, 186, 308]. ▲

Lagrange se matematickým metodám hledání maxima a minima věnoval i v článcích 131–184 v pojednání [344] a právě v článku 167 nalezneme na dvou stranách krátký rozbor úlohy na vázaný extrém se dvěma proměnnými a jedním omezením. Po úspěšné rozvoji metody multiplikátorů v problematice rovnovážné polohy se Lagrange pustil do jejího využití i v jiné oblasti –

[63]

Evangelista Torricelli (15. října 1608 – 25. října 1647) byl italský fyzik a matematik. Byl studentem Galilea a je znám především jako vynálezce barometru. Jeho jméno také nese fyzikální zákon objasňující rychlost proudění kapaliny z otvoru na boku nádoby. Také se podílel na rozvoji optiky a matematické „metody nedělitelnosti", která je známa spíše jako *Cavalieriho princip*: v trojrozměrné případě říká, že tělesa se stejně velkými podstavami a výškami mají stejný objem, pokud mají řezy rovnoběžné s podstavami a vedené ve stejné vzdálenosti od podstav stejné obsahy, viz Obrázek 1.4. Ve dvourozměrném případě tento princip tvrdí rovnost obsahu dvou rovinných obrazců, pokud úsečky rovnoběžné s osou souřadné soustavy, které je protínají ve shodné výšce, mají vždy stejnou délku. Jedná se vlastně o intuitivní pojetí základní myšlenky integrálního počtu a z dnešního pohledu to je důsledek Fubiniho věty.

[64] V českém překladu (asi) *Geometrické věštění toho největšího a nejmenšího v páté kuželosečce Apollónia z Pergy stále chybí.* Originál byl (v roce 2024) dostupný za cca 10000 dolarů.

Obrázek 1.4: Ilustrace Cavalieriho principu na mincích.

[65] Alfred Weber (30. července 1868 – 2. května 1958) byl německý národohospodář, geografa, sociolog a teoretik kultury, jehož práce zásadně ovlivnila moderní ekonomickou geografii. V roce 1954 ho němečtí komunisté nominovali bez jeho vědomí jako kandidáta na spolkového prezidenta.

ve variačním počtu, na jehož rozvoji má zásadní podíl a jemuž se budeme věnovat v Kapitole 9. Nyní ale již je nejvyšší čas na moderní formulaci Lagrangeova základního výsledku týkajícího se řešení úlohy (1.2.5). Ještě doplňme, že toto a následující tvrzení pro jednoduchost vyžadují funkce třídy C^1 či C^2 na nějaké otevřené nadmnožině $\mathscr{U} \supseteq X$ – avšak tentýž závěr obdržíme i v případě, kdy toto bude splněno jen na nějaké otevřené množině obsahující bod x^*.

Věta 1.2.16 Lagrange

Nechť je dána úloha (1.2.5) s $m \in \{1,\dots,n\}$ a funkcemi $f, g_1, \dots, g_m$ třídy C^1 na otevřené množině $\mathscr{U} \subseteq \mathbb{R}^n$ splňující $X \subseteq \mathscr{U}$. Nechť jsou dále vektory $\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_m(x)$ lineárně nezávislé[77] pro všechna $x \in X$, tj. Jacobiho matice

$$\mathrm{D}G(x) = \mathrm{D}\big(g_1(x),\dots,g_m(x)\big)^\top = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(x) & \dots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{m\times n}$$

má plnou hodnost, tj. $\operatorname{rank} \mathrm{D}G(x) = m$. Je-li bod $x^* \in X$ lokálním extrémem funkce f na množině X, tj. lokálním řešením úlohy (1.2.5), pak existuje jediná m-tice multiplikátorů $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ taková, že

$$\frac{\partial f}{\partial x_j}(x^*) + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*) = 0 \quad \text{pro všechna } j = 1,\dots,n \tag{1.2.6}$$

což lze vektorově zapsat jako

$$\mathrm{D}_x L(x^*,\lambda^*) = \mathrm{D}f(x^*) + \lambda^{*\top} \mathrm{D}G(x^*) = 0, \tag{1.2.7}$$

kde $\mathrm{D}f(x^*) = \operatorname{grad}^\top f(x^*)$ a $\lambda^* := \big(\lambda_1^*,\dots,\lambda_m^*\big)^\top \in \mathbb{R}^m$.

Jsou-li funkce $f, g_1, \dots, g_m$ dokonce třídy C^2 na $\mathscr{U}$, pak Hessova matice Lagrangeovy funkce vzhledem k x je pozitivně semidefinitní na ortogonálním doplňku k podprostoru generovanému vektory $\operatorname{grad} g_1(x^*), \dots, \operatorname{grad} g_m(x^*)$, tj.

$$h^\top \nabla_x^2 L(x^*,\lambda^*)\, h \geq 0 \tag{1.2.8}$$

pro všechna

$$h \in \operatorname{Ker} \mathrm{D}G(x^*) = \big\{h \in \mathbb{R}^n \mid \mathrm{D}G(x^*)\, h = 0\big\} = \\ = \big\{h \in \mathbb{R}^n \mid \langle \operatorname{grad} g_i(x^*), h\rangle = 0 \text{ pro všechna } i = 1,\dots,m\big\}.$$

Poznámka 1.2.17. Při pohledu na soustavu (1.2.6) či její maticový zápis (1.2.7) by se mohlo zdát, že dvojice $[x^*,\lambda^*]$ je klasickým stacionárním bodem Lagrangeovy funkce $L(x,\lambda)$. Avšak tento pohled je zcela chybný! Náš přístup rozhodně není založený na tom, že z úlohy na vázaný extrém učiníme pomocí Lagrangeovy funkce úlohu na volný extrém v $\mathbb{R}^{n+m}$, tj. že bychom m-tici $\lambda_1,\dots,\lambda_m$ považovali za dodatečné proměnné (a ani to v důkazu níže neděláme)! Navíc se dá ukázat, že tím bychom získali zcela špatné řešení, neboť nalezený stacionární bod $[x^*,\lambda^*]$ nikdy nebude lokálním extrémem funkce $L(x,\lambda)$, viz [308, str. 191]. Např. pro minimalizaci funkce $f(x,y) = x^2 + y^2$ na množině vymezené podmínkou $x\,y = 1$ dostaneme Lagrangeovu funkci $L(x,y) = x^2 + y^2 + \lambda\,(x\,y - 1)$ a soustavu

$$2x + \lambda\, y = 0, \qquad 2y + \lambda\, x = 0 \quad \text{a} \quad x\,y = 1,$$

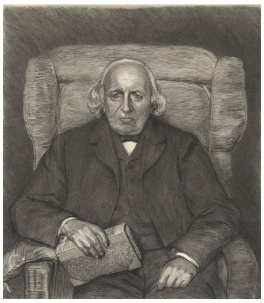
jejíž řešení jsou $[1,1,-2]$ a $[-1,-1,2]$. To jsou skutečná řešení dané úlohy, viz Obrázek 1.7. Avšak nejsou to lokální extrémy funkce $L(x,y,\lambda)$, neboť její Hessova matice

$$\nabla^2 L(x,y,\lambda) = \begin{pmatrix} 2 & \lambda & y \\ \lambda & 2 & x \\ y & x & 0 \end{pmatrix}$$

[66] Tento obecný problém byl poprvé formulován Thomasem Simpsonen v roce 1750. Weber se k němu vrátil v roce 1909 a výrazně přispěl k jeho popularizaci. Tato úloha také souvisí se vznikem tzv. *nové ekonomické geografie* v devadesátých letech 20. století, za níž získal Paul Krugman v roce 2008 *Nobelovu pamětní cenu za ekonomii*.

[67] Thomas Simpson (20. srpna 1710 – 14. května 1761) byl britský matematik. Jeho jméno je spojeno především se *Simpsonovým pravidlem* pro přibližný výpočet určitého integrálu. Avšak jak už to bývá v matematice obvyklé, toto pravidlo bylo nalezeno cca o 100 let dříve Johannem Keplerem, a proto se např. v němčině obvykle hovoří spíše o Keplerově pravidle (*Keplersche Fassregel*).

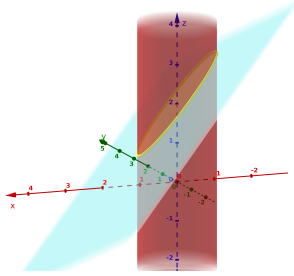
[68]



Karl Theodor Wilhelm Weierstraß (31. října 1815 – 19. února 1897) byl německý matematik a „otec moderní matematické analýzy". Ačkoli opustil univerzitu bez, aniž by získal vysokoškolský diplom, studoval stále matematiku a získal kvalifikaci k vyučování matematiky, fyziky, botaniky a gymnastiky. Později obdržel čestný doktorát a stal se profesorem matematiky na univerzitě v Berlíně. Díky němu přišlo do matematické analýzy „období přesnosti", neboť formalizoval pojmy limity (ε–δ definice) a spojitosti funkce, dokázal větu o střední hodnotě a odhalil sílu stejnoměrné konvergence posloupnosti funkcí.

[69] Máme-li množinu X v libovolném metrickém prostoru, pak ji nazýváme *kompaktní*, pokud z každé posloupnosti bodů v X lze vybrat podposloupnost konvergující v X. V konečně dimenzionálních normovaných prostorech (jako např. $\mathbb{R}^n$) je množina X kom-

paktní právě tehdy, když je uzavřená a ohraničená.



Obrázek 1.6: Množina X (elipsa) vymezená funkcemi $g_1(x,y,z) = x+y+z-2$ (rovina) a $g_2(x,y,z) = x^2+y^2-1$ (válec).

Joseph-Louis Lagrange (25. ledna 1736 – 10. dubna 1813) byl dalším důležitým matematikem a astronomem 18. století, který výrazně ovlivnil především rozvoj matematické analýzy a teorie čísel a také klasickou i nebeskou mechaniku. Narodil se v italském Turíně jako Giuseppe Lodovico Lagrangia, ale vzhledem k jeho činnosti v pozdějších letech je považován spíše za (naturalizovaného) Francouze. Jedním ze 16 matematiků, jejichž jméno je napsáno na Eiffelově věži. V roce 1766 na doporučení Eulera a francouzského matematika Jeana-Baptista le Ronda d'Alemberta stanul v čele Pruské akademie věd v Berlíně – ovšem až poté, co Euler odcestoval z Berlína do Petrohradu. Zde působil více než 20 let, během nichž vykonal mnoho práce a získal řadu ocenění Francouzské akademie věd. V roce 1787 (v 51 letech) se přesunul do Paříže, kde se stal členem tamní akademie věd. Výrazně se podílel na zavedení metrického systému během Velké francouzské revoluce (1789–1799) a v roce 1794 se stal prvním profesorem (matematické) analýzy na věhlasné École polytechnique. Byl také zakládajícím členem „Bureau des Longitudes" („Kancelář pro délky") a v roce 1799 se stal dokonce senátorem.

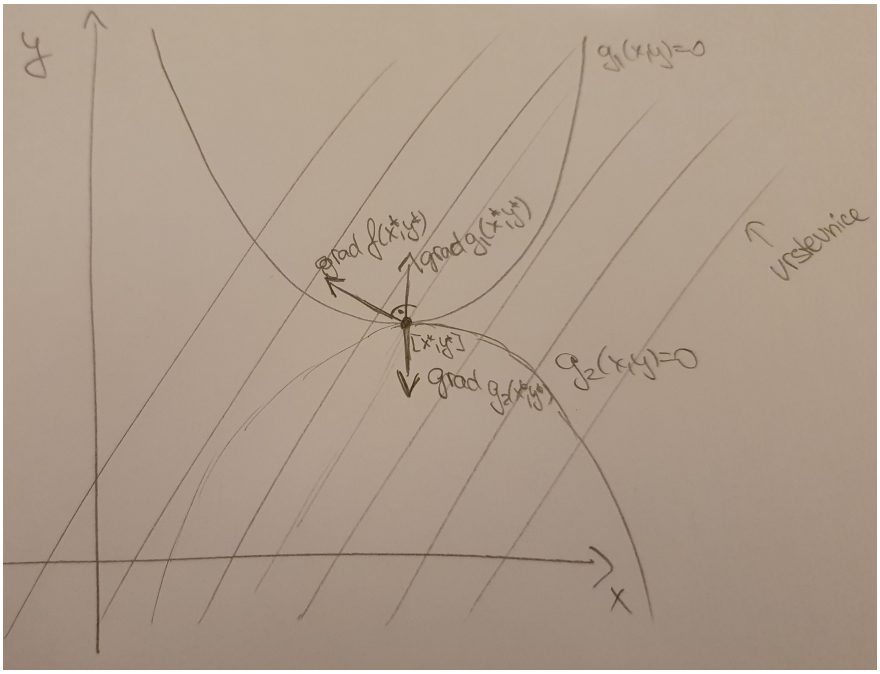
má v těchto bodech podobu

$$\nabla^2 L(1,1,-2) = \begin{pmatrix} 2 & -2 & 1 \\ -2 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{a} \quad \nabla^2 L(-1,-1,2) = \begin{pmatrix} 2 & 2 & -1 \\ 2 & 2 & -1 \\ -1 & -1 & 0 \end{pmatrix},$$

což jsou indefinitní matice (jejich vlastní čísla jsou $4, \pm\sqrt{2}$ a $-4, \pm\sqrt{2}$). To znamená, že není splněna nutná podmínka druhého řádu pro lokální extrém, a tudíž se jedná o sedlové body funkce $L(x,y,\lambda)$. To ovšem není v rozporu s druhou částí Věty 1.2.16, neboť v ní se vyžaduje negativní semidefitnost $n \times n$ matice $\nabla_x^2 L(x^*,\lambda^*)$ pouze na podprostoru $\operatorname{Ker} DG(x^*)$.

Z obecnější formulace Lagrangeova principu ve Větě 8.2.1 v Kapitole 8 uvidíme, že se tato nepřesná interpretace je možná pouze shodou několika okolností a lineární závislostí funkce $L(x,\lambda)$ na jednotlivých multiplikátorech $\lambda_1,\dots,\lambda_m$, takže derivováním vzhledem k λ_i a položením rovného nule dostaneme právě i-té omezení určující množinu X. Ani samotný Lagrange o něčem takovém nehovořil – v předchozí poznámce jsme zmínili jeho komentář týkající se eliminace neurčitých koeficientů přidáním všech omezujících podmínek k soustavě (1.2.6), tj. hledáme dvojici x^* a λ^* takovou, že x^* je stacionárním bodem funkce $L(x,\lambda^*)$, která má pouze n-proměnných $x_1,\dots,x_n$, tj. $\operatorname{grad}_x L(x^*,\lambda^*) = 0$. Tím dostaneme soustavu $n+m$ rovnic o $n+m$ neznámých a požadavek plné hodnosti matice $DG(x^*)$ zaručí jednoznačnost jejího řešení (a pokud lokální extrém neexistuje, nemusí mít tato soustava žádné řešení).

Další možná (a správná) interpretace podmínky (1.2.7) je ta, že hledáme bod x^*, pro který jsou vektory $\operatorname{grad} f(x^*)$ a $\operatorname{grad} g_1(x^*),\dots,\operatorname{grad} g_m(x^*)$ lineárně závislé, přičemž koeficienty určují $\lambda_1^*\dots,\lambda_m^*$ „správnou" lineární kombinaci vektorů $\operatorname{grad} g_1(x^*),\dots,\operatorname{grad} g_m(x^*)$, která odpovídá $-\operatorname{grad} f(x^*)$. K tomu, aby to ale bylo vůbec možné, je opět důležité, aby byl splněn požadavek plné hodnosti matice $DG(x^*)$, viz Obrázek 1.8. Tento požadavek nám svým způsobem říká, že dané *omezení je kvalifikované* pro použití Věty 1.2.16.



Obrázek 1.8: Minimalizace lineární účelové funkce $f(x,y)$ (vrstevnice jsou rovnoběžné přímky) a přípustná množina vymezená dvojicí podmínek $g_1(x,y)=0$ a $g_2(x,y)=0$, které jsou splněny pouze v bodě $[x^*,y^*]$. Tento bod je tedy globální i lokální maximum a minimum funkce f na dané množině. Avšak neexistují takové multiplikátory λ_1^* a λ_2^*, aby byla splněna podmínka (1.2.7). Vektory $\operatorname{grad} g_1(x^*,y^*)$ a $\operatorname{grad} g_2(x^*,y^*)$ jsou totiž lineárně závislé a žádnou jejich lineární kombinací pomocí nedostaneme vektor $\operatorname{grad} f(x^*,y^*)$.

V této souvislosti bychom také měli zdůraznit, že vektor $\operatorname{grad} f(x^*)$ je normálovým vektorem tečny odpovídající vrstevnice procházející bodem x^* a vektory $\operatorname{grad} g_1(x^*),\dots,\operatorname{grad} g_m(x^*)$ jsou

normálovými vektory tečen nadploch $g_1(x)=0,\dots,g_m(x)=0$ v bodě x^*. Proto v případě úlohy s jedním omezením $g(x)=0$ můžeme nabídnou i geometrickou interpretaci podmínky (1.2.7): hledáme bod x^* takový, že tečna k odpovídající vrstevnici funkce f bude splývat s tečnou k nadploše určené daným omezením v tomtéž bodě. ▲

Důkaz Věty 1.2.16. Pro jednoduchost se omezme na hledání minima funkce f na X. Důkaz pro maximalizaci f na X se provede zcela analogicky. Nejdříve uvažme situaci, kdy jsou všechna daná omezení lineární, tj. množina X odpovídá řešení soustavy $Ax=b$ pro matici $A\in\mathbb{R}^{m\times n}$ a vektor $b\in\mathbb{R}^m$. V takovém případě máme $G(x)=Ax-b$ a podmínka plné hodnosti odpovídající Jacobiho matice znamená, že řádky matice A jsou lineárně nezávislé, tj. $\operatorname{rank}A=m$, přičemž bez újmy na obecnosti můžeme předpokládat, že sloupce matice A jsou uspořádány tak, že prvních m sloupců je také lineárně nezávislých (toho lze snadno docílit vhodným uspořádáním proměnných $x_1,\dots,x_n$. Jinými slovy, vedoucí hlavní submatice řádu m matice A je regulární, takže můžeme matici A a vektor b rozdělit do dvou bloků

$$A=\begin{pmatrix}A_r & A_s\end{pmatrix}\quad\text{a}\quad x=\begin{pmatrix}x_r\\ x_s\end{pmatrix}$$

s regulární maticí $A_r\in\mathbb{R}^{m\times m}$, zbývající částí $A_s\in\mathbb{R}^{m\times(n-m)}$ odpovídající složkami $x\in\mathbb{R}^m$ a $x_s\in\mathbb{R}^{n-m}$. Pak můžeme úlohu minimalizace funkce f na X zapsat jako

$$\begin{aligned}f(x)&\to\min\\ A_r x_r+A_s x_s&=b.\end{aligned}$$

Díky regulárnosti matice A_r odtud plyne $x_r=A_r^{-1}\big(b-A_s x_s\big)$, takže můžeme tuto úlohu převést na problém nepodmíněné minimalizace

$$F(x_s):=f\big(A_r^{-1}(b-A_s x_s),x_s\big)\to\min,\quad x_s\in\mathbb{R}^{n-m}.\tag{1.2.9}$$

Je-li $x^*=[x_r^*,x_s^*]$ lokálním minimem funkce f na X, pak x_s^* musí být stacionárním bodem funkce $F(x_s)$, takže

$$0=\operatorname{grad}F(x_s^*)=-A_s^\top A_r^{\top-1}\operatorname{grad}_{x_r}f(x^*)+\operatorname{grad}_{x_s}f(x^*).\tag{1.2.10}$$

Definujeme-li

$$\lambda^*:=-A_r^{\top-1}\operatorname{grad}_{x_r}f(x^*),\tag{1.2.11}$$

pak rovnost (1.2.10) můžeme zapsat ve tvaru

$$\operatorname{grad}_{x_s}f(x^*)+A_s^\top\lambda^*=0,\tag{1.2.12}$$

zatímco z (1.2.11) máme

$$\operatorname{grad}_{x_r}f(x^*)+A_r^\top\lambda^*=0.\tag{1.2.13}$$

Zkombinováním těchto dvou rovností obdržíme

$$\operatorname{grad}f(x^*)+A^\top\lambda^*=0,\tag{1.2.14}$$

což je transponovaná verze rovnosti (1.2.7), neboť $\mathrm{D}G(x)=\mathrm{D}(Ax-b)=A$. Navíc z regulárnosti matice A_r plyne, že multiplikátory $\lambda_1^*,\dots,\lambda_m^*$ jsou skutečně určeny jednoznačně. Připusťme totiž, že existuje další m-tice multiplikátorů $\mu_1^*,\dots,\mu_m^*\in\mathbb{R}$ splňující $\operatorname{grad}f(x^*)+A^\top\mu^*=0$. Pak musí být

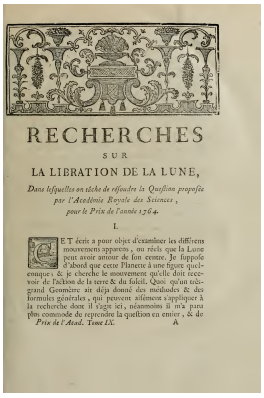
$$0=A^\top(\mu^*-\lambda^*)=\begin{pmatrix}A_r^\top(\mu^*-\lambda^*)\\ A_s^\top(\mu^*-\lambda^*)\end{pmatrix}\quad\text{neboli}\quad A_r^\top(\mu^*-\lambda^*)=0=A_s^\top(\mu^*-\lambda^*),$$

a tudíž díky první rovnosti $\mu^*=\lambda^*$.

[71] Tehdy šlo spíše o *Královskou akademii věd* (Académie Royale des Sciences), jak tehdejší učenou společnost pojmenoval francouzský král Ludvík XIV. při stanovování jejích základních pravidel 20. ledna 1699. On sám tuto učenou společnost založil již o 33 let dříve na návrh svého ministra financí Jeana-Baptista Colberta s cílem podpořit a chránit francouzské vědecké bádání.

[72] Obecně se v astronomii tímto pojmem označuje velmi pomalá oscilace satelitu (ať již reálná či zdánlivá), která se jeví pozorovateli na hmotnějším mateřském tělese, jenž satelit obíhá. Většinou se tento termín používá k označení kývavého pohybu Měsíce vůči pozorovateli na Zemi. Přestože je totiž rotace Měsíce kolem Země synchronní neboli vázaná (tj. Měsíc nám stále ukazuje jednu tzv. přivrácenou stranu), librace umožňuje pozorovatelům na Zemi v průběhu jedné měsíční periody spatřit více než jednu polovinu jeho povrchu – přibližně jde o 59 procent jeho povrchu (ve skutečnosti však jde spíše o něco málo nad 50 procent, neboť v důsledku značného zkreslení v libační zóně je pro praktickou využitelnost hodnota nižší).

[73]


RECHERCHES
SUR
LA LIBRATION DE LA LUNE,
Dans lesquelles on tâche de résoudre la Question proposée par l'Académie Royale des Sciences, pour le Prix de l'année 1764.
L.

[74] *Virtuální rychlostí je třeba rozumět tu, kterou by získalo těleso v rovnovážné poloze, kdyby byla narušena rovnováha; konkrétně jde o rychlost, kterou by těleso skutečně nabylo v prvním okamžiku pohybu; princip spočívá v tomto: síly jsou v rovnováze, pokud jsou v opačném poměru k jejich virtuálním rychlostem ve směrech těchto sil.*
(One must understand by the term *virtual velocity*, the velocity which a body in equilibrium would take if the state of equilibrium ceased to exist, that is, the velocity that the body would have in the first instant of its motion. The principle requires the forces to be in a state of equilibrium if they are to be inversely proportional to their virtual

velocities taken in the direction of these forces.) Viz Odstavec 16 v první sekci [343].

[75] *Cílem této práce je redukovat mechaniku na čistě algebraické operace a vzorec, který jsme právě odvodili, je pro tento účel zcela vyhovující. Je pouze potřeba algebraicky vyjádřit nejobecnějším možným způsobem hodnoty p, q, r atd. ve směrech sil P, Q, R atd. a pak jednoduchým diferencování získáme hodnoty virtuálních rychlostí $\mathrm{d}p, \mathrm{d}q, \mathrm{d}r$.*
(The purpose of this work is to reduce mechanics to purely algebraic operations and the formula which we have just developed is fully adequate for this purpose. It is only necessary to express algebraically, in the most general fashion, the values of p, q, r, etc. taken in the directions of the forces P, Q, R, etc. and the values of the virtual velocities $\mathrm{d}p$, $\mathrm{d}q$, $\mathrm{d}r$, etc. will be obtained by simple differentiation.)

[76] *Z tohoto přístupu plyne mimořádně jednoduché pravidlo pro nalezení podmínek určující rovnovážnou polohu daného systému.*
Můžeme vzít součet momentů všech sil, které musí být v rovnováze, a přidat k nim různé diferenciály funkcí, které musí rovny nule, z podmínek uvažovaného problému poté, co každou z nich vynásobíme neurčitým koeficientem. To celé můžeme položit rovno nule, čímž získáme diferenciální rovnici, s níž můžeme pracovat jako s běžnou rovnicí pro nalezení maxima a minima a z níž může získat tolik jednotlivých konečných rovnic jako je počet proměnných. Tyto rovnice dají po odstranění neurčitých koeficientů nutné podmínky pro rovnovážnou polohu.
Tato diferenciální rovnice bude mít podobu

$$P\,\mathrm{d}p+Q\,\mathrm{d}q+\cdots+\lambda\,\mathrm{d}L+\mu\,\mathrm{d}M+\cdots=0,$$

ve které λ, μ atd. jsou neurčité koeficienty. Tuto rovnici budeme nazývat obecnou rovnicí rovnovážné polohy.
(An extraordinarily simple rule results from this approach to find the equilibrium conditions of an arbitrary, given system.
The sum of the *moments* of all the forces which must be in equilibrium can be gathered and to which can be added the various differential functions which must be equal to zero from the conditions of the problem after having multiplied each of these functions by an undetermined coefficient. The whole can be equated to zero and thus a differential equation results which can be treated as an ordinary equation of *maxima and minima* and from which as many par-

Pro důkaz podmínky druhého řádu (1.2.8) v tomto speciálním případě stačí ukázat její ekvivalenci s nutnou podmínkou druhého řádu pro řešení úlohy nepodmíněné minimalizace (1.2.9), tj. s podmínkou

$$h\,\nabla^2 F(x_s^*)\,h \geq 0 \quad \text{pro všechna } h \in \mathbb{R}^{n-m}.$$

Nechť tedy $f \in \mathrm{C}^2$. Díky (1.2.10) platí

$$\nabla^2 F(x_s) = \mathrm{D}(\operatorname{grad} F(x_s)) = \mathrm{D}\big(-A_s^\top A_r^{\top -1} \operatorname{grad}_{x_r} f((b-A_s x_s), x_s) + \operatorname{grad}_{x_s} f((b-A_s x_s), x_s)\big).$$

Volbou $x_s = x_s^*$ a rozdělením Hessovy matice funkce f do bloků

$$\nabla^2 f(x^*) = \begin{pmatrix} \nabla^2_{x_r} f(x^*) & \nabla^2_{x_r x_s} f(x^*) \\ \nabla^2_{x_r x_s} f(x^*) & \nabla^2_{x_s} f(x^*) \end{pmatrix}$$

dostaneme

$$\nabla^2 F(x_s) = A_s^\top A_r^{\top -1} \nabla^2_{x_r} f(x^*)\, A_r^{-1} A_s - \nabla^2_{x_r x_s} f(x^*)\, A_r^{-1} A_s - A_s^\top A_r^{\top -1} \nabla^2_{x_s, x_r} f(x^*) + \nabla^2_{x_s} f(x^*)$$

neboli

$$\nabla^2 F(x_s) = \begin{pmatrix} -A_r^{-1} A_s \\ 1 \end{pmatrix}^\top \begin{pmatrix} \nabla^2_{x_r} f(x^*) & \nabla^2_{x_r x_s} f(x^*) \\ \nabla^2_{x_r x_s} f(x^*) & \nabla^2_{x_s} f(x^*) \end{pmatrix} \begin{pmatrix} -A_r^{-1} A_s \\ 1 \end{pmatrix}.$$

Tato rovnost společně se skutečnostmi, že nutně musí být $\nabla^2 F(x_s^*) \geq 0$ kvůli lokálnímu minimum v x_s^* a $\nabla^2 G(x^*) = 0$ kvůli linearitě daných omezení, dává pro libovolné $h \in \mathbb{R}^{n-m}$, že

$$0 \leq h^\top \nabla^2 F(x_s)\, h = v^\top \nabla^2 f(x^*)\, v = v^\top \underbrace{\Big(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*)\Big)}_{\nabla^2_x L(x^*, \lambda^*)} v, \tag{1.2.15}$$

kde vektor $v \in \mathbb{R}^n$ je složen z bloků

$$v := \begin{pmatrix} -A_r^{-1} A_s\, h \\ h \end{pmatrix}.$$

Protože $\operatorname{Ker} \mathrm{D}G(x^*)$ je $(n-m)$-dimenzionální podprostor, který můžeme vyjádřit jako

$$\begin{aligned} \operatorname{Ker} \mathrm{D}G(x^*) &= \{h \in \mathbb{R}^n \mid A\,h = 0\} = \{h = (h_r, h_s)^\top \in \mathbb{R}^n \mid A_r\, h_r + A_s\, h_s = 0\} = \\ &= \{h = (h_r, h_s)^\top \in \mathbb{R}^n \mid h_r = -A_r^{-1} A_s\, h_s,\ h_s \in \mathbb{R}^{n-m}\}, \end{aligned}$$

je nerovnost (1.2.15) ekvivalentní s požadavkem (1.2.8).

Důkaz tvrzení v obecném případě je založen na Větě o implicitní funkci[78] aplikované na funkci $G(x)$. Omezíme se pouze na případ, kdy $m < n$, neboť pro $m = n$ je tvrzení triviální – libovolný vektor včetně $-\mathrm{D}f(x^*)$ lze totiž jistě vyjádřit jediným způsobem jako lineární kombinaci n-tice lineárně nezávislých vektorů $\operatorname{grad} g_1(x^*), \ldots, \operatorname{grad} g_m(x^*)$. Podobně jako výše platí, že vhodným přeznačením proměnných docílíme toho, že vedoucí hlavní submatice řádu m v $\mathrm{D}G(x^*)$ je regulární, tj.

$$\text{pro } x = \begin{pmatrix} x_r \\ x_s \end{pmatrix} \quad \text{máme} \quad \mathrm{D}G(x^*) = \big(\mathrm{D}_{x_r} G(x^*)\ \mathrm{D}_{x_s} G(x^*)\big)$$

s bloky $x_r \in \mathbb{R}^m$, $x_s \in \mathbb{R}^{n-m}$, $\mathrm{D}_{x_r} G(x^*) \in \mathbb{R}^{m \times m}$, $\mathrm{D}_{x_s} G(x^*) \in \mathbb{R}^{m \times (n-m)}$ a $\det \mathrm{D}_{x_r} G(x^*) \neq 0$. Proto bez újmy na obecnosti můžeme brát, že tento požadavek je splněn. Omezení $G(x_r, x_s) = 0$ je jistě splněno v bodě $x^* = [x_r^*, x_s^*]$, takže díky Větě o implicitní funkci můžeme x_r vyjádřit pomocí proměnných x_s prostřednictvím spojitě diferencovatelné funkce $\Phi : S \to \mathbb{R}^m$ definované na kouli S se středem v x_s^* a vhodným poloměrem. Zejména platí $x_r^* = \Phi(x_s^*)$ a

$$G(\Phi(x_s), x_s) = 0 \quad \text{a} \quad \mathrm{D}\Phi(x_s) = -\big(\mathrm{D}_{x_r} G\big(\Phi(x_s), x_s\big)\big)^{-1} \mathrm{D}_{x_s} G\big(\Phi(x_s), x_s\big)$$

pro všechna $x_s \in S$. Nyní opět podobně jako v případě lineárních omezení výše vidíme, že x_s^* musí být volným lokálním minimem „redukované“ účelové funkce

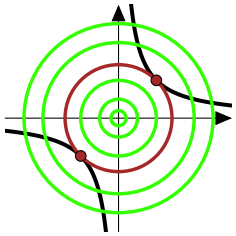
$$F(x_s) := f(\Phi(x_s), x_s), \quad x_s \in \mathbb{R}^{n-m}.$$

ticular finite equations as there are variables can be extracted. These equations from which the undetermined coefficients have been eliminated will give all the necessary conditions for equilibrium.
The differential equation will have the following form

$$P\,\mathrm{d}p+Q\,\mathrm{d}q+\cdots+\lambda\,\mathrm{d}L+\mu\,\mathrm{d}M+\cdots=0$$

in which λ,μ etc. are undetermined quantities. We will call this equation the *General Equation of Equilibrium.*)

[77] A proto nutně musí být $m \le n$!



Obrázek 1.7: Vrstevnice funkce $f(x,y) = x^2+y^2$ (zelená barva) a množina bodů splňující $xy = 1$ (černá). Minimum f na dané množině je vyznačeno červenou barvou (body i vrstevnice).

[78] *Věta o implicitní funkci.* Nechť $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$ je vektorová funkce proměnných $x \in \mathbb{R}^n$ a $y \in \mathbb{R}^m$. Nechť je dále dán bod $[x^*, y^*] \in \mathbb{R}^{n+m}$ takový, že $f(x^*, y^*) = 0$ a je prvkem nějaké otevřené množiny, na které je funkce f spojitá a Jacobiho matice $\mathrm{D}_y f(x,y) \in \mathbb{R}^{m\times m}$ je spojitá a regulární. Pak existuje otevřená množina $S_{x^*} \subseteq \mathbb{R}^n$ obsahující x^* a otevřená množina $S_{y^*} \subseteq \mathbb{R}^m$ obsahující y^* a spojitá funkce $\Phi : S_{x^*} \to S_{y^*}$ taková, že $\Phi(x^*) = y^*$ a $f\big(x,\Phi(x)\big) = 0$ pro všechna $x \in S_{x^*}$. Taková funkce Φ je jediná v tom smyslu, že kdykoli $f(x,y) = 0$ pro nějaké $x \in S_{x^*}$ a $y \in S_{y^*}$, pak nutně $y = \Phi(x)$. Je-li navíc funkce f p-krát spojitě diferencovatelná pro nějaké $p \in \mathbb{N}$, pak totéž platí i o funkci Φ. Zejména,

$$\mathrm{D}\Phi(x) = -\Big(\mathrm{D}_y f\big(x,\Phi(x)\big)\Big)^{-1}\mathrm{D}_x f\big(x,\Phi(x)\big)$$

pro všechna $x \in S_{x^*}$.

[79] Předchozí metodu důkazu můžeme považovat za tzv. *eliminační*, neboť se nám v ní podařilo eliminovat některé proměnné tak, že jsme úlohu podmíněné minimalizace (1.2.5) převedli na úlohu nepodmíněné minimalizace redukované účelové funkce $F(x_s)$. Propojení s úlohou nepodmíněné minimalizace lze ale docílit i s pomocí tzv. *penalizační metody*, což nyní využijeme k alternativnímu důka-

Proto zcela analogickým výpočtem jako v (1.2.11)–(1.2.14) s maticemi

$$A := \mathrm{D}G(x^*), \qquad A_r := \mathrm{D}_{x_r}G(x^*) \quad \text{a} \quad A_s := \mathrm{D}_{x_s}G(x^*) \tag{1.2.16}$$

dostaneme kýženou podmínku prvního řádu.

Konečně dokážeme nutnou podmínku druhého řádu i v obecném případě, čehož docílíme odvozením podobné nerovnosti jako v (1.2.15). Nechť proto $f, g_1, \ldots, g_m \in \mathrm{C}^2$. S pomocí matic z (1.2.16) definujme

$$v := \begin{pmatrix} -A_r^{-1}A_s\,h \\ h \end{pmatrix} \tag{1.2.17}$$

pro $h \in \mathbb{R}^{n-m}$. Označme také

$$G_i(x_s) := g_i(\Phi(x_s), x_s) \quad \text{pro všechna } i \in \{1,\ldots,m\}$$

a jednotlivé složky funkce $\Phi(x_s)$ jako

$$\Phi(x_s) = \begin{pmatrix} \Phi_1(x_s) \\ \vdots \\ \Phi_m(x_s) \end{pmatrix}.$$

Pak dvojnásobným derivováním funkce $F(x_s)$ dostaneme

$$h^\top\nabla^2F(x_s^*)\,h = v^\top\nabla^2f(x^*)\,v + h^\top\left(\sum_{j=1}^m\nabla^2\Phi_j(x_s^*)\,\frac{\partial f(x^*)}{\partial x_j}\right)h. \tag{1.2.18}$$

Podobně dvojnásobným derivováním rovnosti $G_i(x_s) = 0$ získáme

$$0 = h^\top\nabla^2G_i(x_s)\,h = v^\top\nabla^2g_i(x^*)\,v + h^\top\left(\sum_{j=1}^m\nabla^2\Phi_j(x_s^*)\,\frac{\partial g_i(x^*)}{\partial x_j}\right)h.$$

Vynásobením poslední rovnosti číslem λ_i^* a jejich posčítáním pro $i = 1,\ldots,m$ obdržíme

$$0 = \sum_{i=1}^m\lambda_i^*\,v^\top\nabla^2g_i(x^*)\,v + h^\top\left(\sum_{j=1}^m\nabla^2\Phi_j(x_s^*)\sum_{i=1}^m\lambda_i^*\,\frac{\partial g_i(x^*)}{\partial x_j}\right)h. \tag{1.2.19}$$

Jelikož platí (1.2.6) a $h^\top\nabla^2F(x_s^*)\,h \ge 0$ pro libovolné $h \in \mathbb{R}^{n-m}$, sečtením (1.2.18)–(1.2.19) odtud plyne

$$0 \le v^\top\left(\nabla^2f(x^*) + \sum_{i=1}^m\lambda_i^*\nabla^2g_i(x^*)\right)v + h^\top\left[\sum_{j=1}^m\nabla^2\Phi_j(x_s^*)\left(\frac{\partial f(x^*)}{\partial x_j} + \sum_{i=1}^m\lambda_i^*\,\frac{\partial g_i(x^*)}{\partial x_j}\right)\right]h =$$

$$= v^\top\underbrace{\left(\nabla^2f(x^*) + \sum_{i=1}^m\lambda_i^*\nabla^2g_i(x^*)\right)}_{\nabla_x^2L(x^*,\lambda^*)}v$$

pro libovolný vektor v definovaný výše. Avšak stejně jako výše může ukázat, že $v \in \mathrm{Ker}\,\mathrm{D}G(x^*)$ právě tehdy, když v je tvaru (1.2.17) pro libovolné $h \in \mathbb{R}^{n-m}$, čímž je důkaz hotov.[79] ■

A postačující podmínka? K tomu, aby bod x^* byl lokálním vázaným extrémem, rozhodně není nutné (ale postačující), aby byl současně lokálním extrémem funkce $L(x,\lambda^*)$. Zdůrazněme, že pro důkaz následující postačující podmínky není nutné vyžadovat, aby matice $\mathrm{D}G(x^*)$ měla plnou hodnost, jak tomu bylo ve Větě 1.2.16, kde nám tato podmínka zaručovala existenci a jednoznačnost multiplikátorů $\lambda_1,\ldots,\lambda_m$. Tentokrát již máme jejich existenci zahrnutu v předpokladech uvedené věty v části (i). Současně však stále platí, že by omezení nemělo být „příliš mnoho“, neboť zejména v případě $\mathrm{rank}\,\mathrm{D}G(x^*) = n$ bude $\mathrm{Ker}\,\mathrm{D}G(x^*) = \{0\}$, takže splnění podmínky (ii) níže bude nemožné.

Věta 1.2.18

Nechť je dána úloha (1.2.5) s $m \in \{1,\dots,n\}$ a funkcemi $f, g_1, \dots, g_m$ třídy C^2 na otevřené množině $\mathscr{U} \subseteq \mathbb{R}^n$ splňující $X \subset \mathscr{U}$. Jestliže pro bod $x^* \in X$ existuje m-tice multiplikátorů $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ taková, že

(i) Lagrangeova funkce $L(x, \lambda^*)$ má v bodě x^* stacionární bod, tj. $\operatorname{grad}_x L(x^*, \lambda^*) = 0$,

(ii) Hessova matice $\nabla_x^2 L(x^*, \lambda^*)$ je pozitivně definitní na $\operatorname{Ker}(DG(x^*))$, tj.

$$h^\top \nabla_x^2 L(x^*, \lambda^*)\, h > 0 \quad \text{pro všechna } h \in \operatorname{Ker}(DG(x^*)) \setminus \{0\}, \tag{1.2.20}$$

pak funkce f má v bodě x^* ostré vázané lokální minimum na X. Dokonce existují kladná čísla $\varepsilon, \gamma > 0$ taková, že pro libovolné $x \in X$ splňující $\| x - x^* \| < \varepsilon$ platí

$$f(x) \geq f(x^*) + \frac{\gamma}{2} \left\| x - x^* \right\|^2.$$

Důkaz. Nechť dvojice $x^* \in X$, $\lambda^* \in \mathbb{R}^m$ splňuje obě podmínky z tvrzení věty a připusťme, že bod x^* není lokálním minimem funkce f na X, tj. libovolně blízko x^* existují body patřící do X a mající menší funkční hodnotu než $f(x^*)$. Jinými slovy, v X existuje posloupnost $\{x^{[k]}\}_{k=1}^\infty$ splňující $G(x^{[k]}) = 0$, $x^{[k]} \neq x^*$ a

$$f(x^{[k]}) < f(x^*) + \frac{1}{k} \left\| x^{[k]} - x^* \right\|^2 \tag{1.2.21}$$

pro všechna $k \in \mathbb{N}$ a současně $x^{[k]} \to x^*$ pro $k \to \infty$. Položíme-li $\delta^{[k]} := \| x^{[k]} - x^* \|$ a $h^{[k]} := (x^{[k]} - x^*)/\| x^{[k]} - x^* \|$, pak platí $x^{[k]} = x^* + \delta^{[k]} h^{[k]}$ pro všechna $k \in \mathbb{N}$. Jelikož vždy $\| h^{[k]} \| = 1$, je posloupnost $\{h^{[k]}\}$ ohraničená, a tudíž obsahuje konvergentní podposloupnost. Proto bez újmy na obecnosti můžeme předpokládat, že konvergentní je dokonce samotná posloupnost $\{h^{[k]}\}$, a její limitu označíme jako h. Pak z diferencovatelnosti funkce g_i dostaneme

$$0 = \frac{g_i(x^{[k]}) - g_i(x^*)}{\delta^{[k]}} = \frac{g_i(x^* + \delta^{[k]} h^{[k]}) - g_i(x^*)}{\delta^{[k]}} = \langle \operatorname{grad} g_i(x^*), h^{[k]} \rangle + \frac{\omega(\delta^{[k]} h^{[k]})}{\delta^{[k]}},$$

takže pro $k \to \infty$ máme

$$\langle \operatorname{grad} g_i(x^*), h \rangle = 0,$$

tj. $y \in \operatorname{Ker} DG(x^*)$. Podle Taylorova rozvoje druhého řádu také platí

$$0 = g_i(x^{[k]}) - g_i(x^*) = \delta^{[k]} \langle \operatorname{grad} g_i(x^*), h^{[k]} \rangle + \frac{\delta^{[k]2}}{2} h^{[k]\top} \nabla^2 g_i(\xi_i^{[k]})\, h^{[k]}$$

a současně

$$\frac{1}{k} \| x^{[k]} - x^* \|^2 \overset{(1.2.21)}{>} f(x^{[k]}) - f(x^*) = \delta^{[k]} \langle \operatorname{grad} f(x^*), h^{[k]} \rangle + \frac{\delta^{[k]2}}{2} h^{[k]\top} \nabla^2 f(\eta^{[k]})\, h^{[k]}$$

pro nějaká $\xi_i^{[k]}$ a $\eta^{[k]}$ ležící na úsečce spojující $x^{[k]}$ a x^*. Vynásobením první rovnosti číslem λ_i^* a jejím přičtením pro všechna $i = 1, \dots, m$ k nerovnosti níže získáme díky prvnímu předpokladu našeho tvrzení, že

$$\begin{aligned} \frac{1}{k} \| x^{[k]} - x^* \|^2 &> \delta^{[k]} \left\langle \operatorname{grad} f(x^*) + \sum_{i=1}^m \lambda_i^* \operatorname{grad} g_i(x^*), h^{[k]} \right\rangle + \\ &\quad + \frac{\delta^{[k]2}}{2} h^{[k]\top} \left[\nabla^2 f(\eta^{[k]}) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\xi_i^{[k]}) \right] h^{[k]} = \\ &= \frac{\delta^{[k]2}}{2} h^{[k]\top} \left[\nabla^2 f(\eta^{[k]}) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\xi_i^{[k]}) \right] h^{[k]}. \end{aligned}$$

Protože $\delta^{[k]} = \| x^{[k]} - x^* \|$, plyne odtud

$$\frac{2}{k} > h^{[k]\top} \left[\nabla^2 f(\eta^{[k]}) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\xi_i^{[k]}) \right] h^{[k]},$$

důkazu tvrzení Věty 1.2.16. Pro $k = 1, 2, \dots$ si označme funkci

$$F^{[k]}(x) := f(x) + \frac{k}{2} \| G(x) \| + \\ + \frac{\alpha}{2} \| x - x^* \|^2,$$

kde x^* je lokální řešení úlohy (1.2.5) a $\alpha > 0$ je libovolné. Člen $(k/2)\| G(x) \|$ vystupuje jako penalizace za porušení omezení $G(x) = 0$, zatímco role posledního členu $(\alpha/2)\| x - x^* \|^2$ je pouze technická, abychom zajistili, že x^* je *ostré* lokální minimum funkce $f(x) + (k/2)\| G(x) \|$ při omezení $G(x) = 0$.

Jelikož x^* je lokální minimum, existuje $\varepsilon > 0$ takové, že $f(x^*) \leq f(x)$ pro všechna přípustná $x \in X$ patřící do uzavřené koule $S := \{x \in \mathbb{R}^n \mid \| x - x^* \| \leq \varepsilon\}$. Označme jako $x^{[k]}$ řešení úlohy

$$F^{[k]}(x) \to \min, \quad x \in S, \qquad (*)$$

které jistě existuje díky Weierstrassově větě 1.2.11. Ukážeme, že potom posloupnost $\{x^{[k]}\}_{k=1}^\infty$ konverguje k x^*. Pro všechna $k \in \mathbb{N}$ máme

$$F^{[k]}\left(x^{[k]}\right) = f\left(x^{[k]}\right) + \frac{k}{2} \| G(x^{[k]}) \| + \\ + \frac{\alpha}{2} \| x^{[k]} - x^* \|^2 \leq \\ \leq F^{[k]}(x^*) = f(x^*).$$

Protože hodnota $f\left(x^{[k]}\right)$ je ohraničená na S, musí být

$$\lim_{k \to \infty} \| G(x^{[k]}) \| = 0,$$

protože v opačném případě by $F^{[k]}\left(x^{[k]}\right) \to \infty$ pro $k \to \infty$, což by byl spor s předchozí nerovností. Proto každý hromadný bod $\bar{x}$ posloupnosti $\{x^{[k]}\}_{k=1}^\infty$ splňuje omezení $G(\bar{x}) = 0$, tj. $\bar{x} \in X$. Navíc předchozí nerovnost dává

$$f\left(x^{[k]}\right) + \frac{\alpha}{2} \| x^{[k]} - x^* \|^2 \leq f(x^*)$$

pro všechna $k \in \mathbb{N}$, takže pro $k \to \infty$ máme $f(\bar{x}) + \frac{\alpha}{2} \| \bar{x} - x^* \|^2 \leq f(x^*)$. Protože $\bar{x} \in X \cap S$, platí také $f(x^*) \leq f(\bar{x})$, takže

$$f(\bar{x}) + \frac{\alpha}{2} \| \bar{x} - x^* \|^2 \leq f(x^*) \leq f(\bar{x}),$$

což je možné pouze v případě, že $\frac{\alpha}{2} \| \bar{x} - x^* \|^2 = 0$, tj. $\bar{x} = x^*$. To znamená, že posloupnost $\{x^{[k]}\}_{k=1}^\infty$ konverguje k x^*, a tudíž $x^{[k]}$ je jistě vnitřním bodem S pro všechna k dostatečně velká. Proto $x^{[k]}$ je nepodmíněný minimem funkce $F^{[k]}$ pro dostatečně velká k. Pak z nutné podmínky prvního řádu plyne

a tudíž limitním přechodem pro $k \to \infty$ dostáváme

$$0 \geq h^\top \left[\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*) \right] h \quad \text{neboli} \quad h^\top \nabla_x^2 L(x^*, \lambda^*)\, h \leq 0,$$

což je ale ve sporu s druhým předpokladem našeho tvrzení ↯, a tím je důkaz hotov.[80, 81] ■

Otestovat splnění podmínky (1.2.20) můžeme (nejméně) dvěma alternativními způsoby. První možností (přímočarou a vždy funkční) je „analýza“ vektorů splňujících $G^\top(x^*)\, h = 0$, tj. nalezení báze vektorového podprostoru $\operatorname{Ker} DG(x^*)$, díky čemuž můžeme následně zjišťování definitnosti matice $\nabla_x^2 L(x^*, \lambda^*)$ na $\operatorname{Ker} DG(x^*)$ převést na otázku klasické definitnosti matice velikosti $\dim \operatorname{Ker} DG(x^*) \times \dim \operatorname{Ker} DG(x^*)$. Druhá možnost je založena na výpočtu vedoucích hlavních minorů tzv. *rozšířené Hessovy matice*, které se musí chovat v souladu s následujícím tvrzením.

Nechť jsou dány matice $A \in \mathbb{R}^{n\times n}$ a $B \in \mathbb{R}^{m\times n}$, přičemž A je symetrická, $m < n$ a $\operatorname{rank} B = m$ (tj. matice B má plnou hodnost). Potom $A > 0$ na $\operatorname{Ker} B$ neboli $h^\top A\, h > 0$ pro všechna $h \in \operatorname{Ker} B \setminus \{0\}$, jestliže rozšířená matice

$$H := \begin{pmatrix} \overbrace{0}^{m} & \overbrace{B}^{n} \\ B^\top & A \end{pmatrix} \begin{matrix} \} m \\ \} n \end{matrix}$$

má následující vlastnost: posledních $n - m$ vedoucích hlavních minorů matice H počínaje $\det H$ má stejné znaménko jako $(-1)^m$.

Poznámka 1.2.20.

(i) Postačující podmínka pro $A < 0$ na $\operatorname{Ker} B$, tj. $h^\top A\, h < 0$ pro všechna $h \in \operatorname{Ker} B \setminus \{0\}$, vyžaduje otestování posledních $n - m$ vedoucích hlavních minorů počínaje $\det H$, které musí střídat znaménko tak, že $\operatorname{sgn}(\det H) = (-1)^n$.

(ii) Na rozdíl od klasické charakterizace pozitivně definitní matice se zde nejedná o ekvivalentní popis. Problém totiž může být ve struktuře matice B, a bylo by potřeba se vyhnout situacím, kdy např. několik prvních sloupců matice B je nulových. Z podobného důvodu neuvádíme ani podmínky pro semidefinitní matice na $\operatorname{Ker} B$.

(iii) V nejjednodušším případě jednořádkové matice $B \in \mathbb{R}^{1\times n}$ máme rozšířenou matici ve speciálním tvaru

$$H = \left(\begin{array}{c|c} 0_{1\times 1} & \\ | & \mathrm{A} \end{array} \right).$$

Pak pro zjištění definitnosti symetrické matice A na $\operatorname{Ker} B$ pomocí předchozí věty je potřeba otestovat $n - 1$ posledních vedoucích hlavních minorů matice H, přičemž platí:

- jestliže tyto minory mají stejné znaménko jako $(-1)^m = (-1)^1 = -1$ (tj. jsou-li záporné), pak $A > 0$ na $\operatorname{Ker} B$;
- jestliže tyto minory střídají znaménko a $\operatorname{sgn}(\det H) = (-1)^n$, pak $A < 0$ na $\operatorname{Ker} B$.

V případě matice $B = (b_1, \ldots, b_n) \in \mathbb{R}^{1\times n}$ splňující $b_1 \neq 0$, se uvedené postačující podmínky stanou současně nutnými (tj. budeme mít ekvivalentní charakterizaci). ▲

Problém (1.2.5) je velmi speciální případ úlohy matematického programování, ve které jsou všechna omezení pouze ve tvaru rovností a žádná proměnná není omezena na znaménko. V Kapitole 8 si ukážeme, jak řešit zcela obecnou úlohu matematického programování s rovnostními i nerovnostními omezeními a případnými požadavky na znaménko některých (či všech) proměnných. Ovšem existují i i jiné velmi důležité speciální typy úloh matematického programování, a těm se budeme věnovat nejdříve.

$$\left.\begin{aligned} 0 &= \operatorname{grad} F^{[k]}(x^{[k]}) = \\ &= \operatorname{grad} f(x^{[k]}) + \\ &\quad + k\, \mathrm{D}^\top G(x^{[k]})\, G(x^{[k]}) + \\ &\quad + \alpha (x^{[k]} - x^*). \end{aligned}\right\} \quad (**)$$

Protože $\operatorname{rank} DG(x^*) = m$, platí totéž o $\operatorname{rank} DG(x^{[k]})$ pro všechna dostatečně velká k. V takovém případě je čtvercová matice $DG(x^{[k]})\, \mathrm{D}^\top G(x^{[k]})$ regulární (dokonce pozitivně definitní) a vynásobením obou stran rovnosti (**) výrazem $[DG(x^{[k]})\, \mathrm{D}^\top G(x^{[k]})]^{-1} DG(x^{[k]})$ zleva dostaneme

Věta 1.2.19

$$\begin{aligned} k\, G(x^{[k]}) &= \\ &= -[DG(x^{[k]})\, \mathrm{D}^\top G(x^{[k]})]^{-1} \times \\ &\quad \times DG(x^{[k]}) [f(x^{[k]}) + \\ &\quad + \alpha (x^{[k]} - x^*)]. \end{aligned}$$

Limitním přechodem pro $k \to \infty$ vidíme, že posloupnost $\{k\, G(x^{[k]})\}_{k=1}^{\infty}$ konverguje k vektoru

$$\begin{aligned} \lambda^* := -[DG(x^*)\, \mathrm{D}^\top G(x^*)]^{-1} \times \\ \times DG(x^*)\, f(x^*). \end{aligned}$$

Současně tímtéž limitním přechodem v (**) obdržíme

$$\operatorname{grad} f(x^*) + \mathrm{D}^\top G(x^{[k]})\, \lambda^* = 0,$$

což je právě transponovaná verze (1.2.7), čímž je první část tvrzení dokázána. S využitím nutné podmínky druhého řádu pro řešení úlohy (*) bude

$$\begin{aligned} \nabla^2 F^{[k]}(x^{[k]}) &= \\ &= \nabla^2 f(x^{[k]}) + \\ &\quad + k\, \mathrm{D}^\top G(x^{[k]})\, DG(x^{[k]}) + \\ &\quad + k \sum_{i=1}^{m} g_i(x^{[k]}) \nabla^2 g_i(x^{[k]}) + \\ &\quad + \alpha I \geq 0 \end{aligned}$$

pro všechna k dostatečně velká a libovolné $\alpha > 0$. Nechť $h \in \mathbb{R}^n$ je libovolný vektor splňující $DG(x^*)\, h = 0$ a nechť $h^{[k]}$ je projekce tohoto vektoru na $\operatorname{Ker} DG(x^{[k]})$, tj.

$$\begin{aligned} h^{[k]} = h - \mathrm{D}^\top G(x^{[k]}) \times \\ \times [DG(x^{[k]})\, \mathrm{D}^\top G(x^{[k]})]^{-1} \times \\ \times DG(x^{[k]})\, h. \end{aligned}$$

Jelikož $DG(x^{[k]})\, h^{[k]} = 0$ a $\nabla^2 F^{[k]}(x^{[k]}) \geq 0$ platí

$$\left.\begin{aligned} 0 \leq h^{[k]\top} \nabla^2 F^{[k]}(x^{[k]})\, h^{[k]} &= \\ = h^{[k]\top} \Big[\nabla^2 f(x^{[k]}) + \\ + k \sum_{i=1}^{m} g_i(x^{[k]}) \nabla^2 g_i(x^{[k]}) \Big] h^{[k]} \\ + \alpha \| h^{[k]} \|^2. \end{aligned}\right\} \quad (\triangle)$$

Navíc z předchozí části již víme, že $k\,g_i(x^{[k]})$ konverguje k číslu λ_i^* pro $k\to\infty$, a z (△) vidíme, že $h^{[k]}\to h$ pro $k\to\infty$ kvůli $x^{[k]}\to x^*$ a podmínce $\mathrm{D}G(x^*)\,h=0$, takže pravá strana předchozí nerovnosti konverguje pro $k\to\infty$ a tímto limitním přechodem získáme

$$0\le h^\top\Big[\nabla^2 f(x^*)+ \\ +k\sum_{i=1}^m g_i(x^*)\nabla^2 g_i(x^*)\Big]h+ \\ +\alpha\,\|h\|^2.$$

Avšak $\alpha>0$ může být libovolně malé, tudíž nutně musí být dokonce

$$0\le h^\top\Big[\nabla^2 f(x^*)+\textstyle\sum_{i=1}^m\lambda_i^*\nabla^2 g_i(x^*)\Big]h,$$

což je vzhledem k volbě h právě podmínka (1.2.8).

[80] Hlavní myšlenka předchozího důkazu je založena na sporu s případnou existencí *přípustných směrů* $x^{[k]}-x^*$, ve kterých by účelová funkce nabývala menších hodnot než $f(x^*)$. Druhou možností je důkaz založený na tzv. *rozšířené Lagrangeově funkci*

$$L_c(x,\lambda)=f(x)+\sum_{i=1}^m\lambda_i\,g_i(x)+ \\ +\frac{c}{2}\,\|G(x)\|^2,$$

kde $\|\cdot\|$ značí (libovolnou) maticovou normu (všechny jsou ekvivalentní, takže volba konkrétní normy ovlivní hodnotu parametru c) a $c\in\mathbb{R}$. Toto je Lagrangeova funkce minimalizační úlohy

$$f(x)+\tfrac{c}{2}\,\|G(x)\|^2\to\min,\quad x\in X,\quad(\square)$$

která vzhledem k požadavku $G(x)=0$ má stejné řešení jako původní úloha (1.2.5). Gradient funkce $L_c(x,\lambda)$ vzhledem k původním proměnným x je

$$\mathrm{grad}_x L_c(x,\lambda)=\mathrm{grad}\,f(x)+ \\ +(\lambda+c\,G(x))^\top\mathrm{D}G(x)$$

a její Hessova matice

$$\nabla_x^2 L_c(x,\lambda)= \\ =\nabla^2 f(x)+ \\ +\sum_{i=1}^m\big[\lambda_i+c\,g_i(x)\big]\nabla^2 g_i(x)+ \\ +c\,\mathrm{D}^\top G(x)\,\mathrm{D}G(x),$$

takže volbou $x=x^*$ a $\lambda=\lambda^*$ odtud díky předpokladům našeho tvrzení dostaneme

$$\mathrm{grad}_x L_c(x^*,\lambda^*)= \\ =\mathrm{grad}_x L(x^*,\lambda^*)=0$$

a současně

(i) *Lineární programování*: funkce f je lineární a množina X je vymezena soustavou lineárních rovnic a nerovnic (Kapitola 2).

(ii) *Celočíselné programování*: úloha lineárního programování doplněná o požadavek celočíselných hodnot jednotlivých proměnných ((Kapitola 3).

(iii) *Kvadratické programování*: funkce f je kvadratická forma (plus lineární funkce) a množina X je opět vymezena soustavou lineárních rovnic a nerovnic (Kapitola 4).

Nyní se ještě na chvíli zastavme u funkční optimalizace. V ní je přípustná množina $\mathbb{X}$ prostorem funkcí (skalárních nebo vektorových) n-rozměrné proměnné definovaných na nějaké množině $\Omega\subseteq\mathbb{R}^n$, které v závislosti na dané úloze mohou mít jisté kvalitativní vlastnosti (spojitost, diferencovatelnost apod.) a splňovat další (tzv. vedlejší) podmínky v podobě soustav algebraických, transcendentních, diferenciálních nebo integrálních rovnic či nerovnic. Účelové zobrazení $\mathbb{F}$ pak je *funkcionál*, tj. zobrazení přiřazující každé funkci z přípustné množiny $\mathbb{X}$ nějaké reálné číslo. Takový funkcionál může mít mnoho podob, např. tím nejjednodušším by mohlo být vyčíslení každé funkce z $\mathbb{X}$ v jednom konkrétním bodě z množiny Ω, ale mnohem častější je spíše funkcionál ve tvaru nějakého určitého integrálu. Úloha (1.2.1) pak znamená najít funkci $f\in\mathbb{X}$, pro kterou daný funkcionál nabývá největší/nejmenší hodnotu mezi všemi funkcemi z přípustné množiny. Ukážeme si jeden ilustrativní příklad s původním Eulerovým elegantním řešením. Pro historické podrobnosti a souvislosti viz následující sekci.

Příklad 1.2.21

Mezi všemi funkcemi třídy C^1 na intervalu $[a,b]$ najděme tu, jejíž graf je nejkratší spojnicí bodů $A=[a,\alpha]$ a $B=[b,\beta]$, tj.

$$\mathcal{J}[f]=\int_a^b\sqrt{1+[f'(x)]^2}\,\mathrm{d}x\to\min.$$

Řešení. Nejdříve místo speciálního funkcionálu určujícího délku grafu funkce uvažme zcela obecný funkcionál, tj.

$$\mathcal{J}[f]=\int_a^b F(x,f(x),f'(x))\,\mathrm{d}x$$

pro vhodnou funkci $F:\mathbb{R}^3\to\mathbb{R}$. Rozdělme si nyní uzavřený interval $[a,b]$ na $n+1$ stejně dlouhých podintervalů, tj. podintervalů $[x_i,x_{i+1}]$ pro $i\in\{0,\dots,n\}$ s dělícími body

$$a=x_0<x_1<\cdots<x_n<x_{n+1}=b,$$

takže délka každého takového podintervalu je

$$\Delta x=x_{i+1}-x_i=\frac{b-a}{n+1}.$$

Nahradíme-li graf funkce f lomenou čárou s vrcholy v bodech

$$[x_0,f(x_0)]\quad\&\quad[x_1,f(x_1)]\quad\&\quad\dots\quad\&\quad[x_{n+1},f(x_{n+1})].$$

pak můžeme původní funkcionál

$$\mathrm{J}[f]=\sum_{j=0}^n\int_{x_0+j\Delta x}^{x_0+(j+1)\Delta x}F(x,f(x),f'(x))\,\mathrm{d}x$$

aproximovat konečným součtem

$$\mathcal{J}[f]=\mathcal{J}(f_1,\dots,f_n):=\sum_{j=0}^n F\Big(x_i,f_i,\tfrac{f_{i+1}-f_i}{\Delta x}\Big)\Delta x,$$

který je funkcí n-tice proměnných $f_1:=f(x_1)$, …, $f_n:=f(x_n)$[82]. Jaký bude mít efekt zvýšení či snížení hodnoty některého f_i? Zvolme se jednu z nich a označme ji f_k. Jelikož

člen f_k se v sumě $\mathcal{J}(f_1,\dots,f_n)$ objevuje pouze ve sčítancích s indexy $j=k$ a $j=k-1$, dostaneme parciálním derivováním vzhledem k tomuto f_k výraz

$$\frac{\partial \mathcal{J}}{\partial f_k} = F_f\Big(x_k, f_k, \tfrac{f_{k+1}-f_k}{\Delta x}\Big)\Delta x + F_{f'}\Big(x_{k-1}, f_{k-1}, \tfrac{f_k-f_{k-1}}{\Delta x}\Big) - F_{f'}\Big(x_k, f_k, \tfrac{f_{k+1}-f_k}{\Delta x}\Big), \qquad (1.2.22)$$

kde F_f a $F_{f'}$ značí parciální derivace funkce F podle druhé či třetí proměnné. Pro nalezení extrému bychom pak obvykle položili tuto derivaci rovnu nule pro každé k a následně bychom chtěli udělat limitní přechod pro $n\to\infty$. Jenže v takovém případě je $\Delta x\to 0$ a hodnota pravé strany rovnice (1.2.22) půjde k nule, takže bychom dostali rovnost $0=0$, která je ale stejně pravdivá jako neužitečná. Chceme-li získat trochu méně triviální výsledek, zkusme nejdříve vydělit obě strany rovnosti (1.2.22) výrazem Δx, čímž dostaneme

$$\frac{1}{\Delta x}\frac{\partial \mathcal{J}}{\partial f_k} = F_f\Big(x_k, f_k, \tfrac{f_{k+1}-f_k}{\Delta x}\Big) - \frac{1}{\Delta x}\left[F_{f'}\Big(x_k, f_k, \tfrac{f_{k+1}-f_k}{\Delta x}\Big) - F_{f'}\Big(x_{k-1}, f_{k-1}, \tfrac{f_k-f_{k-1}}{\Delta x}\Big)\right].$$

Odtud nyní pro $n\to\infty$ a $\Delta x\to 0$ získáme tzv. *variační derivaci*

$$\frac{\delta \mathcal{J}}{\delta f} = F_f(x,f,f') - \frac{\mathrm{d}}{\mathrm{d}x}F_{f'}(x,f,f'),$$

která hraje pro funkcionály tutéž roli jako parciální derivace pro funkce více proměnných. Proto má-li být funkce f lokální extrém, očekáváme, že její hodnota bude nulová pro každé x, což nás přivádí k *Eulerově–Lagrangeově* rovnici

$$\frac{\partial F}{\partial f} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial f'}\right) = 0 \qquad (1.2.23)$$

Nyní se vraťme k původnímu problému, kde

$$F(x,f,f') = \sqrt{1+[f'(x)]^2}.$$

Jelikož tento výraz nezávisí na f, dostaneme z (1.2.23) požadavek

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f'}{\sqrt{1+[f'(x)]^2}}\right) = 0 \quad \text{neboli} \quad \frac{f'}{\sqrt{1+[f'(x)]^2}} = \text{konst.}$$

To je možné ale pouze v případě, že

$$f' \equiv C$$

pro libovolné $C\in\mathbb{R}$, a tedy hledaná funkce je nutně tvaru

$$f(x) = Cx + D.$$

Dosazením souřadnic krajních bodů pak nalezneme řešení

$$f(x) = \frac{\alpha-\beta}{a-b}(x-b)+\beta.$$

S variační derivací, Eulerovou–Lagrangeovou rovnicí a vlastně se základy celého variačního počtu se více seznámíme v Kapitole 9. Ale do historie jeho počátků se pustíme už nyní.

Cvičení

Řešené příklady s podobnými úlohami jsou k dispozici i v [255, Sekce III. 6]. Úplná řešení níže uvedených úloh naleznete v [527, Sekce I.1].

1.2.1. Určete lokální extrémy funkce

$$f(x,y) = \sqrt{3}x - y + 2 \quad \text{na množině} \quad M = \{[x,y]\in\mathbb{R}^2 \mid x^2+2x+y^2 = 0\}.$$

$$\nabla_x^2 L_c(x^*,\lambda^*) = \\ = \nabla_x^2 L(x^*,\lambda^*) + \\ + c\,\mathrm{D}^\top G(x^*)\,\mathrm{D}G(x^*).$$

Jelikož $h^\top \nabla_x^2 L(x^*,\lambda^*)\,h > 0$ pro všechna $h\in \mathrm{Ker}\,\mathrm{D}G(x^*)$, plyne odtud, že dokonce $\nabla_x^2 L_c(x^*,\lambda^*) > 0$ pro všechna c dostatečně velká (viz následující poznámku „pod čarou“ číslo 81), takže je splněna postačující podmínka pro ostrý lokální extrém a bod x^* je (volným) ostrým lokálním minimem funkce $L_c(\cdot,\lambda^*)$ pro všechna c dostatečně velká. Proto existují čísla $\gamma,\varepsilon>0$ taková, že

$$L_c(x,\lambda^*) > L_c(x^*,\lambda^*)$$

neboli

$$L_c(x,\lambda^*) \ge L_c(x^*,\lambda^*) + \frac{\gamma}{2}\|x-x^*\|^2$$

pro všechna x splňující $0<\|x-x^*\|<\varepsilon$ či ve druhém případě pouze $\|x-x^*\|<\varepsilon$. Zejména pro body splňující navíc $G(x)=0$ se předchozí nerovnost redukuje na

$$f(x) \ge f(x^*,\lambda^*) + \frac{\gamma}{2}\|x-x^*\|^2,$$

což ukazuje, že x^* je skutečně ostrým vázaným lokálním minimem funkce f na X.

[81] Označíme-li $P = \nabla_x^2 L(x^*,\lambda^*)$ a $Q = \mathrm{D}^\top G(x^*)\,\mathrm{D}G(x^*)$, plyne $\nabla_x^2 L_c(x^*,\lambda^*) > 0$ pro všechna c dostatečně velká z tvrzení, že: *pro libovolnou dvojici symetrických $n\times n$ matic P,Q splňujících $Q\ge 0$ a $h^\top P h > 0$ pro všechna $h\ne 0$ s vlastností $h^\top Q h = 0$, existuje číslo $\check{c}$ takové, že $P + cQ > 0$ pro všechna $c > \check{c}$.*

Důkaz. Připusťme opak, tj. že pro libovolné $k\in\mathbb{N}$ existuje vektor $h^{[k]}$ splňující $\|h^{[k]}\| = 1$ a ${h^{[k]}}^\top (P+kQ)\,h^{[k]} \le 0$. Jelikož posloupnost $\{h^{[k]}\}_{k=1}^\infty$ je ohraničená, existuje konvergentní podposloupnost, takže můžeme bez újmy na obecnosti předpokládat, že samotná $\{h^{[k]}\}_{k=1}^\infty$ konverguje k nějakému h, které musí splňovat $\|h\| = 1$. Limitním přechodem pro $k\to\infty$ odtud dostaneme

$$0 \ge \lim_{k\to\infty} {h^{[k]}}^\top (P+kQ)\,h^{[k]} = \\ = h^\top P h + \lim_{k\to\infty} k\,{h^{[k]}}^\top Q\,h^{[k]}.$$

Pak ale z pozitivní semidefinitnosti Q vyplývá, že nutně $\lim_{k\to\infty} {h^{[k]}}^\top Q\,h^{[k]} = 0$, protože jinak by výraz na pravé straně byl zdola neohraničený. Proto $h^\top Q h = 0$, a tudíž $h^\top P h \le 0$, což je ale spor s uvedenými předpoklady $\lightning$.■

[83] Nezapomínejte, že hodnoty $f_0 := f(x_0) = f(a)$ a $f_{n+1} := f(x_{n+1}) = f(b)$ jsou pevně dány.

Řešení: $P_{\max} = [-1+\sqrt{3}/2, -1/2]$ *a* $P_{\min} = [-1-\sqrt{3}/2, 1/2]$

1.2.2. Určete lokální extrémy funkce

$$f(x,y) = 2x - 3y \quad \text{na množině} \quad M = \{[x,y] \in \mathbb{R}^2 \mid x^2 + y^2 = 25\}.$$

Řešení: $P_{\max} = [10/\sqrt{13}, -15/\sqrt{13}]$ *a* $P_{\min} = [-10/\sqrt{13}, 15/\sqrt{13}]$

1.2.3. Určete lokální extrémy funkce

$$f(x,y) = x^2 + 2xy + y^2 \quad \text{na množině} \quad M = \{[x,y] \in \mathbb{R}^2 \mid 3x^2 + y^2 = 9\}.$$

Řešení: $P_{\max} = [\pm\sqrt{3}/2, \pm 3\sqrt{3}/2]$ *a* $P_{\min} = [\pm 3/2, \mp 3/2]$

1.2.4. Pro $a, b > 0$ určete lokální extrémy funkce

$$f(x,y) = \frac{ax^2}{2} + \frac{by^2}{2} \quad \text{na množině} \quad M = \{[x,y] \in \mathbb{R}^2 \mid x^3 + y^3 = 1\}.$$

Řešení: $P_{\max} = [a/\sqrt[3]{a^3+b^3}, b/\sqrt[3]{a^3+b^3}]$ *a* $P_{\min} = [0,1]$, $P_{\min} = [1,0]$,

1.2.5. Určete lokální extrémy funkce

$$f(x,y) = \ln(1+x^2) - y \quad \text{na množině} \quad M = \{[x,y] \in \mathbb{R}^2 \mid (1+x^2)^2 + y^2 = 4\}.$$

Řešení: $P_{\max} = \left[\pm\sqrt{-1+\sqrt{(\sqrt{17}-1)/2}}, (1-\sqrt{17})/2\right]$ *a* $P_{\min} = [0, \pm\sqrt{3}]$

1.2.6. Určete lokální extrémy funkce

$$f(x,y,z) = (x+y)^2 + (y+z)^2 \quad \text{na množině} \quad M = \{[x,y,z] \in \mathbb{R}^3 \mid x + 2y + 3z = 1\}.$$

Řešení: $P_{\min} = [1/2, -1/2, 1/2]$

1.2.7. Určete lokální extrémy funkce

$$f(x,y,z) = xyz \quad \text{na množině} \quad M = \{[x,y,z] \in \mathbb{R}^3 \mid xy + yz + xz = 3\}.$$

Dokážete výsledek geometricky interpretovat?

Řešení: $P_{\max} = [1,1,1]$ *a* $P_{\min} = [-1,-1,-1]$

1.2.8. Pro $a, b, c > 0$ určete lokální extrémy funkce

$$f(x,y,z) = xyz \quad \text{na množině} \quad M = \left\{[x,y,z] \in \mathbb{R}^3 \;\middle|\; \frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1\right\}.$$

Řešení: $P_{\max} = [a/3, b/3, c/3]$

1.2.9. Určete minimium funkce

$$f(x,y,z) = z - xy^2 \quad \text{na množině} \quad M = \{[x,y,z] \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

Řešení: $P_{\min} = [0,0,-1]$

1.2.10. Určete minimium funkce

$$f(x,y,z) = -x^2y^2z^2 \quad \text{na množině} \quad M = \{[x,y,z] \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 3\}.$$

Řešení: $P_{\min}$ *v bodech* $[\pm 1, \pm 1, \pm 1]$ *s libovolnou kombinací znamének*

1.2.11. Určete minimium funkce

$$f(x,y,z) = -xy^2z^3 \quad \text{na množině} \quad M = \{[x,y,z] \in \mathbb{R}^3 \mid x + 2y + 3z = 12\}.$$

Řešení: $P_{\min} = [2,2,2]$

1.2.12. Pro $a > b > c > 0$ určete lokální extrémy funkce

$$f(x,y,z) = x^2 + y^2 + z^2 \quad \text{na množině} \quad M = \left\{[x,y,z] \in \mathbb{R}^3 \;\middle|\; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\right\}.$$

Dokážete výsledek geometricky interpretovat?

Řešení: $P_{\max} = [\pm a, 0, 0]$ *a* $P_{\min} = [0, 0 \pm c]$

1.2.13. Metodou Lagrangeových multiplikátorů řešte následující úlohu: Do elipsoidu

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

umístěte hranol s maximálním objemem. Tento maximální objem určete. *Řešení:* $V_{\max}=8abc/(9\sqrt{3})$ *s jedním vrcholem v bodě* $P=[a/\sqrt{3},b/\sqrt{3},c/\sqrt{3}]$

1.2.14. Pomocí Lagrangeovy funkce odvoďte vzorec pro výpočet vzdálenosti bodu $x^*=[x_1^*,\dots,x_n^*]\in\mathbb{R}^n$ od nadroviny $a_1x_1+\dots a_nx_n=b$ neboli $\langle a,x\rangle=b$. *Řešení:* $d=|b-\langle a,x^*\rangle|/\|a\|$

1.2.15. Nechť $A\in\mathbb{R}^{m\times n}$. Pomocí Lagrangeovy funkce ukažte, že pro tzv. spektrální maticovou normu

$$\|A\|_\sigma:=\max_{\|x\|=1}\|Ax\|,$$

která je indukovaná Euklidovskou vektorovou normou $\|x\|:=\sqrt{x_1^2+\dots+x_n^2}$, platí

$$\|A\|_\sigma=\sqrt{\lambda_{\max}},$$

kde $\lambda_{\max}$ značí největší vlastní číslo matice $A^\top A$.

1.2.16. Určete lokální extrémy funkce

$$f(x,y,z)=\frac{1}{2}(x^2+y^2+z^2)\quad\text{na množině}\quad M=\{[x,y,z]\in\mathbb{R}^3\mid x-y=0\ \&\ x+y+z=1\}.$$

Řešení: $P_{\min}=[1/3,1/3,1/3]$

1.2.17. Určete lokální extrémy funkce $f(x_1,x_2,x_3,x_4,x_5)=(x_1-1)^2+(x_2-x_3)^2+(x_4-x_5)^5$ na množině

$$M=\left\{x=[x_1,\dots,x_5]\in\mathbb{R}^5\;\middle|\;x_1+x_2+x_3+x_4+x_5=5,\ x_3-2(x_4+x_5)+3=0\right\}.$$

Řešení: $P_{\min}=[1,1,1,1,1]$

1.2.18. Pomocí vrstevnic funkce, jejíž extrémy hledáme, určete největší a nejmenší hodnotu funkce:

a) $f(x,y)=-x-2y$ na množině M vymezené nerovnostmi $x-4y\le4$, $-2x+y\le2$, $-3x+4y\le12$, $2x+y\le8$, $x\ge0$ a $y\ge0$; *Řešení:* $P_{\max}=[0,0]$ *a* $P_{\min}=[20/11,48/11]$

b) $f(x,y)=-x-2y$ na množině M vymezené nerovnostmi $x-4y\le4$, $-2x+y\le2$, $-3x+4y\le12$ $x\ge0$ a $y\ge0$; *Řešení:* $P_{\max}=[0,0]$

c) $f(x,y)=x+y$ na množině M vymezené nerovnostmi $x+2y\le4$, $4x+2y\le12$, $-2x+y\le1$, $x\ge0$ a $y\ge0$; *Řešení:* $P_{\max}=[8/3,2/3]$ *a* $P_{\min}=[0,0]$

d) $f(x,y)=x-y$ na množině M vymezené nerovnostmi $x+y\le6$, $3x+y\le15$, $-x+y\le3$, $x\ge0$ a $y\ge0$; *Řešení:* $P_{\max}=[5,0]$ *a minimum na úsečce* $y=3+x$ *pro* $x\in[0,3/2]$

e) $f(x,y)=x^2-2x+y^2-2y+3$ na množině M vymezené nerovnostmi $x\ge0$, $y\ge0$ a $x+y\le1$; *Řešení:* $P_{\max}=[0,0]$ *a* $P_{\min}=[1/2,1/2]$

f) $f(x,y)=xy$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid x^2+y^2\le1\}$; *Řešení:* $P_{\max}=[\pm\sqrt{2}/2,\pm\sqrt{2}/2]$ *a* $P_{\min}=[\pm\sqrt{2}/2,\mp\sqrt{2}/2]$

g) $f(x,y)=\frac{x^2}{4}+y^2$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid x^2+y^2\le1\}$; *Řešení:* $P_{\max}=[0,\pm1]$ *a* $P_{\min}=[0,0]$

h) $f(x,y)=\frac{x^2}{4}+y^2$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid x^2+y^2=1\}$; *Řešení:* $P_{\max}=[0,\pm1]$ *a* $P_{\min}=[\pm1,0]$

i) $f(x,y)=x^2-2x+y^2+4y+1$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid(x-2)^2+(y-1)^2=1\}$; *Řešení:* $P_{\max}=[2+\sqrt{10}/10,1+3\sqrt{10}/10]$, $P_{\min}=[2-\sqrt{10}/10,1-3\sqrt{10}/10]$

j) $f(x,y)=5x^2+y^2-4y-10x$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid5x^2+y^2=25\}$; *Řešení:* $P_{\max}=[-5/3,-10/3]$, $P_{\min}=[5/3,10/3]$

k) $f(x,y)=5x^2+y^2-4y-10x$ na množině $M=\{[x,y]\in\mathbb{R}^2\mid3x^2+y^2=25\}$, nápověda: polynom $12x^4-60x^3+11x^2+500x-625$ má pouze dva reálné kořeny $x_1\approx1,47$ a $x_2\approx-2,74$; *Řešení:* $P_{\max}\approx[-2,74;1,57]$, $P_{\min}\approx[1,47;4,28]$

l) $f(x,y) = 2x^2 - 6xy + 5y^2 - 16x + 40y$ na množině M vymezené nerovnostmi $x + y \geq 3$, $2x - y \geq 2$, $x \geq 0$ a $y \geq 0$; *Řešení:* $P_{\min} = [4,0]$

m) $f(x,y) = x^2 + 3xy - y^2$ na množině $M = \{[x,y] \in \mathbb{R}^2 \mid x + y = 3\}$; *Řešení:* $P_{\max} = [5/2, 1/2]$

n) $f(x,y,z) = x - 2y + 2z$ na množině $M = \{[x,y,z] \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. *Řešení:* $P_{\max} = [1/3, -2/3, 2/3]$, $P_{\min} = [-1/3, 2/3, -2/3]$

1.3 První (ne)optimalizo(tel)ný historický exkurz

Jak už jsme se zmínili, skutečně systematickému studiu různých optimalizačních problémů a rozhodovacích procesů se matematici začali věnovat až ve druhé polovině 20. století. Toto ovšem neplatí pro úlohy variačního počtu, jejichž teorie byla vybudována mnohem dříve. Abychom se seznámili s největšími osobnostmi, které zásadně ovlivnili rozvoj různých oblastí optimalizace, a pochopili některé souvislosti či jejich motivaci, uděláme si v závěrečné části této kapitoly malý historický exkurz, který však bude sledovat spíše časovou souslednost nežli váhu jednotlivých témat v dalších kapitolách.

Před chvílí jsme viděli původní Eulerovo řešení pro nalezení nejkratší křivky v rovině. S tímto problémem také úzce souvisí tzv. *Fermatův princip*, který vznikl při řešení problému geometrické optiky v roce 1662. Podle tohoto principu si světelný paprsek ze všech možných trajektorií mezi dvěma body vždy vybírá právě tu, podél níž se dostane z výchozího bodu do cílového bodu za nejkratší čas (ve skutečnosti spíše za „extrémní čas“, tj. po jakékoli jiné dráze by to vždy trvalo delší/stejnou dobu nebo vždy kratší/stejnou dobu). Jako obvykle i zde si příroda najde to „nejekonomičtější“ řešení – vždyť čas jsou peníze. Máme-li opticky nehomogenní rovinu, pak rychlost světla $c(x,y)$ se mění v každém bodě v závislosti na optických vlastnostech. Protože rychlost je derivace dráhy podle času, je délka křivky $y = u(x)$ určené světelným paprskem rovna

$$c(x,u(x)) = \frac{\mathrm{d}s}{\mathrm{d}t} = \sqrt{1 + [u'(x)]^2}\,\frac{\mathrm{d}x}{\mathrm{d}t}.$$

Integrováním od počátku do cíle pak získáme celkový čas pohybu podél této křivky, tj.

$$T[u] = \int_0^T \mathrm{d}t = \int_a^b \frac{\mathrm{d}t}{\mathrm{d}x}\,\mathrm{d}x = \int_a^b \frac{\sqrt{1 + [u'(x)]^2}}{c(x,u(x))}\,\mathrm{d}x.$$

Fermatův princip (nebo též *princip nejkratšího času*) pak říká, že z bodu $A = (a,\alpha)$ do bodu $B = (b,\beta)$ se světelný paprsek dostane podél křivky $y = u(x)$, která minimalizuje funkcionál $T[u]$ vzhledem k okrajovým podmínkám

$$u(a) = \alpha \quad \& \quad u(b) = \beta.$$

Jestliže je prostředí homogenní (např. vakuum při zanedbání vlivu gravitace vzhledem k teorii relativity), pak $c(x,y) \equiv c$ a $T[u] = \frac{1}{c}\,\ell[u]$, tedy hledaným minimem bude opět vhodná přímka. V případě nehomogenního prostředí již řešení tak zřejmé není a bylo by vhodné mít systematickou metodu pro řešení takového minimalizačního problému. Navíc všechny známé zákony optiky, designu optických čoček, ostření, lámání, odchylky atd. jsou vlastně důsledkem geometrických a analytických vlastností řešení Fermatova principu. Tato úvaha inspirovala vědce k hypotéze, že každý děj v přírodě probíhá tak, aby se během tohoto procesu minimalizovala jistá veličina. Ovšem dlouho se nevědělo, o kterou veličinu se jedná. Až v roce 1760 Joseph Lagrange poprvé přesně zformuloval *princip nejmenší akce* pro konzervativní mechanické systémy a vymezil platnost tohoto principu. Tento princip zobecnil Rowan Hamilton[83] i pro systémy nekonzervativní, kdy transformoval Lagrangeovy rovnice (které jsou diferenciálními rovnicemi 2. řádu) na soustavu diferenciálních rovnic prvního řádu s dvojnásobným počtem rovnic, tzv. *Hamiltonovy kanonické rovnice*, jejichž význam daleko přesahuje rámec klasické mechaniky. V Hamiltonově práci pokračoval Carl Jacobi[84], který rozvinul transformační teorii kanonických rovnic a odstranil řadu obtíží spojených s jejich integrací.

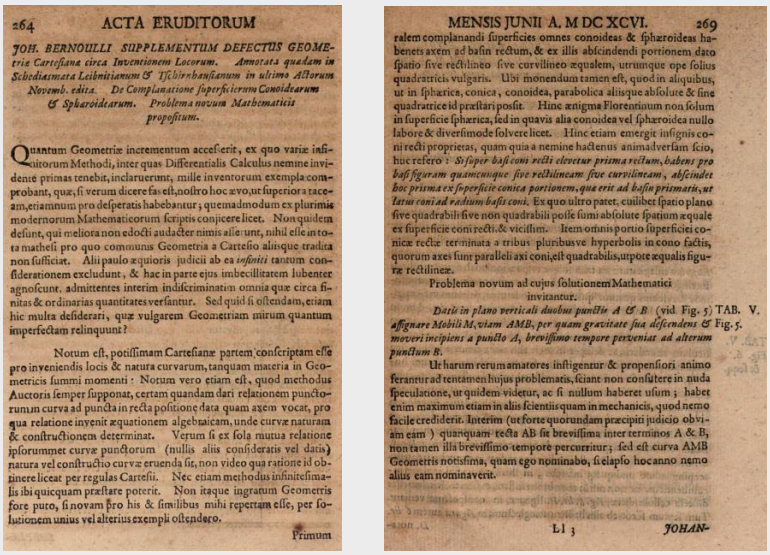
[83] Sir William Rowan Hamilton (4. srpna 1805 – 2. září 1865) byl irský matematik, profesor astronomie na Trinity College v Dublinu, a královský astronom. Věnoval se především optice, klasické mechanice a algebře. Ačkoli nebyl fyzik (sám se považoval za čistého matematika), svoji prací ovlivnil především právě fyziku, a to zejména přeformulováním newtonovské mechaniky, která je nyní spojována především s jeho jménem. Je autorem algebraického přístupu ke komplexním číslům, čímž snad definitivně vyřešil hádanku ohledně významu symbolu i. Tato práce jej přivedla k jeho druhému klíčovému výsledku – zavedení kvaternionů. Mimochodem, Hamilton údajně uměl již v 5 letech latinsky, řecky a hebrejsky a v 9 letech přidal ještě perštinu, arabštinu a sanskrt. Toto skotsko-americký matematik a autor sci-fi Eric Temple Bell trefně okomentoval slovy: „Dobrý Bože! K čemu?!“

[84] Carl Gustav Jacob Jacobi (10. prosince 1804 – 18. února 1851) byl německý matematik, který zásadně přispěl k rozvoji teorie eliptických funkcí, diferenciálních rovnic, determinantů a teorie čísel. Jacobi byl prvním židovským matematikem, který byl jmenován profesorem na německé univerzitě.

[85] Tento název pochází z řečtiny, kde *brachistos* je superlativ od *brachys*=krátký a *chronos*=čas. Leibniz navrhoval označení *Tachystoptotam* z řeckého *tachystos*=nejrychlejší a *piptein*=padat. Tento návrh se ale (naštěstí) nesetkal s velkým pochopením.

Avšak asi tu nejdůležitější roli v rozvoji variačního počtu sehrála *úloha o brachistochroně*[85]. V roce 1696 se Johann Bernoulli[86] v červnovém vydání časopisu Acta Eruditorum snažil ukázat, že právě vznikající diferenciální počet je tím nutným a postačující nástrojem k vyplnění mezer v klasické geometrii, což na závěr doplnil neodmítnutelnou výzvou:

> *Problema Novum ad cujus Solutionem Mathematici invitantur.*
>
> *Datis in plano verticali duobus punctis A & B assignare Mobili M, viam AMB, per quam gravitate sua descendens & moveri incipiens a puncto A, brevissimo tempore perveniat ad alterum punctum B.*[87]



264 ACTA ERUDITORUM

JOH. BERNOULLI SUPPLEMENTUM DEFECTUS GEOMETRIE Cartesiana circa Inventionem Locorum. Annotata quadam in Schediasmata Leibnitianum & Tschirnhausianum in ultimo Actorum Novemb. edita. De Complanatione superficierum Conoidearum & Sphæroidearum. Problema novum Mathematicis propositum.

Quantum Geometriæ incrementum acceserit, ex quo variæ infinitorum Methodi, inter quas Differentialis Calculus nemine invidente primas tenebit, inclaruerunt, mille inventorum exempla comprobant, quæ, si verum dicere fas est, nostro hoc ævo, ut superiora taceam, etiamnum pro desperatis habebantur; quemadmodum ex plurimis modernorum Mathematicorum scriptis conjicere licet. Non quidem desunt, qui meliora non edocti audacter nimis alleguunt, nihil esse in tota matheſi pro quo communis Geometria a Carteſio aliisque tradita non ſufficiat. Alii paulo æquioris judicii ab ea *infiniti* tantum conſiderationem excludunt, & hac in parte ejus imbecillitatem lubenter agnoſcunt, admittentes interim indiſcriminatim omnia quæ circa finitas & ordinarias quantitates verſantur. Sed quid ſi oſtendam, etiam hic multa deſiderari, quæ vulgarem Geometriam mirum quantum imperfectam relinquunt?

Notum eſt, poſſiſſimam Carteſianæ partem conſcriptam eſſe pro inveniendis locis & natura curvarum, tanquam materia in Geometricis ſummi momenti: Notum vero etiam eſt, quod methodus Auctoris ſemper ſupponat, certam quandam dari relationem punctorum curvæ ad puncta in recta poſitione data quam axem vocat, pro qua relatione invenit æquationem algebraicam, unde curvæ naturam & conſtructionem determinat. Verum ſi ex ſola mutua relatione ipſorummet curvæ punctorum (nullis aliis conſideratis vel datis) natura vel conſtructio curvæ eruenda ſit, non video qua ratione id obtinere liceat per regulas Carteſii. Nec etiam methodus infinitesimalis ibi quicquam præſtare poterit. Non itaque ingratum Geometris fore puto, ſi novam pro his & ſimilibus mihi repertam eſſe, per ſolutionem unius vel alterius exempli oſtendero.

Primum

MENSIS JUNII A. M DC XCVI. 269

ralem complanandi ſuperficies omnes conoideas & ſphæroideas habenets axem ad baſin rectum, & ex illis abſcindendi portionem dato ſpatio ſive rectilineo ſive curvilineo æqualem, utrumque ope ſolius quadratricis vulgaris. Ubi monendum tamen eſt, quod in aliquibus, ut in ſphærica, conica, conoidea, parabolica aliisque abſolute & ſine quadratrice id præſtari poſſit. Hinc ænigma Florentinum non ſolum in ſuperficie ſphærica, ſed in quavis alia conoidea vel ſphæroidea nullo labore & diverſimode ſolvere licet. Hinc etiam emergit inſignis conireƈti proprietas, quam quia a nemine hactenus animadverſam ſcio, huc refero: *Si ſuper baſi coni recti elevetur prisma rectum, habens pro baſi figuram quamcunque ſive rectilineam ſive curvilineam, abſcindat hoc prisma ex ſuperficie conica portionem, quæ erit ad baſin prismatis, ut latus coni ad radium baſis coni.* Ex quo ultro patet, cuilibet ſpatio plano ſive quadrabili ſive non quadrabili poſſe ſumi abſolutæ ſpatium æquale ex ſuperficie coni recti & viciſſim. Item omnis portio ſuperficiei conicæ rectæ terminata a tribus pluribusve hyperbolis in cono factis, quorum axes ſunt paralleli axi coni, eſt quadrabilis, utpote æqualis figuræ rectilineæ.

Problema novum ad cujus ſolutionem Mathematici invitantur.

Datis in plano verticali duobus punctis A & B (vid. Fig. 5) TAB. V. *aſſignare Mobili M, viam AMB, per quam gravitate ſua deſcendens* & Fig. 5. *moveri incipiens a puncto A, breviſſimo tempore perveniat ad alterum punctum B.*

Ut harum rerum amatores inſtigentur & propenſiori animo ferantur ad tentamen hujus problematis, ſciant non conſiſtere in nuda ſpeculatione, ut quidem videtur, ac ſi nullum haberet uſum; habet enim maximum etiam in aliis ſcientiis quam in mechanicis, quod nemo facile crediderit. Interim (ut forte quorundam præcipiti judicio obviam eam) quanquam recta AB ſit breviſſima inter terminos A & B, non tamen illa breviſſimo tempore percurritur; ſed eſt curva AMB Geometris notiſſima, quam ego nominabo, ſi elapſo hoc anno nemo alius eam nominaverit.

L l 3 JOHAN-

Obrázek 1.9: Úvodní strana Bernoulliova článku (vlevo) a úloha o brachistochroně (vpravo).

Zároveň aby Johann Bernoulli přitáhl větší pozornost k této úloze, ihned vzápětí zdůrazňuje, že řešení úlohy není úplně zřejmé a že rozhodně hledání *křivky nejkratšího času* není totéž jako hledání nejkratší spojnice bodů A a B:

> *Ut harum rerum amatores instigentur & propensiori animo ferantur ad tentamen hujus problematis, sciant non consistere in nuda speculatione, ut quidem videtur, ac si nullum haberet usum; habet enim maximum etiam in aliis scientiis quam in mechanicis, quod nemo facile crediderit. Interim (ut forte quorundam praecipiti judicio obviam eam) quanquam recta AB sit brevissima inter terminos A & B, non tamen illa brevissimo tempore percurritur; sed est curva AMB Geometris notissima, quam ego nominabo, si elapso hoc anno nemo alius eam nominaverit.*[88]

Johann Bernoulli také stanovil lhůtu na řešení do 31. prosince 1696, což se ale ukázalo být nedostatečné (nutno však říci, že spíše z důvodů nematematických – vzpomeňte si na Mersenna). Proto byla tato úloha krátce zopakována v lednovém vydání časopisu Acta Eruditorum, viz [531, str. 95–96], a další podrobnosti se můžeme dozvědět z Bernoulliho veřejného prohlášení v Groningenu z ledna 1697, jehož pojetí jakoby nás vracelo zpět do časů slavných matematických soubojů, který svedli např. del Fiore a Tartaglia:

> *Jean Bernoulli public professor of mathematics pays his respects to the most acute mathematicians of the entire world.*
>
> *Since it is known with certainty that there is scarcely anything which more greatly excites noble and ingenious spirits to labors which lead to the increase*

[86]

Johann Bernoulli (6. srpna 1667 – 1. ledna 1748) byl švýcarský matematik, fyzik a lékař. Je jedním z prvních členů matematické větve rodu Bernoulliů, která výrazně ovlivnila rozvoj matematiky na přelomu 17. a 18. století. Byl učitelem Leonharda Eulera, jehož matematický rozvoj (zejména v mládí) byl velmi úzce spojen s několika členy rodiny Bernoulliů. Při studiu medicíny na univerzitě se jako jeden z prvních naučil ovládat Leibnizův kalkulus. V roce 1691 odcestoval do Paříže, kde se seznámil l'Hospitalem, se kterým následně uzavřel smlouvu o tom, že jej za určitý roční honorář bude učit novým Leibnizovým metodám. Současně mu ale touto dohodou vzniklo právo užívat Bernoulliho objevy. „Své" poznatky pak l'Hospital publikoval v roce 1696 v první učebnici infinitesimálního počtu *Analyse des Infiniment Petits pour l'Intelligence des Lignes Courbes*, ve které se objevilo i známé pravidlo pro výpočet limity podílu, které dnes nese l'Hospitalovo jméno. Proti tomu se samozřejmě Johann Bernoulli ohradil, neboť jeho jméno nebylo nikde uvedeno. Veřejně své námitky prezentoval zejména až po l'Hospitalově smrti v roce 1704. Nicméně důkazů pro své tvrzení příliš neměl, a tak pravda vyšla najevo až v roce 1922.

[87] *Nová úloha, k jejímuž řešení se vyzývají matematikové.*

Nechť jsou dány dva body A a B ve svislé rovině. Určete křivku vyznačenou hmotným bodem, který startuje v bodě A a pouhým působením gravitace dosáhne bodu B v nejkratším možném čase.

(New Problem Which Mathematicians Are Invited.

If two points A and B are given in a vertical plane, to assign to a mobile particle M the path AMB along which, descending under its own weight, it passes from the point A to the point B in the briefest time.) Viz [530, str. 269].

> *of knowledge than to propose difficult and at the same time useful problems through the solution of which, as by no other means, they may attain to fame and build for themselves eternal monuments among posterity; so I should expect to deserve the thanks of the mathematical world if, imitating the example of such men as Mersenne, Pascal, Fermat, …and others who have done the same before me, I should bring before the leading analysts of this age some problem upon which as upon a touchstone they could test their methods, exert their powers, and, in case they brought anything to light, could communicate with us in order that everyone might publicly receive his deserved praise from us. The fact is that a half a year ago in the June number of the Leipzig Acts [i.e., Acta Eruditorum] proposed such a problem whose usefulness linked with beauty will be seen by all who successfully apply themselves to it… Only the celebrated Leibniz who is so justly famed in the higher geometry has written me that he has by good fortune solved this, as he himself expresses it, very beautiful and hitherto unheard of problem; and he has courteously asked me to extend the time limit to next Easter in order that in the interim the problem might be made public in France and Italy and that no one might have cause to complain of the shortness of time allotted.*[89]

Poté Bernoullli poslal kopii tohoto problému Newtonovi, aby měl jistotu, že o něm ví. Zde je potřeba samozřejmě připomenout vleklý spor mezi Leibnizem a Newtonem ohledně prvenství v základech infinitesimálního počtu, do kterého vstoupil také Bernoulli, neboť v průběhu roku 1696 vyjádřil své přesvědčení, že Newtonova metoda publikovaná ve Wallisově *Opera mathematica* byla „štípnuta" z Leibnizova článku. Bernoulli i Leibniz tak očividně interpretovali ticho od června do prosince jako demonstraci, že tento problém byl pro Newtona záhadou. Prodloužením lhůty měli asi v úmyslu veřejně demonstrovat svou nadřazenost. Ať už si Bernoulli a Leibniz mysleli cokoli, Newton tuto výzvu považoval za namířenou osobně proti němu. No, kdyby to nebylo úplně zřejmé z přímé korespondence, zakomponoval Bernoulli do svého vyjádření i tato slova:

> *Qpod si vero festum paschatis praeterierit nemine deprehenso qui quaesitum nostrum solverit, nos quae ipsi invenimus publico non invidebimus. Incomparabihs enim Leibnitius solutiones tura suam tura nostram ipsi jam pridem commissam protinus, ut spero, in lucem emittet, quas si Geometrae ex penitiori quodam fonte petitas perspexerint, nulli dubitamus quin angustos vulgaris Geometriae limites agnoscant, nostraque proin inventa tanto pluris faciant, quanto pauciores eximiam nostram quaestionem soluturi extiterint etiam inter illos ipsos qui per singulares quas tantopere commendant methodos, interioris Geometriae latibula non solum intime penetrasse, sed etiam ejus pomoeria Theorematis suis aureis, nemini ut putabant cognitis, ab aliis tamen jam longe priiis editis, mirum in modum extendisse gloriantur.*[90]

Zjevně se to Newtonovi vrylo velmi hluboko do paměti, neboť o dva roky později se při jedné z debat s prvním královským astronomem Johnem Flamsteedem na něj ohradil slovy

> *I do not love… to be dunned & teezed by forreigners about Mathematical things…*[91]

Newton výzvu přijal v roce 1697, když si na dopis poznamenal datum doručení: 29. ledna 1697. Už následující den odeslal řešení Charlesi Montagueovi, tehdejšímu prezidentu Královské společnosti, a bylo anonymně zveřejněno ještě v lednovém vydání *Philosophical Transactions of the Royal Society*. Bernoullimu samozřejmě nedalo příliš práce určit „tajného" autora se slovy

[88] *Pro povzbuzení touhy k ujmutí se řešení tohoto problému u těch, kteří takové úlohy milují, je možné zdůraznit, že předložený problém nesestává, jak by se mohlo zdát, z pouhých spekulací, které tudíž nemají žádné využití. Naopak, jak bychom mohli snadno uvěřit, má velký užitek v různých vědních disciplínách jako je mechanika. Mezitím (abychom předešli ukvapenému soudu) [můžeme poznamenat] že ačkoli úsečka AB je vskutku nejkratší spojnicí bodů A a B, rozhodně není křivkou s nejkratším časem pohybu. Avšak křivka AMB, jejíž jméno uvedu já, pokud ji nikdo neobjeví před koncem tohoto roku, je velmi dobře známa geometrům.*
(To arouse in lovers of such things the desire to undertake the solution of this problem, it may be pointed out that the question proposed does not, as might appear, consist of mere speculation having therefore no use. On the contrary, as one would readily believe, it has great usefulness in other branches of science such as mechanics. Meanwhile (to forestall hasty judgment) [it may be remarked that] although the straight line *AB* is indeed the shortest between the points *A* and *B* it nevertheless is not the path traversed in the shortest time. However the curve *AMB* whose name I shall give if no one else has discovered it before the end of this year, is one well known to geometers.)

[89] *Já, Johann Bernoulli, oslovuji nejbystřejší matematiky na světě.*

Jelikož je jistě známo, že není nic přitažlivějšího pro inteligentní lidi něž čestný a náročný problém, jehož jakékoli řešení přinese slávu a zůstane trvalým památníkem. Následujíce řady příkladů Mersenna, Pascala, Fermata,… doufám, že získám vděčnost celé vědecké komunity tím, že před nejlepší matematiky naší doby postavím problém, který otestuje jejich metody a sílu jejich intelektu. Jestliže mi kdokoli sdělí řešení navrhovaného problému, veřejně prohlásím, že je hoden uznání. Ve skutečnosti jsem takový problém, jehož užitečnost spojenou s krásou uvidí každý, kdo se zdárně odhodlají k jeho řešení, zveřejnil před půl rokem v červnovém vydání Acta Eruditorum… Pouze slovutný Leibniz, který je po právu věhlasný vyšší geometrií, mi napsal, že se mu poštěstilo nalézt řešení tohoto, jak se sám vyjádřil, překrásného a doposud nevídaného problému. Zdvořile mě požádal o prodloužení lhůty pro řešení do následujících Velikonoc, aby se mezitím problém mohl rozšířit do Francie a Itálie, takže by nikdo neměl důvod

si stěžovat na nedostatečný čas určený k řešení. Viz [455, str. 646–647].

[90] *Pokud ovšem ani do velikonočních oslav se nenajde nikdo, kdo by byl schopen náš vznešený problém vyřešit, nebudeme se zdráhat zveřejnit to, co jsme sami objevili. Doufám, že jedinečný Leibniz neprodleně zveřejní své i mé řešení, které jsem mu před časem svěřil. Pokud geometři pečlivě prozkoumají tato řešení získaná tak, jak jsou, z toho, co můžeme nazývat hlubokou studnicí (myšlenek), nepochybujeme o tom, že rozpoznají úzké hranice běžné geometrie a ocení naše objevy o to více, že je jen několik těch, kteří jsou schopni vyřešit naše skvělé problémy, ba dokonce, malý počet mezi těmi nejlepšími matematiky, kteří se chlubí, že svými pozoruhodnými metodami, za než jsou chvályhodní, nejenže pronikli hluboko do tajemství esoterické geometrie, ale úžasně rozšířili její hranice díky zlatým větám, které (jak se domnívají) nebyly známy nikomu jinému, avšak které byly ve skutečnosti publikovány již dávno předtím jinými.*
(If however the festival of Easter shall have passed without our having found any who are capable of solving our noble problem, we shall not grudgingly refuse to the public what we have ourselves discovered. For the peerless Leibniz will forthwith, I hope, publish the solutions, his own and mine, which I entrusted to him some time ago. If geometers carefully examine these solutions, drawn as they are from what may be called a deeper well (of thought) we are in no doubt but that they will recognise the narrow limits of the common geometry, and will value our discoveries so much the more as there are fewer who are likely to solve our excellent problems, aye, the small number even among the very mathematicians who boast that by the remarkable methods they so greatly commend, they have not only penetrated deeply the secret places of esoteric geometry but have also wonderfully extended its bounds by means of the golden theorems which (they thought) were known to no one, but which in fact had long previously been published by others.) Viz [442, str. 220–229]

[91] *Nemám rád… když jsem uháněn a popichován cizinci ohledně matematických záležitostí…*

[92] *Podle drápů se pozná lev.*
(We may judge the lion from its claw).

tanquam ex ungue leonem[92]. Newtonův pocit vítězství byl tak velký, že se tento příběh dostal prostřednictvím jeho neteře Catherine dostal až do Conduittovi sbírky anekdot:

> *When the problem in 1697 was sent by Bernoulli – Sir I. N. was in the midst of the hurry of the great recoinage did not come home till four from the Tower very much tired, but did not sleep till he had solved it which was by 4 in the morning.*[93]

Avšak nezanedbatelnou kaňkou na Newtonově úspěchu je fakt, že v článku chybí podrobnější zdůvodnění získaného řešení.

Měli bychom zmínit, že Johann Bernoulli rozhodně není prvním matematikem, který se zabýval takovýmto problémem. Již v roce 1638 se mu věnoval také Galileo Galilei ve své slavné práci *Matematické rozpravy a pokus* (Discourses and Mathematical Demonstrations Relating To Two New Sciences). Nejdříve požadoval bod B pouze na přímce a hledal vhodnou úsečku. Správným řešením je úsečka svírající s přímkou b úhel 45°. Galileo poté také určil (správně), že kratší čas bude dosažen při pohybu po dvou úsečkách AC, BC, kde C je bod ležící na oblouku kružnici procházející body A, B. Ačkoli Galileovy úvahy byly doposud správné, vyvodil odtud, že nejrychlejší bude pohyb po kružnici. Jenže ona existuje ještě rychlejší dráha. Leibniz údajně předpověděl (téměř správně), že existuje pouze pět geometrů, kteří dokáží danou úlohu vyřešit. Měl na mysli Jacoba Bernoulliho, 1'Hospitala, Huygense (kdyby žil), Huddeho (kdyby toho nezanechal) a Newtona (kdyby se takové výzvy chopil). Jednotlivá řešení pak byla zveřejněna v květnovém vydání časopisu Acta Eruditorum z roku 1697 a jejich autory byli:

(0) Johann Bernoulli (on sám řešení znal již v době publikování problému, pouze vyzýval ostatní k jeho řešení), viz [531, str. 206–211];

(1) Jacob Bernoulli, viz [531, str. 211–214];

(2) Isaac Newton, viz [531, str. 223], což je přetisk výše zmíněného „anonymního" článku;

(3) Guillame l'Hospital, viz [531, str. 217–218];

(4) Gottfried Leibniz, viz [531, str. 201–205];

(5) Ehrenfried Walther von Tschirnhaus (ten bývá v této souvislosti velmi často opomíjen!), viz [531, str. 220–223];

Hledanou křivkou byla jedna z těch, které již tehdy byly *velmi dobře známy všem geometrům* – otočená *cykloida* (jméno jí dal v roce 1599 Galileo a objevuje se dokonce v prvním díle Gulliverových cest[93]od anglického spisovatele Jonathana Swifta, který vyšel v roce 1726). Tato křivka byla studována již Pascalem a Huygensem, avšak nikdo z nich netušil, že je to křivka nejrychlejšího spádu. Johann Bernoulli toto okomentoval s nadhledem sobě vlastním

> *… you will be petrified with astonishment when I say that precisely this cycloid… of Huygens is our required brachistochrone.*[94]

Tato úloha vzbudila velký zájem o podobné problémy, kdy se nehledaly extrémy funkcí, ale samotná extremální funkce (ovšem pomocí analytických metod – nikoli geometricky). Dodejme ještě, že cykloida je také současně *tautochronou*, tj. ať kuličku vypustíme z jakéhokoli bodu na brachistochroně mezi A a B, pak dorazí do nejnižšího bodu na dráze vždy za stejný čas.

Navzdory nezměrnému historickému významu úlohy o brachistochroně je nutné zdůraznit, že to rozhodně nebyla jedna z prvních úloh tohoto typu. Již ve starověku se řešili tzv. *isoperimetrické problémy*, jejichž cílem bylo mezi všemi uzavřenými rovinnými křivkami nalézt takovou, která ohraničuje plochu s maximálních obsahem (např. máme-li 10 metrů drátu a chceme s jeho pomocí vymezit největší výběh pro králíky). Nejznámější takovou úlohou je *Problém královny Didó*, díky kterému je královna Didó nedílnou součástí úvodních kapitol drtivé většiny knih věnovaných variačnímu počtu. O této úloze víme díky eposu *Aeneis*, jehož autorem

je římský básník Publius Vergilius Maro. Tento epos vznikl v letech 29–19 př.n.l. a je psán hexametrem (tj. nerýmovaným veršem, zde v překladu K. Vinařického z roku 1851):

> *Punské říše vidíš, Tyrského Agénora město! toť meze Libyčanů, lidu nevlídného ve válce. Z města Tyru přišlá království Didó to řídí, bratrem zahnána. Dlouhá je to křivda a dlouhé okliky; já tedy jen co je hlavnějšího, podotknu. Choť jí byl Sycheys, Foeníčan nejbohatější statkem a od nebohé milován ze celého to srdce panny, jemuž ji otec vydal a v manželstvo prvotní zasnoubil; bratr než se ujal Tyrského vladařství Pygmalion, zločinec horší nade všecky zločince. Rozdvojilo vzteklé oba záští, tak že ve chrámě bezbožný Sychéa tyran, zaslepen zlata vášní, tajně a náhle zabil dýkou, ani lásky neváživ sestřiny; jí pak dlouho tajil to zločinstvo a marnou přetvářen zlostník nadějí manželku klamával. Pohřebu však zbavený ve snách jí sám se ukázal manžel, a vybledlá otvíraje ústa podivně, ukrutný oltář, ocelí proťatá pak i ňádra odhalil a projevil vše tajemství zlosti domácí. K neprodlenému posléz rady z otčiny odchodu dává a v prospěch té pouti pravetché vymkne ze země poklady, neznámé nikomu zlata břímě a stříbra. Pohnula tím chystá i hledá druhy k outěku Didó. Scházeli se, kdo nenáviděli zlostného tyranna neb kdo se báli jeho; zmocnivše se právě korábů ustrojených, naloží je zlatém a statky lakomce Pygmalióna mořem unesou; činu ženština vůdcem. Do kraje přišli polom, kdež náramné teď uhlédáš ohrady a pnoucí se nové Kartháginy bašty. Koupili místo také, a k paměti nazvali Byrsou, tak veliké, co koží volovou zahrnouti se mohlo.*

„Modernější“ (a lépe srozumitelná) verze této legendy by mohla znít takto:

> *Po tom, co manžela a současně strýce fénické princezny Ellisy, později nazývané královnou Didó, zavraždil její bratr Pygmalion, musela sama uniknout a hledat bezpečí. S lidmi, kteří ji byli ochotni následovat, přistála se svou lodí na severoafrickém pobřeží — na území, které se nacházelo v dnešní blízkosti Tunisu — a požádala Iarbase, tehdejšího krále Numidie, o azyl. Iarbas rozhodl, že Ellise daruje území – jeho nabídka ale nebyla příliš velkorysá, neboť své území nové území musela ohraničit jedinou býčí kůží. Doslova jí řekl „Užívej pozemek tak velký, jaký vymezí volská kůže“.*

Ellisa dala rozřezat kůži na velmi tenké pásky, viz Obrázek 1.10, které dohromady utvořili pruh dlouhý 4 km. Tímto pruhem pak nechala vymezit území o tvaru půlkruhu uzavřeného z jedné strany pobřežím. Toto území ohraničené hradbami se nazývalo *Byrsa* (což v překladu znamená „býčí kůže“) a tvořilo centrum Kartága, které vzniklo 72 let před Římem, tj. 825 př.n.l. (někdy se uvádí rok 814 př.n.l). Didó tak vyřešila isoperimetrický problém, neboť se jí podařilo pro své lidi získat největší možné území, viz Obrázek 1.11.

Obrázek 1.10: *Didó a býčí kůže* od italského malíře Gregoria Lazzariniho (1655–1730).

Po zániku antické civilizace nastává dlouhé období vědecké stagnace a úpadku, nicméně i středověcí stavitelé museli vyřešit analogický problém vzhledem k náročnosti výstavby hradebního opevnění. Což se jim mnohdy povedlo, jak můžeme vidět na historických mapách Paříže, Milána a Frankfurtu nad Mohanem níže.

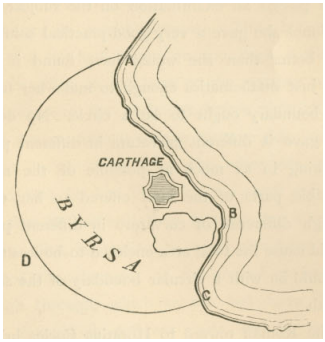
Řešení isoperimetrické úlohy bylo známo již Aristotelovy (cca 384–322 př.n.l.), který věděl, že mezi rovinnými obrazci o stejném obvodu má největší obsah kruh a mezi tělesy stejného povrchu má největší objem koule. Ve starověku lze nalézt i další podobné úlohy s geometrickým základem, např.

(i) *Apollonius* (3. století př.n.l.): jak vést z daného bodu k dané kuželosečce nejkratší a nejdelší úsečku tak, aby jeden její krajní bod byl tím zadaným a druhý krajní bod ležel na dané kuželosečce?

[93] *Když byl ten problém Bernoullim odeslán v roce 1697, Sir I. N. byl zrovna uprostřed shonu kvůli velké měnové reformě a vrátil se domů z londýnského Toweru [kde byla Královská mincovna] velmi unavený až po 16. hodině. Ovšem nešel spát, dokud problém nevyřešil, což se mu povedlo kolem čtvrté hodiny ranní.* Viz [508, str. 582].

[93] *Pokud jsem dovedl posouditi, dával mi král různé otázky a já jsem jej zase oslovoval všemi jazyky, jež jsem znal. Když se ukázalo, že nerozumím ani já jim, ani oni mně, zavedli mě na králův rozkaz do jedné komnaty v paláci (tento panovník totiž vynikal pohostinstvím nad všechny své předchůdce), kde mi byli dáni k obsluze dva sluhové. Přinesli mi oběd a čtyři šlechtici, jež jsem předtím viděl v těsné blízkosti krále, mi prokázali tu čest, že obědvali se mnou. Měli jsme dva chody, každý o třech jídlech. První chod záležel ve skopovém rámínku vykrojeném do rovnoramenného trojúhelníku, z kusu hovězího, připraveného v podobě kosodélníku, a z puddingu v podobě cykloidy.*

[94] *… budete zkamenělí úžasem, když řeknu, že právě tato Huygensova cykloida… je hledanou brachistochronou.* Viz [147, str. 201].

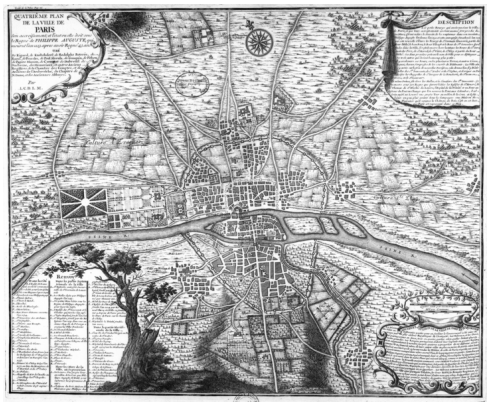


Obrázek 1.11: Situační obrázek městy Byrsa a Tuniského zálivu ve Středozemním moři. Obrázek pochází z přednášky William Thomsona[95] s názvem *Isoperimetrické problémy*, kterou přednesl 12. května 1893 na konferenci Královské společnosti v Londýně, viz [479] a také [16]. Byl to asi právě tento příspěvek lorda Kelvina, který umožnil královně Didó vstup do historie matematiky a učinil z ní nedílnou součást mnoha knih.

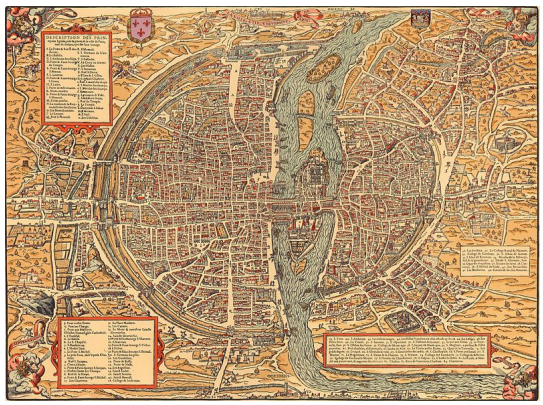
[95] William Thomson (26. června 1824 – 17. prosince 1907) je známý spíše pod svým šlechtickým jménem lord Kelvin (of Largs). Byl irsko-skotský matematický fyzik a vynálezce. Je považován za dovršitele mechaniky ve fyzice.

(ii) *Euklides z Alexandrie* (cca 340–270 př.n.l.): jak vepsat do daného trojúhelníku rovnoběžník maximálního obsahu?

(iii) *Archimédes ze Syrakus* (cca 278–212 př.n.l.): ze všech kulových úsečí stejného povrchu určit tu s maximálním objemem.



Obrázek 1.12.a: Paříž (rok 1223).



Obrázek 1.12.b: Paříž (rok 1575).

Obrázek 1.13.a: Milán (rok 1572).

Obrázek 1.13.b: Frankfurt nad Mohanem (rok 1572).

Později byly isoperimetrické úlohy různě modifikovány a dnes se užívají pro daleko širší třídu extremálních úloh, ve kterých se hledají extrémy integrálních funkcionálů s omezeními v integrálním tvaru, tj. nalézt (alespoň diferencovatelnou) funkci (křivku) $y = y(x)$ splňující

$$y(a) = A \quad \& \quad y(b) = B \quad \& \quad \mathcal{K}[y] := \int_a^b G(x, y, y')\,\mathrm{d}x = \ell,$$

pro kterou nabývá funkcionál

$$\mathcal{J}[y] = \int_a^b f(x, y, y')\,\mathrm{d}x$$

extremální hodnoty pro pevně dané $\ell \in \mathbb{R}$ a funkce $F, G : \mathbb{R}^3 \to \mathbb{R}$. V případě královny Didó při pevných koncích umístěných do bodů $[-1, 0]$ a $[1, 0]$ jde o úlohu

$$\int_{-1}^{1} y(x)\,dx \to \min$$

za podmínek

$$y(-1) = 0 \quad \& \quad y(1) = 0 \quad \& \quad \int_{-1}^{1} \sqrt{1 + [y'(x)]^2}\,\mathrm{d}x = \ell.$$

Řešením je oblouk jisté kružnice, přičemž pro $\ell = \pi$ máme půlkružnici se středem v počátku.

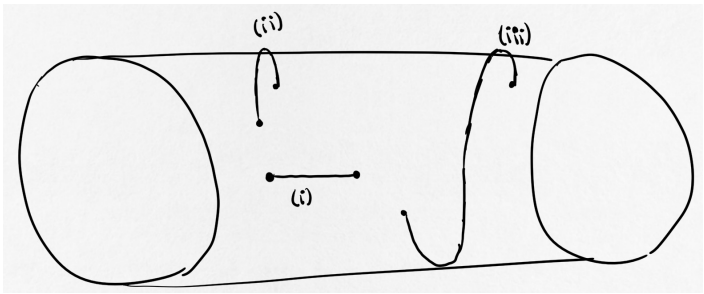
Krátce po vyřešení úlohy o brachistochroně byla řešena řada podobných úloh. Např. *úloha o nejkratších čarách* (neboli *geodetikách*), které se věnoval Euler ve své první publikaci [157]. Podnět k tomuto tématu dostal v roce 1728 od svého školitele Johanna Bernoulliho (ten ji znal již od roku 1698, jak ukazují jeho lekce o integrálním počtu vydané tiskem roku 1742). Mějme

dva body A, B na ploše $S \subseteq \mathbb{R}^3$. Chceme najít nejkratší křivku, která spojuje dané body a současně leží na ploše S. Uvážíme-li např. rotační válec, pak máme tři možné geodetiky (viz Obrázek 1.14 níže):

(i) přímku rovnoběžnou s osou x (pokud body leží na jedné přímce vzniklé řezem rovinou procházející osou válce);

(ii) kružnice kolmou na osu válce (pokud body leží na kružnici vzniklé řezem rovinou kolmou na osou válce);

(iii) šroubovice.



Obrázek 1.14: Geodetiky na válci.

V letectví, aby se minimalizovala dráha letu, letadla „využívají“ právě geodetiky. V ideálním případě spíše koridory, neboť v každém bodě mají jiný kurz – na mapě se zdá, že dráha není nejkratší (není to přímka), ale nezapomínejme, že nežijeme na zeměploše. Pro jednoduchost uvažme, že plocha S je dána jako graf funkce

$$S\colon\ z = F(x,y)$$

(to válec samozřejmě nesplňuje, ale můžeme uvažovat válcové souřadnice; podobně pro kouli sférické souřadnice; v diferenciální geometrii to lze dokonce rozšířit na libovolnou parametrickou plochu). Chtěli bychom najít geodetiku $C \subset S$, která spojuje body

$$A = [a, \alpha, F(a,\alpha)] \quad \& \quad B = [b, \beta, F(b,\beta)].$$

Připusťme, že C může být parametrizována pomocí x jako

$$y = u(x) \quad \& \quad z = v(x) = F(x, u(x)),$$

přičemž poslední rovnost zaručuje to, že C leží na ploše S. Toto zjednodušení ovšem vyžaduje, aby $a \neq b$. Délka křivky je pak dána standardně, takže chceme minimalizovat funkcionál

$$\ell[u] = \int_a^b \sqrt{1 + [u'(x)]^2 + [v'(x)]^2}\,\mathrm{d}x = \int_a^b \sqrt{1 + [u'(x)]^2 + \left[\frac{\partial F}{\partial x}(x,u(x)) + \frac{\partial F}{\partial u}(x,u(x))\,\frac{\mathrm{d}u}{\mathrm{d}x}\right]^2}\,\mathrm{d}x$$

za požadavku $u(a) = \alpha$ a $u(b) = \beta$. Např. pro geodetiky na paraboloidu

$$z = \frac{x^2}{2} + \frac{y^2}{2}$$

chceme

$$\int_a^b \sqrt{1 + [u'(x)]^2 + [x + u(x)\,u'(x)]^2} \to \min.$$

Euler své výsledky v této oblasti shrnul v práci [159]. V ní je odvozena obecná metoda (tzv. *Eulerova přímá metoda konečných diferencí*) pro řešení takovýchto problémů a následně ilustrována na stovce z nich. My jsme si tuto metodu ukázali bez větších detailů již v Příkladě 1.2. Euler tuto svoji metodu opustil poté, co 12. srpna 1755 obdržel od Lagrange dopis s mnohem elegantnější (*přímou*) *metodou variací*, kterou objevil při řešení *problému tautochrony*[96]. Tato metoda uchvátila Eulera natolik, že už v textu [160] z roku 1756 (publikován byl ale až roce 1766) označil na Lagrangeovu počest tuto oblast názvem *variační počet*.

[96] Jde o nalezení křivky, podél níž dorazí hmotný bod do cíle vždy za stejný čas bez ohledu na výchozí bod. Už jsme se zmínili, že řešením i tohoto problému je obrácená cykloida, což ukázal již Huygens v roce 1659.

Zobecněním problému geodetik jsou tzv. *minimální plochy*. Nejjednodušší interpretace je, že máme křivku C v $\mathbb{R}^3$ a cílem je najít plochu s minimálním povrchem mezi všemi plochami, které mají za hranici křivku C. Např. máme-li C jako uzavřenou rovinnou křivku (třeba kružnici), pak minimální plocha je pouze oblast ohraničená touto křivkou. Ovšem pokud je křivka zakřivená i ve třetím rozměru, pak podoba minimální plochy již není tak zřejmá. Fyzikální interpretace by mohla vypadat následovně: pokud ohneme dva dráty do tvaru křivky C a ponoříme do mýdlové vody, pak povrchové napětí výsledné plochy, kterou obdržíme odtažením drátů jako mýdlovou plochu, způsobí to, že máme minimální plochu. Mají-li oba dráty tvar kružnice, pak řešením je tzv. *katenoid*, což je rotační plocha vzniklá rotací *řetězovky*[97] Řetězovka je další zajímavou křivkou – je totiž řešením dalšího isoperimetrického problému hledajícího křivku (v ideálním případě homogenní dokonale ohebné nepružné vlákno – řetěz) dané délky s minimální potenciální energií ve svislé rovině tíhového pole Země (má tedy nejníže položené těžiště). Její tvar zaujímají závěsné mosty (např. lásky přes Svratku v Brně–Komíně) či dráty elektrického vedení. Objevuje se také v architektuře – např. zastřešení Pavilonu A na BVV, Gateway Arch v St. Louis či Casa Milà od Antoniho Gaudího v Barceloně. Mýdlové bubliny a plochy byly po staletí zdrojem velkého zájmu fyzikálního, estetického i matematického, viz tzv. *Plateaův problém*.

[97] Anglicko-matematický vtip: *How do you make cat-enoid? By pulling its tail.* Tohle je bohužel do češtiny asi nepřenositelné. . .

Ovšem parametrická optimalizace rozhodně takovou pozornost neupoutala, ačkoli takových úloh byla v historii řešeno velké množství[98]. Vlastně jeden takový problém dal vzniknout celé naší civilizaci, když Adam a Eva nalezli optimální řešení *jablíčkového dilematu*, které mělo pro hada za následek věčné plazení po břiše a pro lidstvo znamenalo vyhnání z ráje, viz [533]. V Bibli najdeme i možné řešení plánovacího problému týkajícího se *egyptské sedmiletky* či *tenkého programování*:

[98] Velkou inspirací pro některé z následujících historických informací byla kniha [208]

> *25 Josef faraónovi odvětil: „Faraónův sen je jeden a týž. Bůh faraónovi oznámil, co učiní.*
> *26 Sedm pěkných krav, to je sedm let. Také sedm pěkných klasů je sedm let. Je to jeden sen.*
> *27 Sedm vychrtlých a šeredných krav, vystupujících za nimi, je sedm let, stejně jako sedm prázdných a východním větrem sežehlých klasů; to bude sedm let hladu.*
> *28 Když jsem faraónovi řekl: Bůh faraónovi ukázal, co učiní, mínil jsem toto:*
> *29 Přichází sedm let veliké hojnosti v celé egyptské zemi.*
> *30 Po nich však nastane sedm let hladu a všechna hojnost v egyptské zemi bude zapomenuta. Hlad zemi úplně zničí.*
> *31 V zemi nebude po hojnosti ani potuchy pro hlad, který potom nastane, neboť bude velmi krutý.*
> *32 Dvakrát byl sen faraónovi opakován proto, že slovo od Boha je nezvratné a Bůh to brzy vykoná.*
> *33 Ať se tedy farao nyní poohlédne po zkušeném a moudrém muži a dosadí ho za správce egyptské země.*
> *34 Nechť farao ustanoví v zemi dohlížitele a po sedm let hojnosti nechť vybírá pětinu výnosu egyptské země.*
> *35 Ať po dobu příštích sedmi úrodných let shromažďují všechnu potravu a ve městech ať uskladňují pod faraónovu moc obilí a hlídají je.*
> *36 Tato potrava zabezpečí zemi na sedm let hladu, která přijdou na egyptskou zemi. A země nezajde hladem."*
> *37 Tato řeč se faraónovi i všem jeho služebníkům zalíbila.*

viz [534]. V méně vzdálené historii zase najdeme úlohy s geometrickým základem a s geometrickým řešením. Známé jsou (k již dříve uvedeným třem úlohám či Fermatově úloze, kterou lze také formulovat jako problém parametrické optimalizace) např.

(i) *Heronova úloha* z 1. stol. př.n.l.: jsou-li dány dva body A a B na téže straně přímky ℓ, nalezněte bod C na této přímce tak, aby součet vzdáleností z A do C a z C do B byla minimální (populární je její formulace pomocí *kosa a žížal*);

(ii) *Tartagliova úloha*: číslo 8 rozdělte na dvě části tak, aby jejich součin vynásobený jejich rozdílem byl maximální;

(iii) *Keplerova úloha*: do dané koule vložte válec maximálního objemu.

Ve středověku je možné odhalit úlohy parametrické optimalizace zejména v souvislosti se studiem vítězných strategií v různých hrách, tj. s jistým prvkem nejistoty, což vedlo až ke vzniku rigorózní teorie pravděpodobnosti. Za zmínku zde stojí např. Cardanova kniha *Liber de Ludo Aleae*[99] z roku 1564, která ale byla publikována až v roce 1663, nebo práce Blaise Pascala[100] a Fermata. Všichni se věnovali i tzv. *Úloze o rozdělení sázky* (či *Problém s body*), která pochází nejspíše ze 14.–15. století a v tištěné podobě se poprvé objevuje v knize *Summa de arithmetica, geometrica, proportioni et proportionalità* [*Přehled aritmetiky, geometrie, proporcí a proporcionality* (Summary of arithmetic, geometry, proportions and proportionality)] od Lucy Pacioliho[101] a z roku 1494. V této úloze jde o dva hráče A a B hrající několik kol nějaké spravedlivé hry, kteří se dohodli, že celou sázku získá ten, jenž jako první bude mít N výher. Jenže vlivem vnějších okolností je nutné hru přerušit ve chvíli, kdy hráč A má pouze r výher a hráč B pouze s výher. Jak si mají hráči rozdělit sázku? Pacioli navrhoval, aby si sázku rozdělili v poměru výher, tj. $r : s$. Takové řešení se nelíbilo např. Tartagliovi, který trefně namítal, že kdyby se to stalo hned po první hře, tak by si celou sázku vzal vítěz této hry. Tartaglia přišel s jistým návrhem řešení založeným na poměru čísla N a rozdílu $|r - s|$. Částečné řešení přinesl v roce 1539 i Cardano, který přišel s tím, že jde pouze o to, „co by se mohlo stát v budoucnosti, nikoli v minulosti“[102]. Ale ani Cardanovo řešení nebylo správné. S tím přišli až Fermat s Pascalem ve vzájemné korespondenci z roku 1654.

Další zajímavou karetní hrou je *Le Her* (“Jí”), ve které dva hráči A a B losují po jedné kartě z klasického balíčku 52 karet s hodnotami od 1 do 13 ve 4 barvách. Hráč A může přinutit hráče B, aby si karty vyměnili s výjimkou situace, kdy B má číslo 13. Pokud B není spokojen se svoji původní kartou nebo s tou, kterou dostal po výměně s A, může si z balíčku zbývajících 50 karet vylosovat jinou, ovšem pokud si vylosuje číslo 13 není výměna povolena. Poté si oba hráči porovnají karty a vyhrává ten, který má větší číslo, přičemž v případě shody vyhrává B. Tato hra se poprvé objevuje v dopise od Nicholase Bernoulliho[103] z 10. listopadu 1711 určeném Pierru de Montmortovi[104]. Řešení tohoto problému se podařilo najít Francisi Waldegraveovi[105], který hledal strategii maximalizující pravděpodobnost výhry bez ohledu na protihráčovu taktiku. Tím se mu podařilo najít tzv. *minimax* řešení.

Pro některé z těchto úloh lze využít tvrzení, které jsme si uvedli v předchozí části, avšak v minulosti byly obvykle řešeny jinak a zcela individuálně – neexistovala žádná univerzální metoda. Jenže zmíněné nástroje se zdají být nedostatečné např. pro řešení jednoduchého *dopravního problému*. Podnikatel XY má v EU 3 výrobní závody s kapacitami 250, 500 a 700 ks denně. Současně je po Evropě 5 prodejních skladů, z nichž v každém lze prodat až 300 ks denně. Kolik kusů rozeslat do jednotlivých skladů tak, aby ve skladech bylo na začátku každého dne tolik kusů zboží, kolik tam lze prodat, a současně s tím byly spojeny minimální přepravní náklady? Toto vlastně úzce souvisí s výzkumem Charlese Babbageho[106] ohledně nákladů na přepravu a třídění poštovních zásilek, který vedl v roce 1840 ke zřízení *Penny Post* coby jedné ze součástí kompletní reformy Královské pošty. Nebo můžeme připomenout již dříve popsanou *metodu nejmenších čtverců*, která pochází z přelomu 18. a 19. století a jejíž vznik je spojován především s Gaussem a Adrienem–Marie Legendrem[107].

Překvapivě až do první poloviny 20. století nenalezneme příliš mnoho metod k řešení takovýchto problémů (tj. rozhodovacích či řídících procesů). Zájem o tyto problémy stoupal zejména v ekonomii v souvislosti s rozvojem průmyslu a zefektivněním výroby, což pokračovalo i ve 30. letech, i když tentokrát ve spojení s příchodem *Nového údělu*[108] v USA a podobných

[99] *Kniha o hazardních hrách* (The Book on Games of Chances)

[100] Blaise Pascal (19. června 1623 – 19. srpna 1662) byl francouzský matematik, fyzik, vynálezce, spisovatel a katolický teolog. V 19 letech začal pracoval na výrobě mechanického kalkulátoru. Po třech letech, ve kterých sestavil 50 prototypů, úspěšně zkonstruoval tzv. Pascalův kalkulátor či *Pascalinu* a v následujících 10 letech nechal vyrobit 20 kusů. Díky tomu se stal jedním z prvních dvou konstruktérů mechanického kalkulátoru. V 16 letech sepsal krátké pojednání o *mystickém hexagramu*, které nazval *Essai pour les coniques* [*Pojednání o kuželosečkách* (Essay on Conics)], které jakožto svoji první skutečnou matematickou práci odeslal Marinu Mersenneovi do Paříže. Jejím obsahem je dodnes známá Pascalova věta týkající se šestiúhelníku vepsaného do kružnice. Tuto práce ocenila i Pařížská královská akademie a Descartes ji pokládal za práci jeho otce, když to komentoval slovy „Nezdá se mi divné, že předvedl důkazy o kuželosečkách mnohem vhodnější než ty antické, jenže v této souvislosti lze zmínit jiné otázky, které by sotva přišly na mysl šestnáctiletému dítěti“. Pascal byl také v určité době velmi vášnivým hráčem, což jej přivedlo k tomu, že společně s Fermatem položily základy teorie pravděpodobnosti (jím označované jako *matematická naděje*) a vytvořil pojem střední hodnota.

[101] Luca Bartolomeo de Pacioli (někdy v roce 1447 – 19. června 1517) byl italský matematik, františkánský mnich, spolupracovník Leonarda da Vinci a jeden ze zakladatelů účetnictví.

[102] Jinými slovy jde o tzv. *princip úměrnosti*: *při řešení nejisté úlohy v budoucnosti máme uvažovat pouze to, co se může stát, a nikoliv to, co se již stalo.* Na toto si vzpomeňte, až se dostaneme ke Kapitole 5

[103] Nicolaus Bernoulli (21. října 1687 – 29. listopadu 1759) byl švýcarský matematik. Doktorát získal ve 22 letech za práci o teorii pravděpodobnosti v právu. Je autorem věhlasného Petrohradského paradoxu.

plánů v ostatních částech světa, které měly za cíl překonání *Velké hospodářské krize*. Zlom v dosavadním *ad hoc* přístupu přinesla až druhá světová válka, která si žádala řešení i jiných úkolů, a to především těch vojenských. Během této války armádní vedení Velké Británie vyzývalo skupiny vědců z různých oblastí k využití jejich vědeckých znalostí v různých strategických a taktických válečných problémech. Díky tomu vznikla nová oblast matematiky nazývaná *operační výzkum*[109], což je víceméně totéž jako „parametrická optimalizace“, ale toto označení odkazuje na skutečnost, že se jednalo především o *výzkum ve vojenských operacích*. Za skutečný počátek lze proto považovat rok 1936, kdy britský ministr letectví ustanovil v Suffolku tzv. *Bawdsey Research Station* v panství Bawdsey Manor, které dříve patřilo jednomu ze zakladatelů *National Telephone Company*. Úkolem této skupiny bylo hledání vojenského využití nově objevené radarové technologie, např. pro cílené zachytávání nepřátelských letadel. Prvním šéfem této skupiny byl Robert Watson-Watt, jehož v polovině roku 1938 vystřídal Albert Percival Rowe[110].

V roce 1940 byla pod vedením fyzika Patricka M. S. Blacketta[111] ustanovena multioborová skupina (zahrnující tři fyziology, jednoho obecného fyzika, dva matematické fyziky, dva matematiky, jednoho astrofyzika, jednoho armádního úředníka a jednoho geodeta) s organizačním názvem *The Anti-Aircraft Command Research Group, Royal Air Force*, jejímž úkolem bylo zjištění možného využití radarů při protiletadlové střelbě. Tato skupina byla sice označována jako *Blackettův cirkus*, ale podařilo se jim ukázat funkčnost multioborových týmů v operačním výzkumu a zároveň i demonstrovat jejich užitečnost při řešení komplexních (tj. velmi spletitých) reálných problémů, viz mimo jiné [65]. Díky tomu se začaly formovat i další podobné týmy. Tomuto „cirkusu“ se podařilo snížit průměrný počet výstřelů protivzdušné obrany potřebných ke zničení nepřátelského letounu z více než 20000 kusů až na pouhé 4000 při *Bitvě o Británii* v roce 1941. Blackett o sobě říkal, že jeho cílem je nalezení čísel, na nichž by místo nárazových emocí měla být založena vojenská strategie. Proto Blackett např. kritizoval množství zdrojů investovaných do plošných náletů, které podle něj měly jít do jiných armádních složek, jelikož jeho studie ukázaly neefektivnost těchto bombardovacích plánů ve srovnání s důležitostí bojů proti německým ponorkám, které potápěním obchodních lodí výrazně ovlivňovaly válečné úsilí. Tímto názorem samozřejmě tehdejší vojenské velení příliš nenadchl, takže byl po zásluze odměněn tím, že jej odstřihli z různých komunikačních okruhů. Nicméně po válce [sic!] mu bylo ve spojenecké zprávě *United States Strategic Bombing Survey* dáno za pravdu.

Když byl Blackett v prosinci 1941 dotazován britskou admiralitou ohledně vytvoření sekce operačního výzkumu, napsal memorandum s názvem *Scientists at the operational level*, které silně rezonovalo na obou stranách Atlantiku. Díky tomuto podnětu vznikla skupina *U.S. Navy Anti-submarine Warfare Operations Research Group* (ASWORG) a v lednu 1942 se Blackett přesunul do vedení britského námořnictva, aby vytvořil také skupinu věnující se operačnímu výzkumu. Ta dosáhla velkého úspěchu, když se jim podařilo určit optimální velikost obchodního konvoje, při kterém budou minimalizovány ztráty při útoku ponorek v kombinaci s požadavky na doprovod. Nebylo totiž zřejmé, zda jsou výhodnější malé konvoje (které mohou plout rychleji a jsou hůře detekovatelné německými ponorkami) nebo velké konvoje (které mají více válečných lodí schopných bránit náklad). Blackettova skupina ukázala, že mnohem důležitější je počet doprovodných plavidel než velikost samotného konvoje. Závěr proto byl, že několik velkých konvojů je mnohem lépe bránitelných než mnoho malých konvojů, což po implementování do námořní strategie vedlo k výraznému snížení ztrát. Americký ASWORG byl první skupinou s civilními zaměstnanci zapojenou do vojenského operačního výzkumu. Na začátku se jednalo o 15 zaměstnanců, jejichž počet se v průběhu války rozrostl až ke stovce. V říjnu 1942, když vrcholila válka, dorazila první delegace amerických analytiků operačního výzkumu do Anglie, aby spolupracovala s osmým americkým bombardovacím komandem (*VIII Bomber Command*). Jejich prvním úkolem bylo zlepšení přesnosti bombardování. Na základě svých kvantitativních studií předchozích útoků navrhli, že nejlepší bombardér by měl být v hlavním letounu tak, aby mohl zacílit celou dávku bomb vyhozenou v salvě tím, že letadla přesně do-

[104] Pierre Rémond de Montmort (27. října 1678 – 7. října 1719), byl francouzský matematik. V roce 1715 byl zvolen členem Královské společnosti v Londýně a v roce 1716 také Francouzské akademie věd. Je znám především díky knize *Essay d'analyse sur les jeux de hazard* [*Analýza hazardních her* (Analysis of games of chance)] o pravděpodobnosti a hazardních hrách a je také autorem sousloví *Pascalův trojúhelník*, pro který používal označení *Table de M. Pascal pour les combinaisons* [*Tabulka pana Pascala pro kombinace* (Table of Mr. Pascal for combinations)].

[105] Tak o tomhle pánovi nevíme asi vůbec nic. Vlastně bylo docela těžké jej vůbec identifikovat, viz [43, 44].

[106] Charles Babbage (26. prosince 1791 – 18. října 1871) byl anglický matematik, filosof, vynálezce a strojní inženýr. Bývá považován za „otce počítačů“, neboť jako první přišel s myšlenou programovatelných počítačů. Některé části z jeho nedokončených mechanických počítačů jsou vystaveny v Londýnském vědeckém muzeu.

[107] Adrien–Marie Legendre (18. září 1752 – 10. ledna 1833) byl francouzský matematik, který výrazně přispěl k rozvoji statistiky, teorie čísel, abstraktní algebry a matematické analýzy.

[108] *New Deal* je název souboru opatření a ekonomických a sociálních reforem zavedených v průběhu let 1933–1937 v USA během vlády prezidenta Franklina D. Roosevelta s cílem podpořit, ozdravit a zreformovat státní ekonomiku během Velké hospodářské krize.

[109] *Operations research* nebo v původním britském označení *Operational research*.

[110] Právě Roweovi je připisováno autorství termínu „operační výzkum“.

[111] Mj. držitel Nobelovy ceny za fyziku z roku 1948.

držují úzkou formaci, čímž se výrazně sníží rozptyl celé dávky bomb. Na základě této analýzy došlo k nejméně 1000 procentnímu nárůstu počtu bomb dopadajících na stanovené cíle. Pro mnohem více podrobností viz nejenom Kapitoly I.3–4 v [332].

Metody operačního výzkumu během války sice byly stále vyvíjeny přesně na míru konkrétním vojenským problémům, ale jejich výsledky zaznamenaly velký úspěch a všichni si byli dobře vědomi, že je to velmi strategické odvětví, a tudíž je potřeba mít k dispozici vědce, kteří ovládají vojenské plánování a řešení související vládních problémů. Proto vláda USA v prosince 1945 založila *Projekt RAND*[112], ze kterého se v únoru 1948 stala nezávislá nezisková organizace a která sehrál v následném rozvoji operačního výzkumu klíčovou roli. Vznikaly také další neméně významné organizace. Vzhledem ke klesajícímu počtu nepřátelských ponorek a vzniku nových aplikací operačního výzkumu v námořnictví, došlo k přejmenování skupiny ASWORG na *Skupinu operačního výzkumu*[113] a jejímu podstoupení pod Atlantické loďstvo Spojených států amerických. Jenže po válce se velikost této skupiny smrskla cca na čtvrtinu, což se samozřejmě nelíbilo jejímu vedení, takže v listopadu došlo k dalšímu přejmenování na *Skupinu pro vyhodnocování operací*[114] a především k jejímu propojení s *Massachusettským technologickým institutem* (MIT). V červnu 1947 také Pentagon zřídil v rámci letectva výzkumnou skupinu nazývanou *Projekt SCOOP*[115]. Jejím cílem bylo nalezení vhodných (a rychlých) odpovědí na problémy plánování požadavků letectva, např. určení časového harmonogramu materiálových potřeb na podporu válečného plánu. V době vzniku těchto organizací tomu ale ještě stále chyběl jakýsi obecný pohled. Ten však přišel záhy (podzim 1947) díky rozvoji algoritmického řešení úloh *lineárního programování*, tj. úloh

$$f(x) := \langle c, x \rangle = \sum_{i=1}^{n} c_i x_i \to \min \quad \& \quad A_1 x \le b_1 \quad \& \quad A_2 x = b_2 \tag{1.3.1}$$

a případně $x_i \ge 0$ pro $i \in I \subseteq \{1, \dots, n\}$, kde $c \in \mathbb{R}^n$, $A_1 \in \mathbb{R}^{m \times n}$, $A_2 \in \mathbb{R}^{k \times n}$, $b_1 \in \mathbb{R}^m$ a $b_2 \in \mathbb{R}^k$ jsou dané vektory a matice.

[112] **R**esearch **a**nd **D**evelopment

[113] Operations Research Group (ORG)

[114] Operations Evaluation Group (OEG)

[115] Project SCOOP (U.S. Air Force Scientific Computation of Optimum Programs)

Příklad 1.3.1

Jako jednoduchý příklad úlohy lineárního programování může posloužit situace z nedávné doby karanténních opatření kvůli pandemii COVID-19. Ta způsobila, že nám po pár dnech zbyla v lednici pouze syrová mrkev, syrové bílé zelí a zavařené okurky. Internet naštěstí fungoval, takže jsme si na webových stránkách WHO (nebo spíše Blesku pro ženy) našli *nutriční standardy*, z nichž jsme si vybrali tři faktory: vitamín A, vitamín C a vlákninu. Veškeré potřebné informace jsou shrnuté v následující tabulce (cena okurek je ovlivněna nákupem v „předkaranténní“ době ze zbytků armádních zásob, které by za vyšší cenu byly nejspíše zcela neprodejné).

	mrkev	zelí	okurky	minimální požadované množství v 1 porci
vitamín A [mg/kg]	35	0,5	0,5	0,5 mg
vitamín C [mg/kg]	60	300	10	15 mg
vláknina [g/kg]	30	20	10	4 g
cena [Kč/kg]	15	12,50	3,75	

Samozřejmě v této těžké době nemá smysl hrát si na hrdiny ani na milionáře, takže jsme chtěli utrácet za jídlo co nejméně, což nás přivedlo k tomu, že náš jídelníček musí být založen na řešení úlohy (1.3.1) ve tvaru

$$15\,x_M + 12{,}5\,x_Z + 3{,}75\,x_O \to \min$$
$$35\,x_M + 0{,}5\,x_Z + 0{,}5\,x_O \ge 0{,}5$$

$$60\,x_M + 300\,x_Z + 10\,x_O \geq 15$$
$$30\,x_M + 20\,x_Z + 10\,x_O \geq 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0,$$

kde jednotlivé proměnné x_M, x_Z a x_O postupně představují konzumované množství mrkve, zelí a okurek. Nezápornost těchto proměnných je celkem přirozený požadavek, jehož případné porušení by znamenalo, že budeme některé suroviny dále přeprodávat za nákupní ceny. To ale rozhodně nepřicházelo v úvahu – dělit se se sousedy? Ani náhodou! Následující dny jsme tak přežili jenom díky tomu, že náš jídelníček byl založen na konzumaci 9,52634 g mrkve, 38,265 g zelí a 294,891 g okurek s cenou 1,72705 Kč za jednu porci.

Jedná se o nejjednodušší (a zároveň nejstudovanější) typ úlohy matematického programování, což vlastně můžeme také chápat jako rozhodovací proces, ve kterém se má určit efektivní využití omezeného množství zdrojů k dosažení požadovaných cílů. Díky zmechanizování výpočtu řešení úloh lineárního programování došlo k jeho obrovskému rozmachu a následně i jiných typů úloh matematického programování v ekonomii, průmyslu a moha dalších odvětvích současnosti. V této souvislosti je ale zapotřebí zmínit jeden velmi důležitý fakt, který asi není na první pohled úplně zřejmý a který velmi úzce s nerovnostním omezením v (1.3.1). Zatímco teorie řešení soustav lineárních rovnic je velmi dobře prozkoumána nejméně od poloviny 19. století díky intenzivnímu studiu maticové algebry, podobná otázka řešení soustav lineárních nerovnic byla nejméně do roku 1936 na okraji jakéhokoli matematického zájmu. Avšak díky Motzkinovi[116] začala být i této oblasti věnována intenzivnější pozornost, díky čemuž mohlo být lineární programování vybudováno na velmi pevných základech.

Samozřejmě existuje velké množství problémů, které nelze linearizovat. Proto v závislosti na typu účelové funkce f a přípustné množiny X dostaneme další důležité typy úloh matematického programování – souhrnně pojmenovatelné jako *nelineární programování*. Samotné pojmenování „matematické programování“ pochází z roku 1953 od Roberta Dorfmana[117], kterému se označení lineární programování zdálo příliš restriktivní či zavádějící (jelikož ne vše bylo vždy pouze lineární, viz [140]). Nelineární programování vzniká v roce 1951 a jeho základy tvoří zobecnění *Lagrangeova principu* a také *Johnovy* či *Karushovy–Kuhnovy–Tuckerovy podmínky*, čemuž se budeme později velmi intenzivně věnovat i my. Jednou z nejjednodušších úloh nelineárního programování je např. již dříve zmíněné *kvadratické programování*

$$f(x) := \tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \to \min, \quad A_1 x \leq b_1, \quad A_2 x = b_2$$

a případně $x_i \geq 0$ pro $i \in I \subseteq \{1,\dots,n\}$, kde $C \in \mathbb{R}^{n\times n}$, $d \in \mathbb{R}^n$, $A_1 \in \mathbb{R}^{m\times n}$, $A_2 \in \mathbb{R}^{k\in\times n}$, $b_1 \in \mathbb{R}^m$ a $b_2 \in \mathbb{R}^k$ jsou dané vektory a matice, přičemž matice C je navíc symetrická a pozitivně semidefinitní, tj. $C = C^\top \geq 0$. Tyto úlohy hrají důležitou roli např. při tvorbě optimálního portfolia. Avšak kromě kategorie nelineárního programování patří všechny tyto úlohy mezi problémy tzv. *konvexního programování*, tj. úlohy, kde hledáme minimum „konvexní funkce na konvexní množině“. Vznik této oblasti je úzce spojen s Wernerem Fenchelem[118], který v roce 1949 položil základy skutečné (tj. systematické) základy teorie konvexních funkcí. Takové funkce a množiny byly studovány již v první polovině 20. století a za zmínku stojí především Constantin Carathéodory, Hermann Minkowski (ten byl asi první), Ernst Steinitz, Farkas Gyula (nebo Julius Farkas). S většinou z těchto jmen se v našem výkladu ještě setkáme. Fenchelova monografie/referát [170], která vznikla na základě jeho přednášek na Princetonské univerzitě na jaře 1951, výrazně inspirovala mnohé zájemce o tuto problematiku. V roce 1970 vydal Ralph Rockafellar[119] knihu s názvem *Konvexní analýza* [432], což je nejspíše první výskyt takového slovního spojení, které bylo následně obecně přijato a jehož autorem nejspíše je (soudě dle poděkování v předmluvě) již dříve zmíněný Tucker. V čem ale tkví důležitost celého konceptu

116

Theodor Samuel Motzkin (26. března 1908 – 15. prosinec 1970) byl izraelsko-americký matematik. Jeho otec ve svých třinácti letech přestěhoval do Berlína kvůli studiu matematiky. Po univerzitních studiích byl přijat Leopoldem Kroneckerem k doktorskému studiu, avšak toho zanechal před dokončením disertační práce kvůli zapojení do sionistického hnutí. Theodor Motzkin prokázal své matematické nadání již v dětství. Už v patnácti letech nastoupil na univerzitu a v souladu s tehdejšími zvyky prošel univerzitami v Göttingenu, Paříži a Berlíně (světová centra matematiky na začátku 20. století). Disertační práci na téma lineárních nerovností a lineárního programování obhájil v roce 1936.

117 Robert Dorfman (27. října 1916 – 24. června 2002) byl profesorem politické ekonomie na Harvardově univerzitě. Významně přispěl k rozvoji ekonomie, statistiky, skupinových testů a teorie kódování.

118

Moritz Werner Fenchel (3. května 1905 – 24. ledna 1988) byl německý matematik, který kvůli nacistickým protižidovským zákonům emigroval v roce 1933 do Dánska. Do roku 1933 působil na Univerzitě v Göttingenu, odkud se přesunul na Kodaňskou univerzitu. V roce 1946 byl zvolen členem Královské dánské akademie věd a vzdělanosti. V letech 1949–1951 působil na Stanfordově uni-

verzitě a na Princetonské univerzitě. Po návratu byl krátce mezi lety 1952 a 1956 profesorem mechaniky na polytechnice v Kodani, aby se v roce 1956 vrátil zpět na univerzitu v Kodani, kde pracoval až do roku 1974 jako profesor matematiky.

119

Ralph Tyrrell Rockafellar (⋆10. února 1935) je americký matematik a v současnosti působí jako emeritní profesor Oddělení matematiky a aplikované matematiky na Washingtonské univerzitě. Je jedním z předních odborníků na teorii optimalizace a související oblasti analýzy a kombinatoriky. V roce 1999 získal *John von Neumann Theory Prize* od *Institute for Operations Research and the Management Sciences* (INFORMS; Institut pro operační výzkum a vědy o řízení) za „zásadní a trvalé příspěvky k teorii operačního výzkumu a vědám o řízení". V roce 1992 získal nejvyšší ocenění *Society for Industrial and Applied Mathematics* (SIAM; Společnost pro průmyslovou a aplikovanou matematiku), když získal *John von Neumann Prize*. Tato cena je udělována za „vynikající a mimořádné příspěvky do oblasti aplikovaných matematických věd a za efektivní komunikaci těchto myšlenek v rámci komunity".

konvexnosti? To velmi trefně vystihl samotný Rockafellar ve své přednášce při udělování *John von Neumann Prize* v roce 1992:

> *One distinguishing idea which dominates many issues in optimization theory is convexity… An important reason is the fact that when a convex function is minimized over a convex set every locally optimal solution is global. Also, first-order necessary conditions for optimality turn out to be sufficient. A variety of other properties conducive to computation and interpretation of solutions ride on convexity as well. In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity.*[120]

Jinými slovy, díky konvexnosti není potřeba rozlišovat mezi lokálním a globálním minimy a v případě nalezení stacionárního bodu účelové funkce v přípustné množině není nutné ověřovat ještě nějaké dodatečné podmínky druhého řádu. Přidáme-li k úloze lineárního programování (1.3.1) navíc požadavek celočíselnosti jednotlivých proměnných, získáme úlohu *celočíselného programování*, případně *smíšeného programování*, pokud vyžadujeme celočíselnost pouze od některých proměnných. Toto je velmi obsáhlá oblast matematického programování a samozřejmě zde zdůrazníme, že takové úlohy nelze řešit jako klasickou úlohu lineárního programování pouhým zanedbáním podmínek celočíselnosti. Kdyby totiž řešení takové úlohy bylo neceločíselné, pak bychom se následným zaokrouhlením mohli dostat mimo přípustnou množinu nebo do jiného bodu, než je skutečné minimum. Pro „malé" úlohy celočíselného programování dokážeme s pomocí naší intuice a trochou selského rozumu přesně nalézt jejich řešení, což se rozhodně nedá říci o úlohách lineárního programování. V dalších kapitolách uvidíme, že z matematického pohledu je ta situace diametrálně odlišná. Velmi odlišná…

Další důležitou oblastí operačního výzkumu, které se budeme věnovat, je *dynamické programování*. Jedná se o vícekrokový rozhodovací proces, ve kterém rozhodovací proměnná může být mít buď celočíselnou hodnotu nebo libovolnou hodnotu z nějakého předem daného intervalu. I takové úlohy se dají formulovat jako problémy lineárního (či matematického) programování, jenže jejich řešení tímto způsobem nemusí být úplně snadné a efektivní. Naštěstí časová (či jiná) souslednost jednotlivých kroků (tj. dynamika celého procesu) umožňuje i jiný pohled na tuto oblast. Pokud jednotlivé fáze celého procesu vykazují jakousi *separovatelnost*, lze ně nahlížet optikou *dynamického programování*. Navíc se v praxi může celkem snadno stát, že ani neznáme přesné důsledky našich rozhodnutí, což nás přivádí do oblasti *stochastického dynamického programování* nebo v jednodušším případě pouze ke *stochastickému programování*, které se poprvé objevilo v Dantzigově článku z roku 1955, viz [114].

[120] *To, co odlišuje mnoho optimalizačních problémů, je konvexnost…A to zejména proto, že když konvexní funkce je minimalizována na konvexní množině, pak každé lokální minimum je současně globálním minimem. Také nutné podmínky prvního řádu pro lokální extrém se stanou postačujícími. Řada dalších vlastností usnadňující výpočty a interpretaci jejich výsledků také velmi úzce souvisí s konvexností. V této oblasti tedy není klíčové rozlišovat mezi lineární a nelineární úlohou, ale mezi konvexní a nekonvexní úlohou.* Viz [433, str. 184–185].

KAPITOLA 2

Lineární programování

2.1 *Trochu více historie (snad) nikoho nezabije* 62
2.2 *Seznamte se! Úlohy lineárního programování* 80
2.3 *Lineární programování všude kolem nás* 91
2.4 *Základem všeho je slušný náčrtek* 105
2.5 *Teoretické základy aneb být, či nebýt BPB* 116
2.6 *Dualita aneb peníze až na prvním místě* 126
2.7 *Svatý grál lineárního programování – simplexová metoda* 139
2.8 *Teď už si to jen správně zorganizovat* 144
2.9 *Aby začátky nebyly tak těžké* 167
2.10 *Simplexová metoda zprava dolů* 175
2.11 *Dopravní problém* 187
2.12 *Některé aplikace nyní podrobněji* 208
2.13 *Kleeho–Mintyho kostka je vržena* 237
2.14 *Lineární lomené programování* 251

MP: YOU ARE OFTEN CALLED THE *FATHER OF LINEAR PROGRAMMING*. IS THAT A TITLE YOU'RE COMFORTABLE WITH?

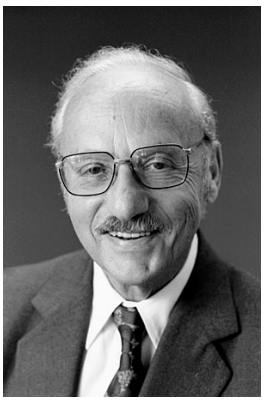
DANTZIG: I HAVE TO TELL YOU A STORY. TWENTY-FIVE YEARS AGO I VISITED JAPAN FOR THE FIRST TIME. WHEN I GOT OFF THE PLANE, THE JAPANESE WHO MET ME WERE VERY SURPRISED AT HOW YOUNG I WAS. SINCE I HAD BEEN BILLED AS THE FATHER OF LINEAR PROGRAMMING, THEY APPARENTLY EXPECTED TO SEE AN OLD MAN WITH WHITE HAIR AND A CANE BEING HELPED DOWN THE RAMP. I WAS 45 AT THE TIME.

MP: THE PREFACE TO YOUR BOOK *LINEAR PROGRAMMING AND EXTENSIONS* OPENS WITH A PROVOCATIVE STATEMENT: "THE FINAL TEST OF A THEORY IS ITS CAPACITY TO SOLVE THE PROBLEMS WHICH ORIGINATED IT."

DANTZIG: DID I SAY THAT? IT'S A GREAT QUOTE. SHOW ME WHERE. . . [1]

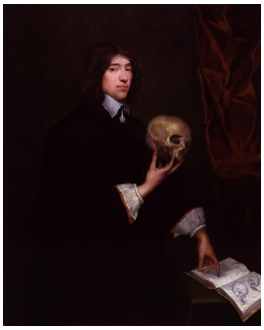
Viz poutavý rozhovor s **GEORGEM DANTZIGEM**[2] v [6] a také [543].

[1] **MP**: *Často jste nazýván „otcem lineárního programování“. Líbí se vám toto označení?*
Dantzig: *Musím vám říct historku. Před 25 lety jsem poprvé navštívil Japonsko. Když jsem vystoupil z letadla, Japonci, se kterými jsem se setkal, byli překvapeni, jak jsem byl mladý. Jelikož jsem byl označován za otce lineárního programování, očividně očekávali, že uvidí starého muže s bílými vlasy a holí pomáhající mu dolů na pojízdných schodech. V té době mi bylo 45 let.*
MP: *Předmluvu vaší knihy „Lineární programování a jeho rozvoj“ začínáte provokativními slovy: „Konečným testem teorie je její schopnost vyřešit problémy, kvůli kterým vznikla.“*
Dantzig: *To jsem řekl? To je úžasný citát. Ukažte mi kde…*

[2]


George Bernard Dantzig (8. listopadu 1914– 13. května 2005) byl americký matematik, který se věnoval průmyslovému inženýrství, operačnímu výzkumu, matematické informatice, ekonomii a statistice. Je považován za *otce lineárního programování*, jemuž vděchl život díky objevu a následnému zveřejnění *simplexového algoritmu*.

2.1 Trochu více historie (snad) nikoho nezabije

[3]


Sir William Petty (26. května 1623 – 16. prosince 1687) byl anglický ekonom, lékař, vědec a filosof. Nejprve se stal prominentem sloužícím Oliveru Cromwellovi a Commonwealthu v Irsku. Vyvinul účinné metody k průzkumu země, která měla být zabavena a dána Cromwellovým vojákům. Petty byl také vědec, vynálezce a obchodník, zakládající člen Královské společnosti (to je britská učená společnost čili akademie pro podporu věd, která byla založena roku 1660 v Londýně) a krátce byl i členem anglického parlamentu. Významné jsou především jeho ekonomické teorie, metoda politické aritmetiky a je mu též připisována filosofie „laissez-faire“, což je heslo klasického liberalismu vyjadřující hlavní ekonomickou zásadu, aby stát nezasahoval do hospodářství a nesnažil se je řídit. V roce 1661 byl pasován na rytíře.

[4]

François Quesnay (4. června 1694 – 16. prosince 1774) byl francouzský ekonom, lékař, filosof a encyklopedista. Ačkoli pocházel ze skromných poměrů (jeho otec byl malým vlastníkem půdy), měl jedenáct sourozenců a ještě v 11 letech neuměl číst, rozhodl se věnovat chirurgii a studoval medicínu. V roce 1749 se stal dvorním lékařem Madame de Pompadour a po vyléčení neštovic syna krále Ludvíka XV (a současně následníka trůnu) byl povýšen do šlechtického stavu. Ludvík XV jej také označoval svými „mluvícími pažemi“ a za svého „myslitele“. Quesnay byl ale především jedním z hlavních představitelů první ekonomické školy (tzv. *fyziokratů*),

Za asi první serioznější pokus o propojení matematiky a ekonomie můžeme považovat práci anglického ekonoma Pettyho[3], který ve své studii národního hospodářství kladl důraz na měření hodnot a schematické modelování vztahů mezi nimi. Jeho vášeň pro přesnost ho přivedla ke slavnému prohlášení, že jeho forma vědy by měla používat pouze měřitelné jevy a hledat kvantitativní přesnost spíše než se spoléhat na srovnávaní nebo superlativy, což vedlo ke vzniku nového předmětu nazvaného „politická aritmetika“. Za další „zlomový“ okamžik pro rozvoj „lineárního programování“ bychom mohli považovat prvním model ekonomického koloběhu, kterým je slavná Quesnayho[4] *Tableau économique* (*Ekonomická tabulka*) z roku 1758, v níž je na příkladu vzájemného vztahu statkáře, rolníka a řemeslníka doplněného o různé vzájemně se protínající čáry ilustrující tok příjmů mezi několika ekonomickými sektory představena jeho ekonomická doktrína, viz Obrázek 31. Tabulka popisuje stacionární stav ekonomiky, kde se explicitně pracuje s rentou a implicitně se mzdou, ale není zde žádný „prostor“ pro zisk. To, že tomu tak skutečně je, lze ukázat i pomocí řešení optimalizační úlohy typu

$$\min\{x_1, x_2\} \to \max$$

doplněné o několik nerovnostních omezení, což lze převést na úlohu lineárního programování

$$z \to \max$$

doplněnou o původní nerovnosti a dvojici podmínek $z \leq x_1$ a $z \leq x_2$. Nicméně Quesnayho tabulka je jen jakýmsi souhrnem jeho různých studií, takže je výrazně zjednodušena, díky čemuž lze v literatuře najít několik jejich různých interpretací. Pro více podrobností doporučujeme [393, Chapter 2] a [412, 458].

Ovšem během dalších více 170 let nalezneme jen velmi málo využití lineárního modelu, ačkoli se objevuje jako část Walrasova modelu ekonomické rovnováhy[5] z roku 1874. Většina matematiků–ekonomů se věnovala analýze teoretických problémů spojených s možností hospodářských rovnováh a efektivností rozdělování výrobních prostředků v konkurenčním nebo monopolním prostředí. V těchto studiích uvažovali klasické konvexní funkce se spojitou derivací, které se hodí na zobrazení podmínek stability více než funkce založené na lineárních nerovnostech. Za zmínku stojí především úsilí skupiny rakouských a německých ekonomů, kteří v roce 1930 pracovali na zobecnění Walrasovy lineární „technologie“. Věnovali se zde např. řešení a jeho jednoznačnosti pro soustavu rovnic, která měla objasnit nutnost nezápornosti cen a míry produkce. Tato soustava také měla ukázat, jaké výrobky jsou nedostatkové a jaké jsou nadbytečné (tj. mají nulovou hodnotu). Uvažovaný model výroby byl však statický ve smyslu, že výrobky mohou být vyráběny jen jedním způsobem a na dané množství výrobků je nutné spotřebovat proporcionální množství primárních zdrojů. Dosažené výsledky vyvolaly pozornost a nejspíše podnítily von Neumanna[6] k článku zaměřeného na model všeobecné rovnováhy k formulaci modelu dynamického lineárního programování, ve kterém uvedl alternativní metody výroby daných výrobků samostatně (odděleně, tj. *diskrétně*) nebo současně (tj. *spojitě*), viz [502]. Díky tomu ukázal nejenom, jaké výrobky jsou nepotřebné, ale také jaké výrobní procesy jsou nevyužity. Navíc v tomto modelu některé komodity mohou být výstupem nějakého výrobního procesu a současně vstupem jiného procesu. Tato myšlenka „cirkulace“ byla rozšířena dokonce i na zboží poptávané spotřebiteli prostřednictvím poněkud vynuceného konceptu činnosti produkující práci absorpcí spotřebního zboží v pevných poměrech. Model se tak stal uzavřeným bez přílivu primárních zdrojů zvenčí i odlivu finálních produktů mimo uvažovaný systém, tj. je hospodářství je plně soběstačné. Předpokládalo se, že jakýkoli nespotřebovaný „přebytek“ bude použit pro tvorbu kapitálu, aby se dosáhlo kontinuálního poměrného růstu všech výrobních činností při neměnné technologii. Von Neumannův model je proto v omezeném smyslu dynamický, neboť změna v čase je popsána jedním skalárním koeficientem stejnoměrného růstu. Ačkoli jeho model neměl žádný explicitní účel, von

kteří byli přesvědčeni, že čistý produkt vzniká pouze v zemědělství. Toto učení později nahradila *klasická ekonomie* (první moderní škola ekonomie vycházející z díla Adama Smithe *Pojednání o podstatě a původu bohatství národů* z roku 1776). V roce 1751 byl Quesnay zvolen do Akademie věd a svými stoupenci byl označován jako „Konfucius Evropy", „moderní Sokrates" či „Mojžíš naší doby".

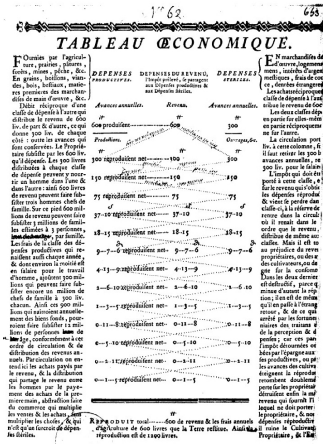


TABLEAU ŒCONOMIQUE.

Obrázek 2.1: Druhé vydání Tableau économique z roku 1759, které bylo doplněno o *Extrait des oeconomies royales de M. de Sully*. Quesnayho tabulka nejspíše vychází ze dvou jeho textů ve věhlasné francouzské *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers* (neboli *Encyklopedie aneb Racionální slovník věd, umění a řemesel*, což je 28 svazkové dílo z let 1751–1766, jehož účelem bylo shrnutí veškerého lidského poznání a za nímž stáli stáli Denis Diderot a Jean le Rond d'Alembert) k heslům „farmiers" (farmář; v 6. svazku z roku 1756) a „grains" (obilí; v 7. svazku z roku 1757).

[5] Marie Esprit Léon Walras (16. prosince 1834 – 5. ledna 1910) byl francouzský matematický ekonom, profesor švýcarské Univerzity v Lausanne a zakladatel lausannské (matematické) školy. Walras zásadně odmítal přístup založený na interakci nabídky a poptávky na oddělených trzích a podle něj byla poptávka po každém zboží závislá nejen na ceně tohoto zboží, ale i na ceně všech ostatních zboží, na ceně všech najímaných výrobních faktorů a konečně i na ceně zboží–úspor. Walrasův model byl určen soustavou lineárních rovnic rovnováhy na jednotlivých trzích a v nejsložitější variantě zahrnoval směnu, výrobu, kapitál a peníze.

Neumann ukázal, že síly trhu maximalizují míru hospodářské expanze, a dokázal, že při maximu se tato míra rovná úrokové míře z kapitálu investovaného do výroby. Nicméně praktický vliv von Neumannova článku asi nebyl příliš velký, neboť obsahoval (podobně jako jiné teoretické články) „jen" zajímavé teoretické věty. Matematici–ekonomové se pravděpodobně zajímali více o získání podobných výsledků pro obecnější model, neboť „pro mnohé ekonomy je výraz *linearity* spojený se zúžením, omezením a nepružností hypotéz". Jinými slovy, toto úsilí patřilo jako mnoho jiných do kvalitativní oblasti národního hospodářství tehdejší doby – světa, ve kterém účelem matematického modelu byl spíše kvalitativní než kvantitativní popis předpokládaných vzájemných vztahů. Jenže takové negativní hodnocení „linearity" by bylo příliš krátkozraké. Neměli bychom totiž opomíjet, odkud samotná linearita pramení. V našem současném pojetí to bude zejména kvůli předpokladu proporcionality vstupů a výstupů v každém základním procesu (výrobním či jiném v závislosti na dané situaci) a kvůli tomu, že výsledek simultánních aktivit je součtem výsledků jednotlivých aktivit. To ale rozhodně neznamená, že by např. příslušná výrobní funkce měla být také lineární, viz [331, str. 1–12] a [117, Podkapitola 2.2]. Možná by bylo potřeba o smysluplnosti tohoto přístupu ekonomy ještě nějaký čas přesvědčovat, ale *špetka praxe je lepší než tuna teorie*, takže zrod lineárního programování přišel zcela nezávisle na rozvoji různých ekonomických teorií – lineární programování vzniklo z čistě praktických a empirických potřeb.

Průkopníkem lineárního programování je bezpochyby sovětský matematik Leonid Kantorovič[7], který se sice na počátku své matematické dráhy věnoval velmi teoretickým problémům, avšak v důsledku různých ideologických represí (obzvláště Stalinem spuštěné tzv. Velké čistce) ve 30. letech 20. století které postihly některé jeho kolegy (zejména Nikolaje Luzina), přestal publikovat v zahraničí a stáhl se z mezinárodní matematické komunity. Avšak velmi brzy jeho pozornost připoutaly zcela jiné problémy.

> *V roce 1936–1937… jsem cítil určitou nespokojenost s matematikou. Ne proto, že by moje práce byla nezajímavá nebo neúspěšná, ale proto, že svět čelil podivnému hrozivému hnědému moru – německému fašismu. Bylo jasné, že několik let bude probíhat velmi tvrdá válka, která ohrozí civilizaci. A cítil jsem zodpovědnost, chápal jsem, že výjimeční lidé musí něco udělat… Jasně jsem vnímal, že stav ekonomických řešení je slabé místo, které snižuje naši průmyslovou a ekonomickou sílu.*[8]

Tato frustrace jej přiměla napsat v roce 1937 článek o přidělování tištěných výrobků, který se zabýval nedostatkem knih a neefektivitou v zásobování různými druhy literatury, a zároveň v něm navrhoval opatření, která by uspokojila poptávku a vymýtila stávající nedostatek. Tento příspěvek z něj později učinil zakladatele lineárního programování a hluboce proměnil roli matematika v Sovětském svazu (a na Západě).

Jako vedoucí Matematické sekce Institutu matematiky a mechaniky na Leningradské univerzitě byl ve spojení i s kolegy–inženýry a v roce 1938 jej požádal o pomoc místní trast výrobců překližek s praktickým výrobním problémem. Při výrobě pěti různých druhů laminované překližky se využívalo osm soustruhů k loupání dýhy z ruských bříz. Jenže se začala zvyšovat poptávka po tomto zboží, které se využívalo při stavbě levných sovětských letadel pro práškování plodin a válečné použití, což trast postupně dostávalo pod nepříjemný politický tlak kvůli nutnosti zvýšení produkce. Trast se snažil zefektivnit svoji produkci, ale nedařilo se mu to, a tak se obrátili na Kantoroviče. Ten předložil tento problém skupině matematiků na svém oddělení, když ale nikdo nedokázal nabídnout odpověď, pustil se do řešení sám. Zpočátku to vypadalo celkem jasně, ale ukázalo se, že to až tak snadné nebude: Kantorovič odhadoval, že pro určení každého možného řešení tohoto problému s osmi stroji a pěti různými překližkami by bylo nutné projít téměř miliardu lineárních rovnic s dvanácti neznámými.

Ekonomika se nachází v rovnováze právě tehdy, když jsou v rovnováze všechny tyto jednotlivé trhy, a tedy všechny uvedené rovnice rovnováhy platí současně. Změna některé veličiny na jakémkoliv trhu pak působí na změny rovnovážných úrovní na ostatních trzích. Walras nehledal podstatu ceny, ale bral cenu jako činitele rovnováhy. Podle něj je rovnováha odvozena z chování spotřebitele. Toto pojetí je ale kritizováno kvůli své statičnosti a nereálným předpokladům. Walrasův vliv a význam pro pozdější vývoj ekonomické teorie tak spočívá především v tom, že byl prvním, kdo se pokusil vytvořit exaktní model zohledňující vzájemné souvislosti veškerých relevantních procesů v ekonomice. J. A. Schumpeter nazval Walrasovu teorii *Magnou Chartou* ekonomie (to je parafráze na slavnou *Magnu chartu libertatum* neboli *Velkou listinu práv a svobod* vydanou v roce 1215 v Anglii) a mnohé ekonomy dodnes oslňuje grandióznost a matematická krása tohoto projektu, kterému se dá vyčítat snad jen to, že nemá mnoho společného s realitou. Ano, Walrasův model je cenný především z historického pohledu, avšak dodnes se používá ve všech současných aplikacích obecné rovnováhy. Viz [220, str. 26] a [335, str. 9].

6

John von Neumann (28. prosince 1903 – 8. února 1957; rodným jménem Margittai Neumann János Lajos) by maďarsko-americký matematik, fyzik, počítačový vědec (informatik) a polyhistor. Významně přispěl k rozvoji funkcionální analýzy, teorie množin, numerické analýzy, statistiky a mnoha dalších matematických disciplín stejně jako kvantové fyziky, ekonomie, informatiky a hydrodynamiky. Byl průkopníkem využití matematické teorie operátorů v kvantové mechanice, což považoval za svůj největší přínos. Klíčovou roli sehrál také v rozvoji teorie her i celulárních automatů a digitálních počítačů. Bývá označován za *posledního představitele velkých*

To je trochu moc, a tak Kantorovič nad tímto problémem stále přemýšlel. Brzy si uvědomil, že nejde jenom o nějaké překližky, ale že se jedná o ukázku mnohem obecnějšího problému: dala by se najít řada dalších aplikací v řízení výroby, inženýrství apod., jenom mít nějaký efektivní způsob, jak to vyřešit. Začal proto s velmi jednoduchým příkladem s jedním výrobkem, který musí být frézován a soustružen řadou různých strojů, a cílem bylo maximalizovat jistou lineární funkci při splnění několika různých podmínek vyjádřených soustavou lineárních rovnic, díky čemuž se výpočet výrazně zjednodušil. Řešení tohoto zjednodušeného problému napočítal ručně a pustil se do zobecňování problému tak, že zvyšoval počet strojů a výrobků. Nakonec se mu podařilo najít obecné matematické řešení, které mohlo maximalizovat produkci různých překližek v poměrech, v jakých je požadovali zákazníci. Toto je vůbec první příklad problému formulovaného jako úloha lineárního programování. Tím to ale neskončilo.

S vědomím, že jeho metoda je využitelná i v mnoha dalších problémech dalece přesahujících sovětský překližkový průmysl, pustil se Kantorovič i do jiných praktických aplikací: nejdříve řešil minimalizaci dřevěného odpadu, pak vyvinul metodu pro optimalizaci střídaní plodin na polích a kvůli omezenému množství dostupného paliva i způsob minimalizace nákladů na přepravu zboží. Díky těmto aplikacím zahrnujícím zemědělství, výrobu a služby mohl Kantorovič poukázat na široké využití lineárního programování. Vlastně mnoho ekonomických problémů je spojeno s nějakou optimalizací za jistých podmínek, a pokud se je podaří matematicky aproximovat v lineárním tvaru, pak se lineární programování ukazuje jako velmi efektivní nástroj. Navíc se sovětská ekonomika ve 30. letech 20. století soustředila na maximalizaci produkce z omezených zdrojů, takže se tento přístup jevil jako velmi slibný. Kantorovič své výsledky včetně jednotlivých aplikací prezentoval na setkání Matematické sekce Institutu matematiky a mechaniky Leningradské univerzity a následně i na zvláštním setkání se zástupci průmyslu kvůli možnému praktickému využití. Všechny to velmi zaujalo a v úvodu rozšířeného přepisu stenografického záznamu z tohoto jednání nalezneme i následující slova odvolávající se soulad jeho metody s usnesením sjezdu komunistické strany SSSR.

> *Obrovské úkoly stanovené v plánu na třetí pětileté období vyžadují, abychom dosáhli co nejvyšší produkce na základě optimálního využití existujících rezerv průmyslu: materiálů, práce a zařízení.*
> *Jsou dvě cesty zvýšení efektivnosti práce v dílně, podniku nebo průmyslovém odvětví. Jedna cesta je pomocí různých zlepšení technologií, tj. nového zlepšení jednotlivých strojů, změn v technologických procesech a objevení nových, lepších druhů surovin. Druhá cesta – doposud mnohem méně používaná – je zlepšení organizace plánování a výroby. Sem patří otázky rozdělení práce mezi jednotlivé stroje podniku nebo mechanismy, rozdělování objednávek mezi podniky, správného rozdělení různých druhů materiálů, paliva a ostatních faktorů. Oba jsou jasně zmíněni v usneseních 18. sjezdu strany. Je tam uvedeno, že „nejdůležitější pro naplnění cílů programu růstu výroby v období třetí pětiletky je… plošný rozvoj práce na propagaci nejmodernější technologie a vědeckou organizaci výroby.“ Tedy jsou vytyčeny obě linie zmíněné výše: stejně jako zavádění nejmodernější technologie je zdůrazněna role vědecké organizace.*[9]

Vypadalo to vskutku nadějně. Předložené výsledky byly publikovány v brožuře vydané Leningradskou univerzitou v nákladu 1000 kusů, které byly rozeslány politikům a akademikům v sovětském svazu. To už není jen o nějakém využití v ad hoc problémech, ale jedná se o první systematickou aplikaci lineárního programování na ekonomické problémy. Vlastně bychom mohli říci, že se zde zrodila nová autonomní matematická disciplína. Kantorovič se vydal na dlouho a leckdy strastiplnou cestu, která vyvrcholila ziskem Nobelovy ceny za jeho roli při rozvoji institucí matematické ekonomie v Sovětském svazu.

Lidé z Lidového komisariátu pro dopravu se ozvali, že mají podobný problém pro minimalizaci najetých kilometrů na železnici v dobách války, ale jinak se žádná výraznější odezva nekonala. Kantorovič později (1942) publikoval samostatný článek týkající se dopravního problému, viz [309]. Vzhledem k válečnému období a utajování čehokoli, co by mohlo nepříteli dát nějakou výhodu, byl článek napsán čistě teoretickým jazykem s náznakem dvou možných aplikací (optimalizace rozmístění skladů na železnici či minimalizace nákladů při nivelizaci půdy), avšak postrádající nějakou efektivní metodu pro jejich řešení.

Jenže Kantorovič se se svými výsledky ocitl ve špatnou dobu (druhá světová válka) na špatném místě (SSSR a jeho komunistický režim). V roce 1938 předseda Rady lidových komisařů Molotov jednoduše zakázal jakoukoli diskusi ekonomů o cenách a brzy došlo i na zdrcující oficiální reakci na Kantorovičovu práci z roku 1939. Kolega Vershik vzpomíná, že vláda totiž velmi rychle zakázala diskusi o Kantorovičově průlomovém objevu.

> *Tento vývoj událostí hrozil zničit celý směr výzkumu a také zničit jeho autora v tom nejpřímějším smyslu. Až o mnoho let později se ukázalo, jak závažná byla obvinění a hrozby vysokých vědeckých a ideologických činitelů. Toto veto existovalo až do roku 1956 a netýkalo se to pouze ekonomických záležitostí, ale dokonce i matematických aspektů prací Leonida Vitalijeviče.*[10]

Kantorovič měl proto velké štěstí, že nebyl poslán někam „pryč“ nebo něco ještě horšího. Jednou z mnoha Stalinových (tehdejší vůdce SSSR) paranoií totiž byla značná nedůvěra v ekonomy a zvláště ty matematické. Řekl, že plánované hospodářství SSSR dosáhlo „závratného úspěchu“, a proto bude jakákoli kritika považována za protisovětskou. Místo toho volal po tom, co označoval za „politickou ekonomii“, která měla poskytovat ideologickou podporu stranické politice. Jenže matematická ekonomie mohla být vnímána jako velmi nebezpečný směr, ve kterém je odstraněna potřeba ideologie a naopak jsou podpořeny samoregulační tržní mechanismy. Navíc optimalizace technických řešení by mohla ohrozit rozhodovací pravomoc představitelů komunistické strany a státu. Každá práce, která používala termín „ceny“, navíc zpochybňovala základní marxisticko-leninské dogma, kde byla práce považována za jediný skutečný zdroj hodnoty. Kantorovičova práce ale byla plná těchto potenciálních chyb a byla odsouzena jako „kapitalistická apologetika“[11]. I když některým chybám se snažil předcházet. Např. se vyhýbal pojmu „ekonomika“ a hovořil raději o organizaci a plánování výroby. Samotná metoda řešení je velmi úzce spojená s metodu Lagrangeových multiplikátorů, které však Kantorovič nazýval *řešící násobitelé*[12]. Jejich skutečný ekonomický význam byl odhalen až o mnoho let později a zatímco Koopmans o nich hovořil jako o *stínové ceně*, sovětští ekonomové používali označení *objektivně určené ocenění*.

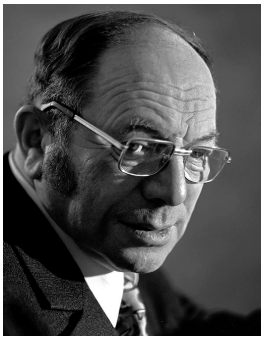
Pro Kantoroviče se tedy zdálo bezpečnější zaměřit svou práci spíše na praktické problémy v průmyslu. Jeho bývalý student řídil Egorovu továrnu na železniční vozy (pojmenovanou po dalším sovětském matematikovi) a dychtil po optimalizaci řezacích technik. Obrátil se na univerzitu s žádostí o radu a Kantorovič sestavil tým, který doporučil využít možnosti lineárního programování pro zlepšení nehospodárných postupů při obrábění kovů. Doporučené nové postupy fungovaly dobře. Jenže závod sloužil jako největší dodavatel železného šrotu pro ocelárny a snížené plýtvání znamenalo, že pro výrobu oceli bylo k dispozici jen málo šrotu (tento typ mezioborového problému byl jedním z těch, na jejichž řešení prostřednictvím svých vnitroprůmyslových nástrojů pracoval Leontief[13] , o čemž však sověti nevěděli.). Kantorovič byl obviňován z toho, že způsobil nedostatek materiálu pro výrobu oceli, a byl předvolán, aby se dostavil do Leningradského regionálního ústředí komunistické strany na základě obvinění ze „spoluviny“. Naštěstí už tehdy měl v armádě pár mocných přátel. Když jiná složka sovětské vlády odhalila, že Kantorovič byl hluboce zapojen do přísně tajného vojenského projektu (stavba tanků a vypracování vzorů pro kladení min), obvinění byla tiše stažena. Problémy však nebyly u konce. Egorovova továrna sice výrazně zlepšila svůj provoz na základě Kantorovi-

matematiků, kteří se věnovali čisté i aplikované matematice a během své kariéry oba tyto směry rozvíjeli, viz [216, str. 89]. Už v průběhu jeho života z něj byla skutečná matematická legenda. V roce 1924 pobýval na univerzitě v Göttingenu a od 30. let je spojen především s univerzitou v Princetonu. Po vypuknutí 2. světové války se stal jednou z hlavních postav *projektu Manhattan*, jehož cílem bylo vyvinutí atomové bomby. Měl také vášeň pro automobily, ovšem ne pro to, aby se v nich „šťoural“, ale aby je řídil, jako by to byly těžkotonážní tanky. Každý rok se v Princetonu objevil s novým kouskem, neboť vzhledem ke způsobu, jakým jezdil, mu žádné auto nemohlo vydržet déle. Pravidelně býval zatčen za překročení povolené rychlosti a některé z jeho nehod se staly legendárními. Jedna z princetonských křižovatek byla po nějakou dobu známá jako *Von Neumannův roh* kvůli množství nehod, které tam způsobil. Jednou vyšel z úplně zdemolovaného auta s jednoduchým vysvětlením: „Jel jsem ulicí. Stromy na pravé straně mě míjely pravidelně rychlostí 60 mil za hodinu. Najednou se mi jeden z nich postavil do cesty. Bum!“, viz [67, str. 96].

Von Neumann byl vskutku geniální a pověstný bleskurychlými mentálními výpočty, takže bylo těžké nad ním vyzrát s nějakou hádankou. Traduje se, že mu jedno dne někdo dal následující hádanku. Dva cyklisté začínají na opačných koncích rovné silnice dlouhé 20 kilometrů. Každý cyklista jede proti druhému rychlostí 10 km/h. Když začnou, moucha sedící na předním kole jednoho z kol vzlétne a uhání rychlostí 15 km/h směrem k druhému kolu. Jakmile se tam dostane, okamžitě se otočí a svišťí zpět k prvnímu kolu, pak zpět k druhému a tak dále. Stále létá tam a zpět, dokud nakonec není zmáčknuta oběma pneumatikami, když se přední kola srazí. Jak dlouhou trasu takto moucha nalétala? Von Neumann okamžitě odpověděl: „15 kilometrů“. To samozřejmě zklamalo tazatele očekávajícího delší přemýšlení a odvětil „Ach, ty znáš ten trik.“ „Jaký trik?“ zeptal se von Neumann, „Jenom jsem sečetl nekonečnou řadu.“

Na druhou stranu pro některé byl von Neumann zároveň spíše blázen a děsil je. Poutavé vyprávění o jeho životě kombinující fakta a fikci nabízí kniha [340], jejíž název odkazuje na první programovatelný počítač MANIAC, u jehož zrodu stál právě von Neumann.

[7]



Leonid Vitalijevič Kantorovič (19. ledna 1912 – 7. dubna 1986) byl sovětský matematik a ekonom. Kantorovičovy první práce se dotýkaly problematiky operací na projektivních množinách a měly zcela teoretický charakter. Ve svém vědeckém vývoji se Kantorovič postupně dostával k „aplikovanějším" otázkám, nikdy však nepřerušil aktivní kontakt s nejabstraktnější matematickou teorií. Skutečnost, že matematik Kantorovič se stal také ekonomem, by bylo nesprávné vidět v tom, že dokázal úspěšně „algoritmizovat" úlohy, vznikající v hospodářské praxi. Podstatné jsou obecnější implikace jeho práce, umožněné jeho mimořádnými schopnostmi výstižně formalizovat ekonomické problémy a adekvátně interpretovat relevantní výsledky. Nezůstal však jen u optimalizace dílčích úloh, snažil se též uplatnit vyvinuté optimalizační procedury a principy na úrovni celého národního hospodářství. Kantorovičovi se dostalo nejvyšších ocenění za jeho výsledky v ekonomickém výzkumu jak v SSSR, tak na mezinárodní úrovni. Jeho cesta k těmto poctám však nebyla lehká a obhajoba a prosazování vědeckých prvků v procesu řízení si vyžádalo hodně úsilí. Je jediným (ryze) sovětským držitelem Nobelovy ceny za ekonomii, viz [499].

[8] *In 1936–1937… I felt some dissatisfaction with mathematics. Not because my work was uninteresting or unsuccessful, but because the world was facing a strange menacing brown plague – German fascism. It was clear that there was going to be a very hard war for some years that would threaten civilization. And I felt a responsibility, understanding that people out of the ordinary must do something… I had a clear perception that the state of economic solutions was a weak spot that was reducing our industrial and economic power.* Viz [71, str. 260].

čovy rady a zvýšila svoji produkci na zhruba 94 procent, jenže v rámci sovětského plánování přišel nový požadavek: každý závod musí každý rok přidat 7 procentních bodů ke své úrovni účinnosti. A tak Kantorovičův institut musel ministerstvu napsat, že 101 procentní úroveň produkce nebude nikdy možná.

Navzdory zuřící válce sepsal Kantorovič v roce 1942 první verzi krátké ekonomicky zaměřené knihy *The Best Use of Economic Resources*, ve které využívá optimalizační metody v širokém spektru ekonomických problémů, viz [311]. Kniha měla napomoci k rozšíření jeho nových idejí a také předpovídala obrovskou výzvu, která měla stát před poválečnou sovětskou ekonomikou ve snaze o efektivitu při absenci mechanizmů západních trhů. Kantorovič navrhl rámec, kde by požadavky na výrobu byly optimálně rozděleny mezi podniky v souladu s nejnižšími možnými výrobními náklady. Každé továrně by byla přidělena výroba toho zboží, které mělo nejvyšší čistý produkt. Poukazoval i na soulad této strategie marxistickou pracovní teorií hodnoty. Později to rozšířil na obecnější přístup k plánování, přičemž rozlišoval mezi krátkodobým plánováním (se současnými metodami výroby) a dlouhodobým plánováním (kde byl prostor pro budování nových továren, infrastruktury a dovedností). Řešení krátkodobého plánu bylo technickým úkolem, ale nastínění cílů dlouhodobého plánu by vyžadovalo politické rozhodování. Také potřeboval znát vzájemné propojení jednotlivých výrobních sektorů, k čemuž v tehdejší době chyběla potřebná data.

Kantorovič v této své práci rozpoznal významné nedostatky v sovětské hospodářské výkonnosti jako důsledek nevyužité pracovní síly, nevyužitého vybavení, plýtvání surovinami a palivem, uspěchaných projektů, jen aby byly splněny termíny, zpoždění dodávek, zbytečného držení přebytečného materiálu na skladě a dlouhotrvající stavební práce. Stěžoval si na nepřesné alokační mechanismy a na nesoulad v objednávkách a požadavcích a vinil je z nedokonalého plánování a účetnictví. Vyhýbal se přímému poukazování na zjevné problémy vládnutí vlastní socialistickému systému, ale přesto to byla odvážná věc – ačkoli Kantorovičovým záměrem bylo postupné zlepšování v rámci současného systému[14]. Jeho naděje do budoucna byla v lepší socialismus s pomocí lepších matematických nástrojů.

V roce 1942 také své výsledky prezentoval, jenže se dočkal velmi negativní reakce. Později dokonce vzpomínal, že na setkání Státní plánovací komise (tzv. Gosplan) v roce 1943 vystoupil ekonom a statistik Boris Jastremskij s komentářem

> *Hlavním úspěchem Kantoroviče je použití matematické analýzy pro optimalizaci. Dala by se použít například pro optimální řezání listů překližky. V jiných případech se však optimální využití materiálu (kovu) doporučeného Kantorovičem setkalo s nepřátelským přijetím, protože závod by měl menší množství využitelného odpadu! Je ale potřeba zjistit, zda je Kantorovičův příspěvek originální, nebo není. A vidíme, že mnoho let před Kantorovičem jej přednesl Pareto. Kdo je Pareto? Byl oblíbeným matematikem Benita Mussoliniho. V důsledku toho je Kantorovič u nás agentem fašismu a měl by být náležitě potrestán. Především by mu mělo být zakázáno šířit jeho teorii a učit studenty.*[15]

Od odeslání do některého z gulagů (tehdejší nápravně–pracovní tábory) jej patrně ochránilo to, že v té době již byl slavným sovětským matematikem, měl válečná vyznamenání a spolupracoval se sovětskou armádou na řešení naléhavých válečných problémů. Nicméně raději dočasně ukončil svou ekonomickou práci s tím, že se cítí depresivně a možná se bude muset ekonomie úplně vzdát. Celou dobu věřil, že jeho práce může zlepšit ekonomický život sovětského lidu. O politicky sporných otázkách psal opatrně a snažil se vyhnout ideologickému střetu. Když ho ale odmítli, Kantorovič, který celý život strávil v atmosféře hospodářského rozvratu, vojenského násilí, náboženského pronásledování a politické represe, věděl, kdy se stáhnout a zůstat zticha.

> *Všichni říkali, že je potřeba tuto práci prozatím opustit. Bylo nebezpečné v tom pokračovat – jak jsem následně zjistil, můj strach nebyl neopodstatněný. Samozřejmě to pro mě byla těžká rána, protože jsem v to hodně věřil. Nějakou dobu jsem dokonce trpěl depresemi.*[16]

Kantorovičovi kolegové jej popisovali jako plného rozporů: laskavého a mírného, přesto houževnatého a rázného. Fyzicky byl malý, plešatý, obrýlený, kulatý a nevýrazného vzhledu. Měl rád kanadské žertíky. V soukromí byl introvert a nebyl dobrý ve společenské konverzaci, ale na veřejnosti dokázal působit velmi extrovertně. Krása Kantorovičových matematických úspěchů možná byla i v překvapivém rozporu s jeho způsobem vystupování. Pro posluchače bylo nesmírně těžké mu porozumět: sám mnohokrát zdůrazňoval, že není schopen přenést na své publikum krásu, kterou cítil. Přesto jej studenti označovali za skvělého přednášejícího[17]. Známý byl také jeho neobvyklý styl při účasti na odborných seminářích. Seděl v první řadě a zřejmě usnul, pak se ale na konci náhle probudil a kladl zkoumavé otázky, které oživili celou seminární místnost. Kolegové s obdivem hovořili o jeho schopnosti vyslechnout všechny argumenty, dojít k tabuli, smazat všechny rovnice a na místě napsat nový důkaz. Jeho vysvětlení mohlo působit matoucím dojmem, ale výsledek byl nevyhnutelně inovativní a přesný.

Po Stalinově smrti v roce 1953 se poměry začaly pomalu uvolňovat a ke slovu se postupně dostávala progresivní sovětská inteligence. Kantorovičovi bylo umožněno opět učit na Leningradské univerzitě, kde vedl kurz Ekonomických výpočtů. V roce 1958 se stal korespondenčním členem Sovětské akademie věd a v roce 1959 mu bylo umožněno znovu publikovat jeho článek z roku 1939 a také rozšířenou verzi jeho průkopnické knihy, i když i v té době musela být doplněna o předmluvu kritizující některé ideologické chyby. V roce 1965 pak byla s Koopmansovou[18] pomocí přeložena i do angličtiny. Post-stalinský sovětský svaz tak konečně přestal přehlížet jeho skvělé výsledky, které mohl spatřit i celý svět. Pozdě, ale přece – chtělo by se až říci *Post bellum auxilium*[19], neboť v té době již „západní svět“ toho o lineárním programování a jeho síle věděl skutečně mnoho. Toho si byl Kantorovič samozřejmě velmi dobře vědom, o čemž nás může přesvědčit i komentář z *The New York Times* z května 1959.

> *The scholar, Professor L. V. Kantorovich, said in a debate that Soviet economists had been inspired by a fear of mathematics that left the Soviet Union far behind the United States in applications of mathematics to economic problems. It could have been a decade ahead.*[20]

Bohužel… nicméně Kantorovič bude navždy tím prvním, kdo si uvědomil, že některé součásti výrobního procesu lze velmi dobře matematicky popsat, a tudíž je lze i prakticky vypočítat a zefektivnit. V roce 1975 společně s Koopmansem obdrželi dokonce *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za jejich „příspěvek k teorii optimálního rozdělení zdrojů“.

Kantorovič se neztratil ani v době druhé světové války, která výrazně změnila směřování hlavního vědeckého zkoumání. I když to rozhodně nebyl vojenský typ a měl problémy s armádní disciplínou. Nastoupil na Vyšší školu námořního inženýrství, kde učil matematiku a později se stal vedoucím katedry. Ačkoli již v té době byl známým vědcem, byl mobilizován v hodnosti svobodníka a musel podstoupit i všechna vojenská cvičení. Nebylo to vůbec lehké období. Měl problémy nosit pušku na rameni. Po 6kilometrovém pochodu nedokázal při střelbě zasáhnout cíl, což velitel komentoval slovy

> *Nemáš probírat integrály, tady musíš přemýšlet!*

Naštěstí to netrvalo příliš dlouho a z Kantoroviče se stal plukovník. Byl součástí tajného projektu Arzamas-16[21], jehož cílem bylo vyvinout jadernou a později vodíkovou bombu. Byl to

[9] *The immense tasks laid down in the plan for the third Five Year Plan period require that we achieve the highest possible production on the basis of the optimum utilization of the existing reserves of industry: materials, labor and equipment.*

There are two ways of increasing the efficiency of the work of a shop, an enterprise, or a whole branch of industry. One way is by various improvements in technology; that is, new attachments for individual machines, changes in technological processes, and the discovery of new, better kinds of raw materials. The other way – thus far much less used – is improvement in the organization of planning and production. Here are included, for instance, such questions as the distribution of work among individual machines of the enterprise or among mechanisms, the correct distribution of orders among enterprises, the correct distribution of different kinds of raw materials, fuel, and other factors. Both are clearly mentioned in the resolutions of the 18th Party Congress. There it is stated that „the most important thing for the fulfillment of the goals of the program for the growth of production in the Third Five Year Plan period is… the widespread development of work to propagate the most up-to-date technology and scientific organization of production.“ Thus the two lines of approach indicated above are specified: as well as the introduction of the most up-to-date technology, the role of scientific organization is emphasized. Viz [310, str. 367–368].

[10] *This turn of events threatened to bury the whole research direction, as well as to bury his author, in the most direct sense. Only many years later it became known how serious were the accusations and threats of high scientific and ideological officials. This veto existed up to 1956. And it applied not only to economic matters, but even partly to the mathematical aspect of L.V.'s works.* Viz [497, str. 2].

[11] *Apologetika* neboli (zaujaté, vyhraněné) obhajování, obhajoba, bránění, obrana nějaké myšlenky, skutečnosti, jevu.

[12] Angl *revolving multipliers.*

[13] Wassily Wassilyevich Leontief (5. srpna 1905 – 5. února 1999), byl rusko-americký ekonom známý především díky jeho „input–output“ analýze a studiu efektu změn v jednom ekonomickém sektoru na jiné. V roce 1973 získal *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za „rozvoj

údajně Kantorovič, který svými výpočty zásadní měrou přispěl k úspěchu tohoto projektu. Zřejmě právě za tyto své výsledky obdržel v roce 1949 Stalinovu státní cenu, ačkoli oficiálně to bylo za jeho práci o aplikacích funkcionální analýzy[22].

Svými výpočty také nepřímo přispěl k záchraně tisíců lidských životů. Navzdory uzavřenému Paktu Ribbentrop–Molotov se Adolf Hitler rozhodl v rámci Operace Barbarossa v červnu 1941 napadnout SSSR, čímž došlo k otevření východní fronty druhé světové války. Tyto boje jsou také známy jako Velká vlastenecká válka a jejich součástí byla i blokáda Leningradu (8. září 1941 – 27. ledna 1944), což byla jedna z klíčových bitev celé druhé světová války a další z důkazů naprosté zrůdnosti nacistického režimu. Nacistické Německo totiž chtělo město nejenom dobít a zničit, ale i vyhladovět jeho obyvatelstvo. Život v Leningradu v té době byl skutečně krutý a z hlediska obětí jde asi o jedno z nejkrvavějších obléhání v dějinách lidstva.

> *Od naprosté blokády uchránila město jen ledová silnice vybudovaná na zamrzlém Ladožském jezeře, která umožňovala dodávky jídla, zbraní a paliva (tzv. Cesta života, anglicky Road of Life a rusky doroga zhizni). Na konci ledna 1942 byly tři desítky tanků přemístěny na východní břeh jezera. Při testovací jízdě ale led 48tunový stroj neudržel a i s posádkou a nákladem šel ke dnu. Proto z nich musely být demontovány věže (tím se tank odlehčil cca o 1/3), které byly taženy lany na vlecích. Praskání ledu, připomínající praskající dělostřelecké granáty, přehlušil hluk motoru, led se pod mnohatunovou hučící ocelovou bednou prohýbal a zdálo se, že tank šplhá do kopce. Psychická i fyzická zátěž pro posádku byla enormní a nedalo se to vydržet více než třicet minut… Tanky jezdily po jednom. První tank se o od vytyčené trasy odchýlil o mnoho kilometrů na sever. Led byl na mnoha místech poničen bombami a v případě havárie by posádka zůstala bez pomoci. Tohle všechno jen zbytečně zvyšovalo riziko celého transportu, které už tak bylo velmi vysoké. Mnohem bezpečnější by bylo jet v koloně. Ale s jakými rozestupy?*
> *To byl úkol, který čekal na Kantoroviče. Dnes je možné potřebné informace dohledat v tabulkách, ale tehdy nic takového nebylo. Kantorovič sestavil algoritmus pro výpočet potřebné tloušťky ledu a rozestupů. Nebylo to nic snadného, neboť povětrnostní podmínky na trase dlouhé 220 kilometrů mohly být velmi rozdílné. Správnost jeho výpočtů znamenala rozdíl mezi životem a smrtí. Kantorovič stanovil potřebnou vzdálenost mezi vozidly jako nejméně 27 metrů a pak sám nasedl do prvního tanku…*

Zájemcům o mnohem více podrobností z Kantorovičova života a o jeho práci doporučujeme např. [72, Chapter 5], [13, str. 163–168] a [71, 199, 540].

Navzdory nezměrnému a nezpochybnitelnému Kantorovičovu přínosu k oblasti parametrické optimalizace (zejména lineárnímu programování) ten nejzásadnější zlom, který skutečně rozpohyboval celou mašinérii, je ale spojen s jiným jménem. Jak jsme se již zmínili v úvodní kapitole, během války se nástroje operačního výzkumu využívaly ke snižování vlastních nákladů a zvyšování ztrát nepřítele. Typickým příkladem může být např. *dopravní problém* pro rozesílání zboží ze zásobovacích skladů až na místo určení, který můžeme využít buď zcela mírově k plánování rozvozu a zamezení dodávek *ad hoc* nebo v dobách války pro demolici klíčových kolejových tratí či efektivnější přesun armádních jednotek, jejich zásobování apod. Tento problém se asi poprvé objevuje v práci z roku 1930 u sovětského matematika A. Tolstoïe[23], na kterou později navázal i Kantorovič, viz [309, 482, 483]. Jakožto úloha lineárního programování byl klasický dopravní problém zformulován v článku [267] z roku 1941 od Hitchcocka[24], a to včetně návrhu možného řešení. Podobnému problému se během druhé světové války věnoval i Koopmans, když pracoval pro *Britsko–americký úřad pro kombinovanou přepravu*[25]. Ale zdá se, že tyto výsledky neměly na skutečný zrod lineárního programování žádný vliv – zde příliš

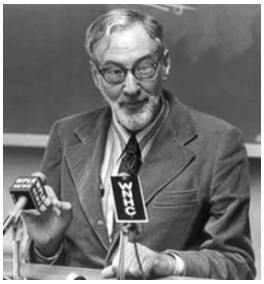
input–output analýzy a její aplikaci v důležitých ekonomických problémech". Tutéž cenu (za jiné aktivity) získali i čtyři jeho doktorští studenti: Paul Samuelson (1970), Robert Solow (1987), Vernon Smith (2002) a Thomas Schelling (2005). Viz také [72, Chapter 6].

[14] V tom byl obrovský rozdíl mezi ním a Leontiefem, který je také úzce spojen se Stalingradem/Petrohradem, kde v mládí žil a studoval. V roce 1923 mu přijali článek v časopise vydávaném Ruskou akademií věd, ale když jej cenzor zamítl jako příliš kontroverzní kvůli odkazům na Hegela a Kanta, rozhodl se emigrovat. Uvědomil si totiž, že když mu zamítnou i takový neškodný teoretický článek, tak by zde nemohl efektivně vědecky pracovat. SSSR opustil v roce 1925. Viz [306, 307].

[15] *The main achievement of Kantorovich is the use of mathematical analysis for optimization. It could be used, for instance, for the optimal cutting of sheets of plywood. However, in other cases the optimal use of material (metal) recommended by Kantorovich met a hostile reception because a plant would have a smaller amount of usable waste! But it is necessary to find out whether Kantorovich's contribution is original or not. And we see that it was presented by Pareto many years before Kantorovich. Who is Pareto? He was a favorite mathematician of Benito Mussolini. Consequently, Kantorovich is an agent of fascism in our country and he should be punished appropriately. First of all, it should be forbidden for him to spread his theory and to teach students.* Viz [13, str. 164].

[16] *Everyone was saying that it was necessary to leave this work for the time being. It was dangerous to continue it – as I subsequently found out, my fear was not unfounded. Of course, this was a severe blow to me as I had put great faith in it. For some time I was even in a state of depression.* Viz [71, str. 263].

[17] Nezapomínejme na slova slavného fyzika Michaela Faradaye, který vytvořil i školu popularizace vědy pro veřejnost: *Přednášky, které jsou opravdu populární, ve skutečnosti nikdy nebudou naučné, zatímco skutečně naučné přednášky, nikdy nebudou populární!*

[18] 

Tjalling Charles Koopmans (28. srpna 1910 – 26. února 1985) byl nizozemsko–americký matematik a ekonom. První tři léta na univerzitě v Utrechtu se věnoval především matematice, poté pak studiu teoretické fyziky. První dvě odborné statě publikoval z oblasti kvantové mechaniky a doktorát získal v roce 1936 v oboru matematické statistiky na univerzitě v Leidenu za práci *Lineární regresní analýza ekonomických časových řad*. Velká deprese počátku 30. let probudila v Koopmansovi zájem o ekonomii, zejména o matematickou ekonomii. V současnosti patrně nejvíce ceněnými jsou výsledky, které Koopmans dosáhl v oblasti modelování výrobního procesu a v makroekonomické teorii ekonomického růstu. Tradiční nástroj ekonomů pro popis výrobního procesu – produkční funkci, založenou na předpokladu jeho efektivního uspořádání a určitého specifického chování ekonomických subjektů, nahradil modelem, který má popisovat výrobní proces nezávisle na institucionální struktuře ekonomiky, na aspiracích ekonomických subjektů. Ukázal souvislost mezi jím zavedeným pojmem výrobní efektivnosti, cenovým systémem a alokací zdrojů v ideální konkurenční ekonomice a tím vytvořil most mezi normativní teorií alokace zdrojů a teorií všeobecné rovnováhy. Významných výsledků dosáhl Koopmans rovněž v oblasti ekonomiky dopravy a ekonometrických metod. Uznávané jsou i jeho zásluhy o rozvoj lineárního programování. Neopomenutelný je také způsob, kterým ve svých pracích uplatňoval matematické modelování. Bývá považován za vzor využívání této metodologie ekonomického výzkumu, za vzor důsledného dodržování pravidel vědeckého postupu, rozhodného odmítnutí argumentů a analogií, kterými je možné přesvědčit a získat, avšak pouze za cenu nepřesnosti nebo skrytých hodnotových postojů. Viz [500].

[19] Doslova *pomoc po válce* (lidově: *s křížkem po funuse*).

nešlo o nalezení nějaké *efektivní* metody řešení daného problému, ale především o vyzdvihnutí a obhájení samotného modelu. Navíc jednotliví autoři o práci ostatních asi ani nevěděli.

Tyto úlohy jsou velmi úzce spojeny s otázkou řešení soustav *lineárních nerovnic*. Všichni velmi dobře známe metody řešení soustav lineárních rovnic, se kterými jsme se setkali už během prvního semestru v základním kurzu lineární algebry a které jsou obvykle spojeny především s Gaussem, Jordanem či Cramerem. Matematická literatura obsahuje tisíce článků s úvahami o technice řešení soustav lineárních rovnic společně s maticovou algebrou, ale co nerovnice? Ty v podstatě až do příchodu *teorie her* v roce 1944 a lineárního programování v roce 1947 nevyvolávaly žádný zájem. Kupodivu. Motzkin se dokonce ve své disertační práci [385] z roku 1936 zmiňuje, že navzdory maximálnímu úsilí byl schopen k tomuto tématu nalézt jenom asi 30 bibliografických odkazů z let 1900–1936 a asi 42 celkově. Např. v roce 1930 se objevily asi pouze 4 články[26], které se věnovaly důkladnější výstavbě teorie lineárních nerovností a které také hodnotily starší vědecké práce. Důkazem toho, že si matematici neuvědomovali důležitost problému hledání řešení soustavy lineárních nerovností, přičemž by se současně minimalizovala i jistá lineární forma, je, že ani v jednom z těchto článků se o takovém problému nehovoří, ačkoli ve starší literatuře zmíněn byl (viz první položku v následujícím seznamu). Do roku 1947 byly publikovány asi pouze 4 články pojednávající o speciálních případech úloh lineárního programování a jejich řešení:

(i) Fourier[27] byl asi první, kdo si uvědomil důležitost řešení soustavy lineárních nerovnic v aplikované matematice (v mechanice a teorii pravděpodobnosti) a systematicky se jim věnoval, i když nešel příliš do hloubky. V článku [183] z roku 1826 je metoda pro jejich řešení, která je vlastně jakýmsi předchůdcem simplexové metody, viz také [280, 511] a Poznámku 2.2.5.

(ii) De la Vallée Poussin[28] zveřejnil v článku [489] z roku 1911 metodu řešení pro Čebyševův problém[29], která je dalším předchůdcem základní metody pro řešení úloh lineárního programování, viz také [172] a pro možné propojení s prací Laplace viz [171].

(iii) Kantorovič a jeho již dříve zmíněná práce z roku 1939.

(iv) Hitchcock a „jeho“ dopravní problém uvedený výše.

Jak ale řešit takové úlohy v případě desítek či stovek proměnných a omezení? Lze k tomu nějak efektivně využít jejich lineární strukturu? Zatímco Kantorovičovi se to už nějak povedlo, západní svět „tápal“. Situace se změnila až díky Dantzigovy a jeho objevu simplexového algoritmu, který zveřejnil v roce 1947.

V černu 1941 před obhajobou své disertační práce přijal Dantzig práci ve Washingtonu jako vedoucího oddělení založeného Pentagonem pro analýzy bitev v útvaru statistické kontroly v rámci armádního letectva Spojených států[30]. Jeho motivací bylo „podílení se na druhé světové válce“, jelikož předpokládal, že USA se brzy do války aktivně zapojí[31]. Vyvinul systém hlášení, díky němuž byly bojové jednotky schopné zaznamenávat počty vzletů, ztracených a zničených letadel, shozených bomb a napadených cílů. Právě zde se Dantzig důvěrně seznámil s přístupem letectva k plánování programů souvisejících činností, což jsou myšlenky, které mu pomohly se sestavením základní struktury lineárních modelů. Mezi jeho kolegy byli např. Robert McNamara (ze kterého se později stal ministr obrany USA a také šéf Světové banky) nebo Warren Hirsch (který se věnoval pravděpodobnosti a v roce 1957 zformuloval domněnku ohledně maximální počtu kroků potřebných pro řešení úloh lineárního programování, viz [522] a Podkapitolu 2.13). Za svoji záslužnou civilní práci v časech války získal Dantzig nejvyšší vojenské ocenění *Exceptional Civilian Service Award* z ministerstva války.

Dantzig se tak stal expertem na *metody programovacího plánování pomocí stolních kalkulátorů*. V roce 1946 se vrátil na Kalifornskou univerzitu v Berkeley, aby formálně splnil všechny studijní požadavky a získal konečně doktorát. V Berkeley mu dokonce na ústavu matematiky nabídli učitelské místo, ale on jej nepřijal z prostého důvodu: „bylo málo placené“. Vrátil se proto do Pentagonu, kde začal pracovat jako matematický poradce v rámci nově ustanoveného

oddělení finanční kontroly letectva USA[32]. Jeho dlouholeté zkušenosti z práce v Pentagonu na lineárním programování z něj učinily silného zastánce reálných problémů jakožto inspirace pro zajímavou a náročnou matematickou práci.

> *Just because my mathematics has its origin in a real problem doesn't make it less interesting to me – just the other way around, I find it makes the puzzle I am working on all the more exciting. I get satisfaction out of knowing that I'm working on a relevant problem. I find that just as much mathematical ingenuity has to go into solving problems from a new developing area as from some old so-called pure math area.*[33]

Své setrvání v Pentagonu ale považoval spíše za dočasné, neboť se přirozeně stále poohlížel i po nějaké akademické pozici. Jenže kolegové ho chtěli udržet, takže přicházeli s různými výzvami ohledně možností mechanizace plánovacích procesů a jednou z nich byla žádost o nalezení způsobu, jak co nejrychleji vypočítat program rozmístění mužstva a vybavení, tréninků a logistické podpory. Díky své matematické průpravě se pustil do sestavování modelu. V tomto ohledu byl inspirován (a zároveň i fascinován) Leontiefovou prací, který v roce 1932 přišel s „Meziprůmyslovým input–output modelem" americké ekonomiky majícím jednoduchou maticovou strukturu.

> *Leontief's great contribution… was his construction of a quantitative model of the American economy, for the purpose of tracing the impact of government policy and consumer trends upon a large number of industries which were imbedded in a highly complex series of interlocking relationships. To appreciate the difference between a purely formal model and an empirical model, it is well to remember that the acquisition of data for a real model requires an organization working many months, sometimes years. After the model has been put together, another obstacle looms – the solution of a very large system of simultaneous linear equations. In the period 1936–1940, there were no electronic computers; the best that one could hope for in general would be to solve twenty equations in twenty unknowns. Finally, there was the difficulty of „marketing" the results of such studies. Hence, from the onset, the undertaking initiated by Leontief represented a triple gamble.*[34]

Dantzig ale s tímto *statickým* modelem příliš spokojen nebyl – postrádal v něm totiž dynamiku umožňující změny v průběhu času. V Leontiefově modelu bylo jednoznačné spojení mezi výrobním procesem a výrobky generovanými těmito procesy (v podstatě totéž v menším měřítku je již v Quesnayho tabulce). Chyběly tomu *alternativní činnosti*! Reálné využití by mělo být ve velkém měřítku se stovkami položek a činností, ale současně to musí být i *spočítatelné*. Jakmile byl model hotov[35], bylo zapotřebí najít praktický způsob výpočtu toho, jaká množství těchto činností je potřeba využít, aby to bylo v souladu s jejich „input–output" charakterem a dostupným množstvím. Jenže na začátku Dantzig *neuvažoval* ve svém modelu žádnou účelovou funkci: explicitní cíle neexistovaly, jelikož plánovači z praxe jednoduše neměli způsob, jak něco takového implementovat. Bez toho se ale dostáváme do skutečně obřích potíží, protože např. při přiřazování každého ze 70 mužů k právě jednomu ze 70 úkolů rozumíme „činností" přidělení i-tého muže k j-tému úkolu s binární hodnotou 0 (úkol není přidělen) nebo 1 (úkol je přidělen), v kterémžto případě dostáváme 2×70 omezení a 70×70 činností. Takový problém nabízí $70! \approx 1{,}12 \times 10^{100}$ možných rozdělení. Ale na základě jakých kritérií z nich vybrat to nejlepší?!? A to ani nemluvíme o tom, jak dlouho by nám vůbec trvalo projít všechny možnosti[36]. Tento jednoduchý příklad ilustruje, proč až do roku 1947 a vlastně i do dnešních dní existuje obrovská propast našimi sny a činy. Můžeme si stanovit naše plány tak, abychom optimalizovali jednotlivé cíle, jenže k jejich dosažení máme nepřeberné množství způsobů (každý

[20] *Učenec, prof. L. V. Kantorovič, řekl v debatě, že sovětští ekonomové byli ovlivněni strachem z matematiky, což zanechalo SSSR daleko za USA ve využití matematiky při řešení ekonomických problémů. Přitom v SSSR mohli být o deset let před USA.* Viz [547].

[21] Jde o uzavřené město, které je jedním z center ruského výzkumu nukleárních technologií a dodnes je nepřístupné pro cizince bez zvláštního povolení. Byla to Sovětská analogie americké národní laboratoře v Los Alamos, a proto o něm často účastníc projektu hovořili jako o „Los Arzamas".

[22] Vždyť o Stalinových skvělých matematických dovednostech nemůže být žádných pochyb. V roce 1965 získat Leninovu cenu sovětské vlády. Toto byla dvě největší státní ocenění, která bylo možné v SSSR získat.

[23] Ten je dalším z těch, o kterých nevíme vůbec nic.

[24] Frank Lauren Hitchcock (6. března 1875 – 31. května 1957) byl americký matematik a fyzik.

[25] *British–American Combined Shipping Board*

[26] Jejich autory byli Stokes, Weyl, Motzkin a Dines společně s McCoyem.

[27] Jean Baptiste Joseph Fourier (21. března 1768 – 16. května 1830) byl francouzský matematik a fyzik, který se nejvíce proslavil zkoumáním Fourierových řad a jejich aplikací v problémech toku tepla. Na jeho počest byla také nazvána jeho jménem Fourierova transformace. Je objevitelem skleníkového efektu (1824). Fourier je jedním ze 72 významných mužů, jejichž jméno je zapsáno na Eiffelově věži v Paříži.

[28] Charles–Jean Étienne Gustave Nicolas, baron de la Vallée Poussin (14. srpna 1866 – 2. března 1962) byl belgický matematik. V roce 1896 jako první (současně a nezávisle na Jacquesu Hadamardovi) podal důkaz *Prvočíselné věty*, která dává hrubý popis rozmístění prvočísel mezi přirozenými čísly.

[29] Jde o minimalizaci maxima absolutní hodnoty jisté odchylky.

[30] Army Air Force's Combat Analysis Branch of Statistical Control

[31] To se stalo 8. prosince 1941, jeden den po japonském útok na Pearl Harbor.

[32] *U.S. Air Force Comptroller*

s různými výhodami a nevýhodami), takže jejich vzájemné porovnávání se zdá být nemožné, natož některé z nich označit jako „to nejlepší“. Nakonec vždy skončíme u intuice a zkušeností (našich či ostatních). Do tohoto stavu se Dantzig dostal v závěru roku 1946, kdy měl sestavený striktně „technicistní“ model zdárně zahrnující všechny běžné souvislosti. Tohle všechno se samozřejmě dělo v době „počítačového pravěku“. Stolní počítače sice k dispozici byly, ale k jejich použití bylo zapotřebí mít přesně formulovaný model, což si vyžádalo *axiomatizaci* celé oblasti. A přesně to udělal i Dantzig v polovině roku 1947. Rozhodl se, že účelová funkce musí být zcela explicitní, a sestavil plánovací problém jako sadu axiomů, které byly dvou druhů: první zahrnovaly vyráběné či spotřebovávané položky a druhá sada obsahovala činnosti či výrobní procesy, do nichž mohly být vkládány nebo z nichž mohly vystupovat položky v pevně daných poměrech, dokud tyto poměry byly vzájemnými nezápornými násobky. Výsledkem pak bylo hledání řešení matematického systému, ve kterém jde o *minimalizaci lineární formy vzhledem k lineárním rovnostem a nerovnostem*. Neotřelost takového přístupu tkvěla právě ve využití lineární formy jakožto účelové funkce. Dantzig tak zavedením *programování v lineární struktuře*[37] zásadně zmodernizoval plánovací proces[38].

Jenže teď byla na řadě odpověď na zcela přirozenou otázku odkazující na Dantzigův rozhovor zmíněný v úvodu této kapitoly: dokážeme takový systém vyřešit? Danztig kvůli tomu dokonce v červenci 1947 vyrazil za Koopmansem, neboť předpokládal, že ekonomové se takovým problémům již věnovali. Koopmans z toho byl samozřejmě nadšený, protože v tom spatřoval velký potenciál směrem k různým ekonomickým aplikacím. Začal rozšiřovat poznatky týkající se modelů založených na lineárním programování mezi mladé ekonomy, což nakonec vedlo až k několika Nobelovým cenám za ekonomii. Jenže Dantzigovi s jeho problémem nepomohl. Dantzig se proto rozhodl zkusit štěstí „sám u sebe“ a vymyslet nějaký vlastní algoritmus. To jej vrátilo zpět k jeho disertační práci na Berkeley sepsané pod vedením Jerzyho Neymana[39], v níž se věnoval dvěma slavným nevyřešeným problémům z matematické statistiky:

> *Jednoho dne [Dantzig] dorazil na přednášku pozdě, a tak si v rychlosti opsal dva problémy ze statistiky napsané na tabuli, neboť se domníval, že jde o domácí úkol. Úlohy se mu však zdály neobvykle obtížné, a tak je vyřešil až po delší době. Když je potom profesorovi [Neymanovi] donesl s omluvou, že domácí úlohu donesl tak pozdě, byl vyzván, aby je hodil na stůl (kde byly kupy papírů, jak už to tak na některých pracovních stolech bývá). Domníval se, že o svých příkladech už neuslyší. Po dlouhé době jej v neděli ráno vzbudilo bouchání na dveře, za nimiž stál jeho profesor se slovy: „Napsal jsem předmluvu k Vašemu článku, přečtěte si to, ať to můžeme zaslat redakci!“ Až tehdy Dantzig zjistil, že vyřešil dva dosud otevřené problémy ze statistiky. A tím byla na světě podstata jeho disertační práce.*[40,41]

Jeden z těchto problémů byl (z dnešního pohledu) o hledání Lagrangeových multiplikátorů pro obecný lineární optimalizační problém s lineárními omezeními v integrálním tvaru. Jeho řešení bylo založeno na jisté geometrizaci související s přístupem skrze sloupce jakési matice místo obvyklých řádků. Toto jej v létě 1947 přivedlo k tzv. *simplexovému algoritmu* a přesvědčení, že to musí fungovat. A vskutku, fungovalo! Viz také [119].

Prvním opravdovým testem simplexového algoritmu bylo určení optimálního jídelníčku s minimálními náklady, díky čemuž snad každá učebnice věnovaná lineárnímu programování začíná obdobným příkladem jako jsme to učinili my v předchozí kapitole. Jenže tady šlo o zcela reálný problém, ve kterém bylo sledováno 9 nutričních hodnot, k jejichž naplnění bylo k dispozici 77 surovin. Jednalo se o klasický *Stiglerův*[42] *problém optimálního jídelníčku* z roku 1945 určený pro průměrně činného muže vážícího 68 kg, jehož cílem bylo nalezení jídelníčku splňujícího doporučení *Národní rady pro výzkum*[43] z roku 1943 uvedená v následující tabulce, viz také [463].

[33] *To, že moje matematika má svůj původ ve skutečném problému, je pro mne nečiní méně zajímavou – právě naopak, zjišťuji, že díky tomu je hádanka, na které pracuji, ještě vzrušující. Jsem spokojený s vědomím, že pracuji na důležitém problému. Zjistil jsem, že do řešení problémů z nově se rozvíjející oblasti musí být vynaloženo právě tolik matematické vynalézavosti jako do problémů z nějaké staré oblasti tzv. čisté matematiky.* Viz [6, str. 305].

[34] *Leontiefovým velkým přínosem… byla jeho konstrukce kvantitativního modelu americké ekonomiky, za účelem sledování dopadu vládní politiky a spotřebitelských trendů na velký počet průmyslových odvětví, která byla zasazena do vysoce komplexní řady vzájemně propojených vztahů. Abychom si uvědomili rozdíl mezi čistě formálním modelem a empirickým modelem, je dobré mít na paměti, že získání dat pro skutečný model vyžaduje organizaci, která pracuje mnoho měsíců, někdy i let. Po sestavení modelu se rýsuje další překážka – řešení velmi rozsáhlého systému lineárních rovnic. V období 1936–1940 nebyly k dispozici žádné elektronické počítače; to nejlepší, v co člověk mohl doufat, bylo vyřešení dvaceti rovnic o dvaceti neznámých. A konečně tu byly potíže s „marketingem“ výsledků takových studií. Projekt iniciovaný Leontiefem tedy od počátku představoval trojí hazard.* Viz [117, str. 17].

[35] Z dnešního pohledu se jednalo o více krokový dynamický lineární model s trojúhelníkovou (či dokonce blokově diagonální) maticí.

[36] Kdybychom každou sekundou stihli projít 1 milion možností a začali bychom s tím už v době Velkého třesku, tak bychom k dnešnímu dni . . . neměli hotovo skoro nic.

[37] Tento název vydržel cca 1 rok. V létě 1948 navštívil Dantzig a Koopmans společnost RAND v kalifornské Santa Monice a při jedné z procházek podél tamní pláže Koopmans řekl „Proč nezkrátit název *programování v lineární struktuře* pouze na *lineární programování*?“. Na to Dantzig zareagoval slovy „To je ono! Od teď se to bude nazývat takto.“ a ještě tentýž den přednesl v RANDu přednášku s názvem *Lineární programování*.

[38] Avšak bylo by chybou připisovat Dantzigovi i „objev“ samotné účelové funkce – ta se objevuje už např. v Lagrangeově práci.

[39] Jerzy Neyman (16. dubna 1894 – 5. srpna 1981), byl polský matematik a statistik, který strávil první část profesní dráhy na různých institucích ve Varšavě a poté na University College London, zatímco druhá část je spojena s Kalifornskou univerzitou v Berkeley. Položil základy moderní teoretické statistiky. Představil nové přístupy k testování statistických hypotéz a je spoluautorem konceptu *nulové hypotézy*, která našla uplatnění v genetice, lékařské diagnostice, astronomii i meteorologii.

[40] Tato historka údajně inspirovala i kázání jistého reverenda v Americe jako ukázka pozitivního myšlení – kdyby Dantzig věděl, že jde o otevřené problémy, pravděpodobně by je tak snadno nevyřešil, neboť by si tolik nevěřil jako v případě domácího úkolu.

[41] Dantzig řešení těchto dvou problémů publikoval ve dvou článcích [111, 127] dělící rozestup 11 let. Proč? Protože byl prý tolik zaneprázdněn prací ve Washingtonu, že zapomněl jednu z částí své disertační práce publikovat. Jednoho dne pak dostal dopis o Walda, ve kterém vysvětloval, že byl přijat k publikaci jeho článek, ale až dodatečně zjistil, že o něčem podobném psal Dantzig již ve své disertaci. Wald se proto ptal Dantziga, co s tím. Dantzig navrhl, aby jej přidal mezi autory, a tím byl problém vyřešen.

[42] George Joseph Stigler (17. ledna 1911 – 1. prosince 1991) byl americký ekonom a společně s Miltonem Friedmanem byli předními představiteli Chicagské ekonomické školy. V roce 1982 získal *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za „významné studie o průmyslové struktuře, fungování trhu a o příčinách a následcích veřejné regulace".

[43] National Research Council

[44] *Byl zmíněn dostatek komplikací, abychom ukázali téměř nekonečnou složitost sofistikovaného a přesného stanovení nutričních hodnot jídelníčku.*

[45] *Ve světle předchozích poznámek snad není nezbytné rozvádět nepřesnost těchto čísel.*

[46] *Další postup už je zkoušením, protože neexistuje žádná přímá metoda pro nalezení minima lineární funkce při lineárních omezeních.*

	doporučená denní dávka
kalorie	3000 jednotek
bílkoviny	70 g
vápník	0.8 g
železo	12 mg
vitamín A	5000 jednotek
tiamin (vitamín B1)	1.8 mg
riboflavin (vitamín B2 nebo G)	2.7 mg
niacin	18 mg
kyselina askorbová	75 mg

Tabulka 2.1: Minimální požadavky pro Stiglerův problém optimálního jídelníčku.

Toho mělo být docíleno pomocí 77 surovin/potravin s cenami z roku 1939 a nutričními hodnotami, které získal z amerického ministerstva zemědělství, viz (dlouhou) tabulku na dalších stranách. Jenže tyto hodnoty samozřejmě nejsou (a snad ani nemohou) být zcela přesné, což Stigler ve svém článku ilustroval na šesti nejmarkantnějších případech a následně okomentoval i těmito slovy.

> *Enough difficulties have been indicated to suggest the almost infinite complexity of a refined and accurate assessment of nutritive value of a diet.*[44]

> *In light of the foregoing remarks it should not be necessary to belabor the tentativeness of the figures.*[45]

Stiglerovým prvním krokem bylo nahrazení některých surovin jinými, „výživnějšími", čímž snížil jejich počet ze 77 na 15. Pak se ale dostal do potíží, neboť jak sám píše

> *Thereafter the procedure is experimental because there does not appear to be any direct method of finding the minimum of a linear function subject to linear conditions.*[46]

Takže metodou pokus–omyl vyzkoušel 510 kombinací a dospěl k celkovým denním nákladům 0,1093 USD (tj. 39,93 USD za rok) pro jídelníček složený z

- pšeničné mouky (1) za 0,0365 USD (tj. 459,9 gramů),
- sušeného mléka (15) za 0,0105134 USD (tj. 63,448 gramů),
- kapusty (46) za 0,011252 USD (tj. 100,694 gramů),
- špenátu (52) za 0,005065 USD (tj. 23,258 gramů),
- fazolí navy (69) za 0,0459959 USD (tj. 353,61648 gramů).

Stigler samozřejmě netvrdil, že nalezl skutečně optimální řešení.

> *There is no reason to believe that the cheapest combination was found... On the other hand the annual cost could not have been reduced by more than a few dollars by a better selection from these commodities.*[47]

ale nabídl několik důvěryhodných argumentů k uvěření, že roční náklady nemohou být sníženy o více něž několik málo dolarů. A on očividně věděl, co dělá a říká!

[47] *Neexistuje žádný důvod k tomu, abychom věřili, že jsme nalezli nejlevnější kombinaci… Na druhou stranu ale roční náklady nemohou být sníženy o více než pár dolarů lepším výběrem z těchto komodit.*

Tabulka 2.2: Dostupné suroviny/potraviny pro Stiglerův problém optimálního jídelníčku.

			množství živin obsažené v porci za 1 USD								
		množství za 1 USD [g]	kalorie [1000 j]	bílkoviny [g]	vápník [g]	železo [mg]	vitamín A [1000 j]	tiamin [mg]	riboflavin [mg]	niacin [mg]	kyselina askorbová [mg]
1.	pšeničná mouka	12600	44,7	1411	2	365		55,4	33,3	441	
2.	makarony	3217	11,6	418	0,7	54		3,2	1,9	68	
3.	pšeničné cereálie	3280	11,8	377	14,4	175		14,4	8,8	114	
4.	kukuřičné lupínky	3194	11,4	252	0,1	56		13,5	2,3	68	
5.	kukuřičná mouka	9861	36	897	1,7	99	30,9	17,4	7,9	106	
6.	kukuřičná kaše	8005	28,6	680	0,8	80		10,6	1,6	110	
7.	rýže	6048	21,2	460	0,6	41		2	4,8	60	
8.	ovesné vločky	6389	25,3	907	5,1	341		37,1	8,9	64	
9.	bílý chleba	5742	15	488	2,5	115		13,8	8,5	126	
10.	celozrnný chleba	4985	12,2	484	2,7	125		13,9	6,4	160	
11.	žitný chleba	4930	12,4	439	1,1	82		9,9	3	66	
12.	pound cake („piškotový dort”)	1829	8	130	0,4	31	18,9	2,8	3	17	
13.	slané krekry	3004	12,5	288	0,5	50					
14.	mléko	8867	6,1	310	10,5	18	16,8	4	16	7	177
15.	sušené mléko	6035	8,4	422	15,1	9	26	3	23,5	11	60
16.	máslo	1473	10,8	9	0,2	3	44,2		0,2	2	
17.	margarín	2817	20,6	17	0,6	6	55,8	0,2			

18.	vejce	1857	2,9	238	1	52	18,6	2,8	6,5	1	
19.	čedar	1874	7,4	448	16,4	19	28,1	0,8	10,3	4	
20.	smetana	1689	3,5	49	1,7	3	16,9	0,6	2,5		17
21.	burákové máslo	2534	15,7	661	1	48		9,6	8,1	471	
22.	majonéza	1198	8,6	18	0,2	8	2,7	0,4	0,5		
23.	Crisco (rostlinný tuk)	2234	20,1								
24.	sádlo	4628	41,7				0,2		0,5	5	
25.	roštěnka	1145	2,9	166	0,1	34	0,2	2,1	2,9	69	
26.	kýta	1246	2,2	214	0,1	32	0,4	2,5	2,4	87	
27.	pečená žebra	1553	3,4	213	0,1	33			2		
28.	hovězí krk	2007	3,6	309	0,2	46	0,4	1	4	120	
29.	hovězí bok	3107	8,5	404	0,2	62		0,9			
30.	hovězí játra	1692	2,2	333	0,2	139	169,2	6,4	50,8	316	525
31.	skopové stehno	1643	3,1	245	0,1	20		2,8	3,9	86	
32.	skopová žebra	1239	3,3	140	0,1	15		1,7	2,7	54	
33.	vepřová žebra	1477	3,5	196	0,2	30		17,4	2,7	60	
34.	vepřová pečeně	1874	4,4	249	0,3	37		18,2	3,6	79	
35.	slanina	1772	10,4	152	0,2	23		1,8	1,8	71	
36.	uzená šunka	1655	6,7	212	0,2	31		9,9	3,3	50	
37.	„solené vepřové”	2835	18,8	164	0,1	26		1,4	1,8		
38.	pečené kuře	1497	1,8	184	0,1	30	0,1	0,9	1,8	68	46

39.	telecí kotlet	1072	1,7	156	0,1	24		1,4	2,4	57	
40.	losos	3489	5,8	705	6,8	45	3,5	1	4,9	209	
41.	jablka	9072	5,8	27	0,5	36	7,3	3,6	2,7	5	544
42.	banány	4932	4,9	60	0,4	30	17,4	2,5	3,5	28	498
43.	citróny	2380	1	21	0,5	14		0,5		4	952
44.	pomeranče	4439	2,2	40	1,1	18	11,1	3,6	1,3	10	1998
45.	zelené fazole	5750	2,4	138	3,7	80	69	4,3	5,8	37	862
46.	kapusta	8949	2,6	125	4	36	7,2	9	4,5	26	5369
47.	mrkev	6080	2,7	73	2,8	43	188,5	6,1	4,3	89	608
48.	celer	3915	0,9	51	3	23	0,9	1,4	1,4	9	313
49.	hlávkový salát	2247	0,4	27	1,1	22	112,4	1,8	3,4	11	449
50.	cibule	11844	5,8	166	3,8	59	16,6	4,7	5,9	21	1184
51.	brambory	1681	14,3	336	1,8	118	6,7	29,4	7,1	198	2522
52.	špenát	4592	1,1	106		138	918,4	5,7	13,8	33	2755
53.	batáty	7649	9,6	138	2,7	54	290,7	8,4	5,4	83	1912
54.	broskve (plechovka)	4894	3,7	20	0,4	10	21,5	0,5	1	31	196
55.	hrušky (plechovka)	4030	3	8	0,3	8	0,8	0,8	0,8	5	81
56.	ananas (plechovka)	3993	2,4	16	0,4	8	2	2,8	0,8	7	399
57.	chřest	1945	0,4	33	0,3	12	16,3	1,4	2,1	17	272
58.	zelené fazole (plechovka)	5386	1	54	2	65	53,9	1,6	4,3	32	431

59.	vepřové s fazolemi (plechovka)	63889	7,5	364	4	134	3,5	8,3	7,7	56	
60.	kukuřice (plechovka)	5452	5,2	136	0,2	16	12	1,6	2,7	42	218
61.	hrášek (plechovka)	4109	2,3	136	0,6	45	84,9	4,9	2,5	37	370
62.	rajčata (plechovka)	6263	1,3	63	0,7	38	53,2	3,4	2,5	36	1253
63.	rajčatová polévka (plechovka)	3917	1,6	71	0,6	43	57,9	3,5	2,4	67	862
64.	broskve (sušené)	2889	8,5	87	1,7	173	86,8	1,2	4,3	55	57
65.	švestky (sušené)	4284	12,8	99	2,5	154	85,7	3,9	4,3	65	257
66.	rozinky	4524	13,5	104	2,5	136	4,5	6,3	1,4	2,4	136
67.	hrášek (sušený)	5742	20	1367	4,2	345	2,9	28,7	18,4	162	
68.	bílé fazole	5097	17,4	1055	3,7	459	5,1	26,9	38,2	93	
69.	fazole navy	7688	26,9	1691	11,4	792		38,4	24,6	217	
70.	káva	2025						4	5,1	50	
71.	čaj	652							2,3	42	
72.	kakao	2637	8,7	237	3	72		2	11,9	40	
73.	čokoláda	1400	8	77	1,3	39		0,9	3,4	14	
74.	cukr	8773	34,9								
75.	kukuřičný sirup	4966	14,7		0,5	74				5	
76.	melasa	3752	9		10,3	244		1,9	7,5	146	
77.	jahodový džem	2213	6,4	11	0,4	7	0,2	0,2	0,4	3	

Na podzim 1947 provedl Jack Laderman[48] z *Projektu matematického tabelování* v *Národním úřadu standardů*[49,50] po 120 člověko-dnech výpočtů na ručně ovládaných stolních kalkulátorech[51,52] nalezl skutečně optimální řešení s denními náklady 0,108662 USD, tj. 39,6888 USD za rok, což se od původního Stiglerova výsledku liší o pouhých 24 centů. Samozřejmě pro domácí vaření nemá taková úloha asi příliš velký smysl (vážně?), ale vraťte se zpět k armádním základům celého operačního výzkumu a uvažte armádu USA s cca 12 000 000 vojáků v roce 1945… Nalezené řešení je založeno na

- pšeničné mouce (1) za 0,0295191 USD (tj. 371,94066 gramů)
- hovězích játrech (30) za 0,00189256 USD (tj. 3,20221152 gramů)
- kapustě (46) za 0,0112144 USD (tj. 100,3576656 gramů)
- špenátu (52) za 0,00500766 USD (tj. 22,99517472 gramů)
- fazolí navy (69) za 0,0610286 USD (tj. 469,1878768 gramů)

s příjmem 3000 kalorií, 147,414 g bílkovin, 0,8 g vápníku, 60,4669 mg železa, 5000 jednotek vitamínu A, 4,12044 mg tiaminu, 2,7 mg riboflavinu, 27,316 mg niacinu a 75 mg kyseliny askorbové. V tomto řešení jsou ale některé doporučené denní dávky překročeny na úkor minimální ceny. Pokud bychom ale chtěli mít optimální jídelníček obsahující zcela přesně doporučené denní dávky, pak by minimální náklady vzrostly na 0,138171 USD (tj. na roční náklady 50,46696 USD) a jídelníček by byl složen z

- pšeničné mouky (1) za 0,0180918 USD (tj. 227,95668 gramů)
- kukuřičné mouky (5) za 0,0170207 USD (tj. 167,8411227 gramů)
- sušeného mléka (15) za 0,0452014 USD (tj. 272,790449 gramů)
- burákového másla (21) za 0,00317017 USD (tj. 8,03321078 gramů)
- sádla (24) za 0,0243964 USD (tj. 112,9065392 gramů)
- hovězích jater (30) za 0,015273 USD (tj. 25,841916 gramů)
- kapusty (46) za 0,00921709 USD (tj. 82,48373841 gramů)
- brambor (51) za 0,0051373 USD (tj. 8,6358013 gramů)
- špenátu (52) za 0,000663078 USD (tj. 3,044854176 gramů)

Samozřejmě ne vždy se hned na poprvé podaří získat smysluplný jídelníček (či výsledek jiného modelu) a správně matematizovat daný problém je skutečné umění – samotný Dantzig o tom ví své. Na začátku 50. let totiž dostal od svého lékaře doporučení, aby upravil svoji stravu. Rozhodl se proto, že si k tomu po vzoru Stiglera sestaví vlastní lineární program a svůj nový jídelníček založí na jeho řešení. Ovšem nešlo mu o minimalizaci ceny. Místo toho se chtěl vyhnout nejčastějšímu úskalí každé diety: neustálý *pocit hladu*, takže cílem bylo maximalizovat celkovou váhu jídla poníženou o množství obsažené vody. Jelikož v té době už Dantzig pracoval ve společnosti RAND (od června 1952), využil k jeho řešení tamní počítač IBM 701, který naštěstí dával výsledky mnohem rychleji než v roce 1947. Jednoho dne své ženě oznámil, že už nastal „ten den“, a cokoli 701 oznámí, to by mu měla nachystat. Vstupní data zahrnovala informace o více než 500 surovinách a jídel. Když mu v 5 hodin odpoledne manželka zavolala, aby zjistila, jaká bude večeře, oznámil ji Dantzig výsledek. Manželka to okomentovala slovy „Je to trochu divné, ale realizovatelné“. Ovšem poté Dantzig dodal poslední položku v seznamu: 1892 litrů octu. To samozřejmě nebyl příliš dobrý nápad, takže další den Dantzig trochu poupravil některé požadavky, což jej přivedlo k novému jídelníčku. Ten byl tentokrát založen na konzumaci 200 kostek bujonu – na to Dantzigova manželka zareagovala řečnickou otázkou, zda „se chystá ovládnout trh s bujónovými kostkami“. Dantzig (jako správný manžel) se ale nedal její poznámkou odradit a druhý den ráno otestoval, kolik dokáže takových kostek sníst. Dal si čtyři do hrnku s horkou vodou a… nedalo se to ani polknout. Bylo to neskutečně slané. Do modelu proto přidal omezení množství přijímané soli, které má lidský organismus přirozeně zabudované, ale matematický model je trochu něco jiného. Třetí pokus jej přivedl k pojídání cca 900 gramů otrub denně a čtvrtý k obdobnému množství třtinové melasy. To už bylo na jeho manželku trochu moc, takže celý tento problém vzala do svých rukou a nasadila mu vlastní jídelníček, díky němuž Dantzig zhubl 10 kg, viz [120].

[48] Další z těch, o kterých se nám nepodařilo nic najít.

[49] *Mathematical Tables Project of the National Bureau of Standards* byla jedna z největších a nejsofistikovanějších výpočetních organizací vytvořená před vynálezem digitálních elektronických počítačů.

[50] V červnu 1947 se Dantzig stal součástí Projektu SCOOP coby tamní šéf matematiků. Tento projekt sponzoroval i dva externí programy: první byl součástí *Národním úřadu standardů*, aby její zaměstnanci mohli experimentovat se simplexovou metodou, a druhý byl spojen s *Graduate School of Industrial Administration* na *Univerzitě Carnegieho–Mellonových* v Pittsburghu s cílem implementovat nástroje lineárního programování do průmyslu.

[51] Abychom ještě více přiblížili tuto dobu:

- v roce 1939 trval výpočet inverze pro 42×42 matici Leontiefova modelu (hodně nulových prvků a ostatní prvky z intervalu $(0,2)$ s přesností na dvě desetinná místa) 56 hodin. V roce 1947 to pro podobnou 38×38 matici bylo 48 hodin, ovšem v roce 1953 již stačilo pouze 45 minut. Viz [379, str. 31].
- V roce 1949 měli v Jižní Kalifornii zcela nový počítač CPC vyžadující 1 sekundu na aritmetickou operaci. Jednoho dne z nedaleké letecké společnosti dorazili s žádostí o výpočet inverzní matice k 16×16 matici s desetimístnými čísly (to je pro ruční výpočty vskutku náročný úkol). S obvyklým počtem mechanických závad, trval výpočet 2 dny! Během kontroly výpočtů se ukázalo, že s přesností na osm platných číslic je nalezená inverzní matice rovna transpozici původní matice. Pak už jen stačilo 10 minut k ověření, že původní matice byla skutečně ortogonální. Viz [3, str. 246].

  Další modely už byly výrazně rychlejší: 2 hodiny pro inverzi 20×20 matice, 30 minut pro inverzi 10×10 s přesností na 7 platných míst atd. Viz [207, str. 62].

Nicméně trvalo téměř rok než si Dantzig a ostatní uvědomili skutečnou sílu simplexového algoritmu. Během té doby, se Dantzig rozhodl navštívit von Neumanna, aby zjistil, zda by nemohl pomoci s metodami řešení. Poprvé ho navštívil 3. října 1947 v Princetonu. Ačkoli už v té době byl von Neumann mnohými považován za největšího žijícího matematika, snažil se mu Dantzig popsat studované problémy jako normálnímu smrtelníkovi. Začal se sestavováním modelu lineárního programování pomocí činností atd. V tom ho von Neumann zastavil se slovy „Jděte k věci.“ A jak si přál, tak se stalo. Během jedné minuty mu Dantzig „plácnul“ na tabuli algebraickou i geometrickou verzi celého problému. Na to von Neumann zareagoval slovy „Aha, toto.“ a následovala jeho cca 90 minutová přednáška o matematické teorie lineárního programování. Dantzig zůstal sedět s vyvalenýma očima a otevřenou pusou, vždyť nic z toho v literatuře nebylo! Von Neumann si toho v jednu chvíli všiml a řekl:

> *I don't want you to think I am pulling all this out of my sleeve on the spur of the moment like a magician. I have recently completed a book with Oscar Morgenstern on the theory of games. What I am doing is conjecturing that the two problems are equivalent. The theory that I am outlining is an analogue to the one we have developed for games.*[53]

Díky tomu se Dantzig poprvé dozvěděl o *Farkasově(–Minkowského) lemmatu* a o *dualitě*. Von Neumann slíbil Dantzigovi, že se nad jeho výpočetním problémem zamyslí a během pár týdnů se ozve. To splnil a navrhl jistý nelineární iterativní algoritmus. Někdy kolem roku 1952 lidé z Hoffmanovy[54] skupiny von Neumannův návrh otestovali na několika problémech a porovnali jej se simplexovým algoritmem a s metodou navrhovanou Motzkinem. Simplexový algoritmus byl jasný vítěz, viz [118]. Mimochodem v roce 2000 byl v časopise *Computing in Science & Engineering* vyhlášen seznam *The Top 10 Algorithms*, které byly vybrány jako ty s *největším vlivem na vývoj a využití vědy a techniky ve 20. století*. Dantzigův algoritmus v něm samozřejmě nemohl chybět, viz [139].

Při další návštěvě v Princetonu v červnu 1948 se Dantzig setkal s Tuckerem[55], který krátce poté začal se svými studenty Kuhnem[56] a Galem[57]. budovat teorii her, nelineárního programování a duality. Výsledky této skupiny z ní učinily ústřední bod matematického výzkumu v této oblasti. O 12 let později se Tucker, který v té době pročítal rukopis Dantzigovy knihy [117], s Dantzigem setkali znovu a společně zabředli i do následujícího rozhovoru.

> **AT**: *Proč přisuzujete dualitu von Neumannovi a ne mojí skupině?*
>
> **GD**: *Protože on byl první, kdo mi ji ukázal.*
>
> **AT**: *To je divné, protože jsme nenašli žádný záznam o tom, co von Neumann udělal. Našli jsme akorát jeho článek „On a Maximizing Problem“.*
>
> **GD**: *To je pravda. Pošlu vám můj článek, který jsem napsal na základě mého prvního setkání s von Neumannem.*

Dantzig poslal Tuckerovi zprávu *A Theorem on Linear Inequalities* datovanou 5. ledna 1948, která obsahovala (zdá se) první rigorózní důkaz duality.

> **AT**: *Proč jste to nepublikoval?*
>
> **GD**: *Protože to nebyly moje výsledky – byly von Neumannovi. Jediné, co jsem udělal, bylo sepsání důkazu naznačeného von Neumannem, aby mohl dále kolovat. Byl to můj způsob vzdělávání lidí z mého úřadu v Pentagonu.*

Dnes je von Neumann obvykle považován za zakladatele teorie duality, zatímco Kuhnovi, Tuckerovi a Galeovi je připisován její první rigorózní důkaz.

[52] Ne všechny úlohy byly takto výpočetně náročné. Např. v roce 1949 se pomocí simplexové metody podařilo vyřešit *dopravní problém* s 25 výchozími místy a 60 cílovými místy během pouhých 9 člověko-dnů ručních výpočtů, viz [112]. To bylo způsobeno tím, že tento typ problémů má velmi speciální strukturu.

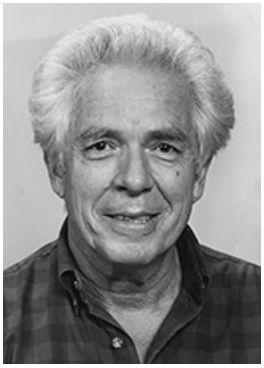
[53] *Nechci, abyste si myslel, že to všechno teď tady tahám bezprostředně z rukávu jako nějaký kouzelník. Nedávno jsem dokončil s Oscarem Morgensternem knihu o teorii her. A já se teď snažím ukázat, že oba problémy jsou ekvivalentní. Teorie, kterou tu nastiňuji, je analogií k té, kterou jsme vybudovali pro hry.*

[54] Alan Jerome Hoffman (⋆30. května 1924) je americký matematik, který významně ovlivnil rozvoj společnosti IBM. V roce 1992 získal společně s Philipem Wolfem *John von Neumann Theory Prize*.

[55]

Albert William Tucker (28. listopadu 1905 – 25. ledna 1995) byl kanadský matematik, který výrazně ovlivnil topologii, teorii her a především teorii nelineárního programování. V roce 1950 poprvé formalizoval nejzákladnější problém teorie her, který následně nazval *vězňovým dilematem*. V roce 1980 získal společně s Kuhnem a Galem *John von Neumann Theory Prize* za jejich příspěvek k teorii optimalizace. Jedním z jeho mnoha doktorských studentů byl i John Nash, který v roce 1994 získal *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela*. V letech 1961–1962 byl prezidentem Americké matematické asociace.

56



Harold William Kuhn (29. července 1925 – 2. července 2014) byl americký matematik a bývalý emeritní profesor v Princetonu, který se věnoval především teorii her. V roce 1980 získal společně s Galem a Tuckerem *John von Neumann Theory Prize* za jejich příspěvek k teorii optimalizace. Po mnoho let byl považován za autora tzv. *maďarského algoritmu* pro řešení přiřazovací úlohy z teorie grafů. Ovšem v roce 2006 se ukázalo, že tentýž algoritmus navrhoval již Jacobi v pracích, které byly posmrtně publikovány v roce 1890.

57

David Gale (13. prosince 1921 – 7. března 2008) byl americký matematik a ekonom. Největší část jeho profesní dráhy je spojena s Kalifornskou univerzitou v Berkeley, kde působil na ústavu matematiky, ekonomie, průmyslového inženýrství a operačního výzkumu. Svůj výzkum zaměřil především na matematickou ekonomii, teorii her a konvexní analýzu. Výrazně ovlivnil také rozvoj lineárního programování a teorie lineárních nerovností. Jeho kniha *The Theory of Linear Economic Models* z roku 1960 patří mezi klasické literární zdroje v této oblasti. Hlavní část této monografie vznikla v době jeho působení jako konzultanta na Oddělení matematické analýzy ve společnosti RAND v letech 1957–58. V roce 1980 získal společně s Kuhem a Tuckerem *John von Neumann Theory Prize* za jejich příspěvek k teorii optimalizace. V roce 1962 s Lloydem Shapleyem napsali

Začátky lineárního programování skutečně nebyly lehké, což dokládá ještě jedna Dantzigova vzpomínka spojená s von Neumannem. Nedlouho po prvním setkání s Tuckerem se ve Wisconsinu konalo společné setkání Institutu matematické statistiky[58] a Ekonometrické společnosti[59], kterého se zúčastnili světoznámí matematici jako von Neumann či Hotelling[60] a ekonomové jako Koopmans. Dantzig byl v té době době spíše „mladý neznámý“, a tak není divu, že byl před takovým publikem celkem vystrašený při své přednášce věnované lineárnímu programování. Po přednášce následovala diskuze. Chvíli bylo ticho, než se přihlásil Hotelling. Vstal, takže to vypadalo, že zabírá asi polovinu místnosti (rád plaval v oceánu a říkalo se o něm, že vždy způsobí znatelný vzestup hladiny), a začal Dantziga pomalu grilovat.

> *But we all know the world is nonlinear.*[61]

Tím vynesl nad Dantzigovou přednáškou devastující verdikt. Než se Dantzig zmohl na nějakou odpověď, zvedla se další ruka se slovy

> *Mr. Chairman, Mr. Chairman, if the speaker doesn't mind, I would like to reply for him.*[62]

Byl to von Neumann a Dantzig samozřejmě souhlasil, načež vstal a pokračoval

> *The speaker titled his talk „linear programming“ and carefully stated his axioms. If you have an application that satisfies the axioms, well use it. If it does not, then don't...*[63,64]

Pak si von Neumann sedl. Když se nad tím zamyslíme, tak Hotelling měl samozřejmě pravdu – svět je značně nelineární. Naštěstí ale systémy lineárních nerovnic (na rozdíl od rovností) nám umožňují aproximovat většinu z těchto nelinearit, které se objevují v praktickém plánování.

V roce 1960 započal Dantzig svoji zářnou akademickou kariéru, když nastoupil na místo profesora na Ústavu průmyslového inženýrství na Kalifornské univerzitě v Berkeley. Poté (1966) se přesunul na Stanfordovu univerzitu, kde se z něj v roce 1985 stal emeritní profesor a aktivnímu výzkumu se věnoval až do roku 1997. Po celou dobu také často navštěvoval *Mezinárodní institut pro aplikovanou systémovou analýzu*[65] (IIASA) v rakouském Laxenburgu, což je mezinárodní výzkumná organizace *provádějící mezioborové vědecké studie v oblastech ochrany životního prostředí, ekonomie, technologií a sociální problematiky v kontextu lidského rozsahu globálních změn.* V letech 1973–1974 zde společně se svoji ženou strávil „sabatikal“ a jeho pobyt bývá spojován se dvěma historkami, viz [17, str. 229].

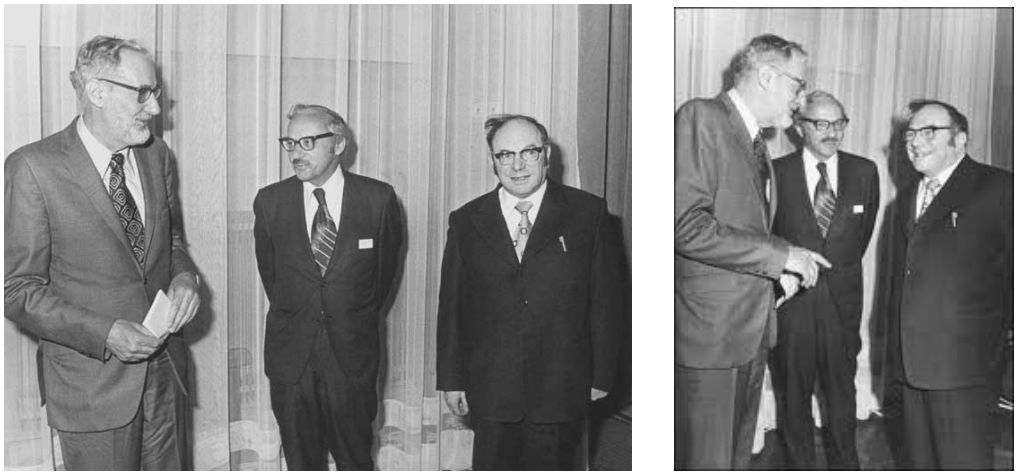
> *Pár dní po svém příchodu do IIASA volal George [Dantzig] Ruth Steinerové, která měla na starosti návštěvníky, s následující žádostí: „Před mou kanceláří stojí velmi dlouhý nákladní vůz. Nedovedu si představit, že je výhodné mít takovou délku. Můžete mi, prosím, zjistit, co je v něm naloženo, odkud přijel, kudy jel a jak projížděl kolem rohů?“ Přišla odpověď: „Nábytek ze Salzburgu, přes dálnici, častokrát musel couvat a vracet se.“ Za nějaký čas informoval Georgeův výzkumný asistent Ruth, že by firma mohla ušetřit 40 % nákladů, pokud by použila čtyři menší nákladní vozy, a navrhoval, aby to firmě oznámila. To také udělala a následně jej informovala o výsledku: „Mysleli si o mně, že jsem se zbláznila.“*

> *V IIASA byl George také proslulý tím, že pil polévku z pivního půllitru. Když se jej jeho asistentka ptala proč, napsal „dlouhý matematický vzorec dokazu-*

jící, že tepelné ztráty při použití půllitru jsou menší ve srovnání s polévkovým talířem“.

Pro další podrobnosti o Dantzigovi viz např. [104–107, 118].

V roce 1975 byla poprvé udělena *John von Neumann Theory Prize* a jejím držitelem se stal Dantzig za „jeho práci v oblasti lineárního programování“. Při udílení ceny se pak Dantzig podělil s publikem o několik osobních historek souvisejících s von Neumannem. Avšak když v roce 1975 obdrželi Kantorovič s Koopmansem Nobelovu cenu za jejich práce související s lineárním programováním, Dantzig mezi oceněnými nebyl. A ani nikdy později tuto cenu za svůj nezměrný přínos k počátkům a následnému rozvoji lineárního programování, jeho využití v praxi a ekonomické interpretaci nikdy nezískal. To je samozřejmě ve světle všeho, co bylo napsáno výše, zcela udivující a nepochopitelné. Jaké důvody k tomu ve skutečnosti vedly, zatím zůstávají zahaleny tajemstvím – snad byl jeho přístup až příliš matematický. Ale možná se to brzy dozvíme, neboť podklady pro rozhodnutí komise mohou být po 50 letech zveřejněny, tj. nejdříve v roce 2025. Koopmans byl „hluboce zklamán“ tím, že nezískali Nobelovu cenu společně, takže alespoň věnoval 40 000 USD institutu IIASA na Dantzigovu počest[66], viz Obrázek 2.2 a pro zajímavost ještě i Obrázek 2.3. Koopmans projevoval svoji nespokojenost s neudělením ceny i Dantzigovy také tím, že navrhoval Kantorovičovi, aby toto ocenění odmítli. Byla to asi jedna z nejtěžších diskuzí mezi nimi – zejména pro Kantoroviče. Jenže vývoj šel trochu jiným směrem, než by se na základě rozhodnutí Nobelova shromáždění mohlo zdát. Práce Kantoroviče i Dantziga věnované lineárnímu programování byly zcela původní a nezávislé, jenže Dantzig rozvinul teorii i mnoha dalšími směry (dualita a zejména simplexový algoritmus), takže většina knih a jiných textů je založena především na Dantzigových výsledcích. Kantorovičův výzkum neměl žádný významný dopad na rozvoj lineárního programování (ať už důvody byly jakékoli; bohužel, politické okolnosti v tom nejdůležitějším období zrodu lineárního programování nebyly pro Kantoroviče zrovna příznivé). Ovšem chcete-li ocenit někoho za prvenství jeho objevu, tak to asi nebude ten, který dokázal uspokojit potřeby všech. Jinými slovy, Kantorovič je jakýsi Kryštof Kolumbus lineárního programování, který je sice považován za „objevitele Ameriky“, ale jejího skutečného území se vlastně nikdy nedotkl, viz [17, str. 231–232].



Obrázek 2.2: Setkání Koopmanse, Dantziga a Kantoroviče (zleva) v roce 1975 v IIASA.

2.2 Seznamte se! Úlohy lineárního programování

Jak již všichni jistě víme/tušíme z úvodní kapitoly, úlohou lineárního programování rozumíme minimalizaci lineární funkce na množině vymezené soustavou lineárních rovnic a/nebo ne-

článek věnovaný *problému stabilního manželství* s algoritmem, který se v dnešní době používá pro rozdělování žáků do veřejných škol v New Yorku a Bostonu, viz také https://ln.math.muni.cz/pz/opt/smp1. Shapley za tuto práci získal v roce 2012 Nobelovu cenu za ekonomii, viz také Příklad 5.3.2.

[58] *Institute of Mathematical Statistics*

[59] *Econometric Society*

[60] Harold Hotelling (29. září 1895 – 26. prosince 1973) byl americký matematik, statistik a vlivný ekonomický teoretik. Jeho jméno je spojováno s Hotellingovým zákonem, Hotellingovým lemmatem a Hotellingovým pravidlem známými z ekonomie.

[61] *Ale my všichni víme, že svět je nelineární.*

[62] *Pane předsedo, pane předsedo, pokud to řečníkovi nevadí, rád bych odpověděl za něj.*

[63] *Řečník nazval svoji přednášku „Lineární programování“ a pečlivě představil své axiomy. Máte-li příklad, který tyto axiomy splňuje, rozhodně to využijte. V opačném případě to nedělejte.*

[64] Toto setkání proběhlo 9. září 1948 a Dantzig svoji přednášku nazval *Programming in a Linear Structure* (Sic! Von Neumann hovořil o něčem trochu jiném...). Jeho příspěvek vyvolal mnohem větší zájem než jeho první veřejné vystoupení na toto téma. To se uskutečnilo 29. prosince 1947 na společném setkání Americké statistické společnosti (*American Statistical Association*) a Institutu matematické statistiky, kde Dantzig vystoupil s příspěvkem nazvaným *Mathematical Techniques of Program Planning*.

[65] *International Institute of Applied Systems Analysis*

[66] Výše Koopmansova daru měla velmi racionální základ. Celková odměna tehdy totiž činila 240 000 USD, takže Koopmansovi zbylo 80 000 USD, které by obdržel, kdyby byl mezi oceněnými i Dantzig.

CONFIDENTIAL

USSR Leonid Vital'yevich KANTOROVICH

Head, Problems Laboratory of Economic-Mathematical Methods and Operations Research, Institute of Management of the National Economy

An internationally recognized creative genius in the fields of mathematics and the application of electronic computers to economic affairs, Academician Leonid Kantorovich (pronounced kahntuhROHvich) has worked at the Institute of Management of the National Economy since 1971. He has been involved in advanced mathematical research since the age of 15; in 1939 he invented linear programming, one of the most significant contributions to economic management in the twentieth century. Kantorovich has spent most of his adult life battling to win acceptance for his revolutionary concept from Soviet academic and economic bureaucracies; the value of linear programming to Soviet economic practices was not really recognized by his country's authorities until 1965, when Kantorovich was awarded a Lenin Prize for his work. International recognition came in October 1975, when the mathematician was awarded the Nobel Prize for Economics jointly with T. C. Koopmans, a Dutch-born American economist who discovered the same concept independently a few years after Kantorovich.

In addition to his mathematical research, Kantorovich has been directly involved in developing improved designs for high-speed digital computers, an activity apparently motivated by the Soviet Union's need for improved computers in solving large economic planning problems.

The Institute of Management of the National Economy

The Institute of Management of the National Economy was established to train high-level economic and industrial administrators in modern methods of management, production organization and the use of economic-mathematical methods and computers in planning. When the institute opened in early 1971, Premier Alekse Kosygin and Party Secretary Andrey Kirilenko attended the ceremonies, thus suggesting the importance that the Soviet Government and Party attach to the application of modern management techniques to Soviet industrial administration and economic planning.

CONFIDENTIAL CR 77-10705

Obrázek 2.3: První ze tří stran z tajné zprávy CIA o Kantorovičovi datované 1. dubnem roku 1977, která byla v roce 1979 ukradena z americké ambasády v íránském Teheránu.

rovnic, tj.

$$\left.\begin{array}{r} \langle c,x\rangle \to \min \\ x\in X := \{x\in\mathbb{R}^n \mid \langle a_i,x\rangle \le b_i,\ i=1,\dots,k\ \ \&\ \langle a_i,x\rangle = b_i,\ i=k+1,\dots,m\ \ \& \\ \&\ x_1\ge 0\ \&\dots\ \&\ x_s\ge 0\}\end{array}\right\} \tag{2.2.1}$$

pro dané vektory $c,a_1,\dots,a_m\in\mathbb{R}^n$, čísla $b_1,\dots,b_m\in\mathbb{R}$ a indexy $k\in\{0,1,\dots\}$, $m\in\{k,k+1,\dots\}$ a $s\in\{0,\dots,n\}$. Rovnostní i nerovnostní podmínky vymezující množinu X lze souhrnně napsat jako dvojici soustav lineárních nerovnic a rovnic

$$A_N x\le b_N \quad \text{a} \quad A_R x = b_R,$$

kde řádky matic $A_N\in\mathbb{R}^{k\times n}$ a $A_R\in\mathbb{R}^{(m-k)\times n}$ jsou postupně tvořeny vektory $a_1,\dots,a_m$, tj.

$$A_N=\begin{pmatrix} a_1^\top \\ a_2^\top \\ \vdots \\ a_k^\top\end{pmatrix} \quad \text{a} \quad A_R=\begin{pmatrix} a_{k+1}^\top \\ a_{k+2}^\top \\ \vdots \\ a_m^\top\end{pmatrix}$$

Pochopitelně jsou povoleny i „mezní“ hodnoty zmíněných indexů, což nás přivádí k následující terminologii[67] týkající se výhradně jenom tvaru množiny X.

(i) Množina X je v *základním tvaru*, pokud $k=m$ a $s=0$ (tj. omezení jsou pouze ve tvaru nerovností a pro žádnou proměnnou nemáme omezení na znaménko).

(ii) Množina X je v *kanonickém tvaru*, pokud $k=0$ a $s=n$ (tj. omezení jsou pouze ve tvaru rovností a pro všechny proměnné máme omezení na znaménko).

(iii) Množina X je ve *standardním tvaru*, pokud $k=m$ a $s=n$ (tj. omezení jsou pouze ve tvaru nerovností a pro všechny proměnné máme omezení na znaménko).

(iv) Situace, kdy $k=0$ a $s=0$ (tj. omezení jsou pouze ve tvaru rovností a pro žádnou proměnnou nemáme omezení na znaménko), je zcela nezajímavá. V takové chvíli je totiž možné eliminovat některé proměnné vyřešením soustavy $A_R x=b_R$ a získat tak úlohu na volný extrém s nižším počtem proměnných.

Tyto speciální případy jsou vzájemně ekvivalentní a jsou zároveň ekvivalentní i s obecnou úlohou (2.2.1). To znamená, že úlohu lineárního programování v libovolném tvaru lze převést na úlohu jiného tvaru, aniž bychom změnili její řešení. Jinými slovy, úloha „A“ je ekvivalentní s úlohou „B“ právě tehdy, když řešení úlohy „A“ lze získat (pokud existuje) z řešení úlohy „B“ pomocí algebraických úprav. Pochopitelně obě úlohy nemusí mít stejný počet omezení či proměnných.

Poznámka 2.2.1.

(i) Z úlohy v kanonickém tvaru uděláme úlohu ve standardním tvaru tak, že každou rovnost $\langle a_i,x\rangle = b_i$ nahradíme dvojicí nerovností $\langle a_i,x\rangle\le b_i$ a $-\langle a_i,x\rangle\le -b_i$, takže počet proměnných se nezmění, ale počet omezení se zdvojnásobí. Naopak z úlohy ve standardním tvaru uděláme úlohu v kanonickém tvaru zavedením *pomocných* (*přídavných, doplňkových, skluzových* atd.) proměnných omezených na znaménko reprezentující rozdíl mezi pravou a levou stranou stranou nerovnice. S jejich pomocí z nerovnostního omezení $\langle a_i,x\rangle\le b_i$ uděláme $\langle a_i,x\rangle + s_i = b_i$ s požadavkem $s_i\ge 0$. Přechod od úlohy v základním tvaru k úloze ve standardním tvaru je snadný, neboť v takovém případě považujeme omezení na znaménko pouze za další nerovnostní omezení. Později navíc uvidíme, že úlohy v kanonickém a základním tvaru jsou v jakémsi „duálním“ vztahu.

(ii) Máme-li proměnnou, která není omezena na znaménko, pak ji můžeme vyjádřit jako rozdíl dvou nezáporných proměnných, tj. místo proměnné $x_i\in\mathbb{R}$ budeme mít dvě nové

[67] Použitá označení nejsou v literatuře zcela pevně ukotvena, takže lze narazit i na jiná rozdělení použitých termínů, viz např. [529]. Zejména v případě „kanonického tvaru“ jde obvykle o to, pro který typ množiny X je následně budován simplexový algoritmus.

proměnné $y_i \geq 0$, $z_i \geq 0$, s jejichž pomocí nahradíme x_i za výraz $y_i - z_i$ (takových dvojic y_i, z_i samozřejmě bude existovat nekonečně mnoho, takže v nové úloze můžeme případně ztratit jednoznačnost řešení). Úlohu bychom tedy s každou proměnnou bez omezení na znaménko museli zvětšit o jednu proměnnou. To však naštěstí není zcela nutné. Nezáleží totiž na tom, kolik proměnných bez omezení na znamínko máme, vždy si vystačíme celkově pouze s jednou proměnnou navíc. Potřebujeme-li např. omezit na znaménko proměnné x_1, x_2, x_3, můžeme využít substituce

$$x_1 = s_1 - z \quad \& \quad x_2 = s_2 - z \quad \& \quad x_3 = s_3 - z$$

a nahradit s jejich pomocí původní proměnné pouze novými nezápornými proměnnými $s_1, s_2, s_3, z \geq 0$. Proč? Mohli bychom místo každé proměnné vzít

$$x_1 = y_1 - z_1 \quad \& \quad x_2 = y_2 - z_2 \quad \& \quad x_3 = y_3 - z_3$$

s novými proměnnými $y_1, y_2, y3, z_1, z_2, z_3 \geq 0$, ovšem položíme-li $z := \max\{z_1, z_2, z_3\} \geq 0$, pak pro každé $i \in \{1,2,3\}$ díky nerovnosti $z - z_i \geq 0$ a $y_i \geq 0$ máme

$$x_i = \underbrace{(y_i + z - z_i)}_{=:s_i} - z = s_i - z$$

a současně kde $s_i \geq 0$. Jelikož stále platí, že tyto nové proměnné nevystupují v účelové funkci samostatně ale vždy pouze ve formě rozdílu $s_i - z$, může i při této efektivnější substituci dojít v nové úloze ke ztrátě jednoznačnosti řešení.

(iii) V případě omezení $x_i \geq \alpha_i \neq 0$ můžeme zavést novou nezápornou proměnnou $s_i \geq 0$ a nahradit x_i výrazem $s_i + b_i$.

(iv) V případě omezení $0 \leq x_i \leq \alpha_i$ opět zavedeme novou nezápornou proměnnou $s_i \geq 0$ a dané omezení nahradíme podmínkou $x_i + s_i = \alpha_i$ doplněnou požadavky na nezápornost x_i a s_i.

(v) Dvojici omezení $0 \neq \alpha_i \leq x_i \leq \beta_i$ můžeme s pomocí dvou nových proměnných s_i, z_i nahradit podmínkou $s_i + z_i = \beta_i - \alpha_i$ s požadavkem $s_i \geq 0$ a $z_i \geq 0$, přičemž platí[68] $x_i = s_i + \alpha_i$. ▲

[68] Nejdříve dostaneme $x_i - s_i = \alpha_i$ pro $s_i \geq 0$, což po dosazení do podmínky $x_i \leq \beta_i$ vede k $s_i + \alpha_i \beta_i$, takže s využitím nové proměnné $z_i \geq 0$ máme $s_i + \alpha_i + z_i = \beta_i$.

Ekvivalenci mezi jednotlivými tvary úloh lineárního programování si podrobněji popíšeme v následujícím příkladě, díky čemuž se v další části budeme moci bez újmy na obecnosti omezit na úlohy lineárního programování pouze v jednom konkrétním tvaru (a my si zvolíme kanonický tvar).

Příklad 2.2.2

(i) Máme-li úlohu lineárního programování v kanonickém tvaru

$$\langle c, x \rangle \to \min, \quad Ax = b, \quad x \geq 0, \tag{2.2.2}$$

pak ji můžeme vyjádřit jako úlohu ve standardním tvaru

$$\langle \widetilde{c}, \widetilde{x} \rangle \to \min, \quad \widetilde{A}\widetilde{x} \leq \widetilde{b}, \quad \widetilde{x} \geq 0, \tag{2.2.3}$$

v níž bereme

$$\widetilde{x} := x \in \mathbb{R}^n, \quad \widetilde{c} = c \in \mathbb{R}^n, \quad \widetilde{A} := \begin{pmatrix} A \\ -A \end{pmatrix} \in \mathbb{R}^{2m \times n} \quad \text{a} \quad \widetilde{b} := \begin{pmatrix} b \\ -b \end{pmatrix} \in \mathbb{R}^{2m}.$$

Je-li x^* řešením (2.2.2), pak $\widetilde{x}^* := x^*$ je také řešením úlohy (2.2.3) a hodnoty obou úloh jsou stejné, tj. $\langle c, x^* \rangle = \langle \widetilde{c}, \widetilde{x}^* \rangle$. Kdyby totiž úloha (2.2.3) měla řešení $\widetilde{x}^{**} \neq \widetilde{x}^*$ s menší funkční hodnotou, pak by $x^{**} := \widetilde{x}^{**}$ bylo řešením úlohy (2.2.2), což je spor ↯, neboť x^{**} je jejím přípustným bodem a opět $\langle c, x^{**} \rangle = \langle \widetilde{c}, \widetilde{x}^* \rangle$.

(ii) Máme-li úlohu lineárního programování ve standardním tvaru

$$\langle c, x\rangle \to \min, \quad Ax \le b, \quad x \ge 0, \tag{2.2.4}$$

pak ji můžeme vyjádřit jako úlohu v kanonickém tvaru

$$\langle \widetilde{c}, \widetilde{x}\rangle \to \min, \quad \widetilde{A}\widetilde{x} = \widetilde{b}, \quad \widetilde{x} \ge 0, \tag{2.2.5}$$

v níž s využitím vektoru pomocných proměnných $s := b - Ax$ bereme

$$\widetilde{x} := \begin{pmatrix} x \\ s \end{pmatrix} \in \mathbb{R}^{n+m}, \quad \widetilde{c} = \begin{pmatrix} c \\ 0 \end{pmatrix} \in \mathbb{R}^{n+m}, \quad \widetilde{A} := \begin{pmatrix} A, & I \end{pmatrix} \in \mathbb{R}^{m\times(n+m)} \quad \text{a} \quad \widetilde{b} := b \in \mathbb{R}^m.$$

Jestliže x^* je řešením úlohy (2.2.4), pak $\widetilde{x}^* := \begin{pmatrix} x^* \\ s^* \end{pmatrix} = \begin{pmatrix} x^* \\ b - Ax^* \end{pmatrix}$ je řešením (2.2.5) a hodnoty obou úloh jsou stejné, tj. $\langle c, x^*\rangle = \langle \widetilde{c}, \widetilde{x}^*\rangle$. Kdyby totiž úloha (2.2.5) měla řešení $\begin{pmatrix} \widetilde{x}^{**} \\ s^{**} \end{pmatrix}$ s $\widetilde{x}^{**} \ne \widetilde{x}^*$ s menší funkční hodnotou, pak by $x^{**} := \widetilde{x}^{**}$ bylo řešením (2.2.4), což je spor ↯, neboť x^{**} je jejím přípustným bodem a opět $\langle c, x^{**}\rangle = \langle \widetilde{c}, \widetilde{x}^{**}\rangle$.

(iii) Je-li úlohu lineárního programování v kanonickém tvaru

$$\langle c, x\rangle \to \min, \quad Ax = b, \quad x \ge 0, \tag{2.2.6}$$

pak ji můžeme vyjádřit jako úlohu v základním tvaru

$$\langle \widetilde{c}, \widetilde{x}\rangle \to \min, \quad \widetilde{A}\widetilde{x} \le \widetilde{b}, \tag{2.2.7}$$

kde bereme

$$\widetilde{x} := x \in \mathbb{R}^n, \quad \widetilde{c} := c \in \mathbb{R}^n, \quad \widetilde{A} := \begin{pmatrix} A \\ -A \\ -I \end{pmatrix} \in \mathbb{R}^{(2m+n)\times n} \quad \text{a} \quad \widetilde{b} := \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}.$$

Jestliže x^* je řešením úlohy (2.2.6), pak $\widetilde{x}^* := x^*$ je řešením (2.2.7) a hodnoty obou úloh jsou stejné, tj. $\langle c, x^*\rangle = \langle \widetilde{c}, \widetilde{x}^*\rangle$. Kdyby totiž úloha (2.2.7) měla řešení $\widetilde{x}^{**} \ne \widetilde{x}^*$ s menší funkční hodnotou, pak by $x^{**} := \widetilde{x}^{**}$ bylo řešením (2.2.6), což je spor ↯, neboť x^{**} je jejím přípustným bodem a opět $\langle c, x^{**}\rangle = \langle \widetilde{c}, \widetilde{x}^{**}\rangle$.

(iv) Naopak, máme-li úlohu lineárního programování v základním tvaru

$$\langle c, x\rangle \to \min, \quad Ax \le b, \tag{2.2.8}$$

pak ji můžeme vyjádřit jako úlohu v kanonickém tvaru

$$\langle \widetilde{c}, \widetilde{x}\rangle \to \min, \quad \widetilde{A}\widetilde{x} = \widetilde{b}, \quad \widetilde{x} \ge 0, \tag{2.2.9}$$

v níž s využitím pomocných nezáporných proměnných $u_1, \dots, u_n, v_1, \dots, v_n, w_1, \dots, w_m$ bereme

$$\widetilde{x} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^{2n+m}, \quad \widetilde{c} := \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix} \in \mathbb{R}^{2n+m}, \quad \widetilde{A} := \begin{pmatrix} A, & -A, & I \end{pmatrix} \in \mathbb{R}^{m\times(2n+m)} \quad \text{a} \quad \widetilde{b} := b \in \mathbb{R}^m.$$

Jestliže x^* je řešením úlohy (2.2.8), pak $\widetilde{x}^* := (u^*, v^*, w^*)$ se složkami

$$u_i^* := \begin{cases} x_i^*, & x_i^* \ge 0, \\ 0, & x_i < 0, \end{cases} \qquad v_i^* := \begin{cases} -x_i^*, & x_i^* \le 0, \\ 0, & x_i > 0, \end{cases} \qquad w := b - Ax^*$$

je řešením (2.2.7) a hodnoty obou úloh jsou stejné, tj. $\langle c, x^*\rangle = \langle \widetilde{c}, \widetilde{x}^*\rangle$. Kdyby totiž úloha (2.2.9) měla řešení $\widetilde{x}^{**}$ s menší funkční hodnotou, pak by $x^{**} := u^{**} - v^{**}$ bylo řešením (2.2.8), což je spor ↯, neboť x^{**} je jejím přípustným bodem[69] a opět $\langle c, x^{**}\rangle = \langle \widetilde{c}, \widetilde{x}^{**}\rangle$.

(v) Ukázali jsme, že všechny tři speciální tvary úlohy lineárního programování jsou skutečně vzájemně ekvivalentní. Zbývá ukázat jejich ekvivalenci s obecnou úlohou lineárního programování. Díky předchozí části se však můžeme omezit pouze na jeden ze speciální tvarů, takže ukážeme přechod od obecné úlohy k úloze v kanonickém tvaru (opačný směr je zcela triviální). Máme-li obecnou úlohu lineárního programování (2.2.1), tj.

$$\langle c, x\rangle \to \min, \quad A_N x \le b_N, \quad A_R x = b_R, \quad x_1, \dots, x_s \ge 0 \tag{2.2.10}$$

pak ji můžeme vyjádřit jako úlohu v kanonickém tvaru

$$\langle \widetilde{c}, \widetilde{x}\rangle \to \min, \quad \widetilde{A}\widetilde{x} = \widetilde{b}, \quad \widetilde{x} \ge 0, \tag{2.2.11}$$

v níž s využitím pomocných nezáporných proměnných $u_1, \dots, u_{n-s}, v_1, \dots, v_{n-s}, w_1, \dots, w_k$ bereme

$$\widetilde{x}_i := \begin{cases} x_i, & i = 1, \dots, s, \\ u_{i-s}, & i = s+1, \dots, n, \\ v_{i-n}, & i = n+1, \dots, 2n-s, \\ w_{i-(2n-s)}, & i = 2n-s+1, \dots, 2n-s+k, \end{cases}$$

$$\widetilde{c}_i := \begin{cases} c_i, & i = 1, \dots, n, \\ -c_{i-(n-s)}, & i = n+1, \dots, 2n-s, \\ 0, & i = 2n-s+1, \dots, 2n-s+k, \end{cases}$$

$$\widetilde{A} := \begin{pmatrix} A_{n,s} & A_{n,n} & -A_{n,n} & I_k \\ A_{r,s} & A_{r,n} & -A_{r,n} & 0 \end{pmatrix} \in \mathbb{R}^{m \times (2n-s+k)}, \qquad \widetilde{b} := \begin{pmatrix} b_N \\ b_R \end{pmatrix} \in \mathbb{R}^m$$

kde $\widetilde{x}, \widetilde{c} \in \mathbb{R}^{2n-s+k}$, I_k je $k \times k$ jednotková matice a $A_{n,s}$ značí submatici matice A_N tvořenou pouze prvními s sloupci, $A_{n,n}$ submatici tvořenou $(s+1)$-ním až n-tým sloupcem a podobně pro submatice $A_{r,s}$ a $A_{r,n}$. Opět platí $\langle c, x\rangle = \langle \widetilde{c}, \widetilde{x}\rangle$ a podobnými argumenty jako v předchozích případech můžeme ukázat těsné spojení mezi řešeními úloh (2.2.10) a (2.2.11).

Převod úlohy lineárního programování do kanonického tvaru si nyní ilustrujeme na dvou konkrétních příkladech.

Příklad 2.2.3

Převeďme úlohu

$$3x_1 - 2x_2 \to \min$$
$$2x_1 - x_2 \le 4 \quad \& \quad x_1 + 3x_2 \ge 5 \quad \& \quad x_2 \ge 0$$

do kanonického tvaru.

Řešení. Nejdříve si zavedeme pomocné proměnné, díky nimž daná nerovnostní omezení převedeme na rovnosti, tj. pomocí $s_1 \ge 0$ a $s_2 \ge 0$ máme

$$2x_1 - x_2 + s_1 = 4 \quad \text{a} \quad x_1 + 3x_2 - s_2 = 5.$$

Jelikož proměnná x_1 může nabývat libovolného znaménka, nahradíme ji pomocí dalších dvou proměnných $s_3 \ge 0$ a $s_4 \ge 0$ výrazem $x_1 = s_3 - s_4$, což nám již dává úlohu v kanonickém tvaru, tj.

$$3s_3 - 3s_4 - 2x_2 \to \min$$

za podmínek

$$2(s_3 - s_4) - x_2 + s_1 = 4 \quad \& \quad s_3 - s_4 + 3x_2 - s_2 = 5$$
$$x_2 \geq 0, \quad s_1 \geq 0, \quad s_2 \geq 0, \quad s_3 \geq 0, \quad s_4 \geq 0.$$

Z původní úlohy se dvěma proměnnými, dvěma nerovnostními omezeními a jedním omezením na znaménko jsme tak získali ekvivalentní úlohu s nezápornými pěti proměnnými a dvěma rovnostními omezeními. Všimněte si také, že první pomocné proměnné s_1 a s_2 nevystupují v účelové funkci.

Příklad 2.2.4

Převeďme úlohu

$$2x_1 - 3x_2 - 2x_3 + 7x_4 \to \max$$
$$-x_1 + 2x_2 + x_3 - x_4 = -6 \quad \& \quad x_1 - x_2 + 3x_3 + x_4 \geq -5$$
$$x_1 \geq -2 \quad \& \quad x_2 \leq 7 \quad \& \quad x_3 \leq 0 \quad \& \quad x_4 \geq 0.$$

do kanonického tvaru.
Řešení. Nejdříve si všimněme, že je tentokrát chceme účelovou funkci maximalizovat. To však jejím vynásobením -1 snadno převedeme na minimalizační úlohu. Dále zavedeme nezáporné pomocné proměnné s_1, s_2 a s_3, abychom nerovnostní omezení převedli na rovnost a vypořádali omezení pro x_1 a x_2, tj.

$$x_1 - x_2 + 3x_3 + x_4 - s_1 = -5 \quad \& \quad x_1 - s_2 = -2 \quad \& \quad x_2 + s_3 = 7,$$

takže všude dosadíme $x_1 = -2 + s_2$ a $x_2 = 7 - s_3$. Místo $x_3 \leq 0$ vezmeme $x_3 = -s_4$ s požadavkem $s_4 \geq 0$, což nás celkem přivádí k úloze lineárního programování v kanonickém tvaru

$$-2s_2 - 3s_3 - 2s_4 - 7x_4 + 25 \to \min$$
$$-s_2 - 2s_3 - s_4 - x_4 = -22 \quad \& \quad -s_1 + s_2 + s_3 - 3s_4 + x_4 = 4$$
$$x_4 \geq 0 \quad \& \quad s_1 \geq 0 \quad \& \quad s_2 \geq 0 \quad \& \quad s_3 \geq 0 \quad \& \quad s_4 \geq 0.$$

Všimněte si také, že absolutní člen $+25$ můžeme v nové účelové funkci zcela ignorovat, neboť nemá žádný vliv na řešení dané úlohy (ovlivní pouze hodnotu hledaného minima).

Už jsme se zmínili, že s jakousi první metodou řešení soustavy lineárních nerovnic přišel Fourier. V následující poznámce si tuto metodu představíme, abychom se mohli přesvědčit, že s postupem času je výpočet poněkud komplikuje a asi není příliš vhodný pro větší soustavy. O to více bychom pak mohli ocenit eleganci, kterou ukrývá simplexová metoda.

Poznámka 2.2.5 (historická, dlouhá a snad zajímavá). Fourierova metoda řešení soustavy lineárních nerovnic je založená na postupné *eliminaci proměnných* v soustavě $A \leq b$, tj. v prvním kroku chceme získat ekvivalentní soustavu $\widetilde{A}\,\widetilde{x} \leq \widetilde{b}$ pouze v proměnných $x_2, \ldots, x_n$, ze které následně eliminujeme x_2 atd. To je samozřejmě podobné jako v případě soustavy lineárních rovnic, avšak tentokrát musíme navíc dávat pozor na znaménka nerovností. Fourier proto navrhuje, abychom uvážili všechny dvojice nerovností s opačnými znaménky a s jejich pomocí eliminovali uvažovanou proměnnou. K tomu budeme potřebovat množiny

$$K(x_i) := \{\text{nerovnosti, ve kterých je u proměnné } x_i \text{ koeficient s kladným znaménkem}\},$$
$$Z(x_i) := \{\text{nerovnosti, ve kterých je u proměnné } x_i \text{ koeficient se záporným znaménkem}\},$$
$$N(x_i) := \{\text{nerovnosti, ve kterých je u proměnné } x_i \text{ nulový koeficient}\}.$$

Pokud každá z množin $K(x_i)$ a $Z(x_i)$ obsahuje zhruba polovinu celkového počtu nerovností, pak eliminací proměnné x_i získáme navíc nerovnosti, jejichž počet je přibližně roven druhé mocnině tohoto počtu. Jsou-li množiny $K(x_i)$ a $Z(x_i)$ prázdné, pak proměnná x_i je již ze soustavy vyeliminována. Je-li pouze jedna z těchto množin prázdná, tj. všechny koeficienty u proměnné x_i mají v dané soustavy stejné znaménko, stačí se omezit pouze na nerovnosti z $N(x_i)$, neboť všechny nerovnosti z $Z(x_i)$ mohou být snadno splněny s pomocí vhodné volby x_i při jakékoli kombinaci ostatních proměnných, která vyhovuje i nerovnostem z $N(x_i)$. Pokud po eliminaci všech n proměnných dostaneme soustavu $\widetilde{A}\widetilde{x} \le \widetilde{b}$ s nezáporným $\widetilde{b}$, pak původní soustava je řešitelná, zatímco v opačném případě alespoň jednoho záporného koeficientu v $\widetilde{b}$ je původní soustava neřešitelná.

Uvažme např. soustavu

$$-x_1 + 2x_2 \le 3 \quad \& \quad 2x_1 - 7x_2 \le -15 \quad \& \quad -x_1 \le 0 \quad \& \quad -x_2 \le 0.$$

Začneme eliminací x_1. Množina $K(x_1)$ je tvořena druhou nerovností, množina $Z(x_1)$ obsahuje první a třetí nerovnost a $N(x_1)$ zahrnuje pouze poslední nerovnost. Jednotlivé nerovnosti také můžeme zapsat jako

$$x_1 \ge 2x_2 - 3 \quad \& \quad x_1 \le \frac{7}{2}x_2 - \frac{15}{2} \quad \& \quad -x_1 \le 0 \quad \& \quad -x_2 \le 0$$

a nyní vytvoříme všechny možné kombinace nerovností s opačnými znaménky (a obsahující x_1), čímž obdržíme různé dolní a horní odhady pro x_1. Začneme zkombinováním první a druhé nerovnosti, z čehož dostaneme

$$2x_2 - 3 \le x_1 \le \frac{7}{2}x_2 - \frac{15}{2} \quad \text{neboli} \quad 2x_2 - 3 \le \frac{7}{2}x_2 - \frac{15}{2}, \tag{2.2.12}$$

a tudíž $x_2 \ge 3$. Podobně zkombinováním druhé a třetí nerovnosti získáme

$$0 \le x_1 \le \frac{7}{2}x_2 - \frac{15}{2} \quad \text{neboli} \quad 0 \le \frac{7}{2}x_2 - \frac{15}{2},$$

a tudíž $x_2 \ge 15/7$. Odtud vidíme, že dané soustavě vyhovují body s $x_2 \ge 3$. Zvolíme-li např. právě $x_2 = 3$, pak z (2.2.12) vidíme, že nutně $x_1 = 3$, zatímco pro $x_2 = 5$ budeme mít $7 \le x_1 \le 10$ atd.

Jak to ale souvisí s lineárním programováním? Nahradíme-li úlohu v základním tvaru

$$c^\top x \to \min \quad \& \quad Ax \le b$$

soustavou

$$c^\top x - z \le 0 \quad \& \quad Ax \le b,$$

pak eliminací proměnných $x_1, \ldots, x_n$ se dopracujeme k soustavě obsahující pouze z, z čehož určením nejmenší možné hodnoty proměnné z získáme řešení. Uvažme např. úlohu

$$\begin{gathered} 4x_1 - 5x_2 - 3x_3 \to \min \\ -x_1 + x_2 - x_3 \le 2 \\ x_1 + x_2 + 2x_3 \le 3 \\ x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \end{gathered}$$

a nahraďme ji soustavou nerovností

$$4x_1 - 5x_2 - 3x_3 - z \le 0 \tag{N0}$$
$$-x_1 + x_2 - x_3 \le 2 \tag{N1}$$
$$x_1 + x_2 + 2x_3 \le 3 \tag{N2}$$
$$-x_1 \le 0 \tag{N3}$$
$$-x_2 \le 0 \tag{N4}$$
$$-x_3 \le 0. \tag{N5}$$

Kombinace první a druhé nerovnosti nás přivede k

$$-2+x_2-x_3 \le x_1 \le (5\,x_2+3\,x_3+z)/4 \quad \text{neboli} \quad -2+x_2-x_3 \le (5\,x_2+3\,x_3+z)/4, \tag{2.2.13}$$

což můžeme přepsat do tvaru

$$-x_2-7\,x_3-z \le 8, \tag{2.2.14}$$

který bychom mohli získat přímo i tak, že bychom k nerovnosti (N0) přičetli čtyřnásobek nerovnosti (N1) – toto ukazuje podobnost s Gaussovou eliminací, akorát je tentokrát potřeba dávat mnohem větší pozor na znaménka jednotlivých koeficientů. Mělo by být zřejmé, že pokud existuje řešení dvojice (N0) a (N1), pak musí existovat i řešení nerovnosti (2.2.14). Naopak, pokud existuje řešení nerovnosti (2.2.14), pak z jejího přepisu do podoby druhé nerovnosti v (2.2.13) plyne, že existuje x_1 vyhovující první nerovnosti v (2.2.13). Odtud také zpětným dosazením povolených hodnot pro ostatní proměnné zjistíme, jaké hodnoty může x_1.

Jelikož se x_1 vyskytuje také v nerovnostech (N2) a (N3), musíme při jeho eliminaci uvážit ještě další možné kombinace nerovností s opačnými znaménky, tj. kombinaci (N0)+(N3), (N1)+(N2) a (N2)+(N3). Kdybychom některou z těchto kombinací vynechali, mohli bychom „zanedbat“ nějakou důležitou informaci o x_1 a získat i nepřípustné řešení. To nás přivede k nové soustavě

$$\begin{aligned} -x_2-7\,x_3-z &\le 8 && \text{(N0+4N1)}\\ -5\,x_2-3\,x_3-z &\le 0 && \text{(N0+4N3)}\\ 2\,x_2+x_3 &\le 5 && \text{(N1+N2)}\\ x_2+2\,x_3 &\le 3 && \text{(N2+N3)}\\ -x_2 &\le 0 && \text{(N4)}\\ -x_3 &\le 0 && \text{(N5)} \end{aligned}$$

kde jednotlivá označení ukazují, jak jsme jednotlivé nerovnosti získali. Tato soustava je ekvivalentní s tou předchozí, tj. pokud existuje nějaké řešení soustavy (N0)–(N5), pak odpovídající hodnoty z, x_2, x_3 vyhovují také předchozí soustavě, a naopak pokud předchozí soustava má nějaké řešení, pak existuje x_1 tak, že tato čtveřice z, x_1, x_2, x_3 vyhovuje původní soustavě.

Nyní eliminujeme proměnnou x_2, k čemuž využijeme celkem 6 možných kombinací předchozích nerovností. Výsledkem pak bude soustava

$$\begin{aligned} -13\,x_3-2\,z &\le 21 && \text{(2(N0+4N1)+(N1+N2))}\\ -x_3-2\,z &\le 25 && \text{(2(N0+4N3)+5(N1+N2))}\\ x_3 &\le 5 && \text{((N1+N2)+2N4)}\\ -5\,x_3-z &\le 11 && \text{((N0+4N1)+(N2+N3))}\\ 7\,x_3-z &\le 15 && \text{((N0+4N3)+5(N2+N3))}\\ 2\,x_3 &\le 3 && \text{((N2+N3)+N4)}\\ -x_3 &\le 0 && \text{(N5)} \end{aligned}$$

Přestože jsme už eliminovali dvě proměnné, stále máme šest nerovností. Jsou ale všechny skutečně potřebné? Nemůžeme některé z nich ignorovat? Naštěstí můžeme. V [327, Section 5] bylo ukázáno, že pokud po eliminaci m proměnných máme omezení závisející na $m+1$ původních omezeních, pak je toto omezení nutně redundantní. V předchozí soustavě se to týká nerovností (2(N0+4N3)+5(N1+N2)) a ((N0+4N1)+(N2+N3)), které vznikly postupnými kombinacemi založenými na 4 původních nerovnostech (N0, N1, N2 a N3). Proto se můžeme v dalších úvahách omezit pouze na soustavu

$$\begin{aligned} -13\,x_3-2\,z &\le 21 && \text{(2(N0+4N1)+(N1+N2))}\\ x_3 &\le 5 && \text{((N1+N2)+2N4)} \end{aligned}$$

$$\begin{aligned} 7x_3 - z &\le 15 && ((\text{N0}+4\text{N3})+5(\text{N2}+\text{N3})) \\ 2x_3 &\le 3 && ((\text{N2}+\text{N3})+\text{N4}) \\ -x_3 &\le 0 && (\text{N5}) \end{aligned}$$

a zbývá eliminovat ještě x_3. Tím získáme soustavu

$$\begin{aligned} -2z &\le 86 && ((2(\text{N0}+4\text{N1})+(\text{N1}+\text{N2}))+13((\text{N1}+\text{N2})+2\text{N4})) \\ -27z &\le 342 && (7(2(\text{N0}+4\text{N1})+(\text{N1}+\text{N2}))+13((\text{N0}+4\text{N3})+5(\text{N2}+\text{N3}))) \\ -4z &\le 81 && (2(2(\text{N0}+4\text{N1})+(\text{N1}+\text{N2}))+13((\text{N2}+\text{N3})+\text{N4})) \\ 0 &\le 5 && (((\text{N1}+\text{N2})+2\text{N4})+\text{N5}) \\ -z &\le 15 && (((\text{N0}+4\text{N3})+5(\text{N2}+\text{N3}))+7\text{N5}) \\ 0 &\le 3 && (((\text{N2}+\text{N3})+\text{N4})+2\text{N5}) \end{aligned}$$

neboli

$$z \ge -43 \quad \& \quad z \ge -38/3 \quad \& \quad z \ge -81/4 \quad \& \quad z \ge -15,$$

takže nejmenší možná hodnota je $z^* = -38/3$. Toto jsme dostali z omezení, které vzniklo kombinací nerovností (2(N0+4N1)+(N1+N2)) a ((N0+4N3)+5(N2+N3)), takže když do nich dosadíme z^*, stanou se z nich rovnosti

$$-13x_3 = -13/3 \quad \& \quad 7x_3 = 7/3,$$

jejichž řešením je $x_3^* = 1/3$. Tato dvě omezení jsme získali z nerovností (N0+4N1), (N1+N2), (N0+4N3) a (N2+N3), což nás po dosazení hodnot z^* a x_3^* přivede ke čtveřici rovnic

$$-x_2 = -5/3 \quad \& \quad 2x_2 = 14/3 \quad \& \quad -5x_2 = -35/3 \quad \& \quad x_2 = 7/3,$$

jejichž jediným řešením je $x_2^* = 7/3$. Všechna tato omezení vzešla z (N0), (N1), (N2) a (N3), která nás po dosazení již známých hodnot přivedou k rovnostem

$$4x_1 = 0 \quad \& \quad -x_1 = 0 \quad \& \quad x_1 = 0 \quad \& \quad -x_1 = 0$$

s jediným řešením $x_1^* = 0$. Řešení daného problému proto je $[x_1^*, x_2^*, x_3^*] = [0, 7/3, 1/3]$ dávající minimální hodnotu účelové funkce $z^* = -38/3$.

Tímto způsobem se nám nejenom podařilo najít řešení dané úlohy, ale z omezení

$$(7(2(\text{N0}+4\text{N1})+(\text{N1}+\text{N2}))+13((\text{N0}+4\text{N3})+5(\text{N2}+\text{N3}))),$$

které nás přivedlo k z^*, dokonce můžeme zjistit, jakou kombinací jednotlivých omezení získáme nejmenší hledanou hodnotu. Protože původní účelová funkce se vyskytuje pouze v (N0) společně s $-z$, pak z omezení vytvořeného jako

$$7(2(N0+4N1)+(N1+N2))+13((N0+4N3)+5(N2+N3))$$

neboli $27N0+63N1+72N2+117N3$ podělením -27 budeme mít

$$-N0-\frac{7}{3}N1-\frac{8}{3}N2-\frac{13}{3}N3,$$

což nás přivede k

$$-\frac{7}{3}(-x_1+x_2-x_3 \le 2)-\frac{8}{3}(x_1+x_2+2x_3 \le 3)-\frac{13}{3}(-x_1 \le 0),$$

odkud v kombinaci s (N0) máme nejlepší dolní odhad

$$z \ge 4x_1-5x_2-3x_3 \ge \frac{38}{3}.$$

Tento výsledek možná není v tuto chvíli až tak zajímavý, ale je to předzvěst velmi důležité *věty o dualitě*, které se budeme věnovat později.

Omezení ((N1+N2)+2N4)+N5) a (((N2+N3)+N4)+2N5) byla triviální a pro naše další úvahy nepotřebné, ale současně jistým způsobem potvrzovaly přípustnost dané úlohy. Kdybychom totiž místo nich měli omezení typu $0 \le -1$, znamenalo by to, že daná úloha je nepřípustná. Z poslední soustavy bychom také poznali případnou neohraničenost zdola dané úlohy, tj. nenalezli bychom žádnou konečnou dolní hranici pro hodnotu z. ▲

Cvičení

2.2.1. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 - x_2 \to \min$$
$$-2x_1 + x_2 \le 2 \quad \& \quad 3x_1 + x_2 \ge 5 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.2. Převeďte úlohu lineárního programování do kanonického tvaru

$$6x_1 + x_2 \to \min$$
$$-5x_1 + 8x_2 \le 80 \quad \& \quad x_1 + 2x_2 \ge 4 \quad \& \quad x_1 \le 10 \quad \& \quad x_2 \ge 0.$$

2.2.3. Převeďte úlohu lineárního programování do kanonického tvaru

$$3x_1 - 2x_2 \to \min$$
$$2x_1 - x_2 \le 4 \quad \& \quad x_1 + 3x_2 \ge 5 \quad \& \quad x_2 \ge 0.$$

2.2.4. Převeďte úlohu lineárního programování do kanonického tvaru

$$4x_1 + x_2 \to \min$$
$$-2x_1 + x_2 \ge 6 \quad \& \quad x_2 + x_3 = 4 \quad \& \quad x_1 \ge -4 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.2.5. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 + x_2 + x_3 \to \min$$
$$2x_1 + x_2 = 3 \quad \& \quad 4x_1 + 5x_2 + 3x_3 = 7 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.6. Převeďte úlohu lineárního programování do kanonického tvaru

$$3x_1 - 4x_2 + x_3 \to \min$$
$$-x_1 + 5x_3 = 50 \quad \& \quad 2x_1 - 3x_2 = 12 \quad \& \quad x_3 \ge 0.$$

2.2.7. Převeďte úlohu lineárního programování do kanonického tvaru

$$-x_1 + 3x_2 + 4x_3 \to \min$$
$$4x_1 - 3x_2 + 5x_3 = 18 \quad \& \quad -x_1 + x_2 - 3x_3 \le 100$$
$$0 \le x_1 \le 10 \quad \& \quad -2 \le x_2 \le 5 \quad \& \quad x_3 \ge 0.$$

2.2.8. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 - 2x_2 + 3x_3 \to \max$$
$$x_1 + x_2 + x_3 \le 7 \quad \& \quad x_1 - x_2 + x_3 \le -2 \quad \& \quad 3x_1 + 2x_3 = 5$$
$$x_2 - x_3 \ge 1 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.9. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 - 2x_2 - 3x_3 \to \min$$
$$x_1 + 2x_2 + x_3 \le 14 \quad \& \quad x_1 + 2x_2 + 4x_3 \ge 12 \quad \& \quad x_1 - x_2 + x_3 = 2 \quad \& \quad x_3 \le -3.$$

2.2.10. Převeďte úlohu lineárního programování do kanonického tvaru

$$6x_1 - 2x_2 + 9x_3 + 300 \to \min$$
$$2x_1 - 6x_2 - x_3 \le 100 \quad \& \quad x_1 + x_2 + 9x_3 \le 200 \quad \& \quad 0 \le x_1 \le 50,$$
$$x_2 \ge -50 \quad \& \quad x_3 \ge 5.$$

2.2.11. Převeďte úlohu lineárního programování do kanonického tvaru

$$6x_1 + 5x_2 - x_3 \to \min$$
$$x_1 + 2x_2 - 7x_3 \le 1 \quad \& \quad |3x_1 - 5x_2 - 20| \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.12. Převeďte úlohu lineárního programování do kanonického tvaru

$$2x_1 + 3x_2 + 5x_3 \to \min$$
$$x_1 + x_2 - x_3 \ge -5 \quad \& \quad -6x_1 + 7x_2 - 9x_3 = 15 \quad \& \quad |19x_1 - 7x_2 + 5x_3| \le 15$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.13. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 - x_2 \to \max$$
$$-x_1 + 2|x_2| \le 0 \quad \& \quad |3x_1 - x_2| \le 3 \quad \& \quad x_1 \ge 0.$$

2.2.14. Převeďte úlohu lineárního programování do kanonického tvaru

$$3x_1 - 2x_2 + 4|x_3| \to \min$$
$$x_1 - 2x_2 \ge 5 \quad \& \quad 3x_2 - x_3 \ge 6 \quad \& \quad 2x_1 + x_3 = 12$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.2.15. Převeďte úlohu lineárního programování do kanonického tvaru

$$|x_1| + |x_2| + |x_3| \to \min$$
$$x_1 + x_2 \le 1 \quad \& \quad 2x_1 + x_3 = 3.$$

2.2.16. Převeďte úlohu lineárního programování do kanonického tvaru

$$3x_1 - x_2 \to \min$$
$$-x_1 + 6x_2 - x_3 + x_4 \ge -3 \quad \& \quad 7x_2 + x_4 = 5 \quad \& \quad x_3 + x_4 \le 2,$$
$$-1 \le x_2, x_3 \le 5 \quad \& \quad -2 \le x_4 \le 2.$$

2.2.17. Převeďte úlohu lineárního programování do kanonického tvaru

$$2x_1 - 2x_2 - x_3 + x_4 \to \max$$
$$4x_1 - x_2 + x_4 \le 6 \quad \& \quad -7x_1 + 8x_2 + x_3 \ge 7 \quad \& \quad x_1 + x_2 + 4x_4 = 12,$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.2.18. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_2 + x_3 + x_4 \to \max$$
$$x_1 + x_2 \ge 6 \quad \& \quad x_2 + x_3 - x_4 \le 1 \quad \& \quad 5x_1 - 6x_2 + 7x_3 - 8x_4 \ge 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \le 0.$$

2.2.19. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 + x_3 - x_4 + 48 \to \min$$
$$-3x_1 + x_2 - x_3 + 2x_4 = -50 \quad \& \quad x_1 - x_2 + x_4 \le 100 \quad \& \quad 2x_2 - x_3 - x_4 \ge 150$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.2.20. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 + 2x_2 + 4x_3 \to \max$$
$$|4x_1 + 3x_2 - 7x_3| \le x_1 + x_2 + x_3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.2.21. Převeďte úlohu lineárního programování do kanonického tvaru

$$x_1 + 6x_2 + 12x_3 \to \max$$
$$-x_1 - x_2 + x_4 \ge \max\{7x_1 + 2x_2, 5x_2 + x_3 + x_4\}$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.2.22. Převeďte úlohu lineárního programování do kanonického tvaru

$$-x_1 - x_2 + 2x_3 + x_5 \to \min$$
$$x_1 + 7x_2 + 16x_3 \le 4x_4 + x_5 \quad \& \quad x_3 + 12x_4 \ge x_1 + 6x_2 \quad \& \quad 9x_5 \le x_2 + 3x_4,$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0.$$

2.2.23. Ukažte, že problém

$$x_1 + 2x_2 - 3x_3 + 4x_4 \to \min$$
$$3x_1 - 2x_2 + 5x_3 - 6x_4 = 20 \quad \& \quad x_1 + 7x_2 - 6x_3 + 9x_4 = 30 \quad \& \quad x_1 \geq 0$$

je ekvivalentní s úlohou

$$x_1 + 2x_2 - 3x_3 + 4x_4 - 3x_5 \to \min$$
$$3x_1 - 2x_2 + 5x_3 - 6x_4 + 3x_5 = 20 \quad \& \quad x_1 + 7x_2 - 6x_3 + 9x_4 - 10x_5 = 30,$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0.$$

2.3 Lineární programování všude kolem nás

K několika již dříve zmíněným úlohám vedoucím na problémy lineárního programování si nyní přidáme některé další, abychom alespoň trochu více naznačili šíři a užitečnost celé této oblasti. Již asi všichni velmi dobře víme (a vlastně jsme se o tom už i několikrát zmínili), že úlohy lineární programování budou užitečné zejména ve chvíli, kdy chceme optimalizovat jistou veličinu závisející lineárně na několika zdrojích, jichž však nemusí být neomezené množství, protože jejich dostupnost je vymezena lineárními podmínkami. Avšak takové problémy mohou mít mnoho podob. Všimněme si také, že ekonomické úlohy bývají z pochopitelných důvodů ve většině případů formulované v maximalizačním tvaru, zatímco s minimalizací se v ekonomii setkáme např. při optimalizování nákladové funkce.

Velkou skupinu problémů lineárního programování tvoří hledání tzv. *optimálního výrobního programu*. Jde o situaci, kdy vyrábíme n různých výrobků s pomocí m-tice různých zdrojů, přičemž na výrobu jednoho kusu j-tého produktu potřebujeme a_{ij} jednotek i-tého zdroje. Zdroje jsou však omezené a my máme k dispozici pouze b_i jednotek i-tého zdroje a zisk z výroby 1 kusu j-tého výrobku je c_j. Vyrobíme-li celkem x_j kusů j-tého výrobku, pak lineární program pro maximalizaci zisku je

$$\sum_{j=1}^{n} x_j c_j \to \max \quad \& \quad Ax \leq b \quad \& \quad x \geq 0.$$

Příklad 2.3.1

(*Úloha o výrobě*) Společnost Tužka & Guma, a.s., vyrábí diáře a stolní kalendáře, přičemž pro zjednodušení připusťme, že k výrobě obou výrobků jsou potřeba pouze 3 zdroje: papír, kroužková vazba a lidská práce. Zisk z jednoho diáře je 100 Kč a zisk z jednoho kalendáře pak 180 Kč. K výrobě jednoho diáře potřebujeme 18 ks papíru, jednu kroužkovou vazbu a 30 minut lidské práce. K výrobě stolního kalendáře využijeme 28 ks papíru, 2 kroužkové vazby a 20 minut práce. Všech zdrojů je omezené množství určené např. pracovní dobou pracovníků nebo množstvím kroužkových vazeb ve skladu firmy. Maximální počet papírů, který můžeme za den využít, je 1000, maximální počet kroužkových vazeb je 80 a maximální čas výroby je 900 minut denně (firma totiž zaměstnává pouze 2 zaměstnance). Jaký je optimální plán denní výroby této firmy, která chce maximalizovat zisk z prodeje svých výrobků?

Řešení. Ze slovního zadání si nejprve sestavíme matematickou úlohu. Označme si d počet diářů a k počet kalendářů, které firma vyrobí za den. Papír, vazbu a práci označíme postupně p, v, ℓ. Firma vlastně hledá maximum funkce

$$f(d,k) = 100\,d + 180\,k$$

za daných podmínek, kterými jsou omezení na zdroje. První omezení se týká využití papíru, kterého potřebujeme 18 na výrobu diáře a 28 na výrobu kalendáře. Naše celková spotřeba papíru je tedy dána $18\,d + 28\,k$. Tato spotřeba je ale omezená naší zásobou papíru, která je 1000. Z těchto informací již můžeme sestavit první omezení, jemuž

naše firma čelí, a to

$$18\,d + 28\,k \leq 1000.$$

S ostatními zdroji postupujeme obdobně. Pro kroužkové vazby tímto způsobem získáme nerovnostní omezení

$$d + 2\,k \leq 80,$$

z omezeného času práce pak dostaneme

$$30\,d + 20\,k \leq 900.$$

Poslední omezení, která k úloze přidáme, jsou omezení

$$d \geq 0 \quad \text{a} \quad k \geq 0,$$

neboť firma chce vyrábět nezáporná množství diářů a kalendářů. Výroba záporného množství možná na první pohled nedává smysl, ekonomicky ji však můžeme chápat jako nakupování výsledných produktů, tj. zde diářů a kalendářů. [Řešením je $d = 0$ a $k = 35\frac{5}{7}$.]

Optimální „výrobní program“ může mít i složitější podobu, a ne nutně musí jít o výrobu nějakých produktů.

Příklad 2.3.2

Zemědělec vlastní 1000 arů[70] (nebo 10 hektarů) víceméně homogenní zemědělské plochy. Na této ploše může chovat skot nebo pěstovat pšenici, kukuřici či rajčata. Jeden kus skotu zabere 4 ary. jedná se o rodinnou farmu, takže k dispozici je 12 000 pracovních hodin, které mohou být využity kdykoli a jakkoli v průběhu celého roku. Následující tabulka ukazuje podrobnosti týkající se zemědělské produkce

	skot	pšenice	kukuřice	rajčata
chovatelné/pěstovatelné množství na 1 aru	1/4	50 bušlů[71]	80 bušlů	500 Kg
potřebná úhrnná roční pracovní doba	40 hod/ks	10 hod/ar	12 hod/ar	25 hod/ar
zisk	32 000 Kč/kus	100 Kč/bušl	120 Kč/bušl	20 Kč/bušl

Navíc je nutné ponechat navíc alespoň část odpovídající 20 % půdy z celkové obhospodařované plochy pro skot a nejvýše 30 % může být využito pro pěstování rajčat, přičemž poměr mezi velikostí půdy určené k pěstování pšenice a neobhospodařované plochy nesmí překročit 2 : 1. Jak rozdělit zemědělskou půdu, aby bylo dosaženo maximálního zisku při splnění všech daných omezení?

Řešení. Označme symbolem x_1 velikost plochy určené pro chov skotu, x_2 pro pěstování pšenice, x_3 pro kukuřici a x_4 pro rajčata. Pak účelová funkce určující dosažený zisk je

$$f(x_1, x_2, x_3, x_4) = 32\,000\,x_1/4 + 100 \cdot 50\,x_2 + 120 \cdot 80\,x_3 + 20 \cdot 500\,x_4$$

při omezení daném velikostí pozemku

$$x_1 + x_2 + x_3 + x_4 \leq 1000,$$

časovou náročností a dostupnému množství pracovních hodin

$$40\,x_1/4 + 10\,x_2 + 12\,x_3 + 25\,x_4 \leq 12\,000,$$

[70] 1 ar je 10×10 m^2

[71] *Bušl* (angl. *bushel*) je stará objemová a hmotnostní míra (se zkratkou bu nebo bsh), která se používá převážně jako měrná jednotka pro obilniny na angloamerických komoditních burzách. Např. 1 bušl pšenice je cca 27,216 kg a v případě kukuřice jde o cca 25,4016 kg.

dodatečnými požadavky na velikost obhospodařované plochy

$$x_1 \geq 0,2\,(x_1 + x_2 + x_3 + x_4)$$
$$x_4 \leq 0,3 \cdot 1000$$
$$x_2 : (1000 - x_1 - x_2 - x_3 - x_4) \leq 2 : 1.$$

Pochopitelně nelze alokovat záporné množství půdy, takže pro našeho zemědělce dostáváme lineární program

$$8000\,x_1 + 5000\,x_2 + 9600\,x_3 + 1000\,x_4 \to \max$$
$$x_1 + x_2 + x_3 + x_4 \leq 1000$$
$$10\,x_1 + 10\,x_2 + 12\,x_3 + 25\,x_4 \leq 12\,000$$
$$8\,x_1 - 2\,x_2 - 2\,x_3 - 2\,x_4 \geq 0$$
$$x_4 \leq 3002\,x_1 + 3\,x_2 + 2\,x_3 + 2\,x_4 \leq 2000$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

[Přesným řešením této úlohy je $x_1 = 200$, $x_2 = 0$, $x_3 = 769,23$ a $x_4 = 30,76$.]

Do téže kategorie spadá v jistém smyslu i tvorba optimálního portfolia z následujícího příkladu. Všimněme si, že zde jsou jasně dány výnosy z jednotlivých komponent portfolia. To však není zcela obvyklé, neboť výnosy mohou fluktuovat a jednotlivé složky se mohou vzájemně ovlivňovat. Proto realističtější modely na tvorbu portfolia mívají spíše podobu úlohy kvadratického programování, kterému se budeme věnovat v Kapitole 4.

Příklad 2.3.3

(*Rozdělení portfolia*) Investor má k dispozici kapitál ve výši 20 milionů Kč, který by chtěl investovat do nějaké kombinace dluhopisů, cenných papírů, termínovaného vkladu, spořícího účtu, nemovitého majetku a zlata. Předpokládaná úroková míra, míra rizika a předpokládaný růst hodnoty investice jsou shrnuty v následující tabulce.

	výnos/úroky [%, počítáno ročně]	riziko	předpokládaný růst [%]
dluhopisy	5	3	0
cenné papíry	2	10	7
termínovaný vklad	4	2	0
spořící účet	3	1	0
nemovitosti	0	5	7
zlato	0	20	11

Např. nemovitosti nemají žádný výnos (nejedná se o pronájem), ale jejich hodnota by měla růst o 7 % ročně. Na druhou stranu cenné papíry se průměrně zúročí o 2 % a navíc jejich cena by měla vzrůst o 7 %, tj. celkový roční výnos je 9 %. Cílem je maximalizovat očekávanou hodnotu investice v průběhu let za podmínek:

(i) z celkového objemu investic musí být nejméně 30 % v dluhopisech, nejvýše 10 % v cenných papírech a nejvýše 10 % v termínovaných vkladech a/nebo spořících účtech;

(ii) nejvýše 50 % z objemu investic do nemovitostí může být získáno zpět formou hypotéky s úrokem 6 %. Nejvýše však 3 miliony Kč.

(iii) průměrné riziko vztažené na objem investovaných peněz nesmí převýšit 9/2.

(iv) průměrný roční úrok vzhledem k objemu investovaných peněz musí být alespoň 2,5 %.

(v) do zlata nelze investovat více než 2 miliony Kč nebo 8 % dostupných financí (podle toho, co je menší).

Řešení. Máme jediný zdroj – peníze, které můžeme rozdělit do 6 druhů investic. Pro $i \in \{1,2,\dots,6\}$ označme x_i jako množství investované do i-tého typu a x_7 bude představovat objem vypůjčených peněz. Potom jde o maximalizaci funkce

$$1{,}05\,x_1 + 1{,}09\,x_2 + 1{,}04\,x_3 + 1{,}03\,x_4 + 1{,}07\,x_5 + 1{,}11\,x_6 - 1{,}06\,x_7$$

za podmínek

$$\begin{gathered}
x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \le 20\,000\,000 + x_7 \\
x_1 \ge 0{,}3 \cdot (x_1 + x_2 + x_3 + x_4 + x_5 + x_6) \\
x_2 \le 0{,}1 \cdot (x_1 + x_2 + x_3 + x_4 + x_5 + x_6) \\
x_3 + x_4 \le 0{,}1 \cdot (x_1 + x_2 + x_3 + x_4 + x_5 + x_6) \\
x_7 \le 0{,}5\,x_5 \\
x_7 \le 3\,000\,000 \\
\frac{3\,x_1 + 10\,x_2 + 2\,x_3 + x_4 + 5\,x_5 + 20\,x_6}{x_1 + x_2 + x_3 + x_4 + x_5 + x_6} \le \frac{9}{2} \\
\frac{5\,x_1 + 2\,x_2 + 4\,x_3 + 3\,x_4}{x_1 + x_2 + x_3 + x_4 + x_5 + x_6} \ge \frac{5}{2} \\
x_6 \le 2\,000\,000 \\
x_6 \le 0{,}08 \cdot (20\,000\,000 + x_7) \\
x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0 \quad \& \quad x_6 \ge 0 \quad \& \quad x_7 \ge 0,
\end{gathered}$$

což už neměl být problém přepsat do tvaru lineárního programu.
[Přesným řešením je $x_1 = 8\,740\,000$, $x_2 = 2\,300\,000 = x_3$, $x_4 = 0$, $x_5 = 9\,568\,000$, $x_6 = 92\,000$ a $x_7 = 3\,000\,000$.]

Nyní si ukážeme další typický učebnicový příklad, který se objevuje především v souvislosti s celočíselným programováním, jemuž se budeme věnovat v příští kapitole. Měli bychom totiž mít na paměti, že ne vždy je možné nalezené řešení skutečně realizovat (jako např. odvážit zcela přesně $26\frac{2}{3}$ kg kávy). Proto se k tomuto příkladu později vrátíme, abychom pro změnu „pomohli“ našemu zloději v knihkupectví, kde již skutečně nedává žádný smysl krást 7/3 knih.

Příklad 2.3.4

(*Problém batohu/zloděje*) Lupič se rozhodl vykrást sklad pražírny kávy. Zde je celkem 6 druhů kávy/kávových bodů. Zloděj má k dispozici batoh se 6 pytli o stejném objemu jako samotný batoh, tj. 40 litrů. Tyto pytle by chtěl naplnit tak, aby se vešly do batohu a jejich hodnota byla maximální vzhledem k ceně jednotlivých druhů uvedených v následující tabulce.

	1	2	3	4	5	6
cena Kč/kg	800	450	900	700	1500	600
objem l/kg	2	0,9	1,5	2	4	1

Řešení. Pro $i \in \{1,2,\ldots,6\}$ označme jako x_i množství i-tého druhu kávy, které zloděj dal do i-tého pytle. Pak jeho úkolem vlastně je nalezení řešení lineárního programu

$$800\,x_1 + 450\,x_2 + 900\,x_3 + 700\,x_4 + 1500\,x_5 + 600\,x_6 \to \max$$
$$2\,x_1 + 0{,}9\,x_2 + 1{,}5\,x_3 + 2\,x_4 + 4\,x_5 + x_6 \le 40$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0 \quad \& \quad x_6 \ge 0.$$

[Přesným řešením tohoto problému je volba $x_3 = 26\frac{2}{3}$ a $x_1 = x_2 = x_4 = x_5 = x_6 = 0$.]

Další významnou skupinou jsou tzv. *směšovací problémy*, v nichž máme n-tici základních surovin, ze kterých chceme namíchat výrobek předepsaného složení a současně minimalizovat surovinové náklady. Proměnná x_j zde reprezentuje spotřebované množství j-té suroviny mající jednotkovou cenu c_j. Jednotkové množství j-té suroviny obsahuje a_{ij} jednotek i-té látky a požadované složení je přesně popsáno vektorem b, jehož složky b_i jsou rovny požadovanému obsahu i-té látky ve výsledné směsi. Cílem pak je nalezení řešení lineárního programu

$$\sum_{j=1}^{n} x_j\,c_j \to \max \quad \& \quad Ax = b \quad \& \quad x \ge 0.$$

Jedním z ukázkových příkladů tohoto typu může být optimalizace výroby krmné směsi[72]. S těmito úlohami se ale lze setkat např. i při míchání oceli, při vytváření strategií reklamních kampaní[73] nebo při hledání optimální léčby s využitím léků různého složení, viz [330]. My si však nyní raději naznačíme využití těchto úloh v petrochemickém průmyslu a vyrazíme tentokrát do rafinérie.

[72] Do této skupiny se obvykle řadí i problém optimálního jídelníčku, kde se ovšem většinou nevyžaduje zcela přesné dodržení všech nutričních požadavků, tj. omezení je spíše ve tvaru $Ax \ge b$ (viz rozdílné výsledky Stiglerova problému uvedeného na začátku této kapitoly). Současně nám v takové úloze nejde o maximalizaci nějakého zisku, ale o minimalizaci ceny jídelníčku. Zde se tedy poněkud stírá rozdíl mezi obecnou formulací optimálního výrobního programu a směšovací úlohy, avšak v duchu této terminologie je asi příhodnější zařadit hledání jídelníčku do té druhé skupiny.

[73] U každé ze strategií známe dosah a také cenu. Pak se snažíme oslovit určitý počet lidí (např. napříč regiony, demografickými skupinami, atp.) s využitím co nejmenšího rozpočtu.

Příklad 2.3.5

(*Směšovací úloha*) Nyní si představme, že jsme poradci prodejců benzínu, kteří chtějí namíchat paliva s různými oktanovými čísly (např. Natural 91, 95 a 98). K dispozici je 5 různých druhů surových paliv, která se liší oktanovými čísly, množstvím, které má prodejce k dispozici, a cenou (v dolarech za barel). Údaje pro konkrétní typy (označené čísly 1-5) jsou zobrazeny v následující tabulce.

palivo	oktanové číslo	prodejní cena	dostupné množství
1	75	10	2200
2	80	12,5	4000
3	85	13	5000
4	90	27,5	5000
5	99	27,5	3000

Oktanové číslo paliva, které chce prodejce namíchat, získáme jako vážený průměr oktanových čísel použitých surových paliv. Za prodej barelu Naturalu 91, 95 a 98, vydělá prodejce postupně 36,5, 39 a 42 USD. Pokud mu zbude nějaké množství surových paliv, které nakoupil, může ho prodat za tržní cenu. Cílem prodejce je maximalizovat zisk. Jakou strategii poradíme prodejci, který disponuje daným množstvím jednotlivých surových paliv, aby dosáhl svého cíle?

Řešení. Stejně jako v předchozích příkladech začneme s formálním zápisem úlohy. Označme množství Naturalu 91, 95 a 98 jako y_1, y_2 a y_3. Dále označme $z_1,\ldots,z_5$ množství surových paliv, které jsme nakoupili, ale nespotřebovali. Zisk z jejich prodeje tedy

také vstoupí do ziskové funkce, kterou můžeme zapsat jako

$$36{,}5\,y_1 + 39\,y_2 + 42\,y_3 + 10\,z_1 + 12{,}5\,z_2 + 13\,z_3 + 27{,}5\,z_4 + 27{,}5\,z_5.$$

Tuto funkci chce prodejce maximalizovat. Čelí však několika omezením. Označme si x_{ij} množství surového paliva i využitého k přípravě prodejného paliva j, přičemž $i \in \{1,\dots,5\}$ a $j \in \{1,2,3\}$. Proměnná x_{23} tak reprezentuje množství paliva 2 využitého k přípravě Naturalu 98. Z tabulky známe maximální dostupná množství jednotlivých paliv, s jejich využitím získáme podmínky

$$\begin{aligned} x_{11} + x_{12} + x_{13} + z_1 &= 2200 \\ x_{21} + x_{22} + x_{23} + z_2 &= 4000 \\ x_{31} + x_{32} + x_{33} + z_3 &= 5000 \\ x_{41} + x_{42} + x_{43} + z_4 &= 5000 \\ x_{51} + x_{52} + x_{53} + z_5 &= 3000. \end{aligned}$$

Tato omezení jsou ve tvaru rovností, neboť surové palivo buď použijeme na výrobu směsi, nebo prodáme dále za tržní cenu. Využijeme tedy vše, co máme. Pro množství jednotlivých paliv, která vyrobíme, platí rovnosti

$$\begin{aligned} y_1 &= x_{11} + x_{21} + x_{31} + x_{41} + x_{51} \\ y_2 &= x_{12} + x_{22} + x_{32} + x_{42} + x_{52} \\ y_3 &= x_{13} + x_{23} + x_{33} + x_{43} + x_{53}. \end{aligned}$$

Posledním faktorem, který prodejce zajímá, je oktanové číslo směsi, kterou vyrobí. Pro y_1 máme podmínku

$$\frac{75\,x_{11} + 80\,x_{21} + 85\,x_{31} + 90\,x_{41} + 99\,x_{51}}{x_{11} + x_{21} + x_{31} + x_{41} + x_{51}} = \frac{75\,x_{11} + 80\,x_{21} + 85\,x_{31} + 90\,x_{41} + 99\,x_{51}}{y_1} \geq 91,$$

kterou můžeme také zapsat jako

$$75\,x_{11} + 80\,x_{21} + 85\,x_{31} + 90\,x_{41} + 99\,x_{51} \geq 91\,y_1.$$

Pro ostatní paliva obdobně získáme další podmínky, kterým při hledání maximálního zisku čelíme. Pro y_2 a y_3 to jsou podmínky

$$\begin{aligned} 75\,x_{12} + 80\,x_{22} + 85\,x_{32} + 90\,x_{42} + 99\,x_{52} &\geq 95\,y_2 \\ 75\,x_{13} + 80\,x_{23} + 85\,x_{33} + 90\,x_{43} + 99\,x_{53} &\geq 98\,y_3. \end{aligned}$$

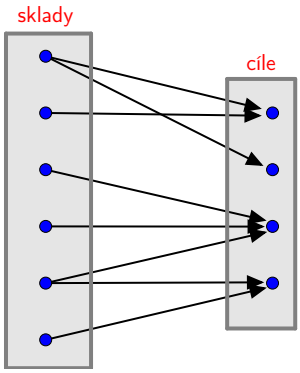
Poslední podmínkou je podmínka nezápornosti všech proměnných, které v úloze uvažujeme, tj.

$$x_{ij},\, y_j,\, z_i \geq 0 \quad \text{pro všechna} \quad i \in \{1,\dots,5\} \text{ a } j \in \{1,2,3\}.$$

Prodejce v této úloze zajímá primárně maximalizace zisku, nejsou zde tedy žádné podmínky na maximální rozpočet, který chce využít. Vzhledem k omezené dostupnosti všech druhů surových paliv však není úloha neomezená.

Náklady na přepravu zboží mezi výrobci a prodejci tvoří nemalou část nákladů. Proto jistě nepřekvapí, že i tato problematika se stala objektem zájmu matematiků. Existuje řada různých scénářů, které vedou na tzv. *dopravní problém*, v němž chceme minimalizovat přepravní náklady na dopravu z m-tice míst do n-tice cílů. Uvažme např. výrobce s m-ticí skladů $S_1,\dots,S_m$, ve kterých jsou dostupná množství $s_1,\dots,s_m$ jistého zboží. Na druhé straně máme n-tici cílů $C_1,\dots,C_n$, v nichž spotřebitelé vyžadují toto zboží v objemech $c_1,\dots,c_n$. Dodávky jsou možné pouze mezi skladem a cílem, přičemž zajížďky nejsou povoleny (tj. při jedné cestě nelze ob-

sloužit více cílů) a dodávky ani nelze kombinovat. Převoz jedné jednotky zboží ze skladu S_i do C_j stojí n_{ij} a předpokládejme, že náklady na přepravu jsou v podstatě lineární. Situace je znázorněna na Obrázku 2.4 vlevo.

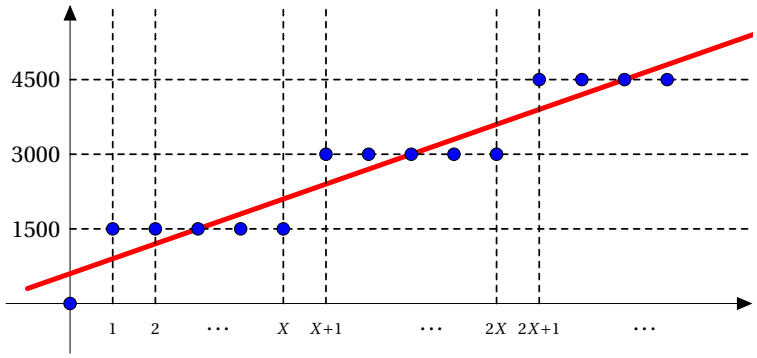


Obrázek 2.4: Sklady a cíle v dopravním problému.

Zatímco znalost stavu skladu s_i je zcela pochopitelná, požadavek znalosti c_j je možné ospravedlnit např. systémem objednávek. Ovšem linearita nákladů na přepravu je již velmi diskutabilní. Uvažme dva jednoduché příklady:

(i) dodavatel bazénů – v tomto případě jsou náklady zcela lineární, neboť např. přivézt 0 bazénů stojí 0 Kč, 1 bazén stojí 1500 Kč (stačí jedna dodávka), 2 bazény 3000 Kč (jsou potřeba dvě dodávky) atd. Tato funkce je sice definována pouze pro $n \in \mathbb{N} \cup \{0\}$, ale lze ji bez problémů dodefinovat na celý interval $[0, \infty)$.

(ii) dodavatel praček – zde již náklady nejsou lineární, neboť přivézt 0 praček stojí 0 Kč, 1 pračku 1500 Kč, 2 pračky 1500 Kč (obě se vejdou se do jedné dodávky), …, X praček stojí stále 1500 Kč, ale $X+1$ už stojí 3000 Kč (k tomu jsou potřeba dvě dodávky). Takže náklady jsou po částech konstantní funkcí, kterou však při vhodně zvolených jednotkách můžeme aproximovat lineární funkcí, aniž bychom se dopustili velkého zkreslení, viz Obrázek 2.5.



Obrázek 2.5: Náklady na rozvoz praček a jejich lineární aproximace (pomocí metody nejmenších čtverců).

Také budeme obvykle požadovat, aby úloha byla *vyrovnaná*, tj. $\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} c_j$, takže dokážeme zcela uspokojit kompletní poptávku a nic ve skladech nezůstane. S možnostmi řešení této úlohy se blíže seznámíme v Podkapitole 2.11, v níž se podíváme i na nevyrovnané problémy. Každopádně naším cílem je minimalizovat náklady na dopravu, což při označení x_{ij} pro množství odeslané z S_i do C_j vede na úlohu lineárního programování

$$\sum_{i=1}^{m}\sum_{j=1}^{n} n_{ij}\,x_{ij} \to \min$$

$$\sum_{j=1}^{n} x_{ij} = s_i \quad \text{pro všechna } i \in \{1,\dots,m\} \text{ (tj. odešleme vše ze skladů)}$$

$$\sum_{i=1}^{m} x_{ij} = c_j \quad \text{pro všechna } j \in \{1,\dots,n\} \text{ (tj. uspokojíme všechny poptávky)}$$

$$x_{ij} \geq 0 \quad \text{pro všechna } i \in \{1,\dots,m\} \text{ a } j \in \{1,\dots,n\}.$$

Máme celkem $m \times n$ proměnných a $m+n$ omezení (vyjma požadavku nezápornosti), z nichž jedno je nadbytečné.

Příklad 2.3.6

Uvažme problém se třemi sklady a čtyřmi cíli. Zásoby jsou postupně 30, 60 a 50 kusů, zatímco poptávky jsou 20, 40, 50 a 30 kusů. Náklady na dopravu jsou zapsány v matici

$$N = \begin{pmatrix} 5 & 6 & 4 & 5 \\ 3 & 4 & 3 & 2 \\ 2 & 7 & 6 & 8 \end{pmatrix}.$$

Řešení. Daný dopravní problém je vyrovnaný, takže odpovídající lineární program má podobu

$$
\begin{aligned}
5x_{11}+6x_{12}+4x_{13}+5x_{14}+3x_{21}+4x_{22}+3x_{23}+2x_{24}+2x_{31}+7x_{32}+6x_{33}+8x_{34} &\to \min\\
x_{11}+x_{12}+x_{13}+x_{14} &= 30\\
x_{21}+x_{22}+x_{23}+x_{24} &= 60\\
x_{31}+x_{32}+x_{33}+x_{34} &= 50\\
x_{11}+x_{21}+x_{31} &= 20\\
x_{12}+x_{22}+x_{32} &= 40\\
x_{13}+x_{23}+x_{33} &= 50\\
x_{14}+x_{24}+x_{34} &= 30\\
x_{ij}\geq 0 \quad &\text{pro všechna } i\in\{1,2,3\} \text{ a } j\in\{1,2,3,4\}.
\end{aligned}
$$

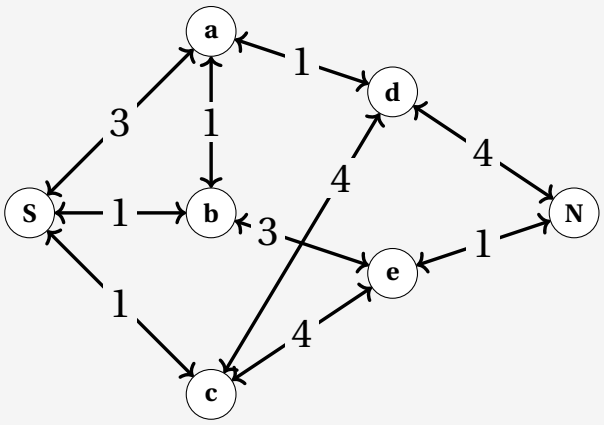
[Tato úloha nemá řešení.]

Ne vždy se však musí jednat o skutečný rozvoz zboží[74].

[74] Následující příklad je převzat z [374].

Příklad 2.3.7

Správce počítačové sítě přesvědčil svého zaměstnavatele ke koupi nového počítače s vylepšeným zvukovým systémem. Chtěl by převést svoji sbírku těch nejlepších skladeb Karla Gotta, Heleny Vondráčkové a Hany Zagorové ze starého počítače do nového pouze s využitím lokální sítě (nemá totiž žádný přenosný datový nosič). Síť je zobrazena na Obrázku 2.6 níže.



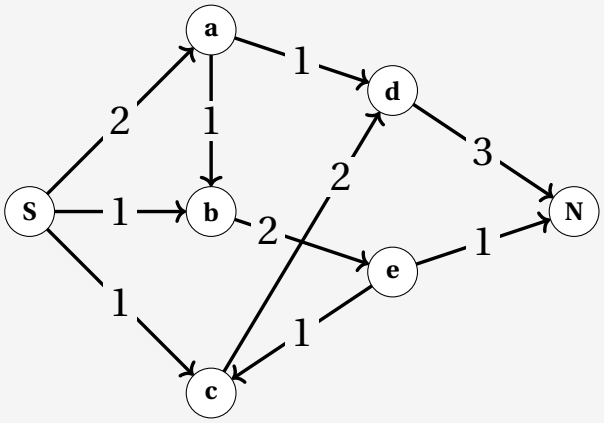
Obrázek 2.6: Lokální počítačová síť a přenosové rychlosti.

Jaká je maximální přenosová rychlost ze starého počítače „S“ do nového počítače „N“? Čísla u datových spojení udávají maximální možnou rychlost přenosu dat mezi oběma uzly (např. v Mbit/s). Navíc se omezme pouze na situaci, kdy každý spoj může přenášet data oběma směry, ale nikoli současně. Např. na spoji ab mohou jít data z a do b jakoukoli rychlostí mezi 0 a 1 Mbit/s, nebo z b do a jakoukoli rychlostí mezi 0 a 1 Mbit/s. Navíc uzly nejsou uzpůsobeny pro uchovávání většího množství dat, proto příchozí data musí okamžitě pokračovat dále. Je tudíž jasné, že není možné dosáhnout maximálního vytížení všech linek současně (viz např. uzel a). Proto potřebujeme najít vhodnou rychlost toku dat na každém spojení tak, aby celková rychlost byla maximální. Pro každý spoj v síti zavedeme jednu proměnnou, např. x_{be} značí přenosovou rychlost z b do e, přičemž x_{be} může nabývat i záporné hodnoty, což značí tok dat opačným směrem, tj. z e do b (a tudíž již není potřeba zavádět proměnnou x_{eb}). Máme tedy celkem 10 proměnných x_{Sa}, x_{Sb}, x_{Sc}, x_{ad}, x_{ab}, x_{be}, x_{cd}, x_{ce}, x_{dN} a x_{eN}, pro které chceme

vyřešit lineární program

$$x_{Sa} + x_{Sb} + x_{Sc} \to \max$$
$$-3 \le x_{Sa} \le 3 \quad \& \quad -1 \le x_{Sb} \le 1 \quad \& \quad -1 \le x_{Sc} \le 1 \quad \& \quad -1 \le x_{ad} \le 1$$
$$-1 \le x_{ab} \le 1 \quad \& \quad -3 \le x_{be} \le 3 \quad \& \quad -4 \le x_{cd} \le 4 \quad \& \quad -4 \le x_{ce} \le 4$$
$$-4 \le x_{dN} \le 4 \quad \& \quad -1 \le x_{eN} \le 1$$
$$x_{Sa} = x_{ab} + x_{ad} \quad \& \quad x_{Sb} + s_{ab} = x_{be} \quad \& \quad x_{ad} + x_{cd} = x_{dN}$$
$$x_{be} + x_{ce} = x_{eN} \quad \& \quad x_{Sc} = x_{cd} + x_{ce}.$$

Účelová funkce $x_{Sa} + x_{Sb} + x_{Sc}$ udává celkovou rychlost, s jakou jsou data odesílána ze starého počítače. Jelikož se data nemohou ukládat ani (doufejme) ztrácet, musí se stejnou rychlostí doputovat i do nového počítače. Prvních 10 omezení popisuje maximální rychlosti přenosu mezi uzly a zbývající omezení popisují nemožnost ukládání dat v uzlech. Optimální řešení je na Obrázku 2.7.



Obrázek 2.7: Lokální počítačová síť a optimální přenosové rychlosti.

Všimněme si, že z uzlu *e* data putují také „nazpět“ do *c*. Optimální hodnota je v tomto případě 4. V tomto případě je zřejmé, že přenosová rychlost nemůže být větší, protože celková kapacita všech spojů spojujících *S* a *a* se zbytkem sítě je 4. Toto je obecně platné tvrzení o tzv. *maximálním toku a minimálním řezu* z teorie grafů. Tento příklad ilustruje využití lineárního programování ve velmi jednoduché situaci. V praxi se jedná často o mnohem propletenější síť s mnoha zdroji a příjemci, např. elektrické sítě (přenos proudu), silniční a železniční sítě (tok dopravy), telefonní sítě, finance (převod peněz) atd. Také se to objevuje například při zpracování obrazu, což nemusí být na první pohled až tak zřejmé. Historicky byly takovéto úlohy studovány nejdříve pro vojenské účely (kvůli hledání efektivních způsobů narušení železniční sítě v sovětském bloku).

V přeneseném smyslu může jít i o úlohu, v níž chceme rozdělit několik činností nebo pracovníků v rámci nějakého týmu. Zde je opět potřeba mít na paměti, že bychom mohli najít řešení, které není prakticky realizovatelné (přiřadit 1/3 pracovníka do jednoho týmu a zbývající 2/3 do druhého týmu nemusí být vždy proveditelné). Avšak některé takovéto úlohy mají velmi speciální strukturu, která zaručí existenci vhodného (např. celočíselného) řešení, viz Kapitolu 3.

Příklad 2.3.8

(*Přiřazovací úloha s binárními proměnnými.*) Představme si čtyřčlennou rodinu, ve které jsou obě děti minimálně patnáctileté a otec ovládá nástroje matematického programování. Rodinu čeká jarní úklid celého domu, a ten je potřeba si pořádně naplánovat. Na úklidu se budou podílet všichni. Činnosti si rodina rozdělila do čtyř kategorií

(můžeme si je představit např. jako umytí oken, umytí podlah, úklid komory a přichystání jarní výzdoby). Každý z rodiny je jinak efektivní při vykonávání jednotlivých činností – efektivita se zde bude měřit časem, jinak předpokládáme, že všichni jsou schopni úkol splnit stejně dobře. Úkolem je přiřadit každému členovi rodiny právě jednu úklidovou činnost tak, aby rodina minimalizovala čas, který uklízením stráví. Čas, který jednotlivým členům zaberou jednotlivé úkony, je zapsán v Tabulce 2.3.

činnost / člen rodiny	1	2	3	4
1	110	55	90	65
2	50	75	100	80
3	50	110	120	95
4	70	150	95	130

Tabulka 2.3: Úklidový čas (v minutách) pro členy domácnosti

Řešení. Rodina, která chce minimalizovat celkový čas úklidu, vlastně stojí před problémem, jenž nyní popíšeme matematicky. Označme c_{ij} čas, který i-tému členu domácnosti trvá provést j-tou činnost. Např. proměnná c_{23} vyjadřuje, kolik času trvá 2. členu rodiny úklidová činnost číslo 3, což z tabulky vidíme $c_{23} = 100$. Dále si označme x_{ij} proměnnou, která vyjadřuje, zda i-tý člověk vykonává j-tou činnost či ne. Proměnné x_{ij} pro všechna $i, j \in \{1,2,3,4\}$ vyjadřují přiřazení člověka a činnosti a nabývají pouze hodnot $\{0,1\}$. Z tohoto důvodu se podobné úlohy nazývají též *úlohami s binárními proměnnými*. Funkce, kterou chce rodina minimalizovat, je ve tvaru

$$\begin{aligned}110\,x_{11} + 50\,x_{21} + 50\,x_{31} + 70\,x_{41} + 55\,x_{12} + 75\,x_{22} + 110\,x_{32} + 150\,x_{42} + \\ + 90\,x_{13} + 100\,x_{23} + 120\,x_{33} + 95\,x_{43} + 65\,x_{14} + 80\,x_{24} + 95\,x_{34} + 130\,x_{44}.\end{aligned}$$

Podle zadání musí každý člen vykonávat právě jednu práci, čehož docílíme pomocí podmínek zaručujících, že součet proměnných, které udávají přiřazení člověka k určité činnosti, je roven jedné, tj.

$$\begin{aligned}x_{11} + x_{12} + x_{13} + x_{14} &= 1,\\ x_{21} + x_{22} + x_{23} + x_{24} &= 1,\\ x_{31} + x_{32} + x_{33} + x_{34} &= 1,\\ x_{41} + x_{42} + x_{43} + x_{44} &= 1.\end{aligned}$$

Podobná omezení naformulujeme i pro jednotlivé činnosti, protože každá činnost musí být vykonána právě jednou. To zajistíme s využitím další čtveřice podmínek

$$\begin{aligned}x_{11} + x_{21} + x_{31} + x_{41} &= 1,\\ x_{12} + x_{22} + x_{32} + x_{42} &= 1,\\ x_{13} + x_{23} + x_{33} + x_{43} &= 1,\\ x_{14} + x_{24} + x_{34} + x_{44} &= 1.\end{aligned}$$

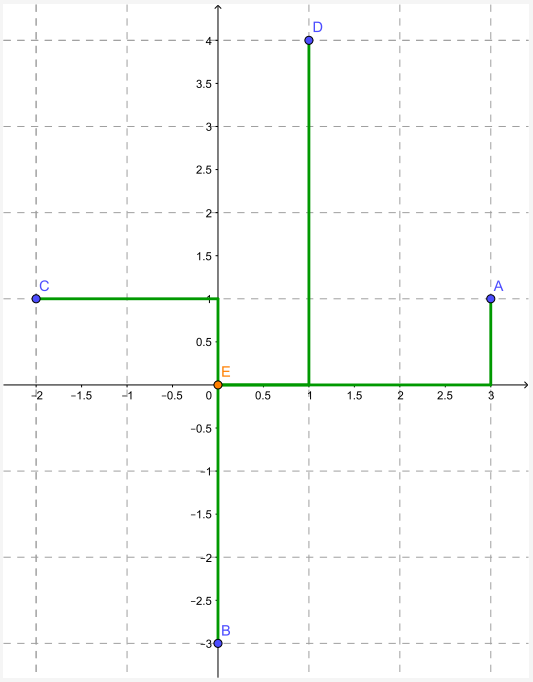
Nesmíme ještě zapomínat na podmínky *bivalence*, tj. $x_{ij} \in \{0,1\}$. Úloha tedy není, na rozdíl od předchozích úloh, tvořena pouze lineární funkcí omezenou afinními funkcemi. Avšak pokud podmínky bivalence nahradíme podmínkami $0 \le x_{ij} \le 1$ pro všechny proměnné, můžeme úlohu řešit jako problém lineárního programování.

Nicméně díky speciální struktuře této úlohy budou výsledné optimální hodnoty proměnných pouze binární, aniž bychom to explicitně vyžadovali, tj. nalezené řešení této přidružené úlohy bude i řešením daného problému.

Na závěr si ukážeme ještě dva příklady, které sice na první pohled nevypadají jako úlohy lineárního programování, avšak jednoduchou úpravou dokážeme i je převést do takové podoby. V obou úlohách totiž v účelové funkci vystupuje absolutní hodnota, kterou je potřeba vhodně nahradit lineární funkcí. Všimněme si také, že v prvním příkladě dokonce původní úloha ani neobsahuje žádná omezení.

Příklad 2.3.9

Uvažme továrnu, ve které jsou umístěny čtyři stroje. Jejich souřadnice jsou $A = [3,1]$, $B = [0,-3]$, $C = [-2,1]$ a $D = [1,4]$. Vedení továrny zakoupilo nový stroj, který chce umístit tak, aby byl součet vzdáleností nového stroje od všech ostatních strojů minimální. Vzdálenost budeme měřit pomocí součtové (či Manhattanské) metriky, ve které je vzdálenost bodu $[x_1, x_2]$ od bodu $[1,3]$ rovna $|x_1 - 1| + |x_2 - 3|$. Pro bod $E = [0,0]$ jsou vzdálenosti od jednotlivých strojů znázorněny na Obrázku 2.8 níže. Naformulujme problém, kterému vedení čelí, jako úlohu lineárního programování.



Obrázek 2.8: Situační obrázek továrny.

Řešení. Označme jako x_1, x_2 souřadnice nového stroje. Ze zadání víme, že chceme minimalizovat součet vzdáleností tohoto stroje od ostatních strojů. Funkce, kterou chceme minimalizovat, je tedy

$$f(x_1, x_2) = |x_1 - 3| + |x_2 - 1| + |x_1| + |x_2 + 3| + |x_1 + 2| + |x_2 - 1| + |x_1 - 1| + |x_2 - 4|.$$

Nyní je potřeba převést tuto funkci na funkci lineární. Označme si pro přehlednost y_1, y_3, y_5 a y_7 horizontální vzdálenosti nového stroje od původních strojů. Dále označme y_2, y_4, y_6 a y_8 příslušné vertikální vzdálenosti. Funkci, kterou chceme minimalizovat, můžeme ekvivalentně vyjádřit jako

$$f(y_1, \ldots, y_8) = \sum_{i=1}^{8} y_i.$$

Všechna y_i vyjadřují vzdálenosti, a musí být proto nezáporné, tj. musíme doplnit podmínky $y_i \geq 0$ pro všechna $i \in \{1,\ldots,8\}$. Nyní už chybí jen propojení původních proměnných x_i s novými proměnnými y_i, které si nyní ukážeme pouze pro proměnnou y_1, neboť pro ostatní proměnné je to analogické. Jelikož pro absolutní hodnotu platí $|x| = \max\{x,-x\}$, máme pro y_1 rovnost

$$y_1 = \begin{cases} x_1 - 3, & \text{pro } x_1 \geq 3, \\ 3 - x_1, & \text{pro } x_1 < 3. \end{cases}$$

Pokud tedy k úloze přidáme podmínky $y_1 \geq x_1 - 3$ a $y_1 \geq 3 - x_1$, zajistíme, že

$$y_1 = \max\{x_1 - 3, 3 - x_1\} = |x_1 - 3|.$$

Podobně pro y_2 musíme přidat podmínky $y_2 \geq x_2 - 1$ a $y_2 \geq 1 - x_2$ a pro ostatní proměnné postupujeme obdobně. Tudíž úloha, kterou musí při umisťování nového stroje vyřešit vedoucí továrny, má podobu[75]

$$\sum_{i=1}^{8} y_i \to \min$$
$$y_1 \geq x_1 - 3 \quad \& \quad y_1 \geq 3 - x_1 \quad \& \quad y_2 \geq x_2 - 1 \quad \& \quad y_2 \geq 1 - x_2$$
$$y_3 \geq x_1 \quad \& \quad y_3 \geq -x_1 \quad \& \quad y_4 \geq x_2 + 3 \quad \& \quad y_4 \geq -3 - x_2$$
$$y_5 \geq x_1 + 2 \quad \& \quad y_5 \geq -2 - x_1 \quad \& \quad y_6 \geq x_2 - 1 \quad \& \quad y_6 \geq 1 - x_2$$
$$y_7 \geq x_1 - 1 \quad \& \quad y_7 \geq 1 - x_1 \quad \& \quad y_8 \geq x_2 - 4 \quad \& \quad y_8 \geq 4 - x_2$$
$$y_1 \geq 0 \quad \& \quad y_2 \geq 0 \quad \& \quad y_3 \geq 0 \quad \& \quad y_4 \geq 0$$
$$y_5 \geq 0 \quad \& \quad y_6 \geq 0 \quad \& \quad y_7 \geq 0 \quad \& \quad y_8 \geq 0.$$

[Řešením je $y_1 = 3$, $y_2 = 0$, $y_3 = 0$, $y_4 = 4$, $y_5 = 2$, $y_6 = 0$, $y_7 = 1$, $y_8 = 3$, takže $x_1 = 0$ a $x_2 = 1$, což dává nejmenší vzdálenost 13.]

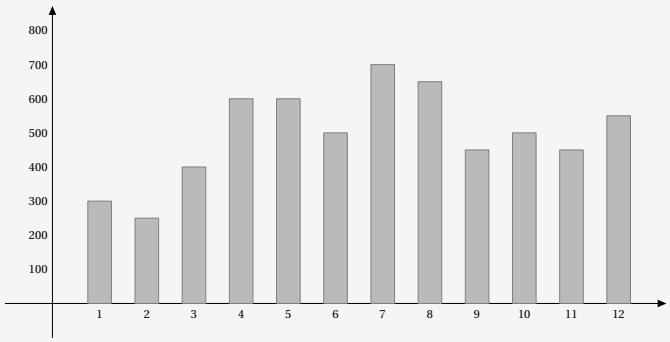
[75] Už výše jsme zmínili nutnost nezápornosti proměnných $y_1,\ldots,y_8$. Avšak při pohledu na další potřebná omezení si čtenář může (snadno?) rozmyslet, že to jinak ani nepůjde, a tudíž nemusí být tento požadavek explicitní součástí formulace celého problému.

Druhý příklad[76] zase obsahuje jídlo, což by vzhledem k jeho důležitosti nemělo být příliš překvapující. Avšak tentokrát jde pouze o obchodní artikl a nebudeme zde již vytvářet žádný jídelníček.

[76] Tento příklad je opět převzat z [374].

Příklad 2.3.10

Výrobce zmrzliny *Rampouchovy závody, s.r.o*, potřebuje sestavit výrobní plán pro příští rok. Vzhledem k historickým zkušenostem, rozsáhlým průzkumům trhu, odpovědím na otázky VM, objednávkám z menzu MU a pohybu andulek nad Brnem v zimních měsících známe přesnou předpověď (sic!) měsíčních prodejů zmrzliny, která je zaznamenána na Obrázku 2.9. A Rampouchovy závody potřebují pomoc se sestavením výrobního plánu tak, aby pokryly tuto poptávku.



Obrázek 2.9: Měsíční předpovědi prodejů zmrzliny (v tunách; 1=leden, 2=únor, ...)

Nejjednodušší by bylo vyrábět „na čas“, tj. zmrzlinu potřebnou v i-tém měsíci vyrobit v i-tém měsíci. To by ovšem znamenalo každý měsíc výrazně měnit objem výroby, což by asi mělo zásadní dopady na výrobní náklady: musely by se najímat brigádníci pro pokrytí zvýšené produkce, a pak zase propouštět, museli by se upravovat stroje atd. Proto je lepší rozložit výrobu v průběhu celého roku: v měsících s nižší poptávkou využít zbývající výrobní kapacitu k vytváření zásob pro měsíce s vyšší poptávkou. Mohli bychom vyrábět zcela rovnoměrně během celého roku. Ovšem taková varianta nemusí být vůbec přípustná, pokud chceme na konci roku skončit s nulovým přebytkem. Dokonce ani v situaci, kde je tato možnost přípustná, nemusí být optimální, neboť skladování zmrzliny také není zcela bez nákladů. Zdá se, že optimální varianta bude ukryta někde mezi těmito dvěma extrémními výrobními plány. Chceme tedy dosáhnout minima nákladů, které tvoří náklady na skladování a změnu výroby.

Nyní se pokusme tento problém popsat matematicky. Pro $i \in \{1,\dots,12\}$ označme poptávku v i-tém měsíci jako $p_i \geq 0$ (opět v tunách) a dále zaveďme nezápornou proměnnou x_i pro produkci v i-tém měsíci a s_i pro celkový stav skladových zásob na konci i-tého měsíce. Abychom uspokojili poptávku v i-tém měsíci, využijeme k tomu výrobu v průběhu i-tého měsíce zásoby z $(i-1)$-ního měsíce, tj. musí platit

$$x_i + s_{i-1} \geq p_i, \quad i \in \{1,\dots,12\},$$

přičemž předpokládáme, že na začátku prvního měsíce je sklad prázdný (jinak bychom uvážili započetli stav skladu na konci předchozího roku), tj. $s_0 := 0$. Současně také chceme $s_{12} := 0$, jelikož pan Rampouch již nehodlá dále pokračovat v podnikání ve zmrzlinovém průmyslu. Hodnota $x_i + s_{i-1} - p_i$ přesně určuje přebytek po uspokojení poptávky v i-tém měsíci, takže

$$x_i + s_{i-1} - p_i = s_i, \quad i \in \{1,\dots,12\}.$$

Mezi všemi nezápornými řešeními těchto nerovnic a rovnic hledáme ta, která minimalizují výrobní náklady. Předpokládejme, že změna úrovně výroby o 1 tunu mezi $(i-1)$-ním a i-tým měsícem stojí 50 € a skladování 1 tuny stojí 20 € měsíčně. Pak celkové náklady lze vyjádřit jako

$$50 \sum_{i=1}^{12} |x_i - x_{i-1}| + 20 \sum_{i=1}^{12} s_i,$$

kde klademe $x_0 := 0$ (avšak mohli bychom snadno započítat i konec minulého roku).

Jenže tato funkce není lineární. Naštěstí tu však je jednoduchý „trik“, s jehož pomocí lze funkci linearizovat za cenu zavedení dalších proměnných. Změna produkce znamená buď její navýšení nebo snížení. Zaveďme proto nezápornou proměnnou y_i pro navýšení produkce mezi $(i-1)$-ním a i-tým měsícem a nezápornou proměnnou pro pokles výroby. Potom

$$x_i - x_{i-1} = y_i - z_i \quad \text{a navíc} \quad |x_i - x_{i-1}| = y_i + z_i.$$

Vskutku, je-li totiž trojice $[s^*, y^*, z^*]$ optimálním řešením, pak nutně $y_i^* z_i^* = 0$ pro každé $i \in \{1,\dots,12\}$. V opačném případě bychom mohli obě hodnoty snížit o stejné číslo a dosáhnout tak lepšího řešení (zvýšení výroby o 20 a snížení o 10 je mnohem méně výhodné než zvýšení o 10 a nulové snížení, jak je patrné z nové účelové funkce níže), tj. $y_i^* + z_i^*$ je skutečně rovno změně výroby mezi $(i-1)$-ním a i-tým měsícem. Takže výrobní plán minimalizující řešení je řešením úlohy

$$50 \sum_{i=1}^{12} y_i + 50 \sum_{i=1}^{12} z_i + 12 \sum_{i=1}^{12} s_i \to \min$$

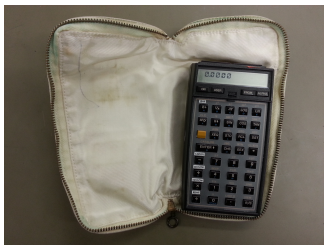
$$x_i + s_{i-1} - s_i = p_i \quad \& \quad x_i + x_{i-1} = y_i - z_i \quad \text{pro } i \in \{1,\dots,12\}$$

$$x_0 = 0 \quad \& \quad s_0 = 0 \quad \& \quad s_{12} = 0 \quad \& \quad x_i, s_i, y_i, z_i \geq 0 \quad \text{pro } i \in \{1,\dots,12\}\,.$$

[Řešením této úlohy je $x_1 = x_2 = 300$, $x_3 = 350$, $x_4 = 600$, $x_5 = 600$, $x_6 = x_7 = x_8 = 616\frac{2}{3}$, $x_9 = x_{10} = x_{11} = x_{12} = 487{,}5$]

S pomocí lineárního programování lze dokonce řešit i některé problémy z oblasti optimálního řízení. Ačkoli Dantzigův simplexový algoritmus je velmi účinným nástrojem pro takové úlohy, využívat jej pouze s pomocí tužky a papíru může být velmi zdlouhavé a rozhodně je zcela neefektivní. Proto rozmach této oblasti (a de facto celé oblasti matematického programování) šel ruku v ruce s rozvojem výpočetní techniky a její dostupností i pro jiné než armádní či vědecké organizace. Dalším zlomovým okamžikem proto byl příchod programovatelných kalkulátorů jako byl např. HP-41C v roce 1979 od společnosti Hewlett Packard, viz Obrázek 2.10. Tento kalkulátor byl nedílnou součástí prvních devíti raketoplánových misí NASA a je stále k vidění v *Národní muzeum letectví a kosmonautiky*[77] ve Washingtonu, viz Obrázek 2.11. Zde sloužily k obyčejným výpočtům, které však byly zcela nezbytné jako např. výpočet změny těžiště v důsledku úbytku paliva. Proto snad nepřekvapí, že v 80. letech bylo na těchto „strojích“ i mezi běžnými uživateli velmi populární řešení úlohy s *přistáním rakety na Měsíci*[78], čehož v marketingu využívala i samotná společnost HP, viz reklamu na Obrázku 2.12. V této úloze máme lunární modul s omezeným množstvím paliva, který klesá v kolmém směru na povrch Měsíce a je ovlivňován pouze jeho gravitací. V daných časových okamžicích můžeme zažehnout jeho motory tak, aby se snížila rychlost jeho klesání (nebo aby dokonce začal stoupat). Naším cílem je přistání na povrchu Měsíce s (téměř) nulovou rychlostí před spotřebováním veškerého paliva. Pokud spotřebujeme veškeré palivo příliš brzy, začne modul klesat volným pádem a dojde k jeho rozbití. Na druhou stranu, pokud budeme čekat se zažehnutím motorů příliš dlouho nebo budou zážehy příliš malé, nabere modul při klesání velkou rychlost a pak jej nedokážeme dostatečně zbrzdit (jsme totiž omezeni také výkonem motorů). Sestavení takového modelu již ponecháme na čtenáři, avšak zdůrazněme, že je potřeba diskretizovat čas. To by ale nemělo být z praktického hlediska na překážku, pokud budou časové okamžiky „dostatečně“ malé.

Obrázek 2.10: Model HP-41CX z roku 1983.



Obrázek 2.11: Model HP-41CV z roku 1983, který NASA využívala v simulátoru letů raketoplánem. Tento model je vybaven i suchým zipem, aby držel v ochranném obalu.

[77] National Air and Space Museum

[78] Angl. *Moon Rocket Landing*.

Diskrétní i spojitá verze simulace přistání rakety na Měsíci jsou k dispozici pro vyzkoušení na autorově webu:
https://ln.math.muni.cz/pz/opt/MRL1
https://ln.math.muni.cz/pz/opt/MRL2

Obrázek 2.12: Reklama na model HP-41.

2.4 Základem všeho je slušný náčrtek[83]

[83] Toto je věta, kterou autor slýchal snad v každé hodině matematiky na druhém stupni základní školy. Jakkoli se to jemu a jeho spolužačkám/spolužákům v té době zdálo pošetilé a přehnané, společně napříč celým textem uvidíme, že v této oblasti matematiky to platí dvojnásob.

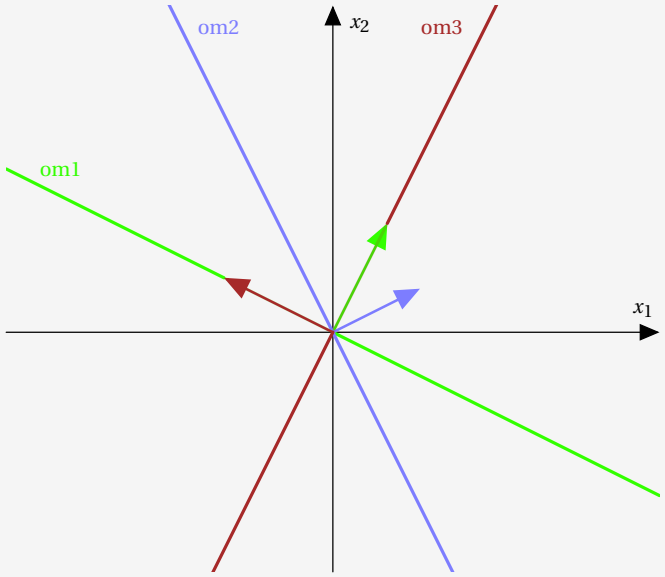
Úlohy v $\mathbb{R}^2$ (a s mnohem větší mírou úsilí a představivosti i v $\mathbb{R}^3$) můžeme snadno vyřešit s pomocí jednoduchého obrázku. V takovém případě jsou obvykle jednotlivá omezení ve tvaru nerovností, neboť případná rovnost by úlohu výrazně zjednodušila. Tyto podmínky vymezují poloroviny s hraničními přímkami s normálovým vektorem a_i a vrstevnice účelové funkce $c_1x_1 + c_2x_2$ odpovídají rovnoběžným přímkám s normálovým vektorem c. Řešení úlohy (2.2.1) pak znamená nalézt vrstevnici na nejnižší možné úrovni, která má ještě neprázdný průnik s množinou X.

Příklad 2.4.1

Uvažme úlohu

$$-x_1 - x_2 \to \min$$
$$x_1 + 2x_2 \le 4 \quad \& \quad 4x_1 + 2x_2 \le 12 \quad \& \quad -2x_1 + x_2 \le 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Přímky a jejich normálové vektory určující první tři omezení jsou zobrazeny na Obrázku 2.13, zatímco pro poslední dvě podmínky dostaneme souřadné osy.



Obrázek 2.13: Přímky určené levými stranami jednotlivých omezení (tj. om1: $x_1 + 2x_2 = 0$, om2: $4x_1 + 2x_2 = 0$, om3: $-2x_1 + x_2 = 0$) a jejich normálové vektory.

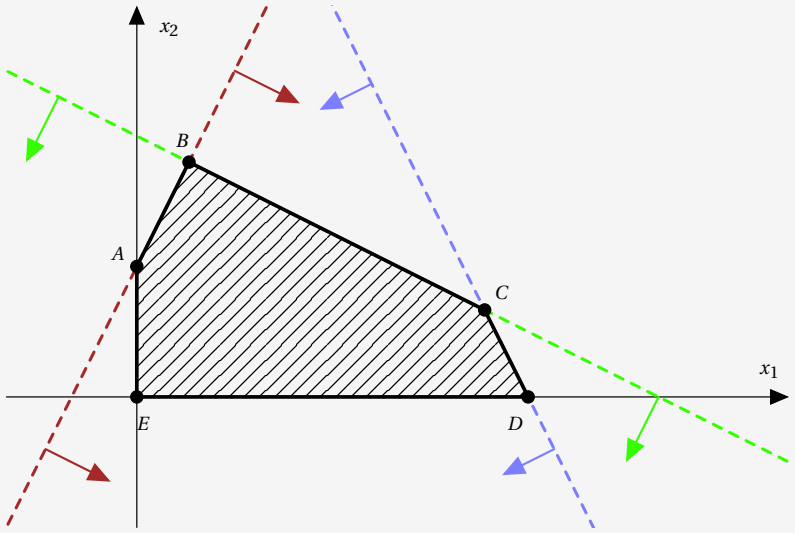
Hodnoty pravých stran pak určují posun těchto přímek (v případě kladného čísla se přímka posune ve směru normálového vektoru[84]. Tím dostaneme hraniční přímky rozdělující $\mathbb{R}^2$ vždy na dvě části a potřebujeme určit, která z nich vyhovuje dané nerovnosti. K tomu stačí vybrat libovolný bod neležící na hraniční přímce[85] a otestovat, zda je pro něj nerovnost splněna či nikoliv. Tyto vyhovující poloroviny jsou na Obrázku 2.14 vyznačeny šipkami. Nesmíme také zapomenout na požadavek nezápornosti obou proměnných, čímž dostaneme přípustnou množinu X jako průnik jednotlivých polorovin a prvního kvadrantu. Výsledkem je pak pětiúhelník z Obrázku 2.14.

Vrstevnice účelové funkce jsou rovnoběžné přímky $-x_1 - x_2 = \alpha$ určené týmž normálovým vektorem, které se pro $\alpha \to \infty$ „pohybují“ jihozápadním směrem (odpovídajícímu právě onomu normálovému vektoru)[86], viz Obrázek 2.15 na další straně. Z tohoto obrázku pak snadno vidíme, že řešením dané úlohy je bod C, který odpovídá průsečíku

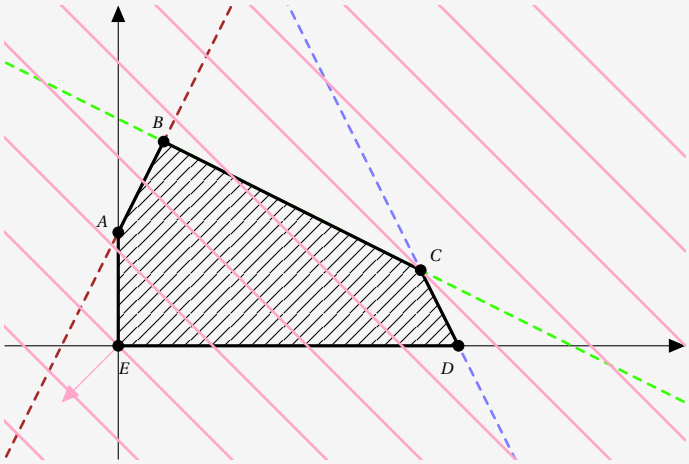
[84] Samozřejmě mnohem praktičtější je spočítat přímo průsečíky těchto přímek se souřadnými osami, čímž dostaneme pro první omezení body $[4,0]$ a $[0,2]$, pro druhé omezení $[3,0]$ a $[0,6]$, pro třetí omezení $[-1/2,0]$ a $[0,1]$.

[85] Je-li to možné, pak je nejvýhodnější vzít bod $[0,0]$.

přímek určující první a druhé omezení, tj. vyhovuje dvojici rovnic $x_1 + 2x_2 = 4$ a $4x_1 + 2x_2 = 12$. Snadným výpočtem zjistíme, že $C = [8/3, 2/3]$ a odpovídající hodnota úlohy je $f^* = -10/3$. Jak by se změnilo řešení a Obrázek 2.15, kdyby účelová funkce byla $x_1 + x_2$?



Obrázek 2.14: Přípustná množina.



Obrázek 2.15: Kompletní vizualizace dané úlohy se zvýrazněním normálového vektoru vrstevnic.

Při méně přesném náčrtku dané úlohy (což je samozřejmě mnohem častější případ) se může stát, že není zcela zřejmé, ve kterém bodě nastává řešení. Obvykle dokážeme identifikovat hranu, na které bude ležet řešení, ale rozhodnutí o konkrétním budě už může být složitější. V takovém případě je možné si pomoci směrnicí vrstevnic, která má vždy hodnotu $-c_1/c_2$, takže odchylka těchto přímek od kladné části osy x_1 je

$$\varphi = \begin{cases} \operatorname{arctg}(-c_1/c_2), & -c_1/c_2 > 0, \\ \pi + \operatorname{arctg}(-c_1/c_2), & -c_1/c_2 < 0. \end{cases}$$

V prvním případě pak bude $\varphi \in (0, \pi/2)$, zatímco ve druhém dostaneme $\varphi \in (\pi/2, \pi)$. Kdyby $c_1 = 0$, pak jsou vrstevnice rovnoběžné s osou x_1, zatímco pro $c_2 = 0$ budou rovnoběžné s osou x_2. Podobným způsobem určíme i sklon hraniční přímky vymezující množinu X a jejich porovnáním získáme hledaný bod, přičemž zde je dobré mít na paměti, že větší hodnota podílu $|c_1/c_2|$ znamená větší sklon neboli úhel φ blíže $\pi/2$. Ovšem nelze *a priori* říci, zda větší sklon je „lepší“ nebo „horší“. Vždy totiž záleží na vzájemné poloze příslušné hrany a vrstevnic a také na směru jejich pohybu. Toto si můžeme ilustrovat na následujících dvou příkladech.

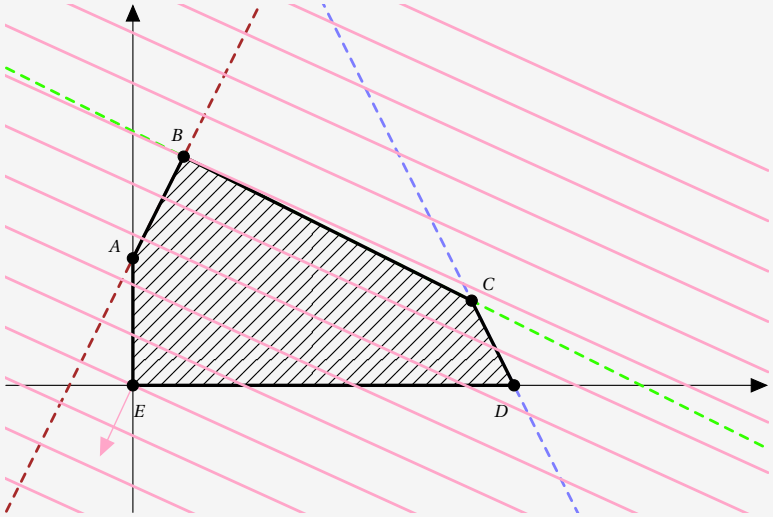
[86] Díky rovnoběžnosti nám stačí zakreslit pouze jednu vrstevnici a ostatní pak pouze „zkopírovat“. Např. pro volbu $\alpha = c_1\, c_2$ (součin koeficientů účelové funkce) bude odpovídající vrstevnice vždy procházet body $[0, c_1]$ a $[c_2, 0]$.

Příklad 2.4.2

Uvažme úlohu

$$-5x_1 - 11x_2 \to \min$$
$$x_1 + 2x_2 \le 4 \quad \& \quad 4x_1 + 2x_2 \le 12 \quad \& \quad -2x_1 + x_2 \le 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Jelikož přípustná množina je stejná jako v předchozím příkladě, můžeme využít Obrázek 2.14 a zakreslit do něj vrstevnice odpovídající současné účelové funkci, viz Obrázek 2.16. Odtud je vidět, že řešením je bod B. Ovšem při běžném náčrtku to až tak zřejmé být nemusí a můžeme být na pochybách, zda řešením je bod B nebo C (tyto jediné dva potenciální kandidáty bychom měli poznat i při velmi hrubém náčrtku). Protože pro vrstevnice máme $c_1 = -5$ a $c_2 = -11$, bude směrnice těchto přímek rovna $\operatorname{arctg}(-5/11) \approx 155,6°$. Hrana přípustné množiny spojující body B a C odpovídá přímce z prvního omezení, takže její směrnice je $\operatorname{arctg}(-1/2) \approx 153,3°$. To znamená, že sklon této hrany je o něco málo strmější než vrstevnic (však také $|-5/11| < |-1/2|$), což potvrzuje náš původní závěr ohledně řešení v bodě B.



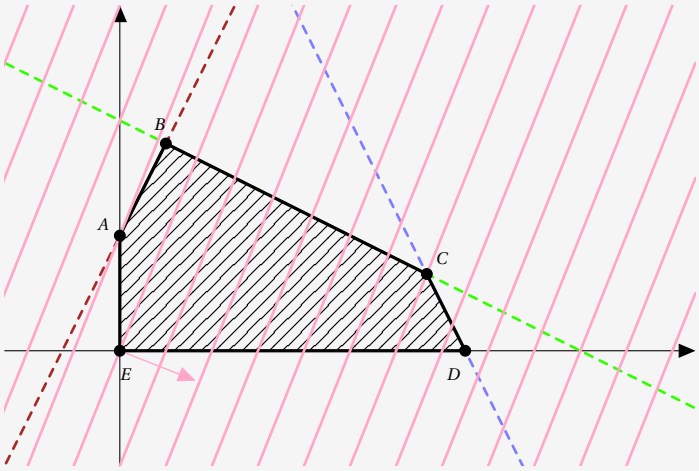
Obrázek 2.16: Kompletní vizualizace dané úlohy se zvýrazněním normálového vektoru vrstevnic.

Příklad 2.4.3

Uvažme ještě také úlohu

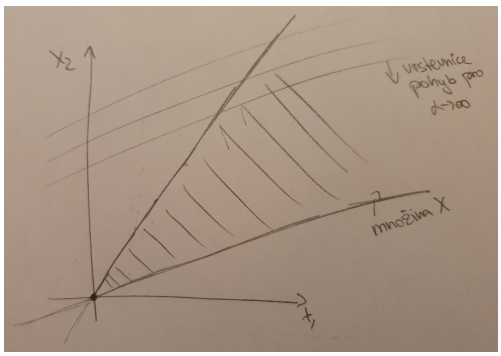
$$5x_1 - 2x_2 \to \min$$
$$x_1 + 2x_2 \le 4 \quad \& \quad 4x_1 + 2x_2 \le 12 \quad \& \quad -2x_1 + x_2 \le 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Opět máme tutéž přípustnou množinu, ale vrstevnice již mají jiný směr, jak můžeme vidět na Obrázku 2.17 níže. V tomto případě je řešením bod A, což si můžeme potvrdit i pomocí výpočtu směrnic odpovídajících přímek. Bez ohledu na preciznost náčrtku by mělo být zřejmé, že řešením je buď bod A nebo B. Tato hrana odpovídá třetímu omezení určeného přímkou se směrnicí $\operatorname{arctg}(2) \approx 63,4°$. Směrnice vrstevnic je $\operatorname{arctg}(5/2) \approx 68,2°$, takže jejich sklon je o trochu strmější než pro hranu přípustné množiny (však také $5/2 > 2$). To potvrzuje, že řešením této úlohy je skutečně bod A.

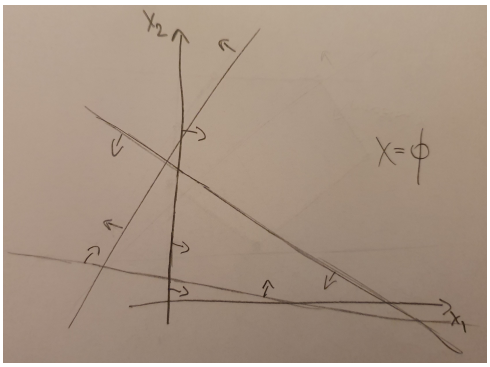


Obrázek 2.17: Kompletní vizualizace dané úlohy se zvýrazněním normálového vektoru vrstevnic.

Všechny předchozí úlohy měly jediná řešení. To samozřejmě neplatí obecně a úplně stejně se může stát, že řešení vůbec neexistuje – a to buď kvůli neohraničenosti účelové funkce zdola nebo kvůli tomu, že $X = \emptyset$, viz Obrázky 2.18.a–2.18.b.

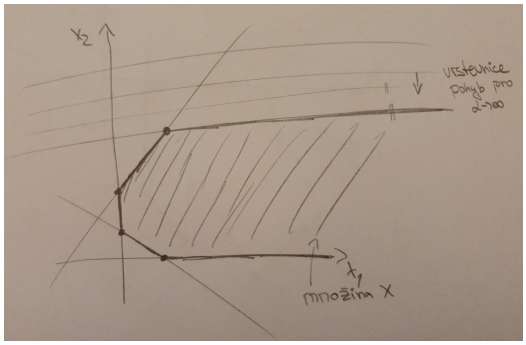


Obrázek 2.18.a: Neohraničená úloha.

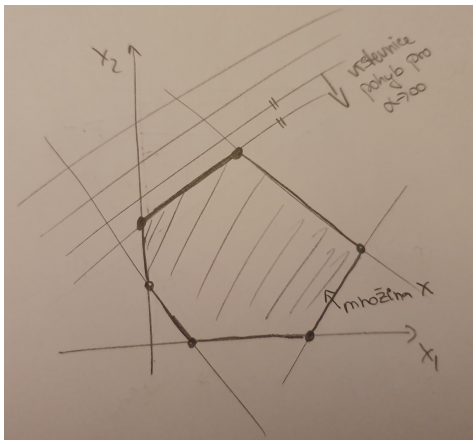


Obrázek 2.18.b: Nepřípustná úloha.

Úloha také může mít nekonečně mnoho řešení. I zde ale mohou nastat dvě principiálně odlišné situace – buď je množina řešení neohraničená (přímka nebo polopřímka) nebo je ohraničená (úsečka), viz Obrázky 2.19.a–2.19.b. K tomuto je ale každopádně nutné, aby normálový vektor vrstevnic a hrany přípustné množiny (tj. jedné z vymezujících přímek) byly lineárně závislé (tj. jeden nějakým nenulovým násobkem druhého).

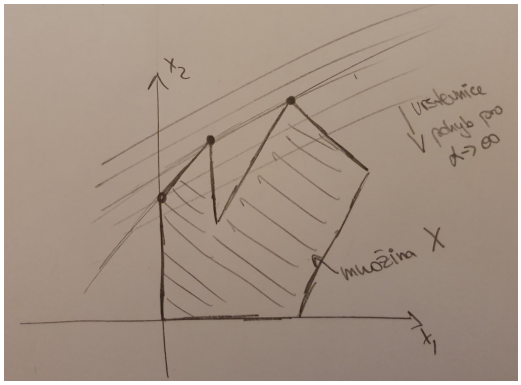


Obrázek 2.19.a: Nekonečně mnoho řešení (polopřímka).



Obrázek 2.19.b: Nekonečně mnoho řešení (úsečka).

V obou předchozích případech byla množina řešení dokonce nespočetná. V této souvislosti se nabízí zcela přirozená otázka: je možné, aby úloha lineárního programování měla pouze konečný počet řešení větší než jedna? Viz např. Obrázek 2.20.



Obrázek 2.20: Úloha lineárního programování může mít právě 2 různá řešení?

Ať už řešení bylo jediné nebo jich bylo nekonečně mnoho, jedna velmi důležitá vlastnost všechny tyto případy spojovala. Vidíte ji na obrázcích? Ano, řešení vždy nastalo v nějakém vrcholu mnohoúhelníku (a bylo-li to dokonce ve dvou z nich, pak řešením byly i body ležící na celé hraně spojující tyto body). Náhoda? Mohlo by to být v $\mathbb{R}^2$ jinak? A v $\mathbb{R}^n$? To zjistíme v další části. Teď se ale podíváme, co je přípustná množina X z (2.2.1) v obecném případě. Tato množina je určena pomocí soustavy rovností a neostrých nerovností daných afinními funkcemi a podmínkami na znaménko některých proměnných. Co to znamená? Každá z rovností v soustavě $A_R x = b_R$, tj. $\langle a_i, x\rangle = b_i$ pro $i = 1,\dots,k$, určuje nadrovinu v $\mathbb{R}^n$ s normálovým vektorem a_i, takže takové soustavě vyhovují body ležící v průniku všech nadrovin. Pokud žádné dvě z nich nejsou rovnoběžné[87] ani „žádná z nich není lineární kombinací ostatních“[88] (viz raději také začátek dalšího odstavce), pak s každou rovnostní podmínkou dojde ke snížení dimenze původního prostoru $\mathbb{R}^n$ o jeden stupeň, takže dané k-tici rovností vyhovuje $n-k$ dimenzionální nadrovina (afinní množina). Oproti tomu každá z daných nerovnostních podmínek v soustavě $A_N x \le b_N$, tj. $\langle a_i, x\rangle \le b_i$ pro $i = k+1,\dots,m$, udává tzv. *poloprostor* v $\mathbb{R}^n$, což je jedna ze dvou částí $\mathbb{R}^n$ ohraničených nadrovinou $\langle a_i, x\rangle = b_i$ s normálovým vektorem a_i. Průnik těchto poloprostorů samozřejmě může být neohraničený (přímka, polopřímka, pás, kužel apod.) nebo ohraničený, v kterémžto případě půjde o nějaký (mnohdy nepravidelný) mnohostěn. Konečně požadavek nezápornosti některých proměnných znamená, že se musíme omezit pouze na část $\mathbb{R}^n$ ohraničenou odpovídajícími osami. Množina X je pak průnikem těchto tří množin, tj. typicky ta část $n-k$ dimenzionální nadroviny ležící v nějakém (ohraničeném/neohraničeném) mnohostěnu. Vrstevnice účelové funkce jsou pro každé $\alpha \in \mathbb{R}$ neprázdné množiny $\{x \in \mathbb{R}^n \mid \langle c, x\rangle = \alpha\}$. Lépe řečeno, jedná se o rovnoběžné nadroviny s normálovým vektorem c, které se pro $\alpha \to \infty$ „pohybují“ ve směru vektoru c. Tyto vrstevnice vyplní celé $\mathbb{R}^n$ a řešení úlohy (2.2.1) opět znamená nalezení bodu (příp. bodů) z X, který leží na vrstevnici s nejmenší možnou hodnotou α, tj. hledáme vrstevnici na nejnižší možné úrovni, která má ještě neprázdný průnik s množinou X. Takový bod může být jediný nebo jich může být nekonečně (nespočetně) mnoho a nebo nemusí existovat žádný.

Tento náš intuitivní pohled na množinu X můžeme popsat také *konvexním* jazykem, který hraje zcela klíčovou roli v celé oblasti matematického programování a my se jej podrobně naučíme později v Kapitole 6. Jelikož každé omezení je dáno pomocí afinní funkce (a tudíž je určeno nějakou nadrovinou), musí být odpovídající množina konvexní[89]. Díky tomu je množina X (jakožto průnik konvexních množin) nutně také konvexní – ovšem ne jen tak ledajaká. Každou rovnost totiž můžeme nahradit dvojicí nerovností a celkový počet všech omezení je konečný, takže se na množinu X můžeme podívat také jako na průnik konečného počtu poloprostorů neboli na *polyedr*[90]. Je-li tento polyedr ohraničený, jedná se o tzv. *polytop* či *simplex*,

[87] Rovnoběžnost obvykle nezahrnuje totožnost (splývání), takže v takovém případě soustava nebude mít řešení.

[88] V takovém případě buď nadrovina splývá (je totožná) s jinou nebo sice nesplývá, ale její průnik s ostatními je stejný jako průnik ostatních mezi sebou, tj. tato nadrovina vlastně žádné nové omezení nepřidává a její případné odebrání nemá žádný vliv na podobu přípustné množiny.

[89] Konvexní množina je taková, že s libovolnými dvěma body obsahuje i celou úsečku určenou těmto body (např. úsečka, trojúhelník, čtverec, kruh, koule atd.)

[90] Velmi specifickým polyedrem je i prázdná množina, kterou a priori z našich úvah nevylučujeme.

[91] Tj. nejmenší konvexní množina obsahující dané body neboli mnohoúhelník s vrcholy v těchto bodech.

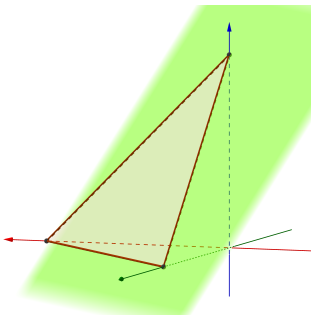
[92] Tj. žádná trojice neleží na jedné přímce, žádná čtveřice bodů neleží v jedné rovině, atd.

což je *konvexní obal*[91] $(k+1)$-tice *afinně nezávislých* bodů[92], které představují vrcholy množiny X, tj.

$$X = \left\{ x \in \mathbb{R}^n \,\middle|\, x = \sum_{i=1}^{k+1} \lambda_i u_i \text{ pro nějaká } \lambda_i \in [0,1] \text{ a } \sum_{i=1}^{k+1} \lambda_i = 1 \right\}, \tag{2.4.1}$$

kde $u_1, \ldots, u_{k+1}$ jsou jednotlivé vrcholy simplexu. Např. množina X na Obrázku 2.21 je tzv. *standardní* (nebo *jednotkový*) 2-*simplex* určený vrcholy $u_1 = [1,0,0]$, $u_2 = [0,1,0]$ a $u_3 = [0,0,1]$. V $\mathbb{R}^2$ bude množina X vždy nějaký konvexní ℓ-úhelník, který má v ohraničeném případě $\ell \leq m+n$ hran a vrcholů[93], zatímco v neohraničeném případě bude mít $\ell - 1$ hran a $\ell - 2$ vrcholů.

[93] V tomto kontextu rozumíme 0-úhelníkem prázdnou množinu, 1-úhelníkem jednobodovou množinu a 2-úhelníkem úsečku



Obrázek 2.21: Množina $X \subseteq \mathbb{R}^3$ pro $A_R = (1,1,1)$ a $b_R = 1$, zatímco $A_N = 0$, $b_N = 0$ (tj. žádné nerovnosti) a $s = 3$ (tj. všechny proměnné jsou nezáporné).

K čemu to vlastně potřebujeme? Možná k ničemu a později to již nevyužijeme, protože naše možnosti grafického znázornění úloh lineární programování (stejně jako kterýchkoliv jiných) končí u $\mathbb{R}^3$. Proto chceme vybudovat algoritmus čistě analyticky (abstraktně) a nezávisle na nějaké vizualizaci. Je navíc asi celkem přirozené očekávat, že praktické úlohy budou obsahovat (mnohem) více než tři proměnné. To by ale bylo až příliš příkré hodnocení. Uvedený geometrický popis má přeci jen velmi důležitou dvojroli

(i) *intuice:* s jeho pomocí můžeme na základě úloh v $\mathbb{R}^2$ či $\mathbb{R}^3$ získat představu o možném řešení obecných úloh, která by se nám mohla hodit při hledání řešení obecných úloh;

(ii) *etymologie:* v roce 1948 popsal Dantzig svůj algoritmus Motzkinovi, který v tom spatřil jakési systematické procházení vrcholů simplexu, díky čemuž získal tento algoritmus svůj název – *simplexový algoritmus*[94].

Než se ale pustíme do hlubší analýzy obecných úloh lineárního programování, zodpovíme si otázku samotné existence řešení takových úloh. K tomu budeme potřebovat jeden důsledek následující velmi důležitého tvrzení, které si v tuto chvíli uvedeme bez důkazu. Ne že bychom jej nezvládli – mohli bychom si zde pro něj *ad hoc* vybudovat potřebný aparát, avšak bude lepší s tím počkat až do Kapitoly 6, kde si jej uvedeme se všemi podrobnostmi a v širších souvislostech, takže samotný důkaz pak už bude „snadný“, viz Větu 6.3.19. Nicméně si ukážeme alespoň jednoduchý obrázek, který velmi dobře ilustruje podstatu tvrzení v $\mathbb{R}^2$, viz Obrázek 2.22.

[94] Viz také poznámku 127 na začátku Podkapitoly 2.7.

Věta 2.4.4 Farkas & Minkowski

Nechť je dána matice $A \in \mathbb{R}^{m \times n}$ a vektor $b \in \mathbb{R}^m$. Pak buď existuje řešení soustavy

$$Ax = b \quad \& \quad x \geq 0, \tag{2.4.2}$$

nebo existuje řešení soustavy

$$A^\top y \geq 0 \quad \& \quad \langle y, b \rangle < 0. \tag{2.4.3}$$

Pro nás je ale v tuto chvíli důležitý především následující důsledek Farkasovy–Minkowského věty, který obvykle bývá připisován Galeovi, viz [196, Theorem 2.7].

Důsledek 2.4.5 Gale

Nechť je dána matice $A \in \mathbb{R}^{m \times n}$ a vektor $b \in \mathbb{R}^m$. Pak buď existuje řešení soustavy

$$Ax \leq b \quad \& \quad x \in \mathbb{R}^n, \tag{2.4.4}$$

nebo existuje řešení soustavy

$$A^\top y = 0 \quad \& \quad \langle y, b \rangle < 0 \quad \& \quad y \geq 0. \tag{2.4.5}$$

Důkaz. Definujme matici $\widetilde{A} := (A,\ b) \in \mathbb{R}^{m \times (n+1)}$ a vektory $\widetilde{b} := (0,\ -1)^\top \in \mathbb{R}^{n+1}$ pro nulový vektor $0 \in \mathbb{R}^n$ a také $\widetilde{x} := \binom{x}{\lambda} \in \mathbb{R}^{n+1}$ pro $\lambda \in \mathbb{R}$. Pak systém (2.4.4) lze zapsat jako

$$\widetilde{A}\widetilde{x} \geq 0 \quad \& \quad \langle \widetilde{x}, \widetilde{b} \rangle < 0 \tag{2.4.6}$$

a systém (2.4.5) je ekvivalentní s

$$\widetilde{A}^\top y = \widetilde{b} \quad \& \quad y \geq 0. \tag{2.4.7}$$

Vskutku, je-li $\widetilde{x}^* = \binom{x^*}{\lambda^*}$ řešením (2.4.6), pak

$$Ax^* + \lambda^* b \geq 0 \quad \& \quad \lambda^* > 0 \quad \text{neboli} \quad -\tfrac{1}{\lambda^*} Ax^* \leq b,$$

tj. $-\frac{1}{\lambda^*} x^*$ je řešením úlohy (2.4.4). Naopak, je-li x^* řešením (2.4.4), pak vektor $\widetilde{x}^* = \binom{-x^*}{1}$ splňuje (2.4.6). Podobně, pokud y^* splňuje (2.4.7), pak $A^\top y^* = 0$, $b^\top y^* = -1$ a $y^* \geq 0$, takže y^* jistě vyhovuje i soustavě (2.4.5). Konečně, je-li y^* řešením (2.4.5), pak pro $-\frac{1}{\langle y^*, b\rangle} y^*$ platí (2.4.7). To tedy znamená (odmyslíme-li si záměnu označení x a y), že úloha (2.4.4) je ekvivalentní s úlohou typu (2.4.3) a problém (2.4.5) s úlohou typu (2.4.2). Pak tvrzení o řešitelnost právě jednoho ze systémů (2.4.4)–(2.4.5) ihned vyplývá z Farkasovy–Minkowského věty 2.4.4. ■

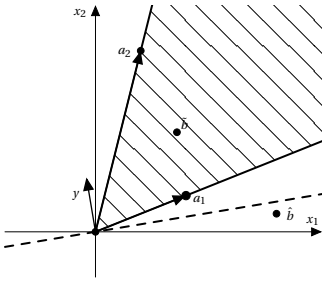
Nyní už máme vše potřebné k tomu, abychom si dokázali větu o řešitelnosti úloh lineárního programování, která říká, že hledání takového *řešení může selhat pouze z očividných důvodů.*

Věta 2.4.6

Je-li v úloze lineárního programování v základním tvaru

$$\langle c, x\rangle \to \min \quad \& \quad x \in X := \{x \in \mathbb{R}^n \mid Ax \geq b\}$$

pro matici $A \in \mathbb{R}^{m\times n}$ a vektory $b \in \mathbb{R}^m$ a $c \in \mathbb{R}^n$ přípustná množina X neprázdná a funkce $f(x) := \langle c, x\rangle$ zdola ohraničená na X, pak má tato úloha alespoň jedno řešení[95].



Obrázek 2.22: Geometrický pohled na Farkasovu–Minkowského větu pro matici $A = (a_1, a_2) \in \mathbb{R}^{2\times 2}$ a vektory $b = \tilde{b}$ a $b = \hat{b}$. Splnění soustavy (2.4.2) pro nějaké $x^* \geq 0$ znamená, že bod $\tilde{b}$ patří do množiny vymezené dvojicí polopřímek vycházejících z počátku ve směrech vektorů určených sloupci $a_1, a_2 \in \mathbb{R}^2$ z matice A. Naopak, je-li pro nějaké y^* splněna soustava (2.4.3), pak přímka procházející počátkem a s normálovým vektorem y^* *odděluje* vyšrafovanou množinu z první části a bod $\hat{b}$. Je tedy zřejmé, že bod b má právě jednu z těchto vlastností, a tedy právě jenom jedna ze soustav (2.4.2) a (2.4.3) může mít řešení.

Důkaz. Nechť $X \neq \emptyset$ a účelová funkce je ohraničená na X, tj. existuje konečné číslo $f^* := \inf_{x\in X} f(x) > -\infty$. Připusťme, že daná úloha není řešitelná. To znamená, že neexistuje řešení soustavy

$$Ax \geq b \quad \& \quad \langle c, x\rangle \leq f^*,$$

což je ekvivalentní s neexistencí řešení soustavy

$$\begin{pmatrix} -A \\ c^\top \end{pmatrix} x \leq \begin{pmatrix} -b \\ f^* \end{pmatrix}, \quad x \in \mathbb{R}^n.$$

Proto podle Důsledku 2.4.5 musí existovat $y \in \mathbb{R}^m$ a $\lambda \in \mathbb{R}$ taková, že dvojice $\binom{y}{\lambda}$ řeší systém

$$\begin{pmatrix} -A^\top & c \end{pmatrix} \begin{pmatrix} y \\ \lambda \end{pmatrix} = 0 \quad \& \quad \left\langle \begin{pmatrix} y \\ \lambda \end{pmatrix}, \begin{pmatrix} -b \\ f^* \end{pmatrix} \right\rangle < 0 \quad \& \quad \begin{pmatrix} y \\ \lambda \end{pmatrix} \geq 0,$$

což je ekvivalentní se soustavou

$$A^\top y = \lambda c \quad \& \quad \langle y, b\rangle > \lambda f^* \quad \& \quad y \geq 0 \quad \& \quad \lambda \geq 0. \tag{2.4.8}$$

V takovém případě pak pro libovolné $x \in X$ platí

$$\lambda \langle c, x\rangle = \langle \lambda c, x\rangle \overset{(2.4.8)}{=} \langle A^\top y, x\rangle = \langle y, Ax\rangle \overset{x\in X}{\geq} \langle y, b\rangle \overset{(2.4.8)}{>} \lambda f^*, \tag{2.4.9}$$

z čehož vyplývá, že[96] $\lambda > 0$ a současně $\langle c, x\rangle \geq \frac{\langle y,b\rangle}{\lambda} > f^*$. Jenže odtud dostáváme spor ↯, neboť vzhledem k libovolnosti $x \in X$ by muselo platit

$$f^* = \inf_{x\in X} \langle c, x\rangle \geq \frac{\langle y, b\rangle}{\lambda} > f^*.$$

■

[95] ⚠ Úloha ve Větě 2.4.6 je sice dána v základním tvaru, ale my již víme (viz Příklad 2.2.2), že ta je ekvivalentní s úlohou v kanonickém/standardním či zcela obecném tvaru, takže totéž tvrzení nutně platí také pro úlohu lineárního programování s množinou X v libovolném tvaru.

[96] Hodnota $\lambda = 0$ totiž není možná vzhledem k ostré nerovnosti v poslední části (2.4.9).

Cvičení

2.4.1. Grafickou metodou určete řešení problému

$$x_1 + x_2 \to \min$$
$$x_1 + 2x_2 \leq 4 \quad \& \quad 4x_1 + 2x_2 \leq 12 \quad \& \quad -2x_1 + x_2 \leq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[0, 0]$.

2.4.2. Grafickou metodou určete řešení problému

$$-x_1 - 2x_2 \to \min$$
$$x_1 - 4x_2 \le 4 \quad \& \quad -2x_1 + x_2 \le 2$$
$$-3x_1 + 4x_2 \le 12 \quad \& \quad 2x_1 + x_2 \le 8$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[20/11, 48/11]$.

2.4.3. Grafickou metodou určete řešení problému

$$-x_1 - 2x_2 \to \min$$
$$x_1 - 4x_2 \le 4 \quad \& \quad -2x_1 + x_2 \le 2 \quad \& \quad -3x_1 + 4x_2 \le 12,$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá řešení.

2.4.4. Grafickou metodou určete řešení problému

$$-3x_1 + x_2 \to \min$$
$$x_1 + x_2 \le 10 \quad \& \quad 3x_1 + x_2 \le 15 \quad \& \quad -x_1 + x_2 \le 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[5, 0]$.

2.4.5. Grafickou metodou určete řešení problému

$$x_1 - x_2 \to \min$$
$$x_1 + x_2 \le 6 \quad \& \quad 3x_1 + x_2 \le 15 \quad \& \quad -x_1 + x_2 \le 3,$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nekonečně mnoho řešení na úsečce spojující body $[0, 3]$ *a* $[3/2, 9/2]$.

2.4.6. Grafickou metodou určete řešení problému

$$-2x_1 + x_2 \to \min$$
$$2x_1 + 3x_2 = 6 \quad \& \quad x_1 - 2x_2 \le 0 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[12/7, 6/7]$.

2.4.7. Grafickou metodou určete řešení problému

$$x_1 + 3x_2 \to \max$$
$$x_1 - 2x_2 \le 0 \quad \& \quad -2x_1 + x_2 \le 4 \quad \& \quad 5x_1 + 3x_2 \le 15$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[3/11, 50/11]$.

2.4.8. Grafickou metodou určete řešení problému

$$2x_1 - x_2 \to \min$$
$$x_1 - 3x_2 \ge -3 \quad \& \quad 2x_1 + x_2 \ge -2$$
$$2x_1 + 3x_2 \le 6 \quad \& \quad 3x_1 - 2x_2 \le 6.$$

Řešení: $[-9/7, 4/7]$.

2.4.9. Grafickou metodou určete řešení problému

$$6x_1 + 5x_2 \to \max$$
$$x_1 + 2x_2 \le 8 \quad \& \quad 3x_1 + x_2 \le 15 \quad \& \quad x_1 + x_2 \le 5$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[5, 0]$.

2.4.10. Grafickou metodou určete řešení problému

$$3x_1 + 2x_2 \to \max$$
$$x_1 - x_2 \le 0 \quad \& \quad x_1 + x_2 \le 10 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[5,5]$.

2.4.11. Grafickou metodou určete řešení problému

$$6x_1 + 3x_2 \to \min$$
$$x_1 + x_2 \ge 1 \quad \& \quad 2x_1 - x_2 \ge 1 \quad \& \quad 3x_2 \le 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[2/3, 1/3]$.

2.4.12. Grafickou metodou určete řešení problému

$$-2x_1 + 3x_2 \to \min$$
$$2x_1 + x_2 \le 10 \quad \& \quad -2x_1 + 3x_2 \le 9 \quad \& \quad 2x_1 + 4x_2 \ge 8$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[5,0]$.

2.4.13. Grafickou metodou určete řešení problému

$$5x_1 + 4x_2 \to \min$$
$$x_1 + x_2 \le 6 \quad \& \quad 2x_1 + x_2 \le 5$$
$$4x_1 + x_2 \ge 2 \quad \& \quad x_1 + x_2 \ge 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[1/3, 2/3]$.

2.4.14. Grafickou metodou určete řešení problému

$$3x_1 + 4x_2 \to \max$$
$$3x_1 + x_2 \le 9 \quad \& \quad -2x_1 + x_2 \le 4 \quad \& \quad x_1 + 2x_2 \ge 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[1,6]$.

2.4.15. Grafickou metodou určete řešení problému

$$x_1 + x_2 \to \max$$
$$x_1 - x_2 \ge 1 \quad \& \quad x_1 + 2x_2 \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá konečné řešení.

2.4.16. Grafickou metodou určete řešení problému

$$x_1 + x_2 \to \min$$
$$x_1 - x_2 \ge 1 \quad \& \quad -x_1 + x_2 \ge 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Úloha není přípustná, tj. $X = \emptyset$.

2.4.17. Grafickou metodou určete řešení problému

$$2x_1 + 4x_2 \to \max$$
$$2x_1 - x_2 \ge -2 \quad \& \quad x_1 + 2x_2 \le 8 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nekonečně mnoho řešení na úsečce spojující body $[8,0]$ *a* $[4/5, 18/5]$.

2.4.18. Grafickou metodou určete řešení problému

$$2x_1 - 6x_2 \to \max$$
$$3x_1 + 2x_2 \le 6 \quad \& \quad x_1 - x_2 \ge -1 \quad \& \quad -x_1 - 2x_2 \ge 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.4.19. Grafickou metodou určete řešení problému

$$x_1 + 2x_2 \to \min$$
$$x_1 - 3x_2 \le 6 \quad \& \quad 2x_1 + 4x_2 \ge 8 \quad \& \quad x_1 - 3x_2 \le -6$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[0,2]$.

2.4.20. Grafickou metodou určete řešení problému

$$3x_1 + 4x_2 \to \max$$
$$x_1 + x_2 \ge 5 \quad \& \quad 3x_1 + x_2 \ge 8 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá konečné řešení.

2.4.21. Grafickou metodou určete řešení problému

$$5x_1 + 4x_2 \to \min$$
$$x_1 + x_2 \ge 2 \quad \& \quad x_1 \le 8 \quad \& \quad x_2 \le 9 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[0,2]$.

2.4.22. Grafickou metodou určete řešení problému

$$10x_1 + 6x_2 \to \max$$
$$5x_1 + 3x_2 \le 30 \quad \& \quad x_1 + 2x_2 \le 18 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nekonečně mnoho řešení na úsečce spojující body $[6,0]$ *a* $[6/7,6/7]$.

2.4.23. Grafickou metodou určete řešení problému

$$2x_1 + 3x_2 \to \min$$
$$2x_1 + 3x_2 \le 30 \quad \& \quad -x_1 + x_2 \le 5$$
$$x_1 + x_2 \ge 5 \quad \& \quad x_1 \le 10$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nekonečně mnoho řešení na úsečce spojující body $[10,10/3]$ *a* $[3,8]$.

2.4.24. Grafickou metodou určete řešení problému

$$3x_1 + 2x_2 \to \max$$
$$x_1 - x_2 \ge 1 \quad \& \quad x_1 + x_2 \ge 3 \quad \& \quad 9x_1 + 6x_2 \ge 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá konečné řešení.

2.4.25. Grafickou metodou určete řešení problému

$$3x_1 - 2x_2 \to \max$$
$$x_1 + x_2 \le 1 \quad \& \quad 2x_1 + 2x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Úloha není přípustná, tj. $X = \emptyset$.

2.4.26. Grafickou metodou určete řešení problému

$$5x_1 + 3x_2 \to \max$$
$$x_1 + x_2 \le 6 \quad \& \quad 2x_1 + 3x_2 \ge 3 \quad \& \quad 0 \le x_1, x_2 \le 3.$$

Řešení: $[3,3]$.

2.4.27. Grafickou metodou určete řešení problému

$$2x_1 - x_2 \to \max$$
$$x_1 - x_2 \le 1 \quad \& \quad 2x_1 + x_2 \ge 6 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá konečné řešení.

2.4.28. Grafickou metodou určete řešení problému

$$3x_1 - 4x_2 \to \max$$
$$x_1 + 2x_2 \geq 4 \quad \& \quad x_1 - x_2 \leq 1 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[2,1]$.

2.4.29. Grafickou metodou určete řešení problému

$$4x_1 + 6x_2 \to \max$$
$$-x_1 + x_2 \leq 11 \quad \& \quad x_1 + x_2 \leq 27 \quad \& \quad 2x_1 + 5x_2 \leq 90.$$

Řešení: $[15,12]$.

2.4.30. Grafickou metodou určete řešení problému

$$5x_1 + 7x_2 \to \min$$
$$2x_1 + 3x_2 \geq 6 \quad \& \quad 3x_1 - x_2 \leq 15$$
$$-x_1 + x_2 \leq 4 \quad \& \quad 2x_1 + 5x_2 \leq 27.$$

Řešení: $[-6/5, 14/5]$.

2.4.31. Grafickou metodou určete řešení problému

$$5x_1 + 7x_2 \to \min$$
$$2x_1 + 3x_2 \geq 6 \quad \& \quad 3x_1 - x_2 \leq 15$$
$$-x_1 + x_2 \leq 4 \quad \& \quad 2x_1 + 5x_2 \leq 27$$
$$x_1 \geq 0.$$

Řešení: $[0,2]$.

2.4.32. Grafickou metodou určete řešení problému

$$3x_1 - 5x_2 \to \max$$
$$4x_1 + 5x_2 \geq 3 \quad \& \quad 6x_1 - 6x_2 = 7 \quad \& \quad x_1 + 8x_2 \leq 20$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[88/27, 113/54]$.

2.4.33. Grafickou metodou určete řešení problému

$$8x_1 - 4x_2 \to \max$$
$$3x_1 + x_2 \leq 7 \quad \& \quad 9x_1 + 5x_2 \leq -2 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Nemá konečné řešení.

2.4.34. Grafickou metodou určete řešení problému

$$4x_1 + 2x_2 \to \max$$
$$x_1 + 2x_2 \geq 4 \quad \& \quad 3x_1 + x_2 \geq 7 \quad \& \quad -x_1 + 2x_2 \leq 7.$$

Nemá konečné řešení.

2.4.35. Uvažme problém

$$x_1 + x_2 \to \max$$
$$a\,x_1 + b\,x_2 \leq 1 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Určete hodnoty $a, b \in \mathbb{R}$ tak, aby úloha (i) měla optimální řešení, (ii) neměla konečné řešení a (iii) neměla přípustné řešení.

2.4.36. Podnikatel by chtěl začít prodávat dva druhy stolních počítačů. Pořizovací cena typu A činí 5000,- a prodává se se ziskem 900,-, typ B stojí 8000,- a zisk činí 1000,-. Základní odhad měsíčního prodeje je nejvýše 250 kusů a počáteční kapitál je 1,4 milionu Kč. Určete, kolik by mělo být na počátku zakoupeno počítačů, aby bylo dosaženo maximálního zisku. *Řešení:* $[200, 50]$.

2.4.37. Jahodárna ve Slatině má k dispozici 150 arů plochy k pěstování dvou odrůd jahod (značme je jako J1 a J2). Údržba 1 aru odrůdy J1 zabere 1 pracovní den a pro J2 to jsou 2 pracovní dny. Celkem je k dispozici 240 pracovních dnů. Sběr jahod na jednom aru s odrůdou J1 zabere 3/10 dne a pro J2 to je pouze 1/10 dne, přičemž sběr probíhá po dobu 30 dnů. Zisk z 1 aru jahod odrůdy J1 činí 2800,- a pro odrůdu J2 to je 4700,-. Jak rozdělit celkovou plochu, aby bylo dosaženo maximálního zisku? *Řešení:* [60, 90].

2.4.38. Výrobce benzínu má na skladě zásoby benzínu s oktanovým číslem 80 a 92, přičemž náklady na 1 litr benzínu s o.č. 80 jsou 16,60,- a s o.č. 92 jsou 19,60,-. Z rozhodnutí vlády č. 231/2015 je možné prodávat benzín pouze s oktanovým číslem alespoň 90. Jak pomocí smíchání dostupných zásob vyrobit prodejný benzín, aby bylo dosaženo minimálních nákladů? *Řešení:* [1/6, 5/6].

2.4.39. Spolek mladých přírodovědců pro PřF MU má roční členský poplatek ve výši 80,- pro studenty SŠ (příp. ZŠ a MŠ) a 200,- pro studenty VŠ a ostatní zájemce. Pro pokrytí ročních nákladů je nutné vybrat alespoň 7920,- Navíc zakladatelé byli značně skeptičtí, proto vyrobili pro první rok pouze 50 kusů členských karet. Současně však vedení fakulty vyžaduje, aby počet členů se sníženým ročním poplatkem byl mezi 1/4 a 1/3 počtu členů se členským poplatkem v plné výši. Jaký nejmenší možný počet členů splní tyto požadavky? *Řešení:* [9, 36].

2.4.40. Halina Pawlowská nabízí dva druhy „samohubnoucích" nápojů. Jedna porce přípravku HP1 obsahuje 60 kalorií, 12 jednotek vitamínu A 10 jednotek vitamínu C. Přípravek HP2 dodá také 60 kalorií, ale 6 jednotek vitamínu A a 30 jednotek vitamínu C. Cena jedné porce HP1 je 2,40,- a HP2 jsou 3,-. Minimální denní příjem činí alespoň 300 kalorií, 36 jednotek vitamínu A a 90 jednotek vitamínu C. Kolik porcí jednotlivých nápojů je potřeba vypít během dne, aby byly splněny minimální požadavky při minimálních nákladech. *Řešení:* [3, 2].

2.5 Teoretické základy aneb být, či nebýt BPB

V dalších částech našeho teoretického výkladu potřebného pro vybudování simplexového algoritmu se zaměříme pouze na úlohu lineárního programování (2.2.1) s množinou X v kanonickém tvaru[97], tj.

$$\langle c, x\rangle \to \min \quad \& \quad x \in X := \{x \in \mathbb{R}^n \mid Ax = b \ \& \ x \geq 0\} \tag{LP}$$

s danou maticí $A \in \mathbb{R}^{m\times n}$ a vektory $c \in \mathbb{R}^n$ a $b \in \mathbb{R}^m$. Samozřejmě se zde má smysl omezit jenom na situaci, kdy soustava $Ax = b$ má alespoň jedno řešení, protože v opačném případě je nutně $X = \emptyset$ a úloha (LP) je *nepřípustná*. Jelikož můžeme bez jakékoli újmy na obecnosti předpokládat, že žádná z rovnic v soustavě $Ax = b$ není nadbytečná, pak podle *Rouchéovy-Capelliho věty*[98] je nutně $\operatorname{rank} A = m \leq n$. Ovšem zajímavý je především případ $m < n$, neboť pro $m = n$ je soustava $Ax = b$ jednoznačně řešitelná, takže množina X bude buď zase prázdná (pokud ono jediné řešení nesplňuje druhou podmínku $x \geq 0$) nebo bude jednobodová (a tento bod bude automaticky naším hledaným řešením daného problému lineárního programování). Nicméně my nebudeme v našem výkladu explicitně vyžadovat ostrou nerovnost mezi m a n, neboť navzdory jednoduchému vzorci $x = A^{-1}b$ může být nalezení jediného přípustného bodu výpočetně velmi náročné[99]. Navíc je potřeba mít na paměti, že úloha může být nepřípustná i pro $m < n$, protože prvky množiny X musí splňovat také podmínku nezápornosti, viz Obrázky 2.23 a 2.24.

Pro náš algoritmus budeme ještě potřebovat, aby složky vektoru b byly nezáporné, tj. $b \geq 0$. Toho ale lze velmi snadno dosáhnout pouhým vynásobením odpovídajících řádků soustavy $Ax = b$ číslem -1. Celkem tedy v úloze (LP) můžeme bez újmy na obecnosti předpokládat, že

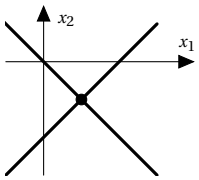
(i) *sloupce* matice A jsou $a_1, \ldots, a_n \in \mathbb{R}^m$ (pouze označení), tj.

$$A = (a_1, a_2, \ldots, a_n);$$

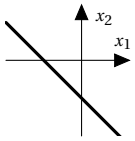
[97] Nezapomeňme, že už umíme do této podoby převést libovolnou úlohu lineárního programování, takže tímto naším „rozhodnutím" vůbec o nic nepřicházíme.

[98] Toto tvrzení je mnohdy nepřesně označováno jako *Frobeniova věta*: soustava $Ax = b$ má alespoň jedno řešení právě tehdy, když hodnost matice soustavy je stejná jako hodnost rozšířené matice soustavy, tj. platí-li rovnost $\operatorname{rank} A = \operatorname{rank}(A \mid b)$. Je-li dokonce $\operatorname{rank} A = n$, pak existuje právě jedno řešení, zatímco pro $\operatorname{rank} A = \operatorname{rank}(A \mid b) < n$ je řešení nekonečně mnoho a tvoří afinní podprostor dimenze $n - \operatorname{rank} A$.

[99] Např. v počítačové tomografii jsou zcela běžné soustavy s $m = n = 10^6$, takže pokud matice A nemá nějakou velmi speciální strukturu (např. velké množství nulových prvků), je nalezení řešení prakticky nemožné.



Obrázek 2.23: Soustava $x_1 - x_2 = 2$ a $x_1 + x_2 = 0$ má jediné řešení $x^* = [1, -1]$, které ovšem nesplňuje podmínku nezápornosti, tj. $X = \emptyset$



Obrázek 2.24: Soustava $x_1 + x_2 = 1$ má nekonečně mnoho řešení, ovšem žádné z nich nesplňuje podmínku nezápornosti, tj. $X = \emptyset$

(ii) $m \le n$ a $\operatorname{rank} A = m$ (tj. řádky matice A jsou lineárně nezávislé[100]);

(iii) $b \ge 0$ (pouze technický požadavek).

A co je množina X v tomto případě? Je to konvexní podmnožina $\mathbb{R}^n$ vytvořená průnikem kladného orthantu v $\mathbb{R}^n$ (kvadrantu v $\mathbb{R}^2$ a oktantu v $\mathbb{R}^3$) určeného podmínkou $x \ge 0$ a m-tice nadrovin určených každou z rovnic v soustavě $Ax = b$. Tato podmnožina může být ohraničená (polytop či simplex) jako na Obrázku 2.21 nebo neohraničená jako na Obrázku 2.25.

Avšak k úspěšnému vybudování zmiňovaného algoritmu je ještě dlouhá cesta lemovaná především několika zajímavými (geometrickými) poznatky o množině X a minimu účelové funkci na této množině, s nimiž si potvrdíme/vyvrátíme náš prvotní dojem ohledně lokalizace hledaného minima nabytý při řešení úloh lineárního programování z $\mathbb{R}^2$ v předchozí podkapitole pomocí vhodných náčrtků.

Nechť je dána úloha (LP). Pro libovolný bod $x = [x_1, \dots, x_n] \in \mathbb{R}^n$ definujeme jeho *nosič* jako množinu

$$J(x) := \{ j \in \{1, \dots, n\} \mid x_j > 0 \},$$

která obsahuje indexy nenulových souřadnic bodu $x \ge 0$. Je-li navíc $x \in X$, pak jej nazveme *bazickým přípustným bodem*[101] (dále obvykle jen *BPB*) úlohy (LP), jestliže vektory $\{a_j \mid j \in J(x)\}$ jsou lineárně nezávislé.

Všimněme si, že v předchozí definici jsme sice vyžadovali úlohu (LP), ale uvedené podmínky pro BPB jsou vztaženy pouze k množině X a účelová funkce $\langle c, x \rangle$ zde nehraje vůbec žádnou roli. Velmi snadno bychom mohli definici přeformulovat pouze s použitím množiny X v kanonickém tvaru, ale z čistě formálních důvodů jsme se rozhodli upřednostnit toto pojetí. Totéž platí i pro následující čtyři tvrzení (až po Větu 2.5.8) a z větší části i pro Definici 2.5.9.

Nejjednodušším BPB úlohy (LP) je $x = 0$ s nosičem $J(x) = \emptyset$, což je však možné pouze pokud $b = 0$. Nicméně rozhodnout, zda i jiné přípustné body množiny X vyhovují podmínkám předchozí definice, je celkem snadné a my si to ilustrujeme v následujícím příkladu.

Příklad 2.5.2

Uvažme úlohu (LP) s množinou X určenou maticí $A = \left(\begin{smallmatrix} 1 & 5 & 3 & 4 & 6 \\ 0 & 1 & 3 & 5 & 6 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 5}$, která očividně splňuje podmínku $\operatorname{rank} A = 2$, a vektorem $b = \binom{14}{7} \in \mathbb{R}^2_+$. Pak například přípustný bod $\widetilde{x} = [0,2,0,1,0] \in X$ je BPB této úlohy, neboť jeho nosič je $J(\widetilde{x}) = \{2,4\}$ a odpovídající vektory $a_2 = \binom{5}{1}$, $a_4 = \binom{4}{5}$ jsou lineárně nezávislé. Na druhou stranu v případě vektoru $b = \binom{27}{27} \in \mathbb{R}^2_+$ přípustný bod $\widehat{x} = [0,0,1,0,4] \in X$ není BPB této úlohy, neboť máme $J(\widehat{x}) = \{3,5\}$ a odpovídající vektory $a_3 = \binom{3}{3}$, $a_5 = \binom{6}{6}$ jsou zjevně lineárně závislé.

Jenže v předchozím příkladě jsme si hledání BPB výrazně usnadnili tím, že jsme testovali přípustné body $\widetilde{x}$ a $\widehat{x}$, které se zjevili jako *Deus ex machina*[102] . To samozřejmě není příliš praktické a při hledání BPB bychom se chtěli vyhnout tomu, že budeme ověřovat podmínky Definice 2.5.1 pro všechny přípustné body množiny X. Jak ale BPB najít snadněji? Mají tyto body nějakou význačnou vlastnost, která by nám jejich hledání usnadnila? A kolik BPB vlastně vůbec může množina X mít? To jsou klíčové otázky, na které se nyní pokusíme najít odpověď. Nejdříve si ukážeme „geometrický" popis BPB. Již jsme se zmínili, že každá konvexní množina obsahuje všechny úsečky určené body z této množiny, tj. pro libovolné $x_1, x_2 \in X$ a $\lambda \in [0,1]$ máme $\lambda x_1 + (1-\lambda) x_2 \in X$. To ale znamená, že (s výjimkou prázdné a jednobodové množiny) většina bodů v množině X leží „uvnitř" nějaké úsečky s krajními body z X. Ovšem to nejsou všechny – spolu s těmito body tam jsou i takové body, pro které žádná taková úsečka neexistuje, tj. nelze ani „o kousek" prodloužit žádnou úsečku vedoucí z množiny X za tento bod, aniž by toto prodloužení neopustilo množinu X. Toto jsou pro nás *extrémně* důležité body.

[100] Tento požadavek ale vůbec není na újmu obecnosti. Kdyby $\operatorname{rank} A = k < m$, pak by bylo lineárně nezávislých pouze k řádků matice A a připusťme, že to jsou řádky $\widetilde{a}_1, \dots, \widetilde{a}_k \in \mathbb{R}^{1\times n}$. Pak ale množinu X můžeme získat i pouze s pomocí těchto vektorů, tj. platí

$$X = Y := \{x \in \mathbb{R}^n \mid \widetilde{a}_1 x = b_1 \ \& \ \dots \ \& \ \widetilde{a}_k x = b_k \ \& \ x \ge 0)\}.$$

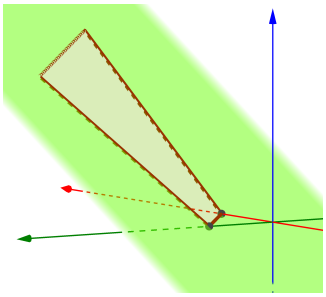
Zjevně je $X \subseteq Y$, neboť libovolné $x \in X$ samozřejmě splňuje i podmínky vymezující množinu Y. Naopak, je-li $y \in Y$, pak

$$\widetilde{a}_i y = \sum_{j=1}^{k} \lambda_{ij} \widetilde{a}_j y = \sum_{j=1}^{k} \lambda_{ij} b_j = b_i,$$

pro všechna $i \in \{1, \dots, m\}$, neboť pro každý řádek $\widetilde{a}_i$ existují čísla λ_{ij} taková, že $\widetilde{a}_i = \sum_{j=1}^{k} \lambda_{ij} \widetilde{a}_j$, a tudíž

$$b_i = \widetilde{a}_i y = \sum_{j=1}^{k} \lambda_{ij} \widetilde{a}_j y = \sum_{j=1}^{k} \lambda_{ij} b_j,$$

tj. $y \in X$.



Obrázek 2.25: Množina X z (LP) pro $A = (1,1,-1)$ a $b = 3$ (ta horní hrana je pouze „virtuální").

[101] Angl. *basic feasible point*. Lze se setkat také s terminologií *bázický přípustný bod*, *bazické/bázické řešení*.

[102] *Deus ex machina* doslova znamená *bůh ze stroje*. Jedná se o literární termín pocházející z antického starořeckého divadla (ἀπό μηχανῆς θεός neboli *apo mechanés theos*). Antičtí dramatici tímto obratem řešili zdánlivě neřešitelnou (sic!) situaci, do které se dostal hlavní hrdina. V takovém momentě byl na jeviště prostřednictvím mechanického zařízení spuštěn nějaký bůh (často v podobě sochy) a situaci rozřešil. Krátce poté divadelní kus zpravidla končil. Ačkoli takovéto rozuzlení zápletky bylo v pozdějších dobách podrobeno kritice, lze se s ním v přenesené podobě setkat i v některých dílech z období klasicismu

DEFINICE 2.5.1

Definice 2.5.3

Nechť množina $Y \subseteq \mathbb{R}^n$ je konvexní. Pak body, které nelze vyjádřit ve tvaru

$$\lambda\widetilde{y} + (1-\lambda)\,\widehat{y} \quad \text{pro nějaká } \widetilde{y}, \widehat{y} \in Y \text{ splňující } \widetilde{y} \neq \widehat{y} \text{ a nějaké } \lambda \in (0,1),$$

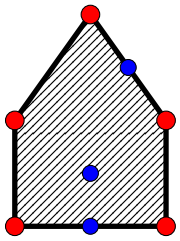
nazýváme *extrémními* (nebo *krajními*) body množiny Y.

Již víme, že jedním z typických příkladů přípustné množiny v obecné úloze lineárního programování je konvexní n-úhelník[103]. Jeho extrémní body jsou pouze jednotlivé vrcholy, jak se můžeme snadno přesvědčit např. na Obrázku 2.26. Extrémní body ale mohou vypadat i trochu jinak. Uvažme např. kruh. Má extrémní body? A jaké to jsou? Ano, je to celý obvod kruhu, tj. hraniční kružnice. Avšak takové přípustné množiny se v úlohách lineárního programování naštěstí nevyskytují.

[103] Je to pravda i v případě množiny X v kanonickém tvaru? Rozmyslete si, co mohou být přípustné množiny úlohy (LP) v $\mathbb{R}^2$ a $\mathbb{R}^3$.

Věta 2.5.4

Nechť je dána úloha (LP). Bod $x \in X$ je BPB úlohy (LP) právě tehdy, když je extrémním bodem množiny X.



Obrázek 2.26: Konvexní n-úhelník a všechny jeho extrémní body (červená barva) společně s některými „neextrémní“ body (modrá).

Důkaz. „$\Longrightarrow$“ Nejdříve předpokládejme, že x je BPB úlohy (LP), ale není extrémním bodem množiny X, tj. existují $\widetilde{x}, \widehat{x} \in X$ a $\lambda \in (0,1)$ taková, že $\widetilde{x} \neq \widehat{x}$ a $x = \lambda\widetilde{x} + (1-\lambda)\,\widehat{x}$. Pak nutně musí být $\widetilde{x}_j = \widehat{x}_j = 0$ pro $j \notin J(x)$, jelikož pro tyto indexy máme $0 = x_j = \lambda\widetilde{x}_j + (1-\lambda)\,\widehat{x}_j$ a na pravé straně jsou pouze nezáporná čísla. Proto platí

$$b = \underbrace{\sum_{j=1}^{n} \widetilde{x}_j\, a_j}_{=A\widetilde{x}} = \sum_{j\in J(x)} \widetilde{x}_j\, a_j = \underbrace{\sum_{j\in J(x)} \widehat{x}_j\, a_j}_{=A\widehat{x}}, \quad \text{tj.} \quad 0 = \sum_{j\in J(x)} (\widetilde{x}_j - \widehat{x}_j)\, a_j. \tag{2.5.1}$$

Jenže x je BPB úlohy (LP), takže vektory $\{a_j \mid j \in J(x)\}$ jsou lineárně nezávislé, a tudíž rovnost (2.5.1) je splněna tehdy a jenom tehdy, když $\widetilde{x}_j - \widehat{x}_j = 0$ neboli $\widetilde{x}_j = \widehat{x}_j$ pro $j \in J(x)$. To ale celkem znamená, že $\widetilde{x} = \widehat{x}$, což je spor $\lightning$.

„$\Longleftarrow$“ Důkaz opačné implikace se provede podobně. Nechť $x \in X$ je extrémním bodem množiny X, ale současně není BPB úlohy (LP), tj. vektory $\{a_j \mid j \in J(x)\}$ jsou lineárně závislé, což vylučuje možnost $x = 0$ (pokud vůbec takový bod patří do X). Musí proto existovat m-tice čísel $\{c_j \in \mathbb{R} \mid j \in J(x)\}$ takových, že alespoň jedno je nenulové a současně platí $\sum_{j\in J(x)} c_j\, a_j = 0$. Položíme-li $c_j := 0$ pro $j \notin J(x)$ a definujeme-li vektor $c := (c_1, \ldots, c_n)^\top \in \mathbb{R}^n$, pak

$$Ac = \sum_{j=1}^{n} c_j\, a_j = \sum_{j\in J(x)} c_j\, a_j + \sum_{j\notin J(x)} c_j\, a_j = 0,$$

a tedy $A(x \pm \alpha\, c) = Ax \pm \alpha\, Ac = b$ pro libovolné $\alpha \geq 0$. Navíc pro dostatečně malá $\alpha \geq 0$ platí také $x \pm \alpha\, c \geq 0$, takže pro tyto hodnoty α máme[104] $x \pm \alpha\, c \in X$. Jenže současně platí

$$x = \tfrac{1}{2}(x + \alpha\, c) + \tfrac{1}{2}(x - \alpha\, c),$$

takže jsme x vyjádřili jako konvexní kombinaci dvou různých bodů z X, což je spor s extrémností bodu x $\lightning$. ∎

[104] Zde totiž perturbujeme pouze kladné souřadnice bodu x, takže stačí vzít $\alpha \in (0, \alpha_0)$ pro

$$0 < \alpha_0 < \min_{\substack{j\in J(x)\\ c_j \neq 0}} \{x_j / |c_j|\}.$$

Zdůrazněme ještě, že $\alpha_0 = 0$ může nastat pouze pro $x = 0$, což jsme z našich úvah vyloučili již na začátku.

Nyní zodpovíme otázku samotné existence BPB úlohy (LP). Hlavní myšlenka v důkazu následujícího tvrzení je podobná té, kterou jsme využili již před chvílí.

Věta 2.5.5

Nechť je dána úloha (LP) s přípustnou množinou $X \neq \emptyset$. Pak množina všech BPB úlohy (LP) je neprázdná a navíc pro libovolné $x \in X$ existuje BPB $\widetilde{x} \in X$ takový, že $J(\widetilde{x}) \subseteq J(x)$.

Toto tvrzení nám zaručuje existenci alespoň jednoho BPB pro každou úlohu (LP) s neprázdnou přípustnou množinou. Ovšem zdůrazněme, že zde je zcela klíčový požadavek množiny X v kanonickém tvaru. Kdybychom např. uvážili množinu X vymezenou pouze nerovností $x_1 - x_2 \leq 2$, dostali bychom poloprostor určený přímkou procházející body $[2,0]$ a $[0,-2]$, což

je množina *nemající* žádné extrémní body (či vrcholy), čímž bychom se s jistým nadhledem mohli dostat do sporu s existencí BPB a Větou 2.5.4. Avšak pokud vyjádříme množinu X v kanonickém tvaru (pomocí dvou nových proměnných, tj. budeme mít množinu v $\mathbb{R}^4$), pak již BPB existovat rozhodně bude[105].

[105] Položíme $x_1 = s_1 - z$ a $x_2 = s_2 - z$. Pak množina X v kanonickém tvaru bude vymezena podmínkami $s_1 - s_2 + s_3 = 2$, $s_1 \geq 0$, $s_2 \geq 0$ a $s_3 \geq 0$. BPB takové množiny je např. $[2,0,0,0]$. Proměnná z sice v novém omezení nevystupuje, ale nelze ji zanedbat, neboť se může objevit v účelové funkci a mít zcela zásadní vliv na řešitelnost úlohy (pokud v účelové funkci proměnná z vystoupí se záporným koeficientem, bude určitě úloha neohraničená; vlastně úloha bude mít řešení pouze, pokud z v účelové funkci nebude vůbec).

Důkaz Věty 2.5.5. Vzhledem k předpokladu $X \neq \emptyset$ je první část tvrzení *de facto* pouze přímým důsledkem druhé části, která zaručuje existenci alespoň jednoho BPB. Dokážeme proto pouze druhou část tvrzení. Nechť $x \in X$ je libovolné, tj. $x \geq 0$ a $b = \sum_{j=1}^{n} x_j a_j$. Naším cílem je najít BPB $\widetilde{x} \in X$ s nejvýše stejným počtem nenulových prvků, tj. $\widetilde{x} \geq 0$ takový, že $b = \sum_{j \in J(x)} \widetilde{x}_j a_j$ a současně $J(\widetilde{x}) \subseteq J(x)$ s lineárně nezávislými vektory $\{a_j \mid j \in J(\widetilde{x})\}$. Je-li $b = 0$, pak všechny tyto požadavky jistě splňuje $\widetilde{x} := 0$. Proto dále uvažujme pouze $b \neq 0$.

(i) Jsou-li vektory $\{a_j \mid j \in J(x)\}$ lineárně nezávislé, pak bod x už je BPB, a volba $\widetilde{x} := x$ bude vyhovovat všem uvedeným požadavkům.

(ii) Jsou-li vektory $\{a_j \mid j \in J(x)\}$ lineárně závislé, pak existují čísla $\{\alpha_j \mid j \in J(x)\}$ taková, že ne všechna jsou nulová a $\sum_{j \in J(x)} \alpha_j a_j = 0$, takže pro libovolné $\gamma \in \mathbb{R}$ platí

$$b = \sum_{j \in J(x)} (x_j - \gamma \alpha_j)\, a_j \tag{2.5.2}$$

Navíc bez újmy na obecnosti můžeme předpokládat, že alespoň pro jedno $j \in J(x)$ je $\alpha_j > 0$, neboť v opačném případě stačí u všech α_j změnit znaménko. Proto existuje

$$\gamma^* := \min\left\{\tfrac{x_j}{\alpha_j} \mid j \in J(x),\ \alpha_j > 0\right\} > 0.$$

V takovém případě jistě $x_j - \gamma^* \alpha_j \geq 0$ pro všechna $j \in J(x)$, přičemž alespoň pro jedno $j \in J(x)$ nastane rovnost. Proto bod $\bar{x}$ se složkami $\bar{x}_j := x_j - \gamma^* \alpha_j$ pro $j \in J(x)$ a $\bar{x}_j := 0$ pro $j \notin J(x)$ je díky volbě γ^* a rovnosti (2.5.2) aplikované pro $\gamma = \gamma^*$ jistě přípustným bodem množiny X a počet prvků v jeho nosiči splňuje $|J(\bar{x})| \leq |J(x)| - 1$. Je-li takový bod $\bar{x}$ už BPB úlohy (LP), jsme hotovi. V opačném případě můžeme tento proces opakovat tak dlouho, dokud

(a) všechny složky $\bar{x}$ nebudou nulové – to by ale vzhledem k jeho přípustnosti bylo možné pouze v případě $b = 0$, což už jsme vyloučili ↯;

(b) vektory $\{a_j \mid j \in J(\bar{x})\}$ budou lineárně nezávislé, v kterémžto případě je důkaz hotov, neboť máme BPB bod splňující dokonce $J(\bar{x}) \subset J(x)$. ■

Už tedy víme, že pro libovolnou přípustnou množinu $X \neq \emptyset$ vždy existuje alespoň jeden BPB. Ale kolik jich je přesně? Kolik vlastně může mít daný polyedr/simplex extrémních bodů? Jde to vůbec *a priori* spočítat? Toto je jedna z naprosto klíčových otázek pro samotnou funkčnost simplexového algoritmu a k její odpovědi budeme ještě jiný (ekvivalentní) popis extrémních bodů z následujícího lemmatu. V něm využijeme značení $A_B := (a_j)_{j \in B}$ pro matici vzniklou z matice A vypuštěním sloupců s indexy mimo danou indexovou množinu[106] B. Podobně budeme používat značení x_B pro vektor vytvořený z x tak, že v něm vynecháme složky na pozicích s indexy $\{1,2,\ldots,n\} \setminus B$.

[106] Při takovéto interpretaci musí být indexy v množině B seřazeny vzestupně. Pro libovolné uspořádání prvků v B, pak A_B znamená matici, ve které první sloupec odpovídá sloupci matice A s index rovným prvnímu prvku v B, druhý sloupec je určen druhým indexem v B atd.

LEMMA 2.5.6

Nechť je dána úloha (LP). Bod x je BPB úlohy (LP) právě tehdy, když existuje m-prvková množina $B \subseteq \{1,2,\ldots,n\}$ taková, že

(i) čtvercová $m \times m$ matice A_B je regulární, tj. vektory $\{a_j \mid j \in B\}$ jsou lineárně nezávislé;

(ii) $x_j = 0$ pro $j \in \{1,2,\ldots,n\} \setminus B$.

Důkaz. „$\Longleftarrow$“ Je-li $x \in X$ a platí-li obě podmínky, pak zjevně $J(x) \subseteq B$, a tudíž vektory a_j jsou lineárně nezávislé i pro $j \in J(x)$, tj. x je BPB úlohy (LP).

„$\Longrightarrow$“ Nechť $x \in X$ je BPB úlohy (LP). Je-li $|J(x)| = m$, pak indexová množina $B := J(x)$ jistě splňuje obě dané podmínky. Je-li $|J(x)| < m$, pak množinu B vytvoříme postupným přidáváním

$m-|J(x)|$ indexů k $J(x)$ tak, že všechny vektory $\{a_j \mid j \in B\}$ budou lineárně nezávislé. To provedeme následujícím postupem. Začneme volbou $B := J(x)$. Jelikož rank $A = m \leq n$ (tj. existuje m lineárně nezávislých sloupců matice A), musí existovat (alespoň jeden) sloupec a_i, který nelze získat jako vhodnou lineární kombinaci sloupců $\{a_j \mid j \in B\}$. Tento index i přidáme do množiny B a celý krok zopakujeme s touto novou množinou B. Ve chvíli, kdy již takový krok není možný, tj. všechny sloupce matice A již jsou v lineárním obalu množiny $\{a_j \mid j \in B\}$ neboli patří do $\operatorname{span}\{a_j \mid j \in B\}$, pak tyto sloupce tvoří bázi množiny $\operatorname{span}\{a_j \mid j \in \{1,\dots,n\}\}$, takže nutně $|B| = m$, jak bylo potřeba. ■

Tato charakterizace nám dává velmi jednoduchý (ale zároveň velmi pracný a neefektivní) návod na hledání *bazických bodů*: stačí vzít m-prvkovou indexovou množinu $B \subseteq \{1,2,\dots,n\}$ určující regulární matici A_B a vypočítat $x_B := A_B^{-1} b$, což po doplnění nulovými hodnotami pro indexy z $\{1,2,\dots,n\} \setminus B$ dává bod $x \in \mathbb{R}^n$ vyhovující podmínkám z Lemmatu 2.5.6. Je-li navíc $x \geq 0$, pak $x \in X$ a máme dokonce BPB úlohy (LP). Zejména jestliže matice A obsahuje všechny sloupce $m \times m$ jednotkové matice (v libovolném pořadí), pak je nalezení alespoň jednoho BPB celkem snadné: vezmeme množinu B tvořenou indexy odpovídajícími sloupcům $e_i := (\underbrace{0,\dots,0}_{i-1},1,0,\dots,0)^\top$ pro $i \in \{1,\dots,m\}$ a položíme $x_j = 0$ pro $j \in \{1,2,\dots,n\} \setminus B$, zatímco pro $j \in B$ s odpovídajícím sloupcem e_i pro nějaké $i \in \{1,\dots,m\}$ bereme x_j jako i-tou složku vektoru b. Ovšem pro matici A_B tvořenou jinými vektory se může stát, že dodatečná podmínka $x_B \geq 0$ nebude splněna, takže pro danou indexovou množinu neexistuje žádný BPB, jak uvidíme v následujícím příkladu.

Příklad 2.5.7

Uvažme úlohu (LP) s množinou X vymezenou podmínkami

$$2x_1 + x_2 + x_3 = 1 \quad \& \quad x_1 + 3x_2 + x_4 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Určeme BPB pro indexové množiny $B = \{3,4\}$, $B = \{1,4\}$ a $B = \{1,2\}$.

Řešení. V tomto případě máme

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix} \quad \& \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Pro množinu $B = \{3,4\}$ máme $A_B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, takže $x_B = A_B^{-1} b = b$, tj. odpovídají BPB je

$$x = [0,0,1,2].$$

Pro množinu $B = \{1,4\}$ máme $A_B = \left(\begin{smallmatrix} 2 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, takže[107] $A_B^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ -1 & 2 \end{smallmatrix}\right)$. Proto $x_B = A_B^{-1} b = \left(\begin{smallmatrix} 1/2 \\ 3/2 \end{smallmatrix}\right)$, tj. odpovídají BPB je

$$x = [1/2,0,0,3/2].$$

Konečně pro $B = \{1,2\}$ máme $A_B = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 3 \end{smallmatrix}\right)$, takže $A_B^{-1} = \frac{1}{5}\left(\begin{smallmatrix} 3 & -1 \\ -1 & 2 \end{smallmatrix}\right)$. Proto $x_B = A_B^{-1} b = \left(\begin{smallmatrix} 1/5 \\ 3/5 \end{smallmatrix}\right)$, tj. odpovídají BPB je

$$x = [1/5,3/5,0,0].$$

Kdybychom upravili druhou rovnost v zadání tak, že by matice A tentokrát byla

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & -3 & 0 & 1 \end{pmatrix},$$

pak bychom pro $B = \{3,4\}$ dostali tentýž BPB, zatímco pro $B = \{1,2\}$ bychom měli $A_B = \left(\begin{smallmatrix} 2 & 1 \\ 1 & -3 \end{smallmatrix}\right)$ a $x_B = \left(\begin{smallmatrix} 5/7 \\ -3/7 \end{smallmatrix}\right)$, takže odpovídající bod je

$$x = [5/7,-3/7,0,0],$$

což je sice bazický bod, ovšem není přípustný! Tj. pro tuto bázi neexistuje BPB.

[107] Zde pouze stačí využít známého vzorce $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^{-1} = \frac{1}{ad-bc}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$.

V dalším tvrzení si pak vztah mezi množinami B a příslušnými BPB ještě upřesníme. Každopádně ale zdůrazněme, že m-prvkových podmnožin n-prvkové množiny existuje pouze konečný počet rovný binomickému číslu $\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$, což dává horní odhad počtu BPB úlohy (LP), a tudíž takových bodů je *pouze konečně mnoho*!

Věta 2.5.8

Nechť je dána úloha (LP). Množina všech BPB úlohy (LP) je jednoznačně určena indexovými množinami B z Lemmatu 2.5.6, tj. pro libovolnou m-prvkovou množinu $B \subseteq \{1,2,\ldots,n\}$ určující regulární matici A_B existuje nejvýše jeden BPB $x \in X$.

Důkaz. Nechť $B \subseteq \{1,2,\ldots,n\}$ je m-prvková množina, pro kterou odpovídající $m \times m$ matice A_B je regulární. Ukážeme, že existuje nejvýše jeden BPB $x \in X$ určený právě touto množinou B. Takový bod x musí splňovat $x_j = 0$ pro $j \in \{1,2,\ldots,n\} \setminus B$ a současně

$$b = Ax = \sum_{j \in B} x_j\, a_j + \sum_{i \in \{1,2,\ldots,n\} \setminus B} x_i\, a_i = \sum_{j \in B} x_j\, a_j = A_B\, x_B. \tag{2.5.3}$$

Ovšem tato soustava má díky regulárnosti matice A_B jediné řešení $x_B = A_B^{-1} b$, s jehož pomocí je bod $x \in \mathbb{R}^n$ odpovídající množině B určen jednoznačně. Pokud navíc $x_B \geq 0$, pak toto x je jediným BPB úlohy (LP) příslušným indexové množině B, zatímco v opačném případě pro takovou množinu B žádný BPB úlohy (LP) neexistuje. ■

Ovšem pozor, toto tvrzení o jednoznačném propojení mezi BPB úlohy (LP) a indexovými množinami $B \subseteq \{1,2,\ldots,n\}$ neříká vůbec nic o *jedno-jednoznačnosti* příslušných BPB, tj. různým indexovým množinám B může příslušet jeden a ten stejný BPB $x \in X$. Nicméně už samotné propojení mezi množinou B a BPB poukazuje na jistou *bazickou* vlastnost množiny B, která je umocněna i tím, že vektory $\{a_j \mid j \in B\}$ tvoří bázi $\mathbb{R}^m$, což nás přivádí k následující definici.

Definice 2.5.9

Nechť je dána úloha (LP). Množinu B popsanou v Lemmatu 2.5.6 nazveme *bází BPB* $x \in X$. V takovém případě hovoříme o tom, že indexy $j \in B$ *jsou v bázi bodu* x a indexy $j \in \{1,2,\ldots,n\} \setminus B$ *jsou mimo bázi bodu* x.

BPB $x \in X$ nazveme *nedegenerovaný*, jestliže $x_j > 0$ pro všechna $j \in B$, zatímco v případě existence alespoň jednoho indexu $j \in B$ s vlastností $x_j = 0$ hovoříme o *degenerovaném* bodě. Jsou-li všechny BPB úlohy (LP) nedegenerované, řekneme, že úloha (LP) je *nedegenerovaná*. V opačném případě nazveme úlohu (LP) *degenerovanou*.

BPB $x \in X$ je tedy nedegenerovaný právě tehdy, když $|J(x)| = m$ neboli $J(x) = B$. Navíc z definice ihned vyplývá, že nedegenerovaný BPB $x \in X$ je vždy jednoznačně určen množinou B, protože různé báze nemohou dát stejný nedegenerovaný BPB vzhledem k druhé podmínce v Lemmatu 2.5.6. V takovém případě pak množina $\{a_j \mid j \in B\}$ skutečně tvoří v jistém smyslu „bázi“ BPB $x \in X$. Samozřejmě zde ale nemá označení *báze* tentýž význam jako v lineární algebře, neboť tentokrát pro *různé* BPB potřebujeme *různé* množiny B. Je-li však BPB $x \in X$ degenerovaný, pak jej *může* být možné získat pomocí dvou (a více[108]) různých množin B. Každopádně platí, že úloha (LP) je degenerovaná právě tehdy, když vektor b lze vyjádřit jako lineární kombinaci s nezápornými koeficienty pomocí nejvýše $m-1$ sloupců matice[109] A. Vskutku, je-li úloha degenerovaná, pak existuje $x \in X$ a báze B tak, že $b = A_B\, x_B$ a $x_j = 0$ pro alespoň jedno $j \in B$, tj. stačí nám nejvýše $m-1$ sloupců matice A. Naopak, stačí-li nejvýše $m-1$ sloupců matice A, pak doplněním indexů na m-prvkovou množinu B při zachování lineární nezávislosti a nulovosti složek bodu x, dostaneme degenerovaný BPB úlohy (LP).

Jak „vypadají“ degenerované BPB? Uvažme úlohu lineárního programování s množinou X ve standardním tvaru, tj. vymezenou soustavou $Ax \leq b$ a podmínkou $x \geq 0$, což při přechodu ke kanonickému tvaru vede na soustavu m rovnic $Ax + Is = b$ pro $n+m$ nezáporných proměnných, která pro libovolný vektor $b \geq 0$ a číslo $m \in \mathbb{N}$ splňuje naše základní požadavky, zejména $\operatorname{rank}(A,\ I) = m$. Vezmeme-li např. $n = 2$, pak degenerované body vzniknou pouze průnikem více než 2 hraničních přímek, tj. pouze v případě existence nadbytečného omezení, které však

[108] Přesný počet je určen „mírou“ lineární (ne-)závislosti vektorů $\{a_j \mid j \in \{1,2,\ldots,n\} \setminus B\}$ doplněných o vektory a_i s indexy $i \in B$, pro něž $x_i = 0$.

[109] Případně naopak, úloha (LP) je nedegenerovaná právě tehdy, když vektor b lze vyjádřit pouze jako lineární kombinaci s kladnými koeficienty pomocí m-tice sloupců matice A.

nemusí být možné odstranit viz Obrázky 2.27–2.28. Pro $n \geq 3$ může být situace ještě komplikovanější a už rozhodně nemusí platit, že degenerovaný bod je důsledkem nadbytečného omezení, jak můžeme vidět např. na Obrázku 2.29. Vrchol či extrémní bod množiny X v $\mathbb{R}^n$ vznikne pouze jako jediný průnik nejméně n hraničních nadrovin v $\mathbb{R}^n$. V takovém případě je pak nedegenerovanost ekvivalentní s tím, že každým přípustným bodem prochází nejvýše n hraničních nadrovin[110,111]. Proto degenerovanost uvažované úlohy můžeme charakterizovat existencí přípustného bodu, v němž se protne více než n hraničních nadrovin.

[110] Jinými slovy, dvě různé n-tice nadrovin vymezujících původní množinu X se nemohou protnout v tomtéž přípustném bodě.

[111] Připusťme, že nějaký přípustný bod je jediným průnikem alespoň $n+1$ hraničních nadrovin a odpovídající úloha v kanonickém tvaru je nedegenerovaná. Tomuto bodu $\widetilde{x} \in X$ pak v kanonickém tvaru odpovídá $[\widetilde{x}, \widetilde{s}] \in \mathbb{R}_+^{n+m}$ s nejméně $n+1$ nulovými složkami (jedno každé nulové $\widetilde{s}_i$ za splnění jedné rovnosti v $A\widetilde{x} = b$ a zbývající počet nulových složek má samotné $\widetilde{x}$), tj. $|J(\widetilde{x}, \widetilde{s})| \leq m-1$. Současně je tento bod BPB, neboť soustava $A\widetilde{x} + I\widetilde{s} = b$ s $m-1$ neznámými ($n+1$ složek už je nulových) má jediné nezáporné řešení (protože $\widetilde{x}$ je jediným průnikem), takže z již dříve zmíněné Rouchéovy–Capelliho věty plyne lineární nezávislost sloupců matice (A, I) odpovídajících jednotlivým neznámým. Tím ale dostáváme spor s nedegenerovaností uvažované úlohy ↯.

Příklad 2.5.10

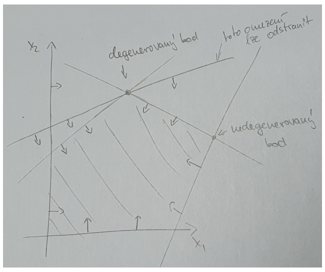
Úloha (LP) uvažovaná v Příkladu 2.5.7 i její modifikace byla nedegenerovaná, neboť vektor b nebyl násobkem žádného ze sloupců matice A (tj. je jej možné získat pouze jako lineární kombinaci nejméně dvou sloupců matice A). Degenerovanou úlohu získáme třeba v případě omezení

$$x_1 - 2x_2 - x_3 + 9x_4 + x_5 = 0$$
$$x_1 - 3x_2 - x_3 + 3x_4 + x_6 = 0 \quad \& \quad x_3 + x_7 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0$$
$$x_5 \geq 0 \quad \& \quad x_6 \geq 0 \quad \& \quad x_7 \geq 0.$$

Pak totiž máme

$$A = \begin{pmatrix} 1 & -2 & -1 & 9 & 1 & 0 & 0 \\ 1 & -3 & -1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \& \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

a BPB je např. $x = [0, \dots, 0, 1]$, pro který platí $J(x) = \{7\}$, zatímco $B = \{5, 6, 7\}$. Je také zřejmé, že pro vektor b nepotřebujeme tři sloupce matice A – stačí nám pouze jeden (ten poslední).



Obrázek 2.27: Množina X určená podmínkami $Ax \leq b$ a $x \geq 0$ v $\mathbb{R}^2$ s degenerovaným bodem, který je důsledkem nadbytečného omezení.

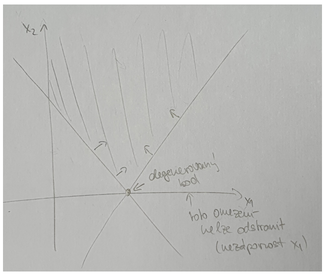
Jak už jsme se zmínili dříve, to, zda nějaký bod je či není BPB, vůbec nepotřebujeme účelovou funkci úlohy (LP), ale jde pouze o „geometrii" samotné přípustné množiny X. Proč se vůbec tomu tak podrobně věnujeme? Protože ve skutečnosti existuje velmi úzká souvislost mezi BPB a řešením úlohy (LP), jak nám ukazuje poslední tvrzení této podkapitoly, jehož důkaz můžeme brát i jako částečnou přípravu na simplexový algoritmus[112].

VĚTA 2.5.11

Jestliže je úloha (LP) řešitelná, pak existuje její BPB, který je řešením.

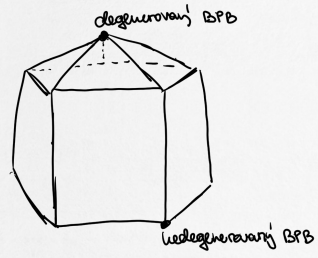
Tato věta ale neříká nic o situaci, kdy úloha (LP) nemá řešení, a rozhodně také neříká, že každé řešení úlohy (LP) je BPB či extrémní bod! Kdyby např. řešením byla celá hrana nějakého polyedru/simplexu, pak BPB jsou pouze její koncové body (vrcholy), ovšem ostatní body a současně řešení již BPB nejsou. Ovšem je-li úloha jednoznačně řešitelná, pak tím řešením nutně musí být nějaký BPB. Odtud a z Věty 2.5.8 dostáváme konečně krokový (ačkoli zcela nepraktický) algoritmus (využívající opravdu velmi hrubé síly) pro řešení úlohy (LP):

> stačí uvážit všechny m-prvkové podmnožiny $B \subseteq \{1, 2, \dots, n\}$, pro každou z nich najít příslušný BPB (pokud existuje), vyčíslit účelovou funkci v těchto bodech a konečně z těchto hodnot vybrat tu nejmenší (odpovídající bod pak je hledaným řešením).



Obrázek 2.28: Množina X určená podmínkami $Ax \leq b$ a $x \geq 0$ v $\mathbb{R}^2$ s degenerovaným bodem, který je sice důsledkem nadbytečného omezení, ale vzhledem k danému tvaru množiny X „nelze" toto omezení odstranit.

Jenže např. pro $n = 2m$ musíme uvážit $\binom{2m}{m} \approx 4^m$ podmnožin. Tato hodnota pro rostoucí m roste *exponenciálně* (tj. s každým zvýšením m o 1 dojde ke zčtyřnásobení počtu množin B), což je z praktického hlediska nepřijatelné. Navíc tento „algoritmus" dává správný výsledek pouze v případě řešitelné úlohy (LP), protože případná neohraničenost dané úlohy na jeho průběhu nic nemění. My ale směřujeme k mnohem efektivnějšímu algoritmu, ve kterém není



Obrázek 2.29: Množina X určená podmínkami $Ax \le b$ a $x \ge 0$ v $\mathbb{R}^3$ s degenerovaným bodem, který není důsledkem nadbytečného omezení.

procházení BPB náhodné a který funguje správně i pro neohraničené úlohy (tj. během výpočtu indikuje případnou neohraničenost účelové funkce na X).

Důkaz Věty 2.5.11. Důkaz si rozdělíme na několik kroků, které se v komplikovanější podobě objeví také v simplexové metodě. Samotné tvrzení Věty 2.5.11 je přímým důsledkem následujícího faktu:

> *Je-li účelová funkce úlohy* (LP) *zdola ohraničená na $X \neq \emptyset$, pak pro každé $x_0 \in X$ existuje BPB $\widetilde{x} \in X$ se stejnou nebo menší funkční hodnotou účelové funkce, tj. $\langle c, \widetilde{x}\rangle \le \langle c, x_0\rangle$.* } (∗)

Potom totiž konečný počet všech BPB úlohy (LP) zaručuje, že v některém z nich musí účelová funkce nabýt minimální hodnoty, a tudíž tento bod musí být řešením úlohy (LP).

Zbývá tak „pouze“ dokázat tvrzení (∗). Uvažme proto libovolné $x_0 \in X$. Mezi všemi $x \in X$ s funkčními hodnotami splňujícími $\langle c, x\rangle \le \langle c, x_0\rangle$ vybereme to, které má největší možný počet nulových složek, a označíme jej jako $\widetilde{x}$ (tento bod samozřejmě nemusí být určen jednoznačně). Jsou-li vektory $\{a_j \mid j \in J(\widetilde{x})\}$ lineárně nezávislé, je $\widetilde{x}$ BPB a my jsme hotovi. V opačném případě musí pro matici $\widetilde{A} := (a_j)_{j\in J(\widetilde{x})}$ existovat vektor[113] $\widetilde{v} \in \mathbb{R}^{|J(\widetilde{x})|}\setminus\{0\}$ takový, že $\widetilde{A}\widetilde{v} = 0$. Následným zkombinováním $\widetilde{v}$ s nulovými hodnotami na pozicích $\{1,2,\dots,n\}\setminus J(\widetilde{x})$ vytvoříme vektor $v \in \mathbb{R}^n$, pro který platí $Av = \widetilde{A}\widetilde{v} = 0$. Připusťme, že v navíc splňuje následující dvě podmínky, jejichž platnost si za chvíli ověříme v poslední části důkazu:

(i) $\langle c, v\rangle \le 0$;

(ii) existuje $j \in J(\widetilde{x})$ takové, že $v_j < 0$.

Pokud potom pro $t \ge 0$ definujeme vektor $x(t) := \widetilde{x} + tv$, tak pro něj vždy platí $x(t) \in X$, neboť $Ax(t) = A\widetilde{x} + tAv = A\widetilde{x} = b$. Navíc $x(0) = \widetilde{x}$ má nenulové (kladné) pouze složky s indexy $j \in J(\widetilde{x})$ a pro j-tou složku vektoru $x(t)$ máme $x_j(t) = \widetilde{x}_j + tv_j$, což v případě platnosti (ii) dává $x_j(t) < 0$ pro dostatečně velké $t > 0$. To znamená, že začneme-li s $t = 0$ a necháme-li t růst, pak hodnoty $x_j(t)$ s indexem odpovídajícím $v_j < 0$ budou klesat a v jistém okamžiku t_1 dosáhnou nulové hodnoty. V tomto okamžiku má ale $x(t_1)$ stále ještě nezáporné hodnoty[114], takže je přípustné a současně má o jednu nulovou složku více než $\widetilde{x}$. Ovšem díky vlastnosti (i) vektor $x(t_1)$ také splňuje

$$\langle c, x(t_1)\rangle = \langle c, \widetilde{x}\rangle + t_1\langle c, v\rangle \le \langle c, \widetilde{x}\rangle \le \langle c, x_0\rangle$$

čímž dostáváme spor s předpokladem, že $\widetilde{x}$ obsahuje největší možný počet nulových složek mezi všemi $x \in X$ splňujícími $\langle c, x\rangle \le \langle c, x_0\rangle$. Proto $\widetilde{x}$ je hledaným BPB.

Nyní již jen zbývá dokázat vlastnosti (i) a (ii). Pokud $\langle c, v\rangle = 0$, pak (i) je určitě splněna a vlastnost (ii) dostaneme případnou změnou všech znamének složek vektoru v, neboť $v \neq 0$. Nechť tedy dále $\langle c, v\rangle \neq 0$. Pak opět případnou změnou znamének (je-li tento součin kladný) dosáhneme $\langle c, v\rangle < 0$, tj. platí (i). Kdyby neplatila vlastnost (ii), tj. bylo by $v \ge 0$, pak $x(t) = \widetilde{x} + tv \ge 0$ pro všechna $t \ge 0$, a tedy $x(t) \in X$ pro všechna $t \ge 0$. Jenže hodnota účelové funkce v $x(t)$ je rovna $\langle c, x(t)\rangle = \langle c, \widetilde{x}\rangle + t\langle c, v\rangle$, takže díky $\langle c, v\rangle < 0$ máme $\langle c, x(t)\rangle \to -\infty$ pro $t \to \infty$, tj. účelová funkce úlohy (LP) není zdola ohraničená na X, což je spor s předpoklady tvrzení (∗). Vzhledem k libovolnosti x_0 a $\widetilde{x}$ je tímto celé tvrzení dokázáno. ■

[112] Podobné tvrzení v trochu jiné terminologii si s využitím zcela jiných nástrojů odvodíme mnohem později také v Kapitole 6, viz Větu 6.2.32.

[113] Ani tento vektor není určen jednoznačně, což ale vůbec nevadí.

[114] Indexy $j \in J(\widetilde{x})$, pro které $v_j \ge 0$, jsou zcela nezajímavé, neboť odpovídající $x_j(t)$ jsou neklesající pro rostoucí t. Pokud ale vlastnost (ii) platí pro více $j \in J(\widetilde{x})$, je pro splnění $x(t_1) \ge 0$ nutné uvažovat pouze index odpovídající nejmenšímu v_j, tj. index s vlastností

$$v_j = \min\{v_i \mid i \in \{1,\dots,n\}\} < 0.$$

Cvičení

2.5.1. Uvažme úlohu

$$f(x) \to \min$$
$$x_1 + x_2 \le 6 \quad \& \quad x_2 \le 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná?

2.5.2. Uvažte úlohu

$$f(x) \to \min$$
$$x_1 + x_2 \le 6 \quad \& \quad x_2 \le 3 \quad \& \quad x_1 + 2x_2 \le 9$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? (Z obrázku je zřejmé, že třetí podmínka je nadbytečná. Ovšem nikoli pouze systémy s redundantními rovnicemi jsou degenerované. Zkuste načrtnout příklady takových množin X v $\mathbb{R}^2$ a $\mathbb{R}^3$.)

2.5.3. Uvažte úlohu

$$f(x) \to \min$$
$$x_1 + 2x_2 \le 2.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? Určete řešení úlohy pro $f_1(x) = x_1 + x_2$ a $f_2(x) = -x_1 - 2x_2$.

Pro $f_1(x)$ neexistuje konečné řešení; pro $f_2(x)$ je řešením přímka spojující body $[2,0]$ a $[0,1]$.

2.5.4. Uvažte úlohu

$$3x_1 + x_2 \to \min / \max$$
$$x_1 + 2x_2 \le 4 \quad \& \quad 3x_1 + x_2 \le 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete její řešení.

2.5.5. Uvažte úlohu

$$x_1 + x_2 \to \min$$
$$2x_1 + x_2 \ge 4 \quad \& \quad x_1 + 2x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete její řešení.

2.5.6. Uvažte úlohu

$$-x_1 + 2x_2 \to \min$$
$$2x_1 + x_2 \le 4 \quad \& \quad 4x_1 + 4x_2 \le 12 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete její řešení.

2.5.7. Uvažte úlohu

$$3x_1 - 2x_2 \to \min$$
$$x_1 + 2x_2 \le 4 \quad \& \quad 2x_1 + x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Převeďte úlohu na kanonický tvar. Určete všechny BPB a znázorněte je v rovině x_1x_2. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete její řešení.

2.5.8. Uvažte úlohu

$$f(x) \to \min$$
$$x_1 + x_2 - 8x_3 + 3x_4 = 2 \quad \& \quad -x_1 + x_2 + x_3 - 3x_4 = 2.$$

(a) Určete nějaké nebazické přípustné řešení, pokud existuje.

(b) Určete BPB s bází $B = \{1,4\}$, pokud existuje.

(c) Existuje BPB, který by odpovídal dvěma různým bázím? Pokud ano, určete jej. Je úloha degenerovaná?

(d) Určete všechny nedegenerované BPB.

2.5.9. Uvažte úlohu

$$120x_1 + 100x_2 \to \min$$
$$2x_1 + 2x_2 + x_3 = 8 \quad \& \quad 5x_1 + 3x_2 + x_4 = 15$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Určete všechny BPB. Je úloha řešitelná? Pokud ano, určete řešení. Interpretujte tuto úlohu geometricky.

2.5.10. Uvažte úlohu

$$f(x) \to \min$$
$$3x_1 + x_3 + x_4 = 5 \quad \& \quad 2x_1 + x_2 = 3 \quad \& \quad 4x_1 + 3x_3 + x_5 = 6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0$$

a pětici bodů $A = [0,3,0,5,6]$, $B = [0,3,5,0,-9]$, $C = [3/2,0,0,1/2,0]$, $D = [1/2,1,1,0,2]$ a $E = [1,1,1/2,3/2,1/2]$. Který za nich je

(a) přípustným bodem,

(b) BPB,

(c) krajním bodem množiny X?

2.5.11. Uvažte úlohu

$$f(x) \to \min$$
$$-x_1 + 2x_2 \leq 5 \quad \& \quad 0 \leq x_1 \leq 4 \quad \& \quad -1 \leq x_2 \leq 4.$$

Určete vrcholy množiny X a BPB v odpovídající úloze v kanonickém tvaru. Je úloha degenerovaná?

2.5.12. Uvažte úlohu

$$f(x) \to \min$$
$$x_1 - x_2 - 2x_3 \leq 1 \quad \& \quad -3x_1 - x_3 + 2x_4 \leq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Určete všechny BPB v odpovídající úloze v kanonickém tvaru. Je úloha degenerovaná?

2.5.13. Uvažte úlohu

$$f(x) \to \min$$
$$x_1 + x_2 + 2x_3 \leq 8 \quad \& \quad x_2 + 6x_3 \leq 12$$
$$x_1 \leq 4 \quad \& \quad x_2 \leq 6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Určete všechny BPB v odpovídající úloze v kanonickém tvaru. Je úloha degenerovaná?

2.5.14. Uvažte úlohu

$$f(x) \to \min$$
$$2x_1 + 6x_2 + 2x_3 + x_4 = 3 \quad \& \quad 6x_1 + 4x_2 + 4x_3 + 6x_4 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Určete všechny BPB. Je úloha degenerovaná?

2.5.15. Uvažte úlohu

$$-2x_1 - 3x_2 + x_3 - 4x_4 - x_5 \to \min$$
$$2x_1 + x_3 + 4x_4 + 2x_5 = 20 \quad \& \quad x_1 + x_2 - x_3 + x_4 + 3x_5 = 10$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0.$$

Určete všechny BPB. Je úloha degenerovaná?

2.5.16. Uvažte úlohu

$$2x_1 + x_2 - x_3 \to \max$$
$$2x_1 + x_2 + 4x_3 \le 6 \quad \& \quad x_1 + 4x_2 - x_3 \le 4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Určete všechny BPB. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete řešení.

2.5.17. Uvažte úlohu

$$2x_1 + x_2 - x_3 \to \max$$
$$2x_1 + x_2 - 4x_3 \le 6 \quad \& \quad x_1 + 4x_2 - x_3 \le 4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Určete vrcholy všechny BPB. Je úloha degenerovaná? Je úloha řešitelná? Pokud ano, určete řešení.

2.5.18. Uvažte úlohu

$$5x_1 + 3x_2 + 3x_3 + x_4 \to \min$$
$$x_1 + x_2 - 2x_3 + 3x_4 = 2 \quad \& \quad -2x_1 + x_3 = 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

(a) Ukažte, že existují právě dva BPB (geometricky i početně).

(b) Určete tyto BPB.

(c) Ukažte, že účelová funkce je ohraničená zdola na X.

(d) Určete řešení úlohy.

2.6 Dualita aneb peníze až na prvním místě

Teď už sice máme vše potřebné k vybudování již několikrát avizovaného algoritmu, ale předtím si ještě uděláme krátkou a z praktického pohledu velmi důležitou odbočku. Na úlohy lineárního programování se dá totiž podívat i z trochu jiného úhlu pohledu. Co tím máme na mysli? Než si to ukážeme na samotné úloze lineárního programování, uvažme klasickou soustavu lineárních rovnic

$$Bx = f \tag{2.6.1}$$

pro nějakou matici B a vektor f (vhodných rozměrů) a připusťme, že nás v tuto chvíli nezajímá celé řešení soustavy (2.6.1), ale pouze nějaká jeho část určená vektorem g, tj.

$$\Phi(f) := g^\top x.$$

Např. pro $g = e_i$ dostaneme i-tou složku řešení nebo volba $g = (1,\dots,1)^\top$ dává součet složek celého řešení atd. Jak ale závisí hodnota $\Phi(f)$ na samotném f? Co se stane při změně f? Uvážíme-li místo soustavy (2.6.1) úlohu

$$B^\top v = g, \tag{2.6.2}$$

pak platí

$$\Phi(f) = g^\top x = (B^\top v)^\top x = v^\top Bx = v^\top f.$$

Tedy se znalostí řešení nové soustavy (2.6.2) máme jasně danou závislost hodnoty $\Phi(f)$ na pravé straně f. Proto úlohu (2.6.2) můžeme považovat za tzv. *duální*[115] (též *adjungovanou*) k původní soustavě (2.6.1). Jak ji ale získat? V tuto chvíli to byl opět spíše *Deus ex machina*, ale v dalším příkladu to již uvidíme se všemi podrobnostmi.

[115] Označení „duální“ je klasický matematický termín, se kterým se můžeme setkat v řadě oblastí. Jak ovšem označit původní úlohu? Tuto otázku si položil William Orchard–Hays někdy kolem roku 1954 a otec zakladatele lineárního programování, Tobias Dantzig, navrhl označení „primární“. Přirozené? Možná...

Příklad 2.6.1

Uvažme nyní úlohu lineárního programování

$$4x_1 + 14x_2 \to \min$$
$$x_1 + 2x_2 \geq 3 \quad \& \quad x_1 + 4x_2 \geq 4 \quad \& \quad 3x_1 + x_2 \geq 3$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

S pomocí náčrtku celkem snadno zjistíme, že řešením je bod $x^* = [2, 1/2]$ s hodnotou $f^* = 15$. O to nám ale v tuto chvíli nejde. Aniž bychom tuto úlohu vyřešili, můžeme pouze s využitím zadání odhadnout hodnotu f^*. Má-li totiž x_1, x_2 být řešením, pak z prvního omezení a podmínky nezápornosti plyne

$$4x_1 + 14x_2 = 4\left(x_1 + 3\tfrac{1}{2}x_2\right) \geq 4(x_1 + 2x_2) \geq 4 \cdot 3 = 12,$$

tj. $f^* \geq 12$. Můžeme dostat i lepší odhad? Ano, druhé a třetí omezení dává

$$4x_1 + 14x_2 \geq 3(x_1 + 4x_2) + \tfrac{1}{3}(3x_1 + x_2) \geq 3 \cdot 4 + \tfrac{1}{3} \cdot 3 = 13,$$

tj. $f^* \geq 13$, a z prvního a druhého omezení dokonce dostaneme

$$4x_1 + 14x_2 \geq 3(x_1 + 4x_2) + 1(x_1 + 2x_2) \geq 3 \cdot 4 + 1 \cdot 3 = 15,$$

tj. $f^* \geq 15$. Jak dobrý odhad vůbec můžeme takovým způsobem získat? A co to vlastně znamená „takovým způsobem“? Začněme od konce. Pomocí jednotlivých omezení se snažíme odvodit nerovnost ve tvaru

$$d_1 x_1 + d_2 x_2 \geq h, \tag{2.6.3}$$

kde $d_1 \leq 4$, $d_2 \leq 14$ a číslo h je největší možné. Pak totiž můžeme tvrdit, že pro všechna $x_1, x_2 \geq 0$ máme

$$4x_1 + 14x_2 \geq d_1 x_1 + d_2 x_2 \geq h,$$

a tudíž h je nějaká dolní hranice minimální hodnoty účelové funkce na X, tj. dolní odhad čísla f^*. Jak takovou nerovnost ale získat? Zkombinujeme všechny tři nerovnosti do tvaru podobnému nerovnosti výše s nějakými nezápornými koeficienty y_1, y_2, y_3, přičemž podmínka nezápornosti je potřeba kvůli správnému využití nerovnostních podmínek ze zadání. Potom

$$y_1(x_1 + 2x_2) + y_2(x_1 + 4x_2) + y_3(3x_1 + x_2) \geq 3y_1 + 4y_2 + 3y_3,$$

což po úpravě výrazu na levé straně vede k nerovnosti

$$(1 \cdot y_1 + 1 \cdot y_2 + 3 \cdot y_3)x_1 + (2 \cdot y_1 + 4 \cdot y_2 + 1 \cdot y_3)x_2 \geq 3y_1 + 4y_2 + 3y_3,$$

která odpovídá (2.6.3) s volbou

$$d_1 := y_1 + y_2 + 3y_3, \quad d_2 := 2y_1 + 4y_2 + y_3, \quad h := 3y_1 + 4y_2 + 3y_3.$$

Už jen zbývá zajistit, aby h bylo největší možné a d_1, d_2 splňovala daná omezení. To nás přivádí k nové úloze lineárního programování

$$3y_1 + 4y_2 + 3y_3 \to \max$$
$$y_1 + y_2 + 3y_3 \leq 4 \quad \& \quad 2y_1 + 4y_2 + y_3 \leq 14$$
$$y_1 \geq 0 \quad \& \quad y_2 \geq 0 \quad \& \quad y_3 \geq 0.$$

Toto je hledaná *duální úloha* k úloze (2.6.1). Tato úloha ohraničuje původní problém zdola v tom smyslu, že každý přípustný bod $[y_1, y_2, y_3]$ duální úlohy dává dolní hranici minima účelové funkce původní úlohy.

Podobným způsobem bychom mohli ukázat, že duální úlohou k obecné úloze lineárního programování[116]

$$\langle c, x\rangle \to \min$$
$$\sum_{j=1}^{n} a_{ij}\, x_j \geq b_i, \quad i = 1,\dots,k$$
$$\sum_{j=1}^{n} a_{ij}\, x_j = b_i, \quad i = k+1,\dots,m$$
$$x_j \geq 0, \quad j = 1,\dots,s$$

je maximalizační úloha lineárního programování tvaru

$$\langle y, b\rangle \to \max$$
$$\sum_{i=1}^{m} a_{ij}\, y_i \leq c_j, \quad j = 1,\dots,s$$
$$\sum_{i=1}^{m} a_{ij}\, y_i = c_j, \quad j = s+1,\dots,n$$
$$y_i \geq 0, \quad i = 1,\dots,k,$$

viz také Příklad 8.3.7. Odtud také vidíme, že vztah mezi primární a duální úlohou je v jistém smyslu *reciproční*. Přepíšeme-li totiž tuto duální úlohu do tvaru odpovídajícího primární úloze výše, pak k ní duální úloha bude shodná s původní úlohou, tj. duální úloha k duální úloze je primární úloha. Proto bychom zde mohli hovořit spíše o *duálně sdružených úlohách* či *dvojici vzájemně duálních úloh*. Tuto recipročnost duality rozhodně obecné úlohy matematického programování nemají a toto je jedna z vlastností, která činí úlohy lineárního programování výjimečnými. Bylo by možné, aby primární a duální úloha byly shodné, tj. aby daná úloha byla *samoduální*[117]? Ano, vezmeme-li $m = n$, $s = k$, $c = -b$ a antisymetrickou matici $A = -A^\top$, pak se snadno můžeme přesvědčit, že obě úlohy splývají[118,119].

Obzvláště duální úlohou k úloze (LP) v kanonickém tvaru je úloha v základním tvaru, tj.

$$\langle y, b\rangle \to \max \quad \& \quad y \in Y := \{y \in \mathbb{R}^m \mid A^\top y \leq c\}. \tag{LP$_\mathrm{D}$}$$

Obě úlohy (LP) a (LP$_\mathrm{D}$) můžeme navíc současně zapsat do velmi elegantní a symetrické *Tuckerovy tabulky*

	$x_1 \geq 0$	$x_2 \geq 0$	$\cdots$	$x_n \geq 0$	
y_1	a_{11}	a_{12}	$\cdots$	a_{1n}	$= b_1$
y_2	a_{21}	a_{22}	$\cdots$	a_{2n}	$= b_2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_m	a_{m1}	a_{m2}	$\cdots$	a_{mn}	$= b_m$
	$\leq c_1$	$\leq c_2$	$\cdots$	$\leq c_n$	

Tabulka 2.4: Úlohy (LP) a (LP$_\mathrm{D}$) v Tuckerově tabulce.

v níž jsou omezení pro primární úlohu „zapsána“ v řádcích a omezení pro duální úlohu ve sloupcích. V případě zcela obecné úlohy lineárního programování dostaneme pro primární a duální úlohu následující Tabulku 2.5.

[116] Nepřehlédněme opačná znaménka nerovnosti v první skupině omezení oproti klasické úloze (2.2.1)!

[117] Angl. *self-dual*.

[118] Jakou úlohu bychom získali, kdyby matice A byla pouze symetrická, tj. $A^\top = A$?

[119] Dá se také ukázat, že libovolná úloha lineárního programování může být transformována do samoduální úlohy s $n+m$ proměnnými a $n+m$ omezeními, přičemž prvních n složek řešení této samoduální úlohy je řešením původního problému a dalších m složek je řešením odpovídající duální úlohy.

Primární úloha									
Duální úloha		$x_1 \geq 0$	$x_2 \geq 0$	$\cdots$	$x_s \geq 0$	x_{s+1}	$\cdots$	x_n	
	$y_1 \geq 0$	a_{11}	a_{12}	$\cdots$	a_{1s}	$a_{1,s+1}$	$\cdots$	a_{1n}	$\geq b_1$
	$y_2 \geq 0$	a_{21}	a_{22}	$\cdots$	a_{2s}	$a_{2,s+1}$	$\cdots$	a_{2n}	$\geq b_2$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
	$y_k \geq 0$	a_{k1}	a_{m2}	$\cdots$	a_{ks}	$a_{k,s+1}$	$\cdots$	a_{kn}	$\geq b_k$
	y_{k+1}	$a_{k+1,1}$	a_{22}	$\cdots$	$a_{k+1,s}$	$a_{k+1,s+1}$	$\cdots$	$a_{k+1,n}$	$= b_{k+1}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	y_m	a_{m1}	a_{m2}	$\cdots$	a_{ms}	$a_{m,s+1}$	$\cdots$	a_{mn}	$= b_m$
		$\leq c_1$	$\leq c_2$	$\cdots$	$\leq c_s$	$= c_{s+1}$	$\cdots$	$= c_n$	$y^\top b \to \max$ / $c^\top x \to \min$

Tabulka 2.5: Obecná dvojice vzájemně duálních úloh lineárního programování v Tuckerově tabulce.

Dříve zmíněné ohraničení dané duální úlohou je obsahem tzv. *Slabé věty o dualitě*. Ačkoli je tento výsledek velmi speciálním případem obecného tvrzení z Podkapitoly 8.3, my si jej ukážeme pro úlohy (LP) a (LP_D) i společně s jeho důkazem, který je poměrně snadný.

Věta 2.6.2 Slabá věta o dualitě

Pro každý přípustný bod $x \in X$ úlohy (LP) a každý přípustný bod $y \in Y$ duální úlohy (LP_D) platí

$$\langle c, x\rangle \geq \langle y, b\rangle. \tag{2.6.4}$$

Jinými slovy, je-li y přípustný bod pro (LP_D), pak hodnota $\langle y, b\rangle$ udává dolní hranici minima účelové funkce úlohy (LP). Zejména je-li (LP) neohraničená (zdola), úloha (LP_D) musí být nepřípustná. Naopak, je-li úloha (LP_D) neohraničená (shora), pak primární úloha (LP) musí být nepřípustná.

Důkaz. Jsou-li $x \in X$ a $y \in Y$, pak díky nezápornosti x a ostatním podmínkám vymezujícím X a Y nutně platí

$$\langle c, x\rangle \geq \langle A^\top y, x\rangle = y^\top A x = \langle y, b\rangle. \tag{2.6.5}$$

Druhá část tvrzení pak je pouze přímým důsledkem rovností $\inf \emptyset = \infty$ a $\sup \emptyset = -\infty$. Je-li totiž primární úloha (LP) zdola neohraničená (tj. $\inf_{x\in X}\langle c, x\rangle = -\infty$), pak nutně $\langle y, b\rangle = -\infty$ pro všechna $y \in Y$, což je možné pouze pro $Y = \emptyset$ neboli duální úloha (LP_D) musí být nepřípustná. ■

Nerovnost (2.6.4) ve skutečnosti ukazuje, že omezení hodnot primární a duální úlohy je vzájemné, tj. hodnota primární úlohy nemůže být menší než hodnota duální úlohy a zároveň naopak. Ovšem jak jsou tyto dvě hodnoty „vzdálené“ neboli jaký je tzv. *duální rozdíl*[120]? Nulový nebo může být i nenulový? Odpověď na tuto otázku dává tzv. *Silná věta o dualitě*, ale ještě před ní si můžeme zformulovat jeden velmi jednoduchý (i když ne příliš praktický) důsledek plynoucí ihned z nerovnosti (2.6.4). Pokud se nám totiž podaří najít dvojici bodů $x^* \in X$ a $y^* \in Y$ takových, že hodnoty odpovídajících účelových funkcí jsou stejné, tj. $\langle c, x^*\rangle = \langle y^*, b\rangle$, pak jsme nalezli dokonce řešení obou úloh, neboť účelové funkce v těchto dvou bodech nabývají své největší a nejmenší hodnoty, kterou mohou nabýt mezi všemi $x \in X$ a $y \in Y$. Z důkazu Věty 2.6.2 také snadno můžeme vidět[121], že se to stane právě tehdy, když x^* a y^* jsou taková, že platí *podmínka komplementarity*

$$\langle x^*, c - A^\top y^*\rangle = 0$$

neboli vzhledem k nezápornosti složek vektorů x^* a $c - A^\top y^*$ musí pro každé $i \in \{1, \ldots, n\}$ být nulová buď i-tá složka vektoru x^* nebo vektoru[122] $c - A^\top y^*$. To znamená, že v případě $x_i^* > 0$ pro nějaké $i \in \{1, \ldots, n\}$, se odpovídající i-tá nerovnost v duálním omezení pro y^* realizuje

[120] Angl. *duality gap*.

[121] Viz nerovnost (2.6.5), která v případě rovnosti výrazů na levé a pravé straně zjevně vede k podmínce $\langle c, x\rangle = y^\top A x$.

[122] Pochopitelně není vyloučena ani nulovost i-tých složek obou vektorů současně.

jako rovnost. Naopak, pokud se i-té duální omezení realizuje pro y^* jako ostrá nerovnost, pak nutně $x_i^* = 0$.

Důsledek 2.6.3 Certifikát optimality[123]

Pokud pro nějaký přípustný bod $x^* \in X$ úlohy (LP) a nějaký přípustný bod $y^* \in Y$ úlohy (LP_D) nastane v (2.6.4) rovnost, tj. $\langle c, x^* \rangle = \langle y^*, b \rangle$, pak x^* a y^* jsou řešeními „svých“ úloh.

[123] Někdy se také hovoří o *kontrolorově* (*šéfově*) *principu* (angl. *supervisor's principle*), které „kontrolorovi“ skýtá velmi účinný (i když trochu časově náročný) nástroj pro potvrzení, že se skutečně podařilo najít požadované řešení – stačí vyřešit i duální úlohu a porovnat odpovídající funkční hodnoty.

Pro každou úlohu matematického programování mohou nastat pouze tři možnosti. Úloha může být

(i) přípustná a ohraničená (tj. řešitelná);
(ii) nepřípustná ($X = \emptyset$);
(iii) neohraničená ($\inf_{x \in X} \langle c, x \rangle = -\infty$ či $\sup_{y \in Y} \langle y, b \rangle = \infty$).

To dohromady dává celkem 9 možností pro dvojici úloh (LP) a (LP_D). Tři z nich nejsou možné kvůli Slabé větě o dualitě (neohraničenost obou úloh i neohraničenost jedné úlohy a současně řešitelnost druhé úlohy) a stejné tvrzení potvrzuje dvě možnosti, jak je uvedeno v Tabulce 2.6. Zbývající čtyři možnosti zůstávají v tuto chvíli otevřené.

PÚ / DÚ	NP	PaO	NO
NP	?	?	✔
PaO	?	?	✖
NO	✔	✖	✖

Tabulka 2.6: Vztah primární a duální úlohy dle Věty 2.6.2. PÚ: primární úloha; DÚ: duální úloha; NP: nepřípustná úloha; NO: neohraničená úloha; PaO: přípustná a ohraničená úloha (tj. existuje řešení).

Avšak pro doplnění Tabulky 2.6 potřebujeme silnější výsledek než ten, který nám poskytuje Věta 2.6.2. Jde o tzv. *Silnou větu o dualitě*, k jejímuž důkazu využijeme následující tvrzení, které víceméně ihned vyplývá z Vět 2.6.2 a 2.4.6

Věta 2.6.4 O dualitě

Je-li $X \neq \emptyset$ a $Y \neq \emptyset$, pak úlohy (LP) a (LP_D) mají řešení.

Důkaz. Jsou-li obě úlohy (LP) a (LP_D) přípustné, pak podle Věty 2.6.2 má primární úloha zdola ohraničenou účelovou funkci a duální úloha shora ohraničenou účelovou funkci, takže Věta 2.4.6 zaručuje jejich řešitelnost. ■

Nyní ukážeme, že primární a duální úloha musí být přípustné (a tedy i řešitelné) současně. V důkazu sice odvodíme pouze jednu část tohoto tvrzení, avšak ta druhá je v podstatě pouze přímým důsledkem toho, že duální úloha je také úlohou lineárního programování. Díky tomu můžeme zaměnit roli obou úloh, takže Tabulka 2.6 musí být po doplnění symetrická.

Věta 2.6.5 Silná věta o dualitě

Úloha (LP) má řešení právě tehdy, když úloha (LP_D) má řešení. V takovém případě je mezi těmito úlohami nulový duální rozdíl, tj. $\inf_{x \in X} \langle c, x \rangle = \sup_{y \in Y} \langle y, b \rangle$.

Důkaz. Dokážeme implikaci pouze jedním směrem, neboť opačné tvrzení lze získat analogickými argumenty či s pomocí vhodné transformace zaměnit roli obou úloh. Nechť tedy primární úloha (LP) je řešitelná, tj. existuje $x^* \in \mathbb{R}^n$ takové, že $Ax^* = b$, $x^* \geq 0$ a $f^* := \langle c, x^* \rangle = \inf_{x \in X} \langle c, x \rangle$. Připusťme, že duální úloha je nepřípustná, tj. neexistuje $y \in \mathbb{R}^m$ splňující $A^\top y \leq c$. Pak záměnou A a $A^\top$, x a y i b a c v Důsledku 2.4.5 odtud vyplývá, že existuje $\widetilde{x} \geq 0$ splňující $A\widetilde{x} = 0$ a $\langle \widetilde{x}, c \rangle < 0$. Jenže v takovém případě je $x^* + \alpha \widetilde{x} \in X$ pro libovolné $\alpha \geq 0$ s funkční hodnotou $\langle c, x^* + \alpha \widetilde{x} \rangle = f^* + \alpha \langle \widetilde{x}, c \rangle \to -\infty$ pro $\alpha \to \infty$, což je spor s řešitelností úlohy (LP) ↯. Proto je duální úloha přípustná, a podle Věty 2.6.4 také řešitelná. Zbývá ukázat nulový duální

rozdíl. To, že hodnota primární úlohy je f^*, znamená, že neexistuje $x \geq 0$ splňující $Ax = b$ a $x^\top c < -f^*$ neboli řešení soustavy

$$\begin{pmatrix} A & -b \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0 \quad \& \quad \begin{pmatrix} x \\ 1 \end{pmatrix}^\top \begin{pmatrix} c \\ -f^* \end{pmatrix} < 0 \quad \& \quad \begin{pmatrix} x \\ 1 \end{pmatrix} \geq 0.$$

Pak ale z Důsledku 2.4.5 vyplývá existence y^* takového, že

$$\begin{pmatrix} A^\top \\ -b^\top \end{pmatrix} y^* \leq \begin{pmatrix} c \\ -f^* \end{pmatrix}$$

neboli $A^\top y^* \leq c$ a $b^\top y^* \geq f^*$. Jelikož ze Slabé věty o dualitě současně vyplývá $f^* \geq b^\top y$ pro každé $y \in Y$, musí být nalezené y^* řešením duální úlohy a platí $f^* = b^\top y^*$, čímž je důkaz hotov. ■

Nulovost duálního rozdílu z předchozí věty mimo jiné v případě řešitelné úlohy (LP) znamená, že musí pro x^* a y^* být vždy splněna podmínka komplementarity. V takovém případě pak platí

$$0 = \langle x^*, c - A^\top y^* \rangle = \sum_{j \in B} x_j^* (c_j - a_j^\top y^*) \quad \text{neboli} \quad y^* = A_B^{\top -1} c_B, \tag{2.6.6}$$

což si ilustrujeme na následujícím příkladě. Ještě poznamenejme, že tato rovnost platí bez ohledu na případnou degenerovanost primární úlohy (LP). V takovém případě je sice $x_j = 0$ pro nějaké $j \in B$, a tudíž hodnota $c_j - a_j^\top y^*$ neovlivní nulovost skalárního součinu[124] v (2.6.6), ale vezmeme-li $y^* := A_B^{\top -1} c_B$, pak určitě platí

[124] Mohla by proto být libovolná nezáporná.

$$\langle y^*, b \rangle = \langle c_B, A_B^{-1} b \rangle = \langle c_B, x_B^* \rangle = \langle c, x^* \rangle,$$

tj. podle Důsledku 2.6.3 je toto y^* řešením úlohy ($\mathrm{LP_D}$). Navíc tento vztah je jistě recipročný, takže ze znalosti řešení úlohy ($\mathrm{LP_D}$) lze určit řešení primární úlohy (LP), k čemuž ale není potřeba zaměňovat role primární a duální úlohy, tj. převádět úlohu ($\mathrm{LP_D}$) do kanonického tvaru. Složky x^* s indexy odpovídajícími omezením v ($\mathrm{LP_D}$), které se pro y^* realizují jako ostré nerovnosti, musí být s souladu s podmínkami komplementarity nulové. To nastane alespoň pro $n - m$ indexů, takže tím se soustava původních m rovností o n neznámých redukuje na soustavu m rovností o nejvýše m neznámých, jejíž řešení pak je současně i řešením (LP). Jinými slovy, je-li $\widehat{B}$ množina indexů omezení duální úlohy realizujících se pro y^* jako rovnosti, pak pro libovolné $x_{\widehat{B}}^*$ splňující $A_{\widehat{B}} x_{\widehat{B}}^* = b$ platí

$$\langle y^*, b \rangle = \langle y^*, A_{\widehat{B}} x_{\widehat{B}}^* \rangle = \langle A_{\widehat{B}}^\top y^*, x_{\widehat{B}}^* \rangle = \langle c_{\widehat{B}}, x_{\widehat{B}}^* \rangle = \langle c, x^* \rangle,$$

tj. podle Důsledku 2.6.3 je odpovídající x^* řešením primární úlohy (LP).

Příklad 2.6.6

Uvažme úlohu lineárního programování

$$2x_1 + 3x_2 + 5x_3 + 2x_4 + 3x_5 \to \min$$
$$x_1 + x_2 + 2x_3 + x_4 + 3x_5 = 4 \quad \& \quad 2x_1 - 2x_2 + 3x_3 + x_4 + x_5 = 3$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0.$$

S pomocí jejího řešení $x^* = [1,0,0,0,1]$ určeme řešení y^* duální úlohy a naopak, tj. s pomocí nalezeného y^* dopočítejte řešení původní úlohy.

Řešení. Duální úloha je

$$4y_1 + 3y_2 \to \max$$
$$y_1 + 2y_2 \leq 2 \quad \& \quad y_1 - 2y_2 \leq 3 \quad \& \quad 2y_1 + 3y_2 \leq 5$$
$$y_1 + y_2 \leq 2 \quad \& \quad 3y_1 + y_2 \leq 3.$$

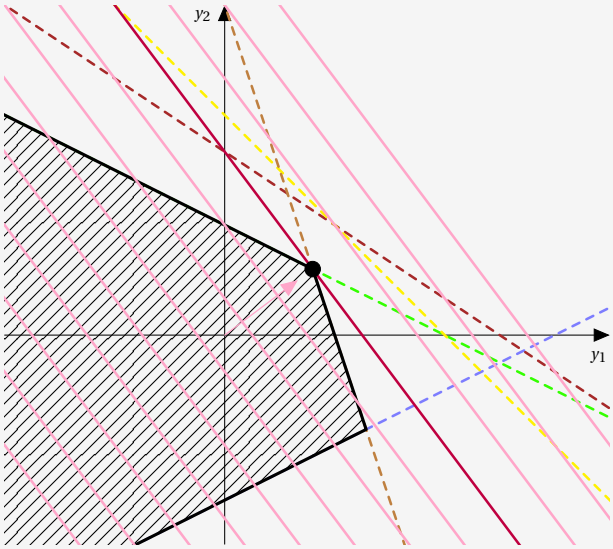
Protože $B(x^*) = \{1,5\}$, je podle (2.6.6) jejím řešením

$$y^* = A_B^{\top -1} c_B = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = -\frac{1}{5} \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}.$$

To si můžeme snadno ověřit z Obrázku 2.30 a také z Důsledku 2.6.3, neboť y^* vyhovuje všem duálním omezením a současně

$$\langle y^*, b \rangle = 5 = \langle c, x^* \rangle.$$

Všimněme si také, že první a poslední duální omezení se v bodě y^* realizují jako rovnosti, zatímco v ostatních omezeních nastávají ostré nerovnosti, což je zcela v souladu s platností podmínek komplementarity.



Obrázek 2.30: Vizualizace duální úlohy z Příkladu 2.6.6.

Naopak, pro $y^* = [4/5, 3/5]$ se první a poslední duální omezení realizují jako rovnosti, tj. $\widehat{B} = \{1,5\}$, zatímco v ostatních omezeních nastávají ostré nerovnosti. Proto $x_2^* = x_3^* = x_4^* = 0$ a pro ostatní složky x^* se původní omezení redukují na soustavu $A_{\widehat{B}} x^*_{\widehat{B}} = b$, tj.

$$\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1^* \\ x_5^* \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad \text{neboli} \quad \begin{pmatrix} x_1^* \\ x_5^* \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

a tedy řešením primární úlohy skutečně je $x^* = [1,0,0,0,1]$.

My jsme se v našich předchozích úvahách bez újmy na obecnosti omezili pouze na dvojici úloh (LP) a ($\mathrm{LP_D}$). Avšak stejné výsledky zůstávají v platnosti i pro libovolnou jinou dvojici duálně sdružených úloh lineárního programování. To je pro přehlednost shrnuto v následujícím důsledku a také v Tabulce 2.7.

DŮSLEDEK 2.6.7

Pro libovolnou dvojici duálně sdružených úloh lineárního programování nastane právě jedna z následujících možností:

(i) obě úlohy jsou nepřípustné;

(ii) primární úloha je neohraničená a duální úloha je nepřípustná;

(iii) primární úloha je nepřípustná a duální úloha je neohraničená;

(iv) obě úlohy jsou přípustné a ohraničené. Pak existují jejich řešení $x^* \in X$ a $y^* \in Y$, která navíc splňují

$$\langle c, x^* \rangle = \langle y^*, b \rangle, \tag{2.6.7}$$

Důsledek 2.6.7

tj. minimum dosažené v primární úloze je stejné jako maximum dosažené v duální úloze neboli hodnoty obou úloh jsou stejné.

PÚ / DÚ	NP	PaO	NO
NP	✔	✖	✔
PaO	✖	✔	✖
NO	✔	✖	✖

Tabulka 2.7: Vztah primární a duální úlohy lineárního programování. PÚ: primární úloha; DÚ: duální úloha; NP: nepřípustná úloha; NO: neohraničená úloha; PaO: přípustná a ohraničená úloha (tj. existuje řešení).

Naše úvahy ohledně duální úlohy ale rozhodně nejsou samoúčelné. Rovnost (2.6.7) nám mimo jiné umožňuje se místo řešení primární úlohy zaměřit na duální úlohu, čehož využívají některé algoritmy. Např. v situaci, kdy úloha (LP) má 9 proměnných a 99 omezení, bude mít odpovídající duální úloha 99 proměnných a jen 9 omezení (plus přídavné proměnné potřebné pro převod do kanonického tvaru), což může výrazně urychlit výpočet. Pomocí duální úlohy také můžeme zkoumat závislost hodnoty primární úlohy na změně parametrů (koeficientů v účelové funkci či omezeních), čemuž se věnuje tzv. *analýza citlivosti*, na kterou se podrobněji zaměříme později v Podkapitole 8.4. Znalost řešení duální úlohy má také značný praktický význam, neboť poskytuje další informace o původní úloze. Uvažme opět problém optimálního jídelníčku ve standardním tvaru

$$\langle c, x\rangle \to \min \quad \& \quad Ax \geq b \quad \& \quad x \geq 0.$$

Pak duální úloha je maximalizační problém

$$\langle y, b\rangle \to \max \quad \& \quad A^\top y \leq c \quad \& \quad y \geq 0.$$

Proměnné $x_1, \ldots, x_n$ z původní úlohy reprezentují množství jednotlivých surovin, ale jaký je význam duálních proměnných $y_1, \ldots, y_m$? To se pokusíme zjistit pomocí tzv. *rozměrové analýzy*. V duálních omezeních $\sum_{i=1}^m a_{ij} y_i \leq c_j$ je c_j jednotková cena j-té suroviny a a_{ij} určuje množství i-té živiny v j-té surovině. Aby obě strany byly porovnatelné, musí y_i udávat jednotkovou cenu i-té živiny. Protože b_i je minimální požadované množství i-té živiny za den, účelová funkce $\langle y, b\rangle$ určuje celkovou cenu požadovaných živin. V duální úloze tedy hledáme hodnoty $y = (y_1, \ldots, y_m)^\top$ určující ceny živin tak, aby se maximalizovala celková hodnota požadovaného příjmu živin za den za podmínky $y \geq 0$, přičemž celková hodnota živin v j-tém jídle $\sum_{i=1}^m a_{ij} y_i$ není vetší než aktuální cena c_j j-tého jídla.

Teď si ale představme, že nějaký obchodník nám nabídne přímý prodej živin bez jídla, např. ve formě vitamínových tablet (vše plně syntetické, nenáročné na přípravu, plnohodnotně nahrazující vše potřebné atd.). Nabízí nám prodej i-té živiny za jednotkovou cenu y_i. Jak by měl správně zvolit cenu, abychom na tento obchod přistoupili? Připusťme, že nás zajímá pouze finanční stránka tohoto obchodu. Pak cena, kterou si účtuje za mix živin nahrazujících j-té jídlo by neměla být vyšší než původní cena j-tého jídla, což je právě omezení z duální úlohy. Je-li toto pravdivé pro všechna j, pak má smysl uzavřít obchod. Obchodník chce samozřejmě vydělat co nejvíce, takže by měl zvolit $y_1, \ldots, y_m$ maximalizující celkový zisk $\langle y, b\rangle$ při daných omezení, tj. právě jako řešení duální úlohy[125]! My sice vzhledem k nulovému duálnímu rozdílu mezi hodnotami primární a duální úlohy žádné peníze neušetříme, ale tento obchod pro nás může mít jiné nehmotné výhody – třeba úsporu času (i když v tomto případě to asi moc velká výhra nebude). V ekonomii se proto hodnoty $y_1, \ldots, y_m$ nazývají *stínovou cenou* jednotlivých živin. Slovo „stínová“ odráží fakt, že takový obchodník asi neexistuje nebo že by se mohlo jednat

[125] Zvolí-li obchodník jiné hodnoty, pak buď jeho potenciální zisk bude nižší nebo jím požadovaná cena bude vyšší než kolik by stál nákup surovin

o komoditu, která je na klasickém trhu neobchodovatelná. Nicméně stále v jistém smyslu vyjadřuje „cenu“ jednotlivých živin určenou tržními cenami jídel a našimi požadavky na skladbu příjmu živin.

Příklad 2.6.8

Uvažme řemeslníka, který k výrobě svých jediných dvou výrobků P a Q využívá dva stroje. Jednotlivé detaily jsou shrnuty v následující tabulce (může jít např. o výrobu stolů a židlí s využitím kotoučové pily a stříkací/lakovací pistole).

	stroj A [h]	stroj B [h]	zisk [Kč]
výrobek P	3	2	300
výrobek Q	2	3	400

Stroj A nesmí běžet déle než 13 hodin za den a stroj B více než 12 hodin za den. Při jakém složení výroby bude dosaženo maximálního zisku? To je řešením úlohy

$$300\,x_1 + 400\,x_2 \to \max$$
$$3\,x_1 + 2\,x_2 \le 13 \quad \& \quad 2\,x_1 + 3\,x_2 \le 12 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešením jsou 3 kusy výrobku P a 2 kusy výrobku Q s maximálním ziskem 1700 Kč.

Teď ale uvažme takovouto situaci: čerstvý absolvent (říkejme mu třeba Werther) studia matematiky na PřF MU by chtěl začít podnikat ve stejném oboru. Ovšem k tomu je potřeba mít počáteční kapitál, který pochopitelně nemá, a vzhledem k jeho věku a majetku (a také dvakrát pozdě vrátil knihy do knihovny, takže už má záznam v registru) nelze očekávat, že by mu nějaká banka poskytla rozumný úvěr. Ale doslechl se, že nás řemeslník zvažuje odchod do důchodu. Náš mladý podnikatel se proto rozhodl chopit příležitosti a navštívit řemeslníka s neodolatelnou nabídkou. Pronajme si oba stroje, přičemž za hodinu provozu stroje A zaplatí y_1 Kč a za hodinu provozu stroje B zaplatí y_2 Kč. Současně zachová stávající druh produkce a výrobků, které bude prodávat sám tak, aby mu to pokrylo alespoň nájem. Nápad se řemeslníkovi pochopitelně líbil, neboť již nebude muset řešit běžné problémy drobného živnostníka (daně, EET atd.) a bude mít čas na vnoučata. Současně je výhodný i pro mladého absolventa, neboť nemusí draze kupovat nové stroje. Nicméně řemeslník je však stále nedůvěřivý, a tak je nutné využít dosažené vzdělání a přesvědčit jej o výhodnosti celé nabídky.

Pro výrobu výrobku P jsou potřeba 3 hodiny na stroji A a 2 hodiny na stroji B, tedy za něj mladý Werther zaplatí nájem

$$3\,y_1 + 2\,y_2.$$

Současně však tento nájem bude alespoň roven jeho aktuálnímu zisku z prodeje tohoto výrobku, tj.

$$3\,y_1 + 2\,y_2 \ge 300.$$

Podobně pro výrobek Q dostaneme

$$2\,y_1 + 3\,y_2 \ge 400.$$

Řemeslníkovi se návrh líbil, ale nelíbí se mu možnost pronájmu pouze na část dne. Chtěl by pronajmout stroje na celou jejich provozní dobu. Jakou nabídku má mladík udělat, aby řemeslník souhlasil (on přitom ví, že jeho cenní zisk je 1700 Kč)?

Mladý absolvent se proto znovu zamyslel. Ví, že bude muset zaplatit $13\,y_1 + 12\,y_2$, ovšem nechce platit více než je nutné. Takže hodnoty y_1 a y_2 musí být řešením úlohy

$$13\,y_1 + 12\,y_2 \to \min$$

za podmínek

$$3y_1 + 2y_2 \geq 300 \quad \& \quad 2y_1 + 3y_2 \geq 400 \quad \& \quad y_1 \geq 0 \quad \& \quad y_2 \geq 0.$$

Toto je právě duální úloha k původnímu problému[126]. Má tato úloha řešení? Ze Důsledku 2.6.7 plyne, že určitě ano. Není těžké se přesvědčit, že tím je volba $y_1 = 20$ a $y_2 = 120$ s platbou 1700 Kč za den.

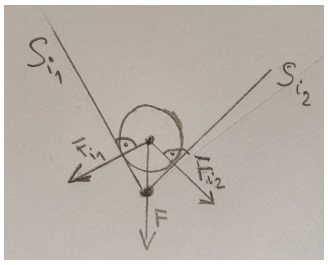
[126] Rozdíl mezi duální a primární je především v typu hledaného extrému. Ovšem již velmi dobře víme, že přechod mezi minimalizací a maximalizací je pouze o změně znaménka.

Dualitu lze přiblížit i z fyzikálního pohledu. Uvažme úlohu

$$\langle c, x \rangle \to \max \quad \& \quad Ax \leq b$$

s $n = 3$ a nechť tato úloha je řešitelná (tj. přípustná a ohraničená). Bod $x \in \mathbb{R}^3$ je prvkem třídimenzionálního prostoru a vektor c chápejme jako *gravitační vektor* (ten tedy směruje dolů – ovšem ne nutně v kolmém směru). Každá nerovnost v daném omezení $Ax \leq b$ určuje nějaký poloprostor a jejich průnik je zdola ohraničený konvexní polyedr. Každá dvoudimenzionální stěna polyedru je určena jednou rovností $\langle a_i, x \rangle = b_i$, kde vektory $a_1, \ldots, a_m$ tentokrát bereme jako transponované řádky matice A. Označme stěnu určenou rovností $\langle a_i, x \rangle = b_i$ jako S_i (ovšem ne každá nerovnost v soustavě $Ax \leq b$ musí odpovídat nějaké stěně, takže S_i nemusí být definováno pro všechna $i = 1, \ldots, m$).

Představme si, že hranice polyedru jsou vyrobeny z lepenky. Vezmeme miniaturní ocelovou kuličku a umístíme ji někam dovnitř polyedru. Kulička dopadne na stěnu a skutálí do nejníže položeného vrcholu (proto potřebujeme ohraničenost polyedru), případně zůstane stát na vodorovné stěně. Tuto pozici označíme x^*. Zde působí největší síla, takže se jedná o řešení dané maximalizační úlohy. V této stabilní pozici se kulička dotýká několika dvoudimenzionálních stěn (obvykle tří). Označme D množinu indexů i takových, že se kulička dotýká S_i. Pro $i \in D$ tedy máme $\langle a_i, x^* \rangle = b_i$. Gravitace vyvíjí na kuličku sílu F, která je úměrná vektoru c. Tato síla je rozložena mezi síly, jimiž kulička působí (tlačí) na jednotlivé stěny, kterých se dotýká. Síla F_i, kterou kulička působí na stěnu S_i je kolmá na S_i a směřuje ven z polyedru (zanedbáme-li tření), viz Obrázek 2.31.



Obrázek 2.31: Duální úloha a polyedr s kuličkou.

Síly působící na kuličku jsou v rovnováze, tj. $F = \sum_{i \in D} F_i$. Normálový vektor stěny S_i nasměrovaný ven z polyedru je právě a_i, takže síle F_i je úměrná a_i. To znamená, že pro nějaká nezáporná čísla y_i^* musí platit

$$\sum_{i \in D} y_i^* a_i = c.$$

Položíme-li $y_i^* = 0$ pro $i \in \{1, 2, \ldots, m\} \setminus D$, můžeme psát $\sum_{i=1}^m y_i^* a_i = A^\top y^* = c$, tj. y^* je přípustným bodem duální úlohy

$$\langle y, b \rangle \to \min \quad \& \quad A^\top y = c \quad \& \quad y \geq 0.$$

Uvažme nyní součin $\langle y^*, Ax^* - b \rangle$. Pro $i \in \{1, 2, \ldots, m\} \setminus D$ je i-tá složka vektoru y^* rovna 0, zatímco pro $i \in D$ je i-tá složka $Ax^* - b$ rovna 0. Tedy tento skalární součin je roven nule, tj. $\langle y^*, b \rangle = \langle y^*, Ax^* \rangle = \langle A^\top y^*, x^* \rangle = \langle c, x^* \rangle$. Celkem tedy x^* je přípustný bod pro primární úlohu, y^* je přípustný bod pro duální úlohu a platí $\langle y^*, b \rangle = \langle c, x^* \rangle$. Proto z Důsledku 2.6.3 vyplývá, že y^* je řešením duální úlohy.

Cvičení

2.6.1. Určete duální úlohu pro

$$3x_1 + 4x_2 \to \min$$
$$x_1 + 4x_2 \geq 8 \quad \& \quad 2x_1 + 3x_2 \geq 12 \quad \& \quad 2x_1 + x_2 \geq 6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

2.6.2. Určete duální úlohu pro

$$7x_1 + x_2 \to \min$$
$$4x_1 + 3x_2 \le 3 \quad \& \quad x_1 - 2x_2 \le 4 \quad \& \quad 5x_1 + 2x_2 \ge -3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.6.3. Určete duální úlohu pro

$$2x_1 + 5x_2 + x_3 \to \min$$
$$3x_1 - x_2 + 4x_3 \ge 2 \quad \& \quad 2x_1 + 2x_2 + x_3 \ge 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.6.4. Určete duální úlohu pro

$$13x_1 + 10x_2 + 6x_3 \to \min$$
$$5x_1 + x_2 + 3x_3 = 8 \quad \& \quad 3x_1 + x_2 = 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.6.5. Určete duální úlohu pro

$$5x_1 + 2x_2 + 6x_3 \to \min$$
$$4x_1 + 2x_2 + x_3 \ge 12 \quad \& \quad 3x_1 + 2x_2 + 3x_3 \le 6$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

2.6.6. Určete duální úlohu pro

$$4x_1 + 8x_2 + 6x_3 \to \min$$
$$x_1 + 2x_2 + x_3 \ge 3 \quad \& \quad 2x_1 - x_2 - x_3 \ge 2 \quad \& \quad -x_1 + x_2 = 1$$
$$x_1 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.6.7. Určete duální úlohu pro

$$2x_1 - 3x_2 + x_4 \to \min$$
$$x_1 + 2x_2 + x_3 \le 7 \quad \& \quad x_1 + 4x_2 - x_4 = 5 \quad \& \quad x_2 + x_3 + 5x_4 \ge 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.6.8. Určete duální úlohu pro

$$6x_1 + 6x_2 + 8x_3 + 9x_4 \to \min$$
$$x_1 + 2x_2 + x_3 + x_4 \ge 3 \quad \& \quad 2x_1 + x_2 + 4x_3 + 9x_4 \ge 0$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.6.9. Pomocí duální úlohy ukažte, že problém

$$6x_1 + x_2 \to \min$$
$$3x_1 + x_2 \ge 2 \quad \& \quad 2x_1 + 2x_2 \le -1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0,$$

nemá přípustné řešení.

2.6.10. Pomocí duální úlohy ukažte, že problém

$$-x_1 - x_2 \to \min$$
$$-x_1 + x_2 \le 1 \quad \& \quad x_1 - x_2 \le 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0,$$

nemá přípustné řešení.

2.6.11. Ukažte, že primární i duální úloha jsou neřešitelné, je-li primární úloha

$$-x_1 - 4x_2 \to \min$$
$$2x_1 - 2x_2 \geq 3 \quad \& \quad -2x_1 + 2x_2 \geq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

2.6.12. Ukažte, že primární i duální úloha jsou neřešitelné, je-li primární úloha

$$x_1 - 2x_2 \to \min$$
$$x_1 - x_2 \geq 1 \quad \& \quad -x_1 + x_2 \geq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

2.6.13. Určete řešení primární i duální úlohy, je-li primární úloha

$$x_1 + x_2 \to \min$$
$$x_1 \geq 2 \quad \& \quad x_2 \geq 1 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Ověřte také platnost tvrzení Silné věty o dualitě.

2.6.14. Určete řešení primární i duální úlohy, je-li primární úloha

$$8x_1 + 18x_2 \to \min$$
$$x_1 + 3x_2 \geq 2 \quad \& \quad 2x_1 + 4x_2 \geq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Ověřte také platnost tvrzení Silné věty o dualitě.

2.6.15. Uvažte úlohu

$$16x_1 + 32x_2 + 12x_3 \to \min$$
$$x_1 + 5x_2 + x_3 \geq 2 \quad \& \quad 4x_1 + 4x_2 - 2x_3 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Pomocí duální úlohy ukažte, že bod $[0, 5/14, 3/14]$ je řešením daného problému.

2.6.16. Uvažte úlohu

$$3x_1 + 5x_2 + 9x_3 \to \max$$
$$4x_1 + 12x_2 + 15x_3 = 900 \quad \& \quad -x_1 + 2x_2 + 3x_3 = 120$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Pomocí duální úlohy ukažte, že bod $[100/3, 0, 460/9]$ je řešením daného problému.

2.6.17. Uvažme problém

$$20x_1 + 30x_2 + 24x_3 \to \min$$
$$2x_1 - 3x_2 \geq -5 \quad \& \quad x_2 + 6x_3 \geq 18$$
$$-x_1 - 2x_2 + 3x_3 \geq 6 \quad \& \quad 3x_1 - x_2 + 4x_3 \geq -1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Pomocí duální úlohy, jejíž přípustný bod je $[10, 9, 0, 0]$, ověřte, že řešením primární úlohy je $[2, 0, 3]$.

2.6.18. Ověřte, že $[0, 17/3, 25/3, 1/3]$ je optimálním řešením úlohy

$$7x_1 + 11x_2 - 3x_3 - x_4 \to \min$$
$$2x_1 + 2x_2 - x_3 - 3x_4 \geq 2 \quad \& \quad -x_1 + 5x_2 - 2x_3 + x_4 \geq 12 \quad \& \quad x_1 - 4x_2 + 3x_3 + 5x_4 \geq 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0,$$

je-li $[7/2, 2, 3/2]$ optimálním řešením duální úlohy.

2.6.19. Uvažte problém

$$26x_1 + 30x_2 + 10x_3 \to \min$$
$$3x_1 + x_2 - 4x_3 \geq 4 \quad \& \quad 2x_1 - 6x_2 - 8x_3 \leq -10$$
$$-7x_1 + x_2 + 2x_3 \leq 3 \quad \& \quad x_1 + 5x_2 - x_3 \geq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

(a) Ukažte, že bod $[7/10, 19/10, 0]$ je přípustným bodem a určete hodnotu účelové funkce.

(b) Určete duální úlohu. Ukažte, že $[54/5, 16/5, 0, 0]$ je jejím přípustným bodem a určete příslušnou hodnotu účelové funkce duální úlohy.

(c) Co můžete na základě těchto informací říci o řešení původní úlohy?

2.6.20. Uvažte úlohu

$$13x_1 + 15x_2 + 12x_3 + 8x_4 \to \min$$
$$4x_1 + 8x_2 - 5x_3 + 3x_4 = 32 \quad \& \quad 3x_1 - 2x_2 + 6x_3 - x_4 \geq 3$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

(a) Rozhodněte, který z bodů $[9, 0, 2, 2]$, $[4, 1, -1, 1]$ a $[5, 1, 1, 3]$ je přípustný. Vyčíslete účelovou funkci v přípustných bodech.

(b) Sestavte duální úlohu a určete, který z bodů $[-1, 1]$, $[0, 2]$ a $[1, 3]$ je přípustný.

(c) Dokážete s pomocí těchto informací najít řešení primární úlohy? Pokud ne, můžete na základě těchto informací něco říci o řešení primární úlohy?

2.6.21. Určete duální úlohu pro

$$x_1 + 2x_2 + x_3 \to \min$$
$$x_1 - 2x_2 + x_3 \geq 2 \quad \& \quad -x_1 + x_2 + x_3 \geq 4$$
$$2x_1 + x_3 \geq 6 \quad \& \quad x_1 + x_2 + x_3 \geq 0$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Ověřte, že $[2/3, 0, 14/3]$ je optimálním řešením primární úlohy a současně $[0, 1/3, 2/3, 0]$ je optimálním řešením duální úlohy.

2.6.22. Uvažte úlohu

$$-2x_1 - 3x_2 - 2x_3 - 3x_4 \to \min$$
$$2x_1 + x_2 + 3x_3 + 2x_4 \leq 8 \quad \& \quad 3x_1 + 2x_2 + 2x_3 + x_4 \leq 7$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Určete duální úlohu a pomocí vhodného obrázku ji vyřešte. Ověřte, že $[3, 0, 2, 0]$ je řešením původní úlohy.

2.6.23. Ačkoli nejsem zrovna nutriční specialista, rozhodl jsem se zhubnout pomocí diety, která obsahuje pouze konzumaci vepřového masa a fazolí (sic!). Zároveň jsem si vědom, že je nutné dodržet alespoň minimální množství denního příjmu proteinů, vitamínu E a uhlovodanů. Detaily jsou v následující tabulce.

	maso	fazole	minimální denní dávka
proteiny [v porci]	5	2	300
vitamín E [v porci]	9	1	510
uhlovodany [v porci]	7	2	800
cena za porci	50	6	

(a) Kolik musím sníst každý den porcí masa a fazolí, abych splnil minimální požadavky při minimálních nákladech?

(b) Sestavte duální úlohu a ověřte, že jejím řešením je $[0, 58/11, 4/11]$.

(c) Poté, co se můj plán rozšířil po Brně, vycítilo vedení nedaleké lékárny možnost snadného zisku a nabídlo mi jednotlivé nutriční složky ve formě tablet. Jak by měla lékárna nastavit ceny, aby dosáhla maximálního zisku a pro mne by bylo ještě racionální jejich nabídku přijmout.

2.6.24. Firma prodávající mobilní telefony a různá příslušenství očekává novou dodávku zboží. Vzhledem k tomu je nutné vyprodat stávající zboží, které čítá 8 mobilních telefonů, 4 hands-free sady a 19 předplacených karet. Zboží bude prodáváno ve dvou balíčcích: balíček A obsahuje telefon a 2 karty, zatímco balíček obsahuje telefon, hands-free sadu a 3 karty. Zisk z prodeje balíčku A je 210,- a z prodeje balíčku B je 270,-.

(a) Konkurence by chtěla jednotlivé výrobky odkoupit. Sestavte úlohu lineárního programování, jehož řešením budou ceny jednotlivých produktů takové, že celková hodnota obchodu bude minimální a současně obchodník dosáhne alespoň požadovaného zisku z jednotlivých balíčků.

(b) Sestavte duální úlohu, vyřešte ji a interpretujte výsledek.

(c) S využitím výsledku části (b) ověřte, že řešením úlohy z části (a) je $[90, 0, 60]$.

2.6.25. Rafinérie má k dispozici tři závody X/Y/Z pro výrobu benzínu v různé kvalitě Q1/Q2/Q3. Produkce [100 galonů za hodinu] společně s náklady na 1 hodinu provozu [ve stovkách Kč] je uvedena v následující tabulce.

	Q1	Q2	Q3	náklady
závod X	3	4	2	32
závod Y	6	6	8	80
závod Z	6	3	4	60

Každý den je nutné vyrobit 3600/2000/3000 galonů benzínu kvality Q1/Q2/Q3.

(a) Sestavte lineární program, jehož řešením bude stanovení takové provozní doby jednotlivých závodů, že při minimálních nákladech bude dosaženo alespoň požadované produkce.

(b) Ověřte, že řešením je $[0, 3/2, 9/2]$, je-li řešením duální úlohy $[20/3, 0, 5]$.

(c) Uvažte, že rafinérie má možnost přímého nákupu od malovýrobce. Určete optimální cenu, kterou by měl malovýrobce požadovat za jednotlivé typy benzínu, aby pro rafinérii ještě bylo výhodné od něj nakupovat zboží pro uspokojení vlastní poptávky.

(d) Vedení rafinérie by chtělo zvýšit aktuální zisk. Ovšem z rozhodnutí Evropské komise č. 2174/2015 je nutné ospravedlnit své ceny před *Výborem pro kontrolu cen pohonných hmot, rohlíků a papíru*. Zvýšení bude povoleno pouze v případě, že aktuální ceny odpovídají provozním nákladům, které musí být na maximu (ovšem víme, že výroba není ztrátová). Při jakých cenách benzínů Q1/Q2/Q3 bude povoleno jejich zvýšení, jsou-li ceny nastaveny tak, aby se maximalizoval aktuální zisk?

2.7 Svatý grál lineárního programování – simplexová metoda

Geometrický základ *simplexového algoritmu* jsme si již naznačili. Nyní se na něj konečně podíváme více analyticky (a mnohem méně geometricky) a zaměříme se na „klíčové okamžiky“ celého algoritmu. Během postupu budeme generovat (obvykle) konečnou posloupnost BPB $x^{[0]}, x^{[1]}, \cdots \in X$, přičemž v p-tém kroku zjistíme, zda:

(i) úloha (LP) je neohraničená;

(ii) existuje BPB $x^{[p+1]}$ takový, že $\langle c, x^{[p+1]} \rangle < \langle c, x^{[p]} \rangle$, tj. nový bod s lepší (menší) funkční hodnotu a souřadnice takového bodu explicitně vypočteme;

(iii) bod $x^{[p]}$ je řešením úlohy (LP).

Předpokládejme, že úloha (LP) je přípustná – v opačném případě se nám ani nepodaří samotný algoritmus spustit, tj. najít výchozí BPB $x^{[0]}$, čemuž se budeme věnovat později. Mějme nyní již posloupnost BPB $x^{[0]}, x^{[1]}, \dots, x^{[p]}$ a pro jednoduchost označme $J := J(x)$ nosič bodu $x := x^{[p]}$ a (nějakou) jeho bázi $B := B(x)$. Pak při značení $x = (x_1, \dots, x_n)^\top$ a $c = (c_1, \dots, c_n)^\top$ platí

$$A\,x = \sum_{j\in B} x_j\,a_j = \sum_{j\in J} x_j\,a_j = b \tag{2.7.1}$$

a označme hodnotu účelové funkce v bodě x jako z_0, tj.

$$z_0 := \sum_{j\in B} c_j\,x_j = \sum_{j\in J} c_j\,x_j. \tag{2.7.2}$$

Jelikož vektory $\{a_j \mid j \in B\}$ jsou lineárně nezávislé, libovolný vektor $a_1, \dots, a_n$ může být vyjádřen jejich pomocí s koeficienty λ_{jk}, tj.

$$a_k = \sum_{j\in B} \lambda_{jk}\,a_j \quad \text{pro všechna } k = 1, \dots, n \tag{2.7.3}$$

a současně definujme čísla[127]

$$z_k := \sum_{j\in B} \lambda_{jk}\,c_j \quad \text{pro všechna } k = 1, \dots, n. \tag{2.7.4}$$

Ve vyjádření (2.7.3) je zřejmé, že $\lambda_{kk} = 1$ pro všechna $k \in B$ a $\lambda_{jk} = 0$ pro $j \in B \setminus \{k\}$. Proto též $z_k = c_k$ pro všechna $k \in B$. Jak ale získáme další koeficienty λ_{jk}? Pro jednoduchost přeznačme pořadí vektorů $a_1, \dots, a_n$ tak, že báze B našeho bodu x je $B = \{1, \dots, m\}$ a definujme matici

$$L := \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \dots & \lambda_{mn} \end{pmatrix} = \big(\lambda_{jk}\big)_{\substack{j=1,\dots,m \\ k=1,\dots,n}} \in \mathbb{R}^{m\times n}. \tag{2.7.5}$$

Rovnost (2.7.3) vlastně znamená

$$A = (a_1, \dots, a_n) = \underbrace{(a_1, \dots, a_m)}_{A_B} \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \dots & \lambda_{mn} \end{pmatrix} = A_B \cdot L,$$

tj. platí[128]

$$L = A_B^{-1} A \quad \text{neboli} \quad \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \dots & \lambda_{mn} \end{pmatrix} = A_B^{-1}\,(a_1, \dots, a_n). \tag{2.7.6}$$

Proto k-tý sloupec matice L dostaneme vynásobením této rovnosti vektorem e_k zprava, tj.

$$\begin{pmatrix} \lambda_{1k} \\ \vdots \\ \lambda_{mk} \end{pmatrix} = A_B^{-1} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} = A_B^{-1}\,a_k,$$

takže z_k definované v (2.7.4) můžeme vyjádřit jako

$$z_k = c_1\,\lambda_{1k} + c_2\,\lambda_{2k} + \dots + c_m\,\lambda_{mk} = (c_1, \dots, c_m) \begin{pmatrix} \lambda_{1k} \\ \vdots \\ \lambda_{mk} \end{pmatrix} = c_B^\top\,A_B^{-1}\,a_k. \tag{2.7.7}$$

Odtud také snadno nahlédneme, že pro libovolné $y \in X$ platí

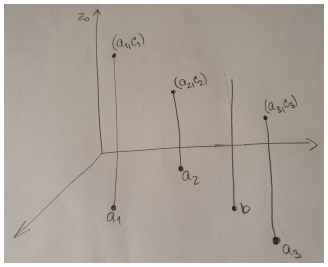
$$\langle c, x\rangle = \langle c, y\rangle + \sum_{j\notin B} (z_j - c_j)\,y_j. \tag{2.7.8}$$

[127] S využitím tohoto nového značení bychom mohli čtenáři parafrázovat Dantzigovo zdůvodnění označení „simplexová metoda“, které nabídl v článku [113] uveřejněném v Koopmansově sborníku z konference o lineárním programování konané v červnu 1949 v Chicagu. Toto zdůvodnění se týká úlohy (LP) s omezeními zahrnujícími i podmínku $x_1 + \dots + x_n = 1$, která bývá označována jako *konvexní omezení* (angl. *convexity constraint*). Ačkoli by se to mohlo zdát jako poměrně restriktivní požadavek, lze ukázat, že libovolnou úlohu lineárního programování s omezenou přípustnou množinou je možné převést do této podoby, viz Cvičení 2.8.35. My se ovšem nevydáme stejnou cestou jako Dantzig, který ve svém komentáři hovoří např. o *těžišti soustavy vektorů* $a_1, \dots, a_n$. Naše terminologie bude trochu více geometrická.

V úloze (LP) s konvexním omezením vlastně chceme minimalizovat hodnotu z_0 pro nezáporná $x \geq 0$ splňující konvexní omezení a rovnost

$$x_1 \begin{pmatrix} a_1 \\ c_1 \end{pmatrix} + \dots + x_n \begin{pmatrix} a_n \\ c_n \end{pmatrix} = \begin{pmatrix} b \\ z_0 \end{pmatrix}. \quad (*)$$

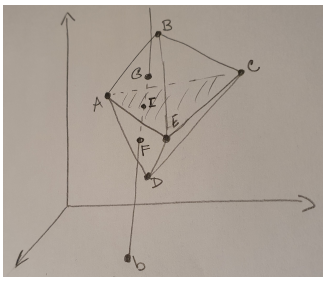
Tento problém můžeme geometricky znázornit v $\mathbb{R}^{m+1}$ tak, že vodorovnou rovinu budeme brát jako prostor $\mathbb{R}^m$ obsahující vektory $a_1, \dots, a_n$ a svislou osu jako jednodimenzionální prostor přidružený s koeficienty účelové funkce $c_1, \dots, c_n$, viz Obrázek 2.32.



Obrázek 2.32: Geometrická vizualizace úlohy (LP).

V tomto pohledu je naším cílem najít konvexní kombinaci (protože $\sum_{i=1}^n x_i = 1$) vektorů $(a_1^\top, c_1)^\top, \dots, (a_n^\top, c_n)^\top$ s nejmenší možnou hodnotou z_0. Ještě zdůrazněme, že body $[b^\top, z_0]$ leží na svislé přímce, která protíná vodorovnou rovinu v bodě b a bývá označována jako *přímka potřeb* (angl. *requirement line*). S terminologií zmíněnou ve druhé části Podkapitoly 2.4 pak můžeme říci, že pokud tato přímka neprotne konvexní obal bodů $[a_1^\top, c_1], \dots, [a_n^\top, c_n]$, pak je daný problém nepřípustný. Pokud je

tento průnik neprázdný, pak je úloha přípustná a její řešení odpovídá nejnižšímu bodu v tomto průniku, viz bod F na Obrázku 2.33.



Obrázek 2.33: Geometrická vizualizace řešení úlohy (LP).

A co jsou v tomto ohledu BPB? Vzhledem k přidanému konvexnímu omezení máme celkem $m+1$ rovnostních podmínek. Proto BPB je přidružen jistě $(m+1)$-tici vektorů tvaru $(a_i^\top, 1)^\top$. Tyto vektory jsou úzce spojeny s body $[a_i^\top, c_i]$, které můžeme označit jako *bazické*, zatímco ty ostatní body jsou *nebazické*. Protože bychom mohli ukázat, že bazické body jsou afinně nezávislé, je jejich konvexní obal jistě m-dimenzionálním simplexem nazývaným *bazickým simplexem*. Pokud průnikem tohoto simplexu s přímkou potřeb je nějaký bod $[b^\top, z_0]$, pak bod složený z $x_1, \dots, x_n$ vystupujících ve vyjádření $[b^\top, z_0]$ jakožto konvexní kombinace bazických bodů v (∗) výše je právě BPB a číslo z_0 určuje odpovídající hodnotu účelové funkce. Např. na Obrázku 2.33 je vyšrafovaný bazický simplex tvořený body AEC a bod I BPB přidružený bazickým bodům A, E a C. Změna báze v jednom bodě pak vede k novému bazickému simplexu a novému průniku s přímkou potřeb, čímž získáme nový bod $[b^\top, z_0]$ a následně i BPB. Cílem je pochopitelně dostat se k bodu F, takže v tomto ohledu je účelné měnit bázi zejména tak, aby nový bazický bod ležel „pod" současným bazickým simplexem. Pro více podrobností viz [60, Podkapitola 3.6].

[128] V případě první rovnosti i bez předchozího přeznačení vektorů.

Vskutku, je-li $y \in X$, pak $Ay = b$ neboli pro $N := \{1,\dots,n\} \setminus B$ máme $A_B\, y_B + A_N\, y_N = b$, takže

$$y_B = A_B^{-1} b - A_B^{-1} A_N\, y_N = A_B^{-1} b - A_B^{-1} \sum_{j \in N} y_j\, a_j,$$

z čehož plyne

$$\begin{aligned}\langle c, y\rangle &= \langle c_B, y_B\rangle + \langle c_N, y_N\rangle = c_B^\top \Big(A_B^{-1} b - A_B^{-1} \sum_{j\in N} y_j\, a_j\Big) + c_N^\top y_N = \\ &= c_B^\top A_B^{-1}\, b - \Big(c_B^\top A_B^{-1} \sum_{j\in N} y_j\, a_j - \sum_{j\in N} c_j\, y_j\Big) = c_B^\top\, x_B - \sum_{j\in N} \Big(c_B^\top A_B^{-1} a_j - c_j\Big) y_j = \\ &= \langle c, x\rangle - \sum_{j\in N} (z_j - c_j)\, y_j.\end{aligned}$$

Proto x bude řešením úlohy (LP) právě tehdy, když $\sum_{j\notin B} (z_j - c_j)\, y_j \le 0$ pro všechna $y \in X$. To vzhledem k nezápornosti složek y_j bude splněno zejména v případě, že $z_j - c_j \le 0$ pro $j \notin B$. Každopádně vzhledem ke znaménku rozdílu $z_k - c_k$ a čísla λ_{jk} a řešitelnosti úlohy (LP) jsou pro nás zajímavé pouze 3 situace, které si nyní rozebereme. Jejich důkazy jsou založeny na rovnostech (2.7.1)–(2.7.4), ze kterých pro libovolné $\alpha \in \mathbb{R}$ odečtením α-násobku rovnosti (2.7.3) od rovnosti (2.7.1) dostaneme

$$\sum_{j\in B} (x_j - \alpha\, \lambda_{jk})\, a_j + \alpha\, a_k = b \quad \text{pro všechna } k = 1,\dots,n \tag{2.7.9}$$

a podobně odečtením α-násobku (2.7.4) od (2.7.2) a přidáním $\alpha\, c_k$ na obě strany získáme

$$\sum_{j\in B} (x_j - \alpha\, \lambda_{jk})\, c_j + \alpha\, c_k = z_0 - \alpha\,(z_k - c_k) \quad \text{pro všechna } k = 1,\dots,n\,. \tag{2.7.10}$$

Věta 2.7.1

Nechť je dána úloha (LP). Jestliže existuje index $s \in \{1,2,\dots,n\} \setminus B$ takový, že $z_s - c_s > 0$ a $\lambda_{js} \le 0$ pro všechna $j \in B$, pak úloha (LP) je neohraničená.

Důkaz. Nechť index $s \in \{1,2,\dots,n\} \setminus B$ je takový, že $z_s - c_s > 0$ a současně $\lambda_{js} \le 0$ pro všechna $j \in B$. Pak pro libovolné $\alpha > 0$ je bod

$$x(\alpha) := \begin{cases} x_j - \alpha\, \lambda_{js}, & j \in B, \\ \alpha, & j = s, \\ 0, & j \notin B \cup \{s\} \end{cases}$$

je přípustný, tj. $x(\alpha) \in X$, neboť $m+1$ jeho složek je kladných pro $B \cup \{s\}$, ostatní jsou nulové a z rovnosti (2.7.9) aplikované pro $k = s$ plyne $A\, x(\alpha) = \sum_{j\in B\cup\{s\}} x_j(\alpha)\, a_j = b$. Jelikož α může být libovolné kladné číslo, dostaneme dále z (2.7.10) pro $k = s$, že

$$\langle c, x(\alpha)\rangle = \sum_{j\in B\cup\{s\}} c_j\, x_j(\alpha) = z_0 - \alpha \underbrace{(z_s - c_s)}_{>0} \to -\infty \quad \text{pro } \alpha \to \infty\,,$$

tj. úloha (LP) je neohraničená. ∎

Další možnost nás přivede k situaci, kdy existuje „lepší" BPB. Uvidíme, že báze tohoto bodu $x^{[p+1]}$ a bodu $x^{[p]}$ se liší pouze v jediném indexu. Takovým bodům se říká *sousední*. Ponecháme na čtenáři, aby si rozmyslel, že v případě degenerované úlohy (LP) existují body, které sousedí samy se sebou.

Věta 2.7.2

Nechť je dána nedegenerovaná úloha (LP). Existuje-li index $s \in \{1,2,\dots,n\} \setminus B$ takový, že $z_s - c_s > 0$ a současně $\lambda_{ks} > 0$ pro nějaké $k \in B$, pak existuje nový BPB $\widehat{x} \in X$, pro který je hodnota účelové funkce menší než v bodě x.

Důkaz. Nechť pro nějaká $s \in \{1,2,\dots,n\} \setminus B$ a $k \in B$ jsou splněny předpoklady věty. Pak díky nedegenerovanosti dané úlohy jistě existuje číslo $\alpha > 0$ takové, že $x_j - \alpha\, \lambda_{js} \ge 0$ pro všechna

$j \in B$. Vskutku, pro $\lambda_{js} \le 0$ je uvedená nerovnost splněna triviálně, zatímco v případě $\lambda_{js} > 0$ je potřeba vzít $\alpha \in (0, x_j/\lambda_{js}]$. Jelikož $\lambda_{js} > 0$ je splněno nejméně pro jeden index j a nejvýše pro n indexů, můžeme definovat číslo

$$\widehat{\alpha} := \min\left\{\frac{x_j}{\lambda_{js}} \;\middle|\; j \in B \,\&\, \lambda_{js} > 0\right\} > 0.$$

Označíme-li index, ve kterém nastává toto minimum, jako $r \in B$, pak pro něj platí $x_r - \widehat{\alpha}\,\lambda_{rs} = 0$. Navíc index r je určen jednoznačně, neboť v opačném případě bychom s využitím (2.7.9) pro $k = s$ a $\alpha = \widehat{\alpha}$ získali vyjádření vektoru b jako nějakou lineární kombinaci nejvýše $(m-1)$-tice vektorů $\{a_j \mid j \in (B \setminus \{r_1, r_2\}) \cup \{s\}\}$, což je spor ↯.

Jestliže se nyní omezíme na $\alpha \in (0, \widehat{\alpha}]$ a definujeme-li bod $x(\alpha)$ se složkami

$$x_j(\alpha) := \begin{cases} x_j - \alpha\,\lambda_{js}, & j \in B, \\ \alpha, & j = s, \\ 0, & j \notin B \cup \{s\}, \end{cases} \tag{2.7.11}$$

pak z první části a rovnosti (2.7.9) pro $k = s$ vyplývá, že $x(\alpha) \in X$. Navíc hodnota účelové funkce v takovém bodě podle (2.7.10) pro $k = s$ splňuje

$$z(\alpha) := \langle c, x(\alpha)\rangle = z_0 - \alpha \underbrace{(z_s - c_s)}_{>0} < z_0, \tag{2.7.12}$$

tj. je menší než v bodě x. Ovšem bod $x(\alpha)$ nemusí být BPB. To se stane pouze a jenom pro $\alpha = \widehat{\alpha}$ vedoucí k $x_r(\widehat{\alpha}) = 0$, neboť v opačném případě bychom vždy měli $x(\alpha)$ s $(m+1)$-ticí nezáporných složek. Proto $\widehat{x} := x(\widehat{\alpha})$ je hledaným „lepším“ BPB[129]. ■

Nastane-li situace popsaná v právě dokázané Větě 2.7.2, pak za další člen naší posloupnosti BPB bereme $x^{[p+1]} := x(\widehat{\alpha})$ s bází $B^{[p+1]} = (B^{[p]} \setminus \{r\}) \cup \{s\}$, přičemž se mluví o tom, že index r (příp. vektor a_r) *vystupuje z báze* a současně index s (příp. vektor a_s) *vstupuje/přibíráme do báze*. Navíc z (2.7.12) lze snadno vidět, že právě tato volby $x(\widehat{\alpha})$ dává největší možné snížení hodnoty účelové funkce mezi všemi body $x(\alpha)$ pro $\alpha \in (0, \widehat{\alpha}]$. Jelikož víme, že BPB je pouze konečný počet, bude tento algoritmus za předpokladu nedegenerovanosti úlohy (LP) vždy konečně krokový. Jestliže je ale úloha (LP) degenerovaná, pak se může stát, že $\widehat{\alpha} = 0$ (pokud by $x_j = 0$ pro nějaké $j \in B$), takže by nastalo $x(\widehat{\alpha}) = x^{[p]}$, a tudíž by došlo pouze ke změně báze BPB $x^{[p]}$. To by mohlo vést až k zacyklení celého algoritmu, čemuž se budeme podrobněji věnovat ve druhé části následující podkapitoly.

[129] V maticovém značení můžeme vyjádření $x_j(\alpha)$ pro $j \in B$ (tj. v původní bázi) z (2.7.11) zapsat jako

$$x_B(\alpha) = x_B - \alpha\, A_B^{-1} A_N\, e_{N_s}$$

neboli

$$x_B(\alpha) - x_B = \\ = -\alpha\, A_B^{-1} A_N\, e_{N_s}, \tag{2.7.13}$$

kde $e_{N_s} \in \mathbb{R}^{m-n}$ a tento vektor je volen tak, aby součin $A_N e_{N_s}$ odpovídal sloupci s indexem vstupujícím do báze (tj. e_{N_s} má nulové prvky s výjimkou pozice pro sloupec odpovídající proměnné x_s, kde má jedničku – kdybychom měli bázi $B = \{1, \dots, m\}$, pak bychom místo N_s mohli psát $s - m$; tento součin bychom také mohli nahradit za Ae_s pro vhodný $e_s \in \mathbb{R}^n$). To znamená, že z bodu x_B vyrazíme směrem $-A_B^{-1} A_N\, e_{N_s}$, který míří do $x_B(\alpha)$. Číslo $\alpha \ge 0$ pak vyjadřuje délku tohoto kroku (záporná hodnota by otočila směr), takže $\widehat{\alpha} > 0$ je jeho maximální možná hodnota, při které neporušíme nezápornost $x_B(\alpha)$ a současně (díky předpokladu nedegenerovanosti) právě jednu souřadnici vynulujeme.

Věta 2.7.3

Nechť je dána nedegenerovaná úloha (LP). Bod x je řešením úlohy (LP) právě tehdy, když $z_k - c_k \le 0$ pro všechna $k \in \{1, \dots, n\} \setminus B$.

Důkaz. „$\Longrightarrow$“ Vzhledem k předpokladu negenerovanosti úlohy (LP) plyne nutnost uvedené podmínky z tvrzení Vět 2.7.1 a 2.7.2.

„$\Longleftarrow$“ Nechť $y = (y_1, \dots, y_n)^\top \ne x$ je jiný přípustný bod množiny X, tj. $Ay = b$ a $y \ge 0$. Protože $z_k - c_k \le 0$ pro všechna $k \in \{1, \dots, n\} \setminus B$ a současně $z_j - c_j = 0$ pro každé $j \in B$ dle (2.7.4), platí $z_k - c_k \le 0$ pro všechna $k \in \{1, \dots, n\}$. Proto

$$\langle y, z\rangle = \sum_{i=1}^{n} y_i\, z_i \le \sum_{i=1}^{n} y_i\, c_i = \langle y, c\rangle =: z^*.$$

Navíc dosazením vyjádření (2.7.3) do rovnosti $Ay = b$ dostaneme

$$b = Ay = \sum_{i=1}^{n} y_i\, a_i = \sum_{i=1}^{n} y_i \Big(\sum_{j \in B} \lambda_{ji}\, a_j\Big) = \sum_{i=1}^{n} y_i\, (\lambda_{j_1 i}\, a_{j_1} + \lambda_{j_2 i}\, a_{j_2} + \dots + \lambda_{j_m i}\, a_{j_m})$$

neboli

$$b = \sum_{j \in B} \Big(\sum_{i=1}^{n} \lambda_{ji}\, y_i\Big)\, a_j. \tag{2.7.14}$$

Podobně dosazením (2.7.4) do nerovnosti $\langle y, z\rangle \le z^*$ získáme

$$\langle y, z\rangle = \sum_{i=1}^{n} y_i z_i = \sum_{i=1}^{n} y_i \Big(\sum_{j\in B} \lambda_{ji}\, c_j \Big) \le z^*$$

neboli

$$\sum_{j\in B} \Big(\sum_{i=1}^{n} \lambda_{ji}\, y_i \Big) c_j \le z^*. \tag{2.7.15}$$

Protože vektory $\{a_j \mid j \in B\}$ tvoří bázi $\mathbb{R}^m$ a protože vyjádření libovolného vektoru pomocí báze musí být jednoznačné, koeficienty odpovídajících vektorů v rovnostech (2.7.1) a (2.7.14) si musí být rovny, tj. $\sum_{i=1}^{n} \lambda_{ji}\, y_i = x_j$ pro všechna $j \in B$. Pak ale z (2.7.15) plyne

$$z^* \ge \sum_{j\in B} \Big(\sum_{i=1}^{n} \lambda_{ji}\, y_i \Big) c_j = \sum_{j\in B} x_j\, c_j = z_0 = \langle c, x\rangle.$$

Jelikož $y \in X$ bylo voleno libovolně, vyplývá z poslední nerovnosti, že x je řešením úlohy (LP). ■

Všímavý čtenář jistě postřehl, že pro důkaz implikace „⇐“ v předchozím tvrzení nebyl požadavek nedegenerovanosti dané úlohy potřeba. To znamená, že podmínka „$z_k - c_k \le 0$ *pro všechna* $k \in \{1,\dots,n\}\setminus B$“ je postačující pro nalezení řešení libovolné úlohy (LP). Ovšem v případě řešitelnosti mohou nastat dvě diametrálně odlišné situace podle jednoznačnosti řešení: řešení se realizuje buď v jediném bodě (ten je nutně BPB) nebo ve více bodech (z nichž alespoň jeden musí být BPB), v kterémžto případě jich bude nutně nespočetně mnoho (neboť množina řešení musí být konvexní, takže řešením bude i každý bod ležící na úsečce spojující dvě řešení). Odlišnost těchto situací není ale zajímavá pouze z matematického pohledu, protože i v praxi jistě mohou nastat situace, kdy přeci jen některé řešení bude výhodnější než jiné, ačkoli dává stejnou minimální hodnotu účelové funkce[130].

[130] Např. kdyby šlo o výrobu založenou na nákupu některých zdrojů od externích dodavatelů, tak bychom mohli preferovat řešení založené na dodavateli, s nímž se nám lépe obchoduje nebo který by nám následně mohl nabídnout i nějakou slevu. To jsou ale faktory, které nelze a priori zahrnout do samotné formulace problému.

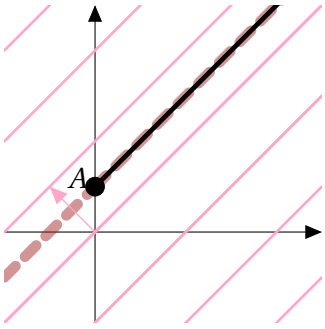
DŮSLEDEK 2.7.4

Nechť je dána úloha (LP). Jestliže $z_k - c_k < 0$ pro všechna $k \in \{1,\dots,n\}\setminus B$, pak x je jediné řešení úlohy (LP).

Důkaz. Jelikož $z_k - c_k < 0$ pro všechny indexy $k \in \{1,\dots,n\}\setminus B$, plyne z Věty 2.7.3, že x je řešením úlohy (LP). Předpokládejme, že řešení dané úlohy není jednoznačné, tj. existuje další přípustný bod $y \in X$ splňující $x \ne y$ a $\langle c, y\rangle = \langle c, x\rangle$. Pak nutně platí $y_j > 0$ pro nějaký index $j \in \{1,\dots,n\}\setminus B$. Kdyby tomu tak nebylo (tj. $y_j = 0$ pro všechna $j \in \{1,\dots,n\}\setminus B$), pak by množina B také byla bází bodu y, což by vzhledem jednoznačnosti BPB nutně znamenalo $x = y$, a tedy bychom dostali spor ↯, viz Větu 2.5.8. Potom ale z (2.7.8) vyplývá

$$\langle c, y\rangle = \langle c, x\rangle - \sum_{j\notin B} \underbrace{(z_j - c_j)}_{\substack{<0 \text{ pro všechna } j\notin B}} \overbrace{y_j}^{\substack{\ge 0 \text{ pro všechna } j\notin B \\ >0 \text{ pro nějaké } j\notin B}} > \langle c, x\rangle,$$

což je opět spor ↯, tj. y nemůže být řešením, čímž je jednoznačnost dokázána. ■



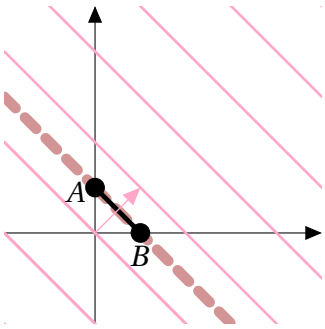
Obrázek 2.34: Úloha $-x_1 + x_2 \to \min$ pro $-x_1 + x_2 = 1$, $x_1 \ge 0$ a $x_2 \ge 0$, která má nespočetně mnoho řešení, ale pouze jedno z nich je BPB (tím je bod A).

Všimněme si, že předchozí tvrzení je platné bez ohledu na degenerovanost úlohy (LP). Na druhou stranu pro následující důsledek je již nedegenerovanost potřeba. V opačném případě bychom se totiž mohli dostat do situace, ve které bychom za alternativní řešení považovali tentýž degenerovaný BPB vyjádřený pouze pomocí různých bází. Z důkazu navíc uvidíme, že v případě rovnosti $z_s - c_s = 0$ pro nějaké $s \in \{1,\dots,n\}\setminus B$ mohou nastat dvě různé situace podle toho, zda alternativní řešení je BPB či nikoli, jak je vidět na Obrázcích 2.34–2.35. Rozdíl mezi těmto situacemi lze popsat také „v řeči bází“: jde totiž o to, zda po nalezení řešení úlohy (LP) ještě existuje nějaký index $r \in B$, na jehož úkor by index s vstoupil do báze.

Důsledek 2.7.5

Nechť je dána nedegenerovaná úloha (LP). Jestliže x je řešením úlohy (LP) a je-li $z_s - c_s = 0$ pro nějaké $s \in \{1,\dots,n\}\setminus B$, pak existuje i jiné (alternativní) řešení této úlohy. Je-li navíc $\lambda_{ks} > 0$ pro nějaké $k \in B$, pak řešením je i jiný BPB, zatímco v opačném případě je x jediným BPB, který řeší úlohu (LP).

Důkaz. Tvrzení víceméně ihned vyplývá z konstrukce bodu $x(\alpha)$ v důkazu Věty 2.7.2, neboť nerovnost $z_s - c_s > 0$ jsme využili až v jeho závěru. Vskutku, je-li $z_s - c_s = 0$ pro nějaké $s \in \{1,\dots,n\}\setminus B$ a současně $\lambda_{ks} > 0$ pro nějaké $k \in B$, pak lze zkonstruovat nový BPB $x(\widehat{\alpha})$ stejně jako v (2.7.11) s tím rozdílem, že tentokrát v (2.7.12) nastane rovnost, tj. hodnoty účelové funkce v BPB x a $x(\widehat{\alpha})$ jsou totožné. Kdyby ale bylo $z_s - c_s = 0$ pro nějaké $s \in \{1,\dots,n\}\setminus B$ a současně $\lambda_{js} \le 0$ pro všechna $j \in B$, pak by bod $x(\alpha)$ z (2.7.11) stále byl přípustným bodem pro všechna $\alpha \ge 0$ a splňoval by $\langle c, x\rangle = \langle c, x(\alpha)\rangle$, ale neexistovala by hodnota $\widehat{\alpha} > 0$, pro kterou by se z něj stal BPB. ∎



Obrázek 2.35: Úloha $x_1 + x_2 \to \min$ pro $x_1 + x_2 = 1$, $x_1 \ge 0$ a $x_2 \ge 0$, která má nespočetně mnoho řešení a dva z nich jsou BPB (body A a B).

2.8 Teď už si to jen správně zorganizovat

Samotný výpočet řešení úlohy (LP) založený na simplexové metodě probíhá prostřednictvím tzv. *simplexové tabulky*, která je postupně systematicky upravována podle pravidel *simplexového algoritmu*. Za chvíli se přesvědčíme, že průběh tohoto algoritmu velmi připomíná *Gaussovu eliminaci* – ovšem s řízeným výběrem *pivota* (či *hlavního/klíčového prvku*) a „drobným" omezením při úpravách dalších řádků. V literatuře lze narazit na různé tvary této tabulky, které se liší především uspořádáním jednotlivých sloupců či řádků a také mírou podrobností uvedených údajů. Vždy z nich ale lze vyčíst všechny informace potřebné pro výpočet. Abychom náš výklad zbytečně nekomplikovali, zdá se nám nejvýhodnější začít v prvním kroku s velmi podrobnou tabulkou o velikosti $(n+4)\times(m+4)$, jak můžeme vidět[131,132] v Tabulce 2.8.

hodnoty bazických složek vektoru c	označení proměnných odpovídajících indexům v bázi	označení proměnných a odpovídající prvky matice L				hodnota bazických složek vektoru x	podíl $\frac{x_B}{\lambda_{js}}$ (pro vhodné s)
c_B	B	x_1	x_2	$\cdots$	x_n	x_B	x_B/λ
c_{b_1}	x_{b_1}	$\lambda_{b_1,1}$	$\lambda_{b_1,2}$	$\cdots$	$\lambda_{b_1,n}$	x_{b_1}	$x_{b_1}/\lambda_{b_1,s}$
c_{b_2}	x_{b_2}	$\lambda_{b_2,1}$	$\lambda_{b_2,2}$	$\cdots$	$\lambda_{b_2,n}$	x_{b_2}	$x_{b_2}/\lambda_{b_2,s}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
c_{b_m}	x_{b_m}	$\lambda_{b_m,1}$	$\lambda_{b_m,2}$	$\cdots$	$\lambda_{b_m,n}$	x_{b_m}	$x_{b_m}/\lambda_{b_m,s}$
	c	c_1	c_2	$\cdots$	c_n		
	z_k	z_1	z_2	$\cdots$	z_n		
	$z_k - c_k$	$z_1 - c_1$	$z_2 - c_2$	$\cdots$	$z_n - c_n$	$\langle c, x\rangle$	

Tabulka 2.8: Výchozí simplexová tabulka

V dalších krocích již není potřeba takto podrobná tabulka a dále se pracuje s tzv. *redukovanou simplexovou tabulkou*, ve které již není potřeba vypisovat sloupec označený c_B a řádky označené c a z_k. Někteří autoři také kvůli úspoře času/barvy či kvůli fyzické námaze ignorují při vyplňování tabulky nulové hodnoty, takže jejich tabulka občas působí neúplně. My ale tento úzus aplikovat nebudeme, neboť by se nám při ručním výpočtu mohla tabulka snadno stát velmi nepřehlednou (a při výpočtu na počítači se tím nemusíme moc trápit).

[131] K ní ale ještě musíme přidat několik vysvětlení:

(i) první řádek je pouze informativní popis a v praktických výpočtech není potřeba (proto jsme jej ani nezahrnuli mezi započítávané řádky do velikosti tabulky);

(ii) červeně jsou uvedena symbolická označení jednotlivých sloupců/řádků, zde nic nedosazujeme;

(iii) zelenou barvou jsou uvedeny proměnné s bazickými indexy, zde místo $x_{b_1},\dots,x_{b_m}$ dosadíme konkrétní hodnoty (viz i následující poznámku);

(iv) černou barvou jsou uvedeny proměnné či složky různých vektorů/matic a při praktickém výpočtu místo nich uvedeme konkrétní číselné hodnoty;

(v) šedé buňky zůstávají prázdné.

[132] Až doposud jsme hovořili výhradně o indexech v bázi. Avšak při převodu úlohy lineárního programování do kanonického tvaru obvykle dojde i k zavedení nových proměnných, takže si už se samotnými indexy nevystačíme (nebo bychom museli vše komplikovaně přeznačovat, aby proměnné byly pouze $x_1, x_2 \dots$). Proto budeme při řešení konkrétních úloh bází rozumět spíše množinu tvořenou několika proměnnými, tj. např. můžeme mít bázi $B = \{x_1, x_3, s_1, w_4\}$ apod.

Příklad 2.8.1

Uvažme úlohu lineárního programování, ve které hledáme

$$x_1 - x_2 + 2x_3 + 3x_4 \to \min$$

za podmínek

$$2x_1 + x_2 + x_3 = 1 \quad \& \quad x_1 + 3x_2 + x_4 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Tato úloha je již v kanonickém tvaru a odpovídá úloze (LP) s $n = 4$, $m = 2$ a volbou

$$c = \begin{pmatrix} 1 & -1 & 2 & 3 \end{pmatrix}^\top \quad \& \quad A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix} \quad \& \quad b = \begin{pmatrix} 1 & 2 \end{pmatrix}^\top.$$

Jelikož vektor b lze získat pouze pomocí dvou (libovolných) sloupců matice A, jedná se o nedegenerovanou úlohu. Zvolíme-li bázi $B = \{3,4\}$, pak $A_B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, takže podle (2.5.3) máme $x_B = A_B^{-1} b = b$ a v souladu s (2.7.6) dostaneme $L = A_B^{-1} A = A$. Proto výchozí simplexová tabulka má následující podobu (poslední sloupec s x_B / λ zatím není doplněný, neboť jeho výpočet úzce souvisí s prvním krokem simplexového algoritmu).

c_B	B	x_1	x_2	x_3	x_4	x_B	x_B / λ
2	x_3	2	1	1	0	1	$x_3 / \lambda_{3,s}$
3	x_4	1	3	0	1	2	$x_4 / \lambda_{4,s}$
	c	1	−1	2	3		
	z_k	7	11	2	3		
	$z_k - c_k$	6	12	0	0	8	

Tabulka 2.9: Výchozí simplexová tabulka k Příkladu 2.8.1

A teď již samotný popis celého algoritmu pro řešení úloh lineárního programování.

Simplexový algoritmus

(1) *Inicializace*: danou úlohu lineárního programování převedeme do kanonického tvaru (LP), tj. vhodnými úpravami z nerovností uděláme rovnosti, proměnné bez omezení na znaménko nahradíme novými nezápornými proměnnými atd.

(2) Pokud nějaká složka vektoru b je záporná, vynásobíme příslušnou rovnici -1 tak, aby všechny složky vektoru b byly nezáporné.

(3) Nalezneme výchozí BPB[133] a určíme jeho bázi B.

(4) Sestavíme výchozí simplexovou tabulku pro zvolený BPB.

(5) *Test optimality*. Jestliže

(i) $z_k - c_k \leq 0$ pro všechna $k \in \{1,\dots,n\} \setminus B$, je výpočet hotov. Zvolený BPB je řešením úlohy. KONEC.

(ii) $z_s - c_s > 0$ pro nějaké $s \in \{1,\dots,n\} \setminus B$ a současně

(ii_a) $\lambda_{js} \leq 0$ pro všechna $j \in B$, pak úloha nemá řešení. KONEC.

(ii_b) $\lambda_{js} > 0$ pro nějaké $j \in B$, pak existuje nový BPB s „lepší“ hodnotou účelové funkce, který najdeme v dalších krocích.

[133] Obsahuje-li matice A všechny sloupce jednotkové matice, může za výchozí BPB vzít $x^{[0]} = b$ s bází tvořenou odpovídajícími indexy (ve vhodném pořadí v bázi či s případnou změnou pořadí složek vektoru b). V opačném případě využijeme některou z modifikací simplexového algoritmu (*metodu velkého M* či *dvoufázovou metodu*), kterým se budeme věnovat v následující Podkapitole 2.9. Zde také zjistíme případnou nepřípustnost úlohy, neboť se nám nepodaří najít žádný BPB, viz Větu 2.5.5.

(6) *Indexy vstupující a opouštějící bázi.* Určíme *klíčový sloupec.*

(i) Pokud $z_s - c_s > 0$ pro jediný index $s \in \{1,\dots,n\}\setminus B$, pak je to sloupec odpovídající proměnné x_s.

(ii) Pokud $z_s - c_s > 0$ pro více indexů $s \in \{1,\dots,n\}\setminus B$, pak je to sloupec odpovídající proměnné x_s, pro kterou má rozdíl $z_s - c_s$ největší hodnotu, tj.

$$z_s - c_s = \max\left\{z_j - c_j \mid j \in \{1,\dots,n\}\setminus B \ \&\ z_j - c_j > 0\right\} > 0.$$

Kdyby i toto nastalo ve více indexech, vybereme ten nejmenší (tj. ten, který je v simplexové tabulce nejvíce vlevo[134]).

Index odpovídající klíčovému sloupci vstupuje do báze. Současně z báze vystupuje ten index r, pro který

$$\frac{x_r}{\lambda_{rs}} = \min\left\{\frac{x_j}{\lambda_{js}} \;\middle|\; j \in B \ \&\ \lambda_{js} > 0\right\} \geq 0 \qquad (2.8.1)$$

(k tomu využijeme poslední sloupec simplexové tabulky), čímž dostaneme tzv. *klíčový řádek*[135]. Takto získáváme novou bázi $B\setminus\{r\}\cup\{s\}$ a prvek λ_{rs} v simplexové tabulce je tzv. *klíčový prvek* nebo *pivot* (průnik klíčového řádku a sloupce).

(7) *Aktualizace simplexové tabulky.* Sestavíme novou simplexovou tabulku (už bez řádků odpovídajících c a z_k) odpovídající bázi $B\setminus\{r\}\cup\{s\}$. Nemusíme ale vše přepočítávat od začátku. Stačí upravit stávající tabulku podle následujících pravidel[136]:

(i) klíčový řádek vynásobíme číslem $1/\lambda_{rs}$, abychom na pozici klíčového prvku dostali 1;

(ii) v ostatních řádcích klíčového sloupce chceme mít 0, takže od každého ze zbývajících $m-1$ řádků pro x_B odečteme klíčový řádek vynásobený číslem $\lambda_{js}/\lambda_{rs}$;

(iii) podobnou úpravu provedeme i pro řádek $z_k - c_k$, tj. klíčový řádek vynásobíme $(z_s - c_s)/\lambda_{rs}$;

(iv) ve sloupci označeném x_B s bazickými proměnnými místo x_r napíšeme x_s a ostatní bazické proměnné ponecháme stejné.

(8) Pokračujeme krokem (5).

[134] Více podrobností o této situaci nalezneme v Poznámce 2.8.9.

[135] Je-li úloha nedegenerovaná, pak je klíčový řádek určen jednoznačně. V případě degenerované úlohy můžeme minimum z podílů $\frac{x_j}{\lambda_{js}}$ nastat pro několik indexů, z nichž si opět můžeme „libovolně“ vybrat. Avšak konkrétní volba může mít vliv na rychlost a úspěšnost samotného algoritmu, viz další část této podkapitoly. Existuje tedy více variant simplexového algoritmu, které se liší právě specifikací pravidel pro výběr klíčového sloupce (ze všech kladných hodnot v řádku $z_k - c_k$) a klíčového řádku (to jenom v případě, kdy ve více řádcích dostaneme tentýž minimální podíl).

[136] Viz Větu 2.8.2 níže.

Jestliže úloha (LP) není neohraničená ani nepřípustná a současně zvolený výchozí BPB není jejím řešením, pak v souladu s Větou 2.7.2 zkonstruujeme nový BPB. Ovšem tento bod již nemusíme počítat pomocí rovnosti (2.5.3), považovat jej za nový výchozí BPB a vracet se až k bodu (4) s konstrukcí nové výchozí simplexové tabulky. Všechny tyto informace totiž dává nová (simplexová) tabulka získaná dle pravidel popsaných v bodě (7), takže můžeme pokračovat přímo bodem (5). Korektnost tohoto postupu zaručuje následující věta, která nevyžaduje nedegenerovanost dané úlohy. Navíc body (6)–(7) lze aplikovat i v případě, kdy $z_s - c_s = 0$ a současně $\lambda_{ks} > 0$ pro nějaké $k \in B$, čímž nalezneme jiný bod se stejnou funkční hodnotou. Ten je buď alternativním řešením nebo součástí tzv. cyklu (je-li úloha degenerovaná), viz později.

Věta 2.8.2

Nechť je dána úloha (LP) a x je jejím BPB s bází B. Jestliže $z_s - c_s \geq 0$ a současně $\lambda_{ks} > 0$ pro nějakou dvojici indexů $s \in \{1,2,\dots,n\}\setminus B$ a $k \in B$, pak tabulka získaná podle pravidel popsaných v bodě (7) simplexového algoritmu je simplexovou tabulkou pro nový BPB $\widehat{x}$ s bází $\widehat{B} := B\setminus\{r\}\cup\{s\}$ získanou v bodě (6).

Důkaz. Z Věty 2.7.2 víme, že novým bazickým bodem je

$$\widehat{x}_j := \begin{cases} x_j - \frac{\lambda_{js}}{\lambda_{rs}} x_r, & j \in B, \\ \frac{x_r}{\lambda_{rs}}, & j = s, \\ 0, & j \notin B \cup \{s\}, \end{cases}$$

pro vhodný index $r \in B$, přičemž $\widehat{x}_r = 0$ a současně $\lambda_{rs} > 0$, viz (2.8.1). Jelikož

$$a_s \overset{(2.7.3)}{=} \sum_{j \in B} \lambda_{js} a_j = \sum_{j \in B \setminus \{r\}} \lambda_{js} a_j + \lambda_{rs} a_r,$$

plyne odtud

$$a_r = \frac{1}{\lambda_{rs}} a_s - \sum_{j \in B \setminus \{r\}} \frac{\lambda_{js}}{\lambda_{rs}} a_j. \tag{2.8.2}$$

Proto můžeme libovolný vektor a_k pro $k \in \{1, \dots, n\}$ vyjádřit jako

$$\begin{aligned} a_k &\overset{(2.7.3)}{=} \sum_{j \in B \setminus \{r\}} \lambda_{jk} a_j + \lambda_{rk} a_r = \\ &\overset{(2.8.2)}{=} \sum_{j \in B \setminus \{r\}} \lambda_{jk} a_j + \lambda_{rk} \Big(\frac{1}{\lambda_{rs}} a_s - \sum_{j \in B \setminus \{r\}} \frac{\lambda_{js}}{\lambda_{rs}} a_j \Big) = \\ &= \sum_{j \in B \setminus \{r\}} \Big(\lambda_{jk} - \lambda_{js} \frac{\lambda_{rk}}{\lambda_{rs}} \Big) a_j + \frac{\lambda_{rk}}{\lambda_{rs}} a_s \end{aligned}$$

neboli dostáváme rovnost

$$a_k = \sum_{j \in \widehat{B}} \hat{\lambda}_{jk} a_j$$

s koeficienty

$$\hat{\lambda}_{jk} = \begin{cases} \lambda_{jk} - \lambda_{js} \frac{\lambda_{rk}}{\lambda_{rs}}, & j \in B \setminus \{r\}, \ \ k \in \{1, \dots, n\}, \\ \frac{\lambda_{rk}}{\lambda_{rs}}, & j = s, \ \ k \in \{1, \dots, n\}, \end{cases}$$

určujícími novou matici $\widehat{L}$ odpovídající bázi $\widehat{B}$, viz (2.7.5). Pak ale $\hat{\lambda}_{ss} = 1$ a $\hat{\lambda}_{js} = 0$ pro $j \in \widehat{B} \setminus \{s\}$, což odpovídá právě krokům popsaným v simplexovém algoritmu v bodě (7)(i)–(ii). Nyní ještě zbývá určit $\widehat{z}_k - c_k$ a $\langle c, \widehat{x} \rangle$. Poněvadž platí

$$\begin{aligned} \widehat{z}_k - c_k &\overset{(2.7.4)}{=} \sum_{j \in \widehat{B}} \hat{\lambda}_{jk} c_j - c_k = \sum_{j \in B \setminus \{r\}} \Big(\lambda_{jk} - \lambda_{js} \frac{\lambda_{rk}}{\lambda_{rs}} \Big) c_j + \frac{\lambda_{rk}}{\lambda_{rs}} c_s - c_k = \\ &= \Big(\sum_{j \in B} \lambda_{jk} c_j - \lambda_{rk} c_r \Big) - \frac{\lambda_{rk}}{\lambda_{rs}} \Big(\sum_{j \in B} \lambda_{js} c_s - \lambda_{rs} c_r \Big) + \frac{\lambda_{rk}}{\lambda_{rs}} c_s - c_k = \\ &= \sum_{j \in B} \lambda_{jk} c_j - c_k - \frac{\lambda_{rk}}{\lambda_{rs}} \Big(\sum_{j \in B} \lambda_{js} c_s - c_s \Big) = (z_k - c_k) - \frac{z_s - c_s}{\lambda_{rs}} \lambda_{rk} \end{aligned}$$

a současně také

$$\begin{aligned} \langle c, \widehat{x} \rangle &= \sum_{j \in \widehat{B}} c_j \widehat{x}_j = \sum_{j \in B \setminus \{r\}} \Big(x_j - \frac{\lambda_{js}}{\lambda_{rs}} x_r \Big) c_j + \frac{x_r}{\lambda_{rs}} c_s = \\ &= \Big(\sum_{j \in B} c_j x_j - c_r x_r \Big) - \frac{x_r}{\lambda_{rs}} \Big(\sum_{j \in B} \lambda_{js} c_j - \lambda_{rs} c_r - c_s \Big) = \\ &= \sum_{j \in B} c_j x_j - \frac{x_r}{\lambda_{rs}} \Big(\sum_{j \in B} \lambda_{js} c_j - c_s \Big) = \langle c, x \rangle - \frac{z_s - c_s}{\lambda_{rs}} x_r, \end{aligned}$$

jsou hodnoty $\widehat{z}_k - c_k$ a $\langle c, \widehat{x} \rangle$ zcela v souladu s výpočtem popsaným v simplexovém algoritmu v bodě (7)(iii). ■

Příklad 2.8.3

Nedegenerovaná úloha v kanonickém tvaru:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp1

Příklad 2.8.4

Nedegenerovaná úloha:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp2
⚠ řádek pro s_4 v čase 7:11 má být $2, 0, 0, 1, 0, -1, 8$.

Příklad 2.8.5

Neohraničená úloha:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp3

Ve chvíli, kdy simplexový algoritmus ukazuje, že daná úloha je neohraničená, nám poslední simplexová tabulka dokonce dává v souladu s důkazem Věty 2.7.1 směr, ve kterém hodnota účelové funkce klesá k $-\infty$. Toto si ukážeme v následujícím příkladu.

Příklad 2.8.6

S využitím simplexového algoritmu vyřešme úlohu lineárního programování

$$x_1 - 2x_2 + x_3 + 3x_4 \to \min$$
$$2x_1 - x_2 + x_3 - x_4 \le 10$$
$$-5x_1 + 2x_2 - 2x_3 + x_4 \le 20$$
$$3x_1 - 4x_2 + 4x_3 - 2x_4 \le 30$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Řešení. Při řešení postupujeme stejně jako v předchozích příkladech. Začneme přidáním pomocných proměnných s_1, s_2 a s_3, čímž získáme úlohu v kanonickém tvaru

$$x_1 - 2x_2 + x_3 + 3x_4 \to \min$$
$$2x_1 - x_2 + x_3 - x_4 + s_1 = 10$$
$$-5x_1 + 2x_2 - 2x_3 + x_4 + s_2 = 20$$
$$3x_1 - 4x_2 + 4x_3 - 2x_4 + s_3 = 30$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0$$
$$s_1 \ge 0 \quad \& \quad s_2 \ge 0 \quad \& \quad s_3 \ge 0.$$

Tímto krokem v matici A vytvoříme jednotkovou matici a snadno určíme výchozí BPB. Sestavíme si výchozí simplexovou tabulku a pokračujeme v souladu se simplexovým algoritmem, viz Tabulku 2.10.

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
0	s_1	2	−1	1	−1	1	0	0	10	✖
0	s_2	−5	(2)	−2	1	0	1	0	20	20/2 →
0	s_3	3	−4	4	−2	0	0	1	30	✖
	c	1	−2	1	3	0	0	0		
	z_k	0	0	0	0	0	0	0		
	$z_k - c_k$	−1	2 ↑	−1	−3	0	0	0	0	
	s_1	−1/2	0	0	−1/2	1	1/2	0	20	
	x_2	−5/2	1	−1	1/2	0	1/2	0	10	
	s_3	−7	0	0	0	0	2	1	70	
	$z_k - c_k$	4	0	1	−4	0	−1	0	−20	

Tabulka 2.10: Simplexová metoda pro Příklad 2.8.6.

Klíčovým sloupcem v prvním kroku je druhý sloupec a klíčovým řádkem je druhý řádek. Pomocí pivota, kterým je zakroužkované číslo 2, provedeme eliminaci podle pravidel simplexového algoritmu, čímž dostaneme druhou část tabulky. V tomto kroku máme dvě kladná čísla v řádku $z_k - c_k$. Zkusíme-li vybrat např. první sloupec, pak ale v něm nenajdeme žádné kladné číslo, tj. nelze určit klíčový řádek. Stejně to dopadne i v případě třetího sloupce. To znamená, že simplexový algoritmus je již u konce a daná úloha je neohraničená.

Avšak z poslední tabulky můžeme vyčíst i další informace, které plynou z důkazu Věty 2.7.1. S pomocí třetího sloupce totiž můžeme definovat bod

$$x(\alpha) := [0, 10+\alpha, \alpha, 0, 20, 0, 70],$$

kde číselné hodnoty odpovídají hodnotám bazických proměnných v poslední simplexové tabulce (čísla 10, 20 a 70) a koeficienty u parametru α jsou u bazických proměnných určeny koeficienty ve třetím sloupci s opačným znaménkem (0 pro s_1 a s_3 a +1 pro x_2). Tento sloupec odpovídá proměnné x_3, takže jí ještě přiřadíme hodnotu α. Bod $x(\alpha)$ splňuje všechna omezení pro libovolné $\alpha \geq 0$ a současně odpovídající hodnota účelové funkce je rovna $-20-\alpha$, což pro $\alpha \to \infty$ konverguje k $-\infty$, tj. účelová funkce je skutečně neohraničená a její hodnoty klesají ve směru určeným přípustným bodem $[0, 10+\alpha, \alpha, 0]$. Analogický výsledek můžeme získat i s pomocí prvního sloupce. Tentokrát ale budeme mít bod $\bar{x}(\alpha) := [\alpha, 10+5\alpha/2, 0, 0, 20+\alpha/2, 0, 70+7\alpha]$ s hodnotou účelové funkce $-20-4\alpha \to -\infty$ pro $\alpha \to \infty$, tj. účelová funkce klesá ve směru určeným přípustným bodem $[\alpha, 10+5\alpha/2, 0, 0]$.

Příklad 2.8.7

Alternativní řešení #1:
Video ☞ 😀 https://ln.math.muni.cz/pz/opt/lp4

Příklad 2.8.8

Alternativní řešení #2:
Video ☞ 😀 https://ln.math.muni.cz/pz/opt/lp5

Poznámka 2.8.9. Volba klíčového sloupce popsaná v kroku (6)(ii) simplexového algoritmu je tzv. *Dantzigovo pravidlo* nebo *pravidlo maximálního relativního ocenění*[137], které pochází z [113, str. 343]. Je motivováno tím, že hodnota $z_k - c_k$ jsou tzv. *redukované náklady* a vyjadřuje jednotkovou míru změny (poklesu) účelové funkce při změně v proměnné x_k. Koeficienty v účelové funkci totiž reprezentují jednotkové náklady „aktivit“ určených proměnnými. Redukované náklady pak indikují, jak moc by se musel redukovat k-tý koeficient účelové funkce, aby současná „aktivita“ reprezentovaná k-tou proměnnou byla *nákladově efektivní* (tzn., že by měla kladnou hodnotu v optimálním řešení). Proto tato volba klíčového sloupce dává největší možné zlepšení hodnoty účelové funkce na jednotku proměnné vstupující do báze. Nevýhodou tohoto pravidla je, že zlepšení hodnoty účelové funkce závisí i na hodnotě vstupující proměnné. Pokles tedy může být poměrně malý i při výběru největší hodnoty rozdílu.

Existují samozřejmě i jiné možnosti výběru klíčového sloupce[138]. Jaká je ta nejlepší? Těžko říct… Náš požadavek by přirozeně mohl být, abychom vždy vybrali takovou proměnnou, že algoritmus proběhne co nejrychleji (neboli s nejmenším počtem kroků). Ovšem zároveň by toto „urychlení“ nemělo být příliš na úkor výpočetní náročnosti celého algoritmu. A to už je velmi obtížná, neřkuli nemožná, kombinace. Při výběru klíčového sloupce se obvykle využívají pravidla založená na hodnotě účelové funkce. My si nyní ukážeme čtyři z nich a později ještě další dvě přidáme, viz také [453, 477].

[137] Angl. *the most negative* (sic!) *reduced cost rule* (v původním Dantzigově algoritmu se totiž místo $z_k - c_k$ počítá $c_k - z_k$) nebo také *the steepest descent rule.*

[138] Taková pravidla pro výběr klíčového sloupce (případně i řádku) se v literatuře obvykle nazývají jako *pivotovací* (angl. *pivot rule*), viz např. [374, Section 5.7]

(i) *Náhodné pravidlo*[139] a *pravidlo nejmenšího indexu*[140] (či *první volby*). Tím nejjednodušším pravidlem, které snad ihned napadne každého z nás, je náhodný výběr (třeba pomocí hodu kostkou). V případě takového výběru jistě dojdeme k optimálnímu řešení, jenže už sám Dantzig zmiňuje, že náhodný výběr vede obecně k většímu počtu kroků, než využití jiných pravidel, viz [113, str. 343]. Výběr sloupce, který má nejmenší index (nebo je v tabulce nejvíce vlevo), je podobný náhodnému výběru. Zlepšení efektivnosti těchto dvou pravidel nabízí všechna další uvedená pravidla.

(ii) *Pravidlo maximálního zmenšení účelové funkce*[141]. Jedná se o další Dantzigovo pravidlo pocházející z [113, str. 343]. Podle něj se klíčový sloupec vybírá tak, abychom v daném kroku zajistili co největší pokles hodnoty účelové funkce. Toho docílíme tím, že nejprve pro všechny indexy splňující $z_k - c_k > 0$ spočítáme podíly $\frac{x_B}{\lambda}$, z nichž poté pro každé x_k vybereme ten nejmenší, který vynásobíme s hodnotou $z_k - c_k$. Do báze pak vstoupí proměnná, pro kterou je tento součin maximální. Tímto pravidlem tedy vybíráme proměnnou s takovým indexem s, že platí

$$(z_s - c_s)\frac{x_j}{\lambda_{js}} = \max\left\{(z_k - c_k)\frac{x_j}{\lambda_{jk}} \;\middle|\; k \in \{1,\dots,n\}\setminus B \;\&\; z_k - c_k > 0 \;\&\; \&\; \frac{x_j}{\lambda_{jk}} = \min\{x_r/\lambda_{rk},\ r \in B,\ \lambda_{rk} > 0\}\right\}.$$

(iii) *Pravidlo nejstrmější hrany*[142]. Toto pravidlo pochází z [133], viz také [182, 223, 253, 336]. Zde vybíráme klíčový sloupec tak, aby došlo k maximálnímu možnému poklesu na jednotku vzdálenosti podél hrany poklesu. Jestliže jsme v p-tém kroku algoritmu, vybíráme další bazický přípustný bod $x^{[p+1]}$ tak, abychom maximalizovali výraz

$$\frac{\left|\langle c, x^{[p+1]} - x^{[p]}\rangle\right|}{\|x^{[p+1]} - x^{[p]}\|}, \tag{2.8.3}$$

kde $\|\cdot\|$ je klasická euklidovská norma. Přesné znění tohoto pravidla se ale v různých zdrojích liší. Např. v [374, str. 71] nalezneme předchozí variantu, zatímco tentýž autor (jeden z nich) v [373, str. 74] toto pravidlo formuluje tak, že se vybírá klíčový sloupec, pro který se minimalizuje výraz

$$\frac{\langle c, x^{[p+1]} - x^{[p]}\rangle}{\|c\| \times \|x^{[p+1]} - x^{[p]}\|},$$

což odpovídá kosinu úhlu sevřeného vektory c a $(x^{[p+1]} - x^{[p]})$. Nicméně obě tyto volby vedou k výběru stejného klíčového sloupce, neboť dělení hodnotou $\|c\|$ výběr maxima neovlivní[143]. Každopádně je nutné se vyhnout tomu, že $\|x^{[p+1]} - x^{[p]}\| = 0$ neboli $x^{[p+1]} = x^{[p]}$, což hrozí v případě degenerované úlohy.[144] ▲

Nyní si ukážeme aplikaci těchto pravidel v jednom konkrétním příkladu. Nebude-li ale řečeno jinak, budeme v dalších úlohách používat simplexový algoritmus s Dantzigovým pravidlem tak, jak jsme jej popsali dříve.

Příklad 2.8.10

Uvažme úlohu

$$-x_1 - 2x_2 - 3x_3 \to \min$$

za podmínek

$$-x_1 + x_2 + x_3 + x_4 = 2 \quad \& \quad x_1 + x_2 + x_3 + x_5 = 4 \quad \& \quad x_3 + x_6 = 1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0 \quad \& \quad x_6 \ge 0.$$

Úloha je již v kanonickém tvaru a snadno můžeme vidět, že za výchozí BPB lze vzít $x^{[0]} = [0,0,0,2,4,1]$ s bází $B(x^{[0]}) = \{4,5,6\}$, což nás přivede k následující výchozí simplexové tabulce.

[139] Angl. the random rule.

[140] Angl. *the least-index rule.*

[141] Angl. *the maximal improvement rule.*

[142] Angl. *the steepest edge rule.* Toto pravidlo velmi úzce souvisí s tzv. *DEVEX algoritmem* z [253], ve kterém dochází pouze k přibližnému hledání nejstrmější hrany, což ale výrazně usnadňuje samotnou implementaci celého výpočtu.

[143] V literatuře lze najít také např. *pravidlo nejvzdálenějšího vrcholu*, ve kterém se chceme z aktuálního BPB přesunout do nejvzdálenějšího BPB apod.

[144] Proč tato volba? Protože pro danou účelovou funkci $f(x) = c^\top x$ je $c = \operatorname{grad} f(x)$, takže výraz uvnitř absolutní hodnoty v (2.8.3) je vlastně

$$\left\langle \operatorname{grad} f(x), \frac{x^{[p+1]} - x^{[p]}}{\|x^{[p+1]} - x^{[p]}\|} \right\rangle,$$

což je rovno normované směrové derivaci účelové funkce v bodě $x^{[p]}$ ve směru $x^{[p+1]} - x^{[p]}$, který z $x^{[p]}$ míří do $x^{[p+1]}$ (tj. jdeme podél odpovídající hrany polyedru). Jelikož v souladu se simplexovým algoritmem je funkce f je v tomto nerostoucí, volbou $x^{[p+1]}$ minimalizujícího tento výraz nebo maximalizujícího (2.8.3) vydáme se od $x^{[p]}$ do $x^{[p+1]}$ ve směru největšího poklesu funkční hodnoty.

Určení vhodného $x^{[p+1]}$ tak, že otestujeme všechny možné kandidáty, může ale být početně náročné. Pro jednoduchost označme $x := x^{[p]}$ s bází B a jejím doplňkem N. Uvažme, že v dalším kroku má do báze vstoupit index s, takže novým BPB $x^{[p+1]}$ bude $x(\widehat{\alpha})$. Pak z (2.7.13) plyne, že rozdíl těchto dvou po sobě jdoucích BPB můžeme vyjádřit jako

$$\Delta x = \begin{pmatrix}\Delta x_B \\ \Delta x_N\end{pmatrix} := x^{[p+1]} - x^{[p]} = \begin{pmatrix} x_B(\widehat{\alpha}) - x_B \\ x_N(\widehat{\alpha}) - x_N\end{pmatrix} = \begin{pmatrix} -\widehat{\alpha}\, A_B^{-1} A_N e_{N_s} \\ \widehat{\alpha}\, e_{N_s}\end{pmatrix}.$$

Proto hodnota výše uvedené normované směrové derivace je rovna

$$\frac{\partial f}{\Delta x} = \left\langle \operatorname{grad} f, \frac{\Delta x}{\|\Delta x\|}\right\rangle = c^\top \frac{\Delta x}{\|\Delta x\|} = \frac{c_B^\top \Delta x_B + c_N^\top \Delta x_N}{\|\Delta x\|}.$$

Výraz v čitateli je díky předchozí-

mu vyjádření Δx a vztahu (2.7.7) roven

$$\begin{aligned} c_B^\top \Delta x_B + c_N^\top \Delta x_N &= \\ &= \widehat{\alpha}\, c_N\, e_{N_s} - \widehat{\alpha}\, c_B^\top A_B^{-1} A_N\, e_{N_s} = \\ &= \widehat{\alpha}\, c_N\, e_{N_s} - \widehat{\alpha}\, c_B^\top A_B^{-1} a_s = \\ &= \widehat{\alpha}\,(c_s - z_s), \end{aligned}$$

zatímco jmenovatel je

$$\begin{aligned} \|\Delta x\|^2 &= \|\Delta x_B\|^2 + 1 = \\ &= |\widehat{\alpha}|^2 \times \left(\| -A_B^{-1} A_N\, e_{N_s}\|^2 + 1\right). \end{aligned}$$

Aktuální hodnotu $\widehat{\alpha}(c_s - z_s)$ snadno vyčteme ze simplexové tabulky, ale výpočet normy $\| -A_B^{-1} A_N\, e_{N_s}\|$ pro všechna $s \in N$ je vlastně totéž jako výpočet celé matice $-A_B^{-1} A_N$, čemuž je dobré se vzhledem k jeho výpočetní náročnosti vyhnout. Naštěstí stačí vypočítat $-A_B^{-1} A_N\, e_{N_s}$ pro všechna $s \in N$ pouze na začátku celého algoritmu, kdy A_B je mnohdy jednotková matice (nebo jí velmi blízká), a pak aktualizovat hodnotu požadované normy pomocí celkem „jednoduchého" vzorce, který si nyní odvodíme. Položme

$$v_k := \| -A_B^{-1} A_N\, e_{N_k}\|^2, \quad k \in N.$$

Připusťme, že všechna tato čísla známe, a využijeme je k provedení dalšího kroku simplexové metody. Nyní jsme na jeho začátku a veličiny aktualizované přechodem k nové bázi budeme značit pomocí symbolu $\widehat{\ }$, tj. např. $\widehat{B}$ značí novou bázi, která vznikne z B výměnou jednou indexu za jiný. Vstoupí-li do báze index s a současně vystoupí index r, pak matici $A_{\widehat{B}}$ můžeme získat pomocí „jednoduchého" vzorce

$$A_{\widehat{B}} = A_B + (a_s - a_r)\, e_r^\top,$$

kde a_s a a_r jsou odpovídající sloupce matice A. Díky regulárnosti matice A_B toto můžeme zapsat i jako

$$A_{\widehat{B}} = A_B\,[I + A_B^{-1}(a_s - a_r)\, e_r^\top].$$

Současně již také víme, že

$$A_B^{-1} a_s = A_B^{-1} A_N\, e_{N_s} = -\frac{\Delta x_B}{\widehat{\alpha}}$$

a

$$A_B^{-1} a_r = e_r.$$

Proto platí

$$\begin{aligned} T &:= I + A_B^{-1}(a_s - a_r)\, e_r^\top = \\ &= I - (\Delta x_B/\widehat{\alpha} + e_r)\, e_r^\top. \end{aligned}$$

Navíc tato matice je v případě nedegenerované úlohy regulární a

$$T^{-1} = I + \frac{(\Delta x_B/\widehat{\alpha} - e_r)\, e_r^\top}{\Delta x_r},$$

což plyne z toho, že jsou-li dány dva vektory $u, v \in \mathbb{R}^n$ splňující $1-$

c_B	B	x_1	x_2	x_3	x_4	x_5	x_6	x_B	x_B/λ
0	x_4	-1	1	1	1	0	0	2	
0	x_5	1	1	1	0	1	0	4	
0	x_6	0	0	1	0	0	1	1	
	c	-1	-2	-3	0	0	0		
	z_k	0	0	0	0	0	0		
	$z_k - c_k$	1	2	3	0	0	0	0	

Tabulka 2.11: Výchozí simplexová tabulka k Příkladu 2.8.10.

V řádku $z_k - c_k$ jsou tři kladná čísla, takže máme tři možné kandidáty pro výběr klíčového sloupce.

(i) Podle klasického Dantzigova pravidla z „našeho" simplexového algoritmu vybereme největší kladné číslo, klíčový sloupec odpovídá proměnné x_3.

(ii) Podle náhodného pravidla bychom vybrali zcela náhodně, např. sloupec odpovídající x_2. Podle pravidla nejmenšího indexu bychom se zvolili z těchto sloupců ten s nejmenším indexem, tj. sloupec odpovídající x_1.

(iii) Pro pravidlo maximálního zmenšení účelové funkce si musíme napočítat podíly x_B/λ pro všechny kandidáty na klíčový sloupec, což je shrnuto v Tabulce 2.12

B	x_B	x_B/λ_{i1}	x_B/λ_{i2}	x_B/λ_{i3}
x_4	2	✖	(2)	2
x_5	4	(4)	4	4
x_6	1	✖	✖	(1)
$z_k - c_k$	0	1	2	3
$(z_k - c_k)\, x_B/\lambda$		$1 \times 4 = 4$	$2 \times 2 = 4$	$3 \times 1 = 3$

Tabulka 2.12: Aplikace pravidla maximálního zmenšení účelové funkce v Příkladu 2.8.10.

V každém sloupci této tabulky je zakroužkován nejmenší podíl, z čehož vidíme, jaká proměnná by opustila bázi v případě zvolení tohoto sloupce za klíčový. Např. při výběru sloupce pro x_1 by bázi opustila proměnná x_5. Tyto vybrané podíly nyní vynásobíme příslušnými čísly v řádku $z_k - c_k$, což je vypočteno v posledním řádku tabulky. Odtud podle pravidla maximálního zmenšení účelové funkce vybereme největší číslo, tj. buď ve sloupci pro x_1 nebo x_2 (toto pravidlo již neříká, co dělat v případě takové shody). Tato čísla nám říkají, o kolik se zmenší hodnota účelové funkce v případě volby takového sloupce za klíčový. Pro první a druhý sloupec se účelová funkce zmenší o 4 (tj. na -4), zatímco pro třetí sloupec by se hodnota zmenšila pouze o 3.

(iv) Využití pravidla nejstrmější hrany vyžaduje nejvíce pomocných výpočtů. Musíme si totiž dopočítat souřadnice všech možných bodů, do kterých se můžeme v dalším kroku dostat. Naštěstí nepotřebujeme nutně celé tabulky, ale postačí nám

$v^\top u \neq 0$, pak matice $I - u v^\top$ je regulární a přímým výpočtem lze ověřit, že $(I - u v^\top)^{-1} = I + \frac{u v^\top}{1 - v^\top u}$. Hodnota ν_k se proto změní na

$$\begin{aligned}\hat{\nu}_k &= a_k^\top A_{\widehat{B}}^{-\top} A_{\widehat{B}}^{-1} a_k = \\ &= a_k^\top A_B^{-\top} T^{\top -1} T^{-1} A_B^{-1} a_k = \\ &= a_k^\top A_B^{-\top} \left(I + \frac{e_r (\Delta x_B / \widehat{\alpha} - e_r)^\top}{\Delta x_r}\right) \times \\ &\quad \times \left(I + \frac{(\Delta x_B / \widehat{\alpha} - e_r) e_r^\top}{\Delta x_r}\right) A_B^{-1} a_k.\end{aligned}$$

Položíme-li

$$v := A_B^{-\top} e_r \quad \text{a} \quad w := A_B^{-\top} \Delta x_B / \widehat{\alpha},$$

pak roznásobením předchozí rovnosti získáme hledaný vzorec

$$\hat{\nu}_k = \nu_k + 2 \frac{a_k^\top v (w - v)^\top a_k}{\Delta x_r} + \\ + (a_k^\top v)^2 \frac{\| \Delta x_B / \widehat{\alpha} - e_r \|^2}{(\Delta x_r)^2},$$

přičemž nezapomínejme, že $\Delta x_B / \widehat{\alpha} = -A_B^{-1} A_N e_{N_s}$. Numerické výpočty využívající tento vzorec ukázaly, že pravidlo nejstrmější hrany je srovnatelné s (ne-li lepší než) ostatními pravidla pro výběr klíčového sloupce, viz také [491, str. 147–149]

vždy pouze výpočet sloupce se souřadnicemi x_B. Z tohoto důvodu jsou v Tabulce 2.13 uvedeny jenom báze a odpovídající sloupec x_B pro všechny možné další BPB.

B	x_B	B	x_B	B	x_B
x_4	6	x_2	2	x_4	1
x_1	4	x_5	2	x_5	3
x_6	1	x_6	1	x_3	1
$z_k - c_k$	-4	$z_k - c_k$	-4	$z_k - c_k$	-3

Tabulka 2.13: Aplikace pravidla nejstrmější hrany v Příkladu 2.8.10.

První část tabulky odpovídá situaci, kdy do báze vstoupí proměnná x_1 a vystoupí x_5. Druhá část odpovídá situaci, ve které proměnná x_2 vystřídá v bázi proměnnou x_4, zatímco v poslední části je proměnná x_6 v původní bázi nahrazena proměnnou x_3. Pak podle vzorce (2.8.3) s $p = 0$ můžeme pro různé volby nového BPB $x^{[1]}$ spočítat (zatím bez absolutní hodnoty)

$$x^{[1]} = [4,0,0,6,0,1]: \quad \frac{(-1,-2,-3,0,0,0)^\top (4,0,0,4,-4,0)}{\|(4,0,0,4,-4,0)^\top\|} = -\frac{4}{4\sqrt{3}} = -\frac{1}{\sqrt{3}},$$

$$x^{[1]} = [0,2,0,0,2,1]: \quad \frac{(-1,-2,-3,0,0,0)^\top (0,2,0,-2,2,0)}{\|(0,2,0,-2,2,0)^\top\|} = -\frac{4}{2\sqrt{3}} = -\frac{2}{\sqrt{3}},$$

$$x^{[1]} = [0,0,1,1,3,0]: \quad \frac{(-1,-2,-3,0,0,0)^\top (0,0,1,-1,-1,-1)}{\|(0,0,1,-1,-1,-1)^\top\|} = -\frac{3}{\sqrt{4}} = -\frac{3}{2}.$$

Výraz v nezkráceném čitateli ukazuje, jaká bude hodnota účelové funkce v novém bazickém přípustném bodu $x^{[1]}$ (tj. je stejná jako příslušné hodnoty v řádku $z_k - c_k$ v Tabulce 2.13). Výraz ve jmenovateli pak vyjadřuje vzdálenost BPB $x^{[0]}$ a $x^{[1]}$, mezi kterými v tomto kroku přecházíme. Chceme, aby se hodnota účelové funkce na jednotkový krok zmenšila co nejvíce, takže podle pravidla nejstrmější hrany vybereme třetí volbu, ve které je dalším BPB $x^{[1]} = [0,0,1,1,3,0]$. Pokles na jednotkový krok je totiž v tomto případě největší možný, neboť

$$\left|-\frac{3}{2}\right| > \left|-\frac{2}{\sqrt{3}}\right| > \left|-\frac{1}{\sqrt{3}}\right|.$$

Až doposud se nám vždy podařilo dané úlohy vyřešit pomocí simplexového algoritmu. Jenže to už nemusí být pravda ve chvíli, kdy je daná úloha degenerovaná, však také viz předpoklady Vět 2.7.2–2.7.3 a 2.8.2.

Příklad 2.8.11

S pomocí simplexového algoritmu vyřešme úlohu

$$-4x_1 - 6x_2 \to \min$$

za podmínek

$$6x_1 + 4x_2 \leq 24 \quad \& \quad 5x_1 + 10x_2 \leq 40 \quad \& \quad x_1 \geq 0 \quad \& \quad 0 \leq x_2 \leq 3.$$

Řešení. Převodem do kanonického tvaru dostaneme úlohu

$$-4x_1 - 6x_2 \to \min$$

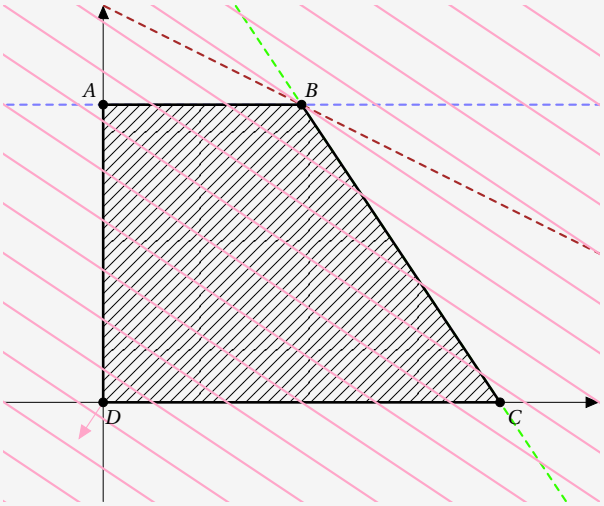
$$6x_1 + 4x_2 + s_3 = 24 \quad \& \quad 5x_1 + 10x_2 + s_4 = 40 \quad \& \quad x_2 + s_5 = 3$$

$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad s_3 \geq 0 \quad \& \quad s_4 \geq 0 \quad \& \quad s_5 \geq 0.$$

Odpovídající matice A a vektor b jsou

$$A = \begin{pmatrix} 6 & 4 & 1 & 0 & 0 \\ 5 & 10 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 24 \\ 40 \\ 3 \end{pmatrix},$$

z čehož vidíme, že úloha je degenerovaná, neboť vektor b lze získat jako lineární kombinaci pouze dvou sloupců matice A, viz také Obrázek 2.36 a omezení určující bod B.



Obrázek 2.36: Kompletní vizualizace dané úlohy se zvýrazněním normálového vektoru vrstevnic.

Za výchozí BPB můžeme vzít $x^{[0]} = [0,0,24,40,3]$ s bází $B(x^{[0]}) = \{s_3, s_4, s_5\}$ a pustíme se do samotného výpočtu.

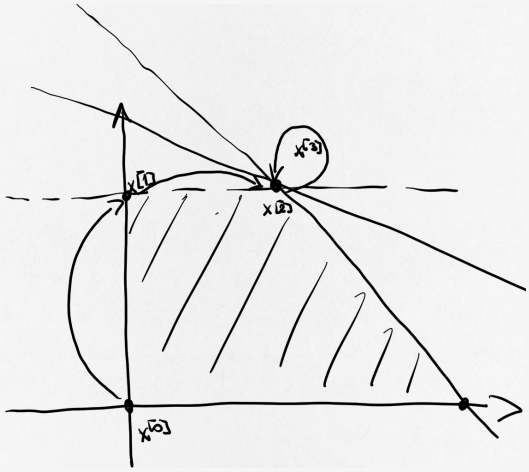
c_B	B	x_1	x_2	s_3	s_4	s_5	x_B	x_B/λ
0	s_3	6	4	1	0	0	24	6
0	s_4	5	10	0	1	0	40	4
0	s_5	0	①1	0	0	1	3	3 →
	c	−4	−6	0	0	0		
	z_k	0	0	0	0	0		
	$z_k - c_k$	4	6 ↑	0	0	0	0	
	s_3	6	0	1	0	−4	12	2
	s_4	①5	0	0	1	−10	10	2 →
	x_2	0	1	0	0	1	3	✖
	$z_k - c_k$	4 ↑	0	0	0	−6	−18	
	s_3	0	0	1	−6/5	①8	0	0 →
	x_1	1	0	0	1/5	−2	2	✖
	x_2	0	1	0	0	1	3	3
	$z_k - c_k$	0	0	0	−4/5	2 ↑	−26	

c_B	B	x_1	x_2	s_3	s_4	s_5	x_B	x_B/λ
	s_5	0	0	1/8	−6/40	1	0	
	x_1	1	0	1/4	−1/10	0	2	
	x_2	0	1	−1/8	3/20	0	3	
	$z_k - c_k$	0	0	−1/4	−1/2	0	−26	

Hned ve druhém kroku jsme dostali možnost výběru pro volbu klíčového řádku, což je důsledek degenerovanosti stejně jako nulová hodnota proměnné s_3 ve třetím kroku. Postupně jsme prošli tyto BPB:

$$x^{[0]} = [0,0,24,40,3] \rightsquigarrow x^{[1]} = [0,3,12,10,0] \rightsquigarrow x^{[2]} = [2,3,0,0,0] \rightsquigarrow x^{[3]} = [2,3,0,0,0].$$

V předposledním kroku jsme se dostali do degenerovaného BPB. Naštěstí jsme ale již po jednom dalším kroku „dorazili do řešení“. Ovšem to je fakticky stejné jako BPB v předposledním kroku, neboť jsme pouze změnili bázi v indexech odpovídajících nulovým hodnotám, viz také Obrázek 2.37.



Obrázek 2.37: Vizualizace průběhu simplexového algoritmu v původních proměnných.

Takto vzniká tzv. *cyklus*[145], který v případě konečné délky (jako před chvílí) způsobí pouze *zdržení*[146] nalezení řešení. Jenže situace může být i mnohem horší než v předchozím příkladě a může dojit i k nekonečné smyčce, skrze niž se „nikdy“ nedostaneme k řešení[147,148]. Toto ale není v rozporu s tvrzením Věty 2.7.2, která *de facto* zaručuje konvergenci simplexové metody. V degenerovaném případě se totiž může stát, že navzdory změně báze a nalezení nového BPB *nedojde* ke „zlepšení“ hodnoty účelové funkce, viz také komentář před Větou 2.8.2. Důkaz konvergence simplexové metody pochopitelně pochází od Dantziga a je (nejspíše) datován roky 1947–1948. Je samozřejmě snadné zkonstruovat degenerovanou úlohu (LP), ovšem docílit současně nějakého cyklu při řešení simplexovým algoritmem již vyžaduje nemalé úsilí. První takový příklad pochází od Hoffmana z roku 1951[149]:

$$\frac{\cos\varphi - 1}{\cos\varphi} x_1 + w\,x_2 + 2w\,x_4 + 4\sin^2\varphi\,x_5 + w\,(2 - 4\cos^2\varphi)\,x_6 + 4\sin^2\varphi\,x_7 + w\,(1 - 2\cos\varphi)\,x_8 \to \min$$

za podmínek

$$\begin{aligned} x_9 &= 1 \\ \cos\varphi\,x_1 - w\cos\varphi\,x_2 + \cos 2\varphi\,x_3 - 2w\cos^2\varphi\,x_4 + \cos 2\varphi\,x_5 + & \\ + 2w\cos^2\varphi\,x_6 + \cos\varphi\,x_7 + w\cos\varphi\,x_8 + x_{10} &= 0 \end{aligned}$$

[145] Angl. *cycle* nebo *cycling*. V Dantzigově knize [117] se hovoří i o *circling*.

[146] Angl. *stalling*.

[147] K cyklu nebo zdržení může dojít vždy jenom v důsledku toho, že jsme se dostali do degenerovaného BPB. Pak obvykle jeden či více následujících BPB bývají také degenerované a jde o to, zda nás některý z následujících degenerovaných BPB „dostane ven z této šlamastyky“ nebo se zamotáme v nějaké smyčce degenerovaných BPB a po několika krocích se vrátíme k některému z předchozích BPB a odpovídající simplexové tabulce.

[148] Vznik cyklu (nekonečné smyčky) je ale jediná situace, kdy simplexový algoritmus může selhat, tj. *pokud simplexový algoritmus neskončí v konečném čase, pak musí dojít k zacyklení.* Vskutku, již totiž víme, že BPB je jen konečný počet, takže v případě zacyklení musí simplexový algoritmus projít alespoň jedním BPB dvakrát. Jelikož z důkazu Věty 2.8.2 a ze vzorců z úvodu Podkapitoly 2.7 je zřejmé, že v takovém případě budou obě odpovídající simplexové tabulky totožné (tj. báze každého BPB jednoznačně určuje celou simplexovou tabulku), plyne odtud, že v algoritmu došlo k zacyklení.

[149] Hofmannova zpráva s tímto příkladem byla „publikována“ v roce 1953, ale ke „skutečnému“ zveřejnění došlo až v roce 2003 v [268, str. 177–181]

$$\frac{\operatorname{tg}\varphi\sin\varphi}{w}x_1+\cos\varphi\,x_2+\frac{\operatorname{tg}\varphi\sin 2\varphi}{w}x_3+\cos 2\varphi\,x_4-$$
$$-\frac{2\sin^2\varphi}{w}x_5+\cos 2\varphi\,x_6-\frac{\operatorname{tg}\varphi\sin\varphi}{w}x_7+\cos\varphi\,x_8+x_{11}=0$$
$$x_1\ge 0\quad\&\quad x_2\ge 0\quad\&\quad x_3\ge 0\quad\&\quad x_4\ge 0\quad\&\quad x_5\ge 0$$
$$x_6\ge 0\quad\&\quad x_7\ge 0\quad\&\quad x_8\ge 0\quad\&\quad x_9\ge 0\quad\&\quad x_{10}\ge 0\quad\&\quad x_{11}\ge 0,$$

kde $\varphi=2\pi/5$ a w je libovolné číslo větší[150] než $\frac{1-\cos 2\varphi}{1-2\cos\varphi}\approx 4{,}74$. Potom dostaneme cyklus délky 10 (tj. první a jedenáctá simplexová tabulka budou totožné) bez toho, aniž by simplexový algoritmus ukázal, že výchozí BPB s $x^{[0]}=[0,\ldots,0,1,0,0]$ a všechny další body dávají minimální (tj. nulovou) hodnotu účelové funkce. Při výpočtu se využívá skutečnost, že pro zvolenou hodnotu φ platí (i) $\cos 2\varphi=\cos 3\varphi$, (ii) $\sin 2\varphi=-\sin 3\varphi$ a (iii)[151] $\cos 2\varphi+\cos\varphi=2\cos\varphi\cos 2\varphi$. Také bychom měli poznamenat, že hlavní rolí prvního omezení $x_9=1$ je výrazné zjednodušení nalezení výchozího BPB. Při degenerovaných úlohách se navíc stává, že máme více možností pro volbu klíčového řádku, takže pro dosažení zmíněného cyklu je nutné použít pravidlo *první volby* zmíněné v Poznámce 2.8.9.

Lee se v článku [351] z roku 1997 snaží objasnit genezi tohoto příkladu z algebraického i geometrického pohledu. Ale nejjednodušší odpověď na otázku „Jak Hoffman na tento příklad došel?“ nalezneme přímo v závěru Hoffmanově článku:

> *Finally, we regret that we are unable at this date to recall any details of the considerations that led to construction of the example, beyond the fact that the geometric meaning of processes I and II in the degenerate case was very much in the foreground. W. Jacobs and S. Gass of the U.S. Air Force have pointed out that if we let...*[152]

Pokračování tohoto komentáře se dotýká toho, že 2×2 matice vytvořená z prvního a druhého sloupce ve druhém a třetím omezení

$$A=\begin{pmatrix}\cos\varphi & -w\cos\varphi\\ \frac{\operatorname{tg}\varphi\sin\varphi}{w} & \cos\varphi\end{pmatrix}$$

tvoří tzv. *cyklickou množinu*: A^2 odpovídá třetímu a čtvrtému sloupci, A^3 pátému a šestému sloupci, A^4 sedmému a osmému sloupci a konečně $A^5=I$. Takové matice konečného řádu lze využít ke konstrukci dalších cyklických problémů, viz také [21, 209, 525, 526].

Po tomto Hoffmanově příkladu byla zveřejněna ještě celá řada dalších (a značně jednodušších) příkladů zacyklení simplexového algoritmu, viz Příklady 2.8.28–2.8.34 ve cvičení k této podkapitole nebo [517]. Avšak žádný z nich *není tak tajemný jako ten Hoffmanův*[153]. Hoffmanův příklad je totiž kromě prvenství unikátní také v jiném ohledu – koeficienty jsou dány pomocí trigonometrických funkcí. Avšak aby to nevyznělo nějak pohrdlivě k ostatním příkladům: chování simplexového algoritmu na degenerované úloze nelze a priori odhadnout (závisí to i na zvolených pravidlech), tj. degenerovanost úlohy ještě neznamená zacyklení – k němu dochází jen zřídka, a tudíž *sestrojit úlohu se zacyklením je složitý úkol*[154].

Je samozřejmě zcela přirozenou otázkou, jak dlouhý/krátký cyklus můžeme získat v závislosti na velikosti dané úlohy lineárního programování. V uvedeném příkladě máme 3 omezení (kromě podmínek na znaménko jednotlivých proměnných) a 11 proměnných. Lze jich mít i méně? Ano, dá se ukázat, že pro úlohu ve tvaru

$$\langle c,x\rangle\to\min\quad\&\quad Iy+Ax=b\quad\&\quad x\ge 0\quad\&\quad y\ge 0$$

s $n\times n$ jednotkovou maticí I a $A\in\mathbb{R}^{m\times(n-m)}$, kde n je celkový počet proměnných, $x\in\mathbb{R}^{n-m}$, $y\in\mathbb{R}^n$ a $m<n$, je možné získat cyklus v případě, že $m\ge 2$, $n\ge m+3$ a $n\ge 6$. Pro zacyklení v neoptimálním bodě je dokonce potřeba $m\ge 3$, $n\ge m+3$ a $n\ge 7$. Úloha (LP) s pouhými

[150] V Hoffmanově článku můžeme pro parametr ω najít požadavek, že *ω je libovolné číslo větší než* $\frac{1-\cos\varphi}{1-2\cos\varphi}\approx 1{,}81$. Tato podmínka se pak v souvislosti s Hoffmanovým příkladem objevuje po dobu více než 50 let snad v každém článku a monografii či učebnici. Jenže to je chybné! V článku [242] z roku 2007 (a o rok dříve již v [210] opravující chybu z předchozím článku [209], k čemuž došlo ale až po upozornění od prvního autora [242]) je totiž ukázáno, že tento odhad je pouze nutnou podmínkou pro cyklus, ale rozhodně není postačující. Už se asi nikdy nedozvíme, jak k této chybě došlo, ale bude jistě pěkné věřit, že vše bylo způsobeno pouze obyčejným psacím strojem, na kterém Hoffman zkrátka zapomněl v čitateli napsat 2. Situaci příliš neulehčil ani Dantzigův přepis Hoffmanova příkladu v [117, str. 229], kde se v simplexové tabulce objevuje výraz $\operatorname{tg}\varphi/\omega$. Jak jej interpretovat? Je tím míněno $(\operatorname{tg}\varphi)/\omega$ nebo $\operatorname{tg}(\varphi/\omega)$? Ta první možnost je správně...

[151] Tyto tři rovnosti jsou zmíněny v [126, Exercise 5.1, str. 150] jako užitečné pomůcky pro výpočet jednotlivých simplexových tabulek. Jako třetí je uvedena rovnost $\cos 2\varphi+\cos\varphi=\cos\varphi\cos 2\varphi$ avšak není těžké ověřit, že tato rovnost neplatí.

[152] *Na závěr je nám líto, že si v tuto chvíli nemůžeme vzpomenout na žádné podrobnosti z našich úvah, které vedly ke vzniku tohoto příkladu kromě toho, že geometrický význam procesů I a II v degenerovaném případě stál výrazně v popředí. W. Jacobs a S. Gass z U.S. Air Force upozornili na to, že když necháme...* Viz [268, str. 181].

[153] *None of these examples is as mysterious as Hoffman's.* Viz [351, str. 99].

[154] *Cycling is a rare phenomenon. In fact, constructing an LP problem on which the simplex method may cycle is difficult.* Viz [292, str. 33].

2 nebazickými proměnnými se nikdy nezacyklí. Pro dosažení cyklu v takové úloze je potřeba mít alespoň 6 proměnných, 2 rovnice a alespoň 3 nebazické proměnné. Pak nejkratší délka cyklu je 6. My si tuto situaci ilustrujeme na příkladu z roku 1955 od Bealeho[155], ve kterém jsou 3 rovnice, 7 proměnných a dojde v něm právě k cyklu délky 6 v neoptimálním bodě (tedy vzhledem k předchozímu komentáři máme „nejmenší“ možnou úlohu s takovým chováním), viz také [36, 206, 209, 369].

Příklad 2.8.12

Degenerovaná úloha:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp6

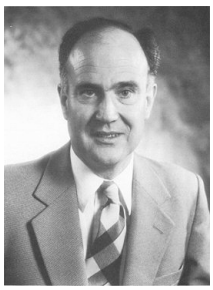
Zacyklení simplexového algoritmu lze naštěstí zabránit vhodnou volbou pravidel pro výběr klíčového prvku – jde o tzv. *anticyklická pravidla*[156]. Kromě již dříve zmíněného pravidla náhodného výběru můžeme použít např.

(i) *Blandovo*[157] *pravidlo* (1977, viz [68]): v případě rovnosti bereme nejnižší možný index pro určení klíčového řádku[158] a v případě klíčového sloupce volíme ten, který je nejvíce vlevo a splňuje $z_s - c_s > 0$ (tj. ignorujeme Dantzigovo pravidlo ohledně výběru nejvyšší možné hodnoty $z_s - c_s$). Toto pravidlo sice není úplně efektivní a konvergence může být velmi pomalá, ale dá se ukázat, že s tímto pravidlem nedojde k zacyklení.[159]

(ii) *Lexikografické pravidlo* (1955 navrženo Dantzigen, Ordenem[160] a Wolfem[161], viz [124]): na rozdíl od Blandova pravidla neudává kritérium pro výběr klíčového sloupce (doporučuje opět využít Dantzigovo pravidlo, tj. vybírat největší kladnou hodnotu v řádku $z_k - c_k$), ale při výběru klíčového řádku (v případě shody podílů) vyžaduje volbu toho, jehož složky z výchozí báze vytvoří v nové bázi ten *lexikograficky nejmenší*: ze dvou řádkových vektorů je lexikograficky menší ten, který má na první pozici s různými hodnotami menší číslo, tj. např. vektor $u = (0,3,1)$ je lexikograficky menší než vektor $v = (0,5,0)$, neboť první složka vektorů je stejná (rovna 0) a ve druhé složce má vektor u menší číslo neboli první nenulová složka rozdílu $u - v$ je záporná, což zapisujeme jako $u \prec v$, viz také Definici 3.5.9.

Přesněji řečeno, tentokrát nás nezajímá pouze hodnota podílu pravé strany a odpovídajícího koeficientu v klíčovém sloupci, tj. číslo ve sloupci x_B/λ, ale uvažujeme (téměř) celé řádky vyhovující podmínce na výběr klíčového řádku, které podělíme příslušným koeficientem v klíčovém sloupci λ_{ik}, čímž na pozici potenciálních pivotů dostaneme jedničku. Mezi těmto řádky pak vybereme ten lexikograficky nejmenší, přičemž pro funkčnost takto upravené simplexové metody je nutné uvažovat pouze sloupce odpovídající proměnným ve výchozí bázi B_0, a to v takovém pořadí, v jakém příslušné sloupce ve výchozí simplexové tabulce tvoří jednotkovou matici (tj. nejdříve bereme bazickou proměnnou, jejíž sloupec ve výchozí simplexové tabulce je e_1, atd.)[162]. Pak postupujeme tak, že nejdříve jako obvykle určíme podíly ve sloupci x_B/λ. Nastane-li toto minimum v jednom řádku, pak jej zvolíme jako klíčový. V případě, že minimum nastane ve více řádcích, podělíme tyto řádky odpovídajícími hodnotami v klíčovém sloupci a porovnáme hodnoty ve sloupci odpovídajícím první proměnné ve výchozí bázi B_0. Určíme minimum z těchto hodnot. Pokud minimum nastane v jednom řádku, zvolíme jej jako klíčový, zatímco v případě shody se u těchto řádků podíváme na hodnoty odpovídající druhé proměnné v bázi B_0 atd. Lineární nezávislost sloupců určených výchozí bází zaručuje, že tímto pravidlem vždy jednoznačně určíme klíčový řádek po nejvýše m krocích[163]. Nicméně už z toho stručného popisu je zřejmé, že implementace a početní náročnost tohoto pravidla bude mnohem vyšší než v případě předchozího Blandova pravidla.

(iii) *Perturbační metoda*: (navrženo Charnesem[164] v roce 1952, viz [282]): hlavní myšlenka se opírá o sporadičnost případné degenerovanosti, takže i malá změna pravé strany soustavy $Ax = b$ způsobí, že vznikne nedegenerovaná úloha (na pravé straně může být jakékoli číslo, a to že se tam objeví zrovna 0, je ve světě reálných čísel prakticky nemožné).

[155]


Evelyn Martin Lansdowne Beale (8. září 1928 – 23. prosince 1985) byl anglický matematik a statistik a jeden z prvních průkopníků matematického programování. V roce 1979 získal za „za jeho využití matematických a statistických dovedností v průmyslových problémech a za jeho přínos k teorii matematického programování“ ocenění *Fellowship of the Royal Society*, čímž se stal doživotním členem Královské společnosti v Londýně.

[156] Angl. *anti-cycling rules*. Není ale nezbytně nutné je aplikovat ihned – jsou totiž početně trochu náročnější. Je možné začít s klasickým simplexovým algoritmem a až ve chvíli, kdy budeme mít pro výběr klíčového řádku dva nebo více kandidátů, můžeme začít některé z těchto pravidel uplatňovat.

[157]

Robert Gary Bland (⋆25. února 1948) je americký matematik, který v současnosti působí jako profesor operačního výzkumu a informačního inženýrství na Cornellově univerzitě.

[158] Pokud v úloze (LP) nebudou všechny proměnné označené pouze jako $x_1, \ldots, x_n$, pak můžeme jejich pořadí určit tak, jak jsou zapsány v záhlavní výchozí simplexové tabulky (tj. více vlevo v záhlaví je totéž jako nižší index).

[159] Důkaz se může zdát být poněkud spletitý vzhledem k potřebě manipulace s větším množství různých indexů. Ale věříme, že pokud čtenář věnuje dostatečný čas jeho pochopení, nakonec jeho jistou eleganci a rafinovanost ocení.
Připusťme, že i s použitím Blandova pravidla došlo k zacyklení, tj. postupně jsme prošli BPB s bá-

zemi $B_0 \to B_1 \to B_2 \to B_3 \to \cdots \to B_\ell \to B_0$. Již víme, že pokud došlo k zacyklení, pak nedochází ke změně funkční hodnoty a jednotlivé BPB mají stejné souřadnice, ale různé báze, tj. i na bazických pozicích jsou nulové hodnoty, které se nezmění přechodem k další bázi. Proměnnou x_j nazveme *vrtkavou*, pokud je nebazickou pro jednu z těchto bází a bazickou pro jinou, tj. $j \in B_k$ a současně $j \notin B_s$ pro nějakou dvojici $k, s \in \{1,\dots,\ell\}$. Tyto vrtkavé proměnné jsou právě těmi, které v průběhu výpočtu simplexového algoritmu opouštějí báze a znovu do nich vstupují.

Nechť x_r je vrtkavá proměnná s nejvyšším indexem a $x_r \in B_i$, ale $x_r \notin B_{i+1}$ pro nějaké $i \in \{0,\dots,\ell-1\}$. Bez újmy na obecnosti můžeme vzít $i = 0$, tj. $r \in B_0$ a současně $r \notin B_1$. Současně nechť x_s je proměnná vstupující do báze při přechodu od B_0 k B_1, tj. $s \notin B_0$ a $s \in B_1$. Pak x_s je také vrtkavá proměnná a nutně

$$s < r.$$

V simplexové tabulce odpovídající B_0 máme

$$x_k = b_k - \sum_{j \notin B_i} \lambda_{kj}\, x_j \qquad (2.8.4)$$

pro $k \in B_0$ a místo účelové funkce můžeme psát

$$v + \sum_{j \notin B_i} \Delta_j\, x_j,$$

kde v je odpovídající hodnota pro aktuální BPB a $\Delta_j := z_j - c_j$ jsou odpovídající hodnoty z posledního řádku simplexové tabulky. Jelikož x_s vstupuje do báze, musí být

$$\Delta_s > 0,$$

a protože x_r vystupuje z báze a je to vrtkavá proměnná, musí být

$$\lambda_{rs} > 0 \quad \& \quad b_r = 0.$$

Druhá rovnost plyne z toho, že došlo k zacyklení a všechny odpovídající BPB musí být stejné, takže hodnota x_r musí být stejná ve chvíli, kdy je i není v bázi, tj. nulová.

Nechť nyní x_r znovu vstoupí do báze B^* odpovídající některé z $B_1,\dots,B_\ell$. Pak odpovídající $\hat{z}_r - \hat{c}_r > 0$ a účelová funkce je

$$v + \sum_{j \notin B^*} \hat{\Delta}_j\, x_j$$

s toutéž hodnotou v, $\hat{\Delta}_j := \hat{z}_j - \hat{c}_j$ jsou opět koeficienty z posledního řádku simplexové tabulky, které můžeme dodefinovat jako $\hat{\Delta}_j := 0$ pro $j \in B^*$. Jelikož jsme toto vyjádření získali pomocí algebraických úprav rovnic v simplexové tabulce pro B_0, musí tutéž hodnotu dá-

Metoda postupuje tak, že ke každému z prvků b_i přičteme ε^i pro nějaké dostatečně malé $\varepsilon > 0$, tj. místo vektoru x_B počítáme v simplexové tabulce s vektorem $x_B + (\varepsilon, \varepsilon^2 \dots, \varepsilon^m)^\top$, čímž ale původní úlohu příliš nezměníme. Touto perturbací získáme tzv. ϵ-úlohu, při jejímž řešení nemůže díky její nedegenerovanosti dojít k zacyklení[165]. Jinými slovy, s každým omezením „nepatrně pohneme“, čímž zaručíme to, že se v žádném bodě v $\mathbb{R}^n$ neprotne více než n nadrovin určujících jednotlivá omezení. Po vyřešení ϵ-úlohy (s využitím simplexové metody) získáme odstraněním perturbací (tj. volbou $\varepsilon = 0$) řešení původní úlohy.

Avšak tato metoda ve skutečnosti není třetím anticyklickým pravidlem – je to jen alternativní způsob pro implementaci lexikografického pravidla a/nebo naopak[166]. Říká se, že *perturbace a lexikografické uspořádání jsou dvě strany téže mince*, což si ilustrujeme v následujícím příkladě. Obě tato pravidla tedy dají tentýž průběh simplexového algoritmu. Perturbační metoda se sice může zdát být poněkud rafinovanější, ale její přesná implementace by vyžadovala symbolické výpočty s abstraktními veličinami, což je asi nepřekonatelný problém pro běžný výpočetní software. Proto může buď jednotlivým parametrům $\varepsilon_1,\dots,\varepsilon_m$ přiřadit nějaké velmi malé konkrétní hodnoty (pak je ale nedokážeme eliminovat, takže nejspíš dostaneme jen nějaké přibližné řešení, i když třeba velmi přesné) nebo můžeme raději využít lexikografické pravidlo.

Využití těchto pravidel si nyní vyzkoušíme na jednom klasickém příkladu degenerované úlohy lineárního programování, kterou v roce 1983 zveřejnil Václav Chvátal[167] ve své populární učebnici lineárního programování, viz [292, str. 30].

Příklad 2.8.13

Uvažme úlohu

$$-10\,x_1 + 57\,x_2 + 9\,x_3 + 24\,x_4 \to \min$$
$$\frac{1}{2}\,x_1 - \frac{11}{2}\,x_2 - \frac{5}{2}\,x_3 + 9\,x_4 \le 0$$
$$\frac{1}{2}\,x_1 - \frac{3}{2}\,x_2 - \frac{1}{2}\,x_3 + x_4 \le 0$$
$$0 \le x_1 \le 1 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Vyřešme ji standardním simplexovým algoritmem doplněným o pravidlo první volby a využijme také výše zmíněná anticyklická pravidla.

Řešení. Úlohu nejprve převedeme na kanonický tvar přidáním tří nezáporných doplňkových proměnných označených s_1, s_2, s_3. Pak odpovídající matice A obsahuje všechny sloupce jednotkové matice a můžeme snadno sestavit výchozí simplexovou tabulku pro bázi tvořenou těmito doplňkovými proměnnými. Tato tabulka je společně s dalšími kroky zobrazena níže. V ní vidíme, že po šesti krocích simplexového algoritmu dospějeme ke stejné tabulce jako na začátku. Navíc ani v jednom kroku se algoritmus „nepohne“ z výchozího BPB $[0,0,0,0,0,0,1]$, ale dochází pouze k obměně báze u proměnných s nulovou hodnotou.

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
0	s_1	(1/2)	−11/2	−5/2	9	1	0	0	0	0 →
0	s_2	1/2	−3/2	−1/2	1	0	1	0	0	0
0	s_3	1	0	0	0	0	0	1	1	1
	c	−10	57	9	24	0	0	0		
	z_k	0	0	0	0	0	0	0		
	$z_k - c_k$	10 ↑	−57	−9	−24	0	0	0	0	

vat i libovolný bod vyhovující soustavě (2.8.4), který získáme např. zvětšením hodnoty x_s a ponecháním všech ostatních nebazických hodnot nulových (nehledě na to, že by mohlo dojít k porušení přípustnosti)

$$\widehat{x}_k := \begin{cases} y, & k = s \notin B_0, \\ 0, & k \notin B_0 \ \&\ k \neq s, \\ b_k - \lambda_{ks}\, y, & k \in B_0 \end{cases}$$

pro libovolné y. Pak dosazením do obou vyjádření účelové funkce dostaneme rovnost

$$v + \Delta_s\, y = v + \hat{\Delta}_s\, y + \\ + \sum_{k \in B_0} \hat{\Delta}_k\, (b_k - \lambda_{ks}\, y),$$

takže po úpravě máme

$$(\Delta_s - \hat{\Delta}_s + \sum_{k \in B_0} \hat{\Delta}_k\, \lambda_{ks})\, y = \\ = \sum_{k \in B_0} \hat{\Delta}_k\, b_k.$$

Jelikož pravá strana této rovnosti nezávisí na y, musí být

$$\Delta_s - \hat{\Delta}_s + \sum_{k \in B_0} \hat{\Delta}_k\, \lambda_{ks} = 0.$$

Protože x_s nevstupuje do báze B^*, je $\hat{\Delta}_s \leq 0$. Současně víme, že $\Delta_s > 0$, takže nutně

$$\sum_{k \in B_0} \hat{\Delta}_k\, \lambda_{ks} < 0,$$

a tudíž musí existovat index $q \in B_0$ takový, že

$$\hat{\Delta}_q\, \lambda_{qs} < 0. \qquad (2.8.5)$$

Tj. $\hat{\Delta}_q \neq 0$, takže $q \notin B^*$, a tedy x_q je také vrtkavá proměnná, takže $q \leq r$. Dokonce platí $q < r$, protože $\hat{\Delta}_r\, \lambda_{rs} > 0$, což plyne z toho, $\hat{\Delta}_r$ kvůli proměnné x_r vstupující do báze B^* a $\lambda_{rs} >$ kvůli x_r opouštějící bázi B_0. Nerovnost $q < r$ implikuje, že $\hat{\Delta}_r \leq 0$, protože v opačném případě by v souladu s Blandovým pravidlem musela proměnná x_q vstoupit do báze. Z nerovnosti (2.8.5) proto vyplývá, že $\lambda_{qs} > 0$. Navíc protože x_q je také vrtkavá proměnná, musí být $b_q = 0$. To ale znamená, že $b_q / \lambda_{qs} = 0$, a tedy i proměnná x_q byla kandidátem na proměnnou opouštějící bázi B_0. Nejenom to, vzhledem k nerovnosti $q < r$ a Blandovu pravidlu to dokonce měla být ona, čímž jsme docílili požadovaného sporu $\lightning$. Viz také [491, Theorem 3.3] nebo [292, Theorem 3.3].

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
	x_1	1	−11	−5	18	2	0	0	0	✖
	s_2	0	(4)	2	−8	−1	1	0	0	0 →
	s_3	0	11	5	−18	−2	0	1	1	1/11
	$z_k - c_k$	0	53 ↑	41	−204	−20	0	0	0	
	x_1	1	0	(1/2)	−4	−3/4	11/4	0	0	0 →
	x_2	0	1	1/2	−2	−1/4	1/4	0	0	0
	s_3	0	0	−1/2	4	3/4	−11/4	1	1	✖
	$z_k - c_k$	0	0	29/2 ↑	−98	−27/4	−53/4	0	0	
	x_3	2	0	1	−8	−3/2	11/2	0	0	✖
	x_2	−1	1	0	(2)	1/2	−5/2	0	0	0 →
	s_3	1	0	0	0	0	0	1	1	✖
	$z_k - c_k$	−29	0	0	18 ↑	15	−93	0	0	
	x_3	−2	4	1	0	(1/2)	−9/2	0	0	0 →
	x_4	−1/2	1/2	0	1	1/4	−5/4	0	0	0
	s_3	1	0	0	0	0	0	1	1	✖
	$z_k - c_k$	−20	−9	0	0	21/2 ↑	−141/2	0	0	
	s_1	−4	8	2	0	1	−9	0	0	✖
	x_4	1/2	−3/2	−1/2	1	0	(1)	0	0	0 →
	s_3	1	0	0	0	0	0	1	1	✖
	$z_k - c_k$	22	−93	−21	0	0	24 ↑	0	0	
	s_1	(1/2)	−11/2	−5/2	9	1	0	0	0	0 →
	s_2	1/2	−3/2	−1/2	1	0	1	0	0	0
	s_3	1	0	0	0	0	0	1	1	1
	$z_k - c_k$	10 ↑	−57	−9	−24	0	0	0	0	

Tabulka 2.14: Chvátalův příklad zacyklení a jeho řešení klasickým simplexovým algoritmem.

Již v prvním kroku jsme měli na výběr pro proměnné opouštějící báze – buď s_1 nebo s_2, a my jsme v souladu s pravidlem první volby zvolili s_1 a takto pokračovali dále. Kdybychom využili Blandovo pravidlo, tak by prvních pět kroků proběhlo úplně stejně. Avšak další krok (tj. sedmá simplexová tabulka) by již byl jiný a tato změna nás po jednom dalším kroku přivede k cíli, viz Tabulku 2.15 níže. Odtud vidíme, že celkově po sedmi krocích (a osmi tabulkách) simplexový algoritmus s Blandovým pravidlem dospěje k řešení $[1,0,1,0,2,0,0]$, čemuž odpovídá řešení původní úlohy $x^* = [1,0,1,0]$ s minimální hodnotou -1.

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
	s_1	−4	8	2	0	1	−9	0	0	✖
	x_4	(1/2)	−3/2	−1/2	1	0	1	0	0	0 →
	s_3	1	0	0	0	0	0	1	1	1
	$z_k - c_k$	22 ↑	−93	−21	0	0	24	0	0	

[160]

Alex Orden (9. srpna 1916 – 9. února 2008) byl americký matematik, který patří k průkopníkům lineárního programování. Doktorát získal v roce 1948 na MIT. Po šesti letech strávených v Burroughs Corporation na oddělení počítačového programování a rozvoje jeho využití nastoupil v roce 1958 na University of Chicago Graduate School of Business jako profesor aplikované matematiky se zaměřením na lineární a nelineární programování, kde zůstal až do roku 1987. V letech 1947–1952 pracoval na Projektu SCOOP, kde se zabýval transformací analytických technik letectva Air Force ze stolních kalkulátorů do doby digitálních počítačů. Když pracoval v Národním úřadu pro standardy (již dříve zmíněný National Bureau of Standards), zasazoval se o převod finančních prostředků letectva ve výši 400 000 USD do tohoto úřadu na zhotovení návrhu a následnou konstrukci prvního programovatelného počítače na světě. Díky tomu se posunul jeho vývoj o několik let a Air Force se tím stal prvním vlastníkem komerčního počítače dodaného mimo továrnu. Tento počítač byl v roce 1952 instalován v suterénu Pentagonu.

[161] O něm si necháme podrobnosti na později.

[162] Alternativně by bylo možné přeskládat sloupce v simplexové tabulce tak, že nejdříve budeme mít sloupec x_B, pak sloupce odpovídající výchozím bazickým proměnným z B_0 ve stejném pořadí jako je máme uspořádány v řádcích (tj. pořadí sloupců bude po celou dobu stejné jako ve sloupci B ve výchozí bázi B_0), díky čemuž bude první část výchozí simplexové tabulky tvořena $m \times m$ jednotkovou maticí. Nakonec umístíme sloupce odpovídající nebazickým proměnným. Takto bude celá výchozí tabulka tvořena lexikograficky kladnými řádky (kromě řádku $z_k - c_k$).

[163] Jinými slovy, jestliže x_s má podle Dantzigova pravidla vstoupit do báze, pak podle simplexového algoritmu může z báze vystoupit kterákoli proměnná s inde-

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
	s_1	0	−4	−2	8	1	−1	0	0	✖
	x_1	1	−3	−1	2	0	2	0	0	✖
	s_3	0	3	(1)	−2	0	−2	1	1	1 →
	$z_k - c_k$	0	−27	1 ↑	−44	0	−20	0	0	
	s_1	0	2	0	4	1	−5	2	2	
	x_1	1	0	0	0	0	0	1	1	
	x_3	0	3	1	−2	0	−2	1	1	
	$z_k - c_k$	0	−30	0	−42	0	−18	−1	−1	

Tabulka 2.15: Chvátalův příklad zacyklení a jeho řešení s pomocí Blandova pravidla.

Při řešení Chvátalovy úlohy s pomocí lexikografického pravidla musíme již v prvním kroku zjistit, který z řádků odpovídajících proměnným s_1 a s_2 je lexikograficky menší. Vynásobením řádků pro s_1 a s_2 číslem 2 a přeskládáním sloupců počínaje x_B a dále dle výchozí báze tvořené postupně s_1, s_2, s_3 dostaneme dvojici vektorů z Tabulky 2.16. Hodnoty označené symbolem „✖“ jsou pro lexikografické pravidlo nepodstatné, takže není nutné je zde uvádět. Pro jistotu jsme navíc v záhlaví zdůraznili, že všechny koeficienty byly poděleny odpovídajícím prvkem v klíčovém sloupci. V tabulce je pro úplnost uveden i řádek pro s_3, ze kterého vidíme, že tato část simplexové tabulky je skutečně tvořena lexikograficky kladnými řádky a současně že zvolený řádek pro s_2 je opravdu lexikograficky nejmenší.

	x_B/λ	s_1/λ	s_2/λ	s_3/λ	x_1/λ	x_2/λ	x_3/λ	x_4/λ
s_1	0	2	0	0	✖	✖	✖	✖
s_2	0	0	2	0	✖	✖	✖	✖
s_3	1	0	0	1	✖	✖	✖	✖

Tabulka 2.16: Výpočet klíčového řádku v prvním kroku řešení Chvátalova příkladu pomocí lexikografického pravidla.

Další průběh výpočtu bude tudíž jiný než v předchozích případech. V dalším kroku již není nutné lexikografické pravidlo uplatňovat a ve třetí simplexové tabulce již nalezneme řešení. Celý výpočet můžeme vidět v Tabulce 2.17.

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
0	s_1	1/2	−11/2	−5/2	9	1	0	0	0	0
0	s_2	(1/2)	−3/2	−1/2	1	0	1	0	0	0 →
0	s_3	1	0	0	0	0	0	1	1	1
	c	−10	57	9	24	0	0	0		
	z_k	0	0	0	0	0	0	0		
	$z_k - c_k$	10 ↑	−57	−9	−24	0	0	0	0	
	s_1	0	−4	−2	8	1	−1	0	0	✖
	x_1	1	−3	−1	2	0	2	0	0	✖
	s_3	0	3	(1)	−2	0	−2	1	1	1 →
	$z_k - c_k$	0	−27	1 ↑	−44	0	−20	0	0	

xem $r \in B$, pro kterou má podíl $\frac{x_r}{\lambda_{rs}}$ nejmenší hodnotu, tj. volíme mezi indexy z množiny

$$I_0 := \{r \in B \mid r \text{ vyhovuje } (2.8.1)\}.$$

Je-li tato množina I_0 jednoprvková, pak volba vystupující proměnné je jasná. V opačném případě otestujeme hodnoty podílů těchto řádků s hodnotami v klíčovém sloupci. Nejdříve uvážíme podíly ve sloupci určeném první bazickou proměnnou ve výchozí bázi B_0, tj. sestavíme množinu

$$I_1 := \left\{ r \in I_0 \,\middle|\, \frac{\lambda_{r\ell_1}}{\lambda_{rs}} = \min_{k \in I_0} \left\{ \frac{\lambda_{k\ell_1}}{\lambda_{ks}} \right\} \right\},$$

kde $\ell_1 \in B_0$ je první bazická proměnná, tj. taková, že sloupec pro x_{ℓ_1} ve výchozí simplexové tabulce je tvořen vektorem e_1. Je-li již množina I_1 jednoprvková, tak odpovídající proměnná vystoupí z báze. V opačném případě sestavíme množinu I_2 atd., přičemž pro konstrukci těchto dalších množin platí

$$I_j := \left\{ r \in I_{j-1} \,\middle|\, \frac{\lambda_{r\ell_j}}{\lambda_{rs}} = \min_{k \in I_{j-1}} \left\{ \frac{\lambda_{k\ell_j}}{\lambda_{ks}} \right\} \right\},$$

kde $\ell_j \in B_0$ je j-tá proměnná ve výchozí bázi B_0, tj. taková, že sloupec x_{ℓ_j} ve výchozí simplexové tabulce je tvořen vektorem e_j.

K určení klíčového řádku budeme v nejhorším případě potřebovat $(m+1)$-tici množin $I_0, I_1, \ldots, I_m$. Kdyby totiž i množina I_m obsahovala více než jeden index, znamenalo by to, že řádkové vektory tvořené postupně složkami ve sloupcích pro $x_{\ell_1}, \ldots, x_{\ell_m}$ jsou proporcionální, a tedy i lineárně závislé, neboť toto je možné pouze v případě, že podělením odpovídajícími hodnotami v klíčovém sloupci dostaneme stejné vektory. Avšak tyto vektory vznikly algebraickými úpravami vektorů, které na začátku tvořili jednotkovou matici, což je očividný spor $\lightning$.

Při tomto přeskládání jednotlivých sloupců budeme při použití lexikografického pravidla vždy mít lexikograficky kladné řádky (vyjma řádku $z_k - c_k$). Vskutku, označme $m+1$ složek těchto řádků (počínaje sloupcem x_B a následovaným sloupci ve výše zmíněném pořadí pro bázi B_0) v bázi B jako $\boldsymbol{\alpha}_i$, tj.

$$\boldsymbol{\alpha}_i := (b_i, \lambda_{i\ell_1}, \ldots, \lambda_{i\ell_m}) \quad \text{pro } i \in B,$$

kde b_i je příslušná hodnota ve sloupci x_B a $\lambda_{i\ell_j}$ jsou hodnoty v i-tém řádku a ℓ_j-tém sloupci simplexové tabulky pro $\ell_1, \ldots, \ell_m \in B_0$. Nechť jsou všechny tyto řádkové

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
	s_1	0	2	0	4	1	−5	2	2	
	x_1	1	0	0	0	0	0	1	1	
	x_3	0	3	1	−2	0	−2	1	1	
	$z_k - c_k$	0	−30	0	−42	0	−18	−1	−1	

Tabulka 2.17: Chvátalův příklad zacyklení a jeho řešení s pomocí lexikografického pravidla.

Řešení jsme očividně nalezli výrazně rychleji než v případě Blandova pravidla, avšak hledání lexikograficky nejmenšího řádku je zase výpočetně poněkud náročnější. V Tabulkách 2.18– 2.19 jsou uvedeny jednotlivé přeskládané řádky pro bazické proměnné ze druhé a třetí simplexové tabulky, abychom názorně viděli, že všechny jsou stále lexikograficky kladné a že řádek pro s_3 ve druhé tabulce je skutečně lexikograficky nejmenší (v uvedené podobě). Je zde uveden také řádek $z_k - c_k$, abychom mohli snadno ověřit, že je lexikograficky klesající.

	x_B/λ	s_1/λ	s_2/λ	s_3/λ	x_1/λ	x_2/λ	x_3/λ	x_4/λ
s_1	0	1	−1	0	✖	✖	✖	✖
x_1	0	0	2	0	✖	✖	✖	✖
s_3	1	0	−2	1	✖	✖	✖	✖
$z_k - c_k$	0	0	−20	0	✖	✖	✖	✖

Tabulka 2.18: Ověření lexikografického uspořádání po prvním kroku řešení Chvátalova příkladu pomocí lexikografického pravidla.

	x_B/λ	s_1/λ	s_2/λ	s_3/λ	x_1/λ	x_2/λ	x_3/λ	x_4/λ
s_1	2	1	−5	2	✖	✖	✖	✖
x_1	1	0	0	1	✖	✖	✖	✖
x_3	1	0	−2	1	✖	✖	✖	✖
$z_k - c_k$	−1	0	−18	−1	✖	✖	✖	✖

Tabulka 2.19: Ověření lexikografického uspořádání po druhém kroku řešení Chvátalova příkladu pomocí lexikografického pravidla.

Při řešení Chvátalovy úlohy perturbační metodou nahradíme vektor x_B v prvním kroku vektorem

$$b + \epsilon_0 = \begin{pmatrix} \varepsilon \\ \varepsilon^2 \\ 1 + \varepsilon^3 \end{pmatrix}.$$

Do výchozí simplexové tabulky doplníme tento vektor a „spustíme“ simplexový algoritmus. Tato změna výchozí tabulky oproti začátku Tabulky 2.14 ovlivní výběr klíčového řádku. Samotný výpočet pak nalezneme v Tabulce 2.20.

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
0	s_1	1/2	−11/2	−5/2	9	1	0	0	ε	2ε
0	s_2	(1/2)	−3/2	−1/2	1	0	1	0	ε^2	$2\varepsilon^2 \rightarrow$
0	s_3	1	0	0	0	0	0	1	$1+\varepsilon^3$	$1+\varepsilon^3$
	c	−10	57	9	24	0	0	0		
	z_k	0	0	0	0	0	0	0		
	$z_k - c_k$	10 ↑	−57	−9	−24	0	0	0	0	

vektory $\boldsymbol{\alpha}_i$ lexikograficky kladné. Jestliže proměnná x_s má v dalším kroku vstoupit do báze a současně má podle lexikografického pravidla vystoupit proměnná x_r, pak v nové tabulce budeme mít místo řádku $\boldsymbol{\alpha}_r$ řádek

$$\widetilde{\boldsymbol{\alpha}}_s = \frac{1}{\lambda_{rs}}\boldsymbol{\alpha}_r,$$

který je opět lexikograficky kladný díky $\lambda_{rs} > 0$. Všechny ostatní řádky pro $i \in B\setminus\{r\}$ se změní na

$$\widetilde{\boldsymbol{\alpha}}_i = \boldsymbol{\alpha}_i - \frac{\lambda_{is}}{\lambda_{rs}}\boldsymbol{\alpha}_r = \\ = \lambda_{is}\Big(\frac{\boldsymbol{\alpha}_i}{\lambda_{is}} - \frac{\boldsymbol{\alpha}_r}{\lambda_{rs}}\Big).$$

Je-li $\lambda_{is} \le 0$, pak v první rovnosti přičítáme k lexikograficky kladnému vektoru $\boldsymbol{\alpha}_i$ lexikograficky nezáporný vektor, takže jistě $\widetilde{\boldsymbol{\alpha}}_i > 0$. Je-li $\lambda_{is} > 0$, pak $\widetilde{\boldsymbol{\alpha}}_i > 0$ plyne ze druhé rovnosti, ve které je díky použitému pravidlu lexikograficky kladný vektor $\frac{\boldsymbol{\alpha}_i}{\lambda_{is}} - \frac{\boldsymbol{\alpha}_r}{\lambda_{rs}}$ násobený kladným číslem – prvních j složek tohoto vektoru bude nulových (podle toho, která z množin I_j je jednoprvková) a první nenulová složka bude kladná díky volbě lexikograficky nejmenšího $\boldsymbol{\alpha}_r$.

Konečnost (tj. „nezacyklovatelnost“) simplexového algoritmu doplněného o lexikografické pravidlo plyne z toho, že část řádku $z_k - c_k$ upravená stejně jako v předchozím případě (včetně aktuální hodnoty účelové funkce na první pozici) je přechodem k nové bázi lexikograficky klesající, tj. přechodem k nové bázi dostaneme na jeho pozici lexikograficky menší řádek. Vskutku, označme přeskládanou část řádku $z_k - c_k$ jako $\boldsymbol{\gamma}$, tj.

$$\boldsymbol{\gamma} := (z, z_{\ell_1} - c_{\ell_1}, \ldots, z_{\ell_m} - c_{\ell_m}),$$

kde z je hodnota účelové funkce v aktuálním BPB. Pak přechodem k nové simplexové tabulce získané pomocí pivota λ_{rs} odtud dostaneme nový řádek

$$\widetilde{\boldsymbol{\gamma}} = \boldsymbol{\gamma} - \frac{z_s - c_s}{\lambda_{rs}}\boldsymbol{\alpha}_r$$

neboli

$$\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma} = -\frac{z_s - c_s}{\lambda_{rs}}\boldsymbol{\alpha}_r.$$

Protože $\boldsymbol{\alpha}_r$ je lexikograficky kladný, bude v případě $z_s - c_s > 0$ řádek na pravé straně lexikograficky záporný, takže

$$\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma} < 0 \quad \text{neboli} \quad \widetilde{\boldsymbol{\gamma}} \prec \boldsymbol{\gamma}.$$

To ale znamená, že řádek $\boldsymbol{\gamma}$ se v průběhu výpočtu lexikograficky zmenšuje, a tudíž není možné se v průběhu výpočtu vrátit k téže bázi, což vzhledem ke konečnému počtu BPB ukazuje, že tento upravený simplexový algoritmus sku-

c_B	B	x_1	x_2	x_3	x_4	s_1	s_2	s_3	x_B	x_B/λ
	s_1	0	-4	-2	8	1	-1	0	$\varepsilon-\varepsilon^2$	✖
	x_1	1	-3	-1	2	0	2	0	$2\varepsilon^2$	✖
	s_3	0	3	(1)	-2	0	-2	1	$1-\varepsilon^2(2+\varepsilon)$	$1-\varepsilon^2(2+\varepsilon)\rightarrow$
	z_k-c_k	0	-27	$1\uparrow$	-44	0	-20	0	$-20\varepsilon^2$	
	s_1	0	2	0	4	1	-5	2	$2(1+\varepsilon^3)+\varepsilon-5\varepsilon^2$	
	x_1	1	0	0	0	0	0	1	$1+\varepsilon^3$	
	x_3	0	3	1	-2	0	-2	1	$1-2\varepsilon^2+\varepsilon^3$	
	z_k-c_k	0	-30	0	-42	0	-18	-1	$-1-18\varepsilon^2-\varepsilon^3$	

Tabulka 2.20: Chvátalův příklad zacyklení a jeho řešení s pomocí perturbační metody.

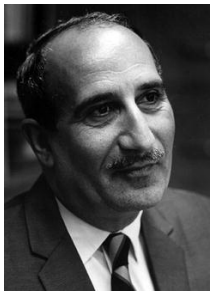
V prvním kroku do báze vstoupí proměnná x_1 a bázi opustí proměnná s_2, neboť díky volbě dostatečně malého $\varepsilon \in (0,1)$ platí $1+\varepsilon^3 > 2\varepsilon > 2\varepsilon^2$. Ve druhém kroku máme při výběru pivota pouze jednu možnost, do báze musí vstoupit x_3 a vystoupit musí s_3. Avšak tímto postupem jsme již v pouhých dvou krocích (a třech tabulkách) dospěli k řešení ε-úlohy. Následným vynecháním perturbací získáme řešení původní úlohy, tj. „vynulováním“ ε v poslední simplexové tabulce obdržíme stejnou tabulku jako při využití Blandova a lexikografického pravidla. Ba co víc, tímto vynulováním ε budou všechny tabulky z perturbační metody zcela shodné s tabulkami získanými lexikografickým pravidlem, jak jsme již avizovali v úvodním komentáři.

V literatuře lze nalézt i řadu dalších anticyklických pravidel, např. pravidlo anglicky označované jako *last in, first out & last out, first in*, ve kterém v případě více kandidátů vstupuje do báze ta proměnná, která se stala nebazickou nejpozději, a totéž platí i pro proměnné vystupující z báze (opět to bude ta, která tam vstoupila nejpozději) – v případě shody můžeme zvolit jakkoli, viz [523] a také [168, 476, 524]. Dá se přirozeně očekávat, že chytřejší výběr klíčového řádku a/nebo klíčového sloupce v případě více kandidátů, povede ke snížení celkového počtu potřebných kroků simplexového algoritmu. Takže jaká pravidla jsou nejlepší? *Žádná.* Kromě předchozích pravidel (a jiných) zaručujících konečnost celého algoritmu neexistují žádná jiná kritéria, která by prokazatelně a univerzálně zefektivnila celý výpočet, zejména až k tzv. polynomiálnímu počtu kroků, viz Podkapitolu 2.13. Navíc žádné z těchto pravidel univerzálně nezabrání zdržení simplexového algoritmu jako jsme viděli v prvním příkladě, viz např. [108] pro zdržení i při využití Blandova pravidla. Toto zdržení je sice konečné, ale může být velmi dlouhé a simplexovému algoritmu může trvat mnoho kroků, než se posune z degenerovaného BPB do jiného bodu. Na druhou stranu v [302, 403] bylo ukázáno, že v některých speciálních typech úlohy lineárního programování existuje jakési omezení pro délku zdržení (tj. počet po sobě jdoucích degenerovaných BPB) při využití lexikografického pravidla.

Ačkoli by se mohlo zdát, že degenerovanost úloh lineárního programování je poměrně vzácný jev[168], je většina úloh modelujících (skutečně) reálné situace degenerovaná. Avšak cykly se při jejich řešení simplexovým algoritmem objevují jen sporadicky – např. již v roce 1952 byl publikován problém jisté rafinérie, který byl degenerovaný, ale při jeho řešení simplexovým algoritmem žádný problém s konvergencí *nenastal*[169], viz [285]. Většina komerčních programů dokonce anticyklická pravidla přímo ani nevyužívá, protože buď je jejich implementace výpočetně složitá (jako v případě lexikografického pravidla) nebo jsou výpočetně neefektivní s ohledem na počet kroků (jako v případě Blandova pravidla). Navíc v námi uvažovaných pravidlech byly koeficienty pouze racionální čísla a naše výpočty byly vždy „přesné“, ale je zřejmé, že ruční výpočty jsou jiné než výpočty prováděné počítačem. Ty jsou obvykle prováděny pouze s jistým počtem desetinných míst, takže dochází k nepřesnostem v důsledku zaokrouhlova-

tečně skončí po konečně mnoha krocích. Viz také [163, str. 89–91] a [33, str. 178–184]

164



Abraham Charnes (4. září 1917 – 19. prosince 1992) byl americký matematik, který se věnoval především oblasti operačního výzkumu. Je autorem více než dvouset odborných článků a sedmi knih včetně *An Introduction to Linear Programming* z roku 1953 (společně s W. W. Cooperem a A. Hendersonem). Výrazně ovlivnil rozvoj metod analýzy obalu dat (*Data envelopment analysis*). V roce 1975 byl mezi nominovanými na Nobelovu cenu za ekonomii. V roce 1982 získal společně s William W. Cooperem and Richardem Duffinem *John von Neumann Theory Prize*, v roce 1989 obdržel *Harold Larnder Prize* udělované kanadskou společností pro operační výzkum a v roce 2006 společně s Cooperem získali *INFORMS Impact Prize*. Je také držitelem *Distinguished Public Service medal* za jeho práci pro americké námořnictvo (U.S. Navy) jakožto fyzika a operačního analytika během druhé světové války. Jako další střípek do skládanky s počátky a rozvojem (nejen) lineárního programování doporučujeme ke čtení [102].

165 Skutečně po přidání jednotlivých $\varepsilon,\dots,\varepsilon^m$ vznikne nedegenerovaná úloha? Označme vektor $\varepsilon_0 := (\varepsilon,\varepsilon^2,\dots,\varepsilon^m)^\top$. V jednotlivých krocích simplexové algoritmu dostaneme BPB s

$$x_B = A_B^{-1}(b+\varepsilon_0),$$

což jsou polynomy v ε, tj.

$$x_j = \beta_{j0}+\beta_{j1}\varepsilon+\cdots+\beta_{jm}\varepsilon^m,\ \ j\in B.$$

Tento polynom je pro každé $j\in B$ jistě nenulový, neboť alespoň pro jedno $k\in\{1,\dots,m\}$ platí $\beta_{jk}\neq 0$ vzhledem k regulárnosti matice A_B. Samozřejmě nelze zaručit, že tento polynom je kladný pro všechny $\varepsilon>0$, ale existuje číslo $\delta>0$ takové, že to v případě kladného prvního nenulového koeficientu platí pro všechna $\varepsilon\in(0,\delta)$. Vskutku, nechť $k\in\{0,1,\dots,m\}$ je nejmenší index, pro který $\beta_{jk}>0$.

cích chyb (obzvláště v případě iracionálních koeficientů) a jen výjimečně budeme mít přesně nulovou hodnotu ve sloupci x_B. To nás opět vrací k poznámce o eleminaci degenerovanosti u perturbační metody, která vede k využití jakési „praktické“ formy anticyklického pravidla v komerčním softwaru, které však ve skutečnosti nedokáže zcela zabránit vzniku cyklů, viz [215, 249]. Může totiž docházet k tzv. *výpočetním cyklům*, tj. k zacyklení v důsledku nastavených pravidel, zvolené přesnosti výpočtů a zaokrouhlovacích chyb, viz [204, 209]. Nicméně dá se očekávat, že pokud úloha projevuje klasické zacyklení, *nebude* docházet k výpočetním cyklům a naopak.

Cvičení

2.8.1. Řešte s pomocí simplexové metody úlohu

$$-3x_1+x_2\to\min$$
$$x_1+x_2\le 10\quad\&\quad 3x_1+x_2\le 15\quad\&\quad -x_1+x_2\le 13$$
$$x_1\ge 0\quad\&\quad x_2\ge 0.$$

Řešení: $[5,0]$.

2.8.2. Řešte s pomocí simplexové metody úlohu

$$x_1-x_2\to\min$$
$$x_1+x_2\le 6\quad\&\quad 3x_1+x_2\le 15\quad\&\quad -x_1+x_2\le 3$$
$$x_1\ge 0\quad\&\quad x_2\ge 0.$$

Řešení je nekonečně mnoho, leží na úsečce spojující body $[0,3]$ *a* $[3/2,9/2]$.

2.8.3. Uvažte úlohu

$$-x_1-2x_2\to\min$$
$$x_1-4x_2\le 4\quad\&\quad -2x_1+x_2\le 2$$
$$-3x_1+4x_2\le 12\quad\&\quad 2x_1+x_2\le 8$$
$$x_1\ge 0\quad\&\quad x_2\ge 0.$$

(a) Určete všechny extrémní body množiny X. Je některý z nich degenerovaný?

(b) Řešte úlohu s pomocí simplexové metody.

(c) Ignorujte čtvrté omezení. Určete nyní extrémní body a řešení úlohy.

Řešení: (b) $[20/11,48/11]$, *(c) nemá konečné řešení.*

2.8.4. Uvažte úlohu

$$2x_1+3x_2\to\max$$
$$x_1+2x_2\le 10\quad\&\quad -x_1+2x_2\le 6\quad\&\quad x_1+x_2\le 6$$
$$x_1\ge 0\quad\&\quad x_2\ge 0.$$

(a) Vyřešte úlohu graficky a ukažte, že řešením je degenerovaný BPB.

(b) Řešte úlohu s pomocí simplexové metody.

(c) Určete omezení, které způsobuje degenerovanost nalezeného řešení v první části. Odstraňte toto omezení a vyřešte tuto úlohu s pomocí simplexového algoritmu.

2.8.5. Řešte s pomocí simplexové metody úlohu

$$x_1+3x_2\to\max$$
$$-2x_1+x_2\le 4\quad\&\quad 5x_1+3x_2\le 15$$
$$x_1\ge 0\quad\&\quad x_2\ge 0.$$

Řešení: $[3/11,50/11]$.

Pak

$$\begin{aligned} x_j(\varepsilon) &= \varepsilon^k\big(\beta_{jk} + \beta_{j,k+1}\,\varepsilon + \\ &\quad + \cdots + \beta_{jm}\,\varepsilon^{m-k}\big) \geq \\ &\geq \varepsilon^k\big(\beta_{jk} - M\,\varepsilon - \ldots - M\,\varepsilon\big) \\ &= \varepsilon^k\big[\beta_{jk} - M\,(m-k)\,\varepsilon\big], \end{aligned}$$

protože $\beta_{j\ell}\,\varepsilon^p \geq -M\,\varepsilon^p \geq -M\,\varepsilon$ pro všechna $\ell = k+1,\ldots,m$, $p \geq 1$ a $\varepsilon \in (0,1)$, kde

$$M := \max\big\{0, \max_{j=k+1,\ldots,m}\{-a_j\}\big\}.$$

Položíme-li

$$\delta := \min\left\{1, \frac{a_k}{M\,(m-k)}\right\},$$

přičemž volíme $\delta = 1$ i pro $M = 0$ nebo $m = k$, pak odtud plyne $x_j(\varepsilon) > 0$ pro všechna $\varepsilon \in (0,\delta)$. Dá se navíc ukázat i opačná implikace, tj. platí $x_j(\varepsilon) > 0$ *pro všechna* $\varepsilon \in (0,\delta)$ *právě tehdy, když polynom* $x_j(\varepsilon)$ *je nenulový a první nenulový koeficient je kladný*. To znamená, že kdyby první nenulový koeficient byl záporný, pak $x_j(\varepsilon) < 0$ pro dostatečně malá ε, takže tato volba báze by nevedla k BPB. Pokud ale pro zvolenou bázi existuje BPB, pak výběrem toho nejmenšího δ pro všechny komponenty aktuálního x_B, zaručíme, že všechny složky x_B budou kladné, a tedy tento bod nebude degenerovaný. Jelikož současně existuje pouze konečný počet BPB, můžeme zvolit δ dokonce jako to nejmenší napříč všemi bázemi. Tím zaručíme, že žádný BPB nebude degenerovaný, a tudíž se simplexová metoda nezacyklí. Viz také [117, Section 10-2] nebo [468, Section 3.5.6].

[166] Stačí místo $\varepsilon, \varepsilon^2, \ldots, \varepsilon^m$ použít na pravé straně rovnice vektor $(\varepsilon_1, \ldots, \varepsilon_m)^\top$, kde $\varepsilon_1, \ldots, \varepsilon_m \in (0,1)$ jsou menší než všechna kladná čísla vyskytující se v úloze, přičemž platí $\varepsilon_1 \gg \varepsilon_2 \gg \cdots \gg \varepsilon_m$, tj. jsou postupně „mnohem větší“ (řádově větší) než jejich předchůdci, např. $\varepsilon_1 = 10^{-5}$, $\varepsilon_2 = 10^{-10}$, $\varepsilon_3 = 10^{-15}$. Proč? Protože kdybychom použili jenom jedno ε ve všech rovnicích, mohlo by se stát, že se odečtou a dostaneme se zpět do degenerovaného případu. To se ale při jakékoli smysluplné manipulaci s řádově odlišnými hodnotami nestane. Není ale nutné tyto hodnoty zcela konkretizovat. Mnohem výhodnější je v simplexové tabulce pracovat s abstraktními symboly $\varepsilon_1, \ldots, \varepsilon_m$, o nichž víme jen to, jak jsou velikostně uspořádány. Pak postup perturbační metody bude zcela kopírovat lexikografické pravidlo, aplikované na simplexovou tabulku rozšířenou o sloupce pro $\varepsilon_1, \ldots, \varepsilon_m$, které tvoří výchozí bázi

2.8.6. Řešte s pomocí simplexové metody úlohu

$$3x_1 + 2x_2 \to \max$$
$$x_1 - x_2 \leq 0 \quad \& \quad x_1 + x_2 \leq 10 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[5,5]$.

2.8.7. Řešte s pomocí simplexové metody úlohu

$$6x_1 + 5x_2 \to \max$$
$$x_1 + 2x_2 \leq 8 \quad \& \quad 3x_1 + x_2 \leq 15 \quad \& \quad x_1 + x_2 \leq 5$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[5,0]$.

2.8.8. Řešte s pomocí simplexové metody úlohu

$$2x_1 - x_2 \to \min$$
$$x_1 - 3x_2 \geq -3 \quad \& \quad 2x_1 + x_2 \geq -2$$
$$2x_1 + 3x_2 \leq 6 \quad \& \quad 3x_1 - 2x_2 \leq 6.$$

Řešení: $[-9/7, 4/7]$.

2.8.9. Řešte s pomocí simplexové metody úlohu (za výchozí BPB vezměte $[0,2]$)

$$-2x_1 + x_2 \to \min$$
$$2x_1 + 3x_2 = 6 \quad \& \quad x_1 - 2x_2 \leq 0 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[12/7, 6/7]$.

2.8.10. Řešte s pomocí simplexové metody úlohu

$$3x_1 + 2x_2 + x_3 \to \max$$
$$3x_1 - 3x_2 + 2x_3 \leq 3 \quad \& \quad -x_1 + 2x_2 + x_3 \leq 6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[8,7,0]$.

2.8.11. Řešte s pomocí simplexové metody úlohu

$$30x_1 + 23x_2 + 29x_3 \to \max$$
$$6x_1 + 5x_2 + 3x_3 \leq 26 \quad \& \quad 4x_1 + 2x_2 + 5x_3 \leq 7$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

2.8.12. Řešte s pomocí simplexové metody úlohu

$$-2x_1 - x_2 + x_3 \to \min$$
$$2x_1 + x_2 + 4x_3 \leq 6 \quad \& \quad x_1 + 4x_2 - x_3 \leq 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Nekonečně mnoho řešení tvaru $x_3 = 0$ *a* $x_2 = 6 - 2x_1$ *pro* $x_1 \in [20/3, 3]$ *a* $x_2 \in [0, 2/7]$.

2.8.13. Řešte s pomocí simplexové metody úlohu

$$x_1 - 2x_2 + x_3 \to \max$$
$$x_1 + 2x_2 + 3x_3 \leq 12 \quad \& \quad 2x_1 + x_2 - x_3 \leq 6 \quad \& \quad -x_1 + 3x_2 \leq 9$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[30/7, 0, 18/7]$.

2.8.14. Řešte s pomocí simplexové metody úlohu

$$8x_1 + 9x_2 + 4x_3 \to \max$$
$$x_1 + x_2 + 2x_3 \leq 2 \quad \& \quad 2x_1 + 3x_2 + 4x_3 \leq 3 \quad \& \quad 7x_1 + 6x_2 + 2x_3 \leq 8$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[2/3, 5/9, 0]$.

a ve výchozí simplexové tabulce vytvoří jednotkovou matici.

[167]

Václav Chvátal (⋆20. července 1946) je česko-kanadským matematikem a emeritním profesorem na univerzitě Concordia v kanadském Montrealu. Mezi hlavní obory jeho vědecké činnosti patří teorie grafů, kombinatorika, a kombinatorická optimalizace. Matematické vzdělání získal na Karlově univerzitě, kde studoval pod vedením Zdeňka Hedrlína. Avšak tři dny po invazi sovětských vojsk v roce 1968 emigroval do Kanady a své doktorské studium dokončil na podzim 1970 na University of Waterloo, kde se pod vlivem Jacka Edmondse začal zajímat o lineární programování. V roce 2015 společně s Jeanem Lasserrem získal *John von Neumann Theory Prize* za *jejich zásadní a hluboký přínos k teoretickým základům optimalizace*. Je také znám díky tzv. *Art gallery theorem*, v němž je dána horní hranice minimálního počtu strážců/kamer k „dokonalé" ochraně galerie ve tvaru jednoduchého n-úhelníku.

[168] Kolik náhodně vygenerovaných úloh lineárního programování daného rozměru bude asi degenerovaných? Obzvláště když si to spojíme s pozorováním učiněným ohledně perturbační metody.

[169] V roce 1949 požádal Koopmans Dantziga, aby něco udělal s požadavkem nedegenerovanosti v důkazu konvergence simplexové metody. V ideálním případě aby zkusil dokázat tvrzení bez této podmínky, kterou Dantzig původně považoval za zcela smysluplnou. Vždyť jaká je vlastně pravděpodobnost, že čtyři roviny v $\mathbb{R}^3$ se potkají v jednom bodě? Ale stalo se něco nečekaného. Ukázalo se totiž, že ačkoli pravděpodobnost degenerovanosti problému lineárního programování je nulová, každý praktický problém z oblasti letectva testovaný Dantzigovým oddělením byl degenerovaný. Degenerovanost se neměla stát, ale děla se. A nebyla to výjimka, ale pravidlo! Viz [118].

2.8.15. Řešte s pomocí simplexové metody úlohu

$$2x_1 + x_2 + 3x_3 \to \max$$
$$2x_1 - x_2 + 3x_3 \le 6 \quad \& \quad x_1 + 3x_2 + 5x_3 \le 10 \quad \& \quad 2x_1 + x_3 \le 7$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

2.8.16. Řešte s pomocí simplexové metody úlohu

$$2x_1 + x_2 + 6x_3 - 4x_4 \to \max$$
$$x_1 + 2x_2 + 4x_3 - x_4 \le 6 \quad \& \quad 2x_1 + 3x_2 - x_3 + x_4 \le 12 \quad \& \quad x_1 + x_3 + x_4 \le 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Řešení: $[0, 2/3, 0, 4/3]$.

2.8.17. Řešte s pomocí simplexové metody úlohu

$$-3x_1 + 2x_2 - x_3 + x_4 \to \max$$
$$2x_1 - 3x_2 - x_3 + x_4 \le 0 \quad \& \quad -x_1 + 2x_2 + 2x_3 - 3x_4 \le 1 \quad \& \quad -x_1 + x_2 - 4x_3 + x_4 \le 8$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Nemá konečné řešení.

2.8.18. Řešte s pomocí simplexové metody úlohu

$$x_1 + x_2 + x_3 + x_4 \to \max$$
$$x_1 + 2x_2 - x_3 + 3x_4 \le 12 \quad \& \quad x_1 + 3x_2 + x_3 + 2x_4 \le 8 \quad \& \quad 2x_1 - 3x_2 - x_3 + 2x_4 \le 7$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.8.19. Řešte s pomocí simplexové metody úlohu

$$x_1 + 2x_2 + 3x_3 + x_4 \to \max$$
$$2x_1 + x_2 + x_3 + 2x_4 \le 18 \quad \& \quad 3x_1 + 5x_2 + 2x_3 + 3x_4 \le 24 \quad \& \quad 3x_1 + 2x_2 + x_3 + x_4 \le 12$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.8.20. Řešte s pomocí simplexové metody úlohu

$$2x_1 + 3x_2 + x_3 + x_4 + 2x_5 \to \max$$
$$2x_1 + x_2 - 3x_3 + x_4 + x_5 \le 10 \quad \& \quad x_1 + 4x_2 + x_4 + 2x_5 \le 20 \quad \& \quad 3x_1 + 4x_4 + 2x_5 \le 15$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0.$$

2.8.21. Řešte s pomocí simplexové metody úlohu

$$3x_1 + x_2 + 5x_3 + 4x_4 \to \max$$
$$3x_1 - 3x_2 + 2x_3 + 8x_4 \le 50 \quad \& \quad 4x_1 + 6x_2 - 4x_3 - 4x_4 \le 40 \quad \& \quad 4x_1 - 2x_2 + x_3 + 3x_4 \le 20$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

2.8.22. Řešte s pomocí simplexové metody úlohu

$$x_1 - 2x_2 + x_3 + 3x_4 \to \min$$
$$2x_1 - x_2 + x_3 - x_4 \le 10 \quad \& \quad -5x_1 + 2x_2 - 2x_3 + x_4 \le 20 \quad \& \quad 3x_1 - 4x_2 + 4x_3 - 2x_4 \le 30$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Nemá konečné řešení.

2.8.23. Řešte s pomocí simplexové metody úlohu

$$2x_1 + 2x_2 + 4x_3 + 5x_5 + 3x_6 \to \min$$
$$3x_1 + 6x_2 + 3x_3 + 3x_4 + 3x_5 + 4x_6 \le 60$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0 \quad \& \quad x_6 \ge 0.$$

Řešení (nejen): $[0, 0, 0, 0, 0, 0]$.

2.8.24. Řešte s pomocí simplexové metody úlohu

$$3x_1 + 2x_2 + 4x_3 \to \max$$
$$x_1 + x_2 + 2x_3 \le 4 \quad \& \quad 2x_1 + x_3 \le 5 \quad \& \quad 2x_1 + x_2 + 3x_3 \le 7$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[5/2, 3/2, 0]$.

2.8.25. Řešte s pomocí simplexové metody úlohu

$$5x_1 + 6x_2 + 9x_3 + 8x_4 \to \max$$
$$x_1 + 2x_2 + 3x_3 + x_4 \le 5 \quad \& \quad x_1 + x_2 + 2x_3 + 3x_4 \le 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Řešení: $[1, 2, 0, 0]$.

2.8.26. Řešte s pomocí simplexové metody úlohu

$$2x_1 - x_2 + 4x_3 - 6x_4 \to \max$$
$$x_2 + 3x_3 - x_4 \ge 0 \quad \& \quad 3x_1 - x_2 + 2x_4 \le 16 \quad \& \quad x_1 + 2x_2 - x_3 - 2x_4 \ge -4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Nemá konečné řešení.

2.8.27. Řešte s pomocí simplexové metody doplněné o Blandovo a lexikografické pravidlo úlohu

$$x_1 + 2x_2 + x_3 \to \max$$
$$x_1 + 4x_2 + 3x_3 \le 4 \quad \& \quad -x_1 + x_2 + 4x_3 \le 1 \quad \& \quad x_1 + 3x_2 + x_3 \le 6$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[4, 0, 0]$.

2.8.28. *Kuhnův příklad zacyklení (nedatováno, zveřejněno v článku z roku 1969)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$-2x_4 - 3x_5 + x_6 + 12x_7 \to \min$$
$$x_1 - 2x_4 - 9x_5 + x_6 + 9x_7 = 0$$
$$x_2 + x_4/3 + x_5 - x_6/3 - 2x_7 = 0$$
$$x_3 + 2x_4 + 3x_5 - x_6 - 12x_7 = 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0$$
$$x_5 \ge 0 \quad \& \quad x_6 \ge 0 \quad \& \quad x_7 \ge 0.$$

Ověřte, že v takovém případě vede simplexový algoritmus na cyklus délky 6. Řešením této úlohy je $x_1 = x_4 = x_6 = 2$ a $x_2 = x_3 = x_5 = x_7 = 0$. Tato úloha je považována za jeden z nejjednodušších příkladů zacyklení. V původním Kuhnově příkladu není třetí omezení a úloha je neohraničená – ověřte, že i v takovém případě nastává cyklus délky 6. Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

2.8.29. *Marshallův–Suurbaleův příklad zacyklení (1969)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$2x_1 + 4x_4 + 4x_6 \to \min$$
$$x_1 - 3x_2 - x_3 - x_4 - x_5 + 6x_6 = 0 \quad \& \quad 2x_2 + x_3 - 3x_4 - x_5 + 2x_6 = 0$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_5 \ge 0 \quad \& \quad x_6 \ge 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Řešením je $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0$ *a nastává cyklus délky* 6.

2.8.30. *Solowův příklad zacyklení (1984)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$-2x_3 - 2x_4 + 8x_5 + 2x_6 \to \min,$$
$$x_1 - 7x_3 - 3x_4 + 7x_5 + 2x_6 = 0 \quad \& \quad x_2 + 2x_3 + x_4 - 3x_5 - x_6 = 0$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0 \quad \& \quad x_6 \geq 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Řešením je $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0$ *a nastává cyklus délky* 6.

2.8.31. *Sierksmův příklad zacyklení (1996)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$3x_1 - 80x_2 + 2x_3 - 24x_4 \to \max,$$
$$x_1 - 32x_2 + 4x_3 + 36x_4 + x_5 = 0 \quad \& \quad x_1 - 24x_2 - x_3 + 6x_4 + x_6 = 0$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0 \quad \& \quad x_6 \geq 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Úloha je neohraničená a nastává cyklus délky 6.

2.8.32. *Yudinův–Gol'shteĭnův příklad zacyklení I. (1965)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$x_3 - x_4 + x_5 - x_6 \to \max$$
$$x_1 + 2x_3 - 3x_4 - 5x_5 + 6x_6 = 0$$
$$x_2 + 6x_3 - 5x_4 - 3x_5 + 2x_6 = 0$$
$$3x_3 + x_4 + 2x_5 + 4x_6 + x_7 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0 \quad \& \quad x_6 \geq 0 \quad \& \quad x_7 \geq 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Řešením je $x_1 = 5/2$, $x_2 = 3/2$, $x_5 = 1/2$, $x_3 = x_4 = x_6 = x_7 = 0$ *a nastává cyklus délky* 6.

2.8.33. *Yudinův–Gol'shteĭnův příklad zacyklení II. (1965)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$x_3 - x_4 + x_5 - x_6 \to \max$$
$$x_1 + x_3 - 2x_4 - 3x_5 + 4x_6 = 0$$
$$x_2 + 4x_3 - 3x_4 - 2x_5 + x_6 = 0$$
$$x_3 + x_4 + x_5 + x_6 + x_7 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0 \quad \& \quad x_6 \geq 0 \quad \& \quad x_7 \geq 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Řešením je $x_1 = 3$, $x_2 = 2$, $x_5 = 1$, $x_3 = x_4 = x_6 = x_7 = 0$ *a nastává cyklus délky* 6.

2.8.34. *Neringův–Tuckerův příklad zacyklení (1993)*: pomocí klasického simplexového algoritmu (s Dantzigovým pravidlem doplněného o pravidlo první volby, tj. sloupec s největší hodnotou $z_k - c_k > 0$ nejvíce vlevo a řádek s nejmenším nezáporným podílem x_B/λ nejvýše v tabulce) řešte úlohu

$$-3x_2 + x_3 - 6x_4 - 4x_6 \to \max$$
$$x_1 + x_2 + x_5/3 + x_6/3 = 2$$
$$9x_2 + x_3 - 9x_4 - 2x_5 - x_6/3 + x_7 = 0$$
$$x_2 + x_3/3 - 2x_4 - x_5/3 - x_6/3 + x_8 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0 \quad \& \quad x_6 \geq 0 \quad \& \quad x_7 \geq 0 \quad \& \quad x_8 \geq 0.$$

Vyřešte úlohu také s pomocí některého z anti-cyklických pravidel.

Úloha je neohraničená a nastává cyklus délky 6.

2.8.35. Uvažte úlohu (LP) s omezenou přípustnou množinou, přičemž dokonce známe hodnotu čísla $u \in \mathbb{R}$ takového, že všechny složky libovolného $x \in X$ splňují $x_i \le u$ pro každé $i \in \{1,\dots,n\}$. Ukažte, že tento problém je ekvivalentní s úlohou lineárního programování zahrnující omezení $\sum_{i=1}^{n} x_i = 1$.

2.9 Aby začátky nebyly tak těžké

Simplexovou metodu a její realizaci jsme si popsali a podrobně ilustrovali v předchozích odstavcích. Všechny uvedené příklady však měly jednu společnou vlastnost: matice A z kanonického tvaru vždy obsahovala jednotkovou matici, díky čemuž jsme snadno určili výchozí BPB. V ostatních případech to ale tak snadné být nemusí. Máme sice obecný návod pro výpočet x_B a příslušných koeficientů λ_{ij} pro libovolnou bázi B, jenže takový bod nemusí být přípustný. Při větším počtu proměnných a omezení to může vést k velkému počtu výpočtů, než se nám podaří najít nějaký BPB. Proto potřebujeme nějaký efektivní nástroj pro určení výchozího BPB pro simplexový algoritmus. Nejčastěji se můžeme setkat s dvěma metodami[170]:

(i) *metoda velkého M*[171] (též *penalizační metoda*), ve které se do účelové funkce přidají umělé proměnné s maximálně nevýhodnou cenou M[172].

(ii) *metoda umělé báze* (též *dvoufázová metoda*[173]) založené na řešení pomocné úlohy lineárního programování.

Obě metody mají mnoho společného, zejména:

(i) nejsou potřeba žádné předpoklady na původní úlohu (může být nepřípustná[174], degenerovaná či neohraničená[175]);

(ii) výchozí BPB pro pomocnou úlohu nebo první fázi nalezneme snadno.

Metoda velkého M. Tato metoda je založená na nahrazení původního problému (LP) úlohou

$$\left.\begin{aligned} \langle c, x\rangle + \langle \mu, w\rangle &\to \min \\ Ax + Iw = b \quad \& \quad x \ge 0 \quad \& \quad w &\ge 0, \end{aligned}\right\} \tag{2.9.1}$$

kde I je jednotková $m \times m$ matice, $w \in \mathbb{R}^m$ a $\mu := (M, M, \dots, M)^\top$ je konstantní vektor s $M > 0$ větším než jakékoli jiné číslo vyskytující se v úloze či v objevující se v průběhu výpočtu. Vzhledem k tomu, že dopředu nevíme, jaká čísla se v průběhu výpočtu objeví, počítáme s hodnotou M jako s parametrem. Rozšíření úlohy o jednotkovou matici nám umožní snadno najít výchozí BPB[176], kterým je $x = 0$ a $w = b$. Cílem této metody pak je vynulovat všechny umělé proměnné (tj. najít bod $[\widetilde{x}, \widetilde{w}]$ splňující $\widetilde{w} = 0$), v kterémžto případě bude vlastně úloha (2.9.1) totožná s (LP). Proto $\widetilde{x}$ bude výchozím BPB původní úlohy[177] (LP) a ve výpočtu můžeme plynule pokračovat dalšími kroky simplexového algoritmu, čímž nalezneme řešení úlohy (LP). Je-li původní úloha přípustná a ohraničená, pak přičítání umělé proměnné s velkým koeficientem M zajistí, že umělá proměnná jistě bázi opustí. Jelikož všechny pomocné proměnné jsou součástí výchozího BPB, je jinými slovy potřeba v průběhu výpočtu postupně všechny pomocné proměnné vyloučit z báze[178]. To mimo jiné znamená, že pokud se nám již podaří nějakou umělou proměnnou z báze vyloučit, pak pro nalezení řešení původní úlohy ji v žádném dalším kroku nemusíme vracet zpět (i když by to mohlo vést k porušení Dantzigova pravidla v simplexovém algoritmu). Proto po vyloučení umělé proměnné z báze již ani nemusíme počítat jí příslušný sloupec v simplexové tabulce. Pokud řešení rozšířené úlohy (2.9.1) závisí na parametru M (tj. umělá proměnná zůstane v bázi s nenulovou hodnotou), původní úloha je nepřípustná. Kdyby simplexový algoritmus ukazoval neohraničenost rozšířené úlohy (2.9.1), pak původní úloha je buď nepřípustná nebo neohraničená[179].

Příklad 2.9.1

S využitím metody velkého M vyřešme úlohu

$$x_1 + 4x_2 + 2x_3 \to \min$$

[170] Za chvíli uvidíme, že obě tyto metody jsou založené na velmi podobném principu. Hlavní rozdíl tkví především v tom, že výpočet metodou velkého M je realizován s využitím jedné simplexové tabulky, ve které ale vystupuje parametr M, zatímco v metodě umělé báze žádný takový parametr nevystupuje, avšak výpočet je rozdělen do dvou kroků, tj. simplexový algoritmus je „spouštěn" na dvakrát. V [125, str. 98] můžeme najít Dantzigův skromný komentář o tom, že otcem obou metod je v podstatě on, a odkazuje se přitom na článek [113]. Vskutku, v [113, Section 2] Dantzig řeší otázku nalezení výchozího BPB, avšak náš současný popis obou metod je již trochu propracovanější. První metoda bývá obvykle spojována s již dříve zmíněným A. Charnesem, jak se můžeme dočíst v [246, str. 120], viz také zajímavý „spor" v [205, 284, 298]. Druhou metodu pak skutečně můžeme přiřknout Dantzigovi ve spolupráci s Ordenem a dalšími z RANDu, viz [246, str. 149] a [368] pro její zobecnění.

[171] Angl. *big-M method*. V [292, str. 128] můžeme najít pro matematiky vtipnou(?) poznámku, že neexistuje žádný důvod se domnívat, že „big M" je zkratkou pro „big mother".

[172] Parametr M tedy nemá konkrétní hodnotu. Víme jen, že *je větší než cokoli jiného*, a tudíž je v této souvislosti zcela přirozené hovořit o „velkém M"

[173] Angl. *two-phase method*, která využívá *umělé proměnné* neboli angl. *artificial variables*. Tyto umělé proměnné využíváme i v první metodě. V úloze pak tedy budeme mít trojici proměnných: původní (strukturní /angl. *original, structural*/), pomocné (volné, přídavné /angl. *slack, surplus*/) a umělé, přičemž podstata druhé a třetí skupiny proměnných je diametrálně odlišná: zatímco pomocné proměnné využíváme pro převod do kanonického tvaru a pracujeme s nimi po celou dobu řešení, tak umělé proměnné jsou potřeba pouze pro nalezení výchozího BPB.

[174] V takovém případě samozřejmě žádný výchozí BPB nenajdeme.

[175] Dokonce by úloha mohla mít i duplicitní podmínky. V takovém případě by ale bylo $\operatorname{rank} A < m$, což by navzdory nalezení výchozího BPB znemožnilo využití simplexového algoritmu při hledání samotného řešení.

$$x_1 - x_2 + x_3 = 0 \quad \& \quad -x_2 + 4x_3 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení. Na první pohled si můžeme všimnout, že proměnná x_1 je obsažena pouze v prvním omezení, navíc s koeficientem 1. Tento sloupec můžeme využít při tvorbě jednotkové matice v rozšířené úloze. Ke druhé rovnosti pak přidáme umělou proměnnou (ozn. w), čímž získáme i druhý sloupec jednotkové matice. Rozšířená úloha je tedy ve tvaru

$$x_1 + 4x_2 + 2x_3 + Mw \to \min$$
$$x_1 - x_2 + x_3 = 0 \quad \& \quad -x_2 + 4x_3 + w = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad w \geq 0.$$

Postup řešení simplexovou metodou je uveden v Tabulce 2.21.

c_B	B	x_1	x_2	x_3	w	x_B	$\frac{x_B}{\lambda}$
1	x_1	1	-1	(1)	0	0	$0 \to$
M	w	0	-1	4	1	1	1/4
	c	1	4	2	M		
	z_k	1	$-1-M$	$1+4M$	M		
	$z_k - c_k$	0	$-5-M$	$-1+4M \uparrow$	0	M	
	x_3	1	-1	1	0	0	✖
	w	-4	(3)	0	1	1	$1/3 \to$
	$z_k - c_k$	$1-4M$	$-6+3M \uparrow$	0	0	M	
	x_3	$-1/3$	0	1	✖	1/3	
	x_2	$-4/3$	1	0	✖	1/3	
	$z_k - c_k$	-7	0	0	✖	2	

Tabulka 2.21: Simplexová metoda pro Příklad 2.9.1.

Po sestavení tabulky je v řádku $z_k - c_k$ pouze jedno kladné číslo, a to ve třetím sloupci. Stejně jako v předchozích příkladech pak vybereme klíčový řádek a provedeme eliminaci dle simplexového algoritmu. Nová báze je tvořena proměnnými x_3 a w. K eliminaci čísla $-1+4M$ přičteme k řádku $z_k - c_k$ klíčový řádek vynásobený číslem $1-4M$. V dalším kroku se tak dostaneme k tabulce, ve které je opět v řádku $z_k - c_k$ jedno kladné číslo. Do báze tedy vstoupí proměnná x_2 a vystoupí umělá proměnná w. Eliminací s využitím pivota (číslo 3) pak získáme poslední tabulku.

Sloupec pro w v poslední tabulce je nevyplněný, neboť tyto hodnoty již dále rozhodně nebudou potřeba (bez ohledu na to, zda jsme již na konci simplexového algoritmu či nikoli). V řádku pro $z_k - c_k$ jsou všechny hodnoty pro nebazické proměnné záporné, a tudíž simplexový algoritmus končí. V bázi nemáme žádnou umělou proměnnou, nalezený bod je tak výchozím BPB původní úlohy. S touto simplexovou tabulkou můžeme nyní pokračovat v hledání samotného řešení. Z posledního kroku algoritmu ale vidíme, že již nelze vybrat klíčový sloupec. Proto nalezený výchozí BPB je zároveň i řešením původní úlohy.

[176] Není nutné vždy zavádět všech m umělých proměnných. Pokud se proměnná x_j objevuje pouze v jedné rovnici (tj. v matici A máme kladný násobek nějakého sloupce jednotkové $m \times m$ matice, vynásobíme tuto rovnost kladným číslem tak, aby koeficient u x_j byl $+1$, a tuto proměnnou x_j zavedeme do báze místo w_j, které vůbec nevyužijeme. Další postup je víceméně stejný, pouze nikde nevystupuje w_j a již neplatí obecný zápis omezení (2.9.1) uvedený výše.

[177] Pro jistotu zdůrazněme, že přidáním umělých proměnných jsme vlastně rozšířili původní přípustnou množinu. Proto x-ová složka bodů $[x, w]$ generovaných simplexovým algoritmem v průběhu řešení úlohy (2.9.1) není v případě $w \neq 0$ přípustným bodem původním úlohy, tj. neplatí $Ax = b$. To se stane právě až v případě, kdy $w = 0$.

[178] Případně je mít v bázi s nulovou hodnotou, což indikuje degenerovanost rozšířené úlohy.

[179] Již víme, že v takovém případě simplexový algoritmus najde směr, ve kterém účelová funkce klesá do $-\infty$, viz Příklad 2.8.6. Výsledek pak záleží na tom, zda tento směr zahrnuje i umělé proměnné či nikoli.

Metoda umělé báze. V literatuře se autorovi kdysi podařilo najít poznámku[180] o tom, že *ačkoli tato metoda velkého M není početně výhodnější než dvoufázová metoda, je implementována ve většině výpočetních systémů.* Jenže takovýto komentář vzbuzuje spíše pochybnosti. Samozřejmě matematici bývají mnohdy mentálními masochisty, takže rádi hledají nástroje, které způsobí jejich mysli ještě větší bolest, avšak lze očekávat, že počítače touto poruchou netrpí. Ano, výpočet pomocí jedné simplexové tabulky s několika kroky činí z metody velkého M vskutku elegantní nástroj pro nalezení výchozího BPB, takže v dohledné budoucnosti asi jen tak nezmizí z našeho povědomí a[181] učebnic lineárního programování[182]. Jenže „vágnost" jejího popisu komplikuje následnou implementaci, neboť počítače s parametrem M majícím *hodnotu větší než jakékoli jiné číslo* pracují jen velmi obtížně. Bude proto asi nutné přiřadit M nějakou konkrétní číselnou hodnotu, např. $M = 1000$. Jenže může se stát, že tento apriorní odhad nebude dostatečně velký. Takže by to chtělo skutečně velké číslo, což ale na druhou stranu může při výpočtu působit numerické potíže... Proto by to chtělo metodu, která nepracuje s abstraktním M – a tou je právě druhá zmíněná metoda, na kterou se nyní zaměříme. V její *první fázi* aplikujeme simplexový algoritmus na problém (LP), v němž ale pracujeme s pomocnou účelovou funkcí. Tím určíme, zda je původní problém přípustný, v kterémžto případě současně nalezneme nějaký BPB původní úlohy. Tento bod pak vezmeme jako výchozí BPB pro simplexový algoritmus, s jehož pomocí ve *druhé fázi* vyřešíme původní problém. velkou výhodou této metody je fakt, že pro spuštění druhé fáze nemusíme dopočítávat koeficienty λ_{ij}, ale můžeme víceméně přímo pokračovat s poslední simplexovou tabulkou z konce první fáze. Abychom byli přesnější:

[180] Autor si již bohužel nepamatuje, kde to přesně bylo. Asi někde vlevo dole...

[181] dobrých (?)

[182] Proto „velké M" asi jen tak nebude zase pouze písmenem abecedy...

(i) Po *inicializaci* řešení, kdy danou úlohu lineárního programování převedeme do kanonického tvaru (LP), zahájíme první fázi. V ní opět rozšíříme[183] soustavu $Ax = b$ o *umělé proměnné* $w_1, \dots, w_m \geq 0$ tak, že platí

$$Ax + Iw = b. \tag{2.9.2}$$

Poté aplikujeme simplexový algoritmus na úlohu

$$\left.\begin{array}{c} w_1 + \cdots + w_m \to \min \\ Ax + Iw = b \\ x_1 \geq 0 \quad \& \quad \dots \quad \& \quad x_n \geq 0 \\ w_1 \geq 0 \quad \& \quad \dots \quad \& \quad w_m \geq 0, \end{array}\right\} \tag{2.9.3}$$

přičemž za výchozí BPB vezmeme opět $x = 0$ a $w = b$. Jelikož $w_1, \dots, w_m \geq 0$, tato úloha nemůže být neohraničená a v závislosti na nalezeném řešení $[\widetilde{x}, \widetilde{w}]$ úlohy (2.9.3) mohou nastat pouze tři možnosti[184]:

[183] Zde platí stejný komentář jako v případě metody velkého M.

[184] Tato metoda indikuje buď nepřípustnost původní úlohy nebo s její pomocí nalezneme výchozí BPB.

(a) jestliže $\sum_{i=1}^{m} \widetilde{w}_i > 0$, pak původní úloha je nepřípustná, tj. pro původní problém v kanonickém tvaru nelze simplexový algoritmus, neboť pro něj nenajdeme výchozí BPB. Vskutku, kdyby totiž nějaké $\widehat{x} \geq 0$ bylo přípustné, pak by dvojice $\widehat{x}$ a $w = 0$ byla přípustným bodem pro úlohu (2.9.3) s nulovou hodnotou účelové funkce, což je spor s nalezeným optimem určeným w^* ↯.

(b) jestliže $\sum_{i=1}^{m} \widetilde{w}_i = 0$ a všechny umělé proměnné jsou mimo bázi, pak nutně $\widetilde{w} = 0$ a bod $\widetilde{x}$ je BPB původní úlohy (LP).

(c) jestliže $\sum_{i=1}^{m} \widetilde{w}_i = 0$ a alespoň jedna umělá proměnná je v bázi (nutně s nulovou hodnotou), pak výchozí BPB pro původní úlohu (LP) dostaneme s těmito proměnnými jako součástí výchozí báze. Toto se stane, má-li původní problém (LP) nadbytečné rovnice a často také má-li degenerované řešení. Pochopitelně je ale možné chtít pokračovat s první fází tak dlouho, dokud se nám nepodaří vyloučit všechny pomocné proměnné z báze (čímž dostaneme předchozí případ), abychom minimalizovali počet proměnných v další fázi. Toto se nám podaří pouze v případě, kdy řádek simplexové tabulky odpovídající umělé proměnné v bázi bude obsahovat alespoň jedno kladné číslo v některém ze sloupců pro původní

proměnné úlohy v kanonickém tvaru. Pak totiž můžeme udělat ještě jeden krok simplexového algoritmu, ve kterém tuto umělou proměnnou nahradíme v bázi nějakou původní proměnnou[185]. Jenže v případě nadbytečných omezení toto nemusí být vůbec možné. Např. kdyby řádek simplexové tabulky odpovídající takové umělé proměnné obsahoval pouze nuly ve sloupcích pro původní proměnné z kanonického tvaru. Ovšem v takovém případě můžeme tento řádek (a odpovídající umělou proměnnou) dále ignorovat, neboť se v další fázi rozhodně nezmění, tj. podaří se nám odstranit jedno omezení, které je pro daný problém zcela nadbytečné.

[185] Umělá proměnná má v tuto chvíli nulovou hodnotu, takže ve sloupci „x_B/λ“ bude nula, a my ji tak v souladu se simplexovým algoritmem a podmínkou (2.8.1) můžeme vyloučit z báze

(ii) Pokud nastal případ (b) nebo (c), následuje *druhá fáze*. Na jejím začátku vymažeme z poslední simplexové tabulky v první fázi všechny sloupce odpovídající nebazickým umělým proměnným, hodnotu pomocné účelové funkce nahradíme hodnotou účelové funkce původní úlohy (LP) v nalezeném výchozím BPB a přepočítáme řádek pro $z_k - c_k$. Tímto dostaneme výchozí simplexovou tabulku pro původní úlohu (LP) a dále pokračujeme simplexovovým algoritmem, přičemž pokud nám zůstaly v bázi ještě nějaké umělé proměnné (s nulovou hodnotou), pak v dalších krocích musíme hlídat, aby jejich hodnota nebyla kladná, což by rozbilo přípustnost aktuálního BPB.

Dvoufázovou metodu si ukážeme v několika následujících příkladech včetně situace s umělou proměnnou ponechanou v bázi pro druhou fázi (viz Příklad 2.9.5).

Příklad 2.9.2

Dvoufázová metoda #1:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp7

⚠ V první fázi jsou počítány i koeficienty ve sloupcích pro umělé proměnné vyloučené z báze – už ale víme, že to není nutné.

Příklad 2.9.3

Dvoufázová metoda #2:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp8

⚠ Vektor c uvedený v čase 0:56 má být $(1,-2,0,0,0)^\top$, ovšem nemá to žádný vliv na další výpočty.
⚠ V první fázi jsou počítány i koeficienty ve sloupcích pro umělé proměnné vyloučené z báze – už ale víme, že to není nutné.

Příklad 2.9.4

Dvoufázová metoda #3:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp9

⚠ V první fázi jsou počítány i koeficienty ve sloupcích pro umělé proměnné vyloučené z báze – už ale víme, že to není nutné.

Příklad 2.9.5

Dvoufázová metoda #4:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/lp10

⚠ V první fázi jsou počítány i koeficienty ve sloupcích pro umělé proměnné vyloučené z báze – už ale víme, že to není nutné.
⚠ V pravém dolním rohu poslední tabulky ze druhé fáze v čase 20:14 být číslo 7/4 místo 3/4, což nemá vliv na správnost výsledku.

Cvičení

2.9.1. Pomocí simplexové metody řešte úlohu

$$-3x_1 - 5x_2 \to \min$$
$$x_1 - 2x_2 \le 6 \quad \& \quad 2x_1 + x_2/2 \le 7 \quad \& \quad x_1 + x_2 \ge 16 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.2. Pomocí simplexové metody řešte úlohu

$$x_1 + x_2 \to \min$$
$$x_1 - x_2 \ge 1 \quad \& \quad -x_1 + x_2 \ge 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.3. Pomocí simplexové metody řešte úlohu

$$2x_1 - 6x_2 \to \max$$
$$3x_1 + 2x_2 \le 6 \quad \& \quad x_1 - x_2 \ge -1 \quad \& \quad -x_1 - 2x_2 \ge 1.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.4. Pomocí simplexové metody řešte úlohu

$$2x_1 + 3x_2 + 4x_3 \to \max$$
$$x_1 + x_2 + x_3 \le 1 \quad \& \quad x_1 + x_2 + 2x_3 = 2 \quad \& \quad 3x_1 + 2x_2 + x_3 \ge 4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.5. Pomocí simplexové metody řešte úlohu

$$3x_1 - 2x_2 - 4x_3 \to \min$$
$$3x_1 - 6x_2 + x_3 \ge 5 \quad \& \quad x_1 - x_2 - 2x_3 \ge 4 \quad \& \quad x_1 + x_2 - x_3 \le 0$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.6. Pomocí simplexové metody řešte úlohu

$$2x_1 - x_2 + x_3 + 5x_4 \to \min$$
$$x_1 + x_2 + x_3 + x_4 = 4 \quad \& \quad 2x_1 + 3x_2 - 4x_3 + 2x_4 \le 5 \quad \& \quad x_1 + 2x_2 - 5x_3 + x_4 \ge 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0.$$

Nemá přípustné řešení, tj. $X = \emptyset$.

2.9.7. Pomocí simplexové metody řešte úlohu

$$x_1 + x_2 \to \max$$
$$x_1 - x_2 \ge 1 \quad \& \quad -x_1 + 2x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Nemá konečné řešení.

2.9.8. Pomocí simplexové metody řešte úlohu

$$x_1 + 3x_2 - x_3 \to \min$$
$$2x_1 + x_2 + 3x_3 \ge 3 \quad \& \quad -x_1 + x_2 \ge 1 \quad \& \quad -x_1 - 5x_2 + x_3 \le 4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Nemá konečné řešení.

2.9.9. Pomocí simplexové metody řešte úlohu

$$-2x_1 + x_2 - x_3 \to \min$$
$$3x_1 + x_2 - 2x_3 \le 8 \quad \& \quad 4x_1 - x_2 + 2x_3 \ge 2 \quad \& \quad 2x_1 + 3x_2 - 2x_3 \ge 4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Nemá konečné řešení.

2.9.10. Pomocí simplexové metody řešte úlohu

$$-4x_1 - 5x_2 + 3x_3 \to \min$$
$$x_1 + 2x_2 + x_3 = 10 \quad \& \quad x_1 - x_2 \geq 6 \quad \& \quad x_1 + 3x_2 + x_3 \leq 14$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[10, 0, 0]$.

2.9.11. Pomocí simplexové metody řešte úlohu

$$-5x_1 + 2x_2 - x_3 \to \min$$
$$2x_1 + 4x_2 + x_3 \leq 6 \quad \& \quad 2x_1 + 2x_2 + 3x_3 \geq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[4, 0, -2]$.

2.9.12. Pomocí simplexové metody řešte úlohu

$$3x_1 + 2x_2 + 3x_3 \to \max$$
$$2x_1 + x_2 + x_3 \leq 2 \quad \& \quad 3x_1 + 4x_2 + 2x_3 \geq 8$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[0, 2, 0]$.

2.9.13. Pomocí simplexové metody řešte úlohu

$$-2x_1 + 3x_2 \to \min$$
$$2x_1 + x_2 \leq 10 \quad \& \quad -2x_1 + 3x_2 \leq 9 \quad \& \quad 2x_1 + 4x_2 \geq 8$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[5, 0]$.

2.9.14. Pomocí simplexové metody řešte úlohu

$$x_1 + x_2 + x_3 \to \min$$
$$x_1 + 2x_2 + 3x_3 = 3 \quad \& \quad -x_1 + 2x_2 + 6x_3 = 2$$
$$4x_2 + 9x_3 = 5 \quad \& \quad 3x_3 + x_4 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

Řešení: $[1/2, 5/4, 0, 1]$.

2.9.15. Pomocí simplexové metody řešte úlohu

$$8x_1 + 4x_2 - 2x_3 \to \min$$
$$2x_1 + 2x_2 - x_3 \geq 1 \quad \& \quad x_1 + 3x_2 + 2x_3 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[0, 4/7, 1/7]$.

2.9.16. Pomocí simplexové metody řešte úlohu

$$2x_1 + x_2 \to \max$$
$$x_1 - 5x_2 - x_3 \leq -2 \quad \& \quad 2x_1 + 3x_2 + 2x_3 \leq 7$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[29/13, 11/13, 0]$.

2.9.17. Pomocí simplexové metody řešte úlohu

$$-x_1 - 2x_2 - x_3 \to \min$$
$$3x_1 + x_2 - x_3 = 15 \quad \& \quad 8x_1 + 4x_2 - x_3 = 50 \quad \& \quad 2x_1 + 2x_2 + x_3 = 20$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[5/2, 15/2, 0]$.

2.9.18. Pomocí simplexové metody řešte úlohu

$$6x_1 + 3x_2 \to \min$$
$$x_1 + x_2 \geq 1 \quad \& \quad 2x_1 - x_2 \geq 1 \quad \& \quad 3x_2 \leq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[2/3, 1/3]$.

2.9.19. Pomocí simplexové metody řešte úlohu

$$5x_1 + 4x_2 \to \min$$
$$x_1 + 2x_2 \leq 6 \quad \& \quad 2x_1 + x_2 \leq 5 \quad \& \quad 4x_1 + x_2 \geq 2 \quad \& \quad x_1 + x_2 \geq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[1/3, 2/3]$.

2.9.20. Pomocí simplexové metody řešte úlohu

$$3x_1 + 4x_2 \to \max$$
$$3x_1 + x_2 \leq 9 \quad \& \quad -2x_1 + x_2 \leq 4 \quad \& \quad x_1 + 2x_2 \geq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[1, 6]$.

2.9.21. Pomocí simplexové metody řešte úlohu

$$3x_1 + 2x_2 + 5x_3 \to \max$$
$$x_1 + 3x_2 + 2x_3 \leq 15 \quad \& \quad 2x_2 - x_3 \geq 5 \quad \& \quad 2x_1 + x_2 - 5x_3 = 10$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[120/23, 65/23, 15/23]$.

2.9.22. Pomocí simplexové metody řešte úlohu

$$2x_1 + 3x_2 + 8x_3 \to \min$$
$$x_1 + 2x_2 + 7x_3 \geq 8 \quad \& \quad x_1 + 3x_2 + 6x_3 \geq 9 \quad \& \quad 2x_1 + 4x_2 + 2x_3 \geq 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení: $[0, 5/3, 2/3]$.

2.9.23. Pomocí simplexové metody řešte úlohu

$$8x_1 + 6x_2 + 11x_3 \to \max$$
$$5x_1 + x_2 + 3x_3 \leq 4 \quad \& \quad 5x_1 + x_2 + 3x_3 \geq 2 \quad \& \quad 2x_1 + 4x_2 + 7x_3 \leq 5$$
$$2x_1 + 4x_2 + 7x_3 \geq 3 \quad \& \quad x_1 + x_2 + x_3 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

2.9.24. Pomocí simplexové metody řešte úlohu

$$5x_1 - 2x_2 + x_3 \to \max$$
$$x_1 + 4x_2 + x_3 \leq 5 \quad \& \quad 2x_1 + x_2 + 3x_3 \geq 2$$
$$x_1 \geq 0 \quad \& \quad x_3 \geq 0.$$

2.9.25. Pomocí simplexové metody řešte úlohu

$$2x_1 + x_2 + 3x_4 \to \max$$
$$9x_1 + 14x_2 - 6x_3 - 6x_4 \leq 2 \quad \& \quad x_1 + x_2 - x_3 - x_4 \geq -1$$
$$-20x_1 - 5x_2 + 5x_3 + 13x_4 = 11 \quad \& \quad 5x_1 + 10x_2 - 2x_3 + 14x_4 = 6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0.$$

2.9.26. Pomocí simplexové metody řešte úlohu

$$-2x_1+3x_2-x_3\to\min$$
$$x_1+2x_2+x_3-x_4=6 \quad \& \quad 2x_1+3x_3+x_4=9 \quad \& \quad 5x_1+6x_2-2x_4=20$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

Řešení: $[0,31/6,7/6,11/2]$.

2.9.27. Pomocí simplexové metody řešte úlohu

$$2x_1+4x_2+4x_3-3x_4\to\max$$
$$2x_1+x_2+x_3=4 \quad \& \quad x_1+4x_2+3x_4=6$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

2.9.28. Pomocí simplexové metody řešte úlohu

$$x_1\to\max$$
$$x_1-x_2+x_3+x_4\leq 8 \quad \& \quad 2x_1+x_2-x_3-x_4\geq 4 \quad \& \quad -x_1+2x_2+x_3-x_4=6$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

Řešení: $[22,14,0,0]$.

2.9.29. Pomocí simplexové metody řešte úlohu

$$3x_1+2x_2+4x_3+8x_4\to\min$$
$$x_1-2x_2+3x_3+6x_4\geq 8 \quad \& \quad -2x_1+5x_2+3x_3-5x_4\leq 3$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

2.9.30. Pomocí simplexové metody řešte úlohu

$$14x_1+x_2+6x_3+5x_4\to\min$$
$$2x_1+x_3\geq 4 \quad \& \quad x_1+x_2+x_3+2x_4\geq 16 \quad \& \quad 2x_1+x_3+x_4\geq 9$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

Řešení: $[0,2,4,5]$.

2.9.31. Pomocí simplexové metody řešte úlohu

$$-2x_1+2x_2+x_3+x_4\to\min$$
$$x_1+2x_2+2x_3+x_4\leq 2 \quad \& \quad x_1-2x_2+x_3+2x_4\geq 3 \quad \& \quad 2x_1-x_2+3x_3\geq 2$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

2.9.32. Pomocí simplexové metody řešte úlohu

$$4x_1+x_2+x_3+3x_4\to\min$$
$$2x_1+x_2+3x_3+x_4\geq 12 \quad \& \quad 3x_1+2x_2+4x_3=5$$
$$2x_1-x_2+2x_3+3x_4=8 \quad \& \quad 3x_1+4x_2+3x_3+x_4\geq 16$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

2.9.33. Pomocí simplexové metody řešte úlohu

$$2x_1+x_2+x_3+x_4\to\min$$
$$x_1+2x_2+x_3+2x_4\leq 7 \quad \& \quad x_1+2x_2+x_3+2x_4\geq 3$$
$$2x_1+3x_2-x_3-4x_4\leq 10 \quad \& \quad x_1+x_2+x_3+x_4=1$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

2.9.34. Pomocí simplexové metody řešte úlohu

$$8x_1+7x_2+17x_3+13x_4\to\min$$
$$2x_1+3x_2+4x_3+4x_4\geq 16 \quad \& \quad 2x_1+x_2+4x_3+3x_4\geq 16 \quad \& \quad x_1+x_2+2x_3+2x_4\geq 9$$
$$x_1\geq 0 \quad \& \quad x_2\geq 0 \quad \& \quad x_3\geq 0 \quad \& \quad x_4\geq 0.$$

Řešení: $[5,0,0,2]$.

2.9.35. Pomocí simplexové metody řešte úlohu

$$-x_1 + x_3 \to \min$$
$$5x_1 + 3x_2 + x_3 = 40 \quad \& \quad x_1 + x_2 + 4x_3 = 10$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

2.9.36. Pomocí simplexové metody řešte úlohu

$$4x_1 + 5x_2 + 7x_3 - x_4 \to \max$$
$$x_1 + x_2 + 2x_3 - x_4 \geq 1 \quad \& \quad 2x_1 - 6x_2 + 3x_3 + x_4 \leq -3 \quad \& \quad -2x_1 + 4x_2 + 2x_3 + 2x_4 = -5$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_4 \geq 0.$$

2.10 Simplexová metoda zprava dolů

Teď již velmi dobře víme, jak funguje simplexový algoritmus, co všechno pro jeho spuštění potřebujeme a jak to případně získat[186]. Klíčovou roli ve vybudování celého algoritmu hrál apriorní požadavek $b \geq 0$, díky kterému se ale občas můžeme dostat do potíží s nalezením výchozího BPB. Jenže simplexová tabulka obsahuje také veškeré informace o duální úloze. Proto asi příliš nepřekvapí, že ji lze využít i pro *duální simplexový algoritmus*. Ten vychází z bodu, který je *duálně přípustný*[187] a s jeho pomocí generujeme další duálně přípustné body s rostoucí (či alespoň neklesající) funkční hodnotou účelové funkce duální úlohy. V souladu se Silnou větou o dualitě (Věta 2.6.5) se pak v případě řešitelné úlohy dopracujeme k tabulce, ve které budeme mít bazické souřadnice *primárně přípustného*[188] bodu a současně v ní nalezneme souřadnice duálně přípustného bodu. Velkou výhodou tohoto postupu je to, že není potřeba explicitně sestavovat duální úlohu a že si vystačíme s toutéž výchozí simplexovou tabulkou jako v případě klasického simplexové metody – i samotný postup je velmi podobný[189]. Počátky této metody sahají do roku 1954, kdy ji nezávisle na sobě zveřejnili Beale v článku [35] také Lemke[190] v [353]. Sedmiletý rozdíl mezi primární a duální simplexovou metodou by ale neměl být příliš překvapující, neboť pochopitelně bylo nutné nejdříve správně porozumět samotné dualitě v lineárním programování a jejímu možnému využití při řešení primární úlohy.

[186] Z nějakého BPB „vyrazíme" vhodným směrem tak, abychom dosáhli snížení funkční hodnoty. Takto postupujeme tak dlouho, dokud v řádku $z_k - c_k$ nebudou jenom nekladné hodnoty.

[187] Tím rozumíme bod $y \in \mathbb{R}^m$ vyhovující všem omezením příslušné duální úlohy.

[188] Tj. bodu vyhovujícího všem omezením primární úlohy.

[189] Měli bychom také zdůraznit, že duální simplexová metoda je totéž jako primární simplexová metoda aplikovaná na duální úlohu. Proto bychom mohli mít jisté pochybnosti o její užitečnosti – avšak tento pohled by byl chybný, neboť pochopitelně řešení primární a duální úlohy toutéž metodou nemusí být stejně výpočetně náročné. Zkrátka existují různé skupiny úloh lineárního programování, pro které jsou některé metody efektivnější než jiné, pro více informací viz [326].

y	c_B	B	x_1	x_2	$\cdots$	x_n	$x_B = A_B^{-1} b$	x_B / λ
y_1	c_{b_1}	x_{b_1}	$\lambda_{b_1,1}$	$\lambda_{b_1,2}$	$\cdots$	$\lambda_{b_1,n}$	x_{b_1}	$x_{b_1} / \lambda_{b_1,s}$
y_2	c_{b_2}	x_{b_2}	$\lambda_{b_2,1}$	$\lambda_{b_2,2}$	$\cdots$	$\lambda_{b_2,n}$	x_{b_2}	$x_{b_2} / \lambda_{b_2,s}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
y_m	c_{b_m}	x_{b_m}	$\lambda_{b_m,1}$	$\lambda_{b_m,2}$	$\cdots$	$\lambda_{b_m,n}$	x_{b_m}	$x_{b_m} / \lambda_{b_m,s}$
		$z_k - c_k$	$z_1 - c_1$	$z_2 - c_2$	$\cdots$	$z_n - c_n$	$\langle c, x \rangle$	

Tabulka 2.22: Simplexová tabulka pro nějaký BPB.

Uvažme (simplexovou) Tabulku 2.22 odpovídající nějakému BPB x v úloze (LP). Do této tabulky jsme navíc přidali[191] jeden sloupec pro duální proměnné $y_1, \ldots, y_m$, viz také Tabulku 2.4. Ze vzorců (2.6.6) a (2.7.7) pak vidíme, že pro $y := A_B^{\top -1} c_B$ a libovolné $k \in \{1, \ldots, n\}$ platí

$$z_k = a_k^\top A_B^{\top -1} c_B = a_k^\top y, \tag{2.10.1}$$

tj. z_k odpovídají levým stranám duálních omezení z ($\mathrm{LP_D}$) pro bod y. To znamená, že tento bod je duálně přípustný právě tehdy, když $z_k - c_k \leq 0$ pro všechna $k \in \{1, \ldots, n\}$. Je-li navíc x primárně přípustný (tj. máme-li $x_B \geq 0$), pak již máme optimální řešení úlohy (LP). Připusťme proto, že $x_j < 0$ pro nějaké $j \in B$. Naším cílem je pomocí vhodného klíčového řádku r a sloupce k určujících pivota $\lambda_{rk} < 0$ docílit primární přípustnosti, aniž bychom porušili duální přípustnost[192]. Pokud se nám to podaří, nalezneme v poslední simplexové tabulce řešení pri-

[190]

Carlton Edward Lemke (11. října 1920 – 12. dubna 2004) by americký matematik, který se věnoval především algebře, matematickému programování, operačnímu výzkumu a statistice. Stejně jako řada jeho vrstevníků musel přerušit své studium kvůli službě v armádě, kam v roce 1940 narukoval jako výsadkář k 82. výsadkové divizi, kde se z něj mimo jiné stala hvězda boxerského týmu. V armádě zůstal do roku 1946. Doktorát získal v roce 1953 na Univerzitě Carnegieho–Mellonových za práci věnovanou extremálním úlohám s lineárními nerovnostmi (*Extremal Problems in Linear Inequalities*). Zde také v letech 1952–1954 pracoval jako lektor, odkud se na další dva roky přesunul do *Knolls Atomic Power Laboratory* provozované společností General Electric ve spolupráci s vládou USA. V letech 1955–1956 byl inženýrem v *Radio Corporation of America* v New Jersey a v roce 1956 nastoupil jako odborný asistent (a později profesor) na technickou vysokou školu *Rensselaer Polytechnic Institute* ve městě Troy v americkém státě New York. V roce 1964 společně s J. T. Howsonem sestrojil algoritmus pro nalezení Nashovy rovnováhy pro konečnou hru dvou hráčů. Za své výsledky v teorii her získal (společně s Johnem F. Nashem) v roce 1978 *John von Neumann Theory Prize* (to je poměrně symbolické vzhledem k tomu, že rozvoj teorie her je snad největším přínosem von Neumanna do oblasti operačního výzkumu). Gomory o Lemkem kdysi prohlásil: „*Carl had about 25 journal publications… but all of them were gems*“. (*Carl měl kolem 25 časopiseckých publikací… ale všechny byly jako drahokam.*)

mární úlohy (LP). Abychom však přechodem od jednoho duálně přípustného bodu k druhému nemuseli přepočítávat i celou simplexovou tabulku, využijeme stejnou modifikovanou Gaussovu eliminaci jako v klasickém simplexovém algoritmu. Zbývá proto už jenom vyřešit otázku vhodné volby klíčového řádku a sloupce, s níž nedojde k porušení duální přípustnosti[193]. Zatímco za klíčový řádek vezmeme libovolný řádek s $x_r < 0$ pro $r \in B$, klíčový sloupec k tentokrát zvolíme tak, aby

$$\frac{z_k - c_k}{\lambda_{rk}} = \min\left\{ \frac{z_j - c_j}{\lambda_{rj}} \;\middle|\; j \in \{1,\dots,n\}\setminus B \;\&\; \lambda_{rj} < 0 \right\} \geq 0. \tag{2.10.2}$$

Potom eliminací s pivotem λ_{rk} budeme v nové simplexové tabulce mít v řádku $z_k - c_k$ hodnoty

$$\Delta_j := z_j - c_j - \frac{\lambda_{rj}}{\lambda_{rk}}(z_k - c_k), \quad j \in \{1,\dots,n\}.$$

Je-li $\lambda_{rj} \geq 0$, pak $\frac{\lambda_{rj}}{\lambda_{rk}}(z_k - c_k) \geq 0$, a tudíž $\Delta_j \leq 0$. Naopak, je-li $\lambda_{rj} < 0$, potom díky volbě (2.10.2) platí

$$\frac{z_j - c_j}{\lambda_{rj}} \geq \frac{z_k - c_k}{\lambda_{rk}} \quad \text{neboli} \quad z_j - c_j \leq \frac{\lambda_{rj}}{\lambda_{rk}}(z_k - c_k),$$

takže opět $\Delta_j \leq 0$. To znamená, že $\Delta_j \leq 0$ pro všechna $j \in \{1,\dots,n\}$, takže tímto postupem získáme nový bod $\widehat{x}$ takový, že odpovídající $\widehat{y} := A_{\hat{B}}^{\top -1} c_B$ bude jistě duálně přípustným bodem. Navíc hodnota účelové funkce duální úlohy se nezmenší, neboť

$$\langle y, b\rangle = c_B^\top A_B^{-1} b = \langle c, x\rangle \leq c_B^\top A_B^{\top -1} b - \frac{x_r}{\lambda_{rk}}(z_k - c_k) = \langle c, \widehat{x}\rangle = \langle \widehat{y}, b\rangle. \tag{2.10.3}$$

Co ale když nelze určit klíčový sloupec, tj. $\lambda_{rj} \geq 0$ pro všechna $j \in \{1,\dots,n\}\setminus B$? V takovém případě r-tý řádek znamená, že $\sum_{j\in B} \lambda_{rj} x_j = x_r$. Jenže my se chceme dopracovat k $x \geq 0$, takže $\sum_{j\in B} \lambda_{rj} x_j \geq 0$, zatímco $x_r < 0$, což je spor ↯, který ukazuje nepřípustnost primární úlohy[194].

Celý postup duální simplexové metody si nyní shrneme a ukážeme na ilustrativním příkladě. Pochopitelně i pro tuto metodu existuje několik variant lišících se různými pravidly pro výběr klíčového řádku či sloupce nebo zabraňujících zacyklení, které se může objevit i zde[195]. Pro spuštění celého algoritmu ale potřebujeme mít nějaký duálně přípustný bod! Jako první (a lehce naivní) pokus lze doplnit matici A o chybějící sloupce jednotkové matice, díky čemuž snadno nalezneme bázi B a odpovídající hodnoty $x_B = b$. Je-li $x_B = b \geq 0$, pak můžeme postupovat klasickým simplexovým algoritmem, zatímco v opačném případě využijeme jeho duální verzi. Ovšem může se stát, že pro nějaké $k \in \{1,\dots,n\}$ není splněna podmínka $z_k - c_k \leq 0$, takže tento bod nebude duálně přípustný. Za chvíli si ukážeme, že tento problém lze vyřešit s pomocí tzv. *umělých omezení*, které jsou duální analogií umělých proměnných z dvoufázové metody pro klasický simplexový algoritmus. Je také užitečné si uvědomit, že kromě duálně přípustných bodů generujeme v průběhu duální simplexové metody posloupnost primárně nepřípustných bodů, tj. až do posledního kroku se po celou dobu výpočtu pohybujeme vně přípustné množiny X, a teprve v posledním kroku získáme primárně přípustný bod (a zároveň řešení). Zatímco v případě klasické simplexové metody můžeme průběh výpočtu přerušit a „spokojit“ se s přípustným bodem jakožto suboptimálním řešením (hodnota účelové funkce v tomto bodě je větší než možné minimum f^*), v případě přerušení výpočtu duální simplexové metody získáme sice bod s menší funkční hodnotou než f^*, ale tento bod je nepřípustný, což je obvykle nepřijatelné! Jinými slovy, pro získání přijatelného výsledku duální simplexovou metodou je nutné, aby výpočet proběhl až dokonce.

Duální simplexový algoritmus

(1) *Inicializace*: danou úlohu převedeme do kanonického tvaru (LP) a doplníme ji tak, aby v matici A byly všechny sloupce jednotkové matice. Proměnné odpovídající těmto sloupcům zavedeme do báze. Je-li $b \geq 0$,

[191] Velmi brzy navíc uvidíme, že v této tabulce budeme potřebovat nahradit sloupec x_B/λ řádkem $(z_k - c_k)/\lambda$.

[192] Jinými slovy se chceme postupem podobným simplexovému algoritmu dopracovat k bázi $\hat{B}$ určující $x_{\hat{B}} \geq 0$, přičemž v každém kroku musí zůstat zachováno $z_k - c_k \leq 0$.

pokračujeme[196] klasickým simplexovým algoritmem, v opačném případě pokračujeme krokem (2).

(2) Nalezneme výchozí duálně přípustný bod a sestavíme si výchozí simplexovou tabulku.

(3) *Index opouštějící bázi.* Určíme *klíčový řádek.* Jestliže $x_r < 0$ pro jediné $r \in B$, vybereme tento index r a příslušná proměnná vystoupí z báze. Jestliže toto nastane pro více indexů $j \in B$, pak vybereme ten, pro který je x_j nejmenší, tj.

$$x_r = \min\{x_j \mid x_j < 0 \ \ \& \ \ j \in B\}.$$

(4) *Index vstupující do báze.* Určíme *klíčový sloupec.*

(i) Platí-li $\lambda_{rj} \geq 0$ pro všechna $j \in \{1,\dots,n\}\setminus B$, je duální úloha neohraničená a primární úloha je nepřípustná. KONEC.

(ii) Je-li $\lambda_{rj} < 0$ alespoň pro jeden index $j \in \{1,\dots,n\}\setminus B$, pak pro tyto hodnoty spočítáme podíly $(z_j - c_j)/\lambda_{rj}$ a za klíčový sloupec k vezmeme takový, že

$$\frac{z_k - c_k}{\lambda_{rk}} = \min\left\{\frac{z_j - c_j}{\lambda_{rj}} \;\middle|\; j \in \{1,\dots,n\}\setminus B \ \ \& \ \ \lambda_{rj} < 0\right\} \geq 0.$$

Realizuje-li se toto minimum ve více indexech k, zvolíme (např.) ten, který je v simplexové tabulce nejvíce vlevo. Odpovídající proměnná vystoupí z báze.

(5) *Aktualizace simplexové tabulky.* Stejnými úpravami jako v kroku (7) pro klasický simplexový algoritmus s využitím pivota λ_{rk} sestavíme novou simplexovou tabulku.

(6) *Test optimality.* Pokud v nové tabulce jsou všechny prvky ve sloupci x_B nezáporné, našli jsme primárně přípustný bod, který je řešením úlohy (LP). Bod x^* získáme s pomocí hodnot ve sloupci x_B doplněním nulovými hodnotami na nebazických pozicích a hodnoty řešení y^* duální úlohy získáme z řádku $z_k - c_k$ ve sloupcích odpovídajících sloupcům jednotkové matice[197] z kroku (1). KONEC. V opačném případě pokračujeme krokem (3).

Příklad 2.10.1

Duální simplexovou metodou vyřešme úlohu

$$x_1 + 4\,x_3 + 3\,x_4 + 2\,x_5 \to \min$$
$$x_1 + x_2 \geq 12 \quad \& \quad x_1 + 2\,x_3 + x_4 \geq 20 \quad \& \quad 2\,x_2 + x_3 + 3\,x_4 + 5\,x_5 \geq 26$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0.$$

Řešení. Danou úlohu nejdříve s využitím pomocných proměnných s_1, s_2 a s_3 převedeme do kanonického tvaru. Navíc jednotlivé rovnice vynásobíme -1 tak, abychom v matici A měli všechny sloupce jednotkové matice, tj.

$$x_1 + 4\,x_3 + 3\,x_4 + 2\,x_5 \to \min$$
$$-x_1 - x_2 + s_1 = -12$$
$$-x_1 - 2\,x_3 - x_4 + s_2 = -20$$

[193] Vzhledem k tomu, co všechno již víme o duální úloze, by nás zde již nemělo překvapit, že tentokrát dojde k výměně rolí klíčového řádku a klíčového sloupce oproti klasickému simplexovému algoritmu.

[194] Tento spor může působit poněkud nepřesvědčivě, takže si tentýž závěr dokážeme exaktnějším způsobem. Uvažme

$$y(\alpha) := A_B^{\top -1} c_B + \alpha\, A_{B,r}^{\top -1},$$

kde $A_{B,r}^{\top -1}$ značí r-tý sloupec $A_B^{\top -1}$, tj. $A_{B,r}^{\top -1} := A_B^{\top -1} e_r$ pro $e_r := (0,\dots,0,1,0,\dots,0)^\top$ představující r-tý prvek standardní kanonické báze $\mathbb{R}^m$. Jelikož

$$0 \geq z_k - c_k = a_k^\top A_B^{\top -1} c_B - c$$

a současně podle (2.7.6) máme $A^\top A_B^{\top -1} = L^\top$, platí

$$A^\top y(\alpha) = A^\top A_B^{\top -1} c_B + \\ + \alpha\, A^\top A_B^{\top -1} e_r \leq \\ \leq c + L^\top e_r = c + \alpha \begin{pmatrix} \lambda_{r1} \\ \vdots \\ \lambda_{rn} \end{pmatrix}.$$

Jelikož pro $r \in B$ je vždy $\lambda_{rj} = 0$ pro všechna $j \in B\setminus\{r\}$ a $\lambda_{rr} = 1$, plyne odtud, že $A^\top y(\alpha) \leq c$ pro libovolné $\alpha \leq 0$. Navíc

$$\langle b, y(\lambda)\rangle = b^\top A_B^{\top -1} c_B + \\ + \alpha\, b^\top A_B^{\top -1} e_r = \\ = x_B^\top c_B + \alpha\, x_B^\top e_r = \\ = \langle x, c\rangle + \alpha\, x_r,$$

což vzhledem k $x_r < 0$ konverguje k ∞ pro $\alpha \to -\infty$. To ukazuje, že duální úloha je (shora) neohraničená, a tudíž primární úloha je skutečně nepřípustná, viz Větu 2.6.2.

[195] To, že se i Lemkeho duální simplexový algoritmus může zacyklit, poprvé ukázal Beale již v roce 1955 na příkladu

$$x_7 \to \min$$
$$x_1 = -\frac{3}{4} + \frac{x_5}{4} + \frac{x_6}{2}$$
$$x_2 = 20 - 8\,x_5 - 12\,x_6$$
$$x_3 = -\frac{1}{2} - x_5 - \frac{x_6}{2} + x_7$$
$$x_4 = 6 + 9\,x_5 + 3\,x_6$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0$$
$$x_4 \geq 0 \quad \& \quad x_5 \geq 0$$
$$x_6 \geq 0 \quad \& \quad x_7 \geq 0$$

a odpovídající „duální úloze“

$$-\frac{3}{4}\,y_1 + 20\,y_2 - \frac{y_3}{2} + 6\,y_4 + \to \min$$
$$y_5 = -\frac{y_1}{4} + 8\,y_2 + y_3 - 9\,y_4$$

$$y_6 = -\frac{y_1}{2} + 12\,y_2 + \frac{y_3}{2} - 3\,y_4$$
$$y_7 = 1 - y_3$$
$$y_1 \geq 0 \quad \& \quad y_2 \geq 0 \quad \& \quad y_3 \geq 0$$
$$y_4 \geq 0 \quad \& \quad y_5 \geq 0$$
$$y_6 \geq 0 \quad \& \quad y_7 \geq 0.$$

Hledaná minimální hodnota obou úloh je $5/4$, která nastane v bodech $x^* = [0,2,0,21/2,0,3/2,5/4]$ a $y^* = [1,0,1,0,3/4,0,0]$. Ve všech krocích duálního simplexového algoritmu volíme jako klíčový řádek ten, ve kterém je hodnota ve sloupci x_B nejmenší, tj. v souladu s pravidlem popsaným v kroku (3) duálního simplexového algoritmu, zatímco v případě shody pro výběr klíčového sloupce volíme ten, který odpovídá proměnné s nejmenším indexem (při standardním uspořádání je v tabulce nejvíce vlevo). Pak v první úloze dojde k cyklu délky 7, tj. po 6 krocích obdržíme stejnou tabulku jako na začátku, přičemž ve všech krocích budeme mít duálně přípustné body s nulovou funkční hodnotou, viz [36]. Toto samozřejmě hrozí pouze v případě degenerované duální úlohy a vyhnout se tomu můžeme např. náhodnou volbou klíčového sloupce v případě shody nebo doplněním duálního algoritmu o lexikografické pravidlo podobně jako v případě primární úlohy, viz závěr Podkapitoly 3.5. V nedegenerované úloze k zacyklení dojít nemůže, neboť v takovém případě se hodnoty účelové funkce duální úlohy v průběhu duálního algoritmu zvětšují, viz (2.10.3).

[196] Právě se zde se dostáváme na rozcestí mezi klasickou (primární) simplexovou metodou a duální simplexovou metodou. Buď budou v matici A všechny sloupce jednotkové matice nebo bude $b \geq 0$. Splnění obou podmínek současně sice umožní přímé využití klasického simplexového algoritmu, ale z praktického pohledu tato situace asi nebude příliš častá (určitě nastane v případě omezení $Ax \leq b \;\&\; x \geq 0$). Nastane-li pouze jedna možnost, můžeme si vybrat: buď preferujeme $b \geq 0$ (to nás nejspíše přivede k dvoufázové metodě pro klasický simplexový algoritmus) nebo chceme sloupce jednotkové matice v A (pak využijeme duální simplexový algoritmus).

[197] V takovém případě máme totiž podle (2.10.1) pro indexy k odpovídající postupně jednotlivým sloupcům jednotkové matice rovnost $z_k = a_k^\top y^* = y_k^*$. Pokud jsme k matici A přidali všechny sloupce jed-

$$-2\,x_2 - x_3 - 3\,x_4 - 5\,x_5 + s_3 = -26$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0$$
$$s_1 \geq 0 \quad \& \quad s_2 \geq 0 \quad \& \quad s_3 \geq 0.$$

Za výchozí bázi vezmeme $B = \{s_1, s_2, s_3\}$ s $A_B = I$ a odpovídajícím bodem $x_B = b = [-12, 20, -26]$, pro který si sestavíme výchozí simplexovou tabulku, viz Tabulku 2.23.

c_B	B	x_1	x_2	x_3	x_4	x_5	s_1	s_2	s_3	x_B
0	s_1	−1	−1	0	0	0	1	0	0	−12
0	s_2	−1	0	−2	−1	0	0	1	0	−20
0	s_3	0	−2	−1	−3	−5	0	0	1	−26
	c	1	0	4	3	2	0	0	0	
	z_k	0	0	0	0	0	0	0	0	
	$z_k - c_k$	−1	0	−4	−3	−2	0	0	0	0
	$\frac{z_k - c_k}{\lambda}$									

Tabulka 2.23: Výchozí tabulka duální simplexové metody pro Příklad 2.10.1.

Odtud vidíme, že pro nebazické pozice jsou v řádku $z_k - c_k$ pouze nekladná čísla, takže bod $y = [-12, -20, -26]$ je duálně přípustný a můžeme spustit duální simplexový algoritmus, viz Tabulku 2.24.

c_B	B	x_1	x_2	x_3	x_4	x_5	s_1	s_2	s_3	x_B
0	s_1	−1	−1	0	0	0	1	0	0	−12
0	s_2	−1	0	−2	−1	0	0	1	0	−20
0	s_3	0	(-2)	−1	−3	−5	0	0	1	−26 →
	c	1	0	4	3	2	0	0	0	
	z_k	0	0	0	0	0	0	0	0	
	$z_k - c_k$	−1	0	−4	−3	−2	0	0	0	0
	$\frac{z_k - c_k}{\lambda}$	✖	$\frac{0}{-2}$ ↑	$\frac{-4}{-1}$	$\frac{-3}{-3}$	$\frac{-2}{-5}$	✖	✖	✖	
	s_1	−1	0	1/2	3/2	5/2	1	0	−1/2	1
	s_2	(-1)	0	−2	−1	0	0	1	0	−20 →
	x_2	0	1	1/2	3/2	5/2	0	0	−1/2	13
	$z_k - c_k$	−1	0	−4	−3	−2	0	0	0	0
	$\frac{z_k - c_k}{\lambda}$	1 ↑	✖	2	3	✖	✖	✖	✖	
	s_1	0	0	5/2	5/2	5/2	1	−1	−1/2	21
	x_1	1	0	2	1	0	0	−1	0	20
	x_2	0	1	1/2	3/2	5/2	0	0	−1/2	13
	$z_k - c_k$	0	0	−2	−2	−2	0	−1	0	20

Tabulka 2.24: Duální simplexový algoritmus pro Příklad 2.10.1.

V první části Tabulky 2.24 vidíme, že ve sloupci x_B jsou tři záporná čísla. Proto jsme jako klíčový řádek zvolili ten, ve kterém je tato hodnota nejmenší, tj. řádek pro proměnnou s_3, která v tomto kroku vystoupí z báze. Nyní pro všechny záporné koeficienty příslušné nebazickým proměnným v řádku pro s_3 vypočteme podíly $(z_k - c_k)/\lambda$. Minimum z těchto hodnot nastane ve sloupci pro x_2, takže tato proměnná vstoupí do báze a prvek $\lambda_{32} = -2$ je pivotem, s jehož pomocí získáme druhou část Tabulky 2.24. Ve druhém kroku postupujeme obdobně, neboť stále nemáme primárně přípustný bod – ve sloupci x_B „zůstala" jedna záporná hodnota, takže odpovídající řádek je nyní klíčovým a proměnná s_2 vystoupí z báze. Následným vypočtením podílů $(z_k - c_k)/\lambda$ zjistíme, že klíčový sloupec odpovídá x_1 a tato proměnná vstoupí do báze. Eliminací s pivotem $\lambda_{21} = -1$ dostaneme poslední část Tabulky 2.24. Jelikož v ní už jsou všechna složky sloupce x_B nezáporné, nalezli jsme primárně přípustný bod a algoritmus končí. Řešením dané úlohy v kanonickém tvaru je $x = [20,13,0,0,0,21,0,0]$, a tedy řešením původní úlohy je $x^* = [20,13,0,0,0]$ s hodnotou $f^* = 20$.
V Tabulce 2.24 také máme duálně přípustné body pro úlohu

$$-12\,y_1 - 20\,y_2 - 26\,y_3 \to \max$$
$$-y_1 - y_2 \le 1 \quad \& \quad -y_1 - 2\,y_3 \le 0 \quad \& \quad -2\,y_2 - y_3 \le 4$$
$$-y_2 - 3\,y_3 \le 3 \quad \& \quad -5\,y_3 \le 2$$
$$y_1 \le 0 \quad \& \quad y_2 \le 0 \quad \& \quad y_2 \le 0,$$

která je duální úlohou k dané primární úloze v kanonickém tvaru. Z poslední části této tabulky navíc vidíme, že řešením uvedené duální úlohy je bod $y^* = [0,-1,0]$, neboť sloupce jednotkové matice v omezeních primární úlohy postupně odpovídají proměnným s_1, s_2 a s_3, pro které máme v posledním řádku $z_k - c_k$ hodnoty 0, -1 a 0. Tímtéž způsobem také můžeme dopočítat, že výchozí duální bod byl $y^{[0]} = [0,0,0]$, ze kterého jsme se přesunuli k $y^{[1]} = [0,0,0]$. To je ale ten stejný bod – pouze jsme změnili bázi, což vlastně velmi úzce souvisí s degenerovaností dané úlohy, kterou snadno můžeme vidět ze souřadnic x^*.
Na závěr ještě dodejme, že duální úloha k původní dané úloze má podobu

$$12\,y_1 + 20\,y_2 + 26\,y_3 \to \max$$
$$y_1 + y_2 \le 1 \quad \& \quad y_1 + 2\,y_3 \le 0 \quad \& \quad 2\,y_2 + y_3 \le 4 \quad \& \quad y_2 + 3\,y_3 \le 3 \quad \& \quad 5\,y_3 \le 2$$
$$y_1 \ge 0 \quad \& \quad y_2 \ge 0 \quad \& \quad y_2 \ge 0,$$

která se od předchozí duální úlohy liší pouze znaménky u y_1, y_2 a y_3, takže jejím řešením je[198] bod $[0,1,0]$.

V předchozím příkladu jsme zavedením pomocných proměnných s_1, s_2, s_3 snadno nalezli výchozí bod, který byl navíc i duálně přípustný. To je velmi „šťastná" situace, ke které ovšem nemusí vždy dojít. Proto si nyní ukážeme, jak najít výchozí duálně přípustný bazický bod v případě, kdy máme bazický bod s alespoň jednou kladnou hodnotou na nebazické pozici v řádku $z_k - c_k$. Základní myšlenka je velmi podobná jako u klasického simplexového algoritmu, pro který jsme si v Podkapitole 2.9 ukázali dvě (velmi podobné) metody. Nyní však místo umělých proměnných (ať již zaváděných společně s velkým M či rozděleným do dvou fází) využijeme tzv. *umělá omezení*[199,200], což dává i název této metodě.

Danou úlohu rozšíříme o omezení ve tvaru[201]

$$\sum_{j \notin B} x_j \le M,$$

kde M je opět velké číslo, se kterým počítáme jako s parametrem. Pokud by M nebylo dostatečně velké, mohli bychom touto úpravou ztratit některé body z původní přípustné mno-

notkové matice, pak odpovídající $c_k = 0$, takže platí dokonce $z_k - c_k = z_k = y_k^*$, tj. souřadnice y_k^* získáme přímo z řádku $z_k - c_k$. Pokud ale bylo možné přidat pouze některé sloupce jednotkové matice, pak příslušnou složku řešení y^* dopočteme z řádku $z_k - c_k$ přičtením c_k, tj. $y_k^* = (z_k - c_k) + c_k$.

[198] To, že $[0,-1,0]$ či $[0,1,0]$ jsou řešeními duálních úloh, plyne ihned z Důsledku 2.6.3, neboť snadno můžeme otestovat přípustnost těchto bodů a odpovídající účelové funkce nabývají v těchto bodech hodnotu 20. Můžete si také sami vyzkoušet vyřešit tyto úlohy přímo pomocí klasického simplexového algoritmu a porovnat průběh výpočtu s Tabulkou 2.24.

[199] Angl. *artificial constraints*.

[200] No, ono to zas tak odlišné není. Ve skutečnosti totiž v této metodě nejde o nic jiného než o metodu velkého M aplikovanou na duální úlohu. Akorát to neděláme explicitně pro duální problém, ale spíše implicitně prostřednictvím duální simplexové metody a simplexové tabulky pro primární úlohu.

[201] Mnohdy bývá uvažováno umělé omezení zahrnující všechny proměnné, tj. $x_1 + \ldots x_n \le M$. Avšak v dalším komentáři si objasníme, proč je využití pouze nebazických proměnných zcela dostačující.

žiny. Toto omezení pak přičtením nové doplňkové nezáporné proměnné s_{n+1} převedeme do tvaru rovnosti a tuto proměnnou vezmeme za součást výchozí báze. V prvním kroku tuto proměnnou vyvedeme z báze[202] a za klíčový sloupec vezmeme ten, pro který je v řádku $z_k - c_k$ největší[203] kladné číslo. Nutnost této volby a zároveň její genialita tkví v tom, že následná eliminace pomocí odpovídajícího pivota v souladu s pravidly simplexového algoritmu nás vždy přivede k duálně přípustnému bodu bez ohledu na (ne-)přípustnost primární úlohy či její nehraničenost na původní množině (nezapomínejme, že umělé omezení nám pro libovolnou volbu M zaručuje ohraničenost jen nebazických proměnných), tj. příslušná duální úloha je vždy přípustná[204]! Dále již pokračujeme klasickým duálním simplexovým algoritmem, pro který mohou nastat následující čtyři situace:

(i) Jestliže duální simplexový algoritmus skončí v důsledku nemožnosti určení klíčového sloupce, je duální úloha neohraničená, a tudíž primární úloha je nepřípustná.

(ii) Jestliže jsme nalezli řešení a hodnota úlohy závisí na M, pak je primární úloha neohraničená[205].

(iii) Jestliže jsme nalezli řešení s hodnotou úlohy nezávisející na M a s proměnnou s_{n+1} v bázi, pak původní úloha má řešení, které snadno získáme z poslední simplexové tabulky.

(iv) Jestliže jsme nalezli řešení s hodnotou úlohy nezávisející na M a s proměnnou s_{n+1} mimo bázi, pak má původní úloha dokonce nekonečně mnoho řešení[206], která získáme libovolnou volbou $M \geq 0$, pro kterou je $x_B \geq 0$.

Všechny možné výsledky pro duální simplexový algoritmus doplněný o umělé omezení si nyní ilustrujeme v následujících čtyřech příkladech.

Příklad 2.10.2

Duální simplexovou metodou doplněnou o umělé omezení vyřešme úlohu lineárního programování

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 \leq 2 \quad \& \quad -3x_1 + x_2 \geq 3$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení. Úlohu nejprve převedeme do kanonického tvaru vynásobením druhého omezení číslem -1 a přidáním dvou umělých proměnných s_1 a s_2. Tím dostaneme úlohu ve tvaru

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 + s_1 = 2$$
$$3x_1 - x_2 + s_2 = -3$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad s_1 \geq 0 \quad \& \quad s_2 \geq 0.$$

Vezmeme-li za výchozí bod $x = [0,0,2,-3]$, který je bazický a nepřípustný, pak odpovídající výchozí simplexová tabulka je znázorněna v Tabulce 2.25.

c_B	B	x_1	x_2	s_1	s_2	x_B
0	s_1	1	1	1	0	2
0	s_2	3	−1	0	1	−3
	c	−2	−1	0	0	
	z	0	0	0	0	
	$z_k - c_k$	2	1	0	0	0

Tabulka 2.25: Výchozí simplexová tabulka pro Příklad 2.10.2.

[202] Řádek pro s_{n+1} je tedy v tuto chvíli klíčovým.

[203] Pokud se to stane pro více sloupců, vybereme např. ten, který je nejvíce vlevo.

[204] Vskutku, přidáním umělého omezení k primární úloze dostaneme (pro jednoduchost se omezme pouze na situaci s bází $B = \{1,\dots,m\}$)

$$\langle c, x\rangle \to \min$$
$$Ax = b$$
$$x_{m+1} + \cdots + x_n + s_{n+1} = M$$
$$x_1 \geq 0 \quad \& \quad \cdots \quad \& \quad s_{n+1} \geq 0.$$

Matici určující levou stranu všech rovnostních omezení tak můžeme zapsat jako

$$\begin{pmatrix} & & & 0 \\ A_B & & A_N & \vdots \\ & & & 0 \\ 0 \;\cdots\; 0 & 1 & \cdots\; 1 & 1 \end{pmatrix} =: \widehat{A},$$

kde $N = \{m+1,\dots,n\}$. Umělé omezení sice nemá žádný vliv na přípustnost primární úlohy, ovšem zaručí přípustnost duální úlohy, která bude mít podobu

$$\langle y, b\rangle + M\, y_{m+1} \to \min$$
$$\widehat{A}^\top \begin{pmatrix} y \\ y_{m+1} \end{pmatrix} \leq \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

Rozepsáním do bloků totiž dostaneme sadu $n+1$ omezení

$$A_B^\top y \leq \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$
$$A_N^\top y + \begin{pmatrix} y_{m+1} \\ \vdots \\ y_{m+1} \end{pmatrix} \leq \begin{pmatrix} c_{m+1} \\ \vdots \\ c_n \end{pmatrix}$$
$$y_{m+1} \leq 0$$

a díky regulárnosti matice A_B bude první část jistě splněna pro

$$\widehat{y} := A_B^{\top -1} \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix},$$

takže volbou dostatečně malého $y_{m+1} \leq 0$ zajistíme i splnění druhé sady nerovností.

Z této tabulky vidíme, že pro obě nebazické proměnné x_1 a x_2 jsou v řádku $z_k - c_k$ kladná čísla. To znamená, že výchozí bod není duálně přípustný. Protože současně je ve sloupci x_B záporné číslo, není tento bod ani primárně přípustný. Využijeme proto metodu umělých omezení, tj. přidáme

$$x_1 + x_2 \le M \quad \text{neboli} \quad x_1 + x_2 + s_3 = M$$

pro pomocnou proměnnou $s_3 \ge 0$. Tuto proměnnou vezmeme jako součást výchozího bazického bodu, sestavíme novou výchozí simplexovou tabulku a v prvním kroku vyvedeme tuto proměnnou z báze. Celý výpočet je zobrazen v Tabulce 2.26.

c_B	B	x_1	x_2	s_1	s_2	s_3	x_B
0	s_1	1	1	1	0	0	2
0	s_2	3	−1	0	1	0	−3
0	s_3	(1)	1	0	0	1	$M \rightarrow$
	c	−2	−1	0	0	0	
	z	0	0	0	0	0	
	$z_k - c_k$	2 ↑	1	0	0	0	0
	s_1	0	0	1	0	−1	$2 - M$
	s_2	0	(-4)	0	1	−3	$-3 - 3M \rightarrow$
	x_1	1	1	0	0	1	M
	$z_k - c_k$	0	−1	0	0	−2	$-2M$
	$\frac{z_k - c_k}{\lambda}$	✖	1/4 ↑	✖	✖	2/3	
	s_1	0	0	1	0	(-1)	$2 - M \rightarrow$
	x_2	0	1	0	−1/4	3/4	$3/4 + 3M/4$
	x_1	1	0	0	1/4	1/4	$-3/4 + M/4$
	$z_k - c_k$	0	0	0	−1/4	−5/4	$3/4 - 5M/4$
	$\frac{z_k - c_k}{\lambda}$	✖	✖	✖	✖	5/4 ↑	
	s_3	0	0	−1	0	1	$-2 + M$
	x_2	0	1	3/4	−1/4	0	9/4
	x_1	1	0	1/4	1/4	0	−1/4
	$z_k - c_k$	0	0	−5/4	−1/4	0	−7/4

Tabulka 2.26: Duální simplexová metoda s využitím umělého omezení pro Příklad 2.10.2.

V prvním kroku jsme vybrali klíčový sloupec odpovídající největší kladné hodnotě v řádku $z_k - c_k$, což je první sloupec s hodnotou 2. Tento sloupec odpovídá proměnné x_1, která nahradí proměnnou s_3 v bázi. Eliminací s pivotem $\lambda_{3,1}$ podle pravidel simplexového algoritmu získáme druhou část této tabulky. Zde jsou v řádku $z_k - c_k$ již pouze nekladná čísla a pro nebazické proměnné dokonce záporná čísla, takže můžeme pokračovat duálním simplexovým algoritmem. Klíčový řádek odpovídá hodnotě $-3 - 3M < 0$, která je nejmenší[207] ze všech tří hodnot ve sloupci x_B. V tomto řádku jsou dvě záporná čísla, takže dopočítáme podíly $(z_k - c_k)/\lambda$, z čehož vidíme, že klíčový sloupec odpovídá proměnné x_2, která se tedy vstoupí do báze na úkor s_2. Po provedení eliminace s pivotem $\lambda_{2,2}$ dostaneme třetí část tabulky a analogickým postupem s pivotem $\lambda_{1,5}$ se dopočítáme k poslední části této tabulky. V ní ještě stále nemáme primární

[205] Nemělo by toto být navíc závislé na znaménku koeficientu u parametru M? Ne. Již se nám podařilo pro všechna dostatečně velká M najít primárně i duálně přípustný bod a máme dvě možnosti, z nichž ale jedna je v přímém rozporu se Silnou (Slabou) větou o dualitě:

(a) Je-li koeficient u M záporný, pak bude hodnota pro $M \to \infty$ klesat k $-\infty$, a tedy primární úloha bude neohraničená. Ovšem pozor, nenechme se zde zmást, toto neznamená nepřípustnost duální úlohy, a tedy spor s tím, že máme duálně i primárně přípustný bod. My jsme nalezli duálně přípustný bod pro úlohu doplněnou o umělé omezení, ale zmíněná nepřípustnost se týká duální úlohy k původní úloze (LP).

(b) Kdyby byl koeficient u M kladný, pak by v případě $M \to \infty$ byla neohraničená duální úloha (máme duálně přípustný bod, jehož souřadnice nezávisí na M), a tudíž primární úloha by musela být nepřípustná, což je ale spor s tím, že jsme již nalezli primárně i duálně přípustný bod ↯ (nezapomínejme, že přidáním umělého omezení jsme docílili nejvýše „nějakého oříznutí“ přípustné množiny pro původní úlohu).

[206] Je-li s_{n+1} mimo bázi, pak nalezené řešení splňuje $\sum_{j \notin B} x_j^* = M$, tj. leží na hranici přípustné množiny určené umělým omezením s $s_{n+1} = 0$. V takovém případě získáme pro každou dostatečně velkou hodnotu M další řešení původní úlohy (stále leží na této hranici, kterou s rostoucím M posouváme).

přípustný bod a musíme pokračovat v duálním algoritmu. Klíčovým řádkem může být pouze ten třetí odpovídající proměnné x_1, která by tedy měla vystoupit z báze. Jenže v tomto řádku není žádný záporný koeficient, takže nelze vybrat klíčový sloupec a duální simplexový algoritmus končí s tím, že duální úloha je neohraničená. To znamená, že primární úloha je nepřípustná. V průběhu duálního simplexového algoritmu jsme vygenerovali duálně přípustné body $y^{[0]} = [0,0,-2]$, $y^{[1]} = [0,-1/4,-5/4]$ a v posledním kroku $y^{[2]} = [-5/4,-1/4,0]$ pro duální úlohu odpovídající původní úloze v kanonickém tvaru s umělým omezením, tj.

$$2\,y_1 - 3\,y_2 + M\,y_3 \to \max$$
$$y_1 + 3\,y_2 + y_3 \le -2 \quad \& \quad y_1 - y_2 + y_3 \le -1$$
$$y_1 \le 0 \quad \& \quad y_2 \le 0 \quad \& \quad y_3 \le 0.$$

Všechna tato omezení splňuje také např. bod $y(\alpha) = [-5/4 - \alpha/4, -1/4 - \alpha/4, 0]$ pro libovolné $\alpha \ge 0$ s hodnotou účelové funkce $-7/4 + \alpha/4$, která pro $\alpha \to \infty$ roste nade všechny meze, tj. duální úloha je skutečně neohraničená.

Příklad 2.10.3

Duální simplexovou metodou doplněnou o umělé omezení vyřešme úlohu lineárního programování

$$4\,x_1 - 5\,x_2 \to \min$$
$$2\,x_1 + 2\,x_2 \ge 4 \quad \& \quad x_1 - x_2 \ge 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. Stejně jako v předchozím příkladě si úlohu nejprve převedeme do kanonického tvaru, čímž získáme omezení

$$-2\,x_1 - 2\,x_2 + s_1 = -4 \quad \& \quad -x_1 + x_2 + s_2 = -3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad s_1 \ge 0 \quad \& \quad s_2 \ge 0.$$

Odpovídající výchozí simplexová tabulka je znázorněna v Tabulce 2.27.

c_B	B	x_1	x_2	s_1	s_2	x_B
0	s_1	−2	−2	1	0	−4
0	s_2	−1	1	0	1	−3
	c	4	−5	0	0	
	z	0	0	0	0	
	$z_k - c_k$	−4	5	0	0	0

Tabulka 2.27: Výchozí simplexová tabulka pro Příklad 2.10.3.

Z této tabulky vidíme, že pro nebazickou proměnnou x_2 je v řádku $z_k - c_k$ kladné číslo. To znamená, že výchozí bod není duálně přípustný. Protože současně je ve sloupci x_B záporné číslo, není tento bod ani primárně přípustný. Využijeme proto metodu umělých omezení, tj. opět přidáme omezení

$$x_1 + x_2 \le M \quad \text{neboli} \quad x_1 + x_2 + s_3 = M$$

pro pomocnou proměnnou $s_3 \ge 0$. Tuto proměnnou vezmeme jako součást výchozího bazického bodu, sestavíme novou výchozí simplexovou tabulku a v prvním kroku vyvedeme tuto proměnnou z báze. Celý výpočet je zobrazen v Tabulce 2.28.

[207] Zdůrazněme, že toto je ta správná volba podle duálního simplexového algoritmu, neboť

$$-3 - 3\,M < 2 - M$$

pro libovolné $M \ge 0$.

c_B	B	x_1	x_2	s_1	s_2	s_3	x_B
0	s_1	-2	-2	1	0	0	-4
0	s_2	-1	1	0	1	0	-3
0	s_3	1	(1)	0	0	1	$M \rightarrow$
	c	4	-5	0	0	0	
	z	0	0	0	0	0	
	$z_k - c_k$	-4	$5\uparrow$	0	0	0	0
	s_1	0	0	1	0	2	$-4+2M$
	s_2	(-2)	0	0	1	-1	$-3-M \rightarrow$
	x_2	1	1	0	0	1	M
	$z_k - c_k$	-9	0	0	0	-5	$-5M$
	$\frac{z_k - c_k}{\lambda}$	$9/2\uparrow$	✖	✖	✖	5	
	s_1	0	0	1	0	2	$-4+2M$
	x_1	1	0	0	$-1/2$	$1/2$	$3/2+M/2$
	x_2	0	1	0	$1/2$	$1/2$	$-3/2+M/2$
	$z_k - c_k$	0	0	0	$-9/2$	$-1/2$	$27/2-M/2$

Tabulka 2.28: Duální simplexová metoda s využitím umělého omezení pro Příklad 2.10.3.

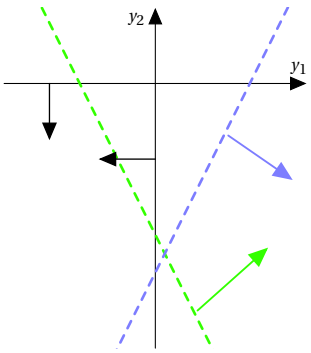
V prvním kroku jsme vybrali klíčový sloupec odpovídající největší (a jediné) kladné hodnotě v řádku $z_k - c_k$, což je druhý sloupec s hodnotou 5. Tento sloupec odpovídá proměnné x_2, která nahradí proměnnou s_3 v bázi. Eliminací s pivotem $\lambda_{3,2}$ podle pravidel simplexového algoritmu získáme druhou část této tabulky. Zde jsou v řádku $z_k - c_k$ již pouze nekladná čísla a pro nebazické proměnné dokonce záporná čísla, takže můžeme pokračovat duálním simplexovým algoritmem. Klíčový řádek odpovídá hodnotě $-3-M<0$, která je nejmenší (a zároveň jediná záporná) ze všech tří hodnot ve sloupci x_B. V tomto řádku jsou dvě záporné hodnoty a vypočteme podíly $(z_k - c_k)/\lambda$, z čehož vidíme, že klíčový sloupec odpovídá proměnné x_1. Po provedení eliminace s pivotem $\lambda_{2,1}$ dostaneme poslední část tabulky. Nyní již máme bod, který je duálně i primárně přípustný, tj. řešením primární úlohy s umělým omezením je $x = [3/2+M/2, -3/2+M/2, -4+2M, 0, 0]$ a odpovídající duální úlohy $y = [0, -9/2, -1/2]$. Avšak hodnota této úlohy $27/2 - M/2$ závisí na M, takže daná primární úloha je neohraničená.

Správnost tohoto výsledku můžeme snadno ověřit bližším pohledem na duální úlohu., která má pro úlohu zahrnující umělé omezení podobu

$$-4\,y_1 - 3\,y_2 + M\,y_3 \rightarrow \max$$
$$-2\,y_1 - y_2 + y_3 \le 4 \quad \& \quad -2\,y_1 + y_2 + y_3 \le -5$$
$$y_1 \le 0 \quad \& \quad y_2 \le 0 \quad \& \quad y_3 \le 0.$$

Tato úloha je nepochybně přípustná, neboť jsme pro ni našli v Tabulce 2.28 přípustný bod $y = [0, -9/2, -1/2]$. Jenže zde je $y_3 \neq 0$, což znamená, že vynecháním této souřadnice dostaneme bod $\bar{y} = [0, -9/2]$, který již není přípustný pro duální úloha k úloze bez umělého omezení. Tato úloha má podobu

$$-4\,y_1 - 3\,y_2 \rightarrow \max$$



Obrázek 2.38: „Přípustná“ množina pro duální úlohu z Příkladu 2.10.3.

s podmínkami

$$-2y_1 - y_2 \le 4 \quad \& \quad -2y_1 + y_2 \le -5$$
$$y_1 \le 0 \quad \& \quad y_2 \le 0$$

a z Obrázku 2.38 vidíme, že je skutečně nepřípustná.

Příklad 2.10.4

Duální simplexovou metodou doplněnou o umělé omezení vyřešme úlohu lineárního programování

$$-x_1 - 6x_2 \to \min$$
$$x_1 + x_2 \le 20 \quad \& \quad x_1 + x_2/2 \ge 1 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. Stejně jako v předchozím příkladě si úlohu nejprve převedeme do kanonického tvaru, čímž získáme omezení

$$x_1 + x_2 + s_1 = 20 \quad \& \quad -x_1 - x_2/2 + s_2 = -1$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad s_1 \ge 0 \quad \& \quad s_2 \ge 0.$$

Odpovídající výchozí simplexová tabulka je znázorněna v Tabulce 2.29.

c_B	B	x_1	x_2	s_1	s_2	x_B
0	s_1	1	1	1	0	20
0	s_2	−1	−1/2	0	1	−1
	c	−1	−6	0	0	
	z	0	0	0	0	
	$z_k - c_k$	1	6	0	0	0

Tabulka 2.29: Výchozí simplexová tabulka pro Příklad 2.10.4.

Z této tabulky vidíme, že pro nebazické proměnné x_1 a x_2 jsou v řádku $z_k - c_k$ kladná čísla. To znamená, že výchozí bod není duálně přípustný. Protože současně je ve sloupci x_B záporné číslo, není tento bod ani primárně přípustný. Využijeme proto metodu umělých omezení, tj. přidáme

$$x_1 + x_2 \le M \quad \text{neboli} \quad x_1 + x_2 + s_3 = M$$

pro pomocnou proměnnou $s_3 \ge 0$. Tuto proměnnou vezmeme jako součást výchozího bazického bodu, sestavíme novou výchozí simplexovou tabulku a v prvním kroku vyvedeme tuto proměnnou z báze. Celý výpočet je zobrazen v Tabulce 2.30.

c_B	B	x_1	x_2	s_1	s_2	s_3	x_B
0	s_1	1	1	1	0	0	20
0	s_2	−1	−1/2	0	1	0	−1
0	s_3	1	(1)	0	0	1	$M \to$
	c	−1	−6	0	0	0	
	z	0	0	0	0	0	
	$z_k - c_k$	1	6 ↑	0	0	0	0

c_B	B	x_1	x_2	s_1	s_2	s_3	x_B
	s_1	0	0	1	0	(-1)	$20-M\rightarrow$
	s_2	$-1/2$	0	0	1	$1/2$	$-1+M/2$
	x_2	1	1	0	0	1	M
	z_k-c_k	-5	0	0	0	-6	$-6M$
	$\frac{z_k-c_k}{\lambda}$	✖	✖	✖	✖	$6\uparrow$	
	s_3	0	0	-1	0	1	$M-20$
	s_2	$-1/2$	0	$1/2$	1	0	9
	x_2	1	1	1	0	0	20
	z_k-c_k	-5	0	-6	0	0	-120

Tabulka 2.30: Metoda umělých omezení pro Příklad 2.10.4.

V prvním kroku jsme vybrali klíčový sloupec odpovídající největší kladné hodnotě v řádku z_k-c_k, což je druhý sloupec s hodnotou 6. Tento sloupec odpovídá proměnné x_2, která nahradí proměnnou s_3 v bázi. Eliminací s pivotem $\lambda_{3,2}$ podle pravidel simplexového algoritmu získáme druhou část této tabulky. Zde jsou v řádku z_k-c_k již pouze nekladná čísla a pro nebazické proměnné dokonce záporná čísla, takže můžeme pokračovat duálním simplexovým algoritmem. Klíčový řádek odpovídá hodnotě $20-M<0$, která je nejmenší ze všech tří hodnot ve sloupci x_B. V tomto řádku je záporná pouze hodnota $\lambda_{1,5}$, takže formálně vypočteme podíly $(z_k-c_k)/\lambda$, z čehož vidíme, že klíčový sloupec odpovídá proměnné s_3, která se tedy vrátí zpět do báze na úkor s_1. Po provedení eliminace s pivotem $\lambda_{1,5}$ dostaneme poslední část tabulky. Jelikož M volíme dostatečně velké, jsou ve sloupci x_B již pouze kladné hodnoty, čímž se nám podařilo najít bod, který je primárně i duálně přípustný, a tudíž i řešením úlohy v kanonickém tvaru. Jelikož navíc hodnota účelové funkce nezávisí na M, je řešením původní úlohy bod $x^*=[0,20]$ s hodnotou $f^*=-120$.
Duální úloha k dané úloze doplněné o umělé omezení má podobu

$$20\,y_1-y_2+M\,y_3\rightarrow\max$$
$$y_1-y_2+y_3\le-1\quad\&\quad y_1-y_2/2+y_3\le-6$$
$$y_1\le0\quad\&\quad y_2\le0\quad\&\quad y_3\le0.$$

Z poslední části Tabulky 2.30 vidíme, že jejím řešením je $y^*=[-6,0,0]$ s hodnotou -120. Všimněme si také, že po prvním kroku jsme měli duálně přípustný bod $[0,0,-6]$ s hodnotou $-6M$, která pro $M>6$ bude menší než -120. Duální úlohu k původní úloze v kanonickém tvaru bez umělého omezení snadno získáme vynecháním proměnné y_3.

Příklad 2.10.5

Duální simplexovou metodou doplněnou o umělé omezení vyřešme úlohu lineárního programování

$$2\,x_1-2\,x_2\rightarrow\min$$
$$2\,x_1+x_2\ge3\quad\&\quad x_1-x_2\ge-3$$
$$x_1\ge0\quad\&\quad x_2\ge0.$$

Řešení. Stejně jako v předchozím příkladě si úlohu nejprve převedeme do kanonického tvaru, čímž získáme omezení

$$-2x_1 - x_2 + s_1 = -3 \quad \& \quad -x_1 + x_2 + s_2 = 3$$

$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad s_1 \geq 0 \quad \& \quad s_2 \geq 0.$$

I v tomto případě bude ve výchozí simplexové tabulce pro bod $x = [0,0,-3,3]$ v řádku $z_k - c_k$ alespoň jedno kladné číslo (konkrétně ve sloupci pro x_2). Proto tento bod není duálně ani primárně přípustný, takže doplníme umělé omezení

$$x_1 + x_2 \leq M \quad \text{neboli} \quad x_1 + x_2 + s_3 = M$$

pro pomocnou proměnnou $s_3 \geq 0$. Tuto proměnnou vezmeme jako součást výchozího bazického bodu, sestavíme novou výchozí simplexovou tabulku a v prvním kroku vyvedeme tuto proměnnou z báze. Celý výpočet je zobrazen v Tabulce 2.31.
V prvním kroku jsme vybrali klíčový sloupec odpovídající jediné kladné hodnotě v řádku $z_k - c_k$, což je druhý sloupec s hodnotou 2. Tento sloupec odpovídá proměnné x_2, která nahradí proměnnou s_3 v bázi. Eliminací s pivotem $\lambda_{3,2}$ podle pravidel simplexového algoritmu získáme druhou část této tabulky. Zde jsou v řádku $z_k - c_k$ již pouze nekladná čísla a pro nebazické proměnné dokonce záporná čísla, takže můžeme pokračovat duálním simplexovým algoritmem. Klíčový řádek odpovídá hodnotě $3 - M < 0$, která je nejmenší ze všech tří hodnot ve sloupci x_B. V tomto řádku jsou dvě

c_B	B	x_1	x_2	s_1	s_2	s_3	x_B
0	s_1	-2	-1	1	0	0	-3
0	s_2	-1	1	0	1	0	3
0	s_3	1	(1)	0	0	1	$M \rightarrow$
	c	2	-2	0	0	0	
	z	0	0	0	0	0	
	$z_k - c_k$	-2	$2\uparrow$	0	0	0	0
	s_1	-1	0	1	0	1	$-3+M$
	s_2	(-2)	0	0	1	-1	$3-M \rightarrow$
	x_2	1	1	0	0	1	M
	$z_k - c_k$	-4	0	0	0	-2	$-2M$
	$\frac{z_k - c_k}{\lambda}$	$2\uparrow$	✖	✖	✖	2	
	s_1	0	0	1	0	$-1/2$	$-9/2+3M/2$
	x_1	1	0	0	$-1/2$	$1/2$	$-3/2+M/2$
	x_2	0	1	0	$1/2$	$1/2$	$3/2+M/2$
	$z_k - c_k$	0	0	0	-2	0	-6

Tabulka 2.31: Duální simplexová metoda s využitím umělého omezení pro Příklad 2.10.5.

záporné hodnoty, takže dopočítáme podíly $(z_k - c_k)/\lambda$, z čehož vidíme, že klíčový sloupec odpovídá např. proměnné x_1 (tentokrát máme na výběr, a tak volíme sloupec více vlevo), která vstoupí do báze na úkor s_2. Po provedení eliminace s pivotem $\lambda_{2,1}$ dostaneme poslední část tabulky. Jelikož M volíme dostatečně velké, jsou ve sloupci x_B

již pouze kladné hodnoty, čímž se nám podařilo najít bod, který je primárně i duálně přípustný, a tudíž i řešením úlohy v kanonickém tvaru. Nyní ovšem souřadnice tohoto bodu závisí na M, zatímco hodnota účelové funkce je stále rovna -6, takže máme nekonečně mnoho řešení (však také s_3 je mimo bázi), která získáme jako $x(M) = [-3/2 + M/2, 3/2 + M/2, -9/2 + 3M/2, 0, 0]$ pro libovolné $M \geq 3$. To znamená, že řešením původní úlohy je libovolný bod $x^*(M) = [-3/2 + M/2, 3/2 + M/2]$ pro $M \geq 3$ a hodnota úlohy je $f^* = -6$. V tomto případě se jedná o polopřímku vycházející z bodu $[0,3]$, který je z těchto řešení jediným BPB.

K procvičení duální simplexové metody a její možné srovnání s primárním metodou lze využít kterýkoli z příkladů uvedený v předchozích podkapitolách[208].

V literatuře lze samozřejmě najít ještě i další varianty simplexového algoritmu. Za zmínku stojí zejména tzv. *revidovaná simplexová metoda*[209] a *primární-duální simplexové metoda*[210]. Revidovaná simplexová metoda se příliš neliší od původní metody – dokonce vyžaduje i stejný počet kroků. Jejím smyslem je především snížení množství prováděných výpočtů[211], takže jsou explicitně určovány pouze skutečně potřebné hodnoty. Algoritmy založené na duální úloze jsou ale jiné. Zatímco v klasické (primární) simplexové metodě využíváme bazické přípustné body a v duální metodě pracujeme s bazickými duálně přípustnými body (jimž odpovídají v původních proměnných body, které nejsou primárně přípustné), v primární–duální simplexové metodě si vystačíme pouze s duálně přípustnými body, které nemusí být bazické. Podrobnější popis obou těchto varianty simplexové metody ovšem již překračuje rámec našeho textu, a tak čtenáře odkážeme pouze např. na [33, Podkapitoly 5.1 a 6.5] či [61, Kapitoly 4 a 5].

[208] Nebo např.

$$3x_1 + x_2 + 2x_3 \to \min$$
$$x_1 + x_2 - 2x_3 \geq 1$$
$$x_1 - 2x_2 + x_3 \geq -1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0$$

či

$$x_5 \to \min$$
$$x_3 + 2x_4 + x_5 \geq 1$$
$$-2x_3 - 3x_4 + 4x_5 \geq 1$$
$$x_3 \geq 0 \quad \& \quad x_4 \geq 0 \quad \& \quad x_5 \geq 0.$$

[209] Angl. *revised simplex method.*

[210] Angl. *primal–dual simplex method.*

[211] A tudíž i snížení výpočetní náročnosti. Pochopitelně pro naše „malé“ úlohy počítané ručně nemá tato modifikace téměř žádný smysl, neboť její možný přínos by mohl zůstat ukryt za několika dalšími potřebnými výpočty na začátku algoritmu. Ovšem v úlohách s velkým počtem proměnných a/nebo omezení řešených touto metodou pomocí počítače už může dojít k výraznému zefektivnění celého výpočtu.

2.11 Dopravní problém

V této podkapitole se podíváme na jeden ze speciálních případů úlohy lineárního programování, kterým je tzv. *dopravní problém* s nímž se můžeme setkat téměř ve všech učebnicích věnovaných lineárnímu programování, např. v [33, 154, 330, 414, 473]. My už jsme se o tomto problému zmínili v Podkapitole 2.3 a v Příkladech 2.3.6–2.3.7. Nicméně než pustíme do jeho podrobnější analýzy a možností jeho řešení, pojďme si ukázat ještě jeden motivační příklad.

Příklad 2.11.1

Představme si, že máme dvě pole, ze kterých dovážíme obilí třem odběratelům, kteří ho dále prodávají. Kvalitu obilí z jednotlivých polí nejsme schopni rozlišit. Na jednom poli vypěstujeme 25 tun a na druhém 15 tun obilí. Odběratelé jsou s námi dohodnutí na dodávce 9, 18 a 13 tun. Náklady na dopravu jedné tuny obilí z obou polí ke všem odběratelům jsou zobrazeny v Tabulce 2.32.

odběratel / pole	O_1	O_2	O_3	nabídka
P_1	5	4	9	25
P_2	8	10	6	15
poptávka	9	18	13	40

Tabulka 2.32: Zadání Příkladu 2.11.1.

V tabulce vidíme též nabídku obilí ze dvou polí a poptávku jednotlivých odběratelů. Naším cílem je najít plán přepravy, který minimalizuje celkové dopravní náklady. For-

mulujme tuto úlohu jako úlohu lineárního programování a vyřešme ji simplexovou metodou.

Řešení. Označme x_{ij} množství obilí, které dopravíme z i-tého pole j-tému odběrateli. Dále označme n_1, n_2 nabídky obilí z našich polí a p_1, p_2, p_3 poptávky jednotlivých odběratelů. Účelovou funkci, jejíž minimum hledáme, můžeme zapsat jako

$$5x_{11} + 4x_{12} + 9x_{13} + 8x_{21} + 10x_{22} + 6x_{23}.$$

V posledním řádku a posledním sloupci Tabulky 2.32 vidíme, že naší nabídkou jsme schopni přesně pokrýt poptávky všech odběratelů, tj. platí rovnost

$$\sum_{i=1}^{2} p_i = \sum_{j=1}^{3} n_j.$$

Dopravní problém s touto vlastností označujeme jako tzv. *vyrovnaný*. V opačném případě bychom o problému hovořili jako o tzv. *nevyrovnaném*. V nevyrovnaném problému dochází k převisu na straně poptávky, případně nabídky. V našem příkladu je součet poptávaného i nabízeného množství roven 40.

Maximální množství, které můžeme dovézt z prvního pole, je 25. Vzhledem k tomu, že jde o vyrovnaný problém, musí být množství, které z prvního pole odvezeme, rovno právě 25. Z této podmínky dostáváme omezení ve tvaru

$$x_{11} + x_{12} + x_{13} = 25.$$

Pro druhé pole obdobně získáme podmínku

$$x_{21} + x_{22} + x_{23} = 15.$$

Podobná omezení platí i pro stranu poptávky. Odběratelé chtějí získat dohodnutá množství, můžeme tedy psát

$$x_{11} + x_{21} = 9 \quad \& \quad x_{12} + x_{22} = 18 \quad \& \quad x_{13} + x_{23} = 13.$$

Poslední podmínky v této úloze jsou podmínky nezápornosti všech proměnných, tj. $x_{ij} \geq 0$ pro všechna $i \in \{1,2\}$ a $j \in \{1,2,3\}$.

Nyní můžeme přejít k samotnému řešení úlohy, která je již v kanonickém tvaru. Matice A však neobsahuje žádný sloupec jednotkové matice, musíme tak využít jednu z metod představených v Podkapitole 2.9. Výchozí BPB najdeme např. s využitím dvoufázové metody. Přidáním pětice umělých proměnných $w_1, \ldots, w_5$, bychom pro první fázi získali následující výchozí simplexovou tabulku.

c_B	B	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	w_1	w_2	w_3	w_4	w_5	x_B	$\frac{x_B}{\lambda}$
1	w_1	1	1	1	0	0	0	1	0	0	0	0	25	
1	w_2	0	0	0	1	1	1	0	1	0	0	0	15	
1	w_3	1	0	0	1	0	0	0	0	1	0	0	9	
1	w_4	0	1	0	0	1	0	0	0	0	1	0	15	
1	w_5	0	0	1	0	0	1	0	0	0	0	1	13	
	c	0	0	0	0	0	0	1	1	1	1	1		
	z_k	2	2	2	2	2	2	1	1	1	1	1		
	$z_k - c_k$	2	2	2	2	2	2	0	0	0	0	0	80	

Tabulka 2.33: První fáze dvoufázové metody pro Příklad 2.11.1.

Již na první pohled vidíme, že tabulka má „speciální“ strukturu – všechny koeficienty v omezeních jsou buď jedničky nebo nuly a i jejich uspořádání je velmi specifické. Právě díky této speciální struktuře (a také kvůli častému využití tohoto typu úloh v praxi) byly vytvořeny další algoritmy, které jsou při řešení dopravních úloh efektivnější než simplexová metoda. Samozřejmě bychom mohli úlohu vyřešit dvoufázovou metodou, ale to již ponecháme na čtenáři. Po šesti krocích první fáze bychom se měli dostat do situace, ve které je v bázi už jen umělá proměnná w_5 s nulovou hodnotou a nelze ji vyloučit, což je podobné jako při řešení Příkladu 2.9.5. Ve druhé fázi bychom pak získali řešení

$$x^* = [x^*_{11}, x^*_{12}, x^*_{13}, x^*_{21}, x^*_{22}, x^*_{23}] = [7, 18, 0, 2, 0, 13].$$

V předchozím příkladu jsme viděli další jednoduchou ukázku dopravního problém. V obecném případě jde o úlohu

$$\left.\begin{aligned} \sum_{i=1}^{m}\sum_{j=1}^{n} n_{ij}\,x_{ij} &\to \min \\ \sum_{j=1}^{n} x_{ij} = s_i > 0 \quad &\text{pro všechna } i \in \{1,\dots,m\} \\ \sum_{i=1}^{m} x_{ij} = c_j > 0 \quad &\text{pro všechna } j \in \{1,\dots,n\} \\ x_{ij} \geq 0 \quad &\text{pro všechna } i \in \{1,\dots,m\} \text{ a } j \in \{1,\dots,n\}, \end{aligned}\right\} \tag{2.11.1}$$

kde indexem i značíme jednotlivé dodavatele (sklady) a jejich kapacity jako s_i, zatímco indexem j značíme odběratele (cíle) a jejich poptávky jako c_j. Proměnné x_{ij} pak určují množství přepravené od i-tého dodavatele k j-tému odběrateli a n_{ij} vyjadřuje náklady na přepravu jednotkového množství zboží tímto směrem. Je nutné zde ještě zdůraznit, že v dopravní úloze předpokládáme, že zboží je tzv. *homogenní*, tj. je naprosto stejné kvality a nikdo není schopen rozlišit, od jakého dodavatele zboží pochází. Také předpokládáme, že náklady jsou lineární funkcí množství zboží, což je sice jisté zjednodušení, které však z praktické pohledu může být zanedbatelné, a především výrazně usnadňuje samotné řešení takového problému.

Problém (2.11.1) má $m \times n$ proměnných a $m+n$ omezení. V případě vyrovnaného problému však odpovídající matice A nemá plnou hodnost, neboť jedno omezení je lineárně závislé: sečteme-li totiž všech m nabídkových omezení, získáme totéž, jako když sečteme všech n poptávkových omezení[212]. Problém, ve kterém by došlo k převisu nabídky, by měl omezení ve tvaru nerovností – místo znaménka „=“ by v první skupině omezení bylo znaménko „≤“. Podobně problém, ve kterém by došlo k převisu na straně poptávky, by měl znaménko „≤“ ve druhé skupině omezení. V následující části textu se budeme věnovat pouze vyrovnaným dopravním problémům, neboť nevyrovnané problémy můžeme snadno převést na vyrovnané přidáním tzv. *fiktivního odběratele* či *dodavatele*, čímž získáme ve všech omezeních rovnosti. Postup při řešení nevyrovnaného problému si ukážeme v Příkladu 2.11.8.

Výhodou dopravního problému je, že již z jeho zadání poznáme, zda má či nemá řešení, a to díky platnosti následující ekvivalence, která ukazuje důležitost vyrovnanosti dopravního problému.

Věta 2.11.2

Pro dopravní úlohu (2.11.1) jsou následující podmínky ekvivalentní:

(i) úloha má BPB;

(ii) úloha má řešení;

(iii) jedná se o vyrovnaný dopravní problém, tj. platí rovnost

$$\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} c_j.$$

[212] Jinými slovy, báze libovolného BPB obsahuje vždy nejvýše $n+m-1$ proměnných. Navíc za lineárně závislé můžeme považovat kterékoli za daných omezení. To plyne z předchozího stručného popisu, který si můžeme rozepsat i se všemi podrobnostmi: sečtením první skupin m omezení a druhé skupin n omezení obdržíme

$$\sum_{i=1}^{n} s_i = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n} c_i.$$

Pochopitelně sečtením pouze $n-1$ omezení ve druhé skupině získáme

$$\sum_{j=1}^{n-1}\sum_{i=1}^{m} x_{ij} = \sum_{j=1}^{n-1} c_j.$$

Odečtení předchozích dvou rovností pak vede k

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n-1}\sum_{i=1}^{m} x_{ij} = \sum_{j=1}^{n} c_j - \sum_{j=1}^{n-1} c_j$$

neboli

$$\sum_{i=1}^{m}\Big(\sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n-1} x_{ij}\Big)\sum_{i=1}^{m} x_{in} = c_n,$$

což je zbývající n-té omezení. Pomocí $n+m-1$ prvních omezení jsme získali i to poslední, což potvrzuje jeho lineární závislost i tvrzení o velikosti báze. Ještě dodejme, že tímtéž postupem můžeme získat i jakékoli jiné omezení s pomocí $m+n-1$ omezení.

Důkaz. „(i)⟹(ii)“ Přípustná množina je neprázdná a ohraničená, neboť zjevně musí být $x_{ij} \le \min\{s_i, c_j\}$, což implikuje ohraničenost účelové funkce na X, a tudíž v souladu s Větou 2.4.6 i řešitelnost celého problému. Opačná implikace je zřejmá stejně jako „(i)⟹(iii)“. Konečně implikace „(iii)⟹(i)“ plyne z toho, že např. bod x se složkami

$$x_{ij} := \frac{s_i\, c_j}{\sum_{i=1}^{m} s_i} \quad \text{pro všechna } i \in \{1,\dots,m\} \text{ a } j \in \{1,\dots,n\}$$

je BPB dané úlohy. ■

O řešitelnosti úlohy (2.11.1) již víme „vše“ a už jen zbývá toto řešení najít. My se nejprve seznámíme se třemi metodami, pomocí kterých můžeme najít výchozí BPB[213] této úlohy. Samotné řešení dopravního problému i hledání výchozího BPB je opět pro přehlednost organizováno v tabulkách, které jsou ale tentokrát podobné Tabulce 2.34 níže. Sloupce této tabulky znázorňují jednotlivé odběratele (označené písmenem „O“) a řádky jednotlivé prodejce (označené písmenem „P“). Podobně jako v případě simplexové metody se při řešení dopravních úloh běžně vyplňují pouze nenulové buňky tabulky, zatímco buňky odpovídající hodnotám $x_{ij} = 0$ jsou při výpočtech ponechány prázdné. Tyto prázdné buňky tabulky tak jistě odpovídají nebazickým proměnným. S využitím podobných tabulek si nyní ukážeme, jak najít výchozí BPB úlohy.

Postupně si představíme tři nejběžnější metody pro hledání výchozího bazického přípustného bodu, a to *metodu severozápadního rohu*, *metodu minimální cesty* a *Vogelovu aproximační metodu*. Metody si vždy stručně popíšeme a ilustrujeme na jednom příkladu. Podrobnější teoretický popis a další metody lze najít např. v [125, Chapter 8], [126, Chapter 7] a [33, Chapter 10].

[213] Dá se ukázat, že pokud N značí nejmenší počet BPB dopravního problému s m-ticí dodavatelů a n-ticí odběratelů s $m \le n$, pak

$$n^{m-1} \le N \le n^{m-1} m^{n-1}$$

pro nedegenerovanou úlohu. V případě degenerované úlohy (viz později) máme odhad

$$\frac{n!}{(n-m+1)!} \le N \le n^{m-1} m^{n-1},$$

viz [203, str. 248].

	O_1	O_2	O_3	nabídka
P_1	n_{11} x_{11}	n_{12} x_{12}	n_{13} x_{13}	s_1
P_2	n_{21} x_{21}	n_{22} x_{22}	n_{23} x_{23}	s_2
P_3	n_{31} x_{31}	n_{32} x_{32}	n_{33} x_{33}	s_3
poptávka	c_1	c_2	c_3	$c_1 + c_2 + c_3$

Tabulka 2.34: Obecná tabulka pro dopravní úlohu v případě $m = n = 3$.

Metoda severozápadního rohu[214]. Tato metoda představuje velmi jednoduchý a intuitivní způsob, jakým můžeme najít výchozí BPB. Jak naznačuje název, v této metodě začneme v severozápadním rohu tabulky, neboli v buňce vlevo nahoře. Do této buňky napíšeme minimum z hodnot nabídky v příslušném řádku a poptávky v příslušném sloupci. V Tabulce 2.34 by se jednalo o minimum z čísel s_1 a c_1. Tímto krokem jsme vyčerpali buď nabídku nebo poptávku, druhou z těchto hodnot jsme změnili. Tyto změny musíme mít při dalším postupu metody na paměti. Přiřazením čísla $k = min\{s_1, c_1\}$ jsme poptávku změnili na $c_1 - k$ a nabídku na $s_1 - k$, jedno z těchto čísel je jistě nulové. Tento krok můžeme chápat tak, že dodavatel prodal k kusů zboží příslušnému odběrateli. Poté pokračujeme s vyplňováním jedné ze sousedních buněk, a to buňky ve stejném řádku, pokud se minimum nacházelo v daném sloupci, v opačném případě buňky ve stejném sloupci (pouze o řádek níže). Do vybrané buňky opět vepíšeme hodnotu minima příslušné upravené poptávky a nabídky. Algoritmicky můžeme tuto metodu popsat následujícím způsobem.

[214] Angl. *Northwest corner rule.*

Metoda severozápadního rohu

(i) Do buňky v severozápadním rohu (tj. buňky s nejmenšími možnými indexy) přiřadíme největší možné číslo tak, aby bylo splněno příslušné nabídkové i poptávkové omezení, tj. vybereme minimum z těchto dvou čísel. Přepočítáme poptávkové a nabídkové omezení, která jsme tímto krokem změnili.

(ii) Do nejbližší dostupné sousední buňky přiřadíme největší možné číslo tak, že opět dodržujeme zadaná omezení. Řazení blízkosti sousedních buněk je následující: buňka vedle, buňka pod, buňka diagonálně vpravo dole. Omezení opět přepočítáme.

(iii) Krok číslo (ii) opakujeme, dokud nedojdeme k buňce v jihovýchodním rohu (tj. buňce s největšími možnými indexy). Pro vyrovnanou dopravní úlohu v tomto kroku vždy dojde ke splnění všech poptávkových a nabídkových omezení úlohy.

Tuto metodu si nyní ukážeme na konkrétním příkladu.

Příklad 2.11.3

Metodou severozápadního rohu najděme výchozí BPB pro dopravní úlohu danou Tabulkou 2.35, což v souladu s (2.11.1) odpovídá

$$6x_{11}+8x_{12}+10x_{13}+7x_{21}+11x_{22}+11x_{23}+4x_{31}+5x_{32}+12x_{33}\to\min$$
$$x_{11}+x_{12}+x_{13}=15 \quad \& \quad x_{21}+x_{22}+x_{23}=17 \quad \& \quad x_{31}+x_{32}+x_{33}=18$$
$$x_{11}+x_{21}+x_{31}=20 \quad \& \quad x_{12}+x_{22}+x_{32}=10 \quad \& \quad x_{13}+x_{23}+x_{33}=20$$
$$x_{ij}\geq 0 \quad \text{pro všechna } i,j\in\{1,2,3\}.$$

prodejce \ odběratel	O_1	O_2	O_3	nabídka
P_1	6	8	10	15
P_2	7	11	11	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.35: Zadání Příkladů 2.11.3, 2.11.4 a 2.11.5.

Řešení. K zápisu řešení využijeme tabulku podobnou Tabulce 2.34, která je pro tento příklad zobrazena v Tabulce 2.36 níže a ve které jsou všechny buňky pro x_{ij} prázdné.

	O_1	O_2	O_3	nabídka
P_1	6	8	10	15
P_2	7	11	11	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.36: Zadání Příkladu 2.11.3.

Metodu zahájíme tím, že do políčka, které je v severozápadním rohu (tj. x_{11}), napíšeme minimum z nabídky a poptávky. V našem příkladu jde o číslo 15, neboť odběratel O_1 poptává 20, prodejce P_1 však nabízí pouze 15. První krok metody je zobrazen v Tabulce 2.37. Vyčerpání nabídky prodejce P_1 vyznačíme v dalším kroku hvězdičkou u P_1. Dále pokračujeme k jedné ze sousedních buněk, kterými jsou buňky pro x_{12}, x_{21} či x_{22}. V našem příkladu již prodejce P_1 prodal všechno své zboží odběrateli O_1, v prvním řádku tak již nemáme žádné dostupné nabízené zboží. V celém řádku bychom mohli doplnit nuly. Odběratel O_1 již má 15 ks zboží, jeho nová poptávka je tak pouze 5.

	O_1	O_2	O_3	nabídka
P_1	6 **15**	8	10	15
P_2	7	11	11	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.37: První krok Příkladu 2.11.3.

V dalším kroku musíme přejít ke druhému řádku, a to do buňky pro proměnnou x_{21}. Do něj napíšeme minimum z čísel 17 (nabídka) a 5 (nová poptávka odběratele 1), pro x_{21} tak platí $x_{21} = 5$. Poptávka odběratele O_1 je po tomto kroku naplněna, což v další tabulce vyznačíme hvězdičkou u O_1. Ve zbylých buňkách prvního sloupce si můžeme představit nuly. Tento krok je zobrazen v Tabulce 2.38.

	O_1	O_2	O_3	nabídka
P_1^*	6 **15**	8	10	15
P_2	7 **5**	11	11	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.38: Druhý krok Příkladu 2.11.3.

V dalších krocích postupujeme obdobně. Vzhledem k tomu, že dodavatel P_2 zatím prodal pouze 5 ks svého zboží, je jeho nová nabídka rovna 12. Odběratel O_2 poptává 10 ks. Do buňky pro proměnnou x_{22} napíšeme 10, neboť jde o hledané minimum z čísel 10 a 12. Tímto krokem, zobrazeným v Tabulce 2.39, jsme vyčerpali poptávku odběratele O_2, nabídka prodejce 2 se snížila na 2 ks.

	O_1^*	O_2	O_3	nabídka
P_1^*	6 **15**	8	10	15
P_2	7 **5**	11 **10**	11	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.39: Třetí krok Příkladu 2.11.3.

V dalším kroku (Tabulka 2.40) se přesuneme k políčku x_{23}, do kterého napíšeme číslo 2, jež je minimem z nabídky 2 a poptávky 20. Tím vyčerpáme nabídku prodejce P_2.

	O_1^*	O_2^*	O_3	nabídka
P_1^*	6 | 15	8	10	15
P_2	7 | 5	11 | 10	11 | 2	17
P_3	4	5	12	18
poptávka	20	10	20	50

Tabulka 2.40: Čtvrtý krok Příkladu 2.11.3.

V posledním kroku (Tabulka 2.41) jsou již splněny poptávky odběratelů 1 a 2 a vyčerpány nabídky prodejců P_1 a P_2. Do políčka x_{33} tak vepíšeme minimum z čísel 18 (nabídka P_3) a 18 (upravená poptávka O_3). Vidíme, že nyní jsou všechna omezení pro nabídky i poptávky splněna, našli jsme tak výchozí bazický přípustný bod.

	O_1	O_2	O_3	nabídka
P_1	6 | 15	8	10	15
P_2	7 | 5	11 | 10	11 | 2	17
P_3	4	5	12 | 18	18
poptávka	20	10	20	50

Tabulka 2.41: Řešení Příkladu 2.11.3.

Ověření splnění všech rovností uvedených v matematické formulaci příkladu ponecháme na čtenáři. Nalezený bod je zobrazen v Tabulce 2.41. Hodnota účelové funkce v tomto bodě je rovna součtu nákladů na jednotlivé převozy, celkové náklady nalezené dopravní strategie tedy jsou

$$15 \cdot 6 + 5 \cdot 7 + 10 \cdot 11 + 2 \cdot 11 + 18 \cdot 12 = 473.$$

[215] Angl. *The least-cost method.*

Metoda minimální ceny[215]. Tato metoda, jak již název napovídá, postupuje po buňkách s minimální cenou. Metoda je opět velmi intuitivní a algoritmicky ji můžeme popsat následujícím způsobem.

Metoda minimální ceny

(i) Najdeme buňku, ve které jsou nejmenší dopravní náklady n_{ij} a přiřadíme do ní největší možnou hodnotu tak, aby bylo splněno jak nabídkové, tak poptávkové omezení. V případě rovnosti n_{ij} u více buněk si vybereme náhodně.

(ii) Tento postup opakujeme, dokud čísla v tabulce nesplňují všechna omezení.

Při aplikování metody minimální ceny se již při hledání výchozího bazického přípustného bodu snažíme najít bod, ve kterém je hodnota účelové funkce blízká optimální hodnotě. Ve většině případů získáme využitím této metody BPB s menší hodnotou účelové funkce než s využitím metody severozápadního rohu, která koeficienty účelové funkce nebere v potaz. Při následném řešení dopravní úlohy s výchozím BPB, který je blíž optimu, pak lze očekávat, že budeme potřebovat méně iterací. Tuto metodu si nyní ukážeme na stejném příkladu, jako metodu severozápadního rohu.

Příklad 2.11.4

Metodou minimální ceny najděme výchozí BPB pro dopravní úlohu zadanou Tabulkou 2.35.

Řešení. Stejně jako v Příkladu 2.11.3 začneme s prázdnou tabulkou (tj. Tabulkou 2.36), do které postupně doplníme hodnoty x_{ij}. Metodu zahájíme nalezením nejmenších dopravních nákladů. Z Tabulky 2.42 vidíme, že nejmenší cena (4) se nachází v buňce x_{31}. Maximální nabídka ve třetím řádku je 18, poptávka v prvním sloupci je 20. Do buňky x_{31} napíšeme minimum z těchto hodnot, tj. 18. První krok metody je zobrazen v Tabulce 2.42.

	O_1	O_2	O_3	nabídka
P_1	6	8	10	15
P_2	7	11	11	17
P_3	4 18	5	12	18
poptávka	20	10	20	50

Tabulka 2.42: První krok Příkladu 2.11.4.

Tímto krokem jsme vyčerpali veškerou nabídku prodejce P_3 a ve třetím řádku jsou tak ostatní buňky prázdné, což ve druhém kroku označíme hvězdičkou u P_3. Jak vidíme z Tabulky 2.43, nejmenší cena mimo třetí řádek je v buňce x_{11}. Vzhledem ke zbylé poptávce odběratele O_1, která je nyní pouze 2, zde doplníme číslo 2. Poptávka odběratele O_1 je po tomto kroku metody vyčerpána, ve zbývajících buňkách v prvním sloupci již nemohou být nenulové hodnoty.

	O_1	O_2	O_3	nabídka
P_1	6 2	8	10	15
P_2	7	11	11	17
P_3^*	4 18	5	12	18
poptávka	20	10	20	50

Tabulka 2.43: Druhý krok Příkladu 2.11.4.

V dalším kroku se nejnižší „nevyčerpaná“ cena nachází v buňce x_{12}, což vidíme z Tabulky 2.44. Prodejce P_1 má nyní k dipozici 13, odběratel O_2 poptává 10, do buňky tedy vepíšeme číslo 10. Tímto krokem jsme uspokojili poptávku odběratele O_2.

	O_1^*	O_2	O_3	nabídka
P_1	6: 2	8: 10	10	15
P_2	7	11	11	17
P_3^*	4: 18	5	12	18
poptávka	20	10	20	50

Tabulka 2.44: Třetí krok Příkladu 2.11.4.

Jediné dvě buňky, které nyní můžeme doplnit, jsou x_{13} a x_{23}. Nejnižší cenu nyní obsahuje buňka x_{13}, doplníme do ní tak minimum z čísel 3 (zbývající nabídka P_1) a 20 (poptávka O_3). Tento krok je znázorněn v Tabulce 2.45.

	O_1^*	O_2^*	O_3	nabídka
P_1	6: 2	8: 10	10: 3	15
P_2	7	11	11	17
P_3^*	4: 18	5	12	18
poptávka	20	10	20	50

Tabulka 2.45: Čtvrtý krok Příkladu 2.11.4.

V posledním kroku vepíšeme do buňky x_{23} číslo 17, které zaručí splnění omezení prodejce P_2 a odběratele O_3. Výsledná tabulka s výchozím bazickým přípustným bodem nalezeným metodou minimální ceny je zobrazena v Tabulce 2.46. Čísla v tabulce splňují všechna omezení, jedná se tedy opravdu o výchozí bazický přípustný bod. Účelová funkce v něm nabývá hodnotu 381, což je výrazné zlepšení oproti bodu, který jsme našli metodou severozápadního rohu.

	O_1	O_2	O_3	nabídka
P_1	6: 2	8: 10	10: 3	15
P_2	7	11	11: 17	17
P_3	4: 18	5	12	18
poptávka	20	10	20	50

Tabulka 2.46: Řešení Příkladu 2.11.4.

[216] Publikováno v [428].

Vogelova[216] **aproximační metoda (VAM)**. Stejně jako metoda minimální ceny se i Vogelova aproximační metoda snaží najít co nejlepší výchozí bazický přípustný bod. V metodě minimální ceny můžeme dojít do situace, kdy musíme obsadit i buňky s velmi vysokými cenami. Tomu se snaží Vogelova aproximační metoda vyhnout. Metodu bychom mohli označit i za tzv. *penalizační metodu*, neboť její postup spočívá v přidělení *penalizace* každému dodavateli a odběrateli. Tyto penalizace vypočítáme jako absolutní hodnotu rozdílu dvou nejmenších čísel v každém sloupci a řádku. Penalizace nám tak říká, o kolik jsou náklady na zvolení druhé nejlepší cesty větší než náklady na zvolení optimální cesty. Algoritmický popis Vogelovy aproximační metody je následující.

Vogelova aproximační metoda

(i) Vypočítáme penalizace pro všechny dostupné sloupce a řádky v tabulce, ve kterých jsou alespoň dvě dostupné buňky. Penalizace je absolutní hodnota rozdílu dvou nejmenších čísel v každém řádku a sloupci.

(ii) Vybereme řádek nebo sloupec s největší penalizací.

(iii) Do buňky s nejmenší cenou ve vybraném řádku nebo sloupci přiřadíme největší možné číslo splňující poptávkové i nabídkové omezení.

(iv) Opakujeme kroky (i)–(iii) dokud nenajdeme bod, který splní všechna zadaná omezení.

Tato metoda je výpočetně mírně náročnější než předchozí metody, obvykle však dává nejlepší výsledky. V příkladu zadaném Tabulkou 2.35 k tomuto výsledku ovšem nedojde. V některých případech velmi malých úloh se může i stát, že pomocí Vogelovy aproximační metody získáme přímo řešení úlohy. I tuto metodu si nyní ukážeme na příkladu.

Příklad 2.11.5

Vogelovou aproximační metodou najděme výchozí BPB pro dopravní úlohu zadanou Tabulkou 2.35.

Řešení. Stejně jako u předchozích metod začneme i zde s prázdnou tabulkou (tj. Tabulkou 2.36). V prvním kroku metody si vypočítáme všechny penalizace. Postup jejich výpočtu je naznačen v Tabulce 2.47.

	O_1	O_2	O_3	nabídka	penalizace
P_1	6	8	10	15	$\|8-6\|=2$
P_2	7 (17)	11	11	17	$\|11-7\|=4$
P_3	4	5	12	18	$\|5-4\|=1$
poptávka	20	10	20	50	
penalizace	$\|6-4\|=2$	$\|8-5\|=3$	$\|11-10\|=1$		

Tabulka 2.47: První krok Vogelovy metody pro Příklad 2.11.5.

Z Tabulky 2.47 vidíme, že největší penalizace se nachází ve druhém řádku. Do buňky s nejmenší cenou v tomto řádku tak doplníme minimum z čísel 17 a 20, tj. 17. Řádky a sloupce, do kterých již nemůžeme z důvodu vyčerpání nabídky či poptávky doplnit kladná čísla, opět označíme hvězdičkou. Nyní musíme přepočítat penalizace ve všech řádcích a sloupcích, ve kterých není vyčerpána nabídka, případně poptávka. Druhý krok Vogelovy aproximační metody je zobrazen v Tabulce 2.48. Hodnoty penalizací

jsou stejné jako v prvním kroku s výjimkou druhého řádku, pro který již penalizace nepočítáme, neboť je v něm vyčerpaná nabídka. Kvůli tomu se změní i penalizace ve třetím sloupci, neboť buňka s náklady 11 již není dostupná. Nejvyšší penalizací je nyní číslo 3 ve druhém sloupci, který vybereme. Do buňky s nejmenšími náklady v tomto sloupci, tj. x_{32} napíšeme číslo 10, což je minimum z čísel 10 a 18.

	O_1	O_2	O_3	nabídka	penalizace
P_1	6	8	10	15	$\lvert 8-6\rvert = 2$
P_2^*	7 17	11	11	17	
P_3	4	5 10	12	18	$\lvert 5-4\rvert = 1$
poptávka	20	10	20	50	
penalizace	$\lvert 6-4\rvert = 2$	$\lvert 8-5\rvert = 3$	$\lvert 12-10\rvert = 2$		

Tabulka 2.48: Druhý krok Vogelovy metody pro Příklad 2.11.5.

V dalších krocích postupujeme obdobně. Po přepočítání penalizací nyní dostaneme Tabulku 2.49, ve které je největší penalizace ve třetím řádku. Do buňky x_{31} doplníme číslo 3, které je minimem zbývající poptávky odběratele O_1 a zbývající nabídky prodejce P_3.

	O_1	O_2^*	O_3	nabídka	penalizace
P_1	6	8	10	15	$\lvert 10-6\rvert = 4$
P_2^*	7 17	11	11	17	
P_3	4 3	5 10	12	18	$\lvert 12-4\rvert = 8$
poptávka	20	10	20	50	
penalizace	$\lvert 6-4\rvert = 2$		$\lvert 12-10\rvert = 2$		

Tabulka 2.49: Třetí krok Vogelovy metody pro Příklad 2.11.5.

Po přepočtení penalizací se dostaneme k Tabulce 2.50. V této tabulce je již pouze jedna penalizace, neboť poptávky odběratelů O_1 a O_2 jsou již naplněny, stejně tak jako nabídka prodejce P_2. V posledním sloupci tak vybereme buňku s nejnižšími náklady, kterou je x_{13}. Do této buňky vepíšeme číslo 15, což je minimum z čísel 15 a 20.

	O_1^*	O_2^*	O_3	nabídka	penalizace
P_1	6	8	10 15	15	
P_2^*	7 17	11	11	17	
P_3	4 3	5 10	12	18	
poptávka	20	10	20	50	
penalizace			$\lvert 12-10\rvert = 2$		

Tabulka 2.50: Čtvrtý krok Vogelovy metody pro Příklad 2.11.5.

V posledním kroku se dostaneme do situace, ve které není možné spočítat žádnou penalizaci. Nyní máme pouze jednu dostupnou buňku x_{33}, kterou můžeme doplnit. Do této buňky vepíšeme číslo 5, čímž zajistíme splnění všech omezení. Výsledek, který jsme získali pomocí Vogelovy aproximační metody, je zobrazen v Tabulce 2.51. Hodnota účelové funkce v nalezeném výchozím BPB je rovna 391, což je více než pro metodu minimální ceny, ale méně než pro metodu severozápadního rohu.

	O_1	O_2	O_3	nabídka
P_1	6	8	10 15	15
P_2	7 17	11	11	17
P_3	4 3	5 10	12 5	18
poptávka	20	10	20	50

Tabulka 2.51: Řešení Příkladu 2.11.5.

Modifikovaná distribuční metoda (MODI). Nyní již víme, jak najít výchozí bazický přípustný bod. Při postupu řešení bychom tak mohli využít klasickou simplexovou metodu. Avšak vzhledem ke speciální struktuře úlohy existují i efektivnější metody. Jednou z těchto metod je tzv. *modifikovaná distribuční metoda*, kterou si nyní ukážeme. Tuto metodu lze v literatuře najít též pod názvem *dopravní simplexová metoda* (viz [125, Kapitola 8]) či *metoda řádkových a sloupcových čísel* nebo *metoda potenciálů* (viz [414]). Základem této metody jsou následující dva kroky. V prvním kroku ověříme, zda je nalezený BPB optimální či nikoli, neboli zda je nalezené *schéma přepravy* to nejlevnější možné. V případě, že BPB není optimální, pak ve druhém kroku najdeme nový bazický přípustný bod, ve kterém jsou celkové náklady nižší. Tento postup si nyní popíšeme podrobněji a následně si jej ukážeme i na příkladu.

Duální úloha pro problém (2.11.1) má podobu

$$\sum_{i=1}^{m} s_i\, u_i + \sum_{j=1}^{n} c_j\, v_j \to \max$$

$$u_i + v_j \le n_{ij} \quad \text{pro všechna } i \in \{1,\dots,m\} \text{ a } j \in \{1,\dots,n\}.$$

Jestliže máme nedegenerovaný BPB a jestliže pro nějaké x_{ij} v úloze (2.11.1) platí $x_{ij} > 0$ (tj. daná buňka je v bázi), duální omezení $u_i + v_j \le n_{ij}$ se nutně realizuje jako rovnost $u_i + v_j = n_{ij}$, což plyne z *podmínky komplementarity* ve tvaru $(n_{ij} - u_i - v_j)x_{ij} = 0$, viz Podkapitolu 2.6. Celkově máme soustavu $n+m-1$ rovnic (neboť jedno omezení primární úlohy jsme kvůli lineární závislosti odebrali) pro $n+m$ neznámých. Její řešení tedy není jednoznačné. Postupujeme tak, že jednu z duálních proměnných položíme rovnu nule, tj. např. $u_1 = 0$, a poté dopočítáme hodnoty ostatních duálních proměnných. Pro buňky, ve kterých je $x_{ij} = 0$ (tj. buňky nejsou v bázi), poté vypočítáme tzv. *náklady příležitosti* či *redukované cenové koeficienty* jako rozdíl

$$r_{ij} = n_{ij} - u_i - v_j.$$

Tyto koeficienty nám říkají, jaká je cena za přeřazení jednotkového množství do dané buňky. Jejich význam je podobný jako v případě prvků $z_k - c_k$ v simplexové tabulce, liší se akorát svým znaménkem. Zatímco v simplexové metodě do báze mohly vstoupit proměnné s nezáporným

číslem v řádku $z_k - c_k$, v případě MODI mohou do báze vstoupit proměnné s nekladnou hodnotou r_{ij}. Pokud tedy existují indexy i a j, pro které je $r_{ij} \leq 0$, můžeme jimi určenou proměnnou zavést do báze a snížit (v případě $r_{ij} = 0$ nezvýšit) tak hodnotu účelové funkce původní úlohy. V opačném případě jsme našli řešení dopravní úlohy, neboť jsou splněna všechna primární omezení (tj. nabídková a poptávková) i duální omezení (všechna r_{ij} jsou kladná).

V případě, že do báze může vstoupit více proměnných, můžeme využít pravidlo podobné Dantzigovu pravidlu maximálního relativního ocenění. V dopravním problému však vybereme tu proměnnou, pro kterou je r_{ij} minimální. Uvažujme nyní případ, ve kterém chceme do báze zavést proměnnou x_{kl}, neboť platí $r_{kl} \leq 0$. Do této buňky chceme přidělit co největší číslo, které označíme θ. Ale jak tuto hodnotu získat? Podobně jako v klasické simplexové metodě musíme v dalším kroku MODI určit proměnnou, která bázi opustí. V případě, že jsme do buňky pro x_{kl} přiřadili množství θ, musíme množství v ostatních buňkách upravit tak, aby stále platila nabídková i poptávková omezení. V řádku k tak musíme od některé z bazických buněk odečíst θ. Tímto ovšem narušíme některá sloupcová poptávková omezení, musíme proto také ostatní buňky upravit tak, aby byla zachována přípustnost daného bodu. V tabulce tedy musíme najít uzavřenou cestu, která začíná i končí v buňce x_{kl}. To uděláme tak, že v každém řádku a sloupci, jež jsou součástí této cesty, jednou přičteme a jednou odečteme θ. Tato cesta prochází pouze přes x_{kl} a některé bazické buňky a má vždy sudý počet vrcholů. Výsledná cesta navíc může sama sebe protínat a pro nedegenerovaný bod je určena jednoznačně. Jelikož musíme dodržet podmínky nezápornosti x_{ij}, získáme hodnotu θ jako minimum z prvků x_{ij}, od kterých θ odčítáme. Tímto vynulujeme jednu z původních bazických proměnných, neboli určíme proměnnou, která v daném kroku bázi opustí.

Tímto postupem získáme nový BPB. Pro tento bod poté vypočítáme hodnoty r_{ij}, podle kterých určíme, zda se jedná o řešení či nikoliv. Celý tento postup můžeme popsat následujícím algoritmem.

Modifikovaná distribuční metoda

(i) Jednou z dříve popsaných metod (či jinak) najdeme výchozí BPB.

(ii) Jednu z proměnných u_i a v_j položíme rovnu nule. Pro usnadnění výpočtu vynulujeme proměnnou v řádku nebo sloupci, ve kterém je nejvíce obsazených buněk. Poté ze vztahu $u_i + v_j = n_{ij}$ vypočítáme hodnoty duálních proměnných u_i a v_j.

(iii) Pro nebazické proměnné vypočítáme s využitím vztahu $r_{ij} = n_{ij} - u_i - v_j$ náklady příležitosti. Pokud všechna r_{ij} splňují $r_{ij} > 0$, našli jsme řešení dopravní úlohy a algoritmus končí.

(iv) V opačném případě najdeme $r_{kl} = \min\{r_{ij} \mid r_{ij} \leq 0\}$. Pokud je $r_{kl} = 0$, existuje alternativní řešení úlohy se stejnou hodnotou účelové funkce, které můžeme najít pokračováním v algoritmu. Pokud je $r_{kl} < 0$, příslušná proměnná vstoupí do báze.

(v) Do buňky dané indexy k a l přiřadíme největší možnou hodnotu tak, aby byla splněna všechna poptávková i nabídková omezení a omezení na znaménko pro všechna x_{ij}. Proměnná, kterou v tomto kroku vynulujeme, vystupuje z báze. V tomto kroku také přepočítáme celou tabulku.

(vi) Vrátíme se ke kroku (ii) a algoritmus opakujeme, dokud nenajdeme řešení[217].

Využití této metody si nyní ilustrujeme na dopravním problému zadaném Tabulkou 2.35.

[217] Posloupnost bodů generovaná MODI, je stejná jako posloupnost, kterou bychom získali při řešení problému simplexovou metodou s využitím Dantzigova pravidla. MODI pochopitelně není jedinou metodou pro řešení dopravního problému. V literatuře lze např. najít tzv. *stepping stone method*, viz [10, Chapter 19].

Příklad 2.11.6

Pro výchozí BPB nalezený metodou minimální ceny pro dopravní problém zadaný Tabulkou 2.35 najděme modifikovanou distribuční metodou řešení.

Řešení. Pro inicializaci metody využijeme Tabulku 2.46, ve které je výchozí BPB nalezený metodou minimální ceny. K této tabulce doplníme jeden sloupec a jeden řádek. Do nových buněk tabulky pak zapíšeme hodnoty duálních proměnných u_i a v_i. Proměnnou u_1 položíme rovnu nule, neboť v řádku, který jí přísluší, je nejvíce obsazených buněk.

	O_1	O_2	O_3	nabídka	u_i
P_1	6 | 2	8 | 10	10 | 3	15	$u_1 = 0$
P_2	7 |	11 |	11 | 17	17	$u_2 = 1$
P_3	4 | 18	5 |	12 |	18	$u_3 = -2$
poptávka	20	10	20	50	
v_i	$v_1 = 6$	$v_2 = 8$	$v_3 = 10$		381

Tabulka 2.52: První krok modifikované distribuční metody pro Příklad 2.11.6.

Do poslední buňky ve sloupci a řádku pro hodnoty duálních proměnných vepíšeme hodnotu účelové funkce. Hodnoty ostatních proměnných získáme s využitím vztahu $u_i + v_j = n_{ij}$. Řešíme tedy následující soustavu rovnic

$$u_1 + v_1 = 6 \quad \& \quad u_1 + v_2 = 8 \quad \& \quad u_1 + v_3 = 10$$
$$u_2 + v_3 = 11 \quad \& \quad u_3 + v_1 = 4.$$

Z této soustavy získáme hodnoty duálních proměnných $u_1 = 0$, $u_2 = 1$, $u_3 = -2$, $v_1 = 6$, $v_2 = 8$ a $v_3 = 10$. Tyto hodnoty poté zapíšeme do výchozí tabulky. Nyní již můžeme vypočítat rozdíly $r_{ij} = n_{ij} - u_i - v_j$ pro proměnné, které nejsou pro výchozí BPB v bázi (neboli pro neobsazené buňky tabulky). Tímto krokem získáme náklady příležitosti

$$r_{21} = n_{21} - u_2 - v_1 = 7 - 1 - 6 = 0$$
$$r_{22} = n_{22} - u_2 - v_2 = 11 - 1 - 8 = 2$$
$$r_{32} = n_{32} - u_3 - v_2 = 5 - (-2) - 8 = -1$$
$$r_{33} = n_{33} - u_3 - v_3 = 12 - (-2) - 10 = 4.$$

Vypočtené náklady příležitosti nám říkají, jak by se změnily celkové náklady, kdybychom do prázdné buňky přesunuli jednotku zboží. V případě buňky x_{22} by např. náklady vzrostly o 2 na jednotku přesunutého zboží. Pouze v případě zahrnutí buňky x_{32} máme možnost celkové náklady snížit. Bod, který jsme našli metodou minimální ceny, tak není řešením, neboť celkové náklady nejsou minimální. V tomto kroku do báze vstupuje proměnná x_{32}. Do této buňky chceme alokovat největší možné množství zboží, neboť za každou jednotku, kterou do buňky přesuneme, snížíme celkové náklady o 1. Množství, které do buňky přiřadíme, označíme θ. Postup výběru proměnné, která z báze vystoupí, je zachycen v Tabulce 2.53.

	O_1	O_2	O_3	nabídka	u_i
P_1	6 $2+\theta$	8 $10-\theta$	10 3	15	$u_1=0$
P_2	7	11	11 17	17	$u_2=1$
P_3	4 $18-\theta$	5 θ	12	18	$u_3=-2$
poptávka	20	10	20	50	
v_i	$v_1=6$	$v_2=8$	$v_3=10$		$381-\theta$

Tabulka 2.53: Pokračování modifikované distribuční metody pro Příklad 2.11.6.

Přiřadíme-li do buňky x_{32} hodnotu θ, musíme upravit hodnotu v buňce x_{31} tak, aby bylo zachováno nabídkové omezení prodejce P_3. Od čísla 18 v dané buňce musíme nyní odečíst θ. Nyní musíme změnit hodnotu v buňce x_{11} tak, abychom zachovali poptávkové omezení odběratele O_1, což zajistíme přičtením θ. Tímto krokem jsme porušili omezení prodejce P_1, od buňky x_{12} tedy odečteme θ, abychom bylo omezení opět splněno. Tímto krokem jsme zároveň splnili i poptávkové omezení odběratele O_2. Výsledná uzavřená cesta je tedy $x_{32} \rightsquigarrow x_{31} \rightsquigarrow x_{11} \rightsquigarrow x_{12} \rightsquigarrow x_{32}$. V Tabulce 2.53 jsou všechna omezení splněna.

Maximální hodnota, kterou může θ nabýt, je 10, neboť pro $\theta > 10$ by nebylo splněno omezení na znaménko pro proměnnou x_{12}. Tato proměnná nyní vystupuje z báze. Pro hodnotu $\theta = 10$ získáme Tabulku 2.54.

	O_1	O_2	O_3	nabídka	u_i
P_1	6 12	8	10 3	15	$u_1=0$
P_2	7	11	11 17	17	$u_2=1$
P_3	4 8	5 10	12	18	$u_3=-2$
poptávka	20	10	20	50	
v_i	$v_1=6$	$v_2=7$	$v_3=10$		371

Tabulka 2.54: Druhý krok modifikované distribuční metody pro Příklad 2.11.6.

Hodnota účelové funkce v nalezeném bodě je rovna 371, což je o 10 menší než ve výchozím BPB. Tento výsledek odpovídá tomu, že jsme do buňky x_{32} přiřadili množství 10 a na jednotku množství ušetřili 1. Pro tento bod nyní přepočítáme hodnoty duálních proměnných. Opět položíme $u_1 = 0$ a řešíme následující soustavu rovnic

$$u_1 + v_1 = 6 \quad \& \quad u_1 + v_3 = 10 \quad \& \quad u_2 + v_3 = 11$$
$$u_3 + v_1 = 4 \quad \& \quad u_3 + v_2 = 5.$$

Řešení této soustavy je již zapsáno v Tabulce 2.54. Pro nebazické proměnné potom najdeme náklady příležitosti vyřešením soustavy

$$r_{12} = n_{12} - u_1 - v_2 = 8 - 0 - 7 = 1$$
$$r_{21} = n_{21} - u_2 - v_1 = 7 - 1 - 6 = 0$$
$$r_{22} = n_{22} - u_2 - v_2 = 11 - 1 - 7 = 3$$
$$r_{33} = n_{33} - u_3 - v_3 = 12 - (-2) - 10 = 4.$$

Žádný z vypočtených rozdílů není v tomto kroku záporný, což znamená, že jsme již nalezli řešení úlohy. Rozdíl r_{21} je nulový, proto pro tuto úlohu existuje i alternativní řešení. To můžeme modifikovanou distribuční metodou najít stejně, jako další BPB. Do báze by v tomto kroku vstoupila proměnná x_{21}. Do příslušné buňky bychom opět přiřadili hodnotu θ. Nyní bychom museli ověřit splnění omezení a doplnit ostatní buňky tak, aby byla všechna omezení splněna. Tímto krokem bychom získali Tabulku 2.55.

	O_1	O_2	O_3	nabídka	u_i
P_1	6 | 12-θ	8 |	10 | 3+θ	15	$u_1 = 0$
P_2	7 | θ	11 |	11 | 17-θ	17	$u_2 = 1$
P_3	4 | 8	5 | 10	12 |	18	$u_3 = -2$
poptávka	20	10	20	50	
v_i	$v_1 = 6$	$v_2 = 7$	$v_3 = 10$		371

Tabulka 2.55: Nalezení alternativního řešení pro Příklad 2.11.6.

Z hodnot v Tabulce 2.55 vidíme, že největší možná hodnota θ je rovna 12. Čtenář si snadno vypočítá hodnoty, které by doplněním této hodnoty do tabulky získal. Hodnota účelové funkce je opět rovna 371. Našli jsme tak dvě řešení úlohy, které minimalizují celkové přepravní náklady.

V motivačním Příkladu 2.11.1 i v Příkladu 2.11.6 jsme získali celočíselná řešení, aniž bychom v úloze celočíselnost jakkoli vyžadovali. Tento výsledek není náhodný, naopak, jedná se o důležitou vlastnost dopravního problému, která je předmětem následující věty. Její tvrzení nám však zaručuje také celočíselnost řešení i libovolného přiřazovacího problému s celočíselnými pravými stranami (viz např. Příklad 2.3.8), který je obecně pouze speciálním případem dopravního problému.

VĚTA 2.11.7

Jestliže jsou všechny koeficienty s_i a c_j v dopravní úloze (2.11.1) celočíselné, pak jsou všechny BPB také celočíselné, a to nezávisle na celočíselnosti koeficientů n_{ij}.

Důkaz. Nejdříve ukážeme, že tabulka dopravního problému vždy obsahuje řádek nebo sloupec, který má pouze jeden nenulový prvek. Bez újmy na obecnosti připusťme, že[218] $m \leq n$. Předpokládejme, že každý sloupec tabulky dopravního problému obsahuje alespoň dva nenulové prvky. Pak počet nenulových prvků N nutně splňuje nerovnosti

$$N \geq 2n \geq m + n > m + n - 1,$$

což je ale ve sporu s tím, že každá báze BPB obsahuje nejvýše $m+n-1$ prvků ↯. Nyní označme tento sloupec indexem j a odpovídající nenulový prvek jako x_{ij}. Pak $x_{ij} = c_j$, což je jistě ce-

[218] V opačném případě stačí v dalších výpočtech pouze zaměnit roli m a n.

ločíselné. Nyní uvažme zbytek tabulky po vymazání j-tého sloupce a nahrazení s_i hodnotou $\hat{s}_i = s_i - c_j$. Tato nová tabulka má m řádků a $n-1$ sloupců. Nejvýše $m+n-2$ buněk je obsazených a „na okraji“ této tabulky zůstaly kladné celočíselné hodnoty. Proto můžeme i na tuto zmenšenou tabulku aplikovat výsledek z první části důkazu, tj. opět bude existovat sloupec/řádek s pouze jedním nenulovým prvkem atd. Po k krocích budeme mít tabulku s $m+n-k$ řádků a sloupců s nabídkami (kapacity skladů) a poptávkami (kapacita cílů)

$$\begin{aligned}\hat{s}_i &:= s_i + (\text{součet } s_i \text{ odpovídající vymazaným řádkům})\\ &\qquad - (\text{součet } c_j \text{ odpovídající vymazaným sloupcům})\\ \hat{c}_j &:= c_j + (\text{součet } c_j \text{ odpovídající vymazaným sloupcům})\\ &\qquad - (\text{součet } s_i \text{ odpovídající vymazaným řádkům}).\end{aligned}$$

Dokud je možné zmenšovat tabulku, můžeme tímto způsobem pokračovat a vždy se nám podaří najít řádek nebo sloupec obsahující pouze jeden nenulový prvek x_{ij} s hodnotou $\hat{s}_i$ nebo $\hat{c}_j$. To znamená, že množství přiřazené v nějaké buňce je vždy rozdílem mezi částečným součtem nabídky v řádcích a částečným součtem poptávky ve sloupcích, což bude mít jistě celočíselnou hodnotu[219]. ■

[219] Celočíselnost libovolného BPB souvisí také s tzv. *totální unimodulárností* matice A, na kterou se blíže podíváme v Podkapitole 3.3.

Modifikovanou distribuční metodu si nyní ukážeme i na příkladu nevyrovnaného dopravního problému.

Příklad 2.11.8

Vyřešte nevyrovnaný dopravní problém zadaný Tabulkou 2.56.

odběratel / prodejce	O_1	O_2	O_3	nabídka
P_1	6	7	5	17
P_2	4	10	9	18
P_3	11	6	12	20
poptávka	15	15	30	

Tabulka 2.56: Zadání Příkladu 2.11.8.

Řešení. Z Tabulky 2.56 vidíme, že úloha je nevyrovnaná, neboť součet poptávek jednotlivých odběratelů (= 60) převyšuje součet nabídek prodejců (= 55). Prvním krokem proto při řešení této úlohy musí být zavedení tzv. *fiktivního prodejce*, jehož nabídka problém vyrovná. Náklady na převoz od fiktivního prodejce budou vždy rovny nule, neboť množství v těchto buňkách ve skutečnosti nikam nepřevážíme. Proměnné v řádku fiktivního prodejce jsou tak vlastně *doplňkovými proměnnými*, díky kterým úlohu převedeme na kanonický tvar a které nikterak neovlivní výsledek úlohy. Řešení původní (nevyrovnané) úlohy pak získáme vynecháním přidaného řádku, některá poptávková omezení však nemusí být vzhledem k nedostatečné nabídce splněna. V případě nedostatečné poptávky bychom postupovali obdobně, tj. k úloze bychom v tomto případě přidali tzv. *fiktivního odběratele*, jehož poptávka by úlohu vyrovnala. Řešení bychom poté získali vynecháním sloupce pro fiktivního odběratele a celková nabídka by v tomto případě zůstala nevyčerpána.
Tabulka pro vyrovnanou úlohu s fiktivním prodejcem je zobrazena v Tabulce 2.57 níže. Výchozí BPB najdeme např. pomocí metody minimální ceny. Tou postupujeme následujícím způsobem. Jako první do buňky x_{41} vepíšeme číslo 5, což je nabídka fiktivního prodejce. Toto minimum bychom mohli vepsat i do jakékoli jiné buňky s nulovými ná-

klady, ale my jsme náhodně vybrali právě x_{41}. Dostupná buňka s nejmenšími náklady je nyní x_{21}. Do této buňky přiřadíme 10, neboť jde o minimum z upravené poptávky odběratele O_1 a nabídky prodejce P_2.

	O_1	O_2	O_3	nabídka
P_1	6	7	5	17
P_2	4	10	9	18
P_3	11	6	12	20
FP	0	0	0	5
poptávka	15	15	30	60

Tabulka 2.57: Výchozí tabulka pro Příklad 2.11.8.

V dalším kroku dáme do buňky x_{13} hodnotu 17, čímž vyčerpáme nabídku prodejce P_1. Další vybranou buňkou je x_{32}, do které přiřadíme číslo 15, neboť jsme limitováni poptávkou odběratele O_2. V následujícím kroku vepíšeme číslo 8 do buňky x_{32}, čímž zajistíme splnění všech nabídkových omezení s výjimkou omezení prodejce P_3. V posledním kroku přiřadíme do buňky x_{33} číslo 5. Tímto postupem jsme našli bod, který splňuje všechna omezení. Celkové náklady v tomto bodě jsou 347 a nalezený výchozí BPB je zobrazen v Tabulce 2.58.

	O_1	O_2	O_3	nabídka
P_1	6	7	5 17	17
P_2	4 10	10	9 8	18
P_3	11	6 15	12 5	20
FP	0 5	0	0	5
poptávka	15	15	30	60

Tabulka 2.58: Výchozí bazický přípustný bod pro Příklad 2.11.8.

Nyní metodou MODI ověříme, zda se jedná o řešení úlohy, či zda můžeme dopravní náklady zmenšit. V Tabulce 2.59 jsou uvedeny hodnoty duálních proměnných v případě, že v_3 položíme rovno nule. Při výpočtu těchto hodnot jsme řešili následující soustavu rovnic

$$u_1 + v_3 = 5 \quad \& \quad u_2 + v_1 = 4 \quad \& \quad u_2 + v_3 = 9$$
$$u_3 + v_2 = 6 \quad \& \quad u_3 + v_3 = 12 \quad \& \quad u_4 + v_1 = 0.$$

	O_1	O_2	O_3	nabídka	u_i
P_1	6	7	5 17	17	$u_1 = 5$
P_2	4 10	10	9 8	18	$u_2 = 9$
P_3	11	6 15	12 5	20	$u_3 = 12$
FP	0 5	0	0	5	$u_4 = 5$
poptávka	15	15	30	60	
v_i	$v_1 = -5$	$v_2 = -6$	$v_3 = 0$		347

Tabulka 2.59: První krok modifikované distribuční metody pro Příklad 2.11.8.

Pro nebazické proměnné potom s využitím vztahu $r_{ij} = n_{ij} - u_i - v_j$ dostaneme

$$r_{11} = 6 \quad \& \quad r_{12} = 8 \quad \& \quad r_{22} = 7 \quad \& \quad r_{31} = 4 \quad \& \quad r_{42} = 1 \quad \& \quad r_{43} = -5.$$

Podrobnosti k této části řešení ponecháme na čtenáři. Z výsledků pro nebazické proměnné vidíme, že zapojením proměnné x_{43} do báze získáme menší celkové náklady. Do buňky x_{43} tak přiřadíme hodnotu θ a zjistíme, jak se změní ostatní buňky v tabulce. Toto je uvedeno v Tabulce 2.60.

	O_1	O_2	O_3	nabídka	u_i
P_1	6	7	5 17	17	$u_1 = 5$
P_2	4 10+θ	10	9 8-θ	18	$u_2 = 9$
P_3	11	6 15	12 5	20	$u_3 = 12$
FP	0 5-θ	0	0 θ	5	$u_4 = 5$
poptávka	15	15	30	60	
v_i	$v_1 = -5$	$v_2 = -6$	$v_3 = 0$		347 - 5θ

Tabulka 2.60: Pokračování modifikované distribuční metody pro Příklad 2.11.8.

Z Tabulky 2.60 vidíme, že přiřadíme-li do buňky x_{43} hodnotu θ, musíme pozměnit buňky x_{41}, x_{21} a x_{23}. Tímto krokem opět dosáhneme splnění všech omezení. Maximální hodnota θ je rovna 5 a z báze v tomto kroku odchází proměnná x_{41}. Dosazení této hodnoty do tabulky již ponecháme na čtenáři. V dalším kroku bychom opět vypočítali hodnoty duálních proměnných a následně rozdíly r_{ij} pro nebazické proměnné. V tomto kroku bychom získali pouze kladné rozdíly, což znamená, že bod nalezený jedním krokem metody MODI je optimální. Minimální náklady na přepravu jsou 322, což můžeme zjistit jak vynásobením převážených hodnot a nákladů na jednotku, tak odečtením hodnoty 25 od celkových nákladů ve výchozím bazickém přípustném bodě.

V novém bodě jsou náklady o 25 menší, neboť jsme do buňky, která přinese pokles nákladů o 5 na jednotku přesunuli množství 5.
Výsledek původní úlohy získáme vynecháním řádku pro fiktivního prodejce. Řešení nevyrovnané úlohy je zobrazeno v Tabulce 2.61. V této tabulce vidíme, že poptávkové omezení odběratele O_3 není kvůli nedostatečné nabídce splněno.

	O_1	O_2	O_3	nabídka
P_1	6	7	5 17	17
P_2	4 15	10	9 3	18
P_3	11	6 15	12 5	20
poptávka	15	15	30	60 55

Tabulka 2.61: Řešení Příkladu 2.11.8.

Podobně jako ostatní úlohy lineárního programování, může i dopravní problém být degenerovaný, k čemuž opět dojde v případě, kdy některá z bazických proměnných je rovna nule. V případě dopravního problému poznáme degenerovaný problém tak, že jsou všechna omezení splněna obsazením méně než $n+m-1$ buněk v tabulce[220]. V případě degenerované úlohy se můžeme dostat do situace, kdy nejsme schopni vypočítat hodnoty duálních proměnných a rozdílů r_{ij}. Tuto situaci můžeme vyřešit tak, že do jedné z prázdných buněk vepíšeme hodnotu nula, neboli zavedeme do báze novou proměnnou s nulovou hodnotou. Poté již můžeme příslušné hodnoty vypočítat a pokračovat v MODI metodě. Degenerovaný dopravní problém si ukážeme v Příkladu 2.11.9 níže. Ještě dodejme, že i v případě dopravního problému může u degenerovaných úloh dojít k zacyklení, čemuž se lze vyhnout podobně jako při využití simplexové metody, tj. např. pomocí náhodného pravidla či perturbační metody[221].

[220] Exaktněji řečeno platí následující tvrzení: *dopravní problém je degenerovaný právě tehdy, když existují indexové množiny* $\emptyset \subsetneq I' \subsetneq \{1,\dots,m\} =: I$ *a* $\emptyset \subsetneq J' \subsetneq \{1,\dots,j\} =: J$ *takové, že*

$$\sum_{i\in I'} s_i = \sum_{j\in J'} c_j.$$

Pro důkaz viz např. [414, str. 199]. Odtud je vidět, že degenerovaných dopravních problémů může být velké množství. Avšak jen pro málokterou z nich dojde při řešení k zacyklení. Jedním takovým uměle sestrojeným příkladem je např.

prodejce \ odběratel	O_1	O_2	O_3	O_4	nabídka
P_1	3	5	5	11	1
P_2	9	7	9	15	1
P_3	7	7	11	13	1
P_4	13	13	13	17	1
poptávka	1	1	1	1	4

Viz opět [414, str. 198] a také [203, str. 253–254].

[221] I zde platí: *pro libovolný dopravní problém existuje číslo* $\varepsilon_1 > 0$ *takové, že pro každé* $\varepsilon \in (0,\varepsilon_1)$ *je odpovídající* ε*-úloha nedegenerovaná.* Viz [414, str. 200].

Příklad 2.11.9

Metodou minimální ceny najděme výchozí BPB dopravního problému zadaného Tabulkou 2.62 a vyřešme jej.

prodejce \ odběratel	O_1	O_2	O_3	nabídka
P_1	3	11	9	10
P_2	4	6	8	30
P_3	8	13	7	20
poptávka	10	25	25	60

Tabulka 2.62: Zadání Příkladu 2.11.9.

Řešení. Při řešení tohoto problému postupujeme obdobně jako v předchozích problémech. Nejprve metodou minimální ceny najdeme výchozí bazický přípustný bod zobrazený v Tabulce 2.63. Posloupnost buněk, které doplňujeme, je x_{11}, x_{22}, x_{33}, x_{23}. Ověření správnosti tohoto výsledku ponecháme na čtenáři. Stejný výchozí bazický přípustný bod bychom získali i metodou severozápadního rohu a Vogelovou aproximační metodou.

	O_1	O_2	O_3	nabídka
P_1	3: 10	11:	9:	10
P_2	4:	6: 25	8: 5	30
P_3	8:	13:	7: 20	20
poptávka	10	25	25	60

Tabulka 2.63: Výchozí bazický přípustný bod pro Příklad 2.11.9.

Chceme-li pro nalezený bod zahájit modifikační distribuční metodu, získáme následující soustavu rovnic

$$u_1 + v_1 = 3 \quad \& \quad u_2 + v_2 = 6$$
$$u_2 + v_3 = 8 \quad \& \quad u_3 + v_3 = 7.$$

Položíme-li u_2 rovno nule, dostaneme $u_3 = -1$, $v_2 = 6$ a $v_3 = 8$. Kvůli degenerovanosti nalezeného bodu (v Tabulce 2.63 jsou vyplněny pouze 4 buňky, což je méně než $3+3-1=5$) nejsme schopni dopočítat hodnoty u_1 a v_1. Do jedné z nebazických buněk v tabulce tak doplníme hodnotu 0, a to např. do buňky x_{21}. V tomto kroku můžeme vybrat jakoukoli nebazickou proměnnou z prvního řádku a prvního sloupce, neboť tyto proměnné nám umožní soustavu vyřešit. Jedno z pravidel doporučuje vybrat buňku s nejnižšími náklady. Doplníme-li proměnnou do buňky x_{21}, která je shodou okolností buňkou s nejnižšími náklady, soustava bude navíc obsahovat rovnici $u_2 + v_1 = 4$. S využitím této rovnice již můžeme dopočítat hodnoty všech duálních proměnných, máme tedy $u_1 = -1$, $v_1 = 4$. Tento výsledek je zobrazen v Tabulce 2.64.

	O_1	O_2	O_3	nabídka	u_i
P_1	3: 10	11:	9:	10	-1
P_2	4: 0	6: 25	8: 5	30	0
P_3	8:	13:	7: 20	20	-1
poptávka	10	25	25	60	
v_j	4	6	8		360

Tabulka 2.64: Řešení Příkladu 2.11.9.

Pro nebazické proměnné dostaneme hodnoty

$$r_{12} = 11 - (-1) - 6 = 6,$$
$$r_{13} = 9 - (-1) - 8 = 2,$$
$$r_{31} = 8 - (-1) - 4 = 5,$$
$$r_{32} = 13 - (-1) - 6 = 8,$$

ze kterých vidíme, že výchozí BPB je zároveň řešením úlohy. Minimální dopravní náklady jsou 360. Ke stejnému výsledku bychom dospěli i zavedením x_{12} či jiné původně nebazické proměnné do báze.

2.12 Některé aplikace nyní podrobněji

V této podkapitole se podrobněji podíváme na šest složitějších úloh lineárního programování. Tyto příklady se sice zaobírají komplexnějšími situacemi než tomu bylo u motivačních příkladů v úvodní Podkapitole 2.3, ale i tak v nich pochopitelně dochází k jistému zjednodušení reality. Současně také zdůrazněme, že vzhledem k jejich rozměrům již tyto úlohy rozhodně nejsou určeny pro řešení pouze pomocí simplexového algoritmu, tužky a papíru, ale je velmi užitečné k tomu využít nějaký výpočetní software[222], což vlastně zcela koresponduje i s historickým vývojem této oblasti.

V některých těchto příkladech je již z jejich zadání zřejmé, že vedou na úlohu lineárního programování a jak odpovídající model sestavit. U jiných to až tak očividné být nemusí. Zde je naše myšlení nepochybně velmi často ovlivněno hledáním analogie s některou ze situací, se kterými jsme se již setkali. Avšak u problémů vzešlých z nějaké skutečně reálné (a nové) situace se může stát, že tento proces není zrovna přímočarý a vyžaduje i jistou míru představivosti. Samotné sestavení celého modelu je totiž jednou z nejkreativnějších a současně možná i nejsložitější částí celého operačního výzkumu[223]. Pro hlubší zamyšlení nad tímto aspektem sestavování matematických modelů v lineárním programování doporučujeme článek [387], ve kterém nalezneme i jakýsi historický vývoj modelů lineárního programování v jeho počátcích.

Problém jídelníčku. S tímto problémem jsme se setkali již v Podkapitolách 1.3, 2.1 a 2.6, kde jsme viděli i klasický Stiglerův problém optimálního jídelníčku, který hrál velmi důležitou roli v počátcích rozvoje lineárního programování. Snad každý správný *lineární programovač* by si měl alespoň jednou v životě zkusit sestavit vlastní problém jídelníčku a nalézt jeho řešení. Jenže celkem snadno se můžeme dostat do podobných potíží jako Dantzig, když se o to sám kdysi pokusil. Asi i proto je tento problém sice nedílnou součástí každé učebnice lineárního programování, ovšem obvykle ve velmi zjednodušené podobě (např. založené pouze na pěti potravinách apod.). Autorka [528] se o to zdařile pokusila ve své diplomové práci, když s využitím třech desítek potravin chtěla najít co nejlevnější osobní jídelníček, který ji umožní ušetřit na vysněnou knihu překypující zajímavostmi z matematické analýzy a jiných (méně zábavných) oblastí matematiky a zároveň zaručí dostatek sil na psaní diplomové práce. Nikdy nevíme, kdy se to čtenářkám či čtenářům může hodit, tak si tento příklad uvedeme i zde.

Příklad 2.12.1

Na začátek je nutné zmínit, že jídelníček je sestaven pro studentku a požadované výživové hodnoty nejsou univerzálně platné. Při sestavování jídelníčku je důležité si stanovit, jaké výživové hodnoty budeme sledovat a kolik bychom jich měli denně přijímat. Pro tento příklad jde o 7 hodnot, a to celkovou energetickou hodnotu (v kJ), množství tuků (g), sacharidů (g), bílkovin (g), vlákniny (g), vitamínu C (mg) a obsahu vody (ml). Otázka kolik čeho jíst je při sestavování jídelníčku velmi důležitá, mírně však přesahuje obsah tohoto textu a našich možností. Pro reálnost hodnot však byly pro zjištění těchto informací využity [297, 418, 535]. Pro vypočtení doporučených hodnot pro celkovou energii byla využita tzv. *Harrisova–Benedictova rovnice*. Z této rovnice na základě věku, výšky, váhy a pohlaví zjistíme, kolik je tzv. *bazální metabolismus*, který udává výdej energie v klidovém stavu. Tento údaj poté upravíme podle typu činností, které daný člověk vykonává. Pro studentku je tak potřebné přijmout přibližně 8060 kJ energie denně. Z výživových doporučení a obsahu energie v jednotlivých typech živin pak vy-

[222] Nabídka je v tomto ohledu poměrně široká. Lze využít např. doplněk Řešitel pro Excelu (viz odkaz) nebo programy LINDO (viz https://ln.math.muni.cz/pz/opt/simplexSOFT2), LiPS (viz https://ln.math.muni.cz/pz/opt/simplexSOFT1) či některý z on-line nástrojů jako jsou např. od kesspess (viz https://ln.math.muni.cz/pz/opt/simplexSOFT4) nebo HMLA (viz https://ln.math.muni.cz/pz/opt/simplexSOFT5).

[223] Toto asi platí i o drtivé většině praktických problémů, které se snažíme nějakým způsobem matematizovat.

počítáme, kolik z toho by měla přijmout ve formě tuku, sacharidů, bílkovin atd. Tímto postupem zjistíme, že doporučený denní příjem pro studentku je 70 gramů tuku, 210 gramů sacharidů, 100 gramů bílkovin, 25 gramů vlákniny a 80 miligramů vitamínu C. Je také nutné, aby vypila alespoň 2 litry vody.

Dalším krokem při sestavování jídelníčku je výběr potravin, které budeme uvažovat. Pro tento příklad jsou potraviny rozděleny do 6 kategorií, a to nápoje, mléčné výrobky, maso, ovoce a zelenina, přílohy a sladkosti. V každé kategorii je vybráno několik potravin, které jsou běžně dostupné a tvoří velkou část obvyklých jídelníčků. U těchto potravin jsou uvedeny také jejich výživové hodnoty, a to z [536, 538, 541], odkud jsou převzaty též ceny jednotlivých výrobků. Výživové hodnoty i ceny všech potravin, s výjimkou vajec, jsou uvedeny pro 100 gramů potraviny. Hodnoty u vejce jsou přepočteny pro jeden kus s tím, že jedno vejce váží přibližně 70 gramů. Seznam vybraných potravin, jejich ceny a výživové hodnoty jsou uvedeny v Tabulce 2.65. V posledním řádku tabulky jsou uvedeny doporučené denní dávky jednotlivých živin.

surovina (100 g)	energie (kJ)	tuk (g)	sacharidy (g)	bílkoviny (g)	vláknina (g)	vitamín C (mg)	voda (ml)	cena (Kč)
neperlivá voda	0	0	0	0	0	0	100	0,8
Coca-Cola	184,8	0	11,2	0	0	0	100	2,79
Poděbradka cit.	75,6	0	4,2	0	0	0	100	1,06
bílý jogurt	280	3,1	5	4,9	0	0	0	6,5
polotučné mléko	192	1,5	4,8	3,2	0	0	0	1,99
eidam (30%)	1215	19	0,3	30	0	0	0	20,9
cottage	497	7	3	11	0	0	0	18,9
kuřecí prsa	486	2,1	0	17,8	0	0	0	14,9
tuňák	461	1	0	24	0	0	0	29,9
vejce	512	8,5	0,53	11,2	0	0	0	2,99
kuřecí šunka	702	11	0	18,2	0	0	0	24,9
losos	902	15,9	0	19,8	0	0	0	34,9
jablko	218,4	0,2	14	0,3	2,4	4,6	0	2,99
banán	369,6	0,3	23	1,1	2,6	8,7	0	2,99
kiwi	252	0,5	15	1,1	3	92,7	0	4,99
mrkev	172,2	0,2	10	0,9	2,8	5,9	0	1,99
rajče	71,4	0,2	3,9	0,9	1,2	13,7	0	5,99
okurka	63	0,1	3,6	0,7	0,5	2,8	0	6
kukuřice (konz.)	309	1,4	11	2,7	4	0	0	7
hrách (konz.)	308	0,7	7,4	5,5	7,8	0	0	6
brambory	336	0,4	19,2	2	2,2	19,7	0	3,2
těstoviny	1603	1,3	78	10,2	3,6	0	0	1,99
ovesné vločky	1571	3	60,3	14,2	2	0	0	4,9
rýže	1489	2,9	79	8,6	0,4	0	0	3,99
kuskus	1563	2,1	73	12,9	4,2	0	0	6,99
chléb kmínový	950	1	46,6	7,6	2,7	0	0	2,99
čočka	1306	1,6	39	24	8	0	0	3,99
Tatranka	2110	28	54	5,8	4,6	0	0	19,8

surovina (100 g)	energie (kJ)	tuk (g)	sacharidy (g)	bílkoviny (g)	vláknina (g)	vitamín C (mg)	voda (ml)	cena (Kč)
Jojo medvídci	1350	0,3	74	5	0,67	0	0	24,9
Milka mléčná	2213	29	59	6,3	1,8	0	0	29,9
denní příjem	8060	70	210	100	25	80	2000	

Tabulka 2.65: Výživové hodnoty vybraných potravin.

Nyní již známe všechny potřebné informace a můžeme tak sestavit a vyřešit konkrétní problém jídelníčku. Po vyřešení se ještě pozastavíme nad situací, kdy studentka uvažuje i některá maximální omezení, a nad případem, ve kterém chce do svého jídelníčku přidat alespoň jednu potravinu z vybrané kategorie. Vzhledem k velikosti této úlohy zde nebudeme uvádět průběh simplexové metody, ale pouze výsledek zadané úlohy.

Řešení. Označme množství potravin v optimálním jídelníčku postupně $x_1, \ldots, x_{30}$ a ceny těchto surovin označíme p_i. Jednotlivé výživové hodnoty označíme podle pořadí jim příslušných sloupců v Tabulce 2.65. Problém, který studentka řeší, můžeme zapsat jako

$$\sum_{i=1}^{30} p_i\, x_i \to \min$$

$$\sum_{i=1}^{30} a_{i1}\, x_i \geq 8060 \quad \& \quad \sum_{i=1}^{30} a_{i2}\, x_i \geq 70 \quad \& \quad \sum_{i=1}^{30} a_{i3}\, x_i \geq 210$$

$$\sum_{i=1}^{30} a_{i4}\, x_i \geq 100 \quad \& \quad \sum_{i=1}^{30} a_{i5}\, x_i \geq 25 \quad \& \quad \sum_{i=1}^{30} a_{i6}\, x_i \geq 80 \quad \& \quad \sum_{i=1}^{30} a_{i7}\, x_i \geq 2000$$

$$x_1 \geq 0 \quad \& \quad \cdots \quad \& \quad x_{30} \geq 0.$$

Řešením této úlohy je situace, ve které studentka pije pouze vodu (2 litry) a jí 7,23 vejce, 86 gramů kiwi a 622,5 gramů těstovin. Za den tak za svoje stravování zaplatí 54,32 Kč. Označme si tento scénář jako jídelníček 1. Výživové hodnoty tohoto jídelníčku jsou zobrazeny v Tabulce 2.66.

surovina (100 g)	energie (kJ)	tuk (g)	sacharidy (g)	bílkoviny (g)	vláknina (g)	vitamín C (mg)	voda (ml)	cena (Kč)
denní příjem	8060	70	210	100	25	80	2000	
jídelníček 1	13899,6	70	502,35	145,45	25	80	2000	54,32

Tabulka 2.66: Výživové hodnoty pro jídelníček 1.

V Tabulce 2.66 vidíme, že jídelníček 1 obsahuje téměř 14 000 kJ energie, což je oproti požadovaným 8 060 o dost více. Studentka sice chce ušetřit, nechce ale přijímat více energie než potřebuje a ztloustnout. Také není spokojená s vysokým příjmem sacharidů. Rozhodneme se tedy přidat k problému další podmínky, a to maximální denní příjem energie na 9 000 kJ a sacharidů na 280 g. Nová omezení jsou ve tvaru

$$\sum_{i=1}^{30} a_{i1} x_i \leq 9000 \quad \& \quad \sum_{i=1}^{30} a_{i3} x_i \leq 280.$$

Vyřešíme-li úlohu s přidanými omezeními, získáme jídelníček, který označíme číslem 2. Jídelníček 2 se skládá z následujících množství jednotlivých potravin: 2 litry vody, 7,59 vejce, 81 gramů kiwi, 82 gramů mrkve, 143,6 gramu těstovin a 188,7 gramu čočky. Tento jídelníček musí být dražší než jídelníček 1, růst ceny není ovšem tak dramatický, jedná se pouze o 45 haléřů, viz Tabulku 2.67. S tímto jídelníčkem je studentka

již spokojenější, nepřijímá totiž zbytečně moc energie, všechny výživové hodnoty (až na bílkoviny) jsou v podstatě na své doporučené hodnotě a navíc je strava pestřejší. Přesto se studentce zdá, že přijímá až moc energie, pro její běžný život by podle výpočtu mělo stačit 8060 kJ. Rozhodne se tedy změnit maximální denní příjem energie na 8 500. Zároveň je právě uprostřed psaní diplomové práce a ráda by si občas dopřála nějakou sladkost, a to alespoň 35 gramů denně. V úloze tedy změníme maximální denní příjem energie na 8500 a přidáme omezení ve tvaru

$$x_{28} + x_{29} + x_{30} \geq 0,35.$$

Vyřešíme-li úlohu s těmito omezeními, získáme následující stravovací plán: 2 litry vody, 6,53 vejce, 73 gramů kiwi, 206,39 gramu mrkve, 133,77 gramu těstovin, 132,5 gramu čočky a 35 gramů Tatranky. Přidáním dalších omezení se stravovací plán opět zdražil, za den by tak studentku stál 58,18 Kč. Na to je ale ochotná přistoupit, neboť si denně může dát téměř celou Tatranku a ostatní živiny přijímá v dostatečném a přiměřeném množství, viz Tabulku 2.67. S tímto jídelníčkem je tak spokojená.

surovina	energie	tuk	sacharidy	bílkoviny	vláknina	vitamín C	voda	cena
(100 g)	(kJ)	(g)	(g)	(g)	(g)	(mg)	(ml)	(Kč)
denní příjem	8060	70	210	100	25	80	2000	
jídelníček 1	13899,6	70	502,35	145,45	25	80	2000	54,32
jídelníček 2	9000	70	210	146,61	25	80	2000	54,77
jídelníček 3	8500	70	210	123,35	25	80	2000	58,18

Tabulka 2.67: Výživové hodnoty pro jídelníčky.

Minimalizace nákladů na zásobování. Nyní se zaměříme na úlohu motivovanou nalezením minimálních nákladů pro určitý proces. Úlohy tohoto typu společně s úlohami, ve kterých se maximalizuje zisk, hrají v ekonomii významnou roli, neboť se dotýkají všech ekonomických subjektů. Námi uvedený příklad je motivován plánováním zásobování v poválečném Berlíně z [7, Exercise 1.4], ale tentokrát půjde o letecké zásobování čínského Wu-chanu[224,225].

Příklad 2.12.2

Na konci ledna 2020 byla kvůli šíření koronaviru pro čínské město Wu-chan vyhlášena karanténa. Jedním ze způsobů, jak do Wu-chanu dovézt potřebné zdravotnické pomůcky, jídlo a další nezbytné věci, je letecká doprava. My se ocitáme v roli poradců státu, kteří plánují dopravit do Wu-chanu ve 4 obdobích postupně 200, 300, 300 a 400 tisíc tun zboží. K tomu máme k dispozici letadla s kapacitou 2 000 tun. Našim cílem je najít plán zásobování, který minimalizuje náklady. Tato operace však bude poměrně náročná, neboť existuje několik aspektů, na které musíme brát ohledy. Také se musíme potýkat s omezeným množstvím pilotů i letadel. Všechna nařízení a omezení si nyní rozebereme podrobněji.

(i) V prvním období máme k dispozici 330 pilotů. Každé letadlo musí obsluhovat 3 piloti. Na začátku operace máme 110 letadel. V průběhu času můžeme najímat a cvičit nové piloty i zakoupit nová letadla, o jejich počtech však musíme rozhodnout vždy o jedno období dříve.

(ii) V každém období přijdeme o 20 % pilotů, a to z důvodu nakažení koronavirem, i letadel, která zůstanou v Číně, neboť piloti jsou v karanténě. Přijdeme o ně vždy až po doručení zásob. Podle nařízení týkajících se letecké bezpečnosti pak piloti, kteří se vrátí zpět, musí alespoň 1 období odpočívat.

[224] Podobný příklad lze nalézt také v [529, str. 37–39]. V tomto textu je i ukázka využití lineárního programování při diagnostice rakoviny prsu.

[225] V oponentském posudku práce [528] se k tomuto příkladu píše, že *je ovšem triviální: proměnné vyjadřující požadované počty letadel a pilotů jsou totiž ve skutečnosti konstanty a vzhledem k monotónně klesajícím cenám je třeba všechny nákupy a výcviky odložit na nejpozdější možnou dobu.* No, posouzení adekvátnosti takového komentáře ponecháme na čtenáři. . .

(iii) Piloti, kteří v daném období neletí do Wu-chanu, mohou být v záloze, odpočívat (pokud minulé období letěli) nebo trénovat nové piloty. Pilot, který se ujme tréninku, je v jednom období schopen vytrénovat 19 pilotů (včetně něho tedy 20 pilotů) schopných pracovat v následujícím období.

(iv) V každém období můžeme nakoupit nová letadla. Jejich ceny jsou společně s náklady na různé aktivity pilotů uvedeny v Tabulce 2.68. Náklady na letadla, která v daném období nevyužijeme, jsou nulové. Všechny ceny jsou uvedeny ve stejných měnových jednotkách.

náklady	období 1	období 2	období 3	období 4
nové letadlo	400	375	350	325
odpočívající pilot	5	4,9	4,8	4,7
pilot v záloze	7	6,9	6,8	6,7
trénující pilot	14,3	11,9	9,8	9,3

Tabulka 2.68: Náklady pro jednotlivá období.

Při plánování zásobování Wu-chanu nebereme v potaz cenu materiálu, který vezeme. Zajímá nás pouze nalezení strategie, která minimalizuje náklady na samotný převoz. S využitím těchto omezení formulujme úlohu, jejímž výsledkem bude nalezení minimálních nákladů na zásobovací akci.

Řešení. Pro přehlednost si nejprve zavedeme značení. Pro celou úlohu platí $i \in \{1,2,3,4\}$ a označíme jako a_i množství zboží, které musíme v i-tém období dopravit do Wu-chanu ve stovkách tisíc tun. Zkratkou ns_i označíme počet letadel, která v i-tém období nepoužijeme, ks_i pak označuje počet nově zakoupených letadel v daném období. Pro piloty zavedeme následující značení: op_i pro odpočívající piloty, zp_i pro piloty v záloze a np_i pro piloty, kteří se v i-tém období věnují tréninku (a to jak trenéři, tak nově trénovaní). Počet trenérů je pak $np_i/20$ a o problémech spojených s neceločíselností tohoto výrazu se zmíníme po vyřešení úlohy.
Ze zadání víme, že pro množství převezeného zboží musí platit

$$a_1 = 2 \quad \& \quad a_2 = 3 \quad \& \quad a_3 = 3 \quad \& \quad a_4 = 4.$$

Nyní si zformulujeme podmínky, které musí splňovat počty letadel. Pro převezení 100 tisíc tun zboží potřebujeme 50 letadel. V každém období tak pro převoz využijeme $50\,a_i$ letadel. Zbytek letadel, která máme v daném období k dispozici, nevyužijeme. Pro první období tak můžeme psát

$$50\,a_1 + ns_1 = 110,$$

neboť na začátku operace máme k dispozici 110 letadel. Pro následující období je situace složitější. Z letadel, která jsme v předchozím období využili, nám jich zůstane 80 %. Letadla ve druhém období mohou být opět využita dvěma způsoby – mohou být nasazeny do akce, nebo mohou být nevyužitá. Počet těchto letadel získáme součtem 80 % letadel použitých v minulém období (tj. $40\,a_1$), počtem nově zakoupených letadel v minulém období (ks_1) a letadel, která jsme v minulém období nevyužili (ns_1). Celkově tedy dostaneme

$$50\,a_2 + ns_2 = 40\,a_1 + ks_1 + ns_1.$$

Pro třetí a čtvrté období postupujeme obdobně, čímž získáme rovnosti

$$50\,a_3 + ns_3 = 40\,a_2 + ks_2 + ns_2,$$
$$50\,a_4 + ns_4 = 40\,a_3 + ks_3 + ns_3.$$

Dále se zaměříme na podmínky, které musí splňovat počty pilotů. K obsluze jednoho letadla potřebujeme vždy 3 piloty. Počet pilotů, kteří v daném období letí do Wu-chanu, je tak třikrát větší, než počet letadel, tj. $150\,a_i$. Další role, kterou mohou piloti mít, je trénování nových pilotů. Abychom pro následující období vytrénovali np_i pilotů, potřebujeme $np_i/20$ pilotů trenérů. Piloti mohou mít ještě dvě činnosti, a to odpočinek po absolvování cesty do Wu-chanu, nebo bytí v záloze. Pro piloty v prvním období tak musí platit rovnost

$$150\,a_1 + 0{,}05\,np_1 + zp_1 = 330,$$

neboť na začátku můžeme využít 330 pilotů. Ve druhém období můžeme využít piloty, kteří byli v předchozím období v záloze nebo trénovali. V dalších obdobích pak budeme moci navíc využít piloty, kteří v předchozím období odpočívali. Pro piloty ve druhém období proto platí rovnost

$$150\,a_2 + 0{,}05\,np_2 + zp_2 = zp_1 + np_1.$$

Pro piloty v následujících obdobích získáme obdobným postupem rovnosti

$$150\,a_3 + 0{,}05\,np_3 + zp_3 = zp_2 + np_2 + op_2,$$
$$150\,a_4 + 0{,}05\,np_4 + zp_4 = zp_3 + np_3 + op_3.$$

Poslední podmínky se týkají pilotů, kteří odpočívají po výkonu. Počet pilotů, kteří ve druhém období odpočívají, můžeme zapsat jako $120\,a_1$, neboť 20 % pilotů, kteří v prvním období letěli, muselo zůstat v karanténě ve Wu-chanu. Pro odpočívající piloty tedy platí rovnosti

$$op_1 = 0 \quad \& \quad op_2 = 120\,a_1$$
$$op_3 = 120\,a_2 \quad \& \quad op_4 = 120\,a_3.$$

Všechny proměnné jsou navíc omezené na znaménko, tj. platí

$$a_i \geq 0 \quad \& \quad ns_i \geq 0 \quad \& \quad ks_i \geq 0 \quad \& \quad op_i \geq 0 \quad \& \quad zp_i \geq 0 \quad \& \quad np_i \geq 0.$$

Nyní již máme všechny podmínky odpovídající požadavkům uvedeným v zadání. Funkce, kterou minimalizujeme, je formálně funkcí 24 proměnných, 8 z nich v ní však vystupuje s nulovým koeficientem (všechna a_i a ns_i), neboť náklady na tyto proměnné jsou nulové. Účelovou funkci tak s využitím nákladů z Tabulky 2.68 zapíšeme jako

$$400\,ks_1 + 375\,ks_2 + 350\,ks_3 + 325\,ks_4 + 5\,op_1 + 4{,}9\,op_2 + 4{,}8\,op_3 + 4{,}7\,op_4 +$$
$$+ 7\,zp_1 + 6{,}9\,zp_2 + 6{,}8\,zp_3 + 6{,}7\,zp_4 + 14{,}3\,np_1 + 11{,}9\,np_2 + 9{,}8\,np_3 + 9{,}3\,np_4.$$

Všechna omezení jsou ve tvaru rovností a všechny proměnné jsou omezené na znaménko, již tedy máme úlohu v kanonickém tvaru. Omezení upravíme tak, aby jednotlivé proměnné byly na levé straně. Pro přehlednost si omezení, která jsme tímto způsobem získali, uvedeme v Tabulce 2.69. V této tabulce jsou zobrazeny i koeficienty účelové funkce. Vzhledem k tomu, že úlohu řešíme jako problém lineárního programování, nemáme zaručeno, že získáme celočíselné výsledky. Výsledky této úlohy (zaokrouhlené na 2 desetinná místa) jsou zobrazeny v Tabulce 2.70.

náklad				letadla navíc				nová letadla				odpočívající piloti				piloti v záloze				trénující piloti				omezení
a_1	a_2	a_3	a_4	ns_1	ns_2	ns_3	ns_4	ks_1	ks_2	ks_3	ks_4	op_1	op_2	op_3	op_4	zp_1	zp_2	zp_3	zp_4	np_1	np_2	np_3	np_4	
0	0	0	0	0	0	0	0	400	375	350	325	5	4,9	4,8	4,7	7	6,9	6,8	6,7	14,3	11,9	9,8	9,3	min
1																								2
	1																							3
		1																						3
			1																					4
50				1																				110
−40	50			−1	1			−1																0
	−40	50			−1	1			−1															0
		−40	50			−1	1			−1														0
150																1				0,05				330
	150															−1	1			−1	0,05			0
		150															−1	1			−1	0,05		0
			150															−1	1			−1	0,05	0
												1												0
−120													1											0
	−120													1										0
		−120													1									0

Tabulka 2.69: Omezení pro Příklad 2.12.2

období \ proměnná	a_i	ns_i	ks_i	op_i	zp_i	np_i
1	2	10	60	0	7,31	453,79
2	3	0	30	240	0	222
3	3	0	80	360	0	240
4	4	0	0	360	0	0

Tabulka 2.70: Výsledky Příkladu 2.12.2

Hodnota účelové funkce v bodě z Tabulky 2.70 je 79380,16. V bodě, ve kterém se realizují minimální náklady, musíme v prvním období nakoupit 60 nových letadel, ve druhém období 30 letadel a ve třetím období 80 letadel. Ve všech obdobích, s výjimkou prvního, využijeme k zásobování všechna dostupná letadla. Počty odpočívajících pilotů jsou jasně dány počty pilotů, kteří pracovali v minulém období.
V případě pilotů v záloze a trénujících pilotů již pro první období dostaneme neceločíselný výsledek. Jeho prostým zaokrouhlením můžeme získat představu o optimálním celočíselném bodě, tato strategie ale nezafunguje vždy. Bod, který tímto způsobem získáme, také nemusí splňovat všechna omezení. Řešení problémů celočíselného programování metodami pro lineární programování může být i tak užitečné, neboť jsou zpravidla rychlejší a výsledek, který získáme, nám může dát dostatečně přesnou informaci o celočíselném výsledku. Hodnota účelové funkce pro celočíselné hodnoty navíc nemůže být menší než hodnota, kterou jsme našli. Zaokrouhlíme-li počet pilotů v záloze, získáme pro první období výsledek 7. U trénujících pilotů je situace obtížnější. Pro 453,79 pilotů potřebujeme $453,79/20 = 22,6895$ trenérů. Jestliže zaokrouhlíme počet trenérů v prvním období na 23, budeme mít ve druhém období k dispozici 460 nově vytrénovaných pilotů. Tímto zásahem se změní hodnoty některých omezení, což ovlivní např. i počty pilotů v záloze (v prvním období jich vycvičíme více, než je optimální, počty pilotů v záloze tak vzrostou). V případě, že bychom chtěli získat celočíselný výsledek, ale využít pouze simplexovou metodu, bychom nyní mohli problém přepočítat pro druhé až čtvrté období, výsledky pro druhé období vhodně zaokrouhlit a pokračovat pro třetí a čtvrté období. Na závěr bychom museli ověřit, že bod, který tímto způsobem získáme, leží v přípustné množině.
I přesto, že jsme řešením tohoto problému simplexovou metodou získali pouze přibližný výsledek, jsme si ukázali další zajímavou aplikaci lineárního programování. Vzhledem ke tvaru omezení této úlohy bychom ji mohli rozdělit na dvě podúlohy – jednu pro letadla, druhou pro piloty. Výsledek bychom poté získali součtem hodnot těchto úloh.

Příklad, který jsme právě viděli, je zajímavý i tím, že optimalizujeme na několik období, která jsou na sobě závislá. Příklad bychom mohli též rozdělit na 4 samostatné příklady, které by se týkaly jednotlivých období. Jedná se tak vlastně o problém tzv. *dynamického programování*, kterému se budeme podrobněji věnovat v Kapitole 5.

I s příklady jako je tento se můžeme setkat i v běžném životě. Aktuální problém, který při začínající epidemii koronaviru řeší běžné domácnosti, je jestli a jak se dobře zásobit. Řeší tak problém podobný tomu, který jsme si zde ukázali, na rozdíl od příkladu však dochází k růstu cen. Ztrátu mezi obdobími si můžeme odůvodnit např. tím, že část potravin se v průběhu času kazí. Do hry by navíc mohla vstupovat množstevní omezení na jeden nákup atd. Další aplikací lineárního programování, kterou si zde uvedeme, je vytváření pracovních rozvrhů.

[226] Angl. *scheduling problems.*

Pracovní rozvrhy[226]. V této části si ukážeme další možné využití lineárního programování v běžném životě, a to při sestavování rozvrhů či plánování pracovních směn. V provozech jako jsou kina, restaurace, továrny či některá oddělení nemocnic je nutné naplánovat směny tak, aby byl vždy přítomen dostatečný počet pracovníků, kteří mohou být homogenní, tj. všichni jsou schopni dělat stejnou práci (s touto možností se můžeme setkat např. v kinech, kde všichni brigádníci dělají všechny úkony), nebo se můžou ve svých schopnostech lišit (např. sestry a lékaři v nemocnicích). Při plánování je též nutné dodržovat předpisy, např. o maximálním počtu odpracovaných hodin atp. V následujícím příkladu se zaměříme na tvorbu pracovního rozvrhu v nemocnici. V tomto příkladu ilustrujeme jak situaci homogenních pracovníků, tak případ, kdy máme dvě různé skupiny zaměstnanců[227].

[227] Více o podobných problémech čtenář najde např. v [8, 52]. Podobný rozvrhový problém bychom mohli motivovat i snahou sestavit např. co nejvýkonnější sportovní tým. Řešením problému bychom pak z konkrétních charakteristik jednotlivých hráčů určili skupiny, které spolu budou na hřišti, či pořadí, v jakém se hráči budou střídat.

Příklad 2.12.3

Představme si nemocnici, ve které na pohotovosti pracuje 60 lékařů. Lékaři pracují v třísměnném provozu, směny jsou osmihodinové a pro jednoduchost předpokládejme, že směny začínají vždy o půlnoci (směna 1), v 8:00 (směna 2) a v 16:00 (směna 3). Vedení nemocnice se chystá vytvořit optimální plán směn tak, aby využilo co nejméně lékařů. Z dlouhodobého pozorování vedení zná minimální počet zaměstnanců, kteří musí být v průběhu týdne přítomni na jednotlivých směnách. Ten je uveden v Tabulce 2.71.

směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
1	9	3	4	4	3	10	12
2	7	9	14	8	7	15	15
3	7	11	10	9	11	12	13

Tabulka 2.71: Zadání Příkladu 2.12.3

Lékaři si u vedení nemocnice za pomoci odborů vydobyli následující podmínky:

(i) pracovat vždy 4 po sobě jdoucí dny v týdnu a poté 3 po sobě jdoucí dny nepracovat;

(ii) ve čtyřech po sobě následujících dnech pracovat vždy na stejné směně (např. 4x za sebou mít směnu 1).

Sestavme a vyřešme lineární úlohu, kterou řeší vedení nemocnice. Vedení chce najít dlouhodobý rozvrh, podle kterého se budou směny řídit, ne pouze rozvrh na jeden týden. Vedení tak při řešení počítá s tím, že na pondělní směně pracují lékaři, kteří začali čtyřdenní směnu minulý týden. Jejich cílem je minimalizovat množství zaměstnanců, kteří pokryjí potřebné minimální počty. Tuto úlohu bychom mohli formulovat i tak, že chceme minimalizovat náklady na platy zaměstnanců – nechceme tedy, aby v době, kdy potřebujeme 12 zaměstnanců, jich na pracovišti bylo např. 20.

Ve druhé části příkladu se podíváme na situaci, ve které jsou zaměstnanci dvou různých typů. Vzhledem k propukající epidemii hlásí nemocnice mírný nárůst počtů pacientů, a tedy i počtů lékařů potřebných na jednotlivých směnách. Tyto počty jsou uvedeny v Tabulce 2.72. Nárůst požadavků na lékaře není prozatím nijak závratný, jejich maximální počet se však snížil na 45, neboť někteří jsou kvůli pobytu v rizikových oblastech nyní uzavřeni v domácí karanténě. Místní univerzita však nemocnici nabídla studenty lékařské fakulty, kteří mohou vypomáhat. Studentů je k dispozici celkem 100. Studenty ale, na rozdíl od stálých zaměstnanců, najímá nemocnice pouze na jednotlivé směny a mohou pracovat bez dalších omezení. Jediné bezpečnostní omezení, které ne-

mocnice vyžaduje, je, aby studenti na všech směnách tvořili maximálně jednu třetinu ze všech potřebných lékařů.

směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
1	9	3	3	5	5	15	12
2	7	9	15	8	7	15	15
3	7	11	10	9	12	12	13

Tabulka 2.72: Zadání druhé části Příkladu 2.12.3.

V této části se již nebudeme snažit minimalizovat celkový počet zaměstnanců, ale celkové náklady na ně. Mzdu všech lékařů i studentů uvádíme ve stejných jednotkách: mzda jednoho zaměstnance je 42 (za 4 směny), studenti dostanou 8 za práci na směně 1 nebo směně 2 a 14 za práci na směně 3. Sestavme úlohu, která odpovídá i druhé části příkladu a vyřešme ji.

Řešení. Pro přehlednost si nejprve zavedeme značení. Značení bude stejné pro první i druhou část příkladu. Označme dny v týdnu čísly 1–7 a směny čísly 1, 2 a 3. Dále označme N_{ij} počet lékařů, kteří začínají svou práci v i-tém dni na j-té směně. Dále označme D_{ij} počty potřebných zaměstnanců v i-tém dni na j-té směně, např. máme $D_{13} = 7$ podle tabulky ze zadání. V první části příkladu chceme minimalizovat počet lékařů, kteří budou v daném týdnu pracovat. Účelová funkce je tedy

$$\sum_{i=1}^{7}\sum_{j=1}^{3} N_{ij}.$$

Prvním naším omezením je maximální možný počet lékařů, které můžeme využít, tj.

$$\sum_{i=1}^{7}\sum_{j=1}^{3} N_{ij} \leq 60.$$

Nyní sestavíme omezení pro jednotlivé směny. Lékaři, kteří začnou první směnou v pondělí, budou pracovat na první směně i v úterý, ve středu a ve čtvrtek. Ve čtvrtek tak pracují zaměstnanci, kteří začali v pondělí, v úterý, ve středu a ve čtvrtek. Pro první směnu ve čtvrtek tak máme omezení ve tvaru

$$N_{11} + N_{21} + N_{31} + N_{41} \geq D_{41} = 4.$$

Stejnou úvahou dostaneme omezení pro ostatní čtvrteční směny. Pro jednotlivé dny tímto způsobem získáme 21 omezení, která pro $j \in \{1,2,3\}$ jsou

$$\begin{aligned}
N_{1j} + N_{5j} + N_{6j} + N_{7j} &\geq D_{1j}\\
N_{1j} + N_{2j} + N_{6j} + N_{7j} &\geq D_{2j}\\
N_{1j} + N_{2j} + N_{3j} + N_{7j} &\geq D_{3j}\\
N_{1j} + N_{2j} + N_{3j} + N_{4j} &\geq D_{4j}\\
N_{2j} + N_{3j} + N_{4j} + N_{5j} &\geq D_{5j}\\
N_{3j} + N_{4j} + N_{5j} + N_{6j} &\geq D_{6j}\\
N_{4j} + N_{5j} + N_{6j} + N_{7j} &\geq D_{7j}.
\end{aligned}$$

Posledním omezením, které uvažujeme, je nezápornost všech N_{ij}. Řešení této úlohy je zobrazeno v Tabulce 2.73.

směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
1	0	0	1	3	6	0	3
2	0	0	7	6	0	2	7
3	0	3	3	5	0	4	4

Tabulka 2.73: Řešení první části Příkladu 2.12.3.

Z této tabulky vidíme například, že v pondělí na směny nenastupují žádní noví lékaři, neboť kapacitu pokryjí lékaři, kteří začaly své směny v pátek, sobotu a neděli. Sečtením všech čísel v této tabulce získáme celkový počet lékařů, které na daný týden potřebujeme – to je 54. Celočíselnost výsledku získaného simplexovou metodou je způsobena specifičností celého zadání, kde na levých stranách omezení vystupují pouze koeficienty 0 a 1, pro více podrobností viz také Kapitolu 3. V Tabulce 2.74 vidíme celkové počty lékařů, kteří jsou přítomni na jednotlivých směnách. Porovnáním této tabulky s Tabulkou 2.71 čtenář snadno ověří, že na všech směnách je dostatek lékařů.

směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
1	9	3	4	4	10	10	12
2	9	9	14	13	13	15	15
3	8	11	10	11	11	12	13

Tabulka 2.74: Řešení první části Příkladu 2.12.3 – celkové počty.

Druhou část úlohy vyřešíme obdobně. Označíme počet studentů vypomáhajících v i-tém dni na j-té směně jako A_{ij}. Účelovou funkci, jejíž minimum nyní hledáme, tak zapíšeme jako

$$42\sum_{i=1}^{7}\sum_{j=1}^{3}N_{ij}+8\sum_{i=1}^{7}(A_{i1}+A_{i2})+14\sum_{i=1}^{7}A_{i3}.$$

Omezení na maximální počet lékařů z první části příkladu se nyní změní na

$$\sum_{i=1}^{7}\sum_{j=1}^{3}N_{ij}\leq 45.$$

Změní se též omezení na minimální počty zaměstnanců, která jsou nyní pro $j\in\{1,2,3\}$ ve tvaru

$$\begin{aligned}
N_{1j}+N_{5j}+N_{6j}+N_{7j}+A_{1j}&\geq D_{1j} \quad \& \quad N_{1j}+N_{2j}+N_{6j}+N_{7j}+A_{2j}\geq D_{2j}\\
N_{1j}+N_{2j}+N_{3j}+N_{7j}+A_{3j}&\geq D_{3j} \quad \& \quad N_{1j}+N_{2j}+N_{3j}+N_{4j}+A_{4j}\geq D_{4j}\\
N_{2j}+N_{3j}+N_{4j}+N_{5j}+A_{5j}&\geq D_{5j} \quad \& \quad N_{3j}+N_{4j}+N_{5j}+N_{6j}+A_{6j}\geq D_{6j}\\
N_{4j}+N_{5j}+N_{6j}+N_{7j}+A_{7j}&\geq D_{7j}.
\end{aligned}$$

Kromě těchto omezení musíme k úloze přidat též omezení na maximální počet studentů, které můžeme v krizové situaci zaměstnat, tj.

$$\sum_{i=1}^{7}\sum_{j=1}^{3}A_{ij}\leq 100.$$

Omezení, která jsou dána bezpečnostním nařízením nemocnice o maximálním poměru studentů na počet potřebných lékařů, jsou pro $i \in \{1,\dots,7\}$ a $j \in \{1,2,3\}$ ve tvaru

$$A_{ij} \le D_{ij}/3 \quad \text{neboli} \quad 3A_{ij} \le D_{ij}.$$

Tímto k úloze přidáme dalších 21 omezení a i v této části úlohy pochopitelně požadujeme nezápornost všech proměnných. Řešení je zobrazeno v Tabulce 2.75. V první části tabulky jsou počty lékařů, kteří začínají pracovat v daný den a pracují 4 dny, ve druhé části pak jsou počty studentů, kteří pracují pouze na konkrétní směně.

	směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
lékaři	1	0	0	2	2	3	3	0
	2	0	0	5	3	0	2	5
	3	0	7	0	2	3	4	0
studenti	1	3	0	1	1	0	5	4
	2	0	2	5	0	0	5	5
	3	0	0	3	0	0	3	4

Tabulka 2.75: Řešení druhé části Příkladu 2.12.3.

V Tabulce 2.76 jsou zobrazeny počty lékařů (první sčítanec) a studentů (druhý sčítanec), kteří jsou přítomni na dané směně. Při tomto sestavování pracovního rozvrhu využijeme 41 zaměstnanců. V průběhu celého týdne také 41x využijeme pomoc studentů. Vzhledem k tomu, že na práci studentů nemáme žádné speciální požadavky (studenti mohou pracovat každý den v týdnu, více směn za den atp.) může se v krajním případě jednat o jakýkoli počet studentů větší než 5, neboť z tabulky vidíme, že nejvyšší počet studentů pracujících najednou je právě 5. Studenti vypomůžou 31x na první i druhé směně a 10x na třetí směně. Minimum nákladů, kterého dosáhneme touto strategií, je 2110. Celočíselnost výsledku je opět důsledkem velmi specifické struktury této úlohy.

směna \ den	PO	ÚT	ST	ČT	PÁ	SO	NE
1	6+3	3+0	2+1	4+1	7+0	10+5	8+4
2	7+0	7+2	10+5	8+0	8+0	10+5	10+5
3	7+0	11+0	7+3	9+0	12+0	9+3	9+4

Tabulka 2.76: Řešení druhé části Příkladu 2.12.3 – celkové počty.

Další možnou aplikací lineárního programování je optimální dělení materiálu. Úlohy tohoto typu najdeme v literatuře obvykle pod názvem *řezné problémy*.

[228] Angl. *cutting stock problem.*

Řezný problém[228]. Tyto úlohy jsou obvykle motivovány tak, že máme k dispozici nějaký materiál určité (standardizované) velikosti, který chceme rozdělit do kusů jiné velikosti. Můžeme si to představit například tak, že výrobce vyrábí dřevěná prkna či ocelové trubky v několika délkách a následně je řeže podle požadavků zákazníků. V řezných problémech pak máme např. omezenou nabídku původního materiálu, či požadovaná množství materiálů různých

velikostí, které chceme po řezání získat. Pro úlohu vždy musíme sestavit též tzv. *řezný plán*, který udává způsoby, jakými je možné materiál rozdělit v závislosti na požadovaných velikostech nařezaných kusů. Účelová funkce pak může vyjadřovat např. množství materiálu, který po řezání zbude, či náklady na jednotlivé kusy materiálu, které musíme před řezáním zakoupit. V obou těchto případech bychom hledali minimum účelové funkce.

Pro $i \in \{1, \ldots, m\}$ označme jako s_i nabídku i-té délky materiálu, kterou máme k dispozici. Je-li celkový počet různých druhů požadovaných délek řezů roven p, pak proměnná d_k pro $k \in \{1, \ldots, p\}$ vyjadřuje poptávku po k-té délce řezu. Při řezání máme k dispozici celkem n řezných plánů, takže proměnnou x_j vyjádříme, kolikrát využijeme j-tý řezný plán, zatímco proměnné z_j budou označovat zbytky materiálu z jednotlivých řezných plánů. Koeficient a_{kj} určuje, kolik kusů požadované délky řezu k získáme využitím j-tého řezného plánu. Konečně ještě budeme potřebovat indexovou množinu T_i všech řezných plánů j, které využívají i-tý materiál. S využitím tohoto značení můžeme řezný problém obecně naformulovat například následujícím způsobem

$$\sum_{j=1}^{n} z_j \to \min$$

$$\sum_{j=1}^{n} a_{ij}\, x_j \geq d_k \quad \text{pro všechna } k \in \{1, \ldots, p\},$$

$$\sum_{j \in T_i} x_j \leq s_i, \quad \text{pro všechna } i \in \{1, \ldots, m\},$$

$$x_1 \geq 0 \quad \& \quad \cdots \quad \& \quad x_n \geq 0.$$

K problému můžeme, stejně jako u problému jídelníčku, dodat i omezení na maximální výrobu jednotlivých délek řezu, označeným b_k pro $k \in \{1, \ldots, p\}$. Takové omezení by bylo ve tvaru

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_k, \qquad k \in \{1, \ldots, p\}.$$

S řeznými problémy se můžeme se setkat ve většině literatury věnované lineárnímu programování jako např. v [154, 292, 374]. Tyto problémy jsou většinou motivovány potřebou nařezat trubky, prkna nebo papírové role. V následujícím příkladu si ukážeme aplikaci motivovanou potřebou nastříhat látky na několik různých délek.

Příklad 2.12.4

V Brně a dalších českých městech se v reakci na nedostatek roušek při první jarní vlně koronavirové epidemie objevili jak dobrovolníci, kteří roušky šijí, tak dárci látek, kteří jsou schopni poskytnout potřebný materiál. Lidé, kteří mají k dispozici různé kusy bavlny, je mohou doručit přímo lidem, kteří roušky šijí. Vzniklo ale i několik skupin a míst, kam mohou dárci látek své látky poslat, aniž by se starali o jejich doručení konkrétním lidem. Lidé, kteří tyto skupiny organizují, mají za úkol doručené látky nastříhat tak, aby se každému, kdo chce šít, dostalo požadované množství kusů, ze kterých může roušky vyrobit. Navíc existuje několik způsobů šití roušek, přičemž každý z nich vyžaduje jiné rozměry použité látky. Dobrovolníci, kterým budeme látku posílat, se přihlásili o následující 4 délky jednotlivých kusů: 20 cm, 21,5 cm, 16 cm a 30 cm. Šířka všech požadovaných kusů je 20 cm.
Předpokládejme, že organizátoři skupiny již látky nastříhali na 20 cm široké pásy. V závislosti na doručených látkách jsou k dispozici pásy ve 4 různých délkách – 70 cm, 45 cm, 38 cm a 64 cm. Organizátoři mají k dispozici postupně 130, 25, 48 a 110 kusů těchto pásů. Tyto pásy nyní potřebují rozstříhat na délky, které požadují jednotliví dobrovolníci k šití. Organizátoři znají celkové poptávky po jednotlivých typech nastříhaných látek, které jsou postupně 300 ks, 250 ks, 175 ks a 125 ks.

Organizátoři akce nyní řeší, jak z látek, které mají k dispozici, nastříhat materiál na jednotlivé roušky. Jejich cílem je minimalizovat zbytky látek, které se již nebudou dát použít. Organizátoři samozřejmě chtějí pokrýt poptávku po látkách na roušky. Vzhledem k náročnosti operace mají však jeden dodatečný požadavek – z každého pruhu látky vystřihovat maximálně 2 různě dlouhé typy kusů. Nemůže se tedy stát, že pás dlouhý 70 cm rozstřihnou na 3 kusy, jeden dlouhý 20 cm, druhý 21,5 cm a třetí 16 cm. Formulujme tuto úlohu jako problém lineárního programování a najděme řešení, které minimalizuje celkový zbytek látek. Organizátory ale také ještě zajímá, jaké maximální množství látek na roušky jednotlivých rozměrů jsou schopni z dostupných látek vyrobit tak, že dodrží poptávková omezení po jednotlivých rozměrech. Zodpovězme i na tuto otázku.

Řešení. Než přistoupíme k samotné matematické formulaci úlohy, musíme si ujasnit, jaké jsou možnosti nastříhání látek o délkách 70 cm, 45 cm, 38 cm a 64 cm na látky o délkách 20 cm, 21,5 cm, 16 cm a 30 cm. Tyto střihy navíc musí splnit omezení, které organizátoři vyžadují – z jednoho kusu vyrobit maximálně 2 různé typy délek. Tyto možnosti, neboli řezné plány, si nyní postupně vypíšeme. Začneme s řeznými plány pro 70 cm dlouhý pruh látky. Nejprve si rozmysleme, jak můžeme tento pruh látky nastříhat na jednotlivé typy délek. Z tohoto pruhu můžeme získat 3 látky o délce 20 cm (se zbytkem 10 cm), 3 látky o délce 21,5 cm (zbytek 5,5 cm), 4 látky o délce 16 cm (zbytek 6 cm) nebo 2 látky o délce 30 cm (zbytek 10 cm). V dalším kroku chceme získat dostupné řezné plány, ve kterých látku o délce 70 cm rozdělíme na výrobu 2 různých délek látek na roušky. Jednou z těchto možností je například ustřihnout 3 kusy dlouhé 16 cm a jeden kus dlouhý 21,5 cm, z čehož získáme pouze 0,5 cm zbytku. Z důvodu přehlednosti si nyní uvedeme všechny řezné plány (označené čísly 1–12) pro látku dlouhou 70 cm v Tabulce 2.77.

řezné plány	1	2	3	4	5	6	7	8	9	10	11	12
20 cm	3					1	2	2	2	1		
21,5 cm		3			1		1			2	2	
16 cm			4		3	3			1		1	2
30 cm				2				1				1
zbytek	10	5,5	6	10	0,5	2	8,5	0	14	7	11	8

Tabulka 2.77: Řezné plány pro 70 cm dlouhý pruh látky.

Prvních pět řezných plánů, které jsme si popsali v předchozím odstavci, najdeme v Tabulce 2.77 pod čísly 1–5. V tabulce vidíme, že možných řezných plánů pro pruh látky dlouhý 70 cm je 12. Tabulku čteme tak, že například řezný plánem číslo 10 vyrobíme z daného pruhu látky 1 kus dlouhý 20 cm a 2 kusy dlouhé 21,5 cm se zbytkem 7 cm.
Řezné plány si obdobně sestavíme i pro látky délek 45, 38 a 64 cm. Vždy si nejprve vypíšeme plány, ze kterých stříháme látky pouze jedné délky a poté plány na stříhání látek dvou různých délek. Tímto postupem získáme sedm řezných plánů pro látku o délce 45 cm, šest řezných plánů pro látku dlouhou 38 cm a jedenáct řezných plánů pro 64 cm dlouhé látky. Všechny dostupné řezné plány pro tuto úlohu jsou zobrazeny v Tabulce 2.78. Pod čísly 1–12 jsou plány pro látku o délce 70 cm, čísla 13–19 označují plány pro látku o délce 45 cm, čísla 20–25 plány pro 38 cm dlouhé látky a čísla 26–36 plány pro látky o délce 64 cm. Při sestavování této tabulky byly záměrně vynechány některé plány, které by nebyly využity, neboť ze zbytku látky by z nich bylo možné vytvořit další roušku. Takový plán je například nastříhání 70 cm dlouhého kusu na 3 kusy

dlouhé 16 cm a vyhození zbytku, atd. Vyřazení těchto plánů úlohu mírně zjednodušuje, je však opodstatněné tím, že organizátoři provádí tuto akci pouze jednou a nechtějí si nechat kusy látek na nějaké jiné použití. V případě, že by se jednalo o opakovaný problém, by mohlo být výhodné využít např. plány, ve kterých z 70 cm dlouhé látky ustřihneme 20 cm dlouhý kus a 50 cm si necháme, čímž vlastně získáme plán beze zbytku.

řezné plány	1	2	3	4	5	6	7	8	9	10	11	12
20 cm	3					1	2	2	2	1		
21,5 cm		3			1		1			2	2	
16 cm			4		3	3			1		1	2
30 cm				2				1				1
zbytek	10	5,5	6	10	0,5	2	8,5	0	14	7	11	8
Řezné plány	13	14	15	16	17	18	19	20	21	22	23	24
20 cm	2				1	1		1				1
21,5 cm		2			1		1		1			
16 cm			2			1	1			2		1
30 cm				1							1	
zbytek	5	2	13	15	3,5	9	7,5	18	16,5	6	8	2
Řezné plány	25	26	27	28	29	30	31	32	33	34	35	36
20 cm		3				2	2	1		1		
21,5 cm	1		2			1		2	2		1	
16 cm	1			4			1		1	2	2	2
30 cm					2							1
zbytek	0,5	4	21	0	4	2,5	8	1	5	12	10,5	2

Tabulka 2.78: Řezné plány pro Příklad 2.12.4.

Nyní víme, jaké všechny řezné plány můžeme využít. Můžeme tak již matematicky naformulovat problém, který organizátoři řeší. Označme x_j počet využití j-tého plánu. Poptávková omezení, která chceme splnit, získáme pro jednotlivé kusy látek z řádků Tabulky 2.78. Tato omezení jsou postupně

$$\begin{aligned}
3x_1 + x_6 + 2x_7 + 2x_8 + 2x_9 + x_{10} + 2x_{13} + x_{17} + x_{18} + x_{20} + \\
+ x_{24} + 3x_{26} + 2x_{30} + 2x_{31} + x_{32} + x_{34} \geq 300 \\
3x_2 + x_5 + x_7 + 2x_{10} + 2x_{11} + 2x_{14} + x_{17} + x_{19} + x_{21} + \\
+ x_{25} + 2x_{27} + x_{30} + 2x_{32} + 2x_{33} + x_{35} \geq 250 \\
4x_3 + 3x_5 + 3x_6 + x_9 + x_{11} + 2x_{12} + 2x_{15} + x_{18} + x_{19} + 2x_{22} + x_{24} + x_{25} + \\
+ 4x_{28} + x_{31} + x_{33} + 2x_{34} + 2x_{35} + 2x_{36} \geq 175 \\
2x_4 + x_8 + x_{12} + x_{16} + x_{23} + 2x_{29} + x_{36} \geq 125.
\end{aligned}$$

Látky pro materiál na roušky jsou k dispozici pouze v omezeném množství, což můžeme s využitím proměnných x_j formulovat jako

$$\sum_{j=1}^{12} x_j \leq 130 \quad \& \quad \sum_{j=13}^{19} x_j \leq 25 \quad \& \quad \sum_{j=20}^{25} x_j \leq 48 \quad \& \quad \sum_{j=26}^{36} x_j \leq 110.$$

První z uvedených omezení říká, že plány 1–12 pro látky dlouhé 70 cm můžeme využít maximálně 130x, neboť máme k dispozici pouze 130 kusů dané látky. Další omezení mají stejný význam pro plány 13–19, které se týkají látky dlouhé 45 cm, atd. Poslední omezení, které k formulaci přidáme, je omezení na nezápornost všech proměnných x_j. Zbytky z provedení jednotlivých plánů (v centimetrech látky) jsou uvedeny v Tabulce 2.78. Účelová funkce je součtem všech těchto zbytků pro provedené plány. Označíme-li písmeny $z_1, \ldots, z_{36}$ zbytky, které získáme provedením jednotlivých řezných plánů, můžeme účelovou funkci zapsat ve tvaru

$$\sum_{j=1}^{36} z_j\, x_j.$$

Nyní již můžeme úlohu vyřešit simplexovou metodou. Tou získáme řešení, ve kterém využijeme pouze plány označené čísly 5, 8, 14, 25, 28 a 32, ostatní plány nevyužijeme (tj. ostatní x_i jsou rovny nule). Pro využité plány získáme následující počty

$$x_5 = 5 \quad \& \quad x_8 = 125 \quad \& \quad x_{14} = 16{,}5 \quad \& \quad x_{25} = 48 \quad \& \quad x_{28} = 28 \quad \& \quad x_{32} = 82.$$

Tato čísla nám říkají, kolikrát daný plán využijeme. První, čeho si můžeme všimnout, je, že výsledný bod obsahuje pouze 6 nenulových proměnných, a je tedy degenerovaný. Setkáváme se tak s praktickým problémem, který vede na degenerovanou úlohu. Dále vidíme, že všechny hodnoty, až na proměnnou x_{14}, jsou celočíselné. Zkontrolujeme-li splněnost omezení, uvidíme, že všechna poptávková omezení kromě prvního se realizují jako rovnosti. Totéž platí pro nabídková omezení, s výjimkou druhého omezení pro látku o délce 45 cm. To znamená, že při této strategii nevyužijeme všechny dostupné látky o délce 45 cm, v nalezeném bodě jich využijeme pouze 16,5. Plán x_{14} tak můžeme udělat 17x, neboť máme k dispozici látku dlouhou 45 cm. Zaokrouhlením x_{14} na hodnotu 17 získáme v tomto případě řešení, které bychom obdrželi s využitím metod pro celočíselné programování. Celkové počty nastříhaných dílků jsou v případě provedení optimálního plánu následující: 332 ks látek dlouhých 20 cm, 250 ks látek dlouhých 21,5 cm, 175 látek o délce 16 cm a 125 kusů délky 30 cm. Z těchto hodnot vidíme, že v případě roušky dlouhé 20 cm vyrobíme při optimálním stříhání více než požadovaný počet. Poptávky po ostatních délkách látek jsou splněny přesně. Provedením tohoto plánu tak získáme látky na 882 roušek. V optimálním případě nám ze střihů zbude pouze 141,5 cm zbytků látky a 8,5 nevyužitých látek o délce 45 cm. Pro zajímavost zde ještě uveďme, že v případě apriorního požadavku celočíselnosti jednotlivých proměnných bychom získali 251 látek dlouhých 21,5 cm, vyrobili látky na 883 roušek, vytvořili 142,5 cm zbytků látky a nevyužili jen 8 látek o délce 45 cm.

Ve druhé části příkladu bychom chtěli najít maximální množství látek na roušky, které můžeme získat z dostupných látek. V této části řešíme problém, který má stejná omezení, neboť organizátoři chtějí splnit všechny poptávky a množství dostupných látek jsou stále stejná. Jediné, co se změní, je účelová funkce. Označme a_{ij} koeficienty z Tabulky 2.78, přičemž prázdné buňky odpovídají koeficientu nula. Účelovou funkci, jež chceme nyní maximalizovat, pak můžeme zapsat ve tvaru

$$\sum_{i=1}^{4} \sum_{j=1}^{36} a_{ij}\, x_j.$$

Řešením této úlohy jsou následující počty látek jednotlivých délek: 300 ks, 250 ks, 237,2 ks a 125 ks. Těchto počtů dosáhneme využitím plánů označených čísly 5, 8, 14, 25, 28, 29 a 32, a to v následujícím počtu opakování

$$x_5 = 14{,}4 \quad \& \quad x_8 = 115{,}6 \quad \& \quad x_{14} = 25 \quad \& \quad x_{25} = 48$$
$$x_{28} = 36{,}5 \quad \& \quad x_{29} = 4{,}7 \quad \& \quad x_{32} = 68{,}8.$$

Čtenář si jistě všimne, že i v tomto případě se jedná o degenerovaný bod. Zbytek je v tomto případě necelých 170 cm látky. V tomto případě jsme schopni vyrobit látku na 912,2 roušek. Vzhledem k tomu, že plány nejde využít neceločíselně-krát, by se toto množství přidáním požadavku celočíselnosti snížilo, a to na 912. I řešení získané pomocí simplexové metody nám ale dává dostatečně přesnou informaci o tom, jaké řezné plány bychom měli využít, a kolik roušek jsme schopni z dostupných látek vyrobit.

Organizátoři se z jasných důvodů nakonec rozhodnou nezvolit plán, který minimalizuje zbytek, ale plán, který umožní vyrobit co největší počet roušek. Na problém šití roušek a jejich rozvozu lidem, kteří je potřebují, se ještě později podíváme v Příkladu 2.12.5.

V uvedeném řezném problému jsme viděli, že nejobtížnější část příkladu nemusí být jeho matematická formulace či řešení simplexovou metodou. Nejobtížnější částí mnoha řezných problémů je formulace jednotlivých řezných plánů, které můžeme využít. V uvedené úloze jsme měli pouze 4 délky látek, ze kterých jsme chtěli stříhat 4 konkrétní délky, a to navíc s omezením ohledně maximálního počtu různých délek vytvořených z jednoho kusu látky. Délky dostupných látek v poměru k požadovaným rozměrům také nebyly nijak velké. I přesto jsme v úloze měli 36 dostupných řezných plánů. Ve většině praktických úloh je však počet řezných plánů velmi velký a tyto problémy mohou být i vícedimenzionální. V případě stříhání roušek bychom např. mohli požadovat jak různou délku, tak různou šířku látek. Formulace dvourozměrného řezného problému by ale z matematického hlediska byla stále stejná. Ovšem o mnoho složitější by bylo najít jednotlivé řezné plány, a to i v případě střihu pravidelných tvarů. Příklad dvourozměrného řezného problému čtenář najde např. v [154]. Speciálním typem dvourozměrného řezného problému je tzv. *gilotinový problém*, ve kterém hledáme způsob, jak z obdélníku za pomoci rovnoběžných a kolmých řezů vyřezat obdélníky zadané velikosti s co nejmenším odpadem. Tento problém je obvykle motivován řezáním tabulek skla a čtenář se o něm dozví více např. v [195].

Když už máme látky nařezané, tak by bylo dobré z nich ušít roušky a ty následně distribuovat lidem. Do toho se pustíme v následujícím příkladu, který navazuje na předchozí příklad a k jehož řešení využijeme nástroje z Podkapitoly 2.11.

Příklad 2.12.5

Organizátoři z předchozího příkladu nyní stojí před dalším problémem, když potřebují z 912 látek co nejrychleji ušít roušky a rozvézt je lidem, kteří je nutně potřebují. Předpokládejme pro jednoduchost, že dobrovolníci, kteří budou roušky šít, si poradí s různými tvary látek, neboli všechny kusy látek nyní považujeme za stejné. Organizátoři mají po Brně celkem 10 míst, ve kterých šijí roušky. Tato místa se liší jak rychlostí šití a následné dopravy k potřebným, tak maximálním množstvím roušek, které mohou ušít. V těchto místech nyní potřebují organizátoři ušít roušky a rozvést je do 18 míst, jako jsou např. domovy důchodců. Organizátoři chtějí provést úkol co nejefektivněji a dopravit do všech míst roušky co nejrychleji. Jejich cílem je najít takový dopravní plán, který minimalizuje celkové časové náklady na šití a následný rozvoz. Všechny potřebné údaje jsou uvedeny v Tabulce 2.79. Domovy důchodců a další zájemci o roušky jsou označeni písmenem „D“, organizátoři písmenem „O“. Pomozme organizátorům s vyřešením tohoto dopravního problému.

Řešení. Buňky budeme pro větší přehlednost označovat symboly O_i D_j. Řešení zahájíme hledáním výchozího BPB pomocí metody minimální ceny[229], viz Tabulku 2.80. V této tabulce jsou již splněny všechny poptávky a vyčerpány všechny nabídky. Nalezli jsme tak BPB. Bod je nedegenerovaný, neboť je v tabulce vyplněno $27 = 18 + 10 - 1$ bu-

[229] To zahájíme buňkou O_3 D_{12} s cenou rovnou jedné, do které doplníme číslo 48, čímž naplníme poptávku dobrovolníka D_{12}. Další buňkou s cenou jedna, kterou vyplníme, je O_4 D_9, do které vepíšeme číslo 29, čímž naplníme poptávku D_9. Pokračujeme dalšími buňkami s cenou 1, a to s buňkou O_5 D_{17}, kterou vyplníme číslem 28, čímž naplníme poptávku D_{17}, a buňkou O_{10} D_8, do které doplníme 19, čímž vyčerpáme nabídku O_{10}.

Další buňky, které vyplníme, jsou ty s cenou rovnou 2. Postupně jde o buňky O_8 D_{13}, jejímž doplněním číslem 105 vyčerpáme nabídku O_8, buňku O_3 D_{14}, do které doplníme číslo 45, čímž naplníme poptávku D_{14} a buňku O_5 D_{18}, do níž vepíšeme 30, čímž naplníme poptávku D_{18}. V dalších krocích postupujeme obdobně, a to přes buňku O_9 D_3, kterou doplníme číslem 24 a vyčerpáme tak poptávku D_3 a buňku O_4 D_{10}, do níž vepíšeme číslo 14 a splníme tím poptávku D_{10}.

Pokračujeme s dostupnými buňkami, které mají dopravní náklady rovny 3. Do buňky O_6 D_8 doplníme číslo 59, čímž splníme poptávku D_8, do buňky O_5 D_1 v dalším kroku vepíšeme 40, čímž naplníme poptávku dobrovolníka D_1. Nejmenší dostupné náklady jsou nyní rovny 4. Postupně doplňujeme buňku O_4 D_{13}, do které vepíšeme 42, čímž vyčerpáme nabídku O_4, buňku O_9 D_4, do které doplníme 41 a vyčerpáme tak nabídku O_9, buňku O_7 D_2, kterou vyplníme číslem 32, čímž vyčerpáme nabídku O_7, a buňku O_1 D_{11}, jejímž doplněním číslem 21 naplníme poptávku dobrovolníka D_{11}.

Nyní jsou nejnižší dostupné náklady rovny 5, a to v buňce O_5 D_{16}, kterou vyplníme číslem 22, čímž vyčerpáme nabídku O_5, a v buňce O_6 D_7, do které vepíšeme 9 a vyčerpáme tak nabídku O_6. Následně najdeme dostupné buňky s nejnižšími náklady, které jsou nyní 6. Těmi jsou buňky: O_2 D_{15}, jež doplníme číslem 76, čímž naplníme poptávku D_{15}, O_1 D_6, kterou vyplníme číslem 42 a naplníme tak poptávku D_6, O_2 D_7, jejímž doplněním číslem 27 pokryjeme poptávku D_7, a O_2 D_{13}, do které vepíšeme 18, čímž naplníme poptávku D_{13}.

V dalším kroku jsou nejnižší dostupné náklady 7, a to v buňce O_3 D_{16}, do které dopíšeme 41 a naplníme tak poptávku D_{16}. Buňka s nejnižšími dostupnými náklady 8 je nyní O_3 D_5, do které vepíšeme 6, čímž vyčerpáme nabídku O_3.

V následujícím kroku doplníme číslem 24 buňku O_1 D_5 s náklady 9, čímž naplníme poptávku D_5. Nyní doplněním čísla 4 do buňky O_2 D_4 s náklady 10 vyčerpáme nabídku O_2. Nejnižší dostupné náklady 13 se v tomto kroku nachází v buňce O_1 D_2, doplněním 21 do této buňky pokryjeme poptávku D_2. V posledním kroku metody doplníme číslo 45 do buňky O_1 D_4 s náklady 14.

něk. Splnění všech omezení můžeme s využitím tabulky snadno ověřit. Celkové časové náklady v tomto bodě jsou 4 028. I přes rozměry této úlohy jsme byli schopni rychle nalézt výchozí bazický přípustný bod metodou minimální ceny. V porovnání se sestavováním simplexových tabulek se jedná o opravdové usnadnění.

	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}	D_{12}	D_{13}	D_{14}	D_{15}	D_{16}	D_{17}	D_{18}	nabídka
O_1	11	13	17	14	9	6	12	8	9	3	4	11	31	12	7	8	10	14	153
O_2	16	18	14	10	10	11	6	5	6	19	8	17	6	5	6	9	11	12	125
O_3	21	24	13	10	8	8	9	6	13	14	6	1	28	2	17	7	3	26	140
O_4	6	7	5	6	11	22	31	4	1	2	18	14	4	19	7	7	8	8	85
O_5	3	6	5	7	17	8	12	27	12	16	6	7	9	15	6	5	1	2	120
O_6	5	25	6	19	18	24	5	3	11	25	22	1	6	17	8	19	3	4	68
O_7	13	4	16	4	7	11	8	21	4	2	14	19	6	7	15	6	5	4	32
O_8	5	14	5	6	27	6	25	5	2	6	17	6	2	6	6	19	2	8	105
O_9	11	7	2	4	4	30	28	17	26	8	7	24	25	13	7	18	7	9	65
O_{10}	6	17	6	7	11	22	21	1	18	3	6	5	12	7	6	26	3	4	19
poptávka	40	53	24	90	30	42	36	78	29	14	21	48	165	45	76	63	28	30	912

Tabulka 2.79: Zadání Příkladu 2.12.5.

[230] V prvním kroku nejprve vypočteme hodnoty duálních proměnných. V případě, že položíme u_1 rovno nule, jsou tyto hodnoty

$$\begin{aligned} u_1 = 0 \quad &\& \quad u_2 = -4 \\ u_3 = -1 \quad &\& \quad u_4 = -6 \\ u_5 = -3 \quad &\& \quad u_6 = -5 \\ u_7 = -9 \quad &\& \quad u_8 = -8 \\ u_9 = -10 \quad &\& \quad u_{10} = -7 \\ v_1 = 6 \quad &\& \quad v_2 = 13 \\ v_3 = 12 \quad &\& \quad v_4 = 14 \\ v_5 = 9 \quad &\& \quad v_6 = 6 \\ v_7 = 10 \quad &\& \quad v_8 = 8 \\ v_9 = 7 \quad &\& \quad v_{10} = 8 \\ v_{11} = 4 \quad &\& \quad v_{12} = 2 \\ v_{13} = 10 \quad &\& \quad v_{14} = 3 \\ v_{15} = 10 \quad &\& \quad v_{16} = 8 \\ v_{17} = 4 \quad &\& \quad v_{18} = 5. \end{aligned}$$

Nyní musíme vypočítat hodnoty nákladů příležitosti $s_{i,j}$, podrobnější řešení této části ponecháme na čtenáři. Při výpočtu získáme 12 záporných hodnot, které jsou

$$\begin{aligned} s_{1,10} = -5 \quad &\& \quad s_{1,15} = -3 \\ s_{3,4} = -3 \quad &\& \quad s_{3,8} = -1 \\ s_{4,3} = -1 \quad &\& \quad s_{4,4} = -2 \\ s_{5,2} = -4 \quad &\& \quad s_{5,3} = -4 \\ s_{5,4} = -4 \quad &\& \quad s_{5,15} = -1 \\ s_{6,3} = -1 \quad &\& \quad s_{7,4} = -1. \end{aligned}$$

Výchozí BPB tak není optimální. Největší záporná hodnota je $s_{1,10} = -5$. Obdobně jako v předchozích příkladech musíme najít uzavřenou cestu, která je nyní

$$\begin{aligned} &O_1D_{10} \to O_4D_{10} \to \\ &\to O_4D_{13} \to O_2D_{13} \to \\ &\to O_2D_4 \to O_1D_4 \to O_1D_{10}. \end{aligned}$$

Do buňky O_1 D_{10} můžeme alokovat maximálně hodnotu 14, z báze v tomto kroku odchází proměnná O_4 D_{10} . Přepočteme-li nyní tabulku, získáme bod, který je zobrazen v Tabulce 2.81. Tímto krokem jsme snížili celkový čas o $14 \cdot 5 = 70$ na hodnotu 3 958, o čemž se čtenář snadno přesvědčí v Tabulce 2.81.

	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}	D_{12}	D_{13}	D_{14}	D_{15}	D_{16}	D_{17}	D_{18}	nabídka
O_1	11	13	17	14	9	6	12	8	9	3	4	11	31	12	7	8	10	14	153
		21		45	24	42					21								
O_2	16	18	14	10	10	11	6	5	6	19	8	17	6	5	6	9	11	12	125
				4			27						18		76				
O_3	21	24	13	10	8	8	9	6	13	14	6	1	28	2	17	7	3	26	140
					6							48		45		41			
O_4	6	7	5	6	11	22	31	4	1	2	18	14	4	19	7	7	8	8	85
									29	14			42						
O_5	3	6	5	7	17	8	12	27	12	16	6	7	9	15	6	5	1	2	120
	40														22		28	30	
O_6	5	25	6	19	18	24	5	3	11	25	22	1	6	17	8	19	3	4	68
							9	59											
O_7	13	4	16	4	7	11	8	21	4	2	14	19	6	7	15	6	5	4	32
		32																	
O_8	5	14	5	6	27	6	25	5	2	6	17	6	2	6	6	19	2	8	105
													105						
O_9	11	7	2	4	4	30	28	17	26	8	7	24	25	13	7	18	7	9	65
			24	41															
O_{10}	6	17	6	7	11	22	21	1	18	3	6	5	12	7	6	26	3	4	19
poptávka	40	53	24	90	30	42	36	78	29	14	21	48	165	45	76	63	28	30	912

Tabulka 2.80: Výchozí BPB pro Příklad 2.12.5.

Při řešení úlohy modifikovanou distribuční metodou musíme v každém kroku vypočítat hodnoty všech duálních proměnných, kterých je 28 a hodnoty všech rozdílů s_{ij},

kterých je v tomto příkladu 153. Pro ilustraci zde ukážeme pouze výsledek prvního kroku[230], viz Tabulku 2.81. Další postup modifikované distribuční metody zde však uvádět nebudeme.

	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}	D_{12}	D_{13}	D_{14}	D_{15}	D_{16}	D_{17}	D_{18}	nabídka
O_1	11	13	17	14	9	6	12	8	9	3	4	11	31	12	7	8	10	14	153
		21		31	24	42				14	21								
O_2	16	18	14	10	10	11	6	5	6	19	8	17	6	5	9	11	12	4	125
				18			27						4		76				
O_3	21	24	13	10	8	8	9	6	13	14	6	1	28	2	17	7	3	26	140
					6							48		45		41			
O_4	6	7	5	6	11	22	31	4	1	2	18	14	4	19	7	7	8	8	85
									29				56						
O_5	3	6	5	7	17	8	12	27	12	16	6	7	9	15	6	5	1	2	120
	40															22	28	30	
O_6	5	25	6	19	18	24	5	3	11	25	22	1	6	17	8	19	3	4	68
							9	59											
O_7	13	4	16	4	7	11	8	21	4	2	14	19	6	7	15	6	5	4	32
		32																	
O_8	5	14	5	6	27	6	25	5	2	6	17	6	2	6	6	19	2	8	105
													105						
O_9	11	7	2	4	4	30	28	17	26	8	7	24	25	13	7	18	7	9	65
			24	41															
O_{10}	6	17	6	7	11	22	21	1	18	3	6	5	12	7	6	26	3	4	19
								19											
poptávka	40	53	24	90	30	42	36	78	29	14	21	48	165	45	76	63	28	30	912

Tabulka 2.81: První krok modifikované distribuční metody pro Příklad 2.12.5.

Vzhledem k velikosti úlohy zde nyní uvedeme řešení, které bychom získali po několika krocích modifikované distribuční metody. Tím je bod, v němž jsou celkové časové náklady rovny 3 683. Výsledný bod uvedený v Tabulce 2.82 má 15 bazických proměnných se stejnými hodnotami jako výchozí bod. Tento bod má dále 7 stejných bazických

proměnných jako výchozí bod, které se ovšem liší svými hodnotami (O_2 D_{13}, O_2 D_{15}, O_3 D_5, O_3 D_{16}, O_4 D_{13}, O_5 D_{16} a O_5 D_{17}), a 5 nových proměnných, které ve výchozím bodě nejsou v bázi (O_1 D_{10}, O_1 D_{15}, O_1 D_{16}, O_4 D_4 a O_5 D_2, které v bázi nahradily proměnné O_1 D_2, O_1 D_4, O_1 D_5, O_2 D_4 a O_4 D_{10}). Stejně jako výchozí bod je i tento bod nedegenerovaný.

	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_{11}	D_{12}	D_{13}	D_{14}	D_{15}	D_{16}	D_{17}	D_{18}	nabídka
O_1	11	13	17	14	9	6	12	14	9	3	4	11	31	12	7	8	10	14	153
						42									31	45			
O_2	16	18	14	10	10	11	6	12	6	19	8	17	6	5	6	9	11	8	125
							27						53		45				
O_3	21	24	13	10	8	8	9	26	13	14	6	1	28	2	17	7	3	26	140
					30							48		45			17		
O_4	6	7	5	6	11	22	31	8	1	2	18	14	4	19	7	7	8	2	85
				49					29				7						
O_5	3	6	5	7	17	8	12	27	12	16	6	7	9	15	6	5	1	30	120
	40	21														18	11		
O_6	5	25	6	19	18	24	5	3	11	25	22	1	6	17	8	19	3	4	68
							9	59											
O_7	13	4	16	4	7	11	8	21	4	2	14	19	6	7	15	6	5	4	32
		32																	
O_8	5	14	5	6	27	6	25	5	2	6	17	6	2	6	6	19	2	8	105
													105						
O_9	11	7	2	4	4	30	28	17	26	8	7	24	25	13	7	18	7	9	65
			24	41															
O_{10}	6	17	6	7	11	22	21	1	18	3	6	5	12	7	6	26	3	4	19
								19											
poptávka	40	53	24	90	30	42	36	78	29	14	21	48	165	45	76	63	28	30	912

Tabulka 2.82: Řešení Příkladu 2.12.5.

V uvedeném příkladu jsme si ukázali, jak postupovat, když považujeme nastříhané látky za homogenní. Kdybychom tento předpoklad upustili, získali bychom čtyři dopravní problémy, jeden pro každý druh nastříhané látky. Postup řešení by však byl opět stejný. Na závěr se ještě vypravíme do banky pro peníze.

Investice – tvorba portfolia. Již jsme viděli, že lineární programování můžeme využít i při plánování investic neboli tvorbě portfolia. Obecně lze problém sestavení portfolia formulovat například následujícím způsobem. Představme si, že máme n investičních příležitostí a známe

jejich výnosy $R_1, \ldots, R_n$. Výnos R_j je obecně náhodná veličina a při tvorbě portfolia se budeme zajímat např. o její střední hodnotu (ozn. $\mathbf{E}(R_j)$). Jsme ochotni investovat finanční prostředky ve výši M. Označíme-li x_j množství peněz vložených do j-té investice, můžeme střední hodnotu zisku (nebo též tzv. *očekávaný výnos*) zapsat ve tvaru

$$\mathbf{E}(\pi) = \sum_{j=1}^{n} x_j \mathbf{E}(R_j).$$

V případě, že by nás zajímal pouze očekávaný výnos, vyřešili bychom problém vyřešili snadno tím, že bychom investovali všechny dostupné prostředky do investice s největším výnosem. Velmi výnosné investice s sebou však obvykle nesou velkou míru rizika (označeného s_j), které můžeme definovat např. pomocí střední hodnoty absolutní odchylky od průměru jako

$$s_j = \mathbf{E}|x_j R_j - x_j \mathbf{E}(R_j)|.$$

Při vytváření portfolia musíme brát v úvahu oba cíle, kterými jsou maximalizace zisku a minimalizace rizika. Jedním ze způsobů, které uvádí např. [491, Chapter 13], je využití parametru μ, jehož přidáním do účelové funkce nastavíme, jak důležitý je pro nás výnos oproti riziku. Účelová funkce proto může být např. ve tvaru

$$\mu \sum_{j=1}^{n} x_j \mathbf{E}(R_j) - \sum_{j=1}^{n} s_j.$$

Pro velké hodnoty μ je pro nás očekávaný výnos mnohem důležitější než riziko, zatímco pro malé hodnoty μ je pro nás důležité především malé riziko. K úloze bychom samozřejmě přidali omezení na maximální hodnotu investovaných prostředků, tj.

$$\sum_{j=1}^{n} x_j \leq M.$$

V literatuře se můžeme setkat i s přístupem, ve kterém x_j označuje podíl celkové částky, kterou investujeme do j-té investice. V tomto případě je předchozí omezení ve tvaru

$$\sum_{j=1}^{n} x_j \leq 1 \quad \text{případně} \quad \sum_{j=1}^{n} x_j = 1.$$

Rovnostní omezení můžeme využít např. v případě, kdy je jedna z investičních možností bez rizika, což si můžeme představit např. jako uložení prostředků do banky. Tuto obecnou formulaci problému sestavování portfolia si nyní ukážeme na příkladu, ve kterém ale budeme pro jednoduchost předpokládat, že výnos není náhodná veličina. V tomto příkladu budeme riziko měřit tzv. *mírou defaultu*, tj. pravděpodobností, že klient, kterému půjčíme peníze, nám je nebude schopen splatit. V následujícím příkladu věnovaném bankovnímu rozhodování o spotřebitelských úvěrech budeme s rizikem pracovat třemi různými způsoby.

Příklad 2.12.6

V tomto příkladu se podíváme na problém, před kterým stojí vedení banky. Banka má momentálně k dispozici 50 milionů Kč, které chce buď investovat do státních dluhopisů nebo vypůjčit klientům, jejichž poptávka po spotřebitelských úvěrech je v současné době velmi vysoká. Investování do státních dluhopisů je v podstatě bezrizikové, úvěrování je však pro banku mnohem výnosnější. Vedení banky chce maximalizovat zisk, ovšem za předpokladu, že nebudou muset podstoupit příliš velké riziko.
Banka svoje klienty řadí podle jejich charakteristik do osmi skupin označených písmeny A–H. V kategorii označené písmenem A jsou pouze nejlepší klienti, se kterými má banka již zkušenost, a ví, že s pravděpodobností 0,999 svůj úvěr splatí i s úroky. Pravděpodobnost, že klient nebude schopen svoji půjčku splatit, zde označíme mírou defaultu. Ze zkušenosti zná vedení také pravděpodobnost toho, že klient, který není

schopen splácet, vrátí bance polovinu hodnoty svého úvěru (bez úroků). Banka zároveň ví, že klienti ve skupině A snadno získají úvěr i u jiné banky, neboť jsou téměř bezrizikoví. Musí jim tak nabídnout nízké úroky, aby si vybrali její produkt. Banka má již vypracovanou strategii, podle které nabídne jednotlivým kategoriím úrokové sazby. Stupnici s úroky, mírou defaultu a pravděpodobností navrácení poloviny úvěru v případě defaultu pro jednotlivé skupiny klientů zobrazuje Tabulka 2.83.

kategorie	A	B	C	D	E	F	G	H	2T repo
úrok [%]	2,9	3,9	4,9	6,9	7,5	10,9	12,9	16,9	1
míra defaultu	0,001	0,009	0,04	0,085	0,1	0,12	0,18	0,2	0
pravd. splacení	0,999	0,991	0,96	0,915	0,9	0,88	0,82	0,8	1
pravd. polovina	0,75	0,75	0,5	0,5	0,5	0,3	0,3	0,25	0

Tabulka 2.83: Rozdělení klientů podle rizikovosti.

V posledním řádku Tabulky 2.83 vidíme, jaká je pravděpodobnost, že klienti, kteří nebudou schopni splácet, vrátí bance polovinu vypůjčené částky bez úroků. Banka tak z těchto klientů ztratí pouze polovinu hodnoty poskytnutého úvěru a úroky. Kromě osmi kategorií klientů je v tabulce též sloupec označený „2T repo“, což je úroková sazba, za kterou si mohou banky na 2 týdny uložit peníze u ČNB. Tento způsob investování je bezrizikový, takže pravděpodobnost splacení je rovna jedné. Stejně jako úrokové sazby pro klienty je i 2T repo sazba roční, v Tabulce 2.83 je uvedena sazba platná k 27. březnu 2020. Tato sazba se v průběhu roku může změnit, avšak v našem příkladu pro jednoduchost považujeme sazbu za neměnnou po celý rok. Banka by tedy své přebytečné prostředky uložila u ČNB na celý rok.

Vedení banky chce maximalizovat svůj zisk. Vzhledem ke striktním pravidlům bankovního systému však nemůže vypůjčit své peníze úplně libovolně. Banka musí kvůli rizikovosti alespoň třetinu hodnoty svých úvěrů vypůjčit klientům z kategorie A a B. Ze stejného důvodu může banka do kategorií F, G a H vypůjčit maximálně polovinu toho, co zapůjčí do kategorií A, B a C. Banka si též váží svých klientů z kategorie A, se kterými má předběžně domluvené úvěry o hodnotě 800 000 Kč. Klientům v kategorii A chce tak vedení nabídnou úvěry minimálně v této hodnotě.

Vzhledem k tomu, že se vedení rozhoduje o hodnotách úvěrů pro jednotlivé kategorie, počítá se ztrátou jako s částkou, kterou jednotlivé kategorie nesplatí, tj. se součinem hodnoty úvěrů v dané kategorii a míry defaultu, od kterého odečte výši částky zaplacenou nesplácejícími klienty. Vedení banky chce uvážit tři situace, které se liší pojetím rizika. Ve všech situacích se však zajímá o úrokový výnos z prvního roku splácení. V první z nich chce maximalizovat výnos (z úroků a návratů částí úvěrů od klientů v defaultu) s tím, že omezí maximální přípustnou ztrátu na částku 2 500 000 Kč. Ve druhé situaci se pak chce podívat na zisk jako rozdíl výnosů a ztráty. Konečně ve třetí situaci je hodnota výnosu dvakrát důležitější než riziko. Jaké pojetí rizika přinese bance větší zisk? Jak se situace změní, jestliže se vedení rozhodne kvůli solidaritě a dobré reputaci vypůjčit alespoň 5 000 000 Kč lidem z kategorií F, G a H? Sestavme a vyřešme úlohu lineárního programování, před kterou stojí vedení banky.

Řešení. Označme $x_1, \ldots, x_8$ prostředky investované do kategorií A–H. Celkovou sumu, kterou banka investuje do bezrizikových dluhopisů za 2T repo sazbu poté vypočítáme jako zbývající částku, kterou neivestuje do úvěrů. Striktní pravidla bankovního systému, kterým banka čelí, můžeme zapsat jako

$$x_1 + x_2 \geq 1/3 \sum_{i=1}^{8} x_i \quad \& \quad x_6 + x_7 + x_8 \leq (x_1 + x_2 + x_3)/2.$$

Banka dále čelí omezení na celkovou výši svých úvěrů, kterým je

$$\sum_{i=1}^{8} x_i \leq 50 \cdot 10^6.$$

Poslední omezení, jež chce banka splnit, je omezení o úvěrech pro klienty z kategorie A, které musí být alespoň ve výši 800 000 Kč, tj. musí platit

$$x_1 \geq 800\,000.$$

Pro všechna $x_1, \ldots, x_8$ platí též omezení nezápornosti. Všechna uvedená omezení jsou společná pro situace, které banka uvažuje.
Jednotlivé situace se však liší ve formulaci účelové funkce či dalších omezeních. V první situaci se banka zajímá pouze o výnos z úroků a navrácenou částku od některých nesplácejících klientů. Označme r_i úrokové míry jednotlivých kategorií, p_i pravděpodobnosti splacení úvěru a h_i pravděpodobnost, že klient nezvládne splácet, ale vrátí bance polovinu zapůjčené částky. Úrokový výnos za první rok splácení, který chce banka maximalizovat, pak v této situaci můžeme zapsat ve tvaru

$$\sum_{i=1}^{8} p_i \cdot x_i \cdot r_i/100 + \sum_{i=1}^{8} (1-p_i) \cdot h_i \cdot x_i/2,$$

kde první suma představuje úrokový výnos a druhá součet navrácené částky od klientů, kteří svůj úvěr nejsou schopni splácet. V této situaci je banka ochotná podstoupit maximálně riziko ztráty 2,5 milionu. Z tohoto důvodu k úloze přidáme omezení

$$\sum_{i=1}^{8} (1-p_i) \cdot x_i - \sum_{i=1}^{8} (1-p_i) \cdot h_i \cdot x_i/2 \leq 2{,}5 \cdot 10^6.$$

První suma v tomto omezení reprezentuje ztrátu kvůli nesplácejícím klientům, z nichž někteří nám však nám nějakou část vrátí, což je vyjádřeno ve druhé sumě. Nyní již máme všechna potřebná omezení a účelovou funkci. Vypočteme-li tuto úlohu lineárního programování, získáme výsledek, který je po zaokrouhlení na dvě desetinná místa uveden v Tabulce 2.84. V řádku označeném vrácení jsou uvedeny hodnoty, které bance vrátí část defaultních klientů.

kategorie	A	B	C	D	E	F	G	H	2T repo
x_i	800000	22967521,37	9565811,97	0	0	0	0	16666666,67	0
úroky	23176,8	887671,73	449975,79	0	0	0	0	2253333,33	0
vrácení	300	77515,38	95658,12	0	0	0	0	416666,67	0
ztráta	500	129192,31	286974,36	0	0	0	0	2083333,33	0

Tabulka 2.84: Výsledky pro první situaci v Příkladu 2.12.6.

Banka se v tomto případě soustředí na klienty z kategorií A, B, C a H. Očekávaný výnos, který banka tímto způsobem získá, dostaneme sečtením řádků pro úroky a vrácení z Tabulky 2.84. Tento výnos je po zaokrouhlení 4 204 298 Kč. Z tabulky zároveň vidíme, že v tomto případě banka čelí velké ztrátě, která je přesně 2 500 000 Kč, což je maximální povolená hodnota. Tento výsledek by nás ale neměl překvapit, neboť čím větší ztrátu jsme ochotni podstoupit, tím více můžeme vydělat na úrocích. Výše ztráty je však velmi důležitá při určování zisku, o který se ve výsledku vždy zajímáme. Zisk banky dostaneme odečtením ztráty od výnosu, a tak v této situaci banka získá přibližně 1 704 298 Kč.
Nyní se podíváme na situaci, ve které chce vedení banky maximalizovat rovnou ziskovou funkci, jež je tvořená příjmy z úroků, příjmy z navrácení poloviny částky některými

nesplácejícími a ztrátou z nesplácení. Účelovou funkci pak můžeme s využitím zadefinovaného značení zapsat ve tvaru

$$\sum_{i=1}^{8} p_i \cdot x_i \cdot r_i/100 + \sum_{i=1}^{8} (1-p_i) \cdot h_i \cdot x_i/2 - \sum_{i=1}^{8} (1-p_i) \cdot x_i.$$

K této úloze nepřidáváme žádné další podmínky. Její výsledek najde čtenář v Tabulce 2.85.

kategorie	A	B	C	D	E	F	G	H	2T repo
x_i	800000	49200000	0	0	0	0	0	0	0
úroky	23176,8	1901530,8	0	0	0	0	0	0	0
vrácení	300	166050	0	0	0	0	0	0	0
ztráta	500	276750	0	0	0	0	0	0	0

Tabulka 2.85: Výsledky pro druhou situaci v Příkladu 2.12.6.

Z Tabulky 2.85 vidíme, že nejvýhodnější strategií je v tomto případě vypůjčit všechny prostředky, až na 800 000 vyhrazených pro kategorii A, do kategorie B, což je dáno tím, že klienti této kategorie poměrně spolehlivě splácí a zároveň jim můžeme dát znatelně vyšší úrok než klientům z kategorie A. Výnos banky je v tomto případě roven 2 091 057,6 Kč. Při ztrátě 277 250 Kč tak banka celkově získá 1 813 807,6 Kč, což je o téměř 110 000 Kč více, než v předchozí situaci. Vedení banky tak prozatím upřednostňuje strategii úvěrování pouze nejlepších klientů. Samozřejmě je však zajímá i výsledek třetí situace. Ve třetím případě chce vedení najít strategii, při které maximalizuje stejnou účelovou funkci jako ve druhém případě s jedinou výjimkou, kterou je vynásobení výnosů koeficientem $\mu = 2$, podobně jako jsme viděli v obecném popisu úlohy. Maximalizujeme tedy funkci

$$2\left(\sum_{i=1}^{8} p_i \cdot x_i \cdot r_i/100 + \sum_{i=1}^{8} (1-p_i) \cdot h_i \cdot x_i/2\right) - \sum_{i=1}^{8} (1-p_i) \cdot x_i$$

na množině zadané omezeními společnými pro všechny případy. V této situaci získáme výsledek zobrazený v Tabulce 2.86.

kategorie	A	B	C	D	E	F	G	H	2T repo
x_i	800000	15866666,67	16666666,67	0	0	0	0	16666666,67	0
úroky	23176,8	613230,8	784000	0	0	0	0	2253333,33	0
vrácení	300	53550	166666,67	0	0	0	0	416666,67	0
ztráta	500	89250	500000	0	0	0	0	2083333,33	0

Tabulka 2.86: Výsledky pro třetí situaci v Příkladu 2.12.6.

Z Tabulky 2.86 vidíme, že se banka, stejně jako v první situaci, zaměří pouze na klienty z kategorií A, B, C a H. Jediný rozdíl je v rozdělení úvěrů mezi kategorie B a C – oproti první situaci se v tomto případě banka ve větší míře soustředí na klienty z kategorie C. Nalezené maximum účelové funkce je nyní rovno 5 948 765,2 Kč. Opět se ale zaměřme spíše na zisk, který je 1 637 840,93 Kč, což je nejméně ze všech uvažovaných situací. Vedení se tak rozhodne využít druhou strategii, ve které maximalizuje přímo ziskovou funkci.

Nyní se podíváme na to, jak se změní optimální rozhodnutí banky v jednotlivých situacích, jestliže k úloze přidáme omezení

$$x_6 + x_7 + x_8 \geq 5 \cdot 10^6.$$

Z Tabulek 2.84 a 2.86 vidíme, že v první a třetí situaci je optimální vypůjčit klientům z kategorie H více než 16 milionů. Nalezená řešení tak jistě splní i nově přidané omezení, které je nijak neovlivní. Ve druhé situaci je tomu ale jinak. Z Tabulky 2.85 vidíme, že banka půjčuje pouze klientům z kategorií A a B. Přidáním dalšího omezení tak jistě změníme výsledek, který je pro novou formulaci úlohy zobrazen v Tabulce 2.87.

kategorie	A	B	C	D	E	F	G	H	2T repo
x_i	800000	44200000	0	0	0	0	0	5000000	0
úroky	23176,8	1708285,8	0	0	0	0	0	676000	0
vrácení	300	149175	0	0	0	0	0	125000	0
ztráta	500	248625	0	0	0	0	0	625000	0

Tabulka 2.87: Výsledky pro změnu ve druhé situaci v Příkladu 2.12.6.

Zisk banky je v tomto případě 1 807 812,6 Kč, takže i s přidáním dalšího omezení je druhý způsob pojetí rizika tím nejvýhodnějším. Vedení banky se tak rozhodne využít při svém rozhodování model, ve kterém maximalizuje přímo ziskovou funkci, a to nezávisle na tom, zda využije poslední uvažované omezení či ne. Ještě zdůrazněme, že čtenář si mohl všimnout toho, že ve všech uvažovaných situacích vypůjčila banka klientům z kategorie A pouze přislíbených 800 000 Kč. I přes vysokou pravděpodobnost splácení těchto klientů se totiž bance nevyplatí jim půjčit více, a to z důvodu velmi nízkých úrokových příjmů.

Příklad na tvorbu portfolia, který jsme právě viděli, obsahoval několik výrazných zjednodušení reality. Banky se totiž běžně rozhodují o každém klientovi zvlášť tím způsobem, že mu podle jeho charakteristik dají hodnocení, podle kterého určí nabízenou úrokovou sazbu. Na tu buď klient přistoupí, nebo ne. V každém případě je však bankovní sektor poměrně regulovaný a banky musí splňovat podmínky pro hodnoty jednotlivých rizik svých investic. A právě příklady takových omezení jsme v úloze využili. Úloha byla sice ilustrativní, avšak v případě obecného rozhodování o investicích bychom úlohu o tvorbě portfolia mohli formulovat obdobně. Problém tvorby portfolia je často formulován též jako úloha kvadratického programování, jež nám dovolí využít nelineární účelovou funkci, viz Kapitolu 4. Riziko pak můžeme definovat např. jako rozptyl. I přesto však můžeme lineární programování ve finančních aplikacích využít, a to nejen při tvorbě portfolia, ale např. i při oceňování opcí. S dalšími příklady využití lineárního programování ve financích se čtenář může setkat např. v [154,491]. Na závěr ještě alespoň zmiňme, že propojení lineárního programování a teorie her můžeme najít např. v [478, 491] a popojení s fyzikou např. v [287]. Kromě problémů týkajících se různých ekonomických potřeb se pochopitelně objevují i jinak motivované úlohy, jako je např. problém vybarvení mapy co nejmenším počtem barev v [60] a další. K dalšímu studiu aplikací lineárního programování lze kromě již uvedené literatury doporučit také např. [7, 125, 154, 292, 330].

Cvičení

2.12.1. Brněnská zbrojovka dodává zbraně po celém světě. Uvažme tři z jejich výrobků. Výrobní proces má tři fáze: odlévání dílů, kompletace a balení. Potřebné informace jsou shrnuty v tabulce:

	výrobek A	výrobek B	výrobek C	dostupný čas [h]
odlévání [h]	1	2	3/2	12000
kompletace [h]	2/3	2/3	1	4600
balení [h]	1/2	1/3	1/2	2400
zisk [tisíc Kč]	11	16	15	

Jak rozdělit výrobu, aby byl zisk maximální?

Řešení: 600 ks výrobků A, 5100 ks výrobků B, 800 ks výrobků C.

2.12.2. (O 10 let později) Přírodovědecká fakulta MU se rozhodla zvýšit svoji propagační činnost tím, že vyrobí různé propagační předměty: trička, mikiny a čepice (vše samozřejmě s potiskem). Následující tabulka ukazuje potřebné informace (údaje jsou v desítkách kusů):

	trička	mikiny	čepice	dostupné množství
látka [m]	28	30	16	1550
niť [m]	36	40	24	2044
potisk [cm^2]	32	50	30	2438
obalový materiál [m]	15	20	10	975
zisk [Kč]	104	136	78	

Vaším úkolem (jakožto úspěšného absolventa ÚMS) je stanovení optimální výrobní strategie.

Řešení: [90, 370, 100] *kusů.*

2.12.3. Při analýzách výsledků přímé volby prezidenta ČR v roce 2013 se ukázalo, že Ing. Miloš Zeman vyhrál pouze díky chybě Ing. Miroslava Kalouska. Ten jako vystudovaný chemik (ač bývalý ministr financí) pochopitelně neovládá základy lineárního programování. Proto nastavil špatnou propagační strategii před druhým kolem, což vedlo k prohře Karla Schwarzenberga. Úkolem bylo rozdělit reklamu mezi TV, rozhlas a noviny. Náklady a účinnost jednotlivých typů jsou uvedeny v tabulce:

	TV	noviny	rozhlas
cena za kus [tisíc Kč]	40	12	6
„zasažených“ voličů za den [v tisících]	100	40	18

Noviny však nepovolují více než 10 ks reklam/reklamních článků za týden. Navíc pro zachování rovnováhy nesmí být více než 1/2 z celkového počtu v rozhlase a nejméně 10 procent musí v TV. Týdenní rozpočet je 364 tisíc Kč. Jakým způsobem rozložit reklamu tak, aby se maximalizoval počet „zasažených“ voličů?

Řešení: 4 v TV, 10 v novinách a 14 v rozhlase.

2.12.4. Tentokrát vlastníte dvě malá knihkupectví – jedno v Brně a druhé na Slovensku. V Anglii se chystá již 27. díl dobrodružství Harryho Pottera. Protože se jedná o první vlnu máte k dispozici pouze omezené množství k objednání, konkrétně až 300 ks pro ČR a 200 ks pro SR (sic!). Knihy mohou být zaslány buď přímo z domu J. K. Rowlingové (JKR) nebo z amerického vydavatelství v USA. Náklady na dopravu jsou shrnuty v tabulce [v Kč za kus, přičemž nelze získat slevu při hromadné objednávce]:

	ČR	SR
JKR	500	600
USA	600	720

Z domu JKR může být zasláno maximálně 400 ks a z USA nejvýše 300 ks. Jak sestavit objednávku, aby náklady byly minimální (přičemž máte jistotu, že bez problému prodáte všechny objednané kusy)?

Řešení: 200 ks od JKR do ČR (a 100 z USA), 200 ks od JKR do SR (a 0 z USA).

2.12.5. Nyní se vraťme do historie. Na počátku 90. let Bill Gates a Steve Jobs zvažovali spojení svých společností do jedné, ve které by se vyráběly počítače s jedním z operačních systémů dle přání

zákazníka. Pracovní týden by nabízel 480 hodin a požadavky na pracovní dobu by byly následující:

	složení počítače [h]	instalace OS [h]
iOS	2.6	2.1
WIN	1.5	1.8

Počítače by také mohly být koupeny přímo od jiných výrobců bez OS za cenu 2200,- pro WIN a 4000,- pro iOS. Kompletní počítač s iOS by se prodával (v 90. letech) za 7000,-. Prvních 70 ks s WIN by se prodávalo za 5000,- a další za 4500,- (jeden z geniálních tahů Billa Gatese), přičemž poptávka bez problému pokryje výrobu. Ovšem ani jeden z velikánů IT neovládal lineární programování tak jako Vy. Proto k fúzi nedošlo, neboť nebyli schopni se dohodnout na optimálním rozdělení práce. Nalezněte správné řešení, chcete-li maximalizovat zisk.

Řešení: 75.32 ks iOS a 70 ks WIN.

2.12.6. Po úspěšném ukončení studia Vás najala ropná společnost (zřejmě Shell nebo Moravské naftové doly), která spravuje mimo jiné dvě rafinérie. První rafinérie má denní náklady na provoz ve výši 400 tisíc Kč a dokáže produkovat až 400 barelů nafty nejvyšší kvality (Q1), 300 barelů nafty nadstandardní kvality (Q2) a 200 barelů nafty průměrné kvality (Q3). Druhá rafinérie je novější a moderněji vybavená. Její náklady jsou 500 tisíc Kč na den a vyrobí až 300/400/500 barelů nafty kvality Q1/Q2/Q3. Firma musí dodat celkem 25000/27000/30000 barelů nafty kvality Q1/Q2/Q3. Kolik dnů provozu každé rafinérie je potřeba, aby byla uspokojena poptávka za minimální náklady?

Řešení: 25 dnů rafinérie A, 50 dnů rafinérie B.

2.12.7. (O dalších 20 let později) Město Brno dokončilo rekonstrukci fotbalového stadionu Za Lužánkami (ZL). Správce fotbalového stadionu se tudíž musí starat o travnatou plochu na novém hřišti i na „starém“ stadiónu v Králově poli (KP). Každé hřiště vyžaduje travnatou směs s různým obsahem jílku vytrvalého (JV) a lipnice luční (LL). K tomuto účelu má správce k dispozici tři druhy směsí. Potřebné informace jsou opět v tabulce:

	JV [%]	LL [%]	cena [Kč/kg]	dostupné množství na trhu [kg]
Směs 1	60	10	80	neomezené
Směs 2	20	50	95	✖
Směs 3	25	15	35	125
požadavek ZL	alespoň 30	alespoň 10	✖	✖
požadavek KP	alespoň 25	alespoň 45	✖	✖

Správce potřebuje právě 200 kg směsi pro ZL a 180 kg pro KP. Jaké množství jednotlivých směsí nakoupit, aby byly náklady minimální?

Řešení pro ZL: $[75, 0, 125]$, *pro KP:* $[22.5, 157.5, 0]$.

2.12.8. Představme si výrobce, který produkuje čtyři druhy ovocných pálenek. Jeho cílem je naplánovat výrobu na následujících šest měsíců tak, aby minimalizoval náklady na výrobu a skladování zásob. Na začátku roku nemá výrobce žádné zásoby. Ze zkušenosti s podnikáním zná výrobce poptávky, kterým v jednotlivých měsících čelí, i výrobní kapacity, jež může v daném měsíci využít. Náklady na výrobu jednotlivých pálenek se v průběhu půlroku mění, neboť se mění cena ovoce potřebného pro jejich výrobu. Maximální produkční možnosti a poptávky po jednotlivých druzích pálenek (označených P_1, P_2, P_3 a P_4) v hektolitrech jsou uvedeny v Tabulce 2.88, ve které sloupce představují jednotlivé měsíce. Produkční možnosti se podobně jako ceny mění, a to z důvodu lišících se dostupných množství ovoce. V poslední části tabulky vidíme náklady na výrobu jednotlivých pálenek a cenu za skladování (ozn. S), jež je pro všechny pálenky stejná. Všechny náklady jsou uvedeny ve sto tisících Kč.

max. výroba	1	2	3	4	5	6
P_1	100	100	110	120	140	60
P_2	120	100	80	80	60	120
P_3	80	90	100	110	100	90
P_4	40	50	90	90	120	110
poptávka	1	2	3	4	5	6
P_1	80	110	50	130	40	80
P_2	60	40	180	90	50	130
P_3	40	100	80	60	170	60
P_4	40	40	40	20	150	90
náklady	1	2	3	4	5	6
P_1	1	1,2	0,9	0,8	0,9	1
P_2	1,10	1	0,8	0,95	1,1	1,2
P_3	1	0,9	0,8	0,9	1	1,1
P_4	0,8	0,8	0,75	0,7	0,85	1
S	0,5	0,5	0,7	0,9	0,8	0,6

Tabulka 2.88: Zadání první části.

Výrobce chce naplánovat výrobu tak, aby vždy uspokojil poptávky a minimalizoval při tom náklady. Náklady na reorganizaci výroby však neuvažuje. Pomozte nyní výrobci najít plán optimální výroby a tvorby zásob, který splní zadané požadavky.

Ve druhé části příkladu se podíváme na situaci, do které se výrobce dostal po dvou měsících výroby podle optimálního plánu z první části příkladu. Vzhledem k pandemii, která na začátku roku 2020 nečekaně zasáhla celý svět, se výrazně změnila očekávaná poptávka po všech druzích pálenky. Náklady na výrobu a skladování však zůstaly stejné. Výrobce chce nyní optimální plán pro zbývající čtyři měsíce upravit tak, aby opět pokryl očekávané poptávky po svých výrobcích. Jeho maximální výrobní kapacity jsou ovšem nyní určeny optimálním výrobním plánem z první části úlohy, neboť podle něho výrobce nakoupil zásoby a přijal zaměstnance. Zároveň chce zachovat alespoň 90 % celkového naplánovaného měsíčního využití výrobních kapacit, neboť nechce propouštět zaměstnance. Zachování výroby chce dosáhnout výrobou dezinfekce (označené D), které je v současné době nedostatek, a k jejíž výrobě může využít líh a další již zakoupené suroviny. Náklady na výrobu dezinfekce jsou spolu s novými poptávkami uvedeny v Tabulce 2.89.

poptávka	3	4	5	6
P_1	40	100	180	40
P_2	120	45	40	95
P_3	40	130	120	30
P_4	110	10	40	45
náklady	3	4	5	6
D	0,6	0,7	0,75	0,8

Tabulka 2.89: Zadání druhé části.

Dezinfekci může výrobce vyrábět s využitím výrobních kapacit pro všechny druhy pálenek. Dezinfekci výrobce nikdy neskladuje, neboť ji chce co nejdříve rozvést na potřebná místa. Zásoby, které máme k dispozici na začátku třetího měsíce, jsou zásoby ze druhého měsíce při provozování původního optimálního výrobního plánu. Najděte nyní nový optimální výrobní plán tak, aby byla splněna všechna omezení. Kolik dezinfekce je výrobce schopen v jednotlivých měsících vyrobit? Jak by se situace změnila, kdyby chtěl výrobce zachovat 90 % celkového (ne měsíčního) využití kapacit podle původního plánu?

2.13 Kleeho[231]–Mintyho[232] kostka je vržena

> *Before attacking a linear programming problem with the simplex algorithm, it can be very comforting to have a good estimate of the number of iterations which may be required in order to reach a solution. For problems of a given size, the number of iterations is nearly proportional to the actual computation time, and it may also give some indication of the extent to which round-off errors will become troublesome (depending on the computer program by means of which the algorithm is implemented).*
> *Of primary interest is the expected number of iterations. This can be defined mathematically in various ways, but the resulting systems appear to be too complex for mathematical analysis and in any case are of doubtful relevance to the practical situation. It seems that the only reliable guide to the expected number of iterations is computational experience such as that reported by Kuhn & Quandt, Wolfe & Cutler or Dantzig.*
> *In addition to the primary interest in the expected number of iterations, there is a strong secondary interest in the maximum number of iterations for problems of a given size. Here the analytical approach seems more relevant as well as more feasible, and the still-unsolved problem of determining the maximum number has been of central interest since the inception of the simplex algorithm...*[233]

Simplexový algoritmus jsem si už popsali velmi podrobně a prozkoumali přitom i jeho různá úskalí, avšak jeden velmi důležitý aspekt jsme zatím opomíjeli – jde o tzv. *časovou* (příp. *výpočetní*) *složitost*[234], která popisuje časovou či výpočetní náročnost nějaké operace či algoritmu v závislosti na velikosti vstupu[235,236], viz také Podkapitolu 3.6. Už jsme si ukázali, že simplexový algoritmus (případně doplněný o nějaké anticyklické pravidlo) je konečně krokový. Dokonce víme, že maximální počet potřebných kroků (ve druhé fázi) je nejvýše roven $\binom{n}{m}$ pro úlohu v kanonickém tvaru. Jenže ne každá báze nutně určuje nějaký BPB, takže kolik to skutečně bude kroků? Samozřejmě v některých situacích bude simplexový algoritmus velmi rychlý a tisícovky úloh řešených s jeho pomocí v prvních 20-25 letech obvykle nepotřebovaly více než $3m$ kroků, což je skvělý výsledek, který neopomenul zmínit ani Dantzig ve své knize.

> *... empirical experience with thousands of practical problems indicates that the number of iterations is usually close to the number of basic variables in the final set which were not present in the initial set. For an m-equation problem with m different variables in the final basic set, the number of iterations may run anywhere from to as a minimum, to $2m$ and rarely to $3m$. The number is usually less than $3m/2$ when there are less than 50 equations and 200 variables (to judge from informal empirical observations). Some believe that for a randomly chosen problem with fixed m, the number of iterations grows in proportion to n.*[237]

231

Victor L. Klee, Jr. (18. září 1925 – 17. srpna 2007) byl americký matematik, ketrý se věnoval především konvexním množinám, funkcionální analýze, analýze algoritmů, optimalizaci a kombinatorice. Téměř celý svůj profesionální život strávil na Washingtonská univerzita v Seattlu.

232

George James Minty (16. září 1929 – 6. srpna 1986) byl americký matematik, který se věnoval především matematické analýze a diskrétní matematice.

[233] *Než se pustíme do řešení problému lineárního programování pomocí simplexového algoritmu, může být velmi povzbudivé mít dobrý odhad počtu iterací, které mohou být nutné k nalezení řešení. U problémů dané velikosti je počet iterací téměř úměrný skutečnému výpočetnímu času, což může také poskytnout určitý vhled do toho, jaké komplikace mohou způsobit zaokrouhlovací chyby (v závislosti na počítačovém programu, pomocí kterého je algoritmus implementován).*

Hlavní otázkou je očekávaný počet iterací. Ten lze definovat různými matematickými způsoby, ale výsledné systémy se ukazují být příliš složité pro jakoukoli matematickou analýzu a rozhodně mají v každém případě pochybný význam z praktického pohledu. Zdá se, že jediným spolehlivým vodítkem k očekávanému počtu iterací jsou zkušenosti s výpočty jako ty, které uvádějí Kuhn & Quandt, Wolfe & Cutler nebo Dantzig.

Navíc vedle primární otázky očekávaného počtu iterací tu je i důležitá druhotná otázka týkající se maximálním počtu iterací pro problémy dané velikosti. Zde se analytický přístup jeví jako relevantnější a také proveditelnější. Tento stále nevyřešený problém s určením maximálního počtu je středem zájmu již od vzniku simplexového algoritmu. Viz úvod ke Kleeho článku [324].

[234] Angl. *computational complexity*.

[235] Toto se zapisuje pomocí *Landauovy notace* (neboli *velkého O*), tj. jako $O(f(N))$, což znamená, že časová náročnost uvažované operace nebo algoritmu je menší než $Af(N)$ pro nějaké $A>0$, kde N popisuje velikost vstupních dat (bity, byty – pro počítač totiž potřebujeme data zapsat do binárních řetězců, nebo např. vrcholy grafu pro grafové algoritmy atd.). Např. $O(N)$ znamená, že časová složitost se zvyšuje stejně jako velikost vstupu (tj. *lineární složitost*), takže zdvojnásobením vstupu se zdvojnásobí výpočetní čas; $O(N^2)$ znamená *kvadratickou složitost*, takže zdvojnásobením velikosti vstupu se zčtyřnásobí výpočetní čas; $O(1)$ dává vždy stejný výpočetní čas bez ohledu na velikost vstupu a v případě $O(2^N)$ pro trojnásobnou velikost vstupu bude potřeba osminásobek výpočetního času.

[236] Např. u sčítání nebo odčítání dvou čísel o N cifrách klasickým (školským) způsobem je výpočetní složitost přímo úměrná počtu cifer, tj. $O(N)$. Klasické násobení dvou čísel o N cifrách má výpočetní složitost $O(N^2)$ a pro násobení dvou $N\times N$ matic to je $O(N^3)$. Řešení soustavy N rovnic o N neznámých Gaussovou eliminací (tj. vytvoření schodovité matice a zpětné řešení pro určení neznámých) vyžaduje $N(N+1)/2$ dělení, $(2N^3+3N^2-5N)/6$ násobení a stejný počet odčítání, což celkem dává přibližně $2N^3/6$ operací neboli máme $O(N^3)$.

[237] *… empirická zkušenost /sic!/ s tisíci praktických problémů ukazuje, že počet iterací se obvykle blíží počtu bazických proměnných ve výsledné bázi, které nebyly v počáteční bázi. Pro problém s m-ticí rovnic a m-ticí různých proměnných ve výsledné bázi může počet iterací jakékoli číslo od minima až po $2m$ a zřídka až do $3m$. Máme-li méně než 50 rovnic a 200 proměnných, je tento počet obvykle menší než $3m/2$ (usuzováno z neformálních empirických pozorování). Někteří se domnívají, že u náhodně zvoleného problému s pevným m roste počet iterací úměrně s n.* Viz [117, str. 160].

[238] *Existuje domněnka, že správnou volbou proměnných vstupujících do báze je možné přejít z jakéhokoli BPB na jakýkoli jiný pomocí m nebo méně kroků, při kterých je každý bazický bod i přípustný. Je známo, že pro případy s $m\le 4$ je tato domněnka pravdivá.* Viz [117, str. 160].

Jenže nás teď zajímá, jak skutečně *pomalý* může tento algoritmus být! Moc ale asi ne, jak se nás dále snaží Dantzig přesvědčit.

> *It has been conjectured that, by proper choice of variables to enter the basic set, it is possible to pass from any basic feasible solution to any other in m or less pivot steps, where each basic solution generated along the way must be feasible. For the cases $m\le 4$ the conjecture is known to be true.* [238]

Toto je tzv. *Hirschova domněnka*, o která se Hirsch[239] zmínil Dantzigovi v dopise z roku 1957. Jinými slovy to znamená, že pro každý problém lineárního programování by měla existovat varianta simplexové metody, s jejíž pomocí nalezneme řešení v nejvýše m krocích. V jádru této domněnky je víra v to, že vhodnou úpravou pravidel simplexového algoritmu je možné zabránit opakovanému vstupu do báze proměnným, které nejsou součástí báze samotného řešení, tj. je-li nějaká proměnná v počáteční (nebo nějaké následné) bázi a není-li v bázi řešení, pak po jejím vystoupení z báze se již v dalším průběhu algoritmu do báze nevrátí. Tenhle „výsledek“ fascinoval řadu matematiků a bylo o něm napsáno mnoho článků. Jenomže v roce 1965 Klee přišel s úlohou lineárního programování (LP) s obecným počtem omezení, ve které je k propojení dvou BPB potřeba *alespoň* $m(n-m-1)+1$ kroků! To nikdo nečekal. Ačkoli Klee se snažil překvapení z tohoto zjištění mírnit následnou domněnkou, že toto je absolutní maximum, byl zde očividný rozpor, který v roce 1969 okomentoval Gale slovy:

> *Thus there is a large and embarrassing gap between what bas been observed and what has been proved. This gap has stood as a challenge to workers in the field for twenty years now and remains in my opinion, the principal open question in the theory of linear computation.* [240]

A to není všechno. O 45 let později dokonce Santos[241] ukázal s pomocí 43-dimenzionální politopu, že Hirschova domněnka je nesprávná. To už je ale jenom pomyslná třešnička na dortu, neboť tento protipříklad toho moc neříká o časové složitosti simplexového algoritmu a navíc v té době už byla správná odpověď dávno známa. Tu totiž skýtá následující příklad od Kleeho a Mintyho z roku 1972, v němž je úloha lineárního programování s obecným počtem proměnných, viz [325]. Pro úplnost poznamenejme, že v literatuře lze najít tento příklad i s jinými koeficienty, ale průběh simplexového algoritmu je vždy stejný (liší se pouze souřadnice jednotlivých bodů).

Příklad 2.13.1

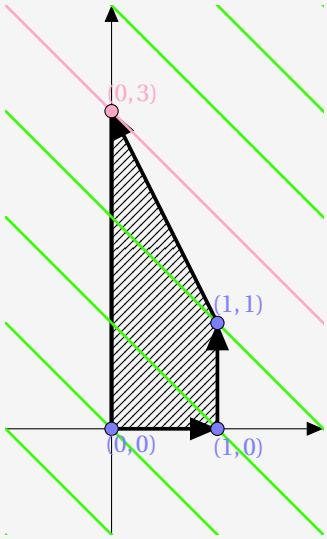
Simplexovým algoritmem vyřešme úlohu

$$x_1+x_2+\cdots+x_n\to\max$$
$$x_1\le 1 \quad\&\quad 2x_1+x_2\le 3$$
$$2x_1+2x_2+x_3\le 9$$
$$\vdots$$
$$2x_1+2x_2+\cdots+2x_{n-1}+x_n\le 3^{n-1}$$
$$x_1\ge 0 \quad\&\quad \cdots \quad\&\quad x_n\ge 0$$

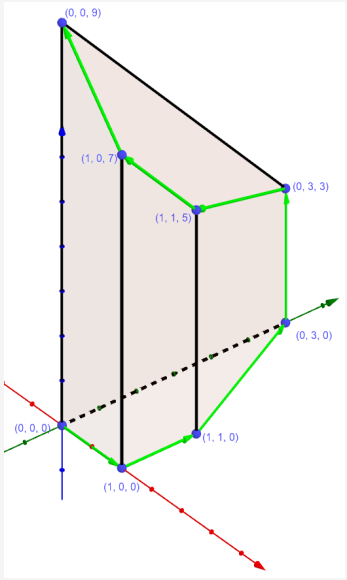
Řešení. Daná omezení určují polytop známý jako *Kleeho–Mintyho kostka*. Pro $n=1$ se jedná pouze o interval $[0,1]$. Pro převod do kanonického tvaru potřebujeme jednu přídavnou proměnnou, takže maximální počet BPB je $\binom{2}{1}=2$. Jestliže za výchozí BPB zvolíme bod odpovídající $x_1=0$, pak simplexový algoritmus nalezne řešení $x_1=1$ hned v dalším kroku. Kleeho–Mintyho kostku pro $n=2$ získáme tak, že vezmeme předchozí

úsečku a v druhém směru k ní pod jistým úhlem „přidáme“ druhou, čímž získáme nepravidelný čtyřúhelník z Obrázku 2.39. Pro převod do kanonického tvaru budeme potřebovat dvě přídavné proměnné, takže maximální počet BPB je $\binom{4}{2} = 6$. Ve skutečnosti jsou pouze 4 (úloha je nedegenerovaná) a simplexový algoritmus s výchozím BPB odpovídajícím $[0,0]$ projde všechny vrcholy daného polytopu, než nalezne řešení v bodě $[0,3]$, tj. budou potřeba 3 kroky. Průběh algoritmu jednotlivými body je na Obrázku 2.39 zvýrazně šipkami.

Pro $n = 3$ získáme Kleeho–Mintyho kostku tak, že vezmeme předchozí čtyřúhelník a ve třetím směru k němu pod jistým úhlem „přidáme“ druhý, čímž získáme nepravidelný šestistěn. Pro převod do kanonického tvaru budeme potřebovat tři přídavné proměnné, takže maximální počet BPB je $\binom{6}{3} = 20$. Ve skutečnosti jich je pouze 8 (úloha je nedegenerovaná) a simplexový algoritmus s výchozím BPB odpovídajícím $[0,0,0]$ opět projde všechny vrcholy daného polytopu, než nalezne řešení v bodě $[0,0,9]$, tj. bude potřeba 7 kroků. Průběh algoritmu jednotlivými body je znázorněn šipkami na Obrázku 2.40.



Obrázek 2.39: Kleeho–Mintyho kostka, vrstevnice a průběh simplexového algoritmu pro $n = 2$.



Obrázek 2.40: Kleeho–Mintyho kostka a průběh simplexového algoritmu pro $n = 3$.

Všimněme si, že algoritmus nejdříve projde všechny vrcholy v dolním čtyřúhelníku odpovídajícímu předchozímu případu s $n = 2$, a to dokonce i ve stejném pořadí (třetí souřadnice je stále nulová). Pak pokračuje do horního čtyřúhelníku. Totéž chování můžeme vidět i v předchozí situaci pro $n = 2$ a také pro libovolné jiné n. Průběh simplexového algoritmu bude vždy stejný: Kleeho–Mintyho kostka bude mít 2^n vrcholů a bude potřeba $2^n - 1$ kroků.

Např. při rychlosti 1000 kroků za vteřinu a $n = 20$ by výpočet trval méně než 9 minut, pro $n = 50$ by to již bylo více než 17000 let a pro $n = 70$ více než $1{,}8 \times 10^{10}$, což přesahuje stáří našeho vesmíru. Tato exponenciální časová složitost je celkem překvapující, neboť už jsme zmínili, že z praktického pohledu je tato metoda (obvykle) velmi rychlá, a tudíž zcela dostačující. Nicméně o pravdivosti následujícího tvrzení nemůže být pochyb.

Věta 2.13.2

Časová složitost simplexového algoritmu je $O(2^N)$, tj. exponenciální.

Samozřejmě mohli bychom namítat, že existují i další varianty simplexové metody (založené na jiných pravidlech pro výběr klíčového sloupce), jenže ani s nimi to nedopadne lépe, jak bylo ukázáno v [9, 20, 167, 224, 295, 388, 519], viz také [457, 520]. V literatuře lze také najít pravidla

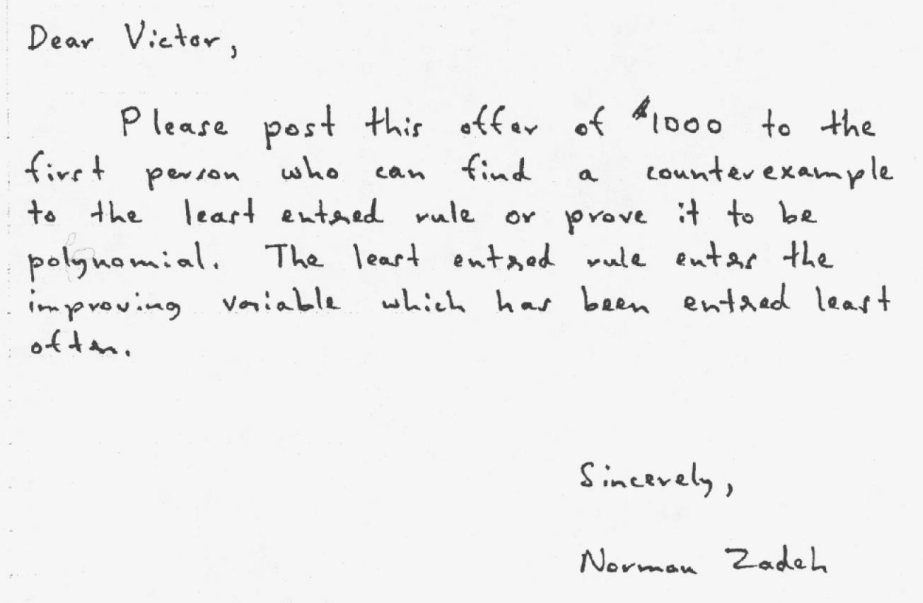
[239] Warren M. Hirsch (1917–2007) byl americký matematik a emeritní profesor matematiky na Courantově institutu matematických věd při Newyorské univerzitě. Věnoval se zejména matematické biologii (zejména přenosu parazitických onemocnění), optimalizaci a operačnímu výzkumu, ekonomii, pravděpodobnosti či statistice. Je znám především díky modelu předpovídajícímu tropickou nemoc schistosomózu, ze kterého se stal základní koncept v matematické epidemiologii. Jednou z jeho prvních prací byl společný článek s Dantzigem ohledně úlohy lineárního programování s „fixními náklady“ (jako při nastupování do taxi, kdy je hned na začátku naúčtována pevná sazba bez ohledu na to, kam jedete), který vyšel v roce 1954 jako zpráva RANDu.

[240] *Existuje tedy velká a trapná propast mezi tím, co bylo pozorováno, a tím, co bylo dokázáno. Tato propast zůstává výzvou pro pracovníky v oboru již dvacet let a zůstává podle mého názoru hlavní otevřenou otázkou v teorii lineárních výpočtů.* Viz [197, str. 590].

[241] Francisco (Paco) Santos Leal (∗ 1968) je španělský matematik, který se proslavil především díky protipříkladu z roku 2010, kterým se mu podařilo vyvrátit Hirschovu domněnku. V roce 2015 získal za tento svůj objev *Fulkerson Prize* udělovanou *Americkou matematickou společností* a *Společností matematického programování* (*Mathematical Programming Society*).

využívající *náhodného výběru* klíčového sloupce z vhodných kandidátů, ale i v jejich případě jsme na hony vzdáleni jakémukoli polynomiálnímu času známého z praxe, neboť jde velmi zjednodušeně řečeno zhruba o $m^{O(\sqrt{n})}$, viz raději [304,375]. Jedním z posledních kandidátů na polynomiální časovou složitost by mohlo být Zadehovo[242] *pravidlo nejraritnější bazické proměnné*[243]. Zadeh sám dokonce nabídl odměnu 1 000 $ za důkaz či protipříklad polynomiální časové složitosti tohoto pravidla, viz Obrázek 2.41 níže a [19,521].

Chtělo by se říci: s chutí do toho! Jenže 20. ledna 2011 vystoupil na workshopu *Efficiency of the Simplex Method: Quo vadis Hirsch conjecture?* na Kalifornské univerzitě v Los Angeles (a 15. června 2011 na mezinárodní konferenci věnované celočíselnému programování a kombinatorické optimalizaci v New Yorku) mladý matematik O. Friedmann s příspěvkem, ve kterém ukázal, že Zadehovo pravidlo je nejméně *subexponenciální* (tj. časová náročnost roste se zvětšující se velikostí vstupu rychleji než jakýkoli polynom neboli je větší než $O(n^k)$ pro jakékoli $k \in \mathbb{N}$.), takže ani tentokrát se teoretické poznatky nepřiblíží k chování známému z praxe, viz [188]. V době, kdy Zadeh svoji nabídku psal, byl „pouze“ univerzitním profesorem, avšak nyní je úspěšným podnikatelem, takže pro něj nebyl žádný problém slíbenou částku zaplatit[244]. Přijal proto pozvání na zmíněný workshopu a slavnostně zde Friedmannovy předal šek se slíbenou odměnou.


Dear Victor,

Please post this offer of $1000 to the first person who can find a counterexample to the least entered rule or prove it to be polynomial. The least entered rule enters the improving variable which has been entered least often.

Sincerely,

Norman Zadeh

Obrázek 2.41: Zadehova nabídka za 1000$ z dopisu někdy kolem roku 1980 a určeném Kleemu. Převzato z [521].

Je ale snad jasné, že se tu jedná spíše o velmi patologické (nebo akademické) případy[245], což potvrzuje i výsledek Spielmana[246] a Tenga[247] z článku [457] z roku 2001 pro jistou verzi simplexového algoritmu: vezme-li se libovolné zadání úlohy lineárního programování (LP) a s každým omezením se nepatrně „pohne“ náhodným směrem (např. trochu změníme všechny složky vektoru b), poběží na výsledné úloze simplexová metoda v polynomiálním čase[248].

Mohlo by se zdát, že tím je příběh lineárního programování a simplexové metody po více než 70 letech u svého konce. Jenže to by nebyl příliš šťastný konec. Naopak, díky těmto poznatkům zažilo lineární programování na přelomu 70. a 80. let minulého století svoji obrodu a opět podnítilo zájem řady matematiků. Zatímco praktici hledali způsoby, jak snížit výpočetní čas simplexového algoritmu o 10 či 20 procent, matematičtí teoretici hledali odpověď na mnohem zásadnější otázku: existuje pro lineární programování algoritmus s polynomiální časovou složitostí, tj. $O(n^k)$ pro nějaké $k \in \mathbb{N}$? Netrvalo to příliš dlouho a první vlaštovka se objevila už v roce 1979. Ruský matematik Chačijan[249] publikoval v ruštině krátký článek (bez důkazu, který doplnil až o rok později) o tom, jak lze pomocí tzv. *metody elipsoidů* řešit

[242] Norman Zada (∗??; původním jménem Norman Askar Zadeh) je bývalým mimořádným profesorem matematiky na Stanfordově univerzitě, Kolumbijské univerzitě, KaKalifornské univerzitě v Los Angeles a Kalifornské univerzitě v Irvine, kde se věnoval aplikované matematice. Je synem Lotfiho Zadeha, který je tvůrcem fuzzy logiky. V letech 1974 a 2020 napsal knihy o výherních strategiích v pokeru. Od 80. let se však věnuje především podnikání. Založil pánský časopis *Perfect 10*, který je zaměřen na ženy bez kosmetické chirurgie, a vede společnost United States Investing Competition.

[243] Zadeh toto pravidlo navrhl někdy kolem roku 1980, viz [520]. Patří do skupiny pravidel založených na „historii“, kdy jsou v průběhu simplexového algoritmu uchovávána data s výskytem jednotlivých proměnných v bázi. Za proměnnou vstupující do báze (tj. klíčový sloupec) je z vhodných kandidátů vybrána ta, která doposud byla v bázi nejméně často. Intuitivně by tím mělo být zajištěno, že proměnné, které by v dlouhodobém horizontu měly přinést podstatné zlepšení, ale pouze malé zlepšení v jednom kroku, budou použity po lineárním počtu kroků. Za zmínku ještě stojí, že Zadehovo pravidlo postrádá dodatečné kritérium pro rozhodnutí o vstupující proměnné v případě shody.

[244] V roce 2016 údajně prodal svoji vilu v Beverly Hills za více než 16 milionů dolarů.

[245] Takové mají matematici v této oblasti nejraději. Skutečně!

[246] Daniel Alan Spielman (∗ 1970) je od roku 2006 profesorem aplikované matematiky a matematické informatiky na Yaleově univerzitě. Doktorát získal v roce 1995 na MIT. V letech 2008 a 2015 získal společně se svým spolupracovníkem Tengem Gödel Prize za práce na analýze hladkosti algoritmů a téměř-v-lineárním-čase řešitelném Laplaciánu.

[247] Shang-Hua Teng (∗ 1964) je čínsko-americký matematický informatik. V letech 2008 a 2015 získal společně se svým spolupracovníkem Speilmanem *Gödel Prize* za práce na analýze hladkosti algoritmů a téměř-v-lineárním-čase řešitelném Laplaciánu. V roce 2009 získal *Fulkerson Prize*.

[248] Vzpomeňme si také na komentář k perturbační metodě a případné nedegenerovanosti.

úlohu lineárního programování s m-ticí nerovností s polynomiálním počtem iterací $O(L\,n^2)$, kde L je celková velikost vstupních dat v bitech, viz později a [278, 279]. Hlavní myšlenka této metody nebyla nová a nikterak nesouvisí s lineárním programováním, neboť ji o několik let dříve publikovali Nemirovski[250] a Yudin[251] (viz [516]) jakožto nástroj pro řešení úloh konvexního (nelineárního) programování. O chvíli později Šor[252] zveřejnil i její první explicitní popis v podobě, v jaké ji víceméně známe dnes, viz [469] a také [69]. Převratné v této souvislosti nebylo ani její využití pro řešení úloh lineárního programování, neboť to explicitně zmiňovali ve svém článku již Nemirovski a Yudin, a nebylo to ani odvození různých vzorců nutných při výpočtu, protože ty můžeme najít již v Šorově článku, o kterém se však Chačijan (překvapivě) nezmiňuje. Chačijanův zásadní[253] přínos tkví v tom, že jeho příspěvek byl posledním dílkem do celé této skládanky, jelikož analyzoval časovou složitost tohoto algoritmu pro úlohy lineárního programování a ukázal, že je polynomiální. Vzbudilo to ale až nečekaně velký rozruch. Avšak nikoli v matematické komunitě, ale ve veřejnosti, neboť navzdory tomu, že se jedná o poměrně složitou matematiku, zaznamenal tento výsledek i velký mediální ohlas a psalo se o něm i v *Science* (2. listopadu 1979, viz Obrázek 2.42 a [328]), *The Guardian* (29. října 1979 s titulkem *Russian Way with the Mathematical Travelling Salesman: Unknown Soviet Mathematician LG Khachian Offers a Solution to a Problem That Has the Computing World Baffled*[254], viz [537]) a také v *The New York Times*, pro který jej napsal americký novinář Browne[255], viz Obrázek 2.43 a také [83].

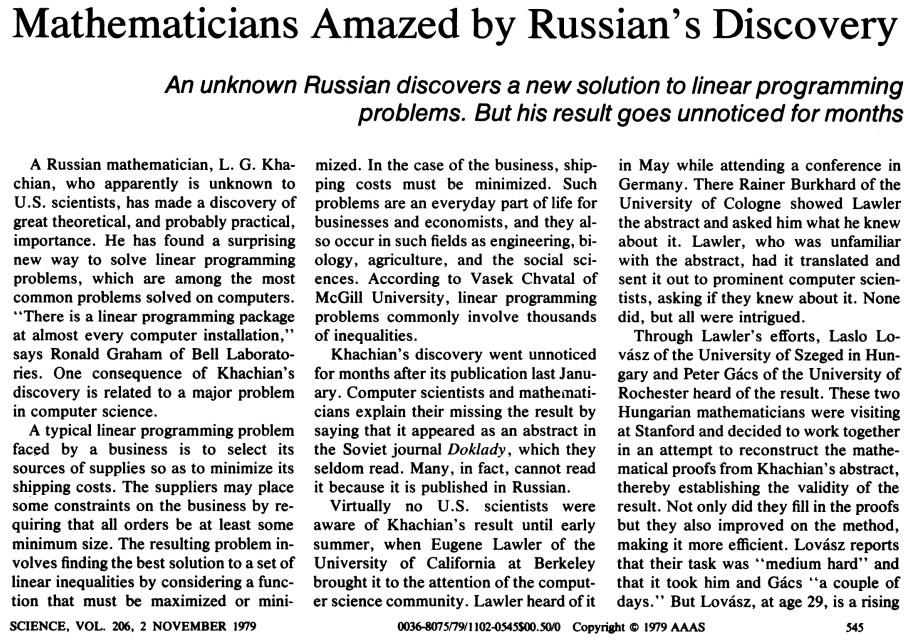
249

Leonid Genrikhovič Chačijan (3. května 1952 –29. dubna 2005) byl sovětský a americký matematik a matematický informatik, který se proslavil především díky metodě elipsoidů pro řešení úloh lineárního programování.

[250] Arkadi Nemirovski (∗ 14. března 1947) je profesorem na Škole průmyslového a systémového inženýrství H. Miltona Stewarta, která je součástí Georgijského technického institutu. Je jedním hlavních představitelů spojité optimalizace a jeho jméno je spojeno především s pracemi ohledně metody elipsoidů, moderní metodě vnitřních prvků a robustní optimalizace. Získal *Fulkerson Prize* (v roce 1982 společně s Grötschelem, Chačijanem, Lovászem, Schrijverem a Yudinem za metodu elipsoidů v lineárním programování a kombinatorické optimalizaci), *Dantzig Prize* (1991 společně s Martinem Grötchlem; udělováno *Mathematical Optimization Society* (Společnost matematické optimalizace) a *Society for Industrial and Applied Mathematics* /SIAM; Společnost pro průmyslovou a aplikovanou matematiku/) a *John von Neumann Theory Prize* (2003 společně s Michaelem Toddem).

[251] David Borisovič Yudin (21. května 1919 – 8. února 2006), byl ruský/sovětský matematik a plukovník sovětské armády. Byl oceněn 2 řády a 15 medailemi. V roce 1982 získal *Fulkerson Prize* (společně s Grötschelem, Chačijanem, Lovászem, Nemirovskim a Schrijverem) a v roce 1992 získal vysoké ruské ocenění v oblasti inženýrství.

[252] Naum Zuselevič Šor (1. ledna 1937 – 26. února 2006) byl sovětský a ukrajinský matematik věnující se především oblasti optimalizace. Svými výsledky významně přispěl k rozvoji nelineárního a stochastického programování, numerických technik nehladké optimalizace, diskrétních optimalizačních problému, maticové optimalizace a dalších.



Mathematicians Amazed by Russian's Discovery

An unknown Russian discovers a new solution to linear programming problems. But his result goes unnoticed for months

A Russian mathematician, L. G. Khachian, who apparently is unknown to U.S. scientists, has made a discovery of great theoretical, and probably practical, importance. He has found a surprising new way to solve linear programming problems, which are among the most common problems solved on computers. "There is a linear programming package at almost every computer installation," says Ronald Graham of Bell Laboratories. One consequence of Khachian's discovery is related to a major problem in computer science.

A typical linear programming problem faced by a business is to select its sources of supplies so as to minimize its shipping costs. The suppliers may place some constraints on the business by requiring that all orders be at least some minimum size. The resulting problem involves finding the best solution to a set of linear inequalities by considering a function that must be maximized or minimized. In the case of the business, shipping costs must be minimized. Such problems are an everyday part of life for businesses and economists, and they also occur in such fields as engineering, biology, agriculture, and the social sciences. According to Vasek Chvatal of McGill University, linear programming problems commonly involve thousands of inequalities.

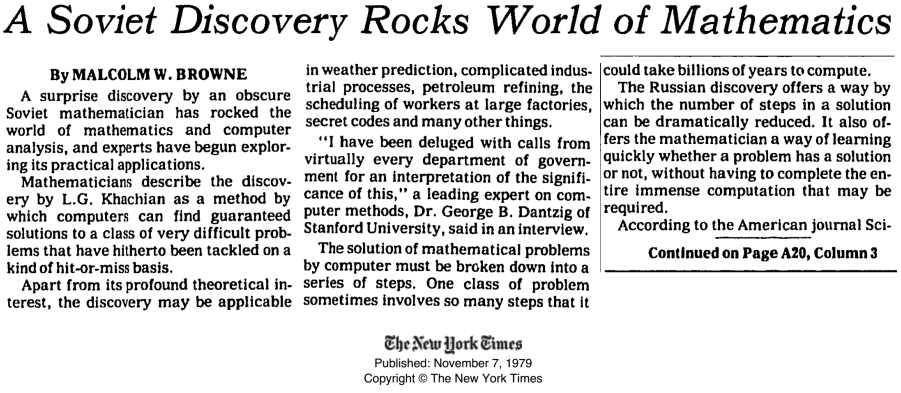
Khachian's discovery went unnoticed for months after its publication last January. Computer scientists and mathematicians explain their missing the result by saying that it appeared as an abstract in the Soviet journal *Doklady*, which they seldom read. Many, in fact, cannot read it because it is published in Russian.

Virtually no U.S. scientists were aware of Khachian's result until early summer, when Eugene Lawler of the University of California at Berkeley brought it to the attention of the computer science community. Lawler heard of it in May while attending a conference in Germany. There Rainer Burkhard of the University of Cologne showed Lawler the abstract and asked him what he knew about it. Lawler, who was unfamiliar with the abstract, had it translated and sent it out to prominent computer scientists, asking if they knew about it. None did, but all were intrigued.

Through Lawler's efforts, Laslo Lovász of the University of Szeged in Hungary and Peter Gács of the University of Rochester heard of the result. These two Hungarian mathematicians were visiting at Stanford and decided to work together in an attempt to reconstruct the mathematical proofs from Khachian's abstract, thereby establishing the validity of the result. Not only did they fill in the proofs but they also improved on the method, making it more efficient. Lovász reports that their task was "medium hard" and that it took him and Gács "a couple of days." But Lovász, at age 29, is a rising

SCIENCE, VOL. 206, 2 NOVEMBER 1979 0036-8075/79/1102-0545$00.50/0 Copyright © 1979 AAAS 545

Obrázek 2.42: Článek z časopisu Science ohledně Chačijanova výsledku (2. listopadu 1979).



A Soviet Discovery Rocks World of Mathematics

By MALCOLM W. BROWNE

A surprise discovery by an obscure Soviet mathematician has rocked the world of mathematics and computer analysis, and experts have begun exploring its practical applications.

Mathematicians describe the discovery by L.G. Khachian as a method by which computers can find guaranteed solutions to a class of very difficult problems that have hitherto been tackled on a kind of hit-or-miss basis.

Apart from its profound theoretical interest, the discovery may be applicable in weather prediction, complicated industrial processes, petroleum refining, the scheduling of workers at large factories, secret codes and many other things.

"I have been deluged with calls from virtually every department of government for an interpretation of the significance of this," a leading expert on computer methods, Dr. George B. Dantzig of Stanford University, said in an interview.

The solution of mathematical problems by computer must be broken down into a series of steps. One class of problem sometimes involves so many steps that it could take billions of years to compute.

The Russian discovery offers a way by which the number of steps in a solution can be dramatically reduced. It also offers the mathematician a way of learning quickly whether a problem has a solution or not, without having to complete the entire immense computation that may be required.

According to the American journal Sci-

Continued on Page A20, Column 3

The New York Times
Published: November 7, 1979
Copyright © The New York Times

Obrázek 2.43: Začátek prvního článku z The New York Times ohledně Chačijanova výsledku (7. listopadu 1979).

Rozruch vyvolaly různé nafouknuté formulace v těchto článcích. Lidé se ptali, jak je možné, že s takový objevem nepřišli Američané, a jak moc to ovlivní jejich životy? Vůbec! Nebylo to nic jiného než další ruský (tentokrát matematický) *sputnik*, což byla první umělá vesmírná družice, která v roce 1957 odstartovala kosmickou éru lidstva a která tehdy šokovala západní svět. Jenže my už víme, že to bylo jenom prázdné gesto, které mělo ilustrovat sovětskou dominanci – sputnik totiž byl jen plechová koule s vysílačkou. Z teoretického pohledu sice je velký rozdíl mezi polynomiálním a exponenciálním výpočetním časem, jenže v praxi se nelze soustředit pouze na tento aspekt. Ukázalo se totiž, že metoda elipsoidů se příliš nehodí k praktickému využití, jelikož je téměř vždy pomalejší než simplexová metoda a výjimkou jsou pouze nejhorší možné případy. Dnes už je velmi těžké říct, jak se to stalo, ale tisk si zkrátka popletl úlohu lineárního programování s mnohem složitějším *problémem obchodního cestujícího* (viz následující Kapitolu 3) a značně zveličil význam celé této metody, jak se můžeme přesvědčit i z Dantzigova rozhovoru s Brownem, který prý probíhal takto:

> "What about the traveling salesman problem?" asked Browne.
> "If there is a connection, I don't know what it is," said Dantzig.
> ("The Russian discovery proposed an approach for [solving] a class of problems related to the 'Traveling Salesman Problem,"' reported Browne.)
> "What about cryptography?" asked Browne.
> "If there is a connection, I don't know what it is," said Dantzig.
> ("The theory of codes could eventually be affected," reported Browne.)
> "Is the Russian method practical?" asked Browne.
> "No," said Dantzig.
> ("Mathematicians describe the discovery ... as a method by which computers can find solutions to a class of very hard problems that has hitherto been attacked on a hit–or–miss basis," reported Browne.)[256]

Matematici volali do The New York Times, psali editorům různé dopisy a dokonce je zvali na přednášky o metodě elipsoidů v IBM, aby informace uvedli na správnou míru. Jenže noviny otiskly ještě tři další články, z nichž ten poslední je na Obrázku 2.44 a sami můžeme posoudit, jak přesné informace obsahuje. Jakási korekce se v The New York Times objevila až 21. března 1980 s titulkem *A Russian's Solution in Math Questioned: Americans Who Studied Khachian Linear Programming Method Express Doubt on Scope*[257]. Nicméně i zde bylo zmíněno (domnělé) propojení s Problémem obchodního cestujícího a možné praktické využití celé metody. Viz také zajímavé články [348, 358].

Jak vlastně metoda elipsoidů funguje? Zatímco simplexová metoda je určena pro řešení jisté třídy optimalizačních úloh (lineárního programování), metoda elipsoidů je určena spíše pro řešení *rozhodovacího problému*[258], ve kterém chceme zjistit, zda nějaká konvexní množina je prázdná či nikoli, přičemž ve druhém případě získáme jako „bonus" i nějaký prvek této množiny. Toho je docíleno tím, že jsou konstruovány elipsoidy $E_0, \ldots, E_k$ se středy $x^{[0]}, \ldots, x^{[k]}$ a v k-tém kroku je testováno, zda odpovídající střed je prvkem dané množiny. Pokud ano, je výpočet ukončen, zatímco v opačném případě je sestrojen nový elipsoid E_{k+1}, který obsahuje pouze polovinu původního elipsoidu E_k a současně má menší objem. Opakováním tohoto iteračního procesu nalezneme po konečném počtu kroků hledaný bod nebo zjistíme, že daná množina je prázdná. Důležité však je, že potřebný počet kroků závisí polynomiálně na velikosti vstupu. Uvažme např. polyedr

$$X := \{x \in \mathbb{R}^n \mid Ax \leq b\}$$

vymezený soustavou nerovnic s maticí $A \in \mathbb{R}^{m \times n}$ a vektorem $b \in \mathbb{R}^m$ s celočíselnými koeficienty a splňující dvojici podmínek (1) polyedr X je ohraničený; (2) polyedr X je buď prázdný nebo $\dim X = n$. První z těchto požadavků je nutný k tomu, abychom dokázali množinu X vnořit do

[253] Vskutku? Popřemýšlejme nad tím.

[254] Ruská cesta s matematickým obchodním cestujícím: Neznámý sovětský matematik L. G. Chačijan nabízí řešení problému, který je pro výpočetní svět záhadou

[255] Malcolm Wilde Browne (17.dubna 1931 – 27. srpna 2012) byl americký novinář a fotograf, který se nejvíce proslavil fotografií sebeupálení buddhistického mnicha Thích Quảng Đức v roce 1963, za kterou v témže roce získal Pulitzerovu cenu.

V roce 1968 nastoupil do The New York Times, kde se z něj stal v roce 1972 jihoamerický dopisovatel. Jelikož dříve pracoval jako chemik, působil od roku 1977 jako vědecký novinář a byl editorem populárně-naučného časopisu Discover. Do NYT se vrátil v roce 1985 a v roce 1991 byl reportérem ve válce v Perském zálivu.

[256] „Co problém obchodního cestujícího?" ptal se Browne. „Je-li tam nějaké spojení, nevím jaké," řekl Dantzig. („Ruský objev skýtá cestu k [řešení] třídy problémů blízkých k 'Problému obchodního cestujícího'," napsal Browne.) „Co kryptografie?" ptal se Browne. „Jestli tam je nějaké spojení, nevím jaké to je," řekl Dantzig („Případně by být ovlivněna i teorie kódování," napsal Browne.) „Je ruská metoda praktická?" ptal se Browne. „Ne," řekl Dantzig. („Matematici popisují tento objev... jako metodu, pomocí níž počítače mohou najít řešení třídy velmi těžkých problémů, jejichž řešení byla až doposud založena na metodě pokus–omyl," napsal Browne. Viz [348, str. 193–194].

[257] *Ruské matematické řešení zpochybněno: Američané, kteří studovali Chačijanovu metodu lineárního programování, vyjádřili pochybnosti.*

[258] Jde o problémy, které lze formulovat jako otázku s jednoznačnou odpovědí ano × ne s nekonečnou množinou vstupů. Tyto úlohy se vyskytují především v informatice, zejména v teorii složitosti a teorii vyčíslitelnosti.

An Approach to Difficult Problems

Mathematicians disagree as to the ultimate practical value of Leonid Khachiyan's new technique, but concur that in any case it is an important theoretical accomplishment.

Mr. Khachiyan's method is believed to offer an approach for the linear programming of computers to solve so-called "traveling salesman" problems. Such problems are among the most intractable in mathematics. They involve, for instance, finding the shortest route by which a salesman could visit a number of cities without his path touching the same city twice.

Each time a new city is added to the route, the problem becomes very much more complex. Very large numbers of variables must be calculated from large numbers of equations using a system of linear programming. At a certain point, the compexity becomes so great that a computer would require billions of years to find a solution.

In the past, "traveling salesmen" problems, including the efficient scheduling of airline crews or hospital nursing staffs, have been solved on computers using the "simplex method" invented by George B. Dantzig of Stanford University.

As a rule, the simplex method works well, but it offers no guarantee that after a certain number of computer steps it will always find an answer. Mr. Khachiyan's approach offers a way of telling right from the start whether or not a problem will be soluble in a given number of steps.

Two mathematicians conducting research at Stanford already have applied the Khachiyan method to develop a program for a pocket calculator, which has solved problems that would not have been possible with a pocket calculator using the simplex method.

Mathematically, the Khachiyan approach uses equations to create imaginary ellipsoids that encapsulate the answer, unlike the simplex method, in which the answer is represented by the intersections of the sides of polyhedrons. As the ellipsoids are made smaller and smaller, the answer is known with greater precision. MALCOLM W. BROWNE

Obrázek 2.44: Další článek z The New York Times od Browneho, viz [84]. Tento vyšel 27. listopadu 1979. Lze s jeho obsahem bez výhrad souhlasit? Rozhodněme (je-li to možné), která tvrzení jsou (a) pravdivá, (b) nepravdivá, (c) zavádějící, (d) ekvivalentní známé domněnce, jejíž řešení nebylo pravděpodobně Browneovi známo. Viz [532].

koule se středem v počátku a poloměrem

$$R := \sqrt{\sum_{i=1}^{n} \max\{|m_i|,|M_i|\}^2}, \tag{2.13.1}$$

kde m_i a M_i jsou nějaké dolní a horní hranice pro hodnoty i-té proměnné, tj. $m_i \le x_i \le M_i$ pro všechna $x = [x_1,\ldots,x_n] \in X$ a $i = 1,\ldots,n$. Pokud tato omezení neznáme, je možné použít kouli s poloměrem

$$R := \sqrt{n}\,2^{\langle A\rangle+\langle b\rangle-n^2}, \tag{2.13.2}$$

kde symboly $\langle A\rangle$ a $\langle b\rangle$ vyjadřují součet počtu znaků (0 a 1, tj. bitů) potřebných pro reprezentaci jednotlivých složek v binárním kódu, tj.

$$\langle A\rangle := \sum_{i,j=1}^{n,m} \langle a_{ij}\rangle \quad \text{a} \quad \langle b\rangle := \sum_{i=1}^{m} \langle b_i\rangle$$

s tím, že počet potřebných bitů pro libovolné $z \in \mathbb{Z}$ je určený vztahem[259]

$$\langle z\rangle := 1 + \lceil \log_2(|z|+1)\rceil. \tag{2.13.3}$$

Druhý předpoklad zaručuje, že v případě $X \neq \emptyset$ je objem množiny X kladný, tj. existuje n-dimenzionální koule s poloměrem $r > 0$, která je podmnožinou X. Dokonce se dá ukázat, že pro tento objem platí

$$\operatorname{vol} X \ge 2^{-(n+1)\langle A\rangle+n^3}, \tag{2.13.4}$$

díky čemuž můžeme omezit počet potřebných kroků.

[259] Pomocí n-tice bitů můžeme v binárním kódu reprezentovat čísla od 0 až po $2^0+2^1+2^2+\cdots+2^{n-1} = 2^n-1$. Proto, chceme-li takto reprezentovat kladné číslo $z \in \mathbb{N}\cup\{0\}$, potřebujeme dostatečný počet bitů, tj.

$$2^n - 1 \ge z$$

neboli

$$n \ge \log_2(z+1),$$

a tudíž potřebný počet bitů je $\lceil \log_2(z+1)\rceil$. Chceme-li rozšířit tuto reprezentaci i o záporná celá čísla, pak potřebujeme ještě jeden bit pro znaménko, jak je uvedeno v (2.13.3).

Základní metoda elipsoidů pro množinu X

(1) *Inicializace*: je dán polyedr X, stanovíme poloměr R a maximální počet iterací $N := 2n\left[(2n+1)\langle A\rangle + n\langle b\rangle - n^3\right]$, položíme $k := 0$, $A_0 := R^2 I \in \mathbb{R}^{n\times n}$ a $x^{[0]} := 0 \in \mathbb{R}^n$.

(2) Je-li $k = N$, pak množina X je prázdná. KONEC. V opačném případě pokračujeme krokem (3).

(3) Otestujeme danou soustavu nerovností pro $x^{[k]}$, tj. zda platí $Ax^{[k]} \le b$ a určíme množinu indexů $I_k \subseteq \{1,\dots,m\}$ odpovídajících těm nerovnostem $a_i^\top x^{[k]} \le b_i$, které nejsou splněny.

(4) Je-li $I_k = \emptyset$, pak $x^{[k]}$ je přípustný bod, $x^{[k]} \in X$. KONEC. V opačném případě pokračujeme krokem (5).

(5) Zvolíme libovolného reprezentanta $i_k \in I_k$ a definujeme

$$x^{[k+1]} := x^{[k]} - \frac{1}{n+1}\,\frac{A_k a_{i_k}}{\sqrt{a_{i_k}^\top A_k a_{i_k}}}$$

$$A_{k+1} := \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1}\,\frac{A_k a_{i_k} a_{i_k}^\top A_k}{a_{i_k}^\top A_k a_{i_k}}\right).$$

Položíme $k := k+1$ a pokračujeme krokem (2).

Navzdory našim předchozím slovům o srovnání metody elipsoidů a simplexové metody neměli bychom opomenout dvě zajímavé skutečnosti: (i) algoritmický popis metody elipsoidů je mnohem jednodušší než v případě simplexové metody; (ii) uvedené vzorce jsou výpočetně celkem snadné, takže už v 80. letech minulého století bylo možné naprogramovat tuto metodu na kapesních kalkulátorech až pro 6 proměnných – tohle rozhodně simplexová metoda neumožňovala! Také bychom měli zdůraznit, že ačkoli v uvedeném algoritmu žádný elipsoid nefiguruje, rozhodně je ukryt v jádru samotné metody a můžeme jej vyjádřit jako

$$E_k := \left\{x \in \mathbb{R}^n \mid (x - x^{[k]})^\top A_k^{-1}(x - x^{[k]}) \le 1\right\}.$$

V kroku (5) je sestrojen nový elipsoid E_{k+1}, který obsahuje polovinu elipsoidu E_k určenou nerovností $a_{i_k}^\top x \le b_{i_k}$, jenže je porušena bodem $x^{[k]}$, tj.

$$E_{k+1} \supseteq E_k \cap \left\{x \in \mathbb{R}^n \mid a_{i_k}^\top x \le a_{i_k}^\top x^{[k]}\right\}, \tag{2.13.5}$$

což zaručuje, že E_{k+1} obsahuje celý daný polytop X. Navíc tento elipsoid E_{k+1} má nejmenší možný objem mezi všemi elipsoidy splňujícími 2.13.5. Klíčem k důkazu toho, že tento algoritmus běží v polynomiální čase, je vztah mezi objemy[260] těchto elipsoidů

$$\frac{\operatorname{vol} E_{k+1}}{\operatorname{vol} E_k} = \sqrt{\left(\frac{n}{n+1}\right)^{n+1}\left(\frac{n}{n-1}\right)^{n-1}} < \exp\left(\frac{-1}{2n}\right) < 1.$$

To znamená, že elipsoid E_{k+1} má vždy menší objem než E_k, i když pro rostoucí n se tento rozdíl zmenšuje. Každopádně můžeme zhora odhadnout objem prvotního elipsoidu E_0 a také máme dolní odhad objemu množiny X, viz (2.13.4). Proto díky postupnému zmenšování objemu jednotlivých elipsoidů víme, že má smysl v tomto procesu pokračovat pouze do té doby, než $\operatorname{vol} E_{k+1}$ bude menší než zmíněný dolní odhad $\operatorname{vol} X$. Pokud se nám to „podaří“, pak není možné, aby $X \subseteq E_{k+1}$, a tudíž nutně $X = \emptyset$. Proto číslo N v inicializaci elipsoidové metody bylo zvoleno jako nejhorší možný případ, neboť po tomto počtu kroků jistě platí

$$\operatorname{vol} E_N < 2^{-(n+1)\langle A\rangle + n^3} \le \operatorname{vol} X.$$

[260] Např. v $\mathbb{R}^2$ pro obsah plochy vymezené kuželosečkou

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

platí vzorec

$$\pi\,\frac{I_1 I_3 - I_4 - 2(I_1^2 - I_2) I_5}{(I_1^2 - I_2)^{3/2}},$$

kde

$$I_1 := A + C$$
$$I_2 := (A-C)^2 + B^2$$
$$I_3 := D^2 + E^2$$
$$I_4 := (A-C)(D^2 - E^2) + 2DEB$$
$$I_5 := F$$

jsou tzv. *invarianty* dané kuželosečky.

Tato hodnota N je odvozena pomocí poměrně obecných nerovností a s trochou pečlivější analýzy by bylo možné ji o něco snížit. Nicméně to nic nezmění na tom, že počet kroků v metodě elipsoidů bude v nejhorším případě řádově $O(n^2(\langle A\rangle+\langle b\rangle))$ neboli dříve zmíněné $O(n^2 L)$, což dokazuje polynomiální časovou složitost této metody.

Věta 2.13.3

Časová složitost metody elipsoidů je $O(n^2 L)$, tj. polynomiální.

Pro více podrobností včetně rozšíření metody elipsoidů i na případ, kdy množina X nesplňuje některý z uvedených předpokladů, doporučujeme [165, Chapter 5], [238, Chapter 3], [312, Chapter 4] nebo [408, Section 8.7].

Příklad 2.13.4

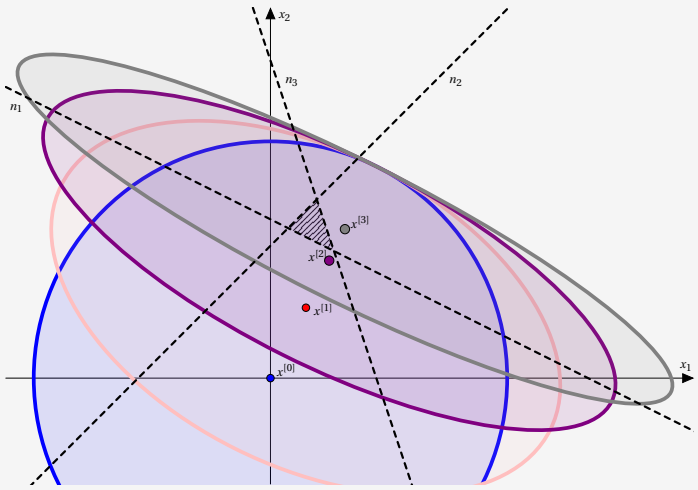
Pomocí metody elipsoidů nalezněme nějaké řešení soustavy nerovností

$$-x_1-2x_2\le -6 \quad \& \quad -2x_1+2x_2\le 5 \quad \& \quad 3x_1+x_2\le 6.$$

Řešení. Ze zadání máme

$$A=\begin{pmatrix}-1 & -2\\ -2 & 2\\ 3 & 1\end{pmatrix} \quad \text{a} \quad b=\begin{pmatrix}-6\\ 5\\ 6\end{pmatrix},$$

takže $\langle A\rangle = 15$ a $\langle b\rangle = 12$, což nás přivede k maximálnímu nutnému počtu kroků $N = 364$ a podle vzorce (2.13.2) k výchozímu elipsoidu jakožto kružnici se středem v počátku a poloměrem $R = 8388608\sqrt{2}$. Vzhledem k tomu, že obsah tohoto elipsoidu je cca $4,42\times 10^{14}$ a každým krokem se zmenšuje s faktorem $\sqrt{2(2/3)^3}\approx 0,7698$ (tj. cca o 23 procent), bude potřeba zhruba 114 kroků[261] (přesný počet může záviset na volbě indexů i_k a zaokrouhlovacích chybách, ale výpočet zabere jen několik sekund). Na to tu však rozhodně nemáme místo, a tak raději využijeme „nápovědu“ s omezením pro jednotlivé proměnné $0\le x_1\le 2$ a $0\le x_2\le 4$, z čehož díky vzorci (2.13.1) dostaneme poloměr výchozí kružnice $R=\sqrt{20}$. V takovém případě nalezneme hledané řešení po pouhých 5 krocích, jejichž průběh je znázorněn Obrázcích 2.45–2.46.

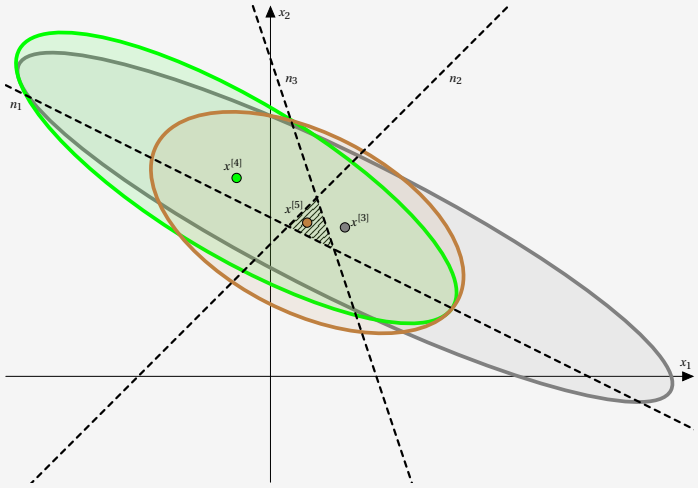


Obrázek 2.45: První tři kroky metody elipsoidů.

Na Obrázku 2.45 jsou zachyceny první tři kroky metody elipsoidů. Modrou barvou je znázorněna výchozí kružnice E_0 s obsahem cca 62,83 a její střed $x^{[0]}$. Růžovou barvou je znázorněn elipsoid E_1 s obsahem cca 48,37 a jeho střed $x^{[1]}$. Bod $x^{[0]}$ nesplňuje první nerovnost a elipsoid E_1 obsahuje tu část E_0, která této nerovnosti vyhovuje. Fialovou barvou je znázorněn elipsoid E_2 s obsahem cca 37,23 a jeho střed $x^{[2]}$. Bod $x^{[1]}$ nesplňuje první nerovnost a elipsoid E_2 obsahuje tu část E_1, která této nerovnosti vyhovuje.

[261] Kolik kroků by to bylo se simplexovou metodou? Vzhledem k absenci omezení na znaménko jednotlivých proměnných bychom po převodu do kanonického tvaru měli polyedr s nejvýše 35 BPB, takže maximální možný počet kroků je výrazně menší (a nebyla by ani potřeba první fáze simplexového algoritmu). Na druhou stranu by v tomto kontextu asi bylo trochu férovější srovnání početní náročnosti jednotlivých kroků, neboť úprava jedné simplexové tabulky rozhodně vyžaduje více početních operací (sčítání a násobení) než jeden krok v metodě elipsoidů (násobení vektorů a matic a jednu odmocninu, která však s velkou pravděpodobností způsobí iracionalitu některých koeficientů). V každém kroku vyžaduje simplexová metoda řádově $O(nm)$ aritmetických operací, zatímco v metodě elipsoidů to je $O(n^2 L)$.

Šedou barvou je znázorněn elipsoid E_3 s obsahem cca 28,66 a jeho střed $x^{[3]}$. Bod $x^{[2]}$ nesplňuje první nerovnost a elipsoid E_3 obsahuje tu část E_1, která této nerovnosti vyhovuje. Za pozornost stojí, že ve všech třech případech byla porušena první nerovnost, což vede k tomu, že elipsoidy se postupně zplošťují a protahují. Toto chování může pro „ploché“ polytopy způsobovat v praxi numerické potíže, jelikož elipsoidy mohou být velmi „placaté“ (třeba dlouhé až několik kilometrů a tenké jen několik milimetrů).



Obrázek 2.46: Další dva kroky metody elipsoidů.

Na Obrázku 2.46 jsou další dva kroky metody elipsoidů. Šedou barvou je opět znázorněn elipsoid E_3 a jeho střed $x^{[3]}$. Zelenou barvou je znázorněn elipsoid E_4 s obsahem cca 22,06 a jeho střed $x^{[4]}$. Bod $x^{[3]}$ nesplňuje třetí nerovnost a elipsoid E_4 obsahuje tu část E_3, která této nerovnosti vyhovuje. Hnědou barvou je znázorněn elipsoid E_5 s obsahem cca 16,99 a jeho střed $x^{[5]}$. Bod $x^{[4]}$ nesplňuje první nerovnost a elipsoid E_5 obsahuje tu část E_4, která této nerovnosti vyhovuje. Bod $x^{[5]}$ již vyhovuje všem daným nerovnostem a výpočet je u konce.

Už bývá jen objasnit, jak tuto metodu využít při řešení úloh lineárního programování. To je založeno na teorii duality, odkud víme, že (i) primární úloha je řešitelná právě tehdy, když duální úloha je řešitelná, a (ii) pro libovolný primárně přípustný bod x a duálně přípustný bod y platí $\langle c, x\rangle \geq \langle b, y\rangle$, přičemž nastane-li rovnost, máme řešení obou úloh. Díky tomu můžeme řešení úlohy lineárního programování

$$\langle c, x\rangle \to \min \quad \& \quad Ax \geq b \quad \& \quad x \geq 0$$

převést na řešení soustavy nerovnic[262]

$$\begin{aligned} -Ax &\leq -b \\ -x &\leq 0 \\ A^\top y &\leq c \\ -y &\leq 0 \\ \langle c, x\rangle - \langle b, y\rangle &\leq 0, \end{aligned}$$

ve které je $n+m$ neznámých a $(2(n+m)+1$ lineárních nerovnic. Tento poměrně vysoký počet neznámých nejspíš způsobí zpomalení konvergence metody elipsoidů. Navíc z nulového duální rozdílu vyplývá, že hledané řešení bude ležet na nadrovině $\langle c, x\rangle - \langle b, y\rangle = 0$, takže pro řešitelné úlohy tato soustava nesplňuje základní požadavek $\dim X = n+m$ a bude nutné využít jisté zobecnění metody elipsoidů. Avšak nic z toho nezmění její polynomiální časovou složitost. Nevýhodou také může být to, že v případě neřešitelnosti této soustavy se nedozvíme, zda primární úloha je neohraničená nebo nepřípustná.

[262] Odpovídající duální úloha je

$$\langle y, b\rangle \to \max$$
$$A^\top y \leq c \quad \& \quad y \geq 0.$$

Matematický svět každopádně mohl být spokojen – otázka existence algoritmu s polynomiální časovou složitostí pro lineární programování byla zdárně vyřešena. Nebo ne? Ne! Nalezená odpověď totiž příliš nesplňovala očekávání, ve kterých bylo implicitně schováno to, že nalezený algoritmus bude i z praktického pohledu lepší než simplexová metoda. A to se nestalo. Proto hledání pokračovalo dále až do roku 1984, kdy Karmarkar[263] tehdy pracující v Bellových laboratořích v New Jersey[264] zveřejnil pro úlohy lineárního programování algoritmus patřící do skupiny *metod vnitřních prvků*[265] a dokázal, že doba jeho běhu je polynomiální, viz [314,315]. Řádově jde o $O(nL)$ iterací a $O(n^{3,5}L^2)$ aritmetických operací v každém kroku, k čemuž je nutné dodat, že metoda funguje pouze pro úlohy lineárního programování v tzv. *Karmarkarově standardním tvaru*, tj.

$$\left.\begin{aligned} \langle c,x\rangle \to \min \quad \& \quad x\in X := P\cap S, \\ \text{kde} \quad P := \{x\in\mathbb{R}^n \mid Ax=0\} \quad \text{a} \quad S := \{x\in\mathbb{R}^n \mid x_1+\cdots+x_n=1 \;\&\; x\geq 0\} \end{aligned}\right\} \tag{2.13.6}$$

s maticí $A\in\mathbb{R}^{n\times m}$ mající plnou hodnost a společně s vektorem $c\in\mathbb{R}^n$ pouze celočíselné koeficienty, tj. přípustná množina je průnik polyedru P a simplexu S, což bude polytop. Navíc pro úlohu (2.13.6) platí:

(a) minimální hodnota této úlohy je 0, tj. $\langle c,x\rangle \geq 0$ pro všechna $x\in X$ a pro nějaké $x^*\in X$ je $\langle c,x^*\rangle = 0$;

(b) $A\mathbf{e}=0$ a $\langle c,\mathbf{e}\rangle \neq 0$ pro vektor $\mathbf{e} := (1,\ldots,1)^\top \in\mathbb{R}^n$, tj. bod $[1/n,\ldots,1/n]$ je přípustný a současně není hledaným řešením.

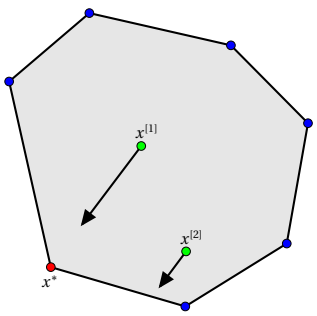
Všimněme si, že duální úlohou k (2.13.6) je nalezení $y\in\mathbb{R}^m$ a $z\in\mathbb{R}$ takových, že

$$z\to\max \quad \& \quad A^\top y + z\mathbf{e}\leq c. \tag{2.13.7}$$

Tato úloha je vždy přípustná, neboť pro libovolné $\widehat{y}\in\mathbb{R}^m$ určitě dokážeme najít $z\in\mathbb{R}$ takové, že dané omezení je splněno – tou nejlepší volbou je $\min_{j=1,\ldots,n}(c-A^\top\widehat{y})_j$. Navíc nulová hodnota úlohy (2.13.6), a tudíž také duální úlohy (2.13.7), znamená, že skutečně hledáme $y^*\in\mathbb{R}$ splňující $A^\top y^*\leq c$, což je přesně úloha, pro kterou byla sestavena metoda elipsoidů.

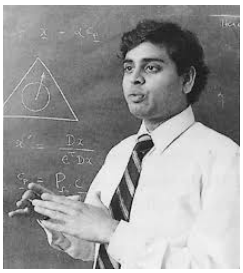
Základní princip Karmarkarovy metody bychom mohli popsat podobně jako v případě simplexové metody (a svým způsobem by tak šlo zpětně interpretovat i metodu elipsoidů), kde se z výchozího (a každého dalšího) BPB posuneme vhodným směrem mířícím do dalšího BPB. Jenže tento pohyb po hranici je velmi limitující, neboť máme jen velmi omezený počet vhodných směrů. Co kdybychom se raději pohybovali uvnitř daného polyedru? To už bychom sice měli na výběr vhodných směrů mnohem více, ale o to hůře by se nám určoval ten správný. Jenže Karmarkar si všiml dvou velmi důležitých vlastností:

(i) Pokud je bod uvnitř polytopu blízko jeho „středu“, má smysl se posunout ve směru největšího poklesu účelové funkce (tj. ve stejném směru, jakým klesají vrstevnice, což je až na znaménko shodné s gradientem účelové funkce), viz Obrázek 2.47.



Obrázek 2.47: Ilustrace prvního Karmarkarova postřehu. Bod $x^{[1]}$ je blízko středu polyedru a hodnotu účelové funkce můžeme výrazně zmenšit, pokud se „posuuneme“ ve směru největšího spádu (bod x^* je skutečným řešením). Bod $x^{[2]}$ je již od středu vzdálen mnohem více, a pokud se „pohneme“ ve směru největšího spádu, tak se brzy dostaneme mimo přípustnou množinu.

263



Narendra Krishna Karmarkar (∗ 1955?) je indický matematik. V roce 1988 obdržel *Fulkerson Prize* za článek věnovaný metodě vnitřních prvků. V roce 2000 získal *ACM Paris Kanellakis Theory and Practice Award* udělovanou *Association for Computing Machinery* (Asociace pro výpočetní techniku, což je nejstarší a největší vědecká a vzdělávací společnost zaměřená na informatiku) za svoji práci na metodě vnitřních prvků pro lineární programování.

[264] Tehdy to byly *AT&T Bell Laboratories*. Zde začal pracovat na podzim roku 1983 a v tomtéž roce absolvoval i postdoktorský výzkumný pobyt v IBM. S afiliací k IBM také Karmarkar svoji metodu prezentoval 11. srpna 1983 na semináři na Stanfordově univerzitě. Jenže v roce 1984 publikoval svůj výsledek v článku, kde již měl jako afiliaci uvedeny Bellovy laboratoře. Vysoká efektivita Karamakarova algoritmu se potvrdila ve chvíli, kdy jej Bellovy laboratoře využili v telefonní síti AT&T, což ihned vedlo k následné žádosti o patent (duben 1985). Bellovy laboratoře jej implementovaly do výpočetního systému, který následně prodávaly za 8,9 milionů dolarů – prvním zákazníkem byl Pentagon. Nicméně tímto se znovu otevřela právní (možná i filosofická) otázka, zda je vůbec možné něco takového patentovat (např. matematické vzorce rozhodně není možné patentovat). Vedla se k tomu dokonce i velmi zevrubná debata v americkém senátu a v květnu 1988 byl patent skutečně udělen (U.S. Patent 4,744,028: „Methods and apparatus for efficient resource allocation“). Platnost patentu vypršela v roce 2006 a od té doby je algoritmus volně k dispozici.

[265] Simplexová metoda i metoda elipsoidů naopak patří mezi *metody vnějších prvků*. Zatímco v simplexové metodě jsou generovány body ležící na hranici polyedru (vrcholy) a metodou elipsoidů body ležící vně přípustné množiny (kromě posledního bodu), tak Karmarkarova metodou získáme posloupnost bodů ležících uvnitř přípustné množiny (kromě posledního bodu).

(ii) Aniž bychom jakkoli podstatně změnili daný problém, je možné jeho vhodnou transformací docílit toho, že zvolený vnitřní bod bude blízko středu transformovaného polytopu.

K druhému poznatku dospěl Karmarkar pomocí tzv. *projektivní transformace*, při které přímky zůstávají přímkami, ale úhly a vzdálenosti se deformují tak, že jakýkoli bod původního polytopu se můžeme stát středem nového polytopu (s trochou představivosti bychom toho mohli snad i docílit pro bod $x^{[2]}$ na Obrázku 2.47). Tato transformace nezmění nic podstatného na dané úloze – pouze se na přípustnou množinu budeme dívat z trochu jiného úhlu pohledu. Tím jsou určeny hlavní kroky Karmarkarova algoritmu: (i) vezmeme přípustný bod uvnitř polytopu, (ii) transformujeme polytop tak, aby tento bod byl v jeho středu, (iii) „posuneme“ se ve směru největšího spádu, ale ne až na hranici – potřebujeme být stále uvnitř polytopu (iv) a inverzní transformací získáme přípustný bod v původním polytopu. Tento proces opakujeme tak dlouho, dokud nenalezneme řešení (případně nějaký jeho velmi kvalitní odhad). Ačkoli se opět jedná o jakýsi iterativní proces, tentokrát je sestaven tak, že nalezený bod můžeme brát současně jako nový výchozí bod. To znamená, že tento algorimus je v jistém smyslu *autokorekční*, tj. případné numerické (či zaokrouhlovací) chyby se nekumulují a v dalším kroku budou opraveny (pouze nesmí být porušena přípustnost uvažovaného bodu). To je přesný opak metody elipsoidů, kde se numerické chyby postupně zvětšovaly.

Karmarkarův algoritmus metody vnitřních prvků pro úlohu (2.13.6)

(1) *Inicializace*: je dána úloha (2.13.6), stanovíme číslo L, zvolíme $\alpha \in (0,1]$, $r := 1/\sqrt{n(n-1)}$, položíme $k := 0$ a definujeme $x^{[0]} := \frac{1}{n}\mathbf{e}$.

(2) Je-li
$$\langle c, x^{[k]} \rangle \le 2^{-L} \langle c, x^{[0]} \rangle,$$
pak $x^* := x^{[k]}$ je hledaným řešením. KONEC. V opačném případě pokračujeme krokem (3).

(3) Definujeme
$$X^{[k]} := \operatorname{diag}\{x_1^{[k]}, \dots, x_n^{[k]}\}$$
a rozšíříme matici $AX^{[k]}$ o řádek obsahující samé jedničky, tj.
$$B^{[k]} := \begin{pmatrix} AX^{[k]} \\ \mathbf{e}^\top \end{pmatrix} \in \mathbb{R}^{(n+1)\times n}.$$
Dále definujeme
$$d^{[k]} := -\left[I - B^{[k]\top}\left(B^{[k]}B^{[k]\top}\right)^{-1}B^{[k]}\right]X^{[k]}c$$
$$y^{[k+1]} := \frac{1}{n}\mathbf{e} + \frac{\alpha r}{\sqrt{d^{[k]\top}d^{[k]}}}d^{[k]}$$
$$x^{[k+1]} := \frac{1}{\mathbf{e}^\top X^{[k]} y^{[k+1]}} X^{[k]} y^{[k+1]}.$$
Položíme $k := k+1$ a pokračujeme krokem (2).

Body $x^{[1]}, x^{[2]}, \dots$ generované tímto algoritmem jsou přípustné a (kromě samotného řešení) leží vždy uvnitř množiny X. Vektor $d^{[k]}$ je směr největšího spádu transformované úlohy a $y^{[k+1]}$ je jejím novým přípustným bodem. Zpětnou transformací tohoto bodu získáme nový přípustný bod původní úlohy $x^{[k+1]}$. Číslo L je voleno jako velikost vstupních dat podobně jako v případě metody elipsoidů. Lze také zvolit nějaký jeho násobek tak, aby $2^{-L} < \varepsilon$ pro danou přesnost ε. Každopádně lze ukázat, že pro všechna k platí
$$\langle c, x^{[k]} \rangle \le \mathrm{e}^{-k/5n} \langle c, x^{[0]} \rangle.$$

To znamená, že po $O(nL)$ krocích nalezneme přípustný bod splňující

$$\langle c, x^{[k]} \rangle = 0 \quad \text{nebo} \quad \langle c, x^{[k]} \rangle / \langle c, x^{[0]} \rangle < 2^{-L},$$

což potvrzuje polynomiální časovou složitost. Chceme-li najít přesné řešení, je nutné provést navíc ještě tzv. očištění[266], které má také polynomiální časovou složitost. Číslo r vyjadřuje poloměr největší koule se středem v $x^{[0]}$ vepsané do simplexu S, zatímco α vyjadřuje délku kroku. Hodnotu α lze volit různými způsoby (např. konstantně jako $(n-1)/(3n)$ nebo v každém kroku jistou minimalizací, aby došlo k největšímu poklesu funkční hodnoty). Karmarkar se zaměřil na $\alpha = 1/4$, pro které „analyzoval“ postupné zlepšování funkčních hodnot. Avšak z praktického pohledu se ukazuje jako výhodnější volit větší hodnoty α, čímž dojde k urychlení výpočtu.

[266] Angl. *purification scheme.*

Věta 2.13.5

Časová složitost Karmarkarovy metody je $O(nL)$, tj. polynomiální.

Jak převést úlohu lineárního programování do tvaru (2.13.6) a mnoho dalších podrobností ohledně Karmarkarovy metody se lze dočíst v původních článcích [314,315] a také např. v [165, Chapter 6], [312, Chapter 4], [225, Sekce 9] nebo [5,138,176,225,274,420,421,447,480].

Příklad 2.13.6

Pomocí Karmarkarova algoritmu vyřešme úlohu

$$x_1 - 3x_3 + 3x_3 \to \min$$
$$x_1 - 3x_2 + 2x_3 = 0 \quad \& \quad x_1 + x_2 + x_3 = 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení. Ze zadání máme

$$c = \begin{pmatrix} 1 & -3 & 3 \end{pmatrix}^\top \quad \text{a} \quad A = \begin{pmatrix} 1 & -3 & 2 \end{pmatrix}$$

a není těžké ověřit, že úloha splňuje všechny výše zmíněné požadavky. Přesným řešením je $x^* = [3/4, 1/4, 0]$. Za výchozí bod vezmeme $x^{[0]} = [1/3, 1/3, 1/3]$ a zvolíme $L = \langle A \rangle + \langle c \rangle + \langle \mathbf{e} \rangle = 16$. Pak v prvním kroku získáme

$$X^{[0]} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}, \quad B^{[0]} = \begin{pmatrix} 1/3 & -1 & 2/3 \\ 1 & 1 & 1 \end{pmatrix}, \quad d^{[0]} \approx \begin{pmatrix} 0,1587 \\ -0,0317 \\ -0,1269 \end{pmatrix},$$

$$y^{[1]} \approx \begin{pmatrix} 0,4383 \\ 0,3123 \\ 0,2493 \end{pmatrix} \quad \text{a} \quad x^{[1]} = y^{[1]} \approx [0,4383; 0,3123; 0,2493].$$

Ve druhém kroku dostaneme

$$X^{[1]} \approx \begin{pmatrix} 0,4383 & 0 & 0 \\ 0 & 0,3123 & 0 \\ 0 & 0 & 0,2493 \end{pmatrix}, \quad B^{[1]} \approx \begin{pmatrix} 0,4383 & -0,9370 & 0,4986 \\ 1 & 1 & 1 \end{pmatrix},$$

$$d^{[1]} \approx \begin{pmatrix} 0,1244 \\ -0,0052 \\ -0,1192 \end{pmatrix}, \quad y^{[2]} \approx \begin{pmatrix} 0,43155 & 0,3292 \\ 0,2392 & \end{pmatrix} \quad \text{a} \quad x^{[2]} \approx [0,5379; 0,2924; 0,1696].$$

Po 21 krocích dostaneme bod

$$x^{[21]} \approx \left[0,749993; 0,2500012; 4,869 \times 10^{-6}\right],$$

což už můžeme považovat za celkem dobrou aproximaci hledaného řešení. Zvolíme-li $\alpha = 8/9$, skončí výpočet již po 6 krocích s hodnotou

$$x^{[6]} \approx \left[0,7499995; 0,2500000997; 3,98 \times 10^{-7}\right].$$

Grafická ilustrace průběhu nám zde příliš nepomůže, jelikož jednotlivé iterace budou ležet na úsečce, která je vznikne průnikem obou rovin určených danými omezeními a prvního oktantu.

Karmarkarův výsledek, o kterém referoval na konferenci *16th Annual ACM Symposium on Theory of Computing* v dubnu 1984, vyvolal velký zájem matematické komunity a neunikl ani pozornosti tisku – tentokrát však po právu (viděno zpětně) a bez chyb, jak se můžeme přesvědčit v článku *A fast way to solve hard problems: A new algorithm to solve linear programming problems is so fast that experts in the field are taken aback*[267] z časopisu *Science* ze září 1984 (viz [329]) nebo z článku na titulní stránce *The New York Times* z 19. listopadu s názvem *Breakthrough in Problem Solving*[268] (viz [217]) nebo *Linear programming made simpler than simplex*[269] z časopisu *New Scientist* z prosince 1984 (viz [413]) a *Folding the perfect corner*[270] z časopisu *Time* z prosince 1984 (viz [11]). Následně Karmarkar vystoupil jako zvaný řečník na dvou mezinárodních konferencích (*ORSA/TIMS-meeting* v Dalasu v listopadu 1984 a *12th International Symposium on Mathematical Programming* v srpnu 1985 na MIT). Karmarkar zde představil svoji metodu a její varianty, vysvětlil některé praktické triky, poukázal na lepší vlastnosti oproti simplexové metodě (v listopadu 1984 tvrdil, že je 10× až 50× rychlejší než simplexový algoritmus; v dubnu 1985 už dokonce šlo o čtyřnásobné až 250-násobné zrychlení) a poskytl několik málo srovnání, avšak odmítl dát k dispozici kompletní informace o testovacích problémech, počítačových programech a době běhu. To vyvolávalo značnou nejistotu a vedlo k mnoha diskuzím mezi „matematickými programovači“ ohledně praktičnosti takové metody, což lze vycítit i z některých komentářů jako např. v článku *Founding father of just a footnote?*[271] zveřejněném 9. srpna 1985 v *Boston Globe*.

> *This week in Cambridge, the 28-year-old Karmarkar came under mounting fire from his colleagues at the 12th International Symposium on Mathematical Programming. They snorted at his scientific manners, scoffed at his claims and derided his results as being everything from 'frisky' to 'majestic'. Mostly, they said that his accounts of super-fast solutions to difficult problems couldn't be replicated…*
> *…"He may have some wonderful method after all, but I habitually mistrust all secret mathematics", said E. M. L. Beale, a pioneer in the commercial applications of linear programming…*
> *…Karmarkar himself didn't advance his cause much in a talk before an unusual plenary session of the MIT meeting of some 800 scientists from around the world. He began by observing that while mathematicians agree on what constitutes convincing proof of a mathematical proposition, there is no corresponding consensus as to what makes a persuasive presentation of experimental results – a contention that was immediately disputed by many of his listeners.*[272]

Zkrátka celý algoritmus byl opředen tajemstvím, o které se Karmarkar a především Bellovy laboratoře nechtěli (zjevně z komerčních důvodů) podělit. Celé této situaci vůbec nepomáhalo ani to, že se ostatním matematikům nedařilo Karmarkarovy výpočty zopakovat, viz také článek [156] s výstižným názvem *Karmarkar's algorithm: A threat to simplex? An elegant new method of solving linear equations has yet to be thoroughly tested*[273]. Nicméně brzy se ukázalo, že tento algoritmus je skutečně velmi efektivní a z praktického hlediska přináší značné úspory nákladů na výpočet. Počátky metod vnitřního bodu lze dohledat až k Frischovým[274] pracím [189, 190] a následně ke snahám o jejich využití při řešení úloh nelineárního programování v 60-tých let 20. století, viz [178]. Jenže metoda byla aplikována na velmi obecné úlohy, pro které leckdy chyběly i některé důležité teoretické poznatky. Fiacco a McCormick ve své knize také přiznávají, že ačkoli jejich metodu lze aplikovat i na úlohy lineárního programo-

[267] *Rychlý způsob řešení těžkých problémů: Nový algoritmus k řešení problémů lineárního programování je tak rychlý, že jsou odborníci v oboru zaskočeni*

[268] *Průlom v řešení problémů*

[269] *Lineární programování učiněno jednodušším než simplex* (tento pokus o slovní hříčku asi nejde do češtiny přenést)

[270] *Skládání dokonalého rohu*

[271] *Otec zakladatel pouhého doplňku?*

[272] *Tento týden v Cambridge se 28letý Karmarkar dostal na 12. mezinárodním sympoziu o matematickém programování pod palbu svých kolegů. Pohrdavě se vyjadřovali o jeho vědeckých způsobech, vysmívali se jeho tvrzením a posmívali se jeho výsledkům, které jsou čímkoli od „živelného“ až po „impozantní“. Většinou říkali, že jeho ukázky superrychlých řešení obtížných problémů nemohou být opakovány.*
„Možná má nějakou úžasnou metodu, ale já obvykle nedůvěřuji žádne matematice opředené tajemnstvím,“ řekl E. M. L. Beale, průkopník v komerčních aplikacích lineárního programování.
Celé záležitosti příliš nepomohl ani sám Karmarkar svým proslovem před neobvyklým valným shromážděním na MIT čítajícím asi 800 vědců z celého světa. Začal poznámkou, že zatímco matematici se shodují na tom, co tvoří přesvědčivý důkaz matematického tvrzení, neexistuje žádná podobná shoda na tom, jak přesvědčivě prezentovat experimentální výsledky – tvrzení, které bylo okamžitě zpochybněno mnoha posluchači. Viz [447, str. 3].

[273] *Karmarkarův algoritmus: simplexy v ohrožení? Elegantní metoda řešení lineárních rovnic ještě musí být důkladně otestována*

vání, nebudou se rozvoji v tomto směru více věnovat – měli totiž pocit, že by se rozhodně nemohla vyrovnat simplexovému algoritmu. Do tohoto rozvoje (ať již v návaznosti na předchozí výsledky či nikoli) včetně analýzy časové náročnosti se pustil až Karmarkar, za což mu nelze odepřít zásluhy. Na druhou stranu počáteční tajemství přispěla k tomu, že se objevily i různé jiné varianty, avšak některé z nich již mohou postrádat polynomiální časovou náročnost. Za zmínku stojí spojení s *Newtonovou hraniční metodou*[275] a především s metodou založenou na tzv. *afinní transformaci*, viz [125, Chapter 4], [165, Chapter 7] nebo [154, Section 7.2]. Tato druhá metoda byla publikována ve dvou nezávislých článcích [29, 493] z roku 1986, po jejichž zveřejnění dorazil do redakce prvního z časopisů mírně nespokojený dopis upozorňující na to, že tato metoda až tak nová rozhodně není, viz [137]. Ozval se totiž bývalý Kantorovičův student Dikin[276], který tento algoritmus navrhl již v roce 1967 a jeho konvergenci k optimálnímu řešení dokázal o sedm let později, viz [134–136, 492]. Jenže po celou tu dobu zůstala jeho práce nepovšimnuta a to i navzdory tomu, že v roce 1967 vyšel její anglický překlad. Dnes už se můžeme jen dohadovat, zda v tom byl nějaký úmysl podobně jako v případě Karmarkarovy metody a Bellových laboratoří nebo zda to jen bylo v duchu Fiaccova a McCormickova tehdejšího postoje. Pokud Dikin skutečně uměl tuto metodu efektivně využít, mohlo to pro sovětský průmysl (nebo nějakou jeho část) znamenat jistou výhodu a úsporu nákladů, o kterou se jistě nechtěli příliš dělit. Pokud ale ne a chyběl tomu ten poslední krok, který udělal až Karmarkar, pak je tehdejší Dikinův příspěvek jen jedním z mnoha a nelze mu přičítat příliš velké zásluhy.

Každopádně na počátku těchto nových metod stál jeden velmi patologický příklad ukazující, že simplexový algoritmus má jeden drobný nedostatek, který se však v praktických výpočtech příliš neprojevil. Na jeho konci máme nový a výrazně efektivnější algoritmus, který snad neočekávali ani Klee a Minty, ale kterým jsme se po 50 letech rozvoje lineárního programování symbolicky vrátili zpět ke kořenům (aneb od Kantoroviče ke Kantorovičovi s „několika“ odbočkami). K dalšímu čtení doporučujeme např. [481], který zahrnuje i další metody pro řešení úloh lineárního programování.

[274] Ragnar Anton Kittil Frisch (3. března 1895 – 31. ledna 1973) byl norským ekonomem, který společně s Janem Tinbergenem získal v roce 1969 *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela*. Je zakladatelem ekonometrie a také autorem termínů makroekonomie/mikroekonomie. Věnoval se především matematizaci ekonomie. Toto jeho úsilí o exaktní charakter ekonomické teorie obsáhlo i širší příbuzné oblasti: operační výzkum včetně matematického programování, kybernetiku, systémové vědy a informatiku, což mu umožnilo aplikovat aparát matematického programování pro makroekonomické účely, viz např. [191].

[275] Angl. *Newton barrier method*

[276] Ilya I. Dikin (1936 – 28. února 2008) byl sovětský matematik. V letech 1963–1966 byl Kantorovičovým doktorským studentem a v roce 1973 získal doktorát. Pracoval na Institutu energetických systémů jako výzkumný pracovník a stal se jedním z předních odborníků v oblasti optimalizace.

2.14 Lineární lomené programování

Pomocí úlohy lineárního programování umíme snadno maximalizovat zisk, tj. rozdíl mezi příjmem a náklady. Např. je-li příjem 1100 Kč a náklady 1000 Kč, pak náš zisk činí 100 Kč, což můžeme chápat jako měřítko *absolutní efektivity*[277] našeho snažení. Ovšem v některých situacích může být mnohem výhodnější sledovat a maximalizovat *relativní efektivitu*, tj. podíl zisku a nákladů, což v uvedeném příkladě bylo 10 %. Reálné situace samozřejmě nejsou takto černobílé, neboť založit podnikání na aktivitě s efektivitou 30 % zní sice hezky, ale taky by se mohlo stát, že jsme tohoto čísla dosáhli při příjmu 100 Kč a nákladech 30 Kč.

[277] tak tohle asi není to nejlepší slovní spojení...

Toto byla ukázka jednoduché úlohy *lineárního lomeného programování*, ve které chceme minimalizovat podíl dvou afinních funkcí na (části daného) polyedru, tj.

$$f(x) := \frac{\langle c,x\rangle+\alpha}{\langle d,x\rangle+\beta} \to \min \quad \& \quad x\in X := \{x\in\mathbb{R}^n \mid Ax\le b \,\&\, \langle d,x\rangle+\beta>0\} \tag{LLP}$$

kde $c,d\in\mathbb{R}^n$, $b\in\mathbb{R}^m$, $A\in\mathbb{R}^{m\times n}$ a $\alpha,\beta\in\mathbb{R}$. Obecně nelze úlohu (LLP) uvažovat na množině $\{x\in\mathbb{R}^n \mid Ax\le b\}$, neboť může obsahovat nulové body jmenovatele. Současně ale nelze vyloučit pouze body splňující $\langle d,x\rangle+\beta=0$, protože bychom v takovém případě rozbili konvexnost přípustné množiny. Proto je v množině X zvolen dodatečný požadavek na kladnost jmenovatele, díky němuž je účelová funkce definována na celém X a současně je tato množina konvexní. Samozřejmě by bylo možné uvažovat v této podmínce i opačnou nerovnost, čímž bychom dostali druhou část polyedru.

Řešení úlohy (LLP) je ekvivalentní s úlohou lineárního programování ve tvaru

$$\langle c,y\rangle+\alpha t\to\min \quad \& \quad Ay\le bt \quad \& \quad \langle d,y\rangle+\beta t=1 \quad \& \quad t\ge 0 \tag{2.14.1}$$

v proměnných y a t. Vskutku, nejdříve si všimněme, že je-li $x \in X$, pak pro dvojici

$$y = \frac{x}{\langle d,x\rangle + \beta} \quad \text{a} \quad t = \frac{1}{\langle d,x\rangle + \beta} \tag{2.14.2}$$

platí, že $t > 0$ a

$$Ay - bt = \underbrace{\frac{1}{\langle d,x\rangle + \beta}}_{>0}\big(\underbrace{Ax - b}_{\le 0}\big) \le 0, \qquad \langle d,y\rangle + \beta t = \frac{\langle d,x\rangle}{\langle d,x\rangle + \beta} + \frac{\beta}{\langle d,x\rangle + \beta} = 1,$$

tj. tato dvojice y, t je přípustný bod úlohy (2.14.1). Jelikož současně máme

$$\langle c,y\rangle + \alpha t = \frac{\langle c,x\rangle}{\langle d,x\rangle + \beta} + \frac{\alpha}{\langle d,x\rangle + \beta} = f(x),$$

tj. hodnota účelové funkce úlohy (2.14.1) pro takovou dvojici y, t je stejná jako hodnota účelové funkce úlohy (LLP) v bodě x, takže nutně (optimální) hodnota úlohy (LLP) je rovna nebo vetší než (optimální) hodnota úlohy (2.14.1). Naopak, je-li dvojice y, t s $t > 0$ přípustná pro (2.14.1), pak $x = y/t$ je přípustný bod úlohy (LLP) s toutéž hodnotou účelové funkce, protože

$$Ax = \tfrac{1}{t}\underbrace{Ay}_{\le bt} \le b \quad \text{a} \quad \langle d,x\rangle + \beta = \tfrac{1}{t}\langle d,y\rangle + \beta = \frac{1}{t} > 0$$

$$f(x) = f(y/t) = \frac{\frac{1}{t}\langle c,y\rangle + \alpha}{\frac{1}{t}\langle d,y\rangle + \beta} = \frac{\langle c,y\rangle + \alpha t}{\underbrace{\langle d,y\rangle + \beta t}_{1}} = \langle c,y\rangle + t\alpha.$$

Je-li ale dvojice y, t s $t = 0$ přípustná pro (2.14.1) a x_0 nějaký přípustný bod úlohy (LLP), pak $x = x_0 + \rho y$ je také přípustný bod úlohy (LLP) pro libovolné $\rho \ge 0$, protože

$$Ax = \underbrace{Ax_0}_{=b} + \rho\,\underbrace{Ay}_{\le 0} \le b \quad \text{a} \quad \langle d,x\rangle + \beta = \big(\underbrace{\langle d,x_0\rangle + \beta}_{>0}\big) + \rho\,\underbrace{\langle d,y\rangle}_{=1} > 0,$$

a navíc pro účelovou funkci $f(x)$ platí

$$\begin{aligned}
\lim_{\rho\to\infty} f(x) &= \lim_{\rho\to\infty} f(x_0 + \rho y) = \lim_{\rho\to\infty} \frac{\langle c, x_0 + \rho y\rangle + \alpha}{\langle d, x_0 + \rho y\rangle + \beta} = \\
&= \lim_{\rho\to\infty} \frac{\langle c,x_0\rangle + \alpha + \rho\langle c,y\rangle}{\langle d,x_0\rangle + \beta + \rho\underbrace{\langle d,y\rangle}_{=1}} = \\
&= \lim_{\rho\to\infty} \left(\frac{\langle c,x_0\rangle + \alpha}{\underbrace{\langle d,x_0\rangle + \beta}_{>0} + \rho} + \frac{\langle c,y\rangle}{\frac{1}{\rho}\big(\underbrace{\langle d,x_0\rangle + \beta}_{>0}\big) + 1} \right) = \\
&= \langle c,y\rangle = \langle c,y\rangle + \alpha t,
\end{aligned}$$

tj. můžeme najít přípustné body úlohy (LLP) s hodnotami účelové funkce libovolně blízkými hodnotě účelové funkce úlohy (2.14.1) v bodě y, t. Odtud vyplývá, že hodnota úlohy (LLP) je menší nebo rovna hodnotě úlohy (2.14.1)[278], což dohromady s první části potvrzuje platnost uvedené ekvivalence mezi úlohami (LLP) a (2.14.1). Transformace (2.14.2) se také někdy nazývá *Charnesova–Cooperova* podle autorů článku z roku 1962, ve kterém byla použita pro důkaz jisté analogie právě uvedené ekvivalence. Toto tvrzení lze rozšířit také na účelovou funkci

$$\max_{i=1,\dots,r} \frac{\langle c_i,x\rangle + \alpha_i}{\langle d_i,x\rangle + \beta_i}.$$

[278] Nastává-li minimum úlohy (2.14.1) v bodě y, t s $t > 0$, pak v X umím najít bod se stejnou funkční hodnotou. Je-li minimum pro $t = 0$, pak v X je posloupnost bodů, pro kterou účelová funkce konverguje k téže hodnotě.

Příklad 2.14.1

(Von Neumannův problém růstu) Uvažme ekonomiku s n-ticí segmentů a úrovní aktivity $x_i > 0$ v i-tém segmentu a v současném časovém období, zatímco v příštím časo-

vém období bude úroveň aktivity rovna $x_i^+ > 0$ (zajímá nás pouze jedno další období). Je spotřebováváno m výrobků, které jsou v závislosti na úrovni aktivity vyráběny. Úroveň aktivity x znamená $Bx \in \mathbb{R}^m$ spotřebovaných výrobků a výrobu Ax výrobků, přičemž v dalším období nelze spotřebovat více než kolik se v současném období vyrobí, tj. $Bx^+ \leq Ax$. *Míra růstu i-tého segmentu* je pak dána podílem x_i^+ / x_i. Von Neumannův problém růstu je úloha pro nalezení aktivity x, která maximalizuje minimum míry růstu mezi všemi segmenty, tj. jedná se o úlohu lineárního lomeného programování ve tvaru

$$\min_{i=1,\ldots,n} \left\{ \frac{x_i^+}{x_i} \right\} \to \max \quad \& \quad Bx^+ \leq Ax \quad \& \quad x^+ \geq 0$$

a doplněná o požadavek $\{[x, x^+] \mid x > 0\}$. Jelikož se jedná o homogenní problém v x a x^+, je možné nahradit implicitní omezení $x > 0$ explicitní podmínkou $x \geq 1$.

Kapitola 3

Celočíselné programování

3.1 *Když reálná čísla nejsou příliš reálná* 256
3.2 *Celočíselnost až na prvním místě* 258
3.3 *Celočíselnost i bez celočíselnosti* 274
3.4 *Metoda větvení a mezí* 280
3.5 *Metoda řezů* 287
3.6 *Chcete být milionářem?* 305

WT: YOU HAD HAD MATHEMATICS, I TAKE IT?

GOMORY: OH I HAD SOME MATHEMATICS, YES. I GUESS I HAD BOTH PLANE GEOMETRY AND ALGEBRA. BUT I HATED IT. I DIDN'T GET THE POINT. I'M AFRAID THAT THROUGHOUT MY LIFE I'VE ALWAYS BEEN SORT OF APPLIED IN THE SENSE THAT I WANTED TO SEE A CONNECTION BETWEEN WHAT WE WERE LEARNING AND SOMETHING IN THE OUTSIDE WORLD.

WT: THAT THERE'S A MOTIVATION BEHIND WHAT YOU HAPPEN TO BE DOING.

GOMORY: YEAH. IT WAS HARD FOR ME TO THINK ABOUT IT IF I COULDN'T SEE THE POINT. EVEN IN MATH, WHICH I DIDN'T DO PARTICULARLY WELL IN, ONCE THEY GOT TO WORD PROBLEMS, YOU KNOW, JOHN CAN MOW THE LAWN IN SIX HOURS BY HIMSELF, AND BILL CAN MOW THE LAWN IN SEVEN HOURS OR EIGHT HOURS. HOW LONG DOES IT TAKE THEM TO MOW IT IF THEY CAN MOW TOGETHER? I COULDN'T MISS THOSE. SO THAT'S JUST MY THING...

GOMORY: SO I BECAME INTERESTED IN MATHEMATICS. AND BECAUSE PHYSICS DIDN'T SEEM TO BE EXPLAINING THE WORLD TO ME ANYMORE, I REACHED THE ERRONEOUS CONCLUSION THAT IT WAS MATHEMATICS THAT EXPLAINED THE WORLD, WHICH I DON'T NOW THINK IS CORRECT BUT I DID AT THE TIME. BECAUSE I FOUND THAT THE SAME DIFFERENTIAL EQUATIONS THAT DESCRIBED SOMETHING IN PHYSICS ALSO WERE BEING USED TO DESCRIBE SAY THE BEATING OF THE HUMAN HEART. SO I GOT VERY INTERESTED IN ORDINARY DIFFERENTIAL EQUATIONS BECAUSE OF THAT. I WAS JUST FASCINATED BY THIS PAPER BY VAN DER POL ENTITLED "THE BEATING OF THE HUMAN HEART DESCRIBED AS A NONLINEAR OSCILLATION". SO I THOUGHT NOW I'VE GOT THE REAL KEY TO UNDERSTANDING THINGS — ITS MATHEMATICS AFTER ALL![1]

RALPH GOMORY[2] o počátcích svého zájmu (cca 1957) o celočíselné programování v zajímavém rozhovoru [546] s Williamem Thomasem (WT). Viz také [545].

[1] **WT**: *Měl jste matematiku, chápu to dobře?*
Gomory: *Oh, ano, měl jsem trochu matematiky. Myslím, že jsem měl jak rovinnou geometrii, tak algebru. Ale nenáviděl jsem to. Nepochopil jsem pointu. Obávám se, že po celý svůj život jsem byl něco jako „aplikovač" v tom smyslu, že jsem chtěl vidět spojení mezi tím, co jsme se učili, a něčím ve vnějším světě.*
WT: *Že za tím, co děláte, je nějaká motivace.*
Gomory: *To jo. Bylo pro mě těžké o něčem přemýšlet, když jsem neviděl význam. Dokonce i v matematice, která mi moc nešla, jakmile jsme se dostali ke slovním úlohám, víte, John dokáže sám posekat trávník za šest hodin a Bill dokáže posekat trávník za sedm nebo osm hodin. Jak dlouho jim potrvá, než to posekají, když mohou sekat společně? To jsem nemohl přehlédnout. To je pro mne to pravé...*
Gomory: *Začal jsem se tedy zajímat o matematiku. A protože se mi zdálo, že fyzika už svět neobjasňuje, dospěl jsem k mylnému závěru, že to je matematika, která vysvětluje svět, což si nyní nemyslím, že je správné, ale v té době jsem si to myslel. Protože jsem zjistil, že stejné diferenciální rovnice, které popisovaly něco ve fyzice, se také používaly třeba k popisu tlukotu lidského srdce. Takže mě kvůli tomu velmi zaujaly obyčejné diferenciální rovnice. Byl jsem zrovna fascinován van der Polovým článkem „The Beating of the Human Heart Described as a Nonlinear Oscillation". Takže jsem si myslel, že teď mám skutečný klíč k pochopení věcí – přeci jen je to matematika!*

3.1 Když reálná čísla nejsou příliš reálná

Všechny úlohy prezentované v předchozí kapitole měly několik společných rysů. Kromě linearity účelové funkce a jednotlivých omezení zde byl ještě jeden velmi důležitý aspekt – byly „spojité“, tj. každá proměnná mohla nabývat libovolné hodnoty z jistého intervalu v $\mathbb{R}$. Jenže snad ani Pejsek s Kočičkou by při pečení svého slavného dortu nedokázali navážit $\frac{31415}{27182}$ kg mouky. Stejně tak pro každého zloděje méně zdatného než Vincenzo Peruggia[3] by krádež $\frac{21}{73}$ vzácné knihy či obrazu byla téměř bezcenná. Podobně nereálné to je i s penězi. Existuje asi jen velmi malá skupina lidí (nepochybně ale zahrnující většinu bývalých ministrů financí ČR), kteří by dokázali při sestavování dokonalého státního rozpočtu přidělit do některé z jeho kapitol přesně $\sqrt{2\pi + e^2}$ milionů Kč. V některých situacích se těmto potížím lze vyhnout vhodně zvolenými jednotkami, díky nimž bude případná odchylka od přesného řešení zanedbatelná[4], např. při rozpočtování projektu s položkami v desítkách milionů korun lze pracovat s jednotkami v korunách a zaokrouhlovat, avšak při práci v jednotkách milionů se při zaokrouhlování již můžeme dopustit poměrně značné nepřesnosti – nehledě na možné porušení podmínek přípustnosti. Dostali jsme se tak na rozcestí, na kterém jsme doposud upřednostňovali snadnost výpočtu před přesností, protože odměna za tuto naši „drobnou“ nepřesnost a případnou nereálnost/nerealizovatelnost získaného řešení byla vskutku úžasná — simplexová metoda[5]. Ovšem práce v malých jednotkách může být (a jistě i bude) početně mnohem náročnější a v některých případech dokonce ani toto zjednodušení společně se zaokrouhlováním nemusí vést k smysluplnému výsledku, např. v problému batohu (viz Příklad 2.3.4, kde však nepůjde o kávu ale např. právě o knihy jako u Peruggii) nebo v dopravním problému (viz Podkapitolu 2.11)[6]. Proto, chceme-li zvládnout i takovéto problémy, musíme přehodnotit naše preference a zaměřit se na skutečně precizní matematickou formulaci těchto úloh. To nás přivádí k úlohám *celočíselného programování*

$$(\text{LP}) \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad \dots \quad \& \quad x_n \in \mathbb{Z} \qquad (\mathbb{Z}\text{LP})$$

ve kterých požadujeme, aby všechny proměnné nabývaly pouze celočíselných hodnot. Přípustná množina takové úlohy pak je průsečíkem množiny X a celočíselné sítě, viz Obrázek 3.1. Pokud dokonce omezíme všechny proměnné pouze na hodnoty $x_j \in \{0,1\}$ pro $j = 1,\dots,n$, pak se hovoří o *binární úloze*. Na druhou stranu pokud celočíselnost bude vyžadována pouze pro některé proměnné, tj. $x_j \in \mathbb{Z}$ pro $j \in A \subsetneq \{1,\dots,n\}$, dostaneme tzv. *úlohu smíšeného celočíselného programování*.

Příklad 3.1.1

Uvažme úlohu

$$-5x_1 - 8x_2 \to \min$$
$$x_1 + x_2 \le 6 \quad \& \quad 5x_1 + 9x_2 \le 45$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

První (naivní) krok řešení by mohl spočívat v odstranění požadavku celočíselnosti a řešení daného problému jako úlohy lineárního programování. Jenže takto nalezené řešení $x^* := [9/4, 15/4]$ nemá celočíselné souřadnice a případným zaokrouhlováním můžeme opustit přípustnou množinu. Pokud totiž souřadnice x^* zaokrouhlíme podle klasických pravidel, získáme bod $\widetilde{x} := [2,4]$, který nesplňuje druhé omezení, tj. $5\widetilde{x}_1 + 9\widetilde{x}_2 \not\le 45$. Jestliže souřadnice bodu x^* zaokrouhlíme tak, aby splňoval obě daná omezení (nezápornost je víceméně zřejmá), získáme $\widehat{x} := [2,3]$ s hodnotou $f(\widehat{x}) = -34$. Avšak je to i řešením dané celočíselné úlohy? Jelikož hodnota úlohy bez požadavku celočíselnosti je $f(x^*) = -41\frac{1}{4}$, může být optimální hodnota celočíselné úlohy nejméně -41 a nejvíce -34 (v právě nalezeném bodě $\widehat{x}$). Existuje v přípustné celočíselné množině

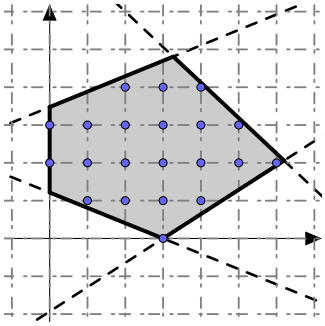
[2] Ralph Edward Gomory (⋆7. května 1929) je americký matematik a vrcholový manažer. Zcela zásadním způsobem ovlivnil rozvoj celočíselného programování, ke kterému se dostal v průběhu své služby u námořnictva USA v letech 1954–1957. Od roku 1959 pracoval v IBM a od roku 1970 zde vedl výzkumnou divizi, díky čemuž došlo ke vzniku nových oblastí aplikované matematiky, rozvoji počítačového průmyslu a i výrobků samotného IBM. Po své kariéře v korporátním světě se Gomory stal prezidentem *Nadace Alfreda P. Sloana*, kde dohlížel na programy věnované rozšíření porozumění veřejnosti ve třech klíčových oblastech: ekonomický význam vědy a výzkumu; dopady globalizace na Spojené státy; a role technologií ve vzdělávání. Gomory (doposud) získal osm čestných doktorátů a několik významných ocenění jako např. *Lanchester Prize* již v roce 1963 udělovanou *Operations Research Society* (Společností pro operační výzkum), *John von Neumann Theory Prize* v roce 1984 či *National Medal of Science* (Národní vyznamenání za vědu) udělované prezidentem USA v roce 1988 za „jeho vědecké příspěvky v matematické oblasti diskrétní optimalizace a jejich dalekosáhlý vliv na zpracování informací, za přivedení jednoho z nejvýznamnějších výzkumných pracovišť v oboru do vedoucí pozice a za jeho zásluhy o veřejné a soukromé vědecké podnikání“.

[3] Vincenzo Peruggia (8. října 1881 – 8. října 1925) byl italský zloděj známý pro krádež obrazu Mony Lisy dne 21. srpna 1911.

[4] Místo nalezeného optimálního řešení, tak využijeme spíše nějaké *suboptimální* řešení, jehož hodnota by se však neměla příliš lišit od toho optimálního (za přijatelnou bychom mohli považovat např. 5% odchylku).

[5] Nebo případně i jiné modernější nástroje pro řešení úloh lineárního programování, viz závěr této kapitoly.

[6] Abychom ale byli zcela korektní (a zmínili jsme se o tom již dříve), u dopravního problému k těmto potížím nedochází, neboť jeho speciální struktura nám zajistí existenci celočíselného řešení i bez explicitního požadavku celočíselnosti na jeho proměnné.

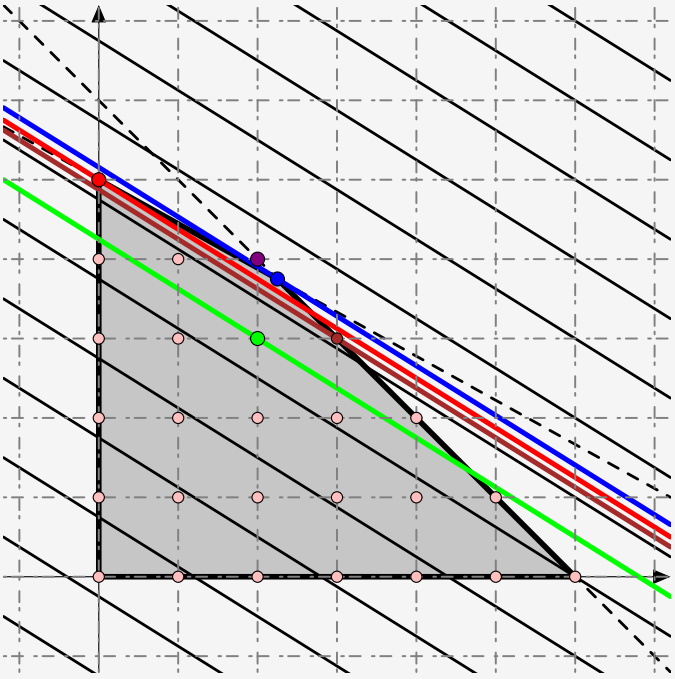


Obrázek 3.1: Ukázka přípustné množiny úlohy celočíselného programování.

bod s funkční hodnotou menší než -34? S trochu nepřesným zaokrouhlováním můžeme z bodu x^* získat $\bar{x} := [3,3]$ s výrazně menší funkční hodnotou $f(\bar{x}) = -39$. Avšak s využitím Obrázku 3.2 vidíme, že v přípustné množině leží dokonce i bod $x^{**} := [0,5]$ s funkční hodnotou $f(x^{**}) = -40$, což je správné celočíselné řešení dané úlohy, viz také Tabulku 3.1.

	řešení v $\mathbb{R}$ (x^*)	zaokrouhlení mimo X ($\widehat{x}$)	zaokrouhlení v X #1 ($\widehat{x}$)	zaokrouhlení v X #2 ($\bar{x}$)	řešení v $\mathbb{Z}$ (x^{**})
x_1	$\frac{9}{4}$	2	2	3	0
x_2	$\frac{15}{4}$	4	3	3	5
$f(x_1, x_2)$	$-41\frac{1}{4}$	-42 & $\notin X$	-34	-39	-40

Tabulka 3.1: Souhrn řešení Příkladu 3.1.1.



Obrázek 3.2: Řešení úlohy z Příkladu 3.1.1. Modrý bod je řešení bez požadavku celočíselnosti x^*, nepřípustný fialový bod $\widehat{x}$ je získán zaokrouhlením souřadnic x^* (dává menší funkční hodnotu než x^*), zelený bod $\widehat{x}$ a hnědý bod $\bar{x}$ jsou získány obdobně (dávají větší funkční hodnotu než x^*) a červený bod x^{**} je řešení dané úlohy.

Tento jednoduchý příklad nám ukazuje, že propojení mezi úlohami lineárního a celočíselného programování nemusí být příliš těsné a že zaokrouhlování obvykle asi nepovede ke správnému výsledku – spíše jen k nějakému jeho „kvalifikovanému“ odhadu, který však může porušovat podmínky přípustnosti. Navíc v případě alespoň jednoho rovnostního omezení v původní úloze takto nikdy přípustný bod nezískáme. To plyne z toho, že při zaokrouhlování měníme hodnoty pouze neceločíselných souřadnic, zatímco s těmi celočíselnými (obzvláště nulovými) se neděje nic.

Jenže my se nechceme tak snadno vzdát simplexového algoritmu. Proto využijeme alespoň některé společné rysy těchto úloh, a s pomocí vhodné dekompozice/restrikce přípustné množiny si vybudujeme dvě metody pro nalezení „přesného“[7] řešení. Neopomeneme však ani příslovečné „výjimky potvrzující pravidlo“, ve kterých nás již samotná přidružená úloha lineárního programování přivede bez jakýchkoli dalších podmínek k hledanému řešení, což jsme vlastně již stručně naznačili v Podkapitole 2.11 věnované dopravnímu problému. Abychom se

[7] Uvozovky jsou zde pro případ, kdy jsou v metody implementovány do prostředí s numerickými výpočty.

ale vyvarovali případným pochybnostem o možném využití úloh celočíselného programování, začneme nejdříve s několika reprezentativními ukázkami jeho aplikací v praxi.

3.2 Celočíselnost až na prvním místě[8]

Na počátku padesátých let 20. století společnost RAND zaměstnávala snad tu *nejpozoruhodnější skupinu matematiků, která se kdy věnovala optimalizačním problémům*: Arrow, Bellman, Dantzig, Flood, Ford, Fulkerson, Gale, Johnson, Nash, Orchard-Hays, Robinsonová, Shapley, Simon, Wagner a mnoho dalších. Jenže takové skupiny potřebují i své výzvy a jednou z těch nejdůležitějších byl *problém obchodního cestujícího* (dále obvykle jen „POC“). Snad každému z nás tento název evokuje osobu projíždějící nějakou oblastí a navštěvující postupně několik míst[9]. Tato nekonkrétní vize jistého matematického problému předcházela dnešnímu pojetí, ve kterém je onou cestující osobou obchodník. Je těžké říci, kde se tento problém zcela objevil poprvé, neboť jeho krása mimo jiné tkví v tom, že jej kromě zjevného propojení se všelijakými doručovacími službami nalezneme i v mnoha jiných různorodých oblastech – od mapování genomu, přes chování mravenců, šimpanzů či holubů při hledání potravy až po sestavování počítačové sítě, viz [2,424]. Kdykoli totiž chcete poskládat nějaké „předměty/činnosti“ do řady za sebou, je možné, že správné řešení skýtá právě jistá variace na POC. Ale i tak se nyní pokusíme zmapovat alespoň část historie tohoto pozoruhodného problému.

Pregola (rusky Pregolja), je nejdelší a nejvodnější řekou v ruské Kaliningradské oblasti (ruské území mezi Litvou a Polskem). Byla by snad jen jednou z položek v dlouhém seznamu řek v Matičce Rusi, kdyby se nevlila i do matematických vod a nevytvořila zde moře, kterému dnes říkáme teorie grafů. Tato řeka totiž protéká Kaliningradem, který byl v 18. století hlavní metropolí Východního Pruska (tehdy se město nazývalo Královec či Königsberg) a vytváří zde dva ostrovy (Kneiphof a Lomse). Tyto ostrovy byly vzájemně a s okolním městem spojeny sedmi mosty (dodnes se zachovaly pouze dva) a traduje se, že místní obyvatelé si během tradičních nedělních procházek městem lámali hlavu s otázkou, zda lze projít všemi částmi města tak, aby při tom využili každý most právě jednou. Tato úloha se dostala i ke slovutnému Eulerovi, který prezentoval 26. srpna 1735 Petrohradské akademii věd své výsledky ukazující nemožnost nalezení řešení tohoto slavného *Problému sedmi mostů města Královce*.

> *V Pruském městě Královec se nachází ostrov Der Kneiphof, kolem kterého proudí dvě větve řeky Pregoly. Obě větve jsou přemostěny pomocí sedmi mostů. V souvislosti s těmito mosty mi byla položena otázku: je možné naplánovat cestu, která projde každým z mostů jednou, ale ne více než jednou? Bylo mi řečeno, že někteří to odmítají a jiní pochybují, že by to bylo možné, ale nikdo to netvrdil s jistotou. To mě přivedlo k tomuto obecnějšímu problému: bez ohledu na rozmístění a rozdělení ramen řeky a bez ohledu na počet mostů, můžete zjistit, zda dokážete projít každým mostem jednou a pouze jednou?*

Tento výsledek byl ale publikován až o pět let později v pojednání [158], kde jsou uvažovány i obecnější problémy téhož typu, pro které se Eulerovi podařilo zformulovat některé nutné a postačující podmínky pro existenci tzv. *eulerovského tahu*[10], viz [275, 440] a [64, str. 3–8]. Počátky Eulerova zájmu o tuto úlohu se zatím nepodařilo vystopovat[11], ale zřejmě v tom sehrál roli pruský matematik a starosta sousedního Gdaňsku (německy Danzig /sic!/) Karl Leonhard Gottlieb Ehler[12,13], který mu 9. března 1736 napsal:

> *Prokázal byste mně a našemu příteli Kühnovi nejcennější službu, čímž bychom se stali Vašimi dlužníky, nejvzdělanější pane, kdybyste nám poslal Vám velmi dobře známé řešení problému sedmi mostů v Královci i s jeho dů-*

[8] První (historická) část této podkapitoly vychází především z [12, 82,101,271,301,449,450,542], kde zvídavý čtenář nalezne řadu dalších (nejen) historických zajímavostí.

[9] Třeba jako legendární ithacký král Odysseus, který podle Homérova eposu navštívil (cca) 15 míst ve Středomoří, aby se po dlouhých 20 letech vrátil zpět domů ke své milované Pénelopě.

[10] Tj. nakreslení cesty jedním tahem (neboli *jednotažky*).

[11] Zdá se, že Euler dokonce ani nikdy Královec nenavštívil.

[12] Za matematika je považován nejspíše kvůli své práci astronoma v Berlíně, se kterou je spojeno velmi důležité setkání s Leibnizem. Ehler mu totiž poslal jak svá vlastní astronomická pozorování, tak i knihu od Isaaca Papina (neplést s jeho bratrancem, Denisem Papinem, který je vynálezcem známého Papinova hrnce). Isaac Papin byl původně kalvínským teologem, avšak v roce 1690 konvertoval ke katolicismu a ve své knize pojednává o skepticistickém argumentu proti vědě, kterým se snaží podpořit svůj postoj ke vztahu vědy a víry. Leibnize pochopitelně tento Papinův postoj nijak zvlášť nepotěšil, a proto „uvítal“, že již na něj kvůli jeho smrti nemusí reagovat. Nicméně Leibniz ve svém dopise Ehlerovi sepsaném pár měsíců před svojí smrtí využil této příležitosti a vysvětlil své názory na to, jak obecně reagovat na skeptickou argumentaci, přičemž zdůraznil roli *dobré logiky* (recta logica) jako *soudce kontroverzí*, viz [129, str. 451–453].

[13] Pravděpodobně zde ale působil spíše jenom jako prostředník mezi Eulerem a jiným gdaňským matematikem Heinrich Kühnem, o němž ale žádné podrobnější informace nelze dohledat.

> *kazem. Ukázalo by se to jako význačný příklad kalkulu pozice (Calculi Situs) hodného Vaší geniality. Přikládám náčrt zmíněných mostů…*[14]

Odpovědi se mu dostalo celkem záhy, již 3. dubna 1736:

> *… Tak vidíte, nejšlechetnější pane, jak malý vztah k matematice má tento typ řešení, a já nechápu, proč očekáváte, že s ním přijde spíše matematik než kdokoli jiný, neboť řešení je založeno pouze na ratiu a jeho objevení nevyžaduje žádný matematický princip. Z tohoto důvodu nevím, proč i otázky, které mají tak malý vztah k matematice, jsou matematiky vyřešeny rychleji než ostatními. Mezitím, nejšlechetnější pane, jste přiřadili tuto otázku ke geometrii polohy, ale já nevím, co je náplní této nové disciplíny a jaké typy problémů očekávali Leibniz a Wolff, že bude možné vyjádřit tímto způsobem.*[15]

Eulerova narážka na Leibnize a Wolffa směřuje na *geometrii polohy* (*analysis situs* nebo *geometria situs*), o níž 8. září 1679 psal Leibniz Huygensovi.

> *Nejsem spokojen s algebrou, protože nepřináší ani nejkratší důkazy, ani nejkrásnější geometrické konstrukce. Z tohoto důvodu se domnívám, že potřebujeme ještě jiný druh analýzy – geometrickou nebo lineární, která se zabývá přímo pozicí tak, jako se algebra zabývá veličinami…*[16]

Z Eulerovi odpovědi možná můžeme cítit mírný despekt k důležitosti této úlohy, ale z Eulerova dopisu datovaného 13. březnem 1736 určeného Marinonimu[17] se dozvíme více.

> *Byl mi předložen problém o ostrově ve městě Královci… Tato otázka je sice banální, ale zdálo si mi hodné pozornosti, že geometrie, algebra ani umění počtů nestačí k jejímu vyřešení. Z tohoto důvodu mě napadlo přemýšlet, jestli to patří do geometrie polohy [geometriam Situs], po které Leibniz kdysi tolik toužil. A tak jsem po čase přemýšlení získal jednoduché a už zcela potvrzené pravidlo, s jehož pomocí se člověk může u jakéhokoli příkladu s libovolně uspořádanými mosty a v jakémkoli počtu okamžitě rozhodnout, zda je okružní trasa možná či nikoli…*[18]

Ať už byl Eulerův postoj k této úloze jakýkoli, jeho článek je dnes řazen mezi ty nejvýznamnější. Je také nutné zmínit, že Euler nesestrojil žádný algoritmus pro nalezení případné trasy – nebylo to potřeba, stačilo pouze zjištění její existence či neexistence. O více než 225 let později (v roce 1960, viz [339]) popsal čínský matematik Guan[19],

který v mládí pracoval jako listonoš, jedno z možných a zajímavých rozšíření Eulerova problému[20] .

> *Představme si listonoše, který potřebuje doručit poštu v jisté oblasti. Tento pošťák je ovšem trochu líný, takže by chtěl najít nejkratší cestu skrze tuto oblast, která splňuje následující dvě podmínky*
>
> *(i) trasu musí začínat a končit v tomtéž místě (zřejmě u pošty, tj. trasa musí být uzavřená);*
>
> *(ii) musí projít každou ulici alespoň jednou.*

Tentokrát však na rozdíl od Eulerovy úlohy se již hledá příslušná trasa, a to dokonce i v případech kdy zde neexistuje eulerovský tah. Jakkoli podobná s abstraktním popisem POC se může tato úloha zdát, není to totéž – v obou problémech sice jde o nalezení jisté trasy (jedná se

[14] *You would render to me and our friend Kühn a most valuable service, putting us greatly in your debt, most learned Sir, if you would send us the solution, which you know well, to the problem of the seven Königsberg bridges, together with a proof. It would prove to be an outstanding example of the calculus of position [Calculi Situs], worthy of your great genius. I have added a sketch of the said bridges…* Viz [440, str. 134].

[15] *… Thus you see, most noble sir, how this type of solution bears little relationship to mathematics and I do not understand why you expect a mathematician to produce it rather than anyone else, for the solution is based on reason alone, and its discovery does not depend on any mathematical principle. Because of this, I do not know why even questions which bear so little relationship to mathematics are solved more quickly by mathematicians than by others. In the meantime most noble sir, you have assigned this question to the geometry of position but I am ignorant as to what this new discipline involves, and as to which types of problem Leibniz and Wolff expected to see expressed this way.* Viz [440, str. 136].

[16] *I am not content with algebra, in that it yields neither the shortest proofs nor the most beautiful constructions of geometry. Consequently, in view of this, I consider that we need yet another kind of analysis, geometric or linear, which deals directly with position, as algebra deals with magnitudes…* Viz [440, str. 134].

[17] Johann Jakob Marinoni či Giovanni Jacopo de Marinoni (1676 – 10. ledna 1755) byl italsko-rakouský matematik a astronom sloužící ve Vídni na dvoře císaře Leopolda I.

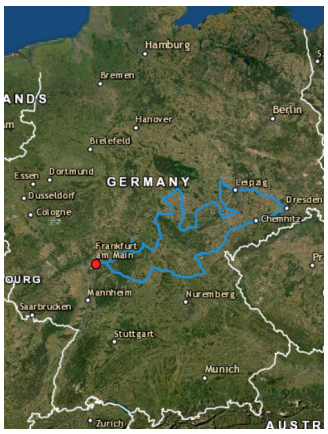
[18] *A problem was posed to me about an island in the city of Königsberg… This question is so banal, but seemed to me worthy of attention in that geometry, nor algebra, nor even the art of counting was sufficient to solve it. In view of this, it occurred to me to wonder whether it belonged to the geometry of position [geometriam Situs], which Leibniz had once so much longed for. And so, after some deliberation, I obtained a simple, yet completely established, rule with whose help one can immediately decide for all examples of this kind, with any number of bridges in any arrangement, whether such a round trip is possible, or not…* Viz [440, str. 136].

vlastně o tzv. trasovací problémy[21]), avšak zde je to podmíněno nutností projít všechny ulice (jako listonoš či při blokovém čištění nebo odklízení sněhu), zatímco u POC je nutné projít všemi stanovenými body (všelijaké doručovací služby, trasa autobusu MHD nebo vyhlídkové jízdy pro turisty apod.).

V německé příručce pro obchodního cestujícího z roku 1832 *Der Handlungsreisende*[22]*: wie er sein soll und was er zu thun hat, um Aufträge zu erhalten und eines glücklichen Erfolgs in seinen Geschäften gewiß zu sein*[23,24] nalezneme následující komentář.

> *Die Geschäfte führen die Handlungsreisenden bald hier, bald dort hin, und es lassen sich nicht füglich Reisetouren angeben, die für alle vorkommende Fälle passend sind; aber es kann durch eine zweckmäßige Wahl und Eintheilung der Tour, manchmal so viel Zeit gewonnen werden, daß wir es nicht glauben umgehen zu dürfen, auch hierüber einige Vorschriften zu geben. Ein Jeder möge so viel davon benutzen, als er es seinem Zwecke für dienlich hält; so viel glauben wir aber davon versichern zu dürfen, daß es nicht wohl thunlich sein wird, die Touren durch Deutschland in Absicht der Entfernungen und, worauf der Reisende hauptsächlich zu sehen hat, des Hin- und Herreisens, mit mehr Oekonomie einzurichten. Die Hauptsache besteht immer darin: so viele Orte wie möglich mitzunehmen, ohne den nämlichen Ort zweimal berühren zu müssen*[25].

Příručka navrhuje pět cest okružních cest po Německu (jedna z nich zahrnuje i část Švýcarska) a na Obrázcích 3.3–3.4 je jedna z nich čítající 47 měst[26]. Zvýrazněná trasa je vytvořena na základě dnešních mapových podkladů a vypadá celkem obstojně, i když nesplňuje onu zmiňovanou „Hauptsache“ a v okolí Frankfurtu a Fuldy obsahuje „zajížďky“[27]. Nicméně před dvěma sty lety byla „dopravní“ situace jistě výrazně odlišná, a tak nezbývá než věřit, že trasa nabízená příručkou skýtala tehdejším obchodním cestujícím skutečně to nejlepší.



Obrázek 3.3: Trasa spojující 47 měst z příručky Der Handlungsreisende v kontextu celého Německa.

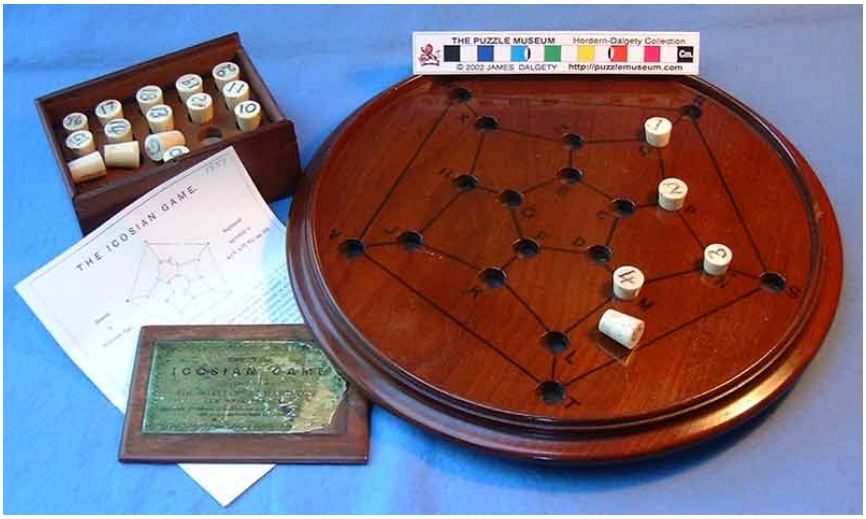
Obrázek 3.4: Trasa spojující 47 měst z příručky Der Handlungsreisende.

Jenže co kdyby obchodník[28] chtěl procestovat jinou oblast? Měl by čekat až někdo napíše vhodného průvodce a usnadní mu plánování? To by mohlo trvat poměrně dlouho… Rychlejší by bylo naplánovat si vlastní trasu. V předchozí tabulce jsme ale viděli, že při tom může přicházet do úvahy skutečně velké možností možností. Proto by to chtělo nějaký „trénink“ v plánování aneb (v parafrázi Komenského slavného výroku) „matematika hrou“. V roce 1856 zveřejnil Thomas Kirkman[29] článek [321] věnovaný mnohostěnům, jimiž se v té době zabýval. Tento článek sice není to nejlepší, čeho se mu v této oblasti podařilo dosáhnout, avšak je to významný především z historického pohledu – *de facto* v něm totiž nabízí první grafickou interpretaci POC, když jej zajímala otázka existence uzavřené lomené čáry procházející

[19] Meigu Guan (též Mei-Ko Kwan nebo Mei-ku Kuan, ⋆1934) je čínský matematik a jeden z předních čínských odborníků na matematické programování. Byl také prezidentem Shandongské pedagogické univerzity (angl. Shandong Normal University).

[20] Shandong je významným místem čínské historie (např. slavný filosof Confucius se narodil ve městě Qufu, které patří právě do provincie Shandong). Během Druhého pětiletého plánu Čínské lidové republiky v letech 1958–1962 pod vedením Mao Ce-tunga byla uplatňována politika s oficiálním názvem „Velký skok vpřed“ (1958–1960), která však vedla až k největšímu lidmi vyvolanému hladomoru v dějinách s desítkami milionů obětí. V tomto období byli čínští vědci pobízení k řešení reálných problémů, aby napomohli svému vůdci k rychlé transformaci země ze zemědělské ekonomiky na průmyslovou a ekonomickou velmoc s moderní komunistickou společností. Proto se řada tehdejších matematiků zaměřila na problémy operačního výzkumu s důrazem na problémy týkající se dopravy či plánování výroby. A jednou z nejdůležitějších oblastí v tomto výzkumu byl právě Shandong, který je rodištěm *Problému čínského listonoše*, viz [239]. Název tohoto problému pochází od Alana Goldmana, který pracoval jako výzkumník v Národní institutu standardů a technologie USA (National Institute of Standards and Technology). Byl pochopitelně inspirován klíčovou postavou této úlohy a národností jeho autora. Samotný Goldman v jednom ze svých soukromých komentářů datovaného 14. prosince 2003 potvrdil, že byl ve výběru jména byl ovlivněn i názvem detektivky *Zá-*

všemi vrcholy daného mnohostěnu právě jednou[30]. A snad možná toto propojení mezi POC a mnohostěny inspirovalo Williama Hamiltona[31], který v roce 1857 přišel s *dvacítkovou hrou*[32] založené na pravidelném dvanáctistěnu[33] mající 20 vrcholů a 30 hran[34]. Cílem bylo nalezení uzavřené trasy procházející všemi vrcholy dvanáctistěnu právě jednou a využívající každou hranu nejvýše jednou[35].



Obrázek 3.5: Rovinná verze „dvacítkové hry". Podle návodu označující písmena různá věhlasná města po celém světě (např. B=Brusel, C=Canton neboli čínské Guangzhou, D=Dillí atd.).

Na Obrázku 3.5 výše je rovinná verze této hry, kde jednotlivé body představují města (vrcholy dvanáctistěnu) a úsečky dostupné spojnice mezi nimi (hrany dvanáctistěnu). Hamilton ukázal, že spektrum možných cest je velmi široké a pestré, neboť bez ohledu na počáteční volbu prvních 5 měst a jejich pořadí, lze vždy hru dokončit. První dochovaný záznam o této hře pochází z Hamiltonova dopisu s datem 7. října 1856, který adresoval svému příteli a irskému matematikovi Johnu Gravesovi. Zde se zmiňuje o novém kalkulu[36] zcela odlišném od kvaternionů. Znovu v této souvislosti a s mnoha podrobnostmi napsal Gravesovi jen o deset dní později a v závěrečné části dokonce píše[37]:

> *I have found that some young persons have been much amused by trying a new mathematical game which the Icosian furnishes, one person sticking five pins in any five consecutive points … and the other player then aiming to insert, which by the theory in this letter can always be done, fifteen other pins, in cyclical succession, so as to cover all the other points, and to end in immediate proximity to the pin wherewith his antagonist had begun.*[38]

O rok později tuto hru Hamilton prezentoval na vědeckém setkání v Dublinu a podle tradiční zprávy [544, druhá část, str. 3] účastníky zaujal:

> *The author stated that this calculus was entirely distinct from that of quaternions, and in it none of the roots concerned were imaginary. He then explained the leading features of the new calculus, and exemplified its use by an amusing game, which he called the Icosian, and which he had been led to invent by it, — a lithograph of which he distributed through the Section, and examples of what the game proposed to be accomplished were lithographed in the margin, the solutions being shown to be exemplifications of the calculus. The figure was the projection on a plane of the regular pentagonal dodecahedron, and at each of the angles were holes for receiving the ivory pins with which the game was played.*[39]

hada čínského pomeranče (The Chinese Orange Mystery) od Ellary Qeen z roku 1934. Goldman to navrhl Jacku Edmondsovi (jednomu z otců kombinatorické optimalizace a držitelem John von Neumann Theory Prize z roku 1985), který tento název rozšířil dále.

[21] Angl. *arc routing problems*. Takto o tom hovořil již Euler, obzvláště v dopise z roku 1757 určeném známému německo-ruskému matematikovi Christianu Goldbachovi, viz [161]. V něm hovoří o *jezdcově procházce*, v níž jde o to, aby jezdec pohybující se po prázdné šachovnici v souladu s šachovými pravidly navštívil každé pole právě jednou. Eulerovo podrobné řešení této úlohy z následujícího roku patří k jeho nejslavnějším dílům a příspěvku k tzv. rekreační matematice. Nicméně zveřejnění se článek dočkal až v roce 1766, viz [162]. Euler svůj výsledek prezentoval Pruské akademii věd v Berlíně, která v roce 1759 vypsala odměnu 4000 franků za nejlepší pojednání o problému jezdcovi procházky. Jenže cena nebyla nikdy udělena. Proč? Nejspíše proto, že Euler v té době byl ředitelem matematiky na akademii, a tudíž s ohledem na svoji pozici jinak by jistě vyhrál. Mimochodem, v šesté partii o titul světového šampiona v šachu v roce 2010 mezi Viswanathanem Anandem a Veselinem Topalovem bylo možné vidět postupně 13 Anandových tahů jezdcem (ačkoli k tomu použil oba své jezdce). Online komentátoři to v žertu komentovali slovy, že se Anand během této partie snaží vyřešit i jezdcovu procházku.

[22] Jakkoli politicky nekorektní to v dnešní době může být, příručka tímto označením má na mysli výhradně muže a varuje před ženami v podnikání i mimo něj.

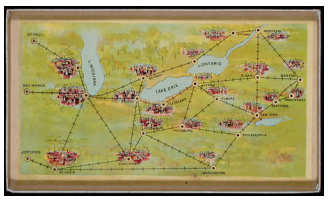
[23] S dodatkem „Von einem alten *Commis-Voyageur*" (od starého obchodního cestujícího). Viz také Obrázek 3.6.

[24] *Obchodní cestující: jaký by měl být a co musí dělat, aby získal zakázky a měl jistotu šťastného úspěchu ve svém podnikání*
(The traveling salesman: how he should be and what he has to do, to obtain orders and to be sure of a happy success in his business – by an old traveling salesman)

Obrázek 3.6: Titulní strana příručky Der Handlungsreisende.

Obrázek 3.7: Cestovní varianta „dvacítkové hry“.

Obrázek 3.8: Cestovní varianta „dvacítkové hry“ (balení).



Obrázek 3.9: Commercial Traveller.

Obrázek 3.10: Worried Woodworm.

Následně Hamilton projevil svého obchodního ducha, když tuto variantu hry prodal společnosti *John Jaques and Son*[40] za £25 a šest exemplářů. Hra byla dokonce vyráběna ve dvou verzích: původní rovinná varianta byla určena do tehdejších „salonů“, zatímco trojrozměrný model sloužil jako cestovní varianta, viz Obrázky 3.7–3.8, jenže se zdá, že pan Jaques svoji investici zpět nezískal – hra byla asi příliš jednoduchá… Což ale nebrání tomu, aby z ní byl unikátní kousek v matematické historii (a muzeu). Pro více podrobností viz [251, Podkapitola 24] a v [64, str. 32–35] je dostupný přetisk původního návodu s patnácti možnými příklady zahájení hry a krátké objasnění v ní ukryté matematiky. Když už jsme u těch her, v roce 1890 společnost McLoughlin Brothers přišla s hrou *Commercial Traveller*, v níž šlo o sestavení vhodné železniční sítě, viz Obrázek 3.9. Tohle by nám mohlo evokovat modernější *Ticket to Ride*[41]. A ještě jednu hru bychom měli zmínit: *Worried Woodworm* od Jamese Dalgetyho z roku 1975, která je další inovací původní Hamiltonovy hry a která již zaznamenala výrazně vyšší počet prodejů. Asi proto, že tentokrát už není tak snadné „uhádnout“ správnou trasu, viz Obrázek 3.10.

První podrobnější matematický pohled na POC nalezneme asi u Karla Mengera[42], který se od roku 1928 věnoval velmi blízkému problému s ryze geometrickým pojetím a ve svém vystoupení 5. února 1930 na vídeňském matematickém [378] oba problémy (téměř) propojil.

> *Wir bezeichnen als Botenproblem (weil diese Frage in der Praxis von jedem Postboten, übrigens auch von vielen Reisenden zu lösen ist) die Aufgabe, für endlichviele Punkte, deren paarweise Abstände bekannt sind, den kürzesten die Punkte verbindenden Weg zu finden. Dieses Problem ist natürlich stets durch endlichviele Versuche lösbar. Regeln, welche die Anzahl der Versuche unter die Anzahl der Permutationen der gegebenen Punkte herunterdrücken würden, sind nicht bekannt. Die Regel, man solle vom Ausgangspunkt erst zum nächstgelegenen Punkt, dann zu dem diesem nächstgelegenen Punkt gehen usw., liefert im allgemeinen nicht den kürzesten Weg.*[43]

Menger sice nepožaduje, aby se poslíček[44] vrátil zpět na výchozí místo, ale to nečiní z jeho problému jiný úkol, neboť přidáním nového bodu s nulovou vzdálenosti od všech ostatních bodů z něj snadno učiníme POC. A tak se POC konečně dostal i do bližší pozornosti matematické komunity, a to zejména těm ve společnosti RAND ve verzi s nalezením nejkratší obchodní cesty z Washingtonu skrze hlavní města dalších 48 států[45] v USA jakožto *problém 48 států Hasslera Whitneyho*[46]. Whitney patřil již od třicátých let 20. století mezi největší příznivce POC a jeho varianta tohoto problému sehrála v dalším rozvoji významnou roli. Proč? To nám mohou přiblížit slova samotného Kuhna.

> *The traveling salesman problem was known by name around Fine Hall by 1949. For instance, it was one of a number of problems for which the RAND corporation offered a money prize. I believe that the list was posted on a bulletin board in Fine Hall in the academic year 1948–49.*[47]

Vážně? Seznam problémů s finanční odměnou od RANDu?! O tomto najdeme zmínku v mnoha zdrojích věnovaných POC, ale bohužel kopii tohoto pozoruhodného dokumentu už asi nenajdeme. Každopádně tohle nepochybně celý výzkum ohledně POC výrazně akcelerovalo. O významu problému 48 států svědčí a asociaci s obchodním cestujícím hovoří slova dalšího z průkopníků operačního výzkumu Flooda[48] v jednom z jeho interview.

> *Developments that started in the 1930s at Princeton have interesting consequences later. For example, Koopmans first became interested in the „48 States Problem“ of Hassler Whitney when he was with me in the Princeton Surveys, as I tried to solve the problem in connection with the work by Bob*

[25] *Obchod přivádí obchodního cestujícího jednou sem, jindy tam a není možné adekvátně určit trasy,*

které jsou vhodné pro všechny možné případy; ovšem někdy vhodnou volbou a uspořádáním zastávek lze získat tolik času, až nelze věřit, že by bylo dovoleno ji obcházet a také opomenout k tomu dát nějaká pravidla. Každý z nich může použít toliko, kolik považuje za užitečné pro svůj cíl; ovšem věříme, že se za ně můžeme na natolik zaručit, že nebude možné uspořádat cestu skrze Německo ekonomičtěji z pohledu vzdálenosti a hlavně, což by mělo každého cestujícího zajímat především, cestu tam a zpět. Hlavní úkol spočívá v návštěvě co nejvíce míst, aniž by cesta vedla dvakrát na téže místo. Viz [449, str. 997].

[26] Seznam měst (s originálními německými názvy), jimiž má procházet požadovaná trasa:
1. Frankfurt am Mein
2. Hanau
3. Aschaffenburg
4. Würzburg
5. Schweinfurt
6. Bamberg
7. Bayreuth
8. Kulmbach
9. Kronach
10. Hof
11. Plauen
12. Greiz
13. Zwickau
14. Chemnitz
15. Freiberg
16. Dresden
17. Meißen
18. Leipzig
19. Halle
20. Merseburg
21. Weißenfels
22. Zeitz
23. Altenburg
24. Gera
25. Naumburg
26. Weimar
27. Rudolstadt
28. Ilmenau
29. Arnstadt
30. Erfurt
31. Greußen
32. Sondershausen
33. Mühlhausen
34. Bad Langensalza
35. Gotha
36. Eisenach
37. Bad Salzungen
38. Meiningen
39. Mellrichstadt
40. Bad Neustadt an der Saale
41. Bischofsheim an der Rhon
42. Gersfeld
43. Bad Brückenau
44. Züntersbach
45. Schlüchtern
46. Fulda
47. Gelnhausen

[27] Tj. trasa v těchto úsecích vede tam a zpět po téže cestě, takže obchodní cestující lokálně prochází dvakrát týmiž místy.

> *Singleton and me on school bus routing for the State of West Virginia. I don't know who coined the peppier name „Traveling Salesman Problem" for Whitney's problem, but that name certainly has caught on, and the problem has turned out to be of very fundamental importance.*[49]

No, ale asi možná víc než o samotný název POC zde jde ve skutečnosti o společnost RAND a její členy, na což poukazují Hoffman a Wolfe.

> *John Williams urged Flood in 1948 to popularize the TSP at the RAND Corporation, at least partly motivated by the purpose of creating intellectual challenges for models outside the theory of games. In fact, a prize was offered for a significant theorem bearing on the TSP. There is no doubt that the reputation and authority of RAND, which quickly became the intellectual center of much of operations research theory, amplified Flood's advertizing.*[50]

Kdo skutečně jako první přišel s označením POC už asi nezjistíme, ale v tištěné podobě můžeme toto prvenství přičíst Julii Robinsonové[51], která ve druhé polovině čtyřicátých letech 20. století strávila cca rok ve společnosti RAND. Zde se věnovala výzkumu v oblasti teorie her, ale její pozornosti neunikl ani POC. O něm sepsala v roce 1949 technickou zprávu, jejíž samotný název *On the Hamiltonian game (a traveling-salesman problem)*[52] obsahuje toto sousloví, které využila i ve stručném popisu samotného problému 48 států, viz [431]:

> *The purpose of this note is to give a method for solving a problem related to the traveling salesman problem. One formulation is to find the shortest route for a salesman starting from Washington, visiting all the state capitals and then returning to Washington. More generally, to find the shortest closed curve containing n given points in the plane.*[53]

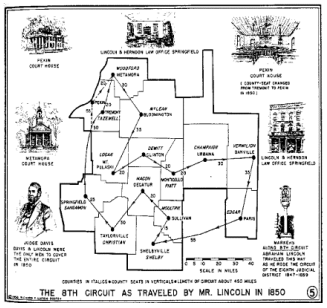
Robinsonová sice ve svém článku iniciovala neúspěšný „lineárně-programovací" přístup k řešení POC skrze speciální případ dopravního problému, ale tato její práce podnítila v RANDu řadu diskuzí ohledně POC, zejména mezi Floodem a jeho výzkumným asistentem Fulkersonem v období let 1950–1952. Toto doposud „nezorané pole" pochopitelně nemohlo uniknout ani Dantzigově pozornosti, který o POC věděl nejméně od konce čtyřicátých let, o čemž svědčí i další Floodova slova:

> *George Dantzig and Tjallings Koopmans met with me in 1948 in Washington, D.C., at the meeting of the International Statistical Institute, to tell me excitedly of their work on what is now known as the linear programming problem and with Tjallings speculating that there was a significant connection with the Traveling Salesman Problem.*[54]

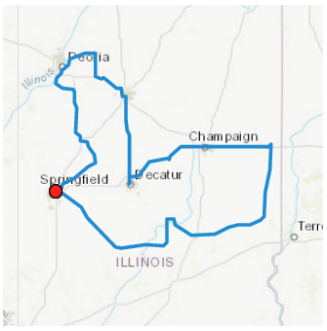
Dantzig se pustil do řešení POC společně s Fulkersonem a Johnsonem. V jejich společném článku [122] popisují, jak se jim během několika týdnů ručních výpočtů podařilo vyřešit problém 48 států s využitím lineárního programování a dodatečných omezení zaručujících skutečnou okružní trasu i samotnou celočíselnost jednotlivých proměnných, viz také [123]. Tento jejich skvělý výsledek nasměřoval další výzkum správným směrem a dostal se i do časopisu *Newsweek*, viz Obrázek 3.11 níže. V úvodu k tomuto článku napsali i toto:

> *The origin of this problem is somewhat obscure. It appears to have been discussed informally among mathematicians at mathematics meetings for many years. Surprisingly little in the way of results has appeared in the*

[28] No, pochopitelně takové cesty mohou vykonávat lidé různých profesí a zmíněný obchodník zde figuruje především jako velmi dobrý marketingový symbol. Za zmínku stojí např. šestnáctý americký prezident Abraham Lincoln, který coby soudní úředník v polovině 19. století musel dvakrát za rok absolvovat několikatýdenní okružní cestu po osmém soudním okruhu státu Illinois. Na této cestě navštívili čtrnáct „okresních" měst, ujeli více než 700 kilometrů a trasu můžeme vidět na Obrázku 3.12. Je optimální? Možná... Ačkoli tím opět můžeme opomenout některá tehdejší specifika, alespoň s ohledem na délky trasy můžeme říci, že existuje i lepší sousloednost jednotlivých měst, viz Obrázek 3.13 s trasou skýtající úsporu cca 75 km.

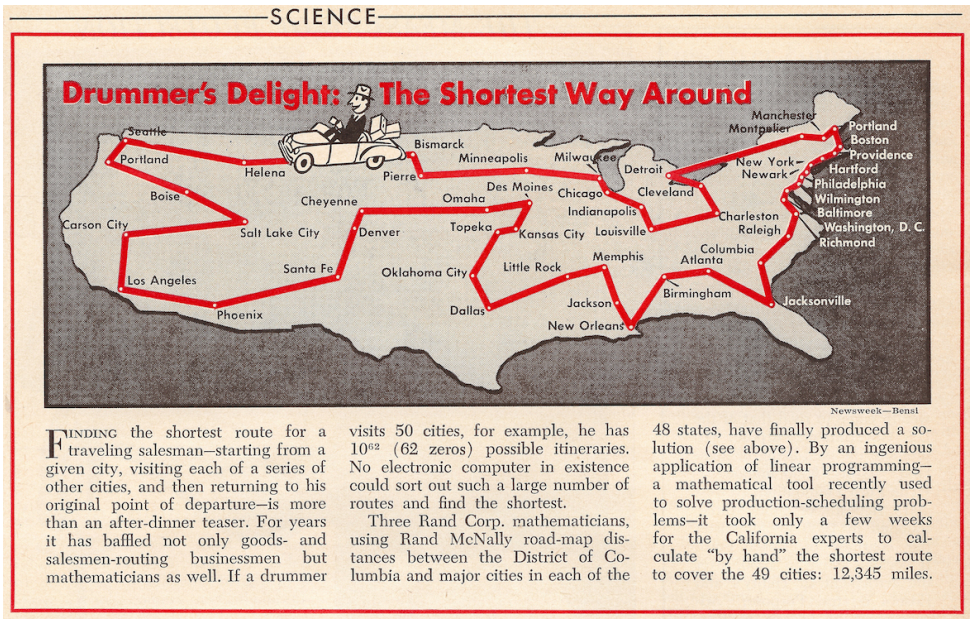


Obrázek 3.12: „Lincolnova" okružní trasa dle Richarda P. Lufkina z jeho vystoupení *Lincoln in 1850* v rámci *Boston Lincoln Group* v roce 1950.



Obrázek 3.13: Optimalizovaná „Lincolnova" okružní trasa.

> *mathematical literature. It may be that the minimal-distance tour problem was stimulated by the so-called Hamiltonian game which is concerned with finding the number of tours possible over a specified network. The latter problem is cited by some as the origin of group theory and has some connections with the famous Four-Color Conjecture. Merrill Flood (Columbia University) should be credited with stimulating interest in the traveling-salesman problem in many quarters. As early as 1937, he tried to obtain near optimal solutions in reference to routing of school buses. Both Flood and A. W. Tucker (Princeton University) recall that they heard the problem first in a seminar talk by Hassler Whitney at Princeton in 1934 (although Whitney, recently queried, does not seem to recall the problem). The relations between the traveling-salesman problem and the transportation problem of linear programming appear to have been first explored by M. Flood, J. Robinson, T. C. Koopmans, M. Beckmann, and later by I. Heller and H. Kuhn.*[55]


SCIENCE

Drummer's Delight: The Shortest Way Around

Newsweek—Bensl

FINDING the shortest route for a traveling salesman—starting from a given city, visiting each of a series of other cities, and then returning to his original point of departure—is more than an after-dinner teaser. For years it has baffled not only goods- and salesmen-routing businessmen but mathematicians as well. If a drummer visits 50 cities, for example, he has 10^{62} (62 zeros) possible itineraries. No electronic computer in existence could sort out such a large number of routes and find the shortest.

Three Rand Corp. mathematicians, using Rand McNally road-map distances between the District of Columbia and major cities in each of the 48 states, have finally produced a solution (see above). By an ingenious application of linear programming—a mathematical tool recently used to solve production-scheduling problems—it took only a few weeks for the California experts to calculate "by hand" the shortest route to cover the 49 cities: 12,345 miles.

Obrázek 3.11: Článek z časopisu Newsweek z 26. července 1954 ohledně Dantzigova, Fulkersonova a Johnsonova řešení problému 48 států z [122].

Vážně si Whitney už nepamatuje na svoji zmínku během některého z vystoupení na semináři? Ve světle předchozích odstavců se to zdá být trochu s podivem. Samozřejmě je to možné, ale naštěstí archív (označovaný jako Pusey Library) Harvardské univerzity obsahuje zhruba 0,1 m^3 Whitneyho dokumentů a různých poznámek včetně přípravy na seminář napsaný krátce po roce 1930, kde se píše:

> *A similar, but much more difficult problem is the following. Can we trace a simple closed curve in a graph through each vertex exactly once? This corresponds to the following problem. Given a set of countries, is it possible to travel through them in such a way that at the end of the trip we have visited each country exactly once?*[56]

Tohle už ale není až tak důležité. Každopádně toto již můžeme definitivně považovat za začátek trnité a zajímavé cesty k řešení POC a i dalších úloh celočíselného programování, z níž alespoň část si společně v následujících podkapitolách projdeme.

29

Thomas Penyngton Kirkman (31. března 1806 – 3. února 1895) byl anglický matematik. Ačkoli byl především knězem anglikánské církve, věnoval se ve volných chvílích matematickému výzkumu. Je znám především díky *problému 15 školaček* zveřejněného jako šestý úkol v časopise pro rekreační matematiku *The Lady's and Gentleman's Diary* (viz Obrázek 3.14) z roku 1850: *Fifteen young ladies in a school walk out three abreast for seven days in succession – it is required to arrange them daily so that no two shall walk twice abreast* (*Patnáct školaček chodí ve trojicích po dobu sedmi po sobě jdoucích dnů – je ale vyžadováno, aby tyto trojice byly uspořádány na jednotlivé dny tak, že žádné dvě dívky nejdou společně v jedné trojici dvakrát*), viz [539, str. 48]. Dnes je Kirkman považován za excentrického amatérského matematika a za muže zmíněného problému. Jenže jeho současníci jej takto rozhodně neviděli a řadili ho mezi silné a originální myslitele, o čemž svědčí i jeho zařazení do seznamu Alexandera Macfarlanea s desítkou předních britských matematiků 19. století. Jenže Kirkman strávil 52 let jako venkovský farář a publikoval většinu své práce v obskurních časopisech s použitím nejasné terminologie a značení, takže je velmi složité hodnotit jeho dílo z dnešního pohledu, obzvláště když svoji nejlepší práci nikdy nepublikoval. To se ale naštěstí již změnilo díky pojednání [63].

Příklad 3.2.1

(*POC*) Jak už jsme několikrát zmínili, v klasickém *problému obchodního cestujícího* jde o někoho, kdo vyjíždí z domu (či odjinud) a chce navštívit $n-1$ dalších měst tak, aby ho to s návratem domů stálo co nejméně (peněz nebo času). Každé město musí navštívit právě jednou a náklady (nebo čas) na cestu z i-tého města do j-tého města jsou určeny hodnotou c_{ij}. Jakou trasu zvolit? Označíme-li

$$x_{ij} = \begin{cases} 1, & \text{vede-li trasa z } i\text{-tého města do } j\text{-tého,} \\ 0, & \text{nevede-li trasa z } i\text{-tého města do } j\text{-tého,} \end{cases}$$

pak můžeme tento problém formulovat jako úlohu celočíselného programování

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\,x_{ij} \to \min$$

s podmínkami

$$\sum_{i=1}^{n} x_{ij} = 1 \quad \text{pro všechna } j \in \{1,\ldots,n\},$$

$$\sum_{j=1}^{n} x_{ij} = 1 \quad \text{pro všechna } i \in \{1,\ldots,n\},$$

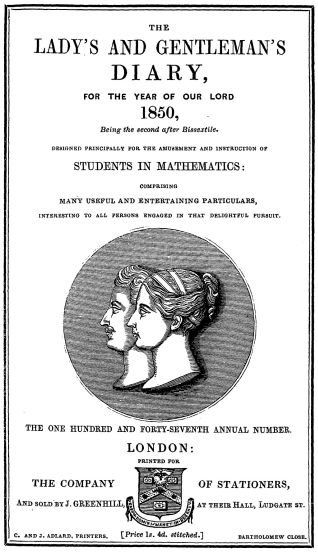
$$x_{ij} \geq 0 \quad \& \quad x_{ij} \in \mathbb{Z} \quad \text{pro všechna } i,j \in \{1,\ldots,n\}.$$

Jednotlivá omezení zaručují, že prodejce každé město navštíví a opustí právě jednou. Jedná se pochopitelně dokonce o binární úlohu, ale vzhledem k uvedeným podmínkám není nutné tento požadavek zmiňovat explicitně (toto záleží spíše na způsobu řešení, ve kterém by případně mohli jít navíc využít binárnost proměnných). Bohužel takto formulovaný problém může vést i k nepřípustnému řešení. Uvážíme-li např. 6 měst, pak řešení může být složeno ze dvou disjunktních tras mezi městy místo jedné trasy, jak můžeme vidět na Obrázku 3.15.
Proto jsou potřeba některé dodatečné podmínky, aby se těmto nepřípustným řešením zabránilo (tzv. *slučovací podmínky/omezení*). V uvedeném příkladě by to mohlo být

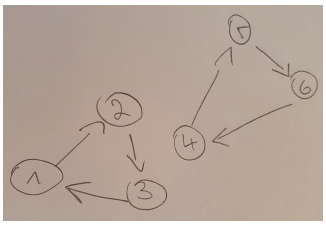
$$x_{14}+x_{15}+x_{16}+x_{24}+x_{25}+x_{26}+x_{34}+x_{35}+x_{36} \geq 1$$

(tj. chceme, aby tam bylo nějaké spojení mezi těmito disjunktními cestami). Toto zaručuje, že bude existovat alespoň jedno spojení mezi městy 1,2,3 a 4,5,6. Mohli bychom zahrnout do našeho problému všechna možná omezení, která by zabraňovala rozdělení měst do dvou skupin. Jenže pro n měst by to znamenalo 2^n-1 dodatečných omezení, takže by takový problém byl velmi rozměrný a početně náročný – exponenciální závislost úlohy na počtu měst je „nepřijatelná“! V praxi bychom mohli postupovat spíše tak, že nejdříve vyřešíme původní problém, a bude-li nalezená optimální trasa založená na rozdělení měst do několika disjunktních skupin, přidáme příslušné slučovací podmínky a zkusíme vyřešit nově získaný problém (tj. a priorní vs. a posteriorní přístup k řešení).

Měli bychom zde ale zdůraznit, že tentokrát jde skutečně o *težký*[57] problém nesrovnatelný s úlohami lineárního programování. Kolik bychom vlastně mohli vytvořit tras pro obchodního cestujícího? V obecném případě čítajícím n-tici míst to je $(n-1)!$ možností, avšak v případě měst lze navíc obvykle očekávat, že úloha je *symetrická*[58], takže to je počet „pouze“ poloviční, tj. $(n-1)!/2$. Tohle číslo se ale s rostoucím počtem měst zvyšuje extrémně rychle, jak můžeme vidět i z Tabulky 3.2. V jejím třetím sloupci je také uveden čas, který by byl potřebný pro projití všech možností v případě, že bychom byli schopni každou vteřinou spočítat 40 milionů tras. Tato čísla mimo jiné ukazují, že *metoda hrubé síly*[59] založená na výpočtu a porovnání všech


THE
LADY'S AND GENTLEMAN'S
DIARY,
FOR THE YEAR OF OUR LORD
1850,
Being the second after Bissextile.
DESIGNED PRINCIPALLY FOR THE AMUSEMENT AND INSTRUCTION OF
STUDENTS IN MATHEMATICS:
COMPRISING
MANY USEFUL AND ENTERTAINING PARTICULARS,
INTERESTING TO ALL PERSONS ENGAGED IN THAT DELIGHTFUL PURSUIT.
THE ONE HUNDRED AND FORTY-SEVENTH ANNUAL NUMBER.
LONDON:
PRINTED FOR
THE COMPANY OF STATIONERS,
AND SOLD BY J. GREENHILL, AT THEIR HALL, LUDGATE ST.
C. AND J. ADLARD, PRINTERS, [Price 1s. 6d. stitched.] BARTHOLOMEW CLOSE.

Obrázek 3.14: Titulní stránka časopisu *The Lady's and Gentleman's Diary* z roku 1850.



Obrázek 3.15: Dvě nežádoucí disjunktní trasy obchodního cestujícího.

Obrázek 3.16: Jeden ze stripů internetového komiksu *Saturday Morning Breakfast Cereal* (SMBC) od Zacha Weinersmitha.

[30] Kirkman v něm pak (v souvislosti s dvojicí tvrzení označených jako A a B) poznamenává: *There is no use in perplexing further the* možných tras nemá v tomto případě žádnou šanci na úspěch[60], viz také strip na Obrázku 3.16.

počet měst	počet možných tras	čas potřebný k otestování
3	1	✖
4	3	✖
5	12	✖
6	60	✖
7	360	✖
8	2 520	✖
9	20 160	✖
10	181 440	0,0045 sekund
11	1 814 400	0,045 sekund
12	19 958 400	0,498 sekund
13	239 500 800	5,98 sekund
17	$1,046 \times 10^{13}$	3,02 dnů
19	$3,201 \times 10^{15}$	2,537 roků
20	$6,082 \times 10^{16}$	48,21 roků
25	$3,102 \times 10^{23}$	$2,45 \times 10^{8}$ roků
47	$2,7513 \times 10^{57}$	$2,181 \times 10^{42}$ roků
48	$1,2931 \times 10^{59}$	$1,025 \times 10^{44}$ roků

Tabulka 3.2: Počet možných tras v symetrickém problému obchodního cestujícího.

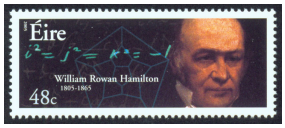
Tohle je skutečně něco jiného než v úlohách lineárního programování. Současně však úspěšné řešení POC by mohlo pro mnoho společností znamenat velké úspory v nákladech na dopravu. V 19. a víceméně i 20. století možná šlo o obchodního cestujícího hledajícího nejkratší cestu především kvůli úspoře času, ale v 21. století jde o zástupce Amazonu, UPS či České pošty doručující velké množství zboží. Proto snad nepřekvapí, že v roce 1962 vyhlásila společnost *The Procter & Gamble* veřejnou soutěž pro hledání řešení POC s 33 městy v USA, o čemž se můžeme dočíst např. v [356], viz také Obrázek 3.17 níže.

> *The traveling salesman problem recently achieved national prominence when a soap company used it as the basis of a promotional contest. Prizes up to $10,000 were offered for identifying the most correct links in a particular 33-city problem. Quite a few people found the best tour. (The tie-breaking contest for these successful mathematicians was to complete a statement of 25 words or less on „I like… because… "). A number of people, perhaps a little over-educated, wrote the company that the problem was impossible – an interesting misinterpretation of the state of the art.*[61]

Jenže ani dnes nemůžeme POC považovat za zcela zvládnutý. Tabulka 3.3 ukazuje, kam jsme se za těch téměř sto let matematického výzkumu dostali s maximálním počtem měst, pro který se podařilo POC vyřešit.

enunciation of B, as its truth follows from, and its extent is parallel with, that of A. Nor is it worth while to trouble the reader with demonstrations of these properties, since they are not properties of all polyedra, a negative which may however be worth the proving. It may suffice, and may perhaps be useful, to show the connexion between these closed polygons and a pleasing mode of representation. (*Nemá smysl se nadále mást tvrzením B, jelikož to vyplývá z A a jeho význam je obdobný. Nestojí ani za to čtenáře obtěžovat důkazy těchto vlastností, jelikož to nejsou vlastnosti všech mnohostěnů – opak, který však může stát to si jej dokázat. Může postačovat a možná může být i užitečnější ukázat spojení mezi těmito uzavřenými čárami a příjemným způsobem znázornění.*) Viz [321, str. 414]. Jinými slovy, Kirkman nás nabádá k tomu, že tvrzení je buď triviální nebo nepravdivé, a zároveň ukazuje poměrně jednoduchou (a správnou) podmínku postačující pro neexistenci takové lomené čáry.

31



Sir William Rowan Hamilton (3. srpna 1805 – 2. září 1865) byl irský matematik, profesor astronomie na Trinity College v Dublinu a irský královský astronom působící jako ředitel observatoře v Dunsinku. Považoval se především za čistého matematika a svoji pozornost upnul zejména k algebře. Podepsal se obzvláště pod algebraický přístup ke komplexním číslům (jakožto uspořádaným dvojicím) abstrahovaný od jejich geometrické interpretace a je autorem tzv. kvaternionů. Avšak zcela zásadním způsobem ovlivnil i fyziku, když svým sofistikovaným matematickým pojetím klasické mechaniky přeformuloval Newtonovy pohybové zákony, čímž dal vzniknout tzv. Hamiltonovské mechanice a tím umožnil moderní studium elektromagnetismu či rozvoj kvantové mechaniky. Jeho otec byl velmi zaneprázdněn obchodem, a tak se mu věnoval především jeho strýc, který byl kurátorem a skvělým lingvistou. Ve třech letech uměl Hamilton již plynule číst anglicky a v pěti letech četl a rozuměl latinským, řeckým i hebrejským textům. V osmi letech ovládal italštinu a francouzštinu. Tím ale výčet jeho jazykových dovedností nekončí. Tři měsíce před desátými

Obrázek 3.17: POC a reklama společnosti The Procter & Gamble.

rok	autoři	počet měst
1954	Dantzig & Fulkerson & Johnson	49
1971	Held & Karp	64
1977	Grötschel	120
1980	Crowder & Padberg	318
1987	Padberg & Rinaldi	532
1987	Grötschel & Holland	666
1987	Padberg & Rinaldi	2 392
1994	Applegate & Bixby & Chvátal & Cook	7 397
1998	Applegate & Bixby & Chvátal & Cook	13 509
2001	Applegate & Bixby & Chvátal & Cook	15 112
2004	Applegate & Bixby & Chvátal & Cook & Helsgaun	24 978
2006	Applegate & Bixby & Chvátal & Cook & Espinoza & Goycoolea & Helsgaun	85 900

Tabulka 3.3: Vývoj rekordního počtu měst ve vyřešeném POC.

narozeninami ovládal hebrejštinu, arabštinu i perštinu a následoval ještě sanskrt či hindustánština. V tomhle výčtu by asi neměla chybět ani čínština, ale získání čínských knih bylo v tehdejší době poměrně složité a nákladné. Co na to říci? Snad jen společně se skotským matematikem a spisovatelem sci-fi, E. T. Bellem: *To which we can only throw up our hands and ejaculate Good God! What was the sense of it all?* (*Nad tím vším můžeme jen sprásknout naše ruce a vykřiknout „Dobrý Bože"! Jaký byl smysl toho všeho?*). Viz [42, str. 375].

[32] Angl. *icosian game*, což má základ ve starořeckém „εἴκοσι" (eíkosi) znamenající dvacet.

[33] Angl. *dodecahedron*.

[34] Jeho duálním protějškem je dvacetistěn (angl. *icosahedron*) s 30 hranami a 12 vrcholy.

[35] Jedná se o tzv. *hamiltonovskou kružnici* (angl. *Hamiltonian cycle*). Kdyby trasa mohla končit jinde než ve výchozím místě, dostali bychom *hamiltonovskou cestu* (angl. *Hamiltonian path*).

[36] Míněn *dvacítkový kalkul* (angl. *icosian calculus*).

[37] Je zde souvislost s Kirkmanovým přístupem? Zdá se, že do jisté míry ano, neboť v dalším dopise z 1. listopadu se o Kirkmanových výsledcích explicitně zmiňuje.

[38] *Zjistil jsem, že některé mladé lidi hodně pobavilo, když vyzkoušeli novou matematickou hru, kterou „dvacítka" poskytuje, přičemž jedna osoba napíchla pět kolíků do pěti po sobě jdoucích bodů … a cílem druhého hráče je pak vložit zbývajících patnáct kolíků, aby v cyklickém sledu pokryly všechny ostatní a skončily v bezprostřední blízkosti prvního soupeřova kolíku, což podle teorie vyložené v tomto dopise je možné vždy.*

[39] *Autor uvedl, že tento kalkul byl zcela odlišný od kvaternionů a že v něm nejsou žádné imaginární kořeny. Poté vysvětlil hlavní rysy nového kalkulu a jeho využití demonstroval zábavnou hrou, kterou nazval „dvacítka" a kterou podle svých slov díky tomu vymyslel — litografii, kterou rozdával mezi členy sekce, a příklady toho, k čemu je hra určena, byly litografovány na okraji, přičemž k ilustraci kalkulu byla ukázána i řešení. Hrací deska byla rovinnou projekcí pravidelného pětiúhelníkového dvanáctistěnu a v každém z rohů byla díra pro vložení jednoho*

Pokud čtenáře při pohledu na předchozí tabulku zarazila potřeba plánovat trasu na procestování 85 900 měst, tak váhá naprosto správně. Alespoň v tuto chvíli. Částečně to je pochopitelně ryze akademická debata a součást souboje (či snad mistrovství světa) v překonávání aktuálních výpočetních hranic[62,63]. Na druhou stranu efektivní řešení úloh POC má široké uplatnění v mnoha oblastech jako je např. tvorba počítačových čipů, viz také [101, Kapitola 3] a [12, Section 2.7].

POC má samozřejmě i řadu různých modifikací. Může jít například o dodatečnou podmínku ohledně časového rámce návštěv jednotlivých měst nebo může být obchodních cestujících více a jde o kompletní pokrytí skupiny měst právě jedním prodejcem. A ani to nemusí být obchodníci, ale třeba může jít o cisterny s benzínem pokrývající svým rozvozem síť čerpacích stanic. Nebo se můžeme zcela oprostit od cestování a přesunout se do továrny, kde bychom mohli chtít seřadit $n-1$ výkonů na jednom stroji, přičemž c_{ij} by značilo náklady na přenastavení stroje na j-tou činnost ve chvíli, kdy byla dokončena i-tá činnost. Jaká posloupnost činností bude mít nejmenší náklady? A jak se změní řešení, pokud budeme navíc požadovat. aby bylo nastavení stroje uvedeno zpět do výchozího stavu (např. zpětné nastavení pro ranní směnu)?

Příklad 3.2.2

(*Umístění skladu*) Při modelování distribučního systému musí být učiněno rozhodnutí ohledně kompromisů mezi přepravními náklady a náklady na provoz distribučních center. Např. uvažme manažera, který musí z n-tice skladů vybrat ty, které uspokojí poptávku m-tice zákazníků po jistém zboží (např. rafinérie a vlastní síť benzinových stanic). Rozhodnutí zahrnuje výběr konkrétních skladů a také množství posílané z jednotlivých skladů k jednotlivým zákazníkům. Označme

$$y_i = \begin{cases} 0, & \text{je-li } i\text{-tý sklad zavřený (nevyužíván),} \\ 1, & \text{je-li } i\text{-tý sklad otevřený (využíván),} \end{cases}$$

a nechť x_{ij} značí množství doručené z i-tého skladu k j-tému zákazníkovi. Relevantní náklady jsou v tomto případě dvojího druhu: f_i představuje provozní náklady i-tého skladu (je-li otevřen) a c_{ij} vyjadřuje jednotkové provozní náklady i-tého skladu plus náklady na dopravu z i-tého skladu k j-tému zákazníkovi. V tomto modelu pak pochopitelně máme i dvojici podmínek:

(i) musí být pokryta poptávka d_j každého zákazníka;

(ii) zboží může být zasláno pouze ze skladu, který je otevřený.

To vše vede k úloze smíšeného celočíselného programování

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}\,x_{ij} + \sum_{i=1}^{n} f_i\,y_j \to \min$$

s podmínkami

$$\sum_{i=1}^{n} x_{ij} = d_j \quad \text{pro všechna } j \in \{1,\dots,m\}$$

$$\sum_{j=1}^{m} x_{ij} - y_i \sum_{j=1}^{m} d_j \le 0 \quad \text{pro všechna } i \in \{1,\dots,n\}$$

$$x_{ij} \ge 0 \quad \text{pro všechna } i \in \{1,\dots,n\} \text{ a } j \in \{1,\dots,m\}$$

$$y_1,\dots,y_n \in \{0,1\}.$$

Účelová funkce zahrnuje fixní náklady plus přepravní a provozní náklady. První omezení znamená, že poptávka každého zákazníka musí být uspokojena. Sumace všech množstevních proměnných x_{ij} v první části druhého omezení určuje množství odesílané z i-tého skladu. Není-li i-tý sklad otevřený, tj. $y_i = 0$, pak toto omezení nutně znamená, že z i-tého skladu nemůže být nic odesláno. Na druhou stranu, je-li i-tý sklad

v provozu, tj. $y_i = 1$, pak toto omezení říká, že množství odesílané z tohoto skladu nemůže převýšit celkovou poptávku (což je trochu něco jiného než v případě první rovnosti). Tedy toto omezení zahrnuje druhý výše uvedený požadavek.
Ačkoli je uvedený příklad velmi zjednodušený oproti praxi, popisuje základní princip sofistikovanějších a realističtějších distribučních modelů s řadou možných všelijakých vylepšení typu:

(i) vícestupňový distribuční systém z továrny do skladu a k zákazníkovi;

(ii) kapacitní omezení na objem výroby a průchodnost skladu;

(iii) úspory vzhledem k velikosti přepravy a provozním nákladům;

(iv) zohlednění kvality služeb jako např. maximální doručovací čas ze skladu k zákazníkovi;

(v) více produktů;

(vi) podmínky zabraňující rozdělení objednávek (v modelu výše je možné, aby zákazník obdržel zboží v různých dávkách z více skladů).

Tato vylepšení je samozřejmě možné zahrnout do modelu různými způsoby. Např. kapacita skladu může být zapracována tak, že ve druhém omezení budeme místo y_i pracovat s $y_i K_i$, kde K_i je kapacita průchodnosti i-tého skladu; vícestupňová distribuce vyžaduje trojité indexování x_{ijk}, což značí množství zaslané z i-té továrny ke k-tému zákazníkovi prostřednictvím j-tého skladu atd.

Abychom však nepůsobili dojmem, že celočíselné programování je pouze o nějakém „obchoďákovi s vysavačem“. S těmito úlohami se můžeme setkat i v jiných podobách, ačkoli nelze vyloučit, že s trochou fantazie by bylo možné i v některých z nich objevit nějakou variaci POC.

Příklad 3.2.3

(*Rozpočtování investic*) V typickém problému investičního rozpočtu jde o výběr několika vhodných investic, např. výběr lokalit pro továrny, nákup (investičního) vybavení nebo výběr projektů výzkumu a vývoje k realizaci. Mnohdy nedává smysl pouze částečné investování do těchto aktivit, takže takové problémy vedou na úlohy celočíselného programování typu ANO×NE, tj. proměnné jsou pouze binární $x_j \in \{0,1\}$ v závislosti na tom, zda j-tá investice je schválena či zamítnuta. Jestliže c_j značí „výnos“ (užitek) z j-té investice a a_{ij} objem i-tého zdroje (např. peníze či pracovní síla) potřebný v j-té investici, přičemž dostupné množství i-tého zdroje je b_i, potom můžeme tuto úlohu zapsat jako

$$\sum_{j=1}^{n} c_j x_j \to \max$$

za podmínek

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \quad \text{pro všechna } i \in \{1,\dots,m\},$$

$$x_j \in \{0,1\} \quad \text{pro všechna } j \in \{1,\dots,n\},$$

tj. maximalizovat celkový výnos investic při dodržení daného omezení zdrojů. Omezení můžeme např. chápat jako podmínky na peněžní tok (cash-flow), tj. nerovnost

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i$$

odráží pravidelnou bilanci v jednotlivých obdobích a číslo a_{ij} značí čistý peněžní tok z j-té investice v i-tém období. Pokud investice vyžaduje dodatečný kapitál v i-tém období, potom $a_{ij} < 0$, zatímco v případě generování kapitálu v i-tém období máme

ze slonovinových kolíků, se kterými se hra hrála.

[40] Tato londýnská firma stále existuje. Na jejich webu se píše, že „Jaques je nejstarší výrobce her a sportovního náčiní na světě. S tradicí předávanou z otce na syna po dobu šesti generací byl Jaques odpovědný za vymýšlení mnoha známých her jako jsou Kroket, Ping Pong, Žebříky a hady (Snakes and Ladders), Ludo, Žabičky (Tiddledy Winks), šachy (The Staunton Chess Set), Happy Families, Snap a mnoho dalších.“

[41] Případně s jiným názvem jako *Jízdenky, prosím* nebo „Zug um Zug“ (německy), *Les Aventuriers du Rail* (francouzsky), *Aventureros al Tren* (španělsky) nebo *Menolippu* (finsky).

[42]

Karl Menger (13. ledna 1902 – 5. října 1985) byl rakouský matematik, který se zabýval algebrou, topologií, teorií křivek a prostoru. Společně s Arthurem Cayleym je považován za jednoho za zakladatelů distanční geometrie, ve které jsou množiny studovány pouze pomocí vzdáleností. Jakožto člen Vídeňského kroužku z období dvacátých let 20. století věnoval také otázce etiky a formálním studiím lidských vztahů. V roce 1934 publikoval důležitý článek ohledně Petrohradského paradoxem a jeho aplikací do teorie užitku v ekonomii. Později společně s Oskarem Morgensternem přispěli k vývoji teorie her. Jeho jméno je spojováno i s tzv. Mengerovou houbou (někdy chybně označovanou jako Sierpińského houba), což je třídimenzionální fraktál zobecňující Cantorovo diskontinuum a Sierpińského koberec.

[43] *Jako poslíčkův problém (protože tuto otázku může v praxi řešit každý listonoš, ostatně i mnoho cestujících) označíme úkol najít nejkratší cestu spojující konečný počet bodů, jejichž vzájemné vzdálenosti známe. Tento problém lze samozřejmě vždy vyřešit konečným počtem pokusů. Pravidla, která by snížila počet pokusů pod celkový počet permutací daných bodů,*

$a_{ij} > 0$. Pravá strana b_i je pravidelný exogenní peněžní tok. Jestliže jsou v i-tém období dostupné dodatečné prostředky, máme $b_i > 0$, zatímco když jsou prostředky odčerpávány (sníženy) v i-tém období, je $b_i < 0$. Tato omezení znamenají, že celkové potřebné náklady na investici musí být menší neb rovny součtu prostředků generovaných před investicí plus dodatečné dostupné prostředky (nebo mínus exogenní odčerpané prostředky).

Investiční model může být samozřejmě mnohem složitější a zohledňovat i některé další přirozené aspekty. Např. uvažme, že investice do nové produktové řady je podmíněna předchozí investicí do nové továrny. Tuto podmíněnost můžeme jednoduše vyjádřit dodatečnou nerovností

$$x_j \geq x_i,$$

což říká, že v případě $x_i = 1$ (tj. je schválen i-tý projekt=vývoj nové produktové řady), musí být také $x_j = 1$ (tj. je schválen j-tý projekt=výstavba nové továrny). Podobnou úvahou se můžeme vypořádat s konfliktními projekty. Např. omezení

$$x_1 + x_2 + x_3 + x_4 \leq 1$$

znamená, že pouze nejvýše jeden z prvních čtyř investičních projektů může být schválen. Takto lze získat omezení „s omezeným výběrem možností“. Zkombinováním těchto podmínek můžeme do úloh tohoto typu zakomponovat řadu složitých interakcí mezi projekty (navíc k problémům s rozdělením zdrojů).

Nejjednodušší úloha tohoto typu má sice pouze jedno omezení, ovšem o to větší pozornost je jí v literatuře věnována. My už jsme se o ní již dříve také zmínili. Jde totiž o maximalizační úlohu

$$\sum_{j=1}^{n} c_j x_j \to \max$$

za podmínek

$$\sum_{j=1}^{n} a_j x_j \leq b$$

$$x_1, \ldots, x_n \in \{0, 1\},$$

což je (binární) *problém batohu* (celočíselná verze problému batohu z Příkladu 2.3.4). Úlohu můžeme snadno rozšířit na *problém kontejneru*, v němž může být od některých (všech) komodit dostupné větší počet kusů, a tudíž jednotlivé proměnné splňují pouze $x_1, \ldots, x_n \in \mathbb{N} \cup \{0\}$. Tento model je důležitý také z toho důvodu, že řada úloh celočíselného programování může být ekvivalentně vyjádřena právě jako problém batohu. Proto při testování metod řešení úloh celočíselné programování leckdy fungoval problém batohu jako vhodný prototyp, jehož úspěšné zvládnutí mohlo motivovat další rozvoj uvažované metody k obecným úlohám (ℤLP). Později (v Kapitole 5) se naučíme řešit tuto úlohu i s pomocí dynamického programování, které nám nabídne poměrně efektivnější způsob.

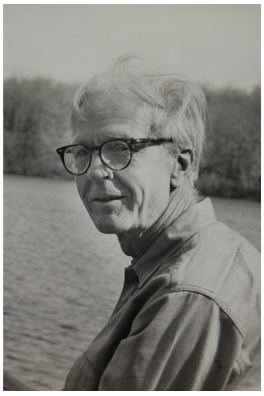
Další důležitou třídu problémů celočíselného programování tvoří řazení, plánování, rozvrhování, trasování (tj. vlastně programování v jeho původním armádním smyslu) apod. Uvažme např. problém se sestavením rozvrhu studentů ÚMS PřF MU(NI). Cílem je najít takový rozvrh předmětů a učeben, aby byl minimalizován počet studentů, pro které je jejich skladba předmětů z registrace nedostupná (tj. konfliktní). Příslušná omezení zahrnují kapacity učeben, jejich počet a dostupnost, časové možnosti vyučujících (a časové preference studentů) atd. Binární proměnné x_{ijn} pak vyjadřují to, zda i-tý student je zařazen do j-tého předmětu v n-tém čase (hodiny během týdne). Matematizace takové úlohy už by ale asi byla poměrně komplikovaná. Variací na toto téma nalezneme celou řadu. Podobným způsobem můžeme přistoupit

nejsou známa. Pravidlo, že by se mělo jít z počátečního bodu nejprve do nejbližšího bodu, pak do bodu nejblíže tomuto bodu atd., obecně neposkytuje nejkratší trasu.

[44] Všimněme si, že stejně jako Guan i Menger zmiňuje listonoše.

[45] Že by nepatrná variace na více než 100 let starou trasu z Der Handlungsreisende se 47 městy?

[46]



Hassler Whitney (23. března 1907 – 10. května 1989) byl americký matematik, který patří mezi zakladatele teorie singularit a výrazně ovlivnil také studium variet, diferenciální geometrie a geometrické integrace. Ve třicátých letech 20. století se věnoval teorii grafů a v článku z roku 1933 položil základy matroidů (abstraktnější a obecnější pojetí lineární nezávislosti ve vektorových prostorech), což je klíčový pojem v moderní kombinatorice a teorii reprezentací (nezávisle na něm přišel s tímto pojetím i Bartel Leendert van der Waerden v polovině třicátých let). V roce 1982 obdržel prestižní *Wolfovu cenu* (mnohými považovanou za „matematickou nobelovku“) za práci v algebraické topologii, diferenciální geometrii a diferenciální topologii. V roce 1985 získal od Americké matematické společnosti také *Cenu Leroye P. Steeleho* za jeho práci na geometrických problémech.

[47] POC byl pod tímto jménem znám lidem z okolí Fine Hall již kolem roku 1949. Například to byl jeden z mnoha problémů, za který společnost RAND nabízela peněžitou odměnu. Myslím, že jejich seznam byl zveřejněn na vývěsce ve Fine Hallu v akademickém roce 1948–49.

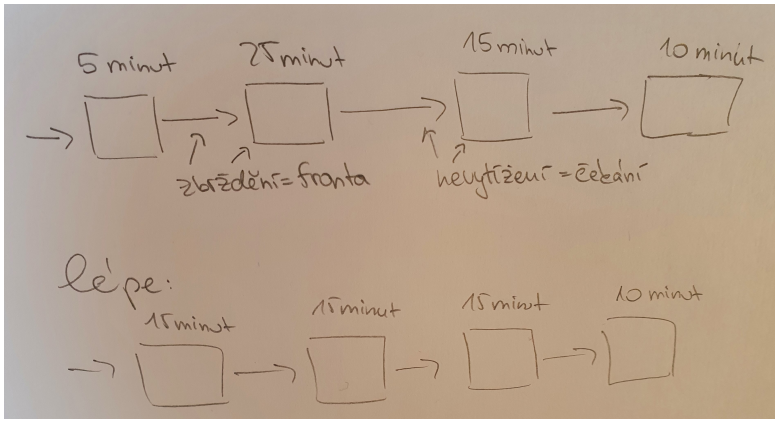
[48]

Merrill Meeks Flood (28. listopadu 1908 – 24. března 1991) byl americký matematik. Největších výsledků se mu podařilo dosáhnout během působení ve společnosti RAND mezi lety 1949–1952, kde pracoval po vedením svého bývalého učitele, kterým byl John von Neumann. Společně pracovali na vybudování základů teorie her, která výrazně ovlivnila vládní jadernou strategii v průběhu studené války. Flood je také znám díky „sestavení" (společně s Melvinem Dresherem) hry *Vězňovo dilema* modelující spolupráci a spor (název hry pochází od Tuckera, který jí dal tuto interpretaci). Jeho práce z roku 1953 dopravním problému v lineárním programování (k tomu jej přivedl Hitchcock) patří k jeho nejcitovanějším pracím z tohoto období. V tomto období se také Flood výrazně věnoval problému obchodního cestujícího a své výsledky zpopularizoval v roce 1956. Flood se výrazně zasloužil i o profesní rozvoj operačního výzkumu a vědy o řízení. Byl hlavním hráčem při zakládání *The Institute of Management Sciences* v USA. Později byl zvolen druhým prezidentem odborné společnosti *Operations Research Society of America.* Za práci pro obě společnosti a revoluční přínos celé disciplíně získal v roce *medaili George E. Kimballa.*

[49] *Vývoj, který začal ve 30. letech v Princetonu, měl později zajímavé důsledky. Například Koopmans se poprvé začal zajímat o „problém 48 států" Hasslera Whitneyho, když byl se mnou na Princeton Surveys a já jsem se pokoušel vyřešit problém související s mojí společnou prací s Bobem Singletonem ohledně trasování školního autobusu ve státě Západní Virginie. Nevím, kdo vymyslel trefnější název „Traveling Salesman Problem" pro Whitneyho problém, ale toto jméno se rozhodně uchytilo a ukázalo se, že tento problém má zásadní význam.*

[50] *John Williams naléhal na Flooda v roce 1948, aby popularizoval POC ve společnosti RAND, k čemuž byl alespoň částečně motivován snahou*

např. k vyvažování výrobních linek (viz Obrázek 3.18), rozvrhování kritické cesty s omezením zdrojů (identifikujeme tzv. „kritické" činnosti projektu, které musí být provedeny, abychom mohli dokončit projekt v minimálním čase, zatímco u těch „nekritických" činností si můžeme dovolit zpoždění, aniž bychom ohrozili celkový čas) nebo k dispečinku (odbavování) vozidel. My se však nyní zaměříme raději na poněkud jednodušší problém.



Obrázek 3.18: Rozvrhování práce na výrobních linkách.

Příklad 3.2.4

(*Rozvrhování*) Uvažme konkrétně pracovní plán palubní posádky letadel. Brněnské aerolinie mají řadu spojů a mezi nejvýznamnější z nich patří let v 10:00 z Brna do Ulánbátaru (Mongolsko) a v 18:00 z Ulánbátaru zpět do Mogadišu (Somálsko). Aerolinie musí sestavit rozvrh cest pro své posádky tak, aby byly obslouženy všechny lety. Jedna posádka může být přidělena na trasu zahrnující oba uvedené spoje. Rozhodovací proměnné pak vyjadřují přidělení posádek na trasy

$$x_j = \begin{cases} 1, & \text{je-li posádka přidělena na } j\text{-tou trasu,} \\ 0, & \text{není-li posádka přidělena na } j\text{-tou trasu.} \end{cases}$$

Označme

$$a_{ij} := \begin{cases} 1, & \text{je-li } i\text{-tý spoj součástí } j\text{-té trasy,} \\ 0, & \text{není-li } i\text{-tý spoj součástí } j\text{-té trasy,} \end{cases}$$

a náklady na zařazení posádky na j-tou trasu jako c_j. Koeficienty a_{ij} určují povolené kombinace spojů a tras, přičemž bereme v úvahu řazení spojů pro možné spojení letů (místo odletu musí být shodné s místem předchozího příletu) a pro zahrnutí spoje do trasy také čas na přípravu (přílet v 11:20 a následný odlet v 11:30 samozřejmě není možný). Tím se dostaneme k úloze

$$\sum_{j=1}^{n} c_j x_j \to \min$$

s podmínkami

$$\sum_{j=1}^{n} a_{ij} x_j = 1 \quad \text{pro všechna } i \in \{1, \ldots, m\}$$

$$x_1, \ldots, x_n \in \{0, 1\},$$

kde i-té omezení $\sum_{j=1}^{n} a_{ij} x_j = 1$ znamená, že právě jedna posádka musí být přiřazena na i-tý spoj. Případně můžeme posádkám povolit přelety, tj. aby letěli jako pasažéři, což by vedlo k omezení

$$\sum_{j=1}^{n} a_{ij} x_j \geq 1.$$

Např. je-li $\sum_{j=1}^{n} a_{1j} x_j = 3$, pak v prvním spoji letí také 2 další posádky jako pasažéři, aby mohly obsloužit jiné spoje, na které byly přiřazeny.

o vytváření intelektuálních výzev pro modely mimo teorii her. Ve skutečnosti byla nabídnuta finanční odměna za významné tvrzení ohledně POC. Není pochyb o tom, že Floodova reklama byla umocněna autoritou společnosti RAND, která se rychle stala intelektuálním centrem velké části teorie operačního výzkumu.

[51]

Julia Hall Bowman Robinson(-ová) (8. prosince 1919 – 30. července 1985) byla americká matematička známá především díky jejím výsledkům v teorii vyčíslitelnosti a teorii složitosti – zejména v souvislosti s rozhodovacími problémy. Její práce na desátém Hilbertově problému týkajícího se algoritmu pro existenci celočíselného řešení libovolné Diofantické rovnice sehrála klíčovou roli v jeho definitivním vyřešení (v roce 1970 se podařilo dokázat neexistenci takového algoritmu a tato tvrzení je dnes známo jako Matijasevičova věta či přesněji Matijasevičova–Robinsonové–Davisova–Putnamova věta). Po tomto úspěchu ji Saunders Mac Lane navrhl za článku Národní akademie věd Spojených států amerických. Alfred Tarski a Jerzy Neyman dokonce letěli do Washingtonu, aby této vážené instituci vysvětlili důležitost a rozsah jejího příspěvku světové matematice – podařilo se to, a v roce 1975 do ní byla zvolena jako první matematička.

[52] Tento název přináší i malou lingvistickou otázku týkající se gramatické správnosti označení „traveling“. Nemělo by být spíše „travelling“? Odpověď je ale celkem prostá. Označení s jedním „l“ je v americké angličtině zcela běžné, a tak najdeme tento problém většinou s tímto názvem.

[53] *Účelem této zprávy je předložení metody pro řešení problému spojeného s problémem obchodního cestujícího. Jednou z jeho formulací je nalezení nejkratší cesty pro obchodníka, který vyrazí z Washingtonu do hlavních měst všech států vrací se zpět do Washingtonu. Obecněji jde o nalezení nejkratší uzavřené křivky procházející n-tic daných bodů v rovině.*

[54] *George Dantzig a Tjallings Koopmans se se mnou setkali v roce 1948 ve Washingtonu na zasedání Mezi-*

Takové binární úlohy jako v předchozím příkladu lze v různých modifikacích nalézt i v jiných oblastech jako jsou ku příkladu doručovací služby (automobilové), počítačové zpracování dat nebo tvorba volebních obvodů, viz např. [25,146,201,272,429]. V případě binární úlohy s m-ticí omezení ve tvaru rovností $\sum_{j=1}^{n} a_{ij}\,x_j = 1$ se označují jako *množinu rozdělující*[64] (neboť spoje jsou rozděleny mezi posádky), pro omezení ve tvaru nerovností $\sum_{j=1}^{n} a_{ij}\,x_j \geq 1$ jako *množinu pokrývající*[65] a pro omezení ve tvaru opačných nerovností $\sum_{j=1}^{n} a_{ij}\,x_j \leq 1$ jako *množinu obalující*[66], přičemž ve všech třech typech úloh platí, že jednotlivé koeficienty v omezeních vždy splňují $a_{ij} \in \{0,1\}$ a v účelové funkci máme pouze $c_i \in \mathbb{Q}$. V případě množinu pokrývajících úloh se jedná o speciální případ tzv. *pokrývajících problémů*[67], ve kterých vystupující omezení $\sum_{j=1}^{n} a_{ij}\,x_j \geq b_i$ a všechny koeficienty jsou nezáporné, tj. $a_{ij} \geq 0$, $b_i \geq 0$ a $c_j \geq 0$. Duální úlohou k tomuto problému pak jsou tzv. *naplňovací problémy*[68], ve kterých jde buď o minimalizaci volného místa při plnění jednoho kontejneru danými předměty nebo o využití co nejmenšího počtu kontejnerů při balení všech daných předmětů.

Takové úlohy lze však objevit i v situacích, kdy o žádnou optimalizaci vůbec nejde. Nicméně to už je údělem matematika, aby nalézal spojení i tam, kde nejsou na první, druhý či nějaký další pohled zřejmá. Prozaičtěji to vyjádřil Stefan Banach[69].

> *Good mathematicians notice the analogies between theories and methods of proof. The very great ones see the analogies between analogies.*[70]

Příklad 3.2.5

Sudoku (příp. magický čtverec) je logická hra s čísly, kterou publikoval Howard Garns v roce 1979 pod názvem „Number Place“. Hra se dočkala obliby zejména v Japonsku, odkud se později vrátila zpět pod názvem sudoku. Cílem hry je doplnit chybějící čísla 1 až 9 v předem dané zčásti vyplněné tabulce, která je rozdělena na 9×9 polí seskupených do 9 čtverců velikosti 3×3, viz Tabulku 3.4. K předem vyplněným číslům je třeba doplnit další čísla tak, aby v každém řádku, v každém sloupci a v každém z devíti čtverců byla použita právě jednou všechna čísla od 1 do 9, tj. na pořadí čísel nezáleží, ale nesmějí se opakovat v žádném sloupci, řádku ani malém čtverci.

	2			3			4	
6								3
		4				5		
			8		6			
8				1				6
			7		5			
		7				6		
4								8
	3			4			2	

Tabulka 3.4: Klasické sudoku.

Celkový počet možností, jak takovým způsobem uvažovaný čtverec vyplnit, je roven[71]

$$6\,670\,903\,752\,021\,072\,936\,960 \approx 6{,}67 \times 10^{21}.$$

To je skutečně velké množství. Naštěstí existují různé návody, jak při hledání správného řešení postupovat. My si nyní ukážeme jeden, který mezi nimi sice asi nenajdeme, protože není zrovna vhodný na „takovéto domácí počítání“, ale o to je matematicky zajímavější, neboť nám umožňuje využít nástroje celočíselného programování.

Pro jednoduchost začněme nejdříve se sudoku velikosti 4×4 (tzv. *shidoku*) a označme jednotlivá pole jejich souřadnicemi, viz Tabulku 3.5.

národního statistického institutu, [sic! zasedání se ve skutečnosti konalo již v září roku 1947] *aby mi s nadšením řekli o jejich práci na tom, co je nyní známo jako problém lineárního programování, a Tjallings při tom uvažoval nad tím, že existuje významné spojení s problémem obchodního cestujícího.*

[55] *Původ tohoto problému je poněkud nejasný. Zdá se, že o tom matematici diskutovali neformálně na matematických setkáních po mnoho let. V matematické literatuře se v tomto směru objevuje překvapivě málo výsledků. Je možné, že problém trasy s minimální vzdáleností byl podnícen takzvanou hamiltonovskou hrou, ve které jde o nalezení počtu možných tras skrze danou síť. Tento problém je některými uváděn jako počátek teorie grup a má určité souvislosti se slavným problémem čtyř barev. Merrillu Floodovi (Kolumbijská univerzita) by měly být připsány zásluhy za podnícení zájmu o problém obchodního cestujícího v mnoha směrech. Již v roce 1937 se snažil získat téměř optimální řešení ohledně trasování školních autobusů. Flood i A. W. Tucker (Princetonská univerzita) si vybavují, že tento problém slyšeli poprvé na přednášce Hasslera Whitneyho v Princetonu v roce 1934 (ačkoli se zdá, že nedávno dotazovaný Whitney si již tento problém nepamatuje). Zdá se, že spojení mezi POC a dopravním problémem lineárního programování bylo poprvé prozkoumáno M. Floodem, J. Robinsonovou, T. C. Koopmansem, M. Beckmannem a později I. Hellerem a H. Kuhnem.*

[56] *Následující problém je sice podobný, ale mnohem obtížnější. Můžeme najít jednoduchou uzavřenou křivku v grafu procházející každým vrcholem právě jednou? To odpovídá následujícímu problému. Je možné danou množinu zemí procestovat takovým způsobem, že na jejím konci bude každá země navštívená právě jednou?*

[57] Toto označení dostane ještě další rozměr v závěrečné Podkapitole 3.6!

[58] Tj. cesta z A do B má stejnou váhu/cenu jako cesta z B do A.

[59] Angl. *the brute force approach* nebo případně *the naive approach.*

[60] Vždyť stáří našeho vesmíru je odhadováno na cca. 14×10^9 roků!

[61] *Problém obchodního cestujícího nedávno dosáhl celostátního významu, když jej drogistická společnost použila jako základ pro-*

[1,1]	[1,2]	[1,3]	[1,4]
[2,1]	[2,2]	[2,3]	[2,4]
[3,1]	[3,2]	[3,3]	[3,4]
[4,1]	[4,2]	[4,3]	[4,4]

Tabulka 3.5: Sudoku velikosti 4×4.

Na každé pole můžeme napsat jedno ze čtyř čísel, takže každému poli přiřadíme čtveřici binárních proměnných určujících, jaké číslo na dané pole napíšeme, tj. celkem máme $4 \times 4 \times 4 = 64$ binárních proměnných

$$\begin{aligned}
[1,1] &\rightsquigarrow x_{1,1,1}, x_{1,1,2}, x_{1,1,3}, x_{1,1,4} \in \{0,1\} \\
[1,2] &\rightsquigarrow x_{1,2,1}, x_{1,2,2}, x_{1,2,3}, x_{1,2,4} \in \{0,1\} \\
&\vdots \\
[4,4] &\rightsquigarrow x_{4,4,1}, x_{4,4,2}, x_{4,4,3}, x_{4,4,4} \in \{0,1\},
\end{aligned}$$

přičemž hodnota $x_{ijk} = 1$ znamená, že na pozici $[i,j]$ napíšeme číslo k. Jelikož jsou na začátku již některá pole obsazena, bude skutečný počet neznámých roven pouze $4 \times (4 \times 4 - \text{počet známých hodnot})$. Pro správné vyplnění celé tabulky musíme ale dodržet čtveřici pravidel, které si můžeme matematizovat následujícím způsobem (nezapomínejme, že některé z podmínek jsou již splněny díky předepsaným hodnotám). To, že do každého pole musíme napsat právě jednu hodnotu, nás přivede k první sadě podmínek

$$\left.\begin{aligned}
x_{1,1,1} + x_{1,1,2} + x_{1,1,3} + x_{1,1,4} &= 1 \\
x_{1,2,1} + x_{1,2,2} + x_{1,2,3} + x_{1,2,4} &= 1 \\
&\vdots \\
x_{4,4,1} + x_{4,4,2} + x_{4,4,3} + x_{4,4,4} &= 1
\end{aligned}\right\} 16\times$$

Další skupinu podmínek dostaneme z požadavku, že každé číslo musí být právě jednou v každém řádku, tj.

$$\left.\begin{aligned}
x_{1,1,1} + x_{1,2,1} + x_{1,3,1} + x_{1,4,1} &= 1 \\
x_{1,1,2} + x_{1,2,2} + x_{1,3,2} + x_{1,4,2} &= 1 \\
&\vdots \\
x_{4,1,4} + x_{4,2,4} + x_{4,3,4} + x_{4,4,4} &= 1
\end{aligned}\right\} 16\times$$

Podobně musí být každé číslo právě jednou v každém ze čtyř sloupců, tj.

$$\left.\begin{aligned}
x_{1,1,1} + x_{2,1,1} + x_{3,1,1} + x_{4,1,1} &= 1 \\
x_{1,1,2} + x_{2,1,2} + x_{3,1,2} + x_{4,1,2} &= 1 \\
&\vdots \\
x_{4,1,4} + x_{4,2,4} + x_{4,3,4} + x_{4,4,4} &= 1
\end{aligned}\right\} 16\times$$

a také v každé 2×2 buňce, tj.

$$\left.\begin{aligned}
x_{1,1,1} + x_{1,2,1} + x_{2,1,1} + x_{2,2,1} &= 1 \\
x_{1,1,2} + x_{1,2,2} + x_{2,1,2} + x_{2,2,2} &= 1 \\
&\vdots \\
x_{3,3,4} + x_{3,4,4} + x_{4,3,4} + x_{4,4,4} &= 1
\end{aligned}\right\} 16\times$$

pagační soutěže. Za nalezení nejlepšího propojení uvedených 33 měst nabízela ceny až do výše 10 000 $. Poměrně dost lidí našlo nejlepší trasu. („Nadstavba" pro tyto úspěšné matematiky spočívala v dokončení prohlášení s pomocí 25 slov nebo méně na téma „Líbí se mi… protože… "). Řada, možná trochu pře-vzdělaných, lidí napsala společnosti, že to není možné – zajímavá dezinterpretace současného stavu poznání.

[62] Existují i jiné varianty POC, v nichž lze dosáhnout celosvětového úspěchu – zejména jde o variantu s nalezením nejdelší cesty s využitím databáze čítající údaje pro 1 904 711 měst a míst ze všech kontinentů (tzv. *World TSP Tour*).

[63] Velká část počítačových odborníků (a možná i matematiků) má značné pochybnosti o existenci efektivního algoritmu pro řešení

DEFINICE 3.3.1

POC s libovolným počtem měst. Nicméně v roce 1976 přišel kyperský matematik Nicos Christofides s celkem přímočarým algoritmem pro nalezení přibližného řešení. Tento algoritmus dokáže najít trasu, která je nejvýše o 50 % delší než ta optimální. Není to moc? Mnoho matematiků a jiných odborníků nepochybně strávili stovky či tisíce hodin ve snaze o vylepšení tohoto algoritmu a zdá se, že se to podařilo v roce 2020. Tehdy se totiž objevil článek (zatím je dostupná pouze jeho nerecenzovaná verze) zlepšující tuto nepřesnost, i když to je „jen" o méně než 0,1 kvintiliardtin (tj. méně než 10^{-34}) procentních bodů. David P. Williamson (profesor operačního výzkumu a informačního inženýrství na Cor-

LEMMA 3.3.2

nellově univerzitě a bývalý vedoucí editor prestižního časopisu SIAM Journal on Discrete Mathematics) o tomto výsledku prohlásil „This is a result I have wanted all my career. It is the first step towards showing that the frontiers of efficient computation are in fact better than what we thought." (Toto je výsledek, který jsem chtěl celou svou kariéru. Je to první krok k ukázání, že hranice efektivních výpočtů jsou ve skutečnosti lepší, než jsme si mysleli.)

[64] Angl. *set partitioning problems.*

[65] Angl. *set covering problems.*

[66] Angl. *set packing problems.*

[67] Angl. *covering problems.*

[68] Angl. *packing problems.*

Celkem tedy máme nejvýše $4 \times 16 = 64$ podmínek. Analogickými úvahami snadno zjistíme, že v klasickém sudoku půjde nejvýše o $9 \times 9 \times 9 = 729$ binárních proměnných a $4 \times 9 \times 9 = 324$ omezení. Co s tím má ale společného celočíselné programování? Sice zde nehledáme žádné optimum, ale vyřešení sudoku je vlastně totéž jako nalezení přípustného bodu úlohy, ve které např. minimalizujeme umělou proměnnou z za výše popsaných omezení.

3.3 Celočíselnost i bez celočíselnosti

V Podkapitole 2.11 jsem se již zmínili o tom, že speciální struktura dopravního problému zaručuje celočíselnost jeho řešení, aniž bychom si to jakkoli explicitně či implicitně vynutili. To se pochopitelně může stát i v jiných úlohách lineárního programování a my se v této podkapitole na onu „speciální strukturu" podíváme blíže. Předně ale zdůrazněme, že toto souvisí pouze s maticí A určující omezení, a my se tím tak dostaneme i do oblasti tzv. *polyedrální kombinatoriky*[72] a *kombinatorické optimalizace*[73].

Nechť je dána matice $A \in \mathbb{R}^{m\times n}$. Jestliže determinant každé její regulární čtvercové submatice je roven ± 1, nazýváme tuto matici *totálně unimodulární.*

Pochopitelně matice A může mít i singulární submatice, tj. s nulovým determinantem, proto z této definice vyplývá, že matice může být totálně unimodulární pouze v případě, že je celočíselná a obsahuje jenom 0 a/nebo ± 1, tj. $A \in \{0, \pm 1\}^{m\times n}$. Opačné tvrzení již samozřejmě neplatí, jak ukazuje např. matice $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ s determinantem rovným -2 a inverzí $\left(\begin{smallmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{smallmatrix}\right)$. Zjištění, zda daná matice je totálně unimodulární či nikoli, může být vzhledem k celkovému počtu submatic poměrně komplikované. Čtvercové submatice jsou totiž zcela určeny volbou řádků a sloupců, takže submatic velikosti $r \times r$ je celkem $\binom{n}{r} \times \binom{m}{r}$. Tudíž pro $m \times n$ matici je počet všech možných submatic roven

$$\sum_{r=1}^{\min\{m,n\}} \binom{n}{r} \times \binom{m}{r},$$

např. pro $m = n = 3$ jde o 19 submatic, zatímco přidáním 4 řádků a 2 sloupců (tj. pro $n = 7$ a $m = 5$) už tento počet vzroste na 791. V tomto rozhodování nám mohou pomoci i následující dvojice tvrzení.

Nechť je matice $A \in \mathbb{R}^{m\times n}$ totálně unimodulární. Pak také matice $-A \in \mathbb{R}^{m\times n}$, $A^\top \in \mathbb{R}^{n\times m}$, $(A, \pm A) \in \mathbb{R}^{m\times 2n}$ a $(A, \pm I) \in \mathbb{R}^{m\times(m+n)}$ jsou totálně unimodulární. Navíc, je-li $m = n$ a $\det A \neq 0$, pak i matice A^{-1} je totálně unimodulární.

Důkaz. Totální unimodularita matic $-A$ a $A^\top$ vyplývá ihned z vlastností determinantu. Matice $(A, \pm A)$ je singulární a determinant její jakékoli čtvercové submatice může být nenulový pouze v případě, kdy bude obsahovat lineárně nezávislé sloupce. Tím ale získáme submatici, kterou můžeme ztotožnit i s vhodnou submaticí samotné matice A, a tedy její determinant bude díky totální unimodularitě roven ± 1.

Předposlední část tvrzení dokážeme indukcí. Je zřejmé, že determinant libovolné 1×1 submatice matice $(A, \pm I)$ bude roven 0 nebo ± 1. Připusťme, že totéž platí pro libovolnou $k \times k$ submatici s $1 \leq k < \min\{m, n\}$. Ukážeme, že pak i determinant libovolné $(k+1) \times (k+1)$ submatice bude buď 0 nebo ± 1. Bude-li tato submatice tvořena pouze sloupci matice A, pak to vyplývá z totální unimodularity matice A. Obsahuje-li tato submatice nějaký sloupec jednotkové matice (případně se záporným znaménkem), pak můžeme determinant vypočítat pomocí Laplaceova rozvoje podle tohoto sloupce, takže determinant bude až na znaménko roven determinantu odpovídající $k \times k$ submatice. Ten ale podle indukčního předpokladu musí být 0 nebo ± 1, což jsme chtěli ukázat.

Konečně, je-li A čtvercová a regulární matice, pak její inverzi můžeme díky Cramerovu pravidlu spočítat s pomocí adjungované matice tvořené jednotlivými kofaktory[74] , neboť platí $A^{-1} = \frac{1}{\det A} A^{\mathrm{adj}}$. Avšak tato rovnost společně totální unimodularitou matice A implikuje pouze celočíselnost všech složek A^{-1}. Pro důkaz našeho tvrzení potřebujeme trochu jiný výsledek: označíme-li pro libovolné dvě neprázdné indexové množiny $I, J \subset \{1,\dots,n\}$ symbolem $A(I,J)$ submatici tvořenou pouze řádky a sloupci odpovídajícími indexům z I (pro řádky) a J (pro sloupce)[75], pak pro determinanty vhodných submatic matic A a A^{-1} platí rovnost

$$\left|\det A^{-1}(I,J)\right| = \frac{1}{\left|\det A\right|}\left|\det A(I^{\mathrm{C}},J^{\mathrm{C}})\right|,$$

kde I^{C} a J^{C} značí doplňky množin I, J v $\{1,\dots,n\}$, viz [490, Lemma 1.34]. Jelikož ale pro všechny povolené volby množin I, J jsou determinanty na pravé straně rovny pouze 0 nebo ± 1, musí totéž platit pro libovolnou submatici matice A^{-1} a současně $\det A^{-1} = (\det A)^{-1} = \pm 1$, což dokazuje totální unimodularitu matice A^{-1}. ■

Samozřejmě platí i opačné tvrzení, tj. je-li kterákoli z pěti matic v Lemma 3.3.2 totálně unimodulární, pak také samotná matice A je totálně unimodulární. V případě některých těchto matic to také úzce souvisí s tím, že totální unimodularita se v rámci jedné matice „dědí“, tj. je-li A totálně unimodulární, pak i každá její submatice je totálně unimodulární. Ještě si ukážeme alespoň jedno trochu sofistikovanější kritérium pro rozhodnutí o totální unimodularitě, viz také [88, 214, 281, 485, 507].

Věta 3.3.3

Nechť je dána matice $A \in \{0,\pm 1\}^{m\times n}$. Jestliže každý sloupec matice A obsahuje nejvýše dva nenulové prvky a její řádky můžeme rozdělit do takových dvou indexových množin $\alpha_1, \alpha_2 \subseteq \{1,\dots,m\}$, že

(i) jsou-li v jednom sloupci dva nenulové prvky se stejným znaménkem, pak řádky obsahující tyto prvky patří do různých množin;

(ii) jsou-li v jednom sloupci dva nenulové prvky s opačnými znaménky, pak řádky obsahující tyto prvky patří do téže množiny,

pak je matice A totálně unimodulární.

Důkaz. Tvrzení dokážeme pomocí indukce. Každá 1×1 submatice má jistě determinant roven 0 nebo ± 1 díky požadavku na jednotlivé prvky matice A. Připusťme, že totéž platí pro libovolnou $k\times k$ submatici s $1 \le k < \min\{m,n\}$. Ukážeme, že pak i determinant libovolné $(k+1)\times(k+1)$ submatice bude buď 0 nebo ± 1. Pokud tato submatice obsahuje nějaký sloupec s nejvýše jedním nenulovým prvkem, pak s pomocí Laplaceova rozvoje podle tohoto sloupce a indukčního předpokladu získáme požadovanou hodnotu determinantu. Jsou-li v každém sloupci právě dva nenulové prvky, pak využijeme rozdělení řádků do množin α_1 a α_2, ze kterého vyplývá, že součet řádků z těchto dvou skupin bude stejný[76], tj. tyto řádky jsou lineárně závislé, a tudíž determinant takové submatice bude nulový. ■

Příklad 3.3.4

Rozhodněme o totální unimodularitě matice

$$A = \begin{pmatrix} -1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 0 & 1 \end{pmatrix}.$$

Řešení. Zkusíme využít Větu 3.3.3. V každém sloupci jsou nejvýše dva nenulové prvky, přičemž pouze ve třetím, čtvrtém a pátém sloupci máme dva nenulové prvky. Protože se vždy jedná o dvojice s opačnými znaménky, budou odpovídající řádky patřit do stejné množiny, tj. $\alpha_1 = \{1,2,3,4\}$, a tudíž zbývá $\alpha_2 = \emptyset$. To ukazuje, že matice A je totálně unimodulární.

[69] Stefan Banach (30. března 1892 – 31. srpna 1945) byl polský matematik. Je považován za zakladatele tzv. *funkcionální analýzy*, což je velmi důležitá oblast v moderní matematice s řadou aplikací (zejména ve fyzice a ekonomii). Bohužel, jeho život fatálně ovlivnila druhá světová válka a zvěrstva nacistického režimu. Po nacistické okupaci Lvovu žil za velmi těžkých podmínek a podobně jako jeho mnozí kolegové a jeho syn byl nucen se živit jako krmič vší vlastní krví na Výzkumném ústavu skvrnitého tyfu a virů prof. Rudolfa Weigla.

[70] *Dobrý matematik si všimne analogií mezi teoriemi a mezi metodami důkazů. Ten naprosto nejlepší vidí analogie mezi analogiemi.* Viz [487, str. 477].

[71] Výpočet této hodnoty je (možná nečekaně) poměrně komplikovaný. Se sudoku je vlastně spojena řada dalších nezodpovězených otázek. Např. jaký je nejmenší počet předvyplněných polí, aby existovalo jediné řešení? Pro sudoku velikosti 4×4 to jsou 4 pole, avšak pro klasické 9×9 sudoku je zatím odpověď neznámá. Zdá se, že je to 17 polí, i když klasický matematický důkaz stále chybí, pro více podrobností viz [377, 436, 471].

[72] Jedná se o oblast matematiky propojující kombinatoriku a diskrétní geometrii, ve které jsou studovány počty a další charakteristiky stěn a vrcholů konvexních polyedrů a polytopů.

[73] Jedná se o část matematické optimalizace blízké k operačnímu výzkumu, teorie algoritmů a teorie (výpočetní) složitosti s využitím např. v umělé inteligenci, strojovém učení, teorii aukcí, softwarovém inženýrství či teoretické informatice. Zahrnuje úlohy, ve kterých jde o nalezení optimálního objektu v diskrétní množině (případně jde o úlohy, které lze do této podoby převést) jako např. problém obchodního cestujícího, problém batohu či hledání minimální kostry grafu.

[74] *Kofaktorem* nebo též *algebraickým doplňkem* ve čtvercové matici rozumíme číslo $\hat{a}_{ij} := (-1)^{i+j}\det A_{ij}$, kde A_{ij} značí submatici vzniklou vynecháním i-tého řádku a j-tého sloupce matice A. Matice $\hat{A} := (\hat{a}_{ij})_{i,j=1,\dots,n}$ složená z těchto kofaktorů se nazývá *ko-*

Předchozí příklad je poněkud specifický, neboť pochopitelně ne vždy to dopadne tak, že jedna z indexových množin bude prázdná. Věta 3.3.3 sice nabízí pouze postačující podmínku a pochází od Hoffmana, ale Galeovi se podařilo ukázat, že pro totální unimodularitu matic mající nejvýše dva nenulové prvky v každém sloupci, jsou tyto podmínky dokonce nutné, viz apendix článku [259]. Kompletní charakterizaci totální unimodularity v podobě poměrně komplikovaných nutných a postačujících podmínek nalezneme v Seymourově[77] článku [443]. Toto kritérium je také jádrem programu dostupného na webu

https://ln.math.muni.cz/pz/opt/TotUnimod

s jehož pomocí lze celkem snadno otestovat totální unimodularitu libovolné matice. S konceptem totálně unimodulárních matic se můžeme setkat (ačkoli bez tohoto pojmenování) již v Poincarého článku [415, str. 366 a 369]. Jak to ale souvisí s úlohami celočíselného programování? Odpověď na tuto otázku nalezneme v Hoffmanově a Kruskalově[78] článku [270] (viz také [301, Chapter 3]), o kterém Hoffman později napsal:

> *In this paper the concept (not the name) of total unimodularity was shown to be a neat explanation (via Cramer's rule) of the fact that some linear programming problems have all their vertices integral. I do not think that this paper would have ever been accepted for publication if we had not fancied it up with a soupçon of generalization: the main idea is too obvious and folklorish. And we also thought we introduced a new class of matrices with the „unimodular property", but Jack Edmonds later found that our new class wasn't really new after all. It is nevertheless true that totally unimodular matrices (as Berge christened them), and unimodular matrices generally are key to understanding how linear programming duality underlies a wide variety of extremal combinatorial analysis. Incidentally, Joe Kruskal, the author of fundamental papers in combinatorics, combinatorial optimization and the theory of ordered sets is now a statistician.*[79]

Tím zmíněným klíčem je následující tvrzení.

Věta 3.3.5 Nechť matice $A \in \mathbb{R}^{m\times n}$ je totálně unimodulární. Pak polyedr $X := \{x \in \mathbb{R}^n \mid Ax = b \ \&\ x \geq 0\}$ je pro jakékoli $b \in \mathbb{Z}^m$ celočíselný, tj. všechny jeho vrcholy (neboli bazické přípustné body) mají pouze celočíselné souřadnice.

Důkaz. Uvažme libovolný BPB $x \in X$. Stačí ukázat, že bazické souřadnice tohoto bodu jsou celočíselné[80]. Jelikož příslušná čtvercová matice A_B je regulární a současně také totálně unimodulární, platí podle Lemma 3.3.2 totéž i pro matici A_B^{-1}, tj. zejména $A^{-1} \in \{0,1\}^{m\times m}$. Proto tvrzení vyplývá z celočíselnosti vektoru b a rovnosti $x_B = A_B^{-1}\, b$, kterou jsme si odvodili v předchozí kapitole. ∎

To znamená, že libovolná úloha lineárního programování v kanonickém tvaru s omezením daným totálně unimodulární maticí bude vždy mít celočíselné řešení, v kterémžto případě můžeme pro jeho nalezení bez obav využít simplexový algoritmus[81]. Navíc odtud společně s Větou 2.13.5 obdržíme následující důsledek.

Důsledek 3.3.6 Úloha celočíselného programování s totálně unimodulární maticí A a celočíselným vektorem b je řešitelná algoritmem s polynomiální časovou složitostí.

Poznámka 3.3.7. Celočíselnost polyedru X jsme již (trochu skrytě) využili při řešení dopravního problému z Podkapitoly 2.11. Ten je totiž ve skutečnosti místo úlohy lineárního programování (2.11.1) spíše úlohou celočíselného programování

$$\langle \mathbf{n}, \mathbf{x}\rangle \to \min \quad \& \quad \mathbf{Ax} = \mathbf{b} \quad \& \quad \mathbf{x} \in (\mathbb{N}\cup\{0\})^{mn},$$

faktorová a jejím transponováním obdržíme *adjungovanou matici*, tj. $A^{\text{adj}} := \hat{A}^\top$.

[75] Tj. v souladu s naším předchozím značením máme $A_{ij} = A(I,J)$ pro $I = \{i\}$ a $J = \{j\}$.

[76] Řádky budou mít na stejných pozicích buď jen 0 nebo 0 a 1 nebo 1 a −1, takže jejich součtem bude vždy řádek obsahující jen 0 a 1.

[77] Paul D. Seymour (⋆26. července 1950) je britský matematik věnující se diskrétní matematice. Je znám především díky svým výsledkům ohledně totálně unimodulárních matic, Větě o čtyřech barvách a dalších výsledků z oblasti teorie grafů. Od roku 1996 je profesorem na Princetonské univerzitě.

[78] Joseph Bernard Kruskal, Jr. (19. ledna 1928 –– 19. září 2010) byl americký matematik, statistik, informatik a psychometrik. V letech 1959–1993 pracoval v Bellových laboratořích. Je znám především díky sestrojení algoritmu hledajícího minimální kostru ve váženém grafu. Jeho matka byla známou popularizátorkou origami v počátcích televizního vysílání.

[79] *V tomto článku bylo ukázáno, že koncept totální unimodularity (nikoli s tímto jménem) je přesným objasněním (pomocí Cramerova pravidla) toho, proč v některých úlohách lineárního programování mají všechny vrcholy* [polyedru, pozn. překl.] *celočíselné souřadnice. Nemyslím si, že by tento článek byl vůbec přijat k publikování, kdybychom jej nevyšperkovali trochou zobecnění: hlavní myšlenka je příliš zřejmá a zlidovělá. Také jsme si mysleli, že jsme zavedli novou třídu matic a „unimodulární vlastností", ale Jack Edmonds později zjistil, že naše nová třída nebyla vůbec nová. Nicméně je pravdou, že totálně unimodulární matice (jak je pokřtil Berge) a celkově unimodulární matice jsou klíčem k porozumění tomu, proč dualita v lineárním programování tvoří základ široké škály extremální kombinatorické analýzy. Mimochodem, Joe Kruskal, autor zásadních článků z kombinatoriky, kombinatorické optimalizace a teorie uspořádaných množin je nyní statistikem.* Viz [268, str. 217].

[80] Pro nebazické souřadnice to platí automaticky, neboť jsou nulové.

[81] Funkčnost simplexového algoritmu i na úlohy z kategorie kombinatorické optimalizace byla pro mnohé překvapující, viz Hofmannovy osobní vzpomínky v [269].

kde bereme

$$\mathbf{x} := \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} & x_{21} & x_{22} & \cdots & x_{2n} & x_{31} & \cdots & x_{mn} \end{pmatrix}^\top \in \mathbb{R}^{nm},$$
$$\mathbf{n} := \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1n} & n_{21} & n_{22} & \cdots & n_{2n} & n_{31} & \cdots & n_{mn} \end{pmatrix}^\top \in \mathbb{R}^{nm},$$
$$\mathbf{b} := \begin{pmatrix} s_1 & \cdots & s_m & c_1 & \cdots & c_n \end{pmatrix}^\top \mathbb{R}^{m+n}$$

a levá strana omezení je určena maticí (na prázdné pozice patří nuly)

$$\mathbf{A} := \begin{pmatrix} 1 & 1 & \cdots & 1 & & & & & & & & & & & \\ & & & & 1 & 1 & \cdots & 1 & & & & & & & \\ & & & & & & & & \ddots & & & & & & \\ & & & & & & & & & & 1 & 1 & \cdots & 1 & \\ & & & & & & & & & & & & \ddots & & \\ 1 & & & & 1 & & & & 1 & & & & 1 & & \\ & 1 & & & & 1 & & & & 1 & & & & 1 & \\ & & \ddots & & & & \ddots & & & \ddots & & & & \ddots & \\ & & & 1 & & & & 1 & & & & 1 & & & 1 \end{pmatrix} \in \mathbb{R}^{(m+n)\times mn}.$$

Všimněme si, že každý $[(i-1)\,n+j]$-tý sloupec matice $\mathbf{A}$ lze vyjádřit jako $\mathrm{e}_i + \mathrm{e}_{m+j}$ pro všechna $i = 1,2,\ldots,m$ a $j = 1,2,\ldots,n$, kde $\mathrm{e}_k \in \mathbb{R}^{mn}$ představuje k-tý jednotkový vektor[82]. Ve Větě 2.11.7 jsme ukázali všechny BPB této úlohy jsou celočíselné, což je ve skutečnosti důsledek toho, že daná matice $\mathbf{A}$ je totálně unimodulární. Tuto matici totiž můžeme chápat jako *matici incidence bipartitního grafu*[83]. V matici incidence $\mathbf{A}$ jednotlivé řádky představují dodavatele a odběratele, zatímco sloupce představují jednotlivá spojení mezi nimi. Je-li tedy v matici $\mathbf{A}$ na nějaké pozici ± 1, znamená to, že odpovídající dodavatel nebo odběratel využívá příslušné spojení. Dá se ukázat, že[84] $\operatorname{rank}\mathbf{A} = n+m-1$, tj. každý BPB má nejvýše $n+m-1$ nenulových prvků. Navíc každá čtvercová submatice má determinant roven 0 nebo ± 1. Vskutku. V každém sloupci matice $\mathbf{A}$ jsou právě dva nenulové prvky, a tudíž v každé submatici budou mít všechny sloupce nejvýše dva nenulové prvky. Obsahuje-li submatice sloupec s pouze nulovými prvky, pak je determinant jistě nulový. Naopak, jsou-li v každém sloupci zvolené submatice dva nenulové prvky, pak první z nich pochází z některého z prvních m řádku a druhý je z posledních n řádků. Potom součet řádků odpovídajících první m-tici řádků je stejný jako součet řádků odpovídajících poslední n-tici řádků, což ukazuje jejich lineární závislost, a tudíž determinant této submatice je opět nulový. Zbývá situace, kdy alespoň jeden sloupec obsahuje právě jeden nenulový prvek. Máme-li $k \times k$ submatici $\mathbf{A}_k$, pak Laplaceovým rozvojem podle tohoto sloupce dostaneme

$$\det \mathbf{A}_k = \pm \det \mathbf{A}_{k-1},$$

kde znaménko $\pm$ závisí na pozici tohoto nenulového prvku. Opakováním předchozích argumentů pro submatici $\mathbf{A}_{k-1}$ získáme požadovaný výsledek $\det \mathbf{A}_k \in \{0, \pm 1\}$, a tedy i totální unimodularitu matice $\mathbf{A}$.

Speciálním případem dopravního problému je tzv. *přiřazovací problém*, kde $m = n$ a $s_1 = \cdots = s_n = c_1 = \cdots = c_n = 1$ a $x_{ij} \in \{0,1\}$ pro všechna $i, j = 1,\ldots,n$. O takové úloze jsme se již zmínili v Příkladě 2.3.8 nebo v něm může jít např. o přidělení n činností mezi n-tici strojů. Každopádně předchozí výsledek ukazuje, že i tuto úlohu lze řešit pouze pomocí simplexového algoritmu. ▲

[82] Tj. s jedničkou na k-té pozici a nulami na všech ostatních pozicích.

[83] To je dvojice vrcholů V a (orientovaných) hran H spojujících tyto vrcholy, přičemž vrcholy lze rozdělit do dvou disjunktních skupin a žádné dva vrcholy z téže skupiny nejsou spojeny hranou. Matice incidence pak obsahuje pouze 0 a ± 1. Řádky této matice odpovídají vrcholům a sloupce hranám, přičemž $a_{ij} = \pm 1$ znamená, že i-tý vrchol je spojen s jiným vrcholem j-tou hranou a v případě $+1$ hrana směřuje do tohoto vrcholu, zatímco pro -1 zde hrana končí. Hodnota $a_{ij} = 0$ znamená, že j-tá hrana není spojena s i-tým vrcholem. Je-li graf neorientovaný (nerozlišujeme, zda hrana z nějakého vrcholu vychází nebo v něm končí), pak prvky matice jsou pouze 0 a 1.

[84] Je jasné, že $\operatorname{rank}\mathbf{A} \le n+m$. Je také snadno vidět, že součet prvních m řádků je stejný jako součet dalších n řádku (tj. $(m+1)$-ho až $(m+n)$-tého), tj. těchto $m+n$ řádku je lineárně závislých, a tudíž $\operatorname{rank}\mathbf{A} \le m+n-1$. Nyní sestavíme regulární $(m+n-1)\times(m+n-1)$ submatici matice $\mathbf{A}$. Vezmeme-li n-tý, $2n$-tý, ... a mn-tý sloupec matice $\mathbf{A}$ společně s prvním, druhým, ..., $(n-1)$-ním sloupcem, pak po odstranění posledního řádku dostaneme $(m+n-1)\times(m+n-1)$ matici

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix},$$

jejíž determinant je roven 1. To ukazuje, že skutečně $\operatorname{rank}\mathbf{A} = n+m-1$.

Příklad 3.3.8

Již dříve jsme se zmínili o společnosti The Procter & Gamble a její soutěže ohledně POC z roku 1962. To ale nebyla její jediná zkušenost s operačním výzkumem. Společ-

nost byla založena kolem roku 1830 a od té doby se postupně rozrůstala. Dnes vyrábí a prodává spotřební zboží více 300 značek ve více než 140 zemích celého světa. Aby se tento růst udržel a urychlil, zahájila společnost v roce 1993 snahy o zesílení globální efektivity, čímž chtěla dosáhnout zefektivnění pracovních a dodavatelských procesů, odstranění duplicit a nákladů bez přidané hodnoty. Klíčovou součástí těchto snah byla studie operačního výzkumu zaměřená na dodavatelský řetěz v Severní Americe (své provozy měla společnost ve více než 58 zemích). To zahrnovalo více než 50 produktových značek, 60 výrobních závodů, 10 distribučních center a stovky zákaznických zón. Studie takových rozměrů do té doby by bylo možné najít snad jen u armády. Výsledkem bylo snížení počtu závodů o téměř 20 procent, což vedlo k úspoře nákladů téměř 200 milionů dolarů za rok (před zdaněním). Hlavní část studie se věnovala dopravním problémům pro jednotlivé kategorie výrobků, které vypadaly takto

$$\sum_{i\in I}\sum_{j\in J} C_{ij}X_{ij} \to \min$$
$$\sum_{i\in I} X_{ij} = 1 \quad \text{pro všechna } j \in J$$
$$X_{ij} \le Y_i \quad \text{pro všechna } i \in I \text{ a } j \in J$$
$$\sum_{i\in I} Y_i = k$$
$$X_{ij}, Y_i \in \{0,1\} \quad \text{pro všechna } i \in I \text{ a } j \in J,$$

kde X_{ij} určuje, zda j-tá zákaznická zóna je obsluhována i-tým distribučním centrem; Y_i určuje, zda je otevřeno i-té distribuční centrum; C_{ij} vyjadřuje veškeré relevantní náklady na uspokojení poptávky i-té zákaznické zóny j-tým distribuční centrem; k je počet povolených distribučních center; I je indexová množina pro distribuční centra a J je indexová množina pro zákaznické zóny. Výsledkem bylo uzavření 12 areálů (tj. 20 procent) a odepsání asi miliardy dolarů za majetek a „lidské zdroje“ zaměstnanců. Na druhou stranu tyto změny přinesly roční úsporu ve výrobních a distribučních nákladech ve výši 250 milionů dolarů (před zdaněním). Pro více podrobností viz [89].

Analogické tvrzení jako ve Větě 3.3.5 platí i pro polyedry s omezeními ve tvaru nerovností, což pochopitelně zahrnuje i duální úlohu[85].

[85] Připomeňme, že pro řešitelné úlohy platí

$$\min\{x^\top c \mid Ax = b \ \&\ x \ge 0\} = \\ = \max\{y^\top b \mid A^\top y \le c\}.$$

DŮSLEDEK 3.3.9

Nechť matice $A \in \mathbb{R}^{m\times n}$ je totálně unimodulární. Pak polyedry $\{x \in \mathbb{R}^n \mid Ax \le b \,\&\, x \ge 0\}$, $\{x \in \mathbb{R}^n \mid Ax \ge b \,\&\, x \ge 0\}$ a $\{x \in \mathbb{R}^n \mid Ax \le b\}$ jsou pro jakékoli $b \in \mathbb{Z}^m$ celočíselné.

Důkaz. Tvrzení vyplývá ihned z Věty 3.3.5, protože první dva dané polyedry můžeme ekvivalentně vyjádřit ve tvaru $\{(x,s)^\top \in \mathbb{R}^{n+m} \mid Ax \pm Is = b \,\&\, x \ge 0\}$ s totálně unimodulárními maticemi $(A, \pm I)$ podle Lemma 3.3.2. Odtud získáme i třetí část, neboť odpovídající polyedr můžeme vyjádřit jako $\{(x^+,x^-)^\top \in \mathbb{R}^{2n} \mid (A,-A)(x^+,x^-)^\top \le b \ \ \&\ \ (x^+,x^-)^\top \ge 0\}$ a $x = x^+ - x^-$, viz také Příklad 2.2.2. ■

Hoffman a Kruskal v dříve zmíněném článku [270] ukázali, že předchozí vlastnost víceméně zcela charakterizuje totálně unimodulární matice, tj. pro celočíselnost polyedru $\{x \in \mathbb{R}^n \mid Ax \le b \,\&\, x \ge 0\}$ z předchozího Důsledku 3.3.9 je totální unimodularita dokonce nutná. My si tento výsledek také ukážeme, avšak nevydáme se stejnou cestou jako Hoffman a Kruskal, neboť jejich původní důkaz je poměrně komplikovaný. Místo toho využijeme tvrzení od Dantziga a Veinotta opírající se o trochu slabší požadavek pro matici A, který je nutný a dostačující pro celočíselnost polyderu X z Věty 3.3.5, viz [488].

DEFINICE 3.3.10

Celočíselnou matici $A \in \mathbb{R}^{m\times n}$ nazveme *unimodulární*, jestliže pro každou $m \times m$ submatici A_B platí $\det A_B \in \{0, \pm 1\}$.

Takové matice byly studovány již ve druhé polovině 19. století, viz [192, 193, 454] a také [496]. V případě regulární matice A můžeme podobně jako v předchozí části ukázat, že matice A je unimodulární právě tehdy, když A^{-1} je také unimodulární. V takovém případě platí též, že sloupce unimodulární matice generují celé $\mathbb{Z}^n$ a naopak, tj. n-tice lineárně nezávislých vektorů generující celé $\mathbb{Z}^n$ tvoří unimodulární matici. Navíc lze ukázat, že každou takovou matici A lze elementárními sloupcovými úpravami[86] převést na jednotkovou matici a naopak, tj. elementárními sloupcovými úpravami jednotkové matice vždy získáme unimodulární matici. Pochopitelně každá totálně unimodulární matice je současně i unimodulární. Navíc se lze snadno přesvědčit, že matice A je totálně unimodulární právě tehdy, když (I, A) je unimodulární[87].

[86] Jde o vynásobení sloupce -1, výměnu sloupců a přičtení celočíselného násobku nějakého sloupce k jinému sloupci.

[87] Implikace „⇒“ plyne ihned z definice. Opačnou implikaci „⇐“ získáme např. pomocí Laplaceova rozvoje, ve kterém $m \times m$ submatice vytvoříme vhodnou kombinací sloupců jednotkové matice a matice A.

Věta 3.3.11 Dantzig & Veinott

Nechť $A \in \mathbb{R}^{m\times n}$ je celočíselná matice s plnou hodností, tj. $\operatorname{rank} A = m$. Pak A je unimodulární právě tehdy, když polyedr $X := \{x \in \mathbb{R}^n \mid Ax = b \;\&\; x \geq 0\}$ je celočíselný pro libovolný celočíselný vektor $b \in \mathbb{R}^m$.

Důkaz. „⇒“ Nechť A je unimodulární, vektor b je celočíselný a x^* libovolný vrchol polyedru X. Jelikož jde o BPB, musí se v tomto bodě n-tice omezení (včetně omezení na znaménko) realizovat jako rovnosti a sloupce matice A odpovídající nenulovým složkám bodu x^* jsou lineárně nezávislé. Tyto sloupce můžeme doplnit na (m-prvkovou) bázi B^* tak, že $A_{B^*}^{-1} b = x_{B^*}$. Poněvadž A je unimodulární a A_{B^*} je regulární $m \times m$ submatice, musí být matice $A_{B^*}^{-1}$ celočíselná, a tudíž x_{B^*} má pouze celočíselné souřadnice, takže po doplnění nulovými hodnotami na nebazických pozicích obdržíme celočíselné x^*.

„⇐“ Naopak, připusťme, že polyedr X je celočíselný pro libovolný celočíselný vektor $b \in \mathbb{R}^m$. Musíme ukázat, že pro libovolnou $m \times m$ submatici A_B vytvořenou vynecháním $n-m$ sloupců je $\det A_B \in \{0, \pm 1\}$. Je-li A_B singulární, pak to je nepochybně splněno. Omezme se proto pouze na situaci, kdy matice A_B je regulární. Pak ale stačí pouze ukázat, že $A_B^{-1} v$ je celočíselný vektor pro libovolný celočíselný vektor $v \in \mathbb{R}^m$. Proč? V takovém případě totiž matice A_B^{-1} nutně musí být celočíselná, a tudíž $\det A_B^{-1} \in \mathbb{Z}$. Ovšem $\det A_B$ je také celočíselný a současně platí rovnost $\det A_B^{-1} = (\det A_B)^{-1}$, takže nutně $\det A_B = \pm 1$. Nechť tedy $v \in \mathbb{R}^m$ je libovolný celočíselný vektor. Pak jistě existuje celočíselný vektor $u \in \mathbb{R}^m$ takový, že platí nerovnost

$$z_B := u + A_B^{-1} v \geq 0.$$

Označme $b := A_B z_B = A_B u + v$, což je nepochybně také celočíselný vektor jakožto součet celočíselných vektorů. Doplňme vektor z_B nulovými složkami na pozice odpovídající vynechaným sloupcům z matice A. Tím získáme celočíselný vektor $z \in \mathbb{R}^n$, pro který platí $Az = A_B z_B = b$. To znamená, že z je vrchol polyedru X (neboli BPB, neboť vyhovuje dané soustavě a zvolená m-tice slupců v A_B splňuje požadavky báze bodu z), takže dle současných předpokladů musí mít celočíselné souřadnice, z čehož vyplývá i celočíselnost vektoru $A_B^{-1} v = z_B - u$. Tím je důkaz opačné implikace hotov. ■

Důsledek 3.3.12 Hoffman & Kruskal

Nechť $A \in \mathbb{R}^{m\times n}$ je celočíselná matice. Pak A je totálně unimodulární právě tehdy, když polyedr $\{x \in \mathbb{R}^n \mid Ax \leq b \;\&\; x \geq 0\}$ je celočíselný pro libovolný celočíselný vektor $b \in \mathbb{R}^m$.

Důkaz. Protože matice A je totálně unimodulární právě tehdy, když (I, A) je unimodulární, plyne tento důsledek z předchozího tvrzení a ekvivalence mezi polyedry $\{x \in \mathbb{R}^n \mid Ax \leq b \;\&\; x \geq 0\}$ a $\{z \in \mathbb{R}^{m+n} \mid (I, A)z = b \;\&\; z \geq 0\}$, která je určena rovností $z = (w, x)^\top$. ■

Jelikož pro libovolnou totálně unimodulární matici A je totálně unimodulární také matice

$$\begin{pmatrix} I \\ -I \\ A \\ -A \end{pmatrix},$$

plyne odtud a z předchozího důsledku, že libovolná celočíselná matice A je totálně unimodulární právě tehdy, když polyder $\{x \in \mathbb{R}^n \mid a \leq Ax \leq b \;\&\; c \leq x \leq d\}$ je celočíselný pro libovolné celočíselné vektory $a, b \in \mathbb{R}^m$ a $c, d \in \mathbb{R}^n$. Podobně můžeme z předchozí důsledku vyvodit to, že celočíselná matice A je totálně unimodulární právě tehdy, když některý z polyedrů

$$\{x \in \mathbb{R}^n \mid Ax \leq b \;\&\; x \leq c\}, \qquad \{x \in \mathbb{R}^n \mid Ax \geq b \;\&\; x \leq c\},$$
$$\{x \in \mathbb{R}^n \mid Ax \leq b \;\&\; x \geq c\}, \qquad \{x \in \mathbb{R}^n \mid Ax \geq b \;\&\; x \geq c\}$$

má všechny vrcholy celočíselné pro všechny celočíselné vektory $b \in \mathbb{R}^m$ a $c \in \mathbb{R}^n$.

Důsledek 3.3.12 lze také následujícím způsobem formulovat v jazyce lineárního programování.

DŮSLEDEK 3.3.13

Nechť $A \in \mathbb{R}^{m \times n}$ je celočíselná matice. Pak A je totálně unimodulární právě tehdy, když pro všechny celočíselné vektory $b \in \mathbb{R}^m$ a $c \in \mathbb{R}^n$ obě strany duální rovnosti

$$\min\{x^\top c \mid Ax \geq b \;\&\; x \geq 0\} = \max\{y^\top b \mid A^\top y \leq c \;\&\; y \geq 0\}$$

jsou dosaženy celočíselnými vektory x a y (pokud existují).

3.4 Metoda větvení a mezí[91]

[91] Angl. *Branch-and-Bound Method*

Pokud ale na rozdíl od předchozí podkapitoly není přípustná množina v úloze celočíselného programování vymezena pomocí (totálně) unimodulární matice, nelze a priori očekávat, že řešením přidružené úlohy lineárního programování získáme hledané celočíselné řešení. Jak tedy najít řešení v takovém případě? Kdyby byla přípustná množina ohraničená, mohli bychom vyzkoušet tzv. *úplný výčet*[92], tj. určit všechny přípustné celočíselné body a porovnáním jednotlivých funkčních hodnot najít jejich minimum. Avšak jakkoli tato myšlenka zní jednoduše, je poměrně naivní a pro větší úlohy je z praktického hlediska téměř nerealizovatelná, neboť vyžaduje velké množství početních operací — např. pro úlohu s n-ticí celočíselných proměnných, z nichž každá může nabývat m celočíselných hodnot, by se jednalo o m^n bodů, u kterých bychom museli nejdříve otestovat přípustnost a následně stanovit funkční hodnotu.

[92] Angl. *complete enumeration*

Jenže simplexový algoritmus se ukázal (navzdory komentářům z Podkapitoly 2.13) být vskutku velmi efektivním nástrojem, a tak by bylo škoda jej nevyužít i při řešení obecných úloh celočíselného programování, které patří mezi ty nejsložitější problémy. Jak ho ale do toho zakomponovat? Třeba tak, že nějakou vhodnou „úpravou“ přípustné množiny se nám podaří udělat z hledaného celočíselného řešení BPB, tj. daný polyedr upravit tak, aby hledané řešení bylo jeho vrcholem. Toho můžeme docílit buď rozdělením přípustné množiny na menší (v extrémním případě možná až na hyperkrychle $1 \times 1 \times \cdots \times 1$) nebo jejím ořezáním, přičemž je ale velmi důležité zmínit, že ani při jedné z těchto úprav nesmíme přijít o žádný přípustný bod s celočíselnými souřadnicemi.

První zmíněná možnost s rozdělování přípustné množiny patří do tzv. skupiny *Divide-and-Conquer*[93]. Postupným rozdělováním přípustné množiny získáme několik „menších“ a „jednodušších“ množin, na kterých jsou následně samostatně řešeny příslušné minimalizační úlohy (obvykle jako úlohy lineárního programování s využitím simplexového algoritmu[94]). Toto je založeno na jednoduchém pozorování, že v případě přípustné množiny

$$X = X^{[1]} \cup \cdots \cup X^{[m]}$$

a příslušných hodnot

$$x^{[k]} := \min\{c^\top x \mid x \in X^{[k]}\}$$

pro $k = 1, \ldots, m$ lze najít minimum na X jako

$$x^* = \min\{x^{[k]} \mid k = 1, \ldots, m\}.$$

[93] To je anglický překlad slavného latinského výroku *Divide et Impera* (*Rozděl a panuj*).

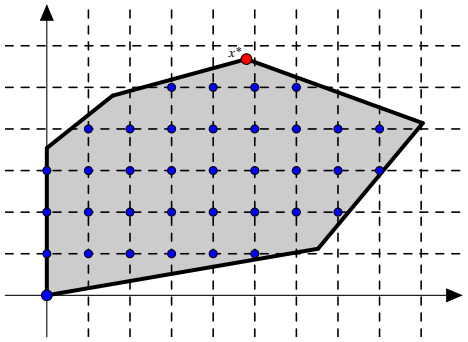
[94] Jedná se o tzv. „rozvolnění“ (příp. relaxaci) úlohy, které může mít mnoho podob. V tomto případě jde o ignorování požadavku celočíselnosti/binárnosti, ale v jiných úlohách můžeme např. rozvolnit či oslabit nějaké jiné omezení (ať již algebraické nebo nealgebraické) nebo nějaká nerovnostní omezení nahradit jejich nezápornou kombinací či můžeme vhodným způsobem změnit účelovou funkci tak, abychom získali nějakou dolní hranici pro hledané minimum, tj. místo funkce f řešit daný problém s účelovou funkcí g, pro kterou na přípustné množině platí $g(x) \leq f(x)$. Každopádně všechny možnosti rozvolnění dané úlohy splňují tři podmínky: přípustná množina rozvolněné úlohy obsahuje všechny přípustné body původní úlohy, hodnota rozvolněné úlohy poskytuje dolní mez pro hodnotu původní úlohy a v některých případech je řešení rozvolněné úlohy shodné s řešením původní úlohy. Pro další možnosti relaxace viz např. [294, Section 24.2.1]

Samotná realizace tohoto principu může být samozřejmě provedena různými způsoby a v literatuře lze najít různé varianty tohoto algoritmu. Jedno (vlastně dvojí) však mají společné. Průběh algoritmu připomíná rozvětvující[95] se strom (tzv. *výčtový strom*) a na koncích jednotlivých větví jsou testovány dolní a horní meze[96] účelové funkce oproti jiným větvím (a tedy potenciálním kandidátům na řešení). Postupem času tak dochází k tomu, že další rozvětvování již není možné (ať už kvůli odpovídajícím funkčním hodnotám nebo omezením přípustné množiny), díky čemuž nalezneme řešení (jistě v konečném čase). Tento algoritmus se poprvé objevuje v článku [346] z roku 1960 od Land(-ové)[97] a Doig(-ové)[98], který byl součástí výzkumu financovaného společností British Petroleum na London School of Economics. My si ukážeme tuto metodu v modifikaci, kterou v roce 1965 představil Dakin[99], viz [109] a pro další modifikace také [37, 74, 75, 145, 213, 244, 349, 484]. Tato metoda se stala nejpoužívanějším nástrojem pro řešení tzv. **NP**-těžkých úloh (viz závěrečnou Podkapitolu 3.6) a je používána i např. programem Maple[100] . Samotné označení „Branch-and-Bound Method“ pochází z článku [356] věnovaného řešení problému obchodního cestujícího.

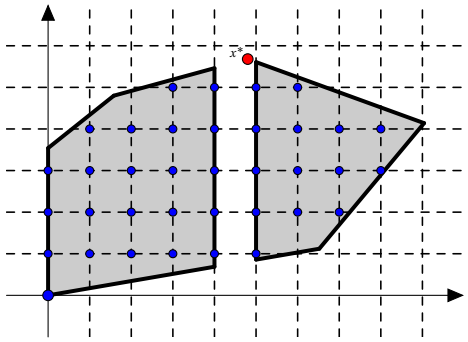
Na začátku máme úlohu (smíšeného) celočíselného programování s celočíselnými koeficienty (případně mohou být i racionální, což ale snadno převedeme na celočíselný případ). Řešením této úlohy bez požadavku celočíselnosti nalezneme bod $x^{[0]}$, jehož odpovídající funkční hodnota $f(x^{[0]})$ udává dolní mez pro hodnotu hledaného minima f^*. Vzhledem k celočíselným koeficientům dokonce nutně platí $f^* \geq \lceil f(x^{[0]}) \rceil$. Splňuje-li $x^{[0]}$ podmínky celočíselnosti, máme řešení. V opačném případě si zvolíme (libovolně) proměnnou $x_i^{[0]} \notin \mathbb{Z}$ a rozdělíme původní přípustnou množinu X na dvě části[101]

$$X^{[1]} := \{x \in X \mid x_i \leq \lfloor x_i^{[0]} \rfloor\} \quad \text{a} \quad X^{[2]} := \{x \in X \mid x_i \geq \lceil x_i^{[0]} \rceil\},$$

kde $\lfloor \cdot \rfloor$ značí *dolní celou část* a $\lceil \cdot \rceil$ *horní celou část*. Tímto jsme z přípustné množiny $\overline{X}$ přidružené úlohy lineárního programování odstranili všechny body splňující $\lfloor x^{[0]} \rfloor < x_i < \lfloor x^{[0]} \rfloor + 1$, takže jistě i $x^{[0]}$, viz Obrázky 3.19–3.20.



Obrázek 3.19: Přípustná množina X (modré body) úlohy celočíselného programování a řešení x^* přidružené úlohy lineárního programování (červený bod, přípustná množina $\overline{X}$ je šedý polyedr).



Obrázek 3.20: Rozdělení přípustné množiny X (modré body) úlohy celočíselného programování na dvě části $X^{[1]}$ a $X^{[2]}$ tak, aby přípustné množiny $\overline{X}^{[1]}$ a $\overline{X}^{[2]}$ přidružených úloh lineárního programování (šedé polyedry) obsahovaly všechny celočíselné body a současně neobsahovaly bod x^*.

Nyní tento proces postupně zopakujeme pro úlohy celočíselného programování na množinách $X^{[1]}$ a $X^{[2]}$, přičemž vše postupně zapisujeme do výčtového stromu. Pochopitelně nemusí být nutné dělit množinu X až na jednotkové hyperkrychle s celočíselnými vrcholy, protože v dalším větvení aktuálního uzlu můžeme přestat, pokud jsme již našli bod splňující všechny podmínky celočíselnosti (to je *kandidát* na hledané řešení, tj. strom je „ořízнut“ z důvodu optimality[102]) nebo pokud současná přípustná množina je *prázdná* (dělením jsme vlastně dostali mimo původní množinu X, tj. strom je „ořízнut“ z důvodu nepřípustnosti[103]) nebo pokud nalezené minimum přidružené úlohy lineárního programování na této množině je větší než hodnota účelové funkce v nějakém bodě splňujícím podmínky celočíselnosti, který jsme nalezli v jiné větvi (tj. v této podmnožině *neexistuje* žádný bod splňující podmínky celočíselnosti, jehož funkční hodnota je menší nebo rovna hodnotě v již dříve nalezeném kandidátovi

[95] *větev=branch*

[96] *omezení=bound*

[97]

Ailsa H. Land (za svobodna Dicken; ∗ 14. června 1927) je britskou matematičkou a emeritní profesorkou operačního výzkumu na London School of Economics. V roce 1994 byla oceněna *Harold Larnder Prize* od *Kanadské společnosti pro operační výzkum* za mezinárodní přesah jejích výsledků v operačním výzkumu. Na London School of Economics je také na její počest udělována *Ailsa Land Prize* nejlepšímu studentovi/studentce v magisterském studijním oboru Operačního výzkumu a analytiky (součástí ocenění je odměna 250 £ a kniha dle výběru Ústavu matematiky.

[98]

Alison Grant Doig (za svobodna Harcourt; ∗ 24. listopadu 1929) je australská matematička a statistička. Byla součástí týmu, který vyvinul *práh chudoby* jako součást Hendersonova zjišťování chudoby v Austrálii v roce 1973 a pomohla zavést metodu dvojitého náhodného výběru uspořádání kandidátů v australských volbách.

[99] R. J. Dakin. Zdá se, že o něm žádné další bližší informace nejsou dostupné. V matematické databázi MathSciNet má pouze dva články. Ve zmíněném článku z roku 1965 je v afiliaci uvedena univerzita v Sydney. V článku z roku 1970, kde se věnuje využití celočíselného programování k řešení struktury krystalů, je v afiliaci uvedena univerzita v Papui Nové Guineji.

[100] Avšak může se stát, že ani Maple hledané řešení nenajde. Pro-

gram má totiž nastavenu výchozí hloubku procházení výčtového stromu, tj. maximální počet povolených kroků, které však nemusí k nalezení řešení stačit.

[101] Toto je ta zmíněná modifikace od Dikina. V původní verzi od Doig(-ové) a Land(-ové) bychom vytvořili $2T+1$ částí, kde T je takové číslo, že pro všechny proměnné platí $-T \le x_i \le T$, k čemuž je nutný a priorní předpoklad ohraničenosti původní přípustné množiny.

[102] Angl. *pruned by optimality*

[103] Angl. *pruned by infeasibility*

[104] Angl. *pruned by bound*

na řešení, tj. strom je „oříznut“ z důvodu ohraničenosti[104]). Každá větev je proto buď ukončena jako „slepá“, neboť jejím dalším rozvíjením nic podstatného nezískáme, nebo je na jejím konci nový kandidát na řešení, který dává lepší (příp. stejnou) funkční hodnotu než ten dosavadní (je-li už nějaký znám). Tento postup si ukážeme v následujících dvou příkladech.

Příklad 3.4.1

Metodou větvení a mezí vyřešme úlohu celočíselného programování

$$-5x_1 - 8x_2 \to \min$$

za podmínek

$$x_1 + x_2 \le 6 \quad \& \quad 5x_1 + 9x_2 \le 45$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Video ☞ 😎 https://ln.math.muni.cz/pz/opt/zlp1.

Příklad 3.4.2

Metodou větvení a mezí vyřešme úlohu smíšeného celočíselného programování

$$3x_1 + 2x_2 - 10 \to \min$$
$$x_1 - 2x_2 + x_3 = 5/2 \quad \& \quad 2x_1 + x_2 + x_4 = 3/2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_4 \ge 0 \quad \& \quad x_2 \in \mathbb{Z} \quad \& \quad x_3 \in \mathbb{Z}.$$

Video ☞ 😎 https://ln.math.muni.cz/pz/opt/zlp2.

Avšak je možné tento postup využít vždy? Ano, pokud se nám podaří začít. Velké úskalí celého procesu totiž může tkvět již v prvním kroku, ve kterém se nám nemusí povést najít řešení $x^{[0]}$ přidružené úlohy lineárního programování. Pokud se to stalo kvůli nepřípustnosti této úlohy, pak je jasné, že i původní úloha je nepřípustná. Pokud to ale nastalo kvůli neohraničenosti přidružené úlohy, pak původní úloha musí být buď nepřípustná nebo neohraničená jako např. v úlohách

$$x_1 + x_2 \to \min \quad \& \quad x_1 = x_2 + 1/2 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}$$
$$x_1 + x_2 \to \min \quad \& \quad x_1 = x_2 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Jiná možnost pro úlohy s celočíselnými (eventuálně racionálními) koeficienty nastat nemůže, protože se dá v takovém případě ukázat, že není možné, aby původní úloha byla řešitelná a přidružená úloha neohraničená, viz [110]. To lze pouze pro úlohy s iracionálními koeficienty jako např. v úloze

$$x_1 + x_2 \to \min \quad \& \quad x_1 = \sqrt{2}\,x_2 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z},$$

ve které je $X = \{[0,0]\}$ a bod $x^* = [0,0]$ je jejím řešením, zatímco přidružená úloha lineárního programování je neohraničená. Avšak ani řešitelnost přidružených úloh lineárního programování nemusí být zárukou, že se nám metodou větvení a mezí podaří nalézt řešení úlohy celočíselného programování. Např. v úloze

$$x_1 + x_2 \to \min \quad \& \quad 1/4 \le x_1 - x_2 \le 3/4$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}$$

budou přidružené úlohy postupně generovat body $x^{[2k]} = [k+1/4, k]$ a $x^{[2k+1]} = [k+1, k+1/4]$ pro $k = 0, 1, \ldots$, aniž bychom zjistili, že původní úloha je nepřípustná. Všechny zmíněné „nepříjemnosti“ ale mohou nastat pouze v případě neohraničené množiny X. Proto v případě ohraničené množiny X zjevně platí následující tvrzení.

VĚTA 3.4.3

Pro úlohu (smíšeného) celočíselného programování s ohraničenou přípustnou množinou je metoda větvení a mezí konečně kroková.

Pro lepší přehlednost jsou možné výsledky úlohy celočíselného programování a přidružené úlohy shrnuty v následující tabulce.

ℤLP / LP	NP	PaO	NO
NP	✔	✖	✖
PaO	✔	✔	✖
NO	✔	✖	✔

Tabulka 3.6: Vztah mezi úlohou celočíselného programování (ℤLP) a přidruženou úlohou lineárního programování (LP). NP: nepřípustná úloha; NO: neohraničená úloha; PaO: přípustná a ohraničená úloha (tj. existuje řešení).

Ještě bychom měli upozornit, že časová složitost metody větvení a mezí není polynomiálně závislá na velikosti vstupních dat, neboť např. pro libovolné $k = 1, 2, \ldots$ mají úlohy

$$-x_1 \to \min \quad \& \quad (2^k + 1)\, x_1 = 2^k\, x_2 \quad \& \quad 0 \le x_2 \le 2^k \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}$$

jediný přípustný bod $[0,0]$, který metodou větvení a mezi nalezne nejdříve po 2^k krocích. V článku [296] bylo také ukázáno, že metoda větvení a mezí potřebuje pro liché $n \in \mathbb{N}$ nejméně $2^{(n+1)/2}$ kroků k tomu, aby ukázala, že binární úloha celočíselného programování

$$-x_1 \to \min \quad \& \quad 2x_1 + 2x_2 + \cdots + 2x_n = n \quad \& \quad x_1, \ldots, x_n \in \{0, 1\}$$

je nepřípustná, viz i [291]. Každopádně jakkoli jednoduše popis základních kroků metody větvení a mezí vypadá, její algoritmický popis je ve srovnání s předchozími metodami poněkud komplikovanější.

Metoda větvení a mezí pro úlohu celočíselného programování

(1) *Inicializace*: je dána úloha (smíšeného) celočíselného programování s celočíselnými koeficienty. Je vyřešena přidružená úloha lineárního programování. Pokud

(i) je tato úloha nepřípustná, pak totéž platí pro původní úlohu. KONEC.

(ii) je tato úloha neohraničená, pak původní úloha je buď také neohraničená nebo nepřípustná. KONEC.

(iii) nalezené řešení splňuje podmínky celočíselnosti, jedná se i o řešení původní úlohy. KONEC.

(iv) nalezené řešení $x^{[0]}$ nesplňuje podmínky celočíselnosti, rozdělíme pomocí nějaké proměnné x_i s hodnotou $x_i^{[0]}$ nesplňující podmínku celočíselnosti a odpovídajících dodatečných podmínek

$$x_i \le \lfloor x_i^{[0]} \rfloor \quad \text{a} \quad x_i \ge \lceil x_i^{[0]} \rceil$$

původní přípustnou množinu na dvě části $X^{[1]}$ a $X^{[2]}$. Následně definujeme „seznam“ $\mathcal{S} := \{X^{[1]}, X^{[2]}\}$.

(2) *Volba uzlu*: je-li $\mathcal{S} = \emptyset$, již jsme nalezli řešení (pokud jsme již nějaký bod prohlásili za kandidáta na řešení $\widehat{x}$ dané úlohy) nebo je úloha nepřípustná. KONEC. V opačném případě zvolíme množinu $X^{[m]} \in \mathcal{S}$.

(3) *Relaxace*: vyřešíme přidruženou úlohu lineárního programování na množině $X^{[m]}$. Je-li tato úloha nepřípustná, pokračujeme Krokem (2) s novým indexem m. V opačném případě označíme nalezené řešení jako $x^{[m]}$.

(4) *Ořezávání a zlepšování mezí*: Jestliže nalezené řešení $x^{[m]}$

(i) splňuje podmínky celočíselnosti a není-li zatím žádný bod označen za kandidáta na řešení $\widehat{x}$, pak definujeme $\widehat{x} := x^{[m]}$ a pokračujeme Krokem (5).

(ii) splňuje podmínky celočíselnosti a je-li již nějaký bod označen za kandidáta na řešení $\widehat{x}$, pak v případě $f(x^{[m]}) \geq f(\widehat{x})$ pokračujeme Krokem (5).

(iii) splňuje podmínky celočíselnosti a je-li již nějaký bod označen za kandidáta na řešení $\widehat{x}$ původní úlohy, pak v případě $f(x^{[m]}) < f(\widehat{x})$ předefinujeme $\widehat{x} := x^{[m]}$ a pokračujeme Krokem (5).

(iv) nesplňuje podmínky celočíselnosti a současně $\lceil f(x^{[m]}) \rceil \geq f(\widehat{x})$ pokračujeme Krokem (5).

(v) nesplňuje podmínky celočíselnosti a současně $\lceil f(x^{[m]}) \rceil < f(\widehat{x})$ pokračujeme Krokem (6).

(5) *Aktualizace*: Odstraníme množinu $X^{[m]}$ z $\mathcal{S}$, tj. aktualizujeme seznam $\mathcal{S} \to \mathcal{S} \setminus \{X^{[m]}\}$, a pokračujeme Krokem (2) s novým indexem m.

(6) *Větvení*: Zvolíme libovolnou proměnnou x_j, pro kterou hodnota $x_j^{[m]}$ nesplňuje podmínku celočíselnosti. Podmínkami

$$x_j \leq \lfloor x_j^{[m]} \rfloor \quad \text{a} \quad x_j \geq \lceil x_j^{[m]} \rceil$$

rozdělíme množinu $X^{[m]}$ na dvě části $X^{[m_1]}$ a $X^{[m_2]}$, které vložíme do seznamu $\mathcal{S}$ místo $X^{[m]}$, tj. $\mathcal{S} \to \left(\mathcal{S} \cup \{X^{[m_1]}, X^{[m_1]}\}\right) \setminus \{X^{[m]}\}$, a pokračujeme Krokem (2) s novým indexem m.

Výhodou této metody je, že u některých uzlů nebude potřeba další větvení kvůli nepřípustnosti nebo horší mezi[105] a že si nemusíme pamatovat všechny výpočty – stačí mít v paměti pouze seznam $\mathcal{S}$ a aktuálního kandidáta $\widehat{x}$. Předchozí algoritmus (a vlastně i samotná metoda větvení a mezí) ale nebude fungovat, pokud neupřesníme, jak volit množinu $X^{[m]}$ ze seznamu $\mathcal{S}$ a jak volit proměnnou x_j pro další větvení. Navíc vhodnou volbou těchto pravidel můžeme celý výpočet urychlit a/nebo snížit výpočetní náročnost. Pro výběr ze seznamu $\mathcal{S}$ lze vyžít princip *FIFO*[106] (volíme množinu, kterou jsme do seznamu přidali nejdříve, tj. rozvíjíme výčtový strom do šířky) nebo *LIFO*[107] (volíme množinu, kterou jsme do seznamu přidali nejpozději, tj. rozvíjíme výčtový strom do hloubky a po zakončení nějaké větve se toutéž cestou vracíme zpět[108]) nebo volba množiny s nejmenší mezí (tj. vybereme tu, pro kterou je dolní odhad $\lceil f(x^{[m]}) \rceil$ hledaného minima nejmenší) nebo pomocí míry „zlomkovosti“ (neceločíselnosti) řešení přidružených úloh lineárního programování s případným přihlédnutím k aktuální hloubce výčtového stromu. Obzvláště strategie LIFO se zdá být v jistých aspektech výhodná, neboť nerozvíjí výčtový strom příliš do šířky, takže seznam $\mathcal{S}$ není zbytečně rozsáhlý (počet členů se zvýší nejvýše o 1, protože odebereme jednu množinu a přidáme dvě nebo žádnou) a s využitím duální simplexové metody lze po přidání nového omezení z větvení navázat na poslední simplexovou

[105] Tomuto se anglicky dohromady říká *fathoming*.

[106] Angl. *first in, first out*.

[107] Angl. *last in, first out*.

[108] Jedná se o tzv. *backtracking*.

tabulku. Pro výběr rozvětvující proměnné se volí např. proměnná se zlomkovou (desetinnou) částí nejblíže 1/2 nebo proměnná s nejmenším koeficientem v účelové funkci nebo proměnná s nejmenším indexem. V některých specifických situacích je možné před zahájením metody větvení a mezí využít některá omezení a požadavek celočíselnosti k získání silnějších dodatečných podmínek (občas to připomíná naše úvahy o duální úloze v lineárním programování). Např. z podmínek

$$5x_1 - 2x_2 + 8x_3 \le 15 \quad \& \quad 0 \le x_1 \le 3 \quad \& \quad 0 \le x_2 \le 1 \quad \& \quad x_3 \ge 1$$

získáme

$$5x_1 \le 15 + 2x_2 - 8x_3 \le 15 + 2 \cdot 1 - 8 \cdot 1 = 9,$$

a tedy při podmínce celočíselnosti pro x_1 musí být $x_1 \le 1$ (bez celočíselnosti to bude $x_1 \le 9/5$). Při výpočtu je také možné zafixovat hodnotu některé z celočíselných proměnných a hledat vhodné hodnoty ostatních proměnných. Nicméně je nutné říci, že každé z těchto pravidel má své výhody i nevýhody a nelze některé z nich označit jako „to nejlepší“, viz také [153, Chapter 6], [395, Kapitola II.4] a [505, Kapitola 4].

Metodu větvení a mezí lze samozřejmě využít i pro binární úlohu celočíselného programování. Poměrně překvapivá ale může být „nutnost“ jejího využití i pro úlohu s jedním omezením jako v případě problému batohu popsaném v Příkladě 3.2.3, tj.

$$\sum_{j=1}^{n} c_j x_j \to \max \quad \& \quad \sum_{j=1}^{n} a_j x_j \le b \quad \& \quad x_1, \dots, x_n \in \{0,1\}. \tag{3.4.1}$$

Zde by se totiž mohlo zdát dostačujícím uspořádat proměnné tak, aby platily nerovnosti

$$\frac{c_1}{a_1} \ge \frac{c_2}{a_2} \ge \cdots \ge \frac{c_n}{a_n}, \tag{3.4.2}$$

a pak postupně přidávat jednotlivé položky (počínaje x_1 atd.), dokud je ještě dostupná nějaká kapacita. Toto nepochybně funguje pro přidruženou úlohu lineárního programování, čehož si všiml již Dantzig v článku [115] z roku 1957. K jejímu řešení nepotřebujeme simplexový algoritmus nebo nějakou jinou metodu, ale stačí postupovat v souladu s následujícím tvrzením.

VĚTA 3.4.4

Nechť je dán takový problém batohu (3.4.1), že jednotlivé koeficienty splňují nerovnosti (3.4.2). Pak existuje index $k \in \{1, \dots, n+1\}$ takový, že optimálním řešením $x^* = [x_1^*, \dots, x_n^*]$ přidružené úlohy lineárního programování je[109]

$$x_1^* = \cdots = x_{k-1}^* = 1 \quad \& \quad x_k^* = \frac{b - \sum_{j=1}^{k-1} a_j}{a_k} \quad \& \quad x_{k+1}^* = \cdots = x_n^* = 0.$$

[109] Pochopitelně v případě $k = 1$ klademe $\sum_{j=1}^{0} \dots := 0$. Na druhou stranu volba $k = n+1$ znamená, že $x_1 = \cdots = x_n = 1$, tj. že do batohu se vejdou úplně všechny dostupné věci.

Důkaz. Zatímco Dantzig ve svém článku nabízí grafický důkaz, my si tvrzení dokážeme více formálně. Je jasné, že hledané řešení přidružené úlohy musí zcela vyčerpat dostupnou kapacitu, tj. $\sum_{j=1}^{n} a_j x_j^* = b$. Bez újmy na obecnosti předpokládejme, že všechny nerovnosti v (3.4.2) jsou ostré. Připusťme, že $\widehat{x}$ je řešením přidružené úlohy a platí $\widehat{x}_j < 1$ pro nějaké $j \in \{1, \dots, k-1\}$. Pak musí být $\widehat{x}_\ell > x_\ell^*$ pro alespoň jeden index $\ell \in \{k, \dots, n\}$. Vezmeme-li libovolně malé $\varepsilon > 0$, pak zvýšením hodnoty $\widehat{x}_j$ o ε a zmenšením x_ℓ o $\varepsilon\, a_j / a_\ell$, zůstane $\sum_{j=1}^{n} a_j \widehat{x}_j = b$, ale dojde ke zvýšení funkční hodnoty o $\varepsilon\,(c_j - c_\ell\, a_j / a_\ell)$, což je kladné díky podmínce[110] $\frac{c_j}{a_j} > \frac{c_\ell}{a_\ell}$. To je ale spor s tím, že $\widehat{x}$ je řešením uvažované úlohy. Podobně můžeme dospět ke sporu, kdyby $\widehat{x}_j > 0$ pro nějaké $j \in \{k+1, \dots, n\}$, čímž je důkaz hotov. ■

[110] Odtud vidíme, že v případě neostré nerovnosti (tj. rovnosti) bychom mohli věci zaměnit, aniž bychom změnili hodnotu účelové funkce.

Ovšem pozor, řešení původní celočíselné úlohy nemusí mít takovouto dichotomní strukturu! Věta 3.4.4 dává řešení přidružené úlohy, a tedy pouze jakýsi horní odhad hodnoty účelové funkce původní úlohy. Jenže řešení celočíselné úlohy může být zcela jiné. Uvažme např., že pro batoh s kapacitou 10 máme na výběr věc A s váhou 9 a cenou 90, pět věcí B s váhou 2 a cenou 19 a deset věcí C s váhou 1 a cenou 1. Věc A má nejlepší poměr mezi cenou a váhou. Když

ji dáme do batohu, tak budeme mít místo ještě na jeden kus C, čímž zcela vyčerpáme kapacitu a dosáhneme hodnoty 91. Jenže to není správné řešení – do batohu se totiž vejde i všech pět věcí B, které budou mít hodnotu 95! Využitím poměrů mezi cenou a váhou nalezneme řešení úlohy (3.4.1) pouze v případě, že tím zcela naplníme dostupnou kapacitu, tj. pouze pokud platí $\sum_{j=1}^{n} a_j x_j^* = b$ pro $x_1^* = \cdots = x_{k-1}^* = 1$ a $x_k^* = \cdots = x_n^* = 0$ pro nějaké $k \in \{1,\ldots,n\}$. A právě tohoto faktu lze využít při řešení úlohy (3.4.1) metodou větvení a mezí. Na začátku přeznačíme proměnné tak, aby platilo (3.4.2), přičemž je dobré mít na paměti, že je-li $a_j > b$ pro nějaké $j \in \{1,\ldots,n\}$, pak nutně $x_j^* = 0$. Poté nalezneme řešení $\widehat{x}$ přidružené úlohy podle Věty 3.4.4. V případě $\widehat{x}_k = 0$ již máme řešení původního problému. V opačném případě dojde k větvení určeném podmínkami $x_k = 0$ a $x_k = 1$ a analogickým postupem (s těmito zafixovanými hodnotami) pokračujeme dále, přičemž hodnotu 1 sice přiřazujeme doposud nezafixovaným proměnným v pořadí dle jejich indexů, ale tentokrát to musíme dělat tak, aby nedošlo k porušení daného omezení, tj. pracujeme s úlohou s menším počtem proměnným a jinou hodnotou pravé strany daného omezení (to platí ve větvi určené podmínkou $x_k = 1$), viz také [370, Kapitola 2].

Příklad 3.4.5

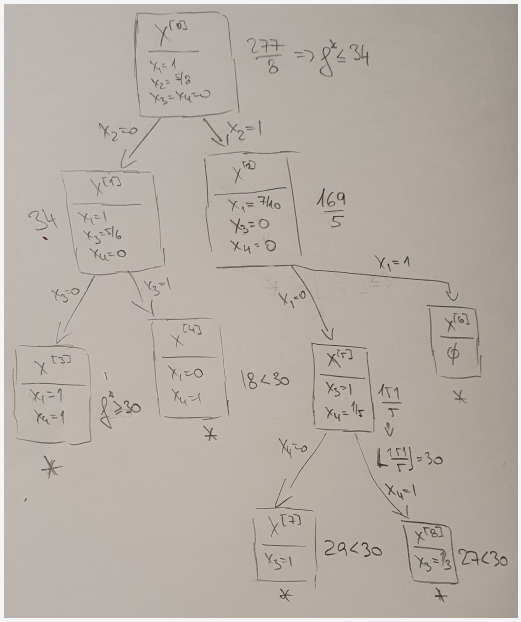
Metodou větvení a mezí vyřešme problém batohu

$$24\,x_1 + 17\,x_2 + 12\,x_3 + 6\,x_4 \to \max$$
$$10\,x_1 + 8\,x_2 + 6\,x_3 + 5\,x_4 \le 15$$
$$x_1 \in \{0,1\} \quad \& \quad x_2 \in \{0,1\} \quad \& \quad x_3 \in \{0,1\} \quad \& \quad x_4 \in \{0,1\}.$$

Řešení. Jednotlivé proměnné již jsou uspořádány v požadovaném pořadí, takže můžeme začít s prvním krokem. Pro přidruženou úlohu lineárního programování nalezneme řešení

$$x_1 = 1 \quad \& \quad x_2 = 5/8 \quad \& \quad x_3 = x_4 = 0$$

s hodnotou 277/8. Přípustnou množinu proto rozdělíme podmínkami $x_2 = 0$ a $x_2 = 1$. Je-li $x_2 = 0$, pak pro řešení přidružené úlohy máme $x_1 = 1$, $x_3 = 5/6$ a $x_4 = 0$ s hodnotou 34. Dalším větvením podle proměnné x_3 dostaneme v případě $x_3 = 0$ celočíselné řešení s $x_1 = x_4 = 1$ s hodnotou 30, takže tuto větev již dále rozvíjet nebudeme, zatímco ve větvi pro $x_3 = 1$ bude $x_1 = 0$ a $x_4 = 1$ s hodnotou 18, což je opět celočíselný bod, ale s hodnotou menší než v předchozí větvi. Vše je podrobně zaznamenáno ve výčtovém stromě na Obrázku 3.21.



Obrázek 3.21: Výčtový strom z Příkladu 3.4.5.

Zbývá větev pro $x_2 = 1$, kde pro přidruženou úlohu získáme řešení $x_1 = 7/10$, $x_3 = x_4 = 0$ s hodnotou $169/5$, která je větší než 30, takže musíme pokračovat v dalším větvení podle x_1. Je-li $x_1 = 0$, pak řešení přidružené úlohy je $x_3 = 1$ a $x_4 = 1/5$ s hodnotou $151/5$, pro kterou platí $\lfloor 151/5 \rfloor = 30$, takže bychom mohli větvení ukončit, neboť zde již lepší řešení než dříve nenajdeme. Ale podíváme se, zda náhodou neexistuje alternativní řešení. Volbou $x_4 = 0$ dostaneme celočíselné řešení s $x_3 = 1$ a hodnotou $29 < 30$ a pro $x_4 = 1$ bude $x_3 = 1/3$ s hodnotou $27 < 30$, takže ani v jednom případě nezískáme celočíselný bod s hodnotou 30. Vrátíme-li se ještě k větvi s $x_1 = 1$, zjistíme, že v takovém případě je přidružená úloha nepřípustná. Tím je celý výpočet u konce a řešením je bod $x^* = [1,0,0,1]$ s hodnotou $f^* = 30$.

Pro binární úlohu (ať již s jedním nebo více omezeními) lze ale metodu větvení a mezí aplikovat i trochu odlišným způsobem, který nevyžaduje řešení přidružených úloh lineárního programování. Tentokrát místo hledání celočíselného bodu mezi těmito řešeními budeme mít hned na začátku celočíselný bod a vhodnou změnou hodnot některých proměnných budeme chtít docílit jeho přípustnosti. Nejdříve upravíme proměnné tak, aby všechny měly v účelové funkci kladný koeficient (tj. v případě záporného koeficientu u x_j nahradíme tuto proměnnou za $x'_j := 1 - x_j$, která bude opět nabývat pouze binárních hodnot), a následně uspořádáme jednotlivé proměnné dle velikosti koeficientů v účelové funkci (pro minimalizaci od nejmenšího, pro maximalizaci od největšího). V prvním kroku zvolíme všechny proměnné rovny 0 (pro minimalizaci) nebo rovny 1 (pro maximalizaci). Tím dostaneme nejmenší/největší možnou hodnotu účelové funkce pro binární proměnné. Pokud tento bod vyhovuje všem omezením, máme řešení. V opačném případě provedeme větvení podle první proměnné (tak aby zvýšení/snížení hodnoty účelové funkce bylo nejmenší možné), přičemž hodnoty ostatních proměnných se nemění. Znovu otestujeme přípustnost tohoto bodu a buď již máme řešení nebo pokračujeme v dalším větvení, přičemž některá větvení nemusí být nutná, protože se může stát, že ani případná změna všech doposud nezafixovaných proměnných nepovede ke splnění všech daných omezení. Tento postup si ukážeme v následujícím příkladě.

Příklad 3.4.6

Metodou větvení a mezí vyřešme úlohu binárního celočíselného programování

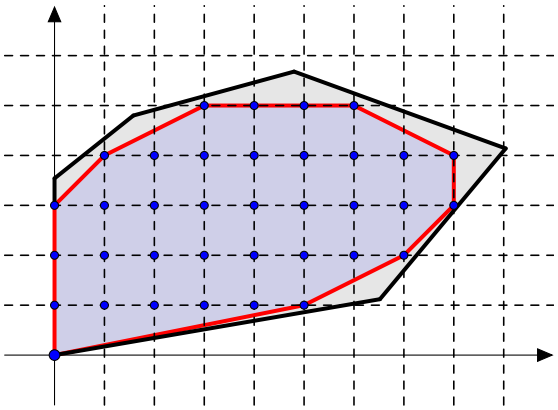
$$8x_1 + 2x_2 + 4x_3 + 7x_4 + 5x_5 - 10 \to \min$$
$$-3x_1 - 3x_2 + x_3 + 2x_4 + 3x_5 \le -2 \quad \& \quad -5x_1 - 3x_2 - 2x_3 - x_4 + x_5 \le -4$$
$$x_1 \in \{0,1\} \quad \& \quad x_2 \in \{0,1\} \quad \& \quad x_3 \in \{0,1\} \quad \& \quad x_4 \in \{0,1\} \quad \& \quad x_5 \in \{0,1\}.$$

Video ☞ 😎 https://ln.math.muni.cz/pz/opt/zlp3.

3.5 Metoda řezů[114]

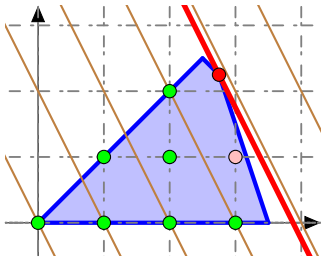
[114] Angl. *Cutting-Plane Method.*

Je sice jasné, že v případě ohraničené množiny docílíme metodou větvení a mezí dříve nebo později toho, že nově získané množiny budou mít celočíselný vrchol, ve kterém má přidružená úloha lineárního programování řešení, jenže celý výpočet je poněkud zdlouhavý a může v něm být velké množství „zbytečných“ výpočtů (a to i při využití duálního simplexového algoritmu a vhodné strategie výběru množin ze seznamu $\mathcal{S}$). Nešlo by dosáhnout celočíselnosti vrcholů přípustné množiny přidružené úlohy jinak? Efektivněji? Ať už tím máme na mysli cokoli… Jako nejpřirozenější se asi jeví zcela ořezat přípustnou množinu přidružené úlohy tak, aby všechny vrcholy měly celočíselné vrcholy a zároveň nedošlo k odstranění žádného přípustného bodu celočíselné úlohy (ℤLP). Takovou množinou by pak byl *konvexní obal* všech přípustných bodů úlohy (ℤLP), tj. nejmenší konvexní množina (mnohoúhelník) obsahující všechny tyto body. To můžeme vidět na Obrázku 3.22.

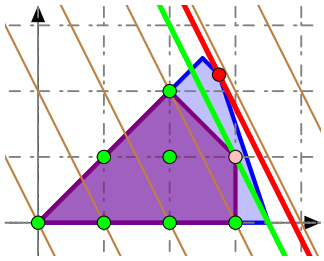


Obrázek 3.22: Přípustná množina přidružené úlohy (šedou barvou a vymezená černými úsečkami) a konvexní obal přípustných celočíselných bodů (modrou barvou a vymezená červenými úsečkami).

Jak ale takový obal najít? Např. množina vymezená podmínkami s unimodulární maticí A splývá se svým konvexním obalem, zatímco pro množinu na Obrázku 3.22 by to znamenalo 8 dodatečných nerovností, které určují jednotlivé červené úsečky (vyjma úsečky na svislé ose, která splývá s hranicí původní množiny). Jak ale tato omezení najdeme? Máme-li k dispozici obrázek či zjeví-li se opět *Deus ex machina*, je to poměrně snadné.



Obrázek 3.23: Přípustná množina z Příkladu 3.5.1.



Obrázek 3.24: Přípustná množina z Příkladu 3.5.1 po přidání dvou dodatečných omezení.

Příklad 3.5.1

Uvažme např. úlohu

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad -x_1 + x_2 \le 0 \quad \& \quad 6x_1 + 2x_2 \le 21$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Přípustná množina přidružené úlohy lineárního programování (tj. bez požadavku celočíselnosti) a přípustné celočíselné body dané úlohy jsou vyznačeny na Obrázku 3.23. Neceločíselným řešením je bod $\widetilde{x} = [11/4, 9/4]$ (červená barva), zatímco hledaným celočíselným řešením je $x^* = [3,1]$ (růžová barva). Přidáním dvou dodatečných omezení

$$x_1 + x_2 \le 4 \quad \& \quad x_1 \le 3$$

ořízneme výchozí přípustnou množinu přidružené úlohy bez požadavku celočíselnosti tak, že její vrcholy budou mít pouze celočíselné souřadnice (body $[0,0]$, $[2,2]$, $[3,1]$ a $[3,0]$), takže řešení této úlohy bude splývat s řešením zadané úlohy, viz Obrázek 3.24 (s oříznutou přípustnou množinou zakreslenou fialovou barvou).

V obecném případě může být počet potřebných omezení poměrně vysoký a jejich určení je de facto nemožný úkol. Naštěstí není zcela nutné, abychom přípustnou množinu přidružené úlohy ihned ořízli, co nejvíce to jde – až na samotnou hranici konvexního obalu. Měla by nám stačit jen ta část konvexního obalu, kde se nachází řešení (z něho potřebujeme udělat vrchol), tj. rozhodně to není potřeba dělat a priori bez ohledu na účelovou funkci. To už by mohlo být realizovatelné, jenže obvykle se jedná o stále poměrně dost složitý úkol. Co kdybychom se ale k této části obalu dopracovali *postupným* ořezáváním přípustné množiny přidružené úlohy? Třeba tak, že bychom (podobně jako v případě metody větvení a mezí) nejdříve nějak odřízli tu část množiny, která obsahuje aktuální řešení přidružené úlohy nesplňující dané podmínky celočíselnosti, a současně neodřízli žádný přípustný bod původní úlohy. To už proveditelné je, ale má-li to fungovat a být konkurenceschopné metodě větvení a mezí, jsou k tomu potřeba ještě jisté znalosti. Tento proces postupného ořezávání pak opakujeme tak dlouho, dokud se nám nepodaří získat hledané řešení. Toto je základní princip tzv. *metody řezů* (nebo

sečných nadrovin či *platných nerovností*). Existují různé možnosti, jak takové řezy v podobě vhodných nerovností získat, a my si ukážeme dvě z nich[115]. Iniciátorem této metody je Gomory, který také stojí za jejím značným rozvojem, viz první článek [227] z roku 1958 a také navazující zprávu [228] či [230]. V následujícím komentáři se Gomory v této souvislosti zmiňuje i o Dantzigovi, který „svoji“ variantu metody řezů zveřejnil o rok později, viz [116].

> *…But the Navy guys, the OR guys, kept me as a consultant. So I went down. And they were using linear programming to design a Navy task force. When you have an aircraft carrier, you've got to have some destroyers with it, and so forth. So you need to figure out the number of ships and things like that, and how far they can go — and do they need tankers.*
>
> *And they use linear programming to minimize the cost of the task force, how it was composed. But the trouble was that the answers came out things like* 2 *and a quarter aircraft carriers. They weren't whole numbers. And it wasn't clear what you did with* 2 *and a quarter aircraft carriers. You could fiddle around with it and decide, did they mean* 2*, or did they mean* 3*?*
>
> *And they said, hey, it'd be really good if we could do linear programming and get whole-number answers. So I listened to that, and I thought, yeah. That's right. And it should be doable. I didn't realize that people had thought about this and thought it was not doable…*
>
> *Greatly excited I set to work and within a few days had discovered how to generate maximal integer forms easily from the rows of the transformed simplex matrix… This was the origin of the "fractional cut."*
>
> *Within a very few days, I had worked out a complete method using the fractional cuts. I thought of this method as the "The Method of Integer Forms." With it I was steadily solving by hand one small numerical example after another and getting the right answer. However, I had no proof of finiteness…*
>
> *Just at this time I ran into Martin Beale in the hall. He was looking for a speaker for the seminar we had on game theory and linear programming. I said I would be glad to give a talk on solving linear programs in integers. Martin said "but that's impossible." That was my first indication that others had thought about the problem.*
>
> *During the exciting weeks that followed, I finally worked out a finiteness proof and then programmed the algorithm on the E101, a pin board computer that was busy during the day but that I could use after midnight. The E101 had only about 100 characters of memory and the board held only 120 instructions at one time, so that I had to change boards after each simplex maximization cycle and put in a new board that generated the cut, and then put the old board back to re-maximize. It was also hard work to get the simplex method down to 120 E101 instructions. But the results were better and more reliable than my hand calculations, and I was able to steadily and rapidly produce solutions to four- and five-variable problems.*
>
> *During these weeks I learned that others had thought about the problem and that George Dantzig had worked on the traveling salesman problem and had applied special handmade cuts to that.*
>
> *Professor Tucker, who was enormously helpful during this period, as during my entire stay at Princeton, gave me the time he had for himself on the program of a mathematical society meeting. There early in 1958 we made the first public presentation of the cutting plane algorithm. This produced*

[115] Zejména s využitím [153, Chapter 5].

[116] *Ale kluci z námořnictva, kluci z OV* (operační výzkum)*, si mě nechali jako konzultanta. Tak jsem šel dolů. Pomocí lineárního programování tam sestavovali plány pro jednotky námořnictva. Když máte letadlovou loď, musíte s ní mít i nějaké torpédoborce atd. Musíte tedy zjistit počet lodí a podobných věcí a jak daleko mohou plout — potřebují totiž i tankery.*

A lineární programování používali k minimalizaci nákladů na jednu skupinu podle jejího složení. Problém však byl v tom, že odpovědi vycházeli jako 2 *a čtvrt letadlové lodi. Nebyla to celá čísla. A nebylo jasné, co s těmi* 2 *a čtvrt loděmi udělat. Mohli byste se v tom vrtat a rozhodnout – znamená to* 2*, nebo* 3 *lodě?*

A oni řekli, hej, bylo by vážně dobré, kdybychom mohli s lineárním programováním dostat celočíselné výsledky. Poslouchal jsem to a myslel jsem si, jo. To je správně. A mělo by to být proveditelné. Neuvědomil jsem si, že lidé už o tom přemýšleli a mysleli si, že to není možné…

S velkým nadšením jsem se pustil do práce a během několika dní jsem objevil, jak snadno získat maximální celočíselné výrazy z řádků transformované simplexové matice… To byl původ „zlomkového řezu“.

Během několika dní jsem vypracoval kompletní metodu využívající zlomkové řezy. Tuto metodu jsem považoval za „Metodu celočíselných forem“. S ní jsem neustále ručně řešil jeden malý numerický příklad za druhým a dostával správné odpovědi. Jenže jsem neměl žádný důkaz konečnosti…

Zrovna v té době jsem v hale narazil na Martina Bealea. Hledal řečníka pro náš seminář o teorii her a lineárním programování. Řekl jsem mu, že bych rád přednesl příspěvek o řešení lineárních programů v celých číslech. Martin řekl „ale to je nemožné“. To pro mně byl první náznak toho, že i ostatní o problému už přemýšleli.

Během následujících vzrušujících týdnů jsem konečně vypracoval důkaz o konečnosti a pak jsem naprogramoval algoritmus na E101, počítači se čtečkou papírových pásek a děrovanými pláty, který byl přes den zaneprázdněn, ale mohl jsem ho používat v nočních hodinách. E101 měl jen asi 100 znaků paměti a deska zvládla pouze 120 instrukcí najednou, takže jsem musel po každém cyklu simplexové maximalizace vyměnit desky a vložit novou desku, která generovala řez, a pak vložit zpět starou desku a znovu maximali-

a great deal of reaction, many people wrote to me, and Rand Corporation invited me to come out to California for the summer.

In the summer of 1958 I flew west to Los Angeles, where Rand was located, carrying the first edition of the manual for Fortran, then a brand new language. I spent one month at Rand and succeeded in producing a working version of the algorithm, written in Fortran, for the IBM 704. During my stay at Rand, I renewed my acquaintance of graduate student days with Lloyd Shapley and Herb Scarf (whom I now count as my oldest friend) and met for the first time George Dantzig, Dick Bellman, and Phil Wolfe. Phil, already well known for his work on quadratic programming, generously took on the assignment of orienting me during my visit at Rand. He helped me in every conceivable way and I am pleased that even today, more than thirty years later, we are still friends.

The Fortran program seemed to be debugged about two days before I left Rand so I was able to do larger examples. Larger meant something like ten to fifteen variables. Most of these problems ran quickly, but one went on and on and producing reams of printout but never reaching a final answer. I thought at the time that perhaps there were still bugs left in the program, but in fact it was the first hint of the computational problems that lay ahead…

In the summer of 1959, I joined IBM Research and was able to compute in earnest. Compute in earnest meant that we had one day turnaround on a 704. But that was enough because I also had a wonderful programmer, Carol Shanesy. We started to experience the unpredictability of the computational results rather steadily.

I then turned my mind to something I had always wanted to have, an all--integer programming algorithm, an algorithm that resembled the simplex method but in which the coefficients always remained integers. In this sense the method I was aiming at resembled the classical diophantine approach.

I was able at the time to invent such a method. To keep all elements integer the pivot element had to be a 1 at all times. I did succeed in generating additional rows (valid inequalities) for the dual simplex method in such a way that there was always a candidate element for pivot element that was 1, and in such a way that it was always chosen by the normal rules of pivot choice. This was what I had more or less imagined and hoped for in the earlier period when I was doing hand calculations, so I was tremendously delighted and excited. In addition I found that when the problem of finding the greatest common divisor of two numbers was formulated as an integer programming problem, the new method did it using the same steps as the classical methods.[116]

zovat. Také bylo těžké dostat simplexovou metodu na 120 instrukcí E101. Ale výsledky byly lepší a spolehlivější než moje ruční výpočty a byl jsem schopen opakovaně a rychle nalézat řešení problému se čtyřmi a pěti proměnnými.

Během těchto týdnů jsem zjišťoval, že o problému přemýšleli i ostatní a že George Dantzig pracoval na problému obchodního cestující a využil k tomu speciální ručně získané řezy.

Profesor Tucker, který mi byl v tomto období stejně jako během celého mého pobytu na Princetonu nesmírně nápomocný, mi věnoval svůj časový slot v rámci programu setkání matematické společnosti. Tam jsem počátkem roku 1958 poprvé veřejně představili algoritmus řezných rovin. Vyvolalo to velkou reakci, mnoho lidí mi psalo a společnost Rand mě pozvala, abych přijel na léto do Kalifornie.

V létě 1958 jsem odletěl na západ do Los Angeles, kde sídlila společnost Rand, a vzal jsem si s sebou první vydání příručky pro Fortran, tehdy zcela nového programovacího jazyku. Strávil jsem jeden měsíc v Randu a podařilo se mi vytvořit pracovní verzi algoritmu napsaného ve Fortranu pro IBM 704. Během mého pobytu v Randu jsem obnovil svá přátelství z vysoké školy s Lloydem Shapleyem a Herbem Scarfem (kterého nyní považuji za svého nejstarší přítele) a poprvé se setkal s Georgem Dantzigem, Dickem Bellmanem a Philem Wolfem. Phil, nyní již velmi dobře známý svou prací na kvadratickém programování, se velkoryse ujal úkolu pomoci mi se zorientovat v Randu během mé návštěvy. Pomáhal mi všemi myslitelnými způsoby a těší mě, že i dnes po více než třiceti letech jsme stále přátelé.

Zdálo se, že asi dva dny před mým odjezdem z Randu byl program ve Fortranu odladěn, takže jsem mohl dělat i větší příklady. Větší znamenalo něco jako deset až patnáct proměnných. Většina těchto problémů běžela rychle, jenže jeden pokračoval a pokračoval a dával stohy výtisků, ale nikdy nedospěl ke konečné odpovědi. V té době jsem si myslel, že v programu možná ještě zůstaly ně-

Dantzigova metoda je založena na poměrně snadném tvrzení, které ukazuje jak lze „odříznout“ neceločíselné řešení přidružené úlohy lineárního programování, aniž bychom přišli o jakýkoli přípustný bod původní (celočíselné) úlohy.

VĚTA 3.5.2
DANTZIGŮV ŘEZ

jaké chyby, ale ve skutečnosti to byl první náznak výpočetních problémů, které byly před námi…

V létě 1959 jsem nastoupil do IBM Research a díky tomu jsem byl schopen doopravdy počítat. To zname-

Uvažme úlohu (ZLP) a řešení $\widehat{x}$ přidružené úlohy lineárního programování, což je BPB s bázi $B(\widehat{x})$. Jestliže tento bod nesplňuje pro nějaký index $i \in \{1,\ldots,n\}$ podmínku celočíselnosti, pak nesplňuje ani podmínku *Dantzigova řezu*

$$\sum_{j \in \{1,\ldots,n\} \setminus B(\widehat{x})} x_j \geq 1, \tag{3.5.1}$$

které však vyhovují všechny přípustné body původní úlohy.

nalo, že jsme měli na jeden den k dispozici 704. Ale to stačilo, protože jsem měl také skvělou programátorku Carol Shanesy. Celkem stabilně jsme začali zaznamenávat nepředvídatelnost výsledků výpočtů.

Pak jsem obrátil svoji pozornost k něčemu, co jsem vždy chtěl mít – algoritmu celočíselného programování, algoritmu, který se podobal simplexové metodě, ale v němž koeficienty vždy zůstávaly celočíselné. V tomto smyslu se metoda, ke které jsem směřoval, podobala klasickému diofantickému přístupu.

Dokázal jsem tehdy takovou metodu vymyslet. Aby všechny prvky zůstaly celé číslo, musel být pivot vždy 1. *Úspěšně se mi podařilo vygenerovat další řádky (platné nerovnosti) pro duální simplexovou metodu tak, že vždy existoval kandidát na pivota, který byl* 1 *a takový pivot mohl vždy být vybrán podle normálních pravidel. To bylo to, co jsem si víceméně představoval a v co jsem doufal v dřívějším období, kdy jsem prováděl ruční výpočty, takže jsem byl nesmírně potěšen a nadšen. Navíc jsem zjistil, že když byl problém nalezení největšího společného dělitele dvou čísel formulován jako problém celočíselného programování, nová metoda to dělala pomocí stejných kroků jako klasické metody…* Viz [232, 546] a Obrázek 3.25.

Obrázek 3.25: Počítač E101 ze společnosti Burroughs, na kterém Gomory prováděl své výpočty.

[117] Totéž platí i v případě degenerované úlohy jen s tím rozdílem, že pro bod $\widehat{x}$ bychom mohli určit více bází.

Důkaz. Jelikož $\widehat{x}_j = 0$ pro všechna $j \in \{1,\dots,n\} \setminus B(\widehat{x})$, je zjevné, že bod $\widehat{x}$ podmínku (3.5.1) nesplňuje. Ovšem na druhou stranu všechny přípustné body s celočíselnými souřadnicemi tuto nerovnost splňují. Vskutku, nejdříve poznamenejme, že z definice báze plyne, že $\widehat{x}$ je jediný přípustný bod mající nulové souřadnice[117] s indexy $j \in \{1,\dots,n\} \setminus B(\widehat{x})$, a tudíž pro všechny ostatní přípustné body přidružené úlohy je

$$\sum_{j \in \{1,\dots,n\} \setminus B(\widehat{x})} x_j > 0.$$

Zejména pro všechny přípustné body s nezápornými celočíselnými souřadnicemi musí být $x_j \in \mathbb{N}$ pro alespoň jeden index $j \in \{1,\dots,n\} \setminus B(\widehat{x})$. Proto všechny takové body nutně vyhovují Dantzigovu řezu, a tudíž přidáním nerovnosti (3.5.1) nedojde k odstranění žádného přípustného bodu původní úlohy. ■

Nicméně jakkoli jednoduše Dantzigovy řezy získáme, je s touto variantou metody řezů spojeno i nemalé úskalí, které si ilustrujeme na následujícím příkladu.

Příklad 3.5.3

Pomocí Dantzigových řezů vyřešme úlohu z Příkladu 3.5.1, tj.

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad -x_1 + x_2 \le 0 \quad \& \quad 6x_1 + 2x_2 \le 21$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Řešení. Přidružená úloha v kanonickém tvaru je

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 + s_1 = 5 \quad \& \quad -x_1 + x_2 + s_2 = 0 \quad \& \quad 6x_1 + 2x_2 + s_3 = 21$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad s_1 \ge 0 \quad \& \quad s_2 \ge 0 \quad \& \quad s_3 \ge 0.$$

Jejím řešením simplexovou metodou získáme $x^{[1]} = [11/4, 9/4, 0, 1/2, 0]$ s hodnotou $f^{[1]} = -31/4 = -7{,}75$. Odpovídající Dantzigův řez v kanonickém tvaru je

$$s_1 + s_3 - s_4 = 1,$$

po jehož přidání dostaneme řešení $x^{[2]} = [5/2, 5/2, 0, 0, 1, 0]$ s hodnotou $f^{[2]} = -15/2 = -7{,}5$. Tentokrát máme v řešení nulových hodnot více než dvě, což souvisí s degenerovaností přidružené úlohy. Ale při bližším pohledu do simplexové tabulky (ty zde pro stručnost vůbec neuvádíme) zjistíme, že báze bodu $x[2]$ je tvořena x_1, x_2, s_2 a s_3. Proto další Dantzigův řez je

$$s_1 + s_4 - s_5 = 1,$$

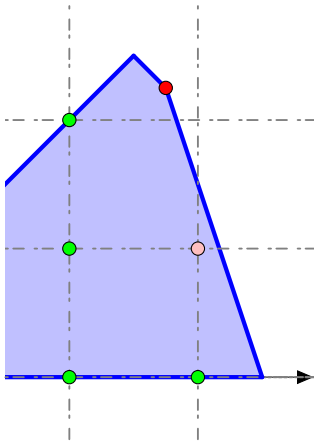
po jehož přidání dostaneme řešení $x^{[3]} = [13/4, 3/4, 1, 5/2, 0, 0, 0]$ s hodnotou $f^{[3]} = -29/4 = -7{,}25$. Dalším Dantzigovým řezem je

$$s_3 + s_5 - s_6 = 1,$$

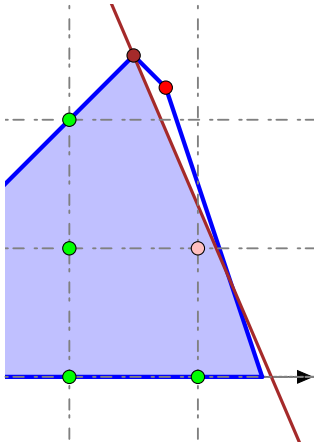
po jehož přidání dostaneme řešení $x^{[4]} = [11/4, 7/4, 1/2, 1, 1, 1/2, 0, 0]$ s hodnotou $f^{[4]} = -29/4 = -7{,}25$. Dalším Dantzigovým řezem je

$$s_5 + s_6 - s_7 = 1,$$

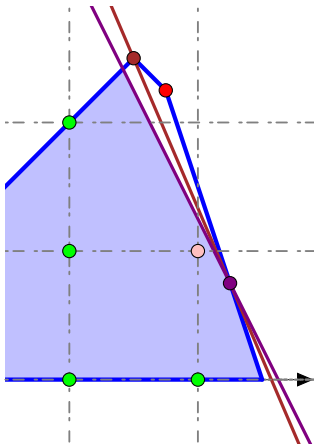
po jehož přidání dostaneme řešení $x^{[5]} = [77/32, 77/32, 3/16, 0, 7/4, 15/16, 1/8, 0]$ s hodnotou $f^{[5]} = -231/32 = -7{,}21875$. Opakováním tohoto postupu dostaneme např. po přidání patnáctého Dantzigova řezu bod $x^{[16]}$ s funkční hodnotou $f^{[16]} = -7{,}06743$. To, zda se nějakou vhodnou volbou dalších Dantzigových řezů dopracujeme k řešení $x^* = [3, 1]$ s hodnotou $f^* = -7$, se nám, bohužel, nepodařilo zjistit.



Obrázek 3.26: Přidružená úloha lineárního programování z Příkladu 3.5.3 a její řešení (červený bod). Růžový bod je řešením původní úlohy.



Obrázek 3.27: Přidružená úloha lineárního programování z Příkladu 3.5.3 s prvním Dantzigovým řezem a její řešení (vše hnědou barvou).



Obrázek 3.28: Přidružená úloha lineárního programování z Příkladu 3.5.3 s prvními dvěma Dantzigovými řezy a její řešení (fialovou barvou).

Ještě dodejme, že náš první Dantzigův řez odpovídá dodatečnému omezení

$$7x_1 + 3x_2 \leq 25 \quad \text{neboli} \quad 0,28x_1 + 0,12x_2 \leq 1.$$

Z druhého řezu získáme podmínku

$$8x_1 + 4x_2 \leq 29 \quad \text{neboli} \quad 0,27586x_1 + 0,137931x_2 \leq 1$$

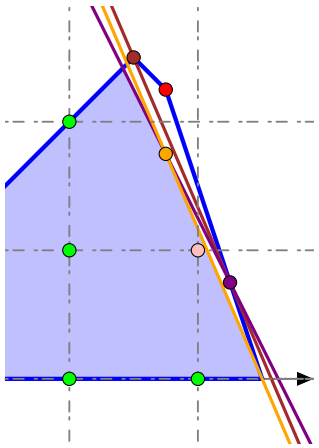
a ze třetího řezu to je

$$14x_1 + 6x_2 \leq 49 \quad \text{neboli} \quad 0,285714x_1 + 0,12244898x_2 \leq 1$$

atd. Na Obrázcích 3.26–3.29 je znázorněno postupné ořezávání přípustné množiny původní přidružené úlohy s detailem na tu její část, kde se nachází řešení. Nový řez a odpovídající řešení přidružené úlohy jsou vždy jinou barvou a aktuální přípustná množina je vybarvena modrou barvou.

V předchozím příkladu jsme názorně viděli, že se s pomocí Dantzigových řezů sice přibližujeme k hledanému řešení, ale celý proces je poměrně zdlouhavý a funkční hodnoty se zlepšují jen velmi pozvolna. To je také největší úskalí celé této metody, neboť jednotlivé řezy jsou takřkajíc *mělké*[118]. Vlastně i samotný Dantzig se ve svém článku [116] zmiňuje o tom, že nemá žádný důkaz konečnosti této metody. Ani jej mít nemohl! Jeho pochybnosti totiž potvrdili Gomory s Hoffmanem a Watsonem, kteří ukázali, že tato metoda povede k řešení v konečném počtu kroků pouze v případě, že hledané celočíselné řešení leží na hraně polyedru přidružené úlohy, viz [233]. Avšak není to úplně k zahození, protože už o pár let později představili Bowman a Nemhauser drobné vylepšení Dantzigových řezů omezující výběr indexů $j \in \{1,\dots,n\} \setminus B(\widehat{x})$, které problém s nekonečným počtem potřebných kroků odstraňuje, viz [78]. Ještě také poznamenejme, že vzhledem ke tvaru nerovností generovaných tímto způsobem je výhodnější (vyjma řešení původní přidružené úlohy) po přidání Dantzigova řezu do poslední simplexové tabulky pokračovat duální simplexovou metodou, přičemž s každým novým řezem dojde ke zvětšení simplexové tabulky o jeden řádek a jeden sloupec. Samozřejmě by bylo možné pokračovat i klasickou simplexovou metodou. To by ale vedlo ke zvětšení tabulky o jeden řádek a dva sloupce a hlavně k nutnosti využít první fázi dvoufázové metody. Proto bývá tato metoda řazena do kategorie tzv. *duálních metod řezů*. Totéž platí i pro Gomoryho metodu, čehož jsme si mohli všimnout i v úryvku výše. Nezapomínejme ale, že tímto způsobem budeme v každém kroku mít během výpočtu pouze duálně přípustné body a primárně přípustný bod získáme až na jeho konci, tj. pokud v daném kroku nedokončíme z nějakého důvodu (zejména kvůli výpočetní náročnosti nebo když už dochází jen k velmi pozvolnému zlepšování funkční hodnoty) celý simplexový algoritmus, nezískáme ani nějakou přípustnou „aproximaci“ hledaného řešení. Budeme mít akorát jakýsi dolní odhad minimální hodnoty účelové funkce (a odpovídající primárně nepřípustný bod). Může se tak stát, že investujeme značné úsilí a množství výpočetního času, aniž bychom dosáhli nějakého užitečného výsledku. Z tohoto důvodu byly vyvinuty i algoritmy, které sice naleznou pouze *suboptimální* řešení, avšak to bude celočíselné a přípustné.

Měli bychom zde také zdůraznit, že na rozdíl od metody větvení a mezí tyto „naše“ současné úvahy fungují pouze pro úlohy celočíselného programování a nelze je *verbatim et litteratim*[119] aplikovat i na úlohy smíšeného celočíselného programování! Proč? K tomu můžeme využít např. obrázky ilustrující řešení Příkladu 3.5.3. Kdybychom vypustili jednu z podmínek celočíselnosti, byla by přípustná množina tvořena úsečkami (případně polopřímkami) s jednou celočíselnou souřadnicí. Řezy by se v průběhu výpočtu nezměnili[120], ale kromě odříznutí nalezeného řešení přidružené úlohy by mohlo dojít (a dojde) i k odříznutí části jedné nebo více těchto úseček/polopřímek, tj. odřízli bychom i některé přípustné body původní úlohy.



Obrázek 3.29: Přidružená úloha lineárního programování z Příkladu 3.5.3 s prvními třemi Dantzigovými řezy a její řešení (oranžovou barvou).

[118] Angl. *shallow cut*.

[119] *Do slova a do písmene*.

[120] Protože přidružená duální úloha lineárního programování by se nezměnila.

[121] Kdyby takové s existovalo, muselo by platit $s + 3x_1 = 13/3 - x_2/2$ pro nějaká celočíselná x_1, x_2. Avšak zatímco výraz na levé straně bude v takovém případě jistě celočíselný, na pravé straně nikdy pro $x_2 \in \mathbb{Z}$ celé číslo nedostaneme ↯.

[122] Angl. *fraction cut* (nebo *f-cut*) *method*.

[123] V takovém případě neexistuje žádná celočíselná volba jednotlivých proměnných, pro kterou bude mít levá strana neceločíselnou hodnotu.

Dodejme také, že požadavek celočíselnosti se samozřejmě vztahuje i na pomocné proměnné, což je potřeba mít na paměti při přechodu do kanonického tvaru, tj. vynásobením vhodnými konstantami je nutné se nejdříve „zbavit“ racionálních koeficientů v účelové funkci i jednotlivých omezeních a mít úlohu pouze s celočíselnými koeficienty. Jinými slovy, při řešení první přidružené úlohy musí být ve výchozí simplexové tabulce pouze celočíselné koeficienty. Např. z podmínky

$$3x_1 + x_2/2 \leq 13/3$$

pro proměnné $x_1, x_2 \in \mathbb{Z}$ bychom po přechodu do kanonického tvaru dostali rovnost

$$3x_1 + x_2/2 + s = 13/3,$$

která však není splněna pro žádné[121] $s \in \mathbb{Z}$, tj. řešením příslušné úlohy celočíselného programování pomocí Dantzigových řezů dospějeme k její nepřípustnosti. Ovšem vynásobíme-li obě strany této nerovnosti číslem 6 jakožto společným jmenovatel obou zlomků, pak po přechodu k rovnosti dostaneme

$$18x_1 + 3x_2 + s = 26,$$

což už nepochybně celočíselné řešení má.

Dantzig svoji metodu prezentoval jako alternativu ke Gomoryho *metodě zlomkových řezů*[122] s cílem snadnějšího (a lépe automatizovatelného) generování jednotlivých řezů. To se Dantzigovi nepochybně povedlo, ale jen za cenu velkých pochybností o samotné praktičnosti navrhovaného postupu. V Gomoryho metodě mohou být jednotlivé řezy generovány pomocí jakéhokoli omezení

$$\sum_{j=1}^{n} a_{kj} x_j = b_k \tag{3.5.2}$$

ze zadání či získaného v průběhu výpočtu mající alespoň jeden neceločíselný koeficient a_{kj} či b_k. Toto omezení můžeme přepsat do tvaru

$$\sum_{j=1}^{n} (\lfloor a_{kj} \rfloor + r_{kj})\, x_j = \lfloor b_k \rfloor + r_k,$$

kde r_{kj} a r_k vyjadřují zlomkové části jednotlivých koeficientů, tj. $a_{kj} = \lfloor a_{kj} \rfloor + r_{kj}$ a $b_k = \lfloor b_k \rfloor + r_k$ pro $\lfloor \cdot \rfloor$ značící dolní celou část. Jelikož hledáme bod s celočíselnými souřadnicemi splňující (mimo jiné) i toto omezení, plyne z celočíselnosti výrazu na levé straně rovnosti

$$\sum_{j=1}^{n} \lfloor a_{kj} \rfloor x_j - \lfloor b_k \rfloor = r_k - \sum_{j=1}^{n} r_{kj} x_j \tag{3.5.3}$$

také celočíselnost výrazu na pravé straně. Ba co víc, jednotlivé složky hledaného řešení musí splňovat podmínku nezápornosti a současně $r_{kj}, r_k \in [0, 1)$, takže odtud získáváme novou nerovnost

$$r_k - \sum_{j=1}^{n} r_{kj} x_j \leq r_k < 1 \tag{3.5.4}$$

neboli

$$r_k - \sum_{j=1}^{n} r_{kj} x_j \leq 0.$$

Toto je onen *Gomoryho zlomkový řez*, který představuje nutnou podmínku, jenž musí splňovat každý přípustný bod (a tedy nutně i řešení) úlohy (ℤLP), přičemž nezapomínejme, že i tentokrát jsme využili předpoklad celočíselnosti všech proměnných, tj. ani tentokrát nelze získané řezy využít při řešení úlohy smíšeného celočíselného programování. A jaký vliv bude mít přidání Gomoryho zlomkového řezu na řešení přidružené úlohy lineárního programování? Na odpověď si budeme muset ještě chvíli počkat – nejdříve si ukážeme ilustrativní příklad.

Můžeme si všimnout i toho, že v případě celočíselných koeficientů na levé straně uvažovaného omezení a neceločíselné hodnoty na jeho pravé straně, bude původní úloha jistě nepřípustná[123], což potvrzuje i odpovídající zlomkový řez

$$0 \geq r_k > 0.$$

Příklad 3.5.4

Přidružená úloha lineárního programování k úloze z Příkladu 3.5.1 má podobu

$$-2x_1 - x_2 \to \min$$
$$x_1 + x_2 + s_1 = 5 \quad \& \quad -x_1 + x_2 + s_2 = 0 \quad \& \quad 6x_1 + 2x_2 + s_3 = 21$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad s_1 \geq 0 \quad \& \quad s_2 \geq 0 \quad \& \quad s_3 \geq 0.$$

Pro tuto úlohu můžeme získat různé zlomkové řezy. Např. podělením první rovnosti dvojkou dostaneme $x_1/2 + x_2/2 + s_1/2 = 5/2$, což po aplikování zlomkového řezu dává dodatečné omezení

$$x_1 + x_2 + s_1 \geq 1,$$

které je však díky prvnímu omezení ekvivalentní s triviální nerovností $5 \geq 1$. Podobně podělením třetí rovnosti dvojkou získáme zlomkový řez

$$s_3 \geq 1$$

neboli v původních proměnných máme podmínku

$$6x_1 + 2x_2 \leq 20,$$

zatímco podělením šestkou obdržíme

$$2x_2 + s_3 \geq 3 \quad \text{neboli} \quad x_1 \leq 3. \tag{3.5.5}$$

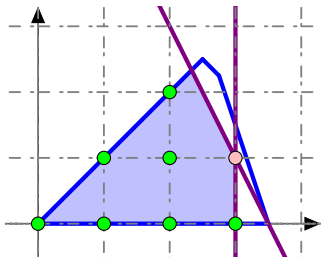
Sečtení prvních dvou rovností a následné podělení dvojkou vede ke zlomkovému řezu

$$s_1 + s_2 \geq 1 \quad \text{neboli} \quad x_2 \leq 2.$$

Konečně odečtení dvojnásobku první rovnosti od třetí rovnosti a následné podělení čtyřkou dává zlomkový řez

$$2s_1 + s_3 \geq 3 \quad \text{neboli} \quad 2x_1 + x_2 \leq 7. \tag{3.5.6}$$

Není těžké ověřit, že řešení $[x_1^*, x_2^*, s_1^*, s_2^*, s_3^*] = [3,1,1,2,1]$ skutečně splňuje všechna tato dodatečná omezení. Pro nalezení takového řešení pomocí přidružené úlohy by mělo stačit pouze přidání podmínek (3.5.5) a (3.5.6), díky kterým se bod $[3,1]$ stane vrcholem ořezaného polyedru, viz Obrázek 3.30 a srovnej s Obrázkem 3.24.



Obrázek 3.30: Přípustná množina z Příkladu 3.5.1 ořezaná s pomocí podmínek (3.5.5) a (3.5.6).

Jak jsme však mohli vidět, postup ilustrovaný v předchozím příkladě není příliš praktický. Takovýmto způsobem bychom totiž mohli generovat nespočet všelijakých zlomkových řezů, takže bychom nakonec obdrželi úlohu s velkým počtem omezení, z nichž některá mohou být zcela nadbytečná. Je potřeba to udělat systematičtěji a efektivněji, tj. pomocí simplexové metody, díky čemuž můžeme ukázat, že přidáním Gomoryho zlomkového řezu dojde k „odříznutí“ řešení přidružené úlohy lineárního programování.

Věta 3.5.5 Gomoryho zlomkový řez

Uvažme úlohu (ZLP) a řešení $\widehat{x}$ přidružené úlohy lineární programování. Jestliže tento bod nesplňuje pro nějaký index $i \in \{1,\ldots,n\}$ podmínku celočíselnosti a řádek v poslední simplexové tabulce odpovídající této bazické proměnné má podobu (3.5.2), pak bod $\widehat{x}$ nesplňuje ani příslušný Gomoryho zlomkový řez

$$\sum_{j \notin B(\widehat{x})} r_{kj} x_j \geq r_k, \tag{3.5.7}$$

kterému však vyhovují všechny přípustné body původní úlohy.

Důkaz. To, že nerovnost (3.5.7) je splněna všemi přípustnými body úlohy (ℤLP), jsme ukázali již v předchozí části. Je-li $\widehat{x}$ řešením přidružené úlohy lineárního programování, pak v poslední simplexové tabulce máme v řádku odpovídajícímu bazické proměnné x_k koeficienty $\lambda_{kk} = 1$ a $\lambda_{kj} = 0$ pro $k \in B(\widehat{x})$ a $j \in B(\widehat{x}) \setminus \{k\}$, v kterémžto případě se omezení (3.5.3) redukuje na

$$x_k + \sum_{j \notin B(\widehat{x})} \lambda_{kj} x_j = b_k.$$

Odtud získáme zlomkový řez (3.5.7) v podobě

$$\sum_{j \notin B(\widehat{x})} r_{kj} x_j \geq r_k.$$

Jestliže $\widehat{x}_k \notin \mathbb{Z}$, pak $r_k \in (0,1)$ a bod $\widehat{x}$ předchozí nerovnost nesplňuje. ■

To nás přivádí k následujícímu návodu pro řešení úlohy (ℤLP).

Gomoryho metoda zlomkových řezů

(1) *Inicializace.* Je dána úloha celočíselného programování s celočíselnými koeficienty. Je vyřešena přidružená úloha lineárního programování. Pokud

(i) je tato úloha nepřípustná, pak totéž platí pro původní úlohu. KONEC.

(ii) je tato úloha neohraničená, pak původní úloha je buď také neohraničená nebo nepřípustná. KONEC.

(iii) nalezené řešení splňuje podmínky celočíselnosti, jedná se i o řešení původní úlohy. KONEC.

(iv) alespoň jedna složka nalezeného řešení není celočíselná, položíme $k = 0$.

(2) *Řezání.* Z poslední simplexové tabulky určující řešení $x^{[k]}$ odvodíme pomocí zlomkového řezu pro řádek odpovídající bazické proměnné s neceločíselnou hodnotou dodatečné omezení, které přidáme k aktuální úloze.

(3) *Relaxace.* Řešíme novou přidruženou úlohu lineárního programování. Jestliže

(i) přidružená úloha je nepřípustná, pak totéž platí pro původní úlohu. KONEC.

(ii) nalezené řešení $x^{[k+1]}$ splňuje podmínky celočíselnosti, jedná se i o řešení původní úlohy. KONEC.

(iii) nalezené řešení $x^{[k+1]}$ nesplňuje podmínky celočíselnosti, položíme $k \to k+1$ a pokračujeme Krokem (2).

V předchozím popisu však není nikterak specifikováno, jak po přidání Gomoryho zlomkového řezu realizovat krok (3). Z důkazu Věty 3.5.5 víme, že musíme využít řádek z poslední simplexové tabulky, který odpovídá bazické proměnné s neceločíselnou hodnotou – jedná se o tzv. *zdrojový řádek*[124]), ze kterého dostaneme Gomoryho zlomkový řez (3.5.7). Nejjednodušší možností je zavedení pomocné proměnné $s \geq 0$ a přidání nového omezení

$$\sum_{j \notin B(\widehat{x})} r_{kj} x_j - s = r_k \tag{3.5.8}$$

k původní přidružené úloze lineárního programování a spuštění celého simplexového algoritmu znovu[125]. Avšak jistě tušíme, že to asi nebude příliš efektivní postup. Nešlo by do toho

[124] Angl. *source row.*

[125] Nezapomínejme, že $r_k > 0$.

nějak zakomponovat již provedené výpočty? Samozřejmě ano. První možností je rozšíření poslední simplexové tabulky o jeden řádek a sloupec, kam bychom vepsali nové omezení (3.5.8) a proměnnou s, kterou přidáme do báze. Pak bychom však kvůli zápornému koeficientu u s museli pokračovat první fází klasického simplexového algoritmu, což je stále poměrně početně náročné. Druhou možností je přepsání zlomkového řezu do tvaru

$$-\sum_{j \notin B(\bar{x})} r_{kj}\, x_j + s = -r_k$$

a přidání tohoto omezení do poslední simplexové tabulky. Přidáním proměnné s sice tentokrát dojde k porušení přípustnosti aktuálního BPB (budeme mít jednu zápornou hodnotu ve sloupci x_B), avšak tento bod bude duálně přípustný[126]. Proto můžeme velmi elegantně pokračovat duální simplexovou metodou[127], přičemž z rovnosti (3.5.3) vyplývá, že i proměnná s musí mít celočíselnou hodnotu, tj. po přidání Gomoryho zlomkového řezu budeme mít opět úlohu celočíselného programování a můžeme pokračovat stejným způsobem.

[126] Hodnoty v posledním $z_k - c_k$ řádku se nezmění, tj. všechny jsou nekladné a přibude k nim nulová hodnota ve sloupci odpovídající nově přidané proměnné s.

[127] Přesně takto Gomory svůj algoritmus koncipoval: po vyřešení první přidružené úlohy lineárního programování klasickou simplexovou metodou je vše dokonale připraveno (v řádku $z_k - c_k$ máme pouze nekladné hodnoty) na pokračování duální simplexovou metodou.

Příklad 3.5.6

Metodou řezů vyřešme úlohu celočíselného programování

$$-x_2 \to \min$$
$$3x_1 + 2x_2 \le 6 \quad \& \quad -3x_1 + 2x_2 \le 0$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Video ☞ 😎 https://ln.math.muni.cz/pz/opt/zlp4.

⚠ Řádek pro a_4 v čase 20:50 (nebo 35:00 v duplicitním čase v obrazu videa) má být $0, 0, 0, 1, 1, -3/2, 1$ (vždyť vznikl pouze vynásobením řádku pro a_4 v předchozí tabulce číslem $-3/2$).

V uvedeném návodu je mechanismus zlomkových řezů aplikován postupně a v každém kroku dojde jen k jednomu oříznutí, při kterém je z přípustné množiny odstraněno neceločíselné minimum přidružené úlohy. Nabízí se otázka, zda by v případě, kdy má bod $x^{[k]}$ více neceločíselných souřadnic, bylo možné odvodit i více odpovídajících zlomkových řezů, a tudíž i přidat více dodatečných omezení současně. Teoreticky jistě ano, avšak takovéto vícenásobné ořezávání nemusí být vždy efektivní a hlavně důkaz zaručující konečnost Gomoryho algoritmu je založen právě na tom (kromě dvou dalších dodatečných požadavků), že se v každém kroku přidá vždy jen jedno nové omezení. Nicméně platí, že z poslední simplexové tabulky můžeme získat různé Gomoryho zlomkové řezy podle toho, jaký zvolíme výchozí řádek: (i) může to být libovolný řádek v této tabulce[128], (ii) libovolný násobek libovolného z těchto řádků nebo (iii) nějaká lineární kombinace těchto řádků. Z výpočetního hlediska by bylo nejlepší získat co „nejsilnější/nejhlubší“ řez, tj. takový, který odstraní největší část přípustné množiny přidružené úlohy, aniž by z ní odstranil jakýkoli bod s celočíselnými souřadnicemi. Jenže takové porovnání řezů by vyžadovalo posouzení i v kontextu ostatních daných omezení, což by bylo poměrně komplikované. Proto se obvykle hloubka řezu měří pouze vzhledem k množině $\mathbb{R}^n_+$.

[128] Pochopitelně by v něm měl být alespoň jeden neceločíselný koeficient.

DEFINICE 3.5.7

Dva zlomkové řezy $rx \ge r_0$ a $\hat{r}x \ge \hat{r}_0$ pro $r, \hat{r} \in \mathbb{R}^n_+$ a $r_0, \hat{r}_0 \in \mathbb{R}_+$ nazveme *ekvivalentní*, pokud existuje číslo $\gamma > 0$ takové, že $r = \gamma \hat{r}$ a $r_0 = \gamma \hat{r}_0$. Pokud tyto dva řezy nejsou ekvivalentní a současně existuje číslo $\mu > 0$ takové, že $r \le \mu \hat{r}$ a $r_0 \ge \mu \hat{r}_0$, pak řekneme, že řez $rx \ge r_0$ je *silnější*[129] než řez $\hat{r}x \ge \hat{r}_0$. Zlomkový řez nazveme *maximální*, jestliže neexistuje žádný silnější zlomkový řez.

[129] Alternativní terminologie je, že řez $rx \ge r_0$ je *dominuje* řezu $\hat{r}x \ge \hat{r}_0$. Případně lze říci, že řez $\hat{r}x \ge \hat{r}_0$ je *slabší* než řez (nebo je *dominován* řezem) $rx \ge r_0$. Každopádně dominance zlomkového řezu znamená, že $r_0 \ge \mu \hat{r}_0$ a $r_j \le \mu \hat{r}_j$ pro všechna $j \in \{1, \dots, n\}$ s tím, že alespoň jedna z těchto nerovností je ostrá.

Předchozí definice je motivována tím, že dva ekvivalentní řezy určují tentýž poloprostor, zatímco v případně dominujícího řezu máme $\{x \in \mathbb{R}^n_+ \mid rx \ge r_0\} \subsetneq \{x \in \mathbb{R}^n_+ \mid \hat{r}x \ge \hat{r}_0\}$, tj. s jeho

pomocí dojde k odstranění větší části $\mathbb{R}^n_+$. Např. vezmeme-li zdrojový řádek odpovídající rovnosti

$$x_1 - \frac{27}{5}x_2 + \frac{23}{5}x_3 = \frac{37}{5},$$

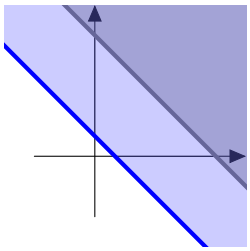
kde x_2 a x_3 jsou nebazické proměnné, pak z něj přímo dostaneme zlomkový řez

$$\frac{3}{5}x_2 + \frac{3}{5}x_3 \geq \frac{2}{5}. \tag{3.5.9}$$

Pokud ale zdrojový řádek vynásobíme 2, pak získáme řez

$$\frac{1}{5}x_2 + \frac{1}{5}x_3 \geq \frac{4}{5}. \tag{3.5.10}$$

Který je lepší? Pro první řez můžeme vzít $\hat{r} = (0,3,3)^\top$ a $\hat{r}_0 = 2$, zatímco druhý řez je určen vektorem $r = (0,1,1)^\top$ a $r_0 = 4$, takže s volbou $\mu = 1$ zjistíme, že druhý řez je silnější než ten první, viz Obrázek 3.31.



Obrázek 3.31: Porovnání zlomkových řezů (3.5.9) a (3.5.10) v proměnných x_2 a x_3. První z nich (slabší) je vymezen modrou přímkou a vyhovují mu všechny body v modré a šedé části. Druhý z nich (silnější) je určen šedou přímkou a vyhovují mu pouze body v šedé části. Odtud vidíme, že tento řez eliminuje větší část prvního kvadrantu.

Tento proces můžeme zobecnit a násobit zdrojový řádek libovolným $\alpha \in \mathbb{Z}$ či dokonce $\alpha > 0$, viz [202, str. 193–194] a [218, 229]. Uvažme proto opět zdrojový řádek

$$x_\ell + \sum_{j \in N} \lambda_{\ell j} x_j = \widehat{b}_\ell.$$

Jeho vynásobením číslem $\alpha \in \mathbb{Z}$ dostaneme

$$\alpha x_\ell + \sum_{j \in N} \alpha \lambda_{\ell j} x_j = \alpha \widehat{b}_\ell$$

neboli

$$\alpha x_\ell + \sum_{j \in N} \alpha \lfloor \lambda_{\ell j} \rfloor x_j - \alpha \lfloor \widehat{b}_\ell \rfloor = \alpha r_\ell - \sum_{j \in N} \alpha r_{\ell j} x_j.$$

Výraz na levé straně je celočíselný pro libovolné $x_\ell \in \mathbb{Z}$, takže totéž platí i pro hodnotu na pravé straně. Tato celočíselnost je zachována, i pokud výraz na pravé straně změníme na

$$\alpha r_\ell - \lfloor \alpha r_\ell \rfloor - \sum_{j \in N} (\alpha r_{\ell j} - \lfloor \alpha r_{\ell j} \rfloor) x_j.$$

Protože $\alpha r_\ell - \lfloor \alpha r_\ell \rfloor < 1$, dostáváme odtud *zobecněný zlomkový řez* (též *zlomkový α-řez*)

$$\alpha r_\ell - \lfloor \alpha r_\ell \rfloor - \sum_{j \in N} (\alpha r_{\ell j} - \lfloor \alpha r_{\ell j} \rfloor) x_j \leq 0,$$

což pro $\alpha = 1$ dává původní zlomkový řez. Vzhledem k tomu, že $\alpha \in \mathbb{Z}$, počet různých zobecněných zlomkových řezů získaných z jakéhokoli zdrojového řádku je *konečný* a je nejvýše roven $D := |\det A_B|$ pro aktuální bázi B, i když některé z nich jsou silnější než jiné[130]. Dá se dokonce ukázat, že celkový počet zobecněných zlomkových řezů, které můžeme z aktuální simplexové tabulky získat, je nejvýše roven D. Konkrétní počet ale záleží na zvoleném zdrojovém řádku – s nějakým řádkem můžeme klidně získat všechny řezy (pomocí nezáporného celého čísla menšího než $D-1$) a s jiným pouze některé.

Samozřejmě bychom nejraději vybrali ty nejsilnější možné řezy. Jak to ale udělat? Mohli bychom zkusit odvodit všechny zobecněné zlomkové řezy a vzájemně je porovnat (pokud to bude možné). To je ale prakticky nerealizovatelné pro větší hodnoty D. Chtělo by to nějaké kritérium určující výběr vhodného zdrojového řádku. Toho můžeme docílit např. volbou takového řádku odpovídajícího bazické proměnné x_ℓ, pro který buď

$$r_\ell = \max_{i \in B} \{ r_i \mid i \in B \}$$

nebo

$$\frac{r_\ell}{\sum_{j \in N} r_{\ell j}} = \max_{i \in B} \left\{ \frac{r_i}{\sum_{j \in N} r_{ij}} \,\middle|\, i \in B \right\}.$$

[130] Vskutku, všechny zobecněné zlomkové řezy mohou být generovány pouze pomocí čísel $\alpha \in \mathbb{Z}$ takových, že $0 \leq \alpha \leq D-1$. Ukážeme, že číslo $\alpha' := \alpha + kD$ pro libovolné $0 \leq \alpha \leq D-1$ a $k \in \mathbb{Z}$ dává stejný zobecněný zlomkový řez jako α. Nejdříve připomeňme, že jednotlivé prvky $\lambda_{\ell j}$ pro $\ell = 1,\ldots,m$ můžeme spočítat pomocí součinů $A_B^{-1} a_j$. Jelikož současně $A_B^{-1} = A_B^{\text{adj}}/(\det A_B)$ a matice A_B je celočíselná, je celočíselná i adjungovaná matice A_B^{adj}, a tudíž $\lambda_{\ell j} = \rho_{\ell j}/D$ pro vhodná $\rho_{\ell j} \in \mathbb{Z}$. Proto jistě $r_{\ell j} = p_{\ell j}/D$ pro nějaké $p_{\ell j} \in \mathbb{Z}$ splňující $0 \leq p_{\ell j} < D$. Díky tomu máme

$$\begin{aligned}
&\alpha' r_{\ell j} - \lfloor \alpha' r_{\ell j} \rfloor = \\
&= (\alpha + kD)\frac{p_{\ell j}}{D} - \left\lfloor (\alpha + kD)\frac{p_{\ell j}}{D} \right\rfloor \\
&= \frac{\alpha p_{\ell j}}{D} + k p_{\ell j} - \left\lfloor \frac{\alpha p_{\ell j}}{D} \right\rfloor - k p_{\ell j} \\
&= \alpha r_{\ell j} - \lfloor \alpha r_{\ell j} \rfloor
\end{aligned}$$

a analogickým způsobem získáme i rovnost $\alpha' r_\ell - \lfloor \alpha' r_\ell \rfloor = \alpha r_\ell - \lfloor \alpha r_\ell \rfloor$

[131] Nejmenší možnou hodnotu jsme našli již v prvním kroku při řešení přidružené úlohy. Pokud odpovídající řešení nesplňuje podmínky celočíselnosti, pak hodnota dané úlohy celočíselného programování nemůže být menší, tj. řešení se musí realizovat v bodě se stejnou nebo větší funkční hodnotou.

Jenže nezapomínejme, že Definice 3.5.7 nikterak nezohledňuje danou přípustnou množinu, takže se může stát, že takto sice zvolíme nejsilnější možný řez, ale s jeho pomocí ořízneme přípustnou množinu ve zcela marginálním směru. Mnohem efektivnější by mohlo být volit řez tak, aby došlo k co největšímu vzrůstu funkční hodnoty[131] (podobně jako u pravidel výběru klíčového sloupce v klasické simplexové metodě, viz Poznámku 2.8.9). Pro zlomkový řez

$$s - \sum_{j \in N} r_{\ell j} x_j = -r_\ell$$

můžeme jakýsi odhad zvětšení funkční hodnoty získat jako

$$P_\ell := r_\ell \max\left\{ \frac{z_j - c_j}{r_{\ell j}} \,\middle|\, j \in N \,\&\, r_{\ell j} > 0 \right\}, \tag{3.5.11}$$

takže největšího možného zvětšení funkční hodnoty bychom mohli dosáhnout pomocí zdrojového řádku, pro který se realizuje minimum

$$\min_{\ell \in B} \left\{ r_\ell \max\left\{ \frac{z_j - c_j}{r_{\ell j}} \,\middle|\, j \in N \,\&\, r_{\ell j} > 0 \right\} \right\}.$$

Ovšem není jisté, že se nám to touto volbou skutečně podaří, protože P_ℓ vyjadřuje jen dolní odhad zvětšení funkční hodnoty, tj. může existovat jiný zdrojový řádek, který by nás přivedl k lepšímu zlomkovému řezu, i když odpovídající odhad vzrůstu by byl menší. Jestliže maximum na pravé straně (3.5.11) nastane pro index $r \in N$, pak

$$P_\ell = (z_r - c_r) \min\left\{ \frac{r_\ell}{r_{\ell r}} \,\middle|\, \ell \in B \,\&\, r_{\ell j} > 0 \right\},$$

což ukazuje i jistou konzistenci mezi volbou zdrojového řádku pomocí P_ℓ a požadavkem silného řezu[132].

[132] Toto je založeno spíše na empirickém pozorování, že hodnota $\max\left\{ \frac{r_\ell}{r_{\ell r}} \,\middle|\, \ell \in B \right\}$ vyjadřuje jakousi míru síly řezu.

Příklad 3.5.8

Metodou řezů vyřešme úlohu celočíselného programování

$$-x_1 + 3x_2 - 3x_3 \to \min$$
$$2x_1 + x_2 - x_3 \le 4 \quad \& \quad 4x_1 - 3x_2 \le 2 \quad \& \quad -3x_1 + 2x_2 + x_3 \le 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z} \quad \& \quad x_3 \in \mathbb{Z}.$$

Řešení. Převodem do kanonického tvaru (a díky celočíselnosti všech koeficientů) získáme úlohu

$$-x_1 + 3x_2 - 3x_3 \to \min$$
$$2x_1 + x_2 - x_3 + s_1 = 4 \quad \& \quad 4x_1 - 3x_2 + s_2 = 2 \quad \& \quad -3x_1 + 2x_2 + x_3 + s_4 = 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0 \quad \& \quad s_1 \ge 0 \quad \& \quad s_2 \ge 0 \quad \& \quad s_3 \ge 0$$
$$x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z} \quad \& \quad x_3 \in \mathbb{Z} \quad \& \quad s_1 \in \mathbb{Z} \quad \& \quad s_2 \in \mathbb{Z} \quad \& \quad s_3 \in \mathbb{Z}$$

Za výchozí BPB můžeme vzít $[x_1, x_2, x_3, s_1, s_2, s_3] = [0,0,0,4,2,3]$ a s ním se pustit do řešení přidružené úlohy lineárního programování, viz Tabulku 3.7.

c_B	B	x_1	x_2	x_3	s_1	s_2	s_3	x_B	x_B/λ
0	s_1	2	1	−1	1	0	0	4	✖
0	s_2	4	−3	0	0	1	0	2	✖
0	s_3	−3	2	①	0	0	1	3	3 →
	c	−1	3	−3	0	0	0		
	z_k	0	0	0	0	0	0		
	$z_k - c_k$	1	−3	3 ↑	0	0	0	0	

c_B	B	x_1	x_2	x_3	s_1	s_2	s_3	x_B	x_B/λ
	s_1	-1	3	0	1	0	1	7	✖
	s_2	(4)	-3	0	0	1	0	2	1/2 →
	x_3	-3	2	1	0	0	1	3	✖
	$z_k - c_k$	10 ↑	-9	0	0	0	-3	-9	
	s_1	0	9/4	0	1	1/4	1	15/2	
	x_1	1	$-3/4$	0	0	1/4	0	1/2	
	x_3	0	$-1/4$	1	0	3/4	1	9/2	
	$z_k - c_k$	0	$-3/2$	0	0	$-5/2$	-3	-14	

Tabulka 3.7: Řešení přidružené úlohy lineárního programování pro Příklad 3.5.8.

Tak nalezneme řešení $[x_1, x_2, x_3] = [1/2, 0, 9/2]$, které ale nesplňuje podmínky celočíselnosti. Z prvního i druhého řádku poslední simplexové tabulky získáme zlomkový řez

$$x_2/4 + s_2/4 \geq 1/2 \quad \text{neboli} \quad -x_1 + x_2 \geq 0, \tag{3.5.12}$$

zatímco ze třetího řádku dostaneme

$$3x_2/4 + 3s_2/4 \geq 1/2 \quad \text{neboli} \quad -3x_1 + 3x_2/2 \geq 0. \tag{3.5.13}$$

Spočítáme-li odhady změny funkční hodnoty, obdržíme

$$P_{s_1} = P_{x_1} = \frac{1}{2}\max\left\{\frac{-3/2}{1/4}, \frac{-5/2}{1/4}\right\} = -3 \quad \text{a} \quad P_{x_3} = \frac{1}{2}\max\left\{\frac{-3/2}{3/4}, \frac{-5/2}{3/4}\right\} = -1,$$

takže největší nárůst funkční hodnoty bychom mohli očekávat u řezu získaného ze zdrojového řádku odpovídajícího proměnné s_1 nebo x_1, tj. z prvního nebo druhého řádku Tabulky 3.7. A to se i skutečně stane. Samotný řez (3.5.12) dokonce stačí k tomu, abychom po jeho přidání nalezli celočíselné řešení $[x_1^*, x_2^*, x_3^*] = [2, 2, 5]$ s hodnotou $f^* = -11$ po jednom kroku duální simplexové metody, jak můžeme vidět v Tabulce 3.8. Na druhou stranu s pomocí řezu (3.5.13) získáme pouze neceločíselné řešení $[1, 2/3, 14/3]$ s hodnotou -13, a tedy bude nutné přidat ještě jeden další řez – ten už by měl stačit, ale tento alternativní výpočet ponecháme na čtenáři.

c_B	B	x_1	x_2	x_3	s_1	s_2	s_3	s_4	x_B
0	s_1	0	9/4	0	1	1/4	1	0	15/2
-1	x_1	1	$-3/4$	0	0	1/4	0	0	1/2
-3	x_3	0	$-1/4$	1	0	3/4	1	0	9/2
0	s_4	0	(-1/4)	0	0	$-1/4$	0	1	$-1/2$ →
	c	-1	3	-3	0	0	0	0	
	z_k	-1	3/2	-3	0	$-5/2$	-3	0	
	$z_k - c_k$	0	$-3/2$	0	0	$-5/2$	-3	0	-14
	$\frac{z_k - c_k}{\lambda}$	✖	6 ↑	✖	✖	10	✖	✖	

c_B	B	x_1	x_2	x_3	s_1	s_2	s_3	s_4	x_B
	s_1	0	0	0	1	−2	1	9	3
	x_1	1	0	0	0	1	0	−3	2
	x_3	0	0	1	0	1	1	−1	5
	x_2	0	1	0	0	1	0	−4	2
	$z_k - c_k$	0	0	0	0	−1	−3	−6	−11

Tabulka 3.8: Řešení přidružené úlohy lineárního programování pro Příklad 3.5.8 po přidání řezu (3.5.12).

Zbývá nám vyřešit poslední problém spojený s metodou řezů: ukázat, že Gomoryho metoda (na rozdíl od té Dantzigovy) je konečně kroková. K tomu jsou nutné dvě drobné modifikace popsaného algoritmu. Tou první je, že mezi jednotlivá omezení je zakomponována i samotná účelová funkce, tj. místo (ℤLP) s minimalizací účelové funkce $c_1 x_1 + \cdots + c_n x_n$ řešíme úlohu

$$x_0 \to \min \quad \& \quad x_0 - c_1 x_1 - \ldots - c_n x_n = 0$$

doplněnou o všechna ostatní původní omezení. Druhou podmínkou je využití tzv. *lexikografické duální metody*, která především zabraňuje možnému vzniku cyklů v duálním simplexovém algoritmu, viz také anticyklická pravidla zmíněná v Podkapitole 2.8 a Podkapitole 2.10. To se sice zdá být v rozporu s našimi řešeními předchozích dvou příkladů, ale není tomu tak. Využití této metody je pouze technickým požadavkem umožňující snazší důkaz konečnosti této metody, jak říká i samotný Gomory na [230, str. 287]:

> *In these proofs we will use the lexicographical dual simplex method described in Section 7. It is not implied that this simplex method need be used in practice or that it is necessary to the proof. It is simply that its use in the proof has reduced the original rather long and tedious proofs to relatively simple ones.*[133]

My si i jistou modifikaci tohoto jednoduššího důkazu za chvíli ukážeme, ale sluší se dodat, že onen zmíněný zdlouhavý a nudný důkaz Gomory asi ani nikde neprezentoval. Na druhou stranu v [230, Section 7] nabízí Gomory dva možné důkazy[134] a oba jsou založeny právě na použití lexikografické duální metody. V literatuře lze najít i další varianty důkazu konečnosti Gomoryho metody, ale většina z nich postrádá přesné vysvětlení některých nuancí, což nepochybně souvisí i s nutností velmi jemného[135] nastavení celé metody, viz např. [98, str. 213–216], [352, str. 165–167], [395, str. 372–373], [411, str. 285–286] a [448, str. 354–358].

V lexikografické simplexové metodě se postupuje velmi podobně jako v klasickém duálním simplexovém algoritmu s tím, že je doplněn o „sloupcové“ lexikografické pravidlo upřesňující volbu klíčového sloupce. Pro správnou funkčnost je však potřeba řádek $z_k - c_k$ nahradit rozdílem $c_k - z_k$ a současně jej přesunout na první pozici pod záhlaví simplexové tabulky, tj. nad řádky odpovídající bazickým proměnným. V případě nedegenerované duálně přípustné úlohy (tj. $z_k - c_k < 0$ pro všechny nebazické proměnné) budou všechny sloupce *lexikograficky kladné*[136] vyjma sloupce označeného x_B.

DEFINICE 3.5.9

Vektor $\alpha \in \mathbb{R}^n$ je *lexikograficky kladný*[137], jestliže první nenulový koeficient je kladný, tj. existuje $k \in \{1,\ldots,n-1\}$ takové, že $\alpha_i = 0$ pro $i \in \{1,\ldots,k-1\}$ a $\alpha_k > 0$. Toto zapisujeme jako $\alpha \succ 0$. Vektor $\beta \in \mathbb{R}^n$ je *lexikograficky záporný*, tj. $\beta \prec 0$, jestliže $-\beta \succ 0$. Řekneme, že vektor α je *lexikograficky větší než* β, jestliže $\alpha - \beta \succ 0$.

Např. $\alpha = (0,0,1,-3,0) \succ 0$, $\beta = (-1,0,2,3,1) \prec 0$ a $\alpha \succ \beta$, protože $\alpha - \beta = (1,0,-1,-6,-1) \succ 0$.

[133] *V těchto důkazech použijeme lexikografickou duální simplexovou metodu popsanou v Sekci 7. Neznamená to, že tuto simplexovou metodu je nutné používat v praxi nebo že je nutná k důkazu. Je to jednoduše proto, že její použití v důkazu zredukovalo původní poměrně dlouhé a nudné důkazy na relativně jednoduché.*

[134] V článku [227] žádný důkaz nenajdeme – v tomto prvním článku o zlomkových řezech Gomory pouze celou metodu představil a slíbil, že důkaz konečnosti takového algoritmu předloží samostatně.

[135] A poměrně nešťastného...

[136] Někdy se také hovoří o tom, že taková tabulka nebo příslušná báze B je *lexikograficky optimální*.

[135] Označení „lexikografický“ odkazuje na analogii s uspořádáním slov ve slovníku. Kdybychom každému slovu přiřadili vektor, jehož složky by odpovídaly pořadí jednotlivých písmen v české abecedě a 0 by značila chybějící písmena, tj. např. slovo „dobrý“ bychom zapsali jako $(5,19,2,22,32,0\ldots)$, pak by lexikografické uspořádání těchto vektorů bylo shodné s pořadím slov ve slovníku (čím lexikograficky menší, tím dříve je slovo ve slovníku).

Zmíněná lexikografická pozitivita nemusí být splněna v případě, kdy duální úloha je degenerovaná, v kterémžto případě se může stát $z_k - c_k = 0$ i pro nebazické proměnné. Naštěstí tomu lze snadno zabránit přidáním podobného umělého omezení jako v případě hledání duálně přípustného výchozího bodu, tj. přidáním dodatečné podmínky

$$\sum_{j \in N_0} x_j \le M$$

pro dostatečně velké číslo M, kde N_0 je množina nebazických indexů, pro které platí[138] $z_k - c_k = 0$. Pokud toto omezení zapíšeme hned jako další řádek po $c_k - z_k$, pak v odpovídající tabulce budou jistě všechny sloupce lexikograficky kladné vyjma x_B. Na takto upravenou simplexovou tabulku pak aplikujeme duální algoritmus s lexikografickým pravidlem pro výběr klíčového sloupce, podle kterého v případě realizace minima

$$-\frac{c_k - z_k}{\lambda_{rk}} = \min\left\{ -\frac{c_j - z_j}{\lambda_{rj}} \;\middle|\; j \in \{1, \dots, n\} \setminus B \;\; \& \;\; \lambda_{rj} < 0 \right\} \ge 0.$$

ve více indexech volíme ten, který po podělení $-\lambda_{rj}$ dává lexikograficky nejmenší sloupec[139]. Díky tomu dostaneme v dalším kroku opět tabulku s lexikograficky kladnými sloupci vyjma x_B, který však bude lexikograficky klesající[140], tudíž nemůže dojít k zacyklení a metoda nalezne řešení v konečně mnoha krocích.

V souladu s popisem Gomoryho metody řešíme v jednotlivých krocích přidruženou úlohu

$$\left.\begin{array}{c} -x_0 \to \min \\ x_0 + c_1 x_1 + \dots + c_n x_n = 0 \quad \& \quad A^{[k]} x^{[k]} = b^{[k]} \quad \& \quad x^{[k]} \ge 0, \end{array}\right\} \tag{3.5.14}$$

kde $x^{[k]}$ je $(n+k+1)$-rozměrný vektor složený z x_0 odpovídající hodnotě účelové funkce s opačným znaménkem, původních proměnných $x_1, \dots, x_n$ a pomocných proměnných $x_{n+1}, \dots, x_{n+k}$, které jsou postupně přidávány v jednotlivých krocích kvůli převedení Gomoryho zlomkových řezů na rovnosti. Podobně $A^{[k]}$ je $(m+k) \times (n+k+1)$ matice, jejíž první sloupec je nulový (x_0 se v těchto omezeních neobjevuje) a následovaný $m \times n$ submaticí odpovídající upravené původní matici A, zatímco ostatní sloupce a řádky odpovídají koeficientům z jednotlivých Gomoryho řezů. Analogickým způsobem získáme vektor $b^{[k]} \in \mathbb{R}^{m+k}$. Index $k = 0$ odpovídá původní přidružené úloze lineárního programování, a tedy platí $A^{[0]} = (0, A) \in \mathbb{R}^{m \times (n+1)}$ a $b^{[0]} = b$.

VĚTA 3.5.10

Uvažme úlohu (ZLP) s celočíselnými koeficienty a nechť její účelová funkce je zdola ohraničená na přípustné množině[141]. Pak Gomoryho metoda zlomkových řezů aplikovaná na ekvivalentní úlohu (3.5.14) a doplněná o následující pravidla

(i) omezení $x_0 + c_1 x_1 + \dots + c_n x_n = 0$ je ve výchozí simplexové tabulce na prvním řádku;

(ii) v každém kroku je přidáno jen jedno omezení z Gomoryho zlomkového řezu, které je do poslední simplexové tabulky z předchozího kroku implementováno na poslední pozici[142];

(iii) první krok je realizován pomocí klasické simplexové metody a v dalších krocích je využita lexikografická duální simplexová metoda;

(iv) jako zdrojový řádek je vždy zvolen ten první (tj. ten nejvýše v simplexové tabulce), pro který je $r_k > 0$ neboli odpovídající bazická proměnná má neceločíselnou hodnotu;

nalezne řešení v konečném počtu kroků.

Důkaz. V průběhu výpočtu jsou v jednotlivých krocích generovány simplexové tabulky s bazickými body $w^{[0]}, w^{[1]}, \dots, w^{[s_0]} =: x^{[0]*}, w^{[s_0+1]}, w^{[s_0+2]}, \dots, w^{[s_1]} =: x^{[1]*}, \dots, w^{[s_1]} \dots$, z nichž zejména $w^{[0]}, w^{[1]}, \dots, w^{[s_0-1]}, x^{[0]*}$ jsou primárně přípustné[143] díky využití klasické simplexové metody v prvním kroku, zatímco dále už jen $x^{[1]*}, x^{[2]*}, \dots$ jsou primárně přípustné jako řešení

[138] Případně sem lze zahrnout všechny nebazické proměnné.

[139] To je podobné jako v případě lexikografického pravidla pro primární metodu z Podkapitoly 2.8. Určíme podíly $-\frac{c_j - z_j}{\lambda_{rj}}$. Pokud se minimum realizuje ve více sloupcích, pak celý sloupec podělíme $-\lambda_{rk}$ a porovnáváme jednotlivé hodnoty v těchto sloupcích od shora dolů. Lexikograficky nejmenší bude ten, který má na první odlišné pozici menší číslo.

[140] Vskutku, nejdříve musíme ukázat, že volba klíčového sloupce podle lexikografického pravidla je určena jednoznačně. K tomu můžeme využít toho, že simplexovou tabulku rozšíříme o triviální omezení $-x_j = 0$ pro nebazické indexy j. Tím rozšíříme „nebazickou část" simplexové tabulky o matici $-I$, což nám zaručí lineární nezávislost příslušných sloupců, a tudíž i jednoznačnou volbu klíčového sloupce.

Nechť výsledným klíčovým prvkem je koeficient λ_{rs}. Pak pro libovolné $k \in \{1, \dots, n\}$ jsou koeficienty v nové simplexové tabulce jsou určeny vztahem

$$\hat{\lambda}_{jk} = \begin{cases} \lambda_{jk} - \lambda_{js} \frac{\lambda_{rk}}{\lambda_{rs}}, & j \in B \setminus \{r\}, \\ \frac{\lambda_{rk}}{\lambda_{rs}}, & j = s, \end{cases}$$

a konstrukce duální simplexové metody zaručuje, že v řádku $c_k - z_k$ budou pouze nezáporné hodnoty (jinak by došlo k porušení duální přípustnosti). Ukážeme nyní, že i v nové simplexové tabulce budou všechny sloupce lexikograficky kladné. To bude nepochybně pravda pro sloupce odpovídající novým bazickým proměnným, tj. pro indexy z $B \setminus \{r\} \cup \{s\}$, kde budou pouze nuly a jedna jednička. Protože $\lambda_{rs} < 0$ a všechny původní sloupce $\boldsymbol{\alpha}_k$ jsou lexikograficky kladné, bude v případě $\lambda_{rk} \ge 0$ celý nový sloupce $\hat{\boldsymbol{\alpha}}_k$ také lexikograficky kladný jakožto součet dvou lexikograficky kladného řádku a lexikograficky nezáporného řádku (on bude také kladný vyjma situace $\lambda_{rk} = 0$). Je-li $\lambda_{rk} < 0$ a hodnota podílu $-\frac{c_k - z_k}{\lambda_{rk}}$ v k-tém sloupci

(i) je větší než možné minimum, pak

$$-\frac{c_k - z_k}{\lambda_{rk}} > -\frac{c_s - z_s}{\lambda_{rs}}$$

neboli

$$0 < (c_k - z_k) - \frac{\lambda_{rk}}{\lambda_{rs}}(c_s - z_s),$$

tj. v prvním řádku k-tého

sloupce v nové tabulce bude kladné číslo neboli $\widehat{\boldsymbol{\alpha}}_k > 0$;

(ii) má stejnou hodnotu jako možné minimum, pak z volby lexikograficky nejmenšího sloupce bud v nové tabulce v k-tém sloupci několik nulových prvků následovaných kladným číslem (podle toho, na kolikáté pozici se sloupce podělené odpovídajícím prvkem z klíčového řádku liší), tj. opět $\widehat{\boldsymbol{\alpha}}_k > 0$.

Zbývá ukázat, že sloupec $\boldsymbol{\beta}$ odpovídající sloupci x_B v simplexové tabulce je lexikograficky klesající. Jednotlivé složky sloupce $\widehat{\boldsymbol{\beta}}$ v nové simplexové tabulce dostaneme ze vztahu

$$\widehat{\beta}_j = \begin{cases} \beta_j - \frac{\beta_r}{\lambda_{rs}}\lambda_{js}, & j \in B\setminus\{r\}, \\ \frac{\beta_r}{\lambda_{rs}}, & j = s. \end{cases}$$

Očividně $\widehat{\beta}_j = \beta_j$ pro všechna $j \in \{1,\dots,n\}$ taková, že $\lambda_{js} = 0$. Současně z $\boldsymbol{\alpha}_s > 0$ plyne, že první nenulový koeficient λ_{js} musí být kladný, takže díky $\lambda_{rs} < 0$ a $\beta_r < 0$ pro tento index j platí $\widehat{\beta}_j < \beta_j$. Proto $\widehat{\boldsymbol{\beta}} \prec \boldsymbol{\beta}$. K otestování této metody může velmi dobře posloužit Bealeho příklady z Poznámky 195 na straně 177.

[141] Tento požadavek je zcela bez újmy na obecnosti. Pokud totiž má úloha řešení, pak se realizuje v konečném bodě, a tudíž se při hledání řešení můžeme omezit pouze na vhodnou ohraničenou podmnožinu přípustné množiny. Pokud úloha není ohraničená, pak to zjistíme ihned při řešení výchozí přidružené úlohy.

[142] Tím je samozřejmě míněno nad řádek $z_k - c_k$.

[143] Bod $w^{[0]}$ je výchozím BPB, který získáme buď přímo ze zadání (je-li to možné) nebo např. pomocí dvoufázové metody

[144] Kromě koeficientu 1 u nově přidané proměnné, jsou v tomto řádku pouze nulové koeficienty nebo záporné násobky zlomkových částí jednotlivých koeficientů ze zdrojového řádku u nebazických proměnných a záporná hodnota $-\zeta(w_0^{[s_q]})$ v posledním sloupci.

[145] Menší být nemůže kvůli tomu, že jde o limitu nerostoucí posloupnosti, a větší být nemůže kvůli platnosti lexikografického uspořádání.

přidružených úloh lineárního programování získaná pomocí lexikografické duální simplexové metody. Pokud v prvním kroku získáme řešením přidružené úlohy (ještě bez přidání jakéhokoli řezu) celočíselné řešení $x^{[0]*}$, pak jsme hotovi a tvrzení je pravdivé. V opačném případě odvodíme Gomoryho zlomkový řez a dále už pokračujeme jenom s využitím lexikografické duální simplexové metody. Díky tomu budou odpovídající sloupce „x_B“ lexikograficky klesající, tj. pro všechna $k \geq s_0$ platí

$$\begin{pmatrix} w_0^{[k]*} & w_1^{[k]*} & \cdots & w_n^{[k]*} \end{pmatrix}^\top \succ \begin{pmatrix} w_0^{[k+1]*} & w_1^{[k+1]*} & \cdots & w_n^{[k+1]*} \end{pmatrix}^\top.$$

Ukážeme, že algoritmus nemůže běžet nekonečně dlouho, tj. že jednotlivé komponenty $x_0^{[k]*}$, $\dots, x_n^{[k]*}$ budou všechny celočíselné pro konečné $k \in \mathbb{N}$. Pro jednoduchost dalšího značení si k tomu zaveďme označení zlomkové části libovolného čísla $a \in \mathbb{R}$ jako $\zeta(a)$, tj. $\zeta(a) := a - \lfloor a \rfloor$.

Uvažme nejdříve posloupnost $\{x_0^{[k]*}\}$. Vzhledem k tomu, že přípustné množiny přidružených úloh jsou postupně zmenšovány a současně vždy obsahují celou přípustnou množinu původní celočíselné úlohy, musí platit

$$-x_0^{[0]*} = \langle c, x^{[0]*}\rangle \leq -x_0^{[1]*} = \langle c, x^{[1]*}\rangle \leq \cdots \leq \langle c, x^*\rangle$$

pro hledané celočíselné řešení x^* neboli

$$x_0^{[0]*} \geq x_0^{[1]*} \geq \cdots \geq -\langle c, x^*\rangle > -\infty.$$

To ukazuje, že posloupnost $\{x_0^{[k]*}\}$ je nerostoucí a zdola ohraničená, a tudíž musí mít konečnou limitu $x_0^{[\infty]*}$. Proto se nepochybně po dostatečně velkém počtu kroků Gomoryho metody dostaneme s hodnotou $x_0^{[k]*}$ tak blízko $x_0^{[\infty]*}$, že budou mít stejnou celočíselnou část, tj. nechť $q \in \mathbb{N}$ je takový index, že řešení $x^{[q]*} := w^{[s_q]}$ splňuje $\lfloor x_0^{[q]*} \rfloor = \lfloor x_0^{[\infty]*} \rfloor$, a tedy

$$0 \leq x_0^{[q]*} - \lfloor x_0^{[\infty]*} \rfloor < 1.$$

Je-li dokonce $x_0^{[q]*} = \lfloor x_0^{[\infty]*} \rfloor$, pak můžeme pokračovat složkou $x_1^{[k]*}$. V opačném případě je $\zeta(x_0^{[q]*}) > 0$ a v souladu s popsanými pravidly přidáme v následujícím kroku Gomoryho zlomkový řez určený zdrojovým řádkem odpovídajícím právě x_0. V důsledku duální simplexové metody na začátku dalšího kroku z nutně báze vystoupí nově přidaná proměnná $x_{n+q+2}^{[q+1]}$, která způsobuje primární nepřípustnost (jako jediná má zápornou hodnotu) a která bude nahrazena nějakou nebazickou proměnnou $x_p^{[q+1]}$ se záporným koeficientem $-\zeta(\lambda_{0p}^{[q]}) < 0$ v tomto klíčovém řádku[144]. Pak hodnotu $w_0^{[s_q+1]}$ získáme ze vzorce

$$w_0^{[s_q+1]} = w_0^{[s_q]} - \zeta(w_0^{[s_q]})\frac{-\lambda_{0p}^{[q]}}{-\zeta(\lambda_{0p}^{[q]})}. \tag{3.5.15}$$

Protože $\zeta(\lambda_{0p}^{[q]}) > 0$, $\zeta(w_0^{[s_q]}) > 0$ a $\lambda_{0p}^{[q]} \neq 0$, musí být $\lambda_{0p}^{[q]} > 0$, neboť jinak bychom měli $w_0^{[s_q+1]} > w_0^{[s_q]}$, čímž bychom dostali spor s lexikografickým uspořádáním. V takovém případě je

$$\lambda_{0p}^{[q]} \geq \zeta(\lambda_{0p}^{[q]}) \quad \text{neboli} \quad \frac{\lambda_{0p}^{[q]}}{\zeta(\lambda_{0p}^{[q]})} \geq 1,$$

a tudíž z (3.5.15) plyne

$$\lfloor x_0^{[\infty]*} \rfloor \leq x_0^{[\infty]*} \leq w_0^{[s_q+1]} \leq w_0^{[s_q]} - \zeta(w_0^{[s_q]}) = \lfloor w_0^{[s_q]} \rfloor = \lfloor x_0^{[\infty]*} \rfloor,$$

což je ale možné pouze v případě, že $w_0^{[s_q+1]} = \lfloor x_0^{[\infty]*} \rfloor$. Tím nerostoucí posloupnost $\{x_0^{[k]*}\}$ dosáhla své limitní (celočíselné) hodnoty v konečně mnoha krocích a vzhledem k lexikografickému pravidlu již ve všech dalších iteracích bude mít tuto hodnotu[145]

Nyní se podíváme na posloupnost $\{x_1^{[k]*}\}$ pro $k \geq s_q + 1$. Opět se jedná o nerostoucí posloupnost, která je tentokrát zdola ohraničená 0, protože kdyby v nějaké simplexové tabulce nabyla

tato proměnná zápornou hodnotu, pak bychom již nikdy v důsledku lexikografického pravidla nedostali primárně přípustný bod. Nutně tedy také tato posloupnost má limitu. Pak s pomocí naprosto stejných argumentů jako výše zjistíme, že po konečně mnoha dalších krocích bude mít proměnná x_1 celočíselnou hodnotu. Analogickým pokračováním v těchto úvahách pro proměnné $x_2, \ldots, x_n$ zjistíme, že po dostatečně velkém počtu kroků Gomoryho metody dosáhneme celočíselných hodnot i pro tyto proměnné. Abychom dokončili celý algoritmus, zbývá už jen dopočítat řešení poslední přidružené úlohy, během níž se nám podařilo získat limitní hodnotu pro x_n. ■

Z historického pohledu představuje Gomoryho metoda první nástroj pro řešení úloh celočíselného programování. Avšak navzdory předchozímu důkazu konečnosti celé metody se sám Gomory stejně jako řada dalších odborníků v této oblasti v 50. letech minulého století domnívala, že není příliš praktická vzhledem k numerické nestabilitě[146] a také možnému vysokému počtu řezů potřebných pro nalezení řešení. Proto Dantzig nabídl svoji dříve zmíněnou alternativní metodu řezů, která však už vůbec nezaručovala konvergenci. Jako mnohem efektivnější se ale posléze ukázal postup založený na metodě větvení a mezí. Ovšem v polovině 90. let 20. století se podařilo ukázat, že skutečně účinný nástroj je jakýmsi kompromisem obou těchto metod, který je znám jako *Branch-and-Cut* a poprvé se pod tímto názvem objevuje nejspíše v článku [406] z roku 1987, ve kterém jsou testovány možnosti řešení symetrické úlohy obchodního cestujícího s 532 městy a 141 246 binárními proměnnými, viz také [407]. V této metodě jsou řezy využity k získání lepších odhadů minima než dává samotná metoda větvení a mezí a zároveň dochází k potlačení numerické nestability, neboť řezy jsou využity pouze při řešení subproblémů získaných větvením a v případě obtíží s konvergencí simplexové metody dochází k dalšímu větvení aktuálního subproblému. Za potvrzení (teoretické) efektivnosti této metody vděčíme zejména Cornuéjolsovi[147] a jeho spolupracovníkům, viz [23, 24, 94, 407]. Úspěch této kombinace neunikl pozornosti ani zakladatelkám metody větvení a mezí a v článku z roku 2007 k tomu poznamenaly

> *While branch and bound began to be built into computer codes, the cutting plane approach was obviously more elegant, and we spent a great deal of time experimenting with it… Work was done, but it was not published because as a method to solve problems branch and bound resoundingly won.*[148]

> *It is gratifying that the combination, „branch and cut", is now often successful in dealing with real problems.*[149]

Viz také [100] a pro další metody řešení úloh celočíselného programování nebo alternativní možnosti řezných rovin či jejich hlubší souvislosti viz např. [98, 286, 395, 408, 448, 472].

To, že představená metoda Gomoryho zlomkových řezů nefunguje pro úlohy smíšeného celočísleného programování, jsme již zmínili. Situace však není zcela bezvýchodná. Uvažme tentokrát úlohu

$$\left.\begin{array}{c} \langle c, x\rangle \to \min \\ Ax = b \quad \& \quad x \geq 0 \\ x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z} \quad \& \quad \cdots \quad \& \quad x_s \in \mathbb{Z} \end{array}\right\} \qquad \text{(sZLP)}$$

pro nějaké[150] $s \in \{1, \ldots, n-1\}$. Stejně jako v čistě celočíselném případě začneme se zdrojovým řádkem

$$x_\ell + \sum_{j \in N} \lambda_{\ell j} x_j = \widehat{b}_\ell \qquad (3.5.16)$$

pro nějakou proměnnou x_ℓ s indexem $\ell \in \{1, \ldots, s\}$ mající neceločíselnou hodnotu v řešení přidružené úlohy lineárního programování. Rozdělíme-li množinu nebazických proměnných N na dvě části $N_{\mathbb{Z}} := N \cap \{1, \ldots, s\}$ (zahrnující indexy nebazických celočíselných proměnných)

[146] Ta je způsobena zaokrouhlovacími chybami, které mohou mimo jiné činit potíže i s rozhodnutím, zda nalezená hodnota proměnné je celočíselná či nikoli. Pro překonání tohoto nedostatku byly vyvinuty také metody, které lze označit za *celo-celočíselné* (angl. *all-integer*). V nich jsou simplexové tabulky upravovány tak, aby odpovídaly primárně přípustným bodům a současně všechny koeficienty v tabulkách zůstaly celočíselné po celou dobu výpočtu. Tím se při upravování vyhneme zlomkům, a tudíž již nebude docházet k potížím způsobeným zaokrouhlováním, takže v případě zastavení výpočtu získáme přípustný celočíselný bod, který (snad) nebude příliš vzdálen od skutečného minima. Viz např. [49, 50, 219, 231, 514, 515].

[147] Gérard Pierre Cornuéjols (∗ 16. listopadu 1950) je profesorem operačního výzkumu na Univerzitě Carnegieho–Mellonových. Jeho výzkum je zaměřen především na „problém rozmístění" (facility location), celočíselné programování, „vyvážené matice" (balanced matrices) a perfektní grafy. V letech 1999–2003 byl šéfredaktorem časopisu *Mathematics of Operations Research*. V roce 2011 obdržel *John von Neumann Theory Prize* za „zásadní a výrazné příspěvky k diskrétní optimalizaci, včetně hlubokého výzkumu vyvážených a ideálních matic, dokonalých grafů a řezných rovin pro úlohy smíšeného celočíselného programování".

[148] *Zatímco metoda větvení a mezí začala být implementována do počítačových kódů, přístup založený na řezných rovinách byl samozřejmě elegantnější a my jsme experimentováním s ním strávily spoustu času… Práce byla hotová, ale nebyla publikována, protože jako nástroj k řešení problémů drtivě vyhrála metoda větvení a mezí.* Viz [347, str. 149].

[149] *Je potěšující, že kombinace „větvení a řezů" je nyní často úspěšná při řešení skutečných problémů.* Viz [347, str. 149].

[150] Bez újmy na obecnosti požadujeme celočíselnost pro prvních s proměnných. Volba $s = 0$ by vedla na normální úlohu lineárního programování, zatímco $s = n$ by byla klasická úloha celočíselného programování.

a $N_{\mathbb{R}} := N \setminus N_{\mathbb{Z}}$ (zahrnující indexy nebazických proměnných bez požadavku na celočíselnost). Rozdělením koeficientů u požadovaných celočíselných proměnných na jejich celočíselné a zlomkové části dostaneme

$$\sum_{j\in N_{\mathbb{Z}}} \zeta(\lambda_{\ell j})\, x_j + \sum_{j\in N_{\mathbb{R}}} \lambda_{\ell j}\, x_j - \zeta(\widehat{b}_\ell) = \lfloor \widehat{b}_\ell \rfloor - \sum_{j\in N_{\mathbb{Z}}} \lfloor \lambda_{\ell j} \rfloor\, x_j - x_\ell \tag{3.5.17}$$

kde opět využíváme označení $\zeta(a) := a - \lfloor a \rfloor$. Pro libovolný přípustný bod úlohy (sZLP) je hodnota pravé strany rovnosti (3.5.17) nutně celočíselná, a tudíž i levá strana musí být celočíselná. Zejména platí buď

$$\sum_{j\in N_{\mathbb{Z}}} \zeta(\lambda_{\ell j})\, x_j + \sum_{j\in N_{\mathbb{R}}} \lambda_{\ell j}\, x_j - \zeta(\widehat{b}_\ell) \geq 0$$

nebo

$$\sum_{j\in N_{\mathbb{Z}}} (-\zeta(\lambda_{\ell j}))\, x_j + \sum_{j\in N_{\mathbb{R}}} (-\lambda_{\ell j})\, x_j + \zeta(\widehat{b}_\ell) \geq 1.$$

Vzhledem k požadavku $x \geq 0$ pro libovolný přípustný bod nemohou členy s nulovými či zápornými koeficienty na levých stranách těchto nerovnosti „pomoci“ k jejich splnění. Je patrné, že toto se týká obzvláště všech nebazických celočíselných proměnných ve druhé nerovnosti a některých koeficientů u proměnných bez požadavku na celočíselnost, pro které si zavedeme označení

$$N_{\mathbb{R}}^{+} := \{j \in N_{\mathbb{R}} \mid \lambda_{\ell j} > 0\} \quad \text{a} \quad N_{\mathbb{R}}^{-} := \{j \in N_{\mathbb{R}} \mid \lambda_{\ell j} < 0\}.$$

To nás přivede k nové dvojici nerovností

$$\sum_{j\in N_{\mathbb{Z}}} \zeta(\lambda_{\ell j})\, x_j + \sum_{j\in N_{\mathbb{R}}^{+}} \lambda_{\ell j}\, x_j \geq \zeta(\widehat{b}_\ell) \quad \text{a} \quad \sum_{j\in N_{\mathbb{R}}^{-}} (-\lambda_{\ell j})\, x_j \geq 1 - \zeta(\widehat{b}_\ell). \tag{3.5.18}$$

U druhé nerovnosti vyžijeme fakt, že $\zeta(\widehat{b}_\ell) \in (0,1)$, takže ji můžeme upravit do tvaru

$$\frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)} \sum_{j\in N_{\mathbb{R}}^{-}} (-\lambda_{\ell j})\, x_j \geq \zeta(\widehat{b}_\ell). \tag{3.5.19}$$

To ukazuje, že pro libovolný přípustný bod úlohy (sZLP) je buď levá strana nerovností v (3.5.18) větší než nebo rovna hodnotě $\zeta(\widehat{b}_\ell)$ a/nebo to splňuje levá strana nerovnosti (3.5.19). Ať už platí jedno nebo druhé, určitě tuto vlastnost bude mít součet levých stran obou těchto nerovností, čímž jsme dokázali následující tvrzení, které dává nutnou podmínku pro všechny přípustné body (sZLP). Gomory tento výsledek prezentoval ve zprávě [228] z roku 1960.

Věta 3.5.11 Gomoryho smíšený zlomkový řez

Uvažme úlohu (sZLP) a řešení $\widehat{x}$ přidružené úlohy lineární programování. Jestliže tento bod nesplňuje pro nějaký index $\ell \in \{1,\ldots,s\}$ podmínku celočíselnosti a řádek v poslední simplexové tabulce odpovídající této bazické proměnné má podobu (3.5.16), pak bod $\widehat{x}$ nesplňuje ani příslušný Gomoryho smíšený zlomkový řez

$$\sum_{j\in N_{\mathbb{Z}}} \zeta(\lambda_{\ell j})\, x_j + \sum_{j\in N_{\mathbb{R}}^{+}} \lambda_{\ell j}\, x_j + \frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)} \sum_{j\in N_{\mathbb{R}}^{-}} (-\lambda_{\ell j})\, x_j \geq \zeta(\widehat{b}_\ell), \tag{3.5.20}$$

kterému však vyhovují všechny přípustné body původní úlohy.

Nerovnost (3.5.20) je zobecněním původního Gomoryho zlomkového řezu, neboť se v případě $N_{\mathbb{R}} = \emptyset$ redukuje přesně na (3.5.7). Výpočet Gomoryho smíšeného zlomkového řezu si ilustrujeme v následujícím příkladě, viz [411, str. 288–290].

Příklad 3.5.12

Uvažme, že v úloze (sZLP) máme $n = 20$ a $s = 13$. Při řešení přidružené úlohy lineárního programování jsme nalezli řešení s bazickou proměnnou $x_{12} = 17/5$ a odpovídajícím zdrojovým řádkem

$$-\frac{11}{5}x_3 + \frac{3}{5}x_7 + 5x_8 + \frac{6}{5}x_{10} + x_{12} - 4x_{14} + \frac{6}{5}x_{15} - \frac{2}{5}x_{20} = 17/5.$$

Pak[151] $N_{\mathbb{Z}} = \{3,7,8,10\}$, $N^+_{\mathbb{R}} = \{15\}$ a $N^-_{\mathbb{R}} = \{14,20\}$. Odpovídající Gomoryho smíšený zlomkový řez (3.5.20) má podobu

$$\frac{4}{5}x_3 + \frac{3}{5}x_7 + \frac{1}{5}x_{10} + \frac{6}{5}x_{15} + \frac{2/5}{1-2/5}\left(4x_{14} + \frac{2}{5}x_{20}\right) \geq \frac{2}{5}.$$

neboli

$$\frac{4}{5}x_3 + \frac{3}{5}x_7 + \frac{1}{5}x_{10} + \frac{6}{5}x_{15} + \frac{8}{3}x_{14} + \frac{4}{15}x_{20} \geq \frac{2}{5}.$$

[151] Nezapomínejme, že v tomto řádku mohou být nenulové koeficienty pouze u nebazických proměnných a u x_{12}.

Smíšený Gomoryho zlomkový řez (a vlastně i ten klasický) lze ještě i mírně zesílit. Při odvození (3.5.20) jsme využili dolní celou část koeficientu $\lambda_{\ell j}$ a rovnost $\lambda_{\ell j} = \lfloor \lambda_{\ell j} \rfloor + \zeta(\lambda_{\ell j})$. Avšak vezmeme-li místo toho horní celou část, pak $\lambda_{\ell j} = \lceil \lambda_{\ell j} \rceil - \big(1 - \zeta(\lambda_{\ell j})\big)$ a rovnost (3.5.17) se změní na

$$-\sum_{j \in N_{\mathbb{Z}}} \big(1-\zeta(\lambda_{\ell j})\big)x_j + \sum_{j \in N_{\mathbb{R}}} \lambda_{\ell j} x_j - \zeta(\widehat{b}_\ell) = \lfloor \widehat{b}_\ell \rfloor - \sum_{j \in N_{\mathbb{Z}}} \lceil \lambda_{\ell j} \rceil x_j - x_\ell.$$

Stejnými úvahami jako v předchozí části potom dostaneme nerovnost

$$\frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)} \sum_{j \in N_{\mathbb{Z}}} \big(1-\zeta(\lambda_{\ell j})\big)x_j + \sum_{j \in N^+_{\mathbb{R}}} \lambda_{\ell j} x_j + \frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)} \sum_{j \in N^-_{\mathbb{R}}} (-\lambda_{\ell j}) x_j \geq \zeta(\widehat{b}_\ell). \tag{3.5.21}$$

Srovnáním (3.5.20) a (3.5.21) vidíme, že rozdíl je pouze u indexů $j \in N_{\mathbb{Z}}$. Také nám nic nebrání ve využití vztahu $\lambda_{\ell j} = \lceil \lambda_{\ell j} \rceil - \big(1 - \zeta(\lambda_{\ell j})\big)$ pro některá $j \in N_{\mathbb{Z}}$ a $\lambda_{\ell j} = \lfloor \lambda_{\ell j} \rfloor + \zeta(\lambda_{\ell j})$ pro ostatní. Tím můžeme získat různé řezy se stejnou vlastností jako v případě (3.5.20) a zvolit ten nejlepší z nich, tj.

$$\sum_{j \in N_{\mathbb{Z}}} \min\left\{\zeta(\lambda_{\ell j}), \frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)}\big(1-\zeta(\lambda_{\ell j})\big)\right\} x_j + \sum_{j \in N^+_{\mathbb{R}}} \lambda_{\ell j} x_j + \frac{\zeta(\widehat{b}_\ell)}{1-\zeta(\widehat{b}_\ell)} \sum_{j \in N^-_{\mathbb{R}}} (-\lambda_{\ell j}) x_j \geq \zeta(\widehat{b}_\ell). \tag{3.5.22}$$

Jeho využitím v Příkladě 3.5.12 bychom dostali

$$\min\left\{\frac{4}{5}, \frac{2}{3}\times\frac{1}{5}\right\}x_3 + \min\left\{\frac{3}{5}, \frac{2}{3}\times\frac{2}{5}\right\}x_7 + \min\left\{\frac{1}{5}, \frac{2}{3}\times\frac{4}{5}\right\}x_{10} + \frac{6}{5}x_{15} + \frac{2/5}{1-2/5}\left(4x_{14} + \frac{2}{5}x_{20}\right) \geq \frac{2}{5}$$

neboli

$$\frac{2}{15}x_3 + \frac{4}{15}x_7 + \frac{1}{5}x_{10} + \frac{6}{5}x_{15} + \frac{8}{3}x_{14} + \frac{4}{15}x_{20} \geq \frac{2}{5}.$$

Tím došlo ke zmenšení koeficientů u proměnných x_3 a x_7, a získali jsme tak silnější řez[152].

[152] Vyhovuje mu totiž méně bodů s nezápornými koeficienty, viz také Definici 3.5.7.

Aplikováním řezu (3.5.22) ve druhém kroku Gomoryho metody zlomkových řezů dostaneme algoritmus pro úlohy (sZLP). Tentokrát však již pomocná proměnná využitá pro převod nerovnosti (3.5.22) na rovnost nebude celočíselná, a bude tedy prvkem množiny $N_{\mathbb{R}}$ místo $N_{\mathbb{Z}}$. Navíc již nemusí být možné zaručit konečnou krokovost tohoto algoritmu. Bude-li možné předpokládat celočíselnost hodnoty x_0 nahrazující účelovou funkci[153], pak lze konečnou krokovost algoritmu ukázat zcela analogickými argumenty jako v případě Věty 3.5.10. Avšak toto nemusí být pravda vzhledem k využití i neceločíselných hodnot některých proměnných, což může způsobit nekonečnou krokovost celého algoritmu.

[153] Tj., budeme-li moci a priori očekávat, že hledaná minimální hodnota bude celočíselná (třeba když v účelové funkci budou vystupovat pouze celočíselné proměnné).

3.6 Chcete být milionářem?

Jak už jsme dříve naznačili, úlohy spojené s celočíselným programováním mají skutečně velmi široké uplatnění v praxi, a proto snad nepřekvapí, že znalosti a dovednosti spojené s touto oblastí optimalizace mohou mít i (mnoha) milionovou hodnotu. Z tohoto důvodu je toto téma velmi atraktivní pro mnohé společnosti z různých odvětví, avšak trvalo to téměř dekádu, než se podařilo matematické modely celočíselného programování a metody jejich řešení dostat až do reálných průmyslových problémů. Přesné zmapování těchto počátků je ale téměř nemožné, neboť si je pochopitelně jednotlivé společnosti uchovávaly jako součást svého obchodního

tajemství, a my se o nich můžeme něco dozvědět jen z osobních vzpomínek těch, kteří se na tom podíleli. Mezi první společnosti určitě patří British Petroleum a jejich podpora výzkumu věnovaného metodě větvení a mezí. Další společností je Philips Electronics a optimalizace rozmístění jejich továren ve Španělsku, o čemž se zmiňoval Max Shaw. Shaw také jako první publikoval článek o takovém průmyslovém využití – konkrétně šlo o úlohu optimálního rozmístění letadel a lodí armády Velké Británie, což vedlo na problém s 10-20 proměnnými a 850 omezeními, viz [444] a [66].

Výčet dalších aplikací by mohl být velmi dlouhý a skvělým zdrojem k těmto výsledkům je např. časopis *INFORMS Journal on Applied Analytics*, ve kterém lze najít např. následující články popisující

- optimalizaci trasování lodí a přidělení personálu pro nábor do námořnictva v Thajsku, viz [288];
- zvýšení výkonu a flexibility ve společnosti Jeppesen Sanderson, viz [316];
- optimalizaci harmonogramu posádek v Air New Zealand, viz [87];
- automatizaci pracovního harmonogramu pro více než 3000 zaměstnanců ve více než 400 obchodů a skladů s alkoholem v kanadském Quebecu (SAQ – Société des alcools du Québec), která by měla přinést úsporu ve výši 1 000 000 CAD, viz [212];
- optimalizaci rozvrhu na Ohijské univerzitě, viz a [371];
- zlepšení globálního distribučního procesu ve společnosti Zara, viz [92].

Pro další články odkazujeme na webové stránky časopisu

https://ln.math.muni.cz/pz/opt/ZLP-aplikace

Každý rok je nejlepší šestice takovýchto projektů oceňována *Franz Edelman Award*, kterou od roku 1972 uděluje INFROMS. Na jejich webových stránkách se dokonce uvádí, že za dobu trvání (do roku 2021) přesáhly kumulativní výhody z projektů finalistů tohoto ocenění hranici 292 miliard dolarů!

Druhá cesta k získání 1 000 000 $ díky znalostem lineárního a celočíselného programování je trochu složitější. V poznámce na straně 20 věnované Hilbertovi jsme se zmínili o jeho seznamu čítajícím 23 nevyřešených matematických problémů. Hilbert však matematický svět obohatil i o mnoho dalších výzev a jednou z nich je tzv. *Entscheidungsproblem*[154], kterou v roce 1928 zveřejnil společně s Ackermann[155]. Cílem této výzvy bylo nalezení (či vyvrácení existence) algoritmu, který by dokázal rozhodnout o tom, zda dané matematické tvrzení je univerzálně platné či nikoli, tj. platné v libovolném modelu splňující uvažované axiomy. Jinými slovy, má jít o zcela „mechanický“ proces (tj. algoritmus), který by dokázal rozhodnout o dokazatelnosti daného tvrzení z uvažovaných axiomů na základě pravidel výrokové logiky[156]. Ještě v roce 1930 Hilbert věřil, že nic takového jako neřešitelný problém neexistuje. Poměrně brzy však nezávisle na sobě ukázali Church[157] (v roce 1936) a jeho student Turing[158] (o chvíli později v roce 1937), že takový algoritmus *neexistuje*. Navzdory negativní odpovědi však Hilbertova a Ackermannova výzva pomohla spustit dvě obrovské a vzájemně propojené revoluce, které (alespoň v jistém smyslu) formovaly 20. století: první ve vědě a filosofii, kde se otázka hranic logického uvažování jako takového stala předmětem matematické analýzy (sem pochopitelně patří Gödel, Turing, Church, Post a další), a druhou v technologii, kam až dosáhl elektronický počítač – sice se nenaplnily všechny Hilbertovy představy, ale je to dost na to, aby to změnilo každodenní život většiny lidí na Zemi.

Problém, který stojí v tuto chvíli před námi, je moderní modifikace výše zmíněného Entscheidungsproblem. Poprvé se v jistém smyslu objevuje asi již v roce 1955, kdy Nash[159]

zaslal Národní bezpečnostní agentuře (NSA) několik dopisů, ve kterých mimo jiné nalezneme následující domněnku.

[154] *Rozhodovací problém*, angl. by to byl *decision problem*, ale i zde se využívá spíše původní německý název.

[155] Wilhelm Friedrich Ackermann (29. března 1896 – 24. prosince 1962) byl německý matematik a filosof, který je znám především díky jeho práci v matematické logice a tzv. Ackermannovou funkcí používanou ve výpočetní teorii (jedná se o příklad funkce, která je rekurzivní a přitom není primitivně rekurzivní; hodnota Ackermannovy funkce $A(n)$ roste velmi rychle a už pro velmi malá čísla (4, 5, ...) je nemyslitelné tuto hodnotu spočítat, např. $A(4)$ je tak obrovské číslo, že už počet jeho číslic je vyšší než počet všech atomů v pozorovaném vesmíru; jinými slovy, Ackermannova funkce roste nade všechny rozumně představitelné meze a není omezitelná žádnou běžně používanou funkcí). Byl Hilbertovým studentem v Göttingenu. Pomáhal Hilbertovi při přepisování jeho přednášek o matematické logice z let 1917–1922 do uceleného textu, ze kterého vznikla kniha *Principles of Mathematical Logic*. Navzdory svému vzdělání a vědeckým výsledkům Ackermann učil na gymnáziu. K získání akademické pozici mu nejspíše chyběla větší Hilbertova podpora, o kterou přišel, když se oženil, což Hilbert u mladých vědců nepodporoval. Hilberta to velmi rozzlobilo, a až tak že to údajně komentoval slovy: *Oh, das ist wunderbar. Das sind gute Neuigkeiten für mich. Denn wenn dieser Mann so verrückt ist, daß er heiratet und sogar ein Kind hat, bin ich von jeder Verpflichtung befreit, etwas für ihn tun zu müssen.* (*Oh, to je úžasné. To je pro mě dobrá zpráva. Protože pokud je tento muž natolik blázen, aby se oženil a dokonce měl dítě, jsem zbaven jakékoli povinnosti pro něj cokoliv udělat.*). No nic, nikdo netvrdí, že Hilbert byl světec...

[156] O něčem takovém snil již Leibniz poté, co se mu podařilo úspěšně sestavit mechanický kalkulátor. Chtěl sestavit *symbolický* mechanický kalkulátor, který by dokázal rozhodnout o pravdivosti matematického tvrzení. Leibniz si přitom byl vědom toho, že prvním krokem musí být sestavení čistě formálního jazyka, čemuž věnoval nemalou část své další práce, viz [130, Chapter One].

> *Now my general conjecture is as follows: For almost all sufficiently complex types of enciphering, especially where the instructions given by different portions of the key interact complexly with each other in the determination of their ultimate effects on the enciphering, the mean key computation length increases exponentially with the length of the key, or in other words, with the information content of the key.*
> *The significance of this general conjecture, assuming its truth, is easy to see. It means that it is quite feasible to design ciphers that are effectively unbreakable. As ciphers become more sophisticated the game of cipher breaking by skilled teams, etc., should become a thing of the past.*
> *The nature of this conjecture is such that I cannot prove it, even for a special type of ciphers. Nor do I expect it to be proven. But this does not destroy its significance. The probability of the truth of the conjecture can be guessed at on the basis of experience with enciphering and deciphering.*[160]

Zdá se tedy, že Nash (navzdory absenci jakéhokoli důkazu) se domníval, že množství času potřebného k prolomení dostatečně komplexního kódu se zvětšuje exponenciálně s délkou klíče. Zdánlivě zcela odlišnou otázku adresoval o chvíli později Gödel von Neumannovi v dopise z roku 1956: *má dokazování vět lineární či kvadratickou časovou složitost*[161]? Kdyby ano tak, jak dále uvádí Gödel, by bylo možné automatizovat/mechanizovat hledání matematických důkazů, viz [254]. Bohužel, von Neumannovu odpověď neznáme. Žádný dopis adresovaný zpět Gödelovi se totiž najít nepodařilo. Navíc v té době již byl von Neumann velmi sužován rakovinou a o pár měsíců později zemřel (únor 1957). Zdá se, že alespoň Gödel (na rozdíl od Nashe) byl optimista a věřil, že odpověď na jeho otázku je kladná. Avšak je tomu skutečně tak? To nevíme ani dnes! A nalezení správné odpovědi se zdá být tak složité a zároveň důležité, že za ni Clayův matematický ústav slíbil odměnu ve výši jednoho miliónu USD v rámci jím vyhlášených *Problémů tisíciletí*[162] Jedná se o tzv. *Problém P vs. NP* (neboli $\mathbf{P} \stackrel{?}{=} \mathbf{NP}$) a velmi zjednodušeně řečeno jde o to, zda existuje *rychlý* postup pro řešení všech problémů, které mají krátké odpovědi, jenž mohou být snadno ověřeny mechanicky. Můžeme si představit např. velkou skládanku puzzle s 10^{1000} možných uspořádání dílků nebo zašifrovanou zprávu s podobně obrovským počtem možných variant rozluštění nebo aerolinky s astronomickým počtem způsobů pro vytvoření jejich letového řádu apod. Všechny tyto úkoly mají společné dvě klíčové vlastnosti

(i) sice konečný ale exponenciálně rostoucí počet možných řešení;

(ii) rychlý, mechanický způsob ověření, zda nalezené řešení je správné (v případě puzzle to je „kouknu a vidím“, v případě aerolinií může jít o snadnou kontrolu dosažení požadovaného zisku apod.).

Ptáme se tedy na to, zda za těchto podmínek existuje obecná metoda pro nalezení správného řešení (kdykoli existuje), která by byla výrazně rychlejší než metoda „hrubé síly“, tj. pouhé zkoušení jednotlivých možných variant od teď až do konce světa.

Samozřejmě aby naše diskuze byla rigoróznější, potřebujeme ony poněkud vágní pojmy jako „rychle“, „mechanicky“ či „snadno“ přesně vymezit. Všechny matematizovatelné problémy můžeme rozdělit do dvou skupin: (i) ty, které lze řešit algoritmicky, (ii) a ty ostatní (tj. neexistuje/neznáme pro ně algoritmické řešení). Nás pochopitelně zajímá pouze ta první skupina – rozhodovací problémy, které si můžeme představit jako nekonečnou řadu otázek s odpověď'mi ano–ne, přičemž bychom rádi měli nějaký nástroj k jejich porovnávání. Tím pro nás bude *časová složitost* algoritmu pro nalezení řešení. Ta určuje závislost výpočetního času algoritmu na velikosti vstupu, o čemž jsme se zmínili již v Podkapitole 2.13. V roce 1965 navrhli Cobhan[163] a Edmonds[164] rozlišit dva druhy problémů podle toho, zda pro ně existují algoritmy vyžadující (i) *polynomiální časovou složitost* (tj. algoritmus běží v polynomiálním čase $O(N^k)$ pro nějaké k, což je dobrý/efektivní/zvládnutelný) a (ii) *exponenciální* nebo *superpo-*

[157] Alonzo Church (14. června 1903 – 11. srpna 1995) byl americký matematik a filosof, který významně přispěl k rozvoji matematické logiky a základů teoretické informatiky. Byl jedním z prvních lidí zabývajících se teorií algoritmů. Vystudoval na Princetonské univerzitě, kde se později stal i profesorem. Po boku Turinga je Church považován za jednoho ze zakladatelů počítačové vědy.

[158] Alan Mathison Turing (23. června 1912 – 7. června 1954) byl anglický matematik, logik, kryptoanalytik, filosof a teoretický biolog. Turing je široce považován za otce teoretické informatiky a umělé inteligence. Formalizoval pojem „algoritmu“ a „Turingova stroje“, což je teoretický/abstraktní (matematický) model počítače, který se využívá pro modelování algoritmů v teorii vyčíslitelnosti. Na jeho počest je od roku 1966 udílena *Turingova cena*, což je jedno z nejvýznamnějších ocenění v informatice.

Veřejně známý je především pro své zásluhy při dešifrování Enigmy, kterou nacisté využívali pro zasílání tajných zpráv během 2. světové války. V 50. letech byl Turing odsouzen za homosexuální akt na základě zákona z roku 1885. Až v září 2009 se britský premiér Gordon Brown jménem vlády Turingovi omluvil za příkoří, které mu bylo způsobeno, a v roce 2013 mu britská královna Alžběta II. udělila posmrtnou královskou milost. Turingův život byl námětem pro film *Kód Enigmy* z roku 2014.

[159] John Forbes Nash Jr (13. června 1928 – 23. května 2015) byl americký matematik,který se věnoval zejména teorii her a diferenciální geometrii. V létě roku 1950 (a také v letech 1952 a 1954) pracoval v RANDu. V roce 1978 získali společně s C. Lemkem *John von Neumann Theory Prize* za „jejich mimořádné příspěvky v teorii her“. Nash ocenění získal především za objev tzv. nekooperativní rovnováhy, který je dnes známa spíše jako *Nashova rovnováha*. Ta byla využita např. v analýzách různých konfliktních situací jako jsou války či závody ve zbrojení, viz [446]. V roce 1994 Nash společně s Reinhardem Seltenem a Johnem Harsanyim získali dokonce *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za jejich *průkopnickou analýzu rovnováhy v teorii nekooperativních her*.

Nashovu slibně se rozvíjející matematickou kariéru na takřka třicet roků výrazně zkomplikovala para-

noidní schizofrenie. Navzdory této nemoci vytrval ve své práci, na čemž měla velkou zásluhu i jeho manželka Alicia. Získal povolení využívat prostředky Princetonské univerzity pro svůj výzkum, avšak občas se zde zjevil i o půlnoci a čmáral na tabule tajemné rovnice. Jeho stav se postupem času zlepšoval a 90. léta 20. století byla ve znamení „návratu génia". Souboj jeho matematického nadání s duševní nemocí byl námět i pro film *Čistá duše*. John Nash zemřel společně se svojí manželkou v taxi při automobilové nehodě.

[160] *Moje obecná domněnka je nyní následující: u téměř všech dostatečně složitých typů šifrování, zejména tam, kde instrukce dané různými částmi klíče vzájemně komplexně interagují při určování jejich konečných účinků na šifrování, se průměrná délka výpočtu klíče prodlužuje exponenciálně s délkou klíče, nebo jinými slovy, s informačním obsahem klíče.*

Význam této obecné domněnky, za předpokladu její pravdivosti, je snadno pochopitelný. Znamená, že je docela možné navrhnout šifry, které jsou efektivně nerozlomitelné. Jak se šifry stávají sofistikovanějšími, měla by se hra na lámání šifer zkušenými týmy atd. stát minulostí.

Povaha této domněnky je taková, že ji nemohu dokázat ani pro zvláštní typ šifer. Ani nečekám, že se to dokáže. Ale to nesnižuje její význam. Pravděpodobnost pravdivosti domněnky lze odhadnout na základě zkušeností s šifrováním a dešifrováním. Viz [392].

[161] O těchto pojmech jsme již mluvili na začátku Podkapitoly 2.13.

[162] Viz opět poznámku o Hilbertovi na straně 20.

[163] Alan Cobham je dalším z těch, o kterém se nepodařilo nic dohledat – vyjma toho, že zavedení polynomiální časové složitosti (tj. třídy **P**) se poprvé objevilo v jeho článku z roku 1965 (nezávisle na Edmondsovi), viz [96].

[164] Jack R. Edmonds (∗ 5. dubna 1934) je americký počítačový vědec a matematik, který žil a pracoval převážně v Kanadě. Jeho největší přínos tkví především v oblastech kombinatorické optimalizace, polyedrové kombinatorice, diskrétní matematice a výpočtové teorii. V roce 1985 obdržel *John von Neumann Theory Prize*. V jeho článku [152] se také objevuje třída **P** pro „snadno" řešitelné problémy.

[165] Následující ilustrativní příklady byly převzaty z [291, str. 52].

lynomiální časovou složitost (tj. algoritmus běží v exponenciálním čase $O(k^N)$ pro nějaké k, což je špatný/neefektivní/nezvládnutelný). Avšak jako obvykle je potřeba upozornit, že toto je pouze jedno z teoretických měřítek a v praxi je potřeba uvážit i některá jiné hlediska, neboť např.[165] algoritmus s časovou složitostí $10^{1000}\,N$ bychom považovali za dobrý (ale v praxi asi nepoužitelný, protože úplně pro všechna N bude potřebný čas větší než 10^{1000}), zatímco případ $10^{N/10^9}$ bychom považovali za špatný (ačkoli v praxi může obstojně fungovat pro velmi široké spektrum hodnot N, neboť pro $N \leq 10^9$ bude potřebný čas menší než 10). Nebo podobně můžeme srovnávat např. N^{1000} a $1{,}000001^N$. Navíc se může stát, že ten dobrý algoritmus budeme mít pouze teoreticky (budeme vědet o jeho možné existenci a znát některé jeho vlastnosti), ale reálně jej mít k dispozici nebudeme. To se „povedlo" např. Roberstonovi a Seymourovi v sérii práci z let 1983–2004 o minorech grafů, viz např. [430], když dokázali existenci polynomiálního algoritmu, avšak žádný takový algoritmus nesestrojili (a z praktického hlediska je taková konstrukce možná pouze v případě splnění jisté velmi restriktivní podmínky). Na druhou stranu se přeci jen podařilo ukázat jistou souvislost mezi „řešitelností v polynomiálním čase" a „efektivní řešitelností v praxi"[166]. Často také v průběhu času dochází ke zlepšování algoritmů a navíc v praxi algoritmy mnohdy běží mnohem efektivněji než je garantováno teoretickými výsledky (jako např. v případě simplexové metody).

Dostáváme se tak k rozdělení problémů do dvou tříd. *Třídou složitosti* **P** rozumíme množinu problémů, pro jejichž řešení existuje deterministický algoritmus[167] s polynomiálním časem běhu (na standardním počítači či Turingově stroji)[168]. Zajímavější a složitější je *třída složitosti* **NP**, která obsahuje problémy řešitelné v nedeterministickém polynomiálním čase. Pro tyto problémy neumíme najít řešení v polynomiálním čase, ale existuje pro ně *verifikační* algoritmus s polynomiální dobou běhu. Verifikační algoritmus ověří správnost řešení problému na základě poskytnutých dat, kterým říkáme *certifikát*. To se stane např. při skládání puzzle, které je tedy v **NP**. Nebo si představme bludiště, ve kterém hledáme nejkratší cestu ven. Je známo, že cestu můžeme nalézt v lineárním čase vzhledem k velikosti bludiště pomocí tzv. principu „prohledávání do šířky". Můžeme se ale dostat do situace, kdy naše bludiště bude plné nadchodů a podchodů, a zároveň se v něm budou vyskytovat různá stanoviště a naším úkolem bude najít ideální nejkratší cestu, přičemž budeme muset právě jednou projít všechna předem vyznačená místa. Kdybychom chtěli použít zmíněné prohledávání do šířky, může se nám stát, že se budeme mnohokrát vracet přes jednotlivá stanoviště, neboť naše zvolená posloupnost může být špatná. Tedy v tento moment nebudeme vyznačená místa navštěvovat maximálně jednou, ale každou jejich posloupnost projdeme nejvýše jednou – těch ale může být exponenciálně mnoho vzhledem k velikosti daného bludiště. Jestliže nebudeme mít nápovědu (to je onen certifikát) se správným řešením, což by nám umožnilo najít optimální cestu v lineárním čase (prostým následováním cesty z nápovědy), nalezli jsme problém, jenž nemůžeme zařadit do třídy **P**.

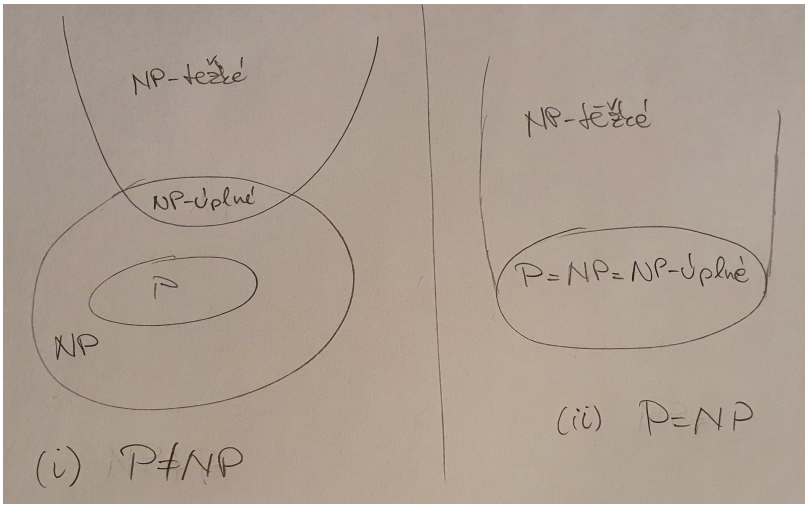
Je zřejmé, že každý problém v **P** je také v **NP**, tj. umím-li problém vyřešit polynomiálním algoritmem, tak jej jistě umím i ověřit v polynomiálním čase. Ovšem platí i opačná implikace a inkluze? Neboli

$$\mathbf{P} \stackrel{?}{=} \mathbf{NP}.$$

Tato otázka byla poprvé přesně zformulována Cookem v článku [99] z roku 1971 a nezávisle také o dva roky později Levinem v [354]. Samozřejmě cílem není jen zodpovězení této otázky. Pokud se třeba ukáže, že problémy z třídy **NP** jsou řešitelné v čase $1{,}00001^N$ nebo N^{1000}, ale už ne v N^{999}, tak ať. Zkrátka to především slouží jako jakýsi ukazatel neznalosti upozorňující nás na to, že pokud neumíme zodpovědět tuto otázku, tak rozhodně nemůžeme očekávat, že dokážeme najít odpovědi na ty mnohem.

Příklad s puzzle podněcuje zápornou odpověď, tj. ověření je sice možné v polynomiálním čase, avšak pro řešení je potřeba exponenciální časová složitost. Jenže komplikovanost této otázky tkví v tom, že je potřeba ukázat, že neexistuje žádný algoritmus s polynomiální časovou

složitostí – tedy musíme uvážit všechny možné algoritmy řešení (i ty, které neznáme) a ukázat, že žádný z nich není efektivní. V množině **NP** je také velké množství důležitých problémů, které jsou tzv. **NP**-*úplné*[169], tj. neexistuje pro ně polynomiální algoritmus ($\notin \mathbf{P}$) a jsou na ně polynomiálně redukovatelné všechny ostatní problémy z **NP**. Tato druhá podmínka určuje tzv. třídu **NP**-*těžkých*[170] problémů, které mohou být i mimo **NP**, tj. problémy z **NP** a mimo **P** jsou v jistém smyslu rovnocenné. Máme tedy dvě možnosti uspořádaní jednotlivých tříd, viz Obrázek 3.32. Jak vlastně vypadá **NP**-těžký problém, který není **NP**-úplný? Je to např. tzv. *problém zastavení*: znáte-li zdrojový kód programu a jeho vstup, rozhodněte, zda program zastaví, nebo zda poběží navždy bez zastavení.



Obrázek 3.32: (i) $\mathbf{P} \neq \mathbf{NP}$ nebo (ii) $\mathbf{P} = \mathbf{NP}$?

Třídu **NP**-úplných úloh tak tvoří v jistém smyslu ty nejtěžší úlohy z **NP**. Např. úlohy celočíselného programování jsou **NP**-úplné (problém batohu/zloděje/kontejneru, problém obchodního cestujícího atd.), zatímco úlohy lineárního programování jsou díky metodě elipsoidů a Karmakarově metodě řešitelné v polynomiálním čase. Pokud by se podařilo najít polynomiální deterministický algoritmus pro nějakou **NP**-úplnou úlohu, znamenalo by to, že všechny nedeterministické polynomiální problémy jsou řešitelné v polynomiálním čase, a tudíž třída **NP** by se „zhroutila“ do **P**, tj. $\mathbf{P} = \mathbf{NP}$. To znamená, že pokud máme nějaký **NP**-úplný problém a víme-li, že to *není protipříklad*, pak *žádný* protipříklad *neexistuje*. Jak ale začít? Prakticky každý problém v **NP**, který někdo vymyslí a není očividně v **P**, se ukáže být **NP**-úplný. Studovaly se tisíce problémů… Pro mnohem více podrobností zahrnujících důležitost tohoto problému, různé námitky proti některým jeho aspektům, historický přehled shrnující více než 40letý vývoj a současný stav doporučujeme čtenáři článek [1].

Cvičení

Pro kontrolu výpočtů můžete využít např.

https://ln.math.muni.cz/pz/opt/simplexSOFT1

Základní manuál k tomuto programu je k dispozici na adrese

https://ln.math.muni.cz/pz/opt/simplexSOFT1_manual

3.6.1. Vhodnou metodou vyřešte problém

$$4x_1 + 3x_2 + 4x_3 + x_4 \to \min$$
$$x_1 - 3x_2 + x_3 + 2x_4 \geq 3$$
$$x_1, x_2, x_3, x_4 \in \{0, 1\}.$$

Řešení: $[0,0,1,1]$ *a* $[1,0,0,1]$.

[166] Zmíněné výrazně kontrastující příklady časové složitosti jsou spíše akademickými úvahami.

[167] Tj. jsou pevně nastavena pravidla jednotlivých kroků algoritmu a dokážeme je přesně predikovat. Kdyby ale některý z kroků závisel např. na hodu kostkou, jednalo by se o nedeterministický algoritmus – jinými slovy, algoritmus by při spuštění na tomtéž problému mohl dávat různé výsledky.

[168] Samozřejmě nějaký problém může být v třídě **P** i díky několika různým algoritmům. To ale nehraje vůbec žádnou roli – důležité je, že alespoň jeden takový existuje. Navíc nám to samozřejmě nebrání v hledání dalších možných vylepšení.

[169] **NP**-*complete*

[170] **NP**-*hard*

3.6.2. Vhodnou metodou vyřešte problém

$$70x_1 + 61x_2 + 54x_3 + 40x_4 + 31x_5 + 25x_6 \to \max$$
$$30x_1 + 27x_2 + 25x_3 + 20x_4 + 18x_5 + 15x_6 \leq 72$$
$$x_1, x_2, x_3, x_4, x_5, x_6 \in \{0, 1\}.$$

Řešení: $[1, 1, 0, 0, 0, 1]$.

3.6.3. Vhodnou metodou vyřešte problém

$$3x_1 + 5x_2 + 6x_3 + 9x_4 + 10x_5 + 10x_6 \to \min$$
$$-2x_1 + 6x_2 - 3x_3 + 4x_4 + x_5 - 2x_6 \geq 2$$
$$-5x_1 - 3x_2 + x_3 + 3x_4 - 2x_5 + x_6 \geq -2$$
$$5x_1 - x_2 + 4x_3 - 2x_4 + 2x_5 - x_6 \geq 3$$
$$x_1, x_2, x_3, x_4, x_5, x_6 \in \{0, 1\}.$$

Řešení: $[0, 1, 1, 0, 0, 0]$.

3.6.4. Vhodnou metodou vyřešte problém

$$-2x_1 + x_2 + x_3 - 10 \to \min$$
$$2x_1 + 3x_2 - x_3 \leq 3$$
$$x_2 + x_3 \geq 1$$
$$3x_1 + 3x_2 + 3x_3 = 6$$
$$x_1, x_2, x_3 \in \{0, 1\}.$$

Řešení: $[1, 0, 1]$.

3.6.5. Vhodnou metodou vyřešte problém

$$-2x_1 - x_2 - 5x_3 \to \min$$
$$3x_1 + 2x_2 + 7x_3 \leq 9$$
$$x_1, x_2, x_3 \in \{0, 1\}.$$

Řešení: $[0, 1, 1]$.

3.6.6. Vhodnou metodou vyřešte problém

$$-10x_1 - 15x_2 - 16x_3 - x_4 \to \min$$
$$20x_1 + 15x_2 + 20x_3 + 5x_4 \leq 25$$
$$x_1, x_2, x_3, x_4 \in \{0, 1\}.$$

Řešení: $[0, 0, 1, 1]$.

3.6.7. Vhodnou metodou vyřešte problém

$$x_1 - 4x_2 + 2x_3 - 3x_4 \to \min$$
$$2x_1 + x_2 - 2x_3 \leq 0$$
$$x_1 + 2x_2 + x_3 + x_4 \leq 2$$
$$x_1, x_2, x_3, x_4 \in \{0, 1\}.$$

Řešení: $[0, 1]$.

3.6.8. Vhodnou metodou vyřešte problém

$$2x_1 - x_2 - 2x_3 + x_4 + 3x_5 - x_6 \to \max$$
$$x_1 + x_2 - 3x_3 + 2x_4 - x_5 \geq 2$$
$$x_1, x_2, x_3, x_4, x_5, x_6 \in \{0, 1\}.$$

Řešení: $[0, 1, 0, 1, 0, 1]$.

3.6.9. Vyřešte „problém batohu“, jehož kapacita je 5 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4
cena	3	4	5	6
váha	2	3	4	5

Řešení: 1 & 2.

3.6.10. Vyřešte „problém batohu“, jehož kapacita je 7 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4
cena	10	7	25	24
váha	2	1	6	5

Řešení: 1 & 4.

3.6.11. Vyřešte „problém batohu“, jehož kapacita je 8 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4
cena	15	10	9	5
váha	1	5	3	4

Řešení: 1 & 3 & 4.

3.6.12. Vyřešte „problém batohu“, jehož kapacita je 6 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4	5
cena	25	20	15	40	50
váha	3	2	1	4	5

Řešení: 3 & 5.

3.6.13. Vyřešte „problém batohu“, jehož kapacita je 11 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4	5
cena	1	6	18	22	28
váha	1	2	5	6	7

Řešení: 3 & 4.

3.6.14. Vyřešte „problém batohu“, jehož kapacita je 15 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] jsou uvedeny v tabulce

věc č.	1	2	3	4	5	6	7	8	9	10
cena	18	20	17	19	25	21	27	23	25	24
váha	1	3	7	4	8	9	6	10	2	5

Řešení: 1 & 2 & 4 & 9 & 10.

3.6.15. Vyřešte „problém batohu“, jehož kapacita je 10 [jednotek] a je-li dostupné libovolné množství každého druhu. Ceny [desítky Kč] jednotlivých druhů jsou uvedeny v tabulce

druh č.	1	2	3	4
cena	7	9	2	15
váha	3	4	1	7

Řešení: 1 & 1 & 2.

3.6.16. Vhodnou metodou vyřešte problém

$$-x_1 - 2x_2 \to \min$$
$$-3x_1 + 4x_2 \le 6 \quad \& \quad 4x_1 + 3x_2 \le 12$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[1,2]$.

3.6.17. Vhodnou metodou vyřešte problém

$$-x_1 - 2x_2 \to \min$$
$$-4x_1 + x_2 \le 0 \quad \& \quad 7x_1 + 4x_2 \le 14$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[1,1]$.

3.6.18. Vhodnou metodou vyřešte problém

$$-x_1 - 4x_2 \to \min$$
$$2x_1 + 4x_2 \le 7 \quad \& \quad 10x_1 + 3x_2 \le 14$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[1,1]$.

3.6.19. Vhodnou metodou vyřešte problém

$$20 - 3x_1 - 4x_2 \to \min$$
$$2x_1 + 5x_2 \le 15 \quad \& \quad 2x_1 - 2x_2 \le 5$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[2,2]$.

3.6.20. Vhodnou metodou vyřešte problém

$$x_1 - 3x_2 \to \min$$
$$x_1 + x_2 \le 15 \quad \& \quad 3x_1 - 2x_2 \ge -5 \quad \& \quad x_1 - x_2 \le 10$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[5, 10]$.

3.6.21. Vhodnou metodou vyřešte problém

$$-x_1 - 5x_2 \to \min$$
$$-4x_1 + 3x_2 \le 6 \quad \& \quad 3x_1 + 2x_2 \le 18$$
$$x_1, x_2 \in \mathbb{N}_0.$$

Řešení: $[3, 4]$.

3.6.22. Vhodnou metodou vyřešte problém

$$-2x_1 + 3x_2 + 4x_3 \to \min$$
$$-x_1 + x_2 + 3x_3 \le 8 \quad \& \quad 3x_1 + 2x_2 - x_3 \le 10$$
$$x_1, x_2, x_3 \in \mathbb{N}_0.$$

Úloha je neohraničená.

3.6.23. Vhodnou metodou vyřešte problém

$$x_1 - 3x_2 + 4x_3 \to \min$$
$$2x_1 + 3x_2 - 5x_3 \le 10 \quad \& \quad 4x_1 - 2x_2 + x_3 \ge 5$$
$$x_1, x_2, x_3 \in \mathbb{N}_0.$$

Řešení: $[3, 3, 1]$.

3.6.24. Vhodnou metodou vyřešte problém

$$-2x_1 + x_2 + x_3 - 10 \to \min$$
$$2x_1 + 3x_2 - x_3 \le 9 \quad \& \quad 2x_2 + x_3 \ge 4 \quad \& \quad 3x_1 + 3x_2 + 3x_3 = 6$$
$$x_1, x_2, x_3 \in \mathbb{N}_0.$$

Řešení: $[0, 2, 0]$.

3.6.25. Vhodnou metodou vyřešte problém

$$5x_1 + x_2 + 2x_3 - 5 \to \min$$
$$-2x_1 + x_2 - x_3 + x_4 = 7/2 \quad \& \quad 2x_1 + x_2 + x_3 + x_5 = 2$$
$$x_1, x_2, x_4 \ge 0 \quad \& \quad x_3, x_5 \in \mathbb{N}_0.$$

Řešení: $[0, 0, 0, 7/2, 2]$.

3.6.26. Vhodnou metodou vyřešte problém

$$5x_1 + x_2 + 2x_3 - 5 \to \min$$
$$-2x_1 + x_2 - x_3 + x_4 = 7/2 \quad \& \quad 2x_1 + x_2 + x_3 + x_5 = 2$$
$$x_1, x_2, x_3 \ge 0 \quad \& \quad x_4, x_5 \in \mathbb{N}_0.$$

Řešení: $[0, 3/4, 1/4, 3, 1]$.

3.6.27. Vhodnou metodou vyřešte problém

$$-6x_1 - 4x_2 - 4x_3 - x_4 - x_5 \to \min$$
$$2x_1 + 2x_2 + 3x_3 + x_4 + 2x_5 = 7$$
$$x_1, x_2 \in \mathbb{N}_0 \quad \& \quad x_3, x_4, x_5 \geq 0.$$

Řešení: $[3, 0, 1/3, 0, 0]$.

3.6.28. Vhodnou metodou vyřešte problém

$$2x_1 + x_2 + 4x_3 - 5x_4 \to \min$$
$$x_1 + 3x_2 + 2x_3 + 5x_4 \leq 8$$
$$2x_1 + x_2 + x_3 + x_4 \geq 5$$
$$3x_1 - 2x_2 - 5x_3 + x_4 \leq 0$$
$$x_1, x_2, x_3, x_4 \in \mathbb{N}_0.$$

Řešení: $[2, 1, 1, 0]$.

3.6.29. Vhodnou metodou vyřešte problém

$$3x_1 + 11x_2 + 5x_3 - x_4 \to \min$$
$$x_1 - 5x_2 + 3x_3 + x_4 \leq 3 \quad \& \quad -x_1 + x_2 + x_3 - x_4 \geq 3$$
$$x_1, x_2, x_3, x_4 \in \mathbb{N}_0.$$

Řešení: $[0, 1, 2, 0]$.

3.6.30. Vhodnou metodou vyřešte problém

$$4x_1 + 2x_2 + x_3 - 2x_4 \to \min$$
$$x_1 - x_2 + 2x_3 + x_4 \leq 11$$
$$x_2 + x_3 + x_4 \geq 17$$
$$3x_1 - x_3 - 3x_4 \geq 5$$
$$x_1, x_2, x_3, x_4 \in \mathbb{N}_0.$$

Řešení: $[9, 9, 1, 7]$.

3.6.31. Vhodnou metodou vyřešte problém

$$4x_1 - 2x_2 + 7x_3 - x_4 \to \max$$
$$x_1 + 5x_3 \leq 10$$
$$x_1 + x_2 - x_3 \leq 1$$
$$6x_1 + 5x_2 \leq 0$$
$$-x_1 + 2x_3 - 2x_4 \leq 3$$
$$x_4 \geq 0$$
$$x_1, x_2, x_3 \in \mathbb{N}_0.$$

Řešení: $[0, 0, 2, 1/2]$.

3.6.32. Vyřešte „problém plnění kontejneru“, jehož kapacita je 11 [jednotek]. Dostupný náklad a jednotlivé ceny [100 Kč] jsou v následující tabulce.

typ č.	1	2	3	4
cena	18	25	30	20
váha	2	4	5	3
dostupné množství	5	2	2	3

Řešení: typ 1=4 ks, t2=0, t3=0, t4=1.

3.6.33. Láďa Hruška začal podnikat v rychlém občerstvení a se svým řetězcem s názvem „Kuřecí kůžičky a bramborové šlupky“ úspěšně konkuruje nadnárodním řetězcům typu McDonald's. Je proto načase, aby otevřel pobočku i v Brně. Jeho tým rozdělil Brno a okolí na 6 oblastí a stanovil průměrnou časovou náročnost jízdy mezi těmito oblastmi. Výsledek je shrnut v následující tabulce, kde prvek na pozici (řádek,sloupec)=(A,B) udává dobu jízdy z A do B (pozor, tabulka není symetrická).

oblast	A	B	C	D	E	F
A	0	10	20	30	30	12
B	17	0	25	35	23	10
C	20	25	0	15	30	20
D	30	40	9	0	15	25
E	20	20	30	15	0	14
F	10	10	20	25	16	0

Samozřejmě není nutné otevírat pobočku v každé oblasti. Ovšem Láďa Hruška ví, že pro úspěch svého podnikání potřebuje, aby doba doručení jeho delikates nepřekročila 15 minut. Zformulujte tuto úlohu jako problém celočíselného programování, jehož řešením bude optimální rozdělení poboček po Brně. *Řešení: otevření poboček v A & D.*

3.6.34. Šestice hráčů basketbalu se přihlásila na turnaj čtyřčlenných družstev. Pro určení základní sestavy se rozhodli maximalizovat průměrnou výšku jejich hráčů, která je daná následující tabulkou.

číslo	jméno	výška vzhledem ke 190 cm	pozice
1	David	+10	centr
2	Jan	+9	centr
3	Marek	+6	útok
4	Richard	+6	útok
5	Karel	+4	obrana
6	Jakub	-1	obrana

Současně je potřeba dodržet tato pravidla

- musí nastoupit alespoň jeden obránce,
- buď Jan nebo Karel musí zůstat na lavičce,
- pouze jeden centr může nastoupit,
- nastoupí-li Jan nebo Richard, tak Jakub zůstane na lavičce.

Zformulujte tento problém jako úlohu ($\mathbb{Z}$LP) a vyřešte ji.

3.6.35. Přírodovědecká fakulta MU otevřela v Botanické zahradě velkoobchodní květinářství. Za tímto účelem byly zakoupeny 4 dodávky. V tuto chvíli má 10 odběratelů s objednávkami o velikosti d_j, $j = 1,\dots,10$. Kapacita dodávek je L_k s denními náklady c_k, $k = 1,\dots,4$. Dodávka nesmí obsloužit více než 5 zákazníků, přičemž z konkurenčních důvodů nelze současně zásobovat tyto dvojice $\{1,7\}$, $\{2,6\}$ a $\{2,9\}$. Sestavte odpovídající úlohu ($\mathbb{Z}$LP) s minimalizací nákladů na doručení všem zákazníkům.

3.6.36. Reklamní agentura Nekecáme, s.r.o., zvažuje množnosti reklamní kampaně pro příští MS v piškvorkách konané v Brně. Na základě předchozích turnajů byly určeny vhodné nástroje a jejich účinek, viz následující tabulku.

	TV	sportovní časopisy	noviny	radio	společenské časopisy	propagační akce	dostupné zdroje
„zasažení odběratelé“	10^6	2×10^5	3×10^5	4×10^5	45×10^4	45×10^5	
náklady [Kč]	5×10^5	15×10^4	3×10^5	25×10^4	25×10^4	10^5	18×10^5
potřební designéři [pracovní hodiny]	700	250	200	200	300	400	1500
potřební obchodní zástupci [pracovní hodiny]	200	100	100	100	100	1000	1200

Cílem je maximalizovat počet „zasažených odběratelů“ jednotlivých nástrojů při dodržení dostupných zdrojů. Navíc je potřeba dodržet požadavky:

(a) má-li se uskutečnit propagační akce, je potřeba ji zviditelnit v rádiu nebo ve společenských časopisech,

(b) nelze využit současně sportovní a společenské časopisy.

Sestavte odpovídající úlohu ($\mathbb{Z}$LP) a vyřešte ji.

3.6.37. Jakub Kryštof Rad býval ředitelem dačické rafinérie a je považován za vynálezce kostkového cukru, který si v roce 1843 nechal patentovat. Aby rozšířil svůj odbyt, rozhodl se podniknout cestu do USA. Na tuto cestu si mohl vzít nejvýše 50 kg cukru s tím, že každý kilogram prodá za 40 zlatých. Náklady na cestu činí 450 zlatých plus 5 zlatých za každý kilogram váhy cukru nad 20 kg. Vyplatí se taková cesta? A kolik cukru by si s sebou měl vzít? Zformulujte tento problém jako úlohu ($\mathbb{Z}$LP) a vyřešte ji.

3.6.38. Spolek přátel PřF MU začal provozovat v knihovně kopírovací služby. Momentálně mají k dispozici jeden kopírovací stroj s kapacitou 1000 kopií za hodinu a náklady 2 Kč za 1.–4. kopii, 1 Kč za 5.–8. kopii a 0,5 Kč za 9. a každou další kopii. Na žádost ÚMS si vyhradili 1 hodinu na kopírování desetistránkového kompletního důkazu Velké Fermatovy věty. Všechny tyto kopie budou prodávány s cenou 10 Kč za kus. Kolik kopií by mělo být uděláno, aby se maximalizoval zisk? Zformulujte tento problém jako úlohu ($\mathbb{Z}$LP).

3.6.39. Uvažme následující hru se slovy (znáte ji?). Máte jistý počet kostiček, z nich každá obsahuje právě jedno z písmen a, …, z. Pro každé písmeno α máte $N_\alpha \in \mathbb{N}_0$ kostiček s písmenem α. Celkové můžete poskládat slova $w_1,\dots,w_n$ (tento seznam obsahuje např. všechna slova ze Slovníku spisovné češtiny. Každé slovo můžete složit nejvýše jednou a libovolná kostička může být také použita nejvýše jednou. Dostanete $p_j \geq 0$ bodů za slovo w_j a navíc bonus $b_{ij} \geq 0$ za současné složení slov w_i a w_j ($i, j = 1,\dots,n$). Zformulujte problém pro určení vítězné strategie jako úlohu ($\mathbb{Z}$LP).

3.6.40. Firma Krach, a.s., dala do „aukce“ N předmětů. Zájemci odeslali své nabídky pro vybrané předměty z nabídky. Aukční dům tak obdržel n nabídek, konkrétně nabídku b_j pro výběr S_j, $j = 1,\dots,n$. Sestavte ($\mathbb{Z}$LP) tak, aby byly určeny vítězné nabídky maximalizující zisk, přičemž každý z N předmětů může být prodán nejvýše jednomu zájemci.

3.6.41. Brněnské flaškárny, a.s., mají 2 stroje na výrobu lahví. Na začátku každého roku je potřeba sestavit plán údržby. Údržba každého stroje trvá 2 měsíce. Navíc v červenci a srpnu je k dispozici pouze polovina zaměstnanců, takže lze využít pouze jeden stroj v tomto období. Měsíční poptávka po lahvích je d_j, $j = 1,\ldots,12$. Stroj k, $k = 1,2$, vyrobí nejvýše a_k lahví měsíčně, k čemuž potřebuje l_k pracovních dní. V j-tém měsíci je k dispozici L_j pracovních dní, $j = 1,\ldots,12$. Zformulujte tento problém jako úlohu (ℤLP) tak, aby byla uspokojena daná poptávka a

(a) minimalizovaná celková suma změn použitých pracovních dní ve dvou po sobě jdoucích měsících,

(b) minimalizovaná největší změna použitých pracovních dní ve dvou po sobě jdoucích měsících.

3.6.42. Firma Bazmek, s.r.o., vyrábí 2 produkty. K tomu má jednu továrnu, dvě distribuční centra a 5 hlavních zákazníků, jejichž poptávka po k-tém výrobku je d_{jk}, $j = 1,\ldots,5$, $k = 1,2$. Ovšem distribuční centrum je schopno pracovat pouze s jedním druhem výrobku. Jak je přiřadit, mají-li být dosaženy minimální náklady, které zahrnují

(a) pevné náklady f_{ik}, je-li k-tý výrobek přiřazen i-tému centru,

(b) pevné náklady f_{ijk}, je-li poptávka j-tého zákazníka po k-tém výrobku uspokojena i-tým centrem,

(c) jednotkové doručovací náklady c_{ijk} pro zaslání k-tého výrobku k j-tému zákazníkovi z i-tého centra.

Sestavte odpovídající úlohu (ℤLP). Jak se úloha změní, pokud bude možné, aby obě centra zajišťovala doručení obou produktů?

3.6.43. Rozdělme rok na n období. Firma Udělátko, v.o.s., vyrábí jediný výrobek s poptávkou d_k v k-tém období. Firma však nemá vlastní výrobní stroj. Jestliže se firma rozhodne spustit výrobu v k-tém období, musí zaplatit inicializační poplatek f_k, jehož výše nezávisí na vyrobeném množství, a jednotkové náklady p_k. V každém období může být vyrobeno libovolné množství výrobku a odběr zboží probíhá vždy na konci měsíce. V případě nadbytku zboží z období k do období $k+1$ musíme zaplatit jednotkový skladovací poplatek s_k. Počáteční stav skladu je s_0. Sestavte odpovídající (smíšený) problém (ℤLP) s plánem výroby minimalizujícím celkové náklady.

3.6.44. Pro MS v piškvorkách z Příkladu 3.6.36 bychom také potřebovali vyrobit ukázkovou sadu medailí, tj. tři různé kusy. K této výrobě se používají 3 různé stroje (odlívání, broušení a leštění). Každá medaile musí projít nejdříve prvním strojem, pak druhým a konečně třetím strojem. Pořadí medailí na strojích se ale může lišit. Čas potřebný pro práci i-tého stroje na j-té medaili je t_{ij} a je celočíselný. Cílem je minimalizovat čas potřebný na výrobu celé sady. Zformulujte tento problém jako úlohu (ℤLP). A co kdybychom chtěli, aby pořadí na jednotlivých strojích bylo stejné?

(Nápověda: označte x_{ij} jako čas začátku práce j-tého stroje na i-té medaili. Musíte zamezit tomu, aby se dvě medale sešly u téhož stroje ve stejnou chvíli. Současně není možné, aby stroj $(j+1)$ pracoval na medaili před dokončením práce j-tého stroje na medaili.)

3.6.45. Metelesku & Blesku je dopravní společnost. Má kamionový vozový park, s nímž přepravuje zboží z depa ($i = 0$) do n distribučních center ($i = 1,\ldots,n$). Označme

- Q_k … nákladní kapacita k-tého vozidla ve vozovém parku ($k = 1,\ldots,m$),
- d_i … počet položek k doručení do i-tého distribučního centra ($i = 1,\ldots,n$),
- t_i^k … čas potřebný pro vyložení zásilky do i-tého distribučního centra z k-tého vozidla ($i = 1,\ldots,n$; $k = 1,\ldots,m$),
- t_{ij}^k … čas pro cestu k-tého vozidla z i-tého do j-tého distribučního centra ($i,j = 0,\ldots,n$; $k = 1,\ldots,m$),

- c_{ij}^k … náklady na cestu k-tého vozidla z i-tého do j-tého distribučního centra ($i, j = 0, \ldots, n$; $k = 1, \ldots, m$).

Navštíví-li vozidlo i-té distribuční centrum, pak sem doručí kompletní zásilku d_i ($i = 1, \ldots, n$). Navíc každé distribuční centrum může navštívit pouze jedno vozidlo, a to pouze jednou. Rozpis cest také musí splňovat kritéria kvality, podle kterých trasa k-tého vozidla trvá nejvýše T_k jednotek času.

Zaveďme proměnné

$$x_{ij}^k = \begin{cases} 1, & \text{má-li } k\text{-té vozidlo jet z } i\text{-tého do } j\text{-tého distribučního centra,} \\ 0, & \text{v ostatních případech.} \end{cases}$$

Zformulujte odpovídající (ZLP) pro minimalizaci celkových nákladů, přičemž musí být splněny požadavky na maximální přepravní čas a vozidla nesmí být přetěžována. Kolik proměnných a omezení by bylo potřeba v případě 400 distribučních center a 20 vozidel.

(Toto je ukázková úloha problému m-tice obchodních cestujících.)

3.6.46. Problémy s plánováním ve sportu spočívají především v určení data a místa, kde se bude hrát každý zápas turnaje či ligy. Při řešení různých variant tohoto problému lze však úspěšně využít nástroje celočíselného programování a lze se s tím setkat v profesionálních ligách různých sportovních disciplín, jako je fotbal, baseball, basketbal, kriket či hokej. Brazilská fotbalová konfederace (CBF) každoročně pořádá fotbalový turnaj, který je nejvýznamnější sportovní událostí v zemi. Jeho hlavním sponzorem je TV Globo, což je největší mediální skupina a televizní síť v Brazílii. Jenže hlavním problémem pro atraktivitu a důvěryhodnost celého turnaje jsou spravedlivé a vyvážené rozpisy pro všechny týmy. Do toho také vstupují televizní sponzoři, kteří podmiňují svou podporu harmonogramy, které umožňují vysílat nejdůležitější zápasy na otevřených kanálech. Nesmíme ale zapomenout ani na to, že ve velkých městech hostících dva nebo více týmů a velký počet fanoušků jsou potřeba některá bezpečnostní omezení, aby se předešlo střetům fanoušků před nebo po zápasech.

Turnaj trvá sedm měsíců (od května do prosince). Hraje jej 20 týmů systémem každý s každým doma–venku, tj. celkem se odehraje 38 kol. Zápasy víkendových kol se hrají v sobotu a neděli, zatímco zápasy kol uprostřed týdne se hrají ve středu a ve čtvrtek. Volné termíny pro hraní zápasů se rok od roku mění a musí být koordinována s jinými soutěžemi, jako jsou regionální turnaje, Brazil's Cup, South America's Cup a Santander Libertadores Cup apod. Platí, že pokud se zápas mezi týmy A a B hraje v místě A a kole $k \in \{1, \ldots, 19\}$, pak druhý zápas se hraje v místě B a v kole $19 + k$. Atraktivnější zápasy a ty, do kterých se zapojí nejsilnější týmy, by se měly hrát co nejvíce o víkendech. Týmy jsou organizovány do dvojic s komplementárními modely zápasy doma–venku. Obvykle týmy ve stejném páru sídlí ve stejném domovském městě. Týmové dvojice jsou definovány CBF před sestavením rozpisu a mohou se rok od roku lišit, protože zúčastněné týmy se pochopitelně mění díky postupu z nižších divizí a sestupu do nižších divizí na konci každé sezóny.

Harmonogram by měl splňovat řadu tvrdých a měkkých omezení, od férovosti k bezpečnostním otázkám a od technických až po kritéria vysílání. Většina z nich odráží strategie pro maximalizaci výnosů a atraktivitu turnajů, zatímco jiné se pokoušejí vyhnout se neférovým situacím, které by mohly prospět jednomu nebo druhému týmu pomocí pohodlnějšího pořadí zápasů. Tyto požadavky byly diskutovány a určeny v průběhu let týmy, federacemi, městskými správci, bezpečnostními agenturami a sponzory.

Pro více podrobností viz

https://ln.math.muni.cz/pz/opt/zlp-fotbal

a také

https://ln.math.muni.cz/pz/opt/zlp-sport.

Kapitola 4

Kvadratické programování

4.1 *Úvod a motivace* 320
4.2 *Když stačí tužka a pravítko…* 328
4.3 *Nepříliš zevrubný teoretický vhled* 332
4.4 *Hildrethova–d'Esopova metoda* 335
4.5 *Wolfeho metoda v krátkém tvaru* 340
4.6 *Wolfeho metoda v dlouhém tvaru* 346
4.7 *Než opustíme kvadratické programování…* 359

THE RISK OF A PORTFOLIO IS NOT A LINEAR FUNCTION OF THE VECTOR OF ITS COMPONENTS. RATHER, THE VARIANCE OF A PORTFOLIO IS A QUADRATIC FUNCTION OF ITS COMPOSITION. THIS THWARTS THE INTUITION OF MOST ANALYSTS AND INVESTORS. INDEED, THE NATURE OF RISK MAY BE THE SINGLE MOST IMPORTANT ARGUMENT FOR THE USE OF QUANTITATIVE ANALYSIS IN INVESTMENT MANAGEMENT. NEITHER INVESTORS NOR ANALYSTS CAN BE BLAMED FOR THIS FACT. NOR CAN HARRY MARKOWITZ. NATURE MADE RISK A QUADRATIC FUNCTION. MARKOWITZ ONLY DISCOVERED IT.[1]

WILLIAM SHARPE[2]

Viz Sharpeho webovou stránku https://ln.math.muni.cz/pz/opt/sharpe1.

[1] *Riziko portfolia není lineární funkcí vektoru jeho složek. Spíše než to je rozptyl portfolia kvadratickou funkcí jeho skladby. To boří intuici většiny analytiků a investorů. Vskutku podstata rizika může být jediným nejdůležitějším argumentem pro použití kvantitativní analýzy v řízení investic. Za tuto skutečnost nelze obviňovat ani investory, ani analytiky. Ani Harry Markowitz za to nemůže. Příroda učinila riziko kvadratickou funkcí. Markowitz to pouze objevil.*

[2] William Forsyth Sharpe (∗ 16. června 1934) je americký ekonom, který působí jako emeritní profesor na Stanfordově univerzitě. V roce 1990 získal společně Harrym Markowitzem a Mertonem Millerem *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za jejich „průkopnickou práci v oblasti finanční ekonomie". Je také jedním z tvůrců modelu oceňování kapitálových aktiv. Je autorem tzv. *Sharpeho poměru* pro měření rizikově očištěného výnosu, přispěl k rozvoji binomiální metody oceňování opcí a také gradientní metody pro optimalizaci alokace aktiv.

4.1 Úvod a motivace

Nejistota. To je aspekt, který jsme v deterministických modelech z předchozích kapitol zcela ignorovali. Sice jsme ukázali, že úlohy s lineární účelovou funkcí mají vskutku velmi široký praktický záběr, ale samozřejmě nemohou být všeobjímající, neboť jejich koeficienty jsou mnohdy dány jen jako střední hodnoty odpovídajících veličin, což je v některých situacích zcela zavádějící. Jedním z takových problémů je otázka výběru cenných papírů pro investiční účely, která vzbuzuje dlouhodobě velký zájem. V Příkladu 2.3.3 jsme využili lineární programování k nalezení maximální očekávané hodnoty zisku vzhledem k daným omezením a dostupnému rozpočtu. Jenže taková formulace zcela ignoruje nejistotu/riziko očekávaného výnosu. Toho si (asi jako první) všiml Markowitz[3] a vybudoval teorii hledání optimálního portfolia založenou na tom, že je sice žádoucí mít vysoký očekávaný výnos, ale současně je nežádoucí mít vysoký rozptyl výnosu (jakožto měřítka jeho rizika), viz [362]. Přičemž ale poměr mezi výší výnosu a mírou rizika je již zcela individuální a závisí pouze na preferencích jednotlivých investorů. Takovéto problémy lze formulovat až jako úlohy kvadratického programování, v nichž jde o nalezení těch portfolií, která představují nejlepší vyvážení výnosu a rizika. Ty tvoří tzv. *efektivní hranici* nebo *E.V. hranici*, kde „E“ značí očekávaný čistý výnos a „V“ rozptyl výnosu. Tímto vyvážením rozumíme taková portfolia, která minimalizují rozptyl při daném výnosu nebo maximalizují výnos při daném rozptylu. E.V. hranici získáme tak, že rozptylu připisujeme různou váhu od nulové hodnoty až po nekonečno, přičemž v případě nulové váhy se celý problém redukuje na úlohu lineárního programování. Uživatel si pak následně může vybrat konkrétní bod na E.V. hranici podle svých osobních preferencí. Avšak Markowitzův model se zdá být ne zcela vhodný pro malé investory, kteří si vzhledem k různým poplatkům a provizím za nákup a prodej jednotlivých položek nemohou dovolit tak diverzifikované portfolio jako podílové fondy. Proto je v jejich případě účelné dodatečně omezit počet nakoupených cenných papírů, což nás přivede až k úlohám (smíšeného) celočíselného kvadratického programování. Navíc zajímavé zjednodušení v tomto směru přináší požadavek i na stejnou výši investic do všech cenných papírů v portfoliu, díky čemuž se podmínka celočíselnosti redukuje na binárnost jednotlivých proměnných.

Úlohou *kvadratického programování* proto rozumíme minimalizaci (příp. maximalizaci) kvadratické funkce na množině vymezené soustavou lineárních rovnic a/nebo nerovnic, tj.

$$\left.\begin{array}{r} \tfrac{1}{2}\langle Cx,x\rangle + \langle d,x\rangle \to \min \\ x\in X := \{x\in P \mid \langle a_i,x\rangle \le b_i,\ i=1,\ldots,k\ \ \&\ \langle a_i,x\rangle = b_i,\ i=k+1,\ldots,m\ \ \& \\ \& \ x_1\ge 0\ \&\ldots\ \&\ x_s\ge 0\}\end{array}\right\} \qquad \text{(QP)}$$

kde $P\subseteq\mathbb{R}^n$ je daná množina[4] a $C\in\mathbb{R}^{n\times n}$, $d\in\mathbb{R}^n$, $a_1,\ldots,a_m\in\mathbb{R}^n$, $b_1,\ldots,b_m\in\mathbb{R}$, $k\in\{0,1,\ldots\}$, $m\in\{k,k+1,\ldots\}$ a $s\in\{0,\ldots,n\}$ jsou dané matice, vektory, čísla a indexy. Navíc nedílnou součástí úlohy (QP) je požadavek symetričnosti a pozitivní semidefinitnosti matice C, tj. $C=C^\top\ge 0$. Díky tomuto požadavku máme i tentokrát úlohu konvexního programování, což nám umožňuje využít některé teoretické výsledky z pozdější Kapitoly 6. Úloha (QP) je pochopitelně zobecněním úlohy (2.2.1), která odpovídá volbě $C=0$.

Ačkoli úlohy kvadratického programování byly ve druhé polovině 50. let a následně i v 60. letech minulého století velmi populární (z této doby jsou také základy algoritmů, které si ukážeme), není jim v dnešní době věnována až tak velká pozornost jako v případě lineárního programování. Důvod je především ten, že pouze málo úloh získaných jako matematické modely optimalizačních problémů z praxe splňuje podmínku konvexnosti/konkávnosti (tj. $C\ge 0$ pro minimalizaci či $C\le 0$ pro maximalizaci). Naším cílem proto bude představení pouze dvou algoritmů pro řešení úlohy (QP): (i) první z nich je založen na duální úloze – bude početně celkem snadný (dokonce bude sám korigovat případně početní chyby), ale bude vyžadovat $C>0$; (ii) druhý bude založen na simplexovém algoritmu a Karushových–Kuhnových–Tuckerových

[3] Harry Max Markowitz (∗ 24. srpna 1927) je americký ekonom. V roce 1989 získal *John von Neumann Theory Prize* za jeho příspěvky k rozvoji tří oblastí: teorie portfolia, metod pro řídké matice v lineárním programování a simulačního programovacího jazyku SIMSCRIPT, viz [362, 363, 365–367]. O rok později obdržel společně Williamem Sharpem a Mertonem Millerem *Cenu Švédské národní banky za rozvoj ekonomické vědy na památku Alfreda Nobela* za jejich „průkopnickou práci v oblasti finanční ekonomie“. V současnosti je profesorem na Kalifornské univerzitě v San Diegu. Je znám především díky jeho práci věnované moderní teorii portfolia. Tu započal články ze šedesátých let, kdy pracoval ve společnosti RAND. Zde se setkal s Dantzigem a s jeho pomocí pokračoval ve studiu optimalizačních metod a dalším rozvoji algoritmu určení kritické přímky pro nalezení portfolia s optimální hodnotou rozptylu, přičemž využíval toho, co je dnes známo jako *Markowitzova hranice*. V roce 1954 obdržel doktorát na Univerzitě v Chicagu s disertační prací věnovanou právě teorii portfolia. Toto téma bylo tak nové, že během obhajoby práce se s ním Milton Friedman dohadoval o tom, zda jeho příspěvek je vůbec o ekonomii. Během let 1955–1956 pobýval na *Cowlesově nadaci* (*Cowles Foundation*) na Yaleově univerzitě. V roce 1956 publikoval článek ohledně kritické přímky a čas strávený v nadaci využil k napsání knihy o teorii portfolia, která vyšla v roce 1959. Během 50. let dospěl Markowitz stejně jako mnoho ostatních k názoru, že mnohé praktické problémy se vymykají analytickému řešení a že jsou potřebné simulační techniky. Ve společnosti RAND se proto účastnil vytváření velkých logistických modelů simulace. Počátkem 60. let vytvořil programovací jazyk, který byl později pojmenovaný SIMSCRIPT. Místo toho, aby popisoval kroky, které musí počítač učinit, aby uskutečnil simulaci, tento jazyk umožnil programátorovi popsat určitým stylizovaným způsobem systém, který měl být simulován, což omezilo čas potřebný na programování.

[4] Množina P se využívá zejména ve chvíli, kdy máme některá omezení, která nechceme/nemůžeme zakomponovat mezi omezující rovnosti či nerovnosti. Typickým pří-

kladem takové množiny je volba

$$P = \{x \in \mathbb{R}^n \mid x_1 \geq 0, \ldots, x_s \geq 0\}.$$

Pokud nic takového není potřeba, volíme $P = \mathbb{R}^n$.

podmínkách – ten bude sice početně náročnější, ale dovolí nám pracovat s libovolnou maticí $C \geq 0$. Ještě předtím si ale ukážeme tři klasické úlohy vedoucí na problémy kvadratického programování.

Regrese (metoda nejmenších čtverců s omezeními)**.** V této úloze chceme najít přímku, která nejlépe (tj. v normě) aproximuje daná data, přičemž o některých parametrech a priori víme, že splňují jistá omezení. Uvažme tedy lineární regresní model

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k,$$

kde $\beta_0, \ldots, \beta_k$ jsou *regresní koeficienty*. Označme y_i jako výsledek i-tého pozorování závislé proměnné při hodnotách x_{ij} nezávislých proměnných, tj.

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + e_i, \quad i = 1, \ldots, n,$$

kde e_i je chyba, která popisuje odchylku pozorování od skutečné regrese (tj. hledané podmíněné střední hodnoty). K tomu, abychom docílili nejlepšího možného modelu, je potřeba odhadnout hodnoty $\beta_0, \ldots, \beta_k$ tak, aby byla chyba nejmenší možná. Toho bychom mohli docílit pouze pomocí absolutní hodnoty[5], tj. řešením úlohy

$$\sum_{i=1}^{n} |e_i| \to \min.$$

[5] Chceme stejně „vážit" kladné i záporné odchylky, proto s absolutní hodnotou. Bez ní by se nám odchylky s opačnými znaménky anulovaly.

Taková funkce je sice konvexní a my bychom dokázali takovou úlohu převést na problém lineárního programování, jenže to by hledání koeficientů $\beta_0, \ldots, \beta_n$ výrazně zkomplikovalo (a v dobách před objevem simplexového algoritmu v podstatě znemožnilo). Obvyklým nástrojem pro řešení této úlohy je derivování, což ale v případě absolutní hodnoty narazí na její nediferencovatelnost. Proto se k určení koeficientů regresní přímky užívá *metoda nejmenších čtverců*[6], ve které chceme minimalizovat sumu druhých mocnin jednotlivých chyb, tj.

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2 \to \min,$$

[6] Kromě eliminace problému s nediferencovatelností účelové funkce tímto dojde i ke zvýšení důrazu na menší odchylky, tj. váha odchylek v rozsahu $(-1, 1)$ se v účelové funkci zmenšuje čím blíže jsou k nule, zatímco váha odchylek s absolutní hodnotou větší než jedna se s rostoucí hodnotou zvětšuje.

což je kvadratická funkce vzhledem k proměnným $\beta_0, \ldots, \beta_k$. S využitím vektorů $\beta \in \mathbb{R}^{k+1}$, $Y \in \mathbb{R}^n$ a matice $X \in \mathbb{R}^{n \times (k+1)}$ daných jako

$$\beta := \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X := \begin{pmatrix} 1 & x_{11} & \ldots & x_{1k} \\ 1 & x_{21} & \ldots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \ldots & x_{nk} \end{pmatrix}$$

lze tuto úlohu zapsat ve formě minimalizace euklidovské normy

$$\| Y - X\beta \|^2 \to \min$$

neboli

$$(Y - X\beta)^\top (Y - X\beta) = \beta^\top X^\top X\beta - \beta^\top X^\top Y - Y^\top X\beta + Y^\top Y = \beta^\top X^\top X\beta - 2\, Y^\top X\beta + Y^\top Y \to \min,$$

což po zanedbání konstantního členu $Y^\top Y$ odpovídá účelové funkci z (QP) s volbou $C = 2\, X^\top X$ a $d = -2\, Y^\top X$, která jistě splňuje podmínku pozitivní semidefinitnosti matice C. V praxi je však mnohdy nutné přidat lineární omezení pro některé koeficienty β_i, tj. např.

$$\sum_{j=0}^{k} a_{rj} \beta_j = b_r, \quad r = 1, \ldots, m$$

$$\beta_0 \geq 0 \quad \& \quad \beta_1 \geq 0 \quad \& \quad \cdots \quad \& \quad \beta_k \geq 0$$

pro daná čísla $a_{10}, \ldots, a_{mk} \in \mathbb{R}$, což je přesně úloha (QP) v kanonickém tvaru. Nebo případně můžeme požadovat, aby hodnoty β_j byly pouze v nějakém omezeném rozsahu, tj. $\beta_j \in [\alpha_j, \gamma_j]$.

Příklad 4.1.1

Uvažme m-tici demografických skupin, které chceme oslovit reklamou, přičemž požadovaný počet zhlédnutí/zaujetí/„líbítek“ v každé skupině je určen vektorem $v^{\text{pož}}$ s kladnými koeficienty. K tomuto účelu využijeme n-tici kanálů vhodných pro šíření reklamy (různé webové platformy, televize, rádio, tisk, …). Matice $\mathcal{R} \in \mathbb{R}^{m\times n}$ vyjadřuje vztah mezi zhlédnutími v jednotlivých skupinách a 1 Kč investovanou do jednotlivých kanálů, tj. hodnota $\mathcal{R}_{ij}$ udává počet zhlédnutí v i-té skupině při investování 1 Kč do do j-tého reklamního kanálu (tyto hodnoty jsou samozřejmě odhadnuty zcela kvalifikovaně a jsou vždy kladné). Pak j-tý sloupec matice $\mathcal{R}$ charakterizuje efektivitu nebo dosah j-tého kanálu za 1 Kč, zatímco i-tý řádek ukazuje, jak jsou které kanály exponované v i-té skupině.

Celkový počet zhlédnutí v každé demografické skupině je dán vektorem $v = Rs \in \mathbb{R}^m$, kde $s \in \mathbb{R}^n$ popisuje investice do jednotlivých kanálů. Cílem pak je nalezení takového vektoru s, aby platilo $v = Rs \approx v^{\text{pož}}$. K tomu můžeme velmi dobře využít metodu nejmenších čtverců, v níž chceme určit vektor s takový, že

$$\|Rs - v^{\text{pož}}\|^2 \to \min.$$

Avšak tímto nemůžeme a priori zaručit nezápornost složek vektoru s a také splnění našeho rozpočtového omezení. Proto je nutné úlohu doplnit o podmínky

$$s_1 + \cdots + s_n = B,$$

kde B určuje celkový dostupný rozpočet[7], tj. celkem dostáváme úlohu

$$\begin{aligned} \|Rs - v^{\text{pož}}\|^2 &\to \min \\ s_1 + \cdots + s_n &= B \\ s_1 \geq 0 \quad \& \quad \cdots \quad \& \quad s_n &\geq 0. \end{aligned}$$

[7] Chceme vlastně optimálně rozdělit naše dostupné zdroje (finance) mezi jednotlivé kanály.

Úloha může mít i výrazně fyzikální charakter.

Příklad 4.1.2

Uvažme 10-rozměrný vektor f, který popisuje posloupnost sil aplikovaných vždy po dobu 1 sekundy na hmotný bod jednotkové hmotnosti na přímce bez tření. Na počátku máme nulovou rychlost a pozici. Podle Newtonových zákonů[8] bude finální rychlost a pozice hmotného bodu

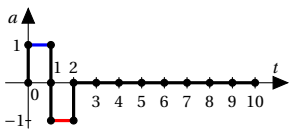
$$\begin{aligned} v^F &= f_1 + \cdots + f_{10} \\ p^F &= \tfrac{19}{2} f_1 + \tfrac{17}{2} f_2 + \cdots + \tfrac{1}{2} f_1. \end{aligned}$$

Naším úkolem je najít takovou posloupnost, aby $v^F = 0$ a $p^F = 1$, tj. posloupnost sil, které zanechají hmotný bod nehybný v pozici 1 m vpravo. Takových sil samozřejmě existuje mnoho, můžeme je např. vyjádřit jako vektor

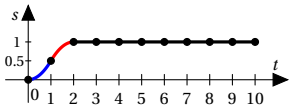
$$f^{bb} = (1, -1, 0, \ldots, 0)^\top,$$

který akceleruje hmotný bod na průměrnou rychlost 0,5 m/s po 1 sekundě a potom jej během druhé sekundy zpomalí na nulovou rychlost, čímž dosáhneme právě pozice 1 m vpravo. Dalším aplikováním nulových sil dosáhneme požadovaného cíle. Horní index „bb“ v označení tohoto vektoru odkazuje na tzv. *bang-bang*, což znamená, že nejdříve aplikujeme velkou sílu, abychom hmotný bod rozpohybovali (tj. první „bang“), a poté jinou velkou silou (druhé „bang“), abychom jej zpomalili na nulovou rychlost, viz Obrázky 4.1–4.2.

[8] Platí $F = ma$, takže vzhledem k jednotkové hmotnosti máme rovnost mezi zrychlením a působící sílou $F = a$. Protože $s(t) = v_0\, t + \frac{1}{2} a\, t^2$ a $v(t) = v_0 + a\, t$, kvůli nulové počáteční rychlosti dostaneme aplikováním síly f_1 v čase $t = 0$ po jedné sekundě rychlost $v(1) = a = f_1$. Aplikováním síly f_2 v čase $t = 1$, kdy už máme rychlost f_1, bude rychlost $v(2) = v_1 + a f_2 = f_1 + f_2$ atd. Současně po první sekundě urazíme dráhu $s(1) = f_1/2$. Po další sekundě to už bude $s(2) = f_1/2 + f_1 + f_2/2 = 3f_1/2 + f_2/2$ a o další sekundu později $s(3) = 3f_1/2 + f_2/2 + (f_1 + f_2) + f_3/2 = 5f_1/2 + 3f_2/2 + f_3/2$ atd.



Obrázek 4.1: Bang-bang akcelerace.



Obrázek 4.2: Bang-bang dráha.

My bychom ale rádi zjistili, jaký „nejmenší“ vektor sil povede k $v^F = 0$ a $p^F = 1$, kde „nejmenší“ je určeno součtem čtverců jednotlivých aplikovaných sil, tj.

$$\|f\|^2 = f_1^2 + \cdots + f_{10}^2.$$

Toto je vlastně *podmíněný problém nejmenší normy*[9]

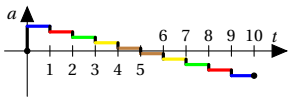
$$\|f\|^2 \to \min$$
$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ 19/2 & 17/2 & \dots & 1/2 \end{pmatrix} f = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

s hledaným vektorem $f = (f_1, f_2, \dots, f_{10})^\top$. Řešením je vektor/posloupnost sil

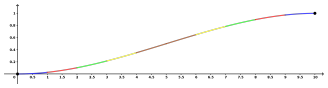
$$f^* = \big(9/165,\ 7/165,\ 5/165,\ 3/165,\ 1/165,\ -1/165,\ -3/165,\ -5/165,\ -7/165,\ -9/165\big)^\top,$$

viz Obrázky 4.3–4.4. Hodnota $\|f^*\|^2 = 2/165$ je ve výrazném kontrastu se čtvercem normy „bang-bang“ vektoru $\|f^{bb}\|^2 = 2$, která je 165-krát větší.

[9] To je speciální případ metody nejmenších čtverců, kde máme $Y = 0$ a $\beta_0 = 0$, v kterémžto případě můžeme vzít $X = I$.



Obrázek 4.3: Optimální akcelerace.



Obrázek 4.4: Optimální dráha.

Ještě si můžeme uvést jednu praktickou úlohu stejného charakteru, ale s jiným obsahem.

Příklad 4.1.3

Masarykova univerzita by ráda odhadla zaměstnanost svých absolventů. Pro jednoduchost se omezme pouze na Přírodovědeckou fakultu a připusťme, že zaměstnání každého absolventa můžeme zařadit do jedné z kategorií: (i) veřejný sektor, (ii) průmysl a (iii) akademická sféra. Zároveň ignorujeme nezaměstnanost, která beztak u absolventů PřF ani není možná (při nejmenším ne z dlouhodobého hlediska). Označme symbolem a_j počet absolventů v j-tém roce pro $j = 1, \dots, n$ a v_j, p_j, z_j počty absolventů v j-tém roce, kteří nastoupí do veřejné sféry (v_j), průmyslu (p_j) nebo na akademickou pozici (a_j), tj. $a_j = v_j + p_j + z_j$.

Jeden z uvažovaných modelů předpokládá, že v každém roce se k jednotlivým zaměstnaneckým kategoriím vždy připojí jisté části populace absolventů. Označme tyto podíly postupně jako λ_1, λ_2 a λ_3. Potom odhadované hodnoty nástupů do jednotlivých kategorií zaměstnání v j-tém roce jsou dány jako

$$\widehat{v}_j = \lambda_1 a_j \quad \& \quad \widehat{p}_j = \lambda_2 a_j \quad \& \quad \widehat{z}_j = \lambda_3 a_j.$$

Smysluplným měřítkem pro platnost modelu pak může být rozdíl mezi aktuálními hodnotami v_j, p_j, z_j a předpovězenými hodnotami $\widehat{v}_j$, $\widehat{p}_j$, $\widehat{z}_j$ jako v metodě nejmenších čtverců, tj.

$$\sum_{j=1}^{n} \left[(v_j - \widehat{v}_j)^2 + (p_j - \widehat{p}_j)^2 + (z_j - \widehat{z}_j)^2 \right] \to \min$$

(pochopitelně stále za předpokladu, že všichni absolventi jsou zaměstnáni v jedné ze tří uvažovaných kategorií). Tím se dostáváme k úloze kvadratického programování v proměnných λ_1, λ_2, λ_3 mající podobu

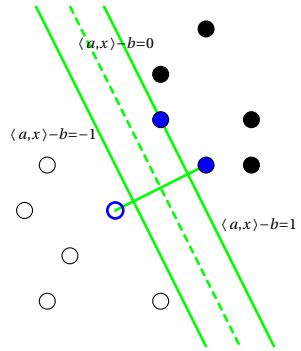
$$\sum_{j=1}^{n} \left[(v_j - \lambda_1 a_j)^2 + (p_j - \lambda_2 a_j)^2 + (z_j - \lambda_3 a_j)^2 \right] \to \min$$
$$\lambda_1 + \lambda_2 + \lambda_3 = 1$$
$$\lambda_1 \geq 0 \quad \& \quad \lambda_2 \geq 0 \quad \& \quad \lambda_3 \geq 0.$$

Support Vector Machines (SVM)[10]**.** Jedná se o metodu z tzv. *strojového učení*[11]. Jedná se o modely učení s učitelem, které analyzují data pro určitou klasifikaci nebo regresní analýzu. SVM je jednou z nejrobustnějších predikčních metod založenou na rámci statistického učení nebo

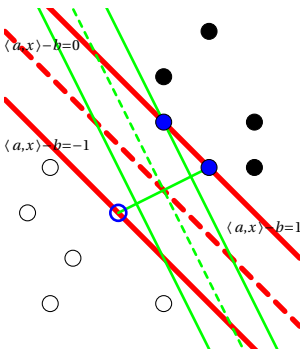
[10] Název souvisí s tzv. *opěrnými nadrovinami* (angl. *supporting hyperplanes*), o kterých se více dozvíme v Kapitole 6.

[11] To je podoblast *umělé inteligence*

tzv. „VC teorie“. Na základě ukázkové sady dat, v nichž je každý prvek přiřazen do jedné ze dvou tříd, SVM tréninkový algoritmus sestaví model, který každý nový příklad zařadí do jedné z tříd, což z něj činí nepravděpodobnostní binární lineární klasifikátor. SVM zobrazuje trénovací data na body v prostoru tak, aby se maximalizovala šířka mezery mezi těmito dvěma kategoriemi. Nová data jsou pak zobrazena do stejného prostoru a předurčena do jedné z tříd podle toho, do které části prostoru spadají. Jinými slovy chceme najít nadrovinu, která prostor „příznaků“ rozděluje tak, že trénovací data z různých tříd leží v jiném poloprostoru. Optimální nadrovinou pak bude ta, která maximalizuje součet vzdáleností obou tříd od oddělující nadroviny, tj. můžeme tam vytvořit nejširší možný pás neobsahující žádná data (to je tzv. *maximální odstup* nebo též *pásmo necitlivosti* či *hraniční pásmo*).

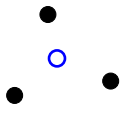


Obrázek 4.5: SVM pro dvě skupiny bodů. Nejbližší body jsou vyznačeny modrou barvou. Zelená skupina přímek je tvořena osou úsečky spojující dva nejbližší body a další dvě přímky vytváří nejširší možný pás (stejně široký na obě strany), který lze mezi obě skupiny bodů vložit (podobně by to dopadlo pro druhou dvojici nejbližších bodů).



Obrázek 4.6: Správné řešení SVM pro dvě skupiny bodů z předchozího obrázku je vyznačeno červenými přímkami (pás je očividně mnohem širší než ten vymezený zelenými přímkami). Body ležící na hranicích červeného pásu jsou tzv. *opěrné vektory* (angl. *support-vectors*.

[11] Pochopitelně v některých situacích nemusí mít úloha SVM řešení, protože jsme výrazně omezeni tím, že hledáme pouze nadrovinu. Jako například pro následující skupinu bodů.



Uvažme m-tici bodů v $\mathbb{R}^n$ a každý z nich přiřaďme do jedné z tříd označených jako $+1$ a -1. Tím získáme množinu dvojic $\{x_i, y_i\} \in \mathbb{R}^n \times \{\pm 1\}$ pro $i = 1, \dots, m$. Naším cílem je nalézt nadrovinu v $\mathbb{R}^n$ s rovnicí $\langle a, x\rangle = b$ pro vhodné $a \in \mathbb{R}^n \setminus \{0\}$ a $b \in \mathbb{R}$, která bude oddělovat body z obou tříd, tj. bude platit

$$\begin{aligned} \langle a, x_i\rangle - b < 0, &\quad \text{je-li odpovídající } y_i = -1, \\ \langle a, x_i\rangle - b > 0, &\quad \text{je-li odpovídající } y_i = +1. \end{aligned}$$

Jelikož se tyto nerovnosti nezmění po vynásobení dvojice (a, b) libovolným kladným číslem, můžeme ekvivalentně říci, že hledáme a, b splňující

$$\begin{aligned} \langle a, x_i\rangle - b \le -1, &\quad \text{je-li odpovídající } y_i = -1, \\ \langle a, x_i\rangle - b \ge 1, &\quad \text{je-li odpovídající } y_i = +1 \end{aligned}$$

neboli

$$y_i\big(\langle a, x_i\rangle - b\big) \ge 1 \quad \text{pro všechna } i = 1, \dots, m.$$

Body jsou odděleny pásem $\{x \in \mathbb{R}^n \mid -1 \ge \langle a, x\rangle - b \ge 1\}$ jako na Obrázku 4.5. V úloze SVM však chceme nalézt nejen oddělující nadrovinu, ale i maximalizovat šířku tohoto pásu, která je rovna $2/\|a\|$. Ovšem to je stejné jako

$$\|a\|^2 = \langle a, a\rangle \to \min$$

za podmínek

$$y_i\big(\langle a, x_i\rangle - b\big) \ge 1 \quad \text{pro všechna } i = 1, \dots, m,$$

což je opět úloha kvadratického programování[12], viz také Obrázek 4.6.

Tvorba portfolia[13]**.** Asi nejdůležitějším příkladem úloh kvadratického programování je tvorba portfolia. Ten získáme zobecněním úlohy lineárního programování, viz Příklad 2.3.3, jestliže jednotlivé koeficienty $c_1, \dots, c_n$ budeme považovat spíše za náhodné veličiny (původně udávaly hodnotu nebo výnos z držení i-tého aktiva). Mají-li tyto náhodné veličiny střední hodnotu

$$d = (d_1, \dots, d_n)^\top$$

a kovarianční matici

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix},$$

pak v takovém případě nemá smysl maximalizovat $\langle c, x\rangle$, neboť to je náhodná veličina. Kdybychom při daných lineárních omezeních maximalizovali střední hodnotu $\langle d, x\rangle$, bylo by výsledné rozhodnutí x^* značně nespolehlivé, neboť by bylo odvozeno bez ohledu na variabilitu veličin $c_1, \dots, c_n$. Proto je mnohem výhodnější snažit se omezit rozptyl $\langle Cx, x\rangle$ náhodné veličiny $\langle c, x\rangle$ a maximalizovat funkci

$$f(x) := a\langle d, x\rangle - b\langle Cx, x\rangle,$$

[13] Portfolio Selection Problem

kde $a, b \geq 0$ jsou váhy, které vyjadřují důležitost, jaká se při rozhodování přikládá výši účelové funkce (koeficient a) a spolehlivosti rozhodnutí (koeficient b). Dokonce se stačí omezit pouze na váhy splňující $a + b = 1$.

Alternativní možností k tvorbě portfolia by mohla být minimalizace rozptylu a při dosažení alespoň požadovaného zisku Z a splnění rozpočtového omezení R, tj.

$$\begin{gathered}\langle Cx, x\rangle \to \min \\ \sum_{i=1}^{n} d_i x_i \geq Z \quad \& \quad \sum_{i=1}^{n} x_i \leq R \\ x_1 \geq 0 \quad \& \quad \ldots \quad \& \quad x_n \geq 0.\end{gathered}$$

Podmínku nezápornosti lze nahradit i specifičtějším požadavkem na investici do jednotlivých (příp. jen do některých) aktiv, tj. $\alpha_i \leq x_i \leq \beta_i$. Pro mnohem více informací viz např. [362, 363].

Příklad 4.1.4

Investor má k dispozici 1 milion korun, který může využít pro nákup 3 druhů akcií. Označme S_i náhodnou veličinu reprezentující roční výnos z 1 Kč investované do i-té akcie, tj. je-li $S_i = 0,12$, pak 1 Kč investovaná do této akcie na začátku roku by měla na konci téhož roku přinést nejhůře 1,12 Kč. Máme následující informace středních hodnot, rozptyl a kovariance náhodných veličin

$$\begin{gathered}\mathrm{E}(S_1) = 0,14 \quad \& \quad \mathrm{E}(S_2) = 0,11 \quad \& \quad \mathrm{E}(S_3) = 0,1 \\ \mathrm{Var}(S_1) = 0,2 \quad \& \quad \mathrm{Var}(S_2) = 0,08 \quad \& \quad \mathrm{Var}(S_3) = 0,18 \\ \mathrm{Cov}(S_1, S_2) = 0,05 \quad \& \quad \mathrm{Cov}(S_1, S_3) = 0,02 \quad \& \quad \mathrm{Cov}(S_2, S_3) = 0,03.\end{gathered}$$

Sestavme úlohu kvadratického programování, jejímž řešením nalezneme skladbu portfolia, které dosáhne alespoň 12% ročního zhodnocení a bude mít minimální rozptyl ročního zisku.

Řešení. Označme jako x_i množství peněz [v tisících Kč] investované do i-té akcie pro $i = 1, 2, 3$. Roční výnos takového portfolia pak je

$$\frac{x_1 S_1 + x_2 S_2 + x_3 S_3}{1000}$$

a očekávaný roční výnos je

$$\frac{x_1 \mathrm{E}(S_1) + x_2 \mathrm{E}(S_2) + x_3 \mathrm{E}(S_3)}{1000},$$

neboť $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$ a $\mathrm{E}(aX) = a\mathrm{E}(X)$. K tomu, aby očekávaný výnos byl alespoň 12 %, musíme požadovat

$$0,14\, x_1 + 0,11\, x_2 + 0,1\, x_3 \geq 1000 \times 0,12 = 120.$$

Současně musí také platit $x_1 + x_2 + x_3 = 1000$ a samozřejmě[14] $x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0$. Naším cílem je za těchto podmínek minimalizovat rozptyl portfolia, tj.

$$\begin{aligned}\mathrm{Var}(x_1 S_1 + x_2 S_2 + x_3 S_3) &= \mathrm{Var}(x_1 S_1) + \mathrm{Var}(x_2 S_2) + \mathrm{Var}(x_3 S_3) + \\ &\quad + 2\,\mathrm{Cov}(x_1 S_1, x_2 S_2) + 2\,\mathrm{Cov}(x_1 S_1, x_3 S_3) + 2\,\mathrm{Cov}(x_2 S_2, x_3 S_3) = \\ &= 0,2\, x_1^2 + 0,08\, x_2^2 + 0,18\, x_3^2 + 0,1\, x_1 x_2 + 0,04\, x_1 x_3 + 0,06\, x_2 x_3,\end{aligned}$$

což je úloha (QP) s $d = 0$.

[14] Nepovolujeme žádný *krátký prodej* (*short sale*), tj. spekulace na pokles ceny.

Ještě můžeme nabídnout trochu jiný pohled na tvorbu portfolia, která vychází z [80]. Chceme rozdělit náš dostupný rozpočet R mezi n-tici aktiv. Výsledek pak můžeme vyjádřit jako vektor

x splňující $x^\top e = x_1 + \cdots + x_n = 1$ pro $e := (1,\ldots,1)^\top$, takže jednotlivé hodnoty $x_1,\ldots,x_n$ popisují zlomkové části našeho rozpočtu investované do jednotlivých aktiv. Např. volba $x = e_j = (0,\ldots,0,1,0,\ldots,0)^\top$ znamená, že vše investujeme do j-tého aktiva. Součin $R\,x_j$ pak vyjadřuje množství peněz investované do j-tého aktiva, přičemž tentokrát můžeme povolit i $x_j < 0$, tj. spekulovat i na pokles ceny[15], v kterémžto případě prodáváme aktivum v celkovém objemu za $|R\,x_j|$. Např. $x = (-0{,}2; 0; 1{,}2)^\top$ znamená, že první aktivum prodáme v hodnotě 1/5 celkového rozpočtu a tuto částku společně s celým rozpočtem investujeme do třetího aktiva. Do druhého aktiva neinvestujeme vůbec nic. *Pákový efekt*[16] celého portfolia je roven

$$L = |x_1| + \cdots + |x_n|.$$

Jsou-li $x_1,\ldots,x_n \geq 0$ máme $L = 1$ a jedná se o tzv. *portfolio s dlouhým prodejem*[17]. Je-li některé z $x_1,\ldots,x_n$ záporné, pak $L > 1$. Např. má-li portfolio pákový efekt 5, znamená to, že na každou 1 Kč v portfoliu máme 3 Kč v dlouhém prodeji a 2 Kč v krátkém prodeji.

Nyní uvažme, že naše investice je na T období, z nichž každé představuje např. jeden den[18]. Výnos z investice můžeme popsat pomocí $T \times n$ matice V, kde V_{tj} představuje zlomkový výnos j-tého aktiva v t-tém období, tj. např. $V_{6,1} = 0{,}02$ znamená 2% výnos z prvního aktiva v šestém období, zatímco $V_{8,2} = -0{,}03$ znamená 3% ztrátu pro druhé aktivum v osmém období. Celý j-tý sloupec matice V představuje časovou řadu výnosů pro j-té aktivum a t-tý řádek popisuje výnos všech aktiv v t-tém období. Bývá obvyklé, že jedním z uvažovaných aktiv jsou peníze, které mají konstantní (kladný) výnos μ^{br}, kde index „br“ zkracuje *bezrizikové*[19]. Pokud tímto bezrizikovým je n-té aktivum, pak poslední řádek matice V je $\mu^{br} e^\top$. Připusťme, že na začátku t-tého období investujeme R_t, takže do j-tého aktiva investujeme $R_t\,x_j$. Na konci tohoto období bude hodnota j-tého aktiva rovna $(1 + V_{tj})\,R_t\,x_j$ a celého portfolia

$$R_{t+1} = \sum_{j=1}^{n} (1 + V_{tj})\,R_t\,x_j = R_t\,(1 + \tilde{v}_t\,x),$$

kde $\tilde{v}_t$ je t-tý řádek matice V. Předpokládejme, že $R_{t+1} > 0$. Kdyby celková hodnota portfolia byla záporná, řekneme, že portfolio „přišlo na buben“[20] a zastavíme obchodování.

Celkový relativní výnos portfolia během t-tého období, tj. relativní nárůst jeho hodnoty, je roven

$$\frac{R_{t+1} - R_t}{R_t} = \frac{R_t\,(1 + \tilde{v}_t\,x) - R_t}{R_t} = \tilde{v}_t\,x.$$

Nezapomínejme, že v tomto případě investujeme v každém období celou hodnotu portfolia rozdělenou podle x. To pak má za následek to, že nakupujeme a prodáváme aktiva tak, že jejich zlomková část v portfoliu je opět dána x. Tomu se říká *vyvažování*[21] portfolia. S využitím maticového značení můžeme výnos portfolia v každém z T období vyjádřit jako

$$v = V\,x,$$

kde v je T-rozměrný vektor výnosů portfolia v T obdobích[22], tj. máme časovou řadu výnosů portfolia. Je-li n-té aktivum bezrizikové a zvolíme-li $x = e_n$, pak $v = V\,e_n = \mu^{br} e$, tj. v každém období máme konstantní výnos μ^{br}.

Celková hodnota portfolia v t-tém období je

$$R_t = R_1\,(1 + v_1)\cdots(1 + v_{t-1}), \tag{4.1.1}$$

kde R_1 představuje celkový objem peněz investovaný v období $t = 1$. Tento součin je důsledkem toho, že v každém období investujeme veškerou aktuální hodnotu portfolia včetně ztrát či zisků v předchozích obdobích. V onom zjednodušeném případě, kdy poslední aktivum je bezrizikové a volíme $x = e_n$, celková hodnota vzroste na $R_t = R_1\,(1 + \mu^{br})^{t-1}$, což odpovídá *složenému úročení*[23] s úrokovou mírou μ^{br}. Když jednotlivé výnosy v_t jsou malé (v řádu jednotek procent) a počet období T není příliš vysoký (řekněme několik málo stovek), můžeme součin (4.1.1) aproximovat buď pomocí součtu nebo průměru jednotlivých výnosů. Roznásobením (4.1.1) dostaneme součet několika členů, z nichž skoro každý obsahuje součin několika

[15] Prodáváme aktivum nyní, ačkoli jej ještě nemáme s tím, že jej v blízké době (což může znamenat leccos) nakoupíme levněji, než za kolik jsme jej nyní prodali, a tedy se ziskem.

[16] Angl. *leverage*.

[17] Angl. *long-only portfolio*. Jde o portfolio založené pouze na spekulaci na růst ceny, tj. nyní nakoupíme a časem dráž prodáme.

[18] Ale může to být cokoli od hodin přes dny, týdny, měsíce atd.

[19] Angl. *risk–free*.

[20] Angl. *gone bust*.

[21] Angl. *re-balancing*.

[22] Zatímco $\tilde{v}_t$ je n-vektor popisující výnosy z n-tice aktiv v t-tém období.

[23] Angl. *compounded interest*.

výnosů – jediný člen bez těchto výnosů je R_1. Dále tam bude $t-1$ členů pouze s jedním výnosem ve tvaru $R_1\, v_s$ pro $s = 1, \ldots, t-1$. Všechny ostatní výrazy budou obsahovat součin alespoň dvou výnosů, a tak je můžeme zanedbat, neboť předpokládáme, že výnosy jsou malé. Toto vede k přibližnému vyjádření

$$R_t \approx R_1 + R_1\,(v_1 + \cdots + v_{t-1}),$$

což pro $t = T+1$ můžeme zapsat jako

$$R_{T+1} \approx R_1 + R\,\mathrm{avg}(r)\,R_1.$$

Tato aproximace ukazuje, že pro maximalizaci výsledné hodnoty portfolia bychom měli maximalizovat výnos, tj. maximalizovat hodnotu $\mathrm{avg}(r)$.

Volba vektoru x rozdělujícího prostředky mezi jednotlivá aktiva je posuzována podle výsledné časové řady výnosů $v = V\,x$. *Střední výnos*[24] portfolia pro T období (často zkráceně označován jen jako výnos) je určen právě hodnotou $\mathrm{avg}(r)$. *Rizikovost*[25] portfolia pro T období je směrodatnou odchylkou výnosu portfolia a značí se $\mathrm{std}(r)$. Veličiny $\mathrm{avg}(r)$ a $\mathrm{std}(r)$ udávají výnos a riziko na jedno období. Obvykle jsou přepočítány na období jednoho roku, čemuž se říká *roční výnos* a *riziko*[26] a jsou uváděny v procentech. Je-li v jednom roce P období, pak tyto roční hodnoty budou rovny

$$P\,\mathrm{avg}(r) \quad \text{a} \quad \sqrt{P}\,\mathrm{std}(r).$$

Např. uvažme situaci, že jedno období je jeden pracovní den. Těch je v roce (obvykle) 252, takže roční výnos a risk budou $252\,\mathrm{avg}(r)$ a $15{,}87\,\mathrm{std}(r)$. Proto při denním výnosu na úrovni 0,05 % a riziku 0,5 % bude roční výnos portfolia 12,6 % a jeho riziko 7,93 %.

Naším cílem je pochopitelně učit x tak, abychom dosáhli vysokého zisku při nízkém riziku. Chceme tedy, aby výnosy portfolia r_t byly dlouhodobě vysoké. V tomto případě máme ale „dvě“ účelové funkce, a budeme proto mít i jistou množinu řešení, které určují jistý kompromis mezi výnosy a rizikem. Např. je-li poslední aktivum bezrizikové, pak volbou $x = e_n$ dosáhneme portfolia s nulovým rizikem (méně to nelze) a výnosem μ^{br}. Je celkem očekávatelné, že jinými volbami x získáme portfolio s vyšším výnosem, avšak i s vyšším rizikem. Rozdělení portfolia, které minimalizuje riziko pro danou hodnotu výnosu (nebo maximalizuje výnos pro danou úroveň rizika) se nazývá *Paretovo*[27] *optimum*. Riziko a výnos této množiny rozdělení portfolia se obvykle znázorňují na grafu riziko-výnos, kde riziko je na vodorovné ose a výnos na svislé ose. Jednotlivá aktiva mohou být uvažována jako velmi jednoduchá portfolia odpovídající $x = e_j$. V takovém případě výnos a riziko portfolia jsou přímo výnosem a rizikem j-tého aktiva.

Jednou z možností je zafixovat výnos portfolia na jisté hodnotě ρ a snažit se minimalizovat riziko napříč portfoliem, abychom dosáhli požadovaného výnosu. Pokud to učiníme pro různé hodnoty ρ dostaneme různá portfolia, která určují vhodný kompromis mezi výnosem a rizikem. Požadavek na výnos portfolia na hodnotě ρ můžeme vyjádřit jako

$$\mathrm{avg}(r) = \frac{1}{T}\,e^\top (V\,x) = \mu^\top x = \rho,$$

kde $\mu := \frac{1}{T}\,R^\top e$ je n-rozměrný vektor průměrných výnosů aktiv. Toto je jednoduchá lineární rovnice pro x. Pokud je splněna, pak druhá mocnina rizika je

$$\mathrm{std}(r)^2 = \frac{1}{T}\,\big\|\, r - \mathrm{avg}(r)\,\big\|^2 = \frac{1}{T}\,\big\|\, r - \rho\, e\,\big\|^2.$$

Proto minimalizace rizika[28] je vlastně úloha nejmenších čtverců s omezeními[29]

$$\left.\begin{array}{c} \big\|\, V\,x - \rho\, e\,\big\|^2 \to \min \\ e^\top x = 1 \quad \& \quad \mu^\top x = \rho. \end{array}\right\} \tag{4.1.2}$$

[24] Angl. *mean return*.

[25] Angl. *risk*.

[26] Angl. *anualized return/risk*.

[27] Vilfredo Federico Damaso Pareto (15. července 1848 – 19. srpna 1923) byl italský civilní inženýr, sociolog, ekonom, politolog a filosof. Výrazně ovlivnil ekonomii např. studiem rozdělení příjmu a analýzou individuálních výběrů/voleb. Je také „zodpovědný“ za rozšíření pojmu elita v sociální analýze. Působil na univerzitě v Lausanne a byl průkopníkem ekonometrie (vědecká disciplína nacházející se na pomezí matematiky, statistiky a ekonomie). Jeho jméno je úzce spojeno s *Paretovým principem* 80:20, který byl založen na jeho pozorování, že 80 % půdy v Itálii patří zhruba 20 % populace.

[28] Nebo spíše jeho druhé mocniny a bez ohledu na počet období, což ale na hledané x nemá žádný vliv.

[29] Řešení lze vyjádřit v podobě

$$\begin{pmatrix} x \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2\,V^\top V & e & \mu \\ e^\top & 0 & 0 \\ \mu^\top & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2\rho\,\mu\,T \\ 1 \\ \rho \end{pmatrix},$$

kde y_1 a y_2 jsou Lagrangeovy multiplikátory odpovídající dané dvojici rovnostních omezení.

Tento model optimalizace portfolia má však poměrně zásadní koncepční nedostatek: abychom dokázali určit optimální rozdělení napříč uvažovanými obdobími potřebujeme znát odpovídající výnosy aktiv. To je ale poněkud pošetilé, protože pokud známe všechny budoucí výnosy, měli bychom být schopni dosáhnout jakkoli vysokého celkového výnosu portfolia vhodnou volbou nákupu aktiv, které budou zdražovat (kladné x_i pro aktivum s kladným výnosem), a krátkého prodeje aktiv, které budou zlevňovat (záporné x_i pro aktivum se záporným výnosem). Celá výzva tedy tkví právě v otázce, jak investovat, když budoucí výnosy *neznáme*.

Předpokládejme, že se právě nacházíme v období $t = T$, takže již známe jednotlivé prvky matice výnosů V. Sestavení portfolia získané řešením úlohy (4.1.2) na základě období $t = 1, \ldots, T$ stále může být užitečné, pokud přijmeme následující (silný) předpoklad:

Budoucí výnosy aktiv jsou podobné minulým výnosům.

Jinými slovy, pokud jsou výnosy pro období $T+1, T+2, \ldots$ v podstatě podobné těm minulým, pak rozdělení portfolia x nalezené řešením (4.1.2) se může ukázat býti rozumnou volbou pro následující období. Nicméně měli bychom mít na paměti, že kdykoli investujeme, nemusí být tento předpoklad pravdivý[30]. Také nemusí být na škodu otestovat si nalezené portfolio na datech z minulosti. Např. známe-li výnosy za celý rok, pak můžeme využít informace pro období leden–září pro sestavení úlohy (4.1.2) a nalezené řešení x otestovat na datech říjen–prosinec. Tomu se v teorii portfolia občas říká *zpětné testování*[31].

[30] Zdá se, že např. trh s kryptoměnami to ukazuje celkem jasně.

[31] Angl. *back testing*.

Kvadratické programování s celočíselnými proměnnými. Mezi další možné aplikace úloh kvadratického programování patří kvadratický problém s více batohy, ve kterém opět rozhodujeme o rozdělení n-tice věcí do několika batohů[32]. Tentokrát však kromě užitku u_i z každé jedné věci vložené do některého z batohů budeme mít i užitek u_{ij} z každého páru[33], přičemž není nutné a priori vyžadovat nezápornost těchto užitků. To nás přivádí k úloze

$$\sum_{i=1}^{n}\sum_{j=1}^{k} u_i\, x_{ij} + \sum_{i=1}^{n-1}\sum_{\ell=i+1}^{n}\sum_{j=1}^{k} u_{ij}\, x_{i\ell}\, x_{j\ell} \to \max$$

za podmínek nepřekročení kapacity K_j každého z batohů, tj.

$$\sum_{i=1}^{n} v_i\, x_{ij} \le K_j, \quad j \in \{1, \ldots, k\},$$

kde $x_{ij} \in \{0,1\}$ je binární proměnná vyjadřující přidělení i-té věci do j-tého batohu a v_i vyjadřuje váhu i-té věci. Pochopitelně není možné tutéž věc do více batohů, takže je nutné přidat ještě požadavek

$$\sum_{j=1}^{k} x_{ij} \le 1.$$

Metody řešení této úlohy byly otestovány v [53] pro rozdělení 80 pasažérů ke 14 stolům v jídelně na palubě lodi Royal Caribbean's Majesty of the Seas.

[32] Doposud jsme uvažovali tento problém pouze s jedním batohem. Přidání dalších batohů charakter úlohy nikterak nezmění. Dojde však ke zvetšení počtu proměnných, neboť tentokrát musíme nejenom určit věci, které dáme do batohů, ale i rozhodnout o příslušném batohu. Máme-li n věcí a k batohů, budeme proto mít $x \times k$ binárních proměnných.

[33] Jaksi automaticky předpokládáme, že $u_{ij} = u_{ji}$ a že tento užitek není nikterak ovlivněn volbou batohů, do kterých jsou tyto věci vloženy.

Další možná aplikace byla ukázána v [423] na problému lokalizace pomocných stanic a stanovení jejich dávkovacích plánů pro udržení dezinfekčního prostředku v distribučních systémech pitné vody. V něm jde o nalezení vhodného kompromisu mezi počtem pomocných stanic a mírou zlepšení kvality pitné vody v systému.

Na závěr ještě dodejme, že kvadratické programování lze využít např. i v zemědělství, veřejné správě, obnově digitálního obrazu, designu oken atd., viz [235, 243, 376].

4.2 Když stačí tužka a pravítko…

Stejně jako úlohy lineárního programování i úlohy (QP) lze v případě $n = 2$ (a s nemalým úsilím i pro $n = 3$) řešit pomocí náčrtku přípustné množiny a vrstevnic účelové funkce. Zatímco zakreslení přípustné množiny je stejné jako v případě úloh lineárního programování, s vrstevnicemi už to je trochu složitější. Jak mohou vrstevnice vypadat? Tentokrát lineární část funkce

f určuje pouze jejich posunutí v $\mathbb{R}^2$ a hlavní roli hraje kvadratická část, takže to bude některá ze známých kuželoseček s kanonickými rovnicemi

- elipsy: $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$ $\quad a, b > 0$
- hyperboly: $\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1$ $\quad a, b > 0$
- paraboly: $x_1^2 + 2\,p\,x_2 = 0$ $\quad p \neq 0$
- imaginární rovnoběžky a bod $[0,0]$: $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 0$ $\quad a, b > 0$
- dvojice různoběžek: $\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 0$ $\quad a, b > 0$
- dvojice rovnoběžek: $x_1^2 - \alpha^2 = 0$ $\quad \alpha > 0$
- dvojnásobná přímka: $x_1^2 = 0$

My hledáme vrstevnice na různých úrovních, tj. $f(x) = \alpha$ pro $\alpha \in \mathbb{R}$, což znamená, že v závislosti na hodnosti matice C může nastat pouze některá z následujících situací:

(i) pro některé hodnoty $\alpha \in \mathbb{R}$ to jsou elipsy, které se „zhroutí“ na bod, a pak už jen prázdné množiny; toto je možné tehdy a jen tehdy, když[34] $C > 0$;

(ii) pro všechna $\alpha \in \mathbb{R}$ to jsou paraboly;

(iii) pro některá $\alpha \in \mathbb{R}$ to jsou dvojice rovnoběžek, které se „zhroutí“ na dvojnásobnou přímku, a pak už jen prázdné množiny.

Vzhledem k požadavku $C \geq 0$ není situace

(iv) pro některé hodnoty $\alpha \in \mathbb{R}$ se jedná o hyperboly, které se „zhroutí“ na bod, a pak už jen prázdné množiny;

možná, neboť pro hyperboly by matice C musela být indefinitní. To ale neznamená, že nemůžeme uvažovat i takovou úlohu kvadratického programování, kde C je indefinitní nebo negativně (semi-)definitní. „Jenom“ funkce f nebude konvexní, takže nebudeme mít už vůbec zaručenu existenci řešení a nebude možné použít metody, které si za chvíli vybudujeme, viz také Podkapitolu 4.7.

[34] K témuž by došlo i v případě $C < 0$, což ale v naší úloze (QP) nedovolujeme.

Pokud je v účelové funkci smíšený člen $x_1\,x_2$ a nejedná se o úplný čtverec, je analýza konkrétního tvaru vrstevnic výrazně komplikovanější než v případě úloh lineárního programování. Máme k tomu víceméně 3 možnosti:

(a) diagonalizujeme kvadratickou část účelové funkce a dopočteme vše potřebné;

(b) jelikož předchozí postup obvykle bývá početně velmi náročný, je mnohem jednodušší pracovat s celou účelovou funkcí jako kvadratickou funkcí v proměnných $(x_1, x_2, 1)^\top$, což nás přivede k matici

$$Q := \begin{pmatrix} \widetilde{Q} & \frac{1}{2}d \\ \frac{1}{2}d^\top & \text{konst.} - \alpha \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)},$$

kde $\widetilde{Q} := \frac{1}{2}C$ je skutečnou maticí kvadratické formy a „konst.“ značí případný absolutní člen ve funkci f. Pak platí následující klasifikace

- $\operatorname{rank}(Q) = 3$ & $\operatorname{rank}(\widetilde{Q}) = 2$: elipsy (hyperboly) s kanonickou rovnicí

$$\lambda_1\,y_1^2 + \lambda_2\,y_2^2 + \frac{\det Q}{\det \widetilde{Q}} = 0;$$

- $\operatorname{rank}(Q) = 3$ & $\operatorname{rank}(\widetilde{Q}) = 1$: paraboly s kanonickou rovnicí

$$\lambda_1 y_1^2 \pm 2\sqrt{-\frac{\det Q}{\lambda_1}}\, y_2 = 0;$$

- $\operatorname{rank}(Q) = 2$ & $\operatorname{rank}(\widetilde{Q}) = 2$: různoběžky s kanonickou rovnicí

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 0;$$

- $\operatorname{rank}(Q) = 2$ & $\operatorname{rank}(\widetilde{Q}) = 1$: rovnoběžky s kanonickou rovnicí

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 0;$$

- $\operatorname{rank}(Q) = 1$ & $\operatorname{rank}(\widetilde{Q}) = 1$: dvojnásobná přímka s kanonickou rovnicí

$$\lambda_1 y_1^2 = 0.$$

(c) úprava kvadratické části tak, abychom odstranili smíšený člen (vhodnou transformací do nových souřadnic, které nemusí být ortogonální – obvykle totiž nepotřebujeme znát zcela přesně tvar vrstevnic v původních souřadnicích).

Navíc zde musíme zdůraznit, že na rozdíl od úloh lineárního programování při řešení úloh kvadratického programování již neplatí to, že hledaný extrém se realizuje v nějakém vrcholu množiny X či na nějaké jeho hraně. Toto si ilustrujeme v několika následujících jednoduchých příkladech.

Příklad 4.2.1

Uvažme

$$x_1^2 + x_2^2 - 4x_1 - 4x_2 + 1 \to \min$$

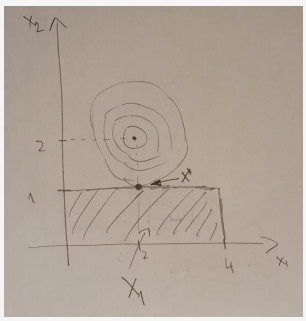
na množinách

$$X_1 := \{x \in \mathbb{R}^2 \mid x_1 \in [0,4] \;\&\; x_2 \in [0,1]\},$$
$$X_2 := \{x \in \mathbb{R}^2 \mid x_1 \in [0,1] \;\&\; x_2 \in [0,1]\},$$
$$X_3 := \{x \in \mathbb{R}^2 \mid x_1 \in [0,4] \;\&\; x_2 \in [0,4]\}.$$

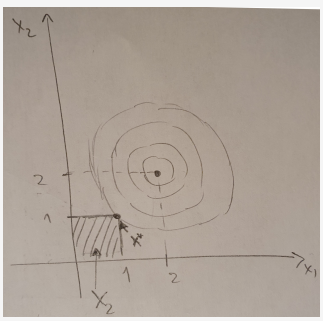
Vrstevnice účelové funkce jsou určeny rovností

$$x_1^2 + x_2^2 - 4x_1 - 4x_2 + 1 = \alpha \quad \text{neboli} \quad (x_1-2)^2 + (x_2-2)^2 = \alpha + 7,$$

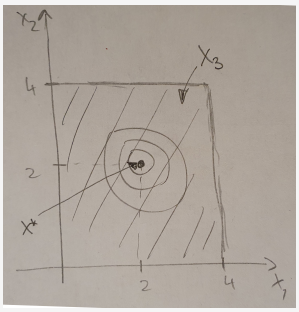
což pro $\alpha < -7$ jsou prázdné množiny, pro $\alpha = -7$ to je bod $[2,2]$ a pro $\alpha > -7$ se jedná o kružnice se středem v bodě $[2,2]$ a poloměrem $\sqrt{\alpha+7}$, které se pro $\alpha \to \infty$ rozpínají. Na množině X_1 je řešením bod $x_1^* = [2,1]$ s hodnotou $f_1^* = -6$, na množině X_2 to je bod $x_2^* = [1,1]$ s hodnotou $f_1^* = -5$ a konečně na množině X_3 to je bod $x_3^* = [2,2]$ s hodnotou $f_1^* = -7$, viz Obrázky 4.7.a–4.7.c. V prvním případě jde o bod ležící na hraně množiny X_1 (ovšem už to není celá hrana), ve druhém případě to je vrchol množiny X_2 a ve třetím případě jde o vnitřní bod přípustné množiny X_3.



Obrázek 4.7.a: Vrstevnice a množina X_1.

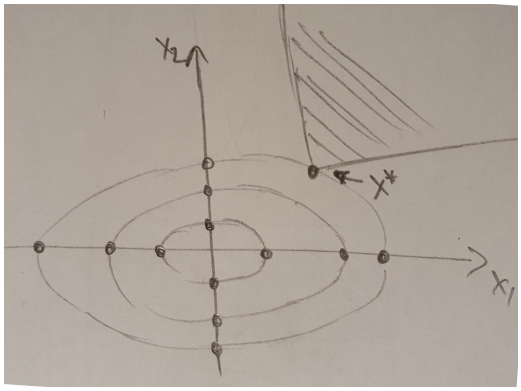


Obrázek 4.7.b: Vrstevnice a množina X_2.



Obrázek 4.7.c: Vrstevnice a množina X_3.

Už jsme se zmínili, že vrstevnice mohou být i elipsy (což vlastně zahrnuje i kružnice jako velmi speciální případ). Jednou z jejich velmi důležitých charakteristik jsou poloosy, které mimo jiné určují vrcholy elipsy (průnik elipsy a poloos). Nicméně je asi celkem zřejmé, že neexistuje žádná souvislost mezi těmito vrcholy a řešením příslušné úlohy (QP), viz Obrázek 4.8.



Obrázek 4.8: Úloha (QP) a vrcholy elipsy.

Ve všech předchozích případech bylo řešení vždy jediné. Mohlo by jich ale být i nekonečně mnoho? Jsou-li vrstevnice elipsy (kružnice), tak to možné není. V takovém případě je totiž matice C pozitivně definitní, a tudíž je účelová funkce f *ostře konvexní*, což implikuje, že existuje nejvýše jediné globální minimum na množině X, viz Větu 6.2.23(ii). Ovšem v případě singulární matice $C \geq 0$ už řešení *může* být nekonečně mnoho, jak ukazuje následující příklad[35].

Příklad 4.2.2

Uvažme

$$x_1^2 + 2\,x_1\,x_2 + x_2^2 \to \min$$

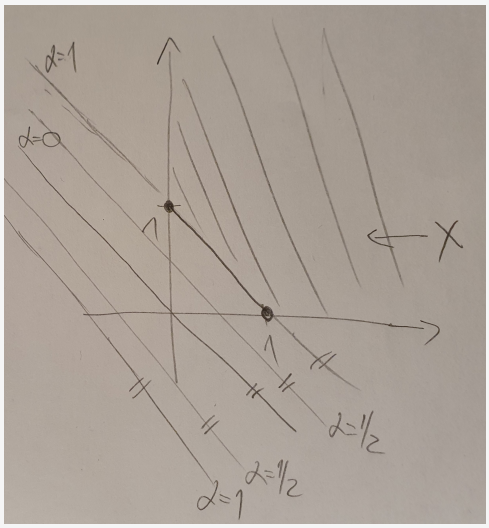
na množině

$$X := \{x \in \mathbb{R}^2 \mid x_1 + x_2 \geq 1 \;\&\; x_1 \geq 0 \;\&\; x_2 \geq 0\}.$$

Účelová funkce odpovídá matici $C = \left(\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right)$ splňující $C \geq 0$ a $\det C = 0$. Vrstevnice jsou dány rovnicí

$$x_1^2 + 2\,x_1\,x_2 + x_2^2 = \alpha \quad \text{neboli} \quad (x_1 + x_2)^2 = \alpha,$$

takže pro $\alpha < 0$ to jsou prázdné množiny, pro $\alpha = 0$ máme přímku $x_2 = -x_1$ a pro $\alpha > 0$ máme dvojici rovnoběžných přímek $x_2 = \pm\sqrt{\alpha} - x_1$. Řešením pak je celá úsečka spojující body $[1,0]$ a $[0,1]$, viz Obrázek 4.9.



Obrázek 4.9: Úloha (QP) a $C \geq 0$ s nekonečně mnoha řešeními.

[35] Obecně platí, že úloha (QP) může mít více řešení pouze v případě, že soustava

$$Cz = 0 \quad \& \quad d^\top z = 0 \qquad (\triangle)$$

má netriviální řešení. Toto je důsledek následujícího tvrzení: Je-li x^* řešením úlohy (QP), pak $\widehat{x} \in X$ je také řešením (QP) právě tehdy, když $z := \widehat{x} - x^*$ splňuje (△).

Důkaz. Je-li $\widehat{x} = x^*$ je tvrzení triviální. Jinak to plyne z Poznámky 140 na straně 496, která se v případě úlohy (QP) redukuje na to, že $\widehat{x} \in X$ je řešením (QP) právě tehdy, když

$$0 = z^\top(C\widehat{x} + d). \qquad (\square)$$

Pokud platí (△), pak je jistě splněna i rovnost (□), a tedy $\widehat{x}$ je řešením (QP). Naopak, pokud $\widehat{x}$ je řešením (QP), pak platí (□) a navíc záměnou x^* a $\widehat{x}$ máme také

$$0 = z^\top(Cx^* + d).$$

Odečtením obou těchto rovností získáme

$$z^\top C z = 0,$$

což díky podmínce $C \geq 0$ je možné pouze v případě $Cz = 0$. Odtud zpětným dosazením do (□) dostaneme, že také nutně

$$d^\top z = 0,$$

a tudíž z je řešením soustavy (△).

Ve zbývajících třech příkladech se podíváme na složitější situace, ve kterých již kvadratická část bude obsahovat i smíšený člen a zároveň se nebude jednat o úplný čtverec.

Příklad 4.2.3

Grafické řešení – řešený příklad #1 (elipsy):
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/qp1

Příklad 4.2.4

Grafické řešení – řešený příklad #2 (paraboly):
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/qp2

Příklad 4.2.5

Grafické řešení – řešený příklad #3 (přímky):
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/qp3

4.3 Nepříliš zevrubný teoretický vhled

V této podkapitole si ukážeme některé základní teoretické výsledky pro úlohy kvadratického programování. Toto je jedno z velmi mála míst dosavadního výkladu, kde už by se mohla hodit znalost některých pojmů z obecné teorie matematického programování prezentované až v Kapitole 8. Nicméně pro tuto chvíli postačí s jistou mírou představivosti alespoň to, co jsme si připomněli ve Větě 1.2.16. Ačkoli tam žádné nerovnosti ani podmínky nezápornosti nebyly.

Napíšeme-li vektory $a_1,\dots,a_m$ z (QP) jako řádky matice A a $b := (b_1,\dots,b_m)^\top$, pak lze jednotlivá omezení získat pomocí součinu Ax a vektoru b, takže k této úloze můžeme přiřadit (*regulární*) *Lagrangeovu funkci*

$$L(x,y) := \tfrac{1}{2}\langle Cx,x\rangle + \langle d,x\rangle + \langle y,Ax-b\rangle = \tfrac{1}{2}\langle Cx,x\rangle + \langle A^\top y+d,x\rangle - \langle y,b\rangle,$$

kde $y \in Q := \{y=(y_1,\dots,y_m)^\top \in \mathbb{R}^m \mid y_1 \ge 0 \;\&\; \cdots \;\&\; y_k \ge 0\}$ jsou *Lagrangeovy multiplikátory*. Funkce $L(x,y)$ je jistě diferencovatelná a není těžké se přesvědčit o tom, že

$$\operatorname{grad}_x L(x,y) = Cx + A^\top y + d.$$

Proto díky požadavku $C \ge 0$ vyplývá z Karushových–Kuhnových–Tuckerových podmínek v diferenciálním tvaru následující tvrzení, viz Větu 8.2.15.

Věta 4.3.1

Bod $x^* \in X$ je řešením úlohy (QP) právě tehdy, když existuje $y^* \in Q$ takové, že platí

$$\langle Cx^* + A^\top y^* + d, x - x^*\rangle \ge 0 \quad \text{pro všechna } x \in P, \tag{4.3.1}$$

$$\langle y, Ax^* - b\rangle = 0. \tag{4.3.2}$$

Zejména je-li $P = \mathbb{R}^n$, podmínka (4.3.1) se redukuje na $Cx^* + A^\top y^* + d = 0$.

Analogicky postupu známému z lineárního programování lze libovolnou úlohu (QP) převést do libovolného z dříve zmíněných speciálních tvarů, přičemž opět platí, že situace $x \in \mathbb{R}^n$ a $Ax = b$ je pro nás nezajímavá. Ovšem tentokrát se tímto přechodem může změnit ta nejdůležitější charakteristika úlohy – definitnost matice kvadratické formy v účelové funkci. Budeme-li se držet popisu z Příkladu 2.2.2, pak ke změně definitnosti matice kvadratické formy dojde např. při přechodu od úlohy ve standardním tvaru

$$\tfrac{1}{2}\langle Cx,x\rangle + \langle d,x\rangle \to \min, \quad Ax \le b, \quad x \ge 0$$

k úloze v kanonickém tvaru

$$\tfrac{1}{2}\langle \widetilde{C}x, x\rangle + \langle \widetilde{d}, x\rangle \to \min, \quad \widetilde{A}\widetilde{x} = \widetilde{b}, \quad \widetilde{x} \geq 0,$$

v níž bereme

$$\widetilde{x} := \begin{pmatrix} x \\ u \end{pmatrix} \in \mathbb{R}^{n+m}, \quad \widetilde{C} := \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}, \quad \widetilde{d} := \begin{pmatrix} d \\ 0 \end{pmatrix} \in \mathbb{R}^{n+m}$$
$$\widetilde{A} := \begin{pmatrix} A & I \end{pmatrix} \in \mathbb{R}^{m\times(n+m)} \quad \text{a} \quad \widetilde{b} := b.$$

Pak bez ohledu na definitnost či semidefinitnost matice C, bude matice $\widetilde{C}$ vždy jen semidefinitní. Podobně to dopadne při přechodu od základního tvaru ke standardnímu tvaru, kde získáme úlohu se semidefinitní maticí kvadratické formy

$$\widetilde{C} := \begin{pmatrix} C & -C \\ -C & C \end{pmatrix} \in \mathbb{R}^{2n\times 2n},$$

nebo od základního tvaru ke kanonickému tvaru se semidefinitní maticí kvadratické formy

$$\widetilde{C} := \begin{pmatrix} C & -C & 0 \\ -C & C & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(2n+m)\times(2n+m)}.$$

A jak v těchto speciálních případech vypadají podmínky (4.3.1)–(4.3.2)?

(i) Ve standardním tvaru volíme $P = \{x \in \mathbb{R}^n \mid x \geq 0\}$ a $X = \{x \in P \mid Ax \leq b\}$, takže podmínka (4.3.1) je splněna právě tehdy, když[36]

$$Cx^* + A^\top y^* + d \geq 0 \tag{4.3.3}$$

a zároveň[37]

$$\langle Cx^* + A^\top y^* + d, x^*\rangle = 0. \tag{4.3.4}$$

Podmínka (4.3.2) zůstává beze změny. Proto řešení úlohy (QP) ve standardním tvaru je ekvivalentní s nalezením řešení soustavy

$$\left.\begin{aligned} Cx + A^\top y + d &= u \\ Ax + v &= b \\ \langle x, u\rangle + \langle y, v\rangle &= 0 \\ x \geq 0 \quad \& \quad u \geq 0 \quad \& \quad y \geq 0 \quad \& \quad v &\geq 0. \end{aligned}\right\} \tag{4.3.5}$$

Vskutku, první rovnost je pouze důsledkem podmínky (4.3.3) a druhá rovnost je analogickým přepisem požadavku $Ax \leq b$. Odtud také plyne nezápornost všech čtyř vektorů x, y, u, v. Zároveň v takovém případě je podmínka (4.3.4) ekvivalentní s rovností $\langle x, u\rangle = 0$ a podmínka (4.3.2) s rovností $\langle y, v\rangle = 0$. Jelikož však zde vystupují pouze vektory s nezápornými složkami, budou oba skalární součiny nulové právě tehdy, když i jejich součet bude nulový, což je třetí uvedená podmínka.

(ii) Podobným způsobem zjistíme, že v případě úlohy (QP) v kanonickém tvaru s $P = \{x \in \mathbb{R}^n \mid x \geq 0\}$ a $X = \{x \in P \mid Ax = b\}$ je podmínka (4.3.1) opět splněna právě tehdy, když platí (4.3.3)–(4.3.4), zatímco podmínka (4.3.2) je splněna vždy. Proto v takovém případě je řešení úlohy (QP) ekvivalentní s nalezením řešení soustavy

$$\left.\begin{aligned} Cx + A^\top y + d &= u \\ Ax &= b \\ \langle x, u\rangle &= 0 \\ x \geq 0 \quad \& \quad u &\geq 0. \end{aligned}\right\} \tag{4.3.6}$$

[36] Kdyby i-tá složka vektoru $Cx^* + A^\top y^* + d$ byla záporná, pak bychom při skalárním součinu tohoto vektoru a vektoru $x \in P$ se složkami $x_j = x_j^* \geq 0$ pro $j \neq i$ a současně $x_i > 2x_i^* \geq 0$ dostali spor s (4.3.1).

[37] Vzhledem k podmínkám $Cx^* + A^\top y^* + d \geq 0$ a $x^* \geq 0$, dostaneme volbou $x = 0$ v (4.3.1) skalární součin $\langle Cx^* + A^\top y^* + d, -x^*\rangle$, což je součet nekladných čísel. Zároveň však tento součet má být nezáporný, takže nutně musí být nulový. Dokonce odtud vidíme, že každý sčítanec v tomto skalárním součinu musí být nulový, tj. buď i-tá složka vektoru $Cx^* + A^\top y^* + d$ je nulová nebo $x_i^* = 0$.

(iii) Konečně v případě úlohy (QP) v základním tvaru je $P = \mathbb{R}^n$ a $X = \{x \in P \mid Ax \le b\}$, takže podmínka (4.3.1) se redukuje pouze na (4.3.3), zatímco podmínka (4.3.2) zůstává beze změny. Proto v takovém případě je řešení úlohy (QP) ekvivalentní s nalezením řešení soustavy

$$\begin{aligned} Cx + A^\top y + d &= 0 \\ Ax + v &= b \\ \langle y, v\rangle &= 0 \\ y \ge 0 \quad \& \quad v &\ge 0. \end{aligned}$$

Tyto ekvivalence ukazují, že místo řešení optimalizační úlohy je možné se omezit na hledání řešení soustavy lineárních rovnic doplněné o podmínku ortogonality (nulovost skalárních součinů) a nezápornost jednotlivých složek. Tento alternativní pohled využijeme při budování druhé z metod pro řešení úloh (QP). Naopak první metoda bude založena na velmi výhodném tvaru jedné z duálních úloh, která má v obecném případě podobu

$$\varphi(y) \to \max \quad \& \quad y \in Q \quad \& \quad \varphi(y) := \inf_{x \in P} L(x, y).$$

My se konkrétně zaměříme na duální úlohu k úloze (QP) v základním tvaru, pro kterou je $Q = \mathbb{R}^m_+$ a $P = \mathbb{R}^n$. V takovém případě se

$$\inf_{x \in \mathbb{R}^n} \left(\tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle + \langle y, Ax - b\rangle \right)$$

díky konvexnosti celé funkce realizuje přímo ve stacionárním bodě vzhledem k x, tj. v bodě splňujícím

$$Cx + A^\top y + d = 0.$$

Proto v tomto případě bude mít duální úloha tvar

$$\tfrac{1}{2}\langle Cx, x\rangle + \langle A^\top y + d, x\rangle - \langle y, b\rangle \to \max \quad \& \quad Cx + A^\top y + d = 0 \quad \& \quad y \ge 0$$

což díky druhé podmínce můžeme zapsat jako[38]

$$\varphi(y) = -\tfrac{1}{2}\langle Cx, x\rangle - \langle y, b\rangle \to \max \quad \& \quad Cx + A^\top y + d = 0 \quad \& \quad y \ge 0. \tag{4.3.7}$$

Kdyby dokonce bylo $C > 0$, lze z druhé podmínky přímo vyjádřit $x =$ a dosadit do účelové funkce, čímž bychom jej odtud vyeliminovali a měli bychom duální úlohu pouze v proměnné y. Toto ale v případě $\det C = 0$ nemusí být možné, takže duální úloha má spíše výše uvedený tvar s proměnnými x a y.

Je také zřejmé, že

(i) duální úloha k úloze (QP) bude vždy opět úloha kvadratického programování;

(ii) volbou $C = 0$ dostaneme dvojici primární a duální úlohy známou z Kapitoly 2.

Vraťme se nyní zpět k úloze (4.3.7) a připusťme, že matice C je regulární. Pak z druhé podmínky dostaneme

$$x = -C^{-1}(A^\top y + d),$$

což po dosazení do účelové funkce v (4.3.7) dává

$$\begin{aligned} \varphi(y) &= -\tfrac{1}{2}\langle A^\top y + d, C^{-1}(A^\top y + d)\rangle - \langle y, b\rangle = \\ &= -\tfrac{1}{2}\langle A^\top y, C^{-1}A^\top y\rangle - \tfrac{1}{2}\langle A^\top y, C^{-1}d\rangle - \tfrac{1}{2}\langle d, C^{-1}(A^\top y + d)\rangle - \langle y, b\rangle = \\ &= -\tfrac{1}{2}\langle AC^{-1}A^\top y, y\rangle - \langle AC^{-1}d + b, y\rangle - \tfrac{1}{2}\langle C^{-1}d, d\rangle = \\ &= -\tfrac{1}{2}\langle Gy, y\rangle - \langle h, y\rangle + \gamma, \end{aligned}$$

kde

$$G := AC^{-1}A^\top, \quad h := AC^{-1}d + b, \quad \gamma := -\tfrac{1}{2}\langle C^{-1}d, d\rangle.$$

[38] Čtenáři tentokrát ponecháme k jeho vlastnímu rozmyšlení, že podobným způsobem bychom mohli dospět k tomu, že v případě úlohy v kanonickém tvaru má duální úloha podobu

$$\begin{aligned} -\tfrac{1}{2}\langle Cx, x\rangle - \langle y, b\rangle &\to \max \\ Cx + A^\top y + d &\ge 0 \end{aligned}$$

a pro úlohu ve standardním tvaru obdržíme

$$\begin{aligned} -\tfrac{1}{2}\langle Cx, x\rangle - \langle y, b\rangle &\to \max \\ Cx + A^\top y + d \ge 0 \quad \& \quad y &\ge 0. \end{aligned}$$

Tentokrát však ani jedna z uvedených dvojic primárních a duálních úloh nevykazuje žádnou symetrii. Té lze docílit pomocí úlohy kvadratického programování ve tvaru

$$\begin{aligned} \tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle + \tfrac{1}{2}\langle Dz, z\rangle &\to \min \\ Ax + Dz \ge -b \quad \& \quad x &\ge 0, \end{aligned}$$

kde C a D jsou symetrické a pozitivně semidefinitní matice vhodných rozměrů, viz [103]. V takovém případě je totiž duální úloha velmi podobného typu, konkrétně

$$\begin{aligned} -\tfrac{1}{2}\langle Cx, x\rangle - \langle y, b\rangle - \tfrac{1}{2}\langle Dy, y\rangle &\to \max \\ Cx - A^\top y \ge -d \quad \& \quad y &\ge 0. \end{aligned}$$

Kdybychom „sečetli“ obě úlohy, dostali bychom samoduální úlohu kvadratického programování

$$\begin{aligned} \langle Cx, x\rangle + \langle d, x\rangle + \\ + \langle Dy, y\rangle + \langle y, b\rangle &\to \min \\ Dy + Ax &\ge -b \\ -A^\top y + Cx &\ge -d \\ x \ge 0 \quad \& \quad y &\ge 0. \end{aligned}$$

Zjevně speciální volbou $C = 0 = D$ dostaneme dvojici symetrických úloh lineárního programování.

Jelikož konstanta v účelové funkci nemá vliv na řešení úlohy, můžeme duální úlohu v tomto případě zapsat jako

$$-\tfrac{1}{2}\langle Gy, y\rangle - \langle h, y\rangle \to \max, \quad y \geq 0$$

neboli

$$\tfrac{1}{2}\langle Gy, y\rangle + \langle h, y\rangle \to \min, \quad y \geq 0. \tag{4.3.8}$$

Jelikož $C > 0$, pak také $C^{-1} > 0$, což implikuje $G \geq 0$ a navíc diagonální prvky matice G splňují $g_{ii} = \langle C^{-1}a_i, a_i\rangle > 0$ pro $i = 1, \ldots, m$.

Ještě než z těchto úvah odvodíme první metodu pro řešení úloh kvadratického programování, potřebujeme znát vztah mezi primární a duální úlohou. Tentokrát si jej dokazovat nebudeme, neboť vyplývá z obecné teorie duality v matematickém programování a platí pro libovolnou dvojici primární a duální úlohy kvadratického programování, viz Věty 8.3.12 a 8.3.16.

VĚTA 4.3.2 SLABÁ VĚTA O DUALITĚ

Uvažme (libovolnou) dvojici primární a duální úlohy kvadratického programování. Je-li $x \in X$ a (z, y) je přípustný bod duální úlohy, pak vždy platí

$$\tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \geq -\tfrac{1}{2}\langle Cz, z\rangle - \langle y, b\rangle.$$

Odtud (opět) vyplývají následující skutečnosti.

(i) Jsou-li obě úlohy přípustné, pak obě mají také řešení.

(ii) Je-li v některé z těchto úloh účelová funkce neomezená (zdola/shora), pak druhá úloha je nepřípustná.

VĚTA 4.3.3 SILNÁ VĚTA O DUALITĚ

Pro (libovolnou) dvojici primární a duální úlohy kvadratického programování platí:

(i) Je-li x^* řešením primární úlohy, pak existuje $y^* \in Q$ takové, že $[x^*, y^*]$ je řešením duální úlohy.

(ii) Je-li $[x^*, y^*]$ řešením duální úlohy a zároveň $C > 0$, pak x^* je řešením primární úlohy.

V obou případech přitom platí rovnost

$$\min\left(\tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle\right) = \max\left(-\tfrac{1}{2}\langle Cx, x\rangle - \langle y, b\rangle\right).$$

Tvrzení Věty 4.3.3(ii) lze ještě více zesílit tak, že vlastně obdržíme podobné tvrzení jako v úloze lineárního programování.

VĚTA 4.3.4

Pro (libovolnou) dvojici primární a duální úlohy kvadratického programování platí:

(i) Má-li jedna z úloh řešení, pak má řešení i druhá úloha.

(ii) Je-li jedna z úloh nepřípustná, pak druhá úloha je buď také nepřípustná nebo účelová funkce je na přípustné množině neohraničená (zdola/shora).

Pro dvojice primární a duální úlohy kvadratického programování tedy platí stejná tvrzení jako v případě lineárního programování. Vlastně lze říci, že teorie duality lineárního programování je speciálním případem teorie duality kvadratického programování.

4.4 Hildrethova[39]–d'Esopova[40] metoda

V 50. letech 20. století představil Hildreth iterativní metodu pro řešení úloh kvadratického programování, která se aplikuje na duální úlohu[41] a v tehdejší době byla výhodná zejména pro využití na programovatelných kalkulačkách, viz [265, 266]. Metoda sice obecně není konečně kroková a ani konvergence není nějak zvláště rychlá, ale tyto nedostatky jsou alespoň částečně kompenzovány výpočetní jednoduchostí jednotlivých iterací a jistou „autokorekční" vlastností. Tuto metodu téměř vzápětí později rozšířil d'Esopo na obecné úlohy konvexního programování, viz [132].

[39] Clifford George Hildreth (8. prosince 1917 – 15. srpna 1995) byl americký ekonometr. Byl také vedoucím Ústavu ekonomie na Michiganské státní univerzitě. V roce 1973 byl prezidentem Americké statistické asociace (American Statistical Association).

[40] Žádné podrobnější informace o něm nemáme. Jeho profesní dráha nejspíše šla jiným než matematickým směrem, neboť se nám pod jeho jménem podařilo dohledat pouze 3 odborné články.

[41] Pro kterou připomíná Gaussovu–Seidelovu metodu určenou pro řešení soustav lineárních rovnic.

Metoda je určena pro úlohu kvadratického programování v základním tvaru

$$\tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \to \min, \quad Ax \le b \tag{4.4.1}$$

s maticí $C > 0$. V předchozí části jsme ukázali, že příslušná duální úloha je ekvivalentní s úlohou (4.3.8), která je vždy přípustná. Potom z části (i) komentáře za Větou 4.3.2 plyne, že je-li také úloha (4.4.1) přípustná, pak obě mají řešení. V takovém případě pak díky Větě 4.3.3(ii) víme, že je-li y^* řešením duální úlohy (4.3.8), pak řešení primární úlohy vypočteme jako

$$x^* := -C^{-1}(A^\top y^* + d). \tag{4.4.2}$$

Navíc díky pozitivní definitnosti matice C je toto jediné řešení úlohy (4.4.1), tj. ať už má duální úloha (4.3.8) jedno nebo více řešení, všechny dávají totéž x^* vypočtené podle (4.4.2). Přechod od primární úlohy (4.4.1) k příslušné duální úloze je výhodný zejména ve chvíli, kdy počet proměnných je větší než počet omezení včetně podmínek nezápornosti, což určuje počet proměnných v duální úloze. Na druhou stranu jednou ze slabých stránek celé metody je vztah pro x^*, neboť při praktické realizaci (tj. při počítání jen s jistou mírou přesnosti) vyžaduje dostatečně přesný výpočet matice C^{-1}, což může být v případě hodnot $\det C$ blízkých nule technicky poměrně náročné[42].

[42] Hildrethova–d'Esopova metoda nám sice umožní najít řešení y^* s jistou mírou přesnosti, avšak tato přesnost se nemusí zachovat při určování x^*, na což bude mít nemalý vliv zejména přesnost výpočtu C^{-1}.

Princip algoritmu je velmi jednoduchý. Začneme v nějakém (de facto libovolném) bodě $y^{[0]} \ge 0$ a bod $y^{[1]} = [y_1^{[1]}, \dots, y_m^{[1]}]$ získáme tak, že postupně řešíme m minimalizačních úloh s podmínkou $y_i^{[1]} \ge 0$, přičemž (obvykle) postupujeme od $i = 1$ až k $i = m$ a v každém kroku považujeme ostatní proměnné za konstanty. Podrobněji si algoritmus popíšeme za chvíli, ale před tím se ještě pozastavme nad jedním problémem souvisejícím s tímto iterativním procesem. Jeho pomocí získáme body $y^{[1]}, \dots$, které jsou přípustnými pro duální úlohu (4.3.8) a aproximují řešení y^*. Nabízí se proto otázka, kdy odpovídající body $x^{[1]} := -C^{-1}(A^\top y^{[1]} + d)$ atd. budou přípustnými body úlohy (4.4.1)?

LEMMA 4.4.1

Je-li $y \ge 0$ a $Gy + h \ge 0$, potom

$$x(y) := -C^{-1}(A^\top y + d) \tag{4.4.3}$$

je přípustným bodem úlohy (4.4.1).

Důkaz. Vždyť platí

$$b - Ax = b + AC^{-1}(A^\top y + d) = Gy + h. \qquad \blacksquare$$

Toto dává jednoduché kritérium přípustnosti přibližného řešení x určeného vztahem (4.4.3) bez toho, aniž bychom samotné x vypočítali. Je-li totiž $(Gy + h)_i = -\varepsilon_i < 0$ pro nějaké $i \in \{1, \dots, m\}$, pak platí $\langle a_i, x\rangle = b_i + \varepsilon_i$, tj. hodnota $\varepsilon_i > 0$ určuje „míru porušení" i-tého omezení primární úlohy (4.4.1). Jinými slovy, jsou-li $\varepsilon_1, \dots, \varepsilon_m$ předem dané tolerance porušení jednotlivých omezení[43] úlohy (4.4.1), pak při řešení úlohy (4.3.8) stačí najít $y \ge 0$ takové, že

$$Gy + h + \varepsilon \ge 0$$

pro $\varepsilon := (\varepsilon_1, \dots, \varepsilon_m)^\top$. Pak odpovídající x bude ε-přípustným bodem úlohy (4.4.1). Ještě jinak řečeno, potřebujeme, aby existovaly takové přípustné body úlohy (4.4.1), že

$$\langle a_i, x\rangle \le b_i - \delta, \quad i = 1, \dots, m \tag{4.4.4}$$

[43] To ale v mnoha praktických úlohách nemusí představovat žádné úskalí, neboť A a b mohou být dány jen s určitým stupněm přesnosti (odhad atd.).

pro nějaké $\delta > 0$. To ale není nic jiného než tzv. *Slaterova podmínka*[44], která zaručuje, že úloha (4.4.1) je tzv. *regulární* úlohou konvexního programování. Díky této podmínce je primární úloha jistě přípustná, a tedy i řešitelná, neboť odpovídající duální úloha je přípustná vždy. Proto i duální úloha je řešitelná a zbývá už jen dokázat, že posloupnost generovaná pomocí Hildrethovy–d'Esopovy metody konverguje k takovému řešení. K tomu se vrátíme na konci této podkapitoly.

[44] Viz Věty 8.2.15(ii) a 8.3.4(i).

Snadno si můžeme všimnout, že volné minimum účelové funkce úlohy (4.4.1) se realizuje v bodě

$$\widetilde{x} = -C^{-1}d.$$

Tedy vztah $x = -C^{-1}(A^\top y + d)$ určuje volné minimum $\widetilde{x}$ v případě $\widetilde{y} = (0,\dots,0)^\top$. Navíc platí

$$b - A\widetilde{x} = b + AC^{-1}(A^\top \cdot 0 + d) = AC^{-1}d + b = h,$$

tj. složky vektoru h určují „odchylku" bodu $\widetilde{x}$ od $Ax = b$, přičemž h_i určuje „odchylku" od hranice i-tého omezení $\langle a_i, x\rangle = b_i$. Pokud $h \geq 0$, pak $\widetilde{y} = 0$ je řešením úlohy (4.3.8), neboť $G \geq 0$, takže $\widetilde{x}$ je řešením úlohy (4.4.1). V opačném případě „nastupuje" následující algoritmus.

Krok 0 (inicializace): Zvolíme nějaké $y^{[0]} \geq 0$. Obvykle se volí $y^{[0]} = 0$, jelikož ve výsledku bývá některá z proměnných nulová (i s ohledem na předchozí komentář). V dalších krocích je potřeba vypočítat $y^{[1]}$ atd., tj. určit *minimalizující posloupnost* $\{y^{[k]}\}_{k=0}^{\bullet}$. Výpočet se zastaví ve chvíli, kdy je dosaženo požadované přesnosti mezi dvěma následujícími členy posloupnosti $\{y^{[k]}\}$, neboť platí $\rho(y^{[k]}) \leq \rho(y^{[k-1]})$, tj. posloupnost $\{\rho(y^{[k]})\}_{k=0}^{\bullet}$ je nerostoucí, kde ρ je účelová funkce úlohy (4.3.8), tj.

$$\rho(y) \coloneqq \tfrac{1}{2}\langle Gy, y\rangle + \langle h, y\rangle.$$

Mějme již vypočteny $y^{[0]}, y^{[1]}, \dots, y^{[k-1]}$. Nyní chceme vypočítat $y^{[k]} \coloneqq (y_1^{[k]}, \dots, y_m^{[k]})^\top$. To se provede pomocí následujícího schématu.

Krok 1 (výpočet $y_1^{[k]}$): Řešíme úlohu

$$\rho(t, y_2^{[k-1]}, \dots, y_m^{[k-1]}) \to \min, \quad t \geq 0.$$

Tato funkce ρ je kvadratickou (konvexní) funkcí jedné proměnné t. Proto minimum nastává buď ve stacionárním bodě nebo pro $t = 0$, tj. $y_1^{[k]} = \max\{0, t_1^{[k]}\}$, kde $t_1^{[k]}$ značí stacionární bod, tj. řešení úlohy

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t, y_2^{[k-1]}, \dots, y_m^{[k-1]}) = 0 \quad \text{neboli} \quad g_{11}t + \sum_{j=2}^{m} g_{1j}\, y_j^{[k-1]} + h_1 = 0,$$

což dává

$$t_1^{[k]} = -\frac{1}{g_{11}}\left(\sum_{j=2}^{m} g_{1j}\, y_j^{[k-1]} + h_1\right).$$

Krok i (výpočet $y_i^{[k]}$, $i = 2,\dots,m$): Řešíme úlohu

$$\rho(y_1^{[k]}, \dots, y_{i-1}^{[k]}, t, y_{i+1}^{[k-1]}, \dots, y_m^{[k-1]}) \to \min, \quad t \geq 0.$$

Opět dostáváme $y_i^{[k]} = \max\{0, t_i^{[k]}\}$, kde $t_i^{[k]}$ je stacionární bod, tj.

$$t_i^{[k]} = -\frac{1}{g_{ii}}\left(\sum_{j=1}^{i-1} g_{ij}\, y_j^{[k]} + h_i + \sum_{j=i+1}^{m} g_{ij}\, y_j^{[k-1]}\right).$$

V jednotlivých iteracích je možné zaměnit pořadí výpočtu $1, 2, \dots, m$-té složky $y^{[\bullet]}$. Princip metody se tím nikterak nezmění, ale může to mít vliv na konvergenci metody, která obvykle není příliš rychlá (pochopitelně to velmi závisí na počáteční volbě $y^{[0]}$). Jistou výhodou této metody je „autokorekce" případných numerických chyb a zaokrouhlování. Jednotlivé body $y^{[k]}$ lze totiž považovat za nové výchozí body bez ohledu na „historii" jejich vzniku $y^{[0]}, \dots, y^{[k-1]}$. Navíc při výpočtu není nutné ihned počítat všechny iterace úplně přesně. Díky zmíněné autokorekční vlastnosti stačí začít s jedním desetinným místem a další přidat až v případě, kdy se hodnota ustálí atd. Toto zjednodušení oceníme obzvláště při ručních výpočtech.

Realizace výpočtu: Danou úlohu převedeme do základního tvaru a určíme (4.3.8), tj. vypočteme matici G a vektor h. Výpočet je pak realizován pomocí jednoduché tabulky. Do záhlaví

napíšeme jednotlivé složky matice G a vektoru h tak, že všechny členy na stejném řádku podělíme diagonálním prvkem, změníme znaménko a místo diagonálního prvku napíšeme 0. Do prvního řádku tabulky zapíšeme výchozí $y^{[0]}$ a číslo 1. Potom počítáme $y_1^{[1]}, y_2^{[1]}, \dots$ tak, že pro jejich i-tou složku potřebujeme i-tý řádek záhlaví, jehož složkami postupně vynásobíme předchozích $m+1$ předchozích prvků (včetně těch, které jsou o řádek výše). Jednotlivé součiny sečteme a přičteme poslední prvek i-tého řádku v záhlaví. Je-li výsledek kladný zapíšeme jej jako $y_i^{[k]}$, v opačném případě zapíšeme 0. Na konci každého řádku pak ještě napíšeme 1. Následující tabulka ilustruje výpočet $y_3^{[k]}$ v případě $m=4$.

	0	$-\frac{g_{12}}{g_{11}}$	$-\frac{g_{13}}{g_{11}}$	$-\frac{g_{14}}{g_{11}}$	$-\frac{h_1}{g_{11}}$
	$-\frac{g_{21}}{g_{22}}$	0	$-\frac{g_{23}}{g_{22}}$	$-\frac{g_{24}}{g_{22}}$	$-\frac{h_2}{g_{22}}$
$i=3$	$-\frac{g_{31}}{g_{33}}$	$-\frac{g_{32}}{g_{33}}$	0	$-\frac{g_{34}}{g_{33}}$	$-\frac{h_3}{g_{33}}$
	$-\frac{g_{41}}{g_{44}}$	$-\frac{g_{42}}{g_{44}}$	$-\frac{g_{43}}{g_{44}}$	0	$-\frac{h_4}{g_{44}}$
$y^{[0]}$	$y_1^{[0]}$	$y_2^{[0]}$	$y_3^{[0]}$	$y_4^{[0]}$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y^{[k-1]}$	$\vdots$	$\vdots$	$y_3^{[k-1]}$	$y_4^{[k-1]}$	1
$y^{[k]}$	$y_1^{[k]}$	$y_2^{[k]}$	⍟		

Tabulka 4.1: Ilustrace jednoho kroku Hildrethovy–d'Esopovy metody.

Číslo ⍟ pak má buď nulovou hodnotu nebo

$$y_3^{[k-1]} \cdot 0 - y_4^{[k-1]} \cdot \frac{g_{34}}{g_{33}} - 1 \cdot \frac{h_3}{g_{33}} - y_1^{[k]} \cdot \frac{g_{31}}{g_{33}} - y_2^{[k]} \cdot \frac{g_{32}}{g_{33}},$$

je-li tento výraz kladný.

Příklad 4.4.2

Hildrethova–D'Esopova metoda – řešený příklad #1:
Video ☞ 😀 https://ln.math.muni.cz/pz/opt/qp4

Zbývá ukázat konvergenci Hildrethovy–d'Esopovy metody v případě přípustné úlohy (4.4.1)[45] splňující dokonce podmínku (4.4.4), tj. že platí

$$\lim_{k\to\infty} \rho(y^{[k]}) = \rho^* = \min_{y\geq 0} \rho(y).$$

Přechod od jednoho přípustného bodu duální úlohy k druhé podle postupu popsaného v algoritmu výše označme jako P, tj. $y^{[k]} = P(y^{[k-1]})$. Pak není těžké ověřit, že zobrazení P je spojité a platí pro něj

$$\rho[P(u)] \leq \rho(u),$$

přičemž rovnost $\rho[P(y)] = \rho(y)$ nastane tehdy a jen tehdy, když $y = y^*$ pro jakékoli u^* představující řešení úlohy (4.3.8). Připusťme nyní, že množina

$$\Omega := \{y \in \mathbb{R}^m \mid y \geq 0 \ \ \& \ \ \rho[P(y)] \leq \rho(y^{[0]})\},$$

která obsahuje všechna $y^{[1]}, \dots$, je ohraničená. Tato množina jistě obsahuje všechny členy posloupnosti $\{y^{[k]}\}$, tudíž je tato posloupnost ohraničená a podle Bolzanovy–Weierstrassovy věty

[45] V opačném případě to ani nelze. Přesněji řešeno, v případě nepřípustné úlohy (4.4.1) bude úloha (4.3.8) neohraničená, a tudíž posloupnost $\{y^{[k]}\}$ bude divergentní.

[46] Nechť $y^{[k]}$ a $y^{[k+1]}$ jsou dvě po sobě jdoucí iterace získané Hildrethovou–d'Esopovou metodou. Jelikož posloupnost $\{y^{[k]}\}$ konverguje k řešení y^*, „zdá se být pravděpodobné", že y^* bude ležet poblíž přímky spojující $y^{[k]}$ a $y^{[k+1]}$. Proto lze očekávat, že minimalizací funkce ρ podél poloprímky

$$y(\lambda) := y^{[k]} + \lambda\,(y^{[k+1]} - y^{[k]}) = \\ = y^{[k]} + \lambda\, v^{[k]} \geq 0$$

nacházející se v přípustné množině duální úlohy dostaneme bod $\bar{y}^{[k+2]}$, který bude blíže k hledanému řešení y^* než následující iterace $y^{[k+2]}$ získaná Hildrethovou–d'Esopovou metodou. Označme jako $\tilde{\lambda}$ největší takové $\lambda \in \mathbb{R}$, pro které je $y(\lambda) \geq 0$, tj.

$$\tilde{\lambda} := \max\{\lambda \in \mathbb{R} \mid y(\lambda) \geq 0\} = \\ = \min\Big\{-y_i^{[k]} / v_i^{[k]} \mid v_i^{[k]} < 0\Big\},$$

Věta 4.4.3 o konvergenci H–D'E metody

přičemž nutně $\tilde{\lambda} \geq 1$ (neboť $\lambda(1) = y^{[k+1]} \geq 0$) a v případě $v^{[k]} \geq 0$ klademe $\tilde{\lambda} = \infty$. Označme nyní jako $\hat{\lambda}$ bod, ve kterém funkce

$$R(\lambda) := \rho(y(\lambda)) = \rho(y^{[k]} + \lambda\, v^{[k]}) = \\ = \rho(y^{[k]}) + \frac{\lambda^2}{2}\, v^{[k]\top} G\, v^{[k]} + \\ + \lambda\, v^{[k]\top}(G\, y^{[k]} + h)$$

nabývá svého lokálního (volného) minima. Z nutné (a současně postačující) podmínky $R'(\lambda) = 0$ plyne

$$\hat{\lambda} = -\frac{v^{[k]\top}(G\, y^{[k]} + h)}{v^{[k]\top} G\, v^{[k]}},$$

přičemž v případě $v^{[k]\top} G\, v^{[k]} = 0$ a $v^{[k]\top}(G\, y^{[k]} + h) < 0$ klademe $\hat{\lambda} = \infty$. Nyní označme jako $\lambda^{[k]}$ hodnotu, pro kterou nabývá funkce $R(\lambda)$ vázaného minima vzhledem k podmínce $y^{[k]} + \lambda\, v^{[k]} \geq 0$. Pak

$$\lambda^{[k]} = \min\{\tilde{\lambda}, \hat{\lambda}\}$$

a

$$\bar{y}^{[k+2]} = y(\lambda^{[k]}) = y^{[k]} + \lambda_k\, v^{[k]}.$$

Např. pro minimalizaci funkce

$$2x_1^2 - 6x_1 x_2 + 5x_2^2 - 16x_1 + 40x_2$$

za podmínek

$$x_1 + x_2 \geq 3 \quad \& \quad 2x_1 - x_2 \geq 2 \\ x_1 \geq 0 \quad \& \quad x_2 \geq 0,$$

dostaneme při výchozí hodnotě $y^{[0]} = [0,0,0,0]$ první tři iterace

$$y^{[1]} = [6;0;0;1],$$

má alespoň jeden hromadný bod $\overline{y}$. Jelikož ρ je spojitá funkce, je $\rho(\overline{y})$ hromadným bodem posloupnosti $\{\rho(y^{[k]})\}$. Současně i P je spojitá funkce, takže $P(\overline{y})$ je hromadným bodem posloupnosti $\{P(y^{[k]})\}$, a tudíž $\rho(P(\overline{y}))$ je hromadným bodem posloupnosti $\{\rho(P(y^{[k]}))\} = \{\rho(y^{[k+1]})\}$, což je ale stejné jako pro posloupnost $\{\rho(y^{[k]})\}$. Protože ale posloupnost $\{\rho(y^{[k]})\}$ je klesající, musí být oba hromadné body stejné, tj.

$$\rho(\overline{y}) = \rho(P(\overline{y})).$$

To ale znamená, že $\overline{y} = y^*$. Jelikož každá klesající posloupnost konverguje k hromadnému bodu, získáváme konečně požadovaný závěr, že posloupnost $\{\rho(y^{[k]})\}$ konverguje k $\rho(y^*) = \rho^*$.

Vraťme se ještě k ohraničenosti množiny Ω. Připusťme, že množina Ω není ohraničená. Pak funkce

$$\rho_\varepsilon(y) := \rho(y) - \delta \sum_{j=1}^{m} y_j$$

je pro libovolné $\delta > 0$ a $y \geq 0$ zdola neohraničená. Avšak funkce $\rho_\varepsilon(y)$ je účelovou funkcí duální úlohy pro modifikovanou primární úlohu, ve které nahradíme každé z omezení $\langle a_i, x\rangle \leq b_i$ omezením (4.4.4). Vzhledem k požadované přípustnosti primární úlohy je přípustná i modifikovaná primární úloha, a tudíž účelová funkce $\rho_\varepsilon(y)$ je ohraničená pro všechna $y \geq 0$. To je ale spor, který dokazuje ohraničenost množiny Ω. Tím jsme dokázali následující tvrzení.

> Posloupnost generovaná pomocí Hildrethovy–d'Esopovy metody pro úlohu (4.4.1) splňující (4.4.4) konverguje k řešení odpovídající duální úlohy.

Odtud plyne, že všechny body posloupnosti $\{x(y^{[k]})\}$ se členy

$$x(y^{[k]}) := -C^{-1}(A^\top y^{[k]} + d)$$

leží v ohraničené množině, která je obrazem množiny Ω skrze lineární transformaci (4.4.3), a konverguje ke svému jedinému hromadnému bodu $x^* = x(y^*)$. Nicméně měli bychom také upozornit, že konvergence Hildrethovy–d'Esopovy metody může být v některých případech poměrně pomalá, a tak se mohou hodit i různé modifikace pro její zrychlení – to už ale ponecháme k prostudování pouze případným zájemcům[46].

Pokud se nám nepoštěstí najít přesné řešení duální úlohy pomocí Hildrethovy–d'Esopovy metody v konečném (větším či menším) počtu kroků, tak asi výpočet po nějaké době ukončíme a spokojíme se jen s nějakou jeho aproximací. Jak přesnou? A naopak, kolik kroků by mělo stačit k tomu, abychom s jistotou dosáhli požadované přesnosti? Stanovme si proto tolerance pro porušení jednotlivých omezení $\varepsilon_1, \ldots, \varepsilon_m$ a požadovanou přesnost Δ určení minimální hodnoty $f(x^*) = \frac{1}{2}\langle Cx^*, x^*\rangle + \langle d, x^*\rangle$.

V první etapě nejdříve budeme počítat iterace $y^{[k]}$, dokud odpovídající bod $x(y^{[k]})$ nebude ε-přípustný, tj. dokud nezískáme $y^{[k]}$ splňující $Gy^{[k]} + h + \varepsilon \geq 0$. Z konvexnosti funkce f vyplývá[47] pro libovolný bod $x(y)$ určený vztahem (4.4.3) pro libovolné $y \geq 0$, že

$$f(x(y)) - f(x^*) \leq \langle \operatorname{grad} f(x(y)), x(y) - x^*\rangle = \big[-CC^{-1}(A^\top y + d) + d\big]^\top (x(y) - x^*) = \\ = -y^\top (Ax^* - Ax(y)) \leq -y^\top (b - Ax(y)) = -y^\top (Gy + h) =: \widehat{\Delta}.$$

Navíc, označíme-li

$$J^* := \{j \in \{1, \ldots, m\} \mid (Gy + h)_j < 0\} \quad \text{a} \quad \widetilde{\Delta} := -\sum_{j \in J^*} y_j^* (Gy + h)_j \geq 0,$$

pak zcela analogickým výpočtem jako před chvílí dostaneme

$$f(x^*) - f(x(y)) \leq \langle \operatorname{grad} f(x^*), x^* - x(y)\rangle \leq y^{*\top}(Gy + h) \leq \widetilde{\Delta}. \tag{4.4.5}$$

Na rozdíl od čísla $\widehat{\Delta}$ ale hodnotu $\widetilde{\Delta}$ dokážeme určit pouze v případě, že známe přesné řešení duální úlohy y^*. Musíme jej proto nahradit nějakým odhadem. Je-li y dostatečně blízko y^*,

$$y^{[2]} = [5,6;0;0;2],$$
$$y^{[3]} = [5,2;0;0;3],$$

přičemž přesné řešení je $y^* = [0;0;0;16]$. V tomto případě máme

$$v^{[1]} = y^{[2]} - y^{[1]} = (-0,4;0;0;1)^\top,$$

a tedy $v^{[1]\top} G v^{[1]} = 0,04$ a $v^{[1]\top}(G y^{[1]} + h) = -1$, takže $\hat\lambda = 25$ a

$$\tilde\lambda = \min\left\{ -y_i^{[1]} / v_i^{[1]} \mid v_i^{[1]} < 0 \right\} = 15.$$

Proto $\lambda_1 = 15$ a $\bar y^{[3]} = [0;0;0;16]$, což je přesně hledané řešení y^*. Nicméně neměli bychom toto extrémní urychlení konvergence nikterak zveličovat (klasická Hildrethova–d'Esopova metoda by v závislosti na přesnosti výpočtu vyžadovala nejméně 24 kroků). V případech, kdy v jednotlivých iteracích máme body s větším počtem nenulových složek, není urychlení tak výrazné. Kdybychom změnili pořadí výpočtu jednotlivých složek $y^{[k]}$ v jednom kroku na y_2, y_3, y_4 a y_1, tak bychom v prvních třech krocích získali

$$y^{[1]} = [0;5,2;0,7;4,8]$$
$$y^{[2]} = [0;2,9;1,2;8,4]$$
$$y^{[3]} = [0;1;1,6;11,6].$$

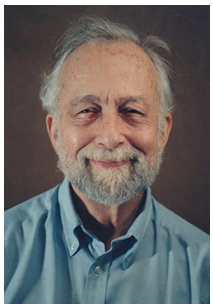
Pak

$$v^{[1]} = (0;-2,3;0,5;3,8),$$

a tedy $v^{[1]\top} G v^{[1]} = 4,205$ a $v^{[1]\top}(G y^{[1]} + h) = -23,18$, takže $\hat\lambda = 5,51$ a $\tilde\lambda = 2,26$. Proto $\lambda_1 = 2,26$ a $\bar y^{[3]} = [0;0;1,8;13,2]$, což je sice není přesné řešení, ale je to o něco lepší než $y^{[3]}$.

[47] Zde (v první nerovnosti následujícího výpočtu) by se nám hodila znalost Věty 6.4.3. Ale snad to zvládneme i bez ní a nedůvěřiví čtenáři si to mohou zkontrolovat.

[48]



Philip Starr „Phil" Wolfe (11. srpna 1927 – 29. prosince 2016) byl americký matematik, který je považován za jednoho z „otců zakladatelů" teorie konvexního programování. Na základní škole patřil mezi vynikající žáky a miloval vědu, ale na střední škole se pro něj stala překážkou na algebře a geometrii. Naštěstí si přečetl Eu-

pak jej můžeme využít pro stanovení přibližného odhadu $\widetilde\Delta$ jako $-\sum_{j\in J^*} y_j (Gy+h)_j \ge 0$, kterou ještě raději vynásobíme nějakým vhodným koeficientem $\alpha(y) \ge 1$, tj. místo nerovnosti (4.4.5) budeme mít

$$f(x^*) \le f(x(y)) + \widetilde\Delta^o$$

pro $\widetilde\Delta^o := -\alpha(y) \sum_{j\in J^*} y_j (Gy+h)_j$. Budeme-li proto ve výpočtu dalších členů $y^{[k]}$ pokračovat ve druhé etapě tak dlouho, dokud

$$\left|\widehat\Delta_k\right| < \Delta \quad \text{a současně} \quad \left|\widetilde\Delta_k^o\right| < \Delta,$$

pak s využitím těchto dvou odhadů získáme

$$f(x^{[k]}) - \widehat\Delta_k \le f(x^*) \le f(x^{[k]}) + \widetilde\Delta_k^o \quad \text{neboli} \quad \left|f(x^{[k]}) - f(x^*)\right| < \Delta,$$

tj. poslední $x^{[k]}$ bude (ε,Δ)-řešením úlohy (4.4.1).

4.5 Wolfeho metoda v krátkém tvaru

Jakkoli je princip předchozí metody snadný, je za něj velmi draze zaplaceno velmi silnými požadavky zahrnující $C > 0$ a Slaterovu podmínku. Je proto potřeba mít k dispozici i algoritmus pro řešení obecnějších úloh (QP), zejména dovolující i singulární matice $C \ge 0$. To nás přivádí k Wolfeho[48] metodě z roku 1959 založené na Karushových–Kuhnových–Tuckerových podmínkách z (4.3.5) a upraveném simplexovém algoritmu, viz [512]. Ačkoli podobně jako v případě lineárního programování dnes už existují efektivnější metody (ať již založené na simplexovém algoritmu nebo na zcela jiných základech), my u této metody zůstaneme kvůli jejímu přirozenému navázaní na simplexový algoritmus a také její názornost (a celkem „snadný" výpočet). Toto jsou vlastně také důvody ospravedlňující velké rozšíření a popularitu tohoto algoritmu v 60. letech. Tato metoda byla původně rozpracována pro úlohu kvadratického programování v kanonickém tvaru a její popis lze pro tento případ nalézt v řadě monografií jako např. [337] a také ve skriptech [143, Sekce 7.5]. Jenže my se touto cestou nevydáme, a to minimálně ze dvou důvodů:

(i) potřebná modifikace simplexové metody ubírá této metodě na „čistotě";

(ii) většina (zajímavých) úloh v $\mathbb{R}^2$, které umíme vizualizovat (a tedy řešit i geometricky), by vyžadovaly pro převod na kanonický tvar zavedení pomocných přídavných proměnných, což by nutně vedlo k použití tzv. *dlouhého tvaru* Wolfeho metody (odpovídající účelová funkce by byla vzhledem k novým proměnným singulární, jak jsme si ukázali za Větou 4.3.1). Toto ale v našem pojetí Wolfeho metody potřeba nebude. Pro úplnost ještě dodejme, že bude-li původní úloha v základním tvaru, pak její řešení Wolfeho metodou bude vždy vyžadovat dlouhý tvar (bez ohledu na to zda použijeme původní algoritmus nebo ten „vylepšený").

Naším výchozím bodem tedy je[49] úloha kvadratického programování

$$\left.\begin{array}{c} \frac12\langle Cx,x\rangle + \langle d,x\rangle \to \min \\ Ax \le b \quad \& \quad x \ge 0 \end{array}\right\} \tag{4.5.1}$$

s maticí $C \ge 0$ a odpovídajícími Karushovými–Kuhnovými–Tuckerovými podmínkami

$$-Cx - A^\top y + u = d \tag{4.5.2a}$$
$$Ax + v = b \tag{4.5.2b}$$
$$x \ge 0 \quad \& \quad y \ge 0 \quad \& \quad u \ge 0 \quad \& \quad v \ge 0 \tag{4.5.2c}$$
$$\langle x,u\rangle = 0 \quad \& \quad \langle y,v\rangle = 0 \tag{4.5.2d}$$

Připusťme pro tuto chvíli, že $b \ge 0$. V opačném případě bude podobně jako v lineárním programování nutná tzv. *první fáze* Wolfeho metody, kterou si ukážeme později. Najdeme-li řešení

klidovy *Základy*, díky čemuž (jak sám Wolfe říkal) se z něj a Euklida „stali kolegové“. V roce 1943 nastoupil na Kalifornskou univerzitu v Berkeley, ale studium musel přerušit kvůli nástupu do armády krátce před koncem druhé světové války. V roce 1947 se na Berkeley vrátil a v roce 1954 zde získal doktorát, když se ve své disertační práci chtěl věnovat především teorii her. Tento zájem jej přivedl k Projektu SCOOP a k setkání s Dantzigem, který jej ponoukl k vyřešení problému s možným zacyklením simplexového algoritmu. To se Wolfemu povedlo, viz [124], s využitím *lexikografického* pravidla, které se stalo součástí jeho disertační práce. Po zisku doktorátu přijal Wolfe pozici na univerzitě v Princetonu, kde se věnoval problémům kvadratického programování. Společně s Marguerite Frank(ovou) navrhli tzv. *Frankové–Wolfeho* algoritmus pro řešení úloh kvadratického programování, který lze aplikovat i na úlohy konvexního programování, viz [184]. Ve stejném čísle časopisu *Naval Research Logistics Quarterly* (vlastně hned na další straně) vyšel i Markowitzův článek [364] ohledně minimalizace kvadratické funkce s využitím kritické přímky. Před nástupem na Princeton dostal i nabídku ze společnosti RAND, kterou ale tehdy odmítl. V roce 1957 RAND svoji nabídku zdvojnásobil, a tak se Wolfe vrátil do Kalifornie, kde společně s Dantzigem a dalšími pracovali na zlepšení simplexového algoritmu. Z této doby je i jeho modifikace simplexového algoritmu pro řešení úloh kvadratického programování z [512]. V roce 1964 zavítal na pozvání svého přítele Gomoryho na půlroční stáž do společnosti IBM, do které následně nastoupil, a strávil v ní zbytek své profesní kariéry. V roce 1992 získal společně s Hoffmanem *John von Neumann Theory Prize* za „jejich příspěvky k intelektuálním základům matematického programování“ (their contributions to the intellectual foundations of mathematical programming). Na webových stránkách *INFORMS* je dokonce dostupný i záznam velmi zajímavého rozhovoru s Wolfem z května 2001, viz https://ln.math.muni.cz/pz/opt/wolfe1.

[49] Stejně jako v [155, 250]. Jiné zdroje s touto úpravou Wolfeho metody se autorovi kupodivu najít nepovedlo. Nicméně rozdíl v popisu i důkazech pro tuto metodu a Wolfeho původní verzi není příliš velký.

$[x^*, y^*, u^*, v^*]$ soustavy (4.5.2), pak odpovídající x^* bude řešením úlohy (4.5.1) podle Věty 4.3.1 a navazujícího komentáře. Podmínky (4.5.2a)–(4.5.2c) jsou lineární, zatímco (4.5.2d) je nelineární. Z poslední podmínky však vyplývá, že z každého páru proměnných x_i, u_i a y_i, v_i je alespoň jedna složka nulová, a tedy nejvýše $n+m$ proměnných v $(x,u,y,v)^\top \in \mathbb{R}^{2(n+m)}$ je nenulových. Avšak to je právě tolik, kolik má soustava (4.5.2a)–(4.5.2b) rovnic, tj. nejvýše tolik proměnných je nenulových, kolik má soustava rovnic. Jelikož platí, že má-li soustava lineárních rovnic nezáporné řešení, pak má i nezáporné bazické řešení, stačí se při řešení (4.5.2a)–(4.5.2d) omezit na nezáporné bazické řešení[50]. Zbývá už jen vhodně doplnit soustavu (4.5.2), abychom mohli vytvořit úlohu lineárního programování (nebo něco podobného), jejímž řešením takový bazický přípustný bod najdeme.

Hlavní myšlenka dalšího postupu je podobná jako v první fázi simplexového algoritmu. Tentokrát to však je tzv. *druhá fáze* Wolfeho metody: zavedením umělých nezáporných proměnných rozšíříme soustavu (4.5.2a)–(4.5.2c) do podoby, ve které bude snadné určit výchozí BPB splňující (4.5.2d). Potom simplexovým algoritmem anulujeme umělé proměnné, přičemž stále máme na paměti podmínku (4.5.2d). Pokud se nám takto podaří vyloučit všechny umělé proměnné z báze, pak konečná simplexová tabulka bude určovat řešení soustavy (4.5.2a)–(4.5.2d).

Nyní si to rozebereme trochu podrobněji. Soustavu (4.5.2a) rozšíříme o umělé proměnné $z \in \mathbb{R}^n$ tak, že dostaneme soustavu

$$\left.\begin{aligned} -Cx - A^\top y + u + Dz &= d \\ Ax + v &= b \\ x \ge 0 \quad \& \quad y \ge 0 \quad \& \quad u \ge 0 \quad \& \quad v \ge 0 \quad \& \quad z \ge 0, & \end{aligned}\right\} \tag{4.5.3}$$

kde $D \in \mathbb{R}^{n\times n}$ je diagonální matice s diagonálními prvky

$$d_{ii} = \begin{cases} -1, & d_i < 0, \\ 1, & d_i \ge 0, \end{cases} \qquad i = 1, \dots, n,$$

tj. $d_{ii} = \operatorname{sgn} d_i$ pro $d_i \ne 0$. Toto lze maticově zapsat jako

$$\begin{matrix} & \overbrace{}^{n\quad m\quad n\quad m\quad n} \\ n\{ \\ m\{ \end{matrix} \begin{pmatrix} -C & -A^\top & I & 0 & D \\ A & 0 & 0 & I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ u \\ v \\ z \end{pmatrix} = \begin{pmatrix} d \\ b \end{pmatrix}. \tag{4.5.4}$$

Potom $[x,y,u,v,z] = \big[0,0,0,b,|d|\big]$ je BPB. Zde symbol $|d|$ značí bod $\big[|d_1|,\dots,|d_n|\big]$ a „bazickost“ zmíněného bodu plyne ihned z toho, že odpovídající část matice na levé straně (4.5.4) je $\left(\begin{smallmatrix} 0 & D \\ I & 0 \end{smallmatrix}\right)$, která má plnou hodnost $n+m$. Nyní za účelem anulování umělých proměnných řešíme úlohu

$$z_1 + \cdots + z_n \to \min \qquad \text{za podmínek (4.5.3).} \tag{4.5.5}$$

Ovšem ne každé řešení úlohy (4.5.5) musí vyhovovat podmínce (4.5.2d), proto je nutné doplnit simplexový algoritmus o *Wolfeho pravidlo*:

> Je-li x_i (nebo y_i) bazickou proměnnou, pak při přechodu k další simplexové tabulce a novému BPB se nesmí u_i (nebo v_i) stát bazickou proměnnou a naopak, tj. je-li u_i (nebo v_i) v bázi, nesmí do báze vstoupit x_i (nebo y_i).

Toto pravidlo nám zaručuje, že v žádném kroku nebudou x_i a u_i (stejně jako y_j a v_j) nikdy současně bazickými proměnnými pro nějaké $i \in \{1,\dots,n\}$ (a $j \in \{1,\dots,m\}$). Pokud se nám pomocí tohoto upraveného simplexového algoritmu podaří získat řešení úlohy (4.5.5) splňující $\sum_{i=1}^n z_i = 0$, potom poslední simplexová tabulka určuje řešení $[x^*, y^*, u^*, v^*]$ soustavy (4.5.3),

[50] To je tzv. *simplexový princip*, na kterém je založen simplexový algoritmus. V soustavě (4.5.2) nemáme žádnou účelovu funkci, zde nám jde jen o nalezení nějakého přípustného bodu (tj. bodu vyhovující soustavě), a tedy stačí najít bazický přípustný bod.

a tedy odpovídající x^* je řešením původní úlohy (4.5.1). Toto s případným doplnění o první fázi je tzv. *krátký tvar Wolfeho metody*[51].

Podaří se nám ale takto vždy najít řešení? Prvním úskalím může být zacyklení simplexového algoritmu, ale to již umíme eliminovat vhodnou volbou pravidel. Mnohem horší nebezpečí je ukryto v situaci, kdy $\sum_{i=1}^n z_i > 0$, avšak kvůli Wolfeho pravidlu není možné pokračovat dále. Kdy se to může stát? Odpověď dává následující věta, která je „ukryta“ i ve Wolfeho článku [512].

Věta 4.5.1 Wolfeho

Je-li $C > 0$ nebo $d = 0$, pak je úloha (4.5.5) řešitelná simplexovým algoritmem doplněným o Wolfeho pravidlo, tj. po konečném počtu kroků dostaneme $\sum_{i=1}^n z_i = 0$.

Důkaz. Bez újmy na obecnosti můžeme předpokládat, že úloha (4.5.5) je nedegenerovaná. V opačném případě bychom museli využít perturbační metodu, abychom se vyhnuli případnému zacyklení a současně nulovým hodnotám pro bazické proměnné[52]. Připusťme, že jsme buď nalezli řešení úlohy (4.5.5) nebo že není v důsledku Wolfeho pravidla možné pokračovat v simplexovém algoritmu. Nechť $[x,y,u,v,z]$ je BPB určený poslední simplexovou tabulkou a označme

$$N_1 := \{i \in \{1,\dots,n\} \mid x_i > 0 \ \&\ u_i = 0\} \qquad M_1 := \{j \in \{1,\dots,m\} \mid y_i > 0 \ \&\ v_i = 0\},$$
$$N_2 := \{i \in \{1,\dots,n\} \mid x_i = 0 \ \&\ u_i > 0\} \qquad M_2 := \{j \in \{1,\dots,m\} \mid y_i = 0 \ \&\ v_i > 0\},$$
$$N_3 := \{i \in \{1,\dots,n\} \mid x_i = 0 \ \&\ u_i = 0\} \qquad M_3 := \{j \in \{1,\dots,m\} \mid y_i = 0 \ \&\ v_i = 0\}.$$

Indexové množiny N_1, N_2, N_3 jsou nepochybně disjunktní a $N_1 \cup N_2 \cup N_3 = \{1,\dots,n\}$, neboť případ $x_i > 0$ a $u_i > 0$ není možný díky Wolfeho pravidlu. Analogické vlastnosti má i druhá trojice indexových množin. Pro jednoduchost přeznačíme indexy i a j tak, aby jednotlivé složky patřily postupně do N_1, N_2 a N_3 a podobně pro M_1, M_2 a M_3, tj.

$$x = \left[x^{[1]}, x^{[2]}, x^{[3]}\right] \quad \& \quad u = \left[u^{[1]}, u^{[2]}, u^{[3]}\right] \quad \& \quad d = \left[d^{[1]}, d^{[2]}, d^{[3]}\right]$$
$$y = \left[y^{[1]}, y^{[2]}, y^{[3]}\right] \quad \& \quad v = \left[v^{[1]}, v^{[2]}, v^{[3]}\right] \quad \& \quad b = \left[b^{[1]}, b^{[2]}, b^{[3]}\right]$$

splňující $x^{[1]} > 0$, $u^{[2]} > 0$, $x^{[2]} = 0$, $u^{[1]} = 0$, $x^{[3]} = 0 = u^{[3]}$ a podobně $y^{[1]} > 0$, $v^{[2]} > 0$, $y^{[2]} = 0$, $v^{[1]} = 0$, $y^{[3]} = 0 = v^{[3]}$. Při tomto přeznačení je poslední BPB získaný při řešení úlohy (4.5.5) simplexovým algoritmem doplněným o Wolfeho pravidlo shodný s řešením úlohy lineárního programování

$$\left.\begin{aligned} z_1 + \cdots + z_n &\to \min \\ -Cx - A^\top y + u + Dz &= d \\ Ax + v &= b \\ x \ge 0 \quad \& \quad y \ge 0 \quad \& \quad u \ge 0 \quad \& \quad v \ge 0 \quad \& \quad z \ge 0 \\ x^{[2]} = 0 \quad \& \quad u^{[1]} = 0 \quad \& \quad y^{[2]} = 0 \quad \& \quad v^{[1]} = 0. \end{aligned}\right\} \tag{4.5.6}$$

Pak poslední simplexová tabulka úlohy (4.5.5) s vynechanými složkami $x^{[2]}, u^{[1]}, y^{[2]}, v^{[1]}$ určuje řešení[53] úlohy (4.5.6). Schematicky lze omezení této úlohy zapsat do následující (neúplné) simplexové tabulky s levým krajním sloupcem pro duální proměnné.

	$x^{[1]}>0$	$x^{[2]}=0$	$x^{[3]}$	$y^{[1]}>0$	$y^{[2]}=0$	$y^{[3]}$	$u^{[1]}=0$	$u^{[2]}>0$	$u^{[3]}$	$v^{[1]}=0$	$v^{[2]}>0$	$v^{[3]}$	z	
δ_1	$-C_{11}$	$-C_{12}$	$-C_{13}$	$-A_{11}^\top$	$-A_{21}^\top$	$-A_{31}^\top$	I	0	0	0	0	0	D_1	$d^{[1]}$
δ_2	$-C_{21}$	$-C_{22}$	$-C_{23}$	$-A_{12}^\top$	$-A_{22}^\top$	$-A_{32}^\top$	0	I	0	0	0	0	D_2	$d^{[2]}$
δ_3	$-C_{31}$	$-C_{32}$	$-C_{33}$	$-A_{13}^\top$	$-A_{23}^\top$	$-A_{33}^\top$	0	0	I	0	0	0	D_3	$d^{[3]}$
β_1	A_{11}	A_{12}	A_{13}	0	0	0	0	0	0	I	0	0	0	$b^{[1]}$
β_2	A_{21}	A_{22}	A_{23}	0	0	0	0	0	0	0	I	0	0	$b^{[2]}$
β_3	A_{31}	A_{32}	A_{33}	0	0	0	0	0	0	0	0	I	0	$b^{[3]}$

Tabulka 4.2: Schematický zápis omezení úlohy (4.5.6).

[51] Výpočet v této fázi lze výrazně urychlit tím, že zavedeme pouze tolik pomocných proměnných z_i, kolik má záporných složek vektor d, tj. pokud $d_i \ge 0$ pro nějaké $i \in \{1,\dots,n\}$, pak místo zavedení z_i ponecháme v bázi u_i. V takovém případě bude výchozím BPB $A := [x,y,u,v,z] = [0,0,\bar{u},b,\bar{z}]$, kde $\bar{u} = [\bar{u}_1,\dots,\bar{u}_n]$ a $\bar{z} = [\bar{z}_1,\dots,\bar{z}_n]$ jsou dána jako

$$\bar{u}_i = \begin{cases} 0 & d_i < 0, \quad i \in \{1,\dots,n\}, \\ d_i & d_i \ge 0, \quad i \in \{1,\dots,n\}, \end{cases}$$

$$\bar{z}_i = \begin{cases} |d_i| & d_i < 0, \quad i \in \{1,\dots,n\}, \\ 0 & d_i \ge 0, \quad i \in \{1,\dots,n\}. \end{cases}$$

Potom hodnota účelové funkce úlohy ve druhé fázi v bodě A je nejvýše rovna hodnotě této účelové funkce v „klasickém“ výchozím BPB popsaném výše. Proto stačí zavést pouze taková z_i, pro která je $\bar{z}_i \ne 0$. Kdybychom měli dokonce dánu úlohu s $d \ge 0$, pak není potřeba zavádět žádné z_i a máme ihned řešení. Jaké? V takovém případě je bod $[0,0,0,b,d]$ řešením soustavy (4.5.3), takže $x^* = 0$ je řešením původní úlohy (4.5.1). To je zřejmé z toho, že $C \ge 0$, $d \ge 0$, $b \ge 0$ a $x \ge 0$ znamená, že nejmenší možná hodnota účelové funkce je 0, kterou dává právě bod $x^* = 0$.

[52] Blandovo pravidlo i lexikografická metoda by proti zacyklení zafungovaly samozřejmě také, ale nulovosti bazických proměnných již nezabrání. Pro jednoduchost by i stačilo požadovat pouze, aby druhá fáze skončila s nedegenerovaným BPB.

[53] V tuto chvíli jsou totiž díky předpokladu nedegenerovanosti v bázi pouze proměnné $x^{[1]}, u^{[2]}, y^{[1]}, v^{[2]}$ doplněné o některé z proměnných $z_1,\dots,z_n$, a tedy v řádku $z_k - c_k$ pro tyto proměnné jsou nulové hodnoty. Pokud jsme nalezli řešení, jsou v řádku $z_k - c_k$ poslední simplexové tabulky pouze nekladné hodnoty. Pokud další krok simplexové metody není možný v důsledku Wolfeho pravidla, musí být v řádku $z_k - c_k$ ve sloup-

cích odpovídajících proměnným $x^{[3]}, u^{[3]}, y^{[3]}$ a $v^{[3]}$ nekladné hodnoty, zatímco ve sloupcích odpovídajících proměnným $x^{[2]}, u^{[1]}, y^{[2]}, v^{[1]}$ musí být alespoň jedna kladná hodnota. Každopádně bude při „ignorování“ sloupců $x^{[2]}, u^{[1]}, y^{[2]}, v^{[1]}$ poslední simplexová tabulka udávat řešení úlohy (4.5.6).

Pak duální úloha je

$$\begin{gathered}
\delta_1^\top d^{[1]} + \delta_2^\top d^{[2]} + \delta_3^\top d^{[3]} + \beta_1^\top b^{[1]} + \beta_2^\top b^{[2]} + \beta_3^\top b^{[3]} \to \max \\
-\delta_1^\top C_{11} - \delta_2^\top C_{21} - \delta_3^\top C_{31} + \beta_1^\top A_{11} + \beta_2^\top A_{21} + \beta_3^\top A_{31} \le 0 \\
-\delta_1^\top C_{13} - \delta_2^\top C_{23} - \delta_3^\top C_{33} + \beta_1^\top A_{13} + \beta_2^\top A_{23} + \beta_3^\top A_{33} \le 0 \\
-\delta_1^\top A_{11}^\top - \delta_2^\top A_{12}^\top - \delta_3^\top A_{13}^\top \le 0 \\
-\delta_1^\top A_{31}^\top - \delta_2^\top A_{32}^\top - \delta_3^\top A_{33}^\top \le 0 \\
\delta_1^\top D_1 + \delta_2^\top D_2 + \delta_3^\top D_3 \le e^\top \\
\delta_2 \le 0 \quad \& \quad \delta_3 \le 0 \quad \& \quad \beta_2 \le 0 \quad \& \quad \beta_3 \le 0,
\end{gathered}$$

kde $e := (1, \dots, 1)^\top$, viz Podkapitolu 2.6. Jelikož pro řešení primární úlohy (4.5.6) jsou $x^{[1]} > 0$, $y^{[1]} > 0$, $u^{[2]} > 0$, $v^{[2]} > 0$ musí se vzhledem k podmínkám komplementarity odpovídající omezení v duální úloze (první, třetí, šesté a osmé) realizovat jako rovnosti, tj. $\delta_2 = 0$ a $\beta_2 = 0$ a

$$-\delta_1^\top C_{11} - \delta_3^\top C_{31} + \beta_1^\top A_{11} + \beta_3^\top A_{31} = 0 \tag{4.5.7a}$$
$$-\delta_1^\top C_{13} - \delta_3^\top C_{33} + \beta_1^\top A_{13} + \beta_3^\top A_{33} \le 0 \tag{4.5.7b}$$
$$-\delta_1^\top A_{11}^\top - \delta_3^\top A_{13}^\top = 0 \tag{4.5.7c}$$
$$-\delta_1^\top A_{31}^\top - \delta_3^\top A_{33}^\top \le 0 \tag{4.5.7d}$$
$$\delta_1^\top D_1 + \delta_3^\top D_3 \le e^\top \tag{4.5.7e}$$
$$\delta_3 \le 0 \quad \& \quad \beta_3 \le 0 \tag{4.5.7f}$$

a pro účelovou funkci musí platit

$$\delta_1^\top d^{[1]} + \delta_3^\top d^{[3]} + \beta_1^\top b^{[1]} + \beta_3^\top b^{[3]} = e\,z = z_1 + \dots + z_n. \tag{4.5.8}$$

Vynásobíme-li zprava rovnost (4.5.7a) vektorem δ_1, nerovnost (4.5.7b) vektorem $\delta_3 \le 0$, rovnost (4.5.7c) vektorem β_1 a konečně nerovnost (4.5.7d) vektorem $\beta_3 \le 0$, pak po jejich sečtení dostaneme

$$-\delta_1^\top C_{11}\delta_1 - \delta_3^\top C_{31}\delta_1 - \delta_1^\top C_{13}\delta_3 - \delta_3^\top C_{33}\delta_3 \ge 0 \quad \text{neboli} \quad -\begin{pmatrix}\delta_1\\ \delta_3\end{pmatrix}^\top \begin{pmatrix}C_{11} & C_{13}\\ C_{31} & C_{33}\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_3\end{pmatrix} \ge 0.$$

Vzhledem k pozitivní semidefinitnosti a symetrii původní matice C může být tato nerovnost splněna pouze v případě

$$\delta_1^\top C_{11} + \delta_3^\top C_{31} = 0 \quad \& \quad \delta_1^\top C_{13} + \delta_3^\top C_{33} = 0. \tag{4.5.9}$$

Pak se díky první rovnosti redukuje (4.5.7a) na

$$\beta_1^\top A_{11} + \beta_3^\top A_{31} = 0. \tag{4.5.10}$$

Nyní vynásobením zleva první části druhého omezení úlohy (4.5.6) vektorem β_1 a třetí části tohoto omezení vektorem β_3 dostaneme (viz Tabulku 4.2)

$$\beta_1^\top A_{11} x^{[1]} + \beta_1^\top A_{13} x^{[3]} = \beta_1^\top b^{[1]} \quad \text{a} \quad \beta_3^\top A_{31} x^{[1]} + \beta_3^\top A_{33} x^{[3]} + \beta_1^\top v^{[3]} = \beta_3^\top b^{[3]},$$

což po jejich sečtení dává

$$\left(\beta_1^\top A_{11} + \beta_3^\top A_{31}\right) x^{[1]} + \left(\beta_1^\top A_{13} + \beta_3^\top A_{33}\right) x^{[3]} + \beta_1^\top v^{[3]} = \beta_1^\top b^{[1]} + \beta_3^\top b^{[3]}.$$

Avšak $x^{[3]} = 0 = v^{[3]}$, takže i s využitím (4.5.10) vidíme, že levá strana předchozí rovnosti je nulová. Proto rovnost (4.5.8) můžeme zapsat jako

$$\delta_1^\top d^{[1]} + \delta_3^\top d^{[3]} = z_1 + \dots + z_n. \tag{4.5.11}$$

Je-li matice C pozitivně definitní, pak soustava (4.5.9) má pouze nulové řešení, tj. $\delta_1 = 0 = \delta_3$, a tedy zjevně $z_1 + \dots + z_n = 0$. Stejný závěr očividně dostaneme i v případě $d = 0$, čímž je věta dokázána. ■

Samozřejmě je možné, že tento algoritmus bude fungovat i v případě, kdy $C \geq 0$, $\det C = 0$ a $d \neq 0$, ovšem pro důkaz konvergence je potřebné $C > 0$ nebo $d = 0$. Alespoň zatím… Možné problémy s Wolfeho pravidlem si ilustrujeme v následujícím krátkém příkladě.

Příklad 4.5.2

Wolfeho metodou v krátkém tvaru vyřešme úlohu

$$x_1^2 + 2x_1 - x_2 \to \min$$
$$x_1 + x_2 \leq 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení. V této úloze máme

$$C = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad d = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad b = 2.$$

Matice C je pouze pozitivně semidefinitní, takže nemáme zaručenu řešitelnost úlohy krátkým tvarem Wolfeho metody. Jelikož $b > 0$, můžeme začít druhou fází s jednou umělou proměnnou díky $d_1 > 0$, tj. chceme vyřešit úlohu

$$z \to \min$$
$$-2x_1 - y + u_1 = 2$$
$$-y + u_2 - z = -1$$
$$x_1 + x_2 + v = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad y \geq 0 \quad \& \quad u_1 \geq 0 \quad \& \quad u_2 \geq 0 \quad \& \quad v \geq 0 \quad \& \quad z \geq 0.$$

Za výchozí BPB můžeme vzít

$$[x_1, x_2, y, u_1, u_2, v, z] = [0, 0, 0, 2, 0, 2, 1]$$

a začít se simplexovou metodou, což je učiněno v následující tabulce.

c_B	B	x_1	x_2	y	u_1	u_2	v	z	x_B	x_B/λ
0	u_1	−2	0	−1	1	0	0	0	2	✖
1	z	0	0	(1)	0	−1	0	1	1	1 →
0	v	1	1	0	0	0	1	0	2	✖
	c_k	0	0	0	0	0	0	1		
	z_k	0	0	1	0	−1	0	1		
	$z_k - c_k$	0	0	1 ↑	0	−1	0	0	1	

Nyní bychom měli do báze zavést proměnnou y, to však není možné, neboť v bázi již je v, a tudíž bychom se dostali do sporu s Wolfeho pravidlem[54]. Zdá se, že úlohu nelze vyřešit Wolfeho metodou, kterou jsme si poměrně pracně vybudovali. Naštěstí v tomto jednoduchém příkladě to možné je, i když za cenu trochu „násilného“ vynucení dalšího kroku simplexové metody. Zavedeme-li totiž do báze místo y proměnnou[55] x_2, opustí bázi v a v dalším kroku ji opustí i z. Tímto vyeliminováním z z báze se dopracujeme k řešení pomocné úlohy, a tudíž i původní dané úlohy $x^* = [0, 2]$. Postup jednotlivých kroků je zapsán v následující tabulce[56].

[51] Všimněme si, že tato tabulka je zcela v souladu se situací popsanou v souvislosti s úlohou 4.5.6, tj. vynecháním nebazických proměnných $x^{[2]} = x_1$ a $y^{[2]} = y$ ($u^{[1]}$ a $v^{[1]}$ „neexistují“) dostaneme simplexovou tabulku určující řešení této pomocné úlohy.

[52] To lze zcela bez problémů díky nulové hodnotě v řádku $z_k - c_k$.

[53] Stejně jako při řešení klasických úloh lineárního programování i zde využíváme toho, že po vyloučení proměnné z z báze již dále není potřeba počítat hodnoty v odpovídajícím sloupci simplexové tabulky.

c_B	B	x_1	x_2	y	u_1	u_2	v	z	x_B	x_B/λ
0	u_1	−2	0	−1	1	0	0	0	2	✖
1	z	0	0	1	0	−1	0	1	1	✖
0	v	1	(1)	0	0	0	1	0	2	2 →
	c_k	0	0	0	0	0	0	1		
	z_k	0	0	1	0	−1	0	1		
	$z_k - c_k$	0	0 ↑	1	0	−1	0	0	1	
	u_1	−2	0	−1	1	0	0	0	2	✖
	z	0	0	(1)	0	−1	0	1	1	1 →
	x_2	1	1	0	0	0	1	0	2	✖
	$z_k - c_k$	0	0	1 ↑	0	−1	0	0	1	
	u_1	−2	0	0	1	−1	0	✖	3	
	y	0	0	1	0	−1	0	✖	1	
	x_2	1	1	0	0	0	1	✖	2	
	$z_k - c_k$	0	0	0	0	0	0	✖	0	

Poznamenejme ještě, že provést takový jednoduchý „trik“ u větších úloh už může být poměrně komplikované.

Příklad 4.5.3

Wolfeho metoda (krátký tvar) – řešený příklad #1:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/qp5

V případě $b \not\geq 0$ je trochu situace komplikovanější a je nutné nejdříve provést *první fázi* Wolfeho metody, abychom našli výchozí BPB pro druhou fázi[57]. Karushovy–Kuhnovy–Tuckerovy podmínky lze schematicky zapsat jako

[57] To je vlastně analogie dvoufázové metody pro lineární programování.

		$n+m$		$n+m$		
	„báze“	x	y	u	v	
$m+n$	u	$-C$	$-A^\top$	I	0	d
	v	A	0	0	I	b

Tabulka 4.3: Schematická tabulka pro omezení (4.5.2a)–(4.5.2b).

Bod $[x, y, u, v] = [0, 0, d, b]$ je bazickým, neboť má $m+n$ bazických proměnných a odpovídající část matice určující omezení v (4.5.2a)–(4.5.2b) je $\left(\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right)$, ovšem pro $b \not\geq 0$ není přípustný[58]. „Rozdělíme“ vektor b na[59] dvě části $\left(\begin{smallmatrix} b_1 \\ b_2 \end{smallmatrix}\right)$, kde $b_1 \geq 0$ a $b_2 < 0$. Analogicky rozdělíme také matici A, tj. máme

$$A_1 x \leq b_1, \quad b_1 \geq 0,$$
$$A_2 x \leq b_2, \quad b_2 < 0.$$

Stejným způsobem rozdělíme také vektory y, v a zavedeme umělé proměnné $w_1, w_2, \ldots$, jejichž počet je dán počtem záporných složek b. Takto dostaneme novou výchozí tabulku

[58] Kdyby tomu tak bylo, pak by $x^* = 0$ muselo být řešením.

[59] Toto provedeme bez újmy na obecnosti – stačí pouze přehodit pořadí jednotlivých omezení v $Ax \leq b$.

„báze“	x	y_1	y_2	u	v_1	v_2	w	
u	$-C$	$-A_1^\top$	$-A_2^\top$	I	0	0	0	d
v_1	A_1	0	0	0	I	0	0	$b_1 \geq 0$
w	$-A_2$	0	0	0	0	$-I$	I	$-b_2 > 0$

Tabulka 4.4: Výchozí tabulka pro první fázi Wolfeho metody.

BPB úlohy (4.5.1) pak dostaneme řešením pomocné úlohy

$$\left.\begin{aligned} w_1 + \cdots + w_r &\to \min \\ A_1 x + v_1 &= b_1 \\ -A_2 x - v_2 + w &= -b_2 \\ x \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad w \geq 0 & \end{aligned}\right\} \tag{4.5.12}$$

s výchozím BPB $[x, v_1, v_2, w] = [0, b_1, 0, -b_2]$. Ve výchozí tabulce ponecháme i první „řádek“ reprezentující omezení (4.5.2a), ale při určování klíčového řádku se omezujeme jen a pouze na proměnné vystupující v pomocné úloze (4.5.12), tj. x, v_1, v_2, w. Jinými slovy, v první fázi žádná z proměnných $y_1, \ldots, y_m$ nevstoupí do báze a žádné $u_1, \ldots, u_n$ nevystoupí z báze. Nicméně úpravy popsané v simplexovém algoritmu provádíme s celou tabulkou[60]. Cílem je vyloučit w z báze. Pokud se nám to nepodaří, znamená to, že původní úloha (4.5.1) je nepřípustná. Pokud se nám podaří vyloučit w z báze, pak pokračujeme druhou fází Wolfeho algoritmu tak, jak jsme ji popsali v předchozí části: zavedeme proměnné z_i podle prvního „bloku“ poslední simplexové tabulky v první fázi, tj. u_i nahradíme z_i a zavedeme matici D. Ovšem tentokrát již neplatí, že z_i zavádíme pouze pro zápornou pravou stranu – musíme totiž respektovat i Wolfeho pravidlo, tj. pokud je pravá strana kladná, ale odpovídající složka vektoru u nesmí být v bázi kvůli Wolfeho pravidlu (tj. v bázi už je x_i), musíme místo ní také zavést z_i.

Příklad 4.5.4

Wolfeho metoda (krátký tvar) – řešený příklad #2:
Video ☞ 😎 https://ln.math.muni.cz/pz/opt/qp6

4.6 Wolfeho metoda v dlouhém tvaru

V případě, kdy nejsou splněny předpoklady Věty 4.5.1, tj. $\det C = 0$ a $d \neq 0$, pak může nastat situace, že při řešení úlohy výše popsaným způsobem již nebude možné kvůli Wolfeho pravidlu provést další krok simplexového algoritmu, ačkoli stále $\sum_{i=1}^n z_i > 0$. Z tohoto důvodu se takové úlohy řeší pomocí tzv. *dlouhého tvaru Wolfeho metody*[61]. Tato verze Wolfeho metody se skládá ze tří fází, přičemž princip prvních dvou je shodný s těmi v krátké variantě[62].

Základní myšlenka je v parametrizaci lineární části účelové funkce $f(x) = \frac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle$ tak, že místo ní budeme pracovat s funkcí

$$\tfrac{1}{2}\langle Cx, x\rangle + \mu\langle d, x\rangle$$

pro $\mu \geq 0$. Pak díky Větě 4.5.1 můžeme najít řešení úlohy odpovídající $\mu = 0$ krátkým tvarem Wolfeho metody. Ovšem řešení původní úlohy odpovídá $\mu = 1$. Proto ve třetí fázi vyjdeme z řešení pro $\mu = 0$ a při stálém zachování Karushových–Kuhnových–Tuckerových podmínek začneme zvětšovat hodnotu μ, dokud nedosáhneme požadované hodnoty $\mu = 1$. Tímto způsobem lze dokonce řešit úlohu (4.5.1), kde účelová funkce $f(x) = \frac{1}{2}\langle Cx, x\rangle + \nu\langle d, x\rangle$ závisí na nějakém parametru ν nabývajícím různých hodnot $0 \leq \nu_1 < \nu_2 < \cdots < \nu_r$. My ale pro jednoduchost zůstaneme u případu[63] $\nu = 1$.

[60] ⚠ Pozor! Nelze se omezit pouze na druhé omezení a pracovat jenom s ním podobně jako v první fázi pro řešení úlohy lineárního programování. Potřebujeme najít hodnoty x, v tak, aby platilo druhé omezení a zároveň byli schopni najít výchozí BPB pro druhou fázi. Není totiž možné upravovat jenom druhé omezení a následně pokračovat druhou fází jako v předchozím případě, neboť ten byl založen na volbě výchozího BPB s $x = 0$, což asi nebude výsledkem první fáze. Nebo ano? Na začátku druhé fáze proto potřebujeme pracovat s novými omezeními, která získáme tím, že v první fázi upravujeme celou soustavu.

[61] Nebo lze zkusit krátký tvar a dostatečně usilovně se pomodlit, aby to tentokrát vyšlo. Nicméně je nutné se rozhodnout již na začátku, protože nelze mezi těmito dvěma tvary Wolfeho metody přecházet uprostřed výpočtů.

[62] Druhou možností je přičíst ke každému prvku na diagonále matice C malé $\varepsilon > 0$, tj. nahradit C za matici $C + \varepsilon I$, díky čemuž se matice stane pozitivně definitní. Pak na takovou úlohu lze aplikovat Wolfeho metodu v krátkém tvaru. Z výpočetních důvodů by ε nemělo být voleno příliš velké. Na druhou stranu by nemělo být ani příliš malé, aby nezpůsobovalo numerickou nestabilitu. Tato myšlenka pochází od Bealeho a nalezneme ji již v původním Wolfeho článku. V [474, str. IV.21] se o této těsné hranici mezi pozitivně definitními a semidefinitními maticemi hovoří jako o *Razor's Edge Property* („vlastnost na hraně břitvy“).

[63] Vše by fungovalo zcela analogicky s tím, že chceme získat $\mu = \nu_i$, tj. zvyšovali bychom hodnotu μ až k ν_i.

Uvažme tedy úlohu

$$\tfrac{1}{2}\langle Cx,x\rangle+\mu\langle d,x\rangle\to\min,\quad Ax\le b,\quad x\ge 0, \tag{4.6.1}$$

s odpovídajícími Karushovými–Kuhnovými–Tuckerovými podmínkami

$$-Cx-A^{\top}y+u-\mu d=0 \tag{4.6.2a}$$

$$Ax+v=b \tag{4.6.2b}$$

$$x\ge 0\quad\&\quad y\ge 0\quad\&\quad u\ge 0\quad\&\quad v\ge 0\quad\&\quad \mu\ge 0 \tag{4.6.2c}$$

$$\langle x,u\rangle=0\quad\&\quad\langle y,v\rangle=0. \tag{4.6.2d}$$

Pokud $b\not\ge 0$, je nutné začít s první fází a pokračovat do druhé fáze, po jejímž skončení chceme mít vyeliminovány všechny proměnné $z_1,\dots,z_n$ z báze, tj. v případě, že již máme řešení splňující $\sum_{i=1}^{n} z_i=0$, ale z důvodu degenerovanosti této pomocné úlohy zůstaly některé proměnné z_i v bázi, je nutné ve druhé fázi pokračovat tak dlouho, dokud všechna z_i nebudou vyloučena z báze. To může vést až k poněkud paradoxním situacím a značnému „znásilnění“ simplexového algoritmu, jak bychom si mohli ukázat v[64] Příkladě 4.7.25. Až se nám to povede, můžeme zahájit třetí fázi s výchozím BPB určeným poslední simplexovou tabulkou ze druhé fáze. Kdyby $b\ge 0$, pak lze ihned přistoupit ke třetí fázi s výchozím BPB $[x,y,u,v,\mu]=[0,0,0,b,0]$ a bází tvořenou proměnnými u,v. Na začátku třetí fáze začneme zvětšovat hodnotu[65] μ, přičemž musíme dbát na všechny podmínky (4.6.2), tj. ve třetí fázi řešíme úlohu

$$-\mu\to\min\quad\&\quad\text{(4.6.2a)–(4.6.2c)} \tag{4.6.3}$$

a současně dodržujeme Wolfeho pravidlo. Při tomto postupu mohou nastat 3 možnosti:

(i) Nalezneme BPB úlohy (4.6.3) s kladným číslem v řádku z_k-c_k (tj. není řešením), přičemž příslušný sloupec lze v souladu s Wolfeho pravidlem zavést do báze, ovšem tento sloupec s výjimkou řádku z_k-c_k neobsahuje žádné další kladné číslo, tj. nelze provést další krok simplexového algoritmu. V takovém případě není účelová funkce úlohy (4.6.3) ohraničená zdola, tj. $\mu\to\infty$.

(ii) Nalezneme řešení úlohy (4.6.3), tj. řádek z_k-c_k neobsahuje kladné číslo.

(iii) Nalezneme BPB úlohy (4.6.3) s kladným číslem v řádku z_k-c_k (tj. není řešením), ale kvůli Wolfeho pravidlu nelze pokračovat dále[66], viz Příklad 4.6.3.

Jaký mají tyto situace vliv na řešení?

Věta 4.6.1

(a) Jestliže při řešení úlohy (4.6.3) simplexovým algoritmem doplněným o Wolfeho pravidlo nastanou situace (ii) nebo (iii) popsané výše, potom hodnota proměnné μ v poslední simplexové tabulce je nutně nulová.

(b) Jestliže při řešení úlohy (4.6.3) simplexovým algoritmem doplněným o Wolfeho pravidlo nastane situace (ii), potom hodnota účelové funkce úlohy

$$Q(x,\mu):=\tfrac{1}{2}\langle Cx,x\rangle+\mu\langle d,x\rangle\to\min,\quad Ax\le b,\quad x\ge 0$$

není zdola ohraničená pro žádné $\mu>0$, tj. $Q(x,\mu)\to-\infty$.

Důkaz. Princip důkazu první části je velmi analogický tomu pro Větu 4.5.1. Tentokrát však v (4.5.6) budeme místo $z=(z_1,\dots,z_n)^{\top}$ mít pouze $-\mu$ a[67] dále položíme $d=0$ a $D=d$. Ať již tentokrát nastala situace (ii) nebo (iii), tj. buď jsme našli řešení nebo je další výpočet zablokován kvůli Wolfeho pravidlu, levá strana rovnosti (4.5.11) musí být vždy nulová kvůli volbě $d=0$, a tudíž i pravá strana $-\mu$ musí být nulová.

Bohužel, navzdory autorovu nezměrnému úsilí upravit původní Wolfeho důkaz části (b) do podoby odpovídající úloze (4.6.1) s nerovnostními omezeními se to nepovedlo[68,69]. Důkaz je sice založen na podobných úvahách jako v případě Věty 4.5.1, nicméně tentokrát se zdá být potřebné uvážit úlohu kvadratického programování v kanonickém tvaru, tj.

$$\tfrac{1}{2}\langle Cx,x\rangle+\mu\langle d,x\rangle\to\min,\quad Ax=b,\quad x\ge 0. \tag{4.6.4}$$

[64] Doporučujeme čtenáři si tento příklad spočítat. Ve druhé fázi díky degenerovanosti příslušné úlohy budeme mít hned v prvním kroku možnost volby klíčového řádku. Jedna volba povede k poměrně rychlému ukončení druhé fáze, zatímco v případně opačné volby budeme mít dvě varianty: buď celý výpočet vzdát, protože se nám nepodařilo vyloučit umělou proměnnou z báze a zároveň simplexový algoritmus neumožňuje pokračovat, nebo si pravidla trochu „přizpůsobit“ a kvůli vyeliminování umělé proměnné v jednom kroku zavést do báze proměnnou se zápornou hodnotou v řádku z_k-c_k, tj. sice změnit bázi, ale současně zvýšit funkční hodnotu... Zkuste si to!

[65] Kéž by... Samozřejmě v případě degenerované úlohy může dojít k tomu, že proměnná μ zůstane nulová, i když vstoupí do báze.

[66] Toto je komplikace, kterou překvapivě Wolfe ve svém původním článku vůbec nezmiňuje.

[67] Toto může působit trochu nekonzistentním dojmem, neboť stále požadujeme $\mu\ge 0$. Pokud ale účelovou funkci v (4.5.6) vezmeme jako součin ez, pak nahrazujeme μ za z a -1 za e, odkud už lépe poznáme, jak se změní v tomto případě duální úloha.

[68] Snad nám to laskavý čtenář promine...

[69] Není příliš mnoho výsledků, které striktně závisí na zvoleném tvaru daných omezení. Např. pro Větu 2.4.6 o řešitelnosti úloh lineárního programování byl výhodný

základní tvar, neboť umožňoval použití Farkasovy–Minkowského věty. Ale asi by se s trochou snahy dal odvodit tentýž výsledek i při použití kanonického tvaru. Tentokrát je však situace mnohem komplikovanější. Autorovi se podařilo najít jen 3 (slovy: tři!) zdroje s důkazem předloženého tvrzení pro dlouhý tvar Wolfeho metody. Všechny ale byly založeny na omezeních v kanonickém tvaru. Na druhou stranu dlouhý tvar Wolfeho metody pro úlohu (4.6.1) se podařilo najít jen ve dvou (!) zdrojích, avšak bez důkazu. Nicméně autor je i nadále přesvědčen, že by mělo být možné Wolfeho důkaz upravit tak, aby jej bylo možné použít i pro úlohu (4.6.1). Řečeno ale slovy anglického matematika Rogera Cotese: *toto už ponechávám k podrobnějšímu prozkoumání těm, kterým to za to stojí.*

Potom i s přihlédnutím k (4.3.6) zjistíme, že ve třetí fázi se řeší následující úloha (v první podmínce jsme u všech členů změnili znaménko)

$$\begin{gathered} -\mu \to \min \\ Cx + A^\top y - u + \mu d = 0 \\ Ax = b \\ x \geq 0 \quad \& \quad u \geq 0 \quad \& \quad \mu \geq 0 \\ \langle x, u \rangle = 0. \end{gathered}$$

Připusťme, že jsme nalezli její řešení o souřadnicích $X^* := [x^*, y^*, u^*, \mu^*]$. Díky předpokladu nedegenerovanosti dané úlohy má každý BPB právě $m+n$ kladných složek. Jelikož z předchozí části plyne, že nutně $\mu^* = 0$, musí být v bázi všech m složek vektoru y a složky $x^{[1]}$ a $u^{[2]}$, které jsou vytvořené podobně jako v důkazu Věty 4.5.1 (množina N_3 je prázdná a množiny M_1, M_2, M_3 v této úpravě nepotřebujeme). Potom podobně jako v případě krátkého tvaru Wolfeho metody získáme odpovídající duální úlohu

$$\left.\begin{gathered} \beta^\top b \to \max \\ \delta_1^\top C_{11} + \delta_2^\top C_{21} + \beta_1^\top A_{11} + \beta_2^\top A_{21} = 0 \\ \delta^\top A^\top = 0 \\ -\delta^{[2]} = 0 \\ \delta^\top d \leq -1 \end{gathered}\right\} \tag{4.6.5}$$

přičemž její řešení $[\delta^*, \beta^*]$ splňuje ${\beta^*}^\top b = 0$ a $C\delta^* = 0$. Navíc musí platit $\delta^{[1]} \geq 0$, neboť v opačném případě by pro $X^*(\varepsilon) := X^* + \varepsilon\,[\delta^*, 0, 0, 0]$ existovalo $\varepsilon^* > 0$ takové, že $x_i^{[1]*} + \varepsilon \delta_i^{[1]*} = 0$ pro nějaký index i, tj. BPB $X^*(\varepsilon^*)$ by měl pouze $m+n-1$ nezáporných složek, což je spor s požadavkem nedegenerovanosti. Proto $x^*(\varepsilon) := x^* + \varepsilon\delta^*$ je pro libovolné $\varepsilon \geq 0$ přípustným bodem úlohy (4.6.4) s hodnotou účelové funkce

$${x^*}^\top(\varepsilon)\, C\, x^*(\varepsilon) + \mu d^\top x^*(\varepsilon) = {x^*}^\top C x^* + \mu d^\top x^* + \mu\varepsilon d^\top \delta^*,$$

což vzhledem k poslední nerovnosti v (4.6.5) ukazuje její neohraničenost zdola pro libovolné $\mu > 0$ a $\varepsilon \to \infty$, neboť

$${x^*}^\top(\varepsilon)\, C\, x^*(\varepsilon) + \mu d^\top x^*(\varepsilon) \leq {x^*}^\top C x^* + \mu d^\top x^* - \mu\varepsilon.$$

Tvrzení v části (b) proto plyne z ekvivalence mezi úlohami (4.6.4) a (4.6.1), o které jsme hovořili v Podkapitole 4.3. ■

Z uvedené věty vyplývá, že původní úloha (4.5.1) může mít řešení pouze, nastane-li situace (i) nebo (iii), zatímco v situaci (ii) je původní úloha neohraničená. Z části (a) také vidíme, že pokud při řešení úlohy (4.6.3) je v některém kroku $\mu > 0$, pak nutně nastane situace (i), tj. μ vstoupí do báze a nabude nenulovou hodnotou, kterou již není možné anulovat. Avšak situace (iii) zůstává nejasná a obecně v takové chvíli nemůžeme rozhodnout, zda úloha (4.6.3) má řešení či nikoli[70]. Tato situace určitě nemůže nastat, je-li $C > 0$, neboť v takovém případě by v důkazu druhé části Věty 4.6.1 muselo být $\delta^* = 0$, což je ve sporu s poslední podmínkou v (4.6.5). Nicméně my jsme se k dlouhému tvaru Wolfeho metody dopracovali kvůli možnému „zaseknutí“ krátkého tvaru této metody v případě singulární matice C a tyto potíže přetrvávají i u dlouhého tvaru?! To není příliš pozitivní zjištění obzvláště, když k tomu může dojít už i u úloh mající jen dvě proměnné, jak je popsáno v řešení Příkladu 4.6.3.

Naštěstí tomu lze zabránit použitím tzv. *Shettyho pravidla*[71], které lze s trochou úsilí odvodit z [445] a které vylučuje situaci (iii). Ve výchozí tabulce pro třetí fázi je z každé dvojice x_i, u_i a y_i, v_i právě jedna proměnná v bázi a druhá mimo bázi[72]. Proto v prvním kroku této fáze

[70] Má-li řešení, tak ho ale v tuto chvíli kvůli $\mu = 0$ nemůžeme určit (potřebujeme se dopracovat k $\mu = 1$ nebo něčemu takovému...

[71] Není to pravidlo v tom klasickém smyslu – simplexový algoritmus doplněný o Wolfeho pravidlo nám totiž nic moc jiného nedovolí, než podle něj postupovat.

[72] Vyloučili jsme všechna z_i z báze a kvůli Wolfeho pravidlu tam zmíněné dvojice proměnných nemohou být současně. Jenže v bázi musí být právě $n+m$ prvků, takže tam z každé dvojice musí být právě jeden.

máme jedinou možnost: zavést do báze proměnnou μ. Označíme-li si dvojice x_i, u_i a y_i, v_i pomocí

$$\widetilde{x}_i := \begin{cases} x_i, & i = 1,\dots,n, \\ v_{i-n}, & i = n+1,\dots,n+m, \end{cases} \qquad \widetilde{u}_i := \begin{cases} u_i, & i = 1,\dots,n, \\ y_{i-n}, & i = n+1,\dots,n+m, \end{cases}$$

pak na začátku třetí fáze je v bázi právě jedna proměnná z dvojice $\widetilde{x}_i, \widetilde{u}_i$ a v prvním kroku některá z nich bázi opustí. Ovšem kvůli Wolfeho pravidlu existují pouze dvě proměnné, které mohou v dalším kroku vstoupit do báze: buď bude vrácena zpět proměnná vystoupivší z báze, nebo tam vstoupí druhá proměnná z dvojice $\widetilde{x}_i, \widetilde{u}_i$, což platí i pro všechny další kroky třetí fáze. Tyto dvě proměnné tvoří tzv. *nebazickou dvojici*. Označíme-li odpovídající hodnotu v řádku $z_k - c_k$ nové simplexové tabulky pro tyto proměnné jako Δ a $\widehat{\Delta}$, pak pro proměnnou vystupující z báze je nutně[73] $\Delta \le 0$, zatímco pro druhou proměnnou z této nebazické dvojice lze ukázat, že $\widehat{\Delta} \ge 0$. To znamená, že v každém kroku třetí fáze lze vždy určit proměnnou, která má vstoupit do báze, a tudíž celý tento proces musí skončit buď situací (i) nebo (ii). Toto chování můžeme označit jako ono avizované *Shettyho pravidlo*:

> Opustilo-li bázi $\widetilde{x}_i$ (nebo $\widetilde{u}_i$), pak v dalším kroku do báze musí vstoupit $\widetilde{u}_i$ (nebo $\widetilde{x}_i$) bez ohledu na to, zda odpovídající hodnota v řádku $z_k - c_k$ je kladná či nulová.

[73] To plyne ze vzorce pro výpočet $\widehat{z}_k - c_k$ v důkazu Věty 2.8.2. V tomto případě totiž máme $z_k - c_k = 0$ a $\lambda_{rk} = 1$, protože proměnná byla v bázi, zatímco $\lambda_{rs} > 0$ dle kritéria pro výběr klíčového řádku a $z_s - c_s \ge 0$ dle kritéria pro výběr klíčového sloupce, což dohromady dává $\Delta = \widehat{z}_k - c_k \le 0$.

Nyní se podíváme na situaci (i), kdy $\mu \to \infty$, tj. když účelová funkce úlohy (4.6.3) není zdola ohraničená, a tedy úloha (4.5.1) má řešení. V takovém případě dostaneme při řešení (4.6.3) upraveným simplexovým algoritmem konečnou posloupnost BPB soustavy (4.6.2a)–(4.6.2c), tj. body

$$X^{[k]} := \left[x^{[k]}, y^{[k]}, u^{[k]}, v^{[k]}, \mu^{[k]}\right], \quad k = 0, 1, \dots, r$$

a nakonec polopřímku

$$X^{[r]} + \tau\, \overline{X}, \quad \tau \ge 0, \tag{4.6.6}$$

na které nabývá μ libovolně velké hodnoty, tj. $\mu \to \infty$. Směr $\overline{X} = \left(\overline{x}, \overline{y}, \overline{u}, \overline{v}, \overline{\mu}\right)^\top$ této polopřímky je určený tím sloupcem poslední simplexové tabulky, který má v řádku $z_k - c_k$ kladné číslo, tzv. *kritický sloupec*. Takový sloupec je nutně jediný v důsledku Shettyho pravidla a ostatní prvky tohoto sloupce jsou nekladné, protože jsme v situaci (i). Tento směr proto dostaneme tak, že proměnné odpovídající tomuto sloupci (tzv. *kritická proměnná*, kterou bychom v posledním kroku chtěli zavést do báze) přiřadíme hodnotu 1, ostatním nebazickým proměnným kritického sloupce přiřadíme hodnotu 0 a bazickým proměnným přiřadíme hodnoty odpovídající číslům z kritického sloupce vynásobeného -1, tj. $\overline{X}$ obsahuje pouze nezáporná čísla[74] .

[74] Přemýšlíte proč právě tato volba? Odpověď je ukryta v důkazu Věty 2.7.1 s volbou $\alpha = 1$.

Wolfeho pravidlo nám zaručuje nejenom nulovost skalárních součinů $\langle x^{[k]}, u^{[k]} \rangle$ a $\langle y^{[k]}, v^{[k]} \rangle$ v každém kroku, ale také součinů

$$\begin{aligned} \langle x^{[k]}, u^{[k+1]} \rangle = 0 = \langle x^{[k+1]}, u^{[k]} \rangle, \quad k = 0, 1, \dots, r-1, \\ \langle y^{[k]}, v^{[k+1]} \rangle = 0 = \langle y^{[k+1]}, v^{[k]} \rangle, \quad k = 0, 1, \dots, r-1, \\ \langle x^{[r]}, \overline{u} \rangle = 0 = \langle \overline{x}, u^{[r]} \rangle, \\ \langle y^{[r]}, \overline{v} \rangle = 0 = \langle \overline{y}, v^{[r]} \rangle, \end{aligned}$$

tedy podmínky ortogonálnosti jsou splněny i pro libovolnou konvexní kombinaci dvou po sobě jdoucích BPB a všechny body polopřímky (4.6.6), o čemž se lze snadno přesvědčit vzhledem k linearitě skalárního součinu, tj. např.

$$\begin{aligned} &\langle \alpha\, x^{[k]} + \beta\, x^{[k+1]}, \alpha\, u^{[k]} + \beta\, u^{[k+1]} \rangle = \\ &\quad = \alpha^2 \langle x^{[k]}, u^{[k]} \rangle + \beta^2 \langle x^{[k+1]}, u^{[k+1]} \rangle + \alpha\,\beta \left(\langle x^{[k]}, u^{[k+1]} \rangle + \langle x^{[k+1]}, u^{[k]} \rangle \right) = 0, \end{aligned}$$

kde $\alpha+\beta=1$, $\alpha\ge 0$, $\beta\ge 0$ a $k=0,1,\dots,r-1$. Stejný výpočet můžeme zopakovat i pro dvojici y, v a konečně

$$\langle x^{[r]}+\tau\overline{x}, u^{[r]}+\tau\overline{u}\rangle = \langle x^{[r]}, x^{[r]}\rangle + \tau^2\langle\overline{x},\overline{u}\rangle + \tau\big(\langle x^{[r]},\overline{u}\rangle + \langle\overline{x}, u^{[r]}\rangle\big) = 0.$$

Proto řešení úlohy (4.5.1) nalezneme jako konvexní kombinaci (s výjimkou prvního velmi speciálního případu) dvou posledních BPB nebo BPB a polopřímky tak, abychom docílili kýžené hodnoty $\mu=1$. To znamená, že musíme rozlišit 3 možné případy:

(i) Je-li dokonce $\mu^{[k]}=1$ pro nějaké $k\in\{0,\dots,r\}$, pak odpovídající $X^{[k]}$ je řešením soustavy (4.6.2) s $\mu=1$, a tedy složka $x^{[k]}$ je řešením úlohy (4.5.1), tj. $x^*=x^{[k]}$.

(ii) Je-li $\mu^{[k]}<1<\mu^{[k+1]}$ pro nějaké $k\in\{0,\dots,r-1\}$, pak konvexní kombinace

$$X^* := \lambda^* X^{[k]} + (1-\lambda^*)X^{[k+1]}, \quad \text{kde } 0<\lambda^* := \frac{\mu^{[k+1]}-1}{\mu^{[k+1]}-\mu^{[k]}} < 1,$$

tvoří řešení[75] soustavy (4.6.2) s $\mu=1$, takže $x^*=\lambda^* x^{[k]}+(1-\lambda^*)x^{[k+1]}$ řeší úlohu (4.5.1).

(iii) Je-li i $\mu^{[r]}<1$, pak

$$X^* := X^{[r]} + \tau^* \overline{X}, \quad \text{kde } \tau^* := \frac{1-\mu^{[r]}}{\overline{\mu}},$$

je řešením[76] soustavy (4.6.2) s $\mu=1$, a tedy $x^*=x^{[r]}+\tau^*\overline{x}$ je řešením úlohy (4.5.1).

[75] Proč? Protože právě touto konvexní kombinací dostaneme

$$\mu^* = \lambda^*\mu^{[k]} + (1-\lambda^*)\mu^{[k+1]} = \\ = \frac{\mu^{[k]}\big(\mu^{[k+1]}-1\big)}{\mu^{[k+1]}-\mu^{[k]}} + \\ + \frac{\mu^{[k+1]}\big(1-\mu^{[k]}\big)}{\mu^{[k+1]}-\mu^{[k]}} = 1$$

[76] Tentokrát očividně máme

$$\mu^* = \mu^{[r]} + \tau^*\overline{\mu} = \\ = \mu^{[r]} + \big(1-\mu^{[r]}\big) = 1.$$

Příklad 4.6.2

Vyřešme úlohu

$$x_1^2 - x_1 - x_2 \to \min$$
$$x_1 - 2x_2 \le 1 \quad \& \quad 2x_1 - x_2 \le 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. V této úloze máme

$$C=\begin{pmatrix}2&0\\0&0\end{pmatrix},\quad d=\begin{pmatrix}-1\\-1\end{pmatrix},\quad A=\begin{pmatrix}1&-2\\2&-1\end{pmatrix},\quad b=\begin{pmatrix}1\\2\end{pmatrix}.$$

Matice C je pouze pozitivně semidefinitní a současně $b\ge 0$, takže musíme využít dlouhý tvar Wolfeho metody, ale můžeme začít rovnou její třetí fází. V ní řešíme úlohu

$$-\mu \to \min$$
$$-2x_1 - y_1 - 2y_2 + u_1 + \mu = 0$$
$$2y_1 + y_2 + u_2 + \mu = 0$$
$$x_1 - 2x_2 + v_1 = 1$$
$$2x_1 - x_2 + v_2 = 2$$
$$x_1\ge 0 \quad \& \quad x_2\ge 0 \quad \& \quad y_1\ge 0 \quad \& \quad y_2\ge 0$$
$$u_1\ge 0 \quad \& \quad u_2\ge 0 \quad \& \quad v_1\ge 0 \quad \& \quad v_2\ge 0 \quad \& \quad \mu\ge 0.$$

$\mu^{[0]}=0$

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	x_B	x_B/λ
0	u_1	−2	0	−1	−2	1	0	0	0	(1)	0	0 →
0	u_2	0	0	2	1	0	1	0	0	1	0	0
0	v_1	1	−2	0	0	0	0	1	0	0	1	✖
0	v_2	2	−1	0	0	0	0	0	1	0	2	✖
	c_k	0	0	0	0	0	0	0	0	−1		
	z_k	0	0	0	0	0	0	0	0	0		
	z_k-c_k	0	0	0	0	0	0	0	0	1 ↑	0	

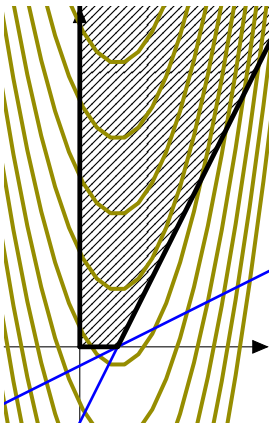
$\mu^{[1]} = 0$

Shetty:
u_1
$\updownarrow$
B
$\updownarrow$
x_1

$\mu^{[2]} = 0$

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	x_B	x_B/λ
	μ	-2	0	-1	-2	1	0	0	0	1	0	✖
	u_2	(2)	0	3	3	-1	1	0	0	0	0	0 →
	v_1	1	-2	0	0	0	0	1	0	0	1	1
	v_2	2	-1	0	0	0	0	0	1	0	2	1
	$z_k - c_k$	2 ↑	0	1	2	-1	0	0	0	0	0	
	μ	0	0	2	1	0	1	0	0	1	0	
	x_1	1	0	3/2	3/2	$-1/2$	1/2	0	0	0	0	
	v_1	0	0	$-3/2$	$-3/2$	1/2	$-1/2$	1	0	0	1	
	v_2	0	-1	-3	-3	1	-1	0	1	0	2	
	$z_k - c_k$	0	0	-2	-1	0	-1	0	0	0	0	

Řádek $z_k - c_k$ již neobsahuje kladné číslo, takže simplexový algoritmus je u konce a nám se podařilo najít řešení úlohy ze třetí fáze. To znamená, že původní úloha je neohraničená, o čemž se lze snadno přesvědčit z Obrázku 4.10 (vrstevnice jsou paraboly a přípustná množina je neohraničená).



Obrázek 4.10: Grafické znázornění úlohy z Příkladu 4.6.2.

V Příkladu 4.5.2 jsme zkoušeli vyřešit pomocí Wolfeho metody v krátkém tvaru úlohu se singulární maticí C. S využitím drobného triku se nám to nakonec podařilo. Nyní se podíváme, jak v takovém případě bude fungovat dlouhý tvar Wolfeho metody. Tento příklad ilustruje „jak je důležité míti Shettyho“.

Příklad 4.6.3

Vyřešme úlohu

$$x_1^2 + 2x_1 - x_2 \to \min$$
$$x_1 + x_2 \le 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. V této úloze máme

$$C = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad d = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad b = 2.$$

Matice C je pouze pozitivně semidefinitní a současně $b \ge 0$, takže musíme využít dlouhý tvar Wolfeho metody, ale můžeme začít rovnou její třetí fází. V ní řešíme úlohu

$$-\mu \to \min$$
$$-2x_1 - y + u_1 - 2\mu = 0$$
$$-y + u_2 + \mu = 0$$
$$x_1 + x_2 + v = 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad y \ge 0 \quad \& \quad u_1 \ge 0 \quad \& \quad u_2 \ge 0 \quad \& \quad v \ge 0 \quad \& \quad \mu \ge 0.$$

Ve druhém kroku se dostaneme do situace, která by se mohla tvářit jako zablokovaná kvůli Wolfeho pravidlu. V řádku $z_k - c_k$ je kladné číslo pouze ve sloupci pro proměnnou y, kterou ovšem kvůli Wolfeho pravidlu nesmíme zavést do báze. Co s tím? Nastala situace (iii) popsaná před Větou 4.6.1? Naštěstí máme Shettyho pravidlo, které nám v tomto případě říká, že pokud jsme z báze vyloučili u_2, musí nyní do báze vstoupit x_2. A to skutečně i uděláme navzdory tomu, že v řádku $z_k - c_k$ je pro tuto proměnnou nulová hodnota.

	c_B	B	x_1	x_2	y	u_1	u_2	v	μ	x_B	x_B/λ
	0	u_1	−2	0	−1	1	0	0	−2	0	✖
$\boxed{\mu^{[0]}=0}$	0	u_2	0	0	−1	0	1	0	(1)	0	0 →
	0	v	1	1	0	0	0	1	0	2	✖
		c_k	0	0	0	0	0	0	−1		
		z_k	0	0	0	0	0	0	0		
		$z_k - c_k$	0	0	0	0	0	0	1 ↑	0	
$\boxed{\mu^{[1]}=0}$		u_1	−2	0	−3	1	2	0	0	0	✖
		μ	0	0	−1	0	1	0	1	0	✖
Shetty:		v	1	(1)	0	0	0	1	0	2	2 →
u_2 ⇕ B ⇕ x_2		$z_k - c_k$	0	0 ↑	1	0	−1	0	0	0	
		u_1	−2	0	−3	0	2	0	0	0	✖
		μ	0	0	−1	0	1	0	1	0	✖
		x_2	1	1	0	0	0	1	0	2	✖
$\boxed{\mu^{[2]}=0}$		$z_k - c_k$	0	0	1	0	−1	0	0	0	

KONEC třetí fáze. Do báze bychom měli nyní zavést proměnnou y, avšak tento sloupec má kladné číslo pouze v řádku $z_k - c_k$, tj. neumíme určit klíčový řádek. Úloha ve třetí fázi je tudíž neohraničená. Protože

$$x^{[2]} = [0,2], \quad \overline{x} = [0,0], \quad \mu^{[0]} = 0, \quad \overline{\mu} = 1, \quad \tau^* = \frac{1-\mu^{[2]}}{\overline{\mu}} = 1,$$

je řešením bod

$$x^* = x^{[2]} + \tau^* \overline{x} = [0,2].$$

Příklad 4.6.4

Vyřešme úlohu

$$x_1^2 - 6x_1x_2 + 9x_2^2 - x_1 + 3x_2 - 3 \to \min$$
$$x_1 + 2x_2 \le 3 \quad \& \quad x_1 - x_2 \le 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. V této úloze máme

$$C = \begin{pmatrix} 2 & -6 \\ -6 & 18 \end{pmatrix}, \quad d = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Matice C je pouze pozitivně semidefinitní a současně $b \ge 0$, takže musíme využít dlouhý tvar Wolfeho metody, ale můžeme začít rovnou její třetí fází. V ní řešíme úlohu

$$-\mu \to \min$$
$$-2x_1 + 6x_2 - y_1 - y_2 + u_1 + \mu = 0$$
$$6x_1 - 18x_2 - 2y_1 + y_2 + u_2 - 3\mu = 0$$
$$x_1 + 2x_2 + v_1 = 3$$

$$x_1 - x_2 + v_2 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad y_1 \geq 0 \quad \& \quad y_2 \geq 0$$
$$u_1 \geq 0 \quad \& \quad u_2 \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad \mu \geq 0.$$

$\mu^{[0]} = 0$

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	x_B	x_B/λ
0	u_1	−2	6	−1	−1	1	0	0	0	(1)	0	0 →
0	u_2	6	−18	−2	1	0	1	0	0	−3	0	✖
0	v_1	1	2	0	0	0	0	1	0	0	3	✖
0	v_2	1	−1	0	0	0	0	0	1	0	2	✖
	c_k	0	0	0	0	0	0	0	0	−1		
	z_k	0	0	0	0	0	0	0	0	0		
	$z_k - c_k$	0	0	0	0	0	0	0	0	1 ↑	0	

$\mu^{[1]} = 0$

	μ	−2	6	−1	−1	1	0	0	0	1	0	✖
	u_2	0	0	−5	−2	3	1	0	0	0	0	✖
	v_1	1	2	0	0	0	0	1	0	0	3	3
	v_2	(1)	−1	0	0	0	0	0	1	0	2	2 →
	$z_k - c_k$	2 ↑	−6	1	1	−1	0	0	0	0	0	

Shetty:
u_1
⇕
B
⇕
x_1

$\mu^{[2]} = 4$

	μ	0	4	−1	−1	1	0	0	2	1	4	
	u_2	0	0	−5	−2	3	1	0	0	0	0	
	v_1	0	3	0	0	0	0	1	−1	0	1	
	x_1	1	−1	0	0	0	0	0	1	0	2	
	$z_k - c_k$	0	−4	1	1	−1	0	0	−2	0	−4	

KONEC třetí fáze, neboť $\mu^{[1]} < 1 < \mu^{[2]}$. Jelikož platí

$$x^{[1]} = [0,0], \quad x^{[2]} = [2,0] \quad \text{a} \quad \lambda^* = \frac{\mu^{[2]} - 1}{\mu^{[2]} - \mu^{[1]}} = \frac{4-1}{4-0} = \frac{3}{4},$$

je hledaným optimálním řešením dané úlohy

$$x^* = \lambda^* x^{[1]} + (1 - \lambda^*)\, x^{[2]} = [1/2, 0].$$

Avšak tutéž úlohu jsme již řešili pomocí obrázku v Příkladě 4.2.5 a výsledkem byla celá úsečka $x_1 - 3x_2 - 1/2 = 0$ pro $x_1 \in [1/2, 1]$. Zdá se ale, že k témuž závěru se s Wolfeho metodou nedopracujeme.

Příklad 4.6.5

Vyřešme úlohu

$$x_1^2 + x_2^2 + 2x_1x_2 + x_1 - 2x_2 \to \min$$
$$x_1 + x_2 \leq 5 \quad \& \quad 2x_1 + x_2 \geq 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení. Přepíšeme-li druhé omezení do tvaru $-2x_1 - x_2 \leq -4$, pak máme

$$C = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad d = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ -4 \end{pmatrix}.$$

Jelikož $b \not\geq 0$ budeme potřebovat první fázi Wolfeho metody. Současně máme $C \geq 0$ a $\det C = 0$, takže je potřeba použít dlouhý tvar Wolfeho metody.

Fáze 1

Máme pouze $b_2 < 0$, takže v této fázi přidáme pouze proměnnou w_2, tj. $w = (0, w_2)^\top$, a řešíme úlohu

$$w_2 \to \min$$
$$x_1 + x_2 + v_1 = 5$$
$$2x_1 + x_2 - v_2 + w_2 = 4$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad w_2 \geq 0.$$

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	w_2	x_B	x_B/λ
	u_1	-2	-2	-1	2	1	0	0	0	-1	0	0	
	u_2	-2	-2	-1	1	0	1	0	0	2	0	0	
0	v_1	1	1	0	0	0	0	1	0	0	0	5	5
1	w_2	$\textcircled{2}$	1	0	0	0	0	0	-1	0	1	4	$2 \to$
	c_k	0	0	0	0	0	0	0	0	0	1		
	z_k	2	1	0	0	0	0	0	-1	0	1		
	$z_k - c_k$	$2\uparrow$	1	0	0	0	0	0	-1	0	0	4	
	u_1	0	-1	-1	2	1	0	0	-1	-1	✖	4	
	u_2	0	-1	-1	1	0	1	0	-1	2	✖	4	
	v_1	0	$1/2$	0	0	0	0	1	$1/2$	0	✖	3	
	x_1	1	$1/2$	0	0	0	0	0	$-1/2$	0	✖	2	
	$z_k - c_k$	0	0	0	0	0	0	0	0	0	✖	0	

KONEC první fáze. Nalezli jsme řešení s $w_2 = 0$, takže úloha je přípustná a můžeme sestrojit výchozí BPB pro druhou fázi.

Fáze 2

Nyní zavedeme proměnné z_1, z_2 a zvolíme $D = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Pak chceme najít řešení úlohy s omezeními z poslední simplexové tabulky z první fáze, tj.

$$\left.\begin{aligned} z_1 + z_2 &\to \min \\ -x_2 - y_1 + 2y_2 + u_1 - v_2 - \mu + z_1 &= 4 \\ -x_2 - y_1 + y_2 + u_2 - v_2 + 2\mu + z_2 &= 4 \\ x_2/2 + v_1 + v_2/2 &= 3 \\ x_1 + x_2/2 - v_2/2 &= 2 \\ x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad y_1 \geq 0 \quad \& \quad y_2 \geq 0 \quad \& \quad u_1 \geq 0 & \\ u_2 \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad \mu \geq 0 \quad \& \quad z_1 \geq 0 \quad \& \quad z_2 \geq 0. & \end{aligned}\right\} \tag{4.6.7}$$

Za výchozí BPB pak vezmeme $x_1 = 2$, $x_2 = y_1 = y_2 = u_1 = u_2 = 0$, $v_1 = 3$, $v_2 = 0$, $\mu = 0$, $z_1 = 4$, $z_2 = 4$, tj. v bázi je x_1, v_1, z_1, z_2. V takovém případě máme následující výchozí simplexovou tabulku.

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	z_1	z_2	x_B	x_B/λ
1	z_1	0	−1	−1	2	1	0	0	−1	−1	1	0	4	
1	z_2	0	−1	−1	1	0	1	0	−1	2	0	1	4	
0	v_1	0	1/2	0	0	0	0	1	1/2	0	0	0	3	
0	x_1	1	1/2	0	0	0	0	0	−1/2	0	0	0	2	
	c_k	0	0	0	0	0	0	0	0	0	1	1		
	z_k	0	−2	−2	3	1	1	0	−2	1	1	1		
	$z_k - c_k$	0	−2	−2	3	1	1	0	−2	1	0	0	8	

Cílem této fáze je vyeliminovat proměnné z_1 a z_2 z báze, což bude vyžadovat nejméně 2 kroky. Ovšem v souladu s úvodní poznámkou to můžeme udělat s pomocí menšího počtu proměnných. V poslední simplexové tabulce v první fázi máme na pravé straně (ve sloupci označeném x_B) kladná čísla, takže za výchozí BPB pro úlohu ve druhé fázi bychom mohli vzít $x_1 = 2$, $x_2 = y_1 = y_2 = 0$ $u_1 = 4$, $u_2 = 4$, $v_1 = 3$, $v_2 = 0$, $\mu = 0$, $z_1 = 0$, $z_2 = 0$, tj. nepotřebujeme z_1, z_2. Mohli bychom tedy tuto fázi přeskočit a pokračovat rovnou třetí fází. Jenže tento BPB je v rozporu s Wolfeho pravidlem – jelikož x_1 nám vyšlo na konci první fáze v bázi, nesmíme u_1 zavést do báze. Proto místo něj zavedeme proměnnou z_1. Současně u_2 v bázi nechat můžeme, takže nepotřebujeme z_2. Proto stačí ve druhé fázi řešit úlohu

$$z_1 \to \min,$$
$$-x_2 - y_1 + 2y_2 + u_1 - v_2 - \mu + z_1 = 4,$$
$$-x_2 - y_1 + y_2 + u_2 - v_2 + 2\mu = 4,$$
$$x_2/2 + v_1 + v_2/2 = 3,$$
$$x_1 + x_2/2 - v_2/2 = 2,$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad y_1 \geq 0 \quad \& \quad y_2 \geq 0 \quad \& \quad u_1 \geq 0$$
$$u_2 \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad \mu \geq 0 \quad \& \quad z_1 \geq 0,$$

s výchozím BPB ve tvaru $x_1 = 2$, $x_2 = y_1 = y_2 = u_1 = 0$ $u_2 = 4$, $v_1 = 3$, $v_2 = 0$, $\mu = 0$, $z_1 = 4$, tj. v bázi je x_1, u_1, v_1, z_2.

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	z_1	x_B	x_B/λ
1	z_1	0	−1	−1	(2)	1	0	0	−1	−1	1	4	2 →
0	u_2	0	−1	−1	1	0	1	0	−1	2	0	4	4
0	v_1	0	1/2	0	0	0	0	1	1/2	0	0	3	✖
0	x_1	1	1/2	0	0	0	0	0	−1/2	0	0	2	✖
	c_k	0	0	0	0	0	0	0	0	0	1		
	z_k	0	−1	−1	2	1	0	0	−1	−1	1		
	$z_k - c_k$	0	−1	−1	2 ↑	1	0	0	−1	−1	0	4	
	y_2	0	−1/2	−1/2	1	1/2	0	0	−1/2	−1/2	✖	2	
	u_2	0	−1/2	−1/2	0	−1/2	1	0	−1/2	5/2	✖	2	
	v_1	0	1/2	0	0	0	0	1	1/2	0	✖	3	
	x_1	1	1/2	0	0	0	0	0	−1/2	0	✖	2	
	$z_k - c_k$	0	0	0	0	0	0	0	0	0	✖	0	

KONEC druhé fáze. Úplně stejnou finální tabulku bychom dostali i při řešení úlohy (4.6.7), ale potřebovali bychom k tomu 2 kroky. Nyní tuto finální tabulku využijeme pro sestrojení úlohy ve třetí fázi.

Fáze 3

Ve třetí fázi řešíme úlohu

$$-\mu \to \min$$
$$-x_2/2 - y_1/2 + y_2 + u_1/2 - v_2/2 - \mu/2 = 2$$
$$-x_2/2 - y_1/2 - u_1/2 + u_2 - v_2/2 + 5\mu/2 = 2$$
$$x_2/2 + v_1 + v_2/2 = 3$$
$$x_1 + x_2/2 + -v_2/2 = 2$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad y_1 \geq 0 \quad \& \quad y_2 \geq 0$$
$$u_1 \geq 0 \quad \& \quad u_2 \geq 0 \quad \& \quad v_1 \geq 0 \quad \& \quad v_2 \geq 0 \quad \& \quad \mu \geq 0.$$

Třetí fázi zahájíme zavedením proměnné μ do báze (jinou možnost ani nemáme).

$\mu^{[0]} = 0$

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	x_B	x_B/λ
0	y_2	0	−1/2	−1/2	1	1/2	0	0	−1/2	−1/2	2	✖
0	u_2	0	−1/2	−1/2	0	−1/2	1	0	−1/2	(5/2)	2	4/5 →
0	v_1	0	1/2	0	0	0	0	1	1/2	0	3	✖
0	x_1	1	1/2	0	0	0	0	0	−1/2	0	2	✖
	c_k	0	0	0	0	0	0	0	0	−1		
	z_k	0	0	0	0	0	0	0	0	0		
	$z_k - c_k$	0	0	0	0	0	0	0	0	1 ↑	0	

$\mu^{[1]} = 4/5$

	y_2	0	−3/5	−3/5	1	2/5	1/5	0	−3/5	0	12/5	✖
	μ	0	−1/5	−1/5	0	−1/5	2/5	0	−1/5	1	4/5	✖
	v_1	0	1/2	0	0	0	0	1	1/2	0	3	6
	x_1	1	(1/2)	0	0	0	0	0	−1/2	0	2	4 →
	$z_k - c_k$	0	1/5 ↑	1/5	0	1/5	−2/5	0	1/5	0	−4/5	

Shetty:
u_2
⇕
B
⇕
x_2

$\mu^{[2]} = 8/5$

	y_2	6/5	0	−3/5	1	2/5	1/5	0	−6/5	0	24/5	
	μ	2/5	0	−1/5	0	−1/5	2/5	0	−2/5	1	8/5	
	v_1	−1	0	0	0	0	0	1	1	0	1	
	x_2	2	1	0	0	0	0	0	−1	0	4	
	$z_k - c_k$	−2/5	0	1/5	0	1/5	−2/5	0	2/5	0	−8/5	

KONEC třetí fáze, neboť $\mu^{[1]} < 1 < \mu^{[2]}$. Jelikož platí

$$x^{[1]} = [2,0], \quad x^{[2]} = [0,4] \quad \text{a} \quad \lambda^* = \frac{\mu^{[2]} - 1}{\mu^{[2]} - \mu^{[1]}} = \frac{8/5 - 1}{8/5 - 4/5} = \frac{3}{4},$$

je hledaným optimálním řešením dané úlohy

$$x^* = \lambda^* x^{[1]} + (1 - \lambda^*)\, x^{[2]} = [3/2, 1].$$

Část předchozích výpočtů lze využít i v případě, kdy vhodně pozměníme lineární částí účelové funkce. To si ukážeme v následujícím příkladě.

Příklad 4.6.6

Vyřešme úlohu

$$x_1^2 + x_2^2 + 2x_1x_2 + 20x_1 - 40x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad 2x_1 + x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení. Daná úloha se od předchozího příkladu liší pouze tím, že tentokrát máme $d = (20, -40)^\top$, což je 20-násobek vektoru d v předchozím příkladu. Proto první a druhá fáze budou zcela totožné s tím, že sloupec pro μ musíme vynásobit 20. Můžeme tedy ihned pokračovat třetí fází.

Fáze 3

Ve třetí fázi řešíme úlohu

$$-\mu \to \min,$$
$$-x_2/2 - y_1/2 + y_2 + u_1/2 - v_2/2 - 10\mu = 2,$$
$$-x_2/2 - y_1/2 - u_1/2 + u_2 - v_2/2 + 50\mu = 2,$$
$$x_2/2 + v_1 + v_2/2 = 3,$$
$$x_1 + x_2/2 + -v_2/2 = 2,$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad y_1 \ge 0 \quad \& \quad y_2 \ge 0$$
$$u_1 \ge 0 \quad \& \quad u_2 \ge 0 \quad \& \quad v_1 \ge 0 \quad \& \quad v_2 \ge 0 \quad \& \quad \mu \ge 0.$$

Třetí fázi opět zahájíme zavedením μ do báze (jinou možnost ani nemáme).

c_B	B	x_1	x_2	y_1	y_2	u_1	u_2	v_1	v_2	μ	x_B	x_B/λ
0	y_2	0	−1/2	−1/2	1	1/2	0	0	−1/2	−10	2	✖
0	u_2	0	−1/2	−1/2	0	−1/2	1	0	−1/2	(50)	2	2/50 →
0	v_1	0	1/2	0	0	0	0	1	1/2	0	3	✖
0	x_1	1	1/2	0	0	0	0	0	−1/2	0	2	✖
	c_k	0	0	0	0	0	0	0	0	−1		
	z_k	0	0	0	0	0	0	0	0	0		
	$z_k - c_k$	0	0	0	0	0	0	0	0	1 ↑	0	

$\mu^{[0]} = 0$

	y_2	0	−3/5	−3/5	1	2/5	1/5	0	−3/5	0	12/5	✖
	μ	0	−1/100	−1/100	0	−1/100	1/50	0	−1/100	1	2/50	✖
	v_1	0	1/2	0	0	0	0	1	1/2	0	3	6
	x_1	1	(1/2)	0	0	0	0	0	−1/2	0	2	4 →
	$z_k - c_k$	0	1/100 ↑	1/100	0	1/100	−1/50	0	1/100	0	−2/50	

$\mu^{[1]} = 2/50$

Shetty: u_2 ↕ B ↕ x_2

	y_2	6/5	0	−3/5	1	(2/5)	1/5	0	−6/5	0	24/5	12 →
	μ	1/50	0	−1/100	0	−1/100	1/50	0	−1/50	1	4/50	✖
	v_1	−1	0	0	0	0	0	1	1	0	1	✖
	x_2	2	1	0	0	0	0	0	−1	0	4	✖
	$z_k - c_k$	−1/50	0	1/100	0	1/100 ↑	−1/50	0	1/50	0	−4/50	

$\mu^{[2]} = 4/50$

Shetty: x_1 ↕ B ↕ u_1

	u_1	3	0	−3/2	5/2	1	1/2	0	−3	0	12	✖
	μ	1/20	0	−1/40	1/40	0	1/40	0	−1/20	1	1/5	✖
	v_1	−1	0	0	0	0	0	1	(1)	0	1	1 →
	x_2	2	1	0	0	0	0	0	−1	0	4	✖
	$z_k - c_k$	−1/20	0	1/40	−1/40	0	−41/200	0	1/20 ↑	0	−1/5	

$\mu^{[3]} = 1/5$

Shetty: y_2 ↕ B ↕ v_2

	u_1	0	0	−3/2	5/2	1	1/2	3	0	0	15	✖
	μ	0	0	−1/40	1/40	0	1/40	1/20	0	1	1/4	✖
	v_2	−1	0	0	0	0	0	1	1	0	1	✖
	x_2	1	1	0	0	0	0	1	0	0	5	✖
	$z_k - c_k$	0	0	1/40 ↑	−1/40	0	−41/200	−1/20	0	0	−1/4	

$\mu^{[4]} = 1/4$

Shetty: v_1 ↕ B ↕ y_1

KONEC třetí fáze. Podle Shettyho pravidla má nyní vstoupit do báze proměnná y_1. To je v souladu s Wolfeho pravidlem, ovšem sloupec pro y_1 neobsahuje kromě řádku $z_k - c_k$ žádné kladné číslo, takže není možné určit, kterou bychom z báze vyloučili. Současně máme $x^{[4]} = [0,5]$, $\mu^{[4]} = 1/4$ a z poslední tabulky určíme

$$\overline{X} = (0,0,1,0,3/2,0,0,0,1/40)^\top,$$

tj. $\overline{x} = [0,0]$ a $\overline{\mu} = 1/40$. Pak $\tau^* = \frac{1-\mu^{[4]}}{\overline{\mu}} = 30$, a tedy hledaným optimálním řešením dané úlohy je

$$x^* = x^{[4]} + \tau^* \overline{x} = [0,5].$$

Kdybychom hledali řešení úlohy (4.6.4) pro jinou hodnotu μ než 1, pak by postup byl zcela shodný, akorát bychom ve vzorcích pro určení x^* uvedených před Příkladem 4.6.2 nahradili číslo 1 požadovanou hodnotou ν, tj. buď se podaří $\mu^{[k]} = \nu$ nebo v čitatelích vzorců pro λ^* a τ^* bude místo 1 vystupovat ν. To nás může přivést až k přirozené otázce *parametrického programování* ohledně optimálních hodnot funkce

$$q(\mu) := Q(x^*(\mu),\mu) = \min_{\substack{Ax\le b\\ x\ge 0}} \{\tfrac{1}{2}\langle Cx,x\rangle + \mu\langle d,x\rangle\},$$

tj. pro jaké hodnoty parametru $\mu \ge 0$ nabude funkce $q(\mu)$ své největší[77] nebo nejmenší hodnoty[78]. Poměrně snadno se dá ukázat, že funkce $q(\mu)$ je konkávní a její optimální hodnotu lze pomocí vektorů

$$\alpha^{[k]} := d^\top x^{[k]} \quad \& \quad \beta^{[k]} := b^\top u^{[k]} \quad \& \quad \overline{\alpha} := d^\top \overline{x} \quad \& \quad \overline{\beta} := b^\top \overline{u}$$

vyjádřit jako:

(i) Je-li $\mu = \mu^{[k]}$ pro nějaké $k \in \{0,\dots,r\}$, pak

$$q(\mu) = Q(x^{[k]},\mu^{[k]}) = \frac{\mu^{[k]}\alpha^{[k]} - \beta^{[k]}}{2}.$$

(ii) Je-li $\mu^{[k]} < \mu < \mu^{[k+1]}$ pro nějaké $k \in \{0,\dots,r-1\}$, pak konvexní kombinace

$$q(\mu) = \frac{\mu^2(\alpha_{k+1}-\alpha_k) + 2\mu(\beta_{k+1}-\beta_k) - (\mu^{[k+1]}\beta_k - \mu^{[k]}\beta_{k+1})}{2(\mu^{[k+1]}-\mu^{[k]})}.$$

(iii) Je-li i $\mu^{[r]} < \mu$, pak

$$q(\mu) = \frac{\mu^2\overline{\alpha} + 2\mu\overline{\beta} - (\overline{\mu}\beta_r - \mu^{[r]}\overline{\beta})}{2\overline{\mu}}.$$

Navíc funkce $q(\mu)$ je diferencovatelná a pro její první derivací platí

(i) Je-li $\mu = \mu^{[k]}$ pro nějaké $k \in \{0,\dots,r\}$, pak

$$q'(\mu) = \alpha^{[k]}.$$

(ii) Je-li $\mu^{[k]} < \mu < \mu^{[k+1]}$ pro nějaké $k \in \{0,\dots,r-1\}$, pak konvexní kombinace

$$q'(\mu) = \lambda^*\alpha_k + (1-\lambda^*)\alpha_{k+1}$$

(iii) Je-li i $\mu^{[r]} < \mu$, pak

$$q'(\mu) = \alpha_r + \tau^*\overline{\alpha}.$$

Z konkávnosti funkce $q(\mu)$ proto plyne, že v případě

$$q'(0) = \alpha_0 \le 0$$

je funkce $q(\mu)$ nerostoucí s maximem pro $\mu^* = 0$. V případě $\alpha_0 > 0$ musíme opět rozlišit tři možné případy podle hodnoty

$$q'(\mu) = \alpha_r + \tau^*\overline{\alpha} \quad \text{pro } \mu \ge \mu^{[r]},$$

přičemž pro $\mu > \mu^{[r]}$ je $\tau^* > 0$.

[77] To tedy bude největší možné minimum, jaké můžeme úpravou hodnoty μ v účelové $Q(x,\mu)$ získat.

[78] Tím určíme nejmenší možné minimum, jaké můžeme získat.

(i) Je-li $\overline{\alpha} > 0$ nebo $\overline{\alpha} = 0$ a $\alpha^{[r]} > 0$, pak $q(\mu)$ je rostoucí funkce s minimem v bodě $\mu^* = 0$.

(ii) Je-li $\overline{\alpha} = 0$ a $\alpha^{[r]} = 0$, pak $q(\mu)$ nabývá minima v bodě $\mu^* = 0$ a maxima v bodech $\mu \geq \mu^{[r]}$.

(iii) Je-li $\overline{\alpha} < 0$ nebo $\overline{\alpha} = 0$ a $\alpha^{[r]} < 0$, potom buď $q'(\mu^{[r]}) = \alpha^{[r]} > 0$ nebo existuje dvojice $\mu^{[k]}$ a $\mu^{[k+1]}$ pro $k \in \{0, \ldots, r-1\}$ s vlastností

$$q'(\mu^{[k]}) = \alpha^{[k]} \geq 0 \quad \text{a} \quad q'(\mu^{[k+1]}) = \alpha^{[k+1]} \leq 0.$$

V prvním případě se minimum funkce $q(\mu)$ realizuje v bodě

$$\mu^* = -\overline{\beta}/\overline{\alpha} \quad \text{s odpovídající hodnotou} \quad q(\mu^*) = -\frac{\beta^{[r]}\overline{\alpha} - \overline{\beta}\,\alpha^{[r]}}{2\overline{\alpha}},$$

zatímco ve druhém případě to nastane v bodě

$$\mu^* = -\frac{\beta^{[k+1]} - \beta^{[k]}}{\alpha^{[k+1]} - \alpha^{[k]}} \quad \text{s odpovídající hodnotou} \quad q(\mu^*) = -\frac{\beta^{[k]}\,\alpha^{[k+1]} - \beta^{[k+1]}\,\alpha^{[k]}}{\alpha^{[k+1]} - \alpha^{[k]}}.$$

Pro více podrobností viz [250, str. 166–170].

Značnou nevýhodou Wolfeho metody je vysoká výpočetní náročnost dlouhého tvaru. Nutnosti jejího použití se lze vyhnout tak, že ke každému prvku na diagonále matice C přičteme malé $\varepsilon > 0$, tj. nahradíme C za matici $C + \varepsilon I$, díky čemuž se matice stane pozitivně definitní. Na takové upravené úloze pak jistě bude fungovat Wolfeho metoda v krátkém tvaru. Z výpočetních důvodů by ε nemělo být voleno příliš velké, ale na druhou stranu by současně nemělo být ani příliš malé, aby nezpůsobovalo numerickou nestabilitu. Tato myšlenka pochází od Bealeho a nalezneme ji již v původním Wolfeho článku[79]. Druhou možností, jak se těmto potížím vyhnout, je využití některé z dalších metod pro řešení úloh kvadratického programování, které mohou/nemusí být opět založeny na simplexovém algoritmu (jako jsou Dantzigova metoda, Whinstonova–van de Panneho metoda, Theilova–van de Panneho metoda[80], Shettyho–Lemkeho metoda, Bealeho metoda, Barankinova–Dorfmanova metoda, Frankové–Wolfova metoda aj. viz [250, 337, 452]).

[79] V [474, str. IV.21] se o této těsné hranici mezi pozitivně definitními a semidefinitními maticemi hovoří jako o *Razor's Edge Property* („vlastnost na hraně břitvy“).

[80] Např. tato metoda se v průběhu času ukázala být poměrně neefektivní.

4.7 Než opustíme kvadratické programování...

A jak je to s časovou složitostí úloh kvadratického programování? Úlohy s maticí $C \geq 0$ patří do třídy **P**, jak můžeme vidět např. s využitím některého z algoritmů popsaného v [48, 222, 236, 289, 333, 382, 383, 397, 518] včetně důkazu polynomiální časové složitosti. Tím však máme vypořádány pouze úlohy *konvexního* kvadratického programování. Jiné jsme doposud ani neuvažovali, ačkoli podmínky z Věty 4.3.1 platí i pro ostatní případy, tj. i pro negativně (semi)-definitní či indefinitní matici C. Vezmeme-li úlohu obecnou úlohu kvadratického programování ve standardním tvaru, tj. (4.5.1) bez jakéhokoli požadavku semidefinitnosti matice C, pak lze také ukázat, že nutnou a postačující podmínkou pro to, aby x^* bylo jejím řešením je

(i) $(Cx^* + d)^\top u \geq 0$ pro všechna $u \in \mathbb{R}^n$ splňující $\widehat{A}d \leq 0$

(ii) a současně $u^\top C u \geq 0$ pro všechna $u \in \mathbb{R}^n$ splňující $\widehat{A}d \leq 0$ a $(Cx^* + d)^\top u = 0$,

kde $\widehat{A}$ značí matici vytvořenou pouze z těch omezení, které se v bodě x^* realizují jako rovnosti. Toto je přímý důsledek Lemma 8.1.1. Také platí, že x^* je řešením (4.5.1) právě tehdy, když je také řešením úlohy lineárního programování

$$(Cx^* + d)^\top u \to \min \quad \& \quad Ax \leq b \quad \& \quad x \geq 0,$$

viz [389, Theorem 1.12].

Řešení nekonvexních úloh kvadratického programování je výrazně komplikovanější, a tudíž i zajímavější. Proč? Částečnou odpověď lze najít v analýze časové složitosti těchto úloh. Ačkoli nám to nabídne spolehlivé výsledky pouze v nejhorších případech, může nás to inspirovat

k možnému rozvoji nějakého efektivního algoritmu. Poměrně brzy převládlo mezi matematiky přesvědčení, že řešení obecné úlohy nekonvexního programování bude obtížné. První výsledek v tomto ohledu zaměřený na kvadratické programování nalezneme v [439], kde se ukázalo, že v případě $C < 0$ se jedná o **NP**-těžkou úlohu, viz také [200]. Později bylo v [494, 495] pomocí převodu na *Problém splnitelnosti booleovské formule*[81] ukázáno, že tyto úlohy současně patří do třídy **NP**, a tudíž se jedná o o **NP**-úplnou úlohu. Zhruba ve stejné době bylo s pomocí převodu na *problém přípustnosti batohu*[82], dokázáno, že **NP**-těžkým problémem je i úloha kvadratického programování s indefinitní maticí, viz [409]. To znamená, že stačí aby matice C měla jedno záporné vlastní číslo, a úloha kvadratického programování bude **NP**-těžká. Tyto výsledky nejsou příliš překvapivé. Překvapivější už může být to, že dokonce i pouhé ověření lokálního extrému je **NP**-úplný problém, viz [390, 495]. Toto je úzce spojeno s lokální ostrou konvexností účelové funkce (jakožto postačující podmínkou druhého řádu), jejíž zjištění je také **NP**-těžký problém s jednou výjimkou[83], viz [410]. Pokud toto navíc zkombinujeme se skutečností, že úlohy kvadratického programování mohou mít více řešení, bude zřejmá nepoddajnost obecné úlohy kvadratického programování. Uvažme např. úlohu

$$-\sum_{i=1}^{n} x_i^2 \to \min \quad \& \quad x_1 \in [-1,1] \quad \& \quad \cdots \quad \& \quad x_n \in [-1,1].$$

Jejím lokálním řešením je libovolný vrchol přípustné množiny, tj. body $[\pm 1, \ldots, \pm 1]$. Takových bodů je 2^n, a tudíž obecně nelze očekávat, že nalezení globálního minima pro tuto úlohu bude jakkoli efektivní – vlastně už jen nalezení lokálního minima (a jeho samotné ověření) může vyžadovat exponenciální množství času. Pro více podrobností viz mimo již uvedené reference také [181] nebo [155, Podkapitola 6.3].

Cvičení

Různými postupy (graficky, Hildrethovou–d'Esopovou metodou s možností různého pořadí výpočtu jednotlivých složek, Wolfeho metodou v krátkém či dlouhém tvaru) vyřešte následující úlohy kvadratického programování.

4.7.1.

$$x_1^2 + 2x_2^2 - 8x_1 - 10x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad x_1 + 2x_2 \le 8 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[3,2]$.

4.7.2.

$$x_1^2 + x_2^2 - 1000x_1 - 1000x_2 \to \min$$
$$3x_1 + x_2 \ge 3 \quad \& \quad x_1 + 4x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.3.

$$2x_1^2 + 2x_2^2 - 4x_1 - 4x_2 \to \min$$
$$2x_1 + 3x_2 \le 6 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[1,1]$.

[81] Angl. *Boolean satisfiability problem* (SAT). Jde o úlohu, kde pro dané $K \in \mathbb{R}$ chceme zjistit, zda existuje x takové, že

$$\tfrac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \le K \quad \& \quad Ax \le b.$$

[82] Angl. *knapsack feasibility problem*. V něm je dáno $n+1$ kladných celých čísel $a_1, \ldots, a_n, b$ a otázkou je existence přípustného bodu pro úlohu

$$\sum_{i=1}^{n} a_i x_i = b \quad \& \quad x_1, \ldots, x_n \in \{0,1\}.$$

[83] Jde o situaci, kdy účelová funkce je konkávní. Pak řešením je některý z vrcholů přípustné množiny a ověření optimality lze realizovat v polynomiálním čase.

4.7.4.

$$3x_1^2 + 3x_2^2 - 10x_1 - 24x_2 \to \min$$
$$2x_1 + x_2 \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.5.

$$x_1^2 + x_2^2 - 8x_1 - 10x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad x_1 + 2x_2 \le 8 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[2,3]$.

4.7.6.

$$x_1^2 + 2x_2^2 + 2x_1 + 8x_2 \to \min$$
$$10x_1 + 16x_2 \ge 72 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[4,2]$.

4.7.7.

$$x_1^2 + 4x_2^2 - 8x_1 - 16x_2 \to \min$$
$$x_1 + x_2 \le 5 \quad \& \quad x_1 \le 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.8.

$$x_1^2 + x_2^2 - 6x_1 - 3x_2 \to \min$$
$$x_1 + x_2 \le 2 \quad \& \quad x_1 - x_2 \le -1 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[1/2, 3/2]$.

4.7.9.

$$(x_1 - 1)^2 + (x_2 - 5/2)^2 \to \min$$
$$x_1 - x_2 \ge -2 \quad \& \quad x_1 + 2x_2 \le 6 \quad \& \quad x_1 - 2x_2 \le 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.10.

$$2x_1^2 - 2x_1x_2 + 2x_2^2 - 6x_1 \to \min$$
$$x_1 + x_2 \le 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[3/2, 1/2]$.

4.7.11.

$$2x_1^2 - 6x_1x_2 + 5x_2^2 - 16x_1 + 40x_2 \to \min$$
$$x_1 + x_2 \ge 3 \quad \& \quad 2x_1 - x_2 \ge 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[4,0]$.

4.7.12.

$$x_1^2 + x_1x_2 + 2x_2^2 + x_1 - 3x_2 \to \min$$
$$-2x_1 + 4x_2 \le -4 \quad \& \quad x_1 - 2x_2 \le -3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Úloha není přípustná, tj. $X = \emptyset$.

4.7.13.

$$3x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 - x_2 \to \min$$
$$2x_1 - x_2 \le 4 \quad \& \quad x_1 + 2x_2 \le 6 \quad \& \quad 3x_1 - 2x_2 \le 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[1/2, 1/2]$.

4.7.14.

$$x_1^2 + x_1x_2 + 2x_2^2 + 3x_1 - 5x_2 \to \min$$
$$x_1 - 3x_2 \le -2 \quad \& \quad 2x_1 + x_2 \ge 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[4/7, 13/7]$.

4.7.15.

$$2x_1^2 - 6x_1x_2 + 5x_2^2 - 16x_1 + 40x_2 \to \min$$
$$x_1 + x_2 \ge 3 \quad \& \quad 2x_1 - x_2 \ge 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[4, 0]$.

4.7.16.

$$x_1^2 - 4x_1x_2 + 5x_2^2 - 3x_1 - 14x_2 \to \min$$
$$3x_1 - 2x_2 \le 5 \quad \& \quad x_1 + 4x_2 \le 6 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $[2, 1]$.

4.7.17.

$$x_1^2 - x_1x_2 + x_2^2 - 3x_1 \to \min$$
$$x_1 + x_2 \le 2 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.18.

$$x_1^2 - x_1x_2 + x_2^2 - 3x_1 \to \min$$
$$x_1 + x_2 \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.19.

$$x_1^2 - 2x_1x_2 + 4x_2^2 - 2x_1 - 28x_2 \to \min$$
$$-x_1 + x_2 \le 2 \quad \& \quad 3x_1 - x_2 \le 0 \quad \& \quad -x_1 - x_2 \le 0.$$

Řešení: $[1, 3]$.

4.7.20.

$$x_1^2 - 2x_1x_2 + 4x_2^2 + 7x_1 - 9x_2 \to \min$$
$$5x_1 - x_2 \geq 2 \quad \& \quad x_1 + x_2 \geq 3 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[1,2]$.

4.7.21.

$$x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 - 6x_2 \to \min$$
$$x_1 + x_2 \leq 2 \quad \& \quad -x_1 + x_2 \leq \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

4.7.22.

$$2x_1^2 + x_2^2 - 2x_1x_2 - 6x_1 - 2x_2 \to \min$$
$$x_1 + x_2 \leq 2 \quad \& \quad 2x_1 - x_2 \leq 2 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

4.7.23.

$$4x_1^2 - 4x_1x_2 + x_2^2 - 3x_1 - x_2 \to \min,$$
$$x_1 - x_2 \leq 1,$$
$$-3x_1 + x_2 \leq 2,$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Nemá konečné řešení.

4.7.24.

$$4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2 \to \max$$
$$x_1 + 2x_2 \leq 2 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[1/3, 5/6]$.

4.7.25.

$$4x_1^2 - 4x_1\,x_2 + x_2^2 + 7x_1 + 7x_2 \to \min$$
$$x_1 - x_2 \leq 6 \quad \& \quad x_1 + 3x_2 \geq 4 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $[1/3, 5/6]$.

4.7.26.

$$x_1^2 + x_2^2 + x_3^2 - 6x_2 - 2x_3 \to \min$$
$$x_1 + x_2 + x_3 \leq 1 \quad \& \quad x_3 - x_2 \geq -1.$$

Řešení: $[-1, 3/2, 1/2]$.

4.7.27.

$$x_1^2 + 2x_2^2 + 4x_3^2 - 4x_2x_3 - 2x_1 - 4x_2 - 2x_3 \to \min$$
$$x_1 + 2x_2 + 4x_3 \le 12 \quad \& \quad 2x_1 + x_2 + 3x_3 \le 9$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[1, 5/2, 3/2]$.

4.7.28.

$$2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 8x_1 - 6x_2 - 4x_3 \to \min$$
$$x_1 + x_2 + x_3 \le 3 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

4.7.29.

$$x_1^2 + 3x_2^2/2 + x_3^2/2 + x_1x_2 + x_2x_3 - x_1 + 5x_2 - 2x_3 \to \min$$
$$2x_1 + 3x_2 - x_3 \le 5 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[2, 1/2, 0]$.

4.7.30.

$$2x_1^2 + x_2^2 + 3x_3^2 + x_1x_3 + 2x_2x_3 \to \min$$
$$2x_1 + 4x_2 + x_3 \le 6 \quad \& \quad 2x_1 + x_2 + 3x_3 \ge 3$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[4/7, 1/7, 4/7]$.

4.7.31.

$$3x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 + x_1x_3 - x_2x_3 + 3x_1 - 4x_3 + x_2 \to \min$$
$$x_1 + 2x_2 - x_3 \ge 3 \quad \& \quad x_1 + 3x_2 + 2x_3 \ge 1 \quad \& \quad 2x_1 - 4x_2 + x_3 \ge 2$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[2, 1/2, 0]$.

4.7.32.

$$x_1^2 + 2x_2^2 + x_3^2/3 + 2x_1x_2 - x_1x_3/2 - 4x_1 + x_2 - x_3 \to \min$$
$$x_1 + 2x_2 - x_3 \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Řešení: $[38/13, 0, 48/13]$.

4.7.33. Pro hodnoty $\alpha = 1, 16, 32, 50$ určete řešení

$$5x_1^2 - 6x_2 + 2x_2^2 + \alpha(-x_1 + x_2) \to \min$$
$$3x_1 - x_2 \ge 3 \quad \& \quad x_1 + x_2 \ge 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Řešení: $\alpha = 1$: $[7/4, 9/4]$,

$\alpha = 16$: $[3, 1]$,

$\alpha = 32$: $[4, 0]$,

$\alpha = 50$: $[5, 0]$.

4.7.34. Pro hodnoty $\alpha = -10, -5, 1$ určete řešení

$$2x_1^2 + 4x_1x_2 - 2x_2 + x_1 + 4x_2^2 + \alpha(x_1 - 2x_2) \to \min$$
$$x_1 + x_2 \geq 1 \quad \& \quad 3x_1 - x_2 \geq 5 \quad \& \quad -x_1 + 2x_2 \leq 1$$
$$x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $\alpha = -10$: $[5/3, 0]$,
$\alpha = -5$: $[1, 0]$,
$\alpha = 1$: $[1/3, 2/3]$.

4.7.35. Pro hodnoty $\alpha = -1, 6, 10$ určete řešení

$$x_1^2 - 2x_1x_2 + 2x_2^2 + \alpha(-x_1 - 4x_2) \to \min$$
$$x_1 + x_2 \leq 8 \quad \& \quad 5x_1 - 2x_2 \leq 9 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $\alpha = -1$: $[0, 0]$,
$\alpha = 6$: $[3, 5]$,
$\alpha = 10$: $[9/5, 31/5]$.

4.7.36. Pro hodnoty $\alpha = -20, -5, 0, 5$ určete řešení

$$x_1^2 + 4x_1x_2 + 4x_2^2 + \alpha(2x_1 - x_2) \to \min$$
$$x_1 + x_2 \geq 2 \quad \& \quad x_1 - x_2 \leq 3 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: $\alpha = -20$: $[28/9, 1/9]$,
$\alpha = -5$: $[3, 0]$,
$\alpha = 0$: $[2, 0]$,
$\alpha = 5$: $[0, 2]$.

4.7.37. Pro hodnoty $\alpha = -1, 0, 7, 10$ určete řešení

$$4x_1^2 - 4x_1x_2 + x_2^2 + \alpha(x_1 + x_2) \to \min$$
$$x_1 - x_2 \leq 6 \quad \& \quad x_1 + 3x_2 \geq 4 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Řešení: pro $\alpha = -1$ *je úloha neohraničená,*
pro $\alpha = 0$ *má úloha nekonečně mnoho řešení tvaru* $x_2 = 2x_1$, $x_1 \geq 4/7$,
$\alpha = 7$: $[1/7, 9/7]$,
$\alpha = 10$: $[0, 4/3]$.

4.7.38. Určete vzdálenost mezi trojúhelníkem s vrcholy v bodech $[0, 0]$, $[2, 0]$, $[0, 1]$ a trojúhelníkem s vrcholy $[0, 3]$, $[1, 2]$ a $[2, 2]$.

Řešení: $3\sqrt{5}/5$.

Kapitola 5

Dynamické programování

5.1 *Motivace* 368
5.2 *Konečně krokový deterministický proces* 370
5.3 *Stochastický proces* 386
5.4 *Nekonečně krokový deterministický proces* 390

AN OPTIMAL POLICY HAS THE PROPERTY THAT WHATEVER THE INITIAL STATE AND THE INITIAL DECISION ARE, THE REMAINING DECISIONS MUST CONSTITUTE AN OPTIMAL POLICY WITH REGARD TO THE STATE RESULTING FROM THE FIRST DECISION.[1]

MY FIRST TASK IN DYNAMIC PROGRAMMING WAS TO PUT IT ON A RIGOROUS BASIS. I FOUND THAT I WAS USING THE SAME TECHNIQUE OVER AND OVER TO DERIVE A FUNCTIONAL EQUATION. I DECIDE TO CALL THIS TECHNIQUE, "THE PRINCIPLE OF OPTIMALITY." OLIVER GROSS SAID ONE DAY, "THE PRINCIPLE IS NOT RIGOROUS." I REPLIED, "OF COURSE NOT, IT IS NOT EVEN PRECISE." A GOOD PRINCIPLE SHOULD GUIDE THE INTUITION.[2]

RICHARD BELLMAN[3]

Viz „bibli" každého dynamického programovače [45, str. 83] a také Bellmanovu autobiografii [47, str. 174].

[1] *Optimální strategie má tu vlastnost, že ať je počáteční stav a počáteční rozhodnutí jakékoli, zbývající rozhodnutí musí tvořit optimální strategii vzhledem ke stavu vyplývajícího z prvního rozhodnutí.*

[2] *Mým prvním úkolem v dynamickém programování bylo vybudovat pro něj rigorózní základy. Zjistil jsem, že pořád dokola využívám stále tu stejnou metodu k odvození funkcionální rovnice. Rozhodl jsem se tuto metodu nazvat „principem optimality". Oliver Gross mi jednoho dne řekl: „Ten princip není rigorózní." Odpověděl jsem mu: „Samozřejmě, že ne. Není dokonce ani přesný." Správný princip by měl nasměrovat intuici.*

[3]

Richard Ernest Bellman (26. srpna 1920 – 19. března 1984) byl americký matematik, v roce 1953 položil základy dynamického programování. V roce 1975 získal *John von Neumann Theory Prize* za „jeho více než dvacet let trvající vůdčí roli v oblasti, kterou sám vytvořil: dynamické programování".

5.1 Motivace

V úvodní části Kapitoly 2 jsme hovořili o tom, že Dantzig postrádal v původním Leontiefově modelu *dynamiku* umožňující změny v průběhu času. O dynamice jsme se zmínili i v souvislosti s von Neumannovou prací z roku 1932. Hledání optimální strategie pro procesy, které se vyvíjejí v čase, jsme si také ukázali v několika příkladech, viz např. Příklady 2.3.7, 2.3.10, 2.12.2. Ovšem ne vždy je řešení takových problémů pomocí lineárního programování možné či efektivní. My si v této kapitole ukážeme jiný možný přístup k takovýmto úlohám, a to s využitím tzv. *dynamického programování*, jehož základy položil Bellman na počátku 50. let 20. století a v průběhu následujících desetiletí jej (společně s mnoha dalšími) úspěšně rozvíjel. Hlavním nástrojem této oblasti je tzv. *Bellmanův princip optimality*, který si zformulujeme za chvíli[4].

Příklad 5.1.1

Uvažme společnost *Sjednocené pískové závody Brno*, které prodávají písek do celého světa (rozuměj mezi Brnem, Rajhradem a Kuřimí). Tento písek je možné těžit buď v Černovicích nebo v Hrušovanech u Brna, přičemž kvalita písku z obou závodů je zcela totožná. Každá koruna investovaná do pískovny v Černovicích na začátku roku přinese na konci roku a v každém dalším roce 2 tuny písku a 0,5 Kč zisku. Podobně každá koruna investovaná v Hrušovanech přinese 1 tunu písku a 0,9 Kč. Naším úkolem je určit rozdělení finančních prostředků, které jsou vždy na počátku roku k dispozici, mezi oba závody tak, aby celková produkce v období následujících pěti let byla maximální. Náš počáteční kapitál činí y Kč.

Řešení. Označme symbolem $x_n^{\check{c}}$ celkové investiční prostředky vynaložené v Černovicích do konce n-tého roku a podobně investice v Hrušovanech označíme jako x_n^{h}. Dále označme volné investiční prostředky na konci $(n-1)$-ho (tj. na začátku n-tého) roku jako v_n a poměrnou část volných prostředků reinvestovaných v Černovicích na začátku n-tého roku jako $0 \le r_n \le 1$. Pak pro investice platí

$$x_n^{\check{c}} = x_{n-1}^{\check{c}} + r_n\, v_n \quad \text{a} \quad x_n^{\mathrm{h}} = x_{n-1}^{\mathrm{h}} + (1-r_n)\, v_n, \quad \text{kde} \quad v_n = 0{,}5\, x_{n-1}^{\check{c}} + 0{,}9\, x_{n-1}^{\mathrm{h}},$$

pro $n = 1,2,\ldots,5$, přičemž bereme $x_0^{\check{c}} = 0 = x_0^{\mathrm{h}}$ a $v_1 = y$. Velikost produkce na konci n-tého roku je $2x_n^{\check{c}} + x_n^{\mathrm{h}}$ tun písku. Cílem proto je stanovení $r_1,\ldots,r_5$ takových, že

$$\sum_{n=1}^{5} (2x_n^{\check{c}} + x_n^{\mathrm{h}}) \to \max \quad \text{neboli} \quad \sum_{n=1}^{5} \left[2(x_{n-1}^{\check{c}} + r_n\, v_n) + x_{n-1}^{\mathrm{h}} + (1-r_n)\, v_n\right] \to \max.$$

To asi nebude úloha lineárního, celočíselného ani kvadratického programování…

Příklad 5.1.2

Uvažme násobení matic $A_1,\ldots,A_n$ vhodných rozměrů. Velmi dobře víme, že tato operace není komutativní, ale pouze asociativní. Navíc početní náročnost tohoto součinu velmi ovlivňuje pořadí jednotlivých součinů, a proto naším cílem je minimalizace celkového počtu potřebných součinů, přičemž neodlišujeme násobení nulou, jedničkou nebo jakýmkoli jiným číslem. Např. pro trojici matic

$$A_1 \in \mathbb{R}^{5\times 4} \quad \& \quad A_2 \in \mathbb{R}^{4\times 6} \quad \& \quad A_3 \in \mathbb{R}^{6\times 2}$$

dostaneme při uzávorkování $(A_1 \cdot A_2)\cdot A_3$ celkem $(5\cdot 6\cdot 4)+(5\cdot 2\cdot 6) = 180$ potřebných součinů, zatímco pro $A_1\cdot(A_2\cdot A_3)$ to je jenom $(5\cdot 2\cdot 4)+(4\cdot 2\cdot 6) = 88$ potřebných součinů. Vzhledem k rovnosti mezi počty řádků a sloupců sousedních matic, není naštěstí nutné si pamatovat kompletní rozměry všech matic, a stačí tak pracovat pouze s posloupností $p_1, p_2, \ldots, p_{n+1}$ určující počty řádků matic $A_1,\ldots,A_n$ a počet sloupců matice A_n. Jak ale pomocí této posloupnosti najít optimální pořadí jednotlivých součinů?

[4] (Tato poznámka vychází především z [438].)
Zdá se, že tento princip byl znám některým matematikům, statistikům či ekonomům již v polovině 40. let 20. století, kteří jej nezávisle na sobě používali při řešení tzv. *sekvenčních rozhodovacích problémů* obsahujících nějaké riziko či nejistotu. Zmiňme např. von Neumanna a Morgensterna, kteří jej použili ve svém průlomovém pojednání o teorii her k získání *subherní perfektní rovnováhy v hrách s rozšířenou formou*, viz [503, str. 126]. Wald, kterému se připisuje vznik *statistické rozhodovací teorie*, tuto teorii rozšířil na sekvenční rozhodování ve své knize [504] z roku 1947. Wald zobecnil problém hráčova zbankrotování z teorie pravděpodobnosti a zavedl *sekvenční podílový test pravděpodobností*, který minimalizuje očekávaný počet pozorování v sekvenčním zobecnění klasického testování hypotéz. Nicméně role zpětné indukce není v práci Walda až tak zjevná. Jasněji ji popsali Arrow, Blackwell a Girshick v jejich článku [15] z roku 1949. Ti zkoumali zobecněnou verzi statistického rozhodovacího problému a formulovali a vyřešili jej způsobem, který je dnes snadno rozpoznatelný jako aplikace moderního dynamického programování. Následujíce Walda charakterizovali optimální rozhodnutí ve statistickém rozhodování (např. přijmout nebo zamítnout hypotézu), přičemž zohlednili náklady na získávání dalších pozorování. V sekci „Nejlepší zkrácený podprogram“ (The Best Truncated Procedure) ukazují, jak lze optimální rozhodnutí aproximovat „mezi všemi sekvenčními postupy, které nevyžadují více než N pozorování…“ a řeší nalezení optimálního zkráceného výběrového podprogramu pomocí „zpětné indukce“, viz [15, str. 217].

Další rané aplikace principu zpětné indukce lze najít v práci Massého (1944) v oblasti statistické hydrologie a řízení přehrad a v analýze optimální zásobovací politiky od Arrowa, Harrise a Marschaka (1951), viz [14, 372]. Bellmanovi se obecně připisuje zásluha za rozpoznání společné struktury, která je základem sekvenčních rozhodovacích procesů, a za ukázání, jak lze pomocí zpětné indukce řešit širokou třídu těchto problémů s nejistotou, tj. v situaci, kdy jsou důsledky rozhodnutí známy pouze s jistou mírou pravděpodobnosti. Většinu své práce v této oblasti vykonal Bellman v době svého působení v RANDu, které začalo v roce 1949. Právě tam vymyslel termín

dynamické programování, který je dnes obecně přijímaným synonymem pro zpětnou indukci. Bellman v [47, str. 159] objasňuje, kde se právě toto označení vzalo. Podle jeho slov nebylo období 50. let pro matematický výzkum dobrým. Ministrem obrany byl v té době jistý Wilson (bývalý generální ředitel společnosti General Motors), který měl „patologický strach a odpor ke slovu *výzkum*" – celý zrudl a rozzuřil se, když někdo toto slovo použil v jeho přítomnosti. Bellman chtěl proto před Wilsonem zakrýt, že provádí v RANDu matematický výzkum. Jaké označení ale použít? Jelikož jeho zájmem bylo plánování (toto slovo mu nepřišlo dobré z několika důvodů) a rozhodovací procesy, rozhodl se pro „programování". Zároveň chtěl zachytit to, že se jedná o něco dynamického, vícestupňového a proměnlivého v čase, což jej přivedlo k označení „dynamický". Tím zabil dvě mouchy jednou ranou, protože ten význam byl naprosto přesný a zároveň má toto přídavné jméno jednu zajímavou vlastnost – nelze ho použít v pejorativním smyslu. Schválně zkuste vymyslet nějakou kombinaci, která by dala slovu „dynamický" nějaký negativní význam. To Bellman považoval za nemožné, a tak se rozhodl používat toto označení, proti kterému nemohl ani kongresman nic namítat.

[5] ⚠ Třetí příklad je velmi zjednodušenou verzí obecné úlohy lineárního programování, pro kterou je aplikovaní Bellmanova principu už „poněkud" složitější a pro praktické účely nepříliš vhodné – ale v závěru Podkapitoly 5.2 si to ukážeme alespoň pro úlohu s dvojicí omezení. Na druhou stranu pro řešení úloh celočíselného programování je využití dynamického programování mnohem efektivnější. Např. pro problém obchodního cestujícího je výpočetní složitost $O(n!)$ při využití hrubé síly a exponenciální s využitím Gomoryho metody řezů, zatímco s pomocí dynamického programování dostaneme $O(n^2 2^n)$ a tzv. *pseudopolynomiální* složitost.

[6] Což lze chápat i v přeneseném smyslu.

Příklad 5.1.3

Při řešení úloh lineárního programování některé z nás možná napadlo, že by se výpočet mohl urychlit, kdybychom nějakým způsobem využili možnost vyjádření vzájemné závislosti proměnných z daných omezení. To by mohlo probíhat např. následujícím způsobem. Uvažme úlohu

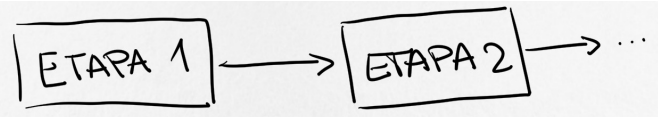
$$x_1 - 2x_2 \to \min \quad \& \quad 3x_1 + 5x_2 \le 10 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Připusťme, že už jsme nějakým způsobem našli vhodné $x_2 = X \ge 0$. Pak pro x_1 musí platit $0 \le x_1 \le (10 - 5X)/3$ a přínos této proměnné k celkové hodnotě účelové funkce je roven přímo x_1. Odtud vidíme, že nejmenšího přínosu docílíme volbou $x_1^* = 0$. V takovém případě máme $0 \le x_2 \le 2$, přičemž přínos této proměnné k účelové funkci je $-2x_2$, což bude minimální pro $x_2^* = 2$. Nalezli jsme tak řešení $x^* = [0,2]$ s hodnotou $f^* = -4$.

Kdybychom nejdříve zvolili $x_1 = X \ge 0$, pak bychom měli $0 \le x_2 \le (10 - 3X)/5$, což by nám minimalizovalo přínos k účelové funkci při volbě $x_2 = (10 - 3X)/5$. Dosazením této hodnoty do účelové funkce obdržíme $X - 2(10 - 3X)/5 = \frac{11}{5}X - 4$, takže minima dosáhneme, pokud X zvolíme nejmenší možné, tj. $X = 0$. Odtud zpětným dosazením do vyjádření x_2 získáme opět řešení $x^* = [0,2]$ s funkční hodnotou -4.

Tyto tři motivační příklady jsou ukázkou[5] tzv. *víceetapových/vícekrokových rozhodovacích procesů*, tj. procesů, při nichž dochází k rozhodnutí v každém kroku. My pak hledáme jednotlivá dílčí rozhodnutí, která nám společně vytvoří *optimální strategii*, která přinese maximální užitek/zisk apod. Tyto procesy obvykle zahrnují změny v čase[6] – dynamičnost, a proto se oblast věnovaná tomuto přístupu k optimalizačním úlohám nazývá *dynamické programování*.

Základní jednotkou každého procesu je *etapa* či *krok* (v předchozích příkladech to byl jeden rok, jedno násobení a jedna proměnná). Každý proces je pak rozdělen do několika dílčích etap a pro každou z nich hledáme vhodné „rozhodnutí", což je ilustrováno na Obrázku 5.1.



Obrázek 5.1: Etapy rozhodovacího procesu.

Dynamické programy můžeme klasifikovat do několika tříd.

(i) Podle rozhodovací proměnné:
  (a) *diskrétní* – rozhodovací proměnná může nabývat pouze diskrétních (a tedy jen spočetně mnoha) hodnot;
  (b) *spojité* – rozhodovací proměnná může nabývat libovolných hodnot z nějakého intervalu (tj. nespočetně mnoha).

(ii) Podle počtu kroků: *konečně krokový* (tj. konečný horizont) nebo *nekonečně krokový* (tj. nekonečný horizont).

(iii) Podle znalosti důsledků konkrétních rozhodnutí:
  (a) *deterministické* – budoucnost je určitá, tj. výsledek každého rozhodnutí je a priori znám;
  (b) *stochastické* – budoucnost je neurčitá, tj. existují i jisté vnější vlivy, které mohou ovlivnit výsledek učiněného rozhodnutí (známe alespoň míru pravděpodobnosti jednotlivých možných výsledků každého rozhodnutí).

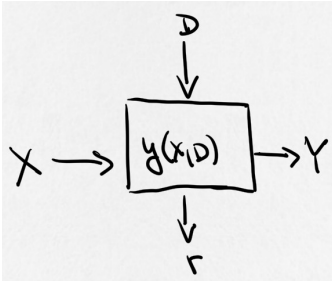
(iv) Podle rozměru rozhodovací proměnné: *jednoparametrický proces* nebo *víceparametrický proces* (např. rozdělení dvou na sobě nezávislých zdrojů velikosti X a Y do n ekonomických aktivit).

Hlavní myšlenkou řešení těchto problémů je tzv. *dekompozice*[7], kdy se původní komplexní problém rozdělí na řadu (z jistého pohledu) jednodušších úloh. Výsledek pak získáme kombinací/složením řešení těchto jednodušších úloh, přičemž

(i) *rozhodovací situace* nastává v případě, kdy existuje více než jedno přípustné řešení;

(ii) *cílem* rozhodování (nebo optimalizačního problému) je určení jednoho rozhodnutí (řešení, politiky, strategie), které dává optimální výsledek (tj. žádná jiná strategie nedává lepší výsledek).

Nyní se podíváme blíže na jednu etapu rozhodovacího procesu, viz Obrázek 5.2.



Obrázek 5.2: Podrobný vhled do jedné etapy rozhodovacího procesu.

V této etapě vystupuje několik veličin a funkcí:

(i) *vstupní veličina* (*vektor*) X představuje popis výchozího (počátečního) stavu systému a obsahuje všechny relevantní informace;

(ii) *výstupní veličina* (*vektor*) Y popisuje systém v konečné úrovni a obsahuje všechny informace o výstupu;

(iii) *rozhodovací proměnná* D (též *vektor řízení*) charakterizuje operace probíhající v průběhu jednotlivých etap;

(iv) *účelová funkce* r je jednorozměrná funkce vstupu, rozhodování a výstupu, tj.

$$r = r(X, D, Y);$$

(v) *transformační funkce* y vyjadřuje každou komponentu výstupní proměnné jako funkci vstupu a rozhodnutí, tj.

$$Y = y(X, D),$$

a ve stochastickém programování bude záviset i na náhodné veličině, tj.

$$Y = y(X, D, Q).$$

S využitím transformační funkce můžeme vyeliminovat Y z účelové funkce, čímž dostaneme[8]

$$r = r\big(X, D, y(X, D)\big) = r(X, D).$$

Jednokrokový optimalizační problém pak spočívá v nalezení maxima/minima účelové funkce jakožto funkce vstupní veličiny. Označíme-li tuto optimální (pro jednoduchost zvolme maximální) hodnotu jako f^* a odpovídající optimální rozhodnutí jako $D^* = D(X)$, pak platí

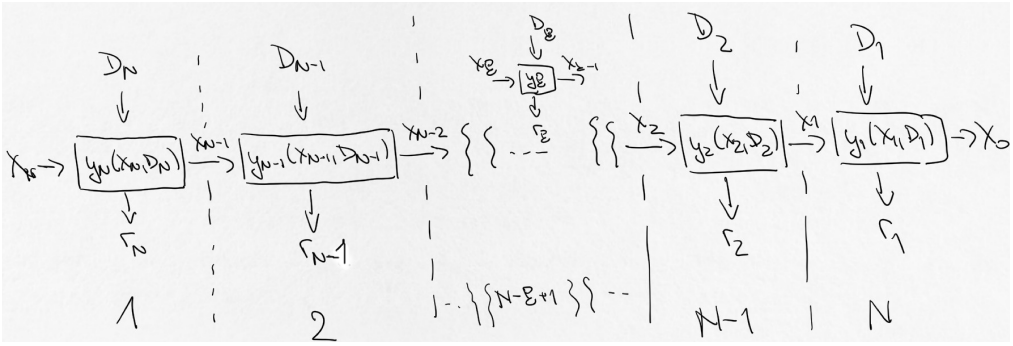
$$f^* = r(X, D^*) = f\big(X, D(X)\big) = \max_D r(X, D) \geq r(X, D) \quad \text{pro všechna přípustná } D.$$

5.2 Konečně krokový deterministický proces

Nyní se blíže podíváme na N-krokový rozhodovací proces, jehož schéma je znázorněno na Obrázku 5.3.

[7] Princip *divide et impera* (*rozděl a panuj*.

[8] ⚠ Pozor na značení: funkce r formálně vyžaduje 3 proměnné, ovšem zápis $r(X, D)$ znamená, že k tomu ve skutečnosti postačuje znát pouze 2 proměnné X a D, tj. vstupní a rozhodovací proměnná.



Obrázek 5.3: Podrobný rozbor všech etap vícekrokového procesu.

Jednotlivé etapy jsou ve schématu úmyslně indexovány od N po 0, tj. X_N představuje výchozí stav a X_0 je cílový stav[9]. Jelikož výstupní veličina k-tého kroku je totožná se vstupní veličinou $(k-1)$-ho kroku, platí

$$X_{k-1} = y_k(X_k, D_k) \quad \text{pro } k = N, N-1, \ldots, 1 \tag{5.2.1}$$

a výnos v k-tém kroku je

$$r_k = r_k(X_k, D_k).$$

Pak ale z (5.2.1) vyplývá, že X_k závisí na rozhodnutích, která předcházejí k-tému kroku, tj. na $D_N, D_{N-1}, \ldots, D_{k+1}$ a X_N neboli[10]

$$\begin{aligned} X_k &= y_{k+1}(X_{k+1}, D_{k+1}) = y_{k+1}(y_{k+2}(X_{k+2}, D_{k+2}), D_{k+1}) = y_{k+1}(X_{k+2}, D_{k+2}, D_{k+1}) = \\ &= y_{k+1}(y_{k+3}(X_{k+3}, D_{k+3}), D_{k+2}, D_{k+1}) = y_{k+1}(X_{k+3}, D_{k+3}, D_{k+2}, D_{k+1}) = \\ &= \cdots = y_{k+1}(X_N, D_N, \ldots, D_{k+1}). \end{aligned}$$

Dosazením do účelové funkce pak získáme

$$r_k = r_k(X_k, D_k) = r_k\big(y_{k+1}(X_N, D_N, \ldots, D_{k+1}), D_k\big) = r_k(X_N, D_N, \ldots, D_k), \tag{5.2.2}$$

tj. výnos v k-tém kroku je ovlivněn pouze rozhodnutími v N-tém, $(N-1)$-ním, …, k-tém kroku. Nemají na něj tedy (pochopitelně) žádný vliv rozhodnutí z budoucnosti. Cílem dynamického programování pak může být nalezení strategie takové, že např. určíme jednu z následujících čtyř optimálních hodnot[11]

$$\sum_{k=1}^{N} g\big(r_k(X_k, D_k)\big) \to \text{opt} \quad \text{nebo} \quad g\Big(\sum_{k=1}^{N} r_k(X_k, D_k)\Big) \to \text{opt} \quad \text{nebo}$$
$$g(X_0) \to \text{opt} \quad \text{nebo} \quad r_k(X_k, D_k) \to \text{opt} \quad \text{pro } k = N, \ldots, 1.$$

Nechť proto celkový „výnos“[12] R_N celého procesu je nějakou funkcí jednotlivých hodnot účelových funkcí, tj.

$$R_N = R_N(X_N, \ldots, X_1, D_N, \ldots, D_1) := g\big[r_N(X_N, D_N), r_{N-1}(X_{N-1}, D_{N-1}), \ldots, r_1(X_1, D_1)\big]. \tag{5.2.3}$$

Nyní ovšem můžeme s pomocí (5.2.2) vyeliminovat z (5.2.3) veličiny $X_{N-1}, \ldots, X_1$, čímž získáme

$$R_N = R_N(X_N, D_N, \ldots, D_1) = g\big[r_N(X_N, D_N), r_{N-1}(X_N, D_N, D_{N-1}), \ldots, r_1(X_N, D_N, \ldots, D_1)\big]$$

neboli optimalizační problém, ve kterém chceme optimalizovat celkový výnos pomocí dílčích rozhodnutí $D_N, \ldots, D_1$ při počátečním stavu X_N. Označíme-li dále jako $f_N(X_N)$ optimální řešení N-krokového procesu s výchozí hodnotou X_N a jako $D_k^* = D_k(X_k)$, $X_{k-1}^* = y_k(X_K, D_k^*)$ optimální rozhodnutí a odpovídající výstupní hodnotu, potom

$$\begin{aligned} f_N(X_N) &:= g\big[r_N(X_N, D_N^*), r_{N-1}(X_{N-1}^*, D_{N-1}^*), \ldots, r_1(X_1^*, D_1^*)\big] = \\ &= \text{opt}_{D_N, \ldots, D_1}\, g\big[r_N(X_N, D_N), r_{N-1}(X_{N-1}, D_{N-1}), \ldots, r_1(X_1, D_1)\big] \end{aligned}$$

[9] V tuto chvíli to možná vypadá jako další zbytečná komplice, ovšem význam této změny oceníme za chvíli. Snad…

[10] Stále mějme na paměti předchozí poznámku ohledně značení a počtu proměnných.

[11] V prvním případě optimalizujeme součet jednotlivých výnosů, které nějakým způsobem modifikujeme (např. jim přiřadíme různé váhy); ve druhém případě modifikujeme až celkový součet jednotlivých výnosů (což může, ale nemusí mít žádný vliv v závislosti na zvolené funkci g; třetí volba je tzv. optimalizace cílového stavu (tj. zajímá nás pouze konečná hodnota, kterou případně nějak modifikujeme); v posledním případě optimalizujeme všechny výnosy v jednotlivých krocích zvlášť. Konkrétní volba optimalizované funkce samozřejmě vždy závisí na konkrétní situaci.

[12] Skutečné označení pochopitelně závisí na interpretaci funkce r.

za podmínky

$$X_{k-1} = y_k(X_k, D_k) \quad \text{pro } k = N, \ldots, 1,$$

a tudíž také

$$f_N(X_N) = \operatorname{opt}_{D_N,\ldots,D_1} g\big[r_N(X_N, D_N), r_{N-1}(X_N, D_N, D_{N-1}), \ldots, r_1(X_N, D_N, \ldots, D_1)\big].$$

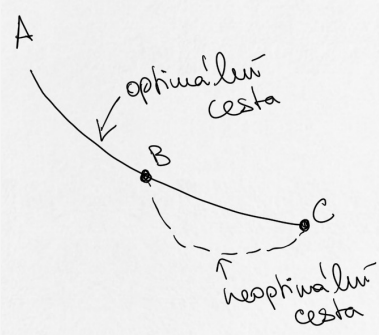
Hlavním nástrojem řešení úloh dynamického programování je *Bellmanův princip optimality* zmíněný na začátku této kapitoly. Zde je ještě několik dalších alternativních formulací:

> *Je-li ke každému subproblému nalezeno optimální řešení, pak existuje i optimální řešení původního problému (a to je z nich složeno).*
>
> *Každá podstrategie optimální strategie je opět optimální.*
>
> *Je-li posloupnost $\{D_N, \ldots, D_k\}$ optimální strategií $(N-k)$-krokového rozhodovacího procesu s počátečním stavem X_N, kde $D_i = D_i(X_i)$ závisí pouze na aktuálním stavu pro $i = N, \ldots, k$, pak posloupnost $\{D_N, \ldots, D_{k+1}\}$ tvoří optimální strategii $(N-k-1)$-krokového rozhodovacího procesu.*

Celou „filosofii“ Bellmanova principu je možné chápat jako používání *zdravého rozumu* a ilustrovat to lze na následujícím velmi jednoduchém příkladu.

Příklad 5.2.1

Mějme za úkol najít optimální cestu z bodu A do C.



Obrázek 5.4: Optimální cesta a Bellmanův princip optimality.

Připusťme, že řešení vede přes bod B. Pak ale cesta z B do C, která je součástí nalezeného řešení, také musí být optimální jakožto nejlepší cesta z B do C. Kdyby totiž existovala výhodnější cesta z B do C, pak by spojení této cesty s cestou z A do B dávalo lepší řešení, než původní nalezené řešení. A to je spor ↯.

Připusťme nyní, že celková účelová funkce je aditivní, tj.

$$\begin{aligned} g\big[r_N(X_N, D_N), r_{N-1}(X_{N-1}, D_{N-1}), \ldots, r_1(X_1, D_1)\big] = \\ = r_N(X_N, D_N) + r_{N-1}(X_{N-1}, D_{N-1}) + \cdots + r_1(X_1, D_1), \end{aligned}$$

a pro určitost se dohodněme, že budeme hledat maximum, tj.

$$f_N(X_N) = \max_{D_N,\ldots,D_1} \big[r_N(X_N, D_N) + r_{N-1}(X_{N-1}, D_{N-1}) + \cdots + r_1(X_1, D_1)\big] \tag{5.2.4}$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \ldots, 1$. Hodnota $r_N(X_N, D_N)$ nezávisí na $D_{N-1}, \ldots, D_1$, proto díky rovnosti

$$\max_{u_1,u_2} \big[h_1(u_1) + h_2(u_2)\big] = \max_{u_1} \big[h_1(u_1) + \max_{u_2} h_2(u_1, u_2)\big] \tag{5.2.5}$$

plyne z (5.2.4), že

$$f_N(X_N) = \max_{D_N}\{r_N(X_N, D_N) + \max_{D_{N-1},\dots,D_1}\big[r_{N-1}(X_{N-1}, D_{N-1}) + \cdots + r_1(X_1, D_1)\big]\}$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \dots, 1$. Pak z definice $f_N(X_N)$ a Bellmanova principu vyplývá

$$f_N(X_N) = \max_{D_N}\big[r_N(X_N, D_N) + f_{N-1}(X_{N-1})\big]$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \dots, 1$, tj.

$$f_N(X_N) = \max_{D_N}\big[r_N(X_N, D_N) + f_{N-1}\big(y_N(X_N, D_N)\big)\big].$$

Položíme-li nyní

$$Q_N(X_N, D_N) := r_N(X_N, D_N) + f_{N-1}\big(y_N(X_N, D_N)\big),$$

pak určení $f_N(X_N)$ a odpovídajícího $D_N^* = D_N(X_N)$ je snadné, pokud již známe $f_{N-1}(X_{N-1})$, tj. máme jednoduchý optimalizační problém se vstupní proměnnou X_N, rozhodnutím D_N a výnosem Q_N neboli

$$f_N(X_N) = \max_{D_N} Q_N(X_N, D_N).$$

Podařilo se nám tak rozdělit N-krokový problém na dva menší subproblémy:

(i) $(N-1)$-krokový optimalizační problém

$$f_{N-1}(X_{N-1}) = \max_{D_{N-1},\dots,D_1}\big[r_{N-1}(X_{N-1}, D_{N-1}) + \cdots + r_1(X_1, D_1)\big]$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N-1, \dots, 1$;

(ii) jednorozměrný problém

$$f_N(X_N) = \max_{D_N} Q_N(X_N, D_N) = \max_{D_N}\big[r_N(X_N, D_N) + f_{N-1}\big(y_N(X_N, D_N)\big)\big].$$

Analogicky můžeme pokračovat i dále a $f_{N-1}(X_{N-1})$, $f_{N-2}(X_{N-2}), \dots f_2(X_2)$ rozdělit na subproblémy, čímž získáme *rekurzivní*[13] *schéma*[14]

$$f_k(X_k) = \max_{D_k} Q_k(X_k, D_k), \quad \text{pro } k = N, \dots, 1, \tag{5.2.6}$$

kde

$$Q_k = \begin{cases} r_k(X_k, D_k), & k = 1, \\ r_k(X_k, D_k) + f_{k-1}\big(y_k(X_k, D_k)\big), & k = 2, \dots, N. \end{cases}$$

Pak pomocí tohoto schématu nalezneme optimální řešení původního problému tak, že jej začneme řešit pro $k = 1$ a pokračujeme[15] až po $k = N$. Tím nalezneme výnos $f_N(X_N)$ v N-tém kroku, odpovídající optimální rozhodnutí $D_N^* = D_N(X_N)$ a *rozhodovací funkci* $D_k = D_k(X_k)$ pro $k = N-1, \dots, 1$. Následně získáme optimální rozhodnutí a aktuální stavy jakožto funkce proměnné X_N, tj.

$$X_{N-1}^* = y_N(X_N, D_N^*) = y_N\big(X_N, D_N(X_N)\big) = y_N(X_N)$$

a

$$D_{N-1}^* = D_{N-1}(X_{N-1}^*) = D_{N-1}\big(y_N(X_N, D_N^*)\big) = D_{N-1}(X_N)$$

a rekurzivně dále

$$X_{k-1}^* = y_k(X_k, D_k^*) = y_k(X_k), \quad D_{k-1}^* = D_{k-1}(X_{k-1}^*) = D_{k-1}(X_k), \quad k = N-1, \dots, 1.$$

Speciálním případem takového rozhodovacího procesu je *optimalizace cílového stavu*. Jedná se o úlohu, jejímž cílem je

$$f_N(X_N) = \operatorname{opt}_{D_N,\dots,D_1} g(X_0)$$

[13] *Rekurze* je stav, kdy je určitý objekt v nějakém smyslu součástí sebe samotného (jeho definice jej obsahuje v nějaké jednodušší podobě nebo v jeho předchozí verzi). V našem případě jde o to, že tento objekt je definován pomocí sebe sama. Definice tohoto typu mají zpravidla dvě základní části: počáteční tvrzení a způsob, jak se z aktuálního stavu odvodí stav následující. Můžeme hovořit také o *rekurentním vztahu*.

[14] Toto si také můžeme představit jako „dominová efekt". Začneme s řešením jednoduchého problému (první kostička domina), pak pokračujeme trochu složitějším (druhá kostka domina), pak o trochu více složitějším atd. až dojdeme úplně na konec a nalezneme celého problému. Trochu patetičtěji řečeno můžeme toto schéma (což je ona *zpětná indukce* zmiňovaná v Poznámce 4 v úvodu této kapitoly) ukrýt do citátu, který už jste nejspíše někde viděli: *Dnes je první den zbytku Tvého života.*

[15] Proto to indexování v opačném pořadí.

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \dots, 1$. V takovém případě můžeme využít předchozí schéma, kde položíme

$$r_N(X_N, D_N) = r_{N-1}(X_{N-1}, D_{N-1}) = \cdots = r_2(X_2, D_2) = 0 \quad \text{a} \quad r_1(X_1, D_1) = g\big[y_1(X_1, D_1)\big] = g(X_0).$$

Potom platí

$$g(X_0) = r_N(X_N, D_N) + \cdots + r_1(X_1, D_1),$$

a vlastně tedy chceme vyřešit problém

$$f_N(X_N) = \operatorname{opt}_{D_N,\dots,D_1}\big[r_N(X_N, D_N) + \cdots + r_1(X_1, D_1)\big] = \operatorname{opt}_{D_N,\dots,D_1} \sum_{k=1}^{N} r_k(X_K, D_k)$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \dots, 1$.

Podobným postupem lze řešit také úlohy, kdy funkce g není jenom aditivní. Ovšem abychom mohli aplikovat Bellmanův princip, nemůže být funkce g úplně zcela libovolná. Musí být tzv. *separovatelná*, tj. musí splňovat

$$g\big[r_N(X_N, D_N), \dots, r_1(X_1, D_1)\big] = g_1\Big(r_N(X_N, D_N), g_2\big[r_{N-1}(X_{N-1}, D_{N-1}), \dots, r_1(X_1, D_1)\big]\Big)$$

a zároveň musí splňovat tzv. podmínku *monotonie*:

g_1 je neklesající funkcí proměnné g_2 pro libovolnou hodnotu r_N,

což jsme viděli již v (5.2.5).

Věta 5.2.2

Nechť funkce $g = g(r_2, r_1)$ je neklesající vzhledem k proměnné r_1 pro všechna r_2. Je-li $X_1 = y_2(X_2, D_2)$, pak platí[16]

$$\begin{aligned} f_2(X_2) &:= \max_{D_2, D_1} g\big[r_2(X_2, D_2), r_1(X_1, D_1)\big] = \\ &= \max_{D_2} g\big[r_2(X_2, D_2), \max_{D_1} r_1(y_2(X_2, D_2), D_1)\big]. \end{aligned} \tag{5.2.7}$$

[16] Stejná podmínka platí také při hledání minima.

Důkaz. Označme výraz na pravé straně (5.2.7) jako $\widehat{f_2}(D_2)$. Pak jistě platí $\widehat{f_2}(X_2) \le f_2(X_2)$, neboť postupným maximalizováním nemůžeme získat více než je globální maximum. Vzhledem k monotonii funkce g ve druhé proměnné, platí pro všechna D_1, D_2, že

$$g\big[r_2(X_2, D_2), \max_{D_1} r_1(y_2(X_2, D_2), D_1)\big] \ge g\big[r_2(X_2, D_2), r_1(y_2(X_2, D_2), D_1)\big],$$

z čehož maximalizací obou stran pro všechna D_2 dostaneme také opačnou nerovnost, a tudíž vskutku platí (5.2.7). ■

Podmínka (5.2.7) zaručuje tzv. *dekompozici* a v takovém případě lze optimalizační problém opět rozložit na jednodušší problémy, čímž získáme rekurzivní schéma

$$f_k(X_k) = \max_{D_k} Q_k(X_k, D_k), \quad \text{pro } k = N, \dots, 1,$$

kde

$$Q_k = \begin{cases} r_k(X_k, D_k), & k = 1, \\ r_k(X_k, D_k) \circ f_{k-1}\big(y_k(X_k, D_k)\big), & k = 2, \dots, N, \end{cases}$$

přičemž $\circ$ je operátor „skládání“ (dříve tam bylo znaménko +), protože díky separovatelnosti můžeme $R_N = g(r_N, \dots, r_1)$ zapsat jako $R_N = g_1\big[r_N, g_2(r_{N-1}, \dots, r_1)\big]$ neboli $R_N = r_N \circ R_{N-1}$. Tím dostaneme

$$f_N(X_N) = \max_{D_N}\Big\{r_N(X_N, D_N) \circ \max_{D_{N-1},\dots,D_1}\big[r_{N-1}(X_{N-1}, D_{N-1}) \circ \cdots \circ r_1(X_1, D_1)\big]\Big\}$$

za podmínky $X_{k-1} = y_k(X_k, D_k)$ pro $k = N, \dots, 1$ neboli

$$\begin{aligned} f_N(X_N) &= \max_{D_N,\dots,D_1}\big[r_N(X_N, D_N) \circ r_{N-1}(X_{N-1}, D_{N-1}) \circ \cdots \circ r_1(X_1, D_1)\big] = \\ &= \max_{D_N}\big[r_N(X_N, D_N) \circ f_{N-1}(X_{N-1})\big]. \end{aligned}$$

Příklad 5.2.3

Např. místo $\sum_{k=1}^{N} r_k$ můžeme mít[17] $\Pi_{k=1}^{N} r_k$ nebo $\ln(r_N + \cdots + r_1)$ apod. Zdůrazněme, že symbol $\circ$ nemusí mít v každém kroku tentýž význam jako v případě sčítání nebo násobení, ale např. můžeme mít funkci $r_3 + r_2 \cdot r_1$ nebo $r_3 \cdot r_2 + r_1$. V prvním případě zvolíme $R_2 = r_2 \cdot r_1$, takže $R_3 = r_3 \circ R_2$, což by ale ve druhém případě nezafungovalo, neboť pro $R_2 = r_2 + r_1$ by bylo $R_3 \neq r_3 \circ R_2$. Druhou situaci by však bylo možné vyřešit obrácením pořadí, tj. zvolit $R_2 = r_3 \cdot r_2$, což pak povede k $R_3 = r_1 \circ R_2$. Ponecháme na čtenáři, aby odvodil pro tyto případy odpovídající rekurzivní schéma. Ukázkou funkce nesplňující podmínku separovatelnosti může být např. $r_4 + r_3 \cdot r_2 + r_1$.

[17] Tento problém bychom pomocí logaritmu také mohli převést na předchozí situaci, neboť logaritmus je monotónní a platí $\ln\left(\Pi_{k=1}^{N} r_k\right) = \sum_{k=1}^{N} \ln r_k$.

Nyní si ilustrujeme využití dynamického programování na několika příkladech. První z těchto úloh je *nalezení optimální cesty v grafu*.

Příklad 5.2.4

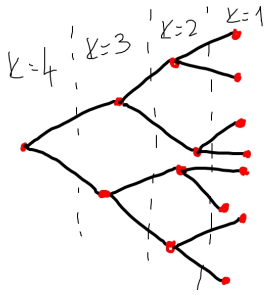
Problém dostavníku:
Video ☞ https://ln.math.muni.cz/pz/opt/dp1_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/dp1_zapisnik

Předchozí příklad můžeme interpretovat také jako *problém kritické cesty*. To je úloha, ve které jednotlivé uzly představují určité činnosti a jsou uspořádány tak, jak musí následovat za sebou (např. při stavbě domu je nutné před instalací oken[18] nejdříve postavit zeď), tj. spojení uzlů pak vyjadřuje časovou souslednost jednotlivých činností. Cílem je nalezení *kritických činností* (tj. těch, jejichž zdržení bude mít za následek zdržení celého projektu)[19] a maximálního nutného času pro dokončení projektu a případně i identifikovat činnosti, které mohou probíhat současně.

[18] Nikoli „windows“ – to by úlohu učinilo neřešitelnou...

[19] V případě *nekritických činností* máme jistou časovou rezervu pro jejich dokončení.

Ve většině úloh řešených pomocí dynamického programování lze nalézt řešení i jinými metodami (viz předchozí příklad). Ovšem využití dynamického programování přináší mnohdy značné zjednodušení (zejména v úlohách variačního počtu). Uvažme například N-krokový rozhodovací proces, kde v každém kroku lze učinit m rozhodnutí. Celkem máme m^N možných variant, jenže při použití dynamického programování máme pouze $m \cdot N$ variant. V situaci ilustrované na Obrázku 5.5 není rozdíl příliš markantní $2^4 = 16$ oproti $2 \cdot 4 = 8$. Jenže např. pro $m = 2$ a $N = 100$ máme $2^{100} \approx 1,267 \cdot 10^{30}$ oproti $2 \cdot 100 = 200$. Na druhou stranu Bellmanův princip *není exaktní matematické tvrzení*, ačkoli je intuitivně přijatelné. Také neexistuje žádná univerzální metoda řešení (na rozdíl např. od lineárního programování a simplexové metody) – každá úloha (nebo typ úloh) má svůj vlastní přístup k řešení a vyžaduje samostatný důkaz potvrzující aplikovatelnost Bellmanova principu.



Obrázek 5.5: Počet všech možností pro $N = 4$ a $m = 2$.

Samozřejmě ne v každé úloze, kterou lze řešit pomocí dynamické programování, musí být dynamika celého procesu způsobena přítomností nějakého časového údaje. Někdy je možné závislost na čase do daného problému implementovat uměle tak, jak to uděláme v následujícím řešení *problému batohu*, s nímž jsme se už několikrát setkali v předchozích kapitolách. To je typický a zároveň nejjednodušší příklad tohoto typu.

Příklad 5.2.5

Problém batohu:
Video ☞ https://ln.math.muni.cz/pz/opt/dp2_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/dp2_zapisnik

Už víme, že problém batohu je velmi speciální případ *alokačního problému*, takže se v následujícím příkladu podíváme, jak lze s pomocí dynamického programování řešit i obecnější úlohy tohoto typu.

Příklad 5.2.6

Alokační problém:
Video ☞ https://ln.math.muni.cz/pz/opt/dp3_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/dp3_zapisnik

Rekurzivní schéma lze pochopitelně využít i při řešení úloh celočíselného programování, což bude opět diskrétní proces. My si toto ilustrujeme v následujícím příkladu s nelineárním celočíselným problémem.

Příklad 5.2.7

Vyřešme úlohu

$$x_1^3 - 3x_1^2 - 5x_2^2 + x_2^3 \to \min$$
$$x_1 + 2x_2 \le 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_1 \in \mathbb{Z} \quad \& \quad x_2 \in \mathbb{Z}.$$

Řešení. Chceme určit optimální hodnotu dvojice proměnných x_1 a x_2, které mohou vzhledem k daným omezením nabývat pouze konečně mnoha hodnot. Rozdělíme si proto celý proces do dvou kroků, ve kterých budeme postupně analyzovat přínos jednotlivých proměnných k hodnotě účelové funkce. Ze zadání také vidíme, že levá strana dané nerovnosti může nabývat pouze 0, 1, 2, 3 a 4. Označme proto jako vstupní proměnnou pravou stranu dané nerovnosti pro zvolená x_1 a x_2, tj. $X_2 = x_1 + 2x_2$. Rozhodovací proměnné budou odpovídat hodnotám x_2 a x_1. Naším cílem je de facto „rozdělit“ X_2 do dvou hodnot tak, abychom minimalizovali účelovou funkci, přičemž platí $X_1 = X_2 - 2x_2$ a $X_0 = X_1 - x_1$. Začneme s prvním krokem, tj. už jsme nějak určili x_2 a chceme najít vhodné x_1, což je zapsáno v Tabulce 5.1.

D_1 / X_1	$f_1(X_1) = x_1^3 - 3x_1^2$					$D_1^* = x_1$	$f_1(X_1)$	$X_0 = X_1 - D_1$
	0	1	2	3	4			
0	0	✖	✖	✖	✖	0	0	0
1	0	−2	✖	✖	✖	1	−2	0
2	0	−2	−4	✖	✖	2	−4	0
3	0	−2	−4	0	✖	2	−4	1
4	0	−2	−4	0	16	2	−4	2

Tabulka 5.1: První krok v řešení Příkladu 5.2.7.

Nyní můžeme pokračovat druhým krokem, který je v Tabulce 5.2 na další straně. Vidíme, že hledaným řešením je volba $x_2^* = 2$ a $x_1^* = 0$ s hodnotou minima −12.
Mohli bychom v Tabulce 5.2 brát rovnou pouze $X_2 = 4$? Intuitivně bychom možná mohli očekávat, že čím větší hodnotu X_2 bude mít, tím více budu mezi x_1 a x_2 „rozdělovat“, a tudíž tím menší funkční hodnotu získám. V takovém případě bychom mohli řádky pro $X_1 = 1$ a $X_1 = 3$ v Tabulce 5.1 proškrtnout, jelikož $X_1 = 4 - 2x_2$ by muselo být jistě sudé číslo. Každopádně ale tato vlastnost nemusí platit obecně a závisí na dané funkci a omezení, neboť hledané minimum se nemusí nutně realizovat v bodě, pro který nastane v daném omezení rovnost.
Ještě si všimněme, že postupným dosazováním z X_1 a X_0 získáme $x_1 + 2x_2 = X_1 - X_0 + X_2 - X_1 = X_2 - X_0 \le 4$, tj. $X_0 \ge X_2 - 4$. To znamená, že X_0 může nabývat pouze hodnot $-4, -3, -2, -1$ a 0. Nicméně tuto informaci jsme k řešení nepotřebovali (podobně jako tomu je v Příkladu 5.2.9 níže) na rozdíl např. od Příkladu 5.2.10 níže. Ponecháme na čtenáři, aby si promyslel, v čem je mezi těmito úlohami rozdíl.

X_2 \ D_2	$f_2(X_2) = -5x_2^2 + x_2^3 + f_1(X_1)$			$D_2^* = x_2$	$f_2(X_2)$	$X_1 = X_2 - 2D_2$
	0	1	2			
0	0	✖	✖	0	0	0
1	−2	✖	✖	0	−2	1
2	−4	−4	✖	0, 1	−4	0
3	−4	−6	✖	1	−6	1
4	−4	−8	−12	2	−12	0

Tabulka 5.2: Druhý krok v řešení Příkladu 5.2.7.

Všechny dosavadní řešené problémy byly ukázkou diskrétních rozhodovacích procesů, ve kterých jsme mohli uvážit „všechny“ dostupné možnosti pomocí jednoduchých výčtových tabulek. Toto ale již není možné v případě *spojitého* procesu a my si v následujícím příkladu ukážeme řešení jednoho z nich.

Příklad 5.2.8

Spojitý rozhodovací proces:
Video ☞ https://ln.math.muni.cz/pz/opt/dp4_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/dp4_zapisnik

Dynamické programování lze využít i při řešení úloh lineárního, kvadratického či obecně konvexního (nebo konkávního) programování s omezeními ve tvaru nerovností. V takovém případě odpovídá stavová veličina X_N hodnotě pravé strany příslušného omezení a rozhodovací proměnné splývají s proměnnými samotné úlohy. Čím více bude omezení, tím složitější celý proces (pochopitelně) bude – stavové veličiny budou totiž m-rozměrné vektory v případě m-tice omezení. Toto si ilustrujeme v následujících příkladech.

Příklad 5.2.9

S pomocí dynamického programování vyřešme úlohu

$$x_1^2 + x_2^2 + x_3^3 \to \min$$
$$x_1 + x_2 + x_3 \geq 3 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0 \quad \& \quad x_3 \geq 0.$$

Řešení. Označme $X_3 := x_1 + x_2 + x_3$. Rekurentní schéma pak vypadá následovně

$$f_1(X_1) := \min_{x_1 = X_1} x_1^2$$
$$f_2(X_2) := \min_{0 \leq x_2 \leq X_2} \{x_2^2 + f_1(X_2 - x_2)\}$$
$$f_3(X_3) := \min_{0 \leq x_3 \leq X_3} \{x_3^2 + f_2(X_3 - x_3)\} \quad \& \quad X_3 \geq 3.$$

Začneme proto s řešením od první úlohy, ze které očividně plyne

$$f_1(X_1) = X_1^2.$$

Dosazením do druhého kroku dostaneme

$$f_2(X_2) = \min_{0 \leq x_2 \leq X_2} \{x_2^2 + (X_2 - x_2)^2\}.$$

Zderivováním funkce na pravé straně obdržíme $2x_2 - 2(X_2 - x_2)$, takže stacionárním budeme je $x_2^* = X_2/2$, a tudíž

$$f_2(X_2) = X_2^2/2,$$

což po dosazení do poslední úlohy dává

$$f_3(X_3) = \min_{0 \le x_3 \le X_3} \{x_3^2 + (X_3 - x_3)^2/2\}.$$

Derivace funkce na pravé straně je $2x_3 - (X_3 - x_3)$, takže stacionárním bodem je $x_3^* = X_3/3$, tudíž $f_3(X_3) = X^2/3$. Máme proto úlohu

$$X_3^2 \to \min \quad \& \quad X_3 \ge 3,$$

jejímž řešením je očividně $X_3^* = 3$. Proto zpětným dosazením dostaneme postupně

$$x_3^* = 1 \quad \& \quad X_2^* = X_3^* - x_3^* = 2 \quad \& \quad x_2^* = X_2^*/2 = 1 \quad \& \quad X_1^* = X_2^* - x_2^* = 1 \quad \& \quad x_1^* = 1,$$

tj. hledaným řešením je $x_1^* = x_2^* = x_3^* = 1$.

Teď zkusíme podobný příklad trochu méně formálně.

Příklad 5.2.10

S pomocí dynamického programování vyřešme úlohu

$$x_1^4 + 2x_2^2 \to \min$$
$$x_1^2 + x_2^2 \ge 2.$$

Řešení. Označíme-li jako $X_2 \ge 2$ možnou hodnotu pravé strany omezení pro všechny možné hodnoty x_1 a x_2, pak $X_1 = X_2 - x_2^2$, tj. X_1 vyjadřuje rozdíl mezi pravou a levou stranou nerovnosti s daným x_2. Podobně budeme mít $X_0 = X_1 - x_1^2$. Potom platí $x_1^2 + x_2^2 = X_1 - X_0 + X_2 - X_1 = X_2 - X_0 \ge 2 - X_0$, takže $X_0 = X_1 - x_1^2 \le 0$, neboť výraz na levé straně musí být roven alespoň 2. Proto v prvním kroku máme úlohu

$$x_1^4 \to \min \quad \& \quad x_1^2 \ge X_1,$$

což má řešení pro $x_1 = \pm\sqrt{X_1}$ s hodnotou $f_1(X_1) = X_1^2$ pro $X_1 \ge 0$, zatímco pro $X_1 \le 0$ bude řešením volba $x_1 = 0$ s $f_1(X_1) = 0$.
Protože $X_1 = X_2 - x_2^2$ a my nemáme žádné omezení na hodnotu X_1, může x_2 nabývat libovolné hodnoty. V tomto kroku pak hledáme řešení úlohy

$$\min_{x_2 \in \mathbb{R}} \left[2x_2^2 + f_1(X_1)\right] = \min\left\{ \min_{\substack{x_2 \in \mathbb{R} \\ X_1 = X_2 - x_2^2 \le 0}} 2x_2^2, \min_{\substack{x_2 \in \mathbb{R} \\ X_1 = X_2 - x_2^2 \ge 0}} 2x_2^2 + X_1^2 \right\} =$$
$$= \min\left\{ \min_{\substack{x_2 \in \mathbb{R} \\ X_1 = X_2 - x_2^2 \le 0}} 2x_2^2, \min_{\substack{x_2 \in \mathbb{R} \\ X_1 = X_2 - x_2^2 \ge 0}} 2x_2^2 + (X_2 - x_2^2)^2 \right\}.$$

V prvním případě je $X_2 \le x_2^2$ neboli $x_2 \ge \sqrt{X_2}$ a $x_2 \le -\sqrt{X_2}$ a hledáme minimum z $2x_2^2$, které se tudíž realizuje pro $x_2 = \pm\sqrt{X_2}$ s hodnotou $f_2(X_2) = 2X_2$. Jelikož $X_2 \ge 2$ je „libovolné“ a my hledáme $\min_{X_2 \ge 2} f_2(X_2)$, nastane toto minimum pro $X_2 = 2$ s hodnotou $f_2(X_2) = 4$. Ve druhém případě máme $X_2 \ge x_2^2$ neboli $-\sqrt{X_2} \le x_2 \le \sqrt{X_2}$ a hledáme minimum z $x_2^4 - 2(X_2 - 1)x_2^2 + X_2^2$. Stacionární body této funkce jsou $x_2 = 0$ a $x_2 = \pm\sqrt{X_2 - 1}$, z čehož pomocí druhé derivace poznáme, že první bod je lokálním maximem a druhé dva body jsou lokálními/globálními minimy se stejnou hodnotou $f_2(X_2) = 2X_2 - 1$. Minimalizací vzhledem k $X_2 \ge 2$ zjistíme, že $f_2(X_2) = 3$ pro $X_2 = 2$. Toto je menší hodnota než v prvním případě, tudíž zpětným dosazením $X_2 = 2$ do předchozích výpočtů ve druhém kroku nalezneme řešení, kterým je čtveřice bodů $x^* = [\pm 1, \pm 1]$ se všemi možnými kombinacemi znamének a hodnotou minima 3.

Nyní se podíváme na obecnější úlohu[20] z nelineárního programování.

[20] Tuto úlohu našel autor označenou jako „Huygensův problém“, avšak bez jakýchkoli podrobností, které se nepodařilo najít ani jinde (snaha o nalezení jiného zdroje, kde by se úloha s tímto označením vyskytovala, byla také neúspěšná).

Příklad 5.2.11

Pro dané $n \in \mathbb{N}$ a konstanty $a, b > 0$ vyřešme úlohu

$$\frac{x_1 x_2 \cdots x_n}{(a+x_1)(x_1+x_2)\cdots(x_{n-1}+x_n)(x_n+b)} \to \max \quad \& \quad 0 < a \le x_1 \le x_2 \le \cdots \le x_n \le b.$$

Řešení. Označme funkci popisující hledané řešení v závislosti na parametrech a, b a hodnotě n jako

$$f_n(a,b) \coloneqq \max_{a \le x_1 \le x_2 \le \cdots \le x_n \le b} \frac{x_1 x_2 \cdots x_n}{(a+x_1)(x_1+x_2)\cdots(x_{n-1}+x_n)(x_n+b)}$$

Pak není těžké ověřit, že platí

$$\begin{aligned} f_n(a,b) &= \max_{a \le x_n \le b} \frac{x_n}{x_n+b} \cdot \max_{a \le x_1 \le x_2 \le \cdots \le x_n} \frac{x_1 x_2 \cdots x_{n-1}}{(a+x_1)(x_1+x_2)\cdots(x_{n-1}+x_n)} = \\ &= \max_{a \le x_n \le b} \left\{ \frac{x_n}{x_n+b} \cdot f_{n-1}(a,x_n) \right\}. \end{aligned}$$

Můžeme proto úlohu začít řešit rekurentně a v prvním kroku najít

$$f_1(a,x_2) = \max_{a \le x_1 \le x_2} \frac{x_1}{(a+x_1)(x_1+x_2)}.$$

Derivováním vzhledem k x_1 dostaneme

$$\frac{\partial}{\partial x_1} f_1(a,x_2) = \frac{(a+x_1)(x_1+x_2) - x_1(a+2x_1+x_2)}{(a+x_1)^2(x_1+x_2)^2} = \frac{-x_1^2 + a x_2}{(a+x_1)^2(x_1+x_2)^2},$$

takže stacionárním bodem a současně globálním maximem je pouze $x_1 = \sqrt{a x_2}$ s hodnotou

$$\begin{aligned} f_1(a,x_2) &= \frac{\sqrt{a x_2}}{(a+\sqrt{a x_2})(\sqrt{a x_2}+x_2)} = \frac{\sqrt{a x_2}}{\sqrt{a}(\sqrt{a}+\sqrt{x_2})\sqrt{x_2}(\sqrt{a}+\sqrt{x_2})} = \\ &= \frac{\sqrt{a x_2}}{\sqrt{a x_2}(\sqrt{a}+\sqrt{x_2})^2} = \frac{1}{(\sqrt{a}+\sqrt{x_2})^2} \end{aligned}$$

Zdůrazněme, že pro $1 < k < n$ hledáme

$$\begin{aligned} f_k(a,x_{k+1}) &= \max_{a \le x_k \le x_{k+1}} \frac{x_k}{x_k+x_{k+1}} \cdot \max_{a \le x_1 \le x_2 \le \cdots \le x_k} \frac{x_1 x_2 \cdots x_{k-1}}{(a+x_1)(x_1+x_2)\cdots(x_{k-1}+x_k)} = \\ &= \max_{a \le x_k \le x_{k+1}} \left\{ \frac{x_k}{x_k+x_{k+1}} \cdot f_{k-1}(a,x_k) \right\} \end{aligned}$$

Proto ve druhém kroku řešíme problém

$$f_2(a,x_3) = \max_{a \le x_2 \le x_3} \left\{ \frac{x_2}{x_2+x_3} \cdot F_1(a,x_n) \right\} = \max_{a \le x_2 \le x_3} \left\{ \frac{x_2}{(x_2+x_3)(\sqrt{a}+\sqrt{x_2})^2} \right\}.$$

Derivováním (jakožto součinu dvou podílů) vzhledem k x_2 odtud dostaneme

$$\begin{aligned} \frac{\partial}{\partial x_2} f_2(a,x_3) &= \frac{x_2+x_3-x_2}{(x_2+x_3)^2} \cdot \frac{1}{(\sqrt{a}+\sqrt{x_2})^2} + \frac{x_2}{x_2+x_3} \cdot \frac{-2\frac{1}{2\sqrt{x_2}}}{(\sqrt{a}+\sqrt{x_2})^3} = \\ &= \frac{1}{(x_2+x_3)(\sqrt{a}+\sqrt{x_2})^2} \left[\frac{x_3}{x_2+x_3} - \frac{\sqrt{x_2}}{\sqrt{a}+\sqrt{x_2}} \right] = \\ &= \frac{1}{(x_2+x_3)(\sqrt{a}+\sqrt{x_2})^2} \cdot \frac{x_3\sqrt{a} - x_2\sqrt{x_2}}{(x_2+x_3)(\sqrt{a}+\sqrt{x_2})} \end{aligned}$$

se stacionárním bodem (a současně globálním maximem) v $x_2 = \sqrt[3]{a x_3^2}$, tudíž

$$\begin{aligned} f_2(a,x_3) &= \frac{\sqrt[3]{a x_3^2}}{\left(\sqrt[3]{a x_3^2}+x_3\right)\left(\sqrt{a}+\sqrt[3]{\sqrt{a}\,x_3}\right)^2} = \frac{a^{1/3} x_3^{2/3}}{x_3^{2/3}\left(a^{1/3}+x_3^{1/3}\right)\left[a^{1/6}\left(a^{1/3}+x_3^{1/3}\right)\right]^2} = \\ &= \frac{1}{\left(\sqrt[3]{a}+\sqrt[3]{x_3}\right)^3}. \end{aligned}$$

Dalším pokračováním v těchto výpočet až po $k = n-1$ bychom zjistili, že

$$f_{n-1}(a, x_n) = \frac{1}{\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^n} \quad \& \quad x_{n-1} = \sqrt[n]{a\, x_n^{n-1}}.$$

V posledním kroku pak budeme mít

$$f_n(a,b) = \max_{a \le x_n \le b} \left\{ \frac{x_n}{x_n + b} \cdot F_{n-1}(a, x_n) \right\} = \max_{a \le x_n \le b} \frac{x_n}{x_n + b} \cdot \frac{1}{\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^n}.$$

Derivováním vzhledem k x_n dostaneme

$$\begin{aligned}
\frac{\partial}{\partial x_n} F_n(a,b) &= \frac{x_n + b - x_n}{(x_n+b)^2} \cdot \frac{1}{\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^n} + \frac{x_n}{x_n+b} \cdot \frac{-n\,\frac{1}{n}\, x_n^{1/n-1}}{\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^{n+1}} = \\
&= \frac{1}{(x_n+b)\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^n} \left[\frac{b}{x_n+b} - \frac{\sqrt[n]{x_n}}{\sqrt[n]{a} + \sqrt[n]{x_n}} \right] = \\
&= \frac{1}{(x_n+b)\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)^n} \cdot \frac{b\sqrt[n]{a} - x_n \sqrt[n]{x_n}}{(x_n+b)\left(\sqrt[n]{a} + \sqrt[n]{x_n}\right)}.
\end{aligned}$$

Stacionárním bodem (a současně globálním maximem) proto je $x_n = \left(b\sqrt[n]{a}\right)^{n/(n+1)} = \sqrt[n+1]{a}\,\sqrt[n+1]{b^n}$ s hodnotou

$$\begin{aligned}
f_n(a,b) &= \frac{\sqrt[n+1]{a}\,\sqrt[n+1]{b^n}}{\left(\sqrt[n+1]{a}\,\sqrt[n+1]{b^n} + b\right)} \cdot \frac{1}{\left(\sqrt[n]{a} + \sqrt[n(n+1)]{a}\,\sqrt[n+1]{b}\right)^n} = \\
&= \frac{\sqrt[n+1]{a}\,\sqrt[n+1]{b^n}}{\sqrt[n+1]{b^n}\left(\sqrt[n+1]{a} + b^{1-n/(n+1)}\right)\left[\sqrt[n(n+1)]{a}\left(a^{1/n-1/(n(n+1))} + \sqrt[n+1]{b}\right)\right]^n} = \\
&= \frac{1}{\left(\sqrt[n+1]{a} + \sqrt[n+1]{b}\right)\left(\sqrt[n+1]{a} + \sqrt[n+1]{b}\right)^n} = \frac{1}{\left(\sqrt[n+1]{a} + \sqrt[n+1]{b}\right)^{n+1}}.
\end{aligned}$$

Zpětným dosazováním bychom pak postupně dopočítali, že

$$x_n^* = \sqrt[n+1]{a}\,\sqrt[n+1]{b^n} \quad \& \quad x_{n-1}^* = \sqrt[n+1]{a^2\, b^{n-1}} \quad \& \quad x_{n-2}^* = \sqrt[n+1]{a^3\, b^{n-2}} \quad \& \quad \ldots \quad \& \quad x_1^* = \sqrt[n+1]{a^n\, b}.$$

Tentokrát sice budeme mít „jen“ úlohu lineárního programování, ale zato se dvěma nerovnostními omezeními a celá situace se trochu zkomplikuje. Z teoretického pohledu v tom de facto žádný rozdíl není, jen je potřeba uvažovat o stavových proměnných X jako o m-rozměrných vektorech a pracovat s jejich složkami. Ovšem z pohledu výpočetní efektivity to už bývá značná komplikace. Místo zvažování přípustných hodnot jedné stavové veličiny, musíme brát do úvahy všechny možné kombinace hodnot některých stavových proměnných. Počet takovýchto kombinací se tak začne s rostoucím počtem omezení zvětšovat a mívá tendenci při zavádění dalších stavových proměnných rychle „vybuchnout“, protože se začne přibližovat součinu počtu všech přípustných hodnot jednotlivých veličin[21]. Tento fenomén dostal výstižný název *prokletí dimenzionality*[22].

[21] Zejména ve chvíli, jsou-li veličiny celočíselné. To už se začneme blížit k použití hrubé síly.

[22] Angl. *Curse of dimensionality*. Termín pochází od samotného Bellmana, viz [45, str. ix].

Příklad 5.2.12

Vyřešme úlohu

$$5x_1 + 4x_2 \to \max/\min \quad \& \quad -x_1 + 3x_2 \le 3 \quad \& \quad 2x_1 + x_2 \le 8 \quad \& \quad x_1 \ge \quad \& \quad x_2 \ge 0.$$

Řešení. Úlohu budeme řešit ve dvou krocích, ve kterých postupně určíme hodnotu x_1 a x_2. Stavová veličina X tentokrát bude dvourozměrná, tj. $X = (X_1, X_2)$, kde X_1 a X_2 budou určovat „kolik zbývá k rovnosti v prvním/druhém omezení“, tj. půjde o rozdíl mezi levou a pravou stranou prvního/druhého omezení (v klasické úloze lineární programování by šlo vlastně o množství dostupných zdrojů). Na začátku celého procesu, kdy jsme ještě nezvolili žádnou proměnnou, budeme mít $X^{[2]} = (3, 8)$. Ve druhém

kroku (nezapomeňme na zpětné indexování) zvolíme x_2 a hodnota stavové proměnné se změní na $X^{[1]} = (3-3x_2, 8-x_2)$, pro kterou budeme hledat vhodné x_1, které bude jakýmsi způsobem záviset na x_2. Podle rekurzivního schématu začneme s řešením od konce, tj. nejdříve se zaměříme na určení vhodného x_1.

Začneme s hledáním maxima. V souladu s předchozím značením a rekurzivním schématem bude funkce

$$f_1(X_1^{[1]}, X_2^{[1]}, x_1) = 5x_1$$

popisovat příspěvek proměnné x_1 k hodnotě účelové funkce, pokud vstupní stavová veličina má hodnotu $X^{[1]} = (X_1^{[1]}, X_2^{[1]})$. Podobně bude hodnota

$$f_2(3, 8, x_2) = 4x_2 + \max_{\substack{-x_1 \le 3-3x_2 \\ 2x_1 \le 8-x_2 \\ x_1 \ge 0}} \{5x_1\}$$

popisovat příspěvek proměnných x_1 a x_2 k hodnotě účelové funkce, pokud má vstupní stavová veličina hodnotu $X^{[2]} = (3, 8)$ a my nejdříve zvolíme x_2, což je následováno optimální volbou x_1. Pro $k = 1, 2$ proto označme

$$f_k^*(X_1^{[k]}, X_2^{[k]}) = \max_{x_k} f_k(X_1^{[k]}, X_2^{[k]}, x_k),$$

kde maximum bereme přes všechny přípustné hodnoty proměnné x_k. Proto

$$f_1^*(X_1^{[1]}, X_2^{[1]}) = \max_{\substack{-x_1 \le X_1^{[1]} \\ 2x_1 \le X_2^{[1]} \\ x_1 \ge 0}} \{5x_1\},$$

z čehož plyne

$$f_2(3, 8, x_2) = 4x_2 + f_1^*(3-3x_2, 8-x_2),$$

a tudíž

$$f_2^*(3, 8) = \max_{\substack{x_2 \le 8 \\ x_2 \ge 0}} \{4x_2 + f_1^*(3-3x_2, 8-x_2)\},$$

kde není podmínka $3x_2 \le 3$, kterou bychom analogickou úvahou jako v ostatních případech dostali z prvního omezení, což však není pravda, neboť x_2 může nabývat i větších hodnot díky členu $-x_1$ na levé straně.

Můžeme se pustit do prvního kroku. Zde máme trojici podmínek $-x_1 \le X_1^{[1]}$, $2x_1 \le X_2^{[1]}$ a $x_1 \ge 0$. Všimněme si, že přípustný bod bude existovat pouze, pokud $X_2^{[1]} \ge 0$ a $-X_1^{[1]} \le X_2^{[1]}/2$, v kterémžto případě získáme maximální hodnotu

$$f_1^*(X_1, X_2) = 5X_2^{[1]}/2, \quad \text{která se realizuje v bodě } x_1^* = X_2^{[1]}/2.$$

Potom ve druhém kroku řešíme problém

$$f_2^*(3, 8) = \max_{\substack{x_2 \le 8 \\ x_2 \ge 0}} \{4x_2 + f_1^*(3-3x_2, 8-x_2)\},$$

ke kterému musíme přidat podmínky $X_2^{[1]} \ge 0$ a $-X_1^{[1]} \le X_2^{[1]}/2$ neboli $8 - x_2 \ge 0$ a $-3 + 3x_2 \le (8-x_2)/2$ z prvního kroku, tj. $7x_2/2 \le 7$. První podmínka je již zahrnutá v daných omezeních, zatímco druhá podmínka dává dodatečné omezení $x_2 \le 2$. Proto s dosazením výsledku z prvního kroku dostáváme úlohu

$$f_2^*(3, 8) = \max_{\substack{x_2 \le 8 \\ x_2 \ge 0 \\ x_2 \le 2}} \left\{ 4x_2 + \frac{5(8-x_2)}{2} \right\} = \max_{\substack{x_2 \le 8 \\ x_2 \ge 0 \\ x_2 \le 2}} \left\{ \frac{3x_2}{2} + 20 \right\},$$

jejímž řešením je $x_2^* = 2$ s hodnotou $f_2^*(3, 8) = 23$. Zpět dosazením ještě dopočítáme $x_1^* = (8-2)/2 = 3$, tj. řešením dané úlohy je $x^* = [3, 2]$ s hodnotou maxima $f^* = 23$.

Podobným způsobem můžeme najít i řešení analogické úlohy, ve které chceme účelovou funkci minimalizovat. V prvním kroku opět máme dodatečné podmínky $X_2^{[1]} \geq 0$ a $-X_1^{[1]} \leq X_2^{[1]}/2$, za kterých hledáme minimální hodnotu

$$f_1^*(X_1, X_2) = \min_{\substack{-x_1 \leq X_1^{[1]} \\ 2x_1 \leq X_2^{[1]} \\ x_1 \geq 0}} \{5x_1\}.$$

Řešení se tentokrát realizuje v bodě $x_1 = \max\{0, -X_1^{[1]}\}$ s hodnotou $f_1^*(X_1, X_2) = -5\max\{0, -X_1^{[1]}\}$. Odtud ve druhém kroku opět získáme

$$f_2^*(3,8) = \min_{\substack{x_2 \leq 8 \\ x_2 \geq 0 \\ x_2 \leq 2}} \left\{ 4x_2 + f_1^*(3 - 3x_2, 8 - x_2) \right\},$$

kde ale nyní musíme rozlišit dvě situace. Buď $-X_1^{[1]} \geq 0$ neboli $-(3 - 3x_2) \geq 0$, tj. $x_2 \geq 1$, což nás přivede k hledání minima

$$\min_{\substack{x_2 \leq 8 \\ x_2 \geq 0 \\ x_2 \leq 2 \\ x_2 \geq 1}} \left\{ 4x_2 - 5(3 - 3x_2) \right\} = \min_{\substack{x_2 \leq 8 \\ x_2 \geq 0 \\ x_2 \leq 2 \\ x_2 \geq 1}} \left\{ 19x_2 - 15 \right\} = 4,$$

které se realizuje v bodě $x_2 = 1$ s odpovídajícím $x_1 = \max\{0, 0\} = 0$. Druhou možností je $-X_1^{[1]} \leq 0$ neboli $-(3 - 3x_2) \leq 0$, tj. $x_2 \leq 1$, což nás přivede k hledání minima

$$\min_{\substack{x_2 \leq 8 \\ x_2 \geq 0 \\ x_2 \leq 2 \\ x_2 \leq 1}} \left\{ 4x_2 + 0 \right\} = 0,$$

které nastane v bodě $x_2 = 0$, a tudíž zpětným dosazením $x_1 = \max\{0, -3\} = 0$. Vidíme, že menší funkční hodnotu obdržíme ve druhém případě, takže řešením minimalizační úlohy je bod $x^* = [0, 0]$ s hodnotou $f^* = 0$.

Na závěr této podkapitoly si ukážeme ještě jeden příklad s reálným podtextem. Může být až s podivem, kolik inspirace matematikům přináší jídlo – tentokrát půjde o narozeninový dort. Obvykle se spory vedou o to, jak jej spravedlivě nakrájet, ale nám teď půjde o to, jak jej sníst, abychom z toho měli co největší požitek, tj. budeme řešit problém *jak (ne)sníst celý dort najednou*[23], abychom si jej mohli vychutnávat co nejdéle a nejlépe.

[23] Angl. *Cake eating problem.*

Příklad 5.2.13

Představme si, že máme jen sami pro sebe celý dort. Můžeme ho samozřejmě hned sníst, ovšem také můžeme teď sníst jenom kousek a o něco později další kousek a pak další. Ale jak by měly být ty kousky velké, abychom z toho měli ten správný požitek? Stojíme tak před otázkou vhodné volby mezi současností a budoucností: jak takový dort sníst, abychom maximalizovali svůj užitek?
Označme velikost dortu jako V a rozdělme časové období ke konzumaci dortu na $T+1$ částí. Pak naše pojídání dortu můžeme reprezentovat pomocí vektoru

$$c = \begin{pmatrix} c_0 & c_1 & \cdots & c_T \end{pmatrix}^\top,$$

jehož jednotlivé složky musí splňovat

$$\sum_{t=0}^{T} c_t = V \quad \& \quad c_0 \geq \quad \& \quad \cdots \quad \& \quad c_T \geq 0.$$

[24] Angl. *policy vector.*

Vektor c se nazývá *vektor rozhodnutí*[24] a popisuje, jak velké části dortu budeme jíst v jednotlivých obdobích. Užitek ze snědení kousku dortu je reprezentováno užitkovou

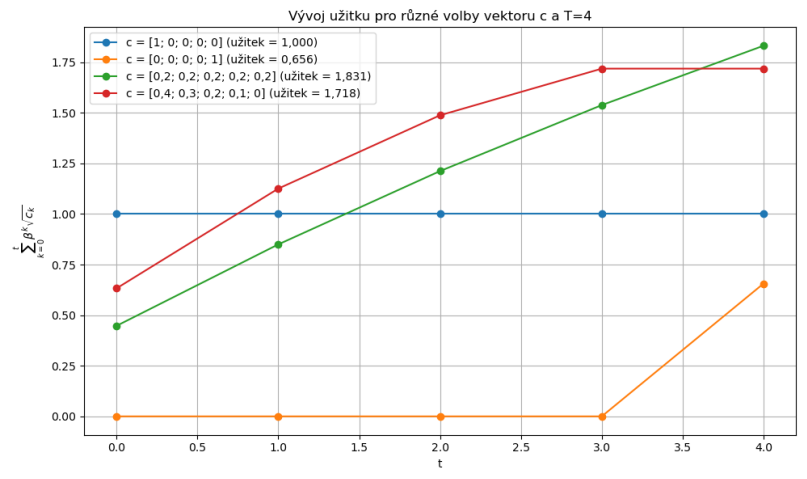
funkcí, tj. konzumací části o velikosti c_t dosáhneme užitku $u(c_t)$. Zároveň můžeme definovat $V_{t+1} := V_t - c_t$ k vyjádření aktuální velikosti dortu na začátku $(t+1)$-ho časového okamžiku s počáteční hodnotou $V_0 := V$ a číslem V_{T+1} vyjadřujícím velikost nezkonzumovaného zbytku dortu. Samozřejmě i zde musí platit $V_0, \dots, V_{T+1} \geq 0$.
V našem rozhodování přirozeně hraje důležitou roli čas, neboť se obvykle rozhodujeme mezi mírou aktuální spotřeby a budoucími úspory. Např. nákup nějakého zboží v tuto chvíli nám může přinést nějaký užitek, zatímco při pozdějším nákupu se „připravíme“ o jisté časové období, ve kterém jsme si mohli tento užitek prožívat, ale zase jsme mohli získat nějaké úroky (i když tohle nám asi nejspíš pokazí inflace…). V případě dortu je ale situace trochu odlišná. Dort se v průběhu času nezvětšuje a nezlepšuje, ale naopak okorává, takže čím dříve jej sníme, tím větší z toho budeme mít užitek. Současně však asi nejsme schopni sníst celý dort naráz, nebo pokud ano, mohlo by to náš celkový užitek dost snížit. Zhoršování kvality dortu můžeme do našeho modelu vnést pomocí násobení tzv. *diskontním faktorem*, což je číslo $\beta \in (0,1)$. Např. pokud v čase $t = 0$ sníme díl o velikosti c_0 a v čase $t = 1$ sníme díl téže velikosti $c_1 = c_0$, pak náš užitek v čase $t = 0$ bude větší než užitek v čase $t = 1$, tj.

$$u(c_0) > \beta\, u(c_1).$$

Proto celková užitková funkce pro naši konzumaci dortu je

$$\sum_{t=0}^{T} \beta^t u(c_t) \tag{5.2.8}$$

a ukázky vývoje užitku pro čtveřici různých strategií v případě $V = 1$, $\beta = 0{,}9$ a $u(c_t) = \sqrt{c_t}$ můžeme vidět na Obrázku 5.6.



Obrázek 5.6: Ukázky vývoje užitku pro čtveřici různých strategií z Příkladu 5.2.13. V tomto případě vychází nejlépe celý dort jíst po stejně velkých kouscích.

Postupným dosazením rovnosti $c_t = V_t - V_{t+1}$ do účelové funkce (5.2.8) dostaneme

$$\begin{aligned} U_{T+1}(V_0) &:= \max_{c_0,\dots,c_T} \left\{u(c_0) + \beta\, u(c_1) + \beta^2 u(c_2) + \dots + \beta^T u(c_T)\right\} = \\ &= \max\left\{u(V_0 - V_1) + \beta\, u(V_1 - V_2) + \beta^2 u(V_2 - V_3) + \dots + \beta^T u(V_T - V_{T+1})\right\}, \end{aligned} \tag{5.2.9}$$

což nás s využitím dekompozice a Bellmanova principu optimality přivede k „jednodušší“ úloze

$$\max_{c_0} \left\{u(c_0) + \beta\, U_T(V_1)\right\} \quad \& \quad V_1 = V_0 - c_0 \tag{5.2.10}$$

pro dané V_0, kde $U_T(V_1)$ představuje optimální užitek, máme-li k dispozici jen T časových okamžiků a z původního dortu jen část o velikosti V_1. Místo určení celého vektoru

c nám teď stačí určit pouze c_0, čímž bude dáno V_1 a náš problém se redukuje na výpočet $U_T(V_1)$. Je-li funkce $U_T(V_1)$ diferencovatelná vzhledem k c_0, potřebujeme k nalezení stacionárního bodu účelové funkce v (5.2.10) vyřešit rovnici

$$u'(c_0) = \beta U_T'(V_1),$$

což lze ekonomicky interpretovat jako to, že mezní užitek z malého snížení spotřeby v čase $t = 0$ je určen hodnotou derivace užitkové funkce $U_T(V_1)$. Současně z pravé strany (5.2.9) derivováním vzhledem k V_t dostaneme

$$-\beta^{t-1} u'(V_{t-1} - V_t) + \beta^t u'(V_t - V_{t+1}),$$

což bude rovno nule právě tehdy, když

$$u'(V_{t-1} - V_t) = \frac{\beta^t}{\beta^{t-1}} u'(V_t - V_{t+1}) \quad \text{neboli} \quad u'(c_{t-1}) = \beta u'(c_t). \tag{5.2.11}$$

Druhá rovnice v (5.2.11) bývá označována jako *Eulerova rovnice* a spojuje mezní užitek z konzumace v časovém okamžiku $t-1$ a v budoucím časovém okamžiku t. Tato rovnice jistým způsobem propojuje dva po sobě jdoucí časové okamžiky a v hledané optimální strategii musí být tato rovnost splněna pro všechna $t = 1, \ldots, T+1$.
Bude-li nyní užitková funkce rovna $u(c_t) = \ln c_t$, pak se Eulerova rovnice redukuje na $1/c_{t-1} = \beta/c_t$ neboli

$$c_{t-1} = \frac{c_t}{\beta} \tag{5.2.12}$$

a můžeme se pustit do rekurzivního řešení našeho problému. Nekonzumovaná část dortu nám zjevně žádný další užitek nepřinese, proto se lze celkem přirozeně omezit na situaci $V_{T+1} = 0$. Pak $c_T = V_T$, což po dosazení do rovnice (5.2.12) dává

$$c_{T-1} = \frac{\overbrace{V_T}^{=V_{T-1}-c_{T-1}}}{\beta} \quad \text{neboli} \quad c_{T-1} = \frac{1}{1+\beta} V_{T-1},$$

což spojuje optimální velikost snědeného dortu v časovém okamžiku $T-1$ a velikost dortu, která je v tomto okamžiku k dispozici. Dále využijeme rovnost $c_{T-2} = \frac{c_{T-1}}{\beta}$, která po dosazení do předchozího vyjádření dává

$$c_{T-2} = \frac{1}{\beta+\beta^2} \overbrace{V_{T-1}}^{=V_{T-2}-c_{T-2}} \quad \text{neboli} \quad c_{T-2} = \frac{1}{1+\beta+\beta^2} V_{T-2}.$$

Dalším pokračováním v tomto postupu bychom nalezli

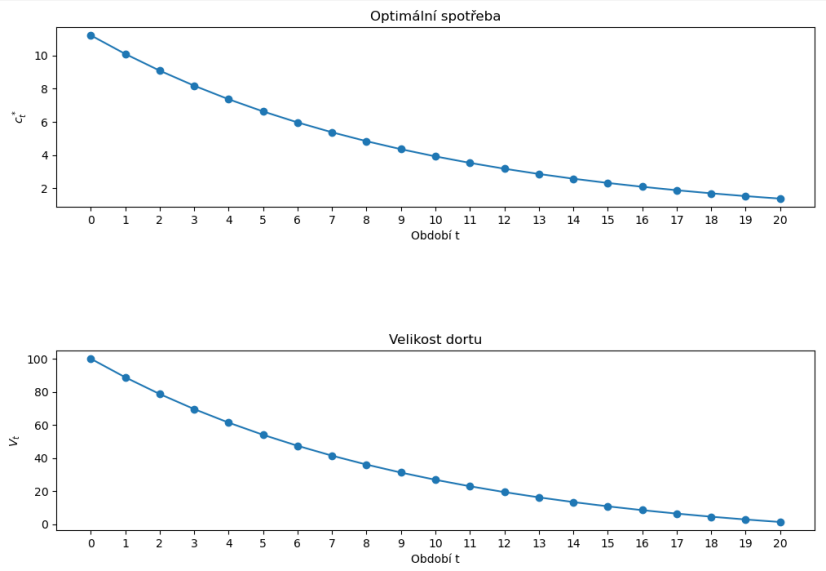
$$c_{T-k} = \frac{1}{1+\beta+\beta^2+\cdots+\beta^k} V_{T-k}$$

až po $k = T$, v kterémžto případě společně se vzorcem pro součet geometrické řady s konečně mnoha členy $\sum_{k=0}^{T} \beta^k = \frac{1-\beta^{T+1}}{1-\beta}$ získáme

$$c_0 = \frac{1-\beta}{1-\beta^{T+1}} V_0.$$

Toto je optimální velikost dílu dortu, který bychom měli sníst v prvním časovém okamžiku. Ve druhém okamžiku pak v souladu s Eulerovou rovnicí sníme díl o velikosti $c_1 = \beta c_0$, potom $c_2 = \beta c_1$ atd.
Např. bude-li výchozí velikost dortu $V = 100$ a $\beta = 0{,}9$, pak v případě $T = 20$, budeme postupně jíst části dortu o velikostech $c_0 = 11{,}2286$, $c_1 = 10{,}1058$, $c_2 = 9{,}0952$, $c_3 = 8{,}1857$, $c_4 = 7{,}3671$, $c_5 = 6{,}6304$, $c_6 = 5{,}9674$, $c_7 = 5{,}3706$, $c_8 = 4{,}8336$, $c_9 = 4{,}3502$, $c_{10} = 3{,}9152$, $c_{11} = 3{,}5237$, $c_{12} = 3{,}1713$, $c_{13} = 2{,}8542$, $c_{14} = 2{,}5687$, $c_{15} = 2{,}3119$, $c_{16} = 2{,}0807$, $c_{17} = 1{,}8726$, $c_{18} = 1{,}6854$, $c_{19} = 1{,}5168$, $c_{20} = 1{,}3651$, viz také Obrázek 5.7. Pro úplnost uvádíme i zdrojový kód skriptu[25], s jehož pomocí bylo toto řešení včetně obrázku získáno.

[25] Skript byl napsán i s využitím ChatGPT.



Obrázek 5.7: Ukázky vývoje užitku optimální strategii konzumace dortu z Příkladu 5.2.13 s volbou $V = 100$, $\beta = 0{,}9$ a $T = 20$.

```
import numpy as np
import matplotlib.pyplot as plt
import math as m
from scipy import stats as st
from scipy import optimize
import time

T = 20
beta = 0.9
kv = np.zeros(T+1,float)
cv = np.zeros(T+1,float)
uv = np.zeros(T+1,float)
kv[0] = 100  # k0
cv[0] = (1.0-beta)/(1.0-beta**(T+1)) * kv[0]  # c0
uv[0] = np.log(cv[0])

for i in range(1,T+1):
    cv[i] = beta * cv[i-1]
    kv[i] = kv[i-1] - cv[i-1]
    uv[i] = beta**i *np.log(cv[i])

print("cv = " + str(cv))
print("kv = " + str(kv))

fig, ax = plt.subplots(2, 1)
plt.subplots_adjust(wspace=0.4, hspace=0.8)

# Spotřeba
ax[0].plot(cv, '-o')
ax[0].set_ylabel(r'$c^*_t$')
ax[0].set_xlabel('Období t')
ax[0].set_title('Optimální spotřeba')
```

```
ax[0].set_xticks(range(T+1))

# Velikost dortu
ax[1].plot(kv, '-o')
ax[1].set_ylabel(r'$V_t$')
ax[1].set_xlabel('Období t')
ax[1].set_title('Velikost dortu')
ax[1].set_xticks(range(T+1))

plt.show()
```

Pro více informací a další rozšíření tohoto problému viz např. [4]. Alternativní přístup k řešení tohoto problému lze nalézt např. na webu [549].

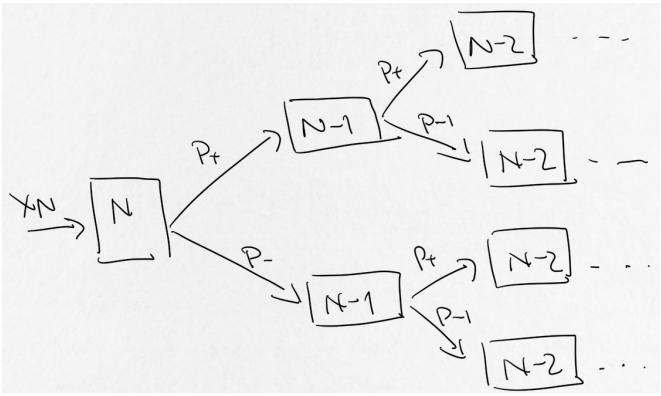
5.3 Stochastický proces

Ještě se krátce pozastavme u *stochastického* rozhodovacího procesu, ve kterém už samozřejmě není možné nalézt jednoznačnou optimální strategii jako v předchozí části. Přesuneme se proto do kasina v Las Vegas. Na počátku si koupíme V žetonů. Budeme hrát celkem N-krát jednoduchou hru, ve které je pravděpodobnost výhry p_+ a prohry přirozeně $p_- = 1 - p_+$. Při výhře se sázka zdvojnásobí, zatímco při prohře naši sázku ztratíme. Naším cílem je nalezení takové optimální strategie (tj. té s největší pravděpodobností), že na konci budeme mít alespoň Z žetonů. Pro $k = N, \ldots, 1$ si označíme jako

X_k ... počet žetonů na začátku k-té hry (tj. ve chvíli, kdy jsme si ještě nevsadili);

D_k ... počet žetonů vsazených v k-té hře.

Jedno kolo hry budeme brát jako jeden krok v dynamickém procesu, jehož začátek je schématicky zakreslen na Obrázku 5.8.



Obrázek 5.8: Začátek stochastického rozhodovacího procesu.

Potom $f_k(X_k, D_k)$ bude vyjadřovat pravděpodobnost, že po k hrách (tolik jich zbývá dokonce při našem zpětném číslování) budeme mít alespoň požadovaných Z žetonů, pokud na začátku k-té hry máme X_k žetonů a do této hry jich vsadíme D_k, přičemž až „doposud“ (tj. v krocích $N, \ldots, k+1$) jsme dělali optimální rozhodnutí. My chceme v tomto kroku najít

$$f_k(X_k) = \max_{D_k=0,\ldots,X_k} f_k(X_k, D_k),$$

přičemž vyjádření $f_k(X_K, D_k)$ musí zohledňovat skutečnost, že je možné mít alespoň Z žetonů i v případě prohry v k-té hře. Pokud v k-té hře prohrajeme, pak budeme mít $X_{k-1} = X_k - D_k$

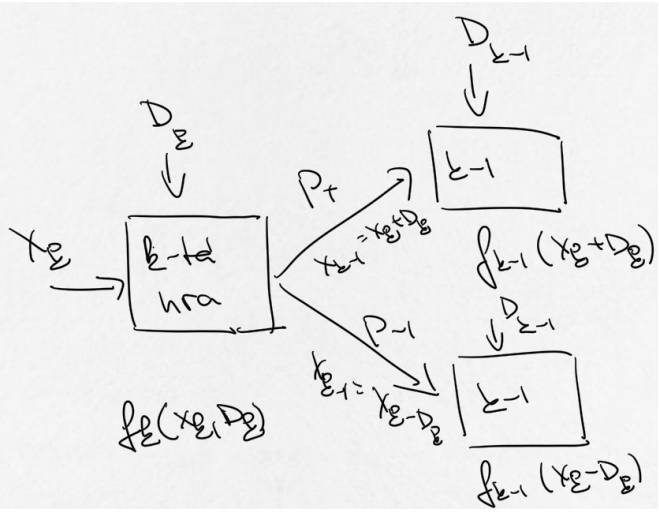
žetonů a pravděpodobnost úspěchu je $f_{k-1}(X_k - D_k)$. Podobně v případě výhry bude $X_{k-1} = X_k + D_k$ s pravděpodobností úspěchu $f_{k-1}(X_k + D_k)$. Proto platí

$$f_k(X_k, D_k) = p_+ f_{k-1}(X_k + D_k) + p_- f_{k-1}(X_k - D_k),$$

přičemž bereme

$$f_0(X_0) := \begin{cases} 0, & \text{je-li } X_0 < Z, \\ 1, & \text{je-li } X_0 \geq Z, \end{cases}$$

viz Obrázek 5.9.



Obrázek 5.9: Jeden krok stochastického rozhodovacího procesu.

Dostáváme tak rekurzivní schéma v podobě

$$f_k(X_k) = \max_{D_k = 0,\ldots,X_k} p_+ f_{k-1}(X_k + D_k) + p_- f_{k-1}(X_k - D_k)$$

pro $k = N, \ldots, 1$ a s $f_0(X_0)$ definovaným výše. V následujícím příkladu si tento výpočet vyzkoušíme s konkrétními hodnotami.

Příklad 5.3.1

Stochastický rozhodovací proces:
Video ☞ https://ln.math.muni.cz/pz/opt/dp5_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/dp5_zapisnik

⚠ Ve funkci $f_2(X_2, D_2)$ v záhlaví tabulky ve druhém kroku jsou chybné indexy (viz slide 2 a video v čase 10:36); správně má být $f_2(X_2, D_2) = \frac{2}{3} f_1(X_2 + D_2) + \frac{1}{3} f_1(X_2 - D_2)$.
⚠ Podobně ve funkci $f_3(X_3, D_3)$ je chybně na pravé straně f_3 (viz slide 3 a video v čase 22:10); správně má být $f_3(X_3, D_3) = \frac{2}{3} f_2(X_3 + D_3) + \frac{1}{3} f_2(X_3 - D_3)$.

Jako motivace pro náš další příklad nám může posloužit jeden klasický vtip[26]. V nejmenované firmě hledají nového zaměstnance. Dnes by měl manažer firmy a jeho kolega z HR oddělení vybrat vhodného kandidáta. Manažer vezme do ruky životopisy všech uchazečů a, aniž by je alespoň prolistoval, hodí polovinu z nich rovnou do koše. Když to kolega vidí, povídá „Co blbneš? To musíme pročíst, ne? Co když v té polovině je někdo, kdo má vystudovaný Harvard a umí přesně to, co my bychom potřebovali?“ „Nene, kašlu na to,“ odpovídá mu manažer, „my potřebujeme hlavně někoho, kdo má štěstí.“

[26] Za tento vtip děkuje autor podcastu *Čermák Staněk Comedy*.

Nemalou skupinu problémů dynamického programování tvoří tzv. *problémy optimálního zastavení*[27], ve kterých je cílem určit kolik času věnovat nějaké aktivitě, abychom dosáhli maximálního očekávaného užitku/zisku. Např. při výběru nového zaměstnance je otázkou, s kolika

[27] Angl. *Optimal stopping problem*.

[28] To už je v dnešní době genderově poněkud nevhodné.

uchazeči bychom měli absolvovat pohovor, než přijmeme toho, se kterým jsme právě pohovor absolvovali. Velmi často bývá v této souvislosti zmiňováno zejména najímání sekretářky[28] nebo lze najít alternativní formulaci v podobě otázky počtu partnerů/partnerek, se kterými bychom měli mít vztah, než se pustíme do manželství. Proto bývají pro tento problém používána označení jako *problém sekretářky*[29], *sňatkový problém*[30], ale také *problém sultánova věna*[31], *problém vybíravého nápadníka, hra googol*[32] či *problém nejlepší volby* a jeho řešení je také známé jako *pravidlo 37 %*. Proč právě toto číslo?

[29] Možná by bylo trefnější problém pojmenovat jako *když si musíte vybrat hned – a správně.* Každopádně anglický název je *Secretary problem.*

[30] Angl. *Marriage problem.*

[31] Úloha má podobnou motivaci jako v předchozím případě – jen tentokrát jde o sultána, který nabízí nápadníkovi jednu ze svých n dcer, které mu jsou postupně představovány i s jejich věny. Kdy má nápadník představování ukončit, pokud ví, že sultán bude se sňatkem souhlasit pouze pod podmínkou, že si vybral dceru s celkově největším věnem (dcery mají různá věna a jsou představovány v náhodném pořadí)?

[32] V této hře máme n karet, na kterých je napsáno n různých libovolných kladných čísel (mohou to být zlomky, iracionální čísla nebo třeba googol). Karty jsou zamíchány a rozloženy na stole otočené čísly dolů. Karty pak jednu po druhé postupně otáčíme v náhodném pořadí. Kdy otáčení karet zastavit, aby se maximalizovala pravděpodobnost, že na poslední otočené kartě bude napsané největší číslo ze všech?

Příklad 5.3.2

Uvažme situaci, kdy máme z kandidátů na jistou pozici vybrat toho nejlepšího. Kandidáti chodí na pohovory postupně a ihned po rozhovoru, než pozveme dalšího uchazeče, musíme rozhodnout, zda toho aktuálního přijmeme nebo nikoliv. Můžeme samozřejmě tohoto kandidáta porovnat s ostatními – ale jen s těmi, se kterými rozhovor již proběhl, nikoli s těmi, se kterými rozhovor teprve proběhne v budoucnosti. Pokud proběhnou pohovory s úplně všemi uchazeči, musí byt přijat ten poslední. Cílem je nalezení strategie, s jejíž pomocí maximalizujeme pravděpodobnost, že vybereme toho nejlepšího uchazeče.

Optimální strategie by měla spočívat v tom, že rozhovor proběhne s jistým počtem uchazečů a od té chvíle vybereme toho, který bude aktuálně nejlepší. Jelikož se ale nemůžeme vracet k odmítnutým uchazečům, musíme při volbě této strategie vyvážit riziko brzkého výběru toho zdánlivě nejlepšího, když ještě lepší může teprve přijít, a pozdního výběru, kdy zjistíme, že jsme toho nejlepšího dříve odmítli. Tímto postupem možná nevybereme toho úplně nejlepšího (třeba pokud by absolvoval pohovor jako první), ale většinou vybereme toho „skoro-nejlepšího“.

Nechť m je počet uchazečů, se kterými rozhovor již proběhl, a n počet zbývajících uchazečů, tj. celkem máme $N := m + n$ uchazečů. Zvolíme m jako naši stavovou veličinu a definujeme si dvojici funkcí popisujících pravděpodobnost výběru toho nejlepšího uchazeče ze všech. Číslo $V(m)$ bude pravděpodobnost toho, že se nám podaří vybrat toho nejlepšího kandidáta ze všech, pokud již proběhl pohovor s m uchazeči a m-tý uchazeč nebyl vybrán, zatímco $U(m)$ bude pravděpodobnost toho, že vybereme toho nejlepšího uchazeče ze všech, pokud proběhl pohovor s m uchazeči a ten m-tý je vybrán. Pokud nevybereme m-tého uchazeče, proces musí pokračovat s $V(m+1)$. Potřebujeme proto rozhodovací pravidlo pro to, kdy m-tý uchazeč může být vybrán, přičemž toto pravidlo musí záviset pouze na m a celkovém počtu N.

Připusťme, že už proběhlo m pohovorů a naše strategie byla doposud optimální. Pak z Bellmanova principu dostaneme rovnici

$$V(m) = \frac{m}{m+1} V(m+1) + \frac{1}{m+1} U(m+1),$$

která ukazuje, že pokud není vybrán m-tý uchazeč, pak s pravděpodobností $\frac{1}{m+1}$ bude vybrán ten $(m+1)$-ní, nebo nebude, a to s pravděpodobností $\frac{m}{m+1}$. Pokud není vybrán $(m+1)$-ní uchazeč, je pravděpodobnost toho, že se nám podaří vybrat toho nejlepšího ze všech $V(m+1)$, protože tento uchazeč určitě nebyl ten nejlepší. V opačném případě odpovídá pravděpodobnost výběru toho nejlepšího hodnotě $U(m+1)$, což je rovno

$$U(m+1) = \max\left\{\frac{m+1}{N}, V(m+1)\right\}.$$

Vskutku, je-li $(m+1)$-ní uchazeč nejlepší z těch, se kterými pohovor již proběhl, pak se musíme rozhodnout, zda proces výběru ukončit či v něm pokračovat. Pokud proces ukončíme, pak s pravděpodobností $\frac{m+1}{N} = 1 - \frac{N-(m+1)}{N}$ jsme vybrali toho úplně nejlepšího. Pokud ve výběru pokračujeme, pak z definice je $U(m+1) = V(m+1)$.

Začněme s $m = N$ a jako obvykle pokračujme rekurzivně. Podle definice je očividně $V(N) = 0$ a $U(N) = 1$. Nyní pro $m = N-1$ platí

$$V(N-1) = \frac{N-1}{N} V(N) + \frac{1}{N} U(N) = \frac{1}{N}$$

a

$$U(N-1) = \max\left\{\frac{N-1}{N}, V(N-1)\right\} = \max\left\{\frac{N-1}{N}, \frac{1}{N}\right\} = \frac{N-1}{N}.$$

Dále,

$$V(N-2) = \frac{N-2}{N-1} V(N-1) + \frac{1}{N-1} U(N-1) = \frac{N-2}{N}\left(\frac{1}{N-2} + \frac{1}{N-1}\right)$$

a

$$U(N-2) = \max\left\{\frac{N-2}{N}, V(N-2)\right\} = \frac{N-2}{N} \max\left\{1, \frac{1}{N-2} + \frac{1}{N-1}\right\}.$$

Indukcí se laskavý čtenář může sám přesvědčit, že

$$V(m) = \frac{m}{N}\left(\frac{1}{m} + \frac{1}{m+1} + \cdots + \frac{1}{N-1}\right)$$

a

$$U(m) = \frac{m}{N} \max\left\{1, \frac{1}{m} + \frac{1}{m+1} + \cdots + \frac{1}{N-1}\right\}.$$

Všimněme si, že

$$V(m) - V(m-1) = \frac{m}{N} \sum_{k=m}^{N-1} \frac{1}{k} - \frac{m-1}{N} \sum_{k=m-1}^{N-1} \frac{1}{k} = \frac{1}{N}\left(\sum_{k=m}^{N-1} \frac{1}{k} - 1\right),$$

což je kladné pro $\sum_{k=m}^{N-1} \frac{1}{k} > 1$ a záporné pro $\sum_{k=m}^{N-1} \frac{1}{k} < 1$, přičemž samotná suma je klesající vzhledem k m. Je-li $N > 3$, pak jistě $\sum_{k=1}^{N-1} \frac{1}{k} > 1$, a tudíž hodnota $V(m)$ nejdříve roste a od jistého m klesá. Optimální strategie je proto taková, že nevybereme kandidáta, který je aktuálně ten nejlepší, pokud

$$\sum_{k=m}^{N-1} \frac{1}{k} < 1,$$

a vybereme jej ve chvíli, kdy $m > m^*$, přičemž hodnota m^* je taková, že

$$\sum_{k=m^*}^{N-1} \frac{1}{k} \leq 1 < \sum_{k=m^*-1}^{N-1} \frac{1}{k}.$$

Pak si přímým výpočtem můžeme ověřit, že $U(0) = V(0) = V(m^* - 1)$, a to je pravděpodobnost, že vybereme toho nejlepšího mezi všemi N uchazeči, pokud postupujeme optimálně. Tato pravděpodobnost je rovna

$$V(0) = V(m^* - 1) = \frac{m^* - 1}{N} \sum_{k=m^*-1}^{N-1} \frac{1}{k} = \frac{m^* - 1}{N} \sum_{k=m^*}^{N-1} \frac{1}{k} + \frac{1}{N}.$$

Je-li N velké, pak rovnici pro m^* můžeme aproximovat jako

$$1 = \sum_{k=m^*}^{N-1} \frac{1}{k} \approx \int_{m^*}^{N} \frac{1}{x}\,\mathrm{d}x = \ln N - \ln m^* = \ln \frac{N}{m^*},$$

což nás přivede k přibližné hodnotě

$$m^* \approx \frac{N}{\mathrm{e}}$$

a podobně

$$V(m^* - 1) \approx \frac{m^* - 1}{N} \ln \frac{N}{m* - 1} \approx \frac{1}{\mathrm{e}} = 0{,}3678794412\ldots,$$

tj. přibližně prvních $\frac{1}{\mathrm{e}} \approx 36,78\,\%$ uchazečů musíme automaticky odmítnout a totéž je také pravděpodobnost, že se nám v takovém případě podaří vybrat nejlepšího možného kandidáta. Pro malé hodnoty N jsou počty odmítnutých kandidátů a odpovídající pravděpodobnost včetně aproximace podílu N/e uvedeny v Tabulkách 5.3–5.4.

N	1	2	3	4	5	6	7	8	9	10
m^*-1	0	0	1	1	2	2	2	3	3	3
N/e	0,3679	0,7358	1,1036	0,4583	1,8394	0,4278	2,5752	2,9430	3,3109	3,6788
$V(0)$	1	0,5	0,5	1,4715	0,4333	2,2073	0,4143	0,4098	0,4060	0,3987

Tabulka 5.3: Vybrané optimální hodnoty pro řešení Příkladu 5.3.2.

N	12	15	20	30	40	50
m^*-1	4	5	7	11	15	18
N/e	4,0467	5,5182	7,3576	11,0364	14,7152	18,3940
$V(0)$	0,3984	0,3894	0,3842	0,3787	0,3757	0,3743

Tabulka 5.4: Vybrané optimální hodnoty pro řešení Příkladu 5.3.2.

Více podrobností k tomuto problému lze nalézt např. v [40, 173, 187, 355].

5.4 Nekonečně krokový deterministický proces[36]

[36] Tato podkapitola vychází především z první kapitoly Bellmanovy knihy [45], kde se lze dočíst i další souvislosti. Viz také např. [46, 54, 55, 394, 456].

Nyní se budeme věnovat úloze dynamického programování s velkým počtem N kroků. Jako úvodní ilustrativní příklad může vzít problém

$$\sum_{k=1}^{N} \{A D_k^2 + B(X_k - D_k)^2\} \to \min, \quad \text{kde} \quad X_{k-1} = b(X_k - D_k) \quad \& \quad 0 \le D_k \le X_k \tag{5.4.1}$$

pro daná $A, B > 0$ a $0 < b < 1$. Využijeme-li k řešení této úlohy podobný postup jako v předchozích podkapitolách, dostaneme v první kroku rekurzivního schématu úlohu

$$f_1(X_1) = \min_{0 \le D_1 \le X_1} \{A D_1^2 + B(X_1 - D_1)^2\},$$

[37] Stačí najít stacionární bod funkce $A D_1^2 + B(X_1 - D_1)^2$ vzhledem k D_1.

jejímž řešením[37] je

$$D_1^* = \frac{B}{A+B} X_1 \quad \text{s hodnotou} \quad f_1(X_1) = \frac{AB}{A+B} X_1^2. \tag{5.4.2}$$

Odtud následně získáme ve druhém kroku problém

$$\begin{aligned} f_2(X_2) &= \min_{0 \le D_2 \le X_2} \left\{ A D_2^2 + B(X_2 - D_2)^2 + f_1(X_1)) \right\} = \\ &= \min_{0 \le D_2 \le X_2} \left\{ A D_2^2 + B(X_2 - D_2)^2 + f_1\big(b(X_2 - D_2)\big) \right\} = \\ &= \min_{0 \le D_2 \le X_2} \left\{ A D_2^2 + B(X_2 - D_2)^2 + \frac{AB}{A+B} b^2 (X_2 - D_2)^2 \right\} \end{aligned}$$

[38] To najdeme podobně jako v prvním kroku.

jehož řešením[38] je

$$D_2^* = \frac{B(A+B) + AB b^2}{(A+B)^2 + AB b^2} X_2 \quad \text{s hodnotou} \quad f_2(X_2) = \frac{AB(A+B) + A^2 B b^2}{(A+B)^2 + AB b^2} X_2^2. \tag{5.4.3}$$

Je vidět, že dalším pokračováním v tomto schématu obdržíme stále komplikovanější a komplikovanější výrazy v jednotlivých krocích. Můžeme si však všimnout, že hodnota f_k závisí na

X_k kvadraticky a D_k závisí na X_k lineárně. Platí to pro libovolné k? To ukážeme s využitím matematické indukce. Připusťme, že pro nějaké k je $f_{k-1}(X_{k-1}) = \alpha_{k-1} X_{k-1}^2 = \alpha_{k-1} b^2 (X_k - D_k)^2$. Pak máme úlohu

$$f_k(X_k) = \min_{0 \le D_k \le X_k} \left\{ A D_k^2 + B (X_k - D_k)^2 + \alpha_{k-1} b^2 (X_k - D_k)^2 \right\},$$

jejímž řešením je

$$D_k^* = \frac{B + \alpha_{k-1} b^2}{A + B + \alpha_{k-1} b^2} X_k \quad \text{s hodnotou} \quad f_k(X_k) = \frac{A B + A \alpha_{k-1} b^2}{A + B + \alpha_{k-1} b^2} X_k^2,$$

což ukazuje, že vskutku i závislost f_k na X_k je kvadratická a u D_k je lineární.

Každopádně obecně budeme v k-tém kroku mít problém

$$f_k(X_k) = \min_{0 \le D_k \le X_k} \left\{ A D_k^2 + B (X_k - D_k)^2 + f_{k-1}(b (X_k - D_k)) \right\}.$$

Je-li zvolený počet kroků skutečně velký, můžeme považovat $N \to \infty$ a v předchozích výpočtech udělat limitní přechod pro $k \to \infty$, přičemž dovolíme záměnu limity a minimalizace, tj. vstoupení s limitou k účelové funkci. Tím se dostaneme k poněkud komplikovanější úloze

$$f(X) = \min_{0 \le D \le X} \left\{ A D^2 + B (X - D)^2 + f(b (X - D)) \right\}, \tag{5.4.4}$$

ve které neznámá funkce f stojí na obou stranách rovnosti. Takto se z klasické optimalizační úlohy, jejíž řešení umíme nalézt pomocí diferenciálního počtu, ale stala spíše funkcionální rovnice[39] zahrnující minimalizaci. Jak dobře řešení této jedné úlohy aproximuje obecný N-krokový proces? My zde naštěstí můžeme využít naši informaci o kvadratické závislosti f na X, a proto budeme řešení dané rovnice hledat ve tvaru $f(X) = \alpha X^2$. Jejím dosazením na obou stranách rovnice obdržíme

$$\alpha X^2 = \min_{0 \le D \le X} \left\{ A D^2 + B (X - D)^2 + \alpha b^2 (X - D)^2 \right\}.$$

Derivováním funkce na pravé straně nalezneme stacionární bod $D = \frac{B + \alpha b^2}{A + B + \alpha b^2} X$, jenž nás po dosazení přivede k rovnici

$$\alpha X^2 = \frac{A (B + \alpha b^2)}{A + B + \alpha b^2} X^2,$$

odkud následně získáme

$$\alpha = \frac{A b^2 - A - B \pm \sqrt{(A + B - A b^2)^2 + 4 b^2 A B}}{2 b^2}.$$

Jelikož všechny koeficienty A a B jsou kladné, je hledaným řešením hodnota

$$\alpha = \frac{A b^2 - A - B + \sqrt{(A + B - A b^2)^2 + 4 b^2 A B}}{2 b^2}. \tag{5.4.5}$$

Zvolíme-li např. $A = 3$, $B = 1$ a $b = \frac{1}{\sqrt{2}}$, pak dosazením do (5.4.5) bude $\alpha = 1$, takže $f(X) = X^2$ a $D = X/3$. Skutečným řešení jsou podle (5.4.2)–(5.4.3) hodnoty $D_1^* = X_1/4$ a $D_2^* = \frac{11}{35} X_2$, což naznačuje, že jsme získali poměrně dobrou aproximaci. Mnohem lépe to je vidět z Tabulky 5.5, kde je hodnota D_N a $f_N(N_X)$ vypočtena pro $N = 2, 3, 4, 5, 7, 10, 15$ a 20.

[39] Ve funkcionální rovnici je neznámou nějaká funkce (obvykle reálné proměnné), která se v rovnici vyskytuje v „různých podobách“. Rovnice může vypadat např. takto

$$f(x+y) + f(x-y) = \\ = f(x) + 6xy \sqrt[3]{f(y)} + x^3$$

a naším cílem může být nalezení všech funkcí $f : \mathbb{R} \to \mathbb{R}$ vyhovujících této rovnici. Takové rovnice se často řeší pomocí *metody separace proměnných*, kdy se volí různé hodnoty proměnných s cílem dostat nějaké užitečné vlastnosti hledané funkce f. V tomto případě volbou $y = 0$ dostaneme

$$f(x) + f(x) = f(x) + x^3,$$

z čehož ihned vyplývá $f(x) = x^3$. Dosazením do dané rovnice se pak přesvědčíme, že tato funkce je skutečně jejím (jediným) řešením.

Nejznámější funkcionální rovnice (tzv. *Cauchyho funkcionální rovnice*) má podobu

$$f(x+y) = f(x) + f(y)$$

a jejím spojitým řešením jsou všechny lineární funkce $f(x) = c\,x$ pro libovolné $c \in \mathbb{R}$. O tom, jak toto řešení najít, a o dalších zajímavých typech funkcionálních rovnic jako např.

$$f(x+y) + f(x-y) = f(x)\,f(y)$$

nebo

$$f(x+y) = \frac{f(x) + f(y)}{1 - f(x)\,f(y)}$$

se lze dočíst např. v [334, 400].

$A=3$ & $B=1$ & $b=\frac{1}{\sqrt{2}}$		
N	D_N	$f_N(X_N)$
1	$\frac{1}{4}X_1$	$\frac{3}{4}X_1^2$
2	$\frac{11}{35}X_2$	$\frac{33}{35}X_2^2$
3	$\frac{103}{313}X_3$	$\frac{309}{313}X_3^2$
4	$\frac{935}{2813}X_4$	$\frac{2805}{2813}X_4^2$
5	$\frac{8431}{25309}X_5$	$\frac{25293}{25309}X_5^2$
7	$\frac{683263}{2049853}X_7$	$\frac{2049789}{2049853}X_7^2$
10	$\frac{498111911}{1494336245}X_{10}$	$\frac{1494335733}{1494336245}X_{10}^2$
15	$0,333333333271440\,X_{15}$	$0,999999999814322\,X_{15}^2$
20	$0,333333333333299\,X_{20}$	$0.999999999999904\,X_{20}^2$

Tabulka 5.5: Postupný výpočet D_N a odpovídající $f_N(X_N)$ pro volbu parametrů $A=3$, $B=1$ a $b=\frac{1}{\sqrt{2}}$ z úlohy (5.4.1).

Z tabulky je zjevné, že získána aproximace je poměrně dobrá a s větším počtem kroků by byla ještě výrazně lepší (kvalitu této aproximace příliš nezhorší ani jiná volba parametrů A, B a b).

Nyní se na tuto úlohu podívejme z poněkud širší perspektivy. Podobně jako v předchozí situaci dostaneme limitním přechodem pro $N\to\infty$ z rekurzivního schématu (5.2.6) s aditivní účelovou funkcí rovnici

$$f(X)=\max_D\big[r(X,D)+f\big(y(X,D)\big)\big]. \tag{5.4.6}$$

[40] Angl. *Method of successive approximations.*

První možností pro řešení této úlohy je tzv. *metoda postupných aproximací*[40], což je vlastně zpětné nahrazení nekonečně krokového procesu jeho konečně krokovou analogií, tj. je-li $f_0(X)$ nějaká počáteční aproximace (např. $f_0(X)\equiv 0$) a pak postupně řešíme

$$f_k(X)=\max_D\big[r(X,D)+f_{k-1}\big(y(X,D)\big)\big]\quad\text{pro } k=1,2,\ldots$$

[41] Angl. *Approximation in the policy space.*

Druhou možností je tzv. *metoda aproximace v prostoru rozhodnutí*[41]. Řešení optimalizační úlohy (5.4.6) se realizuje v bodě D, který závisí na X, tj. $D^*=D^*(X)$, a funkce f tedy splňuje rovnici

$$f(X)=r(X,D^*(X))+f\big(y(X,D^*(X))\big).$$

Tato rovnice v porovnání s (5.4.6) má tu výhodu, že již neobsahuje maximalizaci, a tudíž se jedná o funkcionální rovnici v obvyklém smyslu. To nám dává šanci, že nalezení řešení této rovnice bude jednodušší než v případě rovnice (5.4.6). Proto místo počáteční aproximace $f_0(X)$ se můžeme pokusit „uhodnout“ optimální rozhodnutí $D^*=D^*(X)$ a označit jej jako $D=D_0(X)$. Z této naší počáteční aproximace dostaneme funkcionální rovnici

$$f_0(X)=r(X,D_0(X))+f\big(y(X,D_0(X))\big).$$

Pokud se nám podaří najít její řešení f_0, pak následující aproximaci $D_1(X)$ optimálního rozhodnutí získáme jako řešení úlohy

$$\max_D\big[r(X,D)+f_0\big(y(X,D)\big)\big].$$

S využitím této hodnoty pak obdržíme novou funkcionální rovnici

$$f_1(X)=r(X,D_1(X))+f\big(y(X,D_1(X)\big),$$

jejíž řešení nám poskytne další aproximaci $f_1(X)$ hledané funkce. Analogickým postupem pak pokračujeme dále.

Využití těchto dvou metod si můžeme ilustrovat na naší úvodní „kvadratické“ úloze (5.4.4). Řešení založené na postupných aproximacích s volbou $f_0(X) \equiv 0$ jsme již vlastně využili v (5.4.2)–(5.4.3), kde jsme určili alespoň $f_1(X)$ a $f_2(X)$. Pro metodu aproximací v prostoru rozhodnutí vezměme $D_0(X) = \frac{B}{A+B}\,X$, což by odpovídalo počáteční aproximaci $f_0(X) \equiv 0$, viz (5.4.2). Pak pro f_0 dostaneme rovnici

$$f_0(X) = \frac{AB}{A+B}\,X^2 + f_0\Big(b\,\frac{A\,X}{A+B}\Big).$$

Při jejím řešení využijeme skutečnosti, že by hledaná funkce měla záviset na X kvadraticky, tj. $f_0(X) = \alpha_0\,X^2$, z čehož po dosazení snadno vypočteme

$$\alpha_0 = \frac{A\,B\,(A+B)}{(A+B)^2 - b^2\,A^2}.$$

Hodnotu $D_1(X)$ následně určíme jako řešení minimalizační úlohy

$$A\,D^2 + B\,(X-D)^2 + \alpha_0 b^2\,(X-D)^2 \to \min \quad \& \quad 0 \le D \le X.$$

Derivováním a upravováním zjistíme, že

$$D_1(X) = \frac{B + \alpha_0 b^2}{A + B + \alpha_0 b^2}\,X.$$

To nás přivede ke druhé funkcionální rovnici

$$f_1(X) = \frac{A(B+\alpha_0 b^2) + A^2 B}{(A+B+\alpha_0 b^2)^2}\,X^2 + f_1\Big(\frac{b\,A}{A+B+\alpha_0 b^2}\,X\Big),$$

jejíž řešení opět hledáme v podobě $f_1(X) = \alpha_1\,X^2$. Tudíž

$$f_1(X) = \underbrace{\frac{A(B+\alpha_0 b^2) + A^2 B}{(A+B+\alpha_0 b^2)^2 - b^2 A^2}}_{=\alpha_1}\,X^2.$$

V dalším kroku bychom analogickým způsobem získali

$$D_2(X) = \frac{B + \alpha_1 b^2}{A + B + \alpha_1 b^2}\,X \quad \text{a} \quad f_2(X) = \underbrace{\frac{A(B+\alpha_1 b^2) + A^2 B}{(A+B+\alpha_1 b^2)^2 - b^2 A^2}}_{=\alpha_2}\,X^2.$$

Tato metoda konverguje výrazně rychleji než metoda postupných aproximací, protože zahrnuje jak optimalizaci, tak i „regulaci“ optimální hodnoty f pomocí rovnice pro nekonečně krokový proces. Toto můžeme vidět na úloze (5.4.4) s volbou parametrů $A = 3$, $B = 1$ a $b = \frac{1}{\sqrt{2}}$, viz Tabulku 5.6 a srovnej s Tabulkou 5.5. Jsou zde uvedeny jen první 3 aproximace, protože získaná funkce $f_2(X)$ se od řešení úlohy (5.4.4) liší až na desátém desetinném místě.

$A = 3$ & $B = 1$ & $b = \frac{1}{\sqrt{2}}$		
N	D_N	$f_N(X)$
0	$\frac{1}{4}\,X$	$\frac{24}{23}\,X^2$
1	$\frac{35}{104}\,X$	$\frac{16872}{16871}\,X^2$
2	$\frac{25307}{75920}\,X$	$\frac{8966017032}{8966017031}\,X^2$

Tabulka 5.6: Výpočet D_N a odpovídající $f_N(X)$ pro volbu parametrů $A = 3$, $B = 1$ a $b = \frac{1}{\sqrt{2}}$ z úlohy (5.4.4) pomocí metody aproximací v prostoru rozhodnutí.

Zobecněním funkcionální rovnice (5.4.6) z předchozího příkladu může být následující situace. Máme k dispozici X_1 vstupních prostředků, které chceme rozdělit na dvě investice (např. nákup strojů apod.). První investice je ve výši $0 \le D_1 \le X_1$ a druhá investice $X_1 - D_1$ (tj. celkem investujeme vše), přičemž zisky jsou rovny $g(D_1)$ a $h(X_1 - D_1)$. To vede k hledání řešení úlohy

$$g(D_1) + h(X_1 - D_1) \to \max \quad \& \quad 0 \le D_1 \le X_1.$$

Po určité době obě zařízení prodáme a obdržíme za ně částku aD_1 a $b(X-D_1)$, kde $a,b\in[0,1)$ vyjadřují amortizaci obou zařízení, tj. ve druhém kroku máme k dispozici na rozdělení $X_2 = aD_1 + b(X_1 - D_1)$ prostředků. Tyto prostředky využijeme stejným způsobem jako u první investice, což nás přivede k upravené úloze

$$g(D_1)+h(X_1-D_1)+g(D_2)+h(X_2-D_2)\to\max$$
$$0\le D_1\le X_1 \quad\&\quad 0\le D_2\le X_2 \quad\&\quad X_2 = aD_1+b(X_1-D_1).$$

Odtud se podobnými úvahami jako dříve můžeme dostat k rekurentnímu vztahu

$$f_N(X)=\max_{0\le D\le X}\Big\{g(D)+h(X-D)+f_{N-1}\big(aD+b(X-D)\big)\Big\},$$

z čehož pro $N\to\infty$ získáme funkcionální rovnici

$$f(X)=\max_{0\le D\le X}\Big\{g(D)+h(X-D)+f\big(aD+b(X-D)\big)\Big\}. \tag{5.4.7}$$

Otázku řešitelnosti takovéto rovnice nám pak zodpovídá následující věta.

Věta 5.4.1

Nechť jsou dána čísla $a,b\in[0,1)$ a dvojice funkcí g,h na $\mathbb{R}_+$, které jsou zde spojité a splňují $g(0)=0=h(0)$. Dále označme

$$m(X):=\max_{0\le D\le X}\max\big\{|g(D)|,|h(D)|\big\}$$

a $c:=\max\{a,b\}$. Jestliže

$$\sum_{k=0}^{\infty} m(c^kX)<\infty$$

pro všechna $X\ge 0$, pak existuje jediné řešení funkcionální rovnice (5.4.7), které splňuje $f(0)=0$ a je spojité v $X=0$. Toto jediné řešení je dokonce spojité pro všechna $X\in\mathbb{R}_+$.

Důkaz. Definujme zobrazení[42] $T:\mathrm{C}([0,\infty))\times\mathbb{R}\to\mathrm{C}([0,\infty))$ předpisem

$$T(f,D):=g(D)+h(X-D)+f\big(aD+b(X-D)\big).$$

Zvolme $f_0\equiv 0$ a definujme rekurentně

$$f_{k+1}(X):=\max_{0\le D\le X}T(f_k,D). \tag{5.4.8}$$

Ukážeme, že předpoklady věty zaručují stejnoměrnou konvergenci $f_k\rightrightarrows f$ na intervalu $[0,X_0]$ pro libovolné $X_0>0$. V každém kroku označme jako $D_k=D_k(X)\in[0,X]$ tu hodnotu, pro kterou realizuje maximum v definici f_{k+1}, tj. $f_{k+1}(X)=T(f_k,D_k)$. Pak platí

$$f_{k+1}(X)=T(f_k,D_k)\ge T(f_k,D_{k+1}) \quad\text{a}\quad f_{k+2}(X)=T(f_{k+1},D_{k+1})\ge T(f_{k+1},D_k),$$

tudíž

$$T(f_k,D_{k+1})-T(f_{k+1},D_{k+1})\le f_{k+1}(X)-f_{k+2}(X)\le T(f_k,D_k)-T(f_{k+1},D_k).$$

Kombinací posledních dvou nerovností obdržíme

$$\big|f_{k+1}(X)-f_{k+2}(X)\big|\le\max\Big\{\big|T(f_k,D_{k+1})-T(f_{k+1},D_{k+1})\big|,\big|T(f_k,D_k)-T(f_{k+1},D_k)\big|\Big\}. \tag{5.4.9}$$

Současně z definice operátoru T získáme

$$\big|T(f_k,D_k)-T(f_{k+1},D_k)\big|=\big|f_k\big(aD_k+b(X-D_k)\big)-f_{k+1}\big(aD_k+b(X-D_k)\big)\big| \tag{5.4.10}$$

a podobně

$$\big|T(f_k,D_{k+1})-T(f_{k+1},D_{k+1})\big|=\big|f_k\big(aD_{k+1}+b(X-D_{k+1})\big)-f_{k+1}\big(aD_{k+1}+b(X-D_{k+1})\big)\big|. \tag{5.4.11}$$

[42] Připomeňme, že symbol $\mathrm{C}([0,\infty))$ značí třídu spojitých funkcí na intervalu $[0,\infty)$.

Označme $u_k(X) := \max_{0\le z\le X} \left|f_{k+1}(z) - f_k(z)\right|$. Protože $aD + b(X-D) \le cX$ pro všechna $D \in [0, X]$, z rovností (5.4.10)–(5.4.11) s využitím (5.4.9) plyne

$$\begin{aligned}
u_{k+1}(X) &= \max_{0\le z\le X} \left|f_{k+2}(z) - f_{k+1}(z)\right| \le \\
&\le \max_{0\le z\le X} \max\left\{\left|T(f_k, D_{k+1}(z)) - T(f_{k+1}, D_{k+1}(z))\right|, \left|T(f_k, D_k(z)) - T(f_{k+1}, D_k(z))\right|\right\} \le \\
&\le \max_{0\le z\le X} \max\Big\{\Big|f_k\big(\underbrace{aD_k(z) + b(X - D_k(z))}_{\le cX}\big) - f_{k+1}\big(aD_k(z) + b(X - D_k(z))\big)\Big|, \\
&\qquad\qquad \Big|f_k\big(\underbrace{aD_{k+1}(z) + b(X - D_{k+1}(z))}_{\le cX}\big) - f_{k+1}\big(aD_{k+1}(z) + b(X - D_{k+1}(z))\big)\Big|\Big\} \le \\
&\le \max_{0\le z\le cX} \left|f_k(z) - f_{k+1}(z)\right| = u_k(cX).
\end{aligned}$$

Opakovaným aplikováním této nerovností získáme $u_{k+1}(X) \le u_0(c^{k+1}X)$ a zbývá ještě odhadnout $u_0(X)$. Jelikož[43] $f_0 \equiv 0$, platí

$$\begin{aligned}
u_0(X) &= \max_{0\le D\le X} \left|f_1(D) - f_0(D)\right| = \max_{0\le D\le X} \left|f_1(D)\right| = \max_{0\le D\le X} \left|\max_{0\le z\le D} g(z) + h(X - z)\right| = \\
&= \max_{0\le D\le X} \left|g(D) + h(X - D)\right| \le \max_{0\le D\le X} \left\{|g(D)| + |h(X - D)|\right\} \le \\
&\le \max_{0\le D\le X} |g(D)| + \max_{0\le D\le X} |h(X - D)| \le 2\, m(X).
\end{aligned}$$

Odtud a z předpokladů věty dostáváme

$$\sum_{k=0}^{\infty} \left|f_{k+1}(X) - f_k(X)\right| \le \sum_{k=0}^{\infty} u_k(X) \le 2 \sum_{k=0}^{\infty} m(c^k X) < \infty,$$

takže podle Weierstrassova kritéria stejnoměrné konvergence je řada $\sum_{k=0}^{\infty} \left[f_{k+1}(X) - f_k(X)\right]$ stejnoměrně konvergentní. Označme její součet jako f. Pak $f = \lim_{k\to\infty} f_k(X)$ a tato funkce je spojitá díky stejnoměrné konvergenci a spojitosti jednotlivých funkcí $f_k(X)$. Stejnoměrnost konvergence zaručuje také, že $f(X)$ je hledaným řešením funkcionální rovnice (5.4.7).

Zbývá ukázat jednoznačnost tohoto řešení. Připusťme, že funkce $F(D)$ je také řešením funkcionální rovnice (5.4.7), je spojitá pro $D = 0$ a $F(0) = 0$. Položme $u(X) := \max_{0\le D\le X} \left|F(D) - f(D)\right|$. Podobnými úvahami jako v první části lze získat nerovnost $u(X) \le u(cX)$ a dalším iterováním $u(X) \le u(c^k X)$ pro všechna $k \in \mathbb{N}$. Protože však $c^k X \to 0$ pro $k \to \infty$ a funkce u je spojitá v 0 s hodnotou $u(0) = 0$, dostáváme odtud $u(X) \equiv 0$, tj. $F(D) \equiv f(D)$. ■

V některých speciálních případech lze konvergenci posloupnosti $\{f_k(X)\}_{k=0}^{\infty}$ ukázat i mnohem snadněji jako např. v situaci, kdy funkce g, h jsou nezáporné. To zaručuje, že posloupnost $\{f_k(X)\}_{k=0}^{\infty}$ je navíc neklesající. Proto $f_{k+2}(X) \ge f_{k+1}(X) = \max_{0\le D\le X} T(f_k, D)$ pro všechna $k = 0, 1, 2\ldots$, a tedy nutně také $f(X) \ge \max_{0\le D\le X} T(f_k, D)$. To znamená, že pro každé $D \in [0, X]$ je $f(X) \ge T(f_k, D)$, což pro $k \to \infty$ dává $f(X) \ge T(f, D)$ pro všechna $D \in [0, X]$, a tudíž máme[44] $f(X) \ge \sup_{0\le D\le X} T(f, D)$. Současně z definice f_k plyne $f_{k+1}(X) \le \sup_{0\le D\le D} T(f, D)$ pro všechna $k = 0, 1, \ldots$, a proto $f(X) \le \sup_{0\le D\le D} T(f, D)$, což společně s předchozí nerovností ukazuje, že platí $f(X) = \sup_{0\le D\le D} T(f, D)$, tj. f je řešením mírně upravené funkcionální rovnice (5.4.7), kde maximum nahradíme supremem. Nedostatkem tohoto důkazu založeného pouze na monotónii je právě to, že nám nedokáže zaručit spojitost limitní funkce, a tudíž ani existenci optimálního rozhodnutí, tj. funkce $D = D(X)$, která by dávala maximum v rovnici (5.4.7).

Existence optimálního rozhodnutí pro samotný nekonečný proces nemá přímo žádný zvláštní význam z hlediska výpočtu, nebo co se aplikací týče. Má však velký význam v souvislosti s určením struktury optimálních rozhodnutí pro nekonečný proces. Význam existence spojitých řešení je totiž spíše nepřímý, a to pokud jde o numerické výsledky, jelikož řešení nekonečného procesu lze použít jako aproximaci k řešení konečného.

Předpoklady Věty 5.4.1 zaručují, že jednotlivé funkce v posloupnosti $\{f_k(X)\}_{k=0}^{\infty}$ určené metodou postupných aproximací podle (5.4.8) jsou spojité na $\mathbb{R}_+$. Avšak samotná optimální rozhodnutí D již nutně nemusí na X záviset spojitě, což si ilustrujeme v následujícím příkladě.

[43] Zde není zcela nutné volit $f_0 \equiv 0$. Důkaz by „prošel“ i pro libovolnou funkci f_0, která je spojitá na $\mathbb{R}_+$ a $f(0) = 0$.

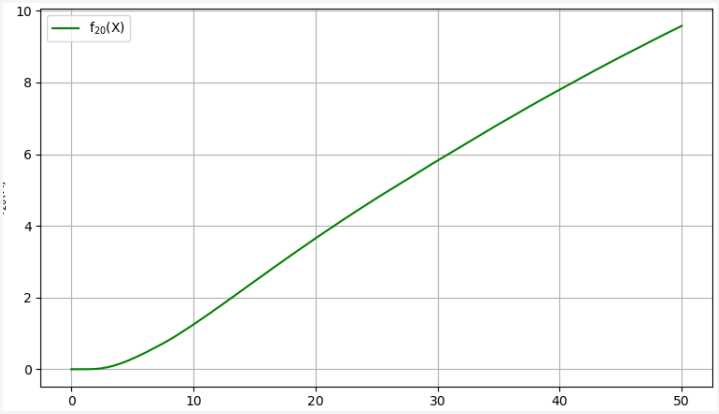
[44] Na pravé straně záměrně nepíšeme $\max_{0\le D\le X}$, abychom nevyužívali předchozí obecný důkaz spojitosti $f(X)$, která by nám tuto záměnu umožnila.

Příklad 5.4.2

(Bellmanův „divnopříklad“) Použijeme-li metodu postupných aproximací na řešení úlohy

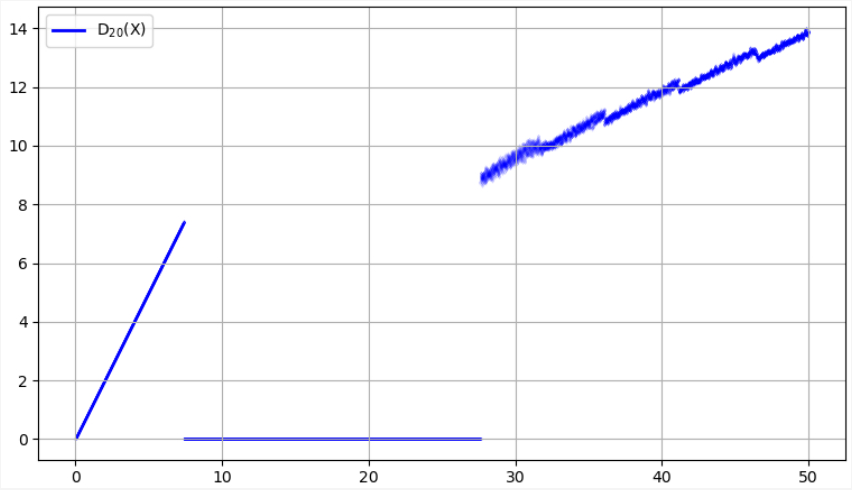
$$f(X) = \max_{0 \le D \le X} \{e^{-10/D} + e^{-15/(X-D)} + f(0{,}8\,D + 0{,}9\,(X-D))\},$$

získáme spojitou funkci $f(X)$ zobrazenou na Obrázku 5.10.



Obrázek 5.10: Funkce $f(X)$ z Příkladu 5.4.2.

Avšak chování funkce odpovídajících optimálních rozhodnutí $D(X)$ už je o dost divočejší, jak můžeme vidět na Obrázku 5.11.



Obrázek 5.11: Funkce $D(X)$ z Příkladu 5.4.2.

Oba obrázky byly získány numericky s pomocí skriptu pro Python[45] uvedeném níže při použití 20 kroků metody postupných aproximací (přidání dalších kroků už na výsledku mnoho nezmění a rozhodně se nepřiblíží k Bellmanově výsledku)[46].

```
import numpy as np
import csv
import matplotlib.pyplot as plt
from multiprocessing import Pool, cpu_count
from matplotlib.lines import Line2D

# Parametry
krok = 0.01
X_values = np.arange(krok, 50 + krok, krok)
j_max = 10
```

[45] Pro
skriptu

[46] Stej
ale sam
hem dě

```python
# Funkce pro výpočet hodnoty F_j(X) pro jednu hodnotu X
def compute_F_for_X(args):
    X, j, prev_F_values = args
    X = round(X, 3)

    # Výpis každých 10 jednotek X
    if abs((X * 1000) % 10000) < 0.001:  # např. 10.000, 20.000, ...
        print(f"      Výpočet pro X = {X:.3f}")

    x_vals = np.arange(krok, X, krok)
    F_temp = []

    for x in x_vals:
        x = round(x, 6)
        diff = X - x
        if x <= 0 or diff <= 0:
            continue

        prev_X = round((0.8 * x + 0.9 * diff) / krok) * krok
        prev_X = round(prev_X, 3)
        prev_val = prev_F_values.get(prev_X, 0)

        val = np.exp(-10 / x) + np.exp(-15 / diff) + prev_val
        F_temp.append((val, x))

    if F_temp:
        max_val, max_x = max(F_temp, key=lambda t: t[0])
        return X, max_val, max_x
    else:
        return X, 0, 0

# Slovníky pro výsledky
F_values = {0: {round(X, 3): 0 for X in X_values}}
D_values = {}

# Výpočet
for j in range(1, j_max + 1):
    print(f"--- Výpočet pro j = {j} ---")
    prev_F = F_values[j - 1]
    D_values[j] = {}
    F_values[j] = {}

    args_list = [(X, j, prev_F) for X in X_values]

    with Pool(processes=cpu_count()) as pool:
        results = pool.map(compute_F_for_X, args_list)

    for X, F_val, D_val in results:
        F_values[j][X] = F_val
```

```
        D_values[j][X] = D_val

# Export do CSV
with open("divnopriklad-paralel.csv", mode="w", newline='') as
 ↪  file:
    writer = csv.writer(file)
    writer.writerow(['X', f'f$_{{{j_max}}}$(X)',
     ↪  f'D$_{{{j_max}}}$(X)'])
    for X in X_values:
        X_rounded = round(X, 3)
        writer.writerow([
            round(X_rounded, 3),
            round(F_values[j_max].get(X_rounded, 0), 6),
            round(D_values[j_max].get(X_rounded, 0), 4)
        ])

# --- GRAFY ---
X_list = [round(X, 2) for X in X_values]
F_list = [F_values[j_max].get(X, 0) for X in X_list]
D_list = [D_values[j_max].get(X, 0) for X in X_list]

plt.figure(figsize=(12, 5))

# Graf F(j_max, X) - spojitá čára
plt.subplot(1, 2, 1)
plt.plot(X_list, F_list, label=f'f$_{{{j_max}}}$(X)',
 ↪  color='green')
plt.xlabel('X')
plt.ylabel(f'f$_{{{j_max}}}$(X)')
plt.title(f'Závislost f$_{{{j_max}}}$(X) na X')
plt.grid(True)
plt.legend()

# Graf D(j_max, X) - bodový graf (scatter)
plt.subplot(1, 2, 2)
plt.scatter(X_list, D_list, label=f'D$_{{{j_max}}}$(X)',
 ↪  color='blue', s=2)
plt.xlabel('X')
plt.ylabel(f'D$_{{{j_max}}}$(X)')
plt.title(f'Bod maxima D$_{{{j_max}}}$(X) při výpočtu
 ↪  f$_{{{j_max}}}$(X)')
plt.grid(True)
# Přidání modré úsečky do legendy
line = Line2D([0], [0], color='blue', lw=2,
 ↪  label=f'D$_{{{j_max}}}$(X)')
plt.legend(handles=[line])
#plt.legend()

plt.tight_layout()
plt.show()
```

Zdůrazněme, že Bellmann ve své knize [45, str. 25] má dost odlišný obrázek pro $D(X)$, i když také je výsledná funkce nespojitá. Autor je nicméně přesvědčen, že jim získané řešení je správné, ale srovnání s Bellmannovým výsledkem chybí, jelikož Bellmann nabízí svůj obrázek bez dalších podrobností ohledně výpočtu. Ještě doplňme, že funkce $e^{-c/D}$ je zde zvolena proto, že se jedná o jednu z nejjednodušších funkcí s inflexním bodem, jehož absence nám umožní ještě přesnější analýzu řešení funkcionální rovnice, jak uvidíme za chvíli.

Nyní obraťme pozornost ke konvergenci druhé základní metody vyšetřování funkcionálních rovnic dynamického programování, tj. metody aproximace v prostoru rozhodnutí, jejíž princip jsme již naznačili dříve. Z funkcionální rovnice (5.4.7) vidíme, že ze znalosti $f(X)$ můžeme určit optimální rozhodnutí $D = D(X)$ a naopak jakákoli volba $D = D(X)$ určuje $f(X)$, které postupně získáme z funkcionální rovnice $f(X) = T(f, D(X))$. Např. kdyby optimální rozhodnutí bylo neustále $D = 0$, pak $f(X)$ by splňovalo funkcionální rovnici

$$f(X) = h(X) + f(bX),$$

což by nás přivedlo k řešení

$$f(X) = \sum_{n=0}^{\infty} h(b^n X).$$

Připomeňme, že naším cílem není určit $f(X)$, které je vskutku jen „vedlejším produktem“ našeho vyšetřování, ale mnohem důležitější je určení struktury optimálních rozhodnutí $D = D(X)$, které mohou pro některá X dokonce nabývat více hodnot, tj. $D(X)$ by vlastně nebyla funkce proměnné X, což však v našich dalších úvahách raději nebudeme rozlišovat.

Toto nás přivádí k důležité a užitečné myšlence. Stejně jako můžeme aproximovat $f(X)$ v prostoru funkcí, můžeme stejně tak aproximovat $D(X)$ v prostoru možných rozhodnutí. V mnoha ohledech se dokonce jedná o přirozenější a jednodušší formu aproximace. Z analytického pohledu je výhodou tohoto typu aproximace, že vždy vede k monotónním aproximacím. Z hlediska aplikací je to daleko přirozenější aproximace, protože to je obvykle ta část problému, k níž máme jisté množství informací.

V následující větě poměrně snadno ve srovnání s předchozím tvrzením ukážeme, že v případě, kdy počáteční aproximaci f_0 bereme jako řešení funkcionální rovnice s nějakou počáteční aproximací $D_0 = D_0(X)$ v prostoru optimálních rozhodnutí, je konvergence posloupnosti postupných aproximací monotónní.

VĚTA 5.4.3

Nechť $D_0 = D_0(X)$ je libovolná spojitá funkce splňující nerovnost $0 \le D_0(X) \le X$ a nechť f_0 je řešením funkcionální rovnice

$$f_0(X) = g(D_0(X)) + h(X - D_0(X)) + f_0\big(aD_0(X) + b(X - D_0(X))\big).$$

Jsou-li splněny předpoklady Věty 5.4.1, pak je posloupnost $\{f_k(X)\}_{k=0}^{\infty}$ definovaná rekurentně jako

$$f_k(X) = \max_{0 \le D \le X} \{g(D) + h(X - D) + f_{k-1}\big(aD + b(X - D)\big)\}$$

monotónní a její konvergence k řešení rovnice (5.4.7) je stejnoměrná.

Důkaz. Z definice funkcí f_1 a f_0 vyplývá

$$\begin{aligned} f_1(X) &= \max_{0 \le D \le X} \{g(D) + h(X - D) + f_0\big(aD + b(X - D)\big)\} \ge \\ &\ge g(D_0(X)) + h(X - D_0(X)) + f_0\big(aD_0(X) + b(X - D_0(X))\big) = f_0(X), \end{aligned}$$

a tudíž také

$$f_2(X) = \max_{0\le D\le X}\{g(D)+h(X-D)+f_1\big(aD+b(X-D)\big)\} \ge$$
$$\ge \max_{0\le D\le X}\{g(D)+h(X-D)+f_0\big(aD+b(X-D)\big)\} = f_1(X).$$

Indukcí pak odtud získáme, že $f_{k+1}(X) \ge f_k(X)$ pro všechna $k = 0,1,2\ldots$ Zbývá ještě ukázat spojitost $f_0(X)$ na $\mathbb{R}_+$. Postupným dosazováním do definice $f_0(X)$ získáme

$$f_0(X) = g(D_0(X)) + h(X - D_0(X)) + g\Big(D_0\big(aD_0(X) + b(X - D_0(X))\big)\Big) + \cdots, \tag{5.4.12}$$

což je díky předpokladům pro funkce g,h stejnoměrně konvergentní řada na libovolném konečném intervalu podle Weierstrassova kritéria. Proto spojitost počáteční volby $D_0(X)$ vzhledem k X zaručuje, že součet řady (5.4.12) je spojitá funkce na $\mathbb{R}_+$. ■

Nyní se podíváme blíže na podobu optimálních rozhodnutí, která v obecném případě může být poměrně komplikovaná. Pokud však zesílíme požadavky na funkce g a h, můžeme získat poměrně přesné informace, jako je tomu v následujících tvrzeních.

Věta 5.4.4

Nechť jsou splněny předpoklady Věty 5.4.1 a funkce g,h jsou navíc konvexní. Pak funkce f, která je řešením rovnice (5.4.7) je také konvexní a pro libovolné X bude optimální rozhodnutí buď $D(X) = 0$ nebo $D(X) = X$, tj. maximum v (5.4.7) bude dosaženo v jednom z krajních bodů intervalu $[0,X]$.

Důkaz. Nejdříve ukážeme, že zmíněnou vlastnost mají všechny funkce $f_0, f_1, \ldots$ Jelikož $f_0 \equiv 0$ a funkce $g(D)+h(X-D)$ je konvexní pro $D \in [0,X]$, platí[47]

$$f_1(X) = \max_{0\le D\le X}\{g(D)+h(X-D)\} = \max\{g(0)+h(X), g(X)+h(0)\} = \max\{h(X), g(X)\},$$

což je konvexní funkce jakožto maximum ze dvou konvexních funkcí[48]. V důsledku toho je konvexní také funkce $g(D)+h(X-D)+f_1\big(aD+b(X-D)\big)$, tudíž platí

$$f_2(X) = \max_{0\le D\le X}\{g(D)+h(X-D)+f_1\big(aD+b(X-D)\big)\} = \max\{h(X)+f_1(bX), g(X)+f_1(aX)\},$$

což je opět konvexní funkce. Dalším pokračováním můžeme ukázat, že f_k je konvexní pro všechna $k \in \mathbb{N}$. Pak i limitní funkce f této posloupnosti musí být konvexní. Vskutku, připusťme naopak, že existují $D_1, D_2 \in [0,X]$ a $\lambda \in [0,1]$ taková, že

$$\lambda f(D_1) + (1-\lambda) f(D_2) - f\big(\lambda D_1 + (1-\lambda) D_2\big) =: -\varepsilon < 0.$$

Pak pro dostatečně velké $k \in \mathbb{N}$ dostaneme ze stejnoměrnosti konvergence a konvexnosti funkce f_k, že platí

$$\begin{aligned}
-\varepsilon &= \lambda\big[f(D_1) - f_k(D_1) + f_k(D_1)\big] + (1-\lambda)\big[f(D_2) - f_k(D_2) + f_k(D_2)\big] - \\
&\quad - f\big(\lambda D_1 + (1-\lambda) D_2\big) + f_k\big(\lambda D_1 + (1-\lambda) D_2\big) - f_k\big(\lambda D_1 + (1-\lambda) D_2\big) = \\
&= \lambda\big[\underbrace{f(D_1) - f_k(D_1)}_{>-\varepsilon/6}\big] + (1-\lambda)\big[\underbrace{f(D_2) - f_k(D_2)}_{>-\varepsilon/6}\big] + \\
&\quad + \underbrace{\big[-f\big(\lambda D_1 + (1-\lambda) D_2\big) + f_k\big(\lambda D_1 + (1-\lambda) D_2\big)\big]}_{>-\varepsilon/6} + \\
&\quad + \underbrace{\lambda f_k(D_1) + (1-\lambda) f_k(D_2) - f_k\big(\lambda D_1 + (1-\lambda) D_2\big)}_{\ge 0} > -\varepsilon/2,
\end{aligned}$$

což je spor. Díky konvexnosti f pak podobně jako výše plyne

$$f(X) = \max_{0\le D\le X}\{g(D)+h(X-D)+f\big(aD+b(X-D)\big)\} = \max\{h(X)+f(bX), g(X)+f(aX)\},$$

čímž je tvrzení dokázáno. ■

V závěru předchozího důkazu jsme sice obdrželi rovnici pro f, která vypadá jednodušeji než původní rovnice (5.4.7), ale stále je[49] poměrně obtížné nalézt její řešení v obecném případě.

[47] „Připomeňme“, že nekonstantní konvexní funkce na kompaktním intervalu I nabývá svého maxima na tomto intervalu v některém z jeho krajních bodů (případně v obou), viz Větu 6.2.30.

[48] Viz Cvičení 6.2.14.

[49] . . . možná překvapivě. . .

Příklad 5.4.5

Ukažme, že spojité řešení rovnice

$$f(X) = \max\{c\,X^d + f(a\,X), e\,X^g + f(b\,X)\} \quad \& \quad f(0) = 0, \quad X \geq 0, \tag{5.4.13}$$

s parametry

$$a, b \in (0,1) \quad \& \quad c, d, e, g > 0 \quad \& \quad 0 < d < g,$$

je tvaru

$$f(X) = \begin{cases} \frac{c\,X^d}{1-a^d}, & 0 \leq X \leq \widehat{X}, \\ e\,X^g + f(b\,X), & X \geq \widehat{X}, \end{cases} \tag{5.4.14}$$

kde

$$\widehat{X} := \left[\frac{c\,(1-b^d)}{e\,(1-a^d)}\right]^{1/(g-d)}. \tag{5.4.15}$$

Jelikož $0 < b < 1$, můžeme $f(X)$ explicitně určit v intervalech

$$[\widehat{X}, \widehat{X}/b], \ldots, [\widehat{X}/b^n, \widehat{X}/b^{n+1}], \ldots \quad \text{pro } n = 0, 1, \ldots$$

Řešení. Označme symbolem A to, že v (5.4.13) vybereme funkci $c\,X^d + f(a\,X)$ a symbolem B, že zvolíme $e\,X^g + f(b\,X)$. Pak řešení S odpovídající posloupnosti optimálních výběrů z A a B můžeme symbolicky zapsat jako

$$S = A^{a_1}\,B^{b_1}\,A^{a_2}\,B^{b_2}\cdots,$$

kde $a_1, b_1, a_2, b_2, \cdots \in \mathbb{N} \cup \{0\}$ a A^{a_i} znamená, že a_i-krát po sobě zvolíme A (podobný význam má symbol B^{b_i}).
Připusťme pro tuto chvíli, že řešení má skutečně podobu popsanou v (5.4.14), tj. existuje bod $\widehat{X}$ takový, že $f(\widehat{X})$ je rovno buď A nebo B, zatímco hodnota $f(X)$ pro $0 < X \leq \widehat{X}$ má uvedený tvar, který získáme opakovanou volbou A (což symbolicky můžeme zapsat jako A^∞), neboť v takovém případě díky předpokladu $a < 1$ platí

$$f(X) = c\,X^d + f(a\,X) = c\,X^d + c\,(a\,X)^d + f(a^2\,X) + \cdots =$$
$$= c\,X^d + c\,(a\,X)^d + c\,(a^2\,X)^d + \cdots = \frac{c\,X^d}{1-a^d}.$$

V bodech $X > \widehat{X}$ volíme B, což nás přivede k tomu, že bod $\widehat{X}$ musí být určen rovností

$$B\,A^\infty = A^\infty,$$

kde výraz $B\,A^\infty$ nás díky předpokladu $b < 1$ přivede k vyjádření

$$f(X) = e\,X^g + f(b\,X) = e\,X^g + c\,(b\,X)^d + f(a\,b\,X) =$$
$$= e\,X^g + c\,(b\,X)^d + \frac{c\,(a\,b\,X)^d}{(1-a^d)} = e\,X^g + \frac{c\,b^d\,X^d}{1-a^d}.$$

Položíme-li oba získané výrazy do rovnosti, pak obdržíme vyjádření $\widehat{X}$ uvedené právě v (5.4.15).
Zbývá ukázat, že funkce $f(X)$ popsaná v (5.4.14) je opravdu hledaným řešením. Jelikož $g > d > 0$ a $b \in (0,1)$ platí

$$\frac{c\,X^d}{1-a^d} = \max\left\{\frac{c\,X^d}{1-a^d}, e\,X^g + \frac{c\,b^d\,X^d}{1-a^d}\right\}$$

pro dostatečně malá X, volíme díky jednoznačnosti řešení[50] vždy A pro malá X. Nyní můžeme hodnotu X postupně zvětšovat. Nechť Z je nejmenší hodnota X, pro kterou je volba B optimální. V tomto bodě je $B\,A^\infty = A^\infty$, což znamená, že $Z = \widehat{X}$. Nyní uvažme

[50] Stri
zení o
aproxi
ve Vět
ji získa
v důka

hodnotu $X > \widehat{X}$ a hledejme Y takové, že volby AB a BA vedou ke stejnému výsledku jakožto množina prvních dvou voleb. Pak máme

$$f_{AB}(X) = c\,X^d + e\,a^g\,X^g + f(a\,b\,X) \quad \& \quad f_{BA}(X) = e\,X^g + c\,b^d\,X^d + f(a\,b\,X),$$

takže hledanou hodnotou je

$$Y := \left[\frac{c\,(1-b^d)}{e\,(1-a^g)}\right]^{1/(g-d)}.$$

Jelikož $g > d$, bude $Y < \widehat{X}$. Proto z nerovnosti $f_{AB}(X) < f_{BA}(X)$ pro $X > Y$ vyplývá, že pro $X > \widehat{X}$ je volba AB plus další pokračování optimálními volbami horší než volba BA plus pokračování optimálními volbami. Proto volba A nemůže být použita pro $X > \widehat{X}$, pokud nenásleduje A^∞, což víme, že není možné. Tím je důkaz hotov.

Podobné tvrzení jako v případě Věty 5.4.4 platí i pro ostře konkávní funkce g a h.

Věta 5.4.6

Nechť jsou splněny předpoklady Věty 5.4.1 a funkce g, h jsou navíc ostře konkávní na $\mathbb{R}_+$. Pak funkce f je také ostře konkávní a optimální rozhodnutí v úloze (5.4.7) je jediné.

Důkaz. Jelikož $f_0 \equiv 0$ a funkce $g(D) + h(X-D)$ je ostře konkávní pro $D \in [0, X]$ má úloha $\max_{0\le D\le X}\{g(D) + h(X-D)\}$ jediné řešení[51], pro libovolné $X \ge 0$, kterým může být kterýkoli bod intervalu $[0, X]$. Nyní ukážeme, že funkce $f_1(X) = \max_{0\le D\le X}\{g(D) + h(X-D)\}$ je konkávní vzhledem k $X \ge 0$. Nechť proto $\widetilde{X}, \widehat{X} \ge 0$ jsou libovolná a označme proměnné $\widetilde{D} \in [0, \widetilde{X}]$ a $\widehat{D} \in [0, \widehat{X}]$, takže $D = \lambda\,\widetilde{D} + (1-\lambda)\,\widehat{D} \in [0, \lambda\,\widetilde{X} + (1-\lambda)\,\widehat{X}]$ pro libovolné $\lambda \in [0,1]$ a $\widetilde{D}, \widehat{D}$. Potom z konkávnosti g a h plyne

$$\begin{aligned}
f_1(\lambda\,\widetilde{X} + (1-\lambda)\,\widehat{X}) &= \max_{0\le D\le \lambda\,\widetilde{X}+(1-\lambda)\,\widehat{X}} \{g(D) + h\big(D - \lambda\,\widetilde{X} - (1-\lambda)\widehat{X}\big)\} = \\
&= \max_{\substack{0\le\widetilde{D}\le\widetilde{X}\\ 0\le\widehat{D}\le\widehat{X}}} \{g\big(\lambda\,\widetilde{D} + (1-\lambda)\,\widehat{D}\big) + h\big(\lambda\,\widetilde{D} + (1-\lambda)\,\widehat{D} - \lambda\,\widetilde{X} - (1-\lambda)\widehat{X}\big)\} \ge \\
&\ge \max_{\substack{0\le\widetilde{D}\le\widetilde{X}\\ 0\le\widehat{D}\le\widehat{X}}} \{\lambda\,g(\widetilde{D}) + (1-\lambda)\,g(\widehat{D}) + \lambda\,h(\widetilde{D}-\widetilde{X}) + (1-\lambda)\,h\big(\widehat{D}-\widehat{X}\big)\big) = \\
&= \lambda \max_{0\le\widetilde{D}\le\widetilde{X}} \{g(\widetilde{D}) + h(\widetilde{D}-\widetilde{X})\} + (1-\lambda) \max_{0\le\widehat{D}\le\widehat{X}} \{g(\widehat{D}) + h\big(\widehat{D}-\widehat{X}\big)\big) = \\
&= \lambda\,f_1(\widetilde{X}) + (1-\lambda)\,f_1(\widehat{X}),
\end{aligned}$$

tj. $f_1(X)$ je skutečně konkávní. Analogickým způsobem můžeme ukázat, že $f_2(X)$ je také konkávní díky konkávnosti $f_1\big(a\,D + b\,(X-D)\big)$ atd. včetně limitní funkce. Ostrost konkávnosti funkce f získáme podobným výpočtem jako v případě f_1 s využitím funkcionální rovnice (5.4.7) s tím rozdílem, že ostrá konkávnost funkcí g a h nám ve skutečnosti umožňuje psát ostrou nerovnost na konci druhého (a začátku třetího) řádku. Odtud také plyne jednoznačnost řešení úlohy (5.4.7). ■

[51] Jeho existence pochopitelně plyne z Weierstrassovy věty.

Předpoklad konkávnosti nám při doplnění dalších vlastností funkcí g a h umožňuje formulovat ještě preciznější výsledek.

Věta 5.4.7

Nechť jsou splněny předpoklady Věty 5.4.1 a funkce g, h jsou navíc ostře konkávní na $\mathbb{R}_+$, rostoucí, spojitě diferencovatelné a splňují

$$\frac{g'(0)}{1-a} > \frac{h'(0)}{1-b} \quad \& \quad h'(0) > \lim_{X\to\infty} g'(X) \quad \& \quad 0 \le a < b < 1. \tag{5.4.16}$$

Je-li $\widetilde{X}$ řešením rovnice

$$h'(0) = g'(X) + (b-a)\,g'(a\,X) + (b-a)\,a\,g'(a^2\,X) + \cdots + (b-a)\,a^k\,g'(a^{k+1}\,X) + \cdots,$$

pak optimální rozhodnutí v rovnici (5.4.7) se realizuje

Věta 5.4.7

(i) buď v bodě $D = X$ pro $0 \le X \le \tilde{X}$ a řešení je tvaru $f(x) = \sum_{k=0}^{\infty} g(a^k X)$;

(ii) nebo v bodě $D = D(X)$ pro $X \ge \tilde{X}$, přičemž $D(X)$ je funkce splňující nerovnost $0 < D(X) < X$ a vyhovující rovnici

$$g'(D) - h'(X-D) + (a-b)\, f'\big(aD + b(X-D)\big) = 0.$$

Podmínka (5.4.16) je jen jedna z několika možných kombinací nerovností mezi čísly $g'(0)$, $h'(0)$, a a b. Další vhodné varianty budou patrné z důkazu této věty uvedeného níže, přičemž jejich počet může být poloviční, jelikož záměna a a b vede jen k záměně D a $X - D$.

Důkaz Věty 5.4.7. Zvolíme-li $f_0 \equiv 0$, pak bude opět $f_1(X) = \max_{0\le D\le X}\{g(D) + h(X-D)\}$. Jelikož g' a h' jsou spojité funkce a z (5.4.16) plyne $g'(0) > h'(0)$, existuje $X > 0$ takové, že platí $\min_{0\le D\le X} g'(D) > \max_{0\le D\le X} h'(D)$. Proto je $g'(D) \ge h'(X-D)$ pro všechna $0 \le D \le X$, což znamená, že $[g(D)+h(D-X)]' > 0$, tj. funkce $g(D)+h(D-X)$ je rostoucí a její maximum se realizuje v pravém krajním bodě intervalu neboli $\max_{0\le D\le X}\{g(D)+h(X-D)\} = g(X)$. Budeme-li X zvětšovat dostaneme se k hodnotě $X = X_1$, která odpovídá jedinému řešení rovnice $g'(X) = h(0)$, jehož existence vyplývá ze druhé nerovnosti v (5.4.16), nerovnosti $g'(0) > h'(0)$ a ostré konkávnosti g (v kterémžto případě je g' klesající funkcí), tj. pro tuto hodnotu $X = X_1$ má rovnice $g'(D) = h'(X-D)$ řešení $D = X_1$. Pro větší hodnoty $X > X_1$ pak bude řešení této rovnice ležet uvnitř intervalu $[0, X]$ a označíme jej jako $D_1 = D_1(X_1)$, přičemž jeho existence plyne ze spojitosti g a h, zatímco jednoznačnost je důsledkem ostré konkávnosti těchto funkcí. Proto platí

$$f_1(X) = \begin{cases} g(X), & X \in [0, X_1], \\ g(D_1) + h(X - D_1), & X \ge X_1, \end{cases}$$

s derivací

$$f_1'(X) = \begin{cases} g'(X), & X \in [0, X_1), \\ \underbrace{\big[g'(D_1) - h'(X-D_1)\big]}_{=0} \frac{\mathrm{d}D_1}{\mathrm{d}X} + h'(X-D_1) = h'(X-D_1), & X > X_1. \end{cases} \tag{5.4.17}$$

Hodnotu $f_1'(X)$ jsme sice pro $X = X_1$ vynechali, avšak zjevně platí $\lim_{X\to X_1^-} f_1'(X) = g'(X_1)$ díky spojitosti g' a současně $\lim_{X\to X_1^+} f_1'(X) = \lim_{X\to X_1^+} h'(X-D_1) = h'(0) = g'(X_1)$ díky volbě X_1, tj. $f_1'(X)$ je spojitá. Navíc funkce f_1 je konkávní, jak jsme ukázali již v předchozí větě.

Nyní se zaměříme na druhou aproximaci

$$f_2(X) = \max_{0\le D\le X}\{g(D) + h(X-D) + f_1\big(aD + b(X-D)\big)\}, \tag{5.4.18}$$

u které nás zajímá chování funkce

$$\partial f_2(D) := \frac{\partial}{\partial D} f_2(D) = g'(D) - h'(X-D) + (a-b)\, f_1'\big(aD + b(X-D)\big).$$

Protože díky (5.4.16)–(5.4.17) a monotonii funkce h platí

$$\begin{aligned} g'(0) - h'(0) + (a-b)\, f_1'(0) &= g'(0) - h'(0) + (a-b)\, g'(0) = \\ &= g'(0)\,(1+a-b) - h'(0) > \left[\frac{(1+a-b)\,(1-a)}{1-b} - 1\right] h'(0) > 0, \end{aligned}$$

existuje $X > 0$ takové, že

$$\min_{0\le D\le X}\{g'(D) + (a-b)\, f_1'\big(aD + b(X-D)\big)\} \ge \max_{0\le D\le X} h(D),$$

a tudíž $\partial f_2(D) > 0$ pro $D \in [0, X]$ a X dostatečně malé. To znamená, že pro tato malá X je maxima v (5.4.18) dosaženo v pravém krajním bodě intervalu, tj. pro $D = X$. S rostoucím X

bude existovat hodnota X_2 taková, že $\partial f_2(D) = \partial f_2(X_2) = 0$. Toto X_2 je řešením rovnice $g'(D) = h'(0) + (b-a)\, f_1'(a\,D)$. Porovnáním rovností

$$g'(X_1) = h(0) \quad \text{a} \quad g'(X_2) = h'(0) + (b-a)\, f_1'(a\,X_2)$$

vidíme, že $0 < X_2 < X_1$. Proto se rovnice pro X_2 redukuje na

$$g'(X_2) = h'(0) + (b-a)\, g'(a\,X_2)$$

a řešení na pravé straně (5.4.18) se realizuje pro $D = X$ v případě $X \in [0, X_2]$, zatímco pro $X \geq X_2$ to je v bodě $D_2 = D_2(X)$, který je jediným řešením rovnice

$$g'(D) = h'(X-D) + (b-a)\, f_1'\big(a\,D + b\,(X-D)\big).$$

Zdůrazněme také, že $D_2 < D_1$, což vyplývá z porovnání předchozí rovnosti určující D_2 a rovnosti $g'(D) = h'(X-D)$ určující hodnotu D_1 společně s monotonií funkce g'. Tímto dostáváme hledanou funkci

$$f_2(X) = \begin{cases} g(X) + f_1(a\,X) = g(X) + g(a\,X), & X \in [0, X_2], \\ g(D_2) + h(X-D_2) + f_1\big(a\,D_2 + b\,(X-D_2)\big), & X > X_2, \end{cases}$$

a její derivaci

$$f_2'(X) = \begin{cases} g'(X) + a\,g'(a\,X), & X \in [0, X_2], \\ h'(X-D_2) + b\,f_1'\big(a\,D_2 + b\,(X-D_2)\big), & X > X_2, \end{cases} \tag{5.4.19}$$

přičemž $f_2'(X)$ je spojitá pro $X = X_2$. Vskutku, to plyne z toho, že

$$\lim_{X \to X_2^-} f_2'(X) = g'(X_2) + a\,f_1'(a\,X_2) \quad \text{a zároveň} \quad \lim_{X \to X_2^+} f_2'(X) = h'(0) + b\,f_1'(a\,X_2),$$

což vzhledem k definici hodnoty X_2 dává

$$\lim_{X \to X_2^-} f_2'(X) - \lim_{X \to X_2^+} f_2'(X) = g'(X_2) - h'(0) + (a-b)\, f_1'(a\,X_2) = 0.$$

Poznamenejme také, že před chvílí zmíněná nerovnost $X_2 < X_1$ je de facto důsledkem toho, že $f_2'(X) \geq f_1'(X)$, na což se nyní ještě zaměříme[52]. Již víme, že body X_1 a X_2 nám vytvoří trojici disjunktních intervalů $[0, X_2]$, $[X_2, X_1]$ a $[X_1, \infty)$. Pro $X > X_1$ je

$$f_1'(X) = h'(X-D_1) \quad \text{a} \quad f_2'(X) = h'(X-D_2) + b\,f_1'\big(a\,D_2 + b\,(X-D_2)\big), \tag{5.4.20}$$

kde D_1 a D_2 jsou jediné hodnoty splňující

$$g'(D_1) = h'(X-D_1) \quad \text{a} \quad g'(D_2) = h'(X-D_2) - (a-b)\, f_1'\big(a\,D_2 + b\,(X-D_2)\big). \tag{5.4.21}$$

S využitím (5.4.21) získáme z (5.4.20) rovnosti

$$f_1'(X) = h'(X-D_1) = \frac{1}{b-a}\Big[b\,\underbrace{h'(D_1)}_{=g'(D_1)} - a\,h'(X-D_1)\Big] = \frac{b\,g'(D_1) - a\,h'(X-D_1)}{b-a}$$

a také

$$\begin{aligned} f_2'(X) &= h'(X-D_2) + b\,f_1'\big(a\,D_2 + b\,(X-D_2)\big) = \\ &= \frac{1}{b-a}\Big[\underbrace{b\,h'(X-D_2) - b\,(a-b)\,f_1'\big(a\,D_2 + b\,(X-D_2)\big)}_{=b\,g'(D_2)} - a\,h'(X-D_2)\Big] = \\ &= \frac{b\,g'(D_2) - a\,h'(X-D_2)}{b-a}. \end{aligned}$$

Jelikož funkce $(b-a)^{-1}\big(b\,g'(D) - a\,h'(X-D)\big)$ je klesající vzhledem k D pro $D \in [0, X]$ díky vlastnostem funkcí g a h, plyne odtud a z faktu $D_2 < D_1$, že skutečně $f_2'(X) > f_1'(X)$. Pro $X \in [0, X_2]$

[52] Samozřejmě zde nebudeme provádět důkazem kruhem, takže nerovnost $X_2 < X_1$ plyne z argumentů výše nikoli z následující části důkazu, která se o tento fakt již opírá.

je nerovnost $f_2'(X) > f_1'(X)$ patrná z vyjádření (5.4.17) a (5.4.19), neboť g je rostoucí funkce. Zbývá interval $[X_2, X_1]$. Zde platí

$$f_1'(X) = g'(X) \quad \text{a} \quad f_2'(X) = \frac{b\,g'(D_2) - a\,h'(X - D_2)}{b - a}.$$

Protože obě funkce g', h' jsou klesající a $g'(X) \geq h'(0)$ jakožto důsledek nerovnosti $g'(D) \geq h'(X - D)$ pro $D \in [0, X]$ a $X \in [0, X_1]$, kterou jsme ukázali v první části důkazu, platí

$$f_2'(X) \geq \frac{b\,g'(X) - a\,h'(0)}{b - a} > g'(X),$$

čímž je důkaz nerovnosti $f_2'(X) \geq f_1'(X)$ hotov.

V tuto chvíli již máme všechny potřebné nástroje k tomu, abychom celou tuto konstrukci pomocí matematické indukce rozšířili pro libovolné $n \in \mathbb{N}$, čímž bychom získali trojici posloupností $\{X_n\}_{n=1}^{\infty}$, $\{f_n'(X)\}_{n=1}^{\infty}$ a $\{D_n(X)\}_{n=1}^{\infty}$ takových, že pro $X \leq X_{n+1}$ se hledané maximum v n-tém kroku

$$f_{n+1}(X) = \max_{0 \leq D \leq X} \{g(D) + h(X - D) + f_n\big(a\,D + b\,(X - D)\big)\}$$

realizuje v pravém krajním bodě $D = X$, zatímco pro $X > X_{n+1}$ to je ve stacionárním bodě $D_{n+1} = D_{n+1}(X)$ ležícím uvnitř intervalu $[0, X]$. Existenci X_{n+1} bychom ukázali podobně jako v předchozí části s tím, že se tentokrát jedná o řešení rovnice

$$g'(D) + h'(0) + (a - b)\,f_n'(a\,D) = 0$$

neboli

$$g'(D) - h'(0) + (a - b)\big(g'(a\,D) + a\,g'(a^2\,D) + \cdots + a^{n-2}\,g'(a^{n-1}\,D)\big) = 0.$$

Následně bychom ukázali, že $X_{n+1} < X_n$, z čehož bychom odvodili $f_n'(X) > f_{n-1}'(X)$ a konečně i $D_n > D_{n+1}$, tj. celkem máme

$$X_1 > X_2 > \cdots > X_n > \cdots > 0, \quad f_1'(X) \leq f_2'(X) \leq f_3'(X) \leq \cdots \leq f_n'(X) \leq \cdots \quad \text{a} \quad D_1 > D_2 > \cdots.$$

Tvrzení věty pak plyne z monotonie těchto posloupností, která zaručuje existenci jejich limit. ■

S využitím předchozího tvrzení můžeme vyřešit následující příklad.

Příklad 5.4.8

Určeme řešení rovnice

$$f(X) = \max_{0 \leq D \leq X} \{c\,D - D^2 + d\,(X - D) - (X - D)^2 + f\big(a\,D + b(X - D)\big)\} \quad \& \quad f(0) = 0 \quad (5.4.22)$$

v intervalu[53] $\big[0, \min\{c/2, d/2\}\big]$, kde $c, d > 0$ a $0 \leq b \leq a \leq 1$.

Řešení. Toto zadání odpovídá volbě $g(D) = c\,D - D^2$, $g'(0) = c$, $\lim_{D\to\infty} g(D) = -\infty$, $h(D) = d\,D - D^2$, $h'(0) = d$ a $\lim_{D\to\infty} h(D) = -\infty$ ve Větě 5.4.7. V závislosti na znaménku výrazu $\frac{c}{1-a} - \frac{d}{1-b}$ dostaneme následující tři případy.

(i) Je-li $\frac{c}{1-a} = \frac{d}{1-b}$, pak

$$f(D) = \frac{(c - d)\,a + d}{1 - b + (b - a)\,a}\,D - \frac{\alpha^2 + (1 - \alpha)^2}{1 - \big[(a - b)\,\alpha + b\big]^2}\,D^2,$$

kde

$$\alpha = \left\{1 + \frac{a^2 - b^2}{2\,(1 - a\,b)} + \sqrt{1 + \left(\frac{a^2 - b^2}{2\,(1 - a\,b)}\right)^2}\right\}^{-1}.$$

(ii) Je-li $\frac{c}{1-a} < \frac{d}{1-b}$, pak

$$f(D) = \frac{d}{1 - b}\,D - \frac{1}{1 - b^2}\,D^2$$

[53] Toto je maximální interval, na kterém jsou odpovídající g a h z Věty 5.4.7 rostoucí.

pro $0 \le D \le \min\{\lambda, c/2, d/2\}$, kde

$$\lambda = \frac{(1+b)\left[d(1-a)-c(1-b)\right]}{2(1-ab)},$$

přičemž pro $\lambda < \min\{c/2, d/2\}$ můžeme využít rovnost (5.4.22) jako rekurzivní vzorec k tomu, abychom získali $f(D)$ na celém požadovaném intervalu.

(iii) Je-li $\frac{c}{1-a} > \frac{d}{1-b}$, pak

$$f(D) = \frac{c}{1-a}D - \frac{1}{1-a^2}D^2$$

pro $0 \le D \le \min\{\mu, c/2, d/2\}$, kde

$$\mu = \frac{(1+a)\left[c(a-b)-d(1-a)+c(1-a)\right]}{2(1-ab)}.$$

Rovnice pro určení $\widetilde{X}$ v případě (iii) je

$$\begin{aligned} d &= c - 2X + (a-b)\left(c - 2aX + ac - 2a^3X + a^2c - 2a^5X + \cdots\right) = \\ &= c - 2X + (a-b)\left(\frac{c}{1-a} - \frac{2aX}{1-a^2}\right) = c - 2X + (a-b)\,\frac{c(1+a)-2aX}{1-a^2}, \end{aligned}$$

z čehož přímým výpočtem získáme

$$\widetilde{X} = \frac{d(1-a^2)-c(1+a)(1-b)}{2(ab-1)}.$$

Pro $X < \min\{\widetilde{X}, c/2, d/2\}$ (neboť předpoklady věty jsou splněny pro $X \le c/2$ a $X \le d/2$) je maximum v pravém krajním bodě X. Odtud plyne

$$\begin{aligned} f(X) &= g(X) + g(aX) + g(a^2X) + \cdots + g(a^kX) + \cdots = \\ &= cX - X^2 + caX - a^2X^2 + ca^2X - a^4X^2 + \cdots = \frac{cX}{1-a} - \frac{X^2}{1-a^2}. \end{aligned}$$

Příklad 5.4.9

Určeme řešení základní rovnice dynamického programování

$$f(X) = \max_{0 \le D \le X}\left\{g(D) + h(X-D) + f\big(aD + b(X-D)\big)\right\},$$

je-li $g(y) = y^2$, $h(y) = 2y^2 - y$ a $a = b = 1/2$.

Řešení. Funkce g a h jsou konvexní, takže maximum je nabyto v 0 nebo v X, tj.

$$f(X) = \max\left\{g(0) + h(X) + f(bX), h(0) + g(X) + f(aX)\right\} = f(aX) + \underbrace{\max\left\{g(X), h(X)\right\}}_{=:F(X)}.$$

Pak

$$f(X) = \sum_{k=0}^{\infty} F(X/2^k),$$

přičemž

$$F(X) = \max\left\{X^2, 2X^2 - X\right\} = \begin{cases} X^2, & X \in [0,1], \\ 2X^2 - X, & X \ge 1. \end{cases}$$

Je-li $X \in [0,1]$, pak $X/2^k \le 1$, takže

$$f(X) = \sum_{k=0}^{\infty} (X/2^k)^2 = \frac{X^2}{1-1/4} = \frac{4}{3}X^2.$$

Je-li $2^m \le X < 2^{m+1}$ pro nějaké $m \in \{0,1,\dots\}$, pak $X/2^m \ge 1$ a současně $X/2^{m+1} < 1$, takže

$$\begin{aligned} f(X) &= \sum_{k=0}^{\infty} F(X/2^k) = \sum_{k=0}^{m} F(\underbrace{X/2^k}_{\ge 1}) + \sum_{k=m+1}^{\infty} F(\underbrace{X/2^k}_{>1}) = \\ &= \sum_{k=0}^{m} \{2(X/2^k)^2 - (X/2^k)\} + \sum_{k=m+1}^{\infty} (X/2^k)^2 = \\ &= \frac{(1-2^{m+1})\,X}{2^m} + \frac{(2^{2m+2}-1/2)\,X^2}{3\cdot 2^{2m-1}}. \end{aligned}$$

Příklad 5.4.10

S pomocí Věty 5.4.7 určeme (tentokrát alespoň částečně) řešení základní rovnice dynamického programování

$$f(X) = \max_{0\le D\le X} \{g(D) + h(X-D) + f\big(aD + b(X-D)\big)\}, \tag{5.4.23}$$

je-li $g(y) = 2y - y^2$, $h(y) = y - y^2$, $a = 1/2$ a $b = 2/3$.

Řešení. Funkce g, h jsou konkávní a platí $g'(0) = 2$, $h'(0) = 1$, takže předpoklady Věty 5.4.7 jsou splněny. Víme proto, že maximum v definici f je pro $X < \widetilde{X}$ dosaženo v $D = X$, kde $\widetilde{X}$ je řešením rovnice

$$h'(0) = g'(X) - (b-a)\sum_{k=1}^{\infty} a^{k-1} g'(a^k X) \quad \text{neboli} \quad 1 = 2 - 2X - \frac{1}{6}\sum_{k=1}^{\infty} \frac{2 - 2\,(1/2)^k\,X}{2^{k-1}}.$$

Výraz na pravé straně můžeme upravit upravíme

$$\begin{aligned} 2 - 2X - \frac{1}{6}\sum_{k=1}^{\infty} \frac{2-2\,(1/2)^k\,X}{2^{k-1}} &= 2 - 2X - \frac{1}{6}\sum_{k=1}^{\infty}\left(\frac{1}{2^{k-2}} - \frac{1}{2^{2k-2}}X\right) = \\ &= 2 - 2X - \frac{1}{6}\left(\frac{2}{1-1/2} - \frac{1}{1-1/4}X\right) = 2 - 2X - \frac{1}{6}\left(4 - \frac{4}{3}X\right) = \frac{4}{3} - \frac{16}{9}X, \end{aligned}$$

čímž dostaneme rovnici

$$1 = \frac{4}{3} - \frac{16}{9}X,$$

jejímž řešením je $\widetilde{X} = 3/16$. Proto řešením rovnice (5.4.23) pro $X < \widetilde{X}$ je

$$f(X) = 2X - X^2 + \sum_{k=1}^{\infty} \frac{1}{2^{k-1}}X - \sum_{k=1}^{\infty}\left(\frac{X}{2^k}\right)^2 = 2X - X^2 + 2X - \frac{X^2}{3} = 4X - \frac{4X^2}{3}.$$

Bylo by samozřejmě žádoucí mít k dispozici explicitní řešení co největšího počtu takovýchto rovnic. Avšak skutečná důležitost explicitních řešení jednoduchých rovnic tkví v jejich použití jako přibližných řešení pro složitější rovnice a v poskytnutí vodítek k nalezení základních vlastností optimálních rozhodnutí pro komplikovanější procesy.

V předchozí části jsme odvodili explicitní řešení pro případ, kdy funkce g a h jsou kvadratické a obecné monomy $c\,X^d$. Zdůrazněme, že aproximace $g(X)$ pomocí $c\,X^d$ je ekvivalentní aproximaci složené funkce $\ln g(\mathrm{e}^X)$ pomocí $d\,X + \ln c$, tj. pomocí přímky, což je mnohem přístupnější. Všimněme si, že jak se X mění, můžeme měnit naše aproximační křivky, abychom získali lepší shody, pokud si přejeme bližší aproximace. Zmiňme také, že obecně je aproximace nejužitečnější jako aproximace v prostoru rozhodnutí než v prostoru funkcí.

Abychom mohli použít aproximační techniky, potřebujeme odhad rozdílu mezi řešeními[54]

[54] Existenci a jednoznačnost řešení těchto rovnic lze zaručit snadnou úpravou Věty 5.4.1. Hovoříme-li o řešení, máme na mysli *spojité* řešení nebo, obecněji řečeno, to řešení, jehož existenci garantuje právě zmíněná věta.

dvou rovnic

$$f(X) = \max_{0\le D\le X}\big[u(D,X) + f\big(aD + b(X-D)\big)\big] \quad \& \quad f(0) = 0, \tag{5.4.24}$$

$$F(X) = \max_{0\le D\le X}\big[v(D,X) + F\big(aD + b(X-D)\big)\big] \quad \& \quad F(0) = 0 \tag{5.4.25}$$

pomocí rozdílu mezi funkcemi $u(D,X)$ a $v(D,X)$, což nás přivede k následující klasické *Větě o stabilitě v dynamickém programování.*

VĚTA 5.4.11 VĚTA O STABILITĚ V DP

Nechť jsou dána čísla $a, b \in (0,1)$ a nechť funkce $u(D,X)$, $v(D,X)$ jsou spojité vzhledem k D i X pro $D, X \ge 0$ a současně platí $\sum_{n=0}^{\infty} m(c^n Z) < \infty$, kde $c := \max\{a,b\}$ a

$$m(Z) := \max_{0\le X\le Z}\Big\{\max_{0\le D\le X}\max\{|u(D,X)|, |v(D,X)|\}\Big\}.$$

Je-li

$$\sum_{n=0}^{\infty} d(c^n Z) < \infty \quad \text{pro} \quad d(Z) := \max_{0\le X\le Z}\Big\{\max_{0\le D\le X}|u(D,X) - v(D,X)|\Big\},$$

pak pro spojitá řešení f a F rovnic (5.4.24)–(5.4.25) platí

$$|f(X) - F(X)| \le \sum_{n=0}^{\infty} d(c^n Z).$$

Důkaz. Definujme postupně

$$f_1(X) := \max_{0\le D\le X} u(D,X) \quad \& \quad f_{N+1}(X) := \max_{0\le D\le X}\big\{u(D,X) + f_N\big(aD + b(X-D)\big)\big\},$$

$$F_1(X) := \max_{0\le D\le X} v(D,X) \quad \& \quad F_{N+1}(X) := \max_{0\le D\le X}\big\{v(D,X) + F_N\big(aD + b(X-D)\big)\big\}.$$

Již víme, že posloupnost $\{f_N(X)\}_{N=0}^{\infty}$ konverguje k $f(X)$ a $\{F_N(X)\}_{N=0}^{\infty}$ k $F(X)$. Odhadneme-li nejdříve rozdíl mezi $f_1(X)$ a $F_1(X)$, pak získáme

$$\big|f_1(X) - F_1(X)\big| \le \max_{0\le D\le X}\big|u(D,X) - v(D,X)\big| \le d(X).$$

Podobně s využitím trojúhelníkové nerovnosti dostaneme

$$\big|f_{N+1}(X) - F_{N+1}(X)\big| \le \max_{0\le D\le X}\big|f_N\big(aD + b(X-D)\big) - F_N\big(aD + b(X-D)\big)\big| + \\ + \max_{0\le D\le X}\big|u(D,X) - v(D,X)\big|.$$

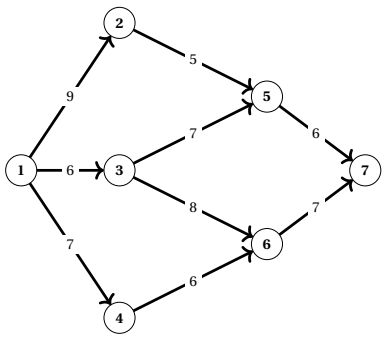
Odtud se s pomocí indukce dopracujeme k nerovnosti

$$\big|f_{N+1}(X) - F_{N+1}(X)\big| \le \sum_{k=0}^{N} d(c^k X),$$

z čehož pro $N \to \infty$ plyne požadované tvrzení. ∎

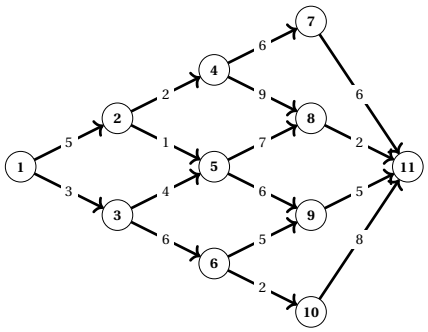
Cvičení

5.4.1. Určete nejkratší cestu spojující uzly 1 a 7 v grafu (ocenění hran=vzdálenost)



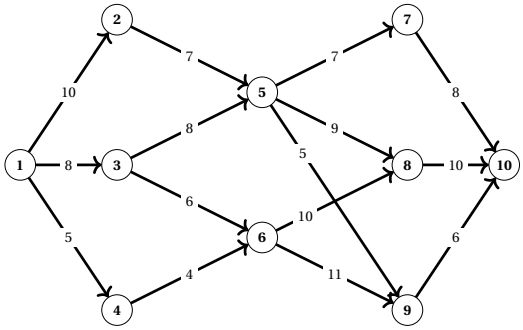
Řešení: 1 ⇝ 3 ⇝ 5 ⇝ 7 *s hodnotou* 19.

5.4.2. Určete nejkratší cestu spojující uzly 1 a 11 v grafu (ocenění hran=vzdálenost).



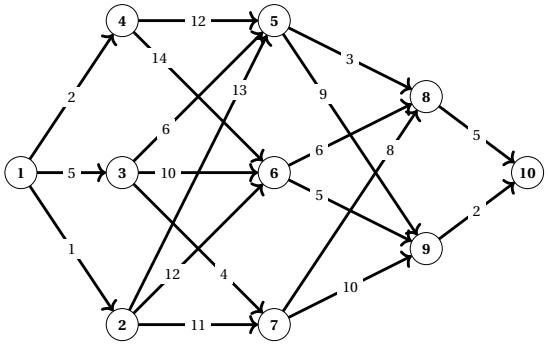
Řešení: 1 ⇝ 2 ⇝ 5 ⇝ 8 ⇝ 11 *s hodnotou* 15.

5.4.3. Určete nejkratší cestu spojující uzly 1 a 10 v grafu (ocenění hran=vzdálenost)



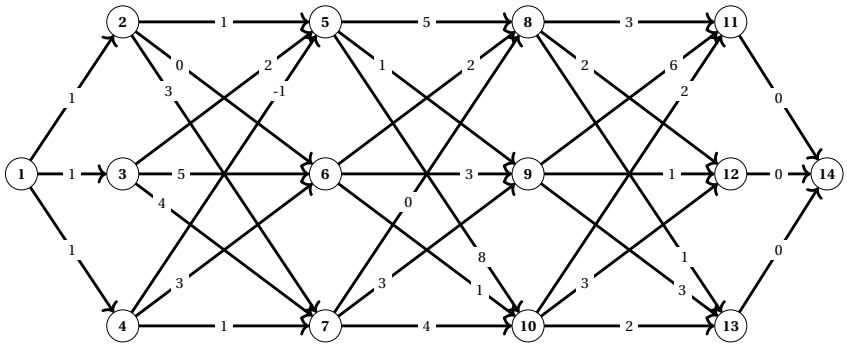
Řešení: 1 ⇝ 4 ⇝ 6 ⇝ 9 ⇝ 10 *s hodnotou* 26.

5.4.4. Určete nejkratší cestu spojující uzly 1 a 10 v grafu (ocenění hran=vzdálenost).



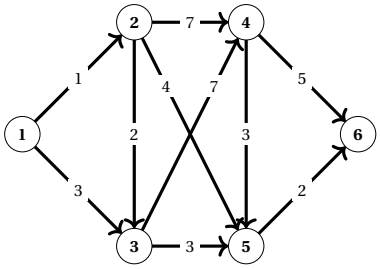
Řešení: 1 ⇝ 3 ⇝ 5 ⇝ 8 ⇝ 10 *s hodnotou* 19.

5.4.5. Určete nejkratší cestu spojující uzly 1 a 14 v grafu (ocenění hran=vzdálenost).



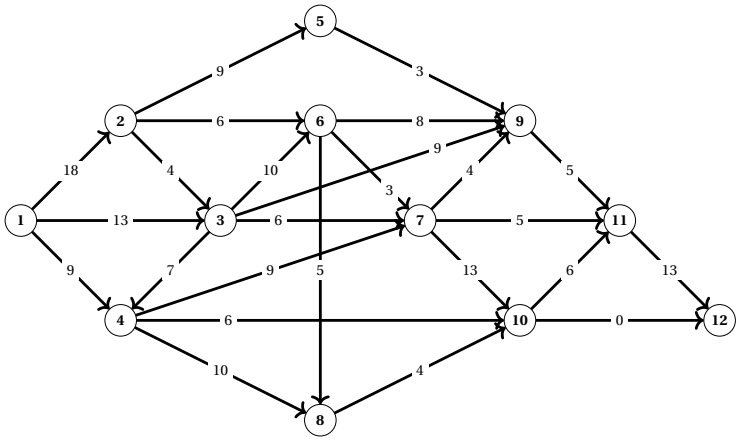
Řešení: 1 ⇝ 4 ⇝ 5 ⇝ 9 ⇝ 12 ⇝ 14 *s hodnotou* 2.

5.4.6. Určete nejkratší cestu spojující uzly 1 a 6 v grafu (ocenění hran=vzdálenost)

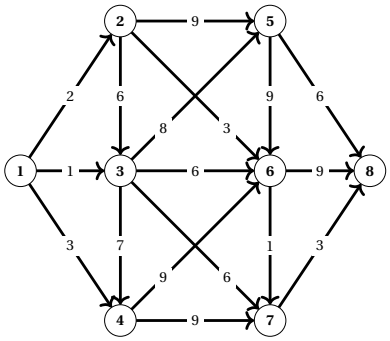


Řešení: 1 ⇝ 2 ⇝ 5 ⇝ 6 *s hodnotou* 7.

5.4.7. Určete nejdelší cestu spojující uzly 1 a 12 v grafu (ocenění hran=zisk).

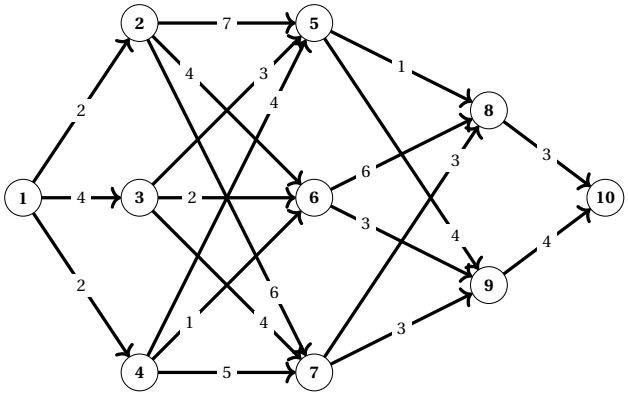


5.4.8. Určete nejkratší cestu spojující uzly 1 a 8 v grafu (ocenění hran=vzdálenost).



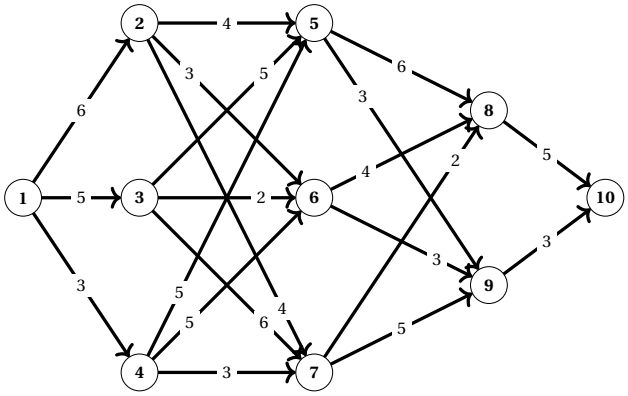
Řešení: 1 ⇝ 2 ⇝ 6 ⇝ 7 ⇝ 8 *s hodnotou* 9.

5.4.9. Určete nejkratší cestu spojující uzly 1 a 10 v grafu (ocenění hran=vzdálenost).



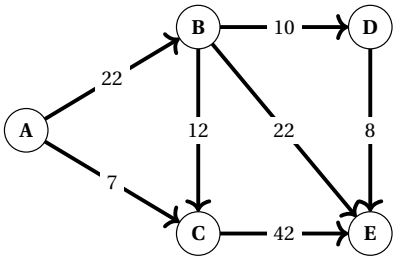
Řešení: 1 ⇝ 4 ⇝ 5 ⇝ 8 ⇝ 10 *nebo* 1 ⇝ 4 ⇝ 6 ⇝ 9 ⇝ 10 *s hodnotou* 10.

5.4.10. Určete nejkratší cestu spojující uzly 1 a 10 v grafu (ocenění hran=vzdálenost).

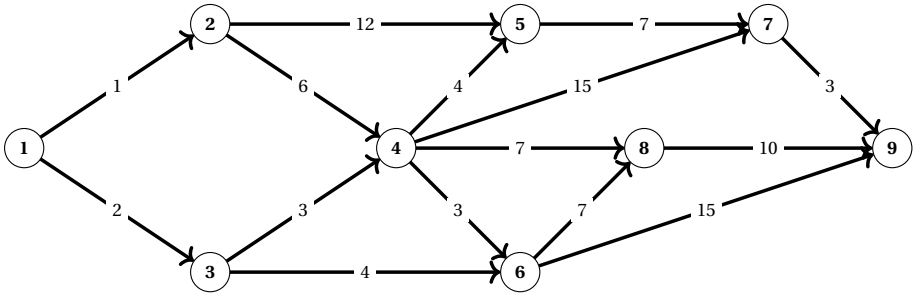


Řešení: 1 ⇝ 3 ⇝ 6 ⇝ 9 ⇝ 10 *nebo* 1 ⇝ 4 ⇝ 7 ⇝ 8 ⇝ 10 *s hodnotou* 13.

5.4.11. Určete nejkratší cestu spojující uzly *A* a *E* v grafu (ocenění hran=vzdálenost).

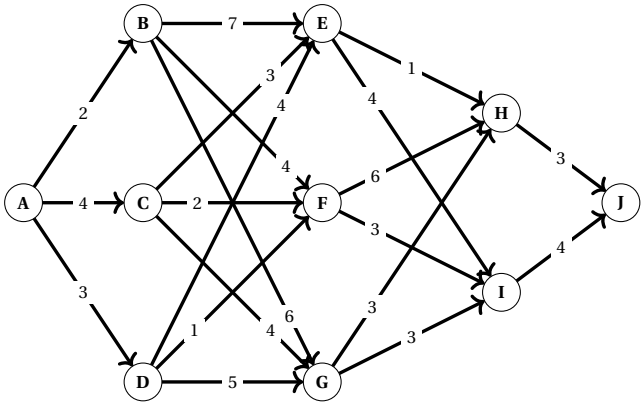


5.4.12. Určete nejkratší cestu spojující uzly 1 a 9 v grafu (ocenění hran=vzdálenost).



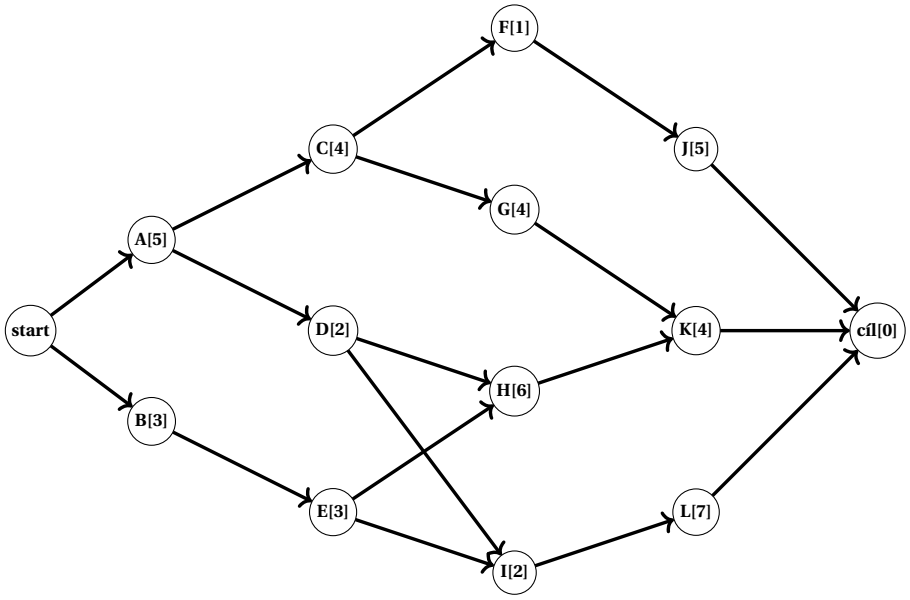
Řešení: 1 ⇝ 3 ⇝ 4 ⇝ 5 ⇝ 7 ⇝ 9 *s hodnotou* 19.

5.4.13. Určete nejkratší cestu spojující uzly *A* a *J* v grafu (ocenění hran=vzdálenost).



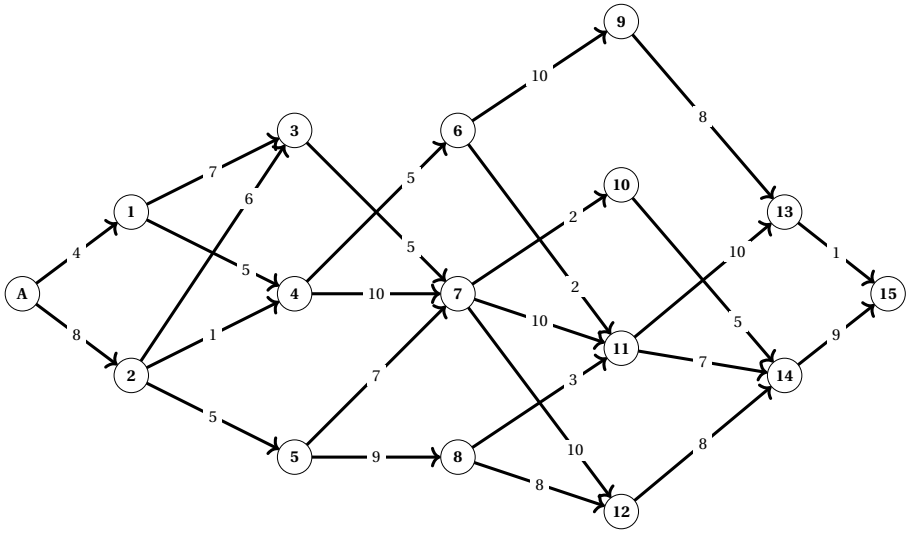
Řešení: $A \rightsquigarrow D \rightsquigarrow E \rightsquigarrow H \rightsquigarrow J$ *nebo* $A \rightsquigarrow D \rightsquigarrow F \rightsquigarrow I \rightsquigarrow J$ *nebo* $A \rightsquigarrow C \rightsquigarrow E \rightsquigarrow H \rightsquigarrow J$ *s hodnotou* 11.

5.4.14. Určete nejdelší cestu spojující uzly „start“ a „cíl“ v grafu (číslo v závorkách značí čas potřebný pro danou aktivitu).



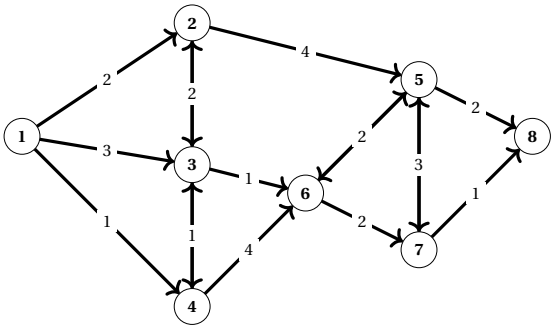
Řešení: start$\rightsquigarrow A \rightsquigarrow C \rightsquigarrow G \rightsquigarrow K \rightsquigarrow$*cíl nebo*
start$\rightsquigarrow A \rightsquigarrow D \rightsquigarrow H \rightsquigarrow K \rightsquigarrow$*cíl*
s hodnotou 17.

5.4.15. Určete nejkratší cestu spojující uzly 1 a 15 v grafu (ocenění hran=vzdálenost).



Řešení: $A \rightsquigarrow 1 \rightsquigarrow 4 \rightsquigarrow 6 \rightsquigarrow 11 \rightsquigarrow 13 \rightsquigarrow 15$ *nebo*
$A \rightsquigarrow 2 \rightsquigarrow 4 \rightsquigarrow 6 \rightsquigarrow 11 \rightsquigarrow 13 \rightsquigarrow 15$
s hodnotou 27.

5.4.16. Určete nejkratší cestu spojující uzly 1 a 8 v grafu (ocenění hran=vzdálenost).



Řešení: $1 \rightsquigarrow 4 \rightsquigarrow 3 \rightsquigarrow 6 \rightsquigarrow 7 \rightsquigarrow 8$ *s hodnotou* 6.

5.4.17. Nadnárodní technologická společnost má detailně propracovaný program vstupního školení, které musí absolvovat každý nový zaměstnanec. Toto školení se skládá ze 4 fází, z nichž každá obsahuje různé moduly. Fáze a jednotlivé moduly s časem uvedeným ve dnech nutných pro jejich dokončení jsou shrnuty v následující tabulce.

Fáze I	Fáze II	Fáze III	Fáze IV
A→13	E→3	H→12	L→10
B→10	F→6	I→6	M→5
C→20	G→5	J→7	N→13
D→17		K→10	

Avšak moduly dokončené v odpovídající fázi mohou být následovány pouze některými moduly z následující fáze. Podrobný soupis možných návazností je uveden v tabulkách níže. Odtud se např. dočteme, že po modulu A dokončeném za 13 dní ve fázi I může následovat pouze modul F nebo G ve fázi II, zatímco po absolvování modulu L, M nebo N školení končí.

modul	možný následující modul
A	F, G
B	F
C	G
D	E, G
E	H, I, J, K
F	H, K
G	J, K

modul	možný následující modul
H	L, M
I	L, M
J	M, N
K	N
L	konec
M	konec
N	konec

(a) Určete posloupnost modulů, která zabere minimální čas pro dokončení.

(b) Pokud jste právě dokončili modul F, jak pokračovat dále, aby bylo potřeba minimum času?

Řešení: (a) $A \rightsquigarrow G \rightsquigarrow J \rightsquigarrow M$*; (b)* $F \rightsquigarrow H$.

5.4.18. Vyřešte „problém batohu“, jehož kapacita je 5 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4
cena	3	4	5	6
váha	2	3	4	5

Řešení: 1 & 2.

5.4.19. Vyřešte „problém batohu“, jehož kapacita je 6 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4
cena	6	3	5	1
váha	2	4	2	3

Řešení: 1 & 3.

5.4.20. Vyřešte „problém batohu“, jehož kapacita je 7 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4
cena	10	7	25	24
váha	2	1	6	5

Řešení: 1 & 4.

5.4.21. Vyřešte „problém batohu“, jehož kapacita je 8 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4
cena	15	10	9	5
váha	1	5	3	4

Řešení: 1 & 3 & 4.

5.4.22. Vyřešte „problém batohu“, jehož kapacita je 10 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4
cena	10	40	30	50
váha	5	4	6	3

Řešení: 2 & 4.

5.4.23. Vyřešte „problém batohu“, jehož kapacita je 6 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4	5
cena	25	20	15	40	50
váha	3	2	1	4	5

Řešení: 3 & 5.

5.4.24. Vyřešte „problém batohu“, jehož kapacita je 11 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4	5
cena	1	6	18	22	28
váha	1	2	5	6	7

Řešení: 3 & 4.

5.4.25. Vyřešte „problém batohu“, jehož kapacita je 12 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4	5	6
cena	1	6	5	15	8	10
váha	3	7	1	6	8	4

Řešení: 3 & 4 & 6.

5.4.26. Vyřešte „problém batohu“, jehož kapacita je 15 [jednotek]. Dostupné věci a jejich ceny [desítky Kč] s váhami jsou uvedeny v následující tabulce.

věc č.	1	2	3	4	5	6	7	8	9	10
cena	18	20	17	19	25	21	27	23	25	24
váha	1	3	7	4	8	9	6	10	2	5

Řešení: 1 & 2 & 4 & 9 & 10.

5.4.27. Vyřešte „problém batohu“, jehož kapacita je 10 [jednotek], přičemž tentokrát je dostupné libovolné množství každého druhu. Ceny [desítky Kč] a váhy jednotlivých druhů jsou uvedeny v následující tabulce.

druh č.	1	2	3	4
cena	7	9	2	15
váha	3	4	1	7

Řešení: 1 & 1 & 2.

5.4.28. Vyřešte „problém plnění kontejneru“, jehož kapacita je 11 [jednotek]. Dostupný náklad a jednotlivé ceny [stovky Kč] s váhami jsou uvedeny v následující tabulce.

typ č.	1	2	3	4
cena	18	25	30	20
váha	2	4	5	3
dostupné množství	5	2	2	3

Řešení: typ 1=4 ks, t2=0, t3=0, t4=1.

5.4.29. Jeden z Vašich kolegů se rozhodl, že dne 1. května (ačkoli/protože je Svátek práce) se začne chystat ke státní závěrečné zkoušce a vypracuje si některé otázky. Na výběr má různé počty otázek z různých oblastí a přípravou by chtěl strávit maximálně 8 hodin. Otázky si ohodnotil způsobem uvedeným v následující tabulce (ohodnocení=užitek z vypracování jedné otázky).

	počet dostupných otázek	počet hodin nutných k přípravě jedné otázky	ohodnocení
matematická analýza (MA)	5	1	3
lineární algebra (LA)	3	2	4
pravděpodobnost a statistika (PaS)	2	3	10

Určete skladbu vypracovaných otázek, které přinesou největší užitek.

Řešení: 2×MA & 0×LA & 2×PaS.

5.4.30. Při odbavování pravidelného létu ČSA Praha – Brno se zjistilo, že je v letadle ještě volná kapacita 2000 kg. ČSA proto nabídly dopravním společnostem v bleskové aukci tuto kapacitu pro transport jejich zásilek. ČSA si mohou vybrat z nabídky zásilek uvedené v následující tabulce.

typ zásilky	dostupné množství	váha jednoho kusu zásilky [stovky kg]	zisk [tisíce Kč]
1	2	8	22
2	2	5	12
3	4	3	7
4	3	2	3

(a) Určete jaké zásilky a v jakém množství by měly ČSA přepravit, aby maximalizovaly svůj zisk.

(b) Těsně před ukončením odbavování prodaly ČSA ještě 2 letenky (zisk z těchto letenek je mnohem vyšší než z přepravy nákladu), čímž se snížila volná kapacita v letadle na 1800 kg. Jak se změní optimální řešení?

Řešení: (a) 2 ks zásilky typu 1 a 1 ks zásilky typu 3;
(b) 2 ks zásilky typu 1 a 1 ks zásilky typu 4.

5.4.31. „Jahodárna“ v Brně – Slatině právě najala 8 nových zaměstnanců. Tito zaměstnanci mohou provádět 4 různé činnosti. Následující tabulka ukazuje zisk [tisíce Kč] za 1 hodinu práce dle jednotlivých činností a počtu přidělených zaměstnanců.

počet zaměstnanců \ činnost	0	1	2	3	4	5	6	7	8
1	22	30	37	44	49	54	58	60	61
2	30	40	48	55	59	62	64	66	67
3	46	52	56	59	62	65	67	68	69
4	5	22	36	48	52	55	58	60	61

(a) Určete optimální rozdělení zaměstnanců k jednotlivým činnostem.

(b) A co kdyby firma najala pouze 6 zaměstnanců?

Řešení: (a) činnost 1 = 3 zaměstnanci, č2=2, č3=0, č4=3
nebo č1=2, č2=3, č3=0, č4=3;
(b) č1=1, č2=2, č3=0, č4=3.

5.4.32. Společnost Dřep, a.s., (=dřevařský podnik) má momentálně na skladě pouze 2000 kusů kulatiny o délce 2 metry. Tyto kulatiny jsou ve zpracovatelském závodě rozřezány na kratší díly a dále prodávány jednotlivým výrobcům. V tuto chvíli má firma objednávky na kusy o délce 30, 70, 110 a 160 cm, přičemž počet objednávek jednotlivých délek zcela překračuje dostupné množství. Prodej různých délek přináší různý zisk dle následující tabulky

délka [cm]	30	70	110	160
zisk [stovky Kč]	1	3	5	8

Určete množství jednotlivých délek kulatiny, které by měly být vyrobeny, aby se maximalizoval zisk.

5.4.33. Investor má k dispozici kapitál 5 ve výši miliónů Kč, které by chtěl investovat do portfolia složeného ze 4 různých aktivit. Zisk z investování do jednotlivých typů aktivit je uveden v následující tabulce, přičemž investor požaduje, aby objem investic do jednotlivých aktivit byl v celočíselných násobcích miliónů Kč.

investice \ aktivita	1	2	3	4
0	0	0	0	0
1	7	5	3	3
2	9	7	4	5
3	9	8	5	8
4	9	8	5	9
5	9	8	5	9

Jak rozdělit jednotlivé investice, aby bylo dosaženo maximálního zisku?

Řešení: aktivita 1 = s investicí 1 mil. Kč, A2=1, A3=1, A4=1 nebo A1=1, A2=2, A3=1, A4=1 nebo A1=2, A2=1, A3=1, A4=1 se ziskem 20 mil. Kč.

5.4.34. Obchodní společnost zaměřená na výrobu jistého výrobku by ráda rozšířila svoji produkci o (a) 100 ks/den nebo (b) 150 ks/den. Toho lze docílit rozšířením současné výrobní kapacity ve 3 výrobních závodech. Následující dvojice tabulek udává míru navýšení zisku [tisíce Kč] z rozšíření současné produkce v jednotlivých závodech, přičemž z technologických důvodů lze v každém závodě rozšířit produkci pouze o násobky 10 ks/den.

(a)

závod \ rozšíření produkce	0	10	20	30	40	50	60	70
A	0	6	10	14	19	22	–	–
B	0	8	12	15	18	20	25	31
C	0	3	8	13	20	–	–	–

(b)

závod \ rozšíření produkce	0	10	20	30	40	50	60	70	80	90
A	0	12	24	35	58	77	93	104	–	–
B	0	5	15	40	80	105	120	125	132	135
C	0	30	48	59	68	73	79	86	–	–

Jak rozšířit stávající výrobu, aby bylo dosaženo maximálního zisku?

Řešení: (b) rozšíření závodu A o 70 ks/den, B=60, C=20 nebo A=60, B=60, C=30 se ziskem 20 tisíc Kč.

5.4.35. Firma vyrábějící komponenty pro automobily plánuje velikost výroby kvartálně, tj. na následující 3 měsíce. Na počátku je stav zásob 10 ks. Potřebná data jsou shrnuta v následující tabulce s cenami v tisících Kč.

měsíc	poptávka	kapacita		náklady	
		výrobní	skladovací	výrobní	skladovací
1	20	30	40	20	3
2	30	20	30	15	3
3	30	30	20	20	2

Určete optimální výrobní program minimalizující náklady.

Řešení: měsíc 1 = výroba 20 ks, M2=20, M3=30 s náklady 133 tisíc Kč.

5.4.36. Obchodní společnost má možnost vybudování 6 nových prodejních míst ve 3 oblastech označených jako A, B, C. V následující tabulce je uveden zisk [tisíce Kč] v závislosti na počtu prodejních míst v jednotlivých oblastech (zisk z nulového počtu nových prodejních míst je způsoben tím, že v jednotlivých oblastech již nějaká prodejní místa jsou).

počet prodejních míst / oblast	0	1	2	3	4	5	6
A	38	41	48	58	66	72	83
B	40	42	50	60	66	75	82
C	60	64	68	78	90	102	109

Jak rozmístit obchody do jednotlivých oblastí, aby se maximalizoval zisk?

5.4.37. Přírodovědecké fakultě MUNI se podařilo velmi výhodně (asi za polovinu cenu) nakoupit 500 ks velmi moderních kalkulaček. Vedení fakulty rozhodlo, že tyto kalkulačky budou nabídnuty studentům matematických, fyzikální a chemických oborů prostřednictvím příslušných sekretariátů. Celá zásilka přišla na fakultu v pěti balících po 100 kusech a pro jednoduchost budou na jednotlivé sekretariáty předány pouze celé balíky. Byl proveden kvalifikovaný odhad zisku [desítky kč] z prodeje v závislosti na počtu kusů dostupných na jednotlivých sekretariátech, který je uveden v následující tabulce (odlišnost cen a jejich nelineární růst jsou důsledkem různé finanční podpory ze strany ústavů).

počet dostupných kusů / obor	100	200	300	400	500
M	60	110	155	170	180
F	50	120	170	200	210
Ch	55	110	150	185	195

Jak by měly být jednotlivé balíky rozděleny ústavům, aby bylo dosaženo maximálního zisku?

5.4.38. Uvažme obchodníka, který prodává zboží pouze jednoho typu. Sklad jeho prodejny má kapacitu 500 ks. Během každého měsíce je schopen prodat libovolné množství, jaké si zvolí, a to až po takové množství, které je ve skladu na začátku měsíce. Na začátku měsíce si může objednat libovolné množství, které mu bude doručeno na konci měsíce, přičemž nesmí přesáhnout volnou kapacitu skladu. Pro následující 4 měsíce má předpověď jednotkových cen [Kč] (změny jsou způsobeny kolísáním kurzu CZK vs. EUR a sezónním výkyvům), která je uveden a v následující tabulce.

měsíc	1	2	3	4
prodejní cena	27	24	26	28
nákupní cena	28	25	25	27

Na počátku má ve skladu 200 ks. Určete nákupní a prodejní strategii, která maximalizuje zisk pro období následujících 4 měsíců.

5.4.39. Nákupní oddělení podniku má zajistit, aby na začátku každého druhého měsíce bylo k dispozici předem dané množství jedné suroviny nutné pro výrobní program dle následující tabulky s údaji pro celý rok.

období (2 měsíce)	1	2	3	4	5	6
spotřeba [ks]	8	5	3	2	7	4
nákupní cena [tisíce Kč]	11	18	13	17	20	10

Počáteční stav zásob jsou 2 ks a na konci má být roven 0 ks. Jaké množství je potřeba nakoupit na začátku každého období, aby se minimalizovaly náklady, je-li kapacita skladu omezena na (a) 9 ks nebo (b) 20 ks, přičemž skladování je zdarma.

Řešení: (a) období 1 = nákup 7 kusů, O2=4, O3=9, O4=3, O5=0, O6=4 s náklady 357 tisíc Kč; (b) O1=18, O2=0, O3=5, O4=0, O5=0, O6=4 s náklady 303 tisíce Kč.

5.4.40. V průběhu nedávné epidemie eboly v Africe bylo rozhodnuto o rozmístění několika lékařských týmů v různých oblastech. V průběhu jednání se dospělo k dvěma možným variantám: buď (a) budou 4 týmy pro 4 možné oblasti nebo (b) 5 týmů pro 3 oblasti s následujícími počty ošetřených pacientů [stovky pacientů za měsíc].

(a)

počet týmů v oblasti \ oblast	1	2	3	4
0	0	0	0	0
1	45	41	25	33
2	78	65	50	48
3	102	80	73	56
4	123	88	90	60

(b)

počet týmů v oblasti \ oblast	A	B	C
0	0	0	0
1	45	20	50
2	70	45	70
3	90	75	80
4	105	110	100
5	120	150	130

Jak rozdělit lékařské týmy, aby bylo dosaženo maximálního počtu ošetřených pacientů?

Řešení: oblast A = 1 tým, B=3, C=1 se 17000 ošetřenými pacienty (první varianta dává maximálně pouze 15200 pacientů při volbě oblast 1 = 2 týmy, O2=1, O3=0, O4=1).

5.4.41. NASA momentálně řeší jistý technologický problém související s lety lidské posádky na Mars. Na řešení pracují 3 týmy využívající různé postupy. Za stávajících okolností byly odhadnuty šance (=pravděpodobnost), že jednotlivé týmy neuspějí: 0,4 pro první tým, 0,6 pro druhý tým a 0,8 pro třetí tým. Vedení NASA se rozhodlo zvýšit šance na úspěch tím, že najme ještě 2 další vědce a začlení je do týmů. Opět byly odhadnuty šance neúspěchu rozšířených týmů, což je shrnuto v následující tabulce.

počet nových členů \ tým	A	B	C
0	0,4	0,6	0,8
1	0,2	0,4	0,5
2	0,15	0,2	0,3

Kam přidělit nové členy, aby byla minimalizována šance na neúspěch?

Řešení: tým A = 1 nový člen, B=0, C=1 s pravděpodobností neúspěchu 0,06.

5.4.42. „Meruňkárna“ v Dobšicích u Znojma prodává svoji sklizeň na různých místech. Tentokrát jí zbylo 5 beden s meruňkami a potřebuje určit vhodná místa k prodeji. Jelikož je potřeba rozvést bedny rychle, byla vytipována 3 nejbližší místa. Zisk [stovky Kč] z prodeje v závislosti na počtu dovezených beden je uveden v následující tabulce (různé ceny jsou způsobeny různými charaktery jednotlivých prodejních míst).

počet dodaných beden \ oblast	A	B	C
0	0	0	0
1	5	6	4
2	9	11	9
3	14	15	13
4	17	19	18
5	21	22	20

Kolik beden bylo mělo být rozvezeno na jednotlivá místa, aby se maximalizoval zisk?

Řešení: místo A = 1 bedna, B=2, C=2 se ziskem 2500 Kč.

5.4.43. V nadcházejícím zkouškovém období Vás čeká zkouška z toho kurzu. Předpokládejme, že jste si na zkoušku vyhradili 7 dní a předmět rozdělili do 4 bloků. Každému bloku se chcete věnovat alespoň 1 den a nejvýše 4 dny. „Užitek“ z přípravy věnované jednotlivým blokům můžeme v závislosti na počtu dnů odhadnout způsobem, který je uveden v následující tabulce.

počet dnů \ blok	LP	QP	DP	VP
1	1	5	4	4
2	3	6	6	4
3	6	8	7	5
4	8	8	9	8

Určete optimální strategii přípravy vedoucí k maximálnímu užitku.

Řešení: oblast LP = 3 dny, QP=1, DP=2, VP=1 nebo
LP = 4 dny, QP=1, DP=1, VP=1 vždy s hodnotou užitku 21.

5.4.44. V závěru jistých voleb se jeden z jejich účastníků rozhodl zvýšit své šance na vítězství reklamou ve 4 celostátních denících. Rozpočet na tuto reklamní kampaň je dostatečný na zveřejnění reklamy ve všech 5 dnech, které zbývají do voleb, přičemž v každém dni bude reklama pouze v jednom deníku. Zisk nových voličů [ve stovkách] z této kampaně v závislosti na deníku a počtu reklamních dní je v následující tabulce.

počet reklamních dnů \ deník	A	B	C	D
0	0	0	0	0
1	4	6	5	3
2	7	8	9	7
3	9	10	11	12
4	12	11	10	14
5	15	12	9	16

Jakým způsobem by měla být reklama rozdělena, aby bylo dosaženo maximálního zisku nových voličů?

Řešení: deník A = 0 dnů, B=1, C=1, D=3 se ziskem 2300 nových voličů.

5.4.45. Vzhledem k rostoucí konkurenci mezi vysokými školami/univerzitami při přijímání nových studentů bakalářských oborů se rozhodlo vedení MUNI investovat 800 tisíc Kč do reklamní kampaně zahrnující noviny, časopisy, rozhlas a televizi. V závislosti na množství investovaných peněz byl učiněn odhad „zasažených" potenciálních uchazečů o studium [ve stovkách], přičemž pro jednoduchost bylo rozhodnuto, že na jednotlivé typy reklamy bude uvolněna částka v celočíselných násobcích stovek tisíc Kč. Tyto údaje jsou shrnuty v následující tabulce.

medium \ investice	1	2	3	4	5	6	7	8
noviny	24	37	46	59	72	80	82	82
časopisy	15	55	70	75	90	95	95	95
rozhlas	20	30	45	55	60	62	63	63
televize	20	40	55	65	70	70	70	70

(a) Určete optimální rozdělení rozpočtu maximalizující počet „zasažených" uchazečů.
(b) Jak se toto optimální rozdělení změní, bude-li dostupný rozpočet zkrácen o 200 tisíc?
(c) Jak se smění předchozí dvě řešení, nebude-li povolena reklama v televizi?

Řešení: (a) noviny = 100 tisíc Kč, Č=3, R=1, TV=3 se 16900 „zasaženými" uchazeči;
(b) N=1, Č=2, R=1, TV=2 se 13900 „zasaženými" uchazeči;
(c) v prvním případě N=2, Č=3, R=3 se 15200 „zasaženými" uchazeči, ve druhém případě N=2, Č=3, R=1 se 12700 „zasaženými" uchazeči.

5.4.46. Uvažme elektronický systém, který se skládá ze 4 komponent zapojených sériově, tj. pro funkčnost celého systému jsou potřeba všechny komponenty. Spolehlivost systému může být zvýšena paralelním zapojením až tří jednotek jedné či více komponent. Pravděpodobnost funkčnosti celého systému a její zvýšení v závislosti na počtu paralelních komponent je popsána v následující tabulce.

počet paralelních jednotek \ komponenta	A	B	C	D
1	0,5	0,6	0,7	0,5
2	0,6	0,7	0,8	0,7
3	0,8	0,8	0,9	0,9

Současně však zapojení těchto paralelních jednotek přináší náklady [tisíce Kč] dle následující tabulky.

počet paralelních jednotek \ komponenta	A	B	C	D
1	1	2	1	2
2	2	4	3	3
3	3	5	4	4

Určete počty paralelně zapojených jednotek jednotlivých komponent maximalizující spolehlivost systému, je-li rozpočet limitován částkou 10 tisíc Kč.

Řešení: komponenta A = 3 paralelní jednotky, B=1, C=1, D=3 s celkovou pravděpodobností funkčnosti systému 0,3024.

5.4.47. Firma na základě stávajících objednávek má spočítáno, že na následující 3 období bude potřebovat počty zaměstnanců uvedené v následující tabulce.

období	1	2	3
počet zaměstnanců	45	40	50

Ovšem propouštění a najímání zaměstnanců má obvykle své náklady stejně jako v případě přezaměstnanosti (podzaměstnanost není povolena). Uvažme, že každý „nadbytečný“ zaměstnanec znamená ztrátu 30 tisíc Kč za příslušné období a změna počtu zaměstnanců znamená ztrátu $2000\times(\text{míra změny})^2$. Jelikož nejsou povoleny částečné úvazky, je možné mít pouze celočíselný počet zaměstnanců. Určete optimální počty zaměstnanců v jednotlivých období, aby bylo dosaženo minimální ztráty v průběhu roku, má-li firma na počátku 50 zaměstnanců.

Řešení: období 1 = 45 zaměstnanců, O2=44, O3=50 s náklady 2440 tisíc Kč.

5.4.48. Míra (ne-)zaměstnanosti v zemědělství nebo stavebnictví je závislá na ročním období. Uvažme zaměstnavatele, který v průběhu roku potřebuje počty zaměstnanců uvedené v následující tabulce.

období	jaro	léto	podzim	zima	jaro
počet zaměstnanců	255	220	240	200	255

Ovšem propouštění a najímání zaměstnanců má obvykle své náklady stejně jako v případě přezaměstnanosti (podzaměstnanost není povolena). Uvažme, že každý „nadbytečný“ zaměstnanec znamená ztrátu 20 tisíc Kč za příslušné období a změna počtu zaměstnanců znamená ztrátu $2000\times(\text{míra změny})^2$. Jelikož jsou povoleny také částečné úvazky, je možné mít neceločíselný počet zaměstnanců. Určete optimální počty zaměstnanců v jednotlivých období, aby bylo dosaženo minimální ztráty v průběhu roku.

Řešení: jaro = 255 zaměstnanců, léto=247,5, podzim=245, zima=247,5, jaro=255 s náklady 1850 tisíc Kč.

5.4.49. Uvažte předchozí příklad s tím, že změna zaměstnanosti ale znamená pouze poloviční ztrátu.

Řešení: jaro = 255 zaměstnanců, léto=242,5, podzim=245, zima=242,5, jaro=255.

5.4.50. Uvažme dva hráče jisté přátelské hry (vrhcáby, dáma, šachy), ve které není povolena remíza. Pro zvýšení napětí je možné v každé hře provést libovolnou sázku v rozmezí od 0 po aktuálně dostupné žetony (tj. „all-in“). Pravděpodobnost výhry je $\frac{1}{2}$. Při počátečním obnosu 75 žetonů je cílem mít po třech hrách alespoň 100 žetonů (žeton=Kč, přátelská hra=není potřeba získat co nejvíce). Určete optimální strategii sázek.

Řešení: sázka v 1. hře = 0 žetonů, S2=25, v případě výhry S3=0, v případě prohry S3=50 nebo S1=25, v případě výhry S2=0=S3, v případě prohry S2=50, v případě prohry S2=50 to vše s pravděpodobností úspěchu $\frac{3}{4}$.

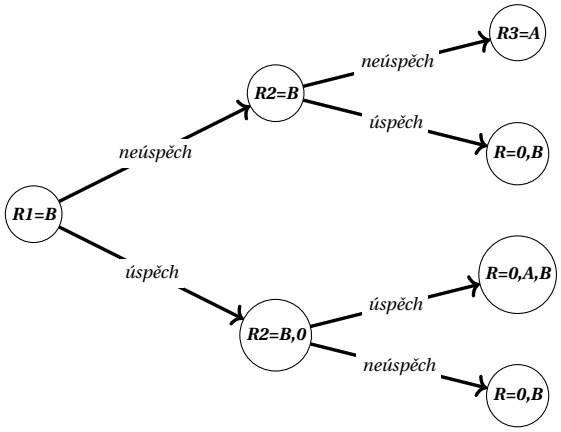
5.4.51. V rámci filosofie „špetka praxe vydá za tunu teorie“ jste získali 10 tisíc Kč s úkolem stanovení optimální investiční strategie na následující 3 roky. Na začátku každého roku je potřeba rozhodnout, zda budeme investovat do produktu A nebo B nebo zda nebudeme investovat nikam. V každém roce je možné provést pouze jednu investici ve výši 10 tisíc Kč (nebo neinvestovat). Výnos [tisíce Kč] z jednotlivých investic je uvedena v následující tabulce.

investice	obnos na konci roku	pravděpodobnost
A	0	0,25
	20	0,25
B	10	0,9
	20	0,1

(a) Jak investovat, aby bylo dosaženo maximálního výnosu po 3 letech?

(b) Jak investovat, aby byla maximální pravděpodobnost, že po 3 letech budete mít nejméně 20 tisíc Kč?

Řešení: (a) je-li to možné, vždy investovat do A s očekávaným výnosem 22,5 tisíc Kč; (b) maximální pravděpodobnost je 0,7975 při následující strategii



5.4.52. Určete optimální strategii pro sázení v Las Vegas, máte-li na počátku 3 žetony a chcete-li mít po 2 hrách alespoň 4 žetony (pravděpodobnost výhry je 0,66).

Řešení: Optimální strategie má pravděpodobnost 0,88.

5.4.53. Určete optimální strategii pro sázení v Las Vegas, máte-li na počátku 2 žetony a chcete-li mít po 5 hrách alespoň 5 žetonů (pravděpodobnost výhry je 0,5).

Řešení: Optimální strategie má pravděpodobnost $\frac{160}{243}$.

5.4.54. Společnost Pokus & Omyl, s.r.o., vyrábí velmi sofistikovaný produkt jednoho druhu. Společnost sice obdržela zakázku na dodání 1 ks tohoto výrobku, ovšem její zadavatel má velmi striktní požadavky na výslednou kvalitu, tj. při jejím nesplnění výrobek neodebere. To vede k tomu, že může být nutné vyrobit více kusů tohoto výrobku, než bude dosaženo požadované kvality. Počet nadbytečných výrobků nevyhovující kvality můžeme označit jako „povolenou zmetkovost“, což bývá běžnou praxí ve výrobním procesu.

Vedení společnosti odhaduje, že je 50% pravděpodobnost vyrobení vyhovujícího kusu a 50% pravděpodobnost vyrobení nevyhovujícího kusu. Proto počet vyhovujících kusů mezi L vyrobenými kusy má binomické rozdělení pravděpodobnosti, tj. pravděpodobnost, že nebude vyroben žádný vyhovující kus je $\left(\frac{1}{2}\right)^L$.

Náklady na výrobu 1 ks (ať již vyhovujícího či nevyhovujícího) jsou 1000 Kč a nevyhovující výrobky jsou bezcenné. Navíc samotné spuštění výrobního procesu stojí 3000 Kč a tyto náklady stojí každé další spuštění výrobního procesu, pokud se ukáže, že mezi vyrobenými kusy v předchozím běhu není žádný vyhovující výrobek. Společnost se rozhodla spustit nejvýše 3 po sobě jdoucí výrobní procesy pro výrobu požadovaného produktu. Pokud ani na konci třetího běhu nebude získán vyhovující kus bude celková ztráta ve výši 16 tisíc Kč včetně pokuty zaplacené zákazníkovi. Vaším úkolem je stanovení optimální velikosti výrobního procesu, tj. kolik má být v každém běhu výrobního procesu vyrobeno kusů (1+ povolená zmetkovost), aby byly minimalizovány náklady.

Řešení: v 1. běhu = 2 ks, B2=2 nebo 3 kusy, B3=3 nebo 4 kusy s očekávanými náklady 6750 Kč.

5.4.55. Pomocí dynamického programování vyřešte úlohu

$$5x_1 + 2x_2^2 + x_3^3/3 + x^4/4 \to \max \quad \& \quad \sum_{i=1}^{4} x_i \le 6 \quad \& \quad x_i \in \{1,2\}.$$

5.4.56. Pomocí dynamického programování vyřešte úlohu

$$18x_1 - x_1^2 + 20x_2 + 10x_3 \to \max,$$
$$2x_1 + 4x_2 + 3x_3 \le 11 \quad \& \quad x_1, x_2, x_3 \in \mathbb{Z} \quad \& \quad x_1, x_2, x_3 \ge 0.$$

5.4.57. Pomocí dynamického programování vyřešte úlohu

$$32x_1 - 2x_2^2 + 30x_2 + 20x_3 \to \max,$$
$$3x_1 + 7x_2 + 5x_3 \le 20 \quad \& \quad x_1, x_2, x_3 \in \mathbb{Z} \quad \& \quad x_1, x_2, x_3 \ge 0.$$

Řešení: $x_1 = 5$, $x_2 = 0$, $x_3 = 1$.

5.4.58. Pomocí dynamického programování vyřešte úlohu

$$36x_1 + 9x_1^2 - 6x_1^3 + 36x_2 - 3x_2^3 \to \max \quad \& \quad x_1 + x_2 \le 3 \quad \& \quad x_1, x_2 \ge 0.$$

Řešení: $x_1 = -2 + \sqrt{13}$, $x_2 = 5 - \sqrt{13}$.

5.4.59. Pomocí dynamického programování vyřešte úlohu

$$2x_1^2 + 2x_2 + 4x_3 - x_3^2 \to \max \quad \& \quad 2x_1 + x_2 + x_3 \le 4 \quad \& \quad x_1, x_2, x_3 \ge 0.$$

Řešení: $x_1 = 0$, $x_2 = 3$, $x_3 = 1$.

5.4.60. Pomocí dynamického programování vyřešte úlohu

$$x_1^2 + x_2^2 + x_3^2 \to \min \quad \& \quad x_1 + x_2 + x_3 \ge 15 \quad \& \quad x_1, x_2, x_3 \ge 0.$$

5.4.61. Pomocí dynamického programování vyřešte úlohu

$$x_1^2 + 2x_2^2 + 4x_3 \to \max \quad \& \quad x_1 + 2x_2 + x_3 \ge 8 \quad \& \quad x_1, x_2, x_3 \ge 0.$$

5.4.62. Pomocí dynamického programování vyřešte úlohu

$$x_1^2 + x_2^2 + x_3^2 \to \max \quad \& \quad x_1 + x_2 + x_3 \le \alpha \quad \& \quad x_1, x_2, x_3 \ge 0.$$

je-li $\alpha \ge 0$.

5.4.63. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} x_k^2 \to \max \quad \& \quad x_1 + \cdots + x_n \le \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$ a $\alpha > 0$.

5.4.64. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} x_k^p \to \max \quad \& \quad x_1 + \cdots + x_n \le \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a $p \ge 1$.

5.4.65. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} x_k^p \to \min \quad \& \quad x_1 + \cdots + x_n \ge \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a $p \ge 1$.

5.4.66. Pomocí dynamického programování vyřešte úlohu

$$\min_{x_1,\ldots,x_n} \{\max[r(x_n), \ldots, r(x_1)]\} \quad \& \quad x_1 + \cdots + x_n \ge \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a r je konvexní, rostoucí a diferencovatelná funkce.

Řešení: $x_1 = \cdots = x_n = \frac{\alpha}{n}$.

5.4.67. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} r(x_k) \to \min \quad \& \quad x_1 + \cdots + x_n \ge \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a r je konvexní, rostoucí a diferencovatelná funkce.

Řešení: $x_1 = \cdots = x_n = \frac{\alpha}{n}$.

5.4.68. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} r(x_k) \to \max \quad \& \quad x_1 \cdot x_2 \cdots x_n \le \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a r je konvexní, rostoucí a diferencovatelná funkce.

5.4.69. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} r(x_k) \to \min \quad \& \quad x_1 \cdot x_2 \cdots x_n \ge \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha > 0$ a r je konvexní, rostoucí a diferencovatelná funkce.

Řešení: $x_1 = \cdots = x_n = \alpha^{1/n}$.

5.4.70. Pomocí dynamického programování vyřešte úlohu

$$\prod_{k=1}^{n} x_k = x_1 \cdot x_2 \cdots x_n \to \max \quad \& \quad x_1 + \cdots + x_n \le \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$ a $\alpha > 0$.

Řešení: $x_1 = \cdots = x_n = \frac{\alpha}{n}$.

5.4.71. Pomocí dynamického programování vyřešte úlohu

$$\max_{0 \le D_n \le X_n} \sum_{n=1}^{N} A D_n + B(X_n - D_n) \quad \& \quad X_{n-1} = a D_n + b(X_n - D_n) \quad n = 1, \ldots, N,$$

pro daná $A, B > 0$ a $a, b \in [0, 1)$. Určete také stacionární řešení, tj.

$$\lim_{N \to \infty} f_N(X_N) \quad \text{a} \quad \lim_{N \to \infty} D_N(X_N)$$

jakožto funkci proměnných A, a, B, b.

5.4.72. Pomocí dynamického programování vyřešte úlohu

$$\prod_{k=1}^{n} x_k = x_1 \cdot x_2 \cdots x_n \to \max \quad \& \quad x_1 + \cdots + x_n = \alpha \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$ a $\alpha > 0$.

5.4.73. Pomocí dynamického programování vyřešte úlohu

$$15x_1 + 10x_2 \to \max \quad \& \quad x_1 + 2x_2 \le 6 \quad \& \quad 3x_1 + x_2 \le 8 \quad \& \quad x_1, x_2 \ge 0.$$

5.4.74. Pomocí dynamického programování vyřešte úlohu

$$3x_1 + 4x_2 \to \max \quad \& \quad x_1 + 6x_2 \le 6 \quad \& \quad 2x_1 + x_2 \le 4 \quad \& \quad x_1, x_2 \ge 0.$$

5.4.75. Pomocí dynamického programování vyřešte úlohu

$$4x_1 + 2x_2 \to \max \quad \& \quad 3x_1 + x_2 \le 42 \quad \& \quad x_1 + x_2 \le 21 \quad \& \quad 3x_1 + 2x_2 \le 48 \quad \& \quad x_1, x_2 \ge 0.$$

5.4.76. Pomocí dynamického programování vyřešte úlohu

$$\sum_{k=1}^{n} r(x_k) \to \min \quad \& \quad x_1 + \cdots + x_n \ge \alpha \quad \& \quad x_1 \cdot x_2 \cdots x_n \ge \beta \quad \& \quad x_1, \ldots, x_n \ge 0,$$

kde $n \in \mathbb{N}$, $\alpha, \beta > 0$ a r je konvexní, rostoucí a diferencovatelná funkce.

Řešení: $x_1 = \cdots = x_n = \frac{\alpha}{n}$ *v případě* $\alpha \ge n\beta^{1/n}$ *a* $x_1 = \cdots = x_n = \beta^{1/n}$ *v případě* $\alpha \le n\beta^{1/n}$.

Kapitola 6

Základy konvexní analýzy

6.1 *Konvexní množiny* 430
6.2 *Konvexní funkce* 452
6.3 *Oddělování konvexních množin a jeho důsledky* 477
6.4 *Vlastnosti konvexních funkcí* 492
6.5 *Subgradient a subdiferenciál* 504
6.6 *Fenchelova transformace* 512

CONVEXITY IS VERY EASY TO DEFINE, TO VISUALIZE AND TO GET AN INTUITION ABOUT... THE MAIN BUILDING BLOCK OF CONVEXITY THEORY IS A STRAIGHT LINE INTERVAL... CONVEXITY IS MORE INTUITIVE THAN, SAY, LINEAR ALGEBRA. IN LINEAR ALGEBRA, THE INTERVAL IS REPLACED BY THE WHOLE STRAIGHT LINE. WE HAVE SOME DIFFICULTY VISUALIZING A STRAIGHT LINE BECAUSE IT RUNS UNCHECKED IN BOTH DIRECTIONS... CONVEXITY HAS AN IMMENSELY RICH STRUCTURE AND NUMEROUS APPLICATIONS. ON THE OTHER HAND, ALMOST EVERY „CONVEX“ IDEA CAN BE EXPLAINED BY A TWO-DIMENSIONAL PICTURE. THERE MUST BE SOME REASON FOR THAT... SIMPLICITY, INTUITIVE APPEAL AND UNIVERSALITY OF APPLICATIONS MAKE TEACHING CONVEXITY (AND WRITING A BOOK ON CONVEXITY) A RATHER GRATIFYING EXPERIENCE.[1]

ALEXANDER BARVINOK[2] v předmluvě ke své knize [30]

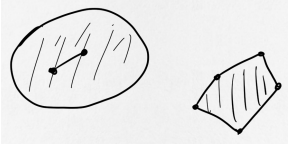
[1] *Konvexnost lze velmi snadno definovat, znázornit a získat o ní jakési povědomí… jejím hlavním stavební kamenem jsou úsečky… Konvexnost je intuitivnější než např. lineární algebra. V lineární algebře je úsečka nahrazena celou přímkou. Máme však jisté potíže si celou úsečku představit, neboť se táhne neomezeně v obou směrech… koncept konvexnosti má nesmírně bohatou strukturu a široké využití. Na druhou stranu ale téměř všechny „konvexní“ pojmy lze vysvětlit pomocí dvoudimenzionálního obrázku. K tomu musí být nějaký důvod… Jednoduchost, intuitivní přitažlivost a všestranné využití dělá z výuky konvexnosti (a psaní knihy o konvexnosti) poněkud potěšujícím zážitkem.*

[2] Alexander I. Barvinok (∗ 27. března 1963) je profesorem matematiky na Michiganské univerzitě. Doktorát získal v roce 1988 na Petrohradské státní univerzitě. V roce 1999 převzal z rukou amerického prezidenta Billa Clintona Cenu prezidenta pro začínající vědce a inženýry a od roku 2012 je spolupracovníkem Americké matematické společnosti.

Cílem této kapitoly je vybudování základů *konvexní analýzy*, tj. oblasti, v níž se vše točí především kolem pojmu konvexnosti a být „nekonvexní“ zde vždy znamená pouze potíže. My jsme na tuto vlastnost již několikrát narazili v předchozích kapitolách a čtenář se s tímto pojmem jistě setkal v souvislosti s funkcemi jedné proměnné v základním kurzu matematické analýzy. To však byla jen nepatrná špička konvexního ledovce, ze kterého se nyní pokusíme poodhalit mnohem výraznější část. Rozšíříme si tento koncept na funkce s libovolným počtem proměnných a uvidíme, že tato vlastnost má pro řešení optimalizačních úloh dalekosáhlé důsledky. Některé z nich jsme již vlastně nepozorovaně (a velmi specifickým způsobem) využili v teoretických základech lineárního a kvadratického programování v předchozích kapitolách. V souladu s úvodním komentářem k této kapitole nabídneme čtenáři/čtenářce řadu ilustrativních obrázků napříč naším výkladem, které by mu/jí mělynapomoci s pochopením mnohých „konvexních vlastností“[3].

[3] Při takovýchto komentářích se mi vždy vybaví výrok ruského premiéra Viktora Černomyrdina z let 1992–1998, který pronesl 6. srpna 1993 během tiskové konference: *Choteli kak luchshe, a poluchilos kak vsegda.* (*Měli jsme ty nejlepší úmysly, ale dopadlo to jako vždy.* [upraveno]).

6.1 Konvexní množiny



Obrázek 6.1: Konvexní množiny.

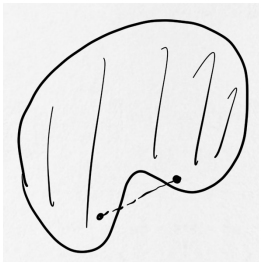
Jenže konvexnost není spojena pouze s funkcemi. Abychom se mohli k funkcím vůbec dostat a celé to správně fungovalo, musíme se nejdříve zaměřit na „úsečkové“ (neboli *konvexní*) *množiny*. S těmi jsme se již setkali v úlohách lineárního a kvadratického programování, kde vystupovaly jako přípustné množiny vymezené systémem nerovnic (ať již s omezením na znaménko či bez něj) – vždyť takové úlohy jsme pak v dvoudimenzionálním případě dokázali snadno řešit pouze pomocí vhodného obrázku.

DEFINICE 6.1.1

Nechť $X \subseteq \mathbb{R}^n$. Množina X se nazývá *konvexní*, jestliže pro všechna $x_1, x_2 \in X$ a libovolné $\lambda \in [0,1]$ platí

$$\lambda x_1 + (1-\lambda) x_2 \in X. \tag{6.1.1}$$

Jaký je geometrický význam podmínky (6.1.1)? Aby množina X byla konvexní, musí s libovolnými dvěma svými body obsahovat i celou úsečku spojující tyto body! Také bychom zde měli zdůraznit, že i množina $X = \emptyset$ je z formálního pohledu konvexní. Tato informace sice nemá příliš praktické využití, ovšem nám to zjednoduší další výklad.



Obrázek 6.2: Nekonvexní množina.

Obrázek 6.3: Nekonvexní množina.

Příklad 6.1.2

Mezi nejjednodušším příklady konvexních množin patří celé $\mathbb{R}^n$, jednobodové množiny $\{a\}$ nebo libovolná úsečka. Zajímavějším příkladem je řešení libovolného (i nekonečného) systému lineárních nerovnic

$$a_i^\top x \le b_i, \quad i \in I,$$

s n neznámými, kde $I \subseteq \mathbb{N}$ je indexová množina a $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$ pro $i \in I$. Jinými slovy se jedná o množinu

$$M := \{x \in \mathbb{R}^n \mid a_i^\top x \le b_i,\ i \in I\}.$$

Vskutku, je-li $x, y \in M$, pak pro libovolné $i \in I$ a $\lambda \in [0,1]$ máme i $1-\lambda \ge 0$, takže

$$a_i^\top [\lambda x + (1-\lambda) y] = \lambda a_i^\top x + (1-\lambda) a_i^\top y \le \lambda b_i + (1-\lambda) b_i = b_i.$$

Tudíž $\lambda x + (1-\lambda) y \in M$, což dokazuje konvexnost množiny M.
V případě konečné množiny I se jedná o řešení soustavy m nerovnic o n neznámých, tj. systému

$$Ax \le b, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m.$$

V takovém případě se odpovídající množina M nazývá *polyedr* (neboli *mnohostěn*), viz např. Obrázky 6.4–6.5. Zejména libovolná rovina jakožto afinní podprostor je řešením

nějakého systému lineárních rovnic $Ax = b$, který je ekvivalentní systému lineárních nerovnic $Ax \le b$ a $Ax \ge b$. Tudíž každá rovina je konvexní (což bychom samozřejmě mohli dokázat i přímo).

Obrázek 6.4: Polyedr.

Jaký vliv mají množinové operace na konvexnost? Např. sjednocení dvou konvexních množin rozhodně nemusí být konvexní, jak lze snadno vidět na Obrázku 6.6. Další operace jsou obsahem následujícího tvrzení. Připomeňme, že *součinem* množiny $X \subseteq \mathbb{R}^n$ a čísla α rozumíme množinu tvořenou α-násobky všech prvků z X, tj.

$$\alpha X := \left\{ y \in \mathbb{R}^n \mid y = \alpha x \text{ pro nějaké } y \in X \right\}. \tag{6.1.2}$$

Podobně *součtem*[4] dvou množin $X_1, X_2 \subseteq \mathbb{R}^n$ rozumíme množinu, v níž každý prvek lze vyjádřit jako součet prvků z X_1 a X_2, tj.

$$X_1 + X_2 := \left\{ y \in \mathbb{R}^n \mid y = x_1 + x_2 \text{ pro nějaká } x_1 \in X_1 \text{ a } x_2 \in X_2 \right\}. \tag{6.1.3}$$

Definici součtu lze indukcí rozšířit na libovolný (konečný) počet množin. Můžeme také tyto dvě operace zkombinovat, čímž získáme lineární kombinaci daných množin.

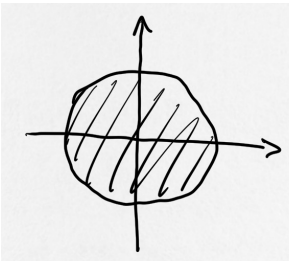
Věta 6.1.3

(i) Nechť I je libovolná indexová množina a množiny $X_i \subseteq \mathbb{R}^n$ jsou konvexní pro všechna $i \in I$. Pak množina $\bigcap_{i\in I} X_i$ je také konvexní.

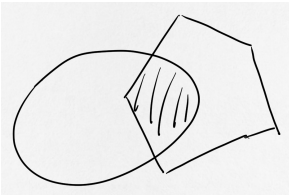
(ii) Nechť $X_1, \dots, X_m \subseteq \mathbb{R}^n$ jsou konvexní množiny a $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ libovolná. Pak množina

$$\alpha_1 X_1 + \cdots + \alpha_m X_m := \left\{ x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \alpha_i x_i \text{ pro nějaká } x_i \in X_i \right\}$$

je opět konvexní.



Obrázek 6.5: Konvexní množina, která není polyedr (kdyby byl na obrázku skutečně kruh).



Obrázek 6.6: Průnik konvexních množin.

[4] Tato operace se také někdy nazývá *Minkowského součtem*.

Důkaz.

(i) Nechť I je konečná nebo nekonečná indexová množina a nechť X_i jsou konvexní množiny pro všechna $i \in I$. Je-li množina $X := \bigcap_{i\in I} X_i = \emptyset$ nebo jednoprvková, pak je tvrzení triviální. Uvažme tedy alespoň dvouprvkovou množinu X. Nechť $x_1, x_2 \in X$ jsou libovolné. Pak nutně $x_1, x_2 \in X_i$ pro všechna $i \in I$. Jenže všechny množiny X_i jsou konvexní, tudíž

$$x := \lambda x_1 + (1-\lambda) x_2 \in X_i \quad \text{pro všechna } \lambda \in [0,1] \text{ a všechna } i \in I.$$

To ale znamená, že také $x \in X$, tj. $\bigcap_{i\in I} X_i$ je konvexní.

(ii) Nechť X je konvexní množina. Nejdříve ukážeme, že také množina αX je konvexní pro libovolné $\alpha \in \mathbb{R}$. Je-li $X = \emptyset$ nebo jednoprvková, je tvrzení triviální. Uvažme tedy alespoň dvouprvkovou množinu X. Jestliže $z_1, z_2 \in \alpha X$ jsou libovolné, pak $z_1 = \alpha x_1$ a $z_2 = \alpha x_2$ pro nějaké $x_1, x_2 \in X$. Proto pro každé $\lambda \in [0,1]$ máme

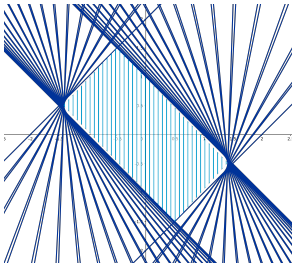
$$\lambda z_1 + (1-\lambda) z_2 = \alpha \underbrace{[\lambda x_1 + (1-\lambda) x_2]}_{\in X} \in \alpha X,$$

tj. množina αX je konvexní.

Nechť nyní jsou X, Y libovolné konvexní množiny. Ukážeme, že také množina $X + Y$ je konvexní. Jestliže X, Y jsou nejvýše jednoprvkové, je tvrzení triviální. Uvažme tedy, že alespoň jedna z těchto množin je nejméně dvouprvková. Jestliže $z_1, z_2 \in X + Y$, pak $z_1 = x_1 + y_1$ a $z_2 = x_2 + y_2$ pro nějaká $x_1, x_2 \in X$ a $y_1, y_2 \in Y$. Proto pro každé $\lambda \in [0,1]$ máme

$$\begin{aligned} \lambda z_1 + (1-\lambda) z_2 &= \lambda (x_1 + y_1) + (1-\lambda)(x_2 + y_2) \\ &= \underbrace{[\lambda x_1 + (1-\lambda) x_2]}_{\in X} + \underbrace{[\lambda y_1 + (1-\lambda) y_2]}_{\in Y} \in X + Y, \end{aligned}$$

tj. $X + Y$ je konvexní. Indukcí pak snadno rozšíříme tyto úvahy na libovolnou lineární kombinaci konvexních množin, čímž obdržíme druhé tvrzení věty. ∎



Obrázek 6.7: Volba $p = \pi/3$.

Příklad 6.1.4

Uvažme např. množinu

$$S_t := \{x \in \mathbb{R}^n \mid -1 \le (\cos t,\ \cos 2t,\ \ldots, \cos nt)\, x \le 1\}.$$

Pro libovolné $t \in \mathbb{R}$ se jedná (s jednou jedinou výjimkou) o pásy vymezené dvojicí rovnoběžných nadrovin s normálovým vektorem $(\cos t,\ \cos 2t,\ \ldots,\ \cos nt)^\top$, tj. pro $n = 1$ a $t \neq \pi/2 + k\pi$, $k \in \mathbb{Z}$, máme interval $\left[-1/|\cos t|, 1/|\cos t|\right]$, zatímco pro $t = \pi/2 + k\pi$ dostaneme celou reálnou osu; pro $n = 2$ vždy dostaneme pás vymezený dvojicí rovnoběžných přímek s normálovým vektorem $(\cos t,\ \cos 2t)^\top$ atd. Pak v souladu s Větou 6.1.3 je průnik množin S_t pro libovolný výběr $t \in \mathbb{R}$ konvexní množinou, viz Obrázky 6.7–6.9 s množinami $\bigcap_{|t| \le p} S_t$ pro tři různé hodnoty parametru $p \in \{\pi/3, \pi/2, \pi\}$ (tyto množiny jsou vyšrafovány světlemodrou barvou).

Náš výklad ovšem nebude pouze o konvexních množinách, a tak si v následující definici zavedeme některé další užitečné typy množin (které vždy zahrnují i $X = \emptyset$).

DEFINICE 6.1.5

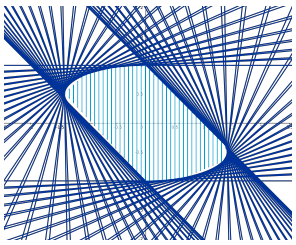
Množina $X \subseteq \mathbb{R}^n$ se nazývá

(i) *kužel*, jestliže pro každé $x \in X$ a pro každé $\lambda \in [0, \infty)$ je také $\lambda x \in X$;

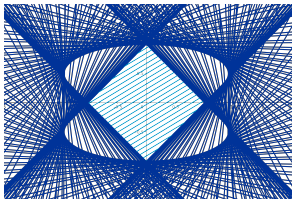
(ii) *konvexní kužel*, jestliže je množina X konvexní a současně kuželem;

(iii) *afinní* (též *lineární varieta*), jestliže pro každé $x_1, x_2 \in X$ a pro každé $\lambda \in \mathbb{R}$ platí

$$\lambda x_1 + (1 - \lambda)\, x_2 \in X.$$

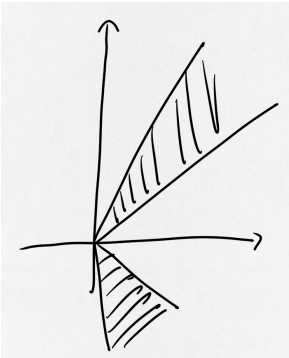


Obrázek 6.8: Volba $p = \pi/2$.



Obrázek 6.9: Volba $p = \pi$.

[5] V literatuře se lze setkat i s definicí kužele vyžadující $\lambda x \in X$ pouze pro $\lambda \in (0, \infty)$, což znamená, že takový kužel nemusí obsahovat počátek.



Obrázek 6.10: Kužel, který není konvexní. Současně je to sjednocení dvou konvexních kuželů, které není konvexní množinou.

Poznámka 6.1.6. Podobně jako v případě konvexních množin lze i požadavky z předchozí definice snadno popsat geometricky.

(i) Množina X je totiž kužel právě tehdy, když obsahuje počátek[5] a s každým dalším bodem $x \in X$ obsahuje i celou polopřímku vycházející z počátku a procházející x. Součet i sjednocení dvou (a více) kuželů je opět kužel, ale pochopitelně ne každý kužel je konvexní, viz Obrázky 6.10–6.11.

(ii) Konvexní kužel musí s libovolnými dvěma body obsahovat i celou „výseč“ určenou polopřímkami vycházejícími z počátku a procházejícími těmito body, takže to v $\mathbb{R}^3$ mnohdy může připomínat spíše neohraničený jehlan. Formální definice konvexního kužele je ekvivalentní s požadavkem, že pro každé $x, y \in X$ a $\lambda \ge 0$ máme i $\lambda x \in X$ a $x + y \in X$. Těmto podmínkám jistě vyhovuje každý vektorový prostor (nebo jeho „polovina“), kladný orthant $\{x \in \mathbb{R}^n \mid x \ge 0\}$, polopřímka $\{x \in \mathbb{R}^n \mid x = \lambda b,\ \lambda \ge 0\}$ pro daný vektor $b \in \mathbb{R}^n$ nebo uzavřený „poloprostor“ $\{x \in \mathbb{R}^n \mid a^\top x \le 0\}$ či otevřený „poloprostor“ $\{x \in \mathbb{R}^n \mid a^\top x < 0\} \cup \{0\}$.

(iii) Afinní množina s libovolnými dvěma body obsahuje také celou přímku určenou těmito body, takže bychom ji mohli označit jako „přímkovou množinu“. Proto je samozřejmě každá taková množina nutně konvexní. S afinními množinami jsme se nejspíše setkali již v základním kurzu lineární algebry. V něm jsou afinní množiny obvykle popsány jako prostor řešení nehomogenní soustavy lineárních rovnic $\mathcal{R}(A, b) := \{x \in \mathbb{R}^n \mid Ax = b\}$ pro nějakou matici $A \in \mathbb{R}^{m \times n}$ a vektor $b \in \mathbb{R}^m$ neboli jako „posunutý“ vektorový podprostor, tj. jako množiny tvaru $x_0 + \mathcal{U} = x_0 + \mathcal{R}(A, 0)$ pro nějaký bod $x_0 \in \mathbb{R}^n$ a vektorový podprostor[6] $\mathcal{U} \subseteq \mathbb{R}^n$. Toto vyjádření je ekvivalentní s naší definicí výše, přičemž vektorový podprostor $\mathcal{U}$ (tzv. *zaměření*) je určen jednoznačně a s jeho pomocí klademe[7] $\dim X := \dim \mathcal{U}$. Odtud mimo jiné plyne, že buď $X = \mathbb{R}^n$ (afinní množina dimenze n) nebo $\operatorname{int} X = \emptyset$, tj. afinní množina dimenze menší než n nemá vnitřní body. ▲

V definicích konvexní a afinní množiny či kužele jsme využili lineární kombinace s dodatečnými podmínkami na koeficienty, což nás přivádí k následující terminologii.

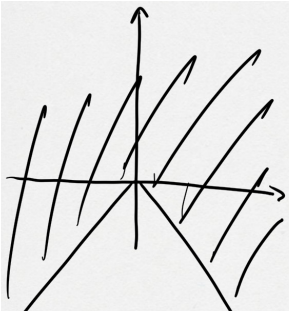
Definice 6.1.7

Nechť $x_1, \dots, x_m \in \mathbb{R}^n$. Lineární kombinace $\lambda_1 x_1 + \cdots + \lambda_m x_m$ se nazývá

(i) *konvexní*, jestliže $\lambda_1, \dots, \lambda_m \geq 0$ a $\sum_{i=1}^m \lambda_i = 1$;

(ii) *nezáporná*, jestliže $\lambda_1, \dots, \lambda_m \geq 0$;

(iii) *afinní*, jestliže $\sum_{i=1}^m \lambda_i = 1$.



Obrázek 6.11: Další ukázka nekonvexního kužele.

Už víme, že konvexní kombinace dvou bodů dává bod na úsečce mezi těmito dvěma body. Co to ale bude v případě konvexní kombinace tří bodů? A m-tice bodů?

Příklad 6.1.8

Konvexní kombinaci tří bodů

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3$$

pro libovolnou trojici $x_1 x_2, x_3 \in \mathbb{R}^n$ a $\lambda_1, \lambda_2, \lambda_3 \geq 0$ splňující $\lambda_1 + \lambda_2 + \lambda_3 = 1$ můžeme zapsat jako

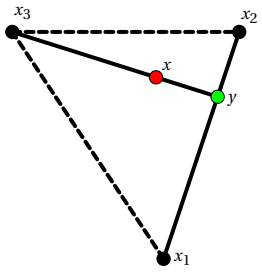
$$x = (\lambda_1 + \lambda_2) \underbrace{\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2 \right)}_{y} + \lambda_3 x_3.$$

Jelikož $(\lambda_1 + \lambda_2) + \lambda_3 = 1$ jedná se o konvexní kombinaci bodů x_3 a y, který je zase konvexní kombinací bodů x_1 a x_2 díky rovnosti $\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\lambda_2}{\lambda_1+\lambda_2} = 1$. Bod y bude proto ležet na spojnici x_1 a x_3, takže bod x leží na úsečce spojující body y a x_3, viz Obrázek 6.12. Samozřejmě bychom si na začátku mohli vybrat jiné dva body, což by ale na samotný postup nemělo moc velký vliv. Navíc je víceméně zřejmé, že bod x musí ležet v trojúhelníku vymezeném body x_1, x_2 a x_3.

V případě konvexní kombinace m-tice bodů bychom tento postup aplikovali postupně na $m-1$ konvexních kombinací dvojic. Výsledkem pak bude bod, který leží uvnitř nebo na hranici m-úhelníku tvořeného danými body, viz také Definici 6.1.10.

[6] Rozdíl mezi afinní množinou a vektorovým podprostorem je „pouze“ v tom, že vektorový podprostor musí obsahovat počátek.

[7] Je-li $X = \mathcal{R}(A, b)$, pak $\mathcal{U} = \ker A$ a platí $\dim X = n - \operatorname{rank} A$.



Obrázek 6.12: Konvexní kombinace trojice bodů z Příkladu 6.1.8.

Již velmi dobře víme, že vektorový podprostor je množina, která obsahuje všechny lineární kombinace libovolných *dvou* svých bodů (tj. s libovolnými dvěma body obsahuje i přímku procházející těmito body a počátek) a naopak. S využitím Definic 6.1.1 a 6.1.7 pak podobně dostaneme ekvivalentní popisy i dalších typů množin:

(i) afinní množina obsahuje všechny afinní kombinace libovolných *dvou* svých bodů;

(ii) konvexní kužel je množina obsahující všechny nezáporné kombinace libovolných *dvou* svých bod;

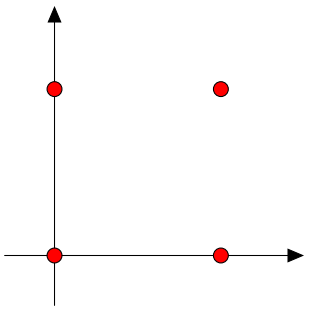
(iii) konvexní množina obsahuje všechny konvexní kombinace libovolných *dvou* svých bodů.

A jak to dopadne v případě kombinací s více než dvěma prvky? Úplně stejně, jak ukazuje následující tvrzení, které nám především dává jakýsi „vnitřní/úplný/technicistní“ popis konvexních množin, tj. tyto množiny musí obsahovat všechny konvexní kombinace svých prvků. Pro uzavřené konvexní množiny lze získat i „vnější/umělecký“ popis, což je jeden z nejdůležitějších výsledků konvexní analýzy: ukážeme, že to jsou množiny řešení (konečných či nekonečných) systémů neostrých lineárních nerovnic, viz Věta 6.3.11.

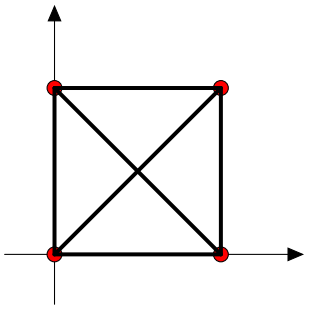
Věta 6.1.9

Nechť $X \subseteq \mathbb{R}^n$. Množina X je konvexní (konvexní kužel/afinní) právě tehdy, když libovolná konvexní (nezáporná/afinní) kombinace prvků z X je opět prvkem množiny X, tj. je uzavřená na všechny konvexní (nezáporné/afinní) kombinace.

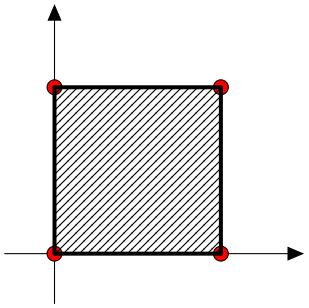
Důkaz. Tvrzení dokážeme pouze pro *konvexní* kombinace, neboť zbývající dvě varianty lze dokázat s použitím analogických argumentů. Protože implikace „$\Leftarrow$“ je zjevně triviální, ukážeme pouze směr „$\Rightarrow$“. Nechť množina X je konvexní. Je-li $X = \emptyset$ nebo jednoprvková je tvrzení opět triviální, zatímco pro ostatní případy lze k důkazu využít matematickou indukci.



Obrázek 6.13



Obrázek 6.14



Obrázek 6.15

Nechť tedy X obsahuje alespoň dva různé body. Z definice konvexnosti ihned vyplývá, že libovolná dvouprvková konvexní kombinace prvků z X je opět v X. Připusťme dále, že konvexní kombinace libovolné m-tice prvků z X je opět v X pro nějaké $m \in \mathbb{N}\setminus\{1\}$. Nyní uvažme konvexní kombinaci $m+1$ prvků z X, tj. bod $x := \sum_{i=1}^{m+1} \lambda_i x_i$ pro libovolná $x_1,\dots,x_{m+1} \in X$ a $\lambda_1,\dots,\lambda_{m+1} \geq 0$ splňující $\sum_{i=1}^{m+1} \lambda_i = 1$. Je-li $\lambda_{m+1} = 1$, pak $\lambda_1 = \dots = \lambda_m = 0$ a jistě $x \in X$. Je-li $\lambda_{m+1} \in [0,1)$, pak můžeme psát

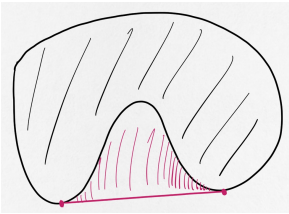
$$x = (1-\lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1}. \tag{6.1.4}$$

Ovšem $\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} = 1$, tudíž podle indukčního předpokladu máme $\sum_{i=1}^{m} \frac{\lambda_1}{1-\lambda_{m+1}} x_i \in X$, jelikož se jedná o konvexní kombinaci m-tice prvků. Pak ale x máme v (6.1.4) vyjádřeno jako konvexní kombinaci dvou prvků z X, z čehož díky konvexnosti X ihned vyplývá $x \in X$ a důkaz je hotov. ■

Ovšem co získáme konvexními kombinacemi prvků z množiny X, která není konvexní? Uvažme například čtveřici bodů $[0,0]$, $[0,1]$, $[1,1]$ a $[1,0]$, viz Obrázek 6.13. Pokud uvážíme konvexní kombinace dvojic vrcholů dostaneme obvod čtverce a jeho úhlopříčky, viz Obrázek 6.14. Další konvexní kombinace (ať již trojice či čtveřice vrcholů nebo dvojic bodů z obvodu čtverce) nám vyplní celý čtverec, viz Obrázek 6.15. Tím jsme nalezli konvexní množinu, která obsahuje všechny čtyři dané body. Existuje menší (vzhledem k množinové inkluzi) konvexní množina obsahující tyto body? Z Obrázků 6.14–6.15 je snad patrné, že nikoli, tj. že nalezený čtverec je vskutku nejmenší takovou množinou. Podařilo se nám tak původní čtyřbodovou množinu „zkonvexnit“ nejmenším možným způsobem. Totéž můžeme vidět také např. na Obrázku 6.16, který vyžaduje trochu odlišnou konstrukci. Podobné úvahy můžeme učinit i pro nalezení nejmenšího konvexního kužele či afinní množiny – tentokrát necháme jenom na čtenáři, aby si sám rozmyslel, jaký bychom v takovém případě dostali výsledek.

Návod pro formální definici tohoto „zkonvexňování“ nám dává již Věta 6.1.3(i): stačí uvážit průnik všech konvexních nadmnožin. Pak máme zajištěno, že výsledek bude jistě konvexní (jakožto průnik konvexních množin) a nejmenší možný (jakožto průnik všech nadmnožin), viz Obrázky 6.16–6.19.

DEFINICE 6.1.10



Obrázek 6.16: Nekonvexní množina (černou barvou) a její „zkonvexnění“ (po doplnění fialové části).

Nechť $X \subseteq \mathbb{R}^n$.

(i) Průnik všech konvexních množin obsahujících množinu X se nazývá *konvexní obal* množiny X a značí se $\operatorname{conv} X$.

(ii) Průnik všech konvexních kuželů obsahujících množinu X se nazývá *kónický* (*kuželový*) *obal* množiny X a značí se $\operatorname{cone} X$.

(iii) Průnik všech afinních množin obsahujících množinu X se nazývá *afinní obal* množiny X a značí se $\operatorname{aff} X$. Jeho zaměření se nazývá *lineární obal* množiny X a značí se $\operatorname{Lin} X$. *Dimenze* afinního obalu množiny X se značí $\dim X$ a klademe $\dim X := \dim \operatorname{Lin} X$.

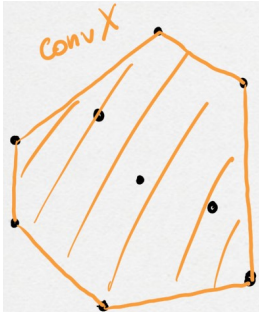
Poznámka 6.1.11. Je-li $X = \{x_1,\dots,x_m\} \subseteq \mathbb{R}^n$, pak $\operatorname{Lin} X$ *není* množina všech lineárních kombinací prvků $x_1,\dots,x_m$, tj. $\operatorname{span} X$! V případě $m = 1$ je samotná množina $X = \{x_1\}$ afinní, takže nutně $X = \operatorname{aff} X$, a tudíž $\operatorname{Lin} X = \{0\}$ a $\dim X = 0$. Pro $m = 2$ je $\operatorname{aff}\{x_1,x_2\}$ přímka procházející body x_1, x_2, což můžeme vyjádřit jako

$$\operatorname{aff} X = \{y \in \mathbb{R}^n \mid y = x_1 + t(x_2 - x_1),\ t \in \mathbb{R}\} = \{x_1\} + \underbrace{\operatorname{span}\{x_2 - x_1\}}_{\operatorname{Lin} X},$$

takže $\operatorname{Lin} X \neq \mathbb{R}^2 = \operatorname{span}\{x_1,x_2\}$, viz Obrázek 6.20. Vezmeme-li trojici bodů v $\mathbb{R}^2$ neležící na jedné přímce, pak již bude platit $\operatorname{aff} X = \operatorname{Lin} X = \mathbb{R}^2$ a přidáváním dalších bodů do X již nic jiného nezískáme, tj. $\operatorname{aff} X$ v $\mathbb{R}^2$ může být pouze[8] jednobodová množina, přímka nebo celé $\mathbb{R}^2$. Podobné

[8] Zde mlčky předpokládáme, že $X \neq \emptyset$, aby to nebylo příliš nudné.

[9] Opět ignorujeme možnost $X = \emptyset$.



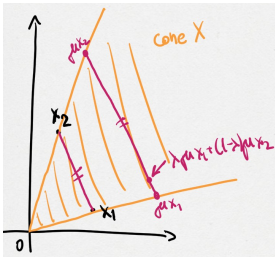
Obrázek 6.17: Konvexní obal množiny tvořené 9 izolovanými body.

úvahy v $\mathbb{R}^3$ nás přivedou k tomu, že $\operatorname{aff} X$ je buď[9] jednobodová množina, přímka, rovina nebo celé $\mathbb{R}^3$. V obecném případě je zaměření $\operatorname{aff} X$ tvořeno lineárními kombinacemi prvků $x_2 - x_1$, $x_3 - x_1, \ldots, x_m - x_1$, což nás přivádí k rovnosti

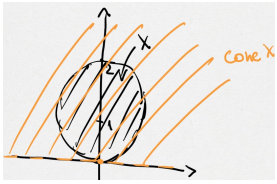
$$\operatorname{aff} X = \{x_1\} + \underbrace{\operatorname{span}\{x_2 - x_1,\ x_3 - x_1,\ \ldots,\ x_m - x_1\}}_{\operatorname{Lin} X}.$$

Pokud dokonce $0 \in X$, pak $\operatorname{aff} X$ je vektorový podprostor a $\operatorname{Lin} X = \operatorname{aff} X$. Podprostor $\operatorname{Lin} X$ můžeme případně získat i s využitím afinního obalu, neboť $\operatorname{Lin} X = \{y \in \mathbb{R}^n \mid y = x - x_0,\ x \in \operatorname{aff} X\}$ pro libovolné $x_0 \in \operatorname{aff} X$. ▲

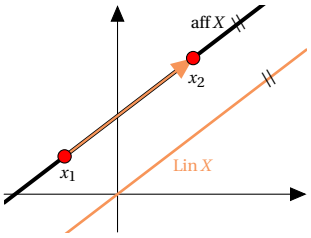
Věta 6.1.12



Obrázek 6.18: Kónický obal dvoubodové množiny $\{x_1, x_2\}$.



Obrázek 6.19: Kónický obal (oranžovou barvou) množiny $\{[x_1, x_2 \in \mathbb{R}^2 \mid x_1^2 + (x_2 - 1)^2 \le 1]\}$.



Obrázek 6.20

Pro libovolné $X \subseteq \mathbb{R}^n$ je $\operatorname{conv} X$ množina tvořená všemi konvexními kombinacemi bodů z X, tj.

$$\operatorname{conv} X = \Big\{x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \lambda_i x_i \text{ pro libovolné } m \in \mathbb{N} \text{ a } x_1, \ldots, x_m \in X,\ \lambda_1, \ldots, \lambda_m \ge 0,\ \sum_{i=1}^{m} \lambda_i = 1\Big\};$$

množina $\operatorname{cone} X$ je tvořená všemi nezápornými kombinacemi bodů z X, tj.

$$\operatorname{cone} X = \Big\{x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \lambda_i x_i \text{ pro libovolné } m \in \mathbb{N} \text{ a } x_1, \ldots, x_m \in X,\ \lambda_1, \ldots, \lambda_m \ge 0\Big\};$$

množina $\operatorname{aff} X$ je tvořená všemi afinními kombinacemi bodů z X, tj.

$$\operatorname{aff} X = \Big\{x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \lambda_i x_i \text{ pro libovolné } m \in \mathbb{N} \text{ a } x_1, \ldots, x_m \in X,\ \sum_{i=1}^{m} \lambda_i = 1\Big\}.$$

Důkaz. Opět dokážeme pouze první část tvrzení, neboť zbývající dvě lze ukázat pomocí analogických argumentů. Je-li množina X prázdná nebo jednobodová je tvrzení triviální. Nechť tedy $X \subseteq \mathbb{R}^n$ je alespoň dvouprvková a položme

$$Z := \Big\{x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \lambda_i x_i \text{ pro libovolné } m \in \mathbb{N} \text{ a } x_1, \ldots, x_m \in X,\ \lambda_1, \ldots, \lambda_m \ge 0,\ \sum_{i=1}^{m} \lambda_i = 1\Big\}.$$

Postupně ukážeme, že $\operatorname{conv} X \subseteq Z$ a $\operatorname{conv} X \supseteq Z$, z čehož ihned vyplyne tvrzení věty.

„$\subseteq$“ Je zřejmé, že $X \subseteq Z$. Proto $\operatorname{conv} X \subseteq \operatorname{conv} Z$. Jenže množina Z je konvexní[10]. Vskutku, vezměme libovolná $z_1, z_2 \in Z$, tj.

$$z_1 = \sum_{i=1}^{m} \alpha_i x_i \quad \& \quad z_2 = \sum_{i=1}^{k} \beta_i y_i$$

pro nějaká $m, k \in \mathbb{N}$, $x_1, \ldots, x_m, y_1, \ldots, y_k \in X$ a čísla $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_k \ge 0$ splňující $\sum_{i=1}^{m} \alpha_i = 1 = \sum_{i=1}^{k} \beta_i$, přičemž bez újmy na obecnosti můžeme předpokládat, že $m \le k$. Pak díky volbě $\alpha_{m+1} = \cdots = \alpha_k := 0$ můžeme brát z_1 a z_2 jako k-prvkové konvexní kombinace a z definice množiny Z plyne pro libovolné $\lambda \in [0, 1]$, že

$$\lambda z_1 + (1 - \lambda) z_2 = \sum_{i=1}^{k} (\lambda \alpha_i x_i + (1 - \lambda) \beta_i y_i) \in Z$$

neboť $\lambda \alpha_i \ge 0$, $(1 - \lambda) \beta_i \ge 0$ a současně

$$\sum_{i=1}^{k} (\lambda \alpha_i + (1 - \lambda) \beta_i) = \lambda + (1 - \lambda) = 1.$$

Proto $Z = \operatorname{conv} Z$, a tudíž $\operatorname{conv} X \subseteq \operatorname{conv} Z = Z$.

„$\supseteq$“ Nechť Y je libovolná konvexní množina taková, že $X \subseteq Y$. Podle Věty 6.1.9 je libovolná konvexní kombinace prvků z Y prvkem Y, což vzhledem k inkluzi $X \subseteq Y$ implikuje[11] $Z \subseteq Y$.

[9] ⚠ Toto je potřeba ukázat přímo a nelze využít Větu 6.1.9, neboť Z obsahuje konvexní kombinace bodů z X nikoli ze Z.

[10] V Z jsou všechny možné konvexní kombinace prvků z X. Protože $X \subseteq Y$ a Y je konvexní, musí být totéž i v Y a možná i něco „navíc“.

Ovšem $Y \supseteq X$ byla zvolena libovolně, takže nutně platí

$$Z \subseteq \bigcap_{\substack{Y \supseteq X \\ Y \text{ konvexní}}} Y \overset{\text{Definice } 6.1.10}{=} \operatorname{conv} X,$$

čímž je důkaz hotov. ∎

Je ale skutečně nutné popisovat $\operatorname{conv} X$ či $\operatorname{cone} X$ pomocí *úplně* všech možných konvexních/nezáporných kombinací prvků z X? Naštěstí ne, jak uvidíme z následujících dvou tvrzení. Nejdříve se podíváme na kónický obal.

Lemma 6.1.13

Nechť $X \subseteq \mathbb{R}^n$ a $x \in \operatorname{cone} X$ je libovolné. Pak existují body $x_1, \dots, x_n \in X$ a nezáporná čísla $\lambda_1, \dots, \lambda_n \geq 0$ taková, že $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$.

Důkaz. Nechť $x \in \operatorname{cone} X$ je libovolné. Podle Věty 6.1.12 existuje $m \in \mathbb{N}$, body $x_1, \dots, x_m \in X$ a čísla $\lambda_1, \dots, \lambda_m \geq 0$ taková, že $x = \sum_{i=1}^m \lambda_i x_i$. Jsou-li body $x_1, \dots, x_m$ „lineárně nezávislé“[12], pak nutně $m \leq n$ a tvrzení věty platí. Je-li $m > n$, pak tyto body jsou „lineárně závislé“. Proto existují čísla $\mu_1, \dots, \mu_m \in \mathbb{R}$ taková, že $\sum_{i=1}^m \mu_i x_i = 0$, přičemž $\mu_i > 0$ pro alespoň jeden index[13] $i \in \{1, \dots, n\}$. Nechť $\alpha \in \mathbb{R}$ je libovolné. Potom

$$x = \sum_{i=1}^m \lambda_i x_i - \alpha \sum_{i=1}^m \mu_i x_i = \sum_{i=1}^m (\lambda_i - \alpha \mu_i) x_i. \tag{6.1.5}$$

Zvolme nyní $\alpha := \min_{\mu_i > 0} \frac{\lambda_i}{\mu_i}$. Pak $\lambda_i - \alpha \mu_i \geq 0$ pro všechny[14] indexy $i \in \{1, \dots, m\}$, přičemž alespoň pro jeden index i platí $\lambda_i - \alpha \mu_i = 0$, tj. v souladu s (6.1.5) lze x vyjádřit jako nezápornou kombinaci nejvýše $m-1$ bodů. Je-li nyní $m - 1 = n$, je důkaz hotov. V opačném případě můžeme celou konstrukci zopakovat ještě $(m-n-1)$-krát, tj. dokud nedostaneme x jako nezápornou kombinaci nejvýše n bodů. ∎

Předchozí tvrzení je klíčem k analogickému (a mnohem důležitějšímu) výsledku pro konvexní obal. Skutečně první výskyt tohoto výsledku se určuje trochu složitěji a obvykle se v této souvislosti hovoří o dvou Carathéodoryho[15] článcích[16] z let 1907 a 1911 věnovaných zcela jinému tématu, viz [90, 91]. Zatímco Carathéodory se ve svém tvrzení omezil (v důsledků zvolené definice konvexnosti) pouze na kompaktní množiny, německý matematik E. Steinitz rozšířil v sérii tří článků z let 1913, 1914 a 1916 Carathéodoryho tvrzení na libovolnou množinu, viz zejména [459, §9–11] a také [460, 461].

Věta 6.1.14 Carathéodory

Nechť $X \subseteq \mathbb{R}^n$. Každý bod konvexního obalu $\operatorname{conv} X$ může být vyjádřen jako konvexní kombinace nejvýše $n+1$ prvků množiny X, tj. pro $x \in \operatorname{conv} X$ existují $x_1 \dots, x_{n+1} \in X$ a $\lambda_1, \dots, \lambda_n \geq 0$ splňující $\sum_{i=1}^m \lambda_i = 1$ taková, že

$$x = \lambda_1 x_1 + \cdots + \lambda_{n+1} x_{n+1}. \tag{6.1.6}$$

Avšak Věta 6.1.14 rozhodně nehovoří o tom, že by v množině X existovala nějaká $(n+1)$-tice, kterou by bylo možné považovat za „bázi“ množiny $\operatorname{conv} X$. O tom se snadno přesvědčíme v Příkladu 6.1.15 níže. Na druhou stranu, v [506] nalezneme výsledek ukazující, že je možné si vynutit, aby $(n+1)$-tice $x_1 \dots, x_{n+1}$ zahrnovala jeden předem daný bod z množiny X, tj. vždy k takovému bodu dokážeme dohledat vhodnou n-tici splňující (6.1.6).

Důkaz Věty 6.1.14. Je-li $X \subseteq \mathbb{R}^n$ prázdná nebo jednobodová množina, je tvrzení triviální. Proto nechť dále je $X \subseteq \mathbb{R}^n$ alespoň dvoubodová množina a uvažme množinu $A := \{[x, 1] \mid x \in X\} \subseteq \mathbb{R}^{n+1}$. Pak platí, že $y = [x, 1] \in \operatorname{cone} A$ právě tehdy, když $x \in \operatorname{conv} X$. Vskutku, nechť $x \in \operatorname{conv} X$, tj. $x = \sum_{i=1}^m \lambda_i x_i$ pro nějaké $m \in \mathbb{N}$, $x_1, \dots, x_m \in X$ a $\lambda_1, \dots, \lambda_m \geq 0$ splňující $\sum_{i=1}^m \lambda_i = 1$. Pak $[x, 1] \in \operatorname{cone} A$, neboť

$$[x, 1] = \left[\sum_{i=1}^m \lambda_i x_i, 1 \right] = \left[\sum_{i=1}^m \lambda_i x_i, \sum_{i=1}^m \lambda_i \right] = \sum_{i=1}^m \lambda_i [x_i, 1],$$

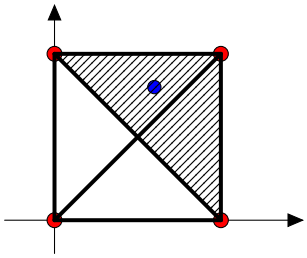
[12] Zde pochopitelně využíváme toho, že příliš nerozlišujeme mezi body a vektory.

[13] Alespoň jedno μ_j musí být nenulové a kdyby všechna $\mu_1, \dots, \mu_m$ byla všechna nezáporná, pak jen u všech stačí změnit znaménko, čímž budeme mít alespoň jedno z těchto čísel kladné.

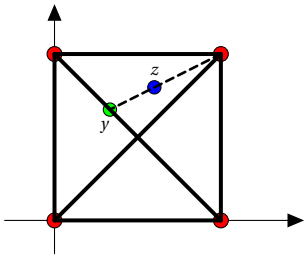
[14] Je-li $\mu_i = 0$, je tato nerovnost jistě splněna díky požadavkům na λ_i. Je-li $\mu_i < 0$, pak $-\alpha \mu_i \geq 0$, takže máme součet dvou nezáporných čísel, který je také nezáporný. Konečně v případě $\mu_i > 0$, musí být dle definice $\alpha \leq \lambda_i / \mu_i$, což dává požadovanou nerovnost.

[15] Constantin Carathéodory (13. září 1873 – 2. února 1950) byl řecký matematik žijící a pracující převážně v Německu. Výrazně ovlivnil rozvoj variačního počtu, teorie míry a teorie funkcí reálné proměnné. Jeho jméno je také spojeno s domněnkou (publikovanou Hansem Ludwigem Hamburgerem v roce 1924 – samotný Carathéodory tuto domněnku nikdy neformuloval, ale napsal článek na podobné téma) ohledně existence alespoň dvou „pupíků“ (angl. *umbilical points*) na každé dostatečně hladké uzavřené konvexní ploše v $\mathbb{R}^3$. Kompletní řešení této domněnky zatím zůstává otevřené.

[16] Nebudeme se zde pouštět do zbytečně komplikované matematicko–lingvistické analýzy nuancí jednotlivých „formulací“ z těchto článků, abychom rozhodli určili, který z těch zdrojů bychom měli uvádět. Smiřme se s tím, že korektnější je kombinace obou... Viz také [73, §2–3] pro více podrobností.



Obrázek 6.21



Obrázek 6.22

[17] Větu 6.1.14 lze dokázat i bez využití Lemma 6.1.13. My si tento důkaz nyní pro zajímavost ukážeme, avšak všímavý čtenář v něm jistě odhalí, že je to vlastně kompilace obou předchozích důkazů – jen se v něm explicitně nehovoří o kónickém obalu.

Přímý důkaz Věty 6.1.14. Je-li $X \subseteq \mathbb{R}^n$ prázdná nebo jednobodová množina, je tvrzení triviální. Nechť tedy $X \subseteq \mathbb{R}^n$ je alespoň dvoubodová množina a $x \in \operatorname{conv} X$ je libovolné. Pak existují $x_1, \dots, x_m \in X$ a čísla $\lambda_1, \dots, \lambda_m \geq 0$ splňující $\sum_{i=1}^m \lambda_i = 1$ taková, že

$$x = \sum_{i=1}^{m} \lambda_i x_i. \qquad (\triangle)$$

Je-li $m \leq n+1$ je důkaz hotov. Je-li $m > n+1$, pak ukážeme, že x lze vyjádřit jako konvexní kombinaci nejvýše $m-1$ bodů. Opakováním této procedury pak dosáhneme po konečném počtu kroků snížení po-

Věta 6.1.16 Bárány

potřebných bodů až na $n+1$. Nechť tedy $m > n+1$ a uvažme $(m-1)$-tici vektorů $\mathbb{R}^n$ tvaru

$$x_1 - x_m,\ x_2 - x_m, \dots,\ x_{m-1} - x_m.$$

Poněvadž $m-1 > n$, musí tyto vektory být lineárně závislé. Tedy existují čísla $\mu_1, \dots, \mu_{m-1} \in \mathbb{R}$, která nejsou všechna současně nulová a splňují

$$\left.\begin{aligned} &\mu_1 (x_1 - x_m) + \dots + \\ &\quad + \mu_{m-1} (x_{m-1} - x_m) = 0. \end{aligned}\right\} \quad (*)$$

Položme $\mu_m := -\sum_{i=1}^{m-1} \mu_i$. Potom

$$\left.\begin{aligned} &\mu_1 x_1 + \dots + \mu_{m-1} x_{m-1} + \\ &\qquad + \mu_m x_m = 0 \\ \text{a} \quad &\sum_{i=1}^{m} \mu_i = 0. \end{aligned}\right\} \quad (\Diamond)$$

což je nezáporná kombinace bodů z A. Naopak, je-li $[x,1] \in \operatorname{cone} A$, pak $[x,1] = \sum_{i=1}^m \lambda_i [x_i, 1]$ pro nějaké $m \in \mathbb{N}$, $x_1, \dots, x_m \in X$ a $\lambda_1, \dots, \lambda_m \geq 0$, tj. $x = \sum_{i=1}^m \lambda_i x_i$ a $\sum_{i=1}^m \lambda_i = 1$ neboli $x \in \operatorname{conv} X$. Jenže díky Lemma 6.1.13 už víme, že nám k takovému vyjádření $[x,1]$ stačí nejvýše $n+1$ bodů z A, potažmo $n+1$ bodů z X pro samotné x, čímž je tvrzení dokázáno[17]. ■

Příklad 6.1.15

Uvažme opět $X = \{[0,0],\ [0,1],\ [1,1],\ [1,0]\}$ jako na Obrázku 6.13. Už jsem ukázali, že $\operatorname{conv} X$ je celý čtverec s vrcholy v těchto bodech. Pak Carathéodoryho věta říká, že libovolný bod v tomto čtverci dokážeme vyjádřit jako konvexní kombinaci nejvýše trojice bodů z X, tj. vrcholů čtverce. Vezmeme-li např. bod $[0,1/2]$, pak nám dokonce stačí jenom vrcholy $[0,0]$ a $[0,1]$, neboť

$$[0,1/2] = \tfrac{1}{2}[0,0] + \left(1 - \tfrac{1}{2}\right)[0,1].$$

Dva body nám stačí pro libovolný bod ležící na obvodu čtverce nebo na některé z jeho uhlopříček, viz Obrázek 6.14. Pro ostatní body v $\operatorname{conv} X$ už jsou ale potřeba tři vrcholy, tj. pro libovolný bod uvnitř čtverce a mimo uhlopříčky vždy existuje (alespoň jeden) trojúhelník s vrcholy v bodech množiny X, v jehož vnitřku daný bod leží. Např. pro $x = [3/5, 4/5]$ může využít vrcholy $[1,0]$, $[0,1]$ a $[1,1]$, neboť

$$[3/5, 4/5] = \tfrac{1}{5}[1,0] + \tfrac{2}{5}[0,1] + \tfrac{2}{5}[1,1],$$

viz Obrázek 6.21. Označíme-li vrcholy (v jakémkoli pořadí) jako x_1, x_2, x_3 a x_4, pak konvexní kombinací prvních dvou z těchto bodů dostaneme bod $y = \lambda x_1 + (1-\lambda) x_2$ na obvodu čtverce nebo některé z jeho úhlopříček. Je-li z uvnitř trojúhelníku s vrcholy v x_1, x_2, x_3, pak jej lze získat jako konvexní kombinaci bodu y s vhodnou volbou λ a vrcholu x_3, tj.

$$z = \mu y + (1-\mu) x_3 = \mu\lambda x_1 + \mu(1-\lambda) x_2 + (1-\mu) x_3, \qquad (6.1.7)$$

přičemž $\mu\lambda + \mu(1-\lambda) + (1-\mu) = 1$, tj. vyjádření (6.1.7) je pro vhodná $\lambda, \mu \in [0,1]$ hledaná konvexní kombinace. Avšak jedna zvolená trojice nám umožní získat pouze polovinu čtverce, což znamená, že ji nelze považovat za „bázi“ $\operatorname{conv} X$.

Existují i různá rozšíření Carathéodoryho věty. Jedním z nich je např. „vícenásobná“ verze od Bárányho z roku 1982, viz [26].

Nechť jsou dány množiny $X_1, \dots, X_{n+1} \subseteq \mathbb{R}^n$ takové, že existuje $p \in \bigcap_{i=1}^{n+1} \operatorname{conv} X_i$. Pak Pak existuje množina $P = \{p_1 \in X_1, p_2 \in X_2, \dots, p_{n+1} \in X_{n+1}\}$ taková, že $x \in \operatorname{conv} P$.

I zde platí, že jeden z bodů $x_1, \dots, x_{n+1}$ může být určen s předstihem, a v případě $X_1 = \dots = X_{n+1}$ dostaneme Větu 6.1.14. Toto zobecnění začalo být později označováno jako *pestrobarevná Carathéodoryho věta*, neboť jednotlivé množiny mohou být reprezentovány různými barvami, a tak bod p bude prvkem konvexního obalu různobarevných bodů[18]. Každopádně opět platí, že počet množin nemůže být menší než $n+1$. Pochopitelně bylo vynaloženo i nemalé úsilí na zjištění, za jakých podmínek pro množinu X by bylo možné toto „Carathéodoryho číslo“ $n+1$ ve Větě 6.1.14 snížit – např. v případě souvislé množiny X bude potřeba nejvýše n bodů apod. – byť třeba za cenu toho, že s tímto menším počtem bodů nalezneme jen nějakou ε-aproximaci zvoleného bodu, což je úzce spojeno s kvadratickým programováním a Frankové–Wolfeho algoritmem zmíněným v Kapitole 4. Kromě obvyklých „hrátek“ s konvexními množinami lze Carathéodoryho větu využít také v důkazu Haarovy věty o jednoznačnosti jisté nejlepší aproximace. Pro více podrobností viz např. [27, 41, 97, 128, 151].

Jako důsledek Carathéodoryho věty je velmi často uváděna další důležitá vlastnost konvexních obalů, kterou je zachování kompaktnosti.

Důsledek 6.1.17

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní množina. Pak její konvexní obal $\operatorname{conv} X$ je také kompaktní.

Důkaz. Zaveďme množiny

$$\Lambda := \{\lambda \in \mathbb{R}^{n+1} \mid \lambda_1, \dots, \lambda_{n+1} \geq 0,\ \sum_{i=1}^{n+1} \lambda_i = 1\} \quad \text{a} \quad Y := \Lambda \times X^{n+1} \subseteq \mathbb{R}^{(n+1)^2}$$

a definujme zobrazení $F : Y \to \mathbb{R}^n$ předpisem

$$F(\lambda, x_1, \dots, x_{n+1}) := \sum_{i=1}^{n+1} \lambda_i x_i.$$

Pak množina Λ je konvexní a kompaktní[19], a tudíž množina Y je také kompaktní jakožto kartézský součin kompaktních množin[20]. Podle Věty 6.1.14 platí

$$\operatorname{conv} X = F(Y) = \{F(\lambda, x_1, \dots, x_{n+1}) \mid \lambda \in \Lambda,\ x_i \in X\}.$$

Jelikož ale F je spojité zobrazení, je také množina $\operatorname{conv} X$ kompaktní jakožto spojitý obraz kompaktní množiny, viz např. [141, Věta 4.10]. ■

Poznámka 6.1.18.

(i) Kompaktnost už je poměrně silná vlastnost. Samotná uzavřenost bez ohraničenosti se ale pro konvexní obal již zachovat nemusí. Uvažme např. dvojici množin

$$X := \{[x_1, x_2] \in \mathbb{R}^2 \mid x_1 > 0,\ x_2 > 0,\ x_1 x_2 \geq 1\}$$

a

$$Y := \{[y_1, y_2] \in \mathbb{R}^2 \mid y_1 > 0,\ y_2 < 0,\ y_1 y_2 \leq -1\},$$

viz Obrázek 6.23. Definujeme-li $Z := X \cup Y$, pak Z je uzavřená jakožto sjednocení dvou uzavřených množin. Ovšem její konvexní obal

$$\operatorname{conv} Z = \{[z_1, z_2] \in \mathbb{R}^2 \mid z_1 > 0\}$$

je otevřená, viz Obrázek 6.24.

(ii) Na druhou stranu, je-li množina X otevřená, pak asi není pochyb o tom, že i její konvexní obal $\operatorname{conv} X$ musí být otevřenou množinou. Je snad také jasné, že ani sjednocení nebude s konvexními obaly fungovat dobře. Avšak to neplatí pro (Minkowského) sčítání množin, které je komutativní s konvexním obalem, tj. platí[21]

$$\operatorname{conv}(X + Y) = \operatorname{conv} X + \operatorname{conv} Y.$$

(iii) *Extrémním bodem* konvexní množiny X rozumíme bod, který neleží mezi koncovými body žádné úsečky určené dvěma jinými body množiny X, tj. nelze jej vyjádřit jako netriviální konvexní kombinaci dvou bodů z množiny X. Díky tomu je jasné, že každý extrémní bod konvexního obalu libovolné množiny X musí být součástí již samotné množiny X, protože jinak by se do $\operatorname{conv} X$ nedostal. Speciálním případem Kreinovy–Milmanovy věty[22] pak je tvrzení, že každá kompaktní konvexní množina je konvexním obalem svým extrémních bodů. ▲

Carathéodoryho větou jsme se dostali až do oblasti, která leží na pomezí diskrétní/kombinatorické[23] a konvexní geometrie[24]. V této oblasti je nesmazatelně zapsána i výrazná stopa s českým původem v podobě následujícího výsledku z roku 1921, viz [425].

Vezmeme-li libovolné $t \in \mathbb{R}$, pak díky rovnostem (△) a (∗) platí

$$x = \sum_{i=1}^{m} (\lambda_i - t\mu_i) x_i. \qquad (\square)$$

Nyní zbývá ukázat, že existuje volba $t = t^*$ taková, že toto je ve skutečnosti nejvýše $(m-1)$-bodová konvexní kombinace bodů z X, tj. $\lambda_i - t^*\mu_i \geq 0$, $\sum_{i=1}^{m}(\lambda_i - t^*\mu_i) = 1$ a alespoň pro jeden index $i \in \{1, \dots, m\}$ nastane rovnost $\lambda_i - t^*\mu_i = 0$. Nechť $I := \{i \mid i \in \{1, \dots, m\} \ \&\ \mu_i > 0\}$. Pak díky nenulovosti alespoň jednoho μ_i a druhé rovnosti v (◊) musí být $I \neq \emptyset$. Proto můžeme definovat

$$t^* := \min\{\lambda_i/\mu_i \mid i \in I\},$$

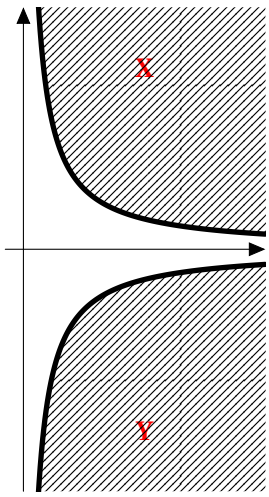
přičemž toto minimum nastane v nějakém indexu $j \in I$, tj. $t^* = \lambda_j/\mu_j$. Pak $t^* \geq 0$ a pro všechna $i \in I$ platí

$$(\lambda_i - t^*\mu_i) = \mu_i \left(\frac{\lambda_i}{\mu_i} - \frac{\lambda_j}{\mu_j}\right) \geq 0,$$

přičemž pro $i = j$ nastane rovnost. Pokud $i \notin I$, pak $\mu_i \leq 0$, a tudíž jistě $\lambda_i - t^*\mu_i \geq 0$. Tímto jsme tedy dostali vyjádření x v (□) jako kónickou kombinaci nejvýše $m-1$ bodů. Avšak současně

$$\sum_{i=1}^{m}(\lambda_i - t^*\mu_i) = \sum_{i=1}^{m}\lambda_i - t^* \underbrace{\sum_{i=1}^{m}\mu_i}_{=0} = 1,$$

takže t^* splňuje všechny tři podmínky zmíněné výše, tj. (□) s $t = t^*$ je hledaná konvexní kombinace nejvýše $m-1$ bodů z X, čímž je důkaz hotov. ■

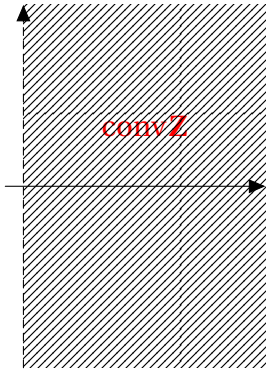


Obrázek 6.23

Věta 6.1.19 Radon[25]

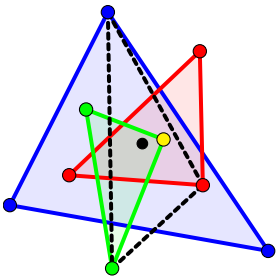
Libovolná množina $X \subseteq \mathbb{R}^n$ s alespoň $n+2$ prvky může být rozdělena do dvou množin, jejichž konvexní obaly mají neprázdný průnik, tj. pro body $a_1, \dots, a_m$ s $m \geq n+2$ existují

Věta 6.1.19 Radon[26]



Obrázek 6.24

[18] Uvažme např. zelený, červený a modrý trojúhelník, jejichž průnik obsahuje počátek. Pak existuje zelený, červený a modrý vrchol, které vytvoří pestrobarevný trojúhelník obsahující počátek, viz Obrázek 6.25.



Obrázek 6.25: Ilustrace pestrobarevné Carathéodory věty (černý bod reprezentuje počátek).

[19] Co je vlastně Λ? Je-li $n=0$, pak $\Lambda=\{1\}$. Pro $n=1$ máme úsečku v $\mathbb{R}^2$ s krajními body $[1,0]$ a $[0,1]$. Pro $n=2$ se jedná o trojúhelník s vrcholy v bodech $[1,0,0]$, $[0,1,0]$ a $[0,0,1]$. V obecném případě jde o n-simplex.

disjunktní indexové množiny $I,J\subseteq\{1,\dots,m\}$ takové, že $I\cup J=\{1,\dots,m\}$ a

$$\operatorname{conv}\{a_i \mid i\in I\}\cap\operatorname{conv}\{a_j \mid j\in J\}\neq\emptyset. \tag{6.1.8}$$

Důkaz. Pro $i\in\{1,\dots,m\}$ definujme body $b_i:=[a_i,1]\in\mathbb{R}^{n+1}$. Avšak nyní máme nejméně $n+2$ bodů v $\mathbb{R}^{n+1}$, což znamená, že jsou lineárně závislé podobně jako v důkazu Lemma 6.1.13, tj. musí existovat čísla $\mu_1,\dots,\mu_m\in\mathbb{R}$ taková, že ne všechna jsou současně nulová a zároveň platí $\sum_{i=1}^m \mu_i b_i=0$. Zvolme indexové množiny

$$I:=\{i\in\{1,\dots,m\}\mid\mu_i\geq 0\}\quad\text{a}\quad J:=\{i\in\{1,\dots,m\}\mid\mu_i<0\}$$

Očividně jsou tyto množiny disjunktní a jejich sjednocení obsahuje každé z čísel $1,\dots,m$. Navíc, protože poslední souřadnice prvků $b_1,\dots,b_m$ rovny 1, jsou obě tyto množiny neprázdné a platí

$$\alpha:=\sum_{i\in I}\mu_i=-\sum_{j\in J}\mu_j\neq 0.$$

Díky tomu je zjevné, že bod

$$\sum_{i\in I}\frac{\mu_i}{\alpha}a_i=-\sum_{j\in J}\frac{\mu_j}{\alpha}a_j$$

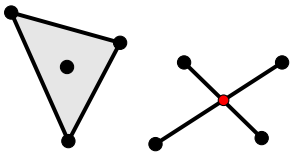
je konvexní kombinací[27] prvků z množin $\{a_i\mid i\in I\}$ a $\{a_j\mid j\in J\}$, takže patří do $\operatorname{conv}\{a_i\mid i\in I\}$ a také $\operatorname{conv}\{a_j\mid j\in J\}$. ■

Bod v průniku konvexních obalů z předchozí věty se nazývá *Radonův bod*. Tento bod pro libovolnou čtveřici bodů je jejich tzv. *geometrickým mediánem*, který minimalizuje součet vzdáleností k ostatním bodům, srovnejte s Příkladem 1.2.9. Pravděpodobnostní algoritmus, ve kterém jsou opakovaně nahrazovány $(n+2)$-bodové množiny jejich Radonovým bodem, se využívá k nalezení přibližného středu libovolné bodové množiny. Tento algoritmus je dokonce polynomiální vzhledem k počtu bodů i dimenzi prostoru. Radonova věta se dočkala mnoha rozšíření a zobecnění jako např. v [150].

Vzhledem k podobnosti důkazů Radonovy věty a Lemma 6.1.13 by nás nemělo překvapit, že lze důkaz Carathéodoryho věty založit i na využití Radonovy věty. To by však nebylo chronologicky správně, a navíc samotný Carathédory svůj důkaz založil na existenci tzv. *opěrných nadrovin*, viz Podkapitolu 6.3 a již dříve zmíněnou monografii [73, §2–3]. A jaký je mezi těmito tvrzeními rozdíl? Pro jednoduchost se omezme pouze na čtyřbodovou množinu $X\subseteq\mathbb{R}^2$. Pak podle Carathédoryho věty musí každý bod z $\operatorname{conv}X$ buď být prvkem X, ležet na úsečce spojující nějaké dva jiné body z X, nebo ležet v trojúhelníku určeném zbývající trojicí bodů z X. Na druhou stranu, z Radonovy věty vyplývá, že buď existuje bod z X, který leží v trojúhelníku vytvořeném zbývajícími třemi body, nebo body z X lze rozdělit do dvou dvojic, které určují protínající se úsečky, viz Obrázek 6.26.

Následující zobecnění Radonovy věty pochází z [486] a jeho překvapivě zjednodušený důkaz byl publikován v [441] s využitím Bárányho věty.

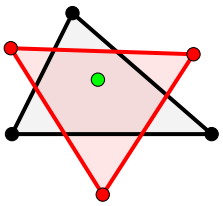
Věta 6.1.20 Tverberg



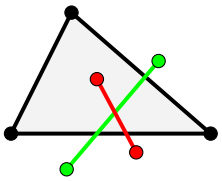
Obrázek 6.26: Ilustrace Radonovy věty pro čtyřbodovou (černé body) množinu v $\mathbb{R}^2$.

Nechť je dáno nejméně $(r-1)(n+1)+1$ bodů v $\mathbb{R}^n$. Tyto body mohou vždy být rozděleny na r částí, jejichž konvexní obaly mají neprázdný průnik.

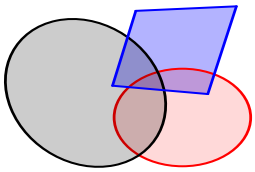
Pro $r=2$ dostaneme právě Radonovu větu, takže tím nejjednodušším (a netriviálním) příkladem posledního tvrzení je volba $r=3$ v $\mathbb{R}^2$. Pak máme 7 bodů, která bez ohledu na jejich rozmístění mohou být vždy rozděleny do 3 množin, jejichž konvexní obaly mají neprázdný průnik, viz Obrázky 6.27–6.28. Důkaz Tverbergovy věty je závislý na existenci množiny P z Bárányho věty. Avšak doposud není (a kdoví, jestli někdy vůbec bude) znám algoritmus s polynomiální časovou složitostí pro její nalezení. Gerard Sierksma dokonce vyslovil v roce 1979 domněnku, že body mohou být dokonce rozděleny nejméně $(r-1)!^n$ způsoby pro splnění závěru



Obrázek 6.27: První ilustrace Tverbergovy věty pro $r = 3$.

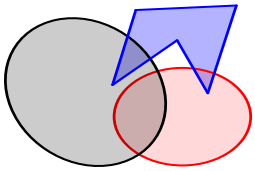


Obrázek 6.28: Druhá ilustrace Tverbergovy věty pro $r = 3$.

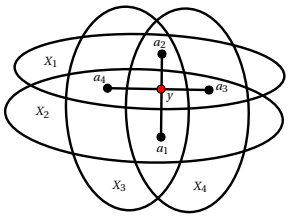


Obrázek 6.29: Ilustrace Hellyho věty pro trojici konvexních množin v $\mathbb{R}^2$ (průnik všech tří množin je neprázdný).

Věta 6.1.21 Helly



Obrázek 6.30: Ilustrace neplatnosti Hellyho věty v případě jedné nekonvexní množiny (průnik všech tří množin je prázdný).



Obrázek 6.31: Ilustrace využití Radonovy věty v důkazu Hellyho věty.

[20] Máme konečný počet kompaktních množin, takže postupným výběrem podposloupností z podposloupností můžeme díky kompaktnosti jednotlivých množin z libovolné posloupnosti v Y vybrat konvergentní podposloupnost. V případě součinu nekonečně mnoha kompaktních množin je toto tvrzení Tverbergovy věty. Nicméně potvrzení nebo vyvrácení této domněnky stále zůstává nezodpovězené. V případě $r = p^k$ pro nějaké prvočíslo p se podařilo Stephanu Hellovi ukázat, že počet těchto možností je nejméně

$$\frac{1}{(r-1)!}\left(\frac{r}{k+1}\right)^{\frac{(n+1)(r-1)}{2}}.$$

Pro mnohem více podrobností viz [28].

Radonova věta je také klíčovým nástrojem pro důkaz následující Hellyho[28] věty. To také bylo hlavní Radonovou motivací pro odvození předchozího výsledku, jak jej najdeme v před chvílí zmíněném Radonově článku [425]. Radon se o tomto tvrzení dozvěděl na Hellyho přednášce pro Vídeňskou matematickou společnost, kterou pronesl ještě v předválečném období. Helly tento výsledek implicitně využíval již v článku [260] v jeho nejjednodušší podobě: Systém $\mathcal{S}$ kompaktních intervalů na reálné ose má neprázdný průnik, jestliže mají neprázdný průnik každá dva intervaly z $\mathcal{S}$. Důkaz je poměrně snadný: je-li $\mathcal{S} := \{[a_i, b_i] \mid i \in I\}$, pak zjevně platí[29] $\sup\{a_i \mid i \in I\} \in \bigcap_{i\in I}[a_i, b_i]$. Tento výsledek využil Helly pro první důkaz (trochu zjednodušené verze) Hahnovy–Banachovy věty, což je ústřední nástroj funkcionální analýzy.

Avšak intervaly jsou současně i konvexní množiny, takže se nabízí otázka, zda to platí i pro jiné takové množiny. V $\mathbb{R}$ už nic dalšího nezjistíme, protože intervaly jsou pochopitelně *jediné* konvexní množiny. Musíme se proto vydat alespoň do $\mathbb{R}^2$, což je znázorněno na Obrázcích 6.29–6.30. K tomuto zobecnění Helly sice dospěl už v roce 1913, ale svůj důkaz publikoval až o 10 let později v [261], tj. až po Radonově článku. Za zmínku ještě stojí, že důkaz Hellyho věty publikoval v roce 1922 i maďarský matematik Dénes König. Ten se o Hellyho větě dozvěděl z Radonova článku a jeho důkaz se příliš od toho Hellyho neliší. Náš důkaz bude (celkem očekávatelně) založen právě na Radonově větě, pro další varianty viz [128, 399].

> Nechť je dán konečný systém konvexních množin $X_1, \dots, X_m \subseteq \mathbb{R}^n$ pro $m \geq n+1$. Pokud průnik libovolné $(n+1)$-tice těchto množin má neprázdný průnik, pak celá m-tice množin má neprázdný průnik, tj. $\bigcap_{i=1}^m X_i \neq \emptyset$.
> Tvrzení zůstane v platnosti i v případě, že je dán dokonce nekonečný systém konvexních množin $X_1, X_2, \dots \subseteq \mathbb{R}^n$, avšak tyto množiny musí být navíc i kompaktní.

Důkaz. Uvažme nejdříve, že počet množin je konečný. K důkazu využijeme matematickou indukci. Je-li $m = n+1$, pak je tvrzení triviální. Nechť tedy $m > n+1$ a připusťme, že tvrzení platí pro $m-1$. Díky tomuto indukčnímu předpokladu jistě pro všechna $i = 1, \dots, m$ existují body $a_i \in \bigcap_{\substack{j=1 \\ j\neq i}}^m X_j$. Podle Věty 6.1.19 můžeme množinu $\{1, \dots, m\}$ rozdělit do dvou disjunktních množin I a J tak, že platí (6.1.8). Označme tento bod v průniku konvexních obalů jako y. Zbývá ukázat, že $y \in X_k$ pro všechna $k = 1, \dots, m$. Bez újmy na obecnosti zvolme index $k \in I$. Pak pro všechna $j \in J$ máme $a_j \in X_k$ podle definice těchto bodů, viz Obrázek 6.31. Avšak množina X_k je konvexní, a tak $\operatorname{conv}\{a_j \mid j \in J\} \subseteq X_k$, což znamená, že $y \in \operatorname{conv}\{a_j \mid j \in J\} \subseteq X_k$. Jelikož index k byl zvolen libovolně, plyne odtud $y \in X_k$ pro všechna $k = 1, \dots, m$, jak jsme chtěli ukázat.

Nechť je nyní dán nekonečný systém množin, kde libovolná $(n+1)$-tice má vždy neprázdný průnik. Pak podle první části má i libovolný konečný počet těchto množin neprázdný průnik. Tomu se říká *vlastnost konečných průniků*[30], která společně s kompaktností všech množin zaručuje neprázdnost průniku $\bigcap_{i=1}^\infty X_i$, viz [386, Theorem 26.9]. ■

Poznámka 6.1.22. K podmínkám Hellyho věty ještě doplňme dvě poznámky.

(i) Nestačí požadovat neprázdný průnik pro menší počet množin než $n+1$. To si snadno můžeme ilustrovat na trojici úseček tvořící obvod trojúhelníku v $\mathbb{R}^2$: libovolné dvě úsečky mají neprázdný průnik, ale průnik všech tří už prázdný je.

(ii) Bez dodatečného předpokladu kompaktnosti jednotlivých množin v případě jejich nekonečného počtu tvrzení neplatí. Uvažme např. množiny $X_j = \{x \in (0, 1/j)\}$ pro $j \in \mathbb{N}$.

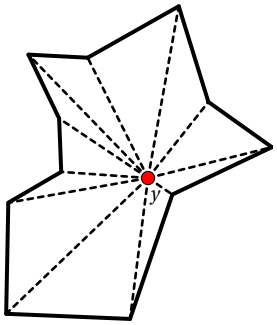
Každá z těchto množin je konvexní a libovolná dvojice má neprázdný průnik, avšak průnik všech těchto množin bude zjevně prázdný. ▲

Hellyho věta se v průběhu času dočkala různých rozšíření a zobecnění. My si některá z nich ukážeme společně s jejich využitím v rozličných kombinatorických úlohách, pro více podrobností viz [128,150,399]. Začneme něčím jednodušším. Na louce je v daný okamžik několik ovcí a vlků. Za jakých rozumných podmínek se nám rovným plotem podaří oddělit ovce od vlků? Odpověď je ukryta v následujícím tvrzení, jehož původně dvacetistránkový důkaz z roku 1902 lze s pomocí Hellyho věty zkrátit cca. na devět řádků, viz [128, Theorem 2.4] a srovnej s [320].

Věta 6.1.23 Kirchberger[31]

Nechť X a Y jsou konečné podmnožiny v $\mathbb{R}^n$. Potom lze tyto množiny ostře oddělit[32] nadrovinou právě tehdy, když pro každou nejvýše $(n+2)$-bodovou množinu $Z \subseteq X \cup Y$ lze ostře oddělit množiny $Z \cap X$ a $Z \cap Y$.

Odtud vidíme, že vlky a ovce oddělíme „přímkovým“ plotem právě tehdy, když takto oddělíme každou čtveřici zvířat. Mezi další zajímavé aplikace Hellyho věty patří *problém hlídače v obrazové galerii*. V ní jde o rozhodnutí, zda jeden hlídač dokáže uhlídat (tj. vidět) současně všechny obrazy umístěné v různých vzájemně propojených sálech (i třeba nepravidelného tvaru)[33]. Dodejme, že (matematicky řečeno) bod $x \in X$ nazveme *viditelným* z bodu y v rámci množiny X, jestliže celá úsečka spojující x a y patří do X. Jestliže existuje bod, z něhož jsou všechny body z X viditelné v X, pak se množina X nazývá *hvězdovitá*[34,35] (nebo též *radiálně konvexní množina*), viz Obrázek 6.32. Hlídač tedy potřebuje vědět, zda příslušná galerie je hvězdovitá, aby ji nemusel zbytečně obcházet. Postačující kombinatorickou podmínku udává následující věta, viz [128, Theorem 2.5] a [350, Theorem 6.10].



Obrázek 6.32: Hvězdovitá množina a vhodné umístění „pozorovatelny“.

Věta 6.1.24 Krasnoselsky[36]

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní podmnožina obsahující alespoň $n+1$ bodů. Jestliže pro každou $(n+1)$-bodovou podmnožinu množiny X existuje bod, z něhož jsou všechny její body viditelné v rámci X, pak množina X je hvězdovitá.

To znamená, že hlídači stačí zkontrolovat, že pro libovolnou trojici obrazů dokáže v galerii najít místo, ze kterého na ni vidí. Existují však i aplikace Hellyho věty, které nemají tento „kombinatorický“ charakter. Např.: *Je-li dáno p bodů v $\mathbb{R}^2$ takových, že libovolné tři jsou obsaženy v jednotkovém kruhu, pak všechny body leží v jednotkovém kruhu.* Vskutku, uvážíme-li jednotkové kruhy $C_1, \ldots, C_p$ se středy v daných bodech, pak jejich libovolná trojice musí mít neprázdný průnik. Proto podle Hellyho věty všechny kruhy mají neprázdný průnik, a jakýkoli bod q v tomto průniku musí mít vzdálenost menší než 1 od středu libovolné kružnice (tj. od těch daných bodů). To znamená, že všechny body leží v jednotkové kružnici se středem v bodě q. Poznamenejme, že číslo 1 v tomto tvrzení může být nahrazeno kterýmkoli jiným kladným číslem.

Toto byla velmi zjednodušená verze *Jungovy*[37] *věty*, viz [299, 300]. Nenechme se zmást tím, že oba uvedené články jsou starší než Hellyho věta – je v nich totiž uvažován pouze rovinný případ. Hellyho věta umožňuje snadné zobecnění pro libovolné $\mathbb{R}^n$. Toto tvrzení se dokonce využívá i v diferenciální geometrii, viz [317].

Věta 6.1.25 Jung

Nechť je dána množina $X \subseteq \mathbb{R}^n$, jejíž průměr[38] je nejvýše 2. Potom existuje koule o poloměru nejvýše $\sqrt{\frac{2n}{n+1}}$ obsahující množinu X. Zejména je-li $n = 2$, pak ten poloměr je nejvýše $2/\sqrt{3}$.

Uvedené číslo v předchozím tvrzení má dokonce i geometrický význam: udává poloměr koule opsané n-rozměrnému pravidelnému simplexu s hranou délky 2. Navíc se dá ukázat, že pokud

[21] Již víme, že součet konvexních množin je konvexní množina, proto $\operatorname{conv} X + \operatorname{conv} Y$ je konvexní, a tudíž nutně $\operatorname{conv}(X+Y) \subseteq \operatorname{conv} X + \operatorname{conv} Y$. Nyní uvažme libovolnou dvojici $x+y \in \operatorname{conv} X + \operatorname{conv} Y$. Pak $x = \sum_i \alpha_i a_i$ a $x = \sum_j \beta_j b_j$ pro nějaká $a_1, \cdots \in X$, $b_1, \cdots \in Y$ a čísla $a_1, \ldots, b_1, \cdots \geq 0$ splňující $\sum_i a_i = 1 = \sum_j b_j$. Jelikož platí

$$x + y = \sum_j \beta_j (x + b_j)$$

a současně

$$x + b_j = \sum_i \alpha_i (a_i + b_j) \in \operatorname{conv}(X+Y),$$

musí být $x+y \in \operatorname{conv}(\operatorname{conv}(X+Y)) = \operatorname{conv}(X+Y)$, tj. $\operatorname{conv}(X+Y) \supseteq \operatorname{conv} X + \operatorname{conv} Y$, čímž je důkaz hotov.

[22] Samotná věta říká toto: Kompaktní konvexní podmnožina Hausdorffova lokálně konvexního topologického vektorového prostoru je rovna uzavřenému konvexnímu obalu svých extrémních bodů.

ní známo jako Tichonovova věta a pro její platnost je vyžadován axiomu výběru. Její speciální případ poprvé dokázal v roce 1930 ruský matematik Andrej Nikolajevič Tichonov. První úplný důkaz obecného tvrzení pochází z roku 1937 a jeho autorem je český matematik Eduard Čech, který je velmi úzce spojen i s Masarykovou univerzitou.

se množina X nevejde do žádné menší koule než s uvedeným poloměrem, pak uzávěr této množiny obsahuje všechny vrcholy takového simplexu.

Tuto část uzavřeme ještě jedním zajímavým výsledkem spojeným s diskrétní geometrií a konvexními množinami. Uvažme kruh v $\mathbb{R}^2$. Na kolik částí jej musíme rozdělit, aby jednotlivé části měly menší průměr než původní množina? Pochopitelně na 3. V případě trojrozměrné koule by to pak byly alespoň 4 části. Polskému matematikovi Karolu Borsukovi se v roce 1933 podařilo ukázat, že pro libovolnou kouli v $\mathbb{R}^n$ to musí být nejméně $n+1$ částí a zároveň ukázal, že pouhých n částí obecně nestačí. To jej přivedlo k následující otázce, která je obvykle označována jako Borsukova domněnka:

> Je pravda, že každá ohraničená množina v $\mathbb{R}^n$ může být rozdělena na $n+1$ částí s menším průměrem?

Tento problém je znám jako *Borsukova domněnka*, ačkoli takové označení je v tomto případě nesprávné. Borsuk nic netvrdil o pravdivosti tohoto výroku a formuloval jej (záměrně?) jako otázku, viz [76]. Nicméně matematici po dlouhá desetiletí věřili, že odpověď je *ano*, a tak se začalo hovořit o domněnce. V průběhu času se postupně objevovaly články potvrzující správnost této domněnky. Avšak zlom nastal v roce 1993. Jeff Kahn a Gil Kalai ukázali, že pro $n = 1325$ a libovolné $n > 2014$ je $n+1$ částí nedostačující! O deset let později se podařilo ukázat, že ani pro $n \geq 298$ to není možné, a o dalších deset let se tato „hranice nemožného dělení“ posunula až na $n \geq 64$. Matematici se také zajímají o chování funkce $\alpha(n)$ vyjadřující největší potřebný počet částí mezi všemi ohraničenými množiny v $\mathbb{R}^n$. Kahn a Kalai se pro dostatečně velká n dopracovali k dolnímu odhadu $\alpha(n) \geq 1{,}2^{\sqrt{n}}$, zatímco Oded Schramm získal horní odhad $\alpha(n) \leq \left(\sqrt{3/2}+\varepsilon\right)^n$ pro libovolné ε a dostatečně velké n. Přesná řádová velikost zatím zůstává neznámá... Pro další podrobnosti viz např. [305, 426].

Další zajímavosti z kombinatorické geometrie už ponechme jen na zvídavých zájemkyních a zájemcích, které tato část zaujala a chtěli by si udělat hlubší přehled o dosavadních poznatcích v této oblasti. My se nyní společně přesuneme k jinému důležitému tématu z teorie konvexních množin. Ze základního kurzu matematické analýzy si jistě všichni pamatujeme, že pro libovolnou množinu $X \subseteq \mathbb{R}^n$ nazveme bod $x \in X$ *vnitřním*, jestliže existuje okolí $\mathcal{O}(x)$ takové, že $\mathcal{O}(x) \subseteq X$. To při euklidovské metrice znamená, že v případě $n = 2$ musí existovat kruh $K(x,\varepsilon$ se středem v x a poloměrem $\varepsilon > 0$ takový, že $K(x,\varepsilon) \subseteq X$. Podobně pro $n = 3$ musí existovat koule $B(x,\varepsilon)$ se středem v bodě x a poloměrem $\varepsilon > 0$ taková, že $B(x,\varepsilon) \subseteq X$. Avšak toto pojetí „vnitřnosti“ může být (a pro nás i bude) v některých situacích příliš restriktivní. Uvažme např. úsečku $\mathbb{R}^2$ či čtverec v $\mathbb{R}^3$. Jaké body jsou vnitřní? Žádné – všechny body leží jen na hranici těchto množin. Intuitivně bychom ovšem asi řekli, že „vnitřními body“ jsou všechny kromě koncových bodů úsečky a obvodu trojúhelníku. Toto více intuitivní pojetí nás přivádí k následujícímu pojmu, který nás bude provázet v dalších částech této kapitoly.

DEFINICE 6.1.26

Nechť $X \subseteq \mathbb{R}^n$. Bod $x^* \in X$ se nazývá *relativně vnitřním* bodem množiny X, jestliže existuje okolí $\mathcal{O}(x^*)$ bodu x^* takové, že

$$\mathcal{O}(x^*) \cap \operatorname{aff} X \subseteq X.$$

Množinu všech relativně vnitřních bodů množiny X nazýváme *relativním vnitřkem* množiny X a značíme $\operatorname{ri} X$. Množina $\operatorname{r}\partial X := \overline{X} \setminus \operatorname{ri} X$ se nazývá *relativní hranice* X.

Je snad zřejmé, že pokud $\operatorname{aff} X = \mathbb{R}^n$, pak se definice relativně vnitřního bodu neliší od definice vnitřního bodu, tj. $\operatorname{ri} X = \operatorname{int} X$. Toto jistě nastane v případě, kdy $\operatorname{int} X \neq \emptyset$, neboť pak X obsahuje nějakou n-rozměrnou kouli $\mathcal{O}(x)$, a tedy nutně $\dim X = n$. Jaký je vlastně vztah mezi $\operatorname{ri} X$ a $\operatorname{int} X$? Jsou možné jen dvě situace: (a) buď $\operatorname{int} X \neq \emptyset$, takže $\operatorname{ri} X = \operatorname{int} X$; (b) nebo

[23] *Diskrétní geometrie* (angl. *discrete geometry*) se věnuje studiu vlastností geometrických objektů generovaných pomocí konečného počtu základních geometrických prvků jako jsou body, přímky, roviny, nadroviny, polyedry, polytopy, kružnice, koule atd. *Kombinatorická geometrie* (angl. *combinatorial geometry*) propojuje diskrétní geometrii s kombinatorikou, když se zaměřuje na kombinatorické vlastnosti geometrických objektů jako např. průniky nebo uspořádání přímek (obecněji nadrovin) v euklidovském či jiném vektorovém prostoru, pokrývání větších objektů menšími, obalování, skládání, obkládání, dělení, rozkládání, obarvování, symetričnost či problémy s osvětlením. Kombinatorická geometrie zahrnuje aspekty topologie, teorie grafů, teorie čísel a dalších disciplín. Ačkoli problémy kombinatorické geometrie se zabývali již Kepler či Euler, výrazné pozornosti se jí dostává až od poloviny 20. století. Samotné slovní spojení „kombinatorická geometrie“ bylo (asi) poprvé použito švýcarským matematikem Hugem Hadwigerem v roce 1954 při přípravě článku [247], viz také monografii viz [248].

[24] Vcelku nečekaně se jedná o oblast geometrie, která se zaměřuje na konvexní množiny, tj. to, co právě teď děláme i my. Viz např. [240, 241].

[25]

Johann Karl August Radon (16. prosince 1887 – 25. května 1956) byl rakouský matematik, který se narodil v Děčíně. Povinnou pětiletou školní docházku absolvoval v Děčíně za čtyři roky. Na gymnázium v Litoměřicích, které navštěvoval v letech 1897–1905, projevil své všeobecné nadání v matematice, fyzice, chemii atd. V roce 1905 se rodina přestěhovala do Vídně a Radon začal studovat na tamní univerzitě matematiku a fyziku. Sice také navštěvoval přednášky z hudební te-

orie a pedagogiky, ale tento „souboj" jednoznačně vyhrála matematika. V roce 1910 zde získal doktorát za práci věnovanou variačnímu počtu. Mezi jeho učitele patřil mimo jiné Hans Hahn, který je akademickým otcem (Němci pro to mají ještě kouzelnější označení *Doktorvater*) Kurta Gödela. Podzimní semestr akademického roku 1910–1911 strávil v Göttingenu, kde navštěvoval i Hilbertovy přednášky.

Dne 2. prosince 1910 odeslal dopis, v němž se ucházel o místo asistence na Německá vysoké škole technické v Brně (ta sídlila na ve dvou budovách na Komenského náměstí a na Joštově ulici, které dnes využívá Masarykova univerzita; škola byla 18. října 1945 zrušena Benešovými dekrety). Uspěl a 18. února byl jmenován asistentem na katedře matematiky pro období 1. dubna 1911 – 30. 8. 1912... Kariéra to byla velmi zajímavá a pestrá (mj. obnovil časopis *Monatshefte für Mathematik* a v roce 1954 byl dokonce zvolen rektorem univerzity ve Vídni), která však mířila jiným směrem, než se zde chceme vydat my, viz také [194]. Vlastně zde zmiňujeme jeho jméno v souvislosti s výsledkem, který rozhodně nebyl tím nejvýznamnějším. Radon je totiž obvykle spojován zejména s tzv. *Radonovu–Nikodýmovou větou* z teorie míry a především s *Radonovou* (integrální) *transformací*, jejíž inverzní podoba je základem *výpočetní tomografie* (angl. *computed tomography* neboli CT) k rekonstrukci obrazu na základě konečného počtu projekcí.

$\operatorname{int} X = \emptyset$, a pak $\operatorname{ri} X \subseteq X$, přičemž to může být i prázdná množina nebo celé X. Odtud také bezprostředně plyne vztah mezi hranicí a relativní hranicí $\operatorname{r}\partial X \subseteq \partial X$ s tím, že buď nastane rovnost nebo $\partial X = \overline{X}$.

Je-li množina X afinní, pak jistě $\operatorname{ri} X = X$, a tudíž také $\operatorname{ri}(\operatorname{aff} X) = \operatorname{aff} X$ a $\operatorname{r}\partial(\operatorname{aff} X) = \emptyset$. Ještě lépe, pro každou množinu $X \subseteq \mathbb{R}^n$ platí[39]

$$\emptyset \subseteq \operatorname{int} X \subseteq \operatorname{ri} X \subseteq X \subseteq \overline{X} \subseteq \operatorname{aff} X.$$

Avšak jednotlivé inkluze mohou být ostré. U prvních dvou to již velmi dobře víme z první části. Pro ostatní uvažme např. množinu $X = \{x \in \mathbb{R} \mid 0 \le x \le 1 \text{ je racionální}\}$. Pak

$$\underbrace{\emptyset}_{\operatorname{ri} X} \subset X \subset \underbrace{[0,1]}_{\overline{X}} \subset \underbrace{\mathbb{R}}_{\operatorname{aff} X},$$

neboť $\operatorname{aff} X = \mathbb{R}$ a z vlastností racionálních čísel plyne $\operatorname{int} X = \operatorname{ri} X = \emptyset$, zatímco $\overline{X} = [0,1] = \operatorname{r}\partial X$, tj. $\operatorname{ri} X$ je „neporovnatelně menší" než X, množina $\overline{X}$ je „neporovnatelně větší" než X a $\operatorname{r}\partial X$ obsahuje celou X.

Příklad 6.1.27

Uvažme v $\mathbb{R}^3$ dvourozměrný čtverec

$$X = \{[x,y,z] \in \mathbb{R}^3 \mid -1 \le x \le 1,\ -1 \le y \le 1,\ z = 1\},$$

viz Obrázek 6.33. Afinní obal této množiny pak je rovina

$$\operatorname{aff} X = \{[x,y,z] \mid z = 1\} = [0,0,1] + \{[x,y,z] \mid z = 0\}.$$

Vnitřek množiny X je prázdný[40], avšak[41]

$$\operatorname{ri} X = \{[x,y,z] \in \mathbb{R}^3 \mid -1 < x < 1,\ -1 < y < 1,\ z = 1\}.$$

Relativní hranice pak je obvod tohoto čtverce, tj.

$$\operatorname{r}\partial X = \{[x,y,z] \in \mathbb{R}^3 \mid \max\{|x|,|y|\} = 1,\ z = 1\}.$$

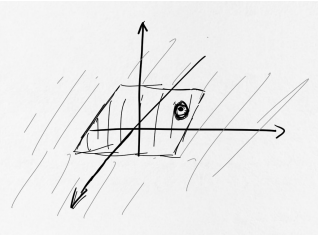
Jednou ze základních vlastností vnitřku je, že zachovává množinovou inkluzi, tj. pro $X \subseteq Y$ bude $\operatorname{int} X \subseteq \operatorname{int} Y$. Avšak pro relativní vnitřky toto již platit nemusí, např. je-li Y krychle v $\mathbb{R}^3$ a X jedna její stěna, pak množiny $\operatorname{ri} X$ a $\operatorname{ri} Y = \operatorname{int} Y$ jsou neprázdné, ale disjunktní. Další vlastnosti relativního vnitřku jsou uvedeny v následující větách[42]. Začneme tzv. *úsečkovým principem*, který kopíruje základní vlastnost konvexních množin.

Věta 6.1.28

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná[43] konvexní množina. Pokud $x \in \operatorname{ri} X$ a $y \in \overline{X}$, pak pro každé $\lambda \in (0,1]$ platí

$$(1-\lambda)\,y + \lambda x \in \operatorname{ri} X,$$

tj. celá úsečka spojující x a y leží v $\operatorname{ri} X$ s možnou výjimkou bodu y.



Obrázek 6.33

[27] Kdyby některá z množin I a J byla jednoprvková, nejednalo by se o konvexní kombinaci dle definice, což by ale nikterak neovlivnilo závěr důkazu.

Důkaz. Nechť $y \in X$. Protože $x \in \operatorname{ri} X$, existuje otevřená koule (tj. okolí) $B(x,\varepsilon) := \{z \in \mathbb{R}^n \mid \|z - x\| < \varepsilon\}$ taková, že $B \cap \operatorname{aff} X \subseteq X$. Pro $\lambda \in (0,1]$ označme $x_\lambda := \lambda x + (1-\lambda)\,y \in X$ a kouli $B_\lambda(x_\lambda,\varepsilon) := \{z \in \mathbb{R}^n \mid \|z - x_\lambda\| < \lambda\varepsilon\}$, viz Obrázky 6.35.a–6.35.b na konci důkazu. Lze vidět, že každé $\widetilde{x} \in B_\lambda \cap \operatorname{aff} X$ je konvexní kombinací bodu $y \in X$ a nějakého bodu $\widehat{x} \in B \cap \operatorname{aff} X \subseteq X$. Ovšem X je konvexní, takže $\widetilde{x} \in X$, a tedy nutně $B_\lambda \cap \operatorname{aff} X \subseteq X$, tj. $x_\lambda \in \operatorname{ri} X$.

Nechť nyní $y \notin X$, tj. $y \in \overline{X} \setminus X$. K tomu, abychom ukázali, že pro všechna $\lambda \in (0,1]$ platí $x_\lambda := \lambda x + (1-\lambda)\,y \in \operatorname{ri} X$, uvažme posloupnost $\{x_k\}_{k=1}^\infty \subseteq X$ takovou, že $x_k \to y$. Označme $x_{k,\lambda} := \lambda x + (1-\lambda)\,x_k \in X$. Podobně jako v předchozí části plyne odtud, že

$$\{z \in \mathbb{R}^n \mid \|z - x_{k,\lambda}\| < \lambda\varepsilon\} \cap \operatorname{aff} X \subseteq X$$

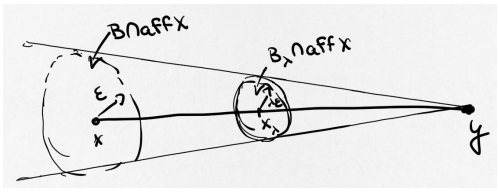
pro všechna k, kde ε je opět takové, že otevřená koule $B(x,\varepsilon) := \{z \in \mathbb{R}^n \mid \|z-x\| < \varepsilon\}$ splňuje $B \cap \operatorname{aff} X \subseteq X$. Protože $x_{k,\lambda} \to x_\lambda$, pro k dostatečně velká (taková, aby $\|x_{k,\lambda} - x_\lambda\| < \lambda\varepsilon/2$) platí

$$\{z \in \mathbb{R}^n \mid \|z - x_\lambda\| < \lambda\varepsilon/2\} \subseteq \{z \in \mathbb{R}^n \mid \|z - x_{k,\lambda}\| < \lambda\varepsilon\},$$

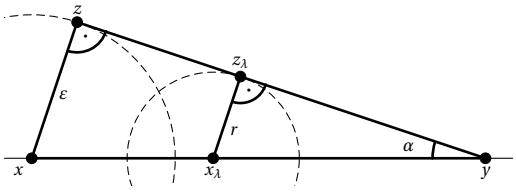
tj. koule o poloměru $\lambda\varepsilon/2$ se středem v x_λ je obsažena ve všech koulích o poloměru $\lambda\varepsilon$ a středy $x_{k,\lambda}$ pro dostatečně velká k. Odtud plyne, že

$$\{z \in \mathbb{R}^n \mid \|z - x_\lambda\| < \lambda\varepsilon/2\} \cap \operatorname{aff} X \subseteq \{z \in \mathbb{R}^n \mid \|z - x_{k,\lambda}\| < \lambda\varepsilon\} \cap \operatorname{aff} X \subseteq X,$$

což dokazuje, že $x_\lambda \in \operatorname{ri} X$.



Obrázek 6.35.a



Obrázek 6.35.b: Množiny $B \cap \operatorname{aff} X$ a $B_\lambda \cap \operatorname{aff} X$ v případě $\dim X = 2$. Z podobnosti pravoúhlých trojúhelníků xyz a $x_\lambda y z_\lambda$ plynou rovnosti $\operatorname{tg}\alpha = \varepsilon/(x-y) = r/|\lambda x + (1-\lambda)y - y|$, což znamená, že $r = \lambda\varepsilon$.

■

Nejvýznamnějším důsledkem předchozího tvrzení je zaručení konvexnosti $\operatorname{ri} X$. Také ukážeme neprázdnost $\operatorname{ri} X$, což ale neplatí pro úplně každou množinu, jak jsme už ukázali dříve.

Věta 6.1.29

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná konvexní množina. Množina $\operatorname{ri} X$ je neprázdná, konvexní a $\operatorname{aff} X = \operatorname{aff}(\operatorname{ri} X)$. Je-li navíc $\dim X = m > 0$, pak existují $z_0, \dots, z_m \in \operatorname{ri} X$ taková, že $z_1 - z_0, \dots, z_m - z_0$ tvoří bázi $\operatorname{Lin} X$.

Důkaz. Bez újmy na obecnosti můžeme předpokládat, že $0 \in X$. V opačném případě totiž stačí aplikovat celý postup na množinu

$$Y := \{y \in \mathbb{R}^n \mid y = x - x_0 \text{ pro všechna } x \in X\}, \quad \text{tj.} \quad Y = X - x_0,$$

kde $x_0 \in X$ je libovolné. Pak totiž jistě $0 \in Y$.

Jelikož $0 \in X$, platí $\operatorname{aff} X = \operatorname{Lin} X$, tj. $\operatorname{aff} X$ je vektorový podprostor. Nechť $\dim \operatorname{Lin} X = m$. Je-li $m = 0$, pak $X = \operatorname{aff} X = \{0\}$ jsou jednobodové množiny a tento bod je současně (jediným) relativně vnitřním bodem. Nechť $m > 0$. Potom můžeme najít takovou m-tici lineárně nezávislých bodů (vektorů) $x_1, \dots, x_m \in X$, že[44] $\operatorname{span}\{x_1, \dots, x_m\} = \operatorname{aff} X = \operatorname{Lin} X$. Tedy $x_1, \dots, x_m$ tvoří bázi $\operatorname{aff} X$.

Označme nyní množiny

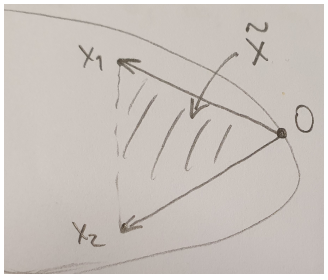
$$\widetilde{\Lambda} := \{\lambda \in \mathbb{R}^m \mid \lambda_1, \dots, \lambda_m > 0 \ \&\ \sum_{i=1}^m \lambda_i < 1\} \quad \text{a} \quad \widetilde{X} := \{\widetilde{x} \in \mathbb{R}^n \mid \widetilde{x} = \sum_{i=1}^m \lambda_i x_i \ \&\ \lambda \in \widetilde{\Lambda}\} \neq \emptyset,$$

tj. je-li $\widetilde{x} \in \widetilde{X}$, pak $\widetilde{x} = A\tilde{\lambda}$, kde $A = (x_1, \dots, x_m) \in \mathbb{R}^{n \times m}$ a $\tilde{\lambda} \in \widetilde{\Lambda}$, viz Obrázek 6.34. Je zřejmé, že $\widetilde{X} \subseteq \operatorname{conv} X = X$. Ukážeme, že množina $\widetilde{X}$ je relativně otevřená, tj. $\operatorname{ri} \widetilde{X} = \widetilde{X}$, tj. pro každé $\widetilde{x} \in \widetilde{X}$ existuje okolí (otevřená koule) $\mathcal{O}(\widetilde{x})$ splňující $\mathcal{O}(\widetilde{x}) \cap \operatorname{aff} X \subseteq X$. Zafixujme $\widetilde{x} \in \widetilde{X}$ a nechť $x \in \operatorname{aff} X \setminus \{\widetilde{x}\}$ je libovolné. Potom

$$\widetilde{x} = A\tilde{\lambda} \quad \& \quad x = A\lambda,$$

kde $\tilde{\lambda} \in \widetilde{\Lambda}$ a $\lambda \in \mathbb{R}^m$, neboť sloupce matice A tvoří bázi $\operatorname{aff} X$. Protože tyto sloupce jsou lineárně nezávislé, je matice $A^\top A$ symetrická a pozitivně definitní. Proto[45] existuje číslo $\gamma > 0$, které nezávisí na x a $\widetilde{x}$ a platí

$$\|x - \widetilde{x}\|^2 = (\lambda - \tilde{\lambda})^\top A^\top A (\lambda - \tilde{\lambda}) \geq \gamma \|\lambda - \tilde{\lambda}\|^2. \tag{6.1.9}$$



Obrázek 6.34: Ukázka množiny $\widetilde{X}$ z důkazu Věty 6.1.29 pro množinu X obsahující počátek. Vektory x_1 a x_2 jsou lineárně nezávislé, patří do X a generují $\operatorname{aff} X$.

Eduard Helly (1. června 1884 – 28. listopadu 1943) byl rakouský matematik, který v roce 1938 uprchl i s rodinou před nacistickým režimem do USA. Jeho jméno je spojeno především s Hellyho větou, Hellyho principem výběru a Hellyho–Brayovou větou. Výrazně ovlivnil především rozvoj funkcionální analýzy, ačkoli jeho jméno zůstává v tomto ohledu mnohdy opomíjeno. Na druhou stranu, jeho jméno potkáme snad v každé monografii věnované konvexním množinám. Pro více podrobností viz [399].

[28]

[29] Ten průnik může být i větší. Podobně totiž platí $\inf\{b_i \mid i \in I\} \in \bigcap_{i\in I}[a_i, b_i]$, takže $\bigcap_{i\in I}[a_i, b_i]$ obsahuje celý interval $\left[\sup_{i\in I} a_i, \inf_{i\in I} b_i\right]$. Pro úplnost poznamenejme, že v případě konečné indexové množiny I můžeme supremum a infimum nahradit maximem a minimem.

[30] Angl. *finite intersection property*.

[31] Paul Kirchberger byl německý matematik. Doktorát získal v roce 1902 na univerzitě v Göttingenu pod Hilbertovým vedením. Jeho jméno je spojeno především s větou z diskrétní geometrie, kterou odvodil ve své disertační práci jako pomocný výsledek při analýze Čebyševovy aproximace.

Z tohoto odhadu plyne, že pokud zvolené $x \in \operatorname{aff} X$ leží uvnitř dostatečně malé koule se středem v $\widetilde{x}$, tj. $x \in \mathcal{O}(\widetilde{x}) \cap \operatorname{aff} X$, pak také $\lambda \in \widetilde{\Lambda}$ (neboť $\widetilde{\Lambda}$ je otevřená množina), z čehož plyne $x \in \widetilde{X}$ (podle definice $\widetilde{X}$). Tudíž

$$\mathcal{O}(\widetilde{x}) \cap \operatorname{aff} X \subseteq \widetilde{X} \subseteq X,$$

tj. $\widetilde{x} \in \operatorname{ri} X$. Jelikož $\widetilde{x} \in \widetilde{X}$ byl zvolen libovolně, platí $\widetilde{X} \subseteq \operatorname{ri} X$, a tedy $\operatorname{ri} X \neq \emptyset$.

Konvexnost $\operatorname{ri} X$ plyne z Věty 6.1.28. Pro důkaz poslední části uvažme body

$$z_0 := \lambda \sum_{i=1}^{m} x_i \quad \text{a} \quad z_i := z_0 + \lambda x_i,$$

kde $i \in \{1,\dots,m\}$ a $\lambda \in (0,1)$ je takové číslo, že $(m+1)\lambda < 1$. Pak $z_0,\dots,z_m \in \widetilde{X}$ (z definice $\widetilde{X}$)[46], a jelikož $\widetilde{X} \subseteq \operatorname{ri} X$ dle předchozí části, máme dokonce $z_0,\dots,z_m \in \operatorname{ri} X$. Navíc platí $z_i - z_0 = \lambda x_i$ pro každé $i \in \{1,\dots,m\}$. Poněvadž $x_1,\dots,x_m$ tvoří bázi $\operatorname{aff} X = \operatorname{Lin} X$ a $\lambda \neq 0$, platí totéž i o vektorech $z_1 - z_0,\dots,z_m - z_0$. ■

Dalším důsledkem Věty 6.1.28 je následující tvrzení „o prodloužitelnosti“, které ukazuje, že i pro body z $\operatorname{ri} X \neq \operatorname{int} X$ platí jedna ze základních vlastností klasických vnitřních bodů charakterizující jejich „vzdálenost od hranice“.

Věta 6.1.30

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná konvexní množina. Platí $x \in \operatorname{ri} X$ právě tehdy, když pro každé $\widetilde{x} \in X$ existuje $\gamma > 1$ takové, že

$$x + (\gamma - 1)(x - \widetilde{x}) = \gamma x + (1-\gamma)\widetilde{x} \in X,$$

tj. každou úsečku v X s krajním bodem $x \in \operatorname{ri} X$ lze prodloužit za x, aniž by „opustila“ množinu[47] X.

Důkaz. Jestliže $x \in \operatorname{ri} X$, pak tvrzení plyne přímo z definice relativně vnitřního bodu v kombinaci s Větou 6.1.28[48]. Naopak, nechť $x \in X$ má vlastnost uvedenou v tvrzení. Ukážeme, že $x \in \operatorname{ri} X$. Podle Věty 6.1.29 existuje $\widetilde{x} \in \operatorname{ri} X$. Předpokládejme, že $x \neq \widetilde{x}$, protože v opačném případě je důkaz hotov. Jelikož $\widetilde{x} \in X$, existuje podle předpokladů $\gamma > 1$ takové, že $y := x + (\gamma-1)(x-\widetilde{x}) \in X$. Tedy platí $y = \gamma x + (1-\gamma)\widetilde{x}$ neboli $\widetilde{x} = (1-\lambda)y + \lambda\widetilde{x}$, kde $\lambda = (1-1/\gamma) \in (0,1)$. Pak ale z Věty 6.1.28 ihned vyplývá, že $x \in \operatorname{ri} X$. ■

Zbývají ještě některé vlastnosti relativního vnitřku v kombinaci s některými dalšími množinovými operacemi. Tady uvedeme jen něco málo pro afinní obaly, uzávěry a porovnání dvou množin se stejnými relativními vnitřky. Další zajímavé výsledky už ponecháme čtenáři ve formě cvičení, viz např. Cvičení 6.1.34.

Věta 6.1.31

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná konvexní množina. Potom platí $\overline{X} = \overline{\operatorname{ri} X}$, $\operatorname{ri} X = \operatorname{ri} \overline{X}$, množina $\overline{X}$ je konvexní, $\operatorname{aff} X = \operatorname{aff} \overline{X}$. Je-li navíc Y jiná konvexní množina, pak následující tvrzení jsou ekvivalentní

(a) $\operatorname{ri} X = \operatorname{ri} Y$,

(b) $\overline{X} = \overline{Y}$,

(c) $\operatorname{ri} X \subseteq Y \subseteq \overline{X}$.

Důkaz. Protože $\operatorname{ri} X \subseteq X$, platí $\overline{\operatorname{ri} X} \subseteq \overline{X}$. Naopak, nechť $\widetilde{x} \in \overline{X}$. Ukážeme, že také $\widetilde{x} \in \overline{\operatorname{ri} X}$. Nechť $x \in \operatorname{ri} X$ je libovolný bod (ten existuje podle Věty 6.1.29) a předpokládejme, že $x \neq \widetilde{x}$ (v opačném případě jsme hotovi). Podle Věty 6.1.28 máme $\lambda x + (1-\lambda)\widetilde{x} \in \operatorname{ri} X$ pro všechna $\lambda \in (0,1]$. Pak ale $\widetilde{x}$ můžeme získat jako limitu posloupnosti $\left\{\frac{1}{k}x + \left(1 - \frac{1}{k}\right)\widetilde{x}\right\}_{k\in\mathbb{N}}$, jejíž prvky leží v $\operatorname{ri} X$. Proto z definice uzávěru množiny vyplývá $\widetilde{x} \in \overline{\operatorname{ri} X}$.

Inkluze $\operatorname{ri} X \subseteq \operatorname{ri} \overline{X}$ plyne z definice relativně vnitřních bodů a ze skutečnosti $\operatorname{aff} X = \operatorname{aff} \overline{X}$ (viz Cvičení 6.1.32). Pro důkaz opačné inkluze uvažme $x \in \operatorname{ri} \overline{X}$. Ukážeme, že $x \in \operatorname{ri} X$. Z Věty 6.1.29

[32] Neděsme se, o tomto se dozvíme více v Podkapitole 6.3. Nyní si stačí pod *ostrou oddělitelností* představit nenulovou vzdálenost obou množin.

[33] Toto je „hellyovská“ modifikace klasického problému hlídače, ve kterém jde o minimální počet hlídačů potřebných pro zabezpečení celé galerie. Tento problém byl poprvé zformulován (asi) Chvátalem a Kleem v roce 1973 a vyřešen prvním jmenovaným o dva roky později v [290].

[34] Angl. *star-shaped set.*

[35] Tuto podmínku nepochybně splňuje každá konvexní množina.

[36] Mark Aleksandrovich Krasnoselsky (27. dubna 1920 – 13. února 1997) byl sovětský a ruský matematik, který je znám především díky své práci v nelineární funkcionální analýze. Je (spolu-)autorem či cca 300 odborných článků a 14 monografií. Mimo jiné získal Humboldtovu cenu, kterou uděluje Nadace Alexandra von Humboldta mezinárodně uznávaným výzkumným pracovníkům na vrcholu jejich kariéry.

[37] Heinrich Wilhelm Ewald Jung (4. května 1876 – 12. března 1953) byl německý matematik, který se věnoval především geometrii a algebraické geometrii.

[38] *Průměrem* množiny X pochopitelně rozumíme číslo $\operatorname{diam}(X) := \sup\{\|x-y\| \mid x,y \in X\}$.

[39] Poslední část je důsledkem toho, že $\operatorname{aff} X$ je uzavřená nadmnožina X

[40] „Do čtverce kouli nedostaneme.“

[41] „Průnik koule s rovinou je kruh, a ten už bychom do čtverce dostat mohli.“

[42] Nezapomínejme, že tímto získáme zároveň tytéž výsledky i pro (klasické) vnitřní body.

[43] Kdyby $X = \emptyset$, bylo by toto i další tvrzení triviální.

[44] Viz Poznámku 6.1.11 a nezapomeňme, že tentokrát $m = \dim \operatorname{Lin} X$.

[45] Pro euklidovskou normu $\|x\| := x^\top x = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ a libovolnou symetrickou a pozitivně definitní matici $A \in \mathbb{R}^{n\times n}$ lze ukázat, že

existují čísla $\gamma, \bar{\gamma} > 0$, pro která platí

$$\gamma \|x\|^2 \le x^\top A x \le \bar{\gamma} \|x\|^2$$

pro každé $x \in \mathbb{R}^n$.

[46] Vezmeme-li $\lambda_1 = \lambda_2 = \cdots = \lambda_m = \lambda_{m+1} = \lambda$, tak $z_0 \in \widetilde{X}$, protože $\sum_{i=1}^m \lambda_i = m\lambda < 1$, zatímco $z_1, \dots, z_m \in \widetilde{X}$, protože $\sum_{i=1}^{m+1} \lambda_i = (m+1)\lambda < 1$.

[47] Toto je klasická formulace tvrzení o prodloužitelnosti v podobě, ve které ji najdeme již v [432, Theorem 6.4]. Avšak není těžké si rozmyslet, že za daných předpokladů nastane dokonce $\gamma x + (1-\gamma)\widetilde{x} \in \operatorname{ri} X$ pro vhodné $\gamma > 1$. Nicméně důvodem pro slabší podobu uvedené věty může být např. to, že ji lze takto využít k alternativní definici relativního vnitřku pro konvexní množiny, tj. jako množiny bodů $x \in X$ takových, že pro všechna $y \in X$ existuje $\gamma > 1$ takové, že $\gamma x + (1-\gamma)\widetilde{x} \in X$.

[48] Podle definice musí pro $x \in \operatorname{ri} X$ existovat dokonce takové okolí $\mathcal{O}(x)$, že $\mathcal{O}(x) \cap \operatorname{aff} X \subseteq \operatorname{ri} X$, a tedy prodloužením úsečky spojující x a $\widetilde{x}$, která pochopitelně leží v $\operatorname{aff} X$, se dostaneme opět do nějakého bodu z $\operatorname{ri} X$. Tudíž podle Věty 6.1.28 musí celá tato prodloužená úsečka s možnou výjimkou bodu $\widetilde{x}$ ležet v $\operatorname{ri} X$.

víme, že určitě existuje $\widetilde{x} \in \operatorname{ri} X$. Nechť $\widetilde{x} \neq x$ (jinak je důkaz hotov). Zvolme $\gamma > 1$ dostatečně blízko 1 tak, že bod $y := x + (\gamma - 1)(x - \widetilde{x}) \in \overline{X}$, což lze vzhledem k Věty 6.1.30. Pak máme $x = (1-\lambda)\, y + \lambda\, \widetilde{x}$, kde $\lambda = (1 - 1/\gamma) \in (0,1)$, a tedy podle Věty 6.1.28 dostáváme $x \in \operatorname{ri} X$, tj. $\operatorname{ri} \overline{X} \subseteq \operatorname{ri} X$.

Pokud $\operatorname{ri} X = \operatorname{ri} Y$, pak (dvojnásobným) aplikováním první části dostaneme $\overline{X} = \overline{\operatorname{ri} X} = \overline{\operatorname{ri} Y} = \overline{Y}$. Podobně pokud $\overline{X} = \overline{Y}$, pak druhá část tvrzení implikuje $\operatorname{ri} X = \operatorname{ri} \overline{X} = \operatorname{ri} \overline{Y} = \operatorname{ri} Y$. Tedy část (a) je ekvivalentní s (b). Konečně části (a) a (b) a vztah $\operatorname{ri} Y \subseteq Y \subseteq \overline{Y}$ implikuje (c), neboť $\operatorname{ri} Y = \operatorname{ri} X \subseteq Y \subseteq \overline{Y} = \overline{X}$. Naopak, „aplikováním uzávěru“ a s využitím první části dostaneme $\overline{X} = \overline{\operatorname{ri} X} \subseteq \overline{Y} \subseteq \overline{\overline{X}}$, tudíž $\overline{Y} = \overline{X}$, tj. platí (b). ■

Poznámka 6.1.32.

(i) Z Věty 6.1.28 snadno obdržíme následující důsledek:

> Je-li $X \subseteq \mathbb{R}^n$ neprázdná konvexní množina, pak každá konvexní kombinace $x = \sum_{i=1}^m \lambda_i x_i$ pro $\lambda_1, \dots, \lambda_m \in [0,1]$ splňující $\sum_{i=1}^m \lambda_i = 1$ a body $x_1, \dots, x_m \in \overline{X}$, přičemž alespoň pro jeden index $i \in \{1, \dots, m\}$ máme $x_i \in \operatorname{ri} X$ a $\lambda_i \neq 0$, je také prvkem $\operatorname{ri} X$.

(ii) Pro pravdivost všech předchozích tvrzení je konvexnost množiny X zcela klíčová. Zejména pro poslední tvrzení uvažme např. $X = [0,1] \cup \{2\}$, v kterémžto případě totiž bude

$$\operatorname{ri} X = (0,1) = \operatorname{int} X \quad \text{a} \quad \overline{\operatorname{ri} X} = [0,1] = \overline{\operatorname{int} X} \neq \overline{X}.$$

Podobně pro $X = (0,1) \cup (1,2)$ dostaneme

$$\operatorname{ri} X = X = \operatorname{int} X, \quad \overline{X} = [0,2] \quad \text{a} \quad \operatorname{ri} \overline{X} = (0,2) \neq \operatorname{ri} X.$$

(iii) Rovnost $\overline{X} = \overline{\operatorname{ri} X}$ ve Větě 6.1.31 mimo jiné znamená, že pro každý bod $x \in \overline{X}$ existuje posloupnost $\{x_k\}_{k=1}^\infty \subseteq \operatorname{ri} X$ taková, že $x_k \to x$ pro $k \to \infty$, tj. pro konvexní množiny je samotný relativní vnitřek dostatečně velký na to, abychom jenom s jeho pomocí „zkonstruovali“ celý $\overline{X}$, a tedy místo posloupnosti v X z definice uzávěru stačí uvážit posloupnost pouze s body z relativního vnitřku.

(iv) Z Věty 6.1.28 dokonce dostáváme i jakýsi návod, jak získat uzávěr konvexní množiny X: stačí uvážit pouze „radiální“ uzávěr, tj. vezmeme nějaký bod $x \in \operatorname{ri} X$ (ten jistě existuje) a z přímek v $\operatorname{aff} X$ procházejících bodem x uděláme polopřímky ℓ vycházející z x a následně zjistíme jejich průnik s množinou X. Tímto získáme konvexní množinu Y, která obsahuje „jednostranné“ okolí x, tj. buď celou polopřímku ℓ nebo úsečku s krajními body x a y_ℓ, přičemž y_ℓ nemusí být prvkem tohoto průniku. V první a druhém případě neuděláme nic, zatímco ve třetím případě přidáme bod y_ℓ do Y. Pokud toto provedeme pro všechny polopřímky ℓ a přidáme všechny chybějící body y_ℓ obdržíme $\overline{X}$. ▲

Cvičení

6.1.1. Ukažte, že množina

$$X = \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^n \frac{x_i^2}{a_i^2} \le 1 \right\}$$

je pro libovolná $a_1, \dots, a_n \in \mathbb{R} \setminus \{0\}$ konvexní.

6.1.2. *Dokažte:* Množina $X \subseteq \mathbb{R}^n$ je konvexní právě tehdy, když její průnik s libovolnou přímkou je konvexní množina.

6.1.3. Rozhodněte, která z následujících množin X je konvexní (a polytop nebo polyedr).

a) *Fošna/traverza* $X = \{x \in \mathbb{R}^n \mid \alpha \le \langle p, x \rangle \le \beta\}$.

b) *Obdélník* (pro $n > 2$ též hyperobdélník) $X = \{x \in \mathbb{R}^n \mid \alpha_i \le x_i \le \beta_i,\ i = 1, \dots, n\}$.

c) *Klín* $X = \{x \in \mathbb{R}^n \mid \langle p_1, x \rangle \le \beta_1,\ \langle p_2, x \rangle \le \beta_2\}$.

d) Množina bodů, které jsou blíže k danému x_0 než množina $Y \subseteq \mathbb{R}^n$, tj.

$$X = \{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ pro všechna } y \in Y\}.$$

e) Množina bodů, které jsou blíže k množině $Y \subseteq \mathbb{R}^n$ než k množině $Z \subseteq \mathbb{R}^n$, tj.

$$X = \{x \in \mathbb{R}^n \mid \rho(x, Y) \leq \rho(x, Z)\},$$

kde $\rho(x, Y) := \inf\{\|x - y\|_2, y \in Y\}$.

f) Množina $X = \{x \in \mathbb{R}^n \mid x + X_1 \subseteq X_2\}$, kde $X_1, X_2 \subseteq \mathbb{R}^n$ a X_2 je konvexní.

g) Množina bodů, jejichž vzdálenost k a nepřekročí pevně daný násobek vzdálenosti k b, tj.

$$X = \{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \theta\|x - b\|_2\},$$

přičemž se můžeme omezit na situaci $a \neq b$ a $0 \leq \theta \leq 1$.

h) Množina $X = \{x \in \mathbb{R}^n \mid x \geq 0, \langle x, y\rangle \leq 1 \text{ pro všechna } y \text{ taková, že } \|y\|_2 = 1\}$.

i) Množina $X = \{x \in \mathbb{R}^n \mid x \geq 0, \langle x, y\rangle \leq 1 \text{ pro všechna } y \text{ taková, že } \sum_{i=1}^n |y_i| = 1\}$.

6.1.4. *Dokažte:* Množina $X \subseteq \mathbb{R}^n$ je konvexní právě tehdy, když pro každé $\lambda, \mu \geq 0$ platí

$$\lambda X + \mu X = (\lambda + \mu) X,$$

kde násobení skalárem a množinové sčítání je definováno jako v (6.1.2)–(6.1.3).

6.1.5. *S využitím předchozího příkladu dokažte:* Jsou-li $X, Y \subseteq \mathbb{R}^n$ konvexní množiny a $\alpha, \beta \geq 0$, pak je konvexní také množina

$$\bigcup_{\alpha,\beta \geq 0} (\alpha X + \beta Y).$$

6.1.6. Dokažte následující vlastnosti kužele.

a) Je-li $X_i \subseteq \mathbb{R}^n$ kužel pro každé $i \in I$, pak $\bigcap_{i \in I} X_i$ je také kužel.

b) Kartézský součin kuželů $X_1, X_2 \subseteq \mathbb{R}^n$, tj. $X_1 \times X_2$, je kužel.

c) Jsou-li $X_1, X_2 \subseteq \mathbb{R}^n$ kužely, pak množina $X_1 + X_2$ je kužel.

d) Je-li $X \subseteq \mathbb{R}^n$ kužel, pak $\overline{X}$ je kužel.

e) Je-li $X \subseteq \mathbb{R}^n$ kužel a $A \in \mathbb{R}^{m \times n}$, pak množina $AX = \{z \in \mathbb{R}^m \mid z = Ax, x \in X\}$ je kužel.

6.1.7. Dokažte následující vlastnosti.

a) Pro libovolnou množinu vektorů $\{a_i \in \mathbb{R}^n, i \in I\}$ je množina $X = \{x \in \mathbb{R}^n \mid a_i^\top x \leq 0, i \in I\}$ uzavřený konvexní kužel.

b) Kužel $X \subseteq \mathbb{R}^n$ je konvexní právě tehdy, když $X + X \subseteq X$.

c) Pro libovolné dva konvexní kužely $X_1, X_2 \subseteq \mathbb{R}^n$ platí

$$X_1 + X_2 = \operatorname{conv}(X_1 \cup X_2), \qquad X_1 \cap X_2 = \bigcup_{\lambda \in [0,1]} \{\lambda X_1 \cap (1 - \lambda) X_2\}.$$

6.1.8. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$ konvexní kužel, $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$, $x_1, \ldots, x_m \in X$, pak

$$\lambda_1 x_1 + \cdots + \lambda_m x_m \in X.$$

6.1.9. Pro libovolné dvě množiny $X, Y \subseteq \mathbb{R}^n$ definujeme množinové odčítání podobně jako sčítání, tj. $X - Y := \{z \in \mathbb{R}^n \mid z = x - y, x \in X, y \in Y\}$. Určete $X - X$, je-li X konvexní kužel $X = \{[x_1, x_2] \in \mathbb{R}^2 \mid x \geq 0, y \geq x\}$.

6.1.10. Jaký je vztah mezi konvexními a afinními množinami?

6.1.11. *Dokažte:* Pro $X \subseteq \mathbb{R}^n$ platí

$$\operatorname{cone} X = \left\{ \sum_{i=1}^m \lambda_i x_i \;\middle|\; \lambda_1, \ldots, \lambda_m \geq 0, x_i \in X \right\}.$$

6.1.12. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$ konvexní množina, potom platí

$$\operatorname{cone} X = \bigcup_{x \in X} \{\gamma x, \gamma \geq 0\}.$$

6.1.13. *Dokažte:* Pro $X \subseteq \mathbb{R}^n$ platí

$$\operatorname{aff} X = \Big\{ \sum_{i=1}^{m} \lambda_i x_i \;\Big|\; \lambda_1, \dots, \lambda_m \in \mathbb{R},\ \sum_{i=1}^{m} \lambda_i = 1,\ x_i \in X \Big\}.$$

6.1.14. *Ilustrujte platnost Lemma 6.1.13:* Určete kónický obal množiny $X = \{[1,2],[5,1]\}$ a ukažte, že jeho libovolný bod lze skutečně vyjádřit pouze pomocí dvou bodů z X.

6.1.15. Nechť $X_i \subseteq \mathbb{R}^n$ jsou konvexní množiny pro $i \in I$, kde I je libovolná indexová množina. Potom

$$\operatorname{conv}\Big(\bigcup_{i\in I} X_i\Big) = \bigcup_{\substack{\tilde{I}\subseteq I \\ \tilde{I} \text{ je konečná}}} \Big\{ \sum_{i\in\tilde{I}} \lambda_i X_i \;\Big|\; \lambda_i \geq 0 \text{ pro všechna } i \in \tilde{I},\ \sum_{i\in\tilde{I}} \lambda_i = 1 \Big\},$$

tj. konvexní obal sjednocení X_i je roven množině všech konvexních kombinací prvků z X_i.

6.1.16. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$, potom

a) $\operatorname{aff} X = \operatorname{aff}(\operatorname{conv} X)$;

b) $\operatorname{cone} X = \operatorname{cone}(\operatorname{conv} X)$;

c) $\operatorname{aff}(\operatorname{conv} X) \subseteq \operatorname{aff}(\operatorname{cone} X)$ a najděte příklad takové množiny X, že nastane ostrá inkluze;

d) je-li $0 \in \operatorname{conv} X$, potom $\operatorname{aff}(\operatorname{conv} X) = \operatorname{aff}(\operatorname{cone} X)$.

6.1.17. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$ otevřená množina, potom množina $\operatorname{conv} X$ je také otevřená. Platí i opačné tvrzení (tj. je-li $X \subseteq \mathbb{R}^n$ uzavřená množina, potom množina $\operatorname{conv} X$ je uzavřená)? Pokud ne, najděte vhodný protipříklad.

6.1.18. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina. Potom množina $\overline{X}$ je také konvexní.

6.1.19. Nechť

$$A := \{[x,y] \in \mathbb{R}^2 \mid y \geq 1 + x^2\} \cup \{[0,0]\}.$$

Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

řešení: $\operatorname{cone} A = \{[x,y] \in \mathbb{R}^2 \mid y \geq 2|x|\}$
a $\operatorname{conv} A = \{[x,y] \in \mathbb{R}^2 \mid y \geq 2|x|, -1 \leq x \leq 1\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \geq 1 + x^2, |x| > 1\}$

6.1.20. Nechť

$$A := \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \geq 2 + x_1^2\} \cup \{[0,1]\}.$$

Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte. Je bod $[1,3]$ vnitřním, hraničním nebo vnějším bodem množiny $\operatorname{conv} A$?

řešení:
$\operatorname{conv} A = \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \geq 2x_1 + 1, x_1 \in [0,1]\} \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \geq -2x_1 + 1, x_1 \in [-1,0]\} \cup$
$\cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \geq 2 + x_1^2, x_1 \in \mathbb{R} \setminus [-1,1]\}$, bod $[1,3]$ je hraničním bodem

6.1.21. Nechť

$$A := \{[x_1,x_2] \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq \mathrm{e}^{x_1}\} \cup \{[0,0]\}.$$

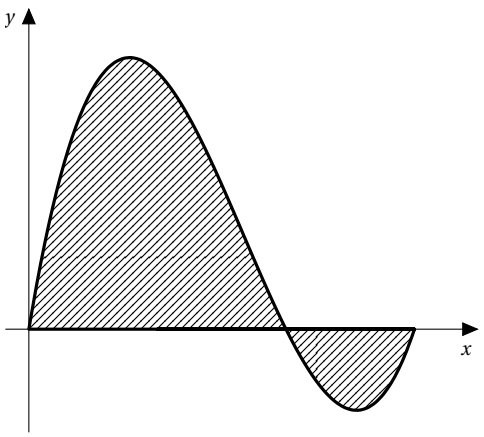
Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

řešení: $\operatorname{cone} A = \{[x_1,x_2] \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq \mathrm{e}\, x_1\}$,
$\operatorname{conv} A = \{[x_1,x_2] \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1, x_2 \geq \mathrm{e}\, x_1\} \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_1 \geq 1,\ x_2 \geq \mathrm{e}^{x_1}\}$

6.1.22. Nechť množina X je vymezena grafem funkce

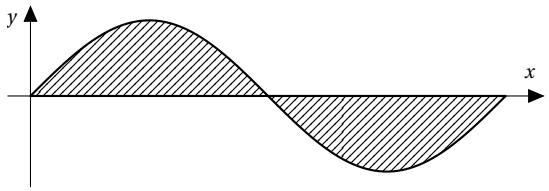
$$f(x) = x(x-2)(x-3) = x^3 - 5x^2 + 6x$$

pro $x \in [1,3]$ a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

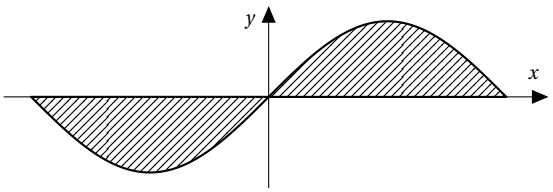


řešení: $\operatorname{cone} X = \{[x,y] \in \mathbb{R}^2 \mid -x/4 \le y \le 6x\}$ a $\operatorname{conv} X = \{[x,y] \in \mathbb{R}^2 \mid x \in [0,1],\ -x/4 \le y \le f(x)\} \cup$
$\cup \{[x,y] \in \mathbb{R}^2 \mid x \in [1,5/2],\ -x/4 \le y \le -x+3\} \cup \{[x,y] \in \mathbb{R}^2 \mid x \in [5/2,3],\ f(x) \le y \le -x+3\}$

6.1.23. Nechť množina X je vymezena grafem funkce $f(x) = \sin x$ pro $x \in [0, 2\pi]$ a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.



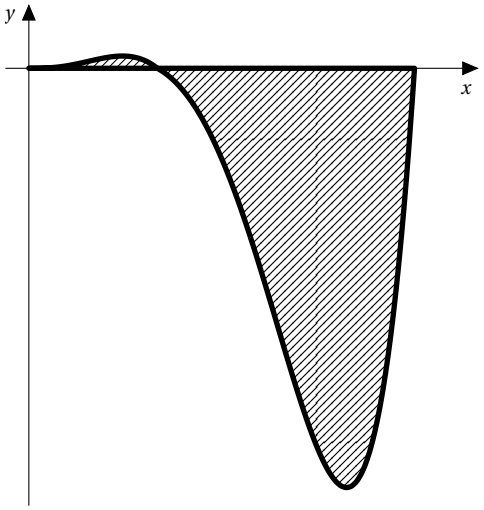
6.1.24. Nechť množina X je vymezena grafem funkce $f(x) = \sin x$ pro $x \in [-\pi, \pi]$ a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.



6.1.25. Nechť množina X je vymezena grafem funkce

$$f(x) = \tfrac{1}{3}x^3(x-1)(x-3) = \tfrac{1}{3}(x^5 - 4x^4 + 3x^3)$$

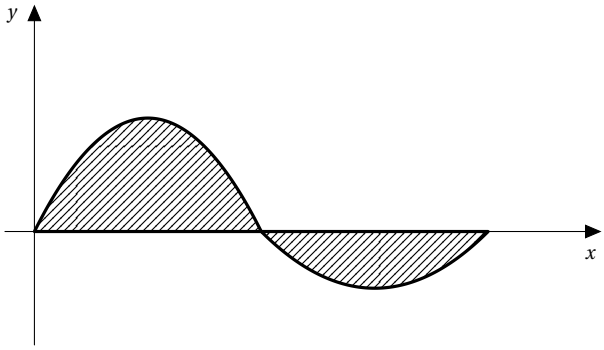
pro $x \in [0,3]$ a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.



6.1.26. Nechť množina A je vymezena grafem funkce

$$f(x) = \begin{cases} -2(x-1)x, & x \in [0,1], \\ (x-1)(x-2), & x \in [1,2], \end{cases}$$

a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.



řešení: $\operatorname{cone} A = \{[x,y] \in \mathbb{R}^2 \mid (2\sqrt{2}-3)x \le y \le 2x, x \ge 0\}$ a $\operatorname{conv} A = \Big(\{[x,y] \in \mathbb{R}^2 \mid y \ge (2\sqrt{2}-3)x, x \in [0,\sqrt{2}]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \ge (x-1)(x-2), x \in [\sqrt{2},2]\}\Big) \cap \Big(\{[x,y] \in \mathbb{R}^2 \mid y \le -2(x-1)x, x \in [0,2-\sqrt{2}]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \le (4\sqrt{2}-6)x + 12 - 8\sqrt{2}, x \in [2-\sqrt{2},2]\}\Big)$

6.1.27. Nechť

$$A = \{[x,y] \in \mathbb{R}^2 \mid y \le -x^2\} \cup \{[0,1]\}.$$

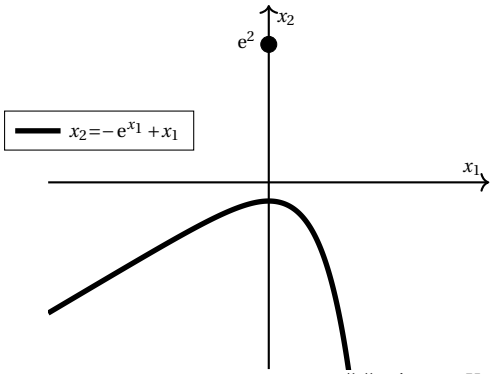
Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

řešení: $\operatorname{cone} A = \mathbb{R}^2$ a $\operatorname{conv} A = \{[x,y] \in \mathbb{R}^2 \mid y + 2x \le 1, x \in [0,1]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y - 2x \le 1, x \in [-1,0]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \le -x^2, x \in (\infty,-1]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \le -x^2, x \in [1,\infty)\}$

6.1.28. Mějme množinu (viz obrázek níže)

$$X = \{[x_1,x_2] \in \mathbb{R}^2 \mid -\mathrm{e}^{x_1} + x_1 \ge x_2,\ x_1 \ge 0\} \cup \{[0,\mathrm{e}^2]\}.$$

Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte. Jak se změní řešení při vypuštění podmínky $x_1 \ge 0$?



řešení: $\operatorname{cone} X = \{[x_1,x_2] \in \mathbb{R}^2 \mid x_1 \ge 0\}$
a $\operatorname{conv} X = X \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \le x_1(1-\mathrm{e}^2) + \mathrm{e}^2,\ x_1 \in [0,2]\}$;
po vypuštění $x_1 \ge 0$ dostaneme $\operatorname{cone} X = \mathbb{R}^2$
a $\operatorname{conv} X = X \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \le x_1(1-\mathrm{e}^2) + \mathrm{e}^2,\ x_1 \in [0,2]\} \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid x_2 \le x_1 + \mathrm{e}^2,\ x_1 \le 0\}$

6.1.29. Nechť

$$A = \{[x,y] \in \mathbb{R}^2 \mid (x-2)^2 + y^2 \le 1\} \cup \{[0,0],[2,2]\}.$$

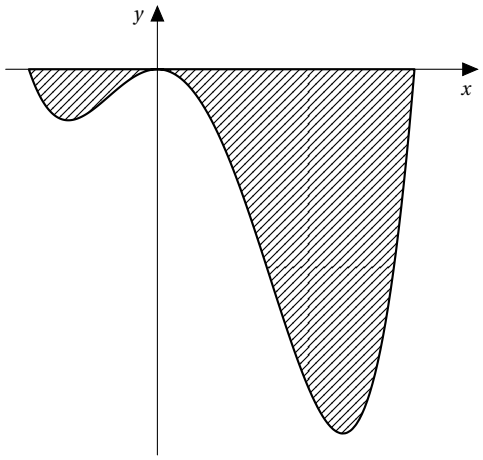
Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

řešení: $\operatorname{cone} A = \{[x,y] \in \mathbb{R}^2 \mid y \le x, x \ge 0, y \ge 0\} \cup \{[x,y] \in \mathbb{R}^2 \mid 3y + \sqrt{3}x \ge 0, x \ge 0, y \le 0\}$
a $\operatorname{conv} A = \{[x,y] \in \mathbb{R}^2 \mid y \le x, y \ge 0, x \in [0,2]\} \cup \{[x,y] \in \mathbb{R}^2 \mid y + \sqrt{3}x - 2 - 2\sqrt{3} \le 0, y \ge 0, x \in [2,\sqrt{3}/2+2]\} \cup \{[x,y] \in \mathbb{R}^2 \mid (x-2)^2 + y^2 \le 1, y \ge 0, x \in [\sqrt{3}/2+2,3]\} \cup \{[x,y] \in \mathbb{R}^2 \mid 3y + \sqrt{3}x \ge 0, y \le 0, x \in [0,3/2]\} \cup \{[x,y] \in \mathbb{R}^2 \mid (x-2)^2 + y^2 \le 1, y \le 0, x \in [3/2,3]\}$

6.1.30. Nechť množina X je vymezena grafem funkce

$$f(x) = x^2(x+1)(x-2) = (x^4 - x^3 - 2)$$

pro $x \in [-1,2]$ a osou x (viz obrázek níže). Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.



6.1.31. Nechť $p > 1$ a

$$A = \left\{ [x,y] \in \mathbb{R}^2 \;\middle|\; y \geq 1 + \frac{|x|^p}{p} \right\} \cup \{[0,0]\}.$$

Určete $\operatorname{conv} A$ a $\operatorname{cone} A$. Výsledné množiny také načrtněte.

řešení: pro $q := p/(p-1)$ je $\operatorname{cone} A = \{[x,y] \in \mathbb{R}^2 \mid y \geq q^{1/q}|x|\}$ a $\operatorname{conv} A = \{[x,y] \in \mathbb{R}^2 \mid y \geq q^{1/q}|x|,\ -q^{1/q} \leq x \leq q^{1/q}\} \cup \{[x,y] \in \mathbb{R}^2 \mid y \geq 1 + |x|^p/p, |x| > q^{1/q}\}$

6.1.32. *Dokažte:* Pro $X \subseteq \mathbb{R}^n$ platí $\operatorname{aff} X = \operatorname{aff} \overline{X}$.

6.1.33. *Dokažte:* Pro množinu $X \subseteq \mathbb{R}^n$ a matici $A \in \mathbb{R}^{m\times n}$ platí

a) $A \operatorname{ri}(X) = \operatorname{ri}(AX)$, kde $AX := \{y \in \mathbb{R}^m \mid y = Ax,\ x \in X\}$ je lineární transformace množiny X;

b) $A\overline{X} \subseteq \overline{AX}$. Je-li navíc množina X ohraničená, potom $A\overline{X} = \overline{AX}$. Uveďte příklad, kdy skutečně nastává ostrá inkluze.

6.1.34. *Dokažte:* Pro konvexní množiny $X, Y \subseteq \mathbb{R}^n$ platí

a) $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$ a $\operatorname{ri} X \cap \operatorname{ri} Y \subseteq \operatorname{ri}(X \cap Y)$. Je-li navíc $\operatorname{ri} X \cap \operatorname{ri} Y \neq \emptyset$, pak $\overline{X \cap Y} = \overline{X} \cap \overline{Y}$ a $\operatorname{ri} X \cap \operatorname{ri} Y = \operatorname{ri}(X \cap Y)$. Uveďte příklad, kdy nastávají ostré inkluze.

b) $\operatorname{ri}(X + Y) = \operatorname{ri} X + \operatorname{ri} Y$ a $\overline{X} + \overline{Y} \subseteq \overline{X + Y}$. Je-li alespoň jedna z množin X nebo Y ohraničená, pak $\overline{X} + \overline{Y} = \overline{X + Y}$. Uveďte příklad, kdy nastává ostrá inkluze.

6.1.35. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$ konvexní množina a $\mathcal{V}$ zaměření afinního obalu $\operatorname{aff} X$, potom

$$\operatorname{ri} X = \operatorname{int}\left(X + \mathcal{V}^\perp\right) \cap X,$$

kde $\mathcal{V}^\perp$ značí obvyklý ortogonální doplněk vektorového podprostoru $\mathcal{V}$, tj. $\mathcal{V}^\perp := \{y \in \mathbb{R}^n \mid \langle x, y\rangle = 0 \text{ pro všechna } x \in \mathcal{V}\}$.

6.1.36. *Dokažte:* Nechť $X, Y \subseteq \mathbb{R}^n$ jsou konvexní množiny takové, že $X \subseteq Y$.

a) Množina $\operatorname{ri} X$ nemusí být podmnožinou $\operatorname{ri} Y$ (stačí najít vhodný protipříklad).

b) Pokud $\operatorname{aff} X = \operatorname{aff} Y$, pak $\operatorname{ri} X \subseteq \operatorname{ri} Y$.

c) Pokud $\operatorname{ri} X \cap \operatorname{ri} Y \neq \emptyset$, $\operatorname{ri} X \subseteq \operatorname{ri} Y$.

d) Pokud $X \cap \operatorname{ri} Y \neq \emptyset$, $\operatorname{ri} X \cap \operatorname{ri} Y \neq \emptyset$.

6.1.37. *Dokažte rozšíření Věty 6.1.30:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina. Potom $x \in \operatorname{ri} X$ právě tehdy, když pro každé $\bar{x} \in \operatorname{aff} X$ existuje $\gamma > 1$ takové, že $x + (\gamma - 1)(x - \bar{x}) \in X$.

6.1.38. *Dokažte:*

a) Je-li $X \subseteq \mathbb{R}^n$ konvexní množina a $0 \in \operatorname{ri} X$, pak $\operatorname{cone} X = \operatorname{aff} X$.

b) Nechť $X \subseteq \mathbb{R}^n$ a $0 \in \operatorname{ri}(\operatorname{conv} X)$, pak $\operatorname{cone} X = \operatorname{aff} X$.

6.1.39. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$.

a) Je-li $\operatorname{int} X \neq \emptyset$, pak $\mathrm{r}\partial!X = \partial X$. Je-li navíc X konvexní množina, pak $\mathrm{r}\partial!X = \partial X$ právě tehdy, když $\operatorname{int} X \neq \emptyset$.

b) Platí $\mathrm{r}\partial X = \emptyset$ právě tehdy, když X je afinní množina,

c) Je-li X konvexní množina, potom $\mathrm{r}\partial X = \mathrm{r}\partial\overline{X} = \mathrm{r}\partial(\operatorname{ri} X)$.

d) Je-li X konvexní a ohraničená množina, pak libovolná polopřímka vycházející z bodu $x \in \operatorname{ri} X$ a ležící v $\operatorname{aff} X$ protíná $\mathrm{r}\partial X$ právě v jednom bodě.

6.1.40. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$.

a) Je-li X konvexní a kompaktní a $0 \notin \mathrm{r}\partial X$, pak $\operatorname{cone} X$ je uzavřená množina. Ukažte na příkladu, že toto tvrzení neplatí, pokud X je neohraničená nebo $0 \in \mathrm{r}\partial X$.

b) Je-li X kompaktní a $0 \notin \mathrm{r}\partial(\operatorname{conv} X)$, pak $\operatorname{cone} X$ je uzavřená množina.

6.1.41. *Dokažte:*

a) Nechť $X \subseteq \mathbb{R}^n$ je konvexní kužel. Pak $\operatorname{ri} X$ je také konvexní kužel s možnou výjimkou počátku, tj. $\operatorname{cone}(\operatorname{ri} X) = \operatorname{ri} X \cup \{0\}$.

b) Nechť $X = \operatorname{cone}(\{x_1, \ldots, x_m\}) \subseteq \mathbb{R}^n$. Pak

$$\operatorname{ri} X = \left\{ y \in \mathbb{R}^n \;\middle|\; y = \sum_{i=1}^m \lambda_i x_i,\ \lambda_i > 0,\ i = 1, \ldots, m \right\}.$$

6.1.42. *Dokažte:* Nechť $X_i \subseteq \mathbb{R}^n$ jsou neprázdné množiny pro $i = 1, \ldots, m$ a uvažme jejich kartézský součin $X = X_1 \times X_2 \times \cdots \times X_m$. Potom

a) $\operatorname{conv} X = \operatorname{conv} X_1 \times \cdots \times \operatorname{conv} X_m$, $\overline{X} = \overline{X_1} \times \cdots \times \overline{X_m}$ a $\operatorname{aff} X = \operatorname{aff} X_1 \times \cdots \times \operatorname{aff} X_m$,

b) pokud $0 \in X_1 \cap \cdots \cap X_m$, tj. všechny množiny obsahují počátek, pak $\operatorname{cone} X = \operatorname{cone} X_1 \times \cdots \times \operatorname{cone} X_m$,

c) jsou-li $X_1, \ldots, X_m$ konvexní množiny, pak $\operatorname{ri} X = \operatorname{ri} X_1 \times \cdots \times \operatorname{ri} X_m$.

6.2 Konvexní funkce

Úlohy matematického programování mohou být vskutku velmi rozličné a tato obšírnost nás pochopitelně může přivést (a mnohem později i přivede) až do značně komplikované situace při jejich řešení, které leckdy probíhá ve dvou krocích: (i) najít „podezřelé“ body[49], (ii) rozhodnout o jejich „extrémnosti“. Zatímco první krok lze převést na soustavu rovnic a nerovnic, druhý krok už je mnohem složitější. Později však uvidíme, že situace se značně zjednoduší, pokud se budeme pohybovat na konvexní množině a v případě minimalizačních úloh pracovat s konvexními[50] funkcemi. Proto se nyní a v Podkapitole 6.4 budeme konvexním funkcím věnovat. Ukážeme si nejdříve několik základních vlastností a pak se zaměříme zejména na ty, které pro nás budou důležité v dalších částech.

S konvexními funkcemi jsme se nepochybně setkali již v základním kurzu matematické analýzy (a možná i na střední škole) v případě jedné proměnné, kde vyjadřovali především jakési zakřivení grafu funkce. Jistě si pamatujeme, že typickými představiteli této třídy funkcí jsou x^2 nebo e^x. V literatuře lze najít několik ekvivalentních definic konvexních funkcí a my začneme následující připomínkou z [142, Definice 6.21].

> Funkce $f : \mathbb{R} \to \mathbb{R}$ se nazývá *konvexní*[51] na intervalu $I \subseteq \mathbb{R}$, jestliže pro libovolná $x, y, z \in I$ platí
>
> $$f(y) \le f(x) + \frac{f(z) - f(x)}{z - x}(y - x). \qquad (6.2.1)$$
>
> Je-li tato nerovnost ostrá pro libovolnou trojici různých bodů, pak je funkce *ostře konvexní.*

[49] Těm obvykle říkáme *stacionární.*

[50] Případně s konkávními funkcemi, pokud by šlo o maximalizaci.

[51] Jakkoli přirozeně a důležitě může konvexnost funkce působit, poprvé se její definice objevuje až v roce 1893 v [464]. Na druhou stranu, již v roce 1821 hledal Cauchy všechny spojité funkce splňující funkcionální rovnici

$$f(x+y) = f(x) + f(y),$$

ze které plyne, že všechna nalezená řešení budou splňovat také konvexnost připomínající nerovnost

$$f\Big(\frac{x+y}{2}\Big) \le \frac{f(x)+f(y)}{2}.$$

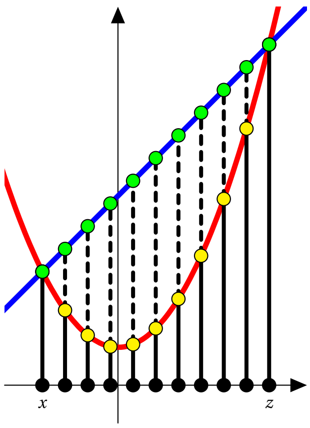
Pro více historických podrobností viz [39, 416].

Co ale nerovnost (6.2.1) ve skutečnosti (tj. nyní geometricky) znamená? Vezmeme-li $x < y < z$ a položíme-li $\lambda := \frac{z-y}{z-x}$, pak $\lambda \in (0,1)$ a podmínka (6.2.1) přejde do tvaru

$$f\big(\lambda x + (1-\lambda) z\big) \le \lambda f(x) + (1-\lambda) f(z), \qquad (6.2.2)$$

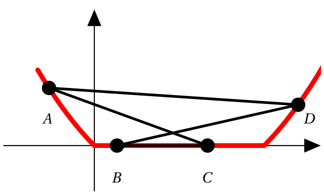
přičemž pro $\lambda \in \{0,1\}$ je tato nerovnost splněna triviálně, tj. lze brát $\lambda \in [0,1]$. Nyní z výrazů na obou stranách této nerovnosti udělejme druhou souřadnici bodů v $\mathbb{R}^2$, jejichž první souřadnice je $\lambda x + (1-\lambda) z$, tj. uvažme body

$$A_\lambda := \big[\lambda x + (1-\lambda) z, f\big(\lambda x + (1-\lambda) z\big)\big] \quad \text{a} \quad B_\lambda := \big[\lambda x + (1-\lambda) z, \lambda f(x) + (1-\lambda) f(z)\big].$$



Obrázek 6.35: Geometrická ilustrace nerovnosti (6.2.2). Body A_λ jsou žlutou barvou pro $\lambda = n/10$ s $n \in \{1, 2, \ldots, 9\}$, zatímco body B_λ pro tytéž hodnoty λ a navíc i pro $n = 0, 1$ (průsečíky přímky a grafu funkce) jsou zelenou barvou. Černé body na ose x odpovídají hodnotám $\lambda x + (1-\lambda) z$.

Pro libovolné $\lambda \in [0,1]$ leží bod A_λ na grafu funkce f, zatímco B_λ se nachází na úsečce s krajními body $[x, f(x)]$ a $[z, f(z)]$. Body A_λ a B_λ vždy mají stejnou první souřadnici, což znamená, že nerovnost (6.2.2) vyjadřuje požadavek, aby *libovolná sečna nadgrafu odpovídající intervalu I ležela nad grafem*, tj. aby sečna mezi body $[x, f(x)]$ a $[z, f(z)]$ nikdy neprotnula graf mimo své krajní body. Toto je znázorněno na Obrázcích 6.35–6.37. Neostrost v (6.2.2) navíc dovoluje, aby sečna i splynula s částí grafu f, nesmí ji ale protnout. Pokud dokonce pro žádné dva různé body sečna nesplyne s grafem funkce, je funkce ostře konvexní.



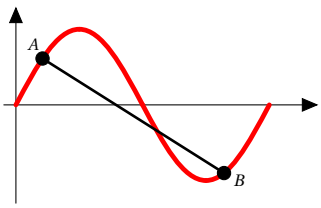
Obrázek 6.36: Geometrická ilustrace nerovnosti (6.2.2). Na obrázku je několik sečen. Funkce je konvexní, není ale ostře konvexní.

Nerovnost (6.2.1) je také ekvivalentní s tím, že pro libovolnou trojici bodů $x, y, z \in I$ splňující bez újmy na obecnosti[52] $x < y < z$ platí[53]

$$\frac{f(y) - f(x)}{y - x} \le \frac{f(z) - f(x)}{z - x} \le \frac{f(z) - f(y)}{z - y}, \qquad (6.2.3)$$

přičemž v případě ostré konvexnosti jsou tyto nerovnosti ostré. Označíme-li odpovídající body na grafu funkce f jako $A := [x, f(x)]$, $B := [y, f(y)]$ a $C := [z, f(z)]$, pak dvojice nerovností v (6.2.3) vyjadřuje

$$\varphi(AB) \le \varphi(AC) \le \varphi(BC),$$

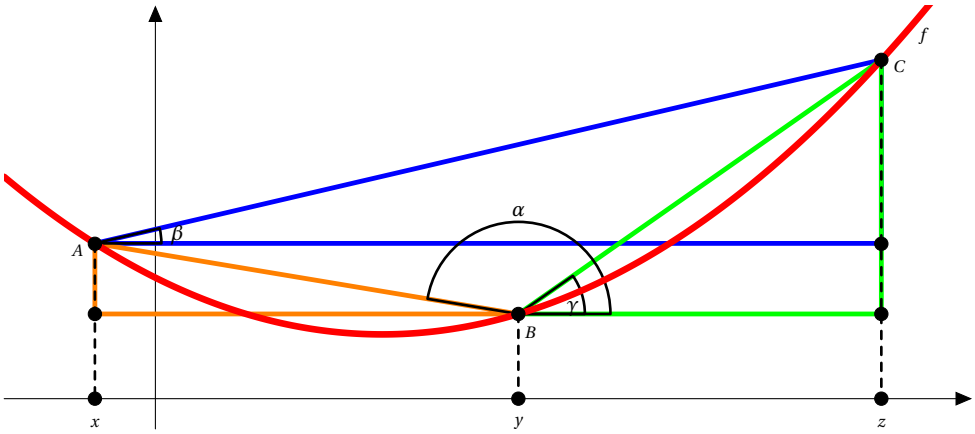


Obrázek 6.37: Ukázka nekonvexní funkce.

kde $\varphi(AB)$ značí směrnici sečny vytyčené body A a B. Viz Obrázek 6.38, ze kterého snadno vidíme, že skutečně

$$\underbrace{\operatorname{tg}\alpha = -\frac{f(x)-f(y)}{y-x}}_{=\varphi(AB)} \le \underbrace{\operatorname{tg}\beta = \frac{f(z)-f(x)}{z-x}}_{=\varphi(AC)} \le \underbrace{\operatorname{tg}\gamma = \frac{f(z)-f(y)}{z-y}}_{=\varphi(BC)}.$$

Tuto vlastnost konvexních funkcí bychom proto mohli označit jako *Větu o třech sečnách.*



Obrázek 6.38: Ilustrace geometrického významu nerovnosti (6.2.3).

My si nyní naše dosavadní poznatky o konvexních funkcích rozšíříme na funkce více proměnných $f:\mathbb{R}^n\to\mathbb{R}$. Ještě před tím ale poznamenejme, že v nemalém množství monografií se lze setkat i s přístupem, kdy $f:\mathbb{R}^n\to\mathbb{R}\cup\{\pm\infty\}$, tj. funkce f může navzdory našim dosavadním zvyklostem nabývat[54] v některých bodech i hodnot $\pm\infty$. Touto cestou se ale nevydáme, neboť bychom celý výklad trochu zkomplikovali, což se jeví býti zbytečné ve světle toho, že je tento přístup vhodný zejména při studiu některých speciálních vlastností konvexních funkcí, které však jdou již nad rámec našeho pojednání. Na základní vlastnosti, kterým se chceme v dalších částech věnovat, toto rozšíření či jeho absence nemá žádný vliv.

[52] Chceme se vyhnout tomu, že dva body budou stejné, pak totiž situace triviální. Zbytek je již jen o vhodném označení jednotlivých bodů.

[53] Vynásobením první nerovnosti jmenovatelem z levé strany obdržíme nerovnost (6.2.1). Naopak, z podmínky $x<y<z$ plyne existence $\lambda\in(0,1)$ takového, že $y=z+\lambda(x-z)$. Je-li f konvexní, pak jistě platí

$$\begin{aligned} f(y)-f(x) &= \\ &= f(\lambda x+(1-\lambda)z)-f(x) \le \\ &\le \lambda f(x)+(1-\lambda)f(z)-f(x) = \\ &= (1-\lambda)(f(z)-f(x)), \end{aligned}$$

takže

$$\begin{aligned} \frac{f(y)-f(x)}{y-x} &= \frac{f(y)-f(x)}{(1-\lambda)(z-x)} \\ &\le \frac{(1-\lambda)(f(z)-f(x))}{(1-\lambda)(z-x)}. \end{aligned}$$

Druhá nerovnost se pak dokáže analogickým způsobem.

Definice 6.2.1

Nechť $X\subseteq\mathbb{R}^n$ je konvexní množina. Funkce $f:X\to\mathbb{R}$ se nazývá

(i) *konvexní* na X, jestliže pro všechna $x,y\in X$ a každé $\lambda\in[0,1]$ platí[55]

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y); \tag{6.2.4}$$

(ii) *ostře* (*ryze*) *konvexní* na X, jestliže nerovnost (6.2.4) je ostrá pro všechna $x,y\in X$, $x\neq y$ a každé $\lambda\in(0,1)$;

(iii) *silně konvexní* na X s *konstantou silné konvexnosti* $\vartheta>0$, jestliže[56,57] pro všechna $x,y\in X$ a každé $\lambda\in[0,1]$ platí

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)-\vartheta\lambda(1-\lambda)\|x-y\|^2. \tag{6.2.5}$$

Funkce f se nazývá *konkávní* (příp. *ostře* nebo *silně*), jestliže $-f$ je konvexní, tj. platí opačné nerovnosti jako v (6.2.4) a (6.2.5).

[54] Nikoli ve smyslu limity, tj. např. je možné, že $f(x^*)=\infty$ pro nějaké x^*.

[55] Množina X musí být nutně konvexní podmnožinou definičního oboru funkce f. Proč? Aby výraz na levé straně nerovnosti (6.2.4) měl smysl pro všechna $\lambda\in[0,1]$.

[56] Je konstanta silné konvexnosti určena jednoznačně? Z nerovnosti (6.2.5) by mělo být patrné, že není, neboť pokud je tato nerovnost splněna pro nějaké $\vartheta=\vartheta^*>0$, pak bude jistě splněna i pro libovolnou jinou hodnotu $\vartheta\in(0,\vartheta^*]$. Proto je užitečné znát především *největší* možnou hodnotu této konstanty, viz kvadratickou formu v Příkladu 6.2.4.

Poznámka 6.2.2. S konceptem konvexnosti (příp. konkávnosti) se můžeme setkat na řadě míst, kde hraje zcela přirozenou roli, např. v teorii užitku (viz později). Nerovnost (6.2.4) aplikovaná pro $\lambda=1/2$ totiž vlastně říká, že

$$f\left(\tfrac{x+y}{2}\right)\le\frac{f(x)+f(y)}{2}$$

neboli funkční hodnota pro průměr dvou hodnot je menší nebo rovna průměru funkčních hodnot – toto vlastně také dává první (velmi primitivní) test konvexnosti. V praxi si koncept

konvexnosti/konkávnosti můžeme představit třeba takto (skutečná teorie užitku přijde opravdu až později):

(i) funkce f vyjadřuje náš užitek ze směsi dvou komodit[58], např. čaj a voda. V takovém případě můj užitek z nějakého množství vody a čajových lístků bude jistě větší než užitek z 20 litrů vody nebo pytle čajových lístků, v čemž můžeme spatřit princip konkávnosti;

(ii) naopak situace s 5 litry kapaliny do ostřikovače a 50 litrů nafty bude zcela opačná, tj. objevuje se zde konvexnost.

Navíc pokud $0 \in X$ a $f(0) = 0$, pak konvexnost garantuje tzv. *neklesající výnosy z rozsahu*[59] , tj.

$$f(\mu x) \geq \mu f(x) \quad \text{pro } \mu \geq 1. \tag{6.2.6}$$

Vskutku, z nerovnosti (6.2.4) volbou $x := \mu z$, $y := 0$ a $\lambda := 1/\mu$ pro libovolné $\mu \geq 1$ dostaneme $\lambda \in (0,1]$ a nerovnost

$$f(z) = f\big(\lambda x + (1-\lambda)\, y\big) \leq \lambda f(x) + (1-\lambda) f(y) = \tfrac{1}{\mu} f(\mu z). \qquad \blacktriangle$$

Příklad 6.2.3

Již velmi dobře víme, že konvexní funkce jsou např.

$$\begin{aligned} &f(x) = \mathrm{e}^{ax}, \quad x \in \mathbb{R}, \quad a \in \mathbb{R}, \\ &f(x) = x^a, \quad x > 0, \quad a \in (-\infty, 0] \cup [1, \infty), \\ &f(x) = |x|^a, \quad x \in \mathbb{R}, \quad a \geq 1, \\ &f(x) = x \ln x, \quad x > 0, \end{aligned}$$

zatímco $f(x) = x^a$ je konkávní pro $a \in [0,1]$. Konstantní funkce $f(x) \equiv a$ je konvexní (i konkávní) pro libovolné $a \in \mathbb{R}$ a libovolný počet proměnných, avšak není ostře a ani silně konvexní, neboť pro libovolná $x, y \in \mathbb{R}^n$ a $\lambda \in [0,1]$ platí

$$f(\lambda x + (1-\lambda) y) = \alpha = \lambda f(x) + (1-\lambda) f(y)$$

a nerovnost (6.2.5) se redukuje na podmínku

$$\alpha \leq \alpha - \vartheta \lambda (1-\lambda) \|x - y\|^2,$$

jenž bude splněna pouze pro $\vartheta \leq 0$, avšak my vyžadujeme $\vartheta > 0$.

Příklad 6.2.4

S využitím Definice 6.2.1 můžeme (s větší či menší mírou námahy) ukázat, že pro $X = \mathbb{R}^n$ nebo libovolnou jeho konvexní podmnožinu jsou konvexní také funkce

(i) $f(x) = \langle c, x \rangle = c^\top x$ (lineární funkce) pro libovolné $c \in \mathbb{R}^n$,

(ii) $f(x) = \langle Ax, x \rangle = x^\top A x$ (kvadratická forma) pro libovolnou matici $A = A^\top \geq 0$,

(iii) $f(x) = \|x\|$ (libovolná norma),

(iv) $f(x) = \|x\|^2$ (čtverec libovolné normy indukované skalárním součinem[60]),

(v) $f(x) = \|x\|_2^4 = (x^\top x)^2$ (čtvrtá mocnina euklidovské normy),

(vi) $f(x) = \|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ (Hölderova norma) pro libovolné $p \geq 1$,

(vii) $f(\beta) = \left\| \mathbf{X}\beta - y \right\|_2^2$ (účelová funkce z metody nejmenších čtverců či spíše lineární regrese), kde $\mathbf{X} = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \in \mathbb{R}^{m \times 2}$ je matice složená z hodnot nezávislé proměnné při měření (nebo experimentu) v regresním modelu a jedniček, $y \in \mathbb{R}^m$ je vektor složený z odpovídajících výsledků měření a $\beta \in \mathbb{R}^2$ jsou hledané parametry regresní přímky.

[57] V nerovnosti (6.2.5) symbolem $\|\cdot\|$ rozumíme obvyklou euklidovskou normu v $\mathbb{R}^n$. Toto nebudeme dále explicitně zmiňovat s výjimkou Příkladu 6.2.4.

[58] Zde „směsí“ myslíme skutečně smíchání těchto komodit.

[59] *Výnosy z rozsahu* (angl. *returns to scale* je ekonomický pojem vyjadřující proporcionalitu změn výstupu poté, co se množství všech vstupů ve výrobě změnilo stejným faktorem. Matematicky zapsáno to znamená, že pro produkční funkci funkci f, libovolné x z vhodné podmnožiny $\mathbb{R}^n$ (tzv. dostupné skladby výroby) a všechna $t \geq 0$ (případně s ostrou nerovností) platí $f(tx) = t^k f(x)$, kde $k \in \mathbb{R}$ vyjadřuje právě onu proporcionalitu. Pro $k = 1$ hovoříme o *konstantních* výnosech z rozsahu (tj. např. zdvojnásobení vstupů znamená zdvojnásobení produkce), zatímco pro $k > 1$ jsou výnosy *rostoucí* a pro $k < 1$ *klesající*. Funkce splňující $f(tx) = t^k f(x)$ se v matematice nazývá *homogenní stupně* k.

[60] To jsou takové normy, které splňují *rovnoběžníkové pravidlo*

$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$.

Řešení. Vskutku, v prvním případě je podmínka (6.2.4) splněna díky linearitě skalárního součinu, neboť

$$f(\lambda x+(1-\lambda)y)=\langle c,\lambda x+(1-\lambda)y\rangle=\lambda\langle c,x\rangle+(1-\lambda)\langle c,y\rangle=\lambda f(x)+(1-\lambda)f(y).$$

Jelikož se požadovaná nerovnost vždy realizuje jako rovnost, nemůže tato funkce být ostře či silně konvexní.

V případě kvadratické formy využijeme toho, že podmínka (6.2.4) je ekvivalentní s nezáporností funkce $\varphi(x,y):=\lambda f(x)+(1-\lambda)f(y)-f(\lambda x+(1-\lambda)y)$ a ostrá konvexnost je ekvivalentní s $\varphi(x,y)>0$ pro libovolná $x\neq y$ a $\lambda\in(0,1)$. Protože však pro libovolná $x,y\in\mathbb{R}^n$ platí

$$\begin{aligned}\varphi(x,y)&=\lambda x^{\top}Ax+(1-\lambda)y^{\top}Ay-[\lambda x+(1-\lambda)y]^{\top}A[\lambda x+(1-\lambda)y]=\\&=\lambda x^{\top}Ax+(1-\lambda)y^{\top}Ay-\lambda^2x^{\top}Ax-\lambda(1-\lambda)x^{\top}Ay-\\&\qquad-\lambda(1-\lambda)y^{\top}Ax-(1-\lambda)^2y^{\top}Ay=\\&=\lambda(1-\lambda)x^{\top}Ax+[1-(1-\lambda)](1-\lambda)y^{\top}Ay-\lambda(1-\lambda)x^{\top}Ay-\lambda(1-\lambda)y^{\top}Ax\\&=\lambda(1-\lambda)x^{\top}Ax+\lambda(1-\lambda)y^{\top}Ay-\lambda(1-\lambda)x^{\top}Ay-\lambda(1-\lambda)y^{\top}Ax=\\&=\lambda(1-\lambda)(x-y)^{\top}A(x-y),\end{aligned}$$

dostáváme odtud

- kvadratická forma je konvexní $\iff$ $\varphi(x,y)\geq 0$ $\iff$ $A=A^{\top}\geq 0$;
- kvadratická forma je ostře konvexní $\iff$ $\varphi(x,y)>0$ $\iff$ $A=A^{\top}>0$.

Je tato funkce i silně konvexní, tj. platí $\varphi(x,y)\geq\vartheta\lambda(1-\lambda)(x-y)^{\top}(x-y)$ pro nějaké $\vartheta>0$? S využitím nerovnosti pro symetrické matice a jejich vlastní hodnoty[61] obdržíme

$$\varphi(x,y)=\lambda(1-\lambda)(x-y)^{\top}A(x-y)\geq\lambda_{\min}(A)\lambda(1-\lambda)(x-y)^{\top}(x-y),$$

tj. kvadratická forma je silně konvexní právě tehdy, když $\lambda_{\min}(A)>0$, tj. matice A je pozitivně definitní. Za konstantu silné konvexnosti pak můžeme vzít libovolné číslo v intervalu $\big(0,\lambda_{\min}(A)\big]$, takže volba $\vartheta=\lambda_{\min}(A)$ dává největší možnou volbu.

Z trojúhelníkové nerovnosti platné pro libovolnou normu ihned pro jakékoli $\lambda\in[0,1]$ a $x,y\in\mathbb{R}^n$ obdržíme

$$\|\lambda x+(1-\lambda)y\|\leq\lambda\|x\|+(1-\lambda)\|y\|,$$

což dokazuje její konvexnost. Avšak žádná norma není ostře (a tudíž ani silně) konvexní. Vezmeme-li $y=tx$, pak na levé straně bude

$$\|\lambda x+(1-\lambda)tx\|=\big|\lambda+(1-\lambda)t\big|x,$$

zatímco na pravé straně

$$\lambda\|x\|+(1-\lambda)\|y\|=\big[\lambda+(1-\lambda)|t|\big]x.$$

Tyto dva výrazy si jsou rovny právě tehdy, když $t\geq 0$.
Dále, díky požadovanému propojení normy a skalárního součinu v části (iv) víme, že

$$\|x\pm y\|^2=\langle x\pm y,x\pm y\rangle=\|x\|^2\pm 2\langle x,y\rangle+\|y\|^2.$$

Odtud podobně jako ve druhé části pro libovolné $x,y\in\mathbb{R}^n$ a $\lambda\in[0,1]$ získáme

$$\begin{aligned}&\|\lambda x+(1-\lambda)tx\|^2-\lambda\|x\|^2-(1-\lambda)\|x\|^2=\\&\qquad=\lambda^2\|x\|^2+2\lambda(1-\lambda)\langle x,y\rangle+(1-\lambda)^2\|y\|^2-\lambda\|x\|^2-(1-\lambda)\|y\|^2=\\&\qquad=\lambda(\lambda-1)\|x\|^2+(1-\lambda)(1-\lambda-1)\|y\|^2+2\lambda(1-\lambda)\langle x,y\rangle=\\&\qquad=-\lambda(1-\lambda)\big(\|x\|^2-2\langle x,y\rangle+\|y\|^2\big)=\\&\qquad=-\lambda(1-\lambda)\|x-y\|^2,\end{aligned}$$

[61] Pro libovolnou symetrickou $n\times n$ matici platí

$$\lambda_{\min}(A)\,I\leq A\leq\lambda_{\max}(A)\,I,$$

kde $\lambda_{\min}(A)$ značí nejmenší vlastní číslo matice A, $\lambda_{\max}(A)$ je největší vlastní číslo A a I je $n\times n$ jednotková matice.

což ukazuje silnou konvexnost čtverců takovýchto norem s největší konstantou[62,63] silné konvexnosti $\vartheta = 1$.
Odvození (ostré/silné) konvexnosti zbývajících funkcí ponecháme čtenáři jako Cvičení 6.4.1 a 6.2.6–6.2.7.

[62] Takto pojatý závěr by vyžadoval, aby v definici silné konvexnosti byla stejná norma. V opačném případě bychom využili ekvivalenci norem, což by mělo za následek pouze jinou maximální hodnotu konstantu silné konvexnosti.

[63] Kdybychom uvažovali jenom euklidovskou normu, pak by tento závěr plynul přímo z části (ii), neboť tato norma odpovídá kvadratické formě s maticí $A = I$.

Z předchozích příkladů bychom možná mohli chybně usoudit, že není rozdíl mezi ostrou a silnou konvexností, neboť uvedené funkce buď splňovaly obě předepsané podmínky nebo žádnou z nich. Taková situace by však pochopitelně byla z matematické hlediska trochu hloupá. Při pohledu na jednotlivé podmínky z Definice 6.2.1 je zřejmé, že silná konvexnost implikuje ostrou konvexnost, a ta zase „obyčejnou“ konvexnost, tj. platí následující dvojice implikací

$$\text{silná konvexnost} \overset{\text{(i)}}{\Longrightarrow} \text{ostrá konvexnost} \overset{\text{(ii)}}{\Longrightarrow} \text{konvexnost.}$$

Navíc v případě implikace (i) se dá ukázat, že funkce f je silně konvexní s konstantou ϑ právě tehdy, když funkce $g(x) := f(x) - \vartheta\|x\|^2$ je konvexní[64], viz Cvičení 6.2.13. V takovém případě totiž silná konvexnost s konstantou ϑ znamená, že

$$f(\lambda x + (1-\lambda)y) - \vartheta\|\lambda x + (1-\lambda)y\|^2 \le \lambda f(x) + (1-\lambda) f(y) - \vartheta\lambda\|x\|^2 - \vartheta(1-\lambda)\|y\|^2,$$

což pro $x \neq y$ a $\lambda \in (0,1)$ společně s ostrou konvexností funkce $\|\cdot\|^2$ vede k nerovnosti

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) + \underbrace{\vartheta\|\lambda x + (1-\lambda)y\|^2 - \vartheta\lambda\|x\|^2 - \vartheta(1-\lambda)\|y\|^2}_{<0 \text{ díky ostré konvexnosti}} <$$

$$< \lambda f(x) + (1-\lambda) f(y)$$

prokazující, že f je skutečně ostře konvexní. Opačné implikace samozřejmě neplatí, jak se snadno přesvědčíme např. pomocí

- funkce $f(x) = \langle c, x\rangle$, která je konvexní, ale není ostře konvexní;
- funkce $f(x) = \|x\|^4$, která je ostře konvexní, ale není silně konvexní na žádné množině X obsahující počátek, viz Cvičení 6.4.1.

[64] Jinými slovy, funkce f „musí“ být konvexní alespoň jako $\|\cdot\|^2$.

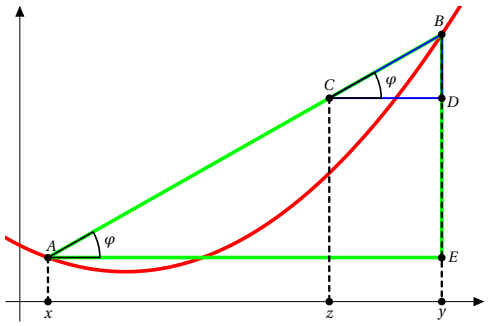
Poznámka 6.2.5. Pro konvexní funkce v $\mathbb{R}^n$ platí analogický vztah mezi grafem a sečnou jako v případě funkce jedné proměnné, tj. funkce f je konvexní právě tehdy, když úsečka spojující libovolné dva body $z_1, z_2 \in \mathbb{R}^{n+1}$ na grafu funkce f je taková, že všechny její body leží na nebo nad tímto grafem. Vskutku, nechť $C := [z, w]$ pro $z := \lambda x + (1-\lambda) y$ pro libovolné $\lambda \in (0,1)$ je bod na úsečce spojující body $A := [x, f(x)]$ a $B := [y, f(y)]$ ležící na grafu funkce f, viz Obrázek 6.39 pro funkci jedné proměnné. Položme $D := [y, w]$ a $E := [y, f(x)]$. Potom díky podobnosti pravoúhlých trojúhelníků AEB a CDB platí

$$\frac{f(y) - w}{\lambda\|y - x\|} = \frac{\|\overline{BD}\|}{\|\overline{CD}\|} = \frac{\|\overline{BE}\|}{\|\overline{AE}\|} = \frac{f(y) - f(x)}{\|y - x\|},$$

což znamená, že

$$f(x) - w = \lambda[f(y) - f(x)] \quad \text{neboli} \quad w = \lambda f(x) + (1-\lambda) f(y).$$

Proto bod C leží nad grafem funkce f právě tehdy, když platí (6.2.4).

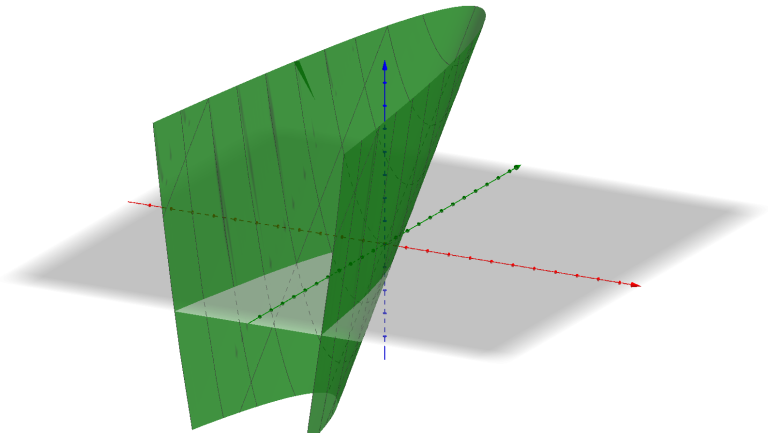


Obrázek 6.39: Konvexnost vs. graf funkce.

Avšak rozhodovat o ostré/silné konvexnosti funkce pouze s pomocí jejího grafu může být zavádějící. Např. z grafu funkce $f(x_1,x_2)=x_1^2+x_2$ by se mohlo zdát, že tato funkce je ostře konvexní, viz Obrázek 6.40. Jenže např. pro dvojici různých bodů $x:=[0,1]$ a $y:=[0,2]$ platí

$$f(\lambda x+(1-\lambda)y)=f(0,\lambda+2(1-\lambda))=2-\lambda=\lambda f(0,1)+(1-\lambda)f(0,2)=\lambda f(x)+(1-\lambda)f(y),$$

tj. f nemůže být ostře konvexní.



Obrázek 6.40: Grafu funkce $f(x_1,x_2)=x_1^2+x_2$.

▲

Pro konvexnost funkce je klíčové její chování na všech úsečkách v definičním oboru, a tak následující trvzení není nikterak překvapivé.

VĚTA 6.2.6

Nechť $X\subseteq\mathbb{R}^n$ je konvexní množina. Funkce $f:X\to\mathbb{R}$ je konvexní právě tehdy, když funkce jedné proměnné $g(t):=f(x+ty)$ je pro libovolné $x\in X$ a $y\in\mathbb{R}^n$ konvexní na intervalu $I\subseteq\mathbb{R}$ takovém, že $x+ty\in X$.

Mnoho numerických metod pro řešení minimalizačních problémů využívá v průběhu svého iteračního procesu minimalizaci dané funkce podél vhodné (polo)přímky. Předchozí tvrzení zaručuje, že v případě konvexní funkce bude i její zúžení konvexní funkcí, a tedy půjde o úlohu konvexního programování.

K charakterizaci konvexnosti funkce f můžeme využít i její nadgraf, jak ukazuje následující tvrzení, které leckdy bývá použito i jako definice konvexnosti.

VĚTA 6.2.7

Nechť $X\subseteq\mathbb{R}^n$ je konvexní množina a nechť $f:X\to\mathbb{R}$. Funkce f je konvexní na X právě tehdy, když její *nadgraf* (epigraf)

$$\operatorname{epi} f:=\{[x,\beta]\in\mathbb{R}^{n+1}\mid x\in X,\ \beta\geq f(x)\}$$

je konvexní množina.

Důkaz. „$\Longrightarrow$“ Nechť f je konvexní funkce na X. Ukážeme, že s libovolnými dvěma body z $\operatorname{epi} f$ obsahuje $\operatorname{epi} f$ také celou úsečku spojující tyto body. Nechť tedy $\lambda\in[0,1]$ a mějme dvojici $[x_1,\beta_1],[x_2,\beta_2]\in\operatorname{epi} f$, tj. $\beta_1\geq f(x_1)$ a $\beta_2\geq f(x_2)$. Potom

$$\lambda[x_1,\beta_1]+(1-\lambda)[x_2,\beta_2]=\big[\underbrace{\lambda x_1+(1-\lambda)x_2}_{\in X},\lambda\beta_1+(1-\lambda)\beta_2\big]$$

a současně

$$\lambda\beta_1+(1-\lambda)\beta_2\geq\lambda f(x_1)+(1-\lambda)f(x_2)\geq f(\lambda x_1+(1-\lambda)x_2),$$

což ale znamená, že také $\lambda[x_1,\beta_1]+(1-\lambda)[x_2,\beta_2]\in\operatorname{epi} f$, tj. množina $\operatorname{epi} f$ je konvexní.

„⟸“ Nechť epi f je konvexní množina. Ukážeme, že v takovém případě pro libovolné dva body z X platí (6.2.4). Nechť $x_1, x_2 \in X$ a $\lambda \in [0,1]$. Pak pro $[x_1, f(x_1)], [x_2, f(x_2)] \in \operatorname{epi} f$ je také

$$\lambda\,[x_1, f(x_1)] + (1-\lambda)\,[x_2, f(x_2)] = \big[\lambda x_1 + (1-\lambda)\,x_2,\, \lambda f(x_1) + (1-\lambda)\,f(x_2)\big] \in \operatorname{epi} f,$$

což ale znamená, že $\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$, tj. f je konvexní na X. ■

Vzhledem k přechodu od konvexní funkce f ke konkávní $-f$, je zřejmé, že množina všech konvexních funkcí na konvexní množině X netvoří vektorový prostor. Nicméně následující z následujícího tvrzení plyne, že tato množina je konvexním kuželem[65].

[65] Dokonce uzavřeným konvexním kuželem, což však už vyžaduje trochu více práce...

Věta 6.2.8

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f_1, \ldots, f_m : X \to \mathbb{R}$ jsou konvexní na X a $\alpha_1, \ldots, \alpha_m \geq 0$ jsou daná čísla. Potom také funkce $F(x) := \alpha_1 f_1(x) + \cdots + \alpha_m f_m(x)$ je konvexní na X.

Důkaz. Tvrzení plyne z přímého výpočtu, přičemž pouze stačí ukázat konvexnost nezáporného násobku konvexní funkce a součtu dvou konvexních funkcí (srovnej s důkazem druhé části Věty 6.1.3). Je-li $f : X \to \mathbb{R}$ konvexní na X, pak funkce αf je také konvexní na X pro libovolné $\alpha \geq 0$, neboť pro $\alpha = 0$ máme nulovou funkci a pro $\alpha > 0$ platí

$$(\alpha f)(\lambda x_1 + (1-\lambda) x_2) = \alpha f(\lambda x_1 + (1-\lambda) x_2) \overset{\alpha>0}{\leq} \alpha\lambda f(x_1) + \alpha(1-\lambda) f(x_2),$$

kde $x_1, x_2 \in X$ a $\lambda \in [0,1]$ jsou libovolná.

Nechť nyní $f_1, f_2 : X \to \mathbb{R}$ jsou konvexní funkce na X. Potom také $f_1 + f_2$ je konvexní funkce na X, neboť pro libovolné $x_1, x_2 \in X$ a $\lambda \in [0,1]$ máme

$$\begin{aligned}(f_1 + f_2)(\lambda x_1 + (1-\lambda) x_2) &= f_1(\lambda x_1 + (1-\lambda) x_2) + f_2(\lambda x_1 + (1-\lambda) x_2)\\ &\leq \lambda f_1(x_1) + (1-\lambda) f_1(x_2) + \lambda f_2(x_1) + (1-\lambda) f_2(x_2)\\ &= \lambda\big[f_1(x_1) + f_2(x_1)\big] + (1-\lambda)\big[f_1(x_2) + f_2(x_2)\big]\\ &= \lambda\,(f_1 + f_2)(x_1) + (1-\lambda)\,(f_1 + f_2)(x_2),\end{aligned}$$

čímž je důkaz hotov. ■

[65] Pro $a_0 = 0$ bude funkce dokonce *lineární*.

Příklad 6.2.9

Funkce

$$f(x) = f(x_1, \ldots, x_n) = a_0 + a_1 x_1 + \cdots + a_n x_n = a_0 + \langle a, x\rangle$$

se nazývá *afinní*[66]. Tato funkce je současně konvexní i konkávní, což můžeme snadno odvodit buď přímo z definice nebo s využitím Příklad 6.2.4 a Věty 6.2.8. Tytéž argumenty nám dávají také konvexnost funkce

$$f(x) = \langle Ax, x\rangle + \langle c, x\rangle + b$$

pro libovolnou pozitivně semidefinitní matici A.

Ve Větě 6.2.7 jsme charakterizovali konvexní funkce pomocí nadgrafu, tj. množiny epi $f \subseteq \mathbb{R}^{n+1}$. V tomto ohledu bychom ale neměli zapomínat ani na vrstevnice funkce. Přesněji řečeno, na *dolní obrysovou množinu*[67]

$$V_K := \big\{x \in X \mid f(x) \leq K\big\}, \tag{6.2.7}$$

což je sjednocení všech vrstevnic $\big\{x \in X \mid f(x) = L\big\}$ na úrovních $L \in (-\infty, K]$. Tuto množinu si lze představit jako „oříznuté pohoří“, kdy z grafu funkce oříznete všechny „výčnělky vyšší než K“. Ještě zdůrazněme, že následující charakterizace konvexních funkcí pomocí množin V_K již nenabízí ekvivalentní popis.

[67] Zde bohužel není česká terminologie (a ani autor) moc kreativní. Bylo by možné použít i označení *podúrovňová množina* nebo *uzavřená vrstevnicová oblast* či *vrstevnicové kontinuum*. Anglický termín *sublevel set* nebo *lower level set* je mnohem výstižnější.

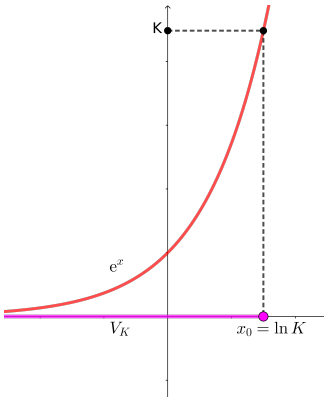
Věta 6.2.10

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$ konvexní funkce na X. Pak pro libovolné $K \in \mathbb{R}$ je odpovídající dolní obrysová množina V_K také konvexní.

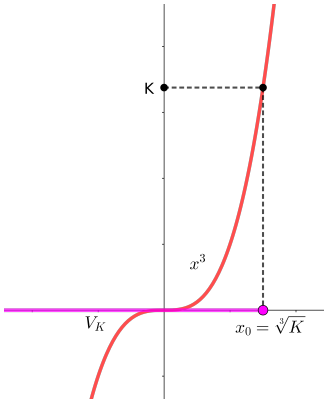
Důkaz. Je-li $V_K = \emptyset$ nebo jednoprvková množina, pak je tvrzení triviální. Nechť proto dále množina V_K obsahuje alespoň dva různé prvky. Zvolme $x_1, x_2 \in V_K$ libovolně. Pak pro každé $\lambda \in [0,1]$ díky konvexnosti f máme

$$f(\lambda x_1 + (1-\lambda) x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) \le \lambda K + (1-\lambda) K = K,$$

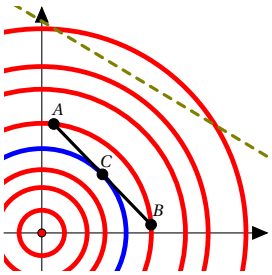
tj. také $\lambda x_1 + (1-\lambda) x_2 \in V_K$, což ukazuje konvexnost množiny V_K. ■



Obrázek 6.41: Funkce $f(x) = e^x$ a množina V_K.



Obrázek 6.42: Funkce $f(x) = x^3$ a množina V_K.



Obrázek 6.43: Ilustrace *axiomu rozmanitosti.*

[68] Johan Ludwig William Valdemar Jensen (8. května 1859 – 5. března 1925) byl dánský matematik–samouk. Základní kurzy matematiky sice absolvoval při studiu na Kodaňské technologické univerzitě, pokročilejší témata se naučil sám a nikdy nezastával žádnou akademickou pozici, takže se svému matematickému výzkumu věnoval výhradně ve volném čase. V letech 1892–1903 byl prezidentem Dánské matematické společnosti.

Poznámka 6.2.11. Podmínka z Věty 6.2.10 je pouze nutnou nikoli postačující. Obě jsou splněny např. funkcemi $f(x) = e^x$ i $g(x) = x^3$ na celém $\mathbb{R}$, přičemž druhá z nich zde rozhodně není konvexní. V obou případech jsou totiž množiny V_K intervaly tvaru $(-\infty, x_0]$ pro vhodné x_0, což jsou vždy konvexní množiny, viz Obrázky 6.41–6.42. Funkce splňující podmínku, že množina V_K je konvexní pro všechna $K \in \mathbb{R}$, se nazývají *kvazikonvexní*. Ekvivalentně lze kvazikonvexní funkce definovat podmínkou

$$f(\lambda x_1 + (1-\lambda) x_2) \le \max\{f(x_1), f(x_2)\} \tag{6.2.8}$$

pro libovolné $\lambda \in [0,1]$ a dvojici x_1, x_2 z konvexní množiny $X \subseteq \mathbb{R}$. V případě ostré nerovnosti pro libovolné $\lambda \in (0,1)$ a $x_1 \neq x_2$ hovoříme o *ostře kvazikonvexní funkci. Kvazikonkávní funkce* jsou určeny podobnou podmínkou

$$f(\lambda x_1 + (1-\lambda) x_2) \ge \min\{f(x_1), f(x_2)\}.$$

Očividně každá konvexní funkce je kvazikonvexní. Avšak i konkávní funkce může být kvazikonvexní jako je tomu např. pro $f(x) = \ln x$. Obecně každá monotónní funkce jedné proměnné je kvazikonvexní a kvazikonkávní současně. Kvazikonvexní bude určitě každá funkce $f : \mathbb{R} \to \mathbb{R}$, která je ostře klesající až do nějakého bodu a od něj naopak ostře rostoucí.

Podobná vlastnost platí také pro konkávní funkce, ovšem místo množiny V_K je potřeba brát *horní obrysové množiny*

$$D_K := \{x \in X \mid f(x) \ge K\}.$$

Právě tato vlastnost činí konkávní funkce důležitými v teorii užitku, o které jsme se zmínili na začátku. Vede to totiž k tzv. *axiomu rozmanitosti*, který říká, že „průměr je lepší než extrém“. Uvažme konkávní funkci f, která určuje náš užitek ze dvou komodit/statků, např. $f(x_1, x_2) = -x_1^2 - x_2^2$. Na Obrázku 6.43 vidíme několik vrstevnic $-x_1^2 - x_2^2 = K$ (v ekonomii nazývaných jako *indiferenční křivky*) na různých úrovních – červené kružnice se středem v počátku, které existují pouze pro nekladná K a s hodnotami K rostoucími (zleva) do nuly se kružnice smršťují. Horní obrysové množiny pak jsou do sebe vnořené kruhy vymezené těmito kružnicemi. Zelená přímka představuje rozpočtové omezení, tj. odpovídající omezená část prvního kvadrantu zahrnuje všechny kombinace množství komodit x_1 a x_2 vyhovující našemu dostupnému rozpočtu. Na obrázku je také zakreslena dvojice bodů A a B, které představují dvě různé kombinace komodit ležící na téže vrstevnici, tj. dávající tentýž užitek neboli $f(A) = f(B)$. Úsečka $\overline{AB}$ však protíná další vrstevnice a vidíme, že libovolná konvexní kombinace A a B dává větší užitek. Obzvláště ve středu této úsečky (bod C) je užitek nejvyšší možný, neboť se zde z úsečky stává tečna nějaké vrstevnice (vyznačené modrou barvou), tj. užitek z průměru daných kombinací A a B je vyšší než užitek v těchto krajních bodech (průměr dokonce dává maximální možných užitek z těchto dvou kombinací). ▲

Dalším velmi důležitým výsledek z teorie konvexních funkcí je *Jensenova*[68] *nerovnost* z roku 1906[69], která rozšiřuje nerovnost (6.2.4) na libovolnou konvexní kombinaci bodů z X, takže ji můžeme brát jako „funkční“ analogii Věty 6.1.9.

Věta 6.2.12 Jensenova nerovnost

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ je konvexní na X. Pak pro libovolné $m \in \mathbb{N}$, $x_1, \dots, x_m \in X$ a čísla $\lambda_1, \dots, \lambda_m \geq 0$ splňující $\sum_{i=1}^m \lambda_i = 1$ platí

$$f\Big(\sum_{i=1}^m \lambda_i x_i\Big) \leq \sum_{i=1}^m \lambda_i f(x_i). \tag{6.2.9}$$

Je-li funkce f navíc ostře konvexní a $\lambda_1, \dots, \lambda_m \in (0,1)$, pak rovnost v (6.2.9) nastane právě tehdy, když $x_1 = \cdots = x_m$.

[69] Speciální tvar se objevuje již v roce 1885 jako *Hölder defect formula*, viz také [273].

Poznámka 6.2.13. První část Věty 6.2.12 lze samozřejmě naformulovat jako ekvivalenci, avšak chybějící opačná implikace je zcela triviální a vyplývá ihned z Definice 6.2.1. Obměnou[70] složeného výroku ze druhé části dostaneme:

> Jestliže existuje $m \in \mathbb{N}$ takové, že buď v (6.2.9) nastane rovnost pro různé body nebo nastane ostrá nerovnost pro stejné body, pak funkce f není ostře konvexní nebo $\lambda_i \in \{0,1\}$ pro nějaké $i \in \{1 \dots, m\}$.

Toto tvrzení asi není úplně přehledné, takže můžeme nabídnout alternativní formulaci:

> Nechť $m \in \mathbb{N}$ je libovolné. Pak pro $\lambda_1, \dots, \lambda_m \in [0,1]$ splňující $\sum_{i=1}^m \lambda_i = 1$ a m-tici $x_1, \dots, x_m \in X$ obsahující alespoň dva různé body s odpovídajícími kladnými koeficienty λ_p a λ_q nastane v (6.2.9) ostrá nerovnost právě tehdy, když funkce f je ostře konvexní[71].

Na anglické verzi *Wikipedie* lze najít druhou část tvrzení Věty 6.2.12 v trochu odlišné podobě:

> Rovnost v (6.2.9) nastane právě tehdy, když $x_1 = \cdots = x_m$ nebo f je lineární *na množině obsahující* $x_1, \dots, x_m$.

Ta poslední část (vyznačená kurzívou) přibyla až v říjnu 2019 – bez ní by bylo tvrzení jistě nepravdivé, jak lze vidět na protipříkladu

$$f(x) = \begin{cases} x^2, & x \in \mathbb{R} \setminus [-1,1], \\ 1, & x \in [-1,1], \end{cases}$$

neboť toto je konvexní funkce a pro $x_1 = 1$, $x_2 = -1$ a $\lambda_1 = \lambda_2 = 1/2$ platí

$$f\big(\tfrac{1}{2}1 + \tfrac{1}{2}(-1)\big) = f(0) = 1 = \tfrac{1}{2} + \tfrac{1}{2} = \tfrac{1}{2}f(1) + \tfrac{1}{2}f(-1),$$

tj. máme různé body a nelineární funkci. Dodatek se tedy vypořádává s tímto protipříkladem — avšak pouze částečně. Vezmeme-li totiž tentokrát trojici bodů $x_1 = 1$, $x_2 = -1$, $x_3 = 2$ s koeficienty $\lambda_1 = \lambda_2 = 1/2$, $\lambda_3 = 0$, pak opět nastává rovnost, ale funkce rozhodně není lineární na množině obsahující x_1, x_2, x_3. Kdybychom se chtěli vyhnout i tomuto protipříkladu, museli bychom vyžadovat $\lambda_1, \dots, \lambda_m \in (0,1)$, tj. zakázat možné nulové koeficienty λ_i. ▲

[70] Máme-li logický výrok $A \implies B$, pak jeho *obměnou* rozumíme tvrzení $\neg B \implies \neg A$ a tento výrok je pravdivý právě tehdy, když je pravdivý původní výrok, tj. $A \implies B \equiv \neg B \implies \neg A$. Obměnou složeného výroku $(A \wedge B) \implies (C \iff D)$ dostaneme ekvivalentní výrok

$$\begin{aligned}&(A \wedge B) \implies (C \iff D) \equiv \\ &\equiv \neg(C \iff D) \implies \neg(A \wedge B) \equiv \\ &\equiv \big((C \wedge \neg D) \vee (\neg C \wedge D)\big) \implies \\ &\qquad \implies (\neg A \vee \neg B)\end{aligned}$$

[71] Implikace „$\implies$“ plyne ihned z Věty 6.2.12 (kdyby f byla ostře konvexní a nastala by rovnost, pak bychom vynecháním bodů s odpovídajícím nulovým koeficientem λ_i dostali spor s Větou 6.2.12). Opačná implikace plyne ihned z definice ostré konvexnosti.

Důkaz Věty 6.2.12. Tvrzení dokážeme pomocí matematické indukce. Pro $m = 1$ je nerovnost triviální, zatímco pro $m = 2$ je ekvivalentní s definicí konvexností funkce f na X. Připusťme nyní, že tvrzení platí pro nějaké $m \in \mathbb{N} \setminus \{1\}$. Ukážeme, že pak platí i pro $m+1$. Je-li $\lambda_{m+1} = 1$, potom $\lambda_1 = \cdots = \lambda_m = 0$ a nerovnost (6.2.9) je splněna triviálně. Uvažme tedy $\lambda_{m+1} < 1$. Pak

$$f(\lambda_1 x_1 + \cdots + \lambda_{m+1} x_{m+1}) = f\left((1-\lambda_{m+1})\left[\frac{\lambda_1}{1-\lambda_{m+1}} x_1 + \cdots + \frac{\lambda_m}{1-\lambda_{m+1}} x_m\right] + \lambda_{m+1} x_{m+1}\right),$$

přičemž $\frac{\lambda_1}{1-\lambda_{m+1}}, \dots, \frac{\lambda_m}{1-\lambda_{m+1}} \in [0,1]$ a $\sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}} = 1$. Proto platí

$$\widetilde{x} := \frac{\lambda_1}{1-\lambda_{m+1}} x_1 + \cdots + \frac{\lambda_m}{1-\lambda_{m+1}} x_m \in X,$$

a tedy z konvexnosti funkce f na X plyne

$$\begin{aligned} f(\lambda_1 x_1 + \cdots + \lambda_{m+1} x_{m+1}) &= f\big((1-\lambda_{m+1})\,\widetilde{x} + \lambda_{m+1} x_{m+1}\big) \leq \\ &\overset{\text{konvexnost}}{\leq} (1-\lambda_{m+1})\, f(\widetilde{x}) + \lambda_{m+1} f(x_{m+1}) \leq \\ &\leq (1-\lambda_{m+1})\, f\left((1-\lambda_{m+1})\left[\frac{\lambda_1}{1-\lambda_{m+1}} x_1 + \cdots + \frac{\lambda_m}{1-\lambda_{m+1}} x_m\right]\right) + \lambda_{m+1} f(x_{m+1}) \leq \end{aligned}$$

$$\overset{\text{„indukce“}}{\leq} (1-\lambda_{m+1})\Big[\frac{\lambda_1}{1-\lambda_{m+1}} f(x_1)+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}} f(x_m)\Big]+\lambda_{m+1} f(x_{m+1}) =$$
$$= \lambda_1 f(x_1)+\cdots+\lambda_{m+1} f(x_{m+1}),$$

což dokazuje platnost nerovnosti (6.2.9) pro libovolné $m \in \mathbb{N}$.

Nechť je nyní navíc f ostře konvexní a $\lambda_1,\dots,\lambda_m \in (0,1)$. Je-li $x_1 = \cdots = x_m =: x \in X$, pak zřejmě platí

$$f(\lambda_1 x_1+\cdots+\lambda_m x_m) = f\Big(x\sum_{i=1}^m \lambda_i\Big) = f(x) = f(x)\sum_{i=1}^m \lambda_i = \sum_{i=1}^m \lambda_i f(x) = \sum_{i=1}^m \lambda_i f(x_i),$$

tedy v (6.2.9) nastává rovnost.

Naopak, nechť $x_1,\dots,x_m \in X$ jsou taková, že v (6.2.9) nastává rovnost. Opět využijeme matematickou indukci. Nechť $m=2$ a $\lambda_1,\lambda_2 \in (0,1)$ splňují $\lambda_1+\lambda_2=1$. Pak rovnost (6.2.9) je tvaru

$$f(\lambda_1 x_1+\lambda_2 x_2) = \lambda_1 f(x_1)+\lambda_2 f(x_2),$$

a tudíž nutně $x_1 = x_2$, protože jinak bychom dostali spor s definicí ostré konvexnosti. Nechť toto platí pro nějaké $m \in \mathbb{N}\setminus\{1\}$ a nechť

$$f\Big(\sum_{i=1}^{m+1} \lambda_i x_i\Big) = \sum_{i=1}^{m+1} \lambda_i f(x_i), \tag{6.2.10}$$

kde $\lambda_1,\dots,\lambda_{m+1} \in (0,1)$ splňují $\sum_{i=1}^{m+1}\lambda_i = 1$. Položme $\widetilde{x} := \sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}} x_i$. Potom

$$f(\lambda_1 x_1+\cdots+\lambda_m x_m+\lambda_{m+1}x_{m+1}) = f\big((1-\lambda_{m+1})\widetilde{x}+\lambda_{m+1}x_{m+1}\big) \leq$$
$$\overset{\text{konvexnost}}{\leq} (1-\lambda_{m+1}) f(\widetilde{x})+\lambda_{m+1} f(x_{m+1}) =$$
$$= (1-\lambda_{m+1}) f\Big(\frac{\lambda_1}{1-\lambda_{m+1}} x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}} x_m\Big)+\lambda_{m+1} f(x_{m+1}) \leq$$
$$\overset{(6.2.9)}{\leq} \lambda_1 f(x_1)+\cdots+\lambda_{m+1} f(x_{m+1}).$$

Jenže vzhledem k indukčnímu předpokladu (6.2.10) je zřejmé, že se nerovnosti v předchozím výpočtu ve skutečnosti realizují jako rovnosti. Pak ale první z nich implikuje $\widetilde{x} = x_{m+1}$ (ze stejného důvodu jako pro $m=2$), zatímco ta poslední dává $x_1 = \cdots = x_m$ (indukční předpoklad). Proto z definice $\widetilde{x}$ plyne

$$x_{m+1} = \widetilde{x} = \sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}} x_i = x_1 \sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}} = \frac{x_1}{1-\lambda_{m+1}}\sum_{i=1}^m \lambda_i = \frac{x_1}{1-\lambda_{m+1}}(1-\lambda_{m+1}) = x_1,$$

což dohromady dává $x_1 = \cdots = x_m = x_{m+1}$ a důkaz je hotov. ■

Poznámka 6.2.14. (i) Ekvidistantní volbou rozdělení m-tice parametrů $\lambda_1,\dots,\lambda_m$ v intervalu $[0,1]$ pro libovolné $m \in \mathbb{N}\setminus\{1\}$, tj. $\lambda_1 = \cdots = \lambda_m = 1/m$, dostaneme pro konvexní funkce nerovnost

$$f\Big(\frac{x_1+\cdots+x_m}{m}\Big) \leq \frac{f(x_1)+\cdots+f(x_m)}{m}$$

ukazující (opět) vztah mezi funkční hodnotou průměru a průměrem funkčních hodnot.

(ii) Z Jensenovy nerovnosti lze odvodit celou řadu zajímavých a užitečných výsledků. Ten nejznámější získáme, pokud ji aplikujeme na (ostře) konvexní funkci $f(x) = -\ln x$ na intervalu $(0,\infty)$. V takovém případě totiž pro libovolná $x_1,\dots,x_m > 0$ a $\lambda_1,\dots,\lambda_m \in [0,1]$ splňující $\sum_{i=1}^n \lambda_i = 1$ obdržíme

$$-\ln\Big(\sum_{i=1}^m \lambda_i x_i\Big) \leq -\sum_{i=1}^m \lambda_i \ln x_i = -\ln\Big(\prod_{i=1}^m x_i^{\lambda_i}\Big),$$

což vzhledem k tomu, že $-\ln x$ je klesající funkce, dává

$$\sum_{i=1}^m \lambda_i x_i \leq \prod_{i=1}^m x_i^{\lambda_i}.$$

[72] Angl. *Inequality of arithmetic and geometric means* nebo jen zkráceně *AM–GM inequality*.

Obzvláště pokud vezmeme $\lambda_1 = \cdots = \lambda_m = 1/m$, dostaneme *nerovnost mezi aritmetickým a geometrickým průměrem* (též *AG-nerovnost*[72])

$$\underbrace{\frac{x_1 + \cdots + x_m}{m}}_{\text{aritmetický průměr}} \geq \underbrace{\sqrt[m]{x_1 \cdots x_m}}_{\text{geometrický průměr}}, \tag{6.2.11}$$

přičemž tato nerovnost dokonce platí pro $x_1, \ldots, x_m \geq 0$ a rovnost nastane tehdy a jen tehdy, když [73] $x_1 = \cdots = x_m$.

Víceméně analogickým postupem lze získat i tzv. *váženou AG-nerovnost*, tj. pro libovolná $x_1, \ldots, x_m \geq 0$, $\lambda_1, \ldots, \lambda_m \geq 0$ a $n := \sum_{i=1}^m \lambda_i$ platí

$$\frac{\lambda_1 x_1 + \cdots + \lambda_m x_m}{n} \geq \left(x_1^{\lambda_1} \cdots x_m^{\lambda_m}\right)^{1/n}, \tag{6.2.12}$$

přičemž rovnost nastane tehdy a jen tehdy, když $x_1 = \cdots = x_m$. Pro $\lambda_1 = \cdots = \lambda_m = 1$ dostaneme předchozí nerovnost (6.2.11), zatímco pro libovolná $a, b \geq 0$ a $p, q > 0$ splňující $1/p + 1/q = 1$ obdržíme volbou $x_1 := a^p$, $x_2 := b^q$, $\lambda_1 := 1/p \geq 0$ a $\lambda_2 := 1/q \geq 0$ *Youngovu nerovnost*

$$\frac{a^p}{p} + \frac{b^q}{q} \geq a\,b.$$

(iii) Předchozí nerovnosti lze ještě více rozšířit pomocí *váženého průměru řádu t*, který je definován pro $x_1, \ldots, x_m > 0$, $\lambda_1, \ldots, \lambda_m > 0$, $n := \sum_{i=1}^m \lambda_i$ a $t \in \mathbb{R} \setminus \{0\}$ jako

$$M_t^{[m]}(x_1, \ldots, x_m; \lambda_1, \ldots, \lambda_m) = M_t := \left(\frac{\lambda_1 x_1^t + \cdots + \lambda_m x_m^t}{n}\right)^{1/t}.$$

Pak volbou $t = 1$ dostaneme *vážený aritmetický průměr*

$$M_1 = \frac{\lambda_1 x_1 + \cdots + \lambda_m x_m}{n},$$

volbou $t = 2$ dostaneme *vážený kvadratický průměr*

$$M_2 = \sqrt{\frac{\lambda_1 x_1^2 + \cdots + \lambda_m x_m^2}{n}}$$

a volbou $t = -1$ obdržíme *vážený harmonický průměr*

$$M_{-1} = \frac{n}{\lambda_1/x_1 + \cdots + \lambda_m/x_m}.$$

Jelikož s využitím l'Hospitalova pravidla můžeme vypočítat limitu

$$\lim_{t\to 0} \ln M_t = \lim_{t\to 0} \frac{\ln \frac{\lambda_1 x_1^t + \cdots + \lambda_m x_m^t}{n}}{t} \left\| \frac{0}{0} \right\| \overset{\text{l'H.p.}}{=} \lim_{t\to 0} \frac{\frac{n}{\lambda_1 x_1^t + \cdots + \lambda_m x_m^t} \cdot \frac{\lambda_1 x_1^t \ln x_1 + \cdots + \lambda_m x_m^t \ln x_m}{n}}{1} =$$
$$= \frac{1}{n}\sum_{i=1}^m \lambda_i \ln x_i = \frac{1}{n}\sum_{i=1}^m \ln x_i^{\lambda_i}, \tag{6.2.13}$$

platí pro vážené průměry, že

$$M_0 := \lim_{t\to 0} M_t \,\|1^\infty\| = \mathrm{e}^{\lim_{t\to 0} \frac{1}{t} \times \ln \frac{\lambda_1 x_1^t + \cdots + \lambda_m x_m^t}{n}} \overset{(6.2.13)}{=} \mathrm{e}^{\frac{1}{n}\sum_{i=1}^m \ln x_i^{\lambda_i}} = \left(\prod_{i=1}^m x_i^{\lambda_i}\right)^{1/n} = \sqrt[n]{x_1^{\lambda_1} \cdots x_m^{\lambda_m}},$$

což vlastně představuje *vážený geometrický průměr*. Navíc, jestliže $x_k = \max\{x_1, \ldots, x_m\}$, pak[74] platí

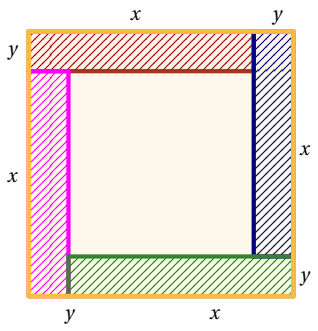
$$M_\infty := \lim_{t\to\infty} \left(\tfrac{\lambda_1 x_1}{n} + \cdots + \tfrac{\lambda_m x_m}{n}\right)^{1/t} = \lim_{t\to\infty} x_k \underbrace{\left[\frac{\lambda_1}{n}\underbrace{\left(\frac{x_1}{x_k}\right)}_{\leq 1}^t + \cdots + \frac{\lambda_k}{n} + \cdots + \frac{\lambda_m}{n}\underbrace{\left(\frac{x_m}{x_k}\right)}_{\leq 1}^t\right]}_{\geq \lambda_k/n \quad \& \quad \leq 1}^{1/t} = x_k$$

a podobně $M_{-\infty} := \lim_{t\to-\infty} M_t = \min\{x_1, \ldots, x_m\}$.

[73] Obzvláště v případě $m = 2$ dostaneme nerovnost

$$(x+y)/2 \geq \sqrt{x\,y},$$

která je přímým důsledkem nerovnosti $(x+y)^2 \geq 4xy$ neboli $0 \leq (x-y)^2 = (x+y)^2 - 4xy$, viz také Obrázek 6.44



Obrázek 6.44: Grafický důkaz nerovnosti $(x+y)^2 \geq 4xy$, ze kterého lze snadno vidět i to, že rovnost nastane právě tehdy, když $x = y$.

[74] Index $k \in \{1, \ldots, m\}$ odpovídá největší hodnotě x_k z čísel $x_1, \ldots, x_m$ a v případě několika takových indexů můžeme zvolit libovolný z nich.

Vážené průměry se obvykle užívají v *normované podobě*, kdy $n = \sum_{i=1}^m \lambda_i = 1$, ale i v *jednotkovém tvaru*, kdy $\lambda_1 = \cdots = \lambda_m = 1$, v němž lze ukázat, že průměry jsou neklesající vzhledem k řádu t, tj. pro libovolná $x_1, \dots, x_m > 0$ a $p > q$ platí $M_p \geq M_q$ neboli[75]

$$\left(\frac{x_1^p + \cdots + x_m^p}{m}\right)^{1/p} \geq \left(\frac{x_1^q + \cdots + x_m^q}{m}\right)^{1/q},$$

přičemž rovnost nastane pouze pro $x_1 = \cdots = x_m$. Zejména platí $M_2 \geq M_1 \geq M_0 \geq M_0$, jak už jsme částečně viděli v (6.2.11). Vskutku, je-li $p > q > 0$ nebo $p > 0 > q$, pak tvrzení plyne z Jensenovy nerovnosti pro funkci $f(x) := x^{p/q}$, která je pro $x > 0$ a $p/q > 1$ nebo $p/q < 0$ konvexní, viz Příklad 6.2.3. Pak totiž volbou $y_i := x_i^q$ pro $i \in \{1, \dots, m\}$ a zároveň $\lambda_1 = \cdots = \lambda_m = 1/m$ obdržíme z (6.2.9) nerovnost

$$\left(\frac{y_1 + \cdots + y_m}{m}\right)^{p/q} \leq \frac{y_1^{p/q} + \cdots + y_m^{p/q}}{m},$$

což umocněním na $1/p > 0$ dává

$$\left(\frac{x_1^q + \cdots + x_m^q}{m}\right)^{1/q} \leq \left(\frac{x_1^p + \cdots + x_m^p}{m}\right)^{1/p}.$$

Je-li $q < p < 0$, pak aplikováním předchozí nerovnosti pro $r := -q$, $s := -p$, tj. $r > s > 0$, a $y_i := 1/x_i$ dostaneme

$$\left(\frac{y_1^s + \cdots + y_m^s}{m}\right)^{1/s} \leq \left(\frac{y_1^r + \cdots + y_m^r}{m}\right)^{1/r}$$

neboli

$$\left(\frac{y_1^{-p} + \cdots + y_m^{-p}}{m}\right)^{-1/p} \leq \left(\frac{y_1^{-q} + \cdots + y_m^{-q}}{m}\right)^{-1/q},$$

a tedy (opět)

$$\left(\frac{x_1^q + \cdots + x_m^q}{m}\right)^{1/q} \leq \left(\frac{x_1^p + \cdots + x_m^p}{m}\right)^{1/p}.$$

Konečně, je-li $q = 0$, tj. $p > q = 0$, pak je nerovnost tvaru

$$\left(\frac{x_1^p + \cdots + x_m^p}{m}\right)^{1/p} \geq \sqrt[m]{x_1 \cdots x_m},$$

která ale platí, neboť z AG-nerovnosti pro x_i nahrazené $x_i^p > 0$ máme

$$\frac{x_1^p + \cdots + x_m^p}{m} \geq \sqrt[m]{x_1^p \cdots x_m^p},$$

což po umocnění na $1/p > 0$ dává požadovanou nerovnost. Konečně pro $p = 0$, tj. $0 = p > q$, máme

$$\sqrt[m]{x_1 \cdots x_m} \geq \left(\frac{x_1^q + \cdots + x_m^q}{m}\right)^{1/q},$$

která opět plyne z AG-nerovnosti pro x_i nahrazené $x_i^q > 0$ a umocněním na $1/q < 0$.

(iv) Pomocí Jensenovy nerovnosti lze dokázat i další velmi užitečný výsledek – *Hölderovu nerovnost*[76]

$$\sum_{i=1}^m |x_i\, y_i| \leq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \times \left(\sum_{i=1}^n |y_i|^q\right)^{1/q}, \tag{6.2.14}$$

kde $p, q > 1$ jsou[77] *konjugovaná čísla*, tj. $1/p + 1/q = 1$. Vskutku, uvažme pouze případ $x_1, \dots, x_m, y_1, \dots, y_m \geq 0$, neboť $|x_i\, y_i| = |x_i| \times |y_i|$, v nerovnosti tedy vystupují pouze nezáporná čísla. Vezměme bez újmy na obecnosti[78] $y_1, \dots, y_m > 0$, pak $\sum_{i=1}^m y_i^q > 0$ a

$$\left(\sum_{i=1}^m x_i\, y_i\right)^p = \left(\frac{\sum_{i=1}^m x_i\, y_i^{1-q}\, y_i^q}{\sum_{j=1}^m y_j^q}\right)^p \times \left(\sum_{j=1}^m y_j^q\right)^p.$$

[75] Bylo by možné uvážit i $x_i = 0$ pro nějaké $i \in \{1, \dots, m\}$, v kterémžto případě je ale nutné vzít $p > q > 0$.

[76] Tato nerovnost se poprvé objevuje (asi) v článku [434] z roku 1888. Inspirován tímto výsledkem podal v roce 1889 Hölder jiný důkaz v práci zaměřené na rozvoj konvexních a konkávních funkcí, viz [273]. Uvedl zde také speciální tvar Jensenovy nerovnosti, na kterou později Jensen navázal.

[77] Volbou $p = q = 2$ dostaneme Cauchyho–Schwarzovu nerovnost $|\langle x, y\rangle| \leq \|x\|\,\|y\|$.

[78] Kdyby $y_j = 0$ pouze pro jedno $j \in \{1, \dots, m\}$, pak bude nerovnost (6.2.14) splněna právě tehdy, když bude bude platit tatáž nerovnost pro $m - 1$ nenulových $y_1, \dots, y_{j-1}, y_{j+1}, \dots, y_m$ a odpovídající $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_m \geq 0$, neboť levá strana a druhý činitel na pravé straně se vynecháním x_j, y_j nezmění, zatímco první činitel na pravé straně se nezvětší. V případě, že $y_j = 0$ pro více indexů $j \in \{1, \dots, m\}$, můžeme argumentovat podobně a v případě $y_1 = \cdots = y_m = 0$ je nerovnost (6.2.14) triviální. Proto stačí dokázat platnost této nerovnosti pouze pro $y_1, \dots, y_m > 0$.

Protože $f(z) = z^p$ je konvexní pro $z \geq 0$ a $p \geq 1$ podle Příkladu 6.2.3, dostaneme aplikováním Jensenovy nerovnosti pro $z_i := x_i\, y_i^{1-q}$ a $\lambda_i := y_i^q / (\sum_{j=1}^m y_j^q)$ nerovnost

$$\left(\sum_{i=1}^m \frac{y_i^q}{\sum_{j=1}^m y_j^q}\, x_i\, y_i^{1-q}\right)^p = f\Big(\sum_{i=1}^m \lambda_i\, z_i\Big) \leq \sum_{i=1}^m \lambda_i\, f(z_i) = \sum_{i=1}^m \frac{y_i^q}{\sum_{j=1}^m y_j^q}\, x_i^p\, y_i^{(1-q)\,p},$$

tudíž

$$\Big(\sum_{i=1}^m x_i\, y_i\Big)^p = \left(\frac{\sum_{i=1}^m x_i\, y_i^{1-q}\, y_i^q}{\sum_{j=1}^m y_j^q}\right)^p \times \Big(\sum_{j=1}^m y_j^q\Big)^p \leq \left(\sum_{i=1}^m \frac{y_i^q}{\sum_{j=1}^m y_j^q}\, x_i^p\, y_i^{(1-q)\,p}\right) \times \Big(\sum_{j=1}^m y_j^q\Big)^p =$$
$$= \left(\sum_{i=1}^m x_i^p\, y_i^{(1-q)\,p+q}\right) \times \Big(\sum_{j=1}^m y_j^q\Big)^{p-1} = \Big(\sum_{i=1}^m x_i^p\Big) \times \Big(\sum_{j=1}^m y_j^q\Big)^{p-1},$$

protože z rovnosti $1/p + 1/q = 1$ plyne $p + q = pq$. Navíc také $q = p/(p-1)$, takže

$$\sum_{i=1}^m x_i\, y_i \leq \Big(\sum_{i=1}^m x_i^p\Big)^{1/p} \times \Big(\sum_{j=1}^m y_j^q\Big)^{\frac{p-1}{p}} = \Big(\sum_{i=1}^m x_i^p\Big)^{1/p} \times \Big(\sum_{j=1}^m y_j^q\Big)^{1/q},$$

což jsme chtěli dokázat. Navíc, máme-li $y_i \neq 0$ pro všechna[79] $i \in \{1,\dots,m\}$, pak rovnost nastane právě tehdy, když $z_1 = \cdots = z_m$, tj.[80]

$$x_1\, y_1^{1-q} = \cdots = x_m\, y_m^{1-q} \quad \text{neboli} \quad x_1^p\, y_1^{(1-q)\,p} = \cdots = x_m^p\, y_m^{(1-q)\,p} \quad \text{neboli} \quad \frac{x_1^p}{y_1^q} = \cdots = \frac{x_m^p}{y_m^q}.$$

▲

[79] Tento požadavek je nutný vzhledem k výše použitým argumentům, avšak je možné $x_j = 0$ pro nějaké $j \in \{1,\dots,m\}$, což však vede k tomu, že rovnost nastane pouze pokud $x_1 = \cdots = x_m = 0$.

[80] Vzpomeňme si na podmínku rovnosti v Cauchyho–Schwarzově nerovnosti.

AG-nerovnost lze velmi elegantně využít i při řešení některých optimalizačních úloh, což si ilustrujeme na následující dvojici příkladů.

Příklad 6.2.15

Určeme (bez derivování) rozměry plechovky o objemu 350 cm^3 s minimálním povrchem (pro jednoduchost ignorujme sváry a různé nepravidelnosti ve tvaru).

Řešení. Víme, že

$$V = \pi\, r^2\, h \quad \text{a} \quad Q = 2\pi\, r\, (r + h),$$

kde r je poloměr podstavy a h výška plechovky. Vzorec pro povrch plechovky lze také vyjádřit jako

$$Q = 2\pi\, r^2 + 2\pi\, r\, h = 2\pi\, r^2 + \frac{2V}{r} = \frac{1}{3}\left(6\pi\, r^2 + \frac{3V}{r} + \frac{3V}{r}\right),$$

což můžeme číst jako aritmetický průměr tří veličin ($6\pi\, r^2$ a dvakrát $3V/r$). Pak aplikováním AG-nerovnosti na tyto tři veličiny obdržíme

$$\underbrace{\sqrt[3]{6\pi\, r^2 \times \frac{3V}{r} \times \frac{3V}{r}}}_{= \sqrt[3]{2\cdot 3^3\, \pi\, V^2} = 3\sqrt[3]{2\pi\, V^2}} \leq \frac{1}{3}\left(6\pi\, r^2 + \frac{3V}{r} + \frac{3V}{r}\right) = Q,$$

přičemž rovnost nastane právě tehdy, když $2\pi\, r^2 = V/r = V/r$. Máme tedy dolní hranici pro hodnotu Q, a chceme-li Q minimalizovat, musí nastat právě zmíněné rovnosti, tj. $\pi\, r^2 = V/(2r) = \pi\, r\, h/2$ neboli $h = 2r$, kdy Q dosáhne právě oné dolní hranice. Vzorec pro objem se pak redukuje na $350 = 2\pi\, r^3$, takže $r \approx 3{,}819\,\text{cm}$, $h \approx 7{,}638\,\text{cm}$ dávají minimální povrch $Q = 3\sqrt[3]{2\pi\, V^2} \approx 274{,}916\,\text{cm}^2$.

Příklad 6.2.16

Z plechu čtvercového tvaru s délkou strany a chceme ve všech rozích odříznout stejné čtvercové díly tak, abychom ohnutím a spojením zbylé části získali krabičku maximálního objemu.

Řešení. Označíme-li obsah původního čtverce jako $S = a^2$ a délku strany odřezku jako x, pak objem krabičky bude

$$V = (a-2x)^2\, x,$$

kde $x \in (0, a/2)$. Aplikováním AG-nerovnosti získáme

$$4V = (a-2x)\,(a-2x)\,4x \le \left(\frac{(a-2x)+(a-2x)+4x}{3}\right)^3 = \frac{8a^3}{27} = \frac{8}{27}\,S^{3/2}$$

s tím, že rovnost nastane právě tehdy, když $a-2x = a-2x = 4x$. Nalezli jsme tak horní odhad pro hodnotu objemu a chceme-li jej maximalizovat, musí se uvedená nerovnost realizovat jako rovnost, tj. maximálního objemu $V = \frac{2}{27}\,a^3$ dosáhneme v případě $a = 6x$ neboli $x = a/6$.

Pochopitelně i Jensenova nerovnost může být velmi užitečná při řešení takovýchto úloh.

Příklad 6.2.17

Pro trojúhelník s vnitřními úhly α, β, γ dokažme, že

$$\frac{1}{\cos\alpha/2} + \frac{1}{\cos\beta/2} + \frac{1}{\cos\gamma/2} \ge 2\sqrt{3}.$$

Kdy zde nastane rovnost, tj. hodnota na levé straně bude minimální?
Řešení. Nejdříve poznamenejme, že funkce $f(x) = \frac{1}{\cos x}$ je na intervalu $(-\pi/2, \pi/2)$ ostře konvexní, neboť zde platí $f''(x) = \frac{1+\sin^2 x}{\cos^3 x} > 0$. Díky tomu můžeme využít Jensenovu nerovnost, která pro tuto funkci s hodnotami $x_1 = \alpha/2$, $x_2 = \beta/2$, $x_3 = \gamma/2$ patřícími do $(0, \pi/2)$ a volbou $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$ dává

$$\underbrace{\frac{1}{\cos(\alpha/6+\beta/6+\gamma/6)}}_{=\frac{1}{\cos(\pi/6)}=\sqrt{3}/2} \le \frac{1}{3\cos\alpha/2} + \frac{1}{3\cos\beta/2} + \frac{1}{3\cos\gamma/2},$$

což je požadovaná nerovnost. Navíc z ostré nerovnosti vyplývá, že rovnost nastane právě tehdy, když $\alpha = \beta = \gamma = \pi/3$.

Ještě si ukážeme příklad z ekonomie, k jehož řešení lze využít Jensenovu nerovnost.

Příklad 6.2.18

Uvažujme vyrábějící po celý rok jen jeden produkt. Výrobní náklady jsou pochopitelně závislé na míře produkce. Bude-li po část roku λ výroba nastavena tak, že bychom tímto tempem vyrobili za celý rok x ks, lze náklady vyjádřit jako $\lambda C(x)$ pro jistou funkci (přirozeně) splňující $C'(x) > 0$ (náklady rostou s rostoucím objemem produkce) a $C''(x) > 0$ na $[0, \infty)$. Má-li být celková produkce rovna Y, jak nastavit objem produkce v průběhu roku, abychom minimalizovali náklady?
Řešení. Připusťme, že rozdělíme rok na m různě dlouhých období $\lambda_1, \dots, \lambda_m$. Pak tato čísla musí být kladná a splňovat $\lambda_1 + \cdots + \lambda_m = 1$. V těchto obdobích je objem produkce postupně roven $x_1, \dots, x_m$, tj. $\lambda_1 x_1 + \cdots + \lambda_m x_m = Y$. Druhý předpoklad pro funkci $C(x)$ zaručuje, že je konvexní, a tudíž z Jensenovy nerovnosti plyne

$$C\Big(\sum_{i=1}^{m} \lambda_i x_i\Big) \le \sum_{i=1}^{m} \lambda_i C(x_i).$$

Na levé straně máme hodnotu $C(Y)$, tj. náklady při konstantním objemu výroby po celý rok, zatímco na pravé straně jsou náklady při objemech $x_1, \dots, x_m$. Naším cílem je tuto hodnotu minimalizovat, čehož docílíme tím, že si obě strany budou rovny, tj. $x_1, \dots, x_m$. Jinými slovy, v tomto případě docílíme minimálních nákladů při konstatním objemu výroby.

Jensenovu nerovnost v podobě představené ve Větě 6.2.12 lze chápat jako její *základní diskrétní* variantu, kterou lze pro konvexní funkci $f : [a,b] \to \mathbb{R}$ pomocí indukce rozšířit např. do podoby *Jensenovy–Steffensenovy nerovnosti*

$$f\left(\frac{\sum_{i=1}^m \lambda_i x_i}{\sum_{i=1}^m \lambda_i}\right) \le \frac{\sum_{i=1}^m \lambda_i f(x_i)}{\sum_{i=1}^m \lambda_i},$$

kde čísla $x_1,\dots,x_m \in [a,b]$ splňují $x_1 \ge x_2 \ge \cdots \ge x_m$ a koeficienty $\lambda_1,\dots,\lambda_m \in \mathbb{R}$ jsou takové, že pro všechna $k = 1,\dots,m-1$ platí $0 \le \sum_{i=1}^k \lambda_i \le \sum_{i=1}^m \lambda_i$ a současně $\sum_{i=1}^m \lambda_i > 0$, viz [405, Theorem 1.2.2]. S její pomocí můžeme získat i *spojitou/integrální* verzi Jensenovy nerovnosti

$$f\left(\frac{\int_a^b h(x)\,g(x)\,\mathrm{d}x}{\int_a^b h(x)\,\mathrm{d}x}\right) \le \frac{\int_a^b h(x)\,f(g(x))\,\mathrm{d}x}{\int_a^b h(x)\,\mathrm{d}x},$$

pro konvexní funkci $f : [a,b] \to \mathbb{R}$ a integrovatelné funkce $h : [a,b] \to (0,\infty)$ a $g : [a,b] \to [0,\infty)$, kdykoli všechny vystupující integrály existují. Zejména pro $h(x) \equiv 1$ dostaneme

$$f\left(\int_a^b g(x)\,\mathrm{d}x\right) \le \int_a^b f(g(x))\,\mathrm{d}x,$$

viz [405, Theorem 1.2.5]. I tato verze Jensenovy nerovnosti má své využití v ekonomii[81].

[81] Tento příklad i jeho řešení pochází z [470, str. 66].

Příklad 6.2.19

Uvažme spotřebitele, který očekává, že bude žít od teď (čas $t = 0$) do času T. Nechť $c(t)$ označuje tok spotřebních výdajů v čase t, funkce $y(t)$ popisuje tok příjmů v čase t a w_0 představuje počáteční majetek. Předpokládejme, že spotřebitel by chtěl zvolit $c(t)$ tak, aby maximalizoval *funkci celoživotního mezičasového užitku*

$$\int_0^T \mathrm{e}^{-\alpha t}\,u(c(t))\,\mathrm{d}t, \tag{6.2.15}$$

kde $\alpha > 0$ je *míra netrpělivosti* nebo *diskontní míra užitku* a $u(c)$ je rostoucí konkávní funkce užitku (jako např. $\ln c$ nebo $-c^{-2}$). Připusťme, že r je okamžitá úroková míra úspor a že spotřebiteli není dovoleno, aby v čase T byl v dluhu. Počáteční majetek společně se současnou diskontovanou hodnotou (SDH) budoucího příjmu je $w_T = w_0 + \int_0^T \mathrm{e}^{-rt}\,y(t)\,\mathrm{d}t$. *Mezičasové rozpočtové omezení* vede k podmínce, že SDH spotřeby nesmí překročit w_T, tj.

$$\int_0^T \mathrm{e}^{-rt}\,c(t)\,\mathrm{d}t \le w_T \tag{6.2.16}$$

pro všechny přípustné $c(t)$. Nalezení optimálního časového průběhu spotřeby obvykle vyžaduje nástroje z teorie optimálního řízení, avšak v tomto speciálním případě, kdy $r = \alpha$, lze úlohu snadno vyřešit pomocí Jensenovy nerovnosti. Označme jako $\bar{c}$ (konstantní) úroveň spotřeby, která splňuje rovnici

$$\int_0^T \mathrm{e}^{-rt}\,\bar{c}\,\mathrm{d}t = w_T = w_0 + \int_0^T \mathrm{e}^{-rt}\,y(t)\,\mathrm{d}t. \tag{6.2.17}$$

Všimněme si, že zejména pro $w_0 = 0$ je $y(t) \equiv \bar{c}$. Ukážeme, že optimálním průběh spotřeby je dán volbou $c(t) \equiv \bar{c}$, což se označuje jako *vyhlazování spotřeby*[82], neboť všechny fluktuace v příjmu jsou urovnány úsporami a půjčkami tak, aby byla zachována konstantní spotřeba.

K důkazu toho tvrzení položme $\tilde{\alpha} := \int_0^T \mathrm{e}^{-rt}\,\mathrm{d}t$. Pak z (6.2.17) plyne, že $\bar{c} = w_T/\tilde{\alpha}$. Nyní aplikováním integrální Jensenovy nerovnosti pro konkávní funkci u s „váhou“ $h(t) = \mathrm{e}^{-rt}/\tilde{\alpha}$ obdržíme

$$u\left(\int_0^T \frac{1}{\tilde{\alpha}}\,\mathrm{e}^{-rt}\,c(t)\,\mathrm{d}t\right) \ge \int_0^T \frac{1}{\tilde{\alpha}}\,\mathrm{e}^{-rt}\,u(c(t))\,\mathrm{d}t = \frac{1}{\tilde{\alpha}}\int_0^T \mathrm{e}^{-rt}\,u(c(t))\,\mathrm{d}t. \tag{6.2.18}$$

[82] Vyhlazování spotřeby je v ekonomii poměrně důležité. John R. Hicks definoval „příjem“ jako míru spotřeby, která může být udržena beze změny. Milton Friedman nazýval podobný rozsah jako „trvalý příjem“ a vyslovil *hypotézu trvalého příjmu*, podle níž rozsah „trvalé spotřeby“ je roven trvalému příjmu.

Nerovnosti (6.2.16) a (6.2.18) společně s rovností $\widetilde{c} = w_T/\tilde{\alpha}$ a definicí $\tilde{\alpha}$ vedou k

$$\int_0^T \mathrm{e}^{-r\,t}\,u(c(t))\,\mathrm{d}t \le \tilde{\alpha}\,u\Big(\int_0^T \frac{1}{\tilde{\alpha}}\,\mathrm{e}^{-r\,t}\,c(t)\,\mathrm{d}t\Big) \le \tilde{\alpha}\,u(w_T/\tilde{\alpha}) = \tilde{\alpha}\,u(\widetilde{c}) = \int_0^T \mathrm{e}^{-r\,t}\,u(\widetilde{c})\,\mathrm{d}t.$$

To ukazuje, že žádný jiný plán spotřeby splňující rozpočtové omezení (6.2.16) nemůže dát vyšší hodnotu celoživotního užitku daného (6.2.15), než jaký dává vyhlazený průběh spotřeby s $c(t) \equiv \widetilde{c}$.

Mnoho dalších nerovností zahrnujících konvexní funkce lze najít např. v [405, Chapter 1]. Spektrum uvedených výsledků je poměrně široké, ale výčet pochopitelně není (a nemůže být) úplný, jak vidíme např. z absence Popoviciuovy nerovnosti z roku 1965, která v nejjednodušším případě ukazuje, že pro konvexní funkci $f : [a,b] \to \mathbb{R}$ a libovolnou trojici $x, y, z \in [a,b]$ platí

$$\frac{f(x)+f(y)+f(z)}{3} + f\Big(\frac{x+y+z}{3}\Big) \ge \frac{2}{3}\Big[f\Big(\frac{x+y}{2}\Big) + f\Big(\frac{y+z}{2}\Big) + f\Big(\frac{z+x}{2}\Big)\Big], \tag{6.2.19}$$

viz [417]. Je-li funkce f spojitá, pak je konvexní právě tehdy, když platí (6.2.19) pro všechny $x, y, z \in [a,b]$. Navíc v případě ostře konvexní funkce f nastane v (6.2.19) rovnost tehdy a jen tehdy, když $x = y = z$.

Jensenovu nerovnost pro $\lambda_1 = \cdots = \lambda_m = \frac{1}{m}$ lze získat i z následujícího obecnějšího výsledku, který byl publikován v roce 1932 v [313] a který bývá spojován s různými jmény, viz [303].

Věta 6.2.20 Karamatova[83] nerovnost

Nechť na intervalu $I \subseteq \mathbb{R}$ nerostoucí posloupnost $\{x_1, \dots, x_m\}$ majorizuje nerostoucí posloupnost $\{y_1, \dots, y_m\}$, tj. $x_1, \dots, x_m, y_1, \dots, y_m \in I$ jsou taková, že platí

$$x_1 \ge x_2 \ge \cdots \ge x_m \quad \text{a} \quad y_1 \ge y_2 \ge \cdots \ge y_m,$$

pro libovolné $i \in \{1, \dots, m-1\}$ je splněna nerovnost

$$x_1 + \cdots + x_i \ge y_1 + \cdots + y_i \quad \text{a současně} \quad x_1 + \cdots + x_m = y_1 + \cdots + y_m. \tag{6.2.20}$$

Je-li funkce $f : I \to \mathbb{R}$ konvexní, pak

$$f(x_1) + \cdots + f(x_m) \ge f(y_1) + \cdots + f(y_m). \tag{6.2.21}$$

Je-li funkce f dokonce ostře konvexní, pak rovnost nastane právě tehdy, když $x_i = y_i$ pro všechna $i \in \{1, \dots, m\}$.

[83] Jovan Karamata (1. února 1902 – 14. srpna 1967) byl srbský matematik, jehož jméno je nejčastěji spojováno s tauberianskou teorií zaměřenou na metody sumaci divergentních řad a s analýzou pomalu měnících se funkcí. Je považován za jednoho z nejvlivnějších srbských matematiků 20. století a je jedním ze zakladatelů Matematického institutu Srbské akademie věd a umění v roce 1946.

Zvolíme-li nerostoucí posloupnost $\{x_1, \dots, x_m\} \subseteq I$ a označíme-li aritmetický průměr těchto čísel jako a, tj.

$$a := \frac{x_1 + \cdots + x_m}{n}$$

pak posloupnosti $\{x_1, \dots, x_m\}$ a $\{a, \dots, a\}$ splňují předpoklady Věty 6.2.20, neboť

$$x_1 + \cdots + x_i \ge i\,\frac{x_1 + \cdots + x_m}{n} \quad \text{neboli} \quad \frac{x_1 + \cdots + x_i}{i} \ge \frac{x_1 + \cdots + x_m}{n} \tag{6.2.22}$$

platí pro libovolné $i \in \{1, \dots, m-1\}$ díky tomu, že aritmetický průměr i-tice největších čísel je určitě větší nebo roven aritmetickému průměru celé m-tice čísel. Navíc pro $i = n$ v (6.2.22) zjevně nastane rovnost. Potom podle Věty 6.2.20 pro konvexní funkci f platí

$$f(x_1) + \cdots + f(x_m) \ge f(a) + \cdots + f(a) = n\,f(a),$$

což odpovídá Jensenově nerovnosti s volbou $\lambda_1 = \cdots = \lambda_m = \frac{1}{m}$.

Důkaz Věty 6.2.20. Nechť posloupnosti $\{x_1, \dots, x_m\}$ a $\{y_1, \dots, y_m\}$ splňují uvedené předpoklady. Je-li $x_i = y_i$ pro všechna $i \in \{1, \dots, m\}$, pak je tvrzení triviální. Nechť proto $x_i \ne y_i$ alespoň pro

jeden index $i \in \{1,\dots,m\}$. Navíc je-li $x_i = y_i$ pro nějaké $i \in \{1,\dots,m\}$, pak se žádná z vlastností nezmění, pokud tyto členy vynecháme. Proto můžeme bez újmy na obecnosti předpokládat, že $x_i \neq y_i$ pro všechna $i \in \{1,\dots,m\}$. Z konvexnosti funkce f plyne

$$c_{i+1} := \frac{f(x_{i+1}) - f(y_{i+1})}{x_{i+1} - y_{i+1}} \leq \frac{f(x_i) - f(y_i)}{x_i - y_i} =: c_i \tag{6.2.23}$$

pro všechna $i \in \{1,\dots,m-1\}$, viz (6.2.3). Označíme-li

$$\alpha_i := x_1 + \cdots + x_i \quad \text{a} \quad \beta_i := y_1 + \cdots + y_i$$

pro $i \in \{1,\dots,m\}$ a $\alpha_0 := 0 =: \beta_0$, pak z (6.2.20) vidíme, že $\alpha_i \geq \beta_i$ pro $i \in \{1,\dots,m-1\}$ a současně $\alpha_m = \beta_m$. Proto

$$\begin{aligned}\sum_{i=1}^{m} \big(f(x_i) - f(y_i)\big) &\overset{(6.2.23)}{\geq} \sum_{i=1}^{m} c_i\,(x_i - y_i) = \sum_{i=1}^{m} c_i \big[\underbrace{\alpha_i - \alpha_{i-1}}_{x_i} - \underbrace{(\beta_i - \beta_{i-1})}_{y_i}\big] = \\ &= \sum_{i=1}^{m} c_i\,(\alpha_i - \beta_i) - \sum_{i=1}^{m} c_i\,(\alpha_{i-1} - \beta_{i-1}) = \\ &= c_m \underbrace{(\alpha_m - \beta_m)}_{=0} + \sum_{i=1}^{m-1} \underbrace{(c_i - c_{i+1})}_{\geq 0} \underbrace{(\alpha_i - \beta_i)}_{\geq 0} - c_1\,(\alpha_0 - \beta_0) \geq 0,\end{aligned} \tag{6.2.24}$$

což dokazuje požadovanou nerovnost.

Z našeho předpokladu $x_i \neq y_i$ pro všechna $i \in \{1,\dots,m-1\}$ plyne, že $x_1 > y_1$. Díky tomu a druhé podmínce v (6.2.20) musí existovat index i takový, že $(x_i, y_i) \neq (x_{i+1}, y_{i+1})$. Nechť j označuje nejmenší takový index. Pak $A_j > B_j$. Z ostré konvexnosti současně plyne, že $c_{j+1} < c_j$, a tedy na levé straně nerovnosti (6.2.24) je alespoň jeden kladný člen, což znamená, že se Karamatova nerovnost nemůže realizovat jako rovnost. ■

Poznámka 6.2.21. Je-li funkce f konvexní a neklesající, může být druhá podmínka v (6.2.20) nahrazena pouhou nerovností $x_1 + \cdots + x_m \geq y_1 + \cdots + y_m$. V takovém případě totiž máme zaručeno, že $c_m \geq 0$ a současně $\alpha_m \geq \beta_m$, což pro platnost nerovnosti (6.2.24) stačí. Je-li f navíc ostře konvexní, pak $c_m > 0$ a oproti předchozímu důkazu musíme navíc uvážit případ, kdy $\alpha_m > \beta_m$. V takovém případě ale na levé straně nerovnosti (6.2.24) očividně alespoň jeden kladný člen, a tedy Karamatova nerovnost se opět nemůže realizovat jako rovnost. ▲

Příklad 6.2.22

Nechť $a_1,\dots,a_m \in \mathbb{R}$ jsou kladná čísla. Dokažme, že

$$(1+a_1)(1+a_2)\cdots(1+a_m) \leq \left(1+\frac{a_1^2}{a_2}\right)\left(1+\frac{a_2^2}{a_3}\right)\cdots\left(1+\frac{a_m^2}{a_1}\right),$$

přičemž rovnost nastane právě tehdy, když $a_1 = \cdots = a_m$.

Řešení. Jelikož díky monotónii funkce $\ln x$ je uvedená nerovnost ekvivalentní s

$$\ln(1+a_1)+\ln(1+a_2)+\cdots+\ln(1+a_m) \leq \ln\left(1+\frac{a_1^2}{a_2}\right)+\ln\left(1+\frac{a_2^2}{a_3}\right)+\cdots+\ln\left(1+\frac{a_m^2}{a_1}\right), \tag{6.2.25}$$

můžeme naši snahu namířit na důkaz této nerovnosti pomocí Karamatova výsledku. Musíme tedy najít vhodnou dvojici posloupností a konvexní funkci, neboť $\ln x$ je konkávní (což by nemuselo vadit), ale především posloupnosti $\{1+a_i\}_{i=1}^m$ a $\{1+a_i^2/a_{i+1}\}_{i=1}^m$ pro $a_{m+1} := a_1$ nesplňují (6.2.20). Bez újmy na obecnosti můžeme předpokládat, že čísla $a_1,\dots,a_m$ tvoří nerostoucí posloupnost. Pak posloupnost $\{y_i\}_{i=1}^m$ se členy $y_i := \ln a_i$ je také nerostoucí. Zvolíme-li druhou posloupnost $\{x_i\}_{i=1}^m$ jako $x_i := 2\ln a_i - \ln a_{i+1}$ pro $i \in \{1,\dots,n\}$ s $a_{m+1} := a_1$, pak i tato posloupnost je nerostoucí a společně s první po-

sloupností splňují podmínky (6.2.20), neboť

$$\begin{aligned} x_1 + x_2 + \cdots + x_i &= 2\ln a_1 - \ln a_2 + 2\ln a_2 - \ln a_3 + \cdots + 2\ln a_i - \ln a_{i+1} = \\ &= 2\ln a_1 + \ln a_2 + \cdots + \ln a_i - \ln a_{i+1} = \\ &= \ln a_1 + \ln a_2 + \cdots + \ln a_i + \underbrace{(\ln a_1 - \ln a_{i+1})}_{\geq 0} \geq \\ &\geq \ln a_1 + \ln a_2 + \cdots + \ln a_i = y_1 + y_2 + \cdots + y_i \end{aligned}$$

pro libovolné $i \in \{1, \ldots, m-1\}$ a současně

$$\begin{aligned} x_1 + x_2 + \cdots + x_m &= 2\ln a_1 - \ln a_2 + 2\ln a_2 - \ln a_3 + \cdots + 2\ln a_m - \ln a_1 = \\ &= \ln a_1 + \cdots + \ln a_m = y_1 + \cdots + y_m. \end{aligned}$$

Zbývá určit vhodnou funkci f. Z právě definovaných posloupností lze vidět, že hodnoty na levé a pravé straně nerovnosti (6.2.25) dostaneme aplikujeme-li na čísla $x_1, \ldots, x_m, y_1, \ldots, y_m$ funkci $f(x) = \ln(1 + e^x)$. Jelikož $f''(x) = \frac{e^x}{(1+e^x)^2} > 0$, je funkce f ostře konvexní, a tak uvedená nerovnost skutečně plyne z Karamatovy nerovnosti včetně rovnosti tehdy a jen tehdy, když $a_1 = \cdots = a_m$.

Pro náš optimalizační výklad však budou nejdůležitější především následující tvrzení o extrémech konvexních funkcí a jejich důsledky.

Věta 6.2.23

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ konvexní. Potom následující tvrzení jsou pravdivá.

(i) Libovolné lokální minimum funkce f na X je současně globálním minimem.

(ii) Množina bodů množiny X, v nichž funkce f nabývá svého minima na X, je konvexní[84]. Je-li funkce f dokonce ostře konvexní, pak je tato množina nejvýše jednoprvková.

(iii) Je-li funkce f diferencovatelná na otevřené množině $\mathscr{U} \supseteq X$ a $x^* \in X$ je jejím stacionárním bodem, tj. $\operatorname{grad} f(x^*) = 0$, pak x^* je bodem globálního minima funkce f na množině X.

[84] Nezapomínejme, že i prázdná množina je konvexní.

Poznámka 6.2.24.

(i) Z části (ii) předchozí věty plyne, že pokud f nabývá minimum ve dvou bodech, pak totéž nutně platí i v libovolném bodě na úsečce spojující tyto body, tj. funkce f je konstantní na takovýchto úsečkách s minimální funkční hodnotou. Proto množina minim funkce f je buď prázdná, jednoprvková nebo nespočetná.

(ii) V části (iii) předchozí věty rozhodně neplatí ekvivalence, jelikož jsme omezeni množinou X a minimum může nastat i na hranici této množiny, tj. konvexní funkce f může mít lokální/globální minimum v bodě x^*, v němž $\operatorname{grad} f(x^*) \neq 0$. Stačí uvážit např. $f(x) = x^2$ a $X = [1,2]$. Pak minimum nastane pro $x = 1$, zatímco $f'(1) = 2 \neq 0$. Ovšem pokud bude množina X otevřená (což nikterak nesouvisí s její konvexností), pak již ekvivalence bude platit. Vrátíme-li se k předchozímu jednoduchému příkladu s $f(x) = x^2$, pak na otevřené množině $X = (1,2)$ neexistuje ani stacionární bod ani lokální/globální minimum[85]. ▲

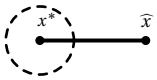
[85] Teď existuje pouze infimum f na X s hodnotou $\inf_{x \in (1,2)} x^2 = 1$.

Důkaz Věty 6.2.23.

(i) Důkaz této části provedeme sporem. Nechť x^* je lokální minimum funkce f na X, tj. existuje $\mathcal{O}(x^*)$ takové, že pro všechna $x \in \mathcal{O}(x^*) \cap X$ platí $f(x) \geq f(x^*)$. Nechť ale současně x^* není globálním minimem funkce f na X, tj. existuje $\widehat{x} \in X$ takové, že $f(\widehat{x}) < f(x^*)$. Pak z konvexnosti funkce f a množiny X plyne

$$f(\underbrace{\lambda x^* + (1-\lambda)\widehat{x}}_{\in X}) \leq \lambda f(x^*) + (1-\lambda) f(\widehat{x}) < \lambda f(x^*) + (1-\lambda) f(x^*) = f(x^*)$$

pro všechny $\lambda \in [0,1)$. To je ale spor s tím, že x^* je lokálním minimem ↯, neboť určitě existuje λ dostatečně blízko 1 takové, že $\lambda x^* + (1-\lambda)\widehat{x} \in \mathcal{O}(x^*) \cap X$, viz Obrázek 6.45.



Obrázek 6.45: Je-li v bodě x^* lokální minimum f na X, pak pro hodnoty λ dostatečně blízko 1 zleva musí být odpovídající část úsečky $\lambda x^* + (1-\lambda)\widehat{x}$ v okolí $\mathcal{O}(x^*) \cap X$ určujícím lokální minimum. Jenže na této úsečce platí $f(\lambda x^* + (1-\lambda)\widehat{x}) < f(x^*)$, což je spor s definicí lokálního minima v x^* ↯.

(ii) Označme jako X^* množinu všech bodů, ve kterých funkce f nabývá svého minima na X, tj. $X^* := \{x \in X \mid f(x) = f^* = \inf_{x\in X} f(x)\}$. Je-li množina X^* prázdná nebo jednoprvková, je tvrzení triviální. Nechť tedy X^* je alespoň dvouprvková, tj. máme různé $x_1, x_2 \in X^*$. Pak pro libovolné $\lambda \in [0,1]$ platí

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) = f^*. \tag{6.2.26}$$

Jenže z definice f^* vyplývá, že také $f^* \le f(\lambda x + (1-\lambda)y)$ pro libovolné $x, y \in X$ a $\lambda \in [0,1]$. To znamená, že v (6.2.26) nastává rovnost, tj. $\lambda x_1 + (1-\lambda)x_2 \in X^*$ pro libovolné $\lambda \in [0,1]$, a tedy množina X^* je konvexní.

Je-li navíc funkce f ostře konvexní, pak pro $\lambda \in (0,1)$ nastane v (6.2.26) rovnost právě tehdy, když $x_1 = x_2$, tj. množina X^* je nejvýše jednoprvková.

(iii) Tato část tvrzení je sice přímým důsledkem Věty 6.4.3 (viz Poznámku 6.4.5(i) později), ale my si ji teď dokážeme přímo. Nechť $x_1, x_2 \in X$ a $\lambda \in [0,1]$ jsou libovolné. Pak z konvexnosti funkce f na X máme

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) \tag{6.2.27}$$

a současně z diferencovatelnosti[86] plyne

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &= f(x_2 + \lambda(x_1 - x_2)) \\ &= f(x_2) + \langle \operatorname{grad} f(x_2), \lambda(x_1 - x_2)\rangle + \omega(\lambda(x_1 - x_2)), \end{aligned} \tag{6.2.28}$$

kde $\lim_{\|h\|\to 0^+} \frac{\omega(h)}{\|h\|} = 0$. Volbou $x_1 = x$, $x_2 = x^*$ a kombinací (6.2.27) a (6.2.28) dostáváme pro $\lambda \in (0,1]$ nerovnost

$$\begin{aligned} f(x) - f(x^*) &\overset{(6.2.27)}{\ge} \frac{1}{\lambda}\big[f(\lambda x + (1-\lambda)x^*) - f(x^*)\big] = \frac{1}{\lambda}\big[f(x^* + \lambda(x - x^*)) - f(x^*)\big] \\ &\overset{(6.2.28)}{=} \frac{1}{\lambda}\big[\langle \operatorname{grad} f(x^*), \lambda(x - x^*)\rangle + \omega(\lambda(x - x^*))\big] \\ &\overset{\operatorname{grad} f(x^*)=0}{=} \frac{\omega(\lambda(x - x^*))}{\lambda}. \end{aligned} \tag{6.2.29}$$

Je-li $x = x^*$, pak zjevně

$$f(x) - f(x^*) \ge 0.$$

Totéž dostaneme v (6.2.29) limitním přechodem pro $\lambda \to 0^+$, neboť

$$\lim_{\lambda\to 0^+} \frac{\omega(\lambda(x-x^*))}{\lambda} = \lim_{\lambda\to 0^+} \underbrace{\frac{\omega(\lambda(x-x^*))}{\lambda\|x-x^*\|}}_{\to 0} \underbrace{\frac{\|x-x^*\|}{1}}_{\text{ohraničené}}.$$

Tedy pro libovolné $x \in X$ platí $f(x) - f(x^*) \ge 0$, tj. x^* je skutečně bodem globálního minima funkce f na X. ■

[86] Připomeňme, že $f : \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce v bodě x, jestliže je definovaná na nějakém jeho okolí a existuje-li lineární funkce $\mathrm{d}f(h)$ proměnné $h \in \mathbb{R}^n$ taková, že

$$f(x+h) = f(x) + \mathrm{d}f(h) + \omega(h)$$

pro nějakou funkci $\omega : \mathbb{R}^n \to \mathbb{R}$ splňující $\omega \in o(h)$ neboli takovou, že $\lim_{h\to 0} \omega(h)/\|h\| = 0$. V takovém případě pak dokonce platí

$$\mathrm{d}f(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} h_i = \langle \operatorname{grad} f(x), h\rangle.$$

Toto je splněno zejména, pokud funkce f má spojité všechny parciální derivace prvního řádu v x.

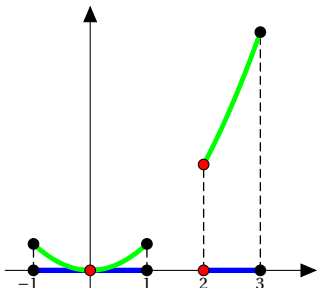
Předchozí věta však neříká vůbec nic o samotné existenci lokálních/globálních minim. Na to je samotná konvexnost „krátká“. Avšak lze toho docílit zkombinováním části (iii) předchozího tvrzení a Weierstrassovy věty (viz Věta 1.2.11).

Důsledek 6.2.25

Je-li $f : X \to \mathbb{R}$ (ostře) konvexní a *spojitá*[87] funkce na konvexní a kompaktní množině $X \subseteq \mathbb{R}^n$, pak f má na X (právě jedno) globální minimum.

[87] Později uvidíme ve Větě 6.4.1, že v případě konvexnosti f na otevřené nadmnožině $\mathcal{U} \supset X$ lze požadavek spojitosti na X vynechat – zejména je-li f konvexní na $\mathbb{R}^n$.

Poznámka 6.2.26. A co kdybychom chtěli minimalizovat konvexní funkci $f : X \to \mathbb{R}$ na nekonvexní množině $Y \subseteq X$? Jednoduchá odpověď zní, že „nevíme nic“. Řešitelnost takového problému dokáže rozlousknout pouze Weierstrassova věta a konvexnost funkce už zde není

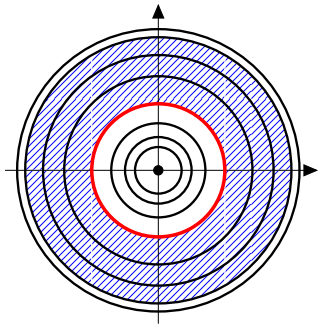


Obrázek 6.46: Graf funkce $f(x) = x^2$ na $Y = [-1,1] \cup [2,3]$ a její globální/lokální minima.

nikterak užitečná. Samozřejmě pokud množina Y obsahuje minimum funkce f na celém X, tak je to snadné. Ovšem v opačném případě už mohou nastat i jiné situace než ty, které jsou popsané ve Větě 6.2.23. Např. pro konvexní funkci $f(x) = x^2$ na množině $Y = [-1,1] \cup [2,3]$ máme globální minimum v bodě $x = 0$ a navíc i lokální minimum pro $x = 2$, což není v souladu s tvrzením Věty 6.2.23(i), viz Obrázek 6.46. Podobně pro ostře konvexní funkci $f(x_1, x_2) = x_1^2 + x_2^2$ na množině $Y = \{[x_1, x_2] \in \mathbb{R}^2 \mid 4 \le x_1^2 + x^2 \le 9\}$ nastane minimum na celé kružnici $\{[x_1, x_2] \in \mathbb{R}^2 \mid x_1^2 + x^2 = 4\}$, což zase nekoresponduje s tvrzením Věty 6.2.23(ii), viz Obrázek 6.47. ▲

Z Věty 6.2.23(ii) vidíme rozdíl mezi konvexní a ostře konvexní funkcí v možném počtu bodů, v nichž nastává globální minimum. Ale co silně konvexní funkce – můžeme říct i něco víc? K odpovědi budeme potřebovat následující lemma.

LEMMA 6.2.27



Obrázek 6.47: Vrstevnice funkce $f(x_1, x_2) = x_1^2 + x_2^2$, množina $Y = \{[x_1, x_2] \in \mathbb{R}^2 \mid 4 \le x_1^2 + x^2 \le 9\}$ (modrou barvou) a globální minima f na Y (červená kružnice).

Je-li funkce $f : X \to \mathbb{R}$ spojitá a silně konvexní s konstantou silné konvexnosti $\vartheta > 0$ na uzavřené konvexní množině X, pak pro libovolné $K \in \mathbb{R}$ jsou dolní obrysové množiny V_K definované v (6.2.7) ohraničené, a tudíž kompaktní[88,89,90].

Důkaz. Nechť $K \in \mathbb{R}$ je libovolné. Je-li $V_K = \emptyset$, pak je tvrzení triviální[91]. Připusťme tedy, že existuje $x^* \in V_K$. Protože funkce f je spojitá a množina X uzavřená existuje podle Weierstrassovy věty (Věta 1.2.11) konstanta L taková, že $f(x) \ge L$ pro všechna $x \in B_1(x^*) \cap X$, kde $B_1(x^*)$ značí uzavřenou kouli se středem v x^* a poloměrem 1. Ukážeme, že

$$\|x - x^*\| \le 1 + \frac{K - L}{\vartheta} \quad \text{pro všechna } x \in V_K, \tag{6.2.30}$$

z čehož ihned vyplyne ohraničenost množiny V_K. Pokud $x \in V_K \cap B_1(x^*)$ je libovolné, tj. $\|x - x^*\| \le 1$, pak (6.2.30) jistě platí. Zbývá uvážit $x \in V_K \setminus B_1(x^*)$. Položme $\lambda := 1/\|x - x^*\|$ a $\widehat{x} := \lambda x + (1-\lambda)x^*$. Pak $\lambda \in (0,1)$ a $\|\widehat{x} - x^*\| = \lambda \|x - x^*\| = 1$, takže díky konvexnosti V_k z Věty 6.2.10 máme $\widehat{x} \in V_K \cap B_1(x^*)$, viz Obrázek 6.49. Proto ze silné konvexnosti funkce f a nerovností $f(x^*) \le K$ a $f(\widehat{x}) \le K$ dostaneme

$$L \le f(\widehat{x}) \le \overbrace{\lambda f(x) + (1-\lambda) f(x^*)}^{\le \lambda K + (1-\lambda) K} - \vartheta \lambda (1-\lambda) \|x - x^*\|^2 \le$$
$$\le K - \vartheta \underbrace{\lambda(1-\lambda)}_{\frac{1}{\|x-x^*\|} \times \frac{\|x-x^*\|-1}{\|x-x^*\|}} \|x - x^*\|^2 = K - \vartheta\big(\|x - x^*\| - 1\big),$$

z čehož už snadno získáme (6.2.30). Jelikož $K \in \mathbb{R}$ bylo zvoleno libovolně, je důkaz hotov. ■

Zkombinujeme-li předchozí lemma s Důsledkem 1.2.12 a Větou 6.2.23(ii) dostaneme následující tvrzení ukazující výjimečnost silně konvexních funkcí.

VĚTA 6.2.28

Nechť funkce $f : X \to \mathbb{R}$ je spojitá a silně konvexní s konstantou silné konvexnosti $\vartheta > 0$ na uzavřené konvexní množině X. Pak existuje jediné globální minimum funkce f na množině X.

Jednotlivé koncepty konvexnosti se tedy podstatně liší i s ohledem na existenci a jednoznačnost globálního minima – v případě ostře konvexní funkce nemáme zaručenu jeho existenci[92] a pro „pouhé“ konvexní funkce postrádáme i jednoznačnost[93]. Proto by se chtělo říci

silně konvexní funkce jsou ideálním objektem pro minimalizační úlohy!

Avšak to bohužel obvykle není/nebude ten případ, který zrovna máme vyřešit... Naštěstí ale silná konvexnost není nutnou podmínkou pro existenci jediného globálního minima[94].

Věta 6.2.28 samozřejmě nevyžaduje, aby samotná množina X byla ohraničená. Proto můžeme vzít dokonce $X = \mathbb{R}^n$, v kterémžto případě dostaneme pro funkci $f : \mathbb{R}^n \to \mathbb{R}$ následující tvrzení:

(i) je-li f konvexní na $\mathbb{R}^n$ a v $x^* \in \mathbb{R}^n$ nastává lokální minimum, pak x^* je globální minimum;

[88] Je funkce $f(x) = e^x$ silně konvexní na $\mathbb{R}$? Toto tvrzení ukazuje, že nikoliv.

[89] Dolní obrysové množiny nemusí být v případě obecné funkce $f : Y \to \mathbb{R}^*$ uzavřené. Pokud všechny množiny V_K jsou uzavřené nebo ekvivalentně všechny množiny $\{x \in Y \mid f(x) > K\}$ jsou otevřené, nazývá se funkce f *zdola polospojitá*. V klasickém $\varepsilon - \delta$ jazyce lze polospojitost zdola v bodě $x^* \in Y$ formulovat takto: pro každé $\varepsilon > 0$ existuje okolí $\mathcal{O}(x^*)$ takové, že $f(x) > f(x^*) - \varepsilon$ pro všechna $x \in \mathcal{O}(x^*) \cap Y$, tj. $\liminf_{x \to x^*} f(x) \ge f(x^*)$, viz Obrázek 6.48. Totéž platí i pro nadgraf funkce f. Vlastně všechna tyto tři vlastnosti jsou úzce propojeny, neboť platí: *funkce f je zdola polospojitá právě tehdy, když všechny dolní obrysové množiny jsou uzavřené, což se stane právě tehdy, když množina* epi f *je uzavřená.*

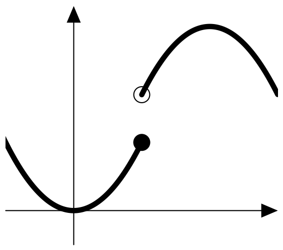
(ii) je-li f ostře konvexní na $\mathbb{R}^n$ a v $x^* \in \mathbb{R}^n$ nastává lokální minimum, pak x^* je jediné globální minimum;

(iii) je-li f silně konvexní na $\mathbb{R}^n$, pak *existuje jediné globální minimum.*

Výše zmíněné příklady také ukazují, že tvrzení Lemma 6.2.27 neplatí pro konvexní či ostře konvexní funkce. Správná analogie Lemma 6.2.27 má pro tyto funkce následující podobu.

LEMMA 6.2.29

Nechť funkce $f : X \to \mathbb{R}$ je spojitá a konvexní na uzavřené konvexní množině X. Pokud pro nějaké $K^* \in \mathbb{R}$ je množina V_{K^*} neprázdná a ohraničená, pak množina V_K je ohraničená pro libovolné $K \in \mathbb{R}$.



Obrázek 6.48: Ukázka zdola polospojité funkce.

Důkaz. Je-li $K < K^*$, pak $V_K \subseteq V_{K^*}$ a ohraničenost V_K je zřejmá. Nechť tedy $K > K^*$ a připusťme, že V_K není ohraničená. Vezmeme-li $x^* \in V_{K^*} \subsetneq V_K$ libovolné, pak díky konvexnosti (Věta 6.2.10) a neohraničenosti V_K musí existovat vektor $h \in \mathbb{R}^n \setminus \{0\}$ takový, že $x^* + \alpha h \in V_K$ pro všechna $\alpha \geq 0$, tj.

$$x^* + \alpha h \in X \quad \text{a současně} \quad f(x^* + \alpha h) \leq K \quad \text{pro všechna } \alpha \geq 0. \tag{6.2.31}$$

Definujme funkci $\varphi(\alpha) := f(x^* + \alpha h)$ pro $\alpha \geq 0$ a připusťme, že $\varphi(\alpha^*) > \varphi(0) = f(x^*)$ pro nějaké $\alpha^* > 0$. Funkce $\varphi(\alpha)$ je konvexní[95] na $[0, \infty)$, takže pro libovolné $\alpha > \alpha^*$ máme

$$\varphi(\alpha^*) = \varphi\left(\left(1 - \frac{\alpha^*}{\alpha}\right)0 + \frac{\alpha^*}{\alpha}\alpha\right) \leq \left(1 - \frac{\alpha^*}{\alpha}\right)\varphi(0) + \frac{\alpha^*}{\alpha}\varphi(\alpha)$$

neboli

$$\varphi(\alpha) \geq \frac{\alpha}{\alpha^*}\big(\varphi(\alpha^*) - \varphi(0)\big) + \varphi(0).$$

Jelikož $\varphi(\alpha^*) > \varphi(0)$, musí $\varphi(\alpha) \to \infty$ pro $\alpha \to \infty$, což je však v rozporu s (6.2.31). Proto nutně

$$f(x^* + \alpha h) = \varphi(\alpha) \leq \varphi(0) = f(x^*) \leq K^*,$$

tj. $x^* + \alpha h \in V_{K^*}$ pro všechna $\alpha \geq 0$, což je ale opět spor – tentokrát s ohraničeností V_{K^*}. Proto množina V_K musí být ohraničená pro všechna $K \in \mathbb{R}$. ■

A co kdybychom chtěli maximalizovat konvexní funkci?! Může se to zdát jako poměrně nezajímavá otázka, ale na závěr se z toho „vyklube" tvrzení, které by mohlo být velmi užitečné pro lineární programování. Nicméně my jsme jej spolu s dalšími důležitými informacemi získali v Kapitole 2 i bez znalosti vlastností konvexních funkcí[96], viz Větu 2.5.11.

VĚTA 6.2.30

Nechť $f : X \to \mathbb{R}$ je konvexní funkce na neprázdné konvexní množině $X \subseteq \mathbb{R}^n$. Jestliže f nabývá svého maxima na množině X v bodě $x^* \in \operatorname{ri} X$, pak f je konstantní funkce, tj. nekonstantní konvexní funkce může nabývat svého globálního maxima pouze v bodech množiny $X \setminus \operatorname{ri} X = \operatorname{r\partial} X \cap X$.[97]

Důkaz. Nechť funkce f nabývá svého maxima na X v bodě $x^* \in \operatorname{ri} X$ a připusťme, že existuje bod $x \in X \setminus \{x^*\}$ takový, že $f(x) < f(x^*)$. Pak podle Věty 6.1.30 existuje $\gamma > 1$ takové, že

$$\widehat{x} := x^* + (\gamma - 1)(x^* - x) = \gamma x^* - (\gamma - 1)x \in X \quad \text{neboli} \quad x^* = \frac{1}{\gamma}\widehat{x} + \frac{\gamma - 1}{\gamma}x.$$

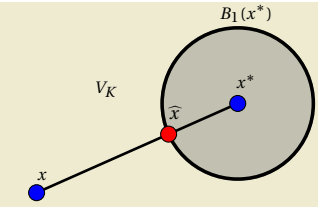
Jelikož $\frac{1}{\gamma}, \frac{\gamma-1}{\gamma} \in (0,1)$ a $\frac{1}{\gamma} + \frac{\gamma-1}{\gamma} = 1$, plyne z konvexnosti f, že´

$$f(x^*) \leq \frac{1}{\gamma}\underbrace{f(\widehat{x})}_{\leq f(x^*)} + \frac{\gamma - 1}{\gamma}\underbrace{f(x)}_{< f(x^*)} < f(x^*),$$

což je spor. neboť $f(x^*) \geq f(x)$ a $f(x^*) \geq f(\widehat{x})$. plyne odtud $f(x) = f(x^*)$, což vzhledem k libovolnosti volby $x \in X$ ukazuje, že f je na X konstantní. ■

[90] Toto také velmi úzce souvisí s tzv. koercivními funkcemi, tj. funkcemi $f : \mathbb{R}^n \to \mathbb{R}^*$ splňujícími $f(x) \to \infty$ pro $\|x\| \to \infty$. Pozor: např. funkce $f(x) = x$ není koercivní! Naopak, libovolná kvadratická funkce $f(x) = x^\top A x + b^\top x + c$ je koercivní právě tehdy, když $A > 0$, viz [38, Lemma 2.42]. Platí, že je-li množina V_K ohraničená pro $K \in \mathbb{N}$, pak f je koercivní. na druhou stranu, je-li f koercivní, pak množiny V_K jsou ohraničené pro všechna $K \in \mathbb{R}$. A jaký je vztah mezi koercivní a konvexními funkcemi? *Mějme funkci $f : X \to \mathbb{R}^*$ splňující $f(x) > -\infty$ a $f \not\equiv \infty$ (tzv. vlastní funkce /angl. proper function/) na konvexní množině X, která je zde navíc zdola polospojitá a konvexní. Jestliže f je dokonce silně konvexní, pak je také koercivní.* Viz [32, Corollary 11.16]. A co minimalizace koercivní funkce? *Je-li $f : \mathbb{R}^n \to \mathbb{R}$ spojitá a koercivní funkce a $X \subseteq \mathbb{R}^n$ neprázdná uzavřená množina, pak f nabývá globálního minima na X*, viz [38, Theorem 2.32]. Uvedené podmínky splňuje např. funkce $f(x_1, x_2) = x_1^2 + x_2^2$ a množina $X = \{[x_1, x_2 \in \mathbb{R}] \mid x_1 + x_2 \leq -1\}$. Množina X není ohraničená, a tudíž nelze využít Weierstrassovu větu. Avšak koercivnost f a uzavřenost X zaručují existenci globálního minima. S trochou námahy můžeme ověřit, že to nastává v bodě $[-1/2, -1/2]$.

[91] Toto se ale nemůže stát pro všechna K, pokud $X \neq \emptyset$, tj. alespoň pro jedno $K \in \mathbb{R}$ je množina $V_K \neq \emptyset$. Kdyby totiž $V_K = \emptyset$ neboli $f(x) \not\leq K$ pro všechna K, pak by muselo být $f \equiv \infty$, což v našich úvahách nedovolujeme.



Obrázek 6.49: Ilustrace polohy bodů x, x^* a $\widehat{x}$.

[92] Např. $f(x) = \mathrm{e}^x$ na $\mathbb{R}$.

[93] Např. $f(x) = \max\{x, 0\}$ na $\mathbb{R}$.

Věta 6.2.31 Základní věta konvexního programování

[94] Např. konvexní funkce $f(x) = x$ nebo ostře konvexní $f(x) = x^4$ nabývají na $X = [-1,1]$ jediného globálního minima, ačkoli nejsou silně konvexní.

[95] Přímým výpočtem s využitím konvexnosti f zjistíme, že pro libovolné $\lambda \in [0,1]$ a $\alpha, \beta \geq 0$ platí

$$\begin{aligned}
&\varphi\big(\lambda\,\alpha + (1-\lambda)\,\beta\big) = \\
&= f\big(x^* + (\lambda\,\alpha + (1-\lambda)\,\beta)\,h\big) = \\
&= f\big(\lambda\,(x^* + \alpha\,h) + (1-\lambda)(x^* + \beta\,h)\big) \leq \\
&\leq \lambda\,f(x^* + \alpha\,h) + (1-\lambda)\,f(x^* + \beta\,h) = \\
&= \lambda\,\varphi(\alpha) + (1-\lambda)\,\varphi(\beta).
\end{aligned}$$

Důsledek 6.2.32 Základní věta lineárního programováni

[96] Což je zcela v souladu i s historickým vývojem.

[97] Zejména je-li množina X otevřená, pak f svého maxima na X nenabývá.

[98] Toto tvrzení lze ale i zobecnit. Platí totiž: *Množina bodů, ve kterých konvexní funkce $f : X \to \mathbb{R}$ nabývá svého maxima na konvexní a kompaktní množině X, je buď prázdná nebo obsahuje extrémní bod* (viz Poznámku 6.1.18(iii))

Předchozí tvrzení však nic neříká o charakteru množiny bodů, v nichž nastává globální maximum. Podobně jako v případě minima, tato množina může být prázdná i v případě kompaktní množiny X. Stačí uvážit např. funkci $f : B_1(0) \to \mathbb{R}$ definovanou na uzavřeném jednotkovém kruhu se středem v počátku, která na vnitřku kruhu nabývá nulovou hodnotu a na hranici kruhu zvolit funkční hodnoty nezáporně ale libovolně, aby zde neexistovalo maximum. Toho lze docílit kupříkladu volbou

$$f(x) := \begin{cases} 0, & \|x\| < 1, \\ 1/\vartheta, & \|x\| = 1, \\ 0, & \vartheta = 0 \ \ \& \ \ \|x\| = 1, \end{cases}$$

kde $\vartheta \in [0, 2\pi]$ je úhel z vyjádření bodu x v polární soustavě souřadnic. Toto však není možné pro funkci jedné proměnné $f : [a,b] \to \mathbb{R}$, která nabývá svého maxima v některém z krajních bodů intervalu $[a,b]$. Ale mohou to být i oba jako v případě funkce $f(x) = x^2$ na $X = [-1,1]$, která nabývá maxima právě v obou bodech ± 1, což jsou jediné prvky relativní hranice. To ale zjevně není konvexní množina jako v případě minima. Navíc, pokud funkce f na X svého maxima nabývá, mají některé body z relativní hranice jistou výhodu. Zejména, je-li X polytop[98]

> Množina bodů, ve kterých konvexní funkce $f : X \to \mathbb{R}$ nabývá svého maxima na polytopu $X := \mathrm{conv}\{x^{[1]}, \ldots, x^{[m]}\} \subseteq \mathbb{R}^n$, je buď prázdná nebo obsahuje alespoň jeden z vrcholů X, tj. z bodů $x^{[1]}, \ldots, x^{[m]}$.

Důkaz. Případ, kdy f svého maxima na polytopu X nenabývá, získáme jednoduchou modifikací příkladu uvedeného výše. Omezme se proto jen na situaci, kdy f svého maxima na X nabývá. Nechť se to stane v nějakém bodě $x^* \in X$ a připusťme, že se to nestane v žádném z vrcholů, tj. $f(x^*) > f(x^{[i]})$ pro všechna $i = 1, \ldots, m$. Jelikož dle definice máme $x^* = \sum_{i=1}^m \lambda_i\, x^{[i]}$ pro vhodná[99] $\lambda_1, \ldots \lambda_m \in [0,1]$ splňující $\sum_{i=1}^m \lambda_i = 1$, vyplývá z Jensenovy nerovnosti

$$f(x^*) \leq \sum_{i=1}^m \lambda_i\, f\big(x^{[i]}\big) \leq \max\big\{f(x^{[1]}), \ldots, f(x^{[m]})\big\} \sum_{i=1}^m \lambda_i = \max\big\{f(x^{[1]}), \ldots, f(x^{[m]})\big\} < f(x^*),$$

což je spor. ∎

Přímým důsledkem předchozího věty je velmi důležitý fakt týkající se minimalizace afinní funkce na polytopu neboli řešení úlohy lineárního programování. Tvrzení plyne z toho, že minimalizace funkce f je totéž jako maximalizace funkce $-f$ a že f i $-f$ jsou současně afinní, a tedy i konvexní[100].

> Afinní funkce $f : X \to \mathbb{R}$ svého minima na polytopu $X := \mathrm{conv}\{x^{[1]}, \ldots, x^{[m]}\} \subseteq \mathbb{R}^n$ nabývá v některém z vrcholů X, tj. z bodů $x^{[1]}, \ldots, x^{[m]}$.

Cvičení

6.2.1. Pomocí definice rozhodněte o konvexnosti funkcí

$$f(x) = \frac{1}{a-x}, \quad x \in (-\infty, a),$$
$$g(x) = \frac{x^2}{a-x}, \quad x \in (-\infty, a),$$
$$h(x) = \frac{x(x-1)}{2}, \quad x \in \mathbb{R}.$$

6.2.2. Pomocí definice dokažte, že funkce $f(x) = \mathrm{e}^x$ je konvexní na $\mathbb{R}$.

množiny X.

Důkaz. Označme jako X^* množina bodů, ve kterých funkce f nabývá svého maxima na X. Je-li $X^* = \emptyset$ nebo je-li f konstantní, je tvrzení triviální. Uvažme proto nyní, že nekonstantní funkce f nabývá svého maxima na X pouze v bodech, které nejsou extrémní, tj. $f(x^*) > f(z)$ pro všechna $x^* \in X^*$ a libovolný extrémní bod $z \in X$. Jelikož $x^* \in X^*$ není extrémní bod, lze jej dle Kreinovy–Milmanovy věty zmíněné v Poznámce 6.1.18(iii) vyjádřit jako konvexní kombinaci extrémních bodů množiny X, tj. $x^* = \sum_{i=1}^m \lambda_i z^{[i]}$ pro vhodná $\lambda_1,\dots,\lambda_m \in (0,1)$ splňující $\sum_{i=1}^m \lambda_i = 1$ a některé (díky Carathédoryho větě nejvýše $n+1$, což ale není v tuto chvíli důležité) extrémní body $z^{[1]},\dots,z^{[m]} \in X$. Pak ale z Jensenovy nerovnosti získáme

$$\begin{aligned} f(x^*) &\le \sum_{i=1}^m \lambda_i f\big(z^{[i]}\big) \le \\ &\le \max\{f(z^{[1]}),\dots,f(z^{[m]})\} < \\ &< f(x^*), \end{aligned}$$

což je spor.

[99] Dokonce díky Carathédoryho větě víme, že by nám zde stačilo nejvýše $n+1$ bodů $x^{[1]},\dots,x^{[m]}$. To ale není v tuto chvíli důležité.

[100] Tento jednoduchý fakt vede k tomu, že: *Afinní funkce $f(x) = p^\top x + q$ pro $p \in \mathbb{R}^n$ a $q \in \mathbb{R}$ svého minima/maxima na konvexní množině $X \subseteq \mathbb{R}^n$ nabývá v bodě $x^* \in \operatorname{ri} X$ právě tehdy, když f je konstantní funkce (tj. $p = 0$).*

6.2.3. V závislosti na hodnotě $p \in \mathbb{R}$ rozhodněte pomocí definice o konvexnosti funkce

$$f(x) = x^p, \quad x \in (0,\infty).$$

řešení: konvexní pro $p \ge 1$ a $p \le 0$, konkávní pro $p \in [0,1]$

6.2.4. Pomocí definice rozhodněte o konvexnosti funkce

$$f(x) = x \ln x, \quad x \in (0,\infty).$$

řešení: je konvexní

6.2.5. Pomocí definice rozhodněte o konvexnosti funkce

$$f(x) = \|x\|^p,$$

kde $x \in \mathbb{R}^n$, $p > 1$ a $\|\cdot\|$ je libovolná norma na $\mathbb{R}^n$.

řešení: je konvexní

6.2.6. Pomocí definice rozhodněte o konvexnosti Hölderovy normy

$$f(x) = \|x\|_p := \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}$$

kde $x \in \mathbb{R}^n$ a $p \ge 1$.

6.2.7. Pomocí definice rozhodněte o konvexnosti funkce

$$f(\beta) = \|X\beta - y\|^2 = (X\beta - y)^\top (X\beta - y),$$

kde $X \in \mathbb{R}^{n\times m}$, $\beta \in \mathbb{R}^m$ a $y \in \mathbb{R}^n$.

6.2.8. Pomocí definice rozhodněte o konvexnosti funkce

$$f(x_1, x_2) = \mathrm{e}^{x_1 + x_2}$$

na $\mathbb{R}^2$. Je zde také ostře a/nebo silně konvexní? V případě silné konvexnosti určete největší konstantu silné konvexnosti ϑ.

řešení: je pouze konvexní

6.2.9. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$ konvexní funkce. Potom kdykoli platí $\lambda x + (1-\lambda) y \in X$ pro nějaké $x, y \in X$ a $\lambda \in \mathbb{R} \setminus [0,1]$, pak

$$f\big(\lambda x + (1-\lambda) y\big) \ge \lambda f(x) + (1-\lambda) f(y).$$

6.2.10. *Dokažte:* Funkce $f : \mathbb{R}^n \to \mathbb{R}$ je současně konvexní i konkávní na $\mathbb{R}^n$ právě tehdy, když f je afinní funkce.

6.2.11. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f : X \to \mathbb{R}$ konvexní a funkce $g : X \to \mathbb{R}$ silně konvexní s konstantou silné konvexnosti ϑ. Potom funkce $f + g : X \to \mathbb{R}$ je opět silně konvexní s konstantou ϑ.

6.2.12. *Dokažte:* Nechť $\alpha \ge 0$, $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ silně konvexní s konstantou silné konvexnosti ϑ. Potom funkce $\alpha f : X \to \mathbb{R}$ je silně konvexní s konstantou $\alpha\vartheta$.

6.2.13. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$. Pak funkce f je silně konvexní s konstantou silné konvexnosti ϑ právě tehdy, když funkce $g(x) := f(x) - \vartheta\|x\|^2$ je konvexní na X.

6.2.14. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, I je libovolná indexová množina, funkce $f_i : X \to \mathbb{R}$ jsou konvexní pro všechna $i \in I$ a pro každé $x \in X$ je množina $\{f_i(x) \mid i \in I\}$ shora ohraničená, tj. existuje funkce $M : X \to \mathbb{R}$ taková, že pro každé $i \in I$ platí $f(x) \le M(x)$ pro každé $x \in X$. Potom funkce

$$f(x) := \sup_{i\in I} \{f_i(x)\}$$

je konvexní na X.

6.2.15. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $g_1, \dots, g_m : X \to \mathbb{R}$ jsou konvexní a funkce $f : \mathbb{R}^m \to \mathbb{R}$ je konvexní a neklesající v tom smyslu, že jsou-li $x = (x_1, \dots, x_m)^\top$ a $y = (y_1, \dots, y_m)^\top$ takové, že $x_i \le y_i$ pro všechna $i = 1, \dots, m$, potom $f(x) \le f(y)$. Potom složená funkce

$$F(x) := f\big(g_1(x), \dots, g_m(x)\big)$$

je také konvexní na X.

6.2.16. Nechť X je konvexní množina. Pak funkce $f : X \to \mathbb{R}_{++} = \{x \in \mathbb{R} \mid x > 0\}$ se nazývá *logaritmicky konvexní* (též *superkonvexní*) na X, jestliže funkce $\ln f(x)$ je konvexní na X.

Dokažte: Každá logaritmicky konvexní funkce je konvexní.

6.2.17. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ spojitá na X. Jestliže pro každé $x, y \in X$, $x \ne y$, existuje $\lambda \in (0,1)$ takové, že

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y),$$

pak funkce f je konvexní na X.

6.2.18. Rozhodněte o konvexnosti množiny

$$X = \big\{[x,y] \in \mathbb{R}^2 \mid (x^3 + \mathrm{e}^y)\ln(x^3 + \mathrm{e}^y) \le 49,\ x \ge 2,\ y \in \mathbb{R}\big\}.$$

řešení: je konvexní

6.2.19. *Dokažte:* Pro $\lambda_1, \dots, \lambda_n \ge 0$ splňující $\sum_{i=1}^n \lambda_i = 1$ a $x_1, \dots, x_n > 0$ platí

$$1 \le \Big(\sum_{i=1}^n \lambda_i x_i\Big)\Big(\sum_{i=1}^n \frac{\lambda_i}{x_i}\Big).$$

6.2.20. *Dokažte:* Pro $\lambda_1, \dots, \lambda_n \ge 0$ splňující $\sum_{i=1}^n \lambda_i = 1$ a $x_1, \dots, x_n \in \mathbb{R}$ platí

$$\sum_{i=1}^n \lambda_i x_i \le \ln\Big(\sum_{i=1}^n \lambda_i \mathrm{e}^{x_i}\Big).$$

6.2.21. Pomocí Jensenovy nerovnosti dokažte nerovnost mezi průměry různých řádů

$$\left(\frac{x_1^p + \dots + x_m^p}{m}\right)^{1/p} \ge \left(\frac{x_1^q + \dots + x_m^q}{m}\right)^{1/q},$$

kde $x_1, \dots, x_m > 0$ a $p > q$.

6.2.22. Dokažte Minkowského nerovnost

$$\Big(\sum_{i=1}^n |a_i + b_i|^p\Big)^{1/p} \le \Big(\sum_{i=1}^n |a_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^n |b_i|^p\Big)^{1/p},$$

kde $p \ge 1$ a $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$.

6.2.23. Pomocí Jensenovy nerovnosti dokažte

$$\big(1 + \sqrt[5]{a}\big)^5 + \big(1 + \sqrt[5]{b}\big)^5 \le 2^6,$$

kde $a, b \ge 0$ a $a + b = 2$.

6.2.24. Pomocí Jensenovy nerovnosti dokažte

$$a^a\, b^b\, c^c \ge \left(\frac{a+b+c}{3}\right)^{a+b+c},$$

kde $a, b, c > 0$.

6.2.25. Pomocí Jensenovy nerovnosti dokažte

$$\sqrt[n]{\sin x_1 \cdots \sin x_n} \le \sin \frac{x_1 + \cdots + x_n}{n},$$

kde $n \in \mathbb{N}$ a $x_1, \ldots, x_n \in [0, \pi/2]$.

6.2.26. Pomocí Jensenovy nerovnosti dokažte

$$\frac{\frac{1}{n}\sum_{i=1}^n x_i}{\sin(\frac{1}{n}\sum_{i=1}^n x_i)} \le \sqrt[n]{\frac{x_1 \cdots x_n}{\sin x_1 \cdots \sin x_n}},$$

kde $x_1, \ldots, x_n \in (0, \pi/2)$.

6.2.27. Pomocí Jensenovy nerovnosti dokažte

$$\frac{a}{a+3b+3c} + \frac{b}{3a+b+3c} + \frac{c}{3a+3b+c} \ge \frac{3}{7},$$

kde $a, b, c > 0$.

6.2.28. Pomocí Jensenovy nerovnosti dokažte

$$\sum_{i=1}^n \frac{1}{1+a_i} \ge \frac{n}{1+\sqrt[n]{a_1 \cdots a_n}},$$

kde $a_1, \ldots, a_n \ge 1$.

6.2.29. Pomocí Jensenovy nerovnosti dokažte

$$\frac{x_1}{\sqrt{1-x_1}} + \frac{x_2}{\sqrt{1-x_2}} + \cdots + \frac{x_n}{\sqrt{1-x_n}} \ge \sqrt{\frac{n}{n-1}},$$

kde $x_1, \ldots, x_n > 0$ a $\sum_{i=1}^n x_i = 1$.

6.2.30. Pomocí Jensenovy nerovnosti dokažte

$$\sqrt{1^2+1} + \sqrt{2^2+1} + \sqrt{3^2+1} + \cdots + \sqrt{n^2+1} \ge \frac{n}{2}\sqrt{n^2+2n+5}.$$

6.2.31. Pomocí Jensenovy nerovnosti dokažte, že pro úhly α, β, γ v trojúhelníku platí

$$\sin\alpha + \sin\beta + \sin\gamma \le \frac{3\sqrt{3}}{2},$$
$$\sqrt{\sin\alpha} + \sqrt{\sin\beta} + \sqrt{\sin\gamma} \le 3\sqrt[4]{\frac{3}{4}},$$
$$\sin\alpha \sin\beta \sin\gamma \le \frac{3\sqrt{3}}{8},$$
$$\cos\alpha \cos\beta \cos\gamma \le \frac{1}{8}.$$

6.2.32. Pomocí AG-nerovnosti dokažte

$$\frac{b+c}{\sqrt{a}} + \frac{c+a}{\sqrt{b}} + \frac{a+b}{\sqrt{c}} \ge \sqrt{a} + \sqrt{b} + \sqrt{c} + 3,$$

kde $abc = 1$ a $a, b, c > 0$.

6.2.33. Pomocí drátu bychom chtěli „vymodelovat“ kvádr (tj. jeho 12 hran) tak, aby model měl objem V a spotřeba drátu byla minimální. Určete rozměry optimálního modelu.

6.2.34. Dokažte, že pro každé $x > 1$ platí

$$\frac{1}{1-x} + \frac{1}{x} + \frac{1}{1+x} \ge \frac{3}{x}.$$

6.2.35. Dokažte, že pro libovolná $a, b \in [-1, 1]$ platí

$$\sqrt{1-a^2} + \sqrt{1-b^2} \le \sqrt{4-(a+b)^2}.$$

6.2.36. Dokažte, že pro libovolná $a, b > 0$ splňující $a + b = 1$ platí

$$\left(a + \frac{1}{a}\right)^2 + \left(b + \frac{1}{b}\right)^2 \geq \frac{25}{2}.$$

6.2.37. Dokažte, že pro libovolná $x, y, z > 0$ splňující $x + y + z = 1$ platí

$$64 \leq \left(1\frac{1}{x}\right)\left(1\frac{1}{y}\right)\left(1\frac{1}{z}\right).$$

6.2.38. Dokažte, že pro každé $n \in \mathbb{N} \setminus \{1\}$, platí

$$(1+n)\cos\frac{\pi}{n+1} - n\cos\frac{\pi}{n} \geq 1.$$

6.2.39. Dokažte, že v každém trojúhelníku s délkami stran a, b, c a obsahem S platí

$$ab + ac + bc \geq 4\sqrt{3}\,S.$$

6.2.40. Ukažte, že ze všech konvexních n-úhelníků vepsaných do dané kružnice má největší obsah právě pravidelný n-úhelník.

6.2.41. Dokažte *Hölderův*[101] *rozdílový vzorec*: Nechť funkce $f : I \to \mathbb{R}$ má vlastní druhou derivaci na intervalu I a nechť pro každé $x \in I$ platí $m \leq f''(x) \leq M$ pro nějaká $m, M \in \mathbb{R}$. Pak pro libovolná $x_1, \ldots, x_n \in I$ a nezáporná $\lambda_1, \ldots, \lambda_n$ splňující $\lambda_1 + \cdots + \lambda_n = 1$ existuje číslo $\mu \in [m, M]$ takové, že

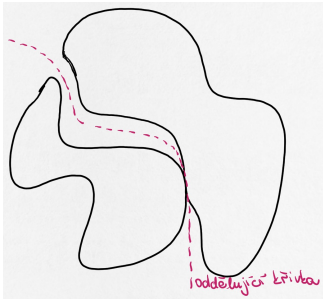
$$\sum_{i=1}^{n} \lambda_i f(x_i) - f\Big(\sum_{i=1}^{n} \lambda_i x_i\Big) = \frac{\mu}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j (x_i - x_j)^2.$$

[101] Tento výsledek také pochází z [273]. Explicitně vyjadřuje rozdíl mezi oběma stranami Jensenovy nerovnosti bez požadavku konvexnosti, a je tedy obecnější než Jensenova nerovnost. V případě konvexní funkce totiž bude $0 \leq m \leq M$, a tudíž $\mu \geq 0$, z čehož již plyne Jensenova nerovnost, srovnej také s definicí silné konvexnosti. Např. pro $m = 2$ a $\lambda_1 = \lambda_2 = \frac{1}{2}$ odtud dostaneme

$$f(x_1) + f(x_2) - 2f\Big(\frac{x_1 + x_2}{2}\Big) = \\ = \frac{\mu}{4}(x_1 - x_2)^2.$$

6.3 Oddělování konvexních množin a jeho důsledky

Nyní se ještě vrátíme zpět ke studiu konvexních množin. S využitím Věty 6.2.23(ii) totiž odvodíme výsledky týkající se oddělování konvexních množin, které patří k základním stavebním kamenům konvexního programování. Klíčovou otázkou by pro následující chvíle/stránky mohlo být, zda je možné oddělit dvě dané množiny $X, Y \subseteq \mathbb{R}^n$, tj. najít takovou „křivku“, aby každá z množin X a Y ležela v jiné části $\mathbb{R}^n$ vymezené onou křivkou, viz Obrázek 6.50.



Obrázek 6.50: Oddělování dvou množin.

Jenže toto pojetí by bylo velmi obšírné a složité[102]. Proto se omezíme pouze na konvexní množiny a využijeme jejich „úsečkový“ charakter, takže nás bude zajímat „přímkové“ oddělování pomocí nadrovin[103].

DEFINICE 6.3.1

Neprázdné množiny $X_1, X_2 \subseteq \mathbb{R}^n$ se nazývají

(i) *oddělitelné*, jestliže existuje vektor $p \in \mathbb{R}^n \setminus \{0\}$ takový, že

$$\langle p, x_1 \rangle \geq \langle p, x_2 \rangle \tag{6.3.1}$$

pro každé $x_1 \in X_1$ a $x_2 \in X_2$;

(ii) *vlastně oddělitelné*, jestliže jsou oddělitelné a zároveň existuje dvojice bodů $x_1^* \in X_1$ a $x_2^* \in X_2$ splňující

$$\langle p, x_1^* \rangle > \langle p, x_2^* \rangle;$$

(iii) *silně oddělitelné*, jestliže existuje $p \in \mathbb{R}^n \setminus \{0\}$ takové, že

$$\inf_{x_1 \in X_1} \langle p, x_1 \rangle > \sup_{x_2 \in X_2} \langle p, x_2 \rangle.$$

Je-li navíc $\beta \in \left[\sup_{x_2 \in X_2} \langle p, x_2 \rangle, \inf_{x_1 \in X_1} \langle p, x_1 \rangle\right]$, nadrovina

$$H_{p,\beta} := \{x \in \mathbb{R}^n \mid \langle p, x \rangle = \beta\}$$

se nazývá *oddělující nadrovinou* množin X_1 a X_2.

[102] Museli bychom si vyjasnit mnoho nuancí. Jak moc obecnou křivku máme na mysli, co jsou ony dvě části, apod.

[103] Připomeňme, že rovina je dvoudimenzionální množina bodů v $\mathbb{R}^3$. Může být určena 3 body, které neleží na jedné přímce, nebo 1 bodem a přímkou neobsahující tento bod (přímka pak je součástí této roviny). Taky ji můžeme brát jako množinu bodů $A \in \mathbb{R}^3$, pro které je příslušný vektor spojující A a daný bod $A_0 = [x_0, y_0, z_0]$ kolmý na daný vektor $p = (a, b, c)^\top \neq (0, 0, 0)^\top$, tj.

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

neboli

$$\langle p, A - A_0 \rangle = \langle p, r - r_0 \rangle = 0,$$

kde r, r_0 jsou vektory určené body

A, A_0. To tedy znamená, že

$$\langle p, r\rangle = \langle p, r_0\rangle = \text{konst.},$$

přičemž p, r_0 jsou pevně dány a hledáme pouze vektory r splňující uvedenou rovnost. Toto pochopitelně platí i obecně, takže nadrovinou v $\mathbb{R}^n$ rozumíme množinu (bod v $\mathbb{R}$, přímku v $\mathbb{R}^2$ a rovinu v $\mathbb{R}^3$)

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid \langle p, x\rangle = b\}$$

pro $b := \langle p, r_0\rangle$, tj.

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid \langle p, x\rangle = \langle p, x_0\rangle\} = \\ = x_0 + \{x \in \mathbb{R}^n \mid \langle p, x\rangle = 0\},$$

což je $(n-1)$-dimenzionální afinní prostor se zaměřením $\operatorname{Lin}\mathcal{R} = \{x \in \mathbb{R}^n \mid \langle p, x\rangle = 0\}$. Tato nadrovina rozděluje prostor $\mathbb{R}^n$ na dva *poloprostory*

$$\{x \in \mathbb{R}^n \mid \langle p, x\rangle \geq b\}$$

a

$$\{x \in \mathbb{R}^n \mid \langle p, x\rangle \leq b\},$$

což motivuje následující Definici 6.3.1.

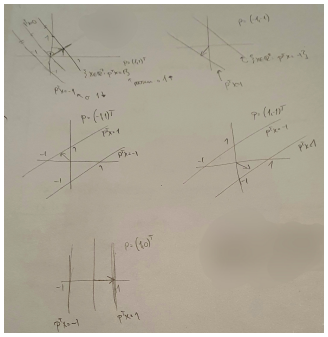
Připomeňme také, že pro vektory $u, v \in \mathbb{R}^n$ je $\langle u, v\rangle$ velikost projekce vektoru v do směru vektoru u vynásobená velikostí u, tj.

$$\langle u, v\rangle = \|u\|\|v\|\cos\alpha.$$

Proto bod x_0 výše je takový, že

$$\langle p, x_0\rangle = \|p\|\|x_0\|\cos\alpha = b.$$

Jaká je role parametru $b \in \mathbb{R}$? Je-li $b = 0$, pak nadrovina $\langle p, x\rangle = b$ prochází počátkem a vektor p je její normálový vektor. Změnou na hodnotu $b \neq 0$ dostaneme rovnoběžnou nadrovinu, která se pro $b > 0$ posune ve směru vektoru p, zatímco pro $b < 0$ to bude v opačném směru. Délka posunutí závisí na hodnotě b a složkách vektoru p. Toto je ilustrováno na následujícím obrázku.



Pochopitelně role množin X_1 a X_2 není v Definici 6.3.1 zafixovaná a můžeme je zaměnit, tj. v oddělování X_1 od X_2 nebo X_2 od X_1 není žádný rozdíl. Samotná oddělitelnost pak znamená, že buď

$$X_1 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \geq b\} \quad \text{a} \quad X_2 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \leq b\}$$

nebo

$$X_1 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \leq b\} \quad \text{a} \quad X_2 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \geq b\}$$

pro nějaké $p \in \mathbb{R}^n \setminus \{0\}$ a $b \in \mathbb{R}$, což je ekvivalentní s (6.3.1). Můžeme nabídnou ještě jiný pohled na podmínku (6.3.1): existují $p \in \mathbb{R}^n \setminus \{0\}$ a $b \in \mathbb{R}$ taková, že afinní funkce $\langle p, x\rangle - b$ je nezáporná na X_1 a nekladná na X_2 (případně naopak). Vlastní oddělitelnost pak vyžaduje, aby navíc množina $X_1 \cup X_2$ nebyla zcela součástí oddělující nadroviny $\{x \in \mathbb{R}^n \mid \langle p, x\rangle = b\}$. Konečně v případě silné oddělitelnosti musí být X_1 a X_2 v disjunktních uzavřených poloprostorech, tj. musí navíc existovat $\varepsilon > 0$ takové, že

$$X_1 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \geq b + \varepsilon\} \quad \text{a} \quad X_2 \subseteq \{x \in \mathbb{R}^n \mid \langle p, x\rangle \leq b\}.$$

Jednotlivé typy oddělitelnosti si nyní ilustrujeme na třech jednoduchých příkladech.

Příklad 6.3.2

(Velmi patologický.) Uvažme

$$X_1 = \{[x, y] \in \mathbb{R}^2 \mid y = 0 \ \&\ 0 \leq x \leq 2\} \quad \text{a} \quad X_2 = \{[x, y] \in \mathbb{R}^2 \mid y = 0 \ \&\ 1 \leq x \leq 3\}.$$

Tyto množiny jsou oddělitelné, neboť pro $p = (0,1)^\top$ a $z = (x, y)^\top$ bude $\langle p, z\rangle = 0$ pro všechna $z \in X_1 \cup X_2$, viz Obrázek 6.51. Toto však asi příliš neodpovídá naší intuici, neboť množiny X_1 a X_2 leží ve stejné nadrovině $\{[x, y] \in \mathbb{R}^2 \mid y = 0\}$, která je tudíž současně i jedinou oddělující nadrovinou.

Příklad 6.3.3

K tomu, aby dvě množiny byly vlastně oddělitelné, nemusí být disjunktní. Např.

$$X_1 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ 0 \leq y \leq 1\} \quad \text{a} \quad X_2 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ -1 \leq y \leq 0\}$$

jsou vlastně oddělitelné nadrovinou $H_{p,\beta}$ pro $p = (0,1)^\top$ a $\beta = 0$, viz Obrázek 6.52. Vskutku, pro libovolné $z \in X_1$ máme $\langle p, z\rangle = y \geq 0$, zatímco pro libovolné $z \in X_2$ máme $\langle p, z\rangle = y \leq 0$. Navíc pro $z_1 = (0,1)^\top \in X_1$ a $z_2 = (0,-1)^\top \in X_2$ bude

$$\langle p, z_1\rangle = 1 > -1 = \langle p, z_2\rangle.$$

Dokonce i v případě $X_1 \subseteq X_2$ mohou být množiny vlastně oddělitelné, např.

$$X_1 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ 0 \leq y \leq 1\} \quad \text{a} \quad X_2 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ y = 0\}$$

jsou vlastně oddělitelné nadrovinou $H_{p,\beta}$ pro $p = (0,1)^\top$ a $\beta = 0$, neboť pro libovolné $z \in X_1$ máme $\langle p, z\rangle = y = 0$, pro libovolné $z \in X_2$ máme $\langle p, z\rangle = y \leq 0$ a pro $z_1 = (1,1)^\top \in X_1$ a $z_2 = (1,0)^\top \in X_2$ bude

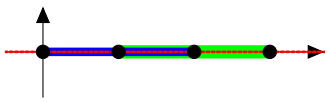
$$\langle p, z_1\rangle = 1 > 0 = \langle p, z_2\rangle,$$

viz Obrázek 6.53.

Příklad 6.3.4

Silně oddělitelné jsou např. množiny

$$X_1 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ 1/2 \leq y \leq 1\},\ X_2 = \{[x, y] \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \&\ -1 \leq y \leq -1/2\},$$



Obrázek 6.51: Množiny X_1, X_2 (modrá a zelená) a oddělující nadrovina (červená) z Příkladu 6.3.2.

viz Obrázek 6.55. Vskutku, např. pro $p = (0,1)^\top$ a $\beta = 0$ bude platit

$$\inf_{z\in X_1} \langle p,z\rangle = \inf_{y\in[1/2,1]} y = 1/2 > 0 \quad \text{a} \quad \sup_{z\in X_2} \langle p,z\rangle = \sup_{y\in[-1,-1/2]} y = -1/2 < 0.$$

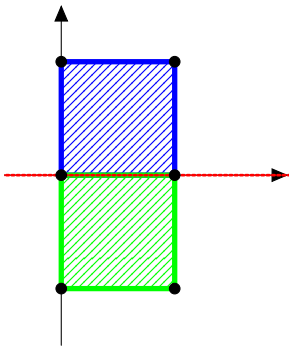
Poznámka 6.3.5. Z uvedených příkladů (a i z Definice 6.3.6) je snad zřejmé, že jakékoli silně oddělitelné množiny jsou i vlastně oddělitelné a podobně z vlastní oddělitelnosti plyne samotná oddělitelnost, zatímco opačné implikace rozhodně neplatí. Navíc vzhledem ke geometrickému významu skalárního součinu v případě jednotkových vektorů, můžeme nerovnost (6.3.1) interpretovat i pomocí odchylky vektorů určených body z X_1 a X_2 od normálového vektoru oddělující nadroviny. Posuneme-li pro jednoduchost množiny X_1 a X_2 tak, že oddělující nadrovina prochází počátkem, tj. je určená rovností $\langle p,x\rangle = 0$, pak nerovnost $\langle p,x\rangle \geq 0$ pro $x \in X_1$ znamená, že vektory určené body z X_1 svírají s vektorem p úhel v rozmezí $[-\pi/2,\pi/2]$, zatímco vektory určené body z X_2 musí svírat úhel v rozmezí $[-\pi,-\pi/2]\cup[\pi/2,\pi]$, viz Obrázky 6.56–6.57. ▲

Pro důkazy některých tvrzení o oddělování konvexních množin budeme potřebovat následující (snad už známý) nástroj.

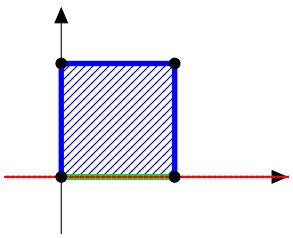
DEFINICE 6.3.6

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná množina a $x \in \mathbb{R}^n$. Bod $x^* \in X$ nazveme *projekcí bodu x na množinu X* a označíme $\Pi_X(x)$, jestliže

$$\|\Pi_X(x) - x\| \leq \|y - x\| \quad \text{pro každé } y \in X. \tag{6.3.2}$$



Obrázek 6.52: Množiny X_1, X_2 (modrá a zelená) a oddělující nadrovina (červená) z první části Příkladu 6.3.3.



Obrázek 6.53: Množiny X_1, X_2 (modrá a zelená) a oddělující nadrovina (červená) ze druhé části Příkladu 6.3.3.

Výraz na pravé straně nerovnosti (6.3.2) je úzce spojen se *vzdáleností* bodu x od množiny X, tj. s hodnotou $\rho(x,X) := \inf_{y\in X}\|x-y\|$, takže projekce $\Pi_X(x)$ musí splňovat $\|\Pi_X(x)-x\| \leq \rho(x,X)$, tj. vzdálenost vzoru od projekce je nejvýše rovna vzdálenosti vzoru od množiny X. Avšak současně musí být $\Pi_X(x) \in X$, což nutně znamená, že $\Pi_X(x)$ je bod, ve kterém se realizuje dokonce $\min_{y\in X}\|y-x\|$, tj.

$$\|\Pi_X(x) - x\| = \min_{y\in X}\|x-y\| = \rho(x,X).$$

Odtud je také zřejmé, že v případě $x \in X$ bude $\Pi_X(x) = x$.

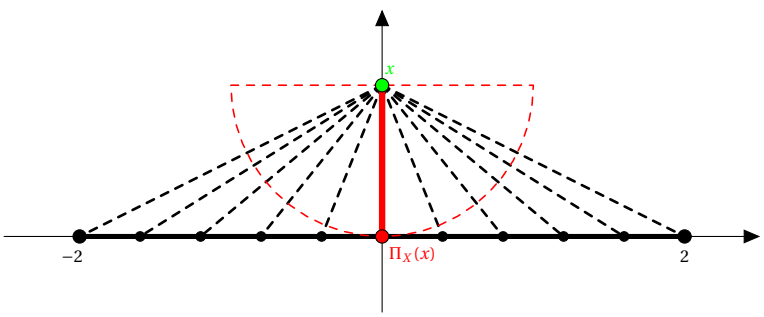
Projekce bodu x na danou množinu však může být ovlivněna možnými „potížemi“ s existencí či jednoznačností minima. Např. pro množinu $X := \{[x_1,x_2] \in \mathbb{R}^2 \mid -2 \leq x_1 \leq 2,\ x_2 = 0\}$ a bod $x = [0,1]$ bude

$$\|y-x\| = \sqrt{y_1^2 + (y_2-1)^2} \overset{y\in X}{=} \sqrt{y_1^2+1} \geq 1 \quad \text{pro všechna } y \in X\,,$$

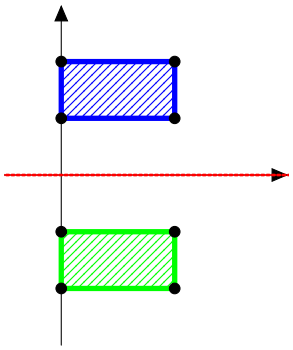
takže hledáme $\Pi_X(x) = [x_1,0] \in X$ takové, že

$$\|\Pi_X(x) - x\| = \sqrt{x_1^2+1} \leq 1,$$

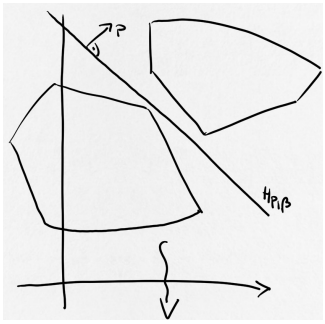
což je možné právě tehdy, když $x_1^2 + 1 \leq 1$ neboli $x_1^2 \leq 0$, takže $x_1 = 0$, tj. $\Pi_X(x) = [0,0]$, viz Obrázek 6.54.



Obrázek 6.54: Projekce bodu $x = [0,1]$ na množinu $X := \{[x_1,x_2] \in \mathbb{R}^2 \mid -2 \leq x_1 \leq 2,\ x_2 = 0\}$.

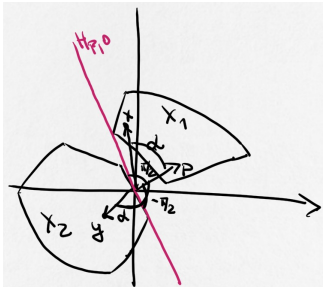


Obrázek 6.55: Množiny X_1, X_2 (modrá a zelená) a oddělující nadrovina (červená) z první části Příkladu 6.3.4.



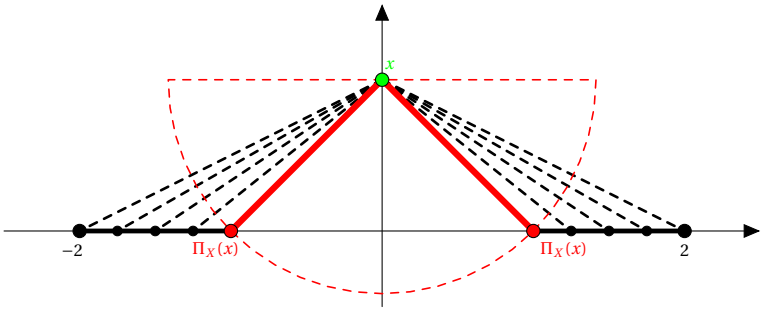
Obrázek 6.56: Geometrický pohled na oddělování na oddělování konvexních množin.

LEMMA 6.3.7

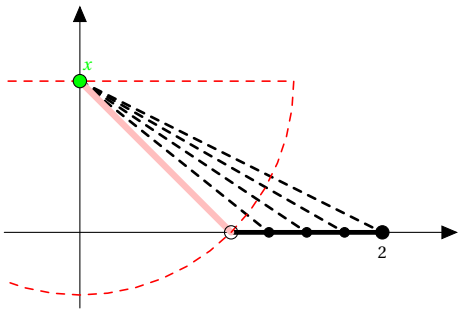


Obrázek 6.57: Geometrický pohled na oddělování na oddělování konvexních množin (pokr.).

Pro daný bod x jsme tedy našli jedinou jeho projekci na danou množinu X. Avšak v případě množiny $Y := \{[x_1,x_2] \in \mathbb{R}^2 \mid -2 \le x_1 \le -1,\ x_2 = 0\} \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid 1 \le x_1 \le 2,\ x_2 = 0\}$ již ztratíme jednoznačnost, neboť $\|y - x\| = \sqrt{y_1^2 + 1} \ge \sqrt{2}$ pro všechna $y \in Y$ a nerovnost $\|\Pi_X(x) - x\| = \sqrt{x_1^2 + 1} \le \sqrt{2}$ bude splněna v případě $x_1 = \pm 1$, tj. $\Pi_Y(x) = [\pm 1, 0]$, viz Obrázek 6.58 níže. Projekce $\Pi_X(x)$ ale pochopitelně ani nemusí existovat. Vezmeme-li např. $Z := \{[x_1,x_2] \in \mathbb{R}^2 \mid 1 < x_1 \le 2,\ x_2 = 0\}$, pak opět $\|y - x\| = \sqrt{y_1^2 + 1} \ge \sqrt{2}$ pro všechna $y \in Z$, avšak neexistuje $\Pi_Z(x) \in Z$ takové, že $\|\Pi_Z(x) - x\| = \sqrt{x_1^2 + 1} \le \sqrt{2}$, viz Obrázek 6.59.



Obrázek 6.58: Projekce bodu $x = [0,1]$ na množinu $Y := \{[x_1,x_2] \in \mathbb{R}^2 \mid -2 \le x_1 \le -1,\ x_2 = 0\} \cup \{[x_1,x_2] \in \mathbb{R}^2 \mid 1 \le x_1 \le 2,\ x_2 = 0\}$.



Obrázek 6.59: Projekce bodu $x = [0,1]$ na množinu $Z := \{[x_1,x_2] \in \mathbb{R}^2 \mid 1 < x_1 \le 2,\ x_2 = 0\}$.

Následující tvrzení ukazuje, že pro existenci a jednoznačnost projekce bodu hraje velmi důležitou roli uzavřenost a konvexnost množiny X, srovnej množiny X, Y a Z.

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná, uzavřená a konvexní množina. Pak pro libovolný bod $x^* \in \mathbb{R}^n$ existuje jediná projekce $\Pi_X(x^*) \in X$ a pro každé $x \in X$ platí

$$\langle \Pi_X(x^*) - x^*, x - \Pi_X(x^*) \rangle \ge 0, \tag{6.3.3}$$

$$\langle \Pi_X(x^*) - x^*, x - x^* \rangle \ge \|\Pi_X(x^*) - x^*\|^2 \ge 0. \tag{6.3.4}$$

Důkaz. Je-li $x^* \in X$, pak tvrzení je triviální, neboť nutně $\Pi_X(x^*) = x^*$. Nechť tedy $x^* \notin X$. Nejdříve si uvědomme, že projekce bodu x^* na množinu X je vlastně řešením úlohy

$$f(x) := \|x - x^*\|^2 \to \min, \quad x \in X, \tag{6.3.5}$$

v níž funkce f odpovídá kvadratické formě s maticí $A = I$ s přidanou afinní část. To znamená, že f je ostře konvexní, a z Věty 6.2.23(ii) vyplývá, že úloha (6.3.5) má nejvýše jedno řešení.

Položme nyní

$$Y := X \cap \{x \in \mathbb{R}^n \mid \|x - x^*\| \le R\},$$

kde $R \in \mathbb{R}$ dostatečně velké. Jak? Aby množina Y byla neprázdná, viz také níže. Pak množina Y je kompaktní, a tedy existuje řešení úlohy

$$f(x) \to \min, \quad x \in Y. \tag{6.3.6}$$

Jenže řešení této úlohy je totožné s řešením (6.3.5). Proč? Je-li množina Y neprázdná, pak zvětšováním čísla R budeme do Y „přidávat“ body s větší hodnotou $\|x-x^*\|$, které neovlivní řešení původní úlohy (6.3.6). Proto projekce $\Pi_X(x^*)$ existuje a je jednoznačně určena. Z konvexnosti množiny X platí pro libovolné $x\in X$ a $\lambda\in(0,1]$, že

$$\begin{aligned}\|\Pi_X(x^*)-x^*\|^2 &\overset{\Pi_X(x^*)\text{ řeší }(6.3.5)}{\leq} \|\underbrace{\lambda x+(1-\lambda)\Pi_X(x^*)}_{\in X\ \&\ \neq\Pi_X(x^*)}-x^*\|^2 = \\ &= \big\|\Pi_X(x^*)-x^*+\lambda\big[x-\Pi_X(x^*)\big]\big\|^2 = \\ &= \big\langle \Pi_X(x^*)-x^*+\lambda\big[x-\Pi_X(x^*)\big],\Pi_X(x^*)-x^*+\lambda\big[x-\Pi_X(x^*)\big]\big\rangle = \\ &= \big\langle \Pi_X(x^*)-x^*,\Pi_X(x^*)-x^*\big\rangle+\lambda\big\langle \Pi_X(x^*)-x^*,x-\Pi_X(x^*)\big\rangle+ \\ &\qquad +\lambda\big\langle x-\Pi_X(x^*),\Pi_X(x^*)-x^*\big\rangle+\lambda^2\big\langle x-\Pi_X(x^*),x-\Pi_X(x^*)\big\rangle \\ &= \|\Pi_X(x^*)-x^*\|^2+2\lambda\big\langle \Pi_X(x^*)-x^*,x-\Pi_X(x^*)\big\rangle+\lambda^2\|x-\Pi_X(x^*)\|^2.\end{aligned}$$

Vydělením číslem $\lambda\in(0,1]$ dostáváme

$$0\leq 2\big\langle \Pi_X(x^*)-x^*,x-\Pi_X(x^*)\big\rangle+\lambda\|x-\Pi_X(x^*)\|^2,$$

z čehož limitním přechodem pro $\lambda\to 0^+$ obdržíme

$$\big\langle \Pi_X(x^*)-x^*,x-\Pi_X(x^*)\big\rangle\geq 0,$$

což je nerovnost (6.3.3). Navíc odtud plyne

$$\begin{aligned}0&\leq\big\langle \Pi_X(x^*)-x^*,x-\Pi_X(x^*)+x^*-x^*\big\rangle \\ &=\big\langle \Pi_X(x^*)-x^*,x-x^*\big\rangle+\big\langle \Pi_X(x^*)-x^*,x^*-\Pi_X(x^*)\big\rangle \\ &=\big\langle \Pi_X(x^*)-x^*,x-x^*\big\rangle-\|\Pi_X(x^*)-x^*\|^2,\end{aligned}$$

tj. platí $\big\langle \Pi_X(x^*)-x^*,x-x^*\big\rangle\geq\|\Pi_X(x^*)-x^*\|^2\geq 0$, což je nerovnost (6.3.4). ■

Díky předchozímu výsledku „snadno“ odvodíme nutnou a postačující podmínku silné oddělitelnosti.

Neprázdné konvexní množiny $X_1,X_2\subseteq\mathbb{R}^n$ jsou silně oddělitelné právě tehdy, když mají nenulovou vzdálenost, tj.

$$\rho(X_1,X_2):=\inf_{x_1\in X_1,x_2\in X_2}\|x_1-x_2\|>0,$$

což je ekvivalentní s podmínkou $0\notin\overline{X_1-X_2}$.

Toto tvrzení ukazuje hlavní rozdíl mezi vlastní a silnou oddělitelností dvou množin: v případě silné konvexnosti lze vzít $\beta\in\big(\sup_{x_2\in X_2}\langle p,x_2\rangle,\inf_{x_1\in X_1}\langle p,x_1\rangle\big)$, tj. obě množiny mají kladnou vzdálenost od odpovídající oddělující nadroviny a jsou tedy nutně i disjunktní, zatímco v případě vlastní oddělitelnosti může být průnik neprázdný.

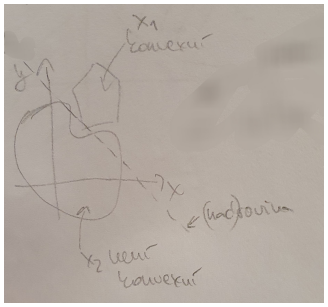
Důkaz Věty 6.3.8. „$\Longrightarrow$“ Nechť X_1,X_2 jsou silně oddělitelné[104], pak

$$\begin{aligned}0<\varepsilon=:\inf_{x_1\in X_1}\langle p,x_1\rangle-\sup_{x_2\in X_2}\langle p,x_2\rangle &\overset{\inf A-\sup B=\inf(A-B)\ [105]}{=}\inf_{(x_1,x_2)\in X_1\times X_2}\langle p,x_1-x_2\rangle \\ &\overset{\text{Cauchy–Schwarz}}{\leq}\inf_{(x_1,x_2)\in X_1\times X_2}\|p\|\times\|x_1-x_2\|=\|p\|\,\rho(X_1,X_2),\end{aligned}$$

a tedy $\rho(X_1,X_2)\geq\varepsilon/\|p\|>0$.

„$\Longleftarrow$“ Nechť X_1,X_2 jsou konvexní množiny a platí $\rho(X_1,X_2)>0$. Označme $X:=\overline{X_1-X_2}$. Pak[106] množina X je konvexní a uzavřená, viz Věty 6.1.3(ii) a 6.1.31. Protože $\rho(X_1,X_2)>0$ právě tehdy, když $\rho(\overline{X_1},\overline{X_2})>0$, máme[107] $0\notin X$. Pak ale $p:=\Pi_X(0)\neq 0$ a ukážeme, že právě tento vektor

[104] V tomto případě dokonce množiny X_1 a X_2 nemusí být ani konvexní. Avšak obecně samozřejmě v případě nekonvexní množiny už tvrzení neplatí, jak ukazuje obrázek níže, tj. množiny s nenulovou vzdáleností nemusí být oddělitelné, pokud alespoň jedna z nich není konvexní.



[105] Jelikož $\inf C=-\sup(-C)$, platí $\inf A-\sup B=\inf A+\inf(-B)=\inf(A-B)$.

Věta 6.3.8 o silně oddělující nadrovině

[106] Pozor, obecně platí pouze $\overline{X_1}-\overline{X_2}\subseteq\overline{X_1-X_2}$. Rovnost $\overline{X_1}-\overline{X_2}=\overline{X_1-X_2}$ platí např. v případě, že alespoň jedna z množin je ohraničená.

[107] Vskutku, připusťme, že $0\in X$. Pak buď $0\in X_1-X_2$ nebo $0\in\overline{X_1-X_2}\setminus(X_1-X_2)$. Jenže v prvním případě musí existovat $x\in X_1\cap X_2$, což je spor s $\rho(X_1,X_2)>0$. Ve druhém případě musí existovat posloupnost $\{x_k\}_{k\in\mathbb{N}}\subseteq X_1-X_2$ taková, že $x_k\to 0\in\overline{X_1-X_2}\setminus(X_1-X_2)$ pro $k\to\infty$. Pak ale $x_k=x_k^{[1]}-x_k^{[2]}$ pro všechna $k\in\mathbb{N}$, kde $x_k^{[1]}\in X_1$ a $x_k^{[2]}\in X_2$, a tudíž že $x_k^{[1]}\to x\in\overline{X_1}$ a současně také $x_k^{[2]}\to x\in\overline{X_2}$. Jenže pak existuje $x\in\overline{X_1}\cap\overline{X_2}$, což je spor s tím, že $0<\rho(X_1,X_2)=\rho(\overline{X_1},\overline{X_2})$.

splňuje požadavky silné oddělitelnosti, tj. p je normálový vektor (silně) oddělující nadroviny. Podle Lemma 6.3.7 je vektor p určen jednoznačně a z nerovnosti (6.3.4) s $x^* := 0$ máme

$$\langle p, x\rangle \geq \|p\|^2 > 0$$

pro všechna $x \in X$, z čehož ihned vyplývá $\inf_{x\in X}\langle p, x\rangle > 0$. Současně pro libovolné $x_1 \in X_1$ a $x_2 \in X_2$ je $x := x_1 - x_2 \in X$ a platí

$$\inf_{x\in X}\langle p, x\rangle = \inf_{(x_1,x_2)\in X_1\times X_2}\langle p, x_1 - x_2\rangle = \inf_{x_1\in X_1}\langle p, x_1\rangle - \sup_{x_2\in X_2}\langle p, x_2\rangle.$$

Tudíž

$$\inf_{x_1\in X_1}\langle p, x_1\rangle - \sup_{x_2\in X_2}\langle p, x_2\rangle > 0,$$

z čehož plyne uvedené tvrzení. ■

Poznámka 6.3.9. Podmínky Věty 6.3.8 jsou splněny zejména v případě, když

(i) $X_1 = \{x\}$ a $x \notin \overline{X_2}$, nebo

(ii) $\overline{X_1} \cap \overline{X_2} = \emptyset$ a X_1 nebo X_2 je ohraničená.

Ovšem pozor! Zde nestačí, aby množiny X_1, X_2 byly pouze konvexní, disjunktní a uzavřené. Uvážíme-li např.

$$X_1 := \{[x_1, x_2] \in \mathbb{R}^2 \mid x_1 > 0,\ x_2 \geq 1/x_1\} \quad \text{a} \quad X_2 := \{[x_1, x_2] \in \mathbb{R}^2 \mid x_1 \leq 0\},$$

viz Obrázek 6.60. Tyto množiny nejsou silně oddělitelné, ale pouze vlastně oddělitelné pouze pomocí nadroviny $H_{p,\beta}$ s $p = (1,0)^\top$ a $\beta = 0$, pro níž platí

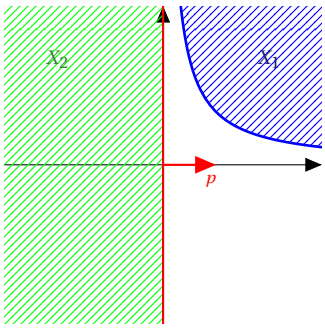
$$\inf_{x\in X_1}\langle p, x\rangle = \inf_{x_1>0} x_1 = 0 \quad \text{a} \quad \sup_{x\in X_2}\langle p, x\rangle = \sup_{x_1\leq 0} x_1 = 0.$$

Avšak bude-li jedna z množin X_1 nebo X_2 dokonce kompaktní (tj. navíc ohraničená), pak je již bude možné silně oddělit, jak ukazuje následující důsledek, který celkem bezprostředně vyplývá[108] z Věty 6.3.8. ▲

[108] Uvedené podmínky totiž garantují splnění předpokladů Věty 6.3.8, neboť platí následující tvrzení: *Nechť (P,ρ) je metrický prostor, množina $A \subseteq P$ je uzavřená a množina $B \subseteq P$ kompaktní. Pokud $A \cap B = \emptyset$, pak $\rho(A,B) > 0$.* Viz [141, Cvičení 3.18(i)].

Důsledek 6.3.10

Nechť $X_1, X_2 \subseteq \mathbb{R}^n$ jsou neprázdné, konvexní a disjunktní množiny. Nechť navíc X_1 je uzavřená a X_2 kompaktní. Potom jsou množiny X_1, X_2 silně oddělitelné.



Obrázek 6.60: „Silně neoddělitelné“ konvexní, disjunktní a uzavřené množiny X_1 a X_2.

Větu 6.3.8 či Důsledek 6.3.10 lze považovat za jistou variaci[109] na tzv. *geometrickou Hahnovu–Banachovu větu* či *Mazurovu* případně *Ascoliho–Mazurovu větu*, která rozšiřuje tento výsledek na obecný (i nekonečně dimenzionální) topologický vektorový prostor: *je-li M vektorový podprostor topologického vektorového prostoru X a $K \subseteq X$ je neprázdná konvexní otevřená podmnožina splňující $M \cap K = \emptyset$, pak existuje uzavřená nadrovina $N \subseteq X$ obsahující M, ale stále disjunktní s K*, viz např. [338, str. 38]. Takovéto výsledky[110] mají široké uplatnění v konvexní geometrii a jsou považovány za její základní princip.

[109] Podobně jako Větu 6.3.16 níže.

Věta 6.3.11

Navíc s pomocí Věty 6.3.8 můžeme podat již dříve slíbený „umělecký/vnější“ popis uzavřených konvexních množin jakožto alternativu k Větě 6.1.9. Její důkaz je založen na silné oddělitelnosti libovolné *uzavřené* konvexní množiny a jednobodové množiny vně tuto množinu, což jsou zjevně množiny s nenulovou vzdáleností.

Libovolná uzavřená konvexní množina $X \subseteq \mathbb{R}^n$ je řešením (konečné/nekonečné) soustavy neostrých lineárních nerovnic.

[110] Už jste dříve potkali Hahnovu-Banachovu větu? Jedná se o klíčový výsledek ve funkcionální analýze, který zaručuje prodloužitelnost libovolného ohraničeného lineárního funkcionálu (tj. lineárního zobrazení $f : X \to \mathbb{R}$ na libo-

Toto tvrzení lze formulovat také geometricky: každá uzavřená konvexní množina $X \subsetneq \mathbb{R}^n$ je průnikem uzavřených poloprostorů (konkrétně všech uzavřených poloprostorů obsahujících X), viz Obrázky 6.61–6.62.

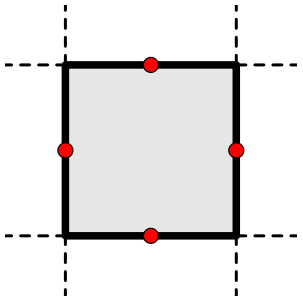
volném vektorovém podprostoru X), aniž by se změnila norma funkcionálu. V našem případě je oním funkcionálem $\langle p,x\rangle$ a neméně důležitou roli hraje tzv. *Minkowského funkcionál*

$$p_K(x) := \inf\{t>0 \mid x/t \in K\}$$

definovaný na konvexní množině $K \subseteq X$ s $0 \in \operatorname{int} K$. Ptáte se, co je p_K? V jistém smyslu měří vzdálenost x od počátku, avšak jednotka vzdálenosti záleží na zvoleném x – přesněji řečeno jednotkou je vzdálenost x od hranice K ve směru určeném bodem x. Uvažme např. elipsu $K = \{[x,y] \in \mathbb{R}^2 \mid \frac{x^2}{4} + y^2 \le 1\}$. Pak $p_K(3,0) = 3/2$, neboť $\frac{2}{3}[3,0] = [2,0]$ je bod ležící na hranici K a současně na polopřímce určené počátkem a bodem $[3,0]$. Podobně $p_K(1,0) = 1/2$, neboť $2\,[1,0] = [2,0]$ je opět vhodný bod ležící na hranici K. Konečně $p_K(0,1) = 1$, neboť $[0,1]$ leží na hranici K. To znamená, že hodnota $p_K(x) = 1$ odpovídá bodům na hranici K, zatímco $p_K(x) < 1$ nastane pro vnitřní body a $p_K(x) > 1$ pro vnější body (nakreslete si obrázek s množinou K a jednotlivými body). Důležitý je také fakt, že každou nadrovinu lze vyjádřit jako

$$f^{-1}(\beta) := \{x \mid f(x) = \beta\}$$

pro vhodný lineární funkcionál f. Více podrobností již uvádět nebudeme a odkážeme čtenáře raději na [422, Podkapitola 5.5] nebo případně [81, Podkapitola 1.2].



Obrázek 6.61: Čtverec jako konvexní množina X a čtveřice přímek určující čtyři odlišné poloprostory, jejichž průnik je právě X (tyto čtyři poloprostory stačí).

Důkaz. Je-li $X = \emptyset$, je tvrzení triviální, neboť se jedná o průnik dvou vhodně zvolených (tj. disjunktních) uzavřených poloprostorů. Je-li $X = \mathbb{R}^n$, pak je tvrzení také triviální, neboť se jedná o řešení prázdného systému, tj. $Ax \le b$ pro $A = 0$ a $b = 0$.

Nechť tedy $\emptyset \ne X \subsetneq \mathbb{R}^n$ je uzavřená konvexní množina. Je-li $x \notin X$ libovolné, pak díky uzavřenosti množiny X má bod x kladnou vzdálenost od X, a tudíž podle Věty 6.3.8 existuje nadrovina, která silně odděluje množiny $\{x\}$ a X, tj. pro každé $x \notin X$ existuje vektor $p_x \in \mathbb{R}^n \setminus \{0\}$ splňující

$$\langle p_x, x\rangle > \alpha_x := \sup_{y\in X}\langle p_x, y\rangle.$$

Současně pro každé $x \notin X$ obsahuje uzavřený poloprostor

$$H_x := \{y \in \mathbb{R}^n \mid \langle p_x, y\rangle \le \alpha_x\}$$

jistě celou množinu X a neobsahuje bod x. Proto X není větší (a samozřejmě ani menší) než průnik všech uzavřených poloprostorů, které obsahují X, tj.

$$X = \bigcap_{x\in X} H_x.$$

■

Mezi všemi uzavřenými poloprostory obsahujícími uzavřenou konvexní a netriviální[111] množinu X jsou velmi zajímavé ty „extrémní“, tj. takové jejichž hraniční nadrovina se dotýká množiny X. Tato terminologie je použitelná pro libovolnou (nikoli nutně uzavřenou) konvexní množinu, což je obsahem následující definice. Následně si s pomocí těchto nadrovin odvodíme postačující podmínku pro oddělitelnost dvou konvexních množin (Věta 6.3.16) a ekvivalentní charakterizaci vlastní oddělitelnosti dvou konvexních množin (Věta 6.3.18).

DEFINICE 6.3.12

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná množina a nechť $a \in \partial X := \overline{X} \setminus \operatorname{int} X$. Nadrovinu $H_{p,\beta}$ nazveme

(i) *opěrnou nadrovinou* množiny X v bodě a, jestliže

$$\langle p,x\rangle \ge \beta = \langle p,a\rangle \quad \text{pro každé } x \in X;$$

(ii) *vlastní opěrnou nadrovinou* množiny X, jestliže je opěrnou nadrovinou množiny X v bodě a a existuje-li $x^* \in X$ takové, že

$$\langle p,x^*\rangle > \beta.$$

Příklad 6.3.13

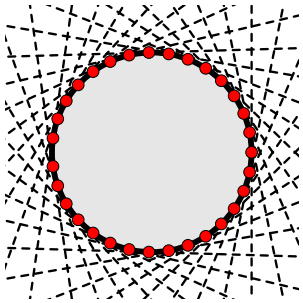
V případě opěrné nadroviny jde vlastně o to, aby množina X ležela v jednom z poloprostorů[112] vymezeném nadrovinou $H_{p,\beta}$ procházející bodem a, viz Obrázky 6.63–6.67. Posuneme-li množinu X tak, aby $a = 0$, potom se podmínka z definice opěrné nadroviny redukuje na $\langle p,x\rangle \ge \langle p,a\rangle = 0$, tj. vektory určené body $x \in X$ musí svírat s vektorem p úhel v rozmezí $[-\pi/2, \pi/2]$, a tedy vektor p musí směřovat do poloprostoru obsahujícího množinu X, viz Obrázek 6.68.

Jak je to s existencí opěrné nadroviny? Uvažme pro jednoduchost úsečku $X = \{[x,0] \mid 0 \le x \le 1\}$ v $\mathbb{R}^2$. Pak $\operatorname{int} X = \emptyset$, $\partial X = X$ a $\mathrm{r}\partial X = \{[0,0],[1,0]\}$. V každém bodě $x \in \partial X \setminus \mathrm{r}\partial X$ existuje jenom jedna opěrná nadrovina, kterou je osa x, zatímco v bodech $x \in \mathrm{r}\partial X$ jich existuje nekonečně mnoho (libovolná úsečka procházející uvažovaným bodem) a kromě osy x jsou všechny vlastními opěrnými nadrovinami. Toto samozřejmě není náhoda, jak potvrzuje následující tvrzení.

VĚTA 6.3.14

Nechť $X \subseteq \mathbb{R}^n$ je neprázdná konvexní množina a nechť $a \in \mathrm{r}\partial X$. Pak v bodě a existuje vlastní opěrná nadrovina množiny X.

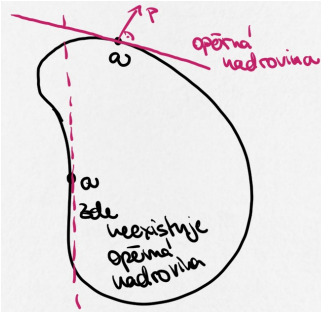
Poznámka 6.3.15. Jelikož buď $\operatorname{ri} X = \operatorname{int} X$ nebo $\operatorname{int} X = \emptyset$, nutně platí $\operatorname{r}\partial X \subseteq \partial X$. Uvážíme-li místo bodů z relativní hranice $\operatorname{r}\partial X$ body ze samotné hranice ∂X, už pro ně nebude možné zaručit existenci vlastní opěrné nadroviny[113], avšak opěrná nadrovina bude stále existovat, tj. můžeme tvrzení Věty 6.3.14 *rozšířit*: pro neprázdnou konvexní množinu X a libovolný bod $a \in \partial X$ existuje[114] opěrná nadrovina množiny X. ▲



Obrázek 6.62: Kruh jako konvexní množina X a několik přímek určující odlišné poloprostory, jejichž průnikem (potřebovali bychom jich nekonečně mnoho) bychom mohli dostat právě X.

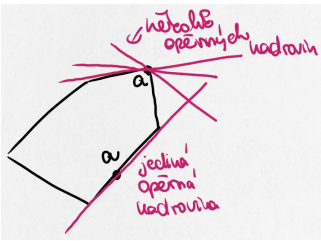
[111] Tj. $\emptyset \neq X \neq \mathbb{R}^n$.

[112] Konkrétní poloprostor je určen volbou vektoru p, tj. přechodem k $-p$ získáme opačný poloprostor.



Obrázek 6.63: Ukázka opěrné nadroviny.

Věta 6.3.16 O oddělující nadrovině[115]



Obrázek 6.64: Ukázka opěrné nadroviny konvexní množiny.

Důkaz Věty 6.3.14. Popíšeme konstrukci normálového vektoru opěrné nadroviny. Nechť $a \in \operatorname{r}\partial X$ je libovolný bod. Pak existuje posloupnost $\{a_k\}_{k\in\mathbb{N}}$ taková, že $a_k \in \operatorname{aff} X \setminus X$ a $a_k \to a$ pro $k \to \infty$, viz Obrázek 6.69. Potom podle Lemma 6.3.7 existuje $\Pi_X(a_k)$ pro každé $k \in \mathbb{N}$ a platí $\Pi_X(a_k) \neq a_k$. Položme

$$p_k := \frac{\Pi_X(a_k) - a_k}{\|\Pi_X(a_k) - a_k\|}.$$

Pak $\|p_k\| = 1$ a $p_k \in \operatorname{Lin} X$ v souladu s Definicí 6.1.10.
Protože množina $\{y \in \mathbb{R}^n \mid \|y\| = 1\}$ je kompaktní, můžeme z posloupnosti $\{p_k\}_{k\in\mathbb{N}}$ vybrat konvergentní podposloupnost. Bez újmy na obecnosti proto předpokládejme, že $p_k \to p$, přičemž $\|p\| = 1$ a $p \in \operatorname{Lin} X$ díky uzavřenosti $\operatorname{Lin} X$. Je toto hledaný normálový vektor? Z nerovnosti (6.3.4) plyne pro libovolné $x \in X$, že

$$\begin{aligned}\langle p_k, x\rangle - \langle p_k, a_k\rangle = \langle p_k, x - a_k\rangle &= \frac{1}{\Pi_X(a_k) - a_k}\langle \Pi_X(a_k) - a_k, x - a_k\rangle \\ &\overset{(6.3.4)}{\geq} \frac{1}{\Pi_X(a_k) - a_k}\|\Pi_X(a_k) - a_k\|^2 = \|\Pi_X(a_k) - a_k\| > 0,\end{aligned}$$

neboť $\Pi_X(a_k) \neq a_k$. Limitním přechodem pro $k \to \infty$ dostáváme

$$\langle p, x\rangle \geq \langle p, a\rangle =: \beta \quad \text{pro každé } x \in X,$$

což znamená, že $H_{p,\beta}$ je hledaná opěrná nadrovina.

A je vlastní? Podle Věty 6.1.29 je $\operatorname{ri} X \neq \emptyset$, a tedy existuje $x_1 \in \operatorname{ri} X$, tj. $\mathcal{O}_\varepsilon(x_1) \cap \operatorname{aff} X \subseteq X$ pro nějaké $\varepsilon > 0$. Položme $\overline{x} := x_1 - \overline{\varepsilon} p$ pro nějaké dostatečně malé $\overline{\varepsilon} > 0$. Protože $p \in \operatorname{Lin} X$, je $\overline{x} \in \operatorname{aff} X$. Z definice $\operatorname{ri} X$ navíc vyplývá, že pro dostatečně malé $\overline{\varepsilon} > 0$ je dokonce $\overline{x} \in \operatorname{ri} X$, a tedy také $\overline{x} \in X$. Proto

$$\beta \leq \langle p, \overline{x}\rangle = \langle p, x_1 - \overline{\varepsilon} p\rangle = \langle p, x_1\rangle - \overline{\varepsilon}\langle p, p\rangle = \langle p, x_1\rangle - \overline{\varepsilon},$$

tj. $\langle p, x_1\rangle \geq \beta + \overline{\varepsilon} > \beta$, což dokazuje, že nadrovina $H_{p,\beta}$ je vlastní opěrnou nadrovinou. ■

Tento výsledek nyní využijeme v důkazech tvrzení o oddělitelnosti konvexních množin. Nejdříve začneme se samotnou oddělitelností, srovnej s Větou 6.3.8.

> Nechť $X_1, X_2 \subseteq \mathbb{R}^n$ jsou neprázdné, konvexní a disjunktní množiny. Pak pro tyto množiny existuje oddělující nadrovina.

Důkaz. Uvažme konvexní množinu

$$X := X_1 - X_2 = \{x \in \mathbb{R}^n \mid x = x_1 - x_2 \text{ pro nějaké } x_1 \in X_1,\ x_2 \in X_2\}.$$

Protože $X_1 \cap X_2 = \emptyset$, máme $0 \notin X$ a mohou tedy nastat pouze dvě možnosti:

(i) Jednak $0 \notin \overline{X}$. Pak množiny $\{0\}$ a X jsou silně oddělitelné, tj. existuje $p \in \mathbb{R}^n \setminus \{0\}$ takový, že

$$\inf_{x\in X} \underbrace{\langle p, x\rangle}_{=\langle p, x_1 - x_2\rangle} > \sup_{y\in\{0\}}\langle p, y\rangle = 0,$$

tj. $\inf_{x_1\in X_1}\langle p, x_1\rangle > \sup_{x_2\in X_2}\langle p, x_2\rangle$, a tedy množiny X_1, X_2 jsou (dokonce silně) oddělitelné.

(ii) Jednak $0 \in \overline{X} \setminus X \subseteq \partial X$. Pak podle Věty 6.3.14 existuje opěrná nadrovina, tj. existuje $p \in \mathbb{R}^n \setminus \{0\}$ takový, že

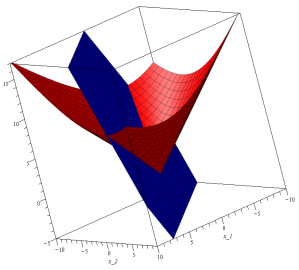
$$\langle p, x\rangle \geq \beta := \langle p, 0\rangle = 0$$

pro všechna $x \in X$. Odtud plyne, že $\langle p, x_1\rangle \geq \langle p, x_2\rangle$ pro všechna $x_1 \in X_1$ a $x_2 \in X_2$, takže množiny X_1 a X_2 jsou oddělitelné. ■

Následující příklad se věnuje situaci, kdy chceme oddělit konvexní množinu a afinní množinu.

Příklad 6.3.17

Uvažme neprázdnou konvexní množinu $X \subseteq \mathbb{R}^n$ a afinní množinu $Y := \{Ax+b \mid x \in \mathbb{R}^m\}$ pro nějakou $A \in \mathbb{R}^{n\times m}$ a $b \in \mathbb{R}^n$. Je nějaký vztah mezi $x \in X$ a vektorem b? Jsou-li tyto dvě množiny disjunktní, pak podle Věty 6.3.16 existuje $p \in \mathbb{R}^n \setminus \{0\}$ a $\beta \in \mathbb{R}$ tak, že $\langle p, x\rangle \le \beta$ pro všechna $x \in X$ a současně $\langle p, y\rangle \ge \beta$ pro všechna $y \in Y$. To ale znamená, že $\langle p, Ax+b\rangle \ge \beta$ pro všechna $x \in \mathbb{R}^m$ neboli $\langle p, Ax\rangle \ge \beta - p^\top b$. Na levé straně máme lineární funkci vzhledem k $x \in \mathbb{R}^m$, která je ohraničená zdola konstantním výrazem $\beta - p^\top b$. Avšak to je možné pouze v případě nulové funkce, tj. vektor p je takový, že $p^\top A = 0$, tj. $p^\top b \ge \beta$. To znamená, že normálový vektor oddělující nadroviny musí být takový, že $p^\top x \le p^\top b$ pro všechna $x \in X$.



Obrázek 6.65: Graf funkce $f(x_1,x_2) = \sqrt{x_1^2 - x_1x_2 + x_2^2}$ a rovina $z = 2(x_1 - x_2)$.

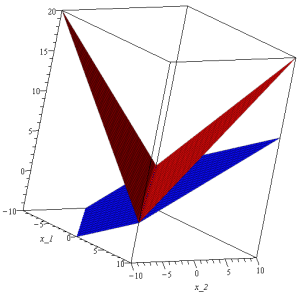
Pochopitelně opačné tvrzení k Větě 6.3.16 neplatí, tj. oddělitelnost dvou konvexních množin neimplikuje jejich disjunktnost, jak jsme mohli vidět v Příkladě 6.3.2. Je tedy potřeba přidat nějaký dodatečný požadavek nad rámec konvexnosti – nejjednodušší možností je otevřenost některé z množin X_1, X_2, tj. jsou-li X_1 a X_2 neprázdné konvexní množiny a X_1 je otevřená, pak jsou tyto množiny disjunktní právě tehdy, když jsou oddělitelné. Ještě lepší tvrzení obdržíme, využijeme-li místo otevřenosti X_1 raději relativní vnitřky množin X_1 a X_2.

Věta 6.3.18 O vlastně oddělující nadrovině

Neprázdné konvexní množiny $X_1, X_2 \subseteq \mathbb{R}^n$ jsou vlastně oddělitelné právě tehdy, když $\operatorname{ri} X_1 \cap \operatorname{ri} X_2 = \emptyset$.

Důkaz. „$\Longrightarrow$“ Nechť X_1, X_2 jsou vlastně oddělitelné nadrovinou $H_{p,\beta}$ a předpokládejme, že existuje $x \in \operatorname{ri} X_1 \cap \operatorname{ri} X_2$. Z definice vlastní oddělitelnosti vyplývá, že existují $\overline{x}_1 \in X_1$ a $\overline{x}_2 \in X_2$ splňující

$$\langle p, \overline{x}_1\rangle > \langle p, \overline{x}_2\rangle, \quad \text{tj. } \langle p, \overline{x}_1 - \overline{x}_2\rangle > 0. \tag{6.3.7}$$



Obrázek 6.66: Graf funkce $f(x_1,x_2) = \sqrt{x_1^2 - 2x_1x_2 + x_2^2}$ a rovina $z = (x_1 - x_2)/2$.

Pro $\alpha \in [-1,0]$ máme

$$\widetilde{x}_1 := x - \alpha(\overline{x}_1 - x) = -\alpha\overline{x}_1 + (1+\alpha)x \in X_1,$$
$$\widetilde{x}_2 := x - \alpha(\overline{x}_2 - x) = -\alpha\overline{x}_2 + (1+\alpha)x \in X_2.$$

a podle Věty 6.1.30 dokonce také $\widetilde{x}_1 \in X_1$ a $\widetilde{x}_2 \in X_2$ pro $\alpha > 0$ dostatečně malé, neboť

$$\widetilde{x}_1 = x + (\gamma - 1)(x - \overline{x}_1) \quad \text{a} \quad \widetilde{x}_2 = x + (\gamma - 1)(x - \overline{x}_2)$$

pro $\gamma := 1 + \alpha$. Pak ale pro toto $\alpha > 0$ dostaneme

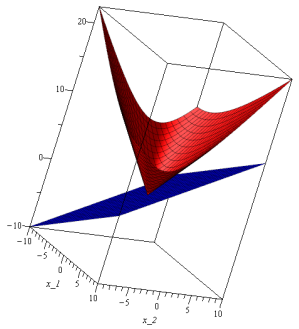
$$\langle p, \widetilde{x}_1\rangle - \langle p, \widetilde{x}_2\rangle = \langle p, \widetilde{x}_1 - \widetilde{x}_2\rangle = \langle p, -\alpha(\overline{x}_1 - x) + \alpha(\overline{x}_2 - x)\rangle = -\alpha\langle p, \overline{x}_1 - \overline{x}_2\rangle \overset{(6.3.7)}{<} 0,$$

což je spor s oddělitelností množin X_1 a X_2.[116]

„$\Longleftarrow$“ Nechť $\operatorname{ri} X_1 \cap \operatorname{ri} X_2 = \emptyset$. Položme $X := \operatorname{ri} X_1 - \operatorname{ri} X_2$. Pak X je konvexní množina[117] a platí $0 \notin X$, přičemž mohou nastat dvě možnosti[118]:

(i) Jednak $0 \notin \overline{X}$. Potom podle Věty 6.3.8 jsou množiny $\{0\}$ a $\overline{X}$ silně oddělitelné a tato oddělující nadrovina vlastně (dokonce i silně) odděluje i množiny X_1 a X_2.

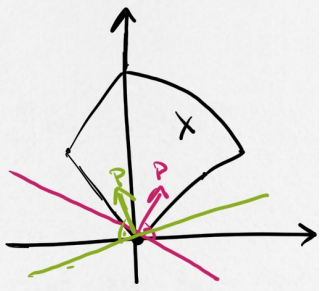
(ii) Jednak $0 \in \overline{X} \setminus X \subseteq \overline{X} \setminus \operatorname{ri} X = \operatorname{r\partial} X$. Potom podle Věty 6.3.14 existuje v bodě 0 vlastní opěrná nadrovina množiny X a právě tato nadrovina vlastně odděluje množiny X_1 a X_2. ■



Obrázek 6.67: Graf funkce $f(x_1,x_2) = \sqrt{x_1^2 + 2x_1x_2 + 2x_2^2}$ a rovina $z = (x_1 + x_2)/2$.

Proč by ale tyto výsledky měly být tak zásadní pro konvexní analýzu/geometrii či optimalizaci? Jeden z důvodů si nyní ukážeme. V Podkapitole 2.1 jsme se již zmínili o nezájmu o soustavy lineárních nerovnic až do příchodu teorie her a lineárního programování, které podnítili rozvoj nových metod pro jejich řešení. Značná část starší literatury se spíše věnovala nalezení podmínek, za kterých je možné vyřešit obecnou homogenní soustavu lineárních nerovnic. Všechny

dosažené výsledky vyjadřují v jedné anebo druhé formě vztah mezi původní (*primární*) soustavou a jinou (*duální*) soustavou, která používá sloupce původní matice koeficientů na vytvoření nových lineárních rovnic anebo nerovnic podle jistých pravidel. Typická je věta od Gordana z roku 1873 dokazující, že homogenní soustava rovnic s nezápornými proměnnými má řešení s alespoň jednou kladnou proměnnou, pokud duální soustava nemá řešení při ostrých nerovnostech, viz [234]. Stiemke v roce 1915 přidal větu o existenci řešení při všech kladných proměnných, viz [462]. I další matematici se věnovali tomuto problému: zejména Motzkin (1936), Ville (1938), von Neumann a Morgenstern (1944), viz [385, 498, 503] a zejména [323]. Tyto věty v podstatě hovoří o tom, že buď původní soustava má netriviální řešení, anebo duální soustava má řešení pro ostré nerovnosti. Pro toto „buď… anebo…“ zavedli von Neumann a Morgenstern označení *věty o alternativě*[119].



Obrázek 6.68: Ukázka opěrné nadroviny konvexní množiny.

[113] Vezmeme-li $X = \{a\}$, pak nikdy nebude existovat vlastní opěrná nadrovina. Podle definice bychom totiž požadovali, aby existovalo $x \in X$ takové, že $\langle p, x\rangle > \langle p, a\rangle$, což je očividně nemožné. Každopádně zde existuje opěrná nadrovina, neboť nerovnost $\langle p, x\rangle \geq \langle p, a\rangle$ je splněna triviálně pro každé $x \in X$ a $p \in \mathbb{R}^n \setminus \{0\}$. Mimochodem, v tomto velmi patologickém případě máme $\operatorname{ri} X = \{a\}$, a tedy $a \notin \operatorname{r\partial} X$, ale $a \in \partial X$.

[114] *Důkaz.* Je-li $\operatorname{int} X = \emptyset$, pak $\operatorname{aff} X \neq \mathbb{R}^n$ a $\operatorname{aff} X$ je průsečík několika nadrovin, neboť $\operatorname{aff} X = \{x \in \mathbb{R}^n \mid Ax = b\}$. Potom ale nutně některá z těchto nadrovin je právě opěrnou (ne nutně vlastní) nadrovinou pro $a \in \partial X \subseteq \overline{X}$. Je-li $\operatorname{int} X \neq \emptyset$, pak $\partial X = \operatorname{r\partial} X$ a tvrzení plyne přímo z Věty 6.3.14.

[115] Angl. *Separating Hyperplane Theorem*. Autorství této věty (příp. v kombinaci s Důsledkem 6.3.10) bývá připisováno Minkowskému, kterého bychom tak mohli považovat za zakladatele této oblasti konvexní geometrie.

Z lineární algebry např. známe tvrzení: *Pokud každé řešení soustavy lineárních rovnic vyhovuje i dané lineární rovnici, je možné tuto rovnici získat jako lineární kombinaci rovnic soustavy.* Ovšem věta od Farkase z roku 1902 říká (překvapivě?), viz [166]:

> Pokud každé řešení soustavy lineárních homogenních nerovností vyhovuje i dané lineární nerovnosti (všechny nerovnosti jsou tvaru ≥ 0), pak je možné tuto nerovnost získat jako nezápornou kombinaci nerovností soustavy.

Farkas se studiu lineárních nerovnic věnoval systematičtěji, ačkoli bez geometrické interpretace. Z lineární algebry nepochybně víme, že pro homogenní soustavy rovnic platí, že obecné řešení je lineární kombinací *bazických řešení*. Podobné tvrzení platí i pro homogenní soustavy lineárních nerovnic[120], jak v roce 1896 ukázal Minkowski, viz [380]:

> Uvažme soustavu $Ax \geq 0$, kde $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ a nerovnost je míněna po složkách. Má-li soustava nenulová řešení, pak tvoří „pyramidu“[121] v $\mathbb{R}^n$

Je zřejmé, že soustava $Ax \geq 0$ má triviální řešení. Také s každým řešením $\widehat{x}$ bude řešením i $\lambda\widehat{x}$ pro libovolné $\lambda \geq 0$, přičemž pro $\lambda > 0$ bude toto řešení netriviální. Minkowski nepovažoval dvě netriviální řešení za *podstatně odlišná*, pokud jedno je kladným násobkem druhého. *Extrémní* či *fundamentální řešení* pak byla definována jako taková netriviální řešení, která nelze napsat jako součet dvou podstatně odlišných řešení. Minkowski ukázal, že existuje pouze konečně mnoho extrémních řešení (bodů na pyramidě) takových, že libovolné řešení je nějakou nezápornou kombinací těchto řešení[122]. Odtud Minkowski vyvodil (tzv. *Farkasovo lemma*):

> Pokud každé řešení soustavy $Ax \leq 0$ splňuje $\langle c, x\rangle \leq 0$, pak $yA = c^\top$ pro nějaké $y \geq 0$.

My si nyní jednu z vět o alternativě odvodíme právě s využitím Věty 6.3.18. S její pomocí lze následně např. zodpovědět otázku řešitelnosti úloh lineárního programování, jak jsem mohli vidět v Kapitole 2, viz Důsledek 2.4.5 a Větu 2.4.6.

Věta 6.3.19 Farkas[123] & Minkowski[124]

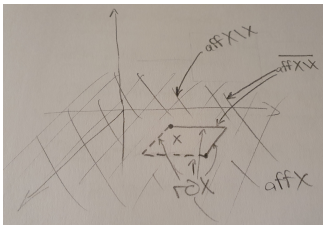
Nechť $A \in \mathbb{R}^{m\times n}$ a $b \in \mathbb{R}^m$. Potom je právě jeden z následujících systémů rovnic a nerovnic řešitelný:

$$Ax = b, \quad x \geq 0, \tag{6.3.8}$$

$$A^\top y \geq 0, \quad \langle y, b\rangle < 0. \tag{6.3.9}$$

Obměnou tvrzení Věty 6.3.19 obdržíme ekvivalentní výsledek:

> Soustava (6.3.8) má řešení právě tehdy, když pro všechna $y \in \mathbb{R}^m$ splňující $A^\top y \geq 0$ platí $\langle y, b\rangle \geq 0$.



Obrázek 6.69: Množinu $\operatorname{aff} X \setminus X$ získáme odebráním celého X z uzavřené nadmnožiny $\operatorname{aff} X$. Pak přidáním (ne nutně celé) $\operatorname{r}\partial X \subseteq \partial X$ získáme $\overline{\operatorname{aff} X \setminus X}$, takže $\operatorname{r}\partial(\overline{\operatorname{aff} X \setminus X}) = \operatorname{r}\partial X$. Proto množina $\operatorname{aff} X \setminus X$ musí obsahovat posloupnost $a_k \to a \in \operatorname{r}\partial X$.

[116] Tuto část tvrzení lze dokázat i s pomocí Věty 6.2.30. Z vlastní oddělitelnosti plyne

$$\inf_{x_1 \in X_1} \langle p, x_1 \rangle \geq \sup_{x_2 \in X_2} \langle p, x_2 \rangle$$

a současně (protože pro nějakou dvojici $x_1 \in X_1$ a $x_2 \in X_2$ musí nastat ostrá nerovnost)

$$\sup_{x_1 \in X_1} \langle p, x_1 \rangle > \inf_{x_2 \in X_2} \langle p, x_2 \rangle.$$

Jestliže existuje $x^* \in \operatorname{ri} X_1 \cap \operatorname{ri} X_2$, pak se v něm první nerovnost realizuje jako rovnost, takže tento bod maximalizuje lineární funkci $\langle p, x \rangle$ na množině X_2 a současně ji minimalizuje na X_1. Proto podle Věty 6.2.30 platí $\langle p, x \rangle \equiv \text{konst.}$ na $X_1 \cup X_2$, což je spor s druhou nerovností.

[117] Viz Věty 6.1.29 a 6.1.3(ii).

[118] Srovnej s důkazem Věty 6.3.16.

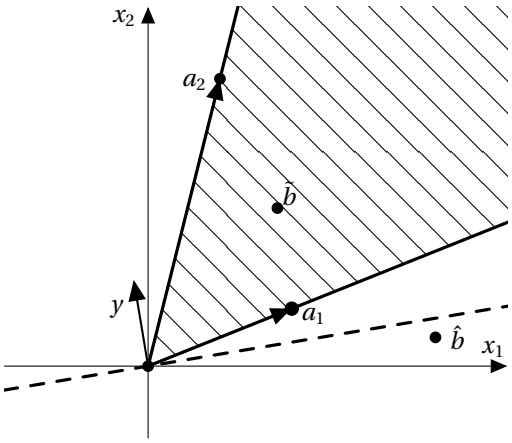
[119] Někdy se také hovoří o (Motzkinově) transpoziční větě vzhledem k přechodu od primární sousty k duální, viz např. [180, str. 2345–2348]

[120] Jak jsme se mimo jiné mohli přesvědčit v Kapitole 2.

[121] V naší terminologii bychom měli mluvit spíše o konvexním kuželu.

[122] Odtud už je jenom „kousek“ k zavedení již dříve zmíněného pojmu *extrémního* bodu konvexní množiny (ten nelze vyjádřit jako konvexní kombinaci bodů z dané množiny, tudíž po jeho odstranění je množina stále konvexní)

Avšak formulace ve formě alternativy má své neoddiskutovatelné kouzlo. Tato věta má také velmi elegantní geometrickou interpretaci, která současně vystihuje hlavní myšlenku jejího důkazu, jenž je uveden níže. Uvažme vektory $a_1, a_2 \in \mathbb{R}^2$ tvořící sloupce matice A a bod $b \in \mathbb{R}^2$. Pak splnění soustavy (6.3.8) pro nějaké $x^* \geq 0$ znamená, že bod b patří do kuželu vymezeného polopřímkami vycházejícími z počátku a se směrovými vektory a_1, a_2 (tj. tyto polopřímky prochází odpovídajícími body a_1 a a_2). Naopak, je-li pro nějaké y^* splněna soustava (6.3.9), pak přímka procházející počátkem a s normálovým vektorem y^* *odděluje* kužel a bod b. Je tedy zřejmé, že bod b má právě jednu z těchto vlastností, viz Obrázek 6.70.



Obrázek 6.70: Geometrický pohled na Farkasovu–Minkowského větu pro $b = \tilde{b}$ a $b = \hat{b}$.

Důkaz Věty 6.3.19. „$\Longrightarrow$“ Nechť $\overline{x} \in \mathbb{R}^n$ splňující $\overline{x} \geq 0$ řeší (6.3.8). Pak pro libovolné $y \in \mathbb{R}^m$ splňující $A^\top y \geq 0$ platí

$$\langle y, b \rangle = \langle y, A\overline{x} \rangle = \langle A^\top y, \overline{x} \rangle \geq 0,$$

neboť $A^\top y \geq 0$ a $\overline{x}$ (po složkách), tj. neplatí (6.3.9).

„$\Longleftarrow$“[125] Nyní předpokládejme, že soustava (6.3.8) nemá řešení. Ukážeme, že pak existuje $y \in \mathbb{R}^m$ splňující (6.3.9). Označme $a_1, \ldots, a_n \in \mathbb{R}^m$ sloupce matice A. Pak vektor $b \notin \operatorname{cone}\{a_1, \ldots, a_n\}$, protože jinak bylo možné vyjádřit b jako nezápornou kombinaci vektorů $a_1, \ldots, a_n$, tj. vyřešit (6.3.8). Protože $\operatorname{cone}\{a_1, \ldots, a_n\}$ je uzavřená množina, jsou množiny $\{b\}$ a $\operatorname{cone}\{a_1, \ldots, a_n\}$ silně oddělitelné, viz Věta 6.3.8. Nechť y je normálový vektor oddělující nadroviny, tj. $\langle y, z \rangle > \langle y, b \rangle$ (bez újmy na obecnosti tvar nerovnosti) pro všechna $z \in \operatorname{cone}\{a_1, \ldots, a_n\}$. Pak pro všechna $x \in \mathbb{R}^n$, $x \geq 0$, platí

$$\langle A^\top y, x \rangle = \langle y, Ax \rangle \overset{Ax \in \operatorname{cone}\{a_1, \ldots, a_n\}}{>} \langle y, b \rangle. \tag{6.3.10}$$

Dosadíme-li do (6.3.10) hodnotu $x = 0$, pak $0 > \langle y, b \rangle$. Pokud naopak budeme jednotlivé složky x zvětšovat nad všechny meze (do ∞), pak nerovnost (6.3.10) bude splněna pouze v případě, že $A^\top y \geq 0$. Kdyby totiž některá složka vektoru $A^\top y$ byla záporná, tj. $(A^\top y)_j < 0$ pro nějaké $j \in \{1, \ldots, n\}$, pak volbou $x_i = 0$ pro $i \in \{1, \ldots, n\} \setminus \{j\}$ a $x_j > 0$ dostaneme $\langle A^\top y, x \rangle = (A^\top y)_j x_j < 0$, což nás pro $x_j > 0$ dostatečně velké přivede ke sporu s nerovností (6.3.10). ■

Poznámka 6.3.20. Tvrzení Věty 6.3.19 lze také ekvivalentně formulovat takto, což se záhy ukáže být velmi užitečným[126]:

Jestliže systém

$$f_0(x) := \langle a_0, x \rangle < 0, \quad f_i(x) := \langle a_i, x \rangle \leq 0, \quad i \in \{1, \ldots, m\}$$

nemá pro daná $a_0, \ldots, a_m \in \mathbb{R}^n$ řešení na $\mathbb{R}^n$, pak existují čísla $y_1, \ldots, y_m \geq 0$ taková, že

$$a_0 + \sum_{i=1}^{m} y_i a_i = 0, \quad \text{tj. } f_0(x) + \sum_{i=1}^{m} y_i f_i(x) = 0$$

pro každé $x \in \mathbb{R}^n$.

▲

V souvislosti s větami o alternativě si v závěrečné části této podkapitoly uvedeme ještě jedno podobné tvrzení a některé jeho „regulární“ modifikace. Tyto výsledky budou hrát velmi důležitou roli při dokazování základní věty matematického programování, např. následující věta bude klíčovým nástrojem v důkazu Věty 8.2.1.

Věta 6.3.21 (Fan[127] & Glicksberg[128] & Hoffman[129])

Nechť množina $X \subseteq \mathbb{R}^n$ je neprázdná a konvexní, funkce $f_1, \dots, f_k : X \to \mathbb{R}$ jsou konvexní a funkce $f_{k+1}, \dots, f_m$ afinní, tj. pro $j \in \{k+1, \dots, m\}$ máme $f_j(x) = \langle a_j, x \rangle + \beta_j$ pro vhodná $a_j \in \mathbb{R}^n$ a $\beta_j \in \mathbb{R}$. Jestliže systém nerovností a rovností

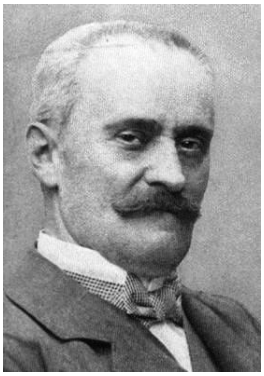
$$\left.\begin{aligned} f_i(x) &< 0, \quad i \in \{1, \dots, k\}, \\ f_j(x) &= 0, \quad j \in \{k+1, \dots, m\}, \end{aligned}\right\} \tag{6.3.11}$$

nemá řešení na X, pak existují takové konstanty

$$y_1, \dots, y_k \geq 0 \quad \text{a} \quad y_{k+1}, \dots, y_m \in \mathbb{R},$$

že alespoň pro jedno $\ell \in \{1, \dots, m\}$ je $y_\ell \neq 0$ a pro všechna $x \in X$ platí

$$\sum_{i=1}^{m} y_i f_i(x) \geq 0. \tag{6.3.12}$$



123 Gyula (nebo Julius) Farkas de Kisbarnak (28. března 1847 – 27. prosince 1930) byl maďarský matematik a fyzik. Jeho jméno je spojeno zejména s „Farkasovým lemmatem“, které je také známo jako základní princip jednoduchých nerovnic.

124 Hermann Minkowski (22. června 1864 – 12. ledna 1909) byl německý matematik německo-polsko-židovského původu, který postupně vyučoval na univerzitách v Bonnu, Göttingenu, Královci a Curychu. Proslul především vytvořením a rozpracováním tzv. geometrické teorie čísel, tj. převedením a vyřešením některých problémů z teorie čísel metodami klasické geometrie. Geometrické nástroje využíval také k řešení problémů matematické fyziky a teorie relativity. Minkowski je asi nejvíce znám díky své fundamentální práci popisující prostor a čas jako čtyřrozměrný prostor, který je nyní známý jako *Minkowského časoprostor* a který usnadnil geometrickou interpretaci Einsteinovy speciální teorie relativity z roku 1905.

Důkaz. Definujme množiny

$$\mathcal{A} := \{u \in \mathbb{R}^m \mid \text{existuje } \overline{x} \in X \text{ tak, že } f_i(\overline{x}) \leq u_i,\ i \in \{1, \dots, k\}\ \&\ f_j(\overline{x}) = u_j,\ j \in \{k+1, \dots, m\}\},$$
$$\mathcal{B} := \{v \in \mathbb{R}^m \mid v_i < 0,\ i \in \{1, \dots, k\}\ \&\ v_j = 0,\ j \in \{k+1, \dots, m\}\}.$$

Obě tyto množiny jsou zjevně neprázdné a množina $\mathcal{B}$ je očividně[130] i konvexní. Navíc, jsou-li $\overline{u}, \widetilde{u} \in \mathcal{A}$ libovolné a $\overline{x}, \widetilde{x} \in X$ nějaké odpovídající prvky z X z definice $\mathcal{A}$, pak pro $\lambda \in [0,1]$ platí[131], že $\lambda\overline{u} + (1-\lambda)\widetilde{u} \in \mathcal{A}$, tj. množina A je také konvexní. Protože systém (6.3.11) nemá řešení na X, platí $\mathcal{A} \cap \mathcal{B} = \emptyset$, což znamená, že množiny $\mathcal{A}, \mathcal{B}$ jsou podle Věty 6.3.16 oddělitelné, tj. existuje $y = (y_1, \dots, y_m)^\top \neq 0$ takové, že $\langle y, u \rangle \geq \langle y, v \rangle$ pro všechna $u \in \mathcal{A}$ a $v \in \mathcal{B}$, tj.

$$\sum_{i=1}^{m} y_i u_i \geq \sum_{i=1}^{k} y_i v_i.$$

Pokud zafixujeme $u \in \mathcal{A}$ a $v_1, \dots, v_k \to -\infty$, vidíme, že předchozí nerovnost může být splněna pouze tehdy, když $y_1, \dots, y_m \geq 0$. Dosadíme-li do této nerovnosti $u_i := f_i(\overline{x})$ pro libovolné $\overline{x} \in X$ a všechna $i \in \{1, \dots, m\}$, pak pro $v_1, \dots, v_k \to 0^-$ dostaneme požadované tvrzení. ■

Zdůrazněme, že tvrzení Věty 6.3.21 rozhodně nevylučuje, že $y_1 = \dots = y_k = 0$! Nyní si ukážeme *regulární modifikace* Věty 6.3.21, jejichž obsahem jsou tzv. *podmínky regularity*, které zajišťují kladnost jistého význačného koeficientu y_i v (6.3.12). Po vydělení tohoto vztahu číslem y_i se příliš nezmění, a tak můžeme bez újmy na obecnosti brát $y_i = 1$. Navíc (opět) bez újmy na obecnosti můžeme brát, že tento význačný index odpovídá $i = 0$.

Důsledek 6.3.22

Nechť množina $X \subseteq \mathbb{R}^n$ je neprázdná a konvexní a funkce $f_0, \dots, f_m : X \to \mathbb{R}$ jsou konvexní. Jestliže systém nerovností

$$f_0(x) < 0, \tag{6.3.13}$$
$$f_i(x) < 0, \quad i \in \{1, \dots, m\}, \tag{6.3.14}$$

nemá řešení na X a podsystém (6.3.14) má řešení na X, pak existují čísla $y_1, \dots, y_m \geq 0$ taková, že pro všechna $x \in X$ platí

$$f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \geq 0. \tag{6.3.15}$$

Důkaz. Využijeme tvrzení Věty 6.3.21, kde $k = m$ (tj. nemáme afinní část) a $i \in \{0,\dots,k\}$. Pak podle této věty existují $y_0,\dots,y_m \geq 0$ taková, že

$$y_0 f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \geq 0 \tag{6.3.16}$$

pro všechna $x \in X$. Je-li $y_0 = 0$, pak nutně alespoň jedno z čísel $y_1,\dots,y_m$ musí být kladné. Je-li $\overline{x} \in X$ řešením (6.3.14), pak v (6.3.16) dostáváme

$$\underbrace{y_0 f_0(\overline{x})}_{=0} + \sum_{i=1}^{m} y_i \underbrace{f_i(\overline{x})}_{<0} < 0,$$

což je ale spor s (6.3.16), tj. nutně $y_0 > 0$. Vydělením nerovnosti (6.3.16) kladným číslem y_0 a přeznačením $y_i / y_0 \rightsquigarrow y_i$ pro $i \in \{1,\dots,m\}$ dostáváme (6.3.15). ■

S využitím ekvivalentní formulace Farkasovy–Minkowského věty uvedené v Poznámce 6.3.20 dostaneme následující tvrzení, které ukazuje, že v případě afinních funkcí lze ostré nerovnosti v podsystému (6.3.14) nahradit neostrými.

DŮSLEDEK 6.3.23

Nechť jsou dány $a_0,\dots,a_m \in \mathbb{R}^n$ a $\beta_0,\dots,\beta_m \in \mathbb{R}$. Jestliže systém lineárních (afinních/nehomogenních) nerovností

$$f_0(x) := \langle a_0, x\rangle + \beta_0 < 0, \tag{6.3.17}$$

$$f_i(x) := \langle a_i, x\rangle + \beta_i \leq 0, \quad i \in \{1,\dots,m\}, \tag{6.3.18}$$

nemá řešení v $\mathbb{R}^n$ a podsystém (6.3.18) zde má řešení, pak existují taková $y_1,\dots,y_m \geq 0$, že platí

$$a_0 + \sum_{i=1}^{m} y_i a_i = 0 \quad \& \quad \beta_0 + \sum_{i=1}^{m} y_i \beta_i \geq 0, \tag{6.3.19}$$

tj. $f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \geq 0$ pro všechna $x \in \mathbb{R}^n$.

Důkaz. Položme $\widetilde{a}_i := (a_i, \beta_i)^\top \in \mathbb{R}^{n+1}$ pro $i \in \{0,1,\dots,m\}$, $a_{m+1} = (0,\dots,0,-1)^\top \in \mathbb{R}^{n+1}$ a $\widetilde{x} := [h,\lambda] \in \mathbb{R}^{n+1}$ pro $h \in \mathbb{R}^n$. Uvažujme homogenní systém nerovností

$$\langle \widetilde{a}_0, \widetilde{x}\rangle < 0 \quad \& \quad \langle \widetilde{a}_i, \widetilde{x}\rangle \leq 0, \quad i \in \{1,\dots,m\}. \tag{6.3.20}$$

Ukážeme, že tento systém nemá řešení na $\mathbb{R}^{n+1}$. Nechť $\widetilde{x}^* = [h^*, \lambda^*]$ je řešením (6.3.20). Z nerovnosti $\langle \widetilde{a}_m, \widetilde{x}^*\rangle = -\lambda^* \leq 0$ plyne, že $\lambda^* \geq 0$. Je-li $\lambda^* > 0$, pak vydělením nerovnosti (6.3.20) číslem λ^* dostáváme, že $x^* := \frac{1}{\lambda^*} h^*$ je řešením soustavy (6.3.17) & (6.3.18), což je spor ↯. Tedy nutně $\lambda^* = 0$. Potom se systém (6.3.20) redukuje na soustavu

$$\langle a_0, h^*\rangle < 0 \quad \& \quad \langle a_i, h^*\rangle \leq 0, \quad i \in \{1,\dots,m\}.$$

Nechť nyní $\overline{x} \in \mathbb{R}^n$ je řešením (6.3.18). Položme $x^* := \overline{x} + t h^*$. Pak pro $t > 0$ dostatečně velké máme

$$\langle a_0, x^*\rangle + \beta_0 = \underbrace{\langle a_0, \overline{x}\rangle + \beta_0}_{<0} + t \underbrace{\langle a_0, h^*\rangle}_{<0} < 0,$$

$$\langle a_i, x^*\rangle + \beta_i = \underbrace{\langle a_i, \overline{x}\rangle + \beta_i}_{\leq 0} + t \underbrace{\langle a_i, h^*\rangle}_{\leq 0} \leq 0, \quad i \in \{1,\dots,m\},$$

což znamená, že x^* je řešením systému (6.3.17) & (6.3.18), což je opět spor ↯.

Tedy neexistuje řešení systému (6.3.20). Pak podle tvrzení uvedeného v Poznámce 6.3.20 existují čísla $y_1,\dots,y_m \geq 0$ taková, že platí $\widetilde{a}_0 + \sum_{i=1}^{m+1} y_i \widetilde{a}_i = 0$, tj.

$$a_0 + \sum_{i=1}^{m} y_i a_i = 0 \quad \text{a současně} \quad \beta_0 + \sum_{i=1}^{m} y_i \beta_i - y_{m+1} = 0,$$

[125] V první části jsme ukázali, že v případě platnosti (6.3.8) neplatí (6.3.9). Nyní bychom měli ukázat opačnou implikaci, tj. že v případě neplatnosti (6.3.9) platí (6.3.8). Místo toho však využijeme obměnu a ukážeme ekvivalentní implikaci, tj. že z neplatnosti (6.3.8) plyne (6.3.9).

[126] *Důkaz.* Pokud v (6.3.9) položíme $b := a_0$, $A := -(a_1,\dots,a_m)$ a zaměníme-li $x \leftrightarrow y$, pak soustava $A^\top x \geq 0$ & $\langle x, a_0\rangle < 0$ nemá podle předpokladů řešení. Proto systém $Ay = a_0$ & $y \geq 0$ má řešení.

[127] Ky Fan (19. září 1914 – 22. března 2010) byl americký matematik čínského původu. Doktorát z matematiky získal pod vedením M. Frécheta na Univerzitě v Paříži v roce 1941. Po druhé světové válce se přestěhoval do USA a od roku 1965 působil jako profesor matematiky na Kalifornské univerzitě v Santa Barbaře. V roce 1999 daroval se svojí ženou jeden milion dolarů Americké matematické společnosti na založení Nadace Kyho a Yu-Fen Fanových.

Jeho jméno je spojeno např. s rozšířením nerovnosti mezi aritmetickým a geometrickým průměrem v podobě

$$\frac{\left(\Pi_1^n x_i\right)^{1/n}}{\left(\Pi_1^n (1-x_i)\right)^{1/n}} \leq \frac{\frac{1}{n}\left(\sum_1^n x_i\right)^{1/n}}{\frac{1}{n}\left(\sum_1^n (1-x_i)\right)^{1/n}}$$

pro $x_1,\dots,x_n \in [0,1/2]$, kde rovnost nastane právě tehdy, když $x_1 = \cdots = x_n$. Na levé straně této rovnosti je podíl geometrického průměru pro $x_1,\dots,x_n$ a pro $1-x_1,\dots,1-x_n$, zatímco na pravé straně je podíl aritmetického průměru pro $x_1,\dots,x_n$ a pro $1-x_1,\dots,1-x_n$. Tato nerovnost se využívá v teorii her a její důkaz je založen na Jensenově nerovnosti pro ostře konkávní funkci $\ln \frac{x}{1-x}$.

[128] Irving Leonard Glicksberg získal doktorát na Kalifornské univerzitě v Los Angeles v roce 1951. Jeho jméno je spojeno s jistým tvrzením o minimaxu v teorii her s nulovým součtem a také s rozšířením Kakutaniho věty o pevném bodě. Na webových stránkách RANDu je uveden seznam s jeho 26 odbornými články z let 1950–1958. Od šedesátých let 20. století působil na Washingtonské univerzitě. V mezinárodní matematické databázi MathSciNet je evidováno 86 jeho publikací.

[129] Toto tvrzení je zobecněním původního výsledku publikovaného v [164], který byl formulován pouze pro $k = m$, tj. bez afinních funkcí $f_{k+1}, \dots, f_m$.

[130] Jsou-li $\overline{v}, \widetilde{v} \in \mathcal{B}$ a $\lambda \in [0,1]$ libovolné, pak jistě pro každé $i \in \{1,\dots,k\}$ a $j \in \{k+1,\dots,m\}$ máme $\lambda\overline{v}_i + (1-\lambda)\widetilde{v}_i < 0$ a $\lambda\overline{v}_j + (1-\lambda)\widetilde{v}_j = 0$.

DŮSLEDEK 6.3.24

[131] Díky konvexnosti množiny X a funkcí $f_1,\dots,f_k$ platí pro libovolné $i \in \{1,\dots,k\}$, že

$$\begin{aligned} f_i(\lambda\overline{x} + (1-\lambda)\widetilde{x}) \le \\ \le \lambda f_i(\overline{x}) + (1-\lambda) f_i(\widetilde{x}) \le \\ \le \lambda\overline{u}_i + (1-\lambda)\widetilde{u}_i, \end{aligned}$$

zatímco z afinnosti funkcí $f_{k+1},\dots,f_m$ dostaneme pro libovolné $j \in \{k+1,\dots,m\}$, že

$$\begin{aligned} f_j(\lambda\overline{x} + (1-\lambda)\widetilde{x}) = \\ = \langle a_j, \lambda\overline{x} + (1-\lambda)\widetilde{x}\rangle + \beta_j = \\ = \lambda\Big(\langle a_j,\overline{x}\rangle + \beta_j\Big) + \\ + (1-\lambda)\Big(\langle a_j,\widetilde{x}\rangle + \beta_j\Big) = \\ = \lambda\overline{u}_j + (1-\lambda)\widetilde{u}_j. \end{aligned}$$

což dává (6.3.19). Odtud také pro libovolné $x \in \mathbb{R}^n$ plyne požadovaná nerovnost

$$f_0(x) + \sum_{i=1}^m y_i f_i(x) = \langle a_0, x\rangle + \beta_0 + \sum_{i=1}^m y_i\langle a_i, x\rangle + \sum_{i=1}^m y_i\beta_i = \underbrace{\langle a_0 + \sum_{i=1}^m y_i a_i, x\rangle}_{=0} + \underbrace{\beta_0 + \sum_{i=1}^m y_i\beta_i}_{=y_{m+1}\ge 0} \ge 0.$$

■

Ještě si ukážeme jedno tvrzení pro případ afinních funkcí na polyedru.

Nechť neprázdná množina $X \subseteq \mathbb{R}^n$ je polyedr a funkce $f_0,\dots,f_k,f_{k+1},\dots,f_m$ jsou afinní. Jestliže systém nerovností

$$\left.\begin{aligned} f_0(x) &< 0, \\ f_i(x) &\le 0, \quad i \in \{1,\dots,k\}, \\ f_j(x) &= 0, \quad i \in \{k+1,\dots,m\}, \end{aligned}\right\} \tag{6.3.21}$$

nemá řešení na X a podsystém (6.3.21) má řešení na X, pak existují čísla taková $y_1,\dots,y_k \ge 0$ a $y_{k+1},\dots,y_m \in \mathbb{R}$, že por všechna $x \in X$ platí

$$f_0(x) + \sum_{i=1}^m y_i f_i(x) \ge 0.$$

Důkaz. Množina X jakožto polyedr může být určena systémem nerovnic $g_1(x) \le 0,\dots,g_s(x) \le 0$ pro vhodné afinní funkce $g_1,\dots,g_s$ a nějaké $s \in \mathbb{N}$. Uvážíme-li nyní na $\mathbb{R}^n$ systém afinních nerovností

$$\begin{aligned} f_0(x) &< 0, \\ f_i(x) &\le 0, \quad i \in \{1,\dots,k\}, \\ f_j(x) \le 0 \quad \& \quad -f_j(x) &\le 0, \quad j \in \{k+1,\dots,m\}, \\ g_\ell(x) &\le 0, \quad \ell \in \{1,\dots,s\}, \end{aligned}$$

pak aplikováním předchozího Důsledku 6.3.23 pro m nahrazené $2m-k+s$ dostaneme existenci čísel $y_1,\dots,y_{2m-k+s} \ge 0$ takových, že

$$f_0(x) + \sum_{i=1}^k y_i f_i(x) + \sum_{i=k+1}^m y_i f_i(x) - \sum_{i=k+1}^m y_{m-k+i} f_i(x) + \sum_{i=1}^s y_{2m-k+i} g_i(x) \ge 0$$

neboli

$$f_0(x) + \sum_{i=1}^k y_i f_i(x) + \sum_{i=k+1}^m (y_i - y_{m-k+i}) f_i(x) \ge \underbrace{-\sum_{i=1}^s \underbrace{y_{2m-k+i}}_{\ge 0}\underbrace{g_i(x)}_{\le 0}}_{\ge 0},$$

tj. platí $f_0(x) + \sum_{i=1}^m \widetilde{y}_i f_i(x) \ge 0$ pro $\widetilde{y} := y_i \ge 0$, $i \in \{1,\dots,k\}$, a $\widetilde{y}_i := y_i - y_{m-k+i} \in \mathbb{R}$, $i \in \{k+1,\dots,m\}$. ■

Cvičení

6.3.1. Pro které hodnoty parametru $\alpha \in \mathbb{R}$ jsou množiny

$$X = \{[x,y] \in \mathbb{R}^2 \mid y \ge x^2 + 1\} \quad \text{a} \quad Y = \{[x,y] \in \mathbb{R}^2 \mid x \ge \alpha y^2\},$$

vlastně/silně oddělitelné? Určete rovnice příslušných oddělujících nadrovin.

řešení: silně oddělitelné pro $\alpha > 3\sqrt{3}/16$, pro $\alpha = 3\sqrt{3}/16$ jsou vlastně oddělitelné a pro $\alpha < 3\sqrt{3}/16$) nejsou oddělitelné, oddělující nadrovina je přímka $y = x/(\sqrt[3]{2\alpha}) + 2^{-5/3}\alpha^{-2/3}$

6.3.2. Určete rovnici oddělující nadroviny pro množiny

$$X = \{[x_1,\dots,x_{n+1}] \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_n^2 \le 1\},$$
$$Y = \{[x_1,\dots,x_{n+1}] \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_n^2 + x_{n+1}^2 + 3 \le 4x_n\}.$$

řešení: nadrovina $H_{p,\beta}$ s vektorem $p = (0,\dots,0,1,0)^\top$ a $\beta = 1$

6.3.3. *Dokažte:* Nechť $\emptyset \neq X_1 \subseteq \mathbb{R}^n$ je konvexní množina a $\emptyset \neq X_2 \subseteq \mathbb{R}^n$ je konvexní kužel.

a) Jsou-li množiny X_1, X_2 vlastně oddělitelné, pak jsou vlastně oddělitelné i pomocí nadroviny procházející počátkem.

b) Jsou-li množiny X_1, X_2 silně oddělitelné, pak existuje taková nadrovina procházející počátkem, že jeden z přidružených poloprostorů obsahuje kužel X_2 a má prázdný průnik s X_1.

6.3.4. *Dokažte:* Nechť $X_1, X_2 \subseteq \mathbb{R}^n$ jsou konvexní množiny takové, že $\operatorname{int} X_1 \neq \emptyset$ a $X_2 \cap \operatorname{int} X_1 = \emptyset$. Pak existuje nadrovina $H \subseteq \mathbb{R}^n$ taková, že X_1 a X_2 leží v opačných uzavřených poloprostorech určených nadrovinou H, tj. existuje $p \in \mathbb{R}^n \setminus \{0\}$ a $\beta \in \mathbb{R}$ tak, že

$$\langle p, x_1 \rangle \le \beta \le \langle p, x_2 \rangle \quad \text{pro všechna } x_1 \in X_1 \text{ a } x_2 \in X_2.$$

6.3.5. Určete rovnici opěrné nadroviny k množině

$$X = \{[x,y] \in \mathbb{R}^2 \mid y \ge \mathrm{e}^x\}$$

v bodě $a = [0,1]$.

řešení: nadrovina $H_{p,\beta}$ s vektorem $p = (-1,1)^\top$ a $\beta = 1$

6.3.6. Určete rovnici opěrné nadroviny k množině

$$X = \{[x,y,z] \in \mathbb{R}^3 \mid z \ge 1 + x^2 + y^2\}$$

v bodě $a = [1,1,3]$.

řešení: nadrovina $H_{p,\beta}$ s vektorem $p = (-2,-2,1)^\top$ a $\beta = -1$

6.3.7. Rozhodněte, zda rovina $z = (x+y)/2$ je opěrnou nadrovinou nadgrafu funkce

$$f(x,y) = \sqrt{2x^2 + xy + y^2}$$

v bodě $[x^*, y^*] = [0,0]$.

řešení: ano

6.3.8. Rozhodněte, zda rovina $z = x + y/\sqrt{2}$ je opěrnou nadrovinou nadgrafu funkce

$$f(x,y) = \sqrt{3x^2 + 2xy + y^2}$$

v bodě $[x^*, y^*] = [0,0]$.

řešení: ano

6.3.9. Vyjádřete (je-li to možné) množinu

$$X = \{[x,y] \in \mathbb{R}^2 \mid xy \ge 1,\ x > 0\}$$

jako průnik uzavřených poloprostorů.

řešení: $X = \bigcap_{x_0 \in (0,\infty)} \{[x,y] \in \mathbb{R}^2 \mid y + x/x_0^2 \ge 2/x_0\}$

6.3.10. *Dokažte:* Je-li $X \subseteq \mathbb{R}^n$ uzavřená množina taková, že $\operatorname{int} X \neq \emptyset$ a v každém bodě ∂X má opěrnou nadrovinu, pak X je konvexní.

6.3.11. Předpoklady konvexnosti a afinnosti hrají ve Větě 6.3.21 klíčovou roli. Uvažte $k = 1$ a $m = 2$ a zkuste najít funkce f_1, f_2 takové, že tvrzení Věty 6.3.21 již nebude splněno, pokud postupně

- f_1 nebude konvexní,
- f_2 nebude afinní,
- množina X nebude konvexní.

6.3.12. Řešitelnost podsystému (6.3.14) hraje v tvrzení Věty 6.3.22 klíčovou roli. Zkuste při $m = 1$ najít f_0 a f_1 tak, že podsystém (6.3.14) nemá řešení a tvrzení věty neplatí.

6.3.13. V tvrzení Věty 6.3.23 hraje klíčovou roli řešitelnost podsystému (6.3.18). Zkuste při $m = 2$ najít afinní funkce f_0, f_1 a f_2 tak, že podsystém (6.3.18) nemá řešení a tvrzení věty neplatí.

6.4 Vlastnosti konvexních funkcí

V této podkapitole si ukážeme několik dalších zajímavých (a důležitých) vlastností konvexních funkcí. Začneme se spojitostí a budeme pokračovat s charakterizací diferencovatelných konvexních funkcí podobně jako v případě funkce jedné proměnné.

Věta 6.4.1

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ je konvexní na X. Pak f je spojitá pro každé $x \in \operatorname{ri} X$.

Důkaz. Nechť bod $x^* \in \operatorname{ri} X$ je libovolný a uvažme funkci $F(x) := f(x^* - x) - f(x^*)$, která je očividně také konvexní. Jelikož pak platí $\lim_{x\to 0} F(x) = \lim_{x\to x^*} f(x) - f(x^*)$, je zřejmé, že funkce f je spojitá v x^* právě tehdy, když F je spojitá v 0. Zaměříme se proto právě na spojitost libovolné konvexní funkce F v počátku.

Uvažme nejdříve, že $\operatorname{int} X \neq \emptyset$. Označme jako K_r (hyper-)krychli s hranou délky $2r$ a středem v počátku, tj.

$$K_r := \{x \in \mathbb{R}^n \mid \|x\|_\infty < r\},$$

kde $\|x\|_\infty := \max_{1\le i\le n} |x_i|$ pro $x = (x_1, \dots, x_n)^\top$ značí „maximální normu“ a číslo $r > 0$ je takové, že $K_r \subsetneq X$. Označme dále jako $X_1, \dots, X_m$ pro $m = 2^n$ vrcholy krychle K_r (tj. souřadnice X_i jsou pouze $\pm r$) a položme

$$\alpha := \max_{1\le i\le m} F(X_i).$$

Poněvadž $K_r = \operatorname{conv}\{X_1, \dots, X_m\}$, lze každý bod $y \in K_r$ vyjádřit jako konvexní kombinaci bodů $X_1, \dots, X_m$, tj. existují čísla[132] $\lambda_1, \dots, \lambda_m \ge 0$ splňující $\sum_{i=1}^m \lambda_i = 1$ a $y = \sum_{i=1}^m \lambda_i X_i$. Současně z konvexnosti funkce F a z Jensenovy nerovnosti (Věta 6.2.12) plyne

$$F(y) = F\Big(\sum_{i=1}^m \lambda_i X_i\Big) \le \sum_{i=1}^m \lambda_i F(X_i) \le \alpha \sum_{i=1}^m \lambda_i = \alpha. \tag{6.4.1}$$

Vezmeme-li nyní $\varepsilon \in (0, 1]$ libovolné a definujeme-li ε-krychli $\widetilde{K}_\varepsilon := \varepsilon K_r$, pak pro každé $x \in \widetilde{K}_\varepsilon$ platí, že $\pm x/\varepsilon \in K_r$, a tedy

$$F(x) = F\big(\varepsilon \tfrac{x}{\varepsilon} + (1-\varepsilon)\,0\big) \overset{\substack{F \text{ je}\\ \text{konvexní}}}{\le} \varepsilon F(\underbrace{x/\varepsilon}_{\in K_r}) + (1-\varepsilon)\,\underbrace{F(0)}_{=0} \overset{(6.4.1)}{\le} \varepsilon\,\alpha.$$

Na druhou stranu máme

$$0 = F(0) = F\big(\tfrac{1}{1+\varepsilon} x + \tfrac{\varepsilon}{1+\varepsilon}(-x/\varepsilon)\big) \overset{\substack{F \text{ je}\\ \text{konvexní}}}{\le} \frac{1}{1+\varepsilon} F(x) + \frac{\varepsilon}{1+\varepsilon} F(\underbrace{-x/\varepsilon}_{\in K_r}) \overset{(6.4.1)}{\le} \frac{1}{1+\varepsilon} F(x) + \frac{\varepsilon}{1+\varepsilon}\,\alpha,$$

což znamená, že současně $F(x) \ge -\varepsilon\,\alpha$. Celkem tedy dostáváme $|F(x)| \le \varepsilon\,\alpha$, což vzhledem k libovolnosti ε znamená[133], že funkce F musí být spojitá v počátku[133].

Pokud $\operatorname{int} X = \emptyset$, tj. $\operatorname{ri} X \neq \operatorname{int} X$, pak postupujeme stejně jako v předchozí části s tím, že místo n-dimenzionální krychle uvažujeme pouze $(\dim X)$-dimenzionální krychli v $\operatorname{aff} X$. ■

Poznámka 6.4.2. (i) Obzvláště je-li funkce $f : \mathbb{R}^n \to \mathbb{R}$ konvexní na $\mathbb{R}^n$, pak je tato funkce spojitá na celém $\mathbb{R}^n$.

(ii) Tvrzení Věty 6.4.1 nelze rozšířit na celé X, tj. zahrnout do něj i body z $\mathrm{r}\partial X$. Např. totiž pro $X := [-1, 1]$ a funkci

$$f(x) := \begin{cases} 0, & x \in (-1, 1), \\ 1, & x = \pm 1, \end{cases}$$

snadno můžeme ukázat její konvexnost na X, avšak spojitost je na $\mathrm{r}\partial X = \{\pm 1\}$ porušena. Na druhou stranu nemůže být spojitost na hranici porušena zcela libovolným způsobem, protože by funkce již nemusela zůstat konvexní jako např. podobná (nekonvexní) funkce

$$g(x) := \begin{cases} 0, & x \in (-1, 1), \\ -1, & x = \pm 1. \end{cases}$$

[132] Dokonce stačí pouze $n + 1$ bodů, viz Větu 6.1.14 – to ale teď nehraje vůbec žádnou roli.

[133] Mimo jiné odtud také plyne, že nutně $\alpha \ge 0$. Je to zřejmé?

[133] Podle definice limity bychom měli ukázat, že pro libovolné $\mathcal{O}_\varepsilon(0)$ existuje $\mathcal{O}_\delta(x^*)$ takové, že $F(x) \in \mathcal{O}_\varepsilon(0)$ pro každé $x \in \mathcal{O}_\delta(x^*)$. Je-li $\alpha = 0$, pak nutně $|F(x)| = 0$ pro každé $x \in \widetilde{K}_\varepsilon$, kde $\varepsilon \in (0, 1]$, a současně platí $F(0) = 0$, takže funkce F je konstantní na každé takové krychli $\widetilde{K}_\varepsilon$, a tudíž spojitost v počátku je zřejmá. Je-li $\alpha \neq 0$, pak pro $\delta := \min\{r, \varepsilon/(\alpha+1)\}$ bude každé $x \in \widetilde{K}_\delta$ splňovat $|F(x)| < \varepsilon$, neboť buď $|F(x)| \le \alpha\,\delta = \alpha\,r < \varepsilon$ pro $r \le \varepsilon/(\alpha+1)$ neboli $\varepsilon \ge (\alpha+1)\,r > \alpha\,r$ nebo $|F(x)| \le \alpha\,\delta = \alpha\,\varepsilon/(\alpha+1) < \varepsilon$ pro $r > \varepsilon/(\alpha+1)$. Každopádně $F(x) \in (-\varepsilon, \varepsilon) = \mathcal{O}_\varepsilon(0)$, a tudíž F je opět spojitá v počátku.

(iii) Opačné tvrzení vyvozující konvexnost funkce z její spojitosti by mohlo vypadat následovně[134]:

> Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ spojitá na X. Jestliže pro každé $x, y \in X$, $x \neq y$, existuje číslo $\lambda \in (0,1)$ takové, že
>
> $$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y),$$
>
> je funkce f konvexní na X.

▲

[134] I když toto je asi poměrně nezajímavý/nudný výsledek...

Z úvodního kurzu matematické analýzy si (snad) pamatujeme několik podmínek zaručujících konvexnost diferencovatelné funkce $f : \mathbb{R} \to \mathbb{R}$:

(i) má-li f vlastní derivaci v otevřeném intervalu I, pak f je (ostře) konvexní na I právě tehdy, když f' je neklesající (rostoucí) na I;

(ii) má-li f vlastní derivaci v otevřeném intervalu I, pak f je (ostře) konvexní na I právě tehdy, když pro každé $x, x^* \in I$ platí (v případě ostré konvexnosti navíc $x \neq x^*$)

$$f(x) \overset{(>)}{\geq} f(x^*) + f'(x^*)(x - x^*),$$

tj. graf funkce f na I leží nad tečnou sestrojenou v libovolném bodě;

(iii) má-li f vlastní druhou derivaci v otevřeném intervalu I, pak f je konvexní na I právě tehdy, když funkce $f''(x) \geq 0$ (a je-li dokonce $f''(x) > 0$ na I, pak je ostře konvexní).

Nyní si tato tvrzení zobecníme pro libovolnou funkci $f : \mathbb{R}^n \to \mathbb{R}$, přičemž při jejich formulaci využijeme toho, že silná konvexnost se při volbě $\vartheta = 0$ redukuje na „obyčejnou“ konvexnost. Začneme s analogií druhého tvrzení.

VĚTA 6.4.3

> Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce f diferencovatelná na otevřené množině[135] $\mathscr{U} \supseteq X$. Pak f je silně konvexní na X s konstantou silné konvexnosti $\vartheta \geq 0$ právě tehdy, když pro každé $x, x^* \in X$ platí
>
> $$f(x) \geq f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle + \vartheta \| x - x^* \|^2. \tag{6.4.2}$$

[135] Potřebujeme mít zaručenu diferencovatelnost f ve všech bodech množiny X.

Poznámka 6.4.4. (i) V tvrzení Věty 6.4.3 jsme opomenuli ostře konvexní funkce. Ty lze ekvivalentně charakterizovat ostrou nerovností v (6.4.2) s $\vartheta = 0$ platnou pro libovolnou dvojici $x, x^* \in X$ splňující $x \neq x^*$.

(ii) Diferencovatelnost funkce f v bodě x^* je ekvivalentní s existencí tečné nadroviny ke grafu funkce f v bodě $[x^*, f(x^*)]$, která má rovnici

$$z = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle.$$

Ze vztahu (6.4.2) s $\vartheta = 0$ pak vidíme, že diferencovatelná funkce f je konvexní na X právě tehdy, když $z \leq f(x)$, tj. její graf leží nad tečnou nadrovinou sestrojenou v libovolném bodě $x \in X$. Navíc, položíme-li

$$p := \left(-\operatorname{grad}^\top f(x^*), 1 \right)^\top \in \mathbb{R}^{n+1} \quad \text{a} \quad \alpha := f(x^*) - \langle \operatorname{grad} f(x^*), x^* \rangle,$$

pak nadrovina $H_{p,\alpha} = \{ x \in \mathbb{R}^n \mid \langle p, x \rangle = \alpha \}$ je opěrnou nadrovinou k nadgrafu funkce f, tj. k množině epi f, v bodě $[x^*, f(x^*)]$ neboli

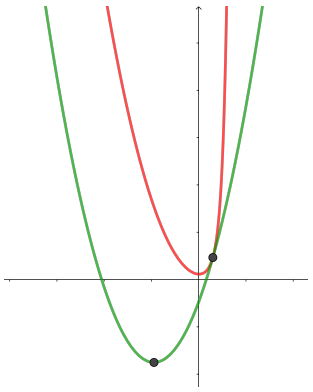
$$\langle p, (x, y)^\top \rangle \geq \alpha = \langle p, \left(x^*, f(x^*) \right)^\top \rangle$$

pro každé $[x, y] \in \operatorname{epi} f$.

(iii) Je-li x^* stacionárním bodem funkce f, tj. grad $f(x^*) = 0$, pak se nerovnost (6.4.2) redukuje na

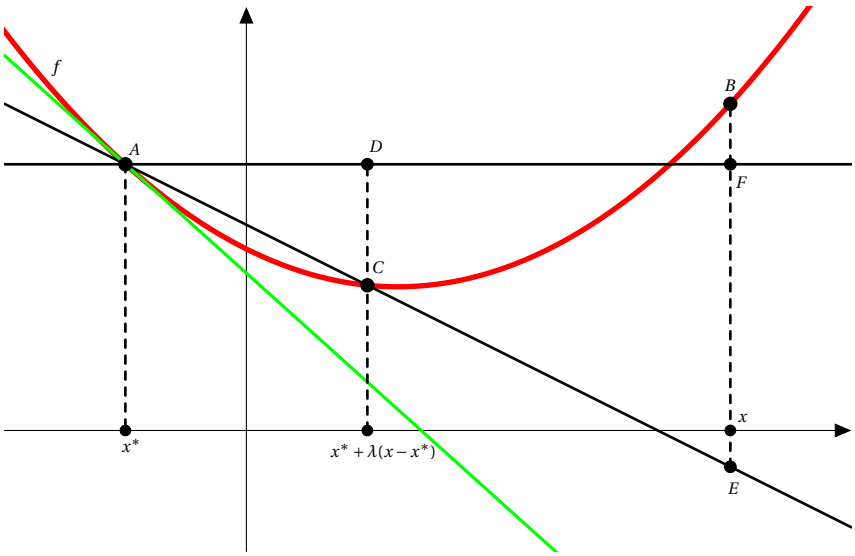
$$f(x) \geq f(x^*) + \vartheta \| x - x^* \|^2.$$

Přitom rovnice $z = f(x^*) + \vartheta \| x - x^* \|^2$ zadává rotační paraboloid v $\mathbb{R}^{n+1}$ s vrcholem v bodě $[x^*, f(x^*)]$, což znamená, že nadgraf silně konvexní funkce f leží uvnitř tohoto rotačního paraboloidu. Totéž bude platit i v případě, kdy máme grad $f(x^*) \neq 0$, avšak s tím rozdílem, že bod $[x^*, f(x^*)]$ nebude vrcholem paraboloidu ale jeho bodem dotyku s epi f, viz Obrázek 6.71. ▲



Obrázek 6.71: Silně konvexní funkce (červená barva) a její „parabolická tečna“ (zelená barva).

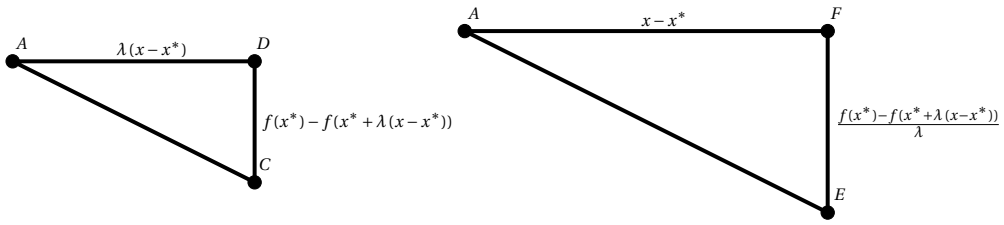
Než si ukážeme analytický důkaz předchozího tvrzení, pokusíme se jej v případě $\vartheta = 0$ (tj. pro konvexní funkce) ilustrovat pomocí jednoduchých obrázků. Nejdříve uvažme, že f je konvexní, a zvolme $x, x^* \in X$ a $\lambda \in (0,1)$ libovolně. Na Obrázku 6.72 je graf funkce f a trojice bodů $A := [x^*, f(x^*)]$, $B := [x, f(x)]$ a $C := [x^* + \lambda(x - x^*), f(x^* + \lambda(x - x^*))]$.



Obrázek 6.72: „Vizualizace“ implikace „⇒“ v důkazu Věty 6.4.3 pro $\vartheta = 0$.

Také je zde vyznačena sečna grafu určená body A a C, jejímž prodloužením získáme bod E. Nyní můžeme vidět dva podobné[136] trojúhelníky $\triangle ADC$ a $\triangle AFE$, takže délka strany FE musí být rovna $\frac{f(x^*) - f(x^* + \lambda(x - x^*))}{\lambda}$, viz Obrázek 6.73.

[136] Trojúhelníky mají všechny tři úhly shodné.



Obrázek 6.73: Trojúhelníky $\triangle ADC$ a $\triangle AEB$ z Obrázku 6.72.

To znamená, že „y-ová“ souřadnice bodu E je

$$f(x^*) - \frac{f(x^*) - f(x^* + \lambda(x - x^*))}{\lambda} = f(x^*) + \frac{f(x^* + \lambda(x - x^*)) - f(x^*)}{\lambda},$$

a zároveň z Obrázku 6.72 vidíme, že

$$f(x) \geq f(x^*) + \frac{f(x^* + \lambda(x - x^*)) - f(x^*)}{\lambda} \tag{6.4.3}$$

pro libovolné $\lambda \in (0,1]$. Pak pro $\lambda \to 0^+$ se bude bod C přibližovat k A, takže se ze sečné nadroviny určené body A a C stane tečná nadrovina (zelená přímka na Obrázku 6.72). Současně bod E bude klesat níže a níže, takže výraz na pravé straně nerovnosti (6.4.3) bude klesající a podíl

[137] Ta existuje díky předpokladu diferencovatelnosti funkce f.

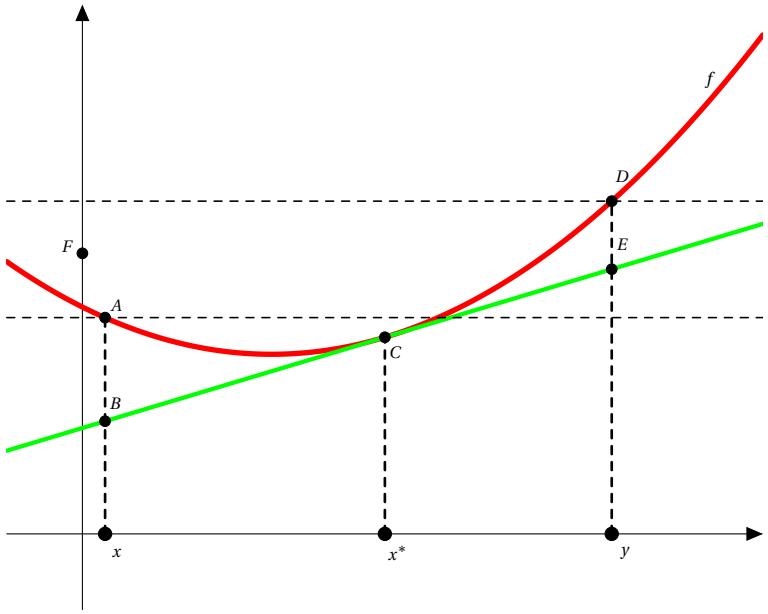
bude konvergovat k hodnotě směrové derivace[137] funkce f v bodě x^* a ve směru $x - x^*$, tj. k číslu $f_{x-x^*}(x^*)$. Jenže víme, že platí

$$f_{x-x^*}(x^*) = \langle \operatorname{grad} f(x^*), x - x^* \rangle,$$

z čehož již snadno odvodíme požadovanou nerovnost

$$f(x) \geq f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle.$$

Na druhou stranu, pokud platí (6.4.2) s $\vartheta = 0$, pak graf funkce f leží nad tečnou nadrovinou sestrojenou v libovolném bodě $x^* \in X$. Vezmeme-li libovolné body $x, y \in X$ a $x^* := \lambda x + (1-\lambda) y$ pro nějaké $\lambda \in (0,1)$, pak s využitím nerovnosti (6.4.2) dostaneme Obrázek 6.74 s tečnou nadrovinou (vyznačenou zelenou barvou) o rovnici $z = f(x^*) + \langle \operatorname{grad} f(x^*), t - x^* \rangle$ pro $t \in \mathbb{R}^n$ a s body $A := [x, f(x)]$, $B := [x, f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle]$, $C := [x^*, f(x^*)]$, $D := [y, f(y)]$, $E := [y, f(x^*) + \langle \operatorname{grad} f(x^*), y - x^* \rangle]$ a $F := [0, \lambda f(x) + (1-\lambda) f(y)]$.



Obrázek 6.74: „Vizualizace“ implikace „⇐“ v důkazu Věty 6.4.3 pro $\vartheta = 0$.

Z tohoto obrázku pak vidíme, že

$$\begin{aligned}\lambda f(x) + (1-\lambda) f(y) &\geq \lambda \left[f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle \right] + (1-\lambda) \left[f(x^*) + \langle \operatorname{grad} f(x^*), y - x^* \rangle \right] = \\ &= f(x^*) + \langle \operatorname{grad} f(x^*), \lambda (x - x^*) + (1-\lambda)(y - x^*) \rangle = \\ &= f(x^*) = f(\lambda x + (1-\lambda) y),\end{aligned}$$

tj. funkce f je konvexní.

Důkaz Věty 6.4.3. „$\Longrightarrow$“ Nechť f je silně konvexní s $\vartheta \geq 0$ a nechť $x, x^* \in X$ jsou libovolná pevně zvolená. Je-li $x = x^*$, je tvrzení zřejmé. Nechť tedy $x \neq x^*$. Pak díky silné konvexnosti platí pro všechna $\lambda \in [0,1]$ nerovnost

$$f(\lambda x + (1-\lambda) x^*) \leq \lambda f(x) + (1-\lambda) f(x^*) - \vartheta \lambda (1-\lambda) \| x - x^* \|^2$$

neboli

$$\lambda f(x) - \vartheta \lambda (1-\lambda) \| x - x^* \|^2 - \lambda f(x^*) \geq f(\lambda x + (1-\lambda) x^*) \underbrace{- (1-\lambda) f(x^*) - \lambda f(x^*)}_{= -f(x^*)}.$$

Odtud pro každé $\lambda \in (0,1]$ dostaneme

$$f(x) - \vartheta (1-\lambda) \| x - x^* \|^2 - f(x^*) \geq \frac{1}{\lambda} \left[f(\underbrace{\lambda x + (1-\lambda) x^*}_{x^* + \lambda (x - x^*)}) - f(x^*) \right] =$$

$$\overset{\text{diferencovatelnost}^{138}}{=} \frac{1}{\lambda} \left[\mathrm{d} f(\lambda (x - x^*)) + \omega(\lambda (x - x^*)) \right] =$$

[138] Viz str. 34.

$$= \frac{1}{\lambda}\big[\langle \operatorname{grad} f(x^*), \lambda(x-x^*)\rangle + \omega(\lambda(x-x^*))\big] =$$
$$= \langle \operatorname{grad} f(x^*), x-x^*\rangle + \frac{\omega(\lambda(x-x^*))}{\lambda}, \tag{6.4.4}$$

přičemž $\lim_{h\to 0}\omega(h)/\|h\| = 0$. Odtud limitním přechodem pro $\lambda \to 0^+$ dostáváme

$$\langle \operatorname{grad} f(x^*), x-x^*\rangle + \underbrace{\frac{\omega(\lambda(x-x^*))}{\|\lambda(x-x^*)\|}}_{\to 0}\underbrace{\|\lambda(x-x^*)\|}_{\text{číslo}} \overset{\lambda\to 0^+}{\to} \langle \operatorname{grad} f(x^*), x-x^*\rangle,$$

což dohromady s (6.4.4) dává (6.4.2).

„$\Longleftarrow$“ Nechť platí nerovnost (6.4.2) a $x, y \in X$ a $\lambda \in [0,1]$ jsou libovolná. Označme $x^* := \lambda x + (1-\lambda) y$. Pak podle (6.4.2) platí

$$f(x) \geq f(x^*) + \langle \operatorname{grad} f(x^*), x-x^*\rangle + \vartheta\|x-x^*\|^2$$

a

$$f(y) \geq f(x^*) + \langle \operatorname{grad} f(x^*), y-x^*\rangle + \vartheta\|y-x^*\|^2.$$

Sečteme-li λ-násobek první nerovnosti a $(1-\lambda)$-násobek druhé nerovnosti, obdržíme

$$\lambda f(x) + (1-\lambda) f(y) \geq f(x^*) + \langle \operatorname{grad} f(x^*), \lambda(x-x^*) + (1-\lambda)(y-x^*)\rangle +$$
$$+ \vartheta\big(\lambda\vartheta\|x-x^*\|^2 + (1-\lambda)\vartheta\|y-x^*\|^2\big),$$

což s využitím $x - x^* = x - \lambda x - (1-\lambda) y = (1-\lambda)(x-y)$ a současně $y - x^* = y - \lambda x - (1-\lambda) y = \lambda(y-x)$ dává

$$\lambda f(x) + (1-\lambda) f(y) \geq f(x^*) + \langle \operatorname{grad} f(x^*), \underbrace{\lambda(1-\lambda)(x-y) + (1-\lambda)\lambda(y-x)}_{=0}\rangle +$$
$$+ \vartheta\big(\lambda(1-\lambda)^2\|x-y\|^2 + (1-\lambda)\lambda^2\|y-x\|^2\big) =$$
$$= f(x^*) + \vartheta\lambda(1-\lambda)\|x-y\|^2,$$

tj.

$$f(x^*) = f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) - \vartheta\lambda(1-\lambda)\|x-y\|^2,$$

což znamená, že funkce f je silně konvexní s konstantou $\vartheta \geq 0$. ■

Z Věty 2.4.2 plyne několik velmi podstatných důsledků – nejdříve si uvedeme ty, které se týkají optimalizačních problémů.

Poznámka 6.4.5.

(i) Je-li $f : \mathbb{R}^n \to \mathbb{R}$ diferencovatelná konvexní funkce a $\operatorname{grad} f(x^*) = 0$, pak nerovnost (6.4.5) znamená, že x^* je globální minimum funkce f na $\mathbb{R}^n$ (srovnej s Větou 6.2.23(iii)). Toto je klasická nutná podmínka[139] pro nepodmíněnou optimalizaci libovolné funkce, která se v případě konvexní funkce stává i postačující.

(ii) Podobně splnění podmínky

$$\langle \operatorname{grad} f(x^*), x-x^*\rangle \geq 0 \quad \text{pro všechna } x \in X$$

implikuje díky (6.4.2), že x^* minimalizuje diferencovatelnou konvexní funkci f na konvexní množině X. Tato postačující podmínka[140] optimality je současně nutná. Vskutku, připusťme, že tomu tak není, tj. x^* minimalizuje f na X a $\langle \operatorname{grad} f(x^*), x-x^*\rangle < 0$ pro nějaké $x \in X$. Potom

$$\lim_{\lambda\to 0^+}\frac{f(x^*+\lambda(x-x^*)) - f(x^*)}{\lambda} = \overbrace{f_{x-x^*}(x^*)}^{\text{směrová derivace}} = \langle \operatorname{grad} f(x^*), x-x^*\rangle < 0,$$

což znamená, že pro dostatečně malé $\lambda > 0$ je $f(x^*+\lambda(x-x^*))$ klesající, což je spor s tím, že x^* minimalizuje f na X (srovnej s Větou 8.1.4 později). ▲

V dalších dvou důsledcích nabídneme i jiné ekvivalentní charakterizace konvexních funkcí.

[139] V případě $n = 1$ můžeme její počátek vystopovat až k Fermatovi a jeho práci *Methodus ad Disquirendam Maximam et Minimam*. Tento text byl nejspíše napsán již kolem roku 1629, ale publikován byl až posmrtně v roce 1679 díky Fermatovu synu Samuelovi. Nicméně na veřejnosti se Fermatovo pojednání objevilo již mnohem dříve, když jej Fermat prostřednictvím Mersenna zaslal Descartesovi. Ten dopis obdržel přibližně 10. ledna 1638 a podrobil Fermatovu práci velké kritice...

[140] Tuto podmínku lze využít také k tomu, že je-li bod x^* řešením úlohy $f(x) \to \min$ pro $x \in X$, pak $\widehat{x} \in X$ bude také řešením právě tehdy, když

$$\langle \operatorname{grad} f(\widehat{x}), \widehat{x} - x^*\rangle = 0.$$

Vskutku, uvedená podmínka společně s Větou 6.4.3 dává

$$f(x^*) - f(\widehat{x}) \geq \langle \operatorname{grad} f(\widehat{x}), x^* - \widehat{x}\rangle = 0,$$

tj. $f(x^*) \geq f(\widehat{x})$, takže nutně $f(x^*) = f(\widehat{x})$, neboť x^* je minimum f na X. Naopak, realizuje-li se minimum f na X v bodech x^* a $\widehat{x}$, pak jistě $f(x^*) = f(\widehat{x})$ a z konvexnosti f máme

$$f\big(\widehat{x} + \lambda(x^* - \widehat{x})\big) = f(x^*) = f(\widehat{x}),$$

a tedy funkce $h(\lambda) := f\big(\widehat{x} + \lambda(x^* - \widehat{x})\big)$ je na intervalu $[0,1]$ konstantní neboli

$$h'(\lambda) = \langle \operatorname{grad} f(\widehat{x}), x^* - \widehat{x}\rangle = 0.$$

tj. uvedená nerovnost platí.

Důsledek 6.4.6

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce f diferencovatelná na otevřené množině $\mathscr{U} \supseteq X$. Pak funkce f je silně konvexní na X s konstantou silné konvexnosti $\vartheta \geq 0$ právě tehdy, když[141] pro každé $x, x^* \in X$ platí

$$\langle \operatorname{grad} f(x) - \operatorname{grad} f(x^*), x - x^* \rangle \geq 2\vartheta \| x - x^* \|^2. \tag{6.4.5}$$

[141] Funkce F je *silně monotónní*, jestliže existuje číslo $c > 0$ takové, že pro každé x, y platí $\langle F(x) - F(y), x - y \rangle \geq c \| x - y \|^2$. Proto nerovnost (6.4.5) znamená, že $\operatorname{grad} f(x)$ je silně monotónní funkce na X.

Podobně jako v případě předchozího tvrzení platí, že funkce f je ostře konvexní právě tehdy, když nerovnost (6.4.5) s $\vartheta = 0$ je ostrá pro libovolnou dvojici $x, x^* \in X$ splňující $x \neq x^*$.

Důkaz. „$\Longrightarrow$“ Nechť f je silně konvexní s konstantou silné konvexnosti $\vartheta \geq 0$. Pak z nerovnosti (6.4.2) pro libovolné $x, x^* \in X$ plyne

$$f(x) - f(x^*) \geq \langle \operatorname{grad} f(x^*), x - x^* \rangle + \vartheta \| x - x^* \|^2,$$
$$f(x^*) - f(x) \geq \langle \operatorname{grad} f(x), x^* - x \rangle + \vartheta \| x^* - x \|^2,$$

což po sečtení ihned dává (6.4.5).

„$\Longleftarrow$“ Nechť platí (6.4.5) a $x, x^* \in X$ jsou libovolné. Nejdříve upravíme výraz $f(x) - f(x^*)$ do vhodnější podoby. Definujme diferencovatelnou funkci $F : [0,1] \to \mathbb{R}$ předpisem

$$F(t) := f(tx + (1-t)x^*) = f(tx_1 + (1-t)x_1^*, \ldots, tx_n + (1-t)x_n^*).$$

Pak očividně $F(0) = f(x^*)$ a $F(1) = f(x)$. Navíc z řetězového pravidla pro derivaci dostaneme

$$F'(t) = \frac{\partial f}{\partial x_1}(tx + (1-t)x^*)(x_1 - x_1^*) + \cdots + \frac{\partial f}{\partial x_n}(tx + (1-t)x^*)(x_n - x_n^*) =$$
$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(tx + (1-t)x^*)(x_i - x_i^*) = \langle \operatorname{grad} f(tx + (1-t)x^*), x - x^* \rangle.$$

Jelikož pro $0 \leq \tilde{t} \leq t \leq 1$ máme

$$F'(t) - F'(\tilde{t}) = \langle \operatorname{grad} f(tx + (1-t)x^*) - \operatorname{grad} f(\tilde{t}x + (1-\tilde{t})x^*), x - x^* \rangle =$$
$$= \frac{1}{t - \tilde{t}} \langle \operatorname{grad} f(x^* + t(x - x^*)) - \operatorname{grad} f(x^* + \tilde{t}(x - x^*)), (t - \tilde{t})(x - x^*) \rangle \overset{(6.4.5)}{\geq} 0,$$

je funkce F' je neklesající (monotónní), a tudíž i integrovatelná (je-li f dokonce *spojitě* diferencovatelná, pak je integrovatelnost F splněna triviálně). Můžeme[142] proto využít Newtonovu–Leibnizovu formuli pro F' na intervalu $[0,1]$, což dává

$$\int_0^1 F'(t)\,\mathrm{d}t = F(1) - F(0) = f(x) - f(x^*).$$

Odtud pak dostaneme

$$f(x) - f(x^*) - \langle \operatorname{grad} f(x^*), x - x^* \rangle =$$
$$= \int_0^1 \langle \operatorname{grad} f(tx + (1-t)x^*), x - x^* \rangle\,\mathrm{d}t - \langle \operatorname{grad} f(x^*), x - x^* \rangle =$$
$$= \int_0^1 \langle \operatorname{grad} f(tx + (1-t)x^*) - \operatorname{grad} f(x^*), x - x^* \rangle\,\mathrm{d}t =$$
$$= \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 \frac{1}{t} \langle \operatorname{grad} f(tx + (1-t)x^*) - \operatorname{grad} f(x^*), t(x - x^*) \rangle\,\mathrm{d}t \geq$$
$$\overset{(6.4.5)}{\geq} \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 \frac{2\vartheta \| t(x - x^*) \|^2}{t}\,\mathrm{d}t = 2\vartheta \| x - x^* \|^2 \underbrace{\int_0^1 t\,\mathrm{d}t}_{=1/2} = \vartheta \| x - x^* \|^2,$$

což s pomocí Věty 6.4.3 znamená, že funkce f je silně konvexní. ∎

Ještě zbývá zobecnění posledního tvrzení pro funkce jedné proměnné – velmi důležité kritérium (ostré/silné) konvexnosti.

[142] Zde je potřeba mít na paměti, že pro funkci $f : \mathbb{R} \to \mathbb{R}$ spojitou na intervalu $[a,b]$ platí

$$\left(\int_a^x f(t)\,\mathrm{d}t \right)' = f(x),$$

avšak pro Riemannův integrál již rovnost při opačném pořadí obecně naplatí, tj. může se stát, že

$$\int_a^b f'(t)\,\mathrm{d}t \neq f(b) - f(a). \tag{6.4.6}$$

Např. pro funkci

$$f(x) = \begin{cases} x \sin(1/x), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

je $f'(x) = \sin(1/x) - \frac{\cos(1/x)}{x}$ neohraničená na $[0,1]$, tj. zjevně f' nemůže být riemannovsky integrovatelná na $[0,1]$. Uvedený vztah nemusí platit ani v případě, že f' je ohraničená, viz např.

$$g(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

nebo tzv. *Volterrovu funkci* získanou „kopírováním“ $g(x)$. Je-li však f' integrovatelná na $[a,b]$, pak již v (6.4.6) nastane rovnost.

Důsledek 6.4.7

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina splňující $\operatorname{int} X \neq \emptyset$. Nechť funkce $f : X \to \mathbb{R}$ je dvakrát spojitě diferencovatelná na otevřené množině $\mathscr{U} \supseteq X$ s Hessovou maticí druhých derivací $\nabla^2 f(x)$. Pak f je silně konvexní na množině X s konstantou silné konvexnosti $\vartheta \geq 0$ právě tehdy, když pro každé $x \in X$ a $h \in \mathbb{R}^n$ platí

$$\langle \nabla^2 f(x)\, h, h\rangle \geq 2\vartheta \|h\|^2, \tag{6.4.7}$$

jinými slovy $\nabla^2 f(x) \geq 2\vartheta I$ pro všechna $x \in X$.

Důkaz. „$\Longrightarrow$“ Nechť funkce f je silně konvexní s konstantou silné konvexnosti $\vartheta \geq 0$ a nechť $x^* \in \operatorname{int} X$ je libovolné. Pak $x = x^* + \lambda h \in X$ pro $\lambda > 0$ dostatečně malé a libovolné $h \in \mathbb{R}^n$. Protože f má spojité parciální derivace druhého řádu, platí[143]

$$f(x) = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^*\rangle + \frac{\lambda^2}{2}\langle \nabla^2 f(x^*)\, h, h\rangle + \omega(\underbrace{x - x^*}_{=\lambda h}), \tag{6.4.8}$$

přičemž $\lim_{\lambda\to 0} \frac{\omega(\lambda h)}{\|\lambda h\|^2} = 0$. Odtud dostáváme

$$\tfrac{1}{2}\langle \nabla^2 f(x^*)\, h, h\rangle + \frac{\omega(\lambda h)}{\lambda^2\|h\|^2}\|h\|^2 \overset{(6.4.8)}{=} \frac{1}{\lambda^2}\Big[f(x) - f(x^*) - \langle \operatorname{grad} f(x^*), x - x^*\rangle\Big] \geq$$
$$\overset{(6.4.2)}{\geq} \frac{1}{\lambda^2}\,\vartheta\|\underbrace{x - x^*}_{=\lambda h}\|^2 = \vartheta\|h\|^2,$$

z čehož limitním přechodem pro $\lambda \to 0^+$ plyne

$$\langle \nabla^2 f(x^*)\, h, h\rangle \geq 2\vartheta\|h\|^2,$$

tj. nerovnost (6.4.7) je splněna.

Je-li nyní $x^* \in X \setminus \operatorname{int} X$, pak $x^* \in \partial X$ a tedy existuje posloupnost[144] $\{x_k\}_{k\in\mathbb{N}}$ taková, že $x_k \in \operatorname{int} X$ pro všechna $k \in \mathbb{N}$ a $x_k \to x^*$ pro $k \to \infty$. Pak z předchozí části plyne, že $\langle \nabla^2 f(x_k)\, h, h\rangle \geq 2\vartheta\|h\|^2$ pro každé $h \in \mathbb{R}^n$ a $k \in \mathbb{N}$. Poněvadž je však $\nabla^2 f(x)$ spojitá, můžeme „jít s limitou dovnitř“ skalárního součinu[145], čímž dostaneme $\langle \nabla^2 f(x^*)\, h, h\rangle \geq 2\vartheta\|h\|^2$ i v tomto případě.

„$\Longleftarrow$“ Nechť platí (6.4.7) pro všechna $x \in X$ a $h \in \mathbb{R}^n$. Nechť jsou $x, x^* \in X$ libovolná a položme $h := x - x^*$. Pak s využitím Taylorova rozvoje druhého řádu dostaneme

$$f(x) - f(x^*) = \langle \operatorname{grad} f(x^*), x - x^*\rangle + \frac{1}{2}\Big\langle \nabla^2 f(\underbrace{x^* + \lambda h}_{\lambda\in(0,1)\Longrightarrow x^*+\lambda h = \lambda x + (1-\lambda)x^* \in X}), h\Big\rangle \geq$$
$$\overset{(6.4.7)}{\geq} \langle \operatorname{grad} f(x^*), x - x^*\rangle + \tfrac{1}{2}\,2\vartheta\|h\|^2 = \langle \operatorname{grad} f(x^*), x - x^*\rangle + \vartheta\|x - x^*\|^2,$$

což s přihlédnutím k (6.4.2) ale znamená, že funkce f je silně konvexní na X s konstantou silné konvexnosti ϑ. ∎

Poznámka 6.4.8. (i) Pro důkaz implikace „$\Longleftarrow$“ jsme vůbec nevyužili předpoklad $\operatorname{int} X \neq \emptyset$, tj. nerovnost (6.4.7) vždy implikuje silnou konvexnost f na X. Avšak opačná implikace již bez tohoto požadavku nemusí být pravdivá, jak lze vidět např. na funkci $f(x_1, x_2) = x_1^2 - x_2^2$ dané na množině $X = \{[x_1, x_2] \in \mathbb{R}^2 \mid x_1 \in \mathbb{R},\ x_2 = 0\}$. Pak totiž

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix},$$

což je indefinitní matice. Jenže současně platí

$$f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda) f(y) = [\lambda x_1 + (1-\lambda) y_1]^2 - \lambda x_1^2 - (1-\lambda) y_1^2 =$$
$$= -\lambda(1-\lambda)(x_1 - y_1)^2,$$

což znamená, že f je silně konvexní funkce s konstantou silné konvexnosti $\vartheta = 1$.

[143] Připomeňme, že pro dvakrát spojitě diferencovatelnou funkci $f : \mathbb{R}^n \to \mathbb{R}$ na množině D a libovolné $x, x^* \in D$ platí *Taylorův rozvoj druhého řádu*

$$f(x) = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^*\rangle + \\ + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*) + \\ + \omega(x - x^*)$$

pro nějakou funkci $\omega : \mathbb{R}^n \to \mathbb{R}$ splňující $\omega \in o(\|x - x^*\|^2)$, tj.

$$\lim_{x\to x^*} \frac{\omega(x - x^*)}{\|x - x^*\|^2} = 0.$$

Obecněji platí

$$f(x^* + h) = f(x^*) + \mathrm{d} f(x^*)(h) + \cdots + \\ + \frac{1}{m!}\mathrm{d}^m f(x^*)(h) + R_m(x^*, h),$$

kde

$$\mathrm{d}^k f(x^*)(h) = \sum_{j_1 + \cdots + j_n = k} \frac{k!}{j_1! \times \cdots \times j_n!} \\ \frac{\partial^k f(x^*)}{\partial x_1^{j_1} \ldots \partial x_n^{j_n}} h_1^{j_1} \ldots h_n^{j_n}$$

a

$$R_m(x^*, h) = \frac{\mathrm{d}^{m+1}}{(m+1)!} f(x^* + \lambda h)(h)$$

pro $h = (h_1, \ldots, h_n)^\top \in \mathbb{R}^n$ a nějaké $\lambda \in (0, 1)$, přičemž platí

$$\lim_{h\to 0} \frac{\|R_m(x^*, h)\|}{\|h\|^m} = 0.$$

Viz str. 34

[144] Mohl by x^* být izolovaným bodem?

[145] Proč?

Podmínku $\operatorname{int} X \neq \emptyset$ lze v tvrzení Důsledku 6.4.7 vypustit, ale pouze když se omezíme jenom na $h \in \operatorname{Lin} X$, aby stále platilo $x = x^* + \lambda h \in X$. V literatuře také bývá mnohdy předpoklad $\operatorname{int} X \neq \emptyset$ nahrazen mnohem silnější podmínkou vyžadující, že X je *otevřená* množina.

(ii) Z Důsledku 6.4.7 bezprostředně plynou následující tři implikace:

(a) jestliže $\nabla^2 f(x) \geq 0$ pro všechna $x \in X$, pak funkce f je konvexní na X;

(b) jestliže $\nabla^2 f(x) > 0$ pro všechna $x \in X$, pak funkce f je ostře konvexní na X;

(c) jestliže $\operatorname{int} X \neq \emptyset$ a f je konvexní na X, pak $\nabla^2 f(x) \geq 0$ pro všechna $x \in X$.

Pro ostrou konvexnost již opačná implikace neplatí, tj. funkce f může být ostře konvexní, i když Hessova matice $\nabla^2 f(x)$ není pozitivně definitní. To můžeme vidět na příkladu z předchozí části této poznámky nebo na funkci $f(x) = \sum_{i=1}^n x_i^4$ na libovolné konvexní množině X obsahující počátek, neboť $\nabla^2 f(x) = 12 \operatorname{diag}\{x_1^2, \dots, x_n^2\}$. ▲

Poznámka 6.4.9. Konvexnost (i silnou) jsme řešili již v Příkladě 6.2.4. Nyní přidáme ještě pár doplňujících poznámek.

(i) Pro $f(x) = \|x\|^2 = \langle x, x \rangle$ máme $\nabla^2 f(x) = 2I$. Pak nerovnost (6.4.7) je tvaru

$$\langle 2h, h \rangle \geq 2\vartheta \|h\|^2,$$

která je pro $h = 0$ splněna triviálně, zatímco pro $h \neq 0$ z ní dostaneme $\vartheta \leq 1$, tj. funkce f je silně konvexní s největší konstantou silné konvexnosti $\vartheta = 1$.

(ii) Uvážíme-li nyní $f(x) = \langle Ax, x \rangle + \langle b, x \rangle + c$ se symetrickou maticí $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ a $c \in \mathbb{R}$ na konvexní množině X, pak[146] $\nabla^2 f(x) = 2A$, a tudíž

(a) pro $A \geq 0$ (tj. $\lambda_{\min}(A) \geq 0$) je f konvexní;

(b) pro $A > 0$ (tj. $\lambda_{\min}(A) > 0$) je f ostře konvexní;

(c) pro konvexní f a $\operatorname{int} X \neq \emptyset$ je $A \geq 0$;

(d) pro $\vartheta := \lambda_{\min}(A) > 0$ je f silně konvexní s konstantou ϑ.

Např. funkce $f(x_1, x_2) = a x_1^2 + b x_1 x_2 + c x_2^2$ odpovídá kvadratické formě

$$f(x_1, x_2) = (x_1, x_2) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

takže

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}.$$

Proto funkce $f(x_1, x_2)$ je konvexní v případě $a \geq 0$, $c \geq 0$ a $4ac - b^2 \geq 0$, zatímco ostrá konvexnost nastane pro $a > 0$ a $4ac - b^2 > 0$. Poněvadž $\nabla^2 f(x)$ nezávisí na x_1, x_2, je tato funkce silně konvexní právě tehdy, když je ostře konvexní. Navíc největší konstanta silné konvexnosti bude mít hodnotu $\vartheta = \lambda_{\min}(A) = a + c - \sqrt{a^2 + c^2 + b^2 - 2ac} > 0$.

(iii) Pro symetrickou matici $A \in \mathbb{R}^{n \times n}$ můžeme pro $h \neq 0$ předpisem

$$\rho_A(h) := \frac{\langle Ah, h \rangle}{\langle h, h \rangle}.$$

definovat[147] tzv. *Rayleighyho* (či *Rayleighyho–Ritzův*) *podíl* $\rho_A : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$. Pak platí tzv. *Min–Max věta* (nebo též *Courantův–Fischerův–Weylův min-max princip*):

> Má-li matice A vlastní čísla $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, pak
>
> $$\lambda_k = \min_{\dim S_k = k} \Big\{ \max_{h \in S_k \setminus \{0\}} \rho_A(h) \Big\} \quad \text{a} \quad \lambda_k = \max_{\dim S_k = n-k+1} \Big\{ \min_{h \in S_k \setminus \{0\}} \rho_A(h) \Big\}.$$
>
> Zejména,
>
> $$\lambda_{\max}(A) = \lambda_n(A) = \max_{h \in \mathbb{R}^n \setminus \{0\}} \rho_A(h) \quad \text{a} \quad \lambda_{\min}(A) = \lambda_1(A) = \min_{h \in \mathbb{R}^n \setminus \{0\}} \rho_A(h).$$

[146] Jak se derivuje kvadratická forma $g(x) := \langle Ax, x \rangle = \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i x_j$? Nejdříve spočteme

$$\frac{\partial g(x)}{\partial x_j} = \sum_{i=1}^n a_{ji} x_i + \\ + a_{1j} x_1 + \dots + a_{nj} x_n = \\ = 2 \sum_{i=1}^n a_{ji} x_i,$$

takže

$$\operatorname{grad} g(x) = 2 \begin{pmatrix} 2\sum_{i=1}^n a_{1i} x_1 \\ \vdots \\ 2\sum_{i=1}^n a_{ni} x_n \end{pmatrix} = \\ = 2 \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \\ = 2Ax$$

díky symetrii matice A – jinak by platilo $\operatorname{grad} g(x) = (A^\top + A)x$. Proto

$$\frac{\partial^2 g(x)}{\partial x_j \partial x_j} = a_{ji} + a_{ji} = 2a_{ij},$$

tudíž

$$\nabla^2 g(x) = 2A.$$

[147] Ekvivalentně by šlo brát $\rho_A(h) = \langle Ah, h \rangle$ pro $\|h\| = 1$.

Důsledek 6.4.7 lze proto zformulovat také takto:

(i) Funkce f je konvexní na X splňující $\operatorname{int} X \neq \emptyset$ právě tehdy, když $\lambda_{\min}(\nabla^2 f(x)) \geq 0$ pro všechna $x \in X$;

(ii) Jestliže $\lambda_{\min}(\nabla^2 f(x)) > 0$ pro všechna $x \in X$, pak funkce f je ostře konvexní na X;

(iii) Funkce f je na X splňující $\operatorname{int} X \neq \emptyset$ silně konvexní s konstantou silné konvexnosti $\vartheta > 0$ právě tehdy, když $\lambda_{\min}(\nabla^2 f(x)) \geq 2\vartheta$ pro každé $x \in X$, tj.

$$\inf_{x \in X} \left(\lambda_{\min}(\nabla^2 f(x))\right) > 0.$$

A jaká bude v takovém případě největší možná hodnota konstanty silné konvexnosti ϑ? ▲

Až doposud jsme pracovali s diferencovatelnými funkcemi. Z kurzu matematické analýzy víme, že tato vlastnost implikuje existenci *směrové derivace* v libovolném směru, avšak opačná implikace již neplatí. Nyní se proto ještě na chvíli podíváme na *jednostranné* směrové derivace ve směru vektoru h a v bodě x^*, čímž pro funkci $f : \mathbb{R}^n \to \mathbb{R}$, bod $x^* \in \mathbb{R}^n$ a vektor $h \in \mathbb{R}^n$ rozumíme

$$f_{h^+}(x^*) := \lim_{t \to 0^+} \frac{f(x^* + th) - f(x^*)}{t}.$$

VĚTA 6.4.10

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f : X \to \mathbb{R}$ konvexní na X a $x^* \in \operatorname{ri} X$. Pak pro všechna $h \in \operatorname{Lin} X$ existuje konečná $f_{h^+}(x^*)$.

Důkaz. Je-li $X = \emptyset$, je tvrzení triviální. Nechť tedy $X \neq \emptyset$. Protože množina X je konvexní, platí také $\operatorname{ri} X \neq \emptyset$. Zvolme proto $x^* \in \operatorname{ri} X$ a $h \in \operatorname{Lin} X$ libovolně. Pak také $x^* + th \in X$ pro $t \in (-\delta, \delta)$ s $\delta > 0$ dostatečně malým (viz Definici 6.1.26). Pro $t > 0$ označme funkci

$$\phi(t) := \frac{f(x^* + th) - f(x^*)}{t}.$$

S využitím konvexnosti funkce f poměrně snadno ukážeme, že $\phi(t)$ je neklesající a zdola ohraničená na intervalu $(0, \delta)$, v kterémžto případě ale nutně existuje konečná $\lim_{t \to 0^+} \phi(t)$ neboli $f_{h^+}(x^*)$. Začneme proto s monotónií. Nechť $0 < t_1 \leq t_2 < \delta$ jsou libovolná. Potom

$$f(x^* + t_1 h) = f\Big(\underbrace{(1 - \tfrac{t_1}{t_2})}_{\in [0,1)} \underbrace{x^*}_{\in X} + \underbrace{\tfrac{t_1}{t_2}}_{\in (0,1]} \underbrace{(x^* + t_2 h)}_{\in X}\Big) \overset{f \text{ je konvexní}}{\leq} \left(1 - \tfrac{t_1}{t_2}\right) f(x^*) + \tfrac{t_1}{t_2} f(x^* + t_2 h),$$

což znamená, že

$$\frac{f(x^* + t_1 h) - f(x^*)}{t_1} \leq \frac{f(x^* + t_2 h) - f(x^*)}{t_2}$$

neboli $\phi(t_1) \leq \phi(t_2)$, tj. funkce ϕ je vskutku neklesající. Nechť nyní $t_0 \in (-\delta, 0)$. Pak pro $t \in (0, \delta)$ platí

$$f(x^*) = f\Big(\underbrace{\tfrac{t}{t - t_0}}_{\in (0,1)} \underbrace{(x^* + t_0 h)}_{\in X} + \underbrace{\tfrac{-t_0}{t - t_0}}_{\in (0,1)} \underbrace{(x^* + th)}_{\in X}\Big) \overset{f \text{ je konvexní}}{\leq} \tfrac{t}{t - t_0} f(x^* + t_0 h) - \tfrac{t_0}{t - t_0} f(x^* + th),$$

z čehož plyne

$$(t - t_0) f(x^*) \leq t f(x^* + t_0 h) - t_0 f(x^* + th)$$

neboli

$$t_0 f(x^* + th) - t_0 f(x^*) \leq t f(x^* + t_0 h) - t f(x^*).$$

Odtud s využitím skutečnosti $t_0 < 0$ dostaneme

$$\frac{f(x^* + t_0 h) - f(x^*)}{t_0} \leq \frac{f(x^* + th) - f(x^*)}{t}, \tag{6.4.9}$$

tj. $\phi(t_0) \leq \phi(t)$ pro všechna $t \in (0, \delta)$. Tedy funkce ϕ je skutečně i zdola ohraničená, čímž je důkaz hotov. ■

Poznámka 6.4.11. (i) Analogickým způsobem bychom mohli ukázat, že funkce ϕ je neklesající a shora ohraničená na intervalu $(-\delta, 0)$ a pro dostatečně malé $t > 0$ platí $\phi(-t) \leq \phi(t)$, viz (6.4.9) s volbou $t_0 = -t$. Pak tedy $\lim_{t\to 0^-} \phi(t) \leq \lim_{t\to 0^+} \phi(t)$.

(ii) Tvrzení Věty 6.4.10 nemusí být pravdivé pro $x^* \in \mathrm{r}\partial X$. Vezmeme-li např. opět $X := [-1, 1]$ a

$$f(x) := \begin{cases} 0, & x \in (-1, 1), \\ 1, & x = \pm 1, \end{cases}$$

pak

$$f_{-1^+}(1) = \lim_{t\to 0^+} \frac{f(1-t) - f(1)}{t} = \lim_{t\to 0^+} \frac{0-1}{t} = -\infty$$

a také

$$f_{1^+}(-1) = \lim_{t\to 0^+} \frac{f(-1+t) - f(-1)}{t} = \lim_{t\to 0^+} \frac{0-1}{t} = -\infty.$$

▲

Cvičení

6.4.1. Pomocí Hessovy matice rozhodněte o (ostré, silné) konvexnosti funkce $f(x) = \|x\|^4$ na $X \subseteq \mathbb{R}^n$.
řešení: vždy konvexní, ostře konvexní pokud $0 \notin X$, silně konvexní pokud $0 \notin \tilde{X}$

6.4.2. Pomocí Hessovy matice rozhodněte o (ostré, silné) konvexnosti funkce $f(x) = \|x\|$ na $X \subseteq \mathbb{R}^n$.
řešení: konvexní vždy, není ostře ani silně konvexní

6.4.3. Rozhodněte, zda funkce

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - x_1 x_2 + x_1 - x_2 + 4$$

je konvexní na $\mathbb{R}^2$. Je zde i ostře nebo silně konvexní? V případě kladné odpovědi určete také největší konstantu silné konvexnosti.
řešení: je silně konvexní s konstantou $\vartheta = (3 - \sqrt{2})/2$

6.4.4. Pomocí Hessovy matice rozhodněte o (ostré, silné) konvexnosti funkce $f(x) = \sqrt{1 + \|x\|^2}$ na $X \subseteq \mathbb{R}^n$, kde $X = \mathbb{R}^n$ a $X = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$.
řešení: je vždy ostře konvexní, není silně konvexní na $\mathbb{R}^n$, v případě $\|x\| \leq 1$ je silně konvexní

6.4.5. Rozhodněte, pro která $[x, y] \in \mathbb{R}^2$ je funkce

$$f(x, y) = 1 - \mathrm{e}^{-x^2-y^2}$$

(ostře) konvexní.
řešení: konvexní na kruhu $x^2 + y^2 \leq 1/2$, ostře konvexní uvnitř tohoto kruhu

6.4.6. Rozhodněte, pro která $[x, y] \in \mathbb{R}^2$ je funkce

$$f(x, y) = -\frac{1}{1 + x^2 + y^2}$$

(ostře) konvexní.
řešení: konvexní na kruhu se středem v počátku a poloměrem $1/\sqrt{3}$, ostře konvexní uvnitř kruhu

6.4.7. Rozhodněte, pro která $[x, y] \in \mathbb{R}^2$ je funkce

$$f(x, y) = \ln(1 + x^2 + y^2).$$

(ostře) konvexní.
řešení: konvexní na kruhu se středem v počátku a poloměrem 1, ostře konvexní uvnitř kruhu

6.4.8. Rozhodněte, zda funkce $f(x, y) = \frac{x^2}{y}$ je konvexní na množině $\{[x, y] \in \mathbb{R}^2 \mid y > 0\}$.
řešení: je konvexní

6.4.9. Rozhodněte, zda funkce

$$f(x,y)=\frac{1}{xy}$$

je konvexní na množině $\mathbb{R}^2_{++}=\{[x,y]\in\mathbb{R}^2 \mid x>0,\ y>0\}$. Je zde silně konvexní?

řešení: je pouze ostře konvexní na $\mathbb{R}^2_{++}$

6.4.10. Nechť

$$f(x,y)=\sqrt{1+x^2+y^2}.$$

Rozhodněte, zda je

a) konvexní na $\mathbb{R}^2$;

b) silně konvexní na $\mathbb{R}^2$;

c) silně konvexní na množině $X=\{[x,y]\in\mathbb{R}^2 \mid x^2+y^2\le 1\}$.

řešení: a) ano; b) ne; c) ano

6.4.11. Rozhodněte, pro které body $[x_1,x_2]\in\mathbb{R}^2$ je funkce

$$f(x_1,x_2)=2x_1\ln x_1-x_1\ln x_2-x_1^2$$

konvexní. Výslednou množinu také načrtněte.

řešení: $x_1\in(0,1/2]$ & $x_2>0$

6.4.12. Rozhodněte, pro které body $[x_1,x_2]\in\mathbb{R}^2$ je funkce

$$f(x_1,x_2)=\frac{x_1^2}{x_2}+\ln\left(\mathrm{e}^{x_1}+\mathrm{e}^{x_2}\right)$$

konvexní. Výslednou množinu také načrtněte a rozhodněte, zda bod $[1,0]$ je vnitřním, hraničním nebo vnějším bodem této množiny.

řešení: konvexní na množině $\{[x_1,x_2\in\mathbb{R}^2 \mid x_1\in\mathbb{R},\ x_2>0]\}$, bod $[1,0]$ je hraničním bodem

6.4.13. Určete maximální množinu (vzhledem k množinové inkluzi) v $\mathbb{R}^2$, na které je splněna postačující podmínka 2. řádu pro ostrou konvexnost funkce

$$f(x_1,x_2)=-\ln(2x_1-x_2^2-2)-\frac{x_2^2}{2}.$$

řešení: $\{[x_1,x_2]\in\mathbb{R}^2 \mid 2x_1-x_2^2-1>0 \ \&\ 2x_1-x_2^2-4<0\}$

6.4.14. Rozhodněte, zda je funkce

$$f(x,y,z)=2x^2+2y^2+z^2+2xy.$$

(ostře, silně) konvexní na $\mathbb{R}^3$. V případě silné konvexnosti určete (největší) konstantu silné konvexnosti ϑ.

řešení: je silně konvexní s $\vartheta=1$

6.4.15. Rozhodněte, zda je funkce

$$f(x,y,z)=x^2+3y^2+z^2+xy-yz.$$

(ostře, silně) konvexní na $\mathbb{R}^3$. V případě silné konvexnosti určete (největší) konstantu silné konvexnosti ϑ.

řešení: je silně konvexní s $\vartheta=2-\sqrt{6}/2$

6.4.16. Rozhodněte, zda je funkce

$$f(x,y,z)=2x^2+y^2+3z^2+2\sqrt{2}xz.$$

(ostře, silně) konvexní na $\mathbb{R}^3$. V případě silné konvexnosti určete (největší) konstantu silné konvexnosti ϑ.

řešení: je silně konvexní s $\vartheta=1$

6.4.17. Rozhodněte, zda je funkce

$$f(x_1,x_2,x_3) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 8x_1 - 6x_2 - 4x_3 + 9.$$

(ostře, silně) konvexní na $\mathbb{R}^3$. V případě silné konvexnosti určete (největší) konstantu silné konvexnosti ϑ.

řešení: je konvexní, není ostře ani silně konvexní

6.4.18. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = 3x_1^2 + x_2^2 + 4x_3^2 + 2\sqrt{2}\,x_1x_2 - 4x_1x_3 - 7x_1 + 8x_2 - x_3 + 6$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní?

řešení: je konvexní, není ostře ani silně konvexní

6.4.19. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = x_1^2 + 3x_2^2 + x_3^2 + 2x_1x_2 - 2x_1x_3 - 2x_2x_3 + 11x_1 + 2x_2 - x_3 + 6$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní?

řešení: je konvexní, není ostře ani silně konvexní

6.4.20. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = 2x_1^2 + 4x_2^2 + 3x_3^2 + 4x_1x_3 - 4x_2x_3 + 6x_1 - 3x_2 + x_3 + 7$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní?

řešení: je konvexní, není ostře ani silně konvexní

6.4.21. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = 4x_1^2 + 3x_2^2 + 2x_3^2 + 4x_1x_2 - 4x_2x_3 + x_1 + 4x_2 - 2x_3 + 9$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní? V případě silné konvexnosti také určete největší konstantu silné konvexnosti ϑ.

řešení: je konvexní, není ostře ani silně konvexní

6.4.22. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = 5x_1^2 + 3x_2^2 + x_3^2 - 2x_1x_2 + 2x_2x_3 + 6x_1 - 7x_2 + 7$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní? V případě silné konvexnosti také určete největší konstantu silné konvexnosti ϑ.

řešení: je silně konvexní s $\vartheta = 3 - \sqrt{6}$

6.4.23. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = 2x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 + 4x_2x_3 + x_1 - 6x_2 + 2x_3 + 4$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní? V případě silné konvexnosti také určete největší konstantu silné konvexnosti ϑ.

řešení: funkce není konvexní

6.4.24. Rozhodněte, zda funkce

$$f(x_1,x_2,x_3) = x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + 8x_1x_3 + 2x_2x_3 + 6x_1 - 11x_2 + 4$$

je konvexní na $\mathbb{R}^3$. Je zde i ostře nebo silně konvexní? V případě silné konvexnosti také určete největší konstantu silné konvexnosti ϑ.

řešení: funkce není konvexní

6.4.25. Rozhodněte, zda funkce n proměnných

$$f(x) = \frac{1}{\langle a, x\rangle}, \quad a \in \mathbb{R}^n_{++},$$

je konvexní pro $x \in \mathbb{R}^n_{++} := \{x = (x_1,\dots,x_n)^\top \in \mathbb{R}^n : x_1 > 0,\dots,x_n > 0\}$.

řešení: je konvexní

6.4.26. Rozhodněte, zda funkce n proměnných

$$f(x) = \ln\Big(\sum_{i=1}^{n} e^{x_i}\Big),$$

je konvexní na $\mathbb{R}^n$.

řešení: je konvexní

6.4.27. Rozhodněte, zda funkce n proměnných

$$f(x) = e^{a\langle Ax,x\rangle}$$

je konvexní na $\mathbb{R}^n$, kde $a \in \mathbb{R}$ je kladné číslo a $A \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně semidefinitní matice.

řešení: je konvexní

6.4.28. Nechť $A \subseteq \mathbb{R}^n$ je konvexní a uzavřená množina. Dokažte, že funkce určující vzdálenost bodu x od množiny A, tj. $f(x) = \rho(x, A) := \inf_{a\in A}\|x - a\|$, je konvexní.

6.4.29. Nechť $X \subseteq \mathbb{R}^n$ je omezená konvexní množina, pro níž $0 \in \operatorname{int} X$. Rozhodněte, zda funkce definovaná předpisem

$$f(x) := \inf\{\alpha > 0 : x \in \alpha X\}$$

je konvexní.

řešení: je konvexní

6.5 Subgradient a subdiferenciál

Pro řešení minimalizačních úloh s nediferencovatelnou funkcí f, tj. pro tzv. *nehladkou optimalizační úlohu*, jsou velmi užitečným nástrojem tzv. *subgradient* a *subdiferenciál* funkce f, které hrají podobně důležitou roli jako gradient v případě diferencovatelných funkcí. Připomeňme, že podle Věty 6.4.3 pro diferencovatelnou konvexní funkci $f : X \to \mathbb{R}$ platí

$$f(x) \geq f(x^*) + \langle \operatorname{grad} f(x^*), x - x^*\rangle \quad \text{pro všechna } x, x^* \in X,$$

tj. vektor $\operatorname{grad} f(x^*)$ určuje *nevertikální* opěrnou nadrovinu k množině $\operatorname{epi} f$ v bodě $[x^*, f(x^*)]$, viz Poznámku 6.4.4(ii). Ovšem co kdyby f nebyla diferencovatelná?

DEFINICE 6.5.1

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina. Vektor $a \in \mathbb{R}^n$ se nazývá *subgradient* funkce $f : X \to \mathbb{R}$ v bodě $x^* \in X$, jestliže

$$f(x) - f(x^*) \geq \langle a, x - x^*\rangle \tag{6.5.1}$$

pro každé $x \in X$. Množina všech subgradientů funkce f v bodě x^* se nazývá *subdiferenciál* funkce f v bodě x^* a značí se $\partial f(x^*)$. Funkce f se nazývá *subdiferencovatelná* v bodě x^*, jestliže $\partial f(x^*) \neq \emptyset$.

Poznámka 6.5.2.

(i) Zjevně platí $\operatorname{grad} f(x^*) \in \partial f(x^*)$ díky Větě 6.4.3, pokud $\operatorname{grad} f(x^*)$ existuje. Lze říci i více? Ano, ale až později ve Větě 6.5.9. Každopádně už nyní by mělo být zřejmé, že na rozdíl od derivace nemusí být přiřazení $x \to \partial f(x)$ zobrazením[148], tj. $\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ je *víceznačná funkce* či *korespondence*.

(ii) Není těžké ukázat rozšíření Fermatova principu na nediferencovatelné funkci, tj. pro funkci $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ platí

$$f(x^*) = \min_{x\in X} f(x)$$

právě tehdy, když $0 \in \partial f(x^*)$.

[148] Může totiž postrádat jednoznačnost, takže jeho grafické znázornění může vypadat velmi odlišně, než jak jsme zvyklí z funkcí.

(iii) Je-li množina $X \subseteq \mathbb{R}^n$ konvexní, funkce $f_1, \dots, f_m : X \to \mathbb{R}$ konvexní na X a $\alpha_1, \dots, \alpha_m \geq 0$, pak

$$\partial\big(\alpha_1 f_1(x^*) + \cdots + \alpha_m f_m(x^*)\big) = \alpha_1 \partial f_1(x^*) + \cdots + \alpha_m \partial f_m(x^*),$$

tj. subdiferenciál je to lineární zobrazení. Avšak jsou-li $f_i : X_i \to \mathbb{R}$ pro $i \in \{1, \dots, m\}$, pak platí pouze

$$\partial\big(\alpha_1 f_1(x^*) + \cdots + \alpha_m f_m(x^*)\big) \supseteq \alpha_1 \partial f_1(x^*) + \cdots + \alpha_m \partial f_m(x^*).$$

Je-li navíc $\bigcap_{i=1}^m \operatorname{ri} X_i \neq \emptyset$, pak pro každé $x \in \bigcap_{i=1}^m \operatorname{ri} X_i = \operatorname{ri}(\bigcap_{i=1}^m X_i)$ nastane rovnost. ▲

Příklad 6.5.3

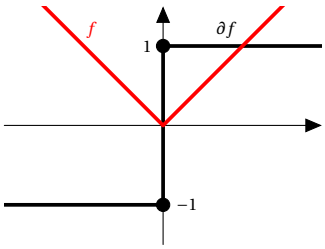
Uvažme $f(x) = |x|$. Pak pro $x^* = 0$ dostaneme z Definice 6.5.1, že $a \in \partial f(0)$ musí splňovat $|x| - 0 \geq a(x - 0)$ pro každé $x \in \mathbb{R}$. Pro $x = 0$ je toto splněno triviálně, zatímco pro $x > 0$ máme $x \geq ax$ neboli $a \leq 1$ a pro $x < 0$ bude $-x \geq ax$ neboli $a \geq -1$, tj. $\partial f(0) = [-1, 1]$. Vezmeme-li nyní $x^* > 0$, pak se nerovnost (6.5.1) redukuje na

$$|x| - x^* \geq a(x - x^*) \quad \text{pro každé } x \in \mathbb{R}. \tag{6.5.2}$$

Opět pro $x = x^*$ je nerovnost splněna triviálně, zatímco pro $x - x^* > 0$ bude $x > x^* > 0$, a tedy (6.5.2) vede na $x - x^* \geq a(x - x^*)$ neboli $a \leq 1$. V případě $x - x^* < 0$ musíme rozlišit dvě různé situace: je-li $x \in [0, x*)$, pak z (6.5.2) získáme zase $x - x^* \geq a(x - x^*)$, což ale tentokrát znamená, že $1 \leq a$. To dohromady s předchozí nerovností znamená, že jediným možným subgradientem je $a = 1$. Zbývá ověřit, zda pro tuto volbu platí nerovnost (6.5.2) i pro $x < 0$, tj. $-x - x^* \geq x - x^*$ neboli $-2x \geq 0$, což je jistě pravdivé. Proto $\partial f(x^*) = \{1\}$ pro libovolné $x^* > 0$. Velmi podobně můžeme ukázat, že $\partial f(x^*) = \{-1\}$ pro libovolné $x^* < 0$. Celkem tedy máme

$$\partial f(x) = \begin{cases} \{1\}, & x > 0, \\ [-1, 1], & x = 0, \\ \{-1\}, & x < 0, \end{cases}$$

viz Obrázek 6.75.



Obrázek 6.75: Grafy f a ∂f z Příkladu 6.5.3.

Poznámka 6.5.4.

(i) Je-li $f : X \subseteq \mathbb{R} \to \mathbb{R}$ konvexní a $x^* \in \operatorname{ri} X$, pak podle Věty 6.4.10 a Poznámky 6.4.11 existují jednostranné derivace $f'_+(x^*)$ a $f'_-(x^*)$, přičemž platí $f'_-(x^*) \leq f'_+(x^*)$. Proto v tomto případě máme $\partial f(x^*) = [f'_-(x^*), f'_+(x^*)]$, což zcela koresponduje s výsledkem Příkladu 6.5.3.

(ii) Podmínka (6.5.1) (podobně jako (6.4.2) ve Větě 6.4.3) znamená, že graf funkce f

$$G_f := \{[x, \beta] \in \mathbb{R}^{n+1} \mid \beta = f(x) \text{ pro } x \in X\}$$

neleží pod grafem

$$H := \{[x, \beta] \in \mathbb{R}^{n+1} \mid \beta = \ell(x) \text{ pro } x \in \mathbb{R}^n\}$$

afinní funkce $\ell(x) := f(x^*) + \langle a, x - x^* \rangle$. Současně $H = H_{p,\beta}$, kde $H_{p,\beta}$ je nadrovina s $p := (-a^\top, 1)^\top \in \mathbb{R}^{n+1}$ a $\beta := f(x^*) - \langle a, x^* \rangle$, je opěrnou nadrovinou k nadgrafu funkce f v bodě $[x^*, f(x^*)]$ neboli

> $a \in \partial f(x^*)$ právě tehdy, když $H_{p,\beta}$ je opěrná nadrovina.

▲

V důkazech některých základních vlastností subdiferencovatelných funkcí bude hrát následující lemma zásadní roli.

Lemma 6.5.5

Nechť $X, Y \subseteq \mathbb{R}^n$ jsou konvexní množiny splňující $\operatorname{ri} X \cap \operatorname{ri} Y \neq \emptyset$, funkce $f : X \to \mathbb{R}$ konvexní na X, funkce $g : Y \to \mathbb{R}$ konkávní na Y a současně platí $f(x) \geq g(x)$ pro všechna $x \in \operatorname{ri} X \cap \operatorname{ri} Y$. Pak existuje afinní funkce $\ell(x) := \langle a, x\rangle + b$ taková, že pro každé $x \in X$ a $y \in Y$ platí

$$f(x) \geq \ell(x) \quad \& \quad \ell(y) \geq g(y). \tag{6.5.3}$$

Důkaz. Nechť

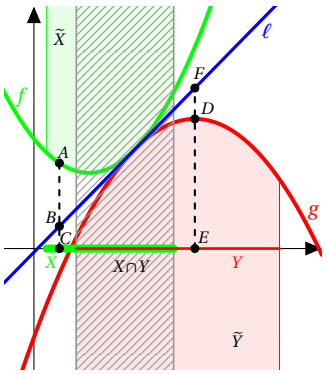
$$\widetilde{X} := \operatorname{epi} f = \{[x,\beta] \in \mathbb{R}^{n+1} \mid f(x) \leq \beta,\ x \in X\} \quad \text{a} \quad \widetilde{Y} := \{[y,\gamma] \in \mathbb{R}^{n+1} \mid g(y) \geq \gamma,\ y \in Y\},$$

tj. $\widetilde{Y}$ je podgraf funkce g, viz Obrázek 6.76. Pak množiny $\widetilde{X}$ a $\widetilde{Y}$ jsou konvexní a podle Věty 6.3.18 vlastně oddělitelné, neboť podmínka $f(x) \geq g(x)$ zaručuje $\operatorname{ri}\widetilde{X} \cap \operatorname{ri}\widetilde{Y} = \emptyset$. Existuje tedy $(p,\lambda)^\top \in \mathbb{R}^{n+1} \setminus \{0\}$ takový, že

$$\big\langle (x,\beta)^\top, (p,\lambda)^\top \big\rangle \geq \big\langle (y,\gamma)^\top, (p,\lambda)^\top \big\rangle \quad \text{pro všechna } (x,\beta)^\top \in \widetilde{X} \text{ a } (y,\gamma)^\top \in \widetilde{Y}, \tag{6.5.4}$$

a současně existují $[\tilde{x},\tilde{\beta}] \in \widetilde{X}$, $[\tilde{y},\tilde{\gamma}] \in \widetilde{Y}$ takové, že pro tyto body nastane v (6.5.4) ostrá nerovnost. Tedy existuje $\delta \in \mathbb{R}$ splňující

$$\langle p, x\rangle + \lambda\beta \geq \delta \geq \langle p, y\rangle + \lambda\gamma. \tag{6.5.5}$$



Obrázek 6.76: Vizualizace tvrzení Lemma 6.5.5 s vyznačenými body $A := [x, f(x)]$, $B := [x, \ell(x)]$, $C := [x, 0]$, $D := [y, g(x)]$, $E := [y, \ell(y)]$, $F := [y, 0]$.

Kdyby $\lambda = 0$, pak by byly vlastně oddělitelné dokonce samotné množiny X a Y, což je ale spor s Větou 6.3.18. Taktéž dostaneme spor i v případě $\lambda < 0$, když vezmeme $\beta \to \infty$ a $\gamma \to -\infty$. Proto nutně $\lambda > 0$. Vydělením nerovnosti (6.5.5) číslem λ, dostaneme

$$\big\langle \tfrac{1}{\lambda} p, x\big\rangle + \beta \geq \delta/\lambda \geq \big\langle \tfrac{1}{\lambda} p, y\big\rangle + \gamma. \tag{6.5.6}$$

Volbou $a := -\frac{1}{\lambda} p$ a $b := \delta/\lambda$ pak $\ell(x) := \langle a, x\rangle + b$ je hledanou afinní funkcí, která splňuje obě nerovnosti v (6.5.3), neboť pro libovolná $x \in X$ a $y \in Y$ vede nerovnost (6.5.6) pro odpovídající dvojice $[x,\beta] = [x, f(x)]$ a $[y,\gamma] = [y, g(y)]$ na

$$f(x) - \langle a, x\rangle \geq b \geq g(y) - \langle a, y\rangle. \qquad \blacksquare$$

Nyní můžeme charakterizovat množinu $\partial f(x)$ pro konvexní funkce.

Věta 6.5.6

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$.

(i) Je-li funkce f konvexní a $x^* \in \operatorname{ri} X$, pak $\partial f(x^*)$ je neprázdná, uzavřená a konvexní množina.

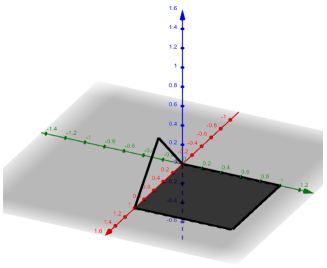
(ii) Je-li $\partial f(x)$ neprázdná pro každé $x \in X$, pak f je konvexní na X.

Poznámka 6.5.7.

(i) Tvrzení ve Větě 6.5.6(i) nelze obrátit, tj. nahradit ve Větě 6.5.6(ii) „každé $x \in X$“ za „každé $x \in \operatorname{ri} X$“. Uvažme např.

$$X = [0,1] \times [0,1] \subseteq \mathbb{R}^2 \quad \& \quad f(x,y) = \begin{cases} 0, & 0 \leq x \leq 1,\ 0 < y \leq 1, \\ \frac{1}{2} - |x - \frac{1}{2}|, & 0 \leq x \leq 1,\ y = 0, \end{cases}$$

jejíž graf je na Obrázku 6.77. Tato funkce f má sice jediný subgradient pro každé $x \in \operatorname{int} X = \operatorname{ri} X$, kterým je vektor[149] $a = (0,0)^\top$, ale f není konvexní[150].



Obrázek 6.77: Graf funkce f z Poznámky 6.5.7(i).

(ii) Požadavek $x^* \in \operatorname{ri} X$ je pro neprázdnost $\partial f(x^*)$ klíčový. Uvažme např. $X = [-1,1]$ a $f(x) = -\sqrt{1-x^2}$, viz Obrázek 6.78. V každém bodě $x^* \in (-1,1)$ existuje subgradient, je jediný a platí

$$\partial f(x^*) = \{f'(x^*)\}.$$

A v bodech $x = \pm 1$? Zvolme $x^* = -1$. Pak subgradientem by mělo být číslo $a \in \mathbb{R}$ takové, že

$$-\sqrt{1-x^2} \geq a(x+1)$$

pro všechna $x \in [-1,1]$. Pro $x=-1$ je nerovnost splněna triviálně, zatímco pro $x \in (-1,1]$ dostaneme po vydělení výrazem $x+1>0$ nerovnost

$$-\frac{\sqrt{1-x^2}}{x+1} \geq a \quad \text{neboli} \quad -\sqrt{\frac{1-x}{1+x}} \geq a.$$

Avšak pro $x \to -1^+$ bude výraz na pravé straně klesat pod jakoukoli hodnotu, což znamená, že neexistuje žádné $a \in \mathbb{R}$, pro které bude tat nerovnost splněna[151].

(iii) Část (i) v předchozí větě vlastně říká, že pro konvexní funkci f existuje pro každé $x^* \in \operatorname{ri} X$ vektor $a \in \mathbb{R}^n$ takový, že

$$f(x) - f(x^*) \geq \langle a, x-x^* \rangle \quad \text{pro všechna } x \in X \setminus \{x^*\}. \tag{6.5.7}$$

Je-li funkce f ostře konvexní, pak dokonce existuje $a \in \mathbb{R}^n$ takové, že nerovnost (6.5.7) je ostrá. Naopak, pokud pro každé $x \in X$ existuje $a \in \partial f(x)$ tak, že

$$f(y) - f(x) > \langle a, y-x \rangle \quad \text{pro všechna } y \in X \setminus \{x\},$$

pak je f ostře konvexní.

(iv) Dá se ukázat, že pro konvexní funkci je $\partial f(x^*)$ dokonce kompaktní množina. ▲

Důkaz Věty 6.5.6. (i) Uzavřenost plyne přímo z definice stejně jako konvexnost, neboť v případě alespoň dvoubodové množiny $\partial f(x^*)$ máme pro $a, b \in \partial f(x^*)$ nerovnosti

$$f(x) - f(x^*) \geq \langle a, x-x^* \rangle \quad \text{a} \quad f(x) - f(x^*) \geq \langle b, x-x^* \rangle.$$

Pak pro libovolné $\lambda \in [0,1]$ dostaneme sečtením λ-násobku první nerovnosti a $(1-\lambda)$-násobku druhé nerovnosti vztah

$$\underbrace{\lambda\big(f(x)-f(x^*)\big) + (1-\lambda)\big(f(x)-f(x^*)\big)}_{f(x)-f(x^*)} \geq \underbrace{\lambda\langle a, x-x^*\rangle + (1-\lambda)\langle b, x-x^*\rangle}_{\langle \lambda a+(1-\lambda)b, x-x^*\rangle},$$

tj. $\lambda a + (1-\lambda) b \in \partial f(x^*)$.

Zbývá ukázat, že $\partial f(x^*) \neq \emptyset$. Nechť $x^* \in \operatorname{ri} X$ je libovolný. Definujme množinu $Y := \{x^*\}$ a funkci $g : Y \to \mathbb{R}$ jako $g(y) := f(x^*)$. Potom $\operatorname{ri} X \cap \operatorname{ri} Y = \{x^*\} \neq \emptyset$ a $f(x) \geq g(y)$ pro $x \in \operatorname{ri} X \cap \operatorname{ri} Y$. Funkce f je konvexní a g je konkávní, takže podle Lemma 6.5.5 existuje afinní funkce $\ell(x) = \langle a, x\rangle + b$ taková, že $f(x) \geq \ell(x)$ pro všechna $x \in X$ a $\ell(y) \geq g(y)$ pro všechna $y \in Y$, tj.

$$\langle a, x^*\rangle + b \geq f(x^*) \quad \text{a} \quad \langle a, x\rangle + b \leq f(x),$$

z čehož vyplývá

$$f(x) - f(x^*) \geq \langle a, x\rangle + b - \langle a, x^*\rangle - b = \langle a, x-x^*\rangle,$$

tudíž $a \in \partial f(x^*)$.

V situaci, kdy $x^* \in \operatorname{int} X$, lze neprázdnost $\partial f(x)$ dokázat i bez využití Lemma 6.5.5. Protože $[x^*, f(x^*)] \in \partial(\operatorname{epi} f)$ a současně $\operatorname{epi} f$ je konvexní množina, dostaneme z Věty 6.3.14 existenci opěrné nadroviny k $\operatorname{epi} f$ v bodě $[x^*, f(x^*)]$, tj. existuje vektor $(p^\top, \beta) \in \mathbb{R}^{n+1} \setminus \{0\}$ s $a \in \mathbb{R}^n$ a $\beta \in \mathbb{R}$ takový, že

$$\left\langle \begin{pmatrix} p \\ \beta \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \leq \left\langle \begin{pmatrix} p \\ \beta \end{pmatrix}, \begin{pmatrix} x^* \\ f(x^*) \end{pmatrix} \right\rangle \tag{6.5.8}$$

pro všechna $[x, y] \in \operatorname{epi} f$, kde $x \in X$ a $y \in \mathbb{R}$. Nerovnost (6.5.8) je ekvivalentní s

$$\langle p, x\rangle + \beta y \leq \langle p, x^*\rangle + \beta f(x^*) \tag{6.5.9}$$

pro každé $x \in X$ a $y \geq f(x)$. Ukážeme, že $\beta < 0$. Připusťme, že $\beta > 0$. Pak výraz na pravé straně (6.5.9) je konstantní, takže v tomto případě snadno obdržíme spor s (6.5.9), pokud vezmeme y dostatečně velké. Jestliže $\beta = 0$, pak $p \neq 0$ a (6.5.9) se redukuje na

$$\langle p, x\rangle \leq \langle p, x^*\rangle \tag{6.5.10}$$

[149] Vezmeme-li $x^* \in \operatorname{ri} X$, pak $f(x^*) = 0$ a hledáme $a = (a_1, a_2)^\top \in \mathbb{R}^2$ takové, že

$$f(x,y) \geq a_1 (x-x^*) + a_2 (y-y^*)$$

pro všechna $[x,y] \in X$. Volbou $y = y^* (\neq 0)$ dostaneme

$$0 = f(x, y^*) \geq a_1 (x-x^*),$$

která je splněna pouze pro $a_1 = 0$, neboť $x^* \in (0,1)$, zatímco $x \in [0,1]$ je libovolné.

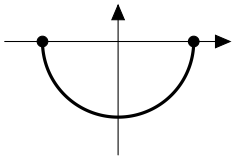
[150] Zvolíme-li body $x_1 = [0,0]$ a $x_2 = [1,0]$, pak

$$f(\lambda x_1 + (1-\lambda) x_2) = f((1-\lambda)x_2) =$$
$$= f(1-\lambda, 0) = \frac{1}{2} - |\frac{1}{2} - \lambda|,$$

zatímco $\lambda f(x_1) + (1-\lambda) f(x_2) =$, tj. konvexnost je ekvivalentní s nerovností

$$\frac{1}{2} - |\frac{1}{2} - \lambda| \leq 0,$$

která není splněna pro žádné $\lambda \in (0,1)$. Případně lze nekonvexnost dané funkce vidět z jejího nadgrafu, neboť ten např. neobsahuje úsečku spojující body $[0,0,0]$ a $[1,0,0]$.



Obrázek 6.78: Graf funkce f z Poznámky 6.5.7(ii).

[151] Jak asi bude vypadat opěrná nadrovina k nadgrafu funkce f v bodě $x^* = 1$? A co se bude dít s tečnami, když je budeme konstruovat stále blíž a blíž k bodu $x^* = 1$?

pro všechna $x \in X$. Ovšem $x^* \in \operatorname{int} X$, takže existuje $\varepsilon > 0$ takové, že $x_1 := x^* + \varepsilon p \in X$. Dosazením x_1 do (6.5.10) dostaneme

$$\langle p, x_1 \rangle = \langle p, x^* \rangle + \varepsilon \langle p, p \rangle \le \langle p, x^* \rangle,$$

tj. $\varepsilon \| p \|^2 \le 0$. Ovšem toto je možné pouze pro $\| p \| = 0$ neboli $p = 0$, což je spor.

Tedy nutně $\beta < 0$. Pokud (6.5.9) vydělíme číslem β a nahradíme-li $y = f(x)$, máme

$$\tfrac{1}{\beta} \langle p, x \rangle + f(x) \ge \tfrac{1}{\beta} \langle p, x^* \rangle + f(x^*)$$

neboli

$$f(x) - f(x^*) \ge \langle \underbrace{\tfrac{1}{\beta} p}_{=:a}, x - x^* \rangle \tag{6.5.11}$$

pro každé $x \in X$. Pak tedy $a := \frac{1}{\beta} p \in \partial f(x^*)$, neboť pak je (6.5.11) tvaru (6.5.1).

(ii) Jestliže X je prázdná nebo jednobodová, tak tvrzení je splněno triviálně. Nechť tedy X je alespoň dvoubodová a zvolme $x_1, x_2 \in X$ a $\lambda \in [0,1]$ libovolně. Položme $x^* := \lambda x_1 + (1-\lambda) x_2 \in X$. Podle předpokladů existuje $a \in \partial f(x^*)$, takže

$$f(x_1) - f(x^*) \ge \langle a, x_1 - x^* \rangle \quad \text{a} \quad f(x_2) - f(x^*) \ge \langle a, x_2 - x^* \rangle.$$

Sečteme-li λ-násobek první nerovnosti a $(1-\lambda)$-násobek druhé nerovnosti dostaneme

$$\lambda f(x_1) + (1-\lambda) f(x_2) - (\lambda + 1 - \lambda) f(x^*) \ge \lambda \langle a, x_1 - x^* \rangle + (1-\lambda) \langle a, x_2 - x^* \rangle = 0,$$

tj. $\lambda f(x_1) + (1-\lambda) f(x_2) \ge f(x^*) = f(\lambda x_1 + (1-\lambda) x_2)$, což dokazuje konvexnost f na X. ■

Subdiferenciál je také velmi úzce spojen s jednostrannou směrovou derivací. Je-li $x^* \in \operatorname{ri} X$, pak nerovnost (6.5.1) je ekvivalentní s

$$\frac{f(x^* + th) - f(x^*)}{t} \ge \langle a, h \rangle$$

pro všechna $h \in \operatorname{Lin} X$ a $t > 0$ dostatečně malé[152]. Pak podíl na levé straně konverguje pro $t \to 0^+$ k $f_{h^+}(x^*)$, viz Větu 6.4.10, takže nerovnost (6.5.1) je ekvivalentní s $f_{h^+}(x^*) \ge \langle a, h \rangle$. Tedy platí

$$a \in \partial f(x^*) \quad \text{právě tehdy, když} \quad f_{h^+}(x^*) \ge \langle a, h \rangle \quad \text{pro všechna } h \in \operatorname{Lin} X,$$

takže nutně také

$$f_{h^+}(x^*) \ge \max_{a \in \partial f(x^*)} \langle a, h \rangle. \tag{6.5.12}$$

Jenže ono platí ještě více.

[152] To jistě díky $x^* \in \operatorname{ri} X$ existuje takové, že $x^* + th \in X$.

Věta 6.5.8

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f : X \to \mathbb{R}$ konvexní na X a $x^* \in \operatorname{ri} X$. Pak platí

$$\partial f(x^*) = \{ a \in \mathbb{R}^n \mid \langle a, h \rangle \le f_{h^+}(x^*) \text{ pro všechna } h \in \operatorname{Lin} X \} \tag{6.5.13}$$

a pro každé $h \in \operatorname{Lin} X$ máme

$$f_{h^+}(x^*) = \max_{a \in \partial f(x^*)} \langle a, h \rangle. \tag{6.5.14}$$

Důkaz. Označme jako A množinu stojící na pravé straně rovnosti (6.5.13). Postupně ukážeme, že $\partial f(x^*) \subseteq A$ a současně $\partial f(x^*) \supseteq A$.

„$\subseteq$“: Nechť $a \in \partial f(x^*)$. Protože $x^* \in \operatorname{ri} X$, je také $x^* + th \in X$ pro libovolné $h \in \operatorname{Lin} X$ a dostatečně malá $t > 0$. Pak podle Definice 6.5.1 platí

$$f(x^* + th) - f(x^*) \ge \langle a, th \rangle,$$

což po vydělení $t>0$ a limitním přechodem pro $t\to 0^+$ dává

$$\underbrace{\frac{f(x^*+th)-f(x^*)}{t}}_{\to f_{h^+}(x^*)} \geq \langle a,h\rangle$$

pro všechna $h\in \operatorname{Lin} X$, tj. $a\in A$.

„$\supseteq$“: Naopak, nechť $a\in A$ a $x\in X$ je libovolné. Označme $h:=x-x^*$. Potom platí

$$f(x)-f(x^*)=\underbrace{\frac{f(x^*+h)-f(x^*)}{1}}_{\substack{=\phi(1)\ \&\ \text{pro každé } h \text{ je } \phi(t) \text{ neklesající}\\ \text{viz důkaz Věty 6.4.10}}} \geq f_{h^+}(x^*)\geq \langle a,h\rangle = \langle a, x-x^*\rangle,$$

což ale znamená, že $a\in \partial f(x^*)$.

Tedy (6.5.13) je dokázáno a zbývá ukázat platnost (6.5.14). Vzhledem k (6.5.12) stačí najít $\widetilde{a}\in \partial f(x^*)$ takové, že

$$f'_{h^+}\leq \langle \widetilde{a},h\rangle \quad \text{pro všechna } h\in \operatorname{Lin} X.$$

Nechť $h\in \operatorname{Lin} X$ je libovolný. Uvažme následující množinu[153]

$$Y:=\{y\in\mathbb{R}^n \mid y=x^*+th,\ t\in\mathbb{R},\ t>0\}.$$

Definujme funkci $g: Y\to\mathbb{R}$ předpisem

$$g(y):=f(x^*)+\frac{\|y-x^*\|}{\|h\|}f_{h^+}(x^*).$$

Dosazením $y=x^*+th$ dostaneme $g(y)=f(x^*)+t\,f_{h^+}(x^*)$, a tudíž g je afinní funkce na Y, což implikuje i její konkávnost. Současně $\operatorname{ri} X\cap \operatorname{ri} Y\neq\emptyset$, neboť $x^*\in \operatorname{ri} X$ implikuje existenci $t>0$ takového, že $x^*+th\in X$, viz Věta 6.1.30. Z důkazu Věty 6.4.10 také víme, že $\phi(t)\geq f_{h^+}(x^*)$ pro $t>0$ malé, takže $f(y)=f(x^*+th)\geq f(x^*)+t\,f_{h^+}(x^*)=g(y)$ pro $y\in Y$, a tedy také pro $y\in \operatorname{ri} Y$. Proto $f(x)\geq g(x)$ pro všechna $x\in \operatorname{ri} X\cap \operatorname{ri} Y$, a tudíž dle Lemma 6.5.5 existuje afinní funkce $\ell(x):=\langle \widetilde{a},x\rangle+b$ taková, že platí (6.5.3), tj.

$$f(x)\geq \langle \widetilde{a},x\rangle+b \quad\text{a}\quad g(y)=f(x^*)+t\,f_{h^+}(x^*)\leq \langle \widetilde{a},x^*+th\rangle+b.$$

Odečtením těchto nerovností získáme

$$f(x)-f(x^*)-t\,f_{h^+}(x^*)\geq \langle \widetilde{a},x\rangle+b-\langle \widetilde{a},x^*+th\rangle-b=\langle \widetilde{a},x-x^*\rangle-t\langle \widetilde{a},h\rangle$$

neboli

$$f(x)-f(x^*)-\langle \widetilde{a},x-x^*\rangle\geq t\big[f_{h^+}(x^*)-\langle \widetilde{a},h\rangle\big].$$

Volbou $t=0$ dostaneme, že $\widetilde{a}\in\partial f(x^*)$. Naopak, vydělením t a přechodem pro $t\to\infty$ obdržíme $0\geq f_{h^+}(x^*)-\langle \widetilde{a},h\rangle$ neboli $\langle \widetilde{a},h\rangle\geq f_{h^+}(x^*)$, čímž je důkaz hotov[154]. ■

Na závěr si ještě upřesníme vztah mezi gradientem a subgradientem, o čemž jsme se částečně zmínili již v první části poznámky Poznámka 6.5.2.

VĚTA 6.5.9

Nechť $X\subseteq\mathbb{R}^n$ je konvexní množina, funkce $f: X\to\mathbb{R}$ konvexní a $x^*\in \operatorname{int} X$.

(i) Je-li f diferencovatelná v x^*, pak $\partial f(x^*)$ je jednoprvková a platí $\partial f(x^*)=\{\operatorname{grad} f(x^*)\}$, tj. gradient je jediným subgradientem.

(ii) Je-li $\partial f(x^*)=\{a\}$ jednoprvková množina, pak je funkce f diferencovatelná v bodě x^* a platí $\operatorname{grad} f(x^*)=a$.

Důkaz. (i) Jestliže existuje $x^*\in\operatorname{int} X$, pak nutně $\operatorname{Lin} X=\mathbb{R}^n$. Již víme, že $\operatorname{grad} f(x^*)\in\partial f(x^*)$. Nechť $a\in\partial f(x^*)$ je libovolný vektor. Protože funkce f je diferencovatelná, pro libovolné $h\in\operatorname{Lin} X$ platí

$$f_{h^+}(x^*)=\langle \operatorname{grad} f(x^*),h\rangle,$$

[153] Jde o polopřímku vycházející z bodu x^* ve směru h.

[154] Pokud bychom se pro „jednoduchost“ omezili na případ $x^*\in\operatorname{int} X$, pak můžeme brát $h\in\operatorname{Lin} X=\mathbb{R}^n$ a nepotřebujeme Lemma 6.5.5. Nechť tedy $x^*\in\operatorname{int} X$ a $h\in\mathbb{R}^n$ jsou libovolná. Zaveďme množiny

$$C_1:=\{[x,y]\in\mathbb{R}^{n+1}\mid f(x)<y\}$$

a

$$C_2:=\{[x,y]\in\mathbb{R}^{n+1}\mid x=x^*+th,\ y=f(x^*)+t\,f_{h^+}(x^*),\ t\geq 0\}.$$

Vzhledem k definici směrové derivace a konvexnosti funkce f, jsou obě tyto množiny neprázdné, konvexní a disjunktní (viz Větu 6.4.10). Pak podle Věty 6.3.16 jsou tyto množiny oddělitelné, tj. existuje vektor $\binom{p}{\mu}\in\mathbb{R}^{n+1}\setminus\{0\}$ takový, že platí

$$\left\langle \binom{p}{\mu},\binom{x}{y}\right\rangle \geq \left\langle \binom{p}{\mu},\binom{x^*+th}{f(x^*)+t\,f_{h^+}(x^*)}\right\rangle$$

neboli

$$\left.\begin{aligned}&\langle p,x\rangle+\mu\,y\geq\\ &\geq\langle p,x^*+th\rangle+\\ &\quad+\mu[f(x^*)+t\,f_{h^+}(x^*)]\end{aligned}\right\}\ (\star)$$

pro všechna $t\geq 0$, $x\in X$ a $y>f(x)$. Jaké znaménko má μ? Kdyby $\mu<0$, pak bychom dostali spor s $(\star)$, neboť pravá strana je konstantní vzhledem k y, zatímco levá strana by v tomto případě mohla být libovolně blízko $-\infty$ pro dostatečně velké y. Pokud $\mu=0$, pak nutně $p\neq 0$ a $(\star)$ se redukuje na

$$\langle p,x\rangle\geq\langle p,x^*+th\rangle.$$

Avšak toto má platit pro každé $x\in X$, takže i pro $x=x^*+t\widetilde{h}\in\operatorname{int} X$, kde $\widetilde{h}\in\mathbb{R}^n$ je takový, že $(\widetilde{h}_i-h_i)\,p_i\leq 0$ pro všechna $i\in\{1,\dots,n\}$, přičemž alespoň jedna nerovnost je ostrá (což lze díky $p\neq 0$). Potom máme

$$\langle p,x^*+t\widetilde{h}\rangle\geq\langle p,x^*+th\rangle$$

neboli

$$\langle p,x^*+t\widetilde{h}-x^*-th\rangle\geq 0$$

neboli

$$t\langle p,\widetilde{h}-h\rangle\geq 0,$$

což je spor (díky volbě $\widetilde{h}$).

což dohromady s (6.5.12) dává

$$\langle a,h\rangle \le f_{h^+}(x^*) = \langle \operatorname{grad} f(x^*), h\rangle.$$

Proto pro každé $h \in \operatorname{Lin} X$ je tedy

$$\langle \operatorname{grad} f(x^*) - a, h\rangle \ge 0,$$

což obzvláště při volbě $h = a - \operatorname{grad} f(x^*) \in \mathbb{R}^n = \operatorname{Lin} X$ dává

$$\|\operatorname{grad} f(x^*) - a\|^2 \le 0,$$

a tedy nutně $a = \operatorname{grad} f(x^*)$.

(ii) Jestliže je $x^* \in \operatorname{int} X$, pak existuje $r > 0$ takové, že $x^* + B(0,r) \subset \operatorname{int} X$. Uvažme nyní funkci

$$\phi(t,h) := \frac{f(x^*+th) - f(x^*)}{t} - \langle a,h\rangle, \tag{6.5.15}$$

kde $t \in (0,1]$ a $h \in B(0,r)$. Jelikož $\partial f(x^*) = \{a\}$, plyne z (6.5.14), že $f_{h^+}(x^*) = \langle a,h\rangle$. Pak podle (důkazu) Věty 6.4.10 pro jakékoli $h \in B(0,r)$ je funkce $\phi(\cdot,h)$ nerostoucí na $(0,1]$ a konverguje k 0 pro $t \to 0^+$. Současně (vzhledem k Větě 6.4.1) je funkce $\phi(t,h)$ spojitá pro pevné $t \in (0,1]$ vzhledem k h z kompaktní množiny $B(0,r)$. To znamená, že konvergence $\phi(t,h) \to 0$ pro $t \to 0^+$ je stejnoměrná díky Diniho větě[155], tj. pro každé $\varepsilon > 0$ existuje $\delta \in (0,1]$ takové, že $0 \le \phi(t,h) \le \varepsilon$ pro všechna $t \in (0,\delta)$ a $h \in B(0,r)$. Zejména vezmeme-li libovolné $h \in B(0,\delta r)$, pak pro $\tilde t := \|h\|/r$ a $\tilde h := rh/\|h\|$ máme $\tilde t \in (0,\delta)$ a $\tilde h \in B(0,r)$, takže

$$0 \le \phi(\tilde t,\tilde h) \le \varepsilon \quad \text{neboli podle (6.5.15) máme} \quad 0 \le \frac{f(x^*+h)-f(x^*)}{\|h\|/r} - \Big\langle a, \tfrac{h}{\|h\|/r}\Big\rangle \le \varepsilon$$

neboli

$$0 \le \frac{f(x^*+h)-f(x^*)-\langle a,h\rangle}{\|h\|} \le \varepsilon/r.$$

Tudíž pro $\|h\| \to 0$ máme[156]

$$\lim_{\|h\|\to 0} \frac{f(x^*+h)-f(x^*)-\langle a,h\rangle}{\|h\|} = 0$$

což znamená, že funkce f je diferencovatelná v bodě x^* s diferenciálem $\mathrm{d}f_h(x^*) = \langle a,h\rangle$ neboli $a = \operatorname{grad} f(x^*)$. ■

Tedy nutně $\mu > 0$. Pak vydělením nerovnosti (⋆) tímto číslem dostaneme

$$\left.\begin{aligned}\left\langle \tfrac{1}{\mu}p, x\right\rangle + y \ge \left\langle \tfrac{1}{\mu}p, x^*+th\right\rangle + \\ + f(x^*) + t f_{h^+}(x^*)\end{aligned}\right\} \quad (\triangle)$$

pro všechna $t \ge 0$, $x \in X$ a $y > f(x)$. Volbou $t = 0$ a limitním přechodem $y \to f(x)^+$ obdržíme

$$\left\langle \tfrac{1}{\mu}p, x\right\rangle + f(x) \ge \left\langle \tfrac{1}{\mu}p, x^*\right\rangle + f(x^*)$$

neboli

$$f(x) - f(x^*) \ge \left\langle -\tfrac{1}{\mu}p, x - x^*\right\rangle$$

pro každé $x \in X$, tj. $-\frac{1}{\mu}p \in \partial f(x^*)$. Konečně volbou $x = x^*$, $t > 0$ a limitním přechodem $y \to f(x)^+$ dostaneme v (△), že

$$\begin{aligned}\left\langle \tfrac{1}{\mu}p, x^*\right\rangle + f(x^*) \ge \\ \ge \left\langle \tfrac{1}{\mu}p, x^*\right\rangle + t\left\langle \tfrac{1}{\mu}p, h\right\rangle + \\ + f(x^*) + t f_{h^+}(x^*)\end{aligned}$$

neboli

$$0 \ge t\left\langle \tfrac{1}{\mu}p, h\right\rangle + t f_{h^+}(x^*),$$

tj. $\left\langle -\frac{1}{\mu}p, h\right\rangle \ge f_{h^+}(x^*)$. Avšak z $-\frac{1}{\mu}p \in \partial f(x^*)$ plyne $f_{h^+}(x^*) \le \max_{a\in\partial f(x^*)}\langle a,h\rangle$, což společně s (6.5.12) dává rovnost (6.5.14). Tím je důkaz hotov i v tomto případě, a to bez použití Lemma 6.5.5.

[155] *Diniho věta*: Je-li $\{f_n\}$ nerostoucí posloupnost spojitých funkcí na kompaktní množině Y, tj. $f_n(x) \ge f_{n+1}(x)$ pro všechna $n \in \mathbb{N}$ a $x \in Y$, která konverguje bodově ke spojité funkci, pak je tato konvergence dokonce stejnoměrná.

[156] Vždyť pro libovolné $\varepsilon/r > 0$ existuje číslo $\delta\, r > 0$ takové, že pro všechna $h \in B(0,\delta\, r)$ platí

$$\left|\frac{f(x^*+h)-f(x^*)-\langle a,h\rangle}{\|h\|}\right| < \varepsilon/r.$$

Cvičení

6.5.1. Ukažte, že funkce f definovaná na intervalu $[0,\infty)$ předpisem

$$f(x) = \begin{cases} 1, & x = 0,\\ 0, & x > 0,\end{cases}$$

není subdiferencovatelná v bodě $x^* = 0$. Jak vypadá opěrná nadrovina v bodě $[0,1]$ k nadgrafu funkce f?

6.5.2. Ukažte, že funkce $f = -\sqrt{x}$ definovaná na intervalu $[0,\infty)$ není subdiferencovatelná v bodě $x^* = 0$. Jak vypadá opěrná nadrovina v bodě $[0,0]$ k nadgrafu funkce f?

6.5.3. Pro funkci $f(x) = \max\{0, \frac{x^2-1}{2}\}$ určete $\partial f(x)$ pro všechna $x \in \mathbb{R}$.
řešení: $\partial f(x) = x$ pro $|x| > 1$, $\partial f(1) = [0,1]$, $\partial f(-1) = [-1,0]$ a $\partial f(x) = 0$ pro $x \in (-1,1)$

6.5.4. Nechť

$$f(x_1,x_2) = \sqrt{x_1^2 + 2x_2^2}.$$

Určete subdiferenciál $\partial f(0,0)$ a tuto množinu znázorněte v rovině.
řešení: $\partial f(0,0) = \{(a_1,a_2)^\top \in \mathbb{R}^2 \mid a_1^2 + a_2^2/2 \le 1\}$

6.5.5. Nechť
$$f(x_1,x_2)=\sqrt{x_1^2+2x_1\,x_2+2x_2^2}.$$
Určete subdiferenciál $\partial f(0,0)$ a tuto množinu znázorněte v rovině.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid 2a_1^2-2a_1\,a_2+a_2^2\le 1\}$

6.5.6. Nechť
$$f(x_1,x_2)=\sqrt{2x_1^2+2x_1\,x_2+3x_2^2}.$$
Určete subdiferenciál $\partial f(0,0)$, tuto množinu znázorněte v rovině a rozhodněte, zda bod $[1,1]$ je vnitřním, hraničním nebo vnějším bodem množiny $\partial f(0,0)$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid 3a_1^2+2a_2^2-2a_1\,a_2\le 5\}$

6.5.7. Nechť
$$f(x_1,x_2)=\sqrt{x_1^2-2x_1\,x_2+x_2^2}.$$
Určete subdiferenciál funkce f v bodě $[0,0]$. Rozhodněte, zda $z=2\,(x_1+x_2)$ je opěrnou nadrovinou k nadgrafu funkce f v bodě $[0,0]$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid a_1\in[-1,1],\ a_1=-a_2\}$ & ne, není opěrnou nadrovinou

6.5.8. Nechť
$$f(x_1,x_2)=\sqrt{3x_1^2-2x_1\,x_2+3x_2^2}.$$
Určete subdiferenciál funkce f v bodě $[0,0]$. Rozhodněte, zda $z=x_1+x_2$ je opěrnou nadrovinou k nadgrafu funkce f v bodě $[0,0]$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid 3a_1^2+2a_1\,a_2+3a_2^2\le 8\}$ & ano, je opěrnou nadrovinou

6.5.9. Nechť
$$f(x_1,x_2)=\sqrt{x_1^2-2x_1\,x_2+2x_2^2}.$$
Určete subdiferenciál funkce f v bodě $[0,0]$. Rozhodněte, zda $z=x_1-x_2$ je opěrnou nadrovinou k nadgrafu funkce f v bodě $[0,0]$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid 2a_1^2+2a_1\,a_2+a_2^2\le 1\}$ & ano, je opěrnou nadrovinou

6.5.10. Nechť
$$f(x_1,x_2)=\sqrt{3x_1^2+2x_1\,x_2+x_2^2}+1.$$
Určete subdiferenciál funkce f v bodě $[0,0]$. Rozhodněte, zda $z=x_1+x_2/2+1$ je opěrnou nadrovinou k nadgrafu funkce f v bodě $[0,0]$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid a_1^2-2a_1\,a_2+3a_2^2\le 2\}$ & ano, je opěrnou nadrovinou

6.5.11. Přímým výpočtem ověřte, že pro funkci
$$f(x_1,x_2)=\sqrt{3x_1^2+2x_1\,x_2+x_2^2+1}.$$
je $\partial f(0,0)=\{\operatorname{grad} f(0,0)\}=\{(0,0)^\top\}$, nikoli pouze $(0,0)^\top\in\partial f(0,0)$.

6.5.12. Nechť
$$f(x_1,x_2)=\sqrt{3x_1^2+2x_1\,x_2+x_2^2+x_1}.$$
Určete subdiferenciál funkce f v bodě $[0,0]$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid a_1\in[0,\sqrt{2}]\ \&\ a_2=0\}$

6.5.13. Nechť
$$f(x_1,x_2)=\sqrt{x_1^2+x_1\,x_2+x_2^2-3}.$$
Určete subdiferenciál funkce f v bodě $[1,1]$.
řešení: $\partial f(1,1)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid a_1=a_2\in[0,\sqrt{3}/2]\}$

6.5.14. Určete subdiferenciál $\partial f(0,0)$ pro funkci $f(x_1,x_2)=|x_1|+|x_2|$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid\max\{|a_1|,|a_2|\}\le 1\}$

6.5.15. Určete subdiferenciály $\partial f(0,0)$, $\partial f(0,1)$ $\partial f(1,1)$ pro funkci $f(x_1,x_2)=\max\{|x_1|,|x_2|\}$.
řešení: $\partial f(0,0)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid |a_1|+|a_2|\le 1\}$, $\partial f(0,1)=\{(0,1)^\top\}$,
$\partial f(1,1)=\{(a_1,a_2)^\top\in\mathbb{R}^2\mid a_1+a_2=1,a_1\ge 0,a_2\ge 0\}$

6.5.16. Určete subdiferenciál $\partial f(0)$ pro funkci $f(x) = \|x\| := \sqrt{\langle x, x\rangle}$.

řešení: $\partial f(0) = \{a \in \mathbb{R}^n \mid \|a\| \leq 1\}$

6.5.17. Nechť $A = \{[x_1, x_2] \in \mathbb{R}^2 \mid x_2 \leq -|x_1|\}$ a uvažme funkci $f(x_1, x_2) = \rho([x_1, x_2], A)$ určující vzdálenost bodu $[x_1, x_2] \in \mathbb{R}^2$ od množiny A. Určete subdiferenciál $\partial f(x_1, x_2)$ pro v libovolném bodě $[x_1, x_2] \in \mathbb{R}^2$.

řešení: $\partial f(x, y) = \operatorname{grad} f(x, y)$ pro $[x, y] \neq [0, 0]$, $\partial f(0, 0) = \{(a, 1/\sqrt{2})^\top \in \mathbb{R}^2 \mid a \in [-1/\sqrt{2}, 1/\sqrt{2}]\}$

6.5.18. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ a $f : X \to \mathbb{R}$. Pak pro $x^* \in X$ platí

$$f(x^*) = \min_{x \in X} f(x)$$

právě tehdy, když $0 \in \partial f(x^*)$.

6.5.19. *Dokažte:* Nechť $f_i : \mathbb{R}^n \to \mathbb{R}$ jsou konvexní funkce pro $i = 1, \ldots, m$. Pak pro každé $x \in \mathbb{R}^n$ platí

$$\partial f(x) = \partial f_1(x) + \cdots + \partial f_m(x).$$

6.5.20. *Dokažte:* Nechť $f : \mathbb{R}^m \to \mathbb{R}$ je konvexní funkce a $A \in \mathbb{R}^{m \times n}$. Položme $F(x) := f(Ax)$. Potom platí

$$\partial F(x) = A^\top \partial f(Ax) = \{z \in \mathbb{R}^n \mid z = A^\top a,\ a \in \partial f(Ax)\}.$$

6.5.21. *Dokažte:* Pro konvexní funkci $f : \mathbb{R} \to \mathbb{R}$ platí

$$\partial f(x) = [f'_-(x), f'_+(x)].$$

6.5.22. *Dokažte:* Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, $\operatorname{int} X \neq \emptyset$ a $f : X \to \mathbb{R}$.

a) Je-li f konvexní a $x^* \in \operatorname{int} X$, pak $\partial f(x^*)$ je neprázdná, uzavřená, ohraničená a konvexní množina.

b) Je-li $\partial f(x)$ neprázdná pro každé $x \in X$, pak f je konvexní na X.

6.5.23. *Dokažte:* Nechť $f : X \to \mathbb{R}$ je konvexní funkce na konvexní množině $X \subseteq \mathbb{R}^n$. Pak pro každé $x^* \in \operatorname{ri} X$ platí:

a) $f'_{\lambda h}(x^*) = \lambda f'_h(x^*)$ pro každé $\lambda \geq 0$ a $h \in \operatorname{Lin} X$;

b) funkce $f'_\cdot(x^*)$ je konvexní na $\operatorname{Lin} X$.

6.6 Fenchelova transformace

Na závěr této kapitoly se podíváme na transformaci, která dané funkci $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ přiřadí konvexní funkci $f^* : \mathbb{R}^n \to \mathbb{R}$. K čemu je to dobré? Již jsme si mohli všimnout, že řešení minimalizačních úloh má ve spojení s konvexními funkcemi značné výhody. Toto pozorování nás při studiu úloh matematického programování přivede k tzv. *duální úloze*, jenž je přirozeně přidružena původní úloze a jejíž studium může být užitečnější/výhodnější/snadnější než v případě původní úlohy[157]. Takovýmto úlohám se budeme podrobněji věnovat v jedné části Kapitoly 8 a v jejím závěru si naznačíme roli Fenchelovy transformace v této oblasti.

[157] Ještě aby ne, když si veškerou dřinu odpracujeme při jejím hledání – leda bychom uměli duální úlohu určit nějak snadno jako v případě lineárního či kvadratického programování.

Uvažme nyní libovolnou funkci $f : \mathbb{R}^n \to \mathbb{R}$ a množinu všech afinních funkcí h, pro které je f majorantou (neboli h minorizují f), tj.

$$h(x) \leq f(x)$$

pro všechna $x \in \mathbb{R}^n$. Jelikož funkce $h(x)$ jsou tvaru $h(x) = \langle y, x\rangle - \alpha$ pro nějaké $y \in \mathbb{R}^n$ a $\alpha \in \mathbb{R}$, požadujeme

$$\langle y, x\rangle - \alpha \leq f(x) \quad \text{pro všechna } x \in \mathbb{R}^n$$

neboli

$$\alpha \geq \langle y, x\rangle - f(x) \quad \text{pro všechna } x \in \mathbb{R}^n,$$

takže při pevně zvoleném $y \in \mathbb{R}^n$ musíme nutně brát

$$\alpha = \sup\{\langle y, x\rangle - f(x) \mid x \in \mathbb{R}^n\},$$

což nás přivádí k následující definici. Možná není na první pohled patrný smysl našich úvah – mnohé se ale objasní, až dorazíme k Definici 6.6.15 a Větě 6.6.16.

DEFINICE 6.6.1

Nechť je dána funkce $f : \mathbb{R}^n \to \mathbb{R}$. Pak funkce

$$f^*(y) := \sup_{x\in\mathbb{R}^n}\{\langle x, y\rangle - f(x)\} \tag{6.6.1}$$

se nazývá *Fenchelovou transformací*[158] (též */konvexně/ konjugovanou funkcí*) funkce f.

[158] Tento pojem a jeho další rozvoj je spojen především s Fenchelem, Legendrem, Moreauem a Rockafellarem, viz [169, 384, 432] a poznámky 118–119 na str. 58–59.

Poznámka 6.6.2. (i) Je zřejmé, že $f^* : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$, tj. je možné, že $f^*(y) = \infty$ pro některé y. Proto je vhodné si označit i tzv. *efektivní definiční obor* $D^*(f^*) := \{y \in \mathbb{R}^n \mid f^*(y) < \infty\}$. Nicméně je-li funkce f konvexní, pak musí být $f^*(y) < \infty$.

(ii) Fenchelovu transformaci jsme zavedli pro funkce definované na celém $\mathbb{R}^n$, avšak toto není nezbytné. Jestliže $D(f) = X \subsetneq \mathbb{R}^n$, pak můžeme upravit (6.6.1) tak, že budeme brát supremum pouze přes $x \in X$. Druhou (ekvivalentní) možností je dodefinování funkce f na $\mathbb{R}^n \setminus X$ jako $f(x) = \infty$, takže pro ni je hodnota suprema přes $\mathbb{R}^n$ totožná[159] s hodnotou suprema pouze přes X.

[159] V bodech $\mathbb{R}^n \setminus X$ pak je totiž $x^\top y - f(x) = -\infty$, což nemůže být shodné se supremem funkce $x^\top y - f(x)$ na X vyjma situace, kdy $f(x) \equiv \infty$ na X. To jsme si ale na začátku Podkapitoly 6.2 „zakázali".

(iii) Uvažme diferencovatelnou funkci f a y takové, že $f^*(y) < \infty$. Pak nutná podmínka pro extrém je

$$\frac{\partial}{\partial x_i}\big[\langle x, y\rangle - f(x)\big] = 0 \quad \text{pro všechna } i \in \{1, \dots, n\},$$

tedy musí platit $y = \operatorname{grad} f(x)$. Je-li navíc stacionární bod x^* funkce $\langle x, y\rangle - f(x)$ zároveň i bodem maxima této funkce[160], pak platí

$$f^*(y) = \langle x^*, y\rangle - f(x^*),$$

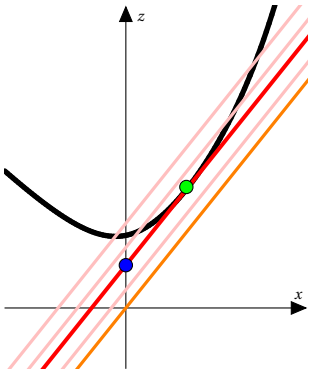
přičemž x^* je takové, že $y = \operatorname{grad} f(x^*)$. V takovém případě je Fenchelova transformace známa jako *Legendreova transformace*.

[160] Kdy se to může stát?

(iv) Z Definice 6.6.1 očividně plyne

$$f^*(0) = \sup_{x\in X}\{-f(x)\} = -\inf_{x\in X} f(x).$$

(v) Jaký je geometrický význam hodnoty $f^*(y)$? Zvolme pevně y. Rozdíl $f(x) - x^\top y$ udává „svislou" vzdálenost grafu funkce f od nadroviny $z = x^\top y$. Budeme-li posouvat tuto nadrovinu směrem k nadgrafu funkce f, bude se tato vzdálenost zmenšovat. Nejmenší hodnota by pak měla nastat v takovém $\widehat{x}$, ve kterém by se z této nadroviny rovnoběžné s $z = x^\top y$ stala tečna. My ale nepožadujeme diferencovatelnost, takže nemáme existenci tečné nadroviny vůbec zaručenu. To naštěstí nevadí, protože nám stačí, aby tato rovnoběžná nadrovina byla opěrnou nadrovinou k epi f v bodě $[\widehat{x}, f(\widehat{x})]$. Ta pak bude mít rovnici $-y^\top x + z + y^\top \widehat{x} - f(\widehat{x}) = 0$, která pro $x = 0$ dává $z = f(x) - y^\top x$, tj. průsečík opěrné nadroviny s normálovým vektorem $(-y, 1)^\top$ k epi f se svislou osou má hodnotu $-f^*(y)$, viz Obrázek 6.79. ▲



Obrázek 6.79: Geometrická ilustrace $f^*(y)$. Na obrázku je přímka $z = y\,x$ (oranžovou barvou), několik rovnoběžek (růžovou barvou) a rovnoběžka (červenou barvou) odpovídající tečně ke grafu funkce f v zeleném bodě. Tato rovnoběžka protíná svislou osu v modrém bodě s hodnotou odpovídající $-f^*(y)$.

Příklad 6.6.3

Ačkoli se Fenchelova transformace může zdát jako čistě matematický objekt, lze snadno nabídnout jednoduchou ekonomickou interpretaci hodnoty $f^*(y)$. Uvažme výrobní podnik, který k výrobě svého produktu spotřebovává n zdrojů. Vektor spotře-

bovaných zdrojů označme $x = (x_1, \dots, x_n)^\top$. Nechť $f(x)$ značí celkový výnos z prodeje výrobků získaných ve výrobním procesu pomocí x-tice zdrojů. Je-li vektor jednotkových cen zdrojů roven $y = (y_1, \dots, y_n)^\top$, pak celkové náklady jsou $\langle y, x \rangle$ a zisk tohoto podniku je $f(x) - \langle y, x \rangle$. Zafixujme nyní ceny zdrojů (tj. žádné množstevní slevy, změny dodavatelů atd.). Jakého maximálního zisku lze dosáhnout vhodnou volbou množství spotřebovaných zdrojů? Odpověď na tuto otázku není nic jiného, než nalezení

$$M(y) := \sup_x \{f(x) - \langle y, x \rangle\}.$$

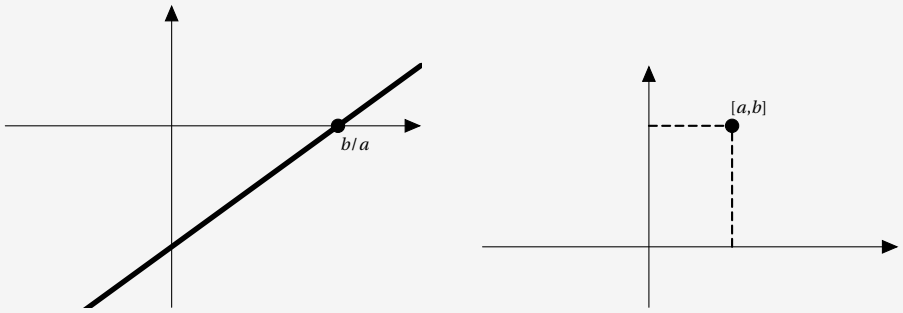
Tato funkce udává maximální zisk dosažitelný jako funkce ceny vstupních zdrojů. S využitím Fenchelovy transformace funkce $-f$ dostáváme

$$M(y) = (-f)^*(-y),$$

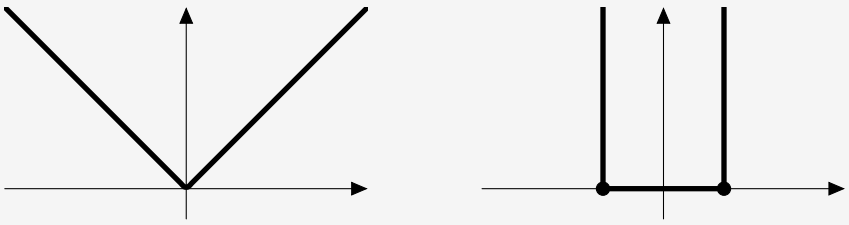
což ukazuje těsný vztah mezi maximálním ziskem a Fenchelovou transformací výnosové funkce (jakožto funkce spotřebovaných zdrojů).

Příklad 6.6.4

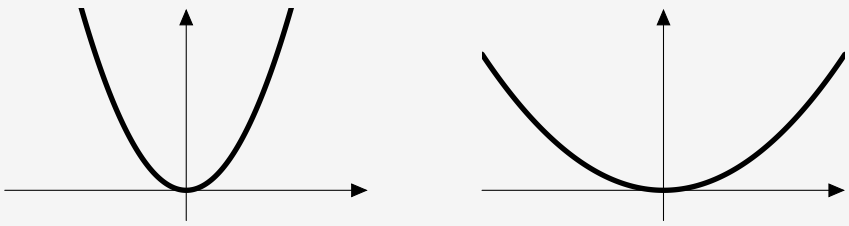
Bez dalších podrobností si zde můžeme uvést několik jednoduchých funkcí f a odpovídajících f^*.



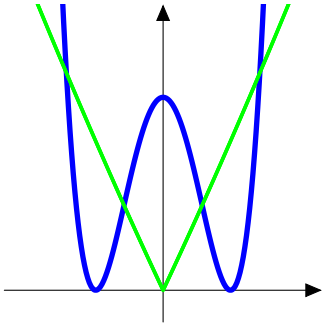
Obrázek 6.80: Funkce $f(x) = ax - b$ (vlevo) a konjugovaná funkce $f*$ s hodnotami $f^*(a) = b$ a $f^*(y) = \infty$ pro $y \neq a$ (vpravo).



Obrázek 6.81: Funkce $f(x) = |x|$ (vlevo) a konjugovaná funkce f^* s hodnotami $f^*(y) = 0$ pro $y \in [-1,1]$ a $f^*(y) = \infty$ pro $y \in (-\infty,-1) \cup (1,\infty)$ (vpravo).



Obrázek 6.82: Funkce $f(x) = \frac{cx^2}{2}$ (vlevo) a konjugovaná funkce $f^*(y) = \frac{y^2}{2c}$ (vpravo).



Obrázek 6.83: Graf funkce $f(x) = \frac{3}{10}x^2(x-3)(x+3) + \frac{243}{40}$ (modrou barvou) a odpovídající f^* (zelenou barvou). Napadá vás, že grafu f^* vypadá podobně jako absolutní hodnota? Možná nejste daleko od pravdy – i když jen s částí tohoto grafu (viz následující poznámku č. 161 a Obrázky 6.81 a 6.84).

[161] Při hledání $\sup xy - f(x)$ narazíme na tři stacionární body. Dva z nich jsou lokální maxima (jeden je dokonce globálním maximem) a třetí je inflexním bodem. Jenže tyto stacionární body jsou kořeny kubického polynomu, z nichž dva máme vyjádřené pomocí komplexních čísel (i když jsou reálné). Konkrétně to jsou tyto body

$$x_1 = \frac{\left(90y+18\sqrt{25y^2-486}\right)^{\frac{2}{3}}+54}{6\left(90y+18\sqrt{25y^2-486}\right)^{\frac{1}{3}}}$$

$$x_2 = \frac{\left(\sqrt{3}\mathrm{i}-1\right)\left(90y+18\sqrt{25y^2-486}\right)^{\frac{2}{3}}-54\mathrm{i}\sqrt{3}-54}{12\left(90y+18\sqrt{25y^2-486}\right)^{\frac{1}{3}}}$$

$$x_3 = -\frac{\left(1+\sqrt{3}\mathrm{i}\right)\left(90y+18\sqrt{25y^2-486}\right)^{\frac{2}{3}}-54\mathrm{i}\sqrt{3}+54}{12\left(90y+18\sqrt{25y^2-486}\right)^{\frac{1}{3}}}.$$

Maximum se pro $y \geq 0$ realizuje v x_1 a pro $y \leq 0$ v x_2. Předpis funkce $f^*(y)$ získaný dosazením těchto bodů do výrazu $xy - f(x)$ je poměrně komplikovaný, ale pro nás v tuto chvíli naštěstí i nepotřebný.

Příklad 6.6.5

Uvažme $f(x) = \mathrm{e}^x$. Pak funkce $xy - \mathrm{e}^x$ je konkávní, takže pro určení suprema/maxima stačí nalézt stacionární bod, tj. vyřešit rovnici $0 = (xy - \mathrm{e}^x)' = y - \mathrm{e}^x$, kde $y \in \mathbb{R}$ považujeme za parametr. Její řešení existuje pouze pro $y > 0$ a je odpovídá $x = \ln y$, což po zpětném dosazení dává $f^*(y) = y\ln y - \mathrm{e}^{\ln y} = y(\ln y - 1)$. Je-li $y = 0$, pak $f^*(0) = \sup_{x\in\mathbb{R}}\{0 - \mathrm{e}^x\} = 0$ pro $x \to -\infty$ a pro $y < 0$ je $f^*(y) = \sup_{x\in\mathbb{R}}\{xy - \mathrm{e}^x\} = \infty$ pro $x \to -\infty$. Celkem tedy máme

$$f^*(y) = \begin{cases} y(\ln y - 1), & y > 0, \\ 0, & y = 0, \\ \infty, & y < 0. \end{cases}$$

Při určování tzv. *bikonjugované funkce* $f^{**}(x)$, se stačí omezit na $y \in \mathbb{R}$, pro které $f^*(y) < \infty$, tj. $y \in D^*(f^*)$. Pak

$$\begin{aligned} f^{**}(x) &= \sup_{y\in\mathbb{R}}\{yx - f^*(y)\} = \sup_{y\in[0,\infty)}\{yx - f^*(y)\} = \\ &= \max\Big\{\sup_{y\in(0,\infty)}\{yx - f^*(y)\}, \sup_{y=0}\{yx - f^*(y)\}\Big\} = \\ &= \max\Big\{\sup_{y\in(0,\infty)}\{yx - y(\ln y - 1)\}, \sup_{y=0}\{0\}\Big\}. \end{aligned}$$

Funkce $yx - y(\ln y - 1)$ je pro libovolné $x \in \mathbb{R}$ a $y > 0$ konkávní, takže stačí najít řešení rovnice $0 = (yx - y(\ln y - 1)' = x - \ln y + 1 - 1 = x - \ln y$, což je $y = \mathrm{e}^x$. Proto

$$f^{**}(x) = \max\left\{x\,\mathrm{e}^x - \mathrm{e}^x(x-1),\ 0\right\} = \mathrm{e}^x = f(x).$$

V předchozím příkladu platilo $f = f^{**}$. Náhoda? Uvidíme… Nicméně už dopředu můžeme říct, že ne vždy to takto dopadne. Např. na Obrázku 6.83 vidíme funkci

$$f(x) = \frac{3}{10}x^2(x-3)(x+3) + \frac{243}{40}$$

a odpovídající f^*, jejíž výpočet je už poměrně komplikovaný[161] vzhledem k nekonvexnosti f. Na dalším Obrázku 6.84 je k těmto funkcím přikreslena i bikonjugovaná funkce f^{**}. Geometrickou interpretaci hledání f^* jsme si již objasnili, jak to ale rozšířit na f^{**}? Pro libovolné $x \in \mathbb{R}^n$ si představme přímku v $\mathbb{R}^{n+1}$ procházející bodem $[x, 0]$ a rovnoběžnou se svislou osou. Současně pro každé $y \in D^*(f^*)$ sestrojme opěrnou nadrovinu k epi f s normálovým vektorem $(-y, 1)^\top$. Již víme, že tato nadrovina prochází bodem $[0, -f^*(y)]$. Průsečík této nadroviny se svislou přímkou bude zjevně v bodě $[x, y^\top x - f^*(y)]$, takže nejvýše položený průsečík pro všechna povolená y leží na grafu f^{**}, tj. $(n+1)$-ní souřadnice těchto bodů odpovídají funkčním hodnotám f^{**}, viz Obrázky 6.85–6.86. Pro úplnost zdůrazněme, že tuto interpretaci nelze použít v případě, že $f^{**}(x) = -\infty$ pro nějaké x.

Nejdříve si ale ukážeme některé základní vlastnosti Fenchelovy transformace.

Lemma 6.6.6

Nechť je dána funkce $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ a f^* je její Fenchelova transformace. Pak následující tvrzení jsou pravdivá.

(i) Funkce f^* je konvexní na množině $Y := D^*(f^*) = \{y \in \mathbb{R}^n \mid f^*(y) < \infty\}$.

(ii) Pro každé $x \in X$ a $y \in \mathbb{R}^n$ platí tzv. *Fenchelova(–Youngova) nerovnost*

$$f(x) + f^*(y) \geq \langle x, y\rangle,$$

přičemž rovnost nastane právě tehdy, když $y \in \partial f(x)$.

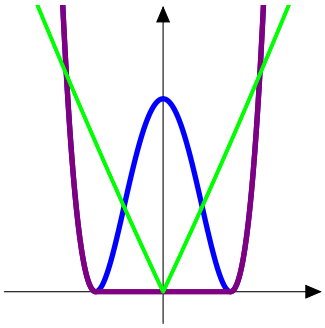
(iii) Je-li $f(x) \geq g(x)$ na X, pak $f^*(y) \leq g^*(y)$ pro všechna $y \in \mathbb{R}^n$.

Poznámka 6.6.7. Aplikováním Fenchelovy nerovnosti na funkci $f(x) = \frac{1}{\alpha}\sum_{i=1}^n |x_i|^\alpha$ pro libovolné $\alpha \in \mathbb{R}\setminus\{0\}$ obdržíme

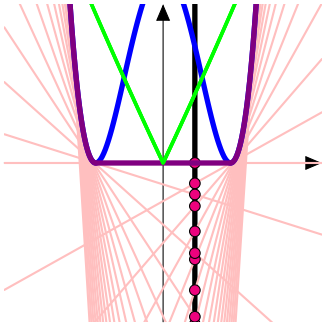
$$\langle x, y\rangle \le \frac{\sum_{i=1}^n |x_i|^\alpha}{\alpha} + \frac{\sum_{i=1}^n |y_i|^\beta}{\beta},$$

kde $\beta \in \mathbb{R}$ je takové, že $1/\alpha + 1/\beta = 1$. Zejména v případě $n = 1$ dostaneme tzv. *Youngovu nerovnost*

$$xy \le \frac{|x|^\alpha}{\alpha} + \frac{|x|^\beta}{\beta}.$$ ▲



Obrázek 6.84: Graf funkce $f(x) = \frac{3}{10}x^2(x-3)(x+3) + \frac{243}{40}$ (modrou barvou), odpovídající f^* (zelenou barvou) a také f^{**} (fialovou barvou).



Obrázek 6.85: Graf funkce $f(x) = \frac{3}{10}x^2(x-3)(x+3) + \frac{243}{40}$ (modrou barvou), odpovídající f^* (zelenou barvou) a také f^{**} (fialovou barvou). Opěrné nadroviny k epi f (ty jsou současně tečnami, ale pouze v částech napravo a nalevo od nulových bodů – mezi nulovými body nejsou tečny opěrnými nadrovinami) a jejich průsečíky s přímkou $x = 1$. Nejvýše z těchto průsečíků leží bod $[1, 0]$, tj. $f^{**}(1) = 0$.

Důkaz Lemma 6.6.6.

(i) Funkce $F_x(y) := \langle x, y\rangle - f(x)$ je afinní vzhledem k y, a tedy také konvexní. Pak podle tvrzení ve Cvičení 6.2.14[162] s celým X jako indexovou množinu je funkce $f^*(y)$ konvexní[163] na Y.

(ii) Toto plyne přímo z Definice 6.6.1, neboť nutně

$$f^*(y) \ge \langle x, y\rangle - f(x),$$

tj. $f^*(y) + f(x) \ge \langle x, y\rangle$. Je-li navíc $y \in \partial f(x)$, pak podle definice máme

$$f(z) - f(x) \ge \langle y, z - x\rangle$$

pro všechna $z \in X$ neboli

$$\langle y, z\rangle - f(z) \le \langle y, x\rangle - f(x),$$

což po aplikování suprema přes $z \in X$ na obě strany nerovnosti dává

$$f^*(y) \le \langle y, x\rangle - f(x).$$

Toto dohromady s předchozí částí vede k rovnosti $f^*(y) + f(x) = \langle y, x\rangle$.

Tento postup lze ale také obrátit. Jestliže tedy pro $x \in X$ a $y \in \mathbb{R}^n$ platí $f^*(y) + f(x) = \langle y, x\rangle$, pak nutně

$$f^*(y) + f(x) \le \langle y, x\rangle$$

neboli

$$\sup_{z\in X}\{\langle z, y\rangle - f(z)\} \le \langle y, x\rangle - f(x).$$

To znamená, že

$$\langle z, y\rangle - f(z) \le \langle y, x\rangle - f(x)$$

pro všechna $z \in X$ neboli

$$\langle y, z - x\rangle \le f(z) - f(x),$$

z čehož vyplývá, že $y \in \partial f(x)$.

(iii) Zjevně platí

$$\langle x, y\rangle - f(x) \le \langle x, y\rangle - g(x),$$

což po aplikování suprema přes $x \in X$ na obě strany dává $f^*(y) \le g^*(y)$. ■

Z Lemma 6.6.6(ii) navíc vyplývají následující důsledky.

Důsledek 6.6.8

Nechť je dána funkce $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ a f^* je její Fenchelova transformace.

(i) Je-li $\partial f(x) \ne \emptyset$, pak $f(x) = f^{**}(x)$ pro každé $x \in X$. Naopak, je-li $f(x) = f^{**}(x)$ pro $x \in X$, pak $\partial f(x) = \partial f^{**}(x)$.

(ii) Je-li $y \in \partial f(x)$, pak $x \in \partial f^*(y)$. Je-li navíc $f(x) = f^{**}(x)$, pak $y \in \partial f(x)$ právě tehdy, když[164] $x \in \partial f^*(y)$.

Důkaz. (i) Pro každé $x, y \in \mathbb{R}^n$ je $f^*(y) + f^{**}(x) \geq \langle x, y\rangle$. Je-li $y \in \partial f(x)$, pak $f^*(y) + f(x) = \langle x, y\rangle$, z čehož plyne $f^{**}(x) \geq f(x)$. Opačná nerovnost platí vždy, neboť pro všechna $x, y \in \mathbb{R}^n$ máme $f(x) \geq \langle y, x\rangle - f^*(y)$, a tudíž $f(x) \geq \sup_{y\in\mathbb{R}^n} \{\langle y, x\rangle - f^*(y)\} = f^{**}(x)$.

Naopak, je-li $f(x) = f^{**}(x)$ pro nějaké $x \in X$, pak z rovnosti $f^*(x) = f^{***}(x)$ (viz Poznámku 6.6.14(ii) později) máme $y \in \partial f(x)$ právě tehdy, když

$$\langle y, x\rangle = f(x) + f^*(y) = f^{**}(x) + f^{***}(y),$$

což je ekvivalentní s $y \in \partial f^{**}(x)$.

(ii) Podmínka $y \in \partial f(x)$ je ekvivalentní s $\langle y, x\rangle = f(x) + f^*(y)$, a tudíž nerovnost $f(x) \geq f^{**}(x)$ implikuje $\langle y, x\rangle \geq f^{**}(x) + f^*(y)$, což je ale ekvivalentní s $x \in \partial f^*(y)$. Platí-li navíc $f(x) = f^{**}(x)$, pak podmínka $y \in \partial f(x)$ je dokonce ekvivalentní s $\langle y, x\rangle \geq f^{**}(x) + f^*(y)$, tj. $x \in \partial f^*(y)$. ■

Následující tvrzení ukazuje chování Fenchelovy transformace při skládání s jednoduchými modifikacemi funkce f.

LEMMA 6.6.9

Nechť je dána funkce $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ a f^* je její Fenchelova transformace. Pak následující tvrzení jsou pravdivá.

(i) Je-li $a \in \mathbb{R}\setminus\{0\}$ a

$$\widetilde{f}(x) := \begin{cases} f(ax), & ax \in X, \\ \infty, & \text{jinak,} \end{cases}$$

potom platí

$$\widetilde{f}^*(y) = f^*(y/a).$$

(ii) Je-li $b \in \mathbb{R}^n$ a

$$\widetilde{f}(x) := \begin{cases} f(x+b), & x+b \in X, \\ \infty, & \text{jinak,} \end{cases}$$

potom platí

$$\widetilde{f}^*(y) = f^*(y) - \langle b, y\rangle.$$

(iii) Pro každé $\alpha > 0$ platí

$$(\alpha f)^*(y) = \alpha f^*(y/\alpha).$$

(iv) Pro každé $\gamma \in \mathbb{R}$ platí

$$(f+\gamma)^*(y) = f^*(y) - \gamma.$$

(v) Je-li $g(x) := \langle a, x\rangle$ pro $a \in \mathbb{R}^n$, potom

$$(f+g)^*(y) = f^*(y-a).$$

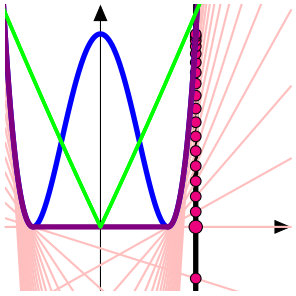
Důkaz. Jednotlivá tvrzení plynou přímo z Definice 6.6.1 a z vlastností suprema, přičemž pro argument mimo X bereme hodnoty funkce f jako ∞.

(i) Platí

$$\widetilde{f}^*(y) = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - \widetilde{f}(x)\} = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - f(ax)\} = \sup_{x\in\mathbb{R}^n} \{\langle ax, y/a\rangle - f(ax)\} = f^*(y/a).$$

(ii) Platí

$$\begin{aligned} \widetilde{f}^*(y) &= \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - \widetilde{f}(x)\} = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - f(x+b)\} = \\ &= \sup_{x\in\mathbb{R}^n} \{\langle x+b, y\rangle - f(x+b) - \langle b, y\rangle\} = \\ &= \sup_{x\in\mathbb{R}^n} \{\langle x+b, y\rangle - f(x+b)\} - \langle b, y\rangle = f^*(y) - \langle b, y\rangle. \end{aligned}$$



Obrázek 6.86: Graf funkce $f(x) = \frac{3}{10}x^2(x-3)(x+3) + \frac{243}{40}$ (modrou barvou), odpovídající f^* (zelenou barvou) a také f^{**} (fialovou barvou). Opěrné nadroviny k epi f (ty jsou současně tečnami, ale pouze v částech napravo a nalevo od nulových bodů – mezi nulovými body nejsou tečny opěrnými nadrovinami) a jejich průsečíky s přímkou $x = 3$. Nejvýše z těchto průsečíků leží bod $[3, f(3)]$, tj. $f^{**}(3) = f(3) = \frac{243}{40}$.

[162] Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, I je libovolná indexová množina, funkce $f_i : X \to \mathbb{R}$ jsou konvexní pro všechna $i \in I$ a pro každé $x \in X$ je množina $\{f_i(x) \mid i \in I\}$ shora ohraničená, tj. existuje funkce $M : X \to \mathbb{R}$ taková, že pro každé $i \in I$ platí $f(x) \leq M(x)$ pro každé $x \in X$. Potom funkce

$$f(x) := \sup_{i\in I} f_i(x)$$

je konvexní na X.

[163] Potřebujeme, aby funkce $F_x(y)$ byly shora ohraničení na Y, a proto požadujeme $f^*(y) < \infty$.

[164] Odtud „můžeme“ získat jiný návod, jak určit graf funkce f^*: sestrojíme graf ∂f, „obrátíme“ jej podle přímky $y = x$ a zpátky „zintegrujeme“, přičemž máme na mysli, že $f^*(0) = -\inf f$. Zkuste si to třeba na $|(|x)$ nebo e^x...

(iii) Platí

$$(\alpha f)^*(y) = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - \alpha f(x)\} = \alpha \{\langle x, y/\alpha\rangle - f(x)\} = \alpha f^*(y/\alpha).$$

(iv) Platí

$$(f + c)^*(y) = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - (f + c)(x)\} = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - f(x)\} - c = f^*(y) - c.$$

(v) Platí

$$(f + g)^*(y) = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - f(x) - \langle a, x\rangle\} = \sup_{x\in\mathbb{R}^n} \{\langle x, y - a\rangle - f(x)\} = f^*(y - a).$$

■

Tyto vlastnosti společně s některými dalšími příklady Fenchelovy transformace jsou shrnuty v Tabulce 6.1 níže.

$f(x)$	$D(f)$	$f^*(y)$	$D^*(f^*)$
$f(ax)$, $a \in \mathbb{R}\setminus\{0\}$	X	$f^*(y/a)$	X^*
$f(x + b)$, $b \in \mathbb{R}^n$	X	$f^*(y) - \langle b, y\rangle$	X^*
$(\alpha f)(x)$, $\alpha > 0$	X	$\alpha f^*(y/\alpha)$	X^*
$f(x) + \gamma$	X	$f^*(y) - \gamma$	X^*
$f(x) + \langle c, x\rangle$, $c \in \mathbb{R}^n$	X	$f^*(y - c)$	X^*
$\alpha f(ax + b) + \langle c, x\rangle + \gamma$ $\alpha > 0$, $a \neq 0$	X	$\alpha f^*\left(\frac{y-c}{\alpha a}\right) - \frac{\langle b, y-c\rangle}{a} - \gamma$	X^*
0	$\mathbb{R}$	0	$\{0\}$
0	$\mathbb{R}_+$	0	$-\mathbb{R}_+$
0	$[-1, 1]$	$\lvert y\rvert$	$\mathbb{R}$
$\lvert x\rvert^p/p$, $p > 1$	$\mathbb{R}$	$\lvert y\rvert^q/q$, $1/p + 1/q = 1$	$\mathbb{R}$
$-x^p/p$, $p \in (0, 1)$	$\mathbb{R}_+$	$-(-y)^q/q$, $1/p + 1/q = 1$	$-\mathbb{R}_{++}$
$\sqrt{1 + x^2}$	$\mathbb{R}$	$-\sqrt{1 - y^2}$	$[-1, 1]$
$-\ln x$	$\mathbb{R}_{++}$	$-1 - \ln(-y)$	$-\mathbb{R}_{++}$
$\cosh x$	$\mathbb{R}$	$y \sinh^{-1}(y) - \sqrt{1 + y^2}$	$\mathbb{R}$
$-\ln(\cos x)$	$(-\pi/2, \pi/2)$	$y \operatorname{tg}^{-1}(y) - \ln\sqrt{1 + y^2}$	$\mathbb{R}$
$\ln(\cosh x)$	$\mathbb{R}$	$y \operatorname{tgh}^{-1}(y) + \ln\sqrt{1 - y^2}$	$(-1, 1)$
e^x	$\mathbb{R}$	$\begin{cases} y(\ln y - 1), & y > 0 \\ 0, & y = 0 \end{cases}$	$\mathbb{R}_+$
$\ln(1 + e^x)$	$\mathbb{R}$	$\begin{cases} y \ln y + (1 - y)\ln(1 - y), & y \in (0, 1) \\ 0, & y \in \{0, 1\} \end{cases}$	$[0, 1]$
$-\ln(1 - e^x)$	$\mathbb{R}_-$	$\begin{cases} y \ln y - (1 + y)\ln(1 + y), & y > 0 \\ 0, & y = 0 \end{cases}$	$\mathbb{R}_+$

Tabulka 6.1: Fenchelova transformace vybraných funkcí a její vlastnosti.

Následující lemma využijeme v důkazu jednoho z nejdůležitějších výsledků této podkapitoly.

Lemma 6.6.10

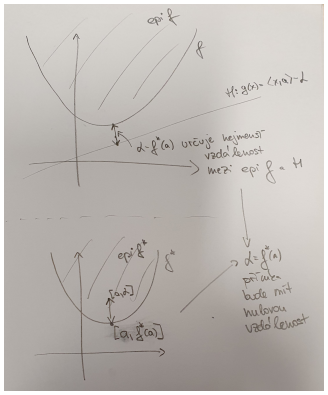
Nechť $g(x) := \langle a, x\rangle - \alpha$ pro $a \in \mathbb{R}^n$ a $\alpha \in \mathbb{R}$. Nechť dále $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$. Pak $f(x) \geq g(x)$ pro každé $x \in X$ právě tehdy, když $[a, \alpha] \in \operatorname{epi} f^*$.

Důkaz. Z definic jednotlivých pojmů plyne, že $[a, \alpha] \in \operatorname{epi} f^*$ právě tehdy, když $\alpha \geq f^*(a)$, což je ekvivalentní s tím, že pro všechna $x \in X$ platí

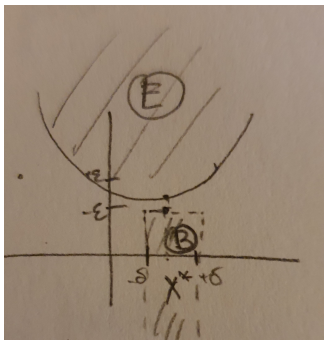
$$\alpha \geq \langle a, x\rangle - f(x) \quad \text{neboli} \quad f(x) \geq \langle a, x\rangle - \alpha = g(x). \qquad \blacksquare$$

Nyní si s využitím Lemma 6.6.10 ukážeme, že Fenchelova transformace nám umožní zpátky rekonstruovat funkci f pomocí vhodných minorizujících přímek, čímž naše úvahy v této podkapitole začaly, viz Obrázek 6.87. Navíc toto tvrzení je základním stavebním kamenem důkazu identity $f = f^{**}$.

Věta 6.6.11



Obrázek 6.87: Nadgraf funkce f, minorizující přímky a funkce f^*.



Obrázek 6.88: Množiny E a B z důkazu Věty 6.6.11.

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f : X \to \mathbb{R}$ konvexní a $x^* \in X$. Je-li funkce f spojitá v bodě x^*, pak platí

$$f(x^*) = \sup\left\{\langle a, x^*\rangle - \alpha \mid (a,\alpha)^\top \in \mathbb{R}^{n+1},\ \langle a, x\rangle - \alpha \leq f(x) \text{ pro všechna } x \in X\right\} = \tag{6.6.2}$$

$$= \sup\left\{\langle a, x^*\rangle - \alpha \mid [a,\alpha] \in \operatorname{epi} f^*\right\}, \tag{6.6.3}$$

tj. graf funkce f je v bodech spojitosti obálkou grafů všech afinních funkcí, které jsou na X majorizovány funkcí f.

Důkaz. Nechť A značí množinu stojící na pravé straně (6.6.2). Volbou $x = x^*$ dostaneme

$$\langle a, x^*\rangle - \alpha \leq f(x^*), \quad \text{tj.} \quad \sup A \leq f(x^*).$$

Ukážeme[165,166], že pro každé $\varepsilon > 0$ existuje $(a,\alpha)^\top \in \mathbb{R}^{n+1}$ splňující $\langle a, x\rangle - \alpha \leq f(x)$ pro $x \in X$ tak, že $\langle a, x^*\rangle - \alpha \geq f(x^*) - \varepsilon$. V kombinaci s předchozí nerovností pak odtud dostaneme $\sup A = f(x^*)$. Protože funkce f je spojitá v x^*, pak ke každému $\varepsilon > 0$ existuje $\delta > 0$ takové, že $f(x) > f(x^*) - \varepsilon$ pro $x \in \mathcal{O}_\delta(x^*)$. Uvažme proto nyní množiny

$$E := \operatorname{epi} f \subseteq \mathbb{R}^{n+1}, \quad B := \mathcal{O}_\delta(x^*) \times (-\infty, f(x^*) - \varepsilon] \subseteq \mathbb{R}^{n+1},$$

viz Obrázek 6.88. Tyto množiny jsou vlastně[167] oddělitelné podle Věty 6.3.18, takže existuje vektor $(p,\lambda)^\top \in \mathbb{R}^{n+1} \setminus \{0\}$ takový, že

$$\left\langle (x,\gamma)^\top, (p,\lambda)^\top \right\rangle \geq \tilde{\alpha} \geq \left\langle (y,\beta)^\top, (p,\lambda)^\top \right\rangle$$

pro každé $(x,\gamma) \in E$ a $(y,\beta) \in B$. Pak stejným způsobem jako v důkazu Lemma 6.5.5 můžeme ukázat, že nutně $\lambda > 0$. Tudíž

$$\left\langle \tfrac{1}{\lambda} p, x \right\rangle + \gamma \geq \tilde{\alpha}/\lambda \geq \left\langle \tfrac{1}{\lambda} p, y \right\rangle + \beta. \tag{6.6.4}$$

Položme $a := -\frac{1}{\lambda} p$ a $\alpha := -\tilde{\alpha}/\lambda$. Jelikož nerovnost (6.6.4) je splněna zejména pro $[x, f(x)] \in E$ a $[x^*, f(x^*) - \varepsilon] \in B$, plyne odtud

$$\langle -a, x\rangle + f(x) \geq -\alpha \geq \langle -a, y\rangle + f(x^*) - \varepsilon,$$

tj. $f(x) \geq \langle a, x\rangle - \alpha$ pro každé $x \in E$ a současně $\langle a, x^*\rangle - \alpha \geq f(x^*) - \varepsilon$. Proto $\sup A = f(x^*)$. $\blacksquare$

Poznámka 6.6.12. Tvrzení Věty 6.6.11 zůstává v platnosti i za předpokladu pouhé *polospojitosti zdola* funkce f v bodě x^*, tj. platí-li

$$\liminf_{x \to x^*} f(x) \geq f(x^*)$$

neboli když ke každému $\varepsilon > 0$ existuje $\delta > 0$ takové, že pro každé $x \in \mathcal{O}_\delta(x^*)$ platí $f(x) > f(x^*) - \varepsilon$, viz Obrázek 6.89. Tato definice je ekvivalentní s požadavkem, že množina $\{x \in X \mid f(x) > \alpha\}$ je otevřená pro každé $\alpha \in \mathbb{R}$. ▲

[165] Připomeňme (ekvivalentní) definici suprema: $\sup_{s \in S} f(s) = K$ právě tehdy, když (i) $f(s) \leq K$ pro všechna $s \in S$ a (ii) existuje $s \in S$ takové, že $f(s) > K - \varepsilon$ pro libovolné $\varepsilon > 0$.

[166] V předchozí ekvivalentní definici suprema je ve druhé části ostrá nerovnost, zatímco my budeme mít nerovnost neostrou. Avšak snadno si můžete sami rozmyslet, že to není v rozporu.

[167] Nebo dokonce silně? Zaručují dané předpoklady neprázdnou vzdálenost množin B a $\mathbb{R}$? Chcete-li znát odpověď, zkuste se nad tím zamyslet.

[168] Jean-Jacques Moreau (31. července 1923 – 9. ledna 2014) byl francouzský matematik a mechanik.

V důkazu prvního tvrzení v Důsledku 6.6.8 jsme ukázali, že pro každé $x \in \mathbb{R}^n$ platí

$$f(x) \geq f^{**}(x).$$

A kdy bude splněna opačná nerovnost?

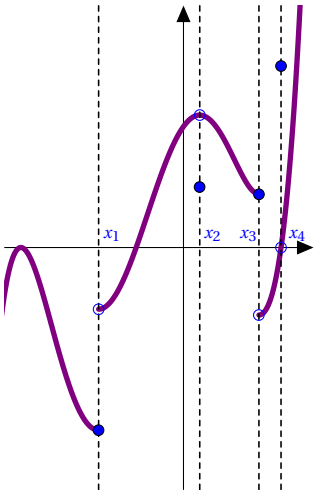
Věta 6.6.13 Fenchel & Moreau[169]

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a funkce $f : X \to \mathbb{R}$ konvexní na X. Pak v každém bodě spojitosti funkce f platí tzv. *Fenchelova rovnost*

$$f^{**}(x) = f(x).$$

Tato rovnost platí zejména[168] pro všechny $x \in \operatorname{ri} X$.

[169] Dá se dokonce ukázat, že Fenchelova rovnost platí právě tehdy, když funkce f je konvexní a polospojitá zdola, viz např. [18, Theorem 1, str. 61].



Obrázek 6.89: Graf funkce, která je polospojitá zdola v bodech x_1 a x_2. Analogicky se definuje polospojitost shora a uvedená funkce je polospojitá shora v bodech x_3 a x_4.

Důkaz. Důkaz je poměrně snadný, neboť podle Věty 6.6.11 platí

$$f(x^*) \overset{(6.6.3)}{=} \sup\{\langle a, x^*\rangle - \alpha \mid [a,\alpha] \in \operatorname{epi} f^*\} = \sup\{\langle a, x^*\rangle - \alpha \mid a \in D(f^*),\ \alpha \geq f^*(a)\} =$$
$$= \sup\{\langle a, x^*\rangle - f^*(a) \mid a \in D(f^*)\} = f^{**}(x^*).$$

Druhá část plyne bezprostředně z Věty 6.4.1. ■

Poznámka 6.6.14.

(i) Zkombinováním Věty 6.6.13 a druhého tvrzení v Důsledku 6.6.8 snadno zjistíme, že pro každou konvexní funkci $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ a libovolné $x \in \operatorname{ri} X$ platí

$$y \in \partial f(x) \quad \text{právě tehdy, když} \quad x \in \partial f^*(y).$$

(ii) Význam opakování různých matematických operací může být odlišný – uzávěr množiny, derivování/integrování funkce atd. A jak je to s Fenchelovou transformací? Předchozí věta říká, že pro spojité konvexní funkce to má smysl nejvýše jednou (podruhé už nic nového nedostaneme). Obecně pak platí, že stačí počítat pouze f^* a f^{**}, neboť dalším konjugováním již nic nového nedostaneme, tj.

$$f^*(y) = f^{***}(y) \quad \text{pro každé } y \in \mathbb{R}^n. \tag{6.6.5}$$

Vskutku, již víme, že $f^{**} \leq f$, takže $f^{***} \geq f^*$ dle Lemma 6.6.6(iii) pro každé $y \in \mathbb{R}^n$. Naopak, z definice f^{***} je

$$f^{***}(y) = \sup_{x \in \mathbb{R}^n}\{\langle y, x\rangle - f^{**}(x)\},$$

zatímco podle Lemma 6.6.6(ii) pro každé $x \in \mathbb{R}^n$ je

$$f^*(y) \geq \langle y, x\rangle - f^{**}(x),$$

což dohromady dává $f^*(y) \geq f^{***}(y)$, a tedy platí (6.6.5). ▲

Na závěr ještě jeden pojem, který nám umožní definitivně propojit konvexní funkce a Fenchelovu transformaci.

Definice 6.6.15

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$. Potom funkce

$$g(x) := \sup\{h(x) \mid h \text{ je konvexní a } h(x) \leq f(x) \text{ pro každé } x \in X\}$$

se nazývá *konvexní obal* (*obálka*) *funkce* f a značí se co f.

Jinými slovy, co f je největší konvexní funkce, která je majorizována funkcí f. Proto je očividné, že co $f = f$ pro libovolnou konvexní funkci f.

Následující věta ukazuje dva ekvivalentní přístupy k definici $\operatorname{co} f$. Motivací pro první z nich je následující jednoduché pozorování: známe-li $\operatorname{epi} f$ pro funkci $f : X \to \mathbb{R}$, pak její předpis můžeme snadno „zrekonstruovat“ pomocí

$$f(x) = \inf\{\alpha \in \mathbb{R} \mid x \in X,\ [x,\alpha] \in \operatorname{epi} f\},$$

přičemž konvexnost $\operatorname{epi} f$ zaručuje konvexnost f. Druhý přístup je mimo jiné založený na Carathéodoryho větě (viz Větu 6.1.14). Jak tedy získat $\operatorname{co} f$?

Věta 6.6.16

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$. Potom

$$\operatorname{co} f(x) = \inf\{\alpha \in \mathbb{R} \mid [x,\alpha] \in \operatorname{conv}(\operatorname{epi} f)\},$$

a tedy

$$\operatorname{co} f(x) = \inf\left\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) \;\middle|\; x_i \in X,\ \sum_{i=1}^{m} \lambda_i x_i = x,\ \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}. \tag{6.6.6}$$

Důkaz. Nechť $g(x) := \inf\{\alpha \mid [x,\alpha] \in \operatorname{conv}(\operatorname{epi} f)\}$. Pak jistě $g(x) \leq f(x)$, neboť (možná) přidáme pouze body, jejichž poslední souřadnice má hodnotu menší než $f(x)$. Množina

$$\operatorname{epi} g = \{[x,\beta] \mid x \in X,\ \beta \geq g(x)\} = \{[x,\beta] \mid x \in X,\ [x,\beta] \in \operatorname{conv}(\operatorname{epi} f)\}$$

je konvexní, a tudíž je podle Věty 6.2.7 funkce g konvexní. Proto platí $g(x) \leq \operatorname{co} f(x)$ pro každé $x \in X$. Na druhou stranu, pro libovolnou funkci $f : X \to \mathbb{R}$ platí $f(x) = \inf\{\alpha \in \mathbb{R} \mid x \in X,\ [x,\alpha] \in \operatorname{epi} f\}$. Navíc pokud pro libovolnou funkci $h : X \to \mathbb{R}$ splňující $h(x) \leq f(x)$ pro všechna $x \in X$ platí $\operatorname{epi} h \supseteq \operatorname{epi} f$, a tedy i pro jejich konvexní obaly platí $\operatorname{conv}(\operatorname{epi} h) \supseteq \operatorname{conv}(\operatorname{epi} f)$. Je-li funkce h konvexní, pak

$$h(x) = \inf\{\alpha \in \mathbb{R} \mid [x,\alpha] \in \operatorname{epi} h = \operatorname{conv}(\operatorname{epi} h)\} \leq \inf\{\alpha \in \mathbb{R} \mid [x,\alpha] \in \operatorname{conv}(\operatorname{epi} f)\} = g(x).$$

Jenže konvexní funkce h splňující $h(x) \leq f(x)$ pro každé $x \in X$ byla zvolena libovolně, takže pro každé $x \in X$ platí nerovnost

$$\sup\{h(x) \mid h(x) \leq f(x) \text{ pro každé } x \in X, \text{ funkce } h \text{ je konvexní}\} \leq g(x)$$

neboli $\operatorname{co} f(x) \leq g(x)$. To dohromady s první částí důkazu implikuje rovnost

$$\operatorname{co} f(x) = g(x) = \inf\{\alpha \in \mathbb{R} \mid [x,\alpha] \in \operatorname{conv}(\operatorname{epi} f)\}.$$

Nyní ukážeme, že platí i (6.6.6). Nejdříve však pro každé $x \in X$ definujme funkci

$$\widetilde{\operatorname{co}} f(x) := \inf\left\{ \sum_{i=1}^{m} \lambda_i f(x_i) \;\middle|\; m \in \mathbb{N},\ \sum_{i=1}^{m} \lambda_i x_i = x,\ \lambda \in \Lambda_m \right\},$$

kde $\Lambda_m := \{(\lambda_1,\dots,\lambda_m) \mid \sum_{i=1}^m \lambda_i = 1,\ \lambda_i \in [0,1]\}$. Protože $f \geq \operatorname{co} f$ a $\operatorname{co} f$ je konvexní funkce, platí pro libovolnou konvexní kombinaci $x = \sum \lambda_i x_i$ podle Jensenovy nerovnosti (viz Věta 6.2.12)

$$\operatorname{co} f(x) = \operatorname{co} f\Big(\sum \lambda_i x_i\Big) \leq \sum \lambda_i \operatorname{co}(x_i) \leq \sum \lambda_i f(x_i),$$

z čehož plyne

$$\operatorname{co} f(x) \leq \inf\Big\{\sum \lambda_i f(x_i)\Big\} = \widetilde{\operatorname{co}} f(x). \tag{6.6.7}$$

Ukážeme, že funkce $\widetilde{\operatorname{co}} f(x)$ je konvexní, tj. pro libovolné $y, z \in X$ a $t \in [0,1]$ máme

$$\widetilde{\operatorname{co}} f(tz + (1-t)y) \leq t\,\widetilde{\operatorname{co}} f(z) + (1-t)\,\widetilde{\operatorname{co}} f(y). \tag{6.6.8}$$

Nechť $\varepsilon > 0$ je libovolné a nechť $m, \ell \geq n+1$, $\lambda \in \Lambda_m$, $\mu \in \Lambda_\ell$ a $y_i, z_i \in X$ jsou takové, že[170]

[170] Taková y_i, z_i lze volit díky definici infima.

$$\varepsilon + \widetilde{\operatorname{co}} f(z) \geq \sum_{i=1}^{m} \lambda_i f(z_i) \quad \& \quad z = \sum_{i=1}^{m} \lambda_i z_i,$$
$$\varepsilon + \widetilde{\operatorname{co}} f(y) \geq \sum_{i=1}^{\ell} \mu_i f(y_i) \quad \& \quad y = \sum_{i=1}^{\ell} \mu_i y_i.$$

Položme

$$\alpha_i := t\lambda_i \quad \& \quad x_i = z_i \quad \text{pro } i = 1,\dots,m,$$
$$\alpha_{m+i} := (1-t)\mu_i \quad \& \quad x_{m+i} = y_i \quad \text{pro } i = 1,\dots,\ell.$$

Pak $\alpha = (\alpha_1,\dots,\alpha_{m+\ell}) \in \Lambda_{m+\ell}$ a sečtením t-násobku první nerovnost s $(1-t)$-násobkem druhé nerovnosti obdržíme

$$\varepsilon + t\widetilde{\operatorname{co}} f(z) + (1-t)\widetilde{\operatorname{co}} f(y) \geq \sum_{i=1}^{m+\ell} \alpha_i f(x_i) \geq \widetilde{\operatorname{co}} f(x), \quad \sum_{i=1}^{m+\ell} \alpha_i x_i = tz + (1-t)y =: x,$$

přičemž druhá nerovnost plyne z definice $\widetilde{\operatorname{co}} f$. Jelikož $\varepsilon > 0$ bylo zvoleno libovolně, skutečně platí nerovnost (6.6.8), a tedy funkce $\widetilde{\operatorname{co}} f$ je konvexní. Navíc platí

$$\operatorname{co} f(x) \overset{(6.6.7)}{\leq} \widetilde{\operatorname{co}} f(x) \leq f(x),$$

neboť každé $x \in X$ můžeme zapsat jako $x = 1 \cdot x + 0$, a tedy buď v x nastává infimum nebo $\sum \lambda_i x_i$ je menší než $f(x)$ pro nějakou konvexní kombinaci. Vzhledem k této nerovnosti a k dokázané konvexnosti funkce $\widetilde{\operatorname{co}} f(x)$ nutně

$$\widetilde{\operatorname{co}} f(x) = \operatorname{co} f(x) \quad \text{pro všechna } x \in X.$$

Nyní ukážeme, že stačí vzít $m = n+1$. Začneme však s tím, že se můžeme omezit na situaci $m \leq n+2$. Uvažme množinu

$$E := \{[x_i, f(x_i)] \in \operatorname{epi} f \mid i = 1,\dots,m\} \subseteq \mathbb{R}^{n+1}.$$

Potom pro $\mu_1,\dots,\mu_m \geq 0$ splňující $\sum_{i=1}^{m} \mu_i = 1$ máme

$$\Big[\sum_{i=1}^{m} \mu_i x_i, \sum_{i=1}^{m} \mu_i f(x_i)\Big] \in \operatorname{conv} E$$

a podle Carathéodoryho věty (viz Věta 6.1.14) existují $y_i \in \{x_1,\dots,x_m\}$, $i = 1,\dots,n+2$ a vektor $(\lambda_1,\dots,\lambda_{n+2}) \in \Lambda_{n+2}$ tak, že

$$\Big[\sum_{i=1}^{m} \mu_i x_i, \sum_{i=1}^{m} \mu_i f(x_i)\Big] = \Big[\sum_{i=1}^{n+2} \lambda_i y_i, \sum_{i=1}^{n+2} \lambda_i f(y_i)\Big].$$

Tedy skutečně stačí nejvýše $n+2$ bodů.

Nyní zbývá ukázat, že dokonce stačí nejvýše $n+1$ bodů, tedy můžeme najít takové $\eta \in \Lambda_{n+2}$, že $\eta_i = 0$ pro alespoň jeden index $i \in \{1,\dots,n+2\}$ (tj. po odstranění tohoto členu vlastně můžeme bez újmy na obecnosti brát $\eta \in \Lambda_{n+1}$) a $y_1,\dots,y_{n+2}$ takové, že

$$\sum_{i=1}^{n+2} \eta_i f(y_i) \leq \sum_{i=1}^{n+2} \lambda_i f(y_i). \tag{6.6.9}$$

Nechť $(\lambda_1,\dots,\lambda_{n+2}) \in \Lambda_{n+2}$ je takové, že $\lambda_1,\dots,\lambda_{n+2} > 0$ (jinak je důkaz hotov). Označme

$$x := \sum_{i=1}^{n+2} \lambda_i y_i.$$

Pak $x \in \operatorname{conv}\{y_1,\dots,y_{n+2}\} \subseteq \mathbb{R}^n$. Aplikujeme-li Carathéodoryho větu (viz Větu 6.1.14) ještě jednou, pak existuje

$$\tilde{\lambda}_1,\dots,\tilde{\lambda}_{n+2} \geq 0 \quad \& \quad \sum_{i=1}^{n+2} \tilde{\lambda}_i = 1$$

a současně alespoň pro jedno $i \in \{1,\dots,n+2\}$ je $\tilde\lambda_i = 0$ a zároveň

$$\sum_{i=1}^{n+2} \tilde\lambda_i y_i = x.$$

Bez újmy na obecnosti předpokládejme, že

$$\sum_{i=1}^{n+2} \tilde\lambda_i f(y_i) > \sum_{i=1}^{n+2} \lambda_i f(y_i), \tag{6.6.10}$$

jinak totiž volbou $\eta_i = \tilde\lambda_i$ plyne (6.6.9). Položme

$$J := \{i \in 1,\dots,n+2 \mid \lambda_i - \tilde\lambda_i < 0\}.$$

Zřejmě je $J \neq \emptyset$, neboť jinak $\lambda_i \geq \tilde\lambda_i \geq 0$ pro všechna $i \in \{1,\dots,n+2\}$ a s využitím skutečnosti, že $\tilde\lambda_i = 0$ pro nějaké i, dostaneme spor s podmínkami $\sum_{i=1}^{n+2} \tilde\lambda_i = 1 = \sum_{i=1}^{n+2} \lambda_i$ a zároveň $\lambda_1,\dots,\lambda_{n+2} > 0$. Definujme

$$\gamma := \min_{i\in J} \{\lambda_i/(\tilde\lambda_i - \lambda_i)\}.$$

Jistě platí $\gamma > 0$. Pro všechna $i \in \{1,\dots,n+2\}$ položme

$$\eta_i := \lambda_i + \gamma(\lambda_i - \tilde\lambda_i).$$

Potom

$$\eta_1,\dots,\eta_{n+2} \geq 0 \quad \& \quad \sum_{i=1}^{n+2} \eta_i = \underbrace{\sum_{i=1}^{n+2} \lambda_i}_{=1} + \gamma\Big(\underbrace{\sum_{i=1}^{n+2} \lambda_i}_{=1} - \underbrace{\sum_{i=1}^{n+2} \tilde\lambda_i}_{=1}\Big) = 1 \quad \& \quad \eta_i = 0$$

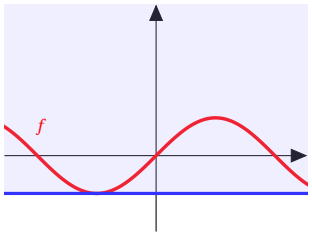
pro takový index $i \in \{1,\dots,n+2\}$, ve kterém nastává minimum v definici množiny J.

Konečně z (6.6.10) plyne

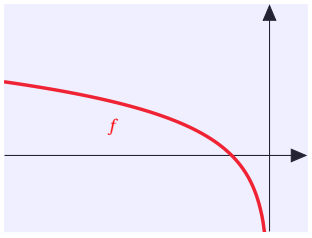
$$\sum_{i=1}^{n+2} \eta_i f(y_i) = \sum_{i=1}^{n+2} \lambda_i f(y_i) + \gamma \underbrace{\sum_{i=1}^{n+2} (\lambda_i - \tilde\lambda_i) f(y_i)}_{\overset{(6.6.10)}{<} 0} < \sum_{i=1}^{n+2} \lambda_i f(y_i),$$

a tudíž dostáváme nerovnost (6.6.9). Tím je důkaz hotov[171]. ■

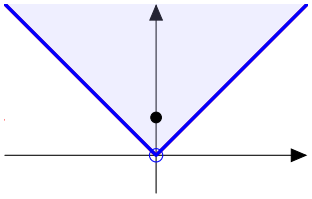
[171] Pro $\operatorname{co} f$ bereme infimum, takže stačí uvažovat pouze kombinace $\sum_{i=1}^{n+2} \eta_i f(y_i)$.



Obrázek 6.90: Funkce $\sin x$ a její konvexní obálka (modrá barva).



Obrázek 6.91: Funkce $\ln(-x)$ a její konvexní obálka (modrá barva).



Obrázek 6.92: Funkce f z Poznámky 6.6.18 a její konvexní obálka.

[172] Už jen proto, že není spojitá v počátku

Příklad 6.6.17

Uvažme $f(x) = \sin x$ pro $X = \mathbb{R}$. Pak $\operatorname{co} f \equiv -1$, neboť $\operatorname{conv}(\operatorname{epi} f) = \mathbb{R} \times [-1,\infty)$, a tudíž

$$\inf\{\alpha \in \mathbb{R} \mid [x,\alpha] \in \mathbb{R} \times [-1,\infty)\} = -1 \quad \text{pro všechna } x \in \mathbb{R},$$

viz Obrázek 6.90. Mohou samozřejmě nastat i velmi patologické situace jako např. pro $f(x) = \ln(-x)$ a $X = (-\infty, 0)$, což dává $\operatorname{co} f \equiv -\infty$, viz Obrázek 6.91.

Poznámka 6.6.18. Platí

$$D(\operatorname{co} f) = \operatorname{conv}(D(f)),$$

což dává $\operatorname{conv}(\operatorname{epi} f) \subseteq \operatorname{epi}(\operatorname{co} f)$. Ale inkluze může být i ostrá. Např. pro funkci

$$f(x) := \begin{cases} |x|, & x \neq 0, \\ 1, & x = 0, \end{cases}$$

která není konvexní[172]. Pak

$$\operatorname{epi}(\operatorname{co} f) = \{[x,y] \in \mathbb{R}^2 \mid y \geq |x|\} \quad \text{a} \quad \operatorname{conv}(\operatorname{epi} f) = \{[x,y] \in \mathbb{R}^2 \mid y \geq |x|\} \setminus \{[0,0]\},$$

viz Obrázek 6.92. ▲

Předchozí věta nám sice dává bližší objasnění co f, ovšem z praktického hlediska stále není příliš užitečná. Jak vlastně spočítat co f? V (6.6.6) dostáváme minimalizační problém, který ale již v případě funkce jedné proměnné obsahuje čtyři proměnné $x_1, x_2 \in X$ a $\lambda_1, \lambda_2 \in [0,1]$ a dvě omezení

$$\lambda_1 x_1 + \lambda_2 x_2 = 1 \quad \& \quad \lambda_1 + \lambda_2 = 1.$$

Chtělo by to něco jednoduššího… Naštěstí máme Fenchelovu transformaci.

Věta 6.6.19

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : X \to \mathbb{R}$. Potom pro každé $x \in \operatorname{ri} X$ platí

$$\operatorname{co} f(x) = f^{**}(x).$$

Důkaz. Z Fenchelovy nerovnosti plyne

$$f(x) \geq \langle x, y \rangle - f^*(y)$$

pro všechna $x, y \in \mathbb{R}^n$ (přičemž pro $x \notin D(f)$ a $y \notin D(f^*)$ klademe $f(x) = \infty = f(y)$), a tedy

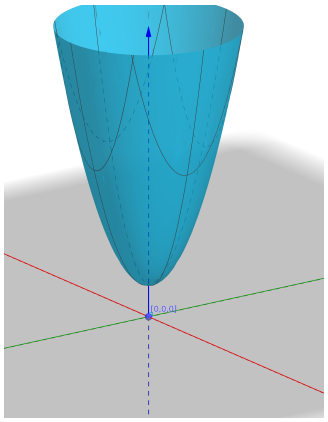
$$f(x) \geq \sup_{y \in \mathbb{R}^n} \{\langle x, y \rangle - f^*(y)\} = f^{**}(y).$$

Proto (s využitím Věty 6.6.16) platí $\operatorname{co} f(x) \geq \operatorname{co} f^{**}(x)$. Jenže podle Lemma 6.6.6(i) je f^{**} konvexní, tj. $\operatorname{co} f^{**}(x) = f^{**}(x)$, což celkem dává $\operatorname{co} f(x) \geq f^{**}(x)$.

Nyní ukážeme opačnou inkluzi. Funkce $\operatorname{co} f(x)$ je konvexní, a tedy podle Věty 6.4.1 je spojitá pro každé $x \in \operatorname{ri} X$. Pak podle Věty 6.6.13 platí $(\operatorname{co} f)^{**}(x) = \operatorname{co} f(x)$ pro každé $x \in \operatorname{ri} X$. Navíc podle definice máme $\operatorname{co} f(x) \leq f(x)$ pro každé $x \in X$, což implikuje $(\operatorname{co} f)^{**}(x) \leq f^{**}(x)$, viz Lemma 6.6.6(iii). Proto v každém bodě $x \in \operatorname{ri} X$ platí

$$\operatorname{co} f(x) = (\operatorname{co} f)^{**}(x) \leq f^{**}(x),$$

což dohromady s první částí dává požadované tvrzení. ■



Obrázek 6.93: Graf funkce f z Příkladu 6.6.20

Příklad 6.6.20

Určeme $\operatorname{co} f$ pro funkci

$$f(x_1, x_2) := \begin{cases} 1 + x_1^2 + x_2^2, & x_1^2 + x_2^2 \neq 0, \\ 0, & x_1^2 + x_2^2 = 0, \end{cases}$$

viz Obrázek 6.93. Funkce f je definována na celém $\mathbb{R}^2$, takže můžeme využít Větu 6.6.19 a určit $\operatorname{co} f$ pomocí bikonjugované funkce f^{**}. Navíc f je tzv. *radiálně symetrická*, tj. funkční hodnoty závisí pouze na vzdálenosti bodu $[x_1, x_2]$ od počátku. Můžeme proto místo f pracovat s funkcí

$$\widetilde{f} := \begin{cases} 1 + z^2, & z \neq 0, \\ 0, & z = 0, \end{cases}$$

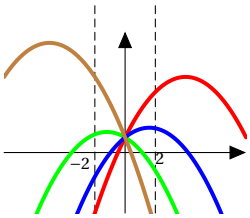
kde $z := \sqrt{x_1^2 + x_2^2}$. Potom

$$\widetilde{f}^*(y) = \sup z \in \mathbb{R}\{zy - \widetilde{f}(z)\} = \max\left\{\sup_{z \neq 0}\{zy - 1 - z^2\}, 0\right\} =$$

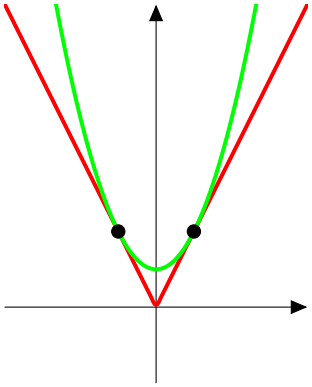
$$= \max\{y^2/4 - 1, 0\} = \begin{cases} y^2/4 - 1, & |y| \geq 2, \\ 0, & |y| < 2, \end{cases}$$

neboť funkce $zy - 1 - z^2$ je konkávní pro libovolné $y \in \mathbb{R}$ a jejím stacionárním bodem je $z^* = y/2$, tj. v tomto bodě má funkce $zy - 1 - z^2$ globální maximum pro $y \neq 0$, zatímco v případě $y = 0$ je hodnota suprema rovna -1. Dále

$$\widetilde{f}^{**}(z) = \sup_{y \in \mathbb{R}}\{zy - \widetilde{f}^*(y)\} = \max\left\{\sup_{|y|<2}\{zy\}, \sup_{|y|\geq 2}\{zy - y^2/4 + 1\}\right\} = \max\{2|z|, g(z)\},$$



Obrázek 6.94: Graf funkce $zy - y^2/4+1$ pro $z=2$ (červená), $z=4/5$ (modrá), $z=-3/5$ (zelená) a $z=-5/2$ (hnědá).



Obrázek 6.95: Graf funkce $2|z|$ (červená) a $1+z^2$ (zelená).

kde

$$g(z) := \begin{cases} 2|z|, & |z| \le 1, \\ 1+z^2, & |z| \ge 1. \end{cases}$$

Vskutku, lineární funkce zy nabývá svého suprema na intervalu $(-2,2)$ v bodě $y=2$ pro $z>0$, v bodě $y=-2$ pro $z<0$ a v $y=0$ pro $z=0$, tj. v bodě $2\operatorname{sgn} z$ s hodnotou $2\operatorname{sgn}(z)\,z = 2|z|$. Kvadratická funkce $h(y) := zy - y^2/4+1$ je konkávní pro libovolné $z\in\mathbb{R}$ s maximem (vrcholem) mezi jejími kořeny, tj. v bodě

$$y^* := (y_1+y_2)/2 = (2z - \sqrt{z^2+1} + 2z + \sqrt{z^2+1})/2 = 2z$$

s odpovídající funkční hodnotou z^2+1. Tento bod vyhovuje podmínce $|y^*|\ge 2$ právě tehdy, když $|z|\ge 1$. V opačném případě leží toto globální minimum mimo povolený interval a funkce $h(y)$ nabude své největší hodnoty v některém z krajních bodů bodů intervalu $(-2,2)$. Funkční hodnota v těchto bodech je $\pm 2z$, z čehož vidíme, že supremum nastává v $2\operatorname{sgn}(z)$ s hodnotou $2|z|$, viz Obrázek 6.94. Tudíž

$$\widetilde{f}^{**}(z) = \begin{cases} 2|z|, & |z| \le 1, \\ 1+z^2, & |z| \ge 1, \end{cases}$$

viz Obrázek 6.95, neboli

$$\operatorname{co} f(x_1,x_2) = f^{**}(x_1,x_2) = \begin{cases} 2\sqrt{x_1^2+x_2^2}, & x_1^2+x_2^2 \le 1, \\ 1+x_1^2+x_2^2, & x_1^2+x_2^2 \ge 1. \end{cases}$$

Cvičení

6.6.1. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} 1, & x \ge 0, \\ -1, & x < 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(0)=1$, $f^*(y)=\infty$ pro $\mathbb{R}\setminus\{0\}$

6.6.2. Určete konjugovanou funkci f^* pro

$$f(x) = \max\{0, 1-x\}, \quad x\in\mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y$ pro $y\in[-1,0]$, $f^*(y)=\infty$ pro $y\in\mathbb{R}\setminus[-1,0]$

6.6.3. Určete konjugovanou funkci f^* pro

$$f(x) = \max\{2x, 3x\}, \quad x\in[-1,1],$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y-3$ pro $y\ge 3$, $f^*(y)=0$ pro $y\in[2,3]$, $f^*(y) = -y+2$ pro $y\le 2$

6.6.4. Určete konjugovanou funkci f^* pro

$$f(x) = |x|, \quad x\in\mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y)=0$ pro $y\in[-1,1]$, $f^*(y)=\infty$ pro $\mathbb{R}\setminus[-1,1]$

6.6.5. Pro libovolné $c>0$ určete konjugovanou funkci f^* pro funkci

$$f(x) = \frac{cx^2}{2}, \quad x\in\mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y^2/(2c)$

6.6.6. Určete konjugovanou funkci f^* pro

$$f(x) = |x|^{3/2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \frac{4}{27}|y|^3$

6.6.7. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} x^2, & x \geq 0, \\ x^2/2, & x \leq 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y^2/4$ pro $y \geq 0$, $f^*(y) = y^2/2$ pro $y \leq 0$

6.6.8. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} 1, & x \in [-1,1], \\ x^2, & |x| \geq 1, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y^2/4$ pro $|y| \geq 2$, $f^*(y) = |y| - 1$ pro $y \in [-2,2]$

6.6.9. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} 1 + x^2, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = 0$ pro $y \in [-2,2]$, $f^*(y) = y^2/4 - 1$ pro $|y| \geq 2$, $f^{**} = 2|x|$ pro $|x| \leq 1$ a $f^{**} = 1 + x^2$ pro $|x| \geq 1$

6.6.10. Určete konjugovanou funkci f^* pro

$$f(x) = x^3 - 5x, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \infty$ pro $y \in \mathbb{R}$, $f^{**}(x) = -\infty$ pro $x \in \mathbb{R}$

6.6.11. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} 3x^3, & x \geq 0, \\ x^2, & x \leq 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \frac{2}{9}y^{3/2}$ pro $y \geq 0$ a $f^*(y) = y^2/4$ pro $y \leq 0$

6.6.12. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} x^2 - 1, & |x| \geq 1, \\ 0, & |x| \leq 1, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = |y|$ pro $y \in [-2,2]$, $f^*(y) = y^2/4 + 1$ pro $|y| \geq 2$

6.6.13. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} (x^2 - 1)^2, & |x| \geq 1, \\ 0, & |x| \leq 1, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

6.6.14. Určete konjugovanou funkci f^* pro

$$f(x) = 1 + \frac{x^4}{4}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -1 + \frac{3}{4}\, y^{4/3}$

6.6.15. Určete konjugovanou funkci f^* pro

$$f(x) = (x^2 - 1)^2, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

6.6.16. Určete konjugovanou funkci f^* pro

$$f(x) = \sqrt{1 - x}, \quad x \le 1,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y$ pro $y \ge 0$, $f^*(y) = \infty$ pro $y < 0$, $f^{**}(x) = \infty$ pro $x > 1$, $f^{**}(x) = 0$ pro $x \le 1$

6.6.17. Určete konjugovanou funkci f^* pro

$$f(x) = \sqrt{1 + x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -\sqrt{1 - y^2}$ pro $y \in [-1, 1]$, $f^*(y) = \infty$ pro $|y| > 1$

6.6.18. Určete konjugovanou funkci f^* pro

$$f(x) = 2\sqrt{1 + x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -2\sqrt{1 - y^2/4}$ pro $y \in [-2, 2]$, $f^*(y) = \infty$ pro $|y| > 2$

6.6.19. Určete konjugovanou funkci f^* pro

$$f(x) = -\sqrt{1 - x^2}, \quad x \in [-1, 1],$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \sqrt{1 + y^2}$

6.6.20. Určete konjugovanou funkci f^* pro

$$f(x) = -\sqrt{2x - x^2}, \quad x \in [-1, 1],$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y + \sqrt{1 + y^2}$

6.6.21. Určete konjugovanou funkci f^* pro

$$f(x) = -\sqrt{1 - 4x^2}, \quad x \in [-1/2, 1/2],$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \frac{1}{2}\sqrt{4 + y^2}$

6.6.22. Určete konjugovanou funkci f^* pro

$$f(x) = \sqrt{1 + 4x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -\frac{1}{2}\sqrt{4 - y^2}$

6.6.23. Určete konjugovanou funkci f^* pro

$$f(x) = x + \sqrt{1+x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -\sqrt{2y - y^2}$

6.6.24. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} \frac{x^2}{4} - 1, & x \geq 0, \\ -\sqrt{1-2x}, & x \leq 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \frac{y^2+1}{2y}$ pro $y \in (0, 1/2]$, $f^*(y) = y^2 + 1$ pro $y \geq 1/2$, $f^*(y) = \infty$ pro $y \leq 0$,
$f^{**}(x) = \frac{x^2}{4} - 1$ pro $x \geq 1$, $f^{**}(x) = \frac{x}{2} - \frac{5}{4}$ pro $x \in [-3/2, 1]$,
$f^{**}(x) = -\sqrt{1-2x}$ pro $x \leq -3/2$

6.6.25. Určete konjugovanou funkci f^* pro

$$f(x) = \mathrm{e}^{|x|}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = |y| \ln |y| - |y|$ pro $|y| \geq 1$, $f^*(y) = -1$ pro $y \in [-1, 1]$

6.6.26. Určete konjugovanou funkci f^* pro

$$f(x) = \mathrm{e}^x + x, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = (y-1)[\ln(y-1) - 1]$ pro $y > 1$, $f^*(1) = 0$, $f^*(y) = \infty$ pro $y < 1$

6.6.27. Určete konjugovanou funkci f^* pro

$$f(x) = \mathrm{e}^x - x, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = (y+1)[\ln(y+1) - 1]$ pro $y > -1$, $f^*(-1) = 0$, $f^*(y) = \infty$ pro $y < -1$

6.6.28. Určete konjugovanou funkci f^* pro

$$f(x) = \begin{cases} \mathrm{e}^x, & x \geq 0, \\ 1, & x \leq 0, \end{cases}$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y(\ln y - 1)$ pro $y \geq 1$, $f^*(y) = -1$ pro $y \in [0, 1]$, $f^*(y) = \infty$ pro $y < 0$

6.6.29. Určete konjugovanou funkci f^* pro

$$f(x) = \sqrt{1 + \mathrm{e}^x}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y \ln(2y^2 + 2y\sqrt{y^2+1}) - \sqrt{1 + 2y^2 + 2y\sqrt{y^2+1}}$ pro $y \geq 0$, $f^*(y) = \infty$ pro $y < 0$

6.6.30. Určete konjugovanou funkci f^* pro

$$f(x) = -\ln x, \quad x > 0,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -1 - \ln(-y)$ pro $y < 0$, $f^*(y) = \infty$ pro $y \geq 0$

6.6.31. Určete konjugovanou funkci f^* pro

$$f(x) = x \ln x, \quad x > 0,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \mathrm{e}^{y-1}$ pro $y \in \mathbb{R}$

6.6.32. Určete konjugovanou funkci f^* pro

$$f(x) = x(\ln x - 1), \quad x > 0,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \mathrm{e}^y$ pro $y \in \mathbb{R}$

6.6.33. Určete konjugovanou funkci f^* pro

$$f(x) = \ln(-x), \quad x < 0,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \infty$ pro $y \in \mathbb{R}$, $f^{**}(x) = -\infty$ pro $x \in \mathbb{R}$

6.6.34. Určete konjugovanou funkci f^* pro

$$f(x) = -\ln(1 - \mathrm{e}^x), \quad x < 0,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y \ln[y/(y+1)] - \ln(y+1)$ pro $y > 0$, $f^*(0) = 0$, $f^*(y) = \infty$ pro $y < 0$

6.6.35. Určete konjugovanou funkci f^* pro

$$f(x) = -\ln(1 - x^2), \quad x \in (-1, 1),$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = -1 + \sqrt{1+y^2} + \ln\left[(-2 + 2\sqrt{1+y^2})/y^2\right]$ pro $y \neq 0$, $f^*(0) = 0$

6.6.36. Určete konjugovanou funkci f^* pro

$$f(x) = \ln(1 + \mathrm{e}^x), \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y \ln[y/(1-y)] + \ln(1-y)$ pro $y \in (0,1)$, $f^*(0) = 0$, $f^*(y) = \infty$ pro $y \in \mathbb{R} \setminus [0,1)$

6.6.37. Určete konjugovanou funkci f^* pro

$$f(x) = x + \ln(1 + \mathrm{e}^x), \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = (y-1)\ln(y-1) + (2-y)\ln(2-y)$ pro $y \in (1,2)$, $f^*(y) = 0$ pro $y \in \{1,2\}$, $f^*(y) = \infty$ pro $y \in \mathbb{R} \setminus [1,2]$

6.6.38. Určete konjugovanou funkci f^* pro

$$f(x) = x \ln(x + \sqrt{1+x^2}) - \sqrt{1+x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \cosh y$ pro $y \in \mathbb{R}$

6.6.39. Určete konjugovanou funkci f^* pro

$$f(x) = -\ln(\cos x), \quad x \in (-\pi/2, \pi/2),$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y \operatorname{arctg} y - \frac{1}{2}\ln(1+y^2)$ pro $y \in \mathbb{R}$

6.6.40. Určete konjugovanou funkci f^* pro

$$f(x) = \sin x, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(0) = 1$, $f^*(y) = \infty$ pro $\mathbb{R} \setminus \{0\}$, $f^{**}(x) = -1$ pro $x \in \mathbb{R}$

6.6.41. Určete konjugovanou funkci f^* pro

$$f(x) = \sin x, \quad x \in [-\pi, 0],$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

6.6.42. Určete konjugovanou funkci f^* pro

$$f(x) = x + \sin x \cos x + 1, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(1) = -1/2$, $f^*(y) = \infty$ pro $y \neq 1$, $f^{**}(x) = x + 1/2$ pro $x \in \mathbb{R}$

6.6.43. Určete konjugovanou funkci f^* pro

(a) $f(x) = x + \sin x, \quad x \in \mathbb{R}$; (b) $f(x) = 1 + 2x + \sin^3 x \cos x, \quad x \in \mathbb{R},$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: (a) $f^*(1) = 1$, $f^*(y) = \infty$ pro $y \neq 1$, $f^{**}(x) = x - 1$ pro $x \in \mathbb{R}$;
(b) $f^*(2) = -1 + 3\sqrt{3}/16$, $f^*(y) = \infty$ pro $y \neq 2$, $f^{**}(x) = 2x + 1 - 3\sqrt{3}/16$ pro $x \in \mathbb{R}$

6.6.44. Určete konjugovanou funkci f^* pro

$$f(x) = \cosh x, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = y \ln(y + \sqrt{1+y^2}) - \sqrt{1+y^2}$ pro $y \in \mathbb{R}$

6.6.45. Určete konjugovanou funkci f^* pro

$$f(x) = x \operatorname{arctg} x - \ln \sqrt{1+x^2}, \quad x \in \mathbb{R},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \ln \sqrt{1 + \operatorname{tg}^2 y}$ pro $y \in \mathbb{R}$

6.6.46. Odvoďte vzorec pro výpočet konjugované funkce k funkci $f(x-a)$ pomocí f^*. S využitím tohoto výsledku a řešení Příkladu 6.6.6.17 určete konjugovanou funkci k funkci $f(x) = \sqrt{x^2 - 2x + 2}$.

řešení: $f^*(y) = -\sqrt{1-y^2} + y$ pro $y \in [-1, 1]$, $f^*(y) = \infty$ pro $|y| > 1$

6.6.47. Odvoďte vzorec pro výpočet konjugované funkce k funkci $c\,f(x)$ pomocí f^*. S využitím tohoto výsledku a řešení Příkladů 6.6.6.17 a 6.6.6.46 určete konjugovanou funkci k funkci $f(x) = 2\sqrt{x^2 + 2x + 2}$.

řešení: $f^*(y) = -y - 2\sqrt{1 - y^2/4}$ pro $y \in [-2, 2]$, $f^*(y) = \infty$ pro $|y| > 2$

6.6.48. Určete konjugovanou funkci f^* pro

$$f(x_1, x_2) = x_1^2 + 2\,x_1\,x_2 + 2x_2^2, \quad [x_1, x_2] \in \mathbb{R}^2,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y_1, y_2) = y_1^2/2 - y_1\,y_2/2 + y_2^2/4$ pro $[y_1, y_2] \in \mathbb{R}^2$

6.6.49. Určete konjugovanou funkci f^* pro

$$f(x_1, x_2) = 1 + 2x_1^2 + 2\,x_1\,x_2 + x_2^2, \quad [x_1, x_2] \in \mathbb{R}^2,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y_1, y_2) = [(y_1 - y_2)^2 + y_2^2 - 4]/4$ pro $[y_1, y_2] \in \mathbb{R}^2$

6.6.50. Určete konjugovanou funkci f^* pro

$$f(x_1, x_2) = -\sqrt{x_1} - \sqrt{x_2}, \quad [x_1, x_2] \in \mathbb{R}^2_+,$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y_1, y_2) = -1/(4y_1) - 1/(4y_2)$ pro $y_1, y_2 < 0$, $f^*(y_1, y_2) = \infty$ ve zbývajících případech

6.6.51. Určete konjugovanou funkci f^* pro

$$f(x_1,x_2)=\begin{cases}1+x_1^2+x_2^2, & [x,y]\neq[0,0],\\ 0, & [x,y]=[0,0].\end{cases}$$

řešení: $f^*(y_1,y_2)=0$ pro $y_1^2+y_2^2\le 4$, $f^*(y_1,y_2)=-1+y_1^2/4+y_2^2/4$ pro $y_1^2+y_2^2\ge 4$,
$f^{**}(x_1,x_2)=2\sqrt{x_1^2+x_2^2}$ pro $x_1^2+x_2^2\le 1$, $f^{**}(x_1,x_2)=1+x_1^2+x_2^2$ pro $x_1^2+x_2^2\ge 1$

6.6.52. Pro libovolná $a,b>0$ určete konjugovanou funkci f^* pro funkci

$$f(x_1,x_2)=\begin{cases}1+\frac{x^2}{a^2}+\frac{y^2}{b^2}, & [x,y]\neq[0,0],\\ 0, & [x,y]=[0,0].\end{cases}$$

řešení: $f^*(y_1,y_2)=0$ pro $a^2y_1^2+b^2y_2^2\le 4$, $f^*(y_1,y_2)=-1+a^2y_1^2/4+b^2y_2^2/4$ pro
$a^2y_1^2+b^2y_2^2\ge 4$,
$f^{**}(x_1,x_2)=2\sqrt{x_1^2/a^2+x_2^2/b^2}$ pro $x_1^2/a^2+x_2^2/b^2\le 1$,
$f^{**}(x_1,x_2)=1+x_1^2/a^2+x_2^2/b^2$ pro $x_1^2/a^2+x_2^2/b^2\ge 1$

6.6.53. Pro libovolné $a\in\mathbb{R}^n$ a $\beta\in\mathbb{R}$ určete konjugovanou funkci f^* pro funkci

$$f(x)=\langle a,x\rangle+\beta,\quad x\in\mathbb{R}^n,$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$.

řešení: $f^*(a)=-b$, $f^*(y)=\infty$ pro $y\neq a$

6.6.54. Určete konjugovanou funkci f^* pro

$$f(x)=\tfrac{1}{2}\|x\|^2=\tfrac{1}{2}x^\top x,\quad x\in\mathbb{R}^n,$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$. Ukažte také, že $f\equiv f^*$ právě tehdy, když $f(x)=\frac{1}{2}\|x\|^2$.

řešení: $f^*(y)=\frac{1}{2}\|y\|^2$ pro $y\in\mathbb{R}^n$

6.6.55. Pro libovolnou symetrickou a pozitivně definitní matici $A\in\mathbb{R}^{n\times n}$ určete konjugovanou funkci f^* pro funkci

$$f(x)=1+x^\top Ax,\quad x\in\mathbb{R}^n,$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$.

řešení: $f^*(a)=-1+\frac{1}{4}y^\top A^{-1}y$ pro $y\in\mathbb{R}^n$

6.6.56. Určete konjugovanou funkci f^* pro

$$f(x)=\|x\|=\sqrt{x^\top x},\quad x\in\mathbb{R}^n,$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$.

řešení: $f^*(y)=0$ pro $\sup_{\|x\|\le 1}x^\top y\le 1$, $f^*(y)=\infty$ pro $\sup_{\|x\|\le 1}x^\top y>1$

6.6.57. Pro libovolné $\alpha>1$ určete konjugovanou funkci f^* pro funkci

$$f(x)=\frac{1}{\alpha}\sum_{i=1}^n|x_i|^\alpha,\quad x\in\mathbb{R}^n,$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$.

řešení: $f^*(y)=\frac{1}{\beta}\sum_{i=1}^n|y_i|^\beta$ pro $y\in\mathbb{R}^n$ a $\beta:=\alpha/(\alpha-1)$

6.6.58. Určete konjugovanou funkci f^* pro

$$f(x)=-\sum_{i=1}^n\ln x_i,\quad x\in\mathbb{R}^n_{++},$$

a ověřte platnost Fenchelovy rovnosti $f^{**}=f$.

řešení: $f^*(y)=-n-\sum_{i=1}^n\ln(-y_i)$ pro $y\in\mathbb{R}^n_{--}$, $f^*(y)=\infty$ v ostatních případech

6.6.59. Pro libovolné $\alpha > 0$ určete konjugovanou funkci f^* pro funkci

$$f(x) = \frac{1}{\alpha} \sum_{i=1}^{n} x_i \ln(x_i), \quad x \in \mathbb{R}^n_{++},$$

a ověřte platnost Fenchelovy rovnosti $f^{**} = f$.

řešení: $f^*(y) = \frac{1}{\alpha} \sum_{i=1}^{n} \mathrm{e}^{\alpha y_i - 1}$ pro $y \in \mathbb{R}^n$

6.6.60. *Dokažte:* Nechť $g : \mathbb{R} \to \mathbb{R}$ je sudá funkce. Definujeme-li $f(x) := g(\|x\|)$, pak platí

$$f^*(y) = g^*(\|y\|_*),$$

kde $\|y\|_* := \sup\{\langle x, y\rangle, x \in \mathbb{R}^n, \|x\| = 1\}$ je tzv. kanonická norma.

Kapitola 7

Numerické metody řešení optimalizačních úloh

7.1 *Není konvergence jako konvergence...* 535
7.2 *Numerické metody nepodmíněné optimalizace v* $\mathbb{R}$ 540
7.2.1 *Metoda prostého dělení intervalu* 542
7.2.2 *Metoda půlení intervalu* 544
7.2.3 *Metoda zlatého řezu* 548
7.2.4 *Fibonacciho metoda* 551
7.2.5 *Metoda půlení intervalu II* 560
7.2.6 *Interpolační metody* 562
7.3 *Numerické metody nepodmíněné optimalizace v* $\mathbb{R}^n$ 569
7.3.1 *Metoda největšího spádu* 572
7.3.2 *Newtonova metoda* 587
7.3.3 *Metoda sdružených gradientů* 593

NUMERICAL ANALYSIS IS OFTEN CONSIDERED NEITHER BEAUTIFUL NOR, INDEED, PROFOUND. PURE MATHEMATICS IS BEAUTIFUL IF YOUR HEART GOES AFTER THE JOY OF ABSTRACTION, APPLIED MATHEMATICS IS BEAUTIFUL IF YOU ARE EXCITED BY MATHEMATICS AS A MEANS TO EXPLAIN THE MYSTERY OF THE WORLD AROUND US. BUT NUMERICAL ANALYSIS? SURELY, WE COMPUTE ONLY WHEN EVERYTHING ELSE FAILS, WHEN MATHEMATICAL THEORY CANNOT DELIVER AN ANSWER IN A COMPREHENSIVE, PRISTINE FORM AND THUS WE ARE COMPELLED TO THROW A PROBLEM ONTO A NUMBER-CRUNCHING COMPUTER AND PRODUCE BORING NUMBERS BY BORING CALCULATIONS. THIS, I BELIEVE, IS NONSENSE.[1]

ARIEH ISERLES na straně ix v předmluvě ke své knize [293]

[1] *Numerická analýza nebývá příliš často považována za krásnou a vlastně ani za hlubokomyslnou. Čistá matematika je krásná, pokud vaše srdce plesá radostí z abstrakce, zatímco aplikovaná matematika je krásná, pokud vás matematika vzrušuje jako nástroj k vysvětlení tajemství světa kolem nás. Ale numerická analýza? Nepochybně jsme donuceni počítat až tehdy, když všechno ostatní selže, když matematická teorie nedokáže poskytnout odpověď ve všeobecné, dokonale čisté podobě, a proto jsme nuceni hodit problém do počítače křoupajícího čísla a s pomocí nudných výpočtů vytvořit nudná čísla. Věřím, že to je nesmysl.*

V této kapitole si vybudujeme několik numerických metod (algoritmů) pro řešení obecné úlohy nepodmíněné optimalizace

$$f(x) \to \min, \quad x \in \mathbb{R}^n. \tag{7.1}$$

Na rozdíl od úloh z Kapitol 2–4 se rozhodně nebudeme omezovat pouze na lineární či kvadratické účelové funkce $f : \mathbb{R}^n \to \mathbb{R}$, ačkoli v takových případech budeme schopni k jednotlivým metodám získat mnohem více teoretických poznatků. Za zmíněnou obecnost však zaplatíme nemalou cenu v podobě nalézání pouze přibližných řešení, tj. v obecném případě budou všechny metody nekonečně krokovými iteračními procesy a my se budeme snažit zajistit alespoň jejich konvergenci k řešení (7.1). Klasické numerické metody jsou totiž pro úlohu (7.1) i případně pro úlohy nepodmíněné optimalizace (obvykle) založeny na následujícím základním principu konstrukce *minimalizující* posloupnosti:

> Při daném bodě $x_0 \in X$ (*počáteční aproximace*) chceme zkonstruovat takovou posloupnost $\{x_k\}_{k=0}^{\infty}$, že
>
> $$\lim_{k\to\infty} f(x_k) = f^* = \inf_{x\in\mathbb{R}^n} f(x).$$

Jednou ze základních otázek pro nás bude *globální/lokální konvergence* jednotlivých algoritmů, tj. závislost konvergence minimalizující posloupnosti na volbě počáteční aproximace x_0. Další důležitou charakteristikou iteračních procesů je *rychlost* jejich konvergence k limitní hodnotě, tj. máme-li minimalizující posloupnost $\{x_k\}_{k=0}^{\infty}$ splňující $x_k \in \mathbb{R}^n$ a současně $x_k \to x^*$, pak se *rychlost konvergence* měří vzhledem k *posloupnosti odchylek* (chyb) $\{e_k\}_{k=0}^{\infty}$, ve které máme $e_k \in [0,\infty)$ a $e_k \to 0$. Obvykle se volí

$$\begin{aligned}
&e_k := \| x_k - x^* \|, \quad \text{kde } x^* \text{ je limitou } \{x_k\}, \text{ což nemusí být řešením (7.1)},\\
&e_k := |f(x_k) - f(x^*)|, \quad \text{kde } x^* \text{ je opět limitou } \{x_k\},\\
&e_k := \ell_k, \quad \text{kde } \ell_k \text{ je délka tzv. } \textit{intervalu lokalizace minima}, \text{ viz později.}
\end{aligned}$$

Posloupnost $\{e_k\}$ je pak srovnávána s geometrickou posloupností $\{h_k\}$ se členy

$$h_k := q\beta^k,$$

kde $q > 0$ a $\beta \in (0,1)$, a případně také s posloupností $\{h_k\}$ se členy

$$h_k := q\beta^{p^k},$$

kde $q > 0$, $\beta \in (0,1)$ a $p > 1$. Proč právě tyto posloupnosti? Neexistuje žádný zásadní důvod pro volbu těchto dvou posloupností kromě toho, že reprezentují poměrně širokou třídu posloupností, která je zcela dostačující pro naše potřeby. Druhým důvodem může být třeba i to, že „to s nimi funguje“ a je to standardní přístup. Žádný jiný hlubokomyslný důvod v tom asi nenalezneme.

Jednotlivé algoritmy lze kategorizovat několika způsoby:

(i) *dle rychlosti/řádu konvergence* minimalizující posloupnosti, kterou tímto způsobem získáme (nebo případně odpovídající posloupnosti funkčních hodnot $\{f(x_k)\}$). Nicméně neměli bychom přecenit význam řádu konvergence a jeho využití při porovnávání jednotlivých metod. Tento tradiční přístup totiž dává pouze jakýsi návod k orientaci mezi metodami — *nic víc*! Vlastně se zdá, že ani neexistuje žádný čistě teoretický způsob k podrobné klasifikaci jednotlivých metod – vše závisí na konkrétních okolnostech a empirických zkušenostech. V praxi hraje velmi důležitou roli také např. *zaokrouhlování*, ale i některé další aspekty, které mohou ovlivnit užitečnost metody (zejména počet potřebných početních operací a jejich náročnost jako např. při výpočtu inverzní matice).

(ii) *dle základních požadavků na diferencovatelnost funkce* f, tj.

- *metoda nultého řádu:* vyžaduje pouze funkci f;

- *metoda prvního řádu:* vyžaduje diferencovatelnost funkce f;
- *metoda druhé řádu:* vyžaduje dvojnásobnou diferencovatelnost funkce f, atd.

Metody druhého řádu mají obvykle rychlejší konvergenci, ovšem nejsou vhodné např. ve chvíli, kdy výpočet druhých (parciálních) derivací je složitý nebo zatížený numerickou chybou. V případě experimentálního určování funkčních hodnot je asi nejvýhodnější použití metod nultého řádu.

(iii) *dle určování bodů* $x_1, x_2, \dots$, tj.

- nezávisí-li volba x_{k+1} na $x_0, \dots, x_k$, pak máme tzv. *pasivní algoritmus*;
- závisí-li volba x_{k+1} na $x_0, \dots, x_k$, pak máme tzv. *aktivní algoritmus*.

Dalším důležitým aspektem těchto algoritmů je jejich zastavení. V tomto ohledu můžeme najít (a vymyslet) širokou škálu pravidel, zejména

$$\|x_{k+1} - x_k\| < \varepsilon, \quad |f(x_{k+1}) - f(x_k)| < \varepsilon, \quad \|\operatorname{grad} f(x_k)\| < \varepsilon$$

nebo jejich kombinace či případně podobná pravidla zohledňující relativní změny, viz později.

7.1 Není konvergence jako konvergence...

Chceme-li porovnávat rychlosti konvergence minimalizující posloupnosti, musíme ji samozřejmě nejdříve umět „správně“ měřit. Proto začneme následující krátkou odbočkou podobně jako v [56, str. 12–16], viz také [404, Kapitola 9].

DEFINICE 7.1.1

Nechť jsou dány dvě posloupnosti $\{e_k\}_{k=0}^{\infty}$ a $\{h_k\}_{k=0}^{\infty}$ takové, že

$$e_k \in [0,\infty), \quad e_k \to 0 \quad \& \quad h_k \in [0,\infty), \quad h_k \to 0.$$

Řekneme, že posloupnost $\{e_k\}$ konverguje *rychleji* (*pomaleji*) než $\{h_k\}$, pokud existuje index $\tilde{k} \in \{0,1,2\dots\}$ takový, že

$$e_k \underset{(\geq)}{\leq} h_k \quad \text{pro všechna } k \in \{\tilde{k}, \tilde{k}+1, \dots\}.$$

Toto porovnávání samozřejmě nemusí být možné vždy, např. posloupnosti $\{e_k\}_{k=0}^{n} = \{\frac{1}{3^k}\}_{k=0}^{n}$ a $\{h_k\}_{k=0}^{n}$ se členy

$$h_k := \begin{cases} \frac{1}{2^k}, & \text{je-li } k \text{ liché}, \\ \frac{1}{4^k}, & \text{je-li } k \text{ sudé}, \end{cases}$$

splňují podmínku $e_k \to 0$ a $h_k \to 0$, ale nelze říci, že by některá z nich konvergovala rychleji.

První možnou klasifikaci dostaneme již avizovaným srovnáním s geometrickou posloupností.

DEFINICE 7.1.2

Nechť je dána posloupnost $\{e_k\}_{k=0}^{\infty}$ splňující $e_k \in [0,\infty)$ a $e_k \to 0$. Řekneme, že posloupnost $\{e_k\}$ konverguje

(i) *alespoň lineárně*[2] *s rychlostí* $\beta \in (0,1)$, pokud existuje $q > 0$ takové, že $\{e_k\}$ konverguje rychleji než geometrická posloupnost se členy tvaru $q\bar{\beta}^k$ pro libovolné $\bar{\beta} \in (\beta, 1)$;

(ii) *nejvýše lineárně s rychlostí* $\beta \in (0,1)$, pokud existuje $q > 0$ takové, že $\{e_k\}$ konverguje pomaleji než geometrická posloupnost se členy $q\bar{\beta}^k$, kde $q > 0$ a $\bar{\beta} \in (0, \beta)$;

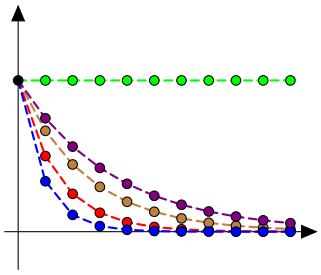
(iii) *lineárně s rychlostí* $\beta \in (0,1)$, pokud konverguje nejvýše a současně alespoň lineárně s rychlostí β;

(iv) *superlineárně* (*sublineárně*), pokud konverguje rychleji (pomaleji) než libovolná geometrická posloupnost se členy tvaru $q\beta^k$, kde $q > 0$ a $\beta \in (0,1)$.

[2] Abychom se vyhnuli neustálému opakování toho, že $\{e_k\}_{k=0}^{\infty}$ srovnáváme s geometrickou posloupností, hovoříme raději o lineární konvergenci místo *geometrické konvergence* (tj. *konvergenci srovnatelné s geometrickou posloupností*). Navíc tato změna se ukáže být velmi užitečná ve chvíli, kdy začneme danou posloupnost srovnávat i se „supergeometrickou“ posloupností $\left\{q\beta^{p^k}\right\}_{k=0}^{\infty}$.

Jaké hodnoty rychlosti β jsou lepší[3]? Máme-li dvě lineárně konvergentní posloupnosti s rychlostmi $0 < \beta_1 < \beta_2 < 0$, pak posloupnost s rychlostí β_1 konverguje rychleji, tj. menší hodnota β znamená rychlejší konvergenci, viz také Obrázek 7.1.

[3] A proč se vlastně vůbec omezujeme pouze na geometrické posloupnosti s $\beta \in (0,1)$? Protože posloupnost $\{q\beta^k\}_{k=0}^{\infty}$ je pro $\beta = 1$ konstantní a pro $\beta > 1$ dokonce divergentní, takže srovnávání s těmito posloupnostmi není příliš užitečné. Navíc posloupnost $\{e_k\}_{k=0}^{\infty}$ konverguje k nule, zatímco v případě lineární konvergence s rychlostí β máme podle Věty 7.1.4(vi) níže, že $e_{k+1} \approx \beta e_k$ pro velká k, takže pro $\beta > 1$ by posloupnost $\{e_k\}$ byla rostoucí.



Obrázek 7.1: Srovnání posloupností $\{q\beta^k\}_{k=0}^{\infty}$ pro $\beta = 1$ (zelená), $\beta = 3/4$ (fialová), $\beta = 2/3$ (hnědá), $\beta = 1/2$ (červená) a $\beta = 1/3$ (modrá).

Příklad 7.1.3

(i) Posloupnosti konvergující lineárně s rychlostí $\beta \in (0,1)$ jsou např.

$$q\beta^k, \quad q(\beta + 1/k)^k, \quad q(\beta - 1/k)^k, \quad q\beta^{k+1/k}, \qquad q > 0.$$

(ii) Je-li $0 < \beta_1 < \beta_2 < 1$ a

$$e_k := \begin{cases} \beta_1^\ell \beta_2^\ell, & k = 2\ell, \\ \beta_1^{\ell+1} \beta_2^\ell, & k = 2\ell + 1, \end{cases}$$

pak posloupnost $\{e_k\}_{k=0}^{\infty}$ konverguje alespoň lineárně s rychlostí β_2 a nejvýše lineárně s rychlostí β_1. Dokonce se dá ukázat, že konverguje lineárně s rychlostí $\sqrt{\beta_1 \beta_2}$, viz Větu 7.1.4.

(iii) Posloupnost $\{e_k\}_{k=0}^{\infty}$ se členy ve tvaru

$$e_k := \begin{cases} 1/2^\ell, & k = 2\ell, \\ 1/(2^\ell + 1), & k = 2\ell + 1. \end{cases}$$

konverguje alespoň sublineárně a nejvýše s rychlostí $1/2$. Při trochu přesnějším zkoumání zjistíme, že tato posloupnost konverguje lineárně s rychlostí $1/\sqrt{2}$.

(iv) Posloupnost $\{1/(k+1)\}_{k=0}^{\infty}$ konverguje sublineárně a každá posloupnost se členy tvaru $q\beta^{p^k}$ konverguje superlineárně pro libovolné $q > 0$, $\beta \in (0,1)$ a $p > 1$. Také např. posloupnost se členy $e_k := 1/k^k$ konverguje superlineárně.

Toto všechno lze snadno ukázat pomocí následujícího tvrzení.

Věta 7.1.4

Nechť je dána posloupnost $\{e_k\}_{k=0}^{\infty}$ splňující $e_k \in [0,\infty)$ a $e_k \to 0$. Potom platí následující tvrzení.

(i) Posloupnost $\{e_k\}$ konverguje alespoň lineárně s rychlostí $\beta \in (0,1)$ právě tehdy, když

$$\limsup_{k\to\infty} \sqrt[k]{e_k} \le \beta. \tag{7.1.1}$$

(ii) Posloupnost $\{e_k\}$ konverguje nejvýše lineárně s rychlostí $\beta \in (0,1)$ právě tehdy, když

$$\liminf_{k\to\infty} \sqrt[k]{e_k} \ge \beta.$$

(iii) Posloupnost $\{e_k\}$ konverguje lineárně s rychlostí $\beta \in (0,1)$ právě tehdy, když

$$\lim_{k\to\infty} \sqrt[k]{e_k} = \beta.$$

(iv) Pokud posloupnost $\{e_k\}$ konverguje rychleji (pomaleji) než nějaká geometrická posloupnost se členy $q\beta^k$, $q > 0$ a $\beta \in (0,1)$, pak konverguje alespoň (nejvýše) lineárně s rychlostí β.

(v) Je-li $e_k > 0$ pro všechna $k \in \mathbb{N}_0$, definujme

$$\beta_1 := \liminf_{k\to\infty} \frac{e_{k+1}}{e_k} \quad \& \quad \beta_2 := \limsup_{k\to\infty} \frac{e_{k+1}}{e_k}.$$

Jestliže $0 < \beta_1 < \beta_2 < 1$, pak $\{e_k\}$ konverguje alespoň lineárně s rychlostí β_1 a nejvýše lineárně s rychlostí β_2.

VĚTA 7.1.4

(vi) Je-li $e_k > 0$ pro všechna $k \in \mathbb{N}_0$, položme

$$\beta := \lim_{k\to\infty} \frac{e_{k+1}}{e_k}.$$

Pak v případě $0 < \beta < 1$ konverguje posloupnost $\{e_k\}$ lineárně s rychlostí β. Je-li $\beta = 0$, pak konverguje superlineárně. V případě $\beta = 1$ se jedná o sublineární konvergenci.

Důkaz. Celé tvrzení plyne bezprostředně z definic lim sup a lim inf. Např. v první části využijeme ekvivalenci nerovnosti (7.1.1) s faktem, že pro každé $\varepsilon \in (0, 1-\beta)$ existuje $n_0 \in \mathbb{N}$ takové, že pro všechna $n \geq n_0$ platí $e_k \leq (\beta+\varepsilon)^k$, tj. $e_k \leq \bar{\beta}^k$ pro libovolné $\bar{\beta} \in (\beta, 1)$. ■

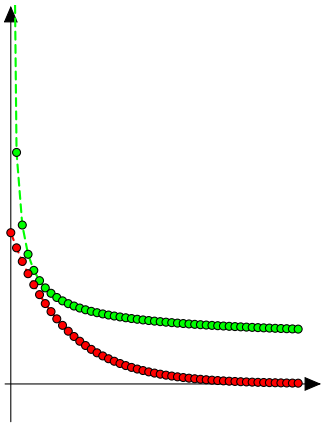
Poznámka 7.1.5. Jelikož velmi dobře víme, že platí

$$\liminf_{k\to\infty} \frac{e_{k+1}}{e_k} \leq \liminf_{k\to\infty} \sqrt[k]{e_k} \leq \limsup_{k\to\infty} \sqrt[k]{e_k} \leq \limsup_{k\to\infty} \frac{e_{k+1}}{e_k},$$

mohou se rychlosti lineární konvergence, které obdržíme s využitím předchozí věty, lišit. Proto se v souvislosti s limitou (případně lim sup nebo lim inf) podílu $\frac{e_{k+1}}{e_k}$ se také hovoří o *Q-lineární konvergenci*[4], zatímco pro limitu (či lim sup nebo lim inf) výrazu $\sqrt[k]{e_k}$ se hovoří o *R-lineární konvergenci*[5]. My toto rozlišování ale uplatňovat *nebudeme*. Navíc v případě, kdy dokonce existuje $\lim_{k\to\infty} \frac{e_{k+1}}{e_k}$, všechny tyto hodnoty zjevně splývají. ▲

[4] Angl. *quotient-linear convergence.*

[5] Angl. *root-linear convergence.*



Obrázek 7.2: Srovnání posloupností $\{2\,(9/10)^k\}_{k=0}^{\infty}$ a $\{1/(\ln(\ln k+3))\}_{k=0}^{\infty}$.

Uvidíme, že většina algoritmů bude konvergovat lineárně a některé dokonce superlineárně. Pokud β *není příliš blízko* 1, lze považovat lineární konvergenci za dostatečnou. Algoritmy se sublineární rychlostí konvergence nejsou obvykle uvažovány (a tak to s jednou výjimkou učiníme i my), neboť nejsou z praktického hlediska příliš efektivní – konvergují příliš pomalu jako např. $e_k = 1/(\ln(\ln k+3))$, viz také Obrázek 7.2. Na druhou stranu je užitečné si superlineární konvergenci klasifikovat ještě podrobněji.

DEFINICE 7.1.6

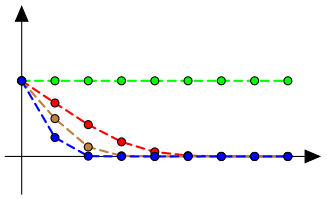
Nechť daná posloupnost $\{e_k\}_{k=0}^{\infty}$ splňuje $e_k \in [0,\infty)$ a $e_k \to 0$, přičemž konvergence je superlineární. Potom $\{e_k\}$ konverguje

(i) *alespoň superlineárně s řádem* $p > 1$, pokud konverguje rychleji než všechny posloupnosti se členy tvaru $q\beta^{\bar{p}^k}$, kde $q > 0$, $\beta \in (0,1)$ a $\bar{p} \in (1,p)$;

(ii) *nejvýše superlineárně s řádem* $p > 1$, pokud konverguje pomaleji než všechny posloupnosti se členy tvaru $q\beta^{\bar{p}^k}$, kde $q > 0$, $\beta \in (0,1)$ a $\bar{p} \in (p,\infty)$;

(iii) *superlineárně s řádem* $p > 1$, pokud konverguje nejvýše a současně alespoň superlineárně s řádem p;

(iv) *superlineárně s řádem* $p = 1$, pokud konverguje pomaleji než všechny posloupnosti se členy tvaru $q\beta^{\bar{p}^k}$, kde $q > 0$, $\beta \in (0,1)$ a $\bar{p} \in (1,\infty)$.



Obrázek 7.3: Srovnání posloupností $\{q\,\beta^{p^k}\}_{k=0}^{\infty}$ pro $\beta = 1/2$ a $p = 1$ (zelená), $p = 3/2$ (červená), $p = 2$ (hnědá), $p = 3$ (modrá).

Oproti rostoucí rychlosti lineární konvergence s klesající hodnotou β znamená tentokrát rychlejší konvergenci větší hodnota řádu p, viz Obrázek 7.3. Analogicky jako v případě Věty 7.1.4 můžeme dokázat následující tvrzení.

VĚTA 7.1.7

Nechť je dána posloupnost $\{e_k\}_{k=0}^{\infty}$ splňující $e_k \in [0,\infty)$ a $e_k \to 0$. Potom platí následující tvrzení.

(i) Posloupnost $\{e_k\}$ konverguje alespoň superlineárně s řádem $p > 1$ právě tehdy, když

$$\lim_{k\to\infty} e_k^{1/\bar{p}^k} = 0 \quad \text{pro všechna } \bar{p} \in (1,p).$$

VĚTA 7.1.7

(ii) Posloupnost $\{e_k\}$ konverguje nejvýše superlineárně s řádem $p > 1$ právě tehdy, když

$$\lim_{k\to\infty} e_k^{1/\bar{p}^k} = 1 \quad \text{pro všechna } \bar{p} \in (p, \infty).$$

(iii) Pokud $\{e_k\}$ konverguje rychleji/pomaleji než nějaká posloupnost se členy ve tvaru $q\beta^{p^k}$ pro nějaké $q > 0$, $\beta \in (0,1)$ a $p > 1$, pak posloupnost $\{e_k\}$ konverguje alespoň/nejvýše superlineárně s řádem p.

(iv) Je-li $e_k > 0$ pro všechna $k \in \mathbb{N}_0$ a pro nějaké $p > 1$ platí

$$\limsup_{k\to\infty} \frac{e_{k+1}}{e_k^p} < \infty,$$

pak posloupnost $\{e_k\}$ konverguje alespoň superlineárně s řádem p.

(v) Je-li $e_k > 0$ pro všechna $k \in \mathbb{N}_0$ a pro nějaké $p > 1$ platí

$$\liminf_{k\to\infty} \frac{e_{k+1}}{e_k^p} > 0,$$

pak posloupnost $\{e_k\}$ konverguje nejvýše superlineárně s řádem p.

Označení *superlineární* se obvykle vynechává a hovoří se pouze o *řádu konvergence*. Zejména v případě $p = 2$ se hovoří o *kvadratické konvergenci* a pro $p = 3$ jde o *kubickou konvergenci*.

Poznámka 7.1.8. Jak určit řád a rychlost konvergence posloupnosti $\{e_k\}$? Je-li $e_k \neq 0$ pro všechna[6] k, pak nejdříve vypočteme hodnotu

$$\gamma := \limsup_{k\to\infty} \frac{e_{k+1}}{e_k}.$$

[6] Nebo lépe: pro všechna k dostatečně velká, protože nás zajímá pouze limitní/asymptotické chování dané posloupnosti.

- Je-li $\gamma \in (0,1)$, pak konvergence je lineární a její rychlost je nejvýše γ.
- Je-li $\gamma = 1$, je konvergence sublineární.
- Je-li $\gamma = 0$, pak konvergence je superlineární a musíme určit její řád. Pokud pro nějaké $q \geq 1$ je $\limsup_{k\to\infty} \frac{e_{k+1}}{e_k^q} = 0$, pak existuje právě jeden exponent p takový, že

$$\limsup_{k\to\infty} \frac{e_{k+1}}{e_k^q} = \begin{cases} 0, & q < p, \\ \alpha, & q = p, \\ \infty, & q > p, \end{cases}$$

přičemž nevylučujeme hodnoty $\alpha \in \{0, \infty\}$. Právě tento exponent p, kde dochází ke změně z vlastní na nevlastní hodnotu, určuje nejnižší možný řád konvergence dané posloupnosti. Jinými slovy, řád superlineární konvergence posloupnosti $\{e_k\}$ je alespoň

$$p := \sup\left\{q \geq 1 \mid \limsup_{k\to\infty} e_{k+1}/e_k^q < \infty\right\}.$$

Obzvláště v případě existence samotné limity lze řád vypočítat i jako

$$p = \lim_{k\to\infty} \frac{\ln e_{k+1}}{\ln e_k} \tag{7.1.2}$$

a platí

$$e_{k+1} \approx \alpha e_k^p \quad \text{neboli} \quad e_{\ell+1} = \alpha e_\ell^p,\ e_{\ell+2} = \alpha^{p+1} e_\ell^{p^2},\ e_{\ell+3} = \alpha^{p(p+1)+1} e_\ell^{p^3}, \ldots$$

pro dostatečně velké ℓ, přičemž číslo α zde hraje roli tzv. *asymptotické odchylky*.

Analogické úvahy lze provést také pro e_k^{1/p^k}. Zejména existuje-li $p \geq 1$ takové, že

$$\lim_{k\to\infty} e_k^{1/q^k} = \begin{cases} 0, & q < p, \\ \alpha, & q = p, \\ 1, & q > p, \end{cases}$$

pak p je řád superlineární konvergence, přičemž hodnoty $\alpha \in \{0,1\}$ jsou povoleny. Jinými slovy, řád superlineární konvergence je roven[7]

$$p := \sup\{q \geq 1 \mid \lim_{k\to\infty} e_k^{1/q^k} < 1\}$$

a v takovém případě je

$$e_k \approx \alpha^{p^k} \quad \text{pro dostatečně velká } k.$$

Podobně jako v případě určování rychlosti lineární konvergence je namístě se ptát, zda tyto přístupy splývají. Obecně samozřejmě nikoli, jak ukazuje např. posloupnost $\{e_k\}$ se členy

$$e_k := \begin{cases} (a/2)^{p^k}, & k \text{ sudé}, \\ (a^{q/p})^{p^k}, & k \text{ liché}, \end{cases}$$

kde $0 < a < 1$ a $1 < q < p$ jsou libovolná. Potom „odmocninový“ přístup dává řád konvergence p, neboť $(a/2)^{p^k/p^k} = a/2 < 1$ a $(a^{p/q})^{p^k/p^k} = a^{p/q} < 1$. Avšak „podílový“ přístup dává, že řád je nejvýše roven q, neboť nerovnost $a^{p^2} < a^{q^2}$ dává

$$\lim_{k\to\infty} \frac{(a/2)^{p^{k+1}}}{(a^{q/p})^{p^k q}} = \lim_{k\to\infty} \left(\frac{a^{p^2}}{a^{q^2}}\right)^{p^{k-1}} \frac{1}{2^{p^{k+1}}} = 0 \quad \text{a} \quad \lim_{k\to\infty} \frac{(a^{q/p})^{p^{k+1}}}{(a/2)^{p^k q}} = \lim_{k\to\infty} \frac{a^{q\,p^k}}{a^{p^k q}} 2^{p^k q} = \infty,$$

takže $\liminf_{k\to\infty} \frac{e_{k+1}}{e_k^q} = 0$. Nicméně v případě existence $0 < \lim_{k\to\infty} e_{k+1}/e_k^p < \infty$ pro nějaké $p \geq 1$ se obě hodnoty musí shodovat, což je obvyklý jev v iteračních procesech a pro naše další úvahy zcela *postačující*. Podobně k (7.1.2) v takovém případě pak platí, že řád konvergence je roven[8]

$$p = \lim_{k\to\infty} \sqrt[k]{|\ln e_k|}, \qquad \blacktriangle$$

[7] Tato rovnost bývá někdy i uváděna jako definice.

[8] Tato rovnost bývá někdy i uváděna jako definice.

Příklad 7.1.9

(a) Uvažme posloupnost $\{e_k\}$ se členy

$$e_k = a^k$$

pro libovolné $0 < a < 1$. Pak

$$\lim_{k\to\infty} \frac{a^{k+1}}{a^{kp}} = \lim_{k\to\infty} a^{k(1-p)+1} = \begin{cases} 0, & p < 1, \\ a, & p = 1, \\ \infty, & p > 1, \end{cases}$$

tj. posloupnost konverguje superlineárně s řádem $p = 1$ a asymptotickou odchylkou a.

(b) Podobně pro

$$e_k := A\,a^{p^k}$$

s libovolnými $0 < a < 1$ a $A > 0$, dostaneme

$$\lim_{k\to\infty} \frac{A\,a^{p^{k+1}}}{\left(A\,a^{p^k}\right)^p} = A^{1-p} \lim_{k\to\infty} \frac{a^{p^{k+1}}}{a^{p^{k+1}}} = A^{1-p},$$

takže posloupnost konverguje superlineárně s řádem p a asymptotickou odchylkou A^{1-p}, takže platí $e_{k+1} = A^{1-p}\,e_k^p$.

(c) Konečně uvažme posloupnost $\{e_k\}$ se členy

$$e_k := (1/k)^k.$$

Pak platí

$$p = \lim_{k\to\infty} \frac{\ln(1/(k+1))^{k+1}}{\ln(1/k)^k} = \lim_{k\to\infty} \frac{-(k+1)\ln(k+1)}{-k\ln k} \overset{\text{l'H.p.}}{=} \lim_{k\to\infty} \frac{\ln(k+1)+1}{\ln k+1} =$$
$$\overset{\text{l'H.p.}}{=} \lim_{k\to\infty} \frac{1/(k+1)}{1/k} = 1.$$

Je to skutečně řád? Ano, neboť

$$\lim_{k\to\infty}\frac{e_{k+1}}{e_k^p}=\lim_{k\to\infty}\frac{1/(k+1)^{k+1}}{1/k^{kp}}=\lim_{k\to\infty}\frac{k^{kp}}{(k+1)^{k+1}}=\lim_{k\to\infty}\frac{k^{kp}\,k^{-k-1}}{(1+1/k)^k\,(1+1/k)}=$$

$$=\lim_{k\to\infty}\frac{k^{k(p-1)-1}}{\underbrace{(1+1/k)^k\,(1+1/k)}_{\to \mathrm{e}}}=\begin{cases}0, & p<1,\\ 0, & p=1,\\ \infty, & p>1.\end{cases}$$

Posloupnost tedy konverguje superlineárně s řádem $p=1$.

7.2 Numerické metody nepodmíněné optimalizace v $\mathbb{R}$

Před budováním numerických metod pro řešení úlohy (7.1) se zaměříme na jednodušší úlohu

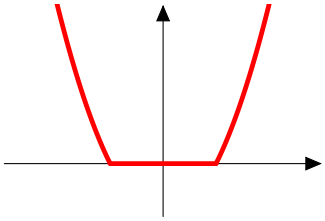
$$f(x)\to\min,\quad x\in I:=[a,b]. \tag{7.2.1}$$

Proč? Protože některé metody v $\mathbb{R}^n$ jsou založené na tom, že v jistém kroku musíme nalézt řešení právě takové úlohy. Teoreticky sice víme, že v případě dostatečně „hezké“ funkce stačí vyřešit rovnici $f'(x)=0$ a odpovídající funkční hodnotu[9] porovnat s $f(a)$ a $f(b)$, jenže to nemusí být zrovna snadné, jak můžeme vidět např. na funkci $f(x)=(x-1)^2+\mathrm{e}^{x-1}$, která nás přivede k transcendentní rovnici $2(x-1)+\mathrm{e}^{x-1}=0$.

My se omezíme pouze na jistou třídu funkcí, čímž ono výše zmíněné „hezké chování“ funkce f vynutíme, a tudíž budeme mít zaručenu jednoznačnou řešitelnost úlohy (7.2.1) či nalezení hodnoty $f^*=\inf_{x\in I}f(x)$.

[9] Nebo hodnoty...

[10] V případě spojité funkce f na I to bude dokonce minimum.



Obrázek 7.4: Ukázka konvexní funkce, která není unimodální.

DEFINICE 7.2.1

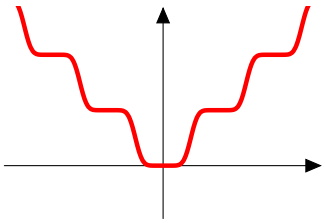
Nechť je dána funkce $f: I\to\mathbb{R}$. Řekneme, že f je *unimodální na I*, jestliže existuje $x^*\in I$ takové, že

$$f(x_1)>f(x_2)\quad\text{pro libovolná } x_1,x_2\in I \text{ splňující } x^*>x_2>x_1,$$
$$f(x_1)<f(x_2)\quad\text{pro libovolná } x_1,x_2\in I \text{ splňující } x^*<x_1<x_2.$$

[11] V této poznámce jsme si definovali pouze *kvazikonvexní* funkce. *Ostrá* kvazikonvexnost znamená, že nerovnost (6.2.8) je ostrá pro všechna $x_1\neq x_2$ a $\lambda\in(0,1)$.

Předchozí definice jinými slovy vyžaduje, aby existoval bod x^* takový, že funkce f je ostře rostoucí napravo od x^* a ostře klesající nalevo od x^*. Zároveň je zcela irelevantní hodnota $f(x^*)$. Taková funkce pak v daném intervalu I nabývá svého infima[10] právě v bodě x^*.

Je-li funkce f *ostře konvexní na I*, pak je nutně i unimodální, neboť pro takové funkce vždy existuje x^* s vlastnostmi z Definice 7.2.1. Buď je to (jediný) stacionární bod ležící uvnitř I nebo některý z krajních bodů intervalu I. Avšak pouhá konvexnost pro unimodalitu funkce f nestačí, viz např. Obrázek 7.4. Na druhou stranu, unimodalita rozhodně neimplikuje ostrou konvexnost či konvexnost, viz např. Obrázek 7.5. Unimodalita je totiž velmi úzce spojena s kvazikonvexností, jak ukazuje následující tvrzení, viz Poznámku[11] 6.2.11.



Obrázek 7.5: Ukázka unimodální funkce, která není konvexní.

LEMMA 7.2.2

Nechť je dána funkce $f: I\to\mathbb{R}$, která na I nabývá svého jediného minima v x^*. Pak f je unimodální na I právě tehdy, když je ostře[12] kvazikonvexní na I.

Důkaz. Je-li f ostře kvazikonvexní a $x_1<x_2<x^*$, pak nutně $f(x_2)<\max\{f(x_1),f(x^*)\}=f(x_1)$. Podobně pro $x_2>x_1>x^*$ máme $f(x_1)<\max\{f(x_2),f(x^*)\}=f(x_2)$, tj. f je unimodální. Naopak, nechť f je unimodální a pro libovolná $\lambda\in(0,1)$ a $x_1,x_2\in I$ splňující $x_1<x_2$ položme $x_0:=\lambda x_1+(1-\lambda)x_2$. Pak z nerovností $x_1<x_0<x^*$ plyne $f(x_0)<f(x_1)\le\max\{f(x_1),f(x_2)\}$ a podobně z $x_2>x_0>x^*$ dostaneme $f(x_0)<f(x_2)\le\max\{f(x_1),f(x_2)\}$, což ukazuje, že f je vskutku ostře kvazikonvexní. ■

[12] Kdybychom v definici unimodality povolili neostré nerovnosti mezi funkčními hodnotami, platilo by analogické tvrzení pro „pouze“ *kvazikonvexní* funkce. Pro více informací o kvazikonvexních a unimodálních funkcích viz [22, Kapitola 3].

Unimodalita funkce f je velmi obecná vlastnost, která nezaručuje ani diferencovatelnost na I a dokonce ani spojitost, viz Obrázky 7.6 a 7.7.

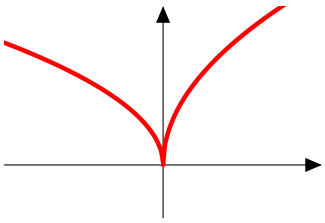
Přímo z Definice 7.2.1 dostaneme následující tvrzení, které nám usnadní „lokalizaci“ bodu x^*.

LEMMA 7.2.3

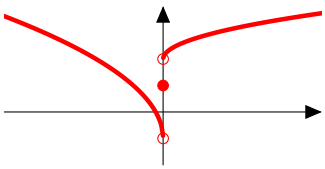
Nechť je dána unimodální funkce $f : I \to \mathbb{R}$ a nechť $x_1, x_2 \in I$ jsou taková, že $x_1 < x_2$. Potom platí:

(i) je-li $f(x_1) \le f(x_2)$, pak $x^* \le x_2$;

(ii) je-li $f(x_1) \ge f(x_2)$, pak $x^* \ge x_1$.



Obrázek 7.6: Ukázka unimodální funkce, která není diferencovatelná.



Obrázek 7.7: Ukázka unimodální funkce, která není spojitá.

Poznámka 7.2.4. Naše využití předchozího tvrzení bude především následující. Bude dána unimodální funkce f, jakýsi interval $[a_k, b_k]$ obsahující x^* a v něm dva body $a_k < \lambda_{k+1} < \mu_{k+1} < b_k$. Pak podle Lemma 7.2.3 v případě

(i) $f(\lambda_{k+1}) \le f(\mu_{k+1})$ je $x^* \in [a_k, \mu_{k+1}] =: [a_{k+1}, b_{k+1}]$, tj. pokud funkční hodnoty rostou (neboli funkční hodnoty „zachovávají uspořádání“), můžeme ignorovat body za μ_{k+1}.);

(ii) $f(\lambda_{k+1}) \ge f(\mu_{k+1})$ je $x^* \in [\lambda_{k+1}, b_k] =: [a_{k+1}, b_{k+1}]$, tj. pokud funkční hodnoty klesají (neboli funkční hodnoty „otáčí uspořádání“), můžeme ignorovat body před λ_{k+1}.)

Takové algoritmy patří do skupiny *Divide-and-Conquer* stejně jako Metoda větvení a mezí v celočíselném programování[13], i když tentokrát bude rozdělování na jednodušší problémy probíhat trochu jiným způsobem[14]. Ještě dodejme, že v případě $f(\lambda_k) = f(\mu_k)$ si můžeme „svobodně“ zvolit kterýkoli z těchto intervalů – my budeme *vždy* uplatňovat druhou volbu. Ve skutečnosti je sice $x^* \in [\lambda_{k+1}, \mu_{k+1}]$, avšak tato „singulární“ volba by nám zbytečně zkomplikovala popis dalších vlastností studovaných metod a po několika krocích algoritmu by už užitek z takového dodatečného zkrácení byl nepodstatný. Navíc taková rovnost funkčních hodnot je z praktického pohledu téměř nemožná při jakkoli sofistikovaném způsobu nalezení bodů λ_{k+1}, μ_{k+1}, obzvláště jsou-li výpočty funkčních hodnot prováděny numericky s omezeným počtem desetinných míst. ▲

My naše úvahy směřujeme pouze na kompaktní interval $I = [a, b]$. Ovšem v případě unimodální funkce na intervalu ohraničeném pouze zdola/shora či dokonce neohraničeném lze pomocí Lemmatu 7.2.3 zkonstruovat konečně krokový proces, s jehož pomocí nalezneme kompaktní interval obsahující x^*: vyčíslíme funkci f v bodech

$$x_1 \quad \text{a} \quad x_2 := x_1 + h_1$$

pro nějaké $h_1 > 0$. Nechť např. $f(x_1) > f(x_2)$. Je-li interval I ohraničený shora číslem c, pak nutně $x^* \in [x_1, c]$. V opačném případě musíme napočítat funkční hodnoty $f(x_i)$ pro $x_i := x_{i-1} + h_{i-1}$ a[15] $i \in \{3, 4, \ldots\}$ tak dlouho, dokud nenalezneme[16] x_j takové, že $f(x_j) \ge f(x_{j-1})$. Pak podle Lemmatu 7.2.3 je $x^* \in [x_{j-2}, x_j]$. V případě $f(x_1) < f(x_2)$ musíme podobné úvahy aplikovat na interval vlevo od x_1. Podobně jako v předchozí poznámce platí, že v případě $f(x_1) = f(x_2)$ je $x^* \in [x_1, x_2]$.

Příklad 7.2.5

Uvažme např. funkci $f(x) = (x-2)^2$ na intervalu $I = [0, \infty)$ s $x_1 := 0$ a $h_i^{[1]} \equiv 0{,}7$ a $h_i^{[2]} \equiv 0{,}8$. Pak v případě první volby máme

$$f(0) = 4 > f(0{,}7) = 1{,}69 > f(1{,}4) = 0{,}36 > f(2{,}1) = 0{,}01 \le f(2{,}8) = 0{,}64,$$

tj. $x^* \in [1{,}4; 2{,}8]$, zatímco pro druhou volbu dostaneme

$$f(0) = 4 > f(0{,}8) = 1{,}44 > f(1{,}6) = 0{,}16 \le f(2{,}4) = 0{,}16,$$

tj. $x^* \in [0{,}8; 2{,}4]$. Víme, že $x^* = 2$, a tento příklad velmi dobře ilustruje, že je skutečně nutné brát interval $[x_{j-2}, x_j]$, neboť v prvním případě vyšlo $x^* \in [x_{j-2}, x_{j-1}]$ a ve druhém $[x_{j-1}, x_j]$.

[13] Viz Podkapitolu 3.4.

[14] A obejdeme se i bez zakreslován rozhodovacího stromu.

[15] Je možné volit $h_i \equiv$ konst., avšak lepší je jistý kompromis mezi příliš „malými“ (sice dostaneme přesnější interval obsahující x^*, ale může to být zdlouhavé) a „velkými“ (rychlejší výpočet, ale výsledný interval může být velmi dlouhý, takže následný odhad x^* může být zatížen velkou nepřesností) kroky, např. $h_{i+1} = 2h_i$ apod.

[16] Vzhledem k unimodálnosti f na I musí takový bod v $[x_1, \infty)$ existovat, neboť jinak by f byla klesající do $-\infty$.

Nyní si již představíme několik numerických metod, s jejichž pomocí se budeme snažit najít přibližné řešení úlohy (7.2.1) s využitím pouze N*-tice funkčních hodnot*, což je *povolený počet vyčíslení funkce* f. *Přesnost* těchto metod by mohla být dána hodnotou chyby zvolené aproximace $|\overline{x} - x^*|$, kde x^* je (přesné) řešení úlohy (7.2.1) a $\overline{x}$ je jeho nalezená aproximace. Avšak toto by byl jen ryze teoretický pohled na uvažovanou metodu, neboť v praxi hodnotu x^* samozřejmě neznáme. Místo toho budeme za přesnost považovat jakýsi *apriorní odhad maximální chyby*, tj. maximální možnou chybu, které si použitím zvolené metody za daných podmínek dopustíme. Skutečná hodnota chyby ale může být cokoli mezi nulou a tímto odhadem. Její konkrétní hodnota už je spíše věcí náhody a závisí na konkrétní poloze budu x^*, tj. drobným posunem funkce f doprava/doleva by mohl výpočet probíhat úplně stejně, ale přesná chyba by byla menší/větší. Jediný způsob, jak ovlivnit hodnotu chyby, je volit parametry uvažované metody tak, aby se apriorní odhad byl dostatečně malý.

WWW

Metody z Podkapitol 7.2.1–7.2.4 jsou dostupné ve formě Mapletu na webu:
☞ https://ln.math.muni.cz/pz/opt/MapletR

7.2.1 Metoda prostého dělení intervalu[20]

[20] *Exhaustive Search Method* nebo *Brute-Force Search Method*

Metoda prostého dělení intervalu (dále už obvykle jen *MPD*) rozhodně není příliš intelektuálně náročným nástrojem a vlastně ani efektivní – vlastně jde spíše o užití *hrubé síly*, avšak může být velmi užitečná ve chvíli, kdy všechny ostatní sofistikované nástroje selžou. Ovšem je ideální pro začátek našich úvah. V závislosti na paritě N určíme dělící body intervalu I. Je-li N liché, pak bereme

$$x_i := a + \frac{b-a}{N+1}\, i, \quad i = 1, \ldots, N = 2k-1,$$

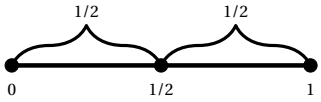
zatímco pro sudé N volíme

$$x_{2i} := a + \frac{b-a}{k+1}\, i \quad \& \quad x_{2i-1} := x_{2i} - \delta, \quad i = 1, \ldots, N = 2k,$$

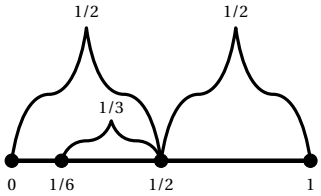
kde δ je vhodné malé kladné číslo. Co znamená „vhodné“? V případě lichého N rozdělíme interval I na $N+1$ intervalů délky $(b-a)/(N+1)$. V případě $N = 2k$ dostaneme také $N+1$ intervalů, které ovšem již nemusí být stejně dlouhé. Stejně dlouhé budou pouze intervaly $[a, x_2]$, $[x_2, x_4]$,... a $[x_{2k}, b]$, tj. budeme mít postupně intervaly délek $(b-a)/(k+1) - \delta$, δ, $(b-a)/(k+1) - \delta$, δ,..., $(b-a)/(k+1)$. Zvolíme-li $\delta := (b-a)/(k+1)$, pak

$$x_{2i-1} = x_{2i} - \delta = a + \frac{b-a}{k+1}\, i - \frac{b-a}{k+1} = x_{2i-2},$$

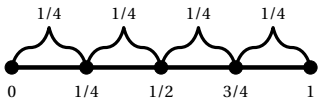
tj. všechny dělící intervaly budou stejné dlouhé – avšak bude jich pouze $k+1$! Toto je také největší možná hodnota δ, která dává smysl, neboť pro $\delta > (b-a)/(k+1)$ bude $x_{2i-1} < x_{2i-2}$. Zejména $x_1 < a$, tj. opustíme interval $[a, b]$ ↯. V případě $\delta = 0$ budeme mít opět pouze $k+1$ intervalů, neboť v takovém případě bude $x_{2i-1} = x_{2i}$, tj. volba $\delta \in \{0, (b-a)/(k+1)\}$ dává ekvidistantní dělení, při kterém využijeme pouze polovinu povolených vyčíslení. Má proto smysl omezit se pouze na hodnoty $\delta \in \big(0, (b-a)/(k+1)\big)$.



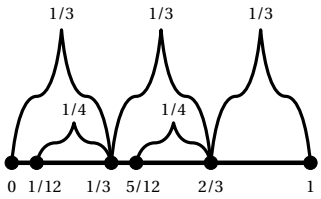
Obrázek 7.8: Dělící body pro MPD v případě $N = 1$.



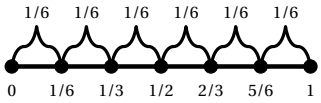
Obrázek 7.9: Dělící body pro MPD v případě $N = 2$ a $\delta = 1/3$.



Obrázek 7.10: Dělící body pro MPD v případě $N = 3$.



Obrázek 7.11: Dělící body pro MPD v případě $N = 4$ a $\delta = 1/4$.



Obrázek 7.12: Dělící body pro MPD v případě $N = 5$.

Příklad 7.2.6

Uvažme interval $I = [0, 1]$. Pak pro

(i) $N = 1$ máme $x_1 = 0 + (1-0)/2 = 1/2$;

(ii) $N = 2$ a $\delta = 1/3$ bude $x_2 = 0 + (1-0)/2 = 1/2$ a $x_1 = x_2 - \delta = 1/6$;

(iii) $N = 3$ získáme $x_1 = 0 + (1-0)4 = 1/4$, $x_2 = 1/2$ a $x_3 = 3/4$;

(iv) $N = 4$ a $\delta = 1/4$ dostaneme $x_2 = 0 + (1-0)/3 = 1/3$ a $x_4 = 2/3$, zatímco $x_1 = x_2 - \delta = 1/12$ a $x_3 = x_4 - \delta = 5/12$;

(v) $N = 5$ jsou $x_1 = 0 + (1-0)/6 = 1/6$, $x_2 = 2/6$, $x_3 = 3/6$, $x_4 = 4/6$ a $x_5 = 5/6$,
viz Obrázky 7.8–7.12.

Následně vyčíslíme $f(x_1), \dots, f(x_N)$ (v případě $N = 2k$ a $\delta \in \{0, (b-a)/(k+1)\}$ máme de facto jenom k vyčíslení). Nechť v x_j nastává nejmenší funkční hodnota, tj.

$$f(x_j) = \min_{1 \le i \le N} f(x_i).$$

Položíme-li $x_0 := a$ a $x_{N+1} := b$, pak z Lemma 7.2.3 plyne, že $x^* \in [x_{j-1}, x_{j+1}]$. Tento interval[21] nazveme *intervalem lokalizace minima* (dále obvykle jen *ILM*) a za aproximaci x^* vezmeme jeho střed, tj.

$$\overline{x} := \frac{x_{j-1} + x_{j+1}}{2}.$$

Ještě poznamenejme, že v případě $f(x_i) = f(x_{i+1})$ pro nějaké $i \in \{1, \dots, N\}$ si můžeme „zcela svobodně“ zvolit[22], zda vezmeme $j = i$ nebo $j = i+1$. Pro délku ILM pak vždy platí

$$\ell_N := \max_{1 \le i \le N} (x_{i+1} - x_{i-1}) = \begin{cases} 2\frac{b-a}{N+1}, & N = 2k-1, \\ \frac{b-a}{N/2+1} + \delta, & N = 2k, \end{cases}$$

což můžeme vyjádřit jako[23] $\frac{b-a}{\lceil (N+1)/2 \rceil} + \delta$, přičemž pro liché N bereme $\delta = 0$. Polovina tohoto čísla pak udává *apriorní odhad maximální chyby* neboli *přesnost* MPD. Dodejme, že skutečná hodnota se

Jaký je řád/rychlost konvergence této metody? Vezmeme-li posloupnost odchylek $\{e_N\}$ určenou délkami ILM pro rostoucí počty vyčíslení, tj.

$$e_N := \begin{cases} \ell_N = 2\frac{b-a}{N+1}, & N \text{ je liché}, \\ \ell_N - \delta = \frac{b-a}{N/2+1}, & N \text{ je sudé}. \end{cases}$$

Pak $e_N \to 0$ pro $N \to \infty$ a

$$\lim_{N\to\infty} \frac{e_{N+1}}{e_N} = \begin{cases} \lim_{N\to\infty} \frac{2\frac{b-a}{N+2}}{\frac{b-a}{N/2+1}} = \lim_{N\to\infty} \frac{N+2}{N+2} = 1, \\ \lim_{N\to\infty} \frac{\frac{b-a}{(N+1)/2+1}}{2\frac{b-a}{N+1}} = \lim_{N\to\infty} \frac{N+1}{N+3} = 1, \end{cases}$$

což znamená, že konvergence je sublineární, tj. posloupnost délek se chová jako $1/N$. Navíc tento algoritmus (jako jediný, který si ukážeme) je pasivní, neboť volba bodů $x_1, \dots, x_N$ je určena pouze hodnotou N a případně i δ, ovšem ne funkčními hodnotami funkce f.

Poznámka 7.2.7. Jak už jsme se zmínili, v případě $N = 2k-1$ je dělení intervalu I ekvidistantní, zatímco pro $N = 2k$ a $\delta \in \big(0, (b-a)/(k+1)\big)$ již nikoli. Proč ale neuvažujeme ekvidistantní dělení i při sudém počtu dělení? Protože při vhodně zvoleném δ lze námi zvoleným postupem dosáhnout kratšího ILM. Vskutku, při ekvidistantním dělení by délka ILM byla rovna $2(b-a)/(N+1)$, což odpovídá maximální možné délce ILM $(b-a)/(N/2+1) + \delta$ získané pomocí MPD pro hodnotu $\delta = 2\frac{b-a}{(N+1)(N+2)} < \frac{b-a}{N/2+1}$. Proto omezíme-li se dokonce na

$$\delta \in \Big(0, \frac{2(b-a)}{(N+1)(N+2)}\Big), \tag{7.2.2}$$

dosáhneme v MPD se sudým počtem vyčíslení jistě kratšího ILM než při ekvidistantním dělení. ▲

Příklad 7.2.8

Uvažme funkci $f(x) = 7x^2 - 6x + 2$ na intervalu $I = [0,1]$. S chybou nejvýše $\varepsilon = 0{,}14$ určeme aproximaci řešení úlohy (7.2.1) s využitím
(a) vhodného lichého počtu vyčíslení;

[21] Podobně jako v Příkladě 7.2.5 musíme vzít dva sousední intervaly sdílející x_j, neboť nevíme, zda funkce f začala růst již před bodem x_j nebo až za ním, tj. zda $x^* \in [x_{j-1}, x_j]$ nebo $x^* \in [x_j, x_{j+1}]$.

[22] V takovém případě sice platí $x^* \in [x_i, x_{i+1}]$, ale my tento fakt dále nebudeme zohledňovat, viz Poznámku 7.2.4.

[23] Zde $\lceil \cdot \rceil$ značí horní celou část.

(b) vhodného sudého počtu vyčíslení s $\delta = 2/25$. Tento výsledek srovnejme s volbou $N = 16$ a $\delta = 1/9$.

Řešení. Funkce f je nepochybně unimodální, protože je ostře konvexní.

(a) Při lichém počtu dělení musí platit $\frac{1}{2}\frac{2(b-a)}{N+1} \le \varepsilon$, což nás v tomto případě přivede k hodnotě $N \ge 6,1$, a tedy volbě $N = 7$. V takovém případě budeme mít body $x_i = i/8$ pro $i = 1,2,\dots,7$ s funkčními hodnotami

$$f(1/8) = \frac{87}{64} \quad \& \quad f(2/8) = \frac{60}{64} \quad \& \quad f(3/8) = \frac{47}{64} \quad \& \quad f(4/8) = \frac{48}{64}$$
$$f(5/8) = \frac{63}{64} \quad \& \quad f(6/8) = \frac{92}{64} \quad \& \quad f(7/8) = \frac{135}{64}.$$

Nejmenší funkční hodnotu má bod x_3, takže ILM $[2/8, 4/8]$ a aproximací řešení $\overline{x} = 3/8$. Není těžké zjistit, že skutečným řešením je $x^* = 3/7$, a tedy nalezená aproximace opravdu splňuje požadovanou přesnost.

(b) V případě sudého počtu vyčíslení požadujeme, aby $\frac{1}{2}\left(\frac{(b-a)}{\frac{N}{2}+1} + \delta\right) \le \varepsilon$, což nás přivede k podmínce $N \ge 8$, a tedy zvolíme $N = 8$. Dělícími body pak jsou

$$x_1 = \frac{3}{25} \quad \& \quad x_2 = \frac{1}{5} \quad \& \quad x_3 = \frac{8}{25} \quad \& \quad x_4 = \frac{2}{5}$$
$$x_5 = \frac{13}{25} \quad \& \quad x_6 = \frac{3}{5} \quad \& \quad x_7 = \frac{18}{25} \quad \& \quad x_8 = \frac{4}{5}$$

s funkčními hodnotami

$$f(x_1) = \frac{863}{625} \quad \& \quad f(x_2) = \frac{675}{625} \quad \& \quad f(x_3) = \frac{498}{625} \quad \& \quad f(x_4) = \frac{450}{625}$$
$$f(x_5) = \frac{483}{625} \quad \& \quad f(x_6) = \frac{575}{625} \quad \& \quad f(x_7) = \frac{818}{625} \quad \& \quad f(x_1) = \frac{1050}{625}.$$

Nejmenší funkční hodnota nasává v bodě x_4, a proto ILM je $[8/25, 13/25]$ s aproximací $\overline{x} = 21/50$.

V případě volby $N = 16$ a $\delta = 1/9$ budeme mít klasické ekvidistantní dělení ale jen s 8 různými body. Řešením pak bude ILM $[3/9, 5/9]$ a $\overline{x} = 4/9$.

Příklad 7.2.9

Uvažme funkci $f(x) = 5x^2 - 9x + 2$ na intervalu $I = [0,1]$. Určeme aproximaci řešení úlohy (7.2.1) při volbě

(a) $N = 4$ a $\delta = 1/15$;

(b) $N = 4$ a $\delta = 1/30$.

Srovnejme tyto výsledky s tím, který bychom při stejném počtu povolených vyčíslení dostali při ekvidistantním dělení.

Řešení. Zde již ponecháme konkrétní výpočty na čtenáři (případně na výše zmíněném Mapletu). V případě (a) je δ voleno tak, aby apriorní odhad maximální chyby vyšel stejně jako při ekvidistantním dělení. V případě (b) bude tento odhad menší, viz (7.2.2).

7.2.2 Metoda půlení intervalu[27]

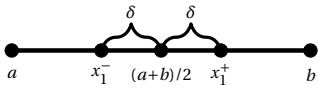
Nechť nyní $N = 2k$. Položíme $x_1^- := \frac{1}{2}(a+b) - \delta$ a $x_1^+ := \frac{1}{2}(a+b) + \delta$, kde $\delta > 0$ je dostatečně malé číslo[28], viz Obrázek 7.13. Porovnáme-li funkční hodnoty $f(x_1^-)$ a $f(x_1^+)$, pak

(i) v případě $f(x_1^-) < f(x_1^+)$ bude podle Lemma 7.2.3 dalším ILM $[a, x_1^+]$;

(ii) v případě $f(x_1^-) \ge f(x_1^+)$ bude podle Lemma 7.2.3 dalším ILM $[x_1^-, b]$.

Tento nově získaný interval opět rozdělíme na polovinu a ve vzdálenosti δ od středu nalezneme body x_2^- a x_2^+. Následným porovnáním obou funkčních hodnot opět nalezneme nový

[27] *Dichotomous Search Method*



Obrázek 7.13: První dělící body pro MPI.

[28] Očividně musí být

$$0 < \delta < \frac{b-a}{2}.$$

Pro $\delta = 0$ máme $x_1^- = x_1^+ = (a+b)/2$ a pro $\delta = (b-a)/2$ bude $x_1^- = a$ a $x_1^+ = b$.

ILM. V *Metodě půlení intervalu* (MPI) tento postup opakujeme celkem k-krát a za *aproximaci* $\overline{x}$ řešení bereme střed posledního ILM, což můžeme zapsat následovně. Odtud také snadno uvidíme, že algoritmus je *aktivní*, neboť rozložení bodů x_i^- a x_i^+ závisí na předchozích bodech a odpovídajících funkčních hodnotách.

Algoritmický popis Metody půlení intervalu

Zadání: Funkce f, interval $[a,b]$, přesnost $\varepsilon > 0$ nebo číslo $N \geq 2$. Stanovení čísel $N/2$ a $\delta \in \big(0,(b-a)/2\big)$.

Krok 1: (*Inicializace*) Položíme $a_0 := a$, $b_0 := b$ a $k := 1$. Vypočteme

$$x_1^- := \frac{a_0+b_0}{2} - \delta \quad \text{a} \quad x_1^+ := \frac{a_0+b_0}{2} + \delta.$$

Krok 2: Vyčíslíme $f(x_k^-)$ a $f(x_k^+)$. Jestliže $f(x_k^-) \geq f(x_k^+)$, pokračujeme Krokem 3. V opačném případě následuje Krok 4.

Krok 3: Položíme $a_k := x_k^-$, $b_k := b_{k-1}$ a pokračujeme Krokem 5.

Krok 4: Položíme $a_k := a_{k-1}$, $b_k := x_k^+$ a pokračujeme Krokem 5.

Krok 5: Je-li $k = N/2$ pokračujeme Krokem 6. Je-li $k < N/2$, vypočteme

$$x_{k+1}^- := \frac{a_k+b_k}{2} - \delta \quad \text{a} \quad x_{k+1}^+ := \frac{a_k+b_k}{2} + \delta.$$

Položíme $k := k+1$ a pokračujeme Krokem 2.

Krok 6: Stanovíme poslední ILM jako $[a_k,b_k]$ a vypočteme $\overline{x} := \frac{a_k+b_k}{2}$.
KONEC

V prvním kroku je délka ILM

$$\ell_1 = b - x_1^- = x_1^+ - a = \frac{b-a}{2} + \delta,$$

ve druhém kroku je délka ILM

$$\ell_2 = \frac{1}{2}\left(\frac{b-a}{2} + \delta\right) + \delta = \frac{b-a}{4} + \frac{3\delta}{2},$$

ve třetím

$$\ell_3 = \frac{1}{2}\left(\frac{b-a}{4} + \frac{3\delta}{2}\right) + \delta = \frac{b-a}{8} + \frac{7\delta}{4},$$

…, v i-tém kroku budeme mít ILM s délkou

$$\ell_i = \frac{\ell_{i-1}}{2} + \delta = \frac{1}{2}\left(\frac{\ell_{i-2}}{2} + \delta\right) + \delta = \cdots = \frac{b-a}{2^i} + \frac{(2^i-1)\delta}{2^{i-1}}. \tag{7.2.3}$$

Tedy po $N = 2k$ vyčíslení budeme mít výsledný ILM délky

$$\ell_k = \frac{b-a}{2^k} + \frac{(2^k-1)\delta}{2^{k-1}}. \tag{7.2.4}$$

Přesnost této metody je tedy nejméně $\frac{1}{2}\ell_k$, což můžeme ovlivnit volbou δ. Mohlo by se zdát, že čím menší zvolíme δ, tím přesnější výsledek dostaneme. Jenže vezmeme-li δ příliš malé, pak hodnoty x_i^- a x_i^+ budou blízké (vždyť jejich „vzdálenost“ je 2δ) a (zejména při numerických výpočtech) může být problém s porovnáním funkčních hodnot $f(x_i^-)$ a $f(x_i^+)$. Budeme-li mít zadanou maximální povolenou přesnost hodnoty $\overline{x}$, pak nejvyšší možný počet ILM k (s počtem vyčíslení $2k$) pro dosažení této přesnosti získáme pomocí řešení nerovnice $\ell_k/2 \leq \varepsilon$, tj.

$$b - a + 2(2^k - 1)\delta \leq 2^{k+1}\varepsilon \quad \text{neboli} \quad \frac{b-a-2\delta}{2(\varepsilon-\delta)} \leq 2^k,$$

což dává $k \geq \left(\ln \frac{b-a-2\delta}{2(\varepsilon-\delta)}\right)/\ln 2$. Proto nejpozději po

$$2\left\lceil \frac{\ln \frac{b-a-2\delta}{2(\varepsilon-\delta)}}{\ln 2} \right\rceil \tag{7.2.5}$$

vyčíslení bude dosaženo požadované přesnosti[29].

[29] Kdyby zadání bylo dáno tak, že ε neudává absolutní povolenou chybu, ale pouze relativní (např. $\varepsilon = 10$ % by znamenalo maximální povolenou odchylku rovnu 10 % z délky počátečního intervalu I), pak bychom museli řešit nerovnost $\ell_k \leq 2\varepsilon(b-a)$, což v případě $I = [0,1]$ splývá s předchozím požadavkem.

Příklad 7.2.10

Pro funkci $f(x) = x(x-3/2)$ na intervalu $I = [0,1]$ určeme nejmenší potřebný počet vyčíslení v MPI s $\delta = 1/1000$, při kterém dosáhneme chyby nejvýše $\varepsilon = 1/10$.
Řešení. Po dosazení do předchozího vzorce obdržíme, že

$$\frac{\ln \frac{b-a-2\delta}{2(\varepsilon-\delta)}}{\ln 2} = \frac{\ln \frac{998}{198}}{\ln 2} \approx 2{,}33, \quad \text{tj.} \quad \left\lceil \frac{\ln \frac{b-a-2\delta}{2(\varepsilon-\delta)}}{\ln 2} \right\rceil = 3$$

Proto bude potřeba alespoň 6 vyčíslení funkce f, což znamená provedení nejméně 3 kroků MPI.

Jaký je řád/rychlost konvergence MPI? S rostoucím počtem vyčíslení dostaneme

$$\ell_k = \frac{b-a}{2^k} + \left(2 - \frac{1}{2^{k-1}}\right)\delta \xrightarrow{k\to\infty} 2\delta. \tag{7.2.6}$$

Vzhledem k (7.2.6) musíme místo posloupnosti odchylek $\{\ell_k\}_{k=0}^{\infty}$ vzít $\{\ell_k - 2\delta\}_{k=0}^{\infty}$, která již konverguje k nule. Potom pro $\ell_0 := b-a$ platí

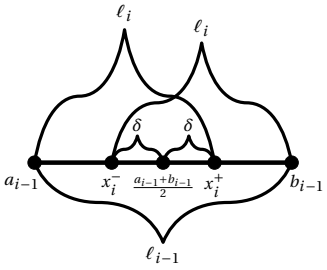
$$\lim_{k\to\infty} \frac{\ell_{k+1} - 2\delta}{\ell_k - 2\delta} = \lim_{k\to\infty} \frac{\ell_0/2^{k+1} - \delta/2^k}{\ell_0/2^k - \delta/2^{k-1}} = 1/2,$$

což znamená, že konvergence je lineární s rychlostí $1/2$ neboli délka ILM klesá exponenciálně s exponentem $1/2$, tj. jako $(1/2)^k = (1/\sqrt{2})^N$.

V úvodu jsme snadno zpozorovali, že má smysl uvažovat pouze $\delta \in (0, (b-a)/2)$. Objeví se ale s rostoucím počtem vyčíslení ještě nějaké omezení pro δ? Z Obrázku 7.14 vidíme, že musí platit $\delta < \ell_{i-1}/2$. Odtud však po dosazení z (7.2.3) dostaneme podmínku

$$\delta < \frac{b-a}{2^i} + \frac{(2^{i-1}-1)\delta}{2^{i-1}}$$

neboli opět $\delta < (b-a)/2$, tj. pro libovolné $\delta \in (0, (b-a)/2)$ se bude MPI vždy chovat „rozumně“. Tato nezávislost δ na počtu vyčíslení velmi úzce souvisí s jednou z velkých výhod MPI. Tou je možnost kdykoli ve výpočtech pokračovat – můžeme přidávat další a další dodatečné ILM, aniž bychom celý výpočet spouštěli od začátku.



Obrázek 7.14: Další dělící body pro MPI.

Příklad 7.2.11

Uvažme funkci $f(x) = 5x^2 - 9x + 2$ na intervalu $I = [0,1]$. S $N = 4$ a $\delta = 1/15$ určeme aproximaci řešení úlohy (7.2.1) s využitím MPI.
Řešení. Začínáme s intervalem $I_0 = I = [0,1]$. Jeho střed je $1/2$, takže v prvním kroku máme body

$$x_1^- = \frac{1}{2} - \frac{1}{15} = \frac{13}{30} \quad \text{a} \quad x_1^+ = \frac{1}{2} + \frac{1}{15} = \frac{17}{30}$$

s funkčními hodnotami $f(x_1^-) = 0{,}96\bar{1}$ a $f(x_1^+) = -1{,}49\bar{4}$, tj. $f(x_1^-) > f(x_1^+)$. Proto v souladu s Lemmatem 7.2.3 dostáváme nový ILM $I_1 = [13/30, 1]$ s délkou $\ell_1 = \frac{1}{2} + \frac{1}{15} = \frac{17}{30}$. Střed intervalu I_1 je v bodě $43/60$, takže ve druhém (a zároveň posledním) kroku máme body

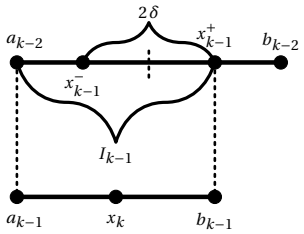
$$x_2^- = \frac{43}{60} - \frac{1}{15} = \frac{13}{20} \quad \text{a} \quad x_1^+ = \frac{43}{60} + \frac{1}{15} = \frac{47}{60}$$

s funkčními hodnotami $f(x_2^-) \approx 1{,}74$ a $f(x_1^+) \approx -1{,}98$, tj. $f(x_2^-) > f(x_2^+)$. Proto v souladu s Lemmatem 7.2.3 dostáváme nový ILM $I_2 = [13/20, 1]$ s délkou $\ell_2 = \frac{1}{4} + \frac{3}{2} \times \frac{1}{15} = \frac{7}{20}$. Hledanou aproximací minima je $\overline{x} = \frac{33}{40}$.

Příklad 7.2.12

Uvažme funkci $f(x) = 7x^2 - 6x + 2$ na intervalu $I = [0,1]$. S chybou nejvýše $\varepsilon = 0{,}14$ a $\delta = 2/25$ určeme aproximaci řešení úlohy (7.2.1) s využitím MPI.

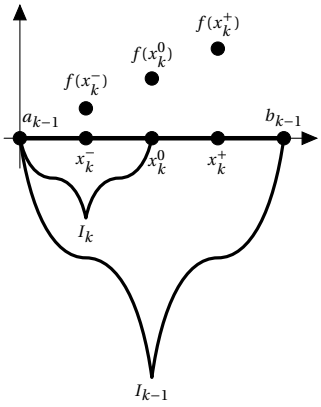
Řešení. Nejdříve určíme potřebný počet vyčíslení. Buď dosazením do vzorce (7.2.5) nebo z rovnosti (7.2.4) dostaneme, že musíme použít alespoň $N = 6$ vyčíslení. Podobnosti jednotlivých kroků ponecháme na čtenáři, ale v jednotlivých krocích dostaneme ILM $I_1 = [0, 29/50]$, $I_2 = [21/100, 29/50]$ a $I_3 = [63/200, 29/50]$. Nalezli jsme tak aproximací minima $\overline{x} = \frac{179}{400}$. Přidáme-li další krok, získáme $I_4 = [63/200, 211/400]$ a $\overline{x} = \frac{337}{800}$.



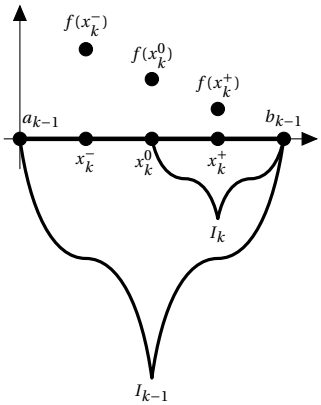
Obrázek 7.15: Poslední krok MPI s lichým počtem vyčíslení.

Mohli bychom MPI upravit tak, aby ji bylo možné aplikovat i s lichým počtem povolených vyčíslení? Ano. Např. tak, že v případě $N = 2k-1$ s $k \geq 2$ nejdříve uděláme $k-1$ kroků stejných jako v MPI. V posledním kroku pak místo x_k^-, x_k^+ uvážíme pouze $x_k := (b_{k-1} - a_{k-1})/2$, jehož funkční hodnotu porovnáme s hodnotou v tom z bodů x_{k-1}^- nebo x_{k-1}^+, který leží uvnitř $[a_{k-1}, b_{k-1}]$, viz Obrázek 7.15. Ovšem tato úprava s sebou nese dodatečné omezení na volbu δ. Nesmí se totiž stát, že $x_k = x_{k-1}^{\pm}$ neboli $\ell_{k-1} = 4\delta$. Po dosazení z (7.2.3) snadno zjistíme, že poslední rovnost je ekvivalentní s $\delta = (b-a)/(2^k+2) < (b-a)/2$, tj. musíme tentokrát volit $\delta \in \big(0, (b-a)/(2^k+2)\big)$. Např. pro $k = 2$ to znamená, že je možné brát pouze $\delta \in \big(0, (b-a)/6\big)$. V tomto případě ale můžeme ztratit možnost přidání dalšího kroku, neboť zvolené δ nemusí vyhovovat odpovídajícímu omezení pro $k+1$. Nicméně myšlenka opakovaného využití některého z již známých bodů je velmi zajímavá. Proto se k ní brzy *vrátíme*, abychom ji zdokonalili a využili k odvození dvou dalších metod.

[30] *Interval Halving Method*



Obrázek 7.16: Ilustrace nového ILM v alternativní metodě půlení intervalu (poprvé).



Obrázek 7.17: Ilustrace nového ILM v alternativní metodě půlení intervalu (podruhé).

Poznámka 7.2.13. Název námi zkonstruované MPI může být trochu zavádějící. V každém kroku totiž dostáváme interval s délkou větší (o hodnotu δ) než je polovina předchozího intervalu ($\ell_i = \ell_{i-1}/2 + \delta$). Kdybychom chtěli docílit *skutečného půlení intervalu*, tj. $\ell_i = \ell_{i-1}/2$, pak bychom to mohli provést třeba takto[30]:

(i) budeme potřebovat lichý počet vyčíslení, tj. $N = 2k+1$;

(ii) v prvním kroku určíme body

$$x_1^- := a + \frac{b-a}{4} = \frac{3}{4}a + \frac{b}{4} \quad \& \quad x_1^0 := a + \frac{b-a}{2} = \frac{b+a}{2} \quad \& \quad x_1^+ := a + \frac{3(b-a)}{4} = \frac{a}{4} + \frac{3}{4}b,$$

vyčíslíme odpovídající funkční hodnoty, a je-li

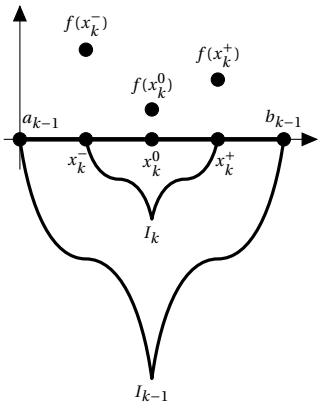
(a) $f(x_1^-) < f(x_1^0) < f(x_1^+)$, pak $x^* \in [a, x_1^0]$ a toto je nový ILM (s bodem x_1^- v jeho středu), viz Obrázek 7.16;

(b) $f(x_1^-) > f(x_1^0) > f(x_1^+)$, pak $x^* \in [x_1^0, b]$ a toto je nový ILM (s bodem x_1^+ v jeho středu), viz Obrázek 7.17;

(c) $f(x_1^-) > f(x_1^0) < f(x_1^+)$, pak $x^* \in [x_1^-, x_1^+]$ a toto je nový ILM (s bodem x_1^0 v jeho středu), viz Obrázek 7.18;

(iii) s tímto novým ILM pokračujeme analogicky dále, přičemž v každém dalším kroku využijeme vždy už jen dvě vyčíslení (neboť funkční hodnotu ve středu intervalu známe z předchozího kroku). Algoritmus skončí ve chvíli, kdy dosáhneme požadované přesnosti nebo povoleného počtu vyčíslení.

Místo ekvidistantního dělení intervalu I_{k-1} pomocí trojice bodů x_k^-, x_k^0 a x_k^+ ve vzdálenosti $\ell_{i-1}/4$ bychom mohli případně uvážit i body x_k^- a x_k^+ ve vzdálenosti (absolutní nebo relativní) určené nějakým vhodným číslem δ. V takovém případě bychom však obecně potřebovali $3k$ vyčíslení a v případě funkčních hodnot uspořádaných jako v (iii) by již neplatilo $\ell_i = \ell_{i-1}/2$, ale $\ell_i = 2\delta$. Každopádně jsme zpět u myšlenky opětovného využití znalosti funkční hodnoty v bodě, který leží uvnitř aktuálního ILM. Proto je načase se na ni podívat podrobněji. ▲



Obrázek 7.18: Ilustrace nového ILM v alternativní metodě půlení intervalu (potření).

[34] *Golden-Section Search Method*

7.2.3 Metoda zlatého řezu[34]

Číslo δ v MPI udává absolutní vzdálenost bodů x_i^- a x_i^+ od středu intervalu I_{i-1}. Toto použití je vlastně jen speciálním případem zcela obecného algoritmu:

(i) máme interval I_{i-1};

(ii) vybereme (jakkoli) dva body $x_i^-, x_i^+ \in I_{i-1}$ a vypočteme odpovídající funkční hodnoty;

(iii) porovnáme funkční hodnoty a s využitím Lemma 7.2.3 stanovíme nový ILM I_i.

Stojíme tak vlastně „jen“ před otázkou, jak *efektivně* zvolit dvojici $x_i^-, x_i^+ \in I_{i-1}$. Místo jednoho pevného δ bychom mohli například použít různé hodnoty v každém kroku, tj. místo δ bychom měli $\{\delta_j\}_{j=1}^k$. Analýza takové metody by však byla o něco komplikovanější než jsme viděli v případě MPI. Každopádně by muselo platit $\delta_j \in \big(0, (b_{j-1}-a_{j-1})/2\big)$.

Mohli bychom ale také vzdálenost bodů x_i^-, x_i^+ od středu intervalu $I_{i-1} = [a_{i-1}, b_{i-1}]$ vyjádřit relativně vzhledem k délce intervalu I_{i-1}, tj.

$$x_i^- = \frac{a_{i-1}+b_{i-1}}{2} - \delta\,(b_{i-1}-a_{i-1}) = a_{i-1} + (1/2-\delta)\,(b_{i-1}-a_{i-1}), \tag{7.2.7}$$

$$x_i^+ = \frac{a_{i-1}+b_{i-1}}{2} + \delta\,(b_{i-1}-a_{i-1}) = a_{i-1} + (1/2+\delta)\,(b_{i-1}-a_{i-1}). \tag{7.2.8}$$

Je zřejmé, že pro $\delta = 0$ oba body splývají se středem intervalu, zatímco pro $\delta = 1/2$ dostaneme krajní body. Proto nutně $\delta \in (0, 1/2)$, a pak z intervalu I_{i-1} délky ℓ_{i-1} dostaneme interval délky

$$\ell_i = \frac{a_{i-1}+b_{i-1}}{2} + \delta\,(b_{i-1}-a_{i-1}) = (1/2+\delta)\,(b_{i-1}-a_{i-1}) = (1/2+\delta)\,\ell_{i-1}.$$

Takže z výchozího intervalu $I = [a, b]$ budeme mít po k krocích (tj. po $N = 2k$ vyčíslení) ILM délky

$$\ell_k = (1/2+\delta)^k\,(b-a).$$

Budeme-li mít zadanou minimální přesnost (nebo maximální odchylku) ε, pak musí platit

$$(1/2+\delta)^k(b-a) \le 2\varepsilon \quad \text{neboli} \quad k \ge \frac{\ln\frac{2\varepsilon}{b-a}}{\ln\frac{1+2\delta}{\delta}}.$$

Místo zkracování délky ILM rychlostí $1/2+\delta$ bychom mohli uvažovat i $\{1/2+\delta_j\}_{j=1}^k$ nebo $\{\alpha_j\}_{j=1}^k$ s $\alpha_j \in (1/2, 1)$, což[35] nás v k-tém kroku přivede k ILM délky

$$\ell_k = (1/2+\delta_1)\,(1/2+\delta_2)\cdots(1/2+\delta_k)\,(b-a) = \prod_{j=1}^{k}(1/2+\delta_j)\,(b-a) = (b-a)\prod_{j=1}^{k}\alpha_j.$$

Avšak k tomu, abychom se dopracovali ke k-tému ILM budeme (obecně) potřebovat $2k$ vyčíslení. Kdyby se nám však nějakým způsobem podařilo snížit potřebný počet vyčíslení, mohli bychom *a priori* očekávat, že naše metoda bude efektivnější při stejné hodnotě N, tj. mohli bychom získat více ILM. Proto naším cílem nyní bude nalezení předpisu pro x_k^- a x_k^+ tak, abychom v každém kroku potřebovali *pouze jedno vyčíslení* funkce f (s výjimkou prvního kroku), tj. po N vyčísleních budeme mít $N-1$ ILM místo pouhých $N/2$. Toho lze docílit jak pomocí konstantní volby δ, tak i s využitím proměnného kroku $\{\delta_j\}_{j=1}^{N-1}$.

Uvažme nejdříve konstantní volbu δ. Mějme v tuto chvíli ILM $I_{i-1} = [a_{i-1}, b_{i-1}]$ s délkou $\ell_{i-1} := b_{i-1} - a_{i-1}$, z něhož chceme pomocí bodů[36] $\lambda_i, \mu_i \in (a_{i-1}, b_{i-1})$ a z Lemmatu 7.2.3 získat nový ILM $I_i = [a_i, b_i]$, tj. hledáme (obecně)

$$\lambda_i = a_{i-1} + \alpha\,\ell_{i-1} \quad \text{a} \quad \mu_i = a_{i-1} + \beta\,\ell_{i-1},$$

kde $\alpha, \beta \in (0,1)$ a $\alpha < \beta$. Potom unimodalita funkce f (potažmo Lemma 7.2.3) opět dává

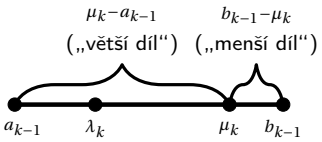
(i) je-li $f(\lambda_i) < f(\mu_i)$, pak $x^* \in [a_{i-1}, \mu_i] =: [a_i, b_i]$;

(ii) je-li $f(\lambda_i) \ge f(\mu_i)$, pak $x^* \in [\lambda_i, b_{i-1}] =: [a_i, b_i]$.

[35] Mohli bychom dokonce vzít $\alpha \in (0,1)$? Raději spíše ne. Kdyby šlo pouze o to, abychom nerozlišovali x_k^- a x_k^+ (tj. x_k^- by vždy bylo určeno pomocí α a x_k^+ by leželo v analogické vzdálenosti ve druhé polovině intervalu I_{k-1}), pak by to přineslo pouze zbytečné komplikace v algoritmickém popisu takové metody – jenže právě takový popis by měl být naším hlavním cílem, neboť ruční výpočty jsou příliš zdlouhavé (a leckdy nepřesné). Kdyby šlo ale o to, aby x_k^- mohlo ležet i v pravé polovině intervalu I_{k-1} a současně by platilo $x_k^- < x_k^+$, pak by situace byla mnohem komplikovanější. Už bychom totiž nedokázali zaručit, že x^* bude ležet v tom kratším intervalu – toto by záleželo na funkci f, což samozřejmě nechceme.

[36] Pro lepší přehlednost a jednoduchost budeme místo x_i^- a x_i^+ používat označení λ_i a μ_i, přičemž bude chtít, aby $\lambda_i < \mu_i$.

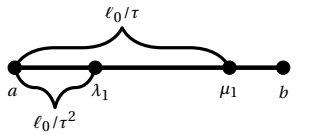
[37] Šlo by to i jinak?



Obrázek 7.19: MZŘ vs. zlatý řez.

[38] Označení právě pomocí řeckého písmene τ je zde inspirováno řeckým slovem *tome* znamenající v překladu *řez*. Toto číslo sice bývá obvykle označováno jako φ (nebo Φ) a zdá se, že sousloví „zlatý řez“ bylo poprvé použito až v knize [401] z roku 1835 (jako *Goldener Schnitt* v poznámce pod čarou na straně 194), ale v historii lze najít několik (cca 7) různých označení. Ovšem se starými Řeky každopádně nemají nic společného, a proto jsme si zavedli symbol τ, když už nejsme prakticky/historicky ničím vázáni.

[39] *Geometrie má dva poklady: Pythagorovu větu a zlatý řez. První má cenu zlata, druhý připomíná spíše drahocenný kámen.*
Johannes Kepler



Obrázek 7.20: První krok MZŘ.

[40] Číslo $\frac{1+\sqrt{5}}{2}$ je známé také jako *zlaté číslo* a bylo výchozím bodem řecké geometrie pravidelných mnohoúhelníků a také všeho, co s nimi souviselo, jako např. dvanáctistěn nebo 20-stěn. Úlohou o rozdělení úsečky v poměru zlatého řezu se zabýval již Eukl(e)ides, který ji využil ve své práci o pravidelných pětiúhelnících. Načrtneme-li totiž pětiúhelník s délkou strany rovnou 1, pak mají jeho dlouhé úhlopříčky délky rovnu právě číslu τ. Proto by nemělo být překvapivé jisté propojení mytologie se zlatým řezem, neboť úhlopříčky tohoto pětiúhelníku tvoří pentagram, u kterého lze u některých úseček také objevit tento poměr, viz Obrázky 7.22–7.23.

Zlatý řez také bývá často spojován s estetikou, harmonií a vkusem. Například jako „nejkrásnější“ obdélník bývá označován takový, jehož strany mají délky v poměru zlatého řezu, viz Obrázek 7.24. Ovšem podklady pro takové tvrzení jsou zjevně poněkud chabé. Na druhou stranu asi nelze popřít, že proporce blízké zlatému číslu (třeba v rozsahu 1,5–1,8) mohou

Jaká jsou nejvhodnější α, β? My navíc požadujeme (z již dříve zmíněných důvodů), aby λ_{i+1} nebo μ_{i+1} odpovídalo některému z předchozích bodů, konkrétně[37]: v případě (i) $\mu_{i+1} = \lambda_i$ a (ii) $\lambda_{i+1} = \mu_i$, tj. v případě (i) máme

$$\mu_{i+1} = a_i + \beta\,\ell_i \overset{\text{(i)}}{=} a_{i-1} + \beta(\mu_i - a_{i-1}) = a_{i-1} + \beta^2\,\ell_{i-1} \quad \text{a} \quad \lambda_i = a_{i-1} + \alpha\,\ell_{i-1},$$

tedy rovnost $\mu_{i+1} = \lambda_i$ nastane právě tehdy, když $\alpha = \beta^2$; v případě (ii) dostaneme

$$\lambda_{i+1} = a_i + \alpha\,\ell_i \overset{\text{(ii)}}{=} \lambda_i + \alpha(b_{i-1} - \lambda_i) = a_{i-1} + \alpha\,\ell_{i-1} + \alpha(\underbrace{b_{i-1} - a_{i-1}}_{=\ell_{i-1}} - \alpha\,\ell_{i-1}) = a_{i-1} + (2\alpha - \alpha^2)\,\ell_{i-1}$$

a současně

$$\mu_i = a_{i-1} + \beta\,\ell_{i-1},$$

takže rovnost $\lambda_{i+1} = \mu_i$ nastane právě tehdy, když $\beta = 2\alpha - \alpha^2$. Celkem tím dostáváme soustavu

$$\alpha = \beta^2 \quad \& \quad \beta = 2\alpha - \alpha^2 \quad \& \quad 0 < \alpha < \beta < 1,$$

tj. hledáme řešení rovnice $2\beta^2 - \beta^4 = \beta$ neboli $\beta(\beta-1)(\beta^2+\beta-1) = 0$ v intervalu $(0,1)$. Tím je pouze hodnota

$$\beta = \frac{\sqrt{5}-1}{2} =: \frac{1}{\tau} \approx 0{,}618, \quad \text{která následně dává} \quad \alpha = \beta^2 = 1 - \beta = \frac{1}{\tau^2} = \frac{3-\sqrt{5}}{2} \approx 0{,}382.$$

Tímto získáváme tzv. *Metodu zlatého řezu* (MZŘ). Proč? A co je vlastně *zlatý řez* a jak souvisí s číslem τ? Z našeho výpočtu plyne, že číslo β lze vyjádřit dvěma způsoby: buď jako

$$\beta = \frac{\beta\,\ell_{k-1}}{\ell_{k-1}} = \frac{\mu_k - a_{k-1}}{b_{k-1} - a_{k-1}}$$

nebo jako

$$\beta = \frac{\beta^2}{\beta} = \frac{1-\beta}{\beta} = \frac{(1-\beta)\,\ell_{k-1}}{\beta\,\ell_{k-1}} = \frac{b_{k-1} - a_{k-1} - \beta\,\ell_{k-1}}{\beta\,\ell_{k-1}} = \frac{b_{k-1} - \mu_k}{\mu_k - a_{k-1}}.$$

Odtud dostáváme rovnost

$$\frac{\mu_k - a_{k-1}}{b_{k-1} - a_{k-1}} = \frac{b_{k-1} - \mu_k}{\mu_k - a_{k-1}},$$

která říká, že body λ_k a μ_k rozdělují interval $[a_{k-1}, b_{k-1}]$ v poměru

$$\frac{\text{větší díl}}{\text{celek}} = \frac{\text{menší díl}}{\text{větší díl}},$$

viz Obrázek 7.19. Jinými slovy, rozdělili jsme interval (nebo chcete-li úsečku) $[a_{k-1}, b_{k-1}]$ v „posvátném“ *poměru zlatého řezu*, který má hodnotu rovnu právě číslu[38,39,40]

$$\tau = \frac{1}{\beta} = \frac{1+\sqrt{5}}{2} \approx 1{,}618$$

a který je řešením kvadratické rovnice

$$\boxed{\tau^2 - \tau - 1 = 0,}$$

a tedy obzvláště platí $1/\tau = \tau - 1$.

A jak MZŘ prakticky použít? Na počátku máme interval $I = [a, b] =: I_0$ délky ℓ_0 a $N \geq 2$. Potom

$$\lambda_1 = a + (1 - 1/\tau)\,\ell_0 = a + \ell_0/\tau^2 \quad \text{a} \quad \mu_1 = a + \ell_0/\tau,$$

viz Obrázek 7.20. Určíme funkční hodnoty $f(\lambda_1)$ a $f(\mu_1)$, jejichž porovnáním následně získáme díky Lemma 7.2.3 nový ILM $I_1 = [a_1, b_1]$. Celý tento postup opakuje tentokrát s intervalem I_1 majícím délku $\ell_1 = \ell_0/\tau$ atd., tj. postupně určujeme body

$$\lambda_i = a_{i-1} + \frac{1}{\tau^2}(b_{i-1} - a_{i-1}) \quad \text{a} \quad \mu_i = a_{i-1} + \frac{1}{\tau}(b_{i-1} - a_{i-1}),$$

s jejichž pomocí prostřednictvím Lemma 7.2.3 získáme nový ILM $I_i = [a_i, b_i]$, jehož jeden krajní bod je právě jeden z bodů λ_i, μ_i, přičemž druhý z těchto bodů leží uvnitř tohoto intervalu (a stane se z něj μ_{i+1} nebo λ_{i+1}). Máme-li povoleno N vyčíslení, pak posledním takto získaným ILM je I_{N-1}, jehož střed opět bereme za aproximaci řešení úlohy (7.2.1). Algoritmický popis MZŘ může vypadat následovně.

mít své estetické kouzlo a opodstatnění. Ale rozhodně bych nečinil z jednoho čísla s nekonečným desetinným rozvojem středobod veškeré dokonalosti. To už se částečně dostáváme někam na pomezí numerologie – budeme-li se opravdu snažit, pak skoro vždy nalezneme podklady pro takové tvrzení. Proto se stává, že se zlatý řez objevuje i v situacích, kam nepatří. Uvádí se, že např. Cheopsova pyramida nebo Parthenon v Akropoli byly navrženy s použitím zlatého řezu, což nejspíše nepodložená domněnka – umíte si představit, že by to skutečně bylo provedeno s využitím takového iracionálního čísla? Nebo se také traduje, že dokonalé lidské tělo má proporce některých částí v poměru zlatého řezu, např. poměr výšky postavy ke vzdálenosti od temene hlavy k prstům na ruce svěšené podél těla nebo poměr předchozí vzdálenosti ke vzdálenosti od temene hlavy k loktům apod. Tento poměr dokonce učaroval také mnohým umělcům při ztvárňování lidského těla – jako třeba ve slavném da Vinciho díle *Vitruviánský muž*, viz Obrázek 7.25. Slavný švýcarsko-francouzský architekt a malíř Le Corbusier s pomocí zlatého řezu dokonce navrhl celý proporcionální systém zvaný *Modulor*, který měl poskytnout standardizovaný systém dodávající harmonické proporce všemu, od klik dveří po výškové budovy. Prostě vše spojené s tímto poměrem je prý v dokonalé harmonii.

Ovšem kdybychom místo úsečky rozdělovali stejným způsobem oblouk jednotkové kružnice, bude už situace zajímavější. Dostali bychom tím úhel cca $137{,}5\ldots°$, kterému se říká *zlatý úhel*. Pro jeho výskyt v přírodě už existují mnohem pádnější argumenty (samozřejmě stále s jistou tolerancí, neboť ne vždy se vše podaří zcela dokonale). Např. uspořádání semen v terči slunečnice bude optimální, pokud k němu využijeme právě zlatý úhel, viz Obrázky 7.26–7.29.

Mnoho dalších zajímavostí jako např. soběpodobnost skrze řetězový zlomek

$$\tau = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1+\ldots}}}$$

nebo o zlaté spirále z Obrázku 7.30 lze najít např. v [262, 357]. Mimochodem traduje se, že schránka loděnky hlubinné vykazuje právě tvar takové zlaté spirály. Není to pravda, je to sice logaritmická spirála, ale zlatý řez v ní nenajdete, viz Obrázek 7.31.

Algoritmický popis Metody zlatého řezu

Zadání: Funkce f, interval $[a,b]$, přesnost $\varepsilon > 0$ nebo číslo $N \geq 2$.

Krok 1: (*Inicializace*) Položíme $a_0 := a$, $b_0 := b$ a $k := 1$. Vypočteme

$$\lambda_1 := a_0 + (b_0 - a_0)/\tau^2 \quad \text{a} \quad \mu_1 := a_0 + (b_0 - a_0)/\tau.$$

Krok 2: Je-li $k = N$, pokračujeme Krokem 7. Jinak následuje Krok 3.

Krok 3: Vyčíslíme $f(\lambda_k)$ a $f(\mu_k)$. Jestliže $f(\lambda_k) \geq f(\mu_k)$, pokračujeme Krokem 4. V opačném případě následuje Krok 5.

Krok 4: Položíme

$$a_k := \lambda_k, \quad b_k := b_{k-1}, \quad \lambda_{k+1} := \mu_k,$$
$$f(\lambda_{k+1}) := f(\mu_k), \quad \mu_{k+1} := a_k + (b_k - a_k)/\tau$$

a pokračujeme Krokem 6.

Krok 5: Položíme

$$a_k := a_{k-1}, \quad b_k := \mu_k, \quad \mu_{k+1} := \lambda_k,$$
$$f(\mu_{k+1}) := f(\lambda_k), \quad \lambda_{k+1} := a_k + (b_k - a_k)/\tau^2$$

a pokračujeme Krokem 6.

Krok 6: Položíme $k := k+1$ a pokračujeme Krokem 2.

Krok 7: Stanovíme poslední ILM jako $[a_{k-1}, b_{k-1}]$ a vypočteme $\overline{x} := \frac{a_{k-1}+b_{k-1}}{2}$.

KONEC

Všimněme si, že vzdálenost bodů λ_1 a μ_1 od krajních bodů intervalu I_0 je stejná neboli tyto body jsou umístěny symetricky podle středu intervalu I_0, viz Obrázek 7.21. Tímto principem pak pokračujeme dále, neboť v každém kroku již jeden bod uvnitř ILM již máme, takže jej „pouze překlopíme“ podle středu tohoto intervalu[41]. V prvním kroku je tedy vzdálenost bodů λ_1 a μ_1 od krajních bodů intervalu rovna ℓ_0/τ^2 (pro bližší krajní bod) a ℓ_0/τ (pro vzdálenější krajní bod) neboli body jsou ve vzdálenosti ℓ_0/τ^2 a ℓ_0/τ od levého krajního bodu. Nový interval lokalizace minima má proto délku ℓ_0/τ. Ve druhém kroku jsou body ve vzdálenosti ℓ_0/τ^3 (pro λ_2) a ℓ_0/τ^2 (pro μ_2) od bodu a_1 a délka nového ILM je ℓ_0/τ^2. Toto je shrnuto v následující tabulce.

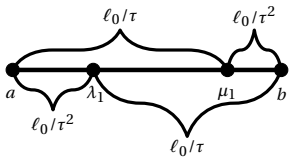
i ($i+1$ vyčíslení)	vzdálenost od a_{i-1} (b_{i-1})		délka ILM
	λ_i (μ_i)	μ_i (λ_i)	
1	ℓ_0/τ^2	ℓ_0/τ	ℓ_0/τ
2	ℓ_0/τ^3	ℓ_0/τ^2	ℓ_0/τ^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$N-1$	ℓ_0/τ^N	ℓ_0/τ^{N-1}	ℓ_0/τ^{N-1}

Tabulka 7.1: Vzdálenosti dělících bodů ILM a jeho délky v MZŘ.

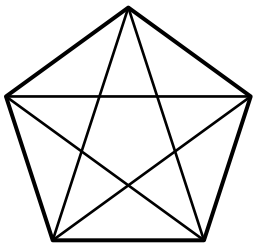
Vezmeme-li posloupnost délek $\{\ell_N\}_{N=1}^{\infty}$ výsledných ILM při N povolených vyčísleních, pak zjevně $\ell_N \to 0$ pro $N \to \infty$ a současně

$$\lim_{N\to\infty} \frac{\ell_{N+1}}{\ell_N} = \lim_{N\to\infty} \frac{\frac{\ell_0}{\tau^{N+1}}}{\frac{\ell_0}{\tau^N}} = \lim_{N\to\infty} \frac{1}{\tau} = \frac{1}{\tau}$$

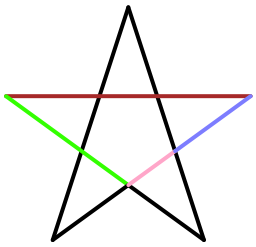
[41] Stačí tedy pouze „nasadit“ λ_1 (nebo μ_1) a dále již vše běží podle tohoto grafického návodu.



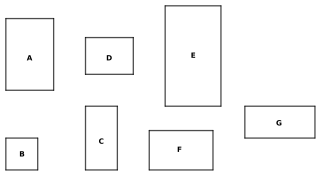
Obrázek 7.21: První krok MZŘ (podruhé).



Obrázek 7.22: Pětiúhelník s jednotkovou stranou a všemi vnitřními úhlopříčkami délky τ.



Obrázek 7.23: Pentagram vytvořený z jednotkového pětiúhelníku. Poměr délky červené a zelené úsečky je roven zlatému číslu. Totéž platí pro poměr délek fialové a modré úsečky.



Obrázek 7.24: Obdélníky s různými poměry stran. Který je nejkrásnější? Jakkoli obskurně může tato otázka znít, prý by to měl být ten pod písmenem „F“...

[42] *Fibonacci Search Method*

[43] Toto dává větší smysl i z praktického hlediska. Je lepší znát vzdálenost od krajního bodu intervalu než od jeho středu, který bychom museli dopočítávat.

takže konvergence je lineární s rychlostí $1/\tau \approx 0{,}618$. To je rychlejší než u MPD, jenže současně pomalejší než pro MPI, u které byla rychlost rovna $1/2$. To ale příliš nekoresponduje s naší původní motivací ohledně nalezení efektivnější metody! Ve skutečnosti zde však žádný rozpor není a pouze to ukazuje, jak zavádějící by mohlo být hodnocení metod pouze podle rychlosti konvergence. Nesmíme totiž zapomínat na to, že v MPI spotřebujeme v každém kroku 2 vyčíslení, zatímco u MZŘ to je pouze 1 (s výjimkou prvního kroku). Srovnáme-li proto odpovídající délky ILM po $2k$ vyčísleních, dostaneme

$$\ell^{\mathrm{MPI}} = \ell_k = \frac{\ell_0}{2^k} + \frac{2^k-1}{2^{k-1}}\delta \quad \text{a} \quad \ell^{\mathrm{MZŘ}} = \ell_{2k-1} = \frac{\ell_0}{\tau^{2k-1}}$$

a pro libovolné $k \geq 2$ bude vždy $\ell^{\mathrm{MPI}}/\ell^{\mathrm{MZŘ}} > 1$, tj. maximální chyba bude vždy (výrazně) menší u MZŘ. Pouze pro $k = 1$ by mohl být ILM v MPI kratší pro dostatečně malé δ, neboť $\ell^{\mathrm{MPI}} = \ell_0/2 + \delta$ a $\ell^{\mathrm{MZŘ}} = \ell_0/\tau$.

Příklad 7.2.14

Uvažme funkci $f(x) = 5x^2 - 4x + 2$ na intervalu $I = [0,1]$. S $N = 5$ určeme aproximaci řešení úlohy (7.2.1) s využitím MZŘ.

Řešení. Začínáme s intervalem $I_0 = I = [0,1]$, takže v prvním kroku máme body

$$\lambda_1 = 0 + \frac{1}{\tau^2} \quad \text{a} \quad \mu_1 = 0 + \frac{1}{\tau}$$

s funkčními hodnotami $f(\lambda_1) \approx 1{,}2016$ a $f(\mu_1) \approx 1{,}4376$, tj. $f(\lambda_1) < f(\mu_1)$. Proto v souladu s Lemmatem 7.2.3 dostáváme nový ILM $I_1 = [0, \mu_1] = [0, 1/\tau]$ s délkou $\ell_1 = \frac{1}{\tau}$.
Ve druhém kroku máme body

$$\lambda_2 = 0 + \frac{1}{\tau^2}\ell_1 = 0 + \frac{1}{\tau^2}\frac{1}{\tau} = \frac{1}{\tau^3} \quad \text{a} \quad \mu_2 = \lambda_1 = \frac{1}{\tau^2}$$

s novou funkční hodnotou $f(\lambda_2) \approx 1{,}3344$, takže $f(\lambda_2) > f(\mu_2)$. Proto dalším ILM je $I_2 = [\lambda_2, 1/\tau] = [1/\tau^3, 1/\tau]$ s délkou $\ell_2 = \frac{1}{\tau} - \frac{1}{\tau^3} = \frac{\tau^2-1}{\tau^3} = \frac{1}{\tau^2}$.
Ve třetím kroku budeme mít body

$$\lambda_3 = \mu_2 = \frac{1}{\tau^2} \quad \text{a} \quad \mu_3 = \lambda_2 + \frac{1}{\tau}\ell_2 = \frac{1}{\tau^3} + \frac{1}{\tau^3} = \frac{2}{\tau^3}$$

s novou funkční hodnotou $f(\mu_3) \approx 1{,}2259$, takže $f(\lambda_3) < f(\mu_3)$. Proto dalším ILM je $I_3 = [1/\tau^3, \mu_3] = [1/\tau^3, 2/\tau^3]$ s délkou $\ell_3 = 1/\tau^3$.
Konečně ve čtvrtém kroku máme body

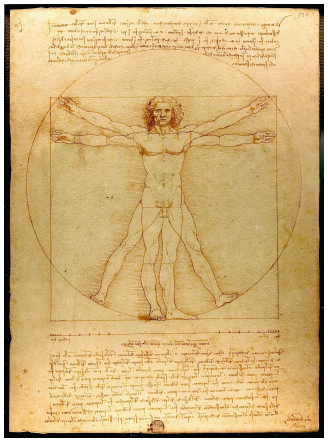
$$\lambda_4 = \frac{1}{\tau^3} + \frac{1}{\tau^2}\ell_3 = \frac{1}{\tau^3} + \frac{1}{\tau^5} = \frac{\tau^2+1}{\tau^5} = \cdots = \frac{7}{\tau} - 4 \quad \text{a} \quad \mu_4 = \lambda_3 = \frac{1}{\tau^2}$$

s novou funkční hodnotou $f(\lambda_4) \approx 1{,}2273$, takže $f(\lambda_4) > f(\mu_4)$. Proto posledním ILM je $I_4 = [\lambda_4, 2/\tau^3] = [(\tau^2+1)/\tau^5, 2/\tau^3]$ s délkou $\ell_4 = \frac{2}{\tau^3} - \frac{\tau^2+1}{\tau^5} = \frac{2\tau^2-\tau^2-1}{\tau^5} = \frac{\tau^2-1}{\tau^5} = \frac{1}{\tau^4}$.
Aproximací hledaného minima je proto bod

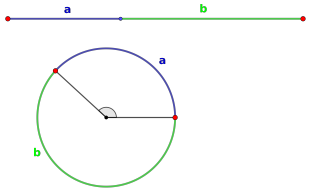
$$\overline{x} = \frac{1}{2}\left(\frac{\tau^2+1}{\tau^5} + \frac{2}{\tau^3}\right) = \frac{3\tau^2+1}{2\tau^5} \approx 0{,}399$$

7.2.4 Fibonacciho metoda[42]

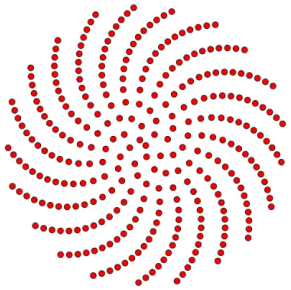
Nyní se vraťme k situaci zmíněna na začátku předchozí podkapitoly, kdy v každém kroku bude možné mít jiné δ určující „relativní zkrácení“ aktuálního ILM. Ovšem podobně jako v MZŘ si místo hodnoty $1/2 - \delta_i$ vystupující v analogických rovnostech jako (7.2.7)–(7.2.8) zavedeme číslo[43] $\alpha_i \in (0, 1/2)$ (číslo β_i zavádět nebudeme). Potom v i-té iteraci (tj. pro výpočet I_i) máme (opět zůstáváme u λ_i a μ_i místo $x_i^{\pm}$)



Obrázek 7.25: Kresba Leonarda da Vinciho s názvem *Vitruviánský muž* (italsky *L'uomo vitruviano*; kolem roku 1490), která znázorňuje proporce lidského těla muže. Snaha o zakomponování zlatého řezu do některých proporcích lidského těla je celkem v souladu s da Vinciho výrokem: *Protože Bůh věčně geometrizuje, příroda je ze své podstaty matematická: „Proporce se nenachází pouze v číslech a měřeních ale také ve zvuku, váze, čase, poloze a vůbec v čemkoli, co existuje."*

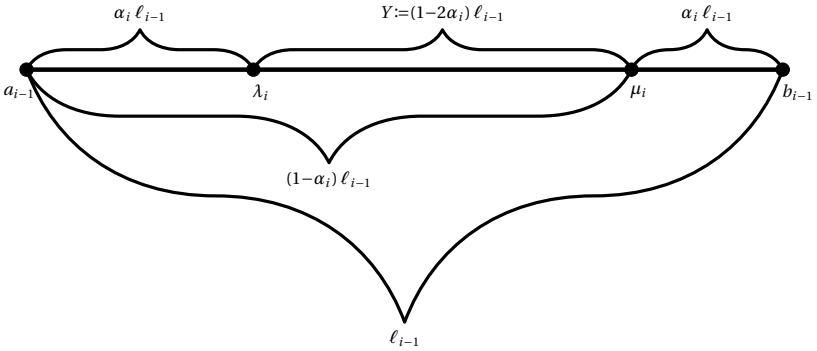


Obrázek 7.26: Rozdělení oblouku jednotkové kružnice v poměru zlatého řezu a zlatý úhel.



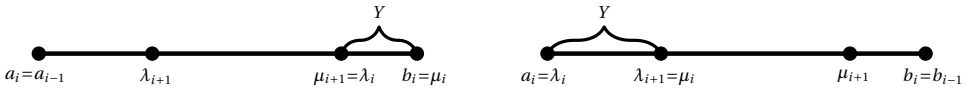
Obrázek 7.27: Rozmístění semen v terči s využitím úhlu 137°.

[44] Pokud bychom se vrátili zpět ke konstantnímu δ potažmo α, dostali bychom rovnici $\alpha = 1-\alpha/(1-\alpha)$ neboli $\alpha^2-3\alpha+1=0$, jenž vede k $\alpha = (3-\sqrt{5})/2 = 1/\tau^2 \in (0,1/2)$, což přesně odpovídá našim výpočtům v MZŘ.



Obrázek 7.32: Cesta k FM.

z čehož podle toho, zda $f(\lambda_i) < f(\mu_i)$ nebo $f(\lambda_i) \geq f(\mu_i)$ dostaneme s využitím Lemma 7.2.3 v dalším kroku jednu z těchto dvou situací:



Obrázek 7.33: Cesta k FM (pokračování).

Takže v obou případech musí platit

$$Y = (1-2\alpha_i)\,\ell_{i-1} \quad \text{a současně} \quad Y = \alpha_{i+1}\,\ell_i = \alpha_{i+1}\,(1-\alpha_i)\,\ell_{i-1},$$

z čehož plyne $1-2\alpha_i = \alpha_{i+1}\,(1-\alpha_i)$ neboli[44]

$$\alpha_{i+1} = 1 - \frac{\alpha_i}{1-\alpha_i} \tag{7.2.9}$$

Budeme-li takto postupovat dále, bude po M krocích (tj. $M+1$ vyčíslení) ILM délky

$$\ell_M = (1-\alpha_1)\,(1-\alpha_2)\cdots(1-\alpha_M)\,\ell_0.$$

Vztah (7.2.9) vlastně ukazuje, jako volit $\alpha_1, \alpha_2, \ldots$, aby celý algoritmus probíhal dle našich představ, tj. stačí vybrat $\alpha_1 \in (0,1/2)$. V takovém případě jsou již hodnoty $\alpha_2, \alpha_3, \ldots$ jasně dány a apriorní odhad maximální možné chyby bude roven $\ell_M/2$. Co kdybychom ale přidali přirozený požadavek, aby naše volba $\alpha_1, \ldots, \alpha_m$ byla taková, že tato hodnota byla nejmenší možná? To by znamenalo, že $\alpha_1, \ldots, \alpha_M$ musí být řešením minimalizační úlohy

$$\left.\begin{aligned} (1-\alpha_1)\,(1-\alpha_2)\cdots(1-\alpha_M) &\to \min, \\ \alpha_{n+1} = 1 - \frac{\alpha_n}{1-\alpha_n}, &\quad n = 1,2,\ldots,M-1, \\ 0 \leq \alpha_n \leq 1/2, &\quad n = 1,\ldots,M, \end{aligned}\right\} \tag{7.2.10}$$

Hodnotu ℓ_0 může vynechat, neboť nemá žádný vliv na volbu $\alpha_1, \ldots, \alpha_M$. Navíc jsme zde povolili i „extrémní" hodnoty $\alpha_n \in \{0, 1/2\}$. Je jasné, že $\alpha_1 \in \{0, 1/2\}$ k řešení nepovede, ale co kdyby to bylo náhodou užitečné v některém z dalších kroků (např. v tom posledním). Hlavním důvodem pro tuto změnu ale je, abychom měli zaručenu existenci řešení dané úlohy, k čemuž využíváme Weierstrassovu větu (viz Věta 1.2.11).

Řešení této úlohy úzce souvisí s tzv. *Fibonacciho*[45] *čísly* F_n, které jsou členy nekonečné posloupnosti přirozených začínající hodnotami 1 a 1 (někdy bývá uváděna s 0 a 1) a každé další číslo je součtem dvou předchozích, tj.

$$F_{n+2} = F_{n+1} + F_n, \quad F_0 = 1, \quad F_1 = 1, \quad n \in \mathbb{N} \cup \{0\}, \tag{7.2.12}$$

což dává posloupnost $1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$

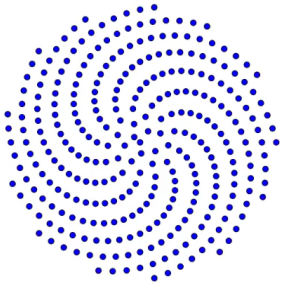
Věta 7.2.15

Řešením naší minimalizační úlohy (7.2.10) jsou čísla

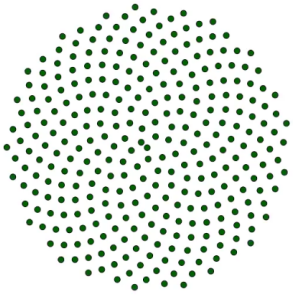
$$\alpha_1 = \frac{F_{M-1}}{F_{M+1}}, \quad \alpha_2 = \frac{F_{M-2}}{F_M}, \quad \ldots \quad , \quad \alpha_i = \frac{F_{M-i}}{F_{M-i+2}}, \quad \ldots, \quad \alpha_M = \frac{1}{2}, \tag{7.2.13}$$

takže délka ILM I_M (po $M+1$ vyčísleních) je

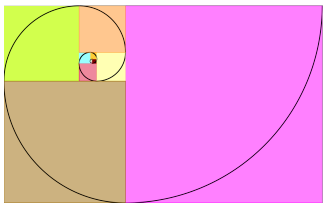
$$\ell_M = \frac{F_M}{F_{M+1}} \cdot \frac{F_{M-1}}{F_M} \cdots \frac{F_2}{F_3} \cdot \frac{F_1}{F_2} \cdot \ell_0 = \frac{\ell_0}{F_{M+1}}.$$



Obrázek 7.28: Rozmístění semen v terči s využitím úhlu 138°.



Obrázek 7.29: Rozmístění semen v terči s využitím zlatého úhlu.



Obrázek 7.30: Zlatá spirála. Největší čtverec má délku strany 1, menší postupně $1/\tau$, $1/\tau^2$ atd. Viz také [221].

Obrázek 7.31: Řez schránkou loděnky hlubinné, na kterém je patrná *logaritmická* spirála.

[45] Leonardo Bonacci (1170–1250, též Leonardo z Pisy nebo Leonardo Pisano Bigollo, což používal, nebo Leonardo Fibonacci, což je označení znamenající Bonacciho syn a pocházející pravděpodobně až z 18. století) byl středověký italský matematik. V roce 1202 vydal knihu *Liber Abaci* (*Kniha počtů*), v níž propagoval hindustánsko-

Důkaz. Pro jednoduchost si označme $\gamma_n := 1 - \alpha_n$ pro $n \in \{1, \ldots, M\}$. Pak máme úlohu

$$\gamma_1 \gamma_2 \cdots \gamma_M \to \min,$$
$$\gamma_{n+1} = \frac{1}{\gamma_i} - 1, \quad n = 1, 2, \ldots, M-1,$$
$$1/2 \le \gamma_n \le 1, \quad n = 1, \ldots, M.$$

Je jasné, že podmínka $\gamma_{n+1} = \frac{1}{\gamma_n} - 1$ zaručuje $\gamma_n \ge 1/2$ právě tehdy, když $\gamma_{n+1} \le 1$. Také odtud plyne, že $\gamma_n \ge 1/2$ právě tehdy, když $\gamma_{n-1} \le 2/3$. Proto druhou podmínku $\gamma_n \le 1$ můžeme odstranit, neboť její splnění je zaručeno požadavkem $\gamma_n \le 1/2$, tj. stačí uvážit pouze podmínky

$$\gamma_{n+1} = \frac{1}{\gamma_n} - 1, \quad n = 1, 2, \ldots, M-1,$$
$$\gamma_n \ge 1/2, \quad n = 1, \ldots, M.$$

S využitím Fibonacciho čísel a volbou $F_{-1} := 0$, můžeme napsat[46]

$$\gamma_2 = \frac{1}{\gamma_1} - 1 = -\frac{1-\gamma_1}{0-\gamma_1} = -\frac{F_0 - F_1\gamma_1}{F_{-1} - F_0\gamma_1}.$$

Připusťme, že pro libovolné $i = 1, \ldots, k$ platí

$$\gamma_n = -\frac{F_{n-2} - F_{n-1}\gamma_1}{F_{n-3} - F_{n-2}\gamma_1}. \tag{7.2.14}$$

Pak také máme

$$\gamma_{k+1} = \frac{1}{\gamma_k} - 1 = -\frac{F_{k-3} - F_{k-2}\gamma_1}{F_{k-2} - F_{k-1}\gamma_1} - 1 = -\frac{(F_{k-3} + F_{k-2}) - (F_{k-2} + F_{k-1})\gamma_1}{F_{k-2} - F_{k-1}\gamma_1} = -\frac{F_{k-1} - F_k\gamma_1}{F_{k-2} - F_{k-1}\gamma_1},$$

což ukazuje, že rovnost (7.2.14) platí pro všechna $n \in \mathbb{N}$.

S využitím právě odvozeného vztahu můžeme ukázat, že všechna $n \in \mathbb{N}$ platí

$$(-1)^n (F_{n-2} - F_{n-1}\gamma_1) > 0. \tag{7.2.15}$$

Pro $n = 1$ je tato triviální, neboť $(-1)^2 (F_0 - F_1\gamma_1) = 1 - \gamma_1$. Ale $\gamma_1 = \frac{1}{\gamma_2 + 1} \le 2/3$, protože $\gamma_2 \ge 1/2$. Tudíž $1 - \gamma_1 \ge 1/3 > 0$. Připusťme, že nerovnost (7.2.15) platí pro $n = 1, \ldots, k$. Pak

$$(-1)^{k+1} (F_{k-1} - F_k\gamma_1) = (-1)^{k+1} \gamma_{k+1} \frac{1}{\gamma_{k+1}} (F_{k-1} - F_k\gamma_1) \overset{(7.2.14)}{=} (-1)^k \gamma_{k+1} (F_{k-2} - F_{k-1}\gamma_1) > 0,$$

a tedy nerovnost (7.2.15) platí pro všechna $n \in \mathbb{N}$.

Jelikož $\gamma_{n+1} = \gamma_n - 1$ a $\gamma_n \ge 1/2$, máme $\gamma_{n+1} \le 1$. Dosazením vyjádření γ_{n+1} z (7.2.14) dostaneme

$$-\frac{F_{n-1} - F_n\gamma_1}{F_{n-2} - F_{n-1}\gamma_1} \le 1 \quad \text{neboli} \quad -\frac{(-1)^n (F_{n-1} - F_n\gamma_1)}{(-1)^n (F_{n-2} - F_{n-1}\gamma_1)} \le 1,$$

což s využitím (7.2.15) dává

$$-(-1)^n (F_{n-1} - F_n\gamma_1) \le (-1)^n (F_{n-2} - F_{n-1}\gamma_1)$$

neboli

$$(-1)^{n+1} (F_{n-1} + F_{n-2}) \le (-1)^{n+1} (F_{n-1} + F_n)\gamma_1.$$

-arabský číselný systém v Evropě a přispěl tak výrazně k rozšíření arabského číselného systému. Fibonacci zkonstruoval tuto posloupnost k popisu růstu populace králíků za následujících (zidealizovaných) podmínek:

- na začátku prvního měsíce si pořídíme 1 nově narozený pár;
- nově narozené páry jsou „produktivní“ od druhého měsíce;
- každý měsíc zplodí každý produktivní pár jeden další pár;
- králíci neumírají a plodí neustále.

Čísla F_n pak vyjadřují velikost populace (v párech) na konci n-tého měsíce a tvoří posloupnost, kterou lze vystopovat až do indické matematiky v souvislosti s básnickou tradicí v sanskrtu. Nicméně toto je samozřejmě velmi naivní přístup a existují mnohem realističtější populační modely (např. Leslieho modely). Avšak přeci jen existuje v přírodě skupina, pro kterou se Fibonacciho čísla hodí mnohem lépe. Jde o včelí rodokmen (toto zjištění bylo poprvé publikováno asi J. Horganem v článku *Bees and mathematics* zveřejněného v únoru 2004 v časopise An Beachaire /Irský včelař/). Včely se totiž rozmnožují pomocí *partenogeneze* (nebo samobřezostí či pannobřezostí), takže platí, že

- je-li nakladeno neoplodněné vejce, pak z něj vznikne samec nebo trubec;
- je-li nakladeno vejce oplozené samcem, pak vznikne samice.

To znamená, že každý samec má jen jednoho rodiče, zatímco samice má rodiče dva. Podíváme-li se potom na rodokmen nějakého samce, tak ten má jednoho rodiče, 2 prarodiče (= 2 včely), 3 pra-prarodiče, 5 pra-pra-prarodičů atd., což jsou právě čísla z Fibonacciho posloupnosti. Již Kepler zjistil, že rychlost růstu Fibonacciho posloupnosti (tj. podíl dvou po sobě jdoucích členů) konverguje k hodnotě zlatého řezu, tj. platí

$$\frac{F_{n+1}}{F_n} \to \tau \quad \text{pro } n \to \infty. \tag{7.2.11}$$

Vskutku? Rekurentní předpis (7.2.12) vlastně zadává diferenční rovnici druhého řádu se dvěma počátečními podmínkami. Příslušný charakteristický polynom $\lambda^2 - \lambda - 1 = 0$ má kořeny $(1 \pm \sqrt{5})/2$, a proto obecné řešení dané diferenční rovnice je

Z definice Fibonacciho čísel pak máme nerovnost

$$(-1)^{n+1} F_n \le (-1)^{n+1} F_{n+1} \gamma_1 \quad \text{neboli} \quad (-1)^{n+1} \frac{F_n}{F_{n+1}} \le (-1)^{n+1} \gamma_1. \tag{7.2.16}$$

Dosazením za $\gamma_1, \dots, \gamma_M$ z (7.2.14) a s využitím (7.2.16) dostaneme

$$\begin{aligned}
\gamma_1 \cdots \gamma_M &= \gamma_1 \times (-1) \frac{F_0 - F_1\gamma_1}{F_{-1} - F_0\gamma_1} \times (-1) \frac{F_1 - F_2\gamma_1}{F_0 - F_1\gamma_1} \times (-1) \frac{F_2 - F_3\gamma_1}{F_1 - F_2\gamma_1} \times \cdots \times (-1) \frac{F_{M-2} - F_{M-1}\gamma_1}{F_{M-3} - F_{M-2}\gamma_1} = \\
&= (-1)^{M-1} \gamma_1 \frac{1}{-\gamma_1} (F_{M-2} - F_{M-1}\gamma_1) = (-1)^M (F_{M-2} - F_{M-1}\gamma_1) = \\
&= (-1)^M F_{M-2} + (-1)^{M+1} F_{M-1}\gamma_1 \ge (-1)^M F_{M-2} + (-1)^{M+1} F_{M-1} \frac{F_M}{F_{M+1}} = \\
&= (-1)^M (F_{M-2} F_{M+1} - F_{M-1} F_M) \frac{1}{F_{M+1}}.
\end{aligned} \tag{7.2.17}$$

Jelikož

$$(-1)^M (F_{M-2} F_{M+1} - F_{M-1} F_M) = 1,$$

plyne[47] odtud, že

$$\gamma_1 \cdots \gamma_M \ge \frac{1}{F_{M+1}},$$

a dokonce

$$\gamma_1 \cdots \gamma_M = \frac{1}{F_{M+1}} \quad \text{právě tehdy, když} \quad \gamma_1 = \frac{F_M}{F_{M+1}},$$

viz nerovnost v (7.2.17). Proto volba

$$\gamma_i = \frac{F_{M-i+1}}{F_{M-i+2}}$$

pro $i = 1, 2, \dots, M$ dává požadovaný minimální součin. Pak zpětným výpočtem $\alpha_1, \dots, \alpha_M$ s využitím definice Fibonacciho čísel získáme hodnoty uvedené v (7.2.13). ■

Máme-li povoleno N vyčíslení, bude $M = N-1$ a výpočet λ_k a μ_k pro $k = 1, 2, \dots, N-1$ probíhá podle vzorců

$$\begin{aligned}
\lambda_i &= a_{i-1} + \frac{F_{N-i-1}}{F_{N-i+1}} \ell_{i-1} = b_{i-1} + \frac{F_{N-i}}{F_{N-i+1}} \ell_{i-1}, \\
\mu_i &= a_{i-1} + \frac{F_{N-i}}{F_{N-i+1}} \ell_{i-1} = b_{i-1} + \frac{F_{N-i-1}}{F_{N-i+1}} \ell_{i-1},
\end{aligned}$$

přičemž $\mu_{i+1} = \lambda_i$ nebo $\lambda_{i+1} = \mu_i$ podle příslušných funkčních hodnot. Tímto postupem dostáváme *Fibonacciho metodu* (FM). Vzdálenost λ_1 od a_0 a μ_1 od b_0 je rovna $\frac{F_{N-2}}{F_N} \ell_0$, zatímco vzdálenost λ_1 od b_0 a μ_1 od a_0 je rovna $\frac{F_{N-1}}{F_N} \ell_0$. Délka nového ILM I_1 pak je $\ell_1 = \frac{F_{N-1}}{F_N} \ell_0$. Podobně vzdálenost λ_2 od a_1 a μ_2 od b_1 je rovna $\frac{F_{N-3}}{F_{N-1}} \ell_1 = \frac{F_{N-3}}{F_N} \ell_0$, zatímco vzdálenost λ_2 od b_1 a μ_2 od a_1 je rovna $\frac{F_{N-2}}{F_{N-1}} \ell_1 = \frac{F_{N-2}}{F_N} \ell_0$, takže délka ILM I_2 je $\ell_2 = \frac{F_{N-2}}{F_N} \ell_0$. Obecně platí, že délka ILM I_k je $\ell_k = \frac{F_{N-k}}{F_N} \ell_0$, vzdálenost λ_k od a_{k-1} a μ_k od b_{k-1} je $\frac{F_{N-k-1}}{F_N} \ell_0$, zatímco pro λ_k od b_{k-1} a μ_k od a_{k-1} to je $\frac{F_{N-k}}{F_N} \ell_0$, viz také Tabulku 7.2.

Jenže k tomu, abychom vůbec našli optimální řešení úlohy (7.2.10) jsme povolili také hodnoty $\alpha_k = 1/2$ (bez toho by řešení ani neexistovalo). Za tento ústupek nakonec přeci jen budeme muset něco „zaplatit“, neboť pro $i = N-1$ (tj. po vyčerpání $N-1$ vyčíslení) je

$$\lambda_{N-1} = a_{N-2} + \frac{1}{2} \ell_{N-2} = \mu_{N-1},$$

takže v posledním kroku nedostaneme žádný nový bod. Co s tím? Jedno vyčíslení nevyužijeme? To by byla škoda. Zvolíme-li číslo $0 < |\delta| < \ell_0 / F_N$, potom

(i) v případě $\delta \in (0, \ell_0/F_N)$ je $\lambda_{N-1} = a_{N-2} + \ell_{N-2}/2$ střed intervalu I_{N-2} a

$$\mu_{N-1} = a_{N-2} + \ell_{N-2}/2 + \delta,$$

viz Obrázek 7.44.

$$z_n = C_1\left(\frac{1+\sqrt{5}}{2}\right)^n + C_2\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Dosazením podmínek $z_0 = 1 = z_1$ dostaneme $C_1 = (\sqrt{5}+1)/(2\sqrt{5})$ a $C_2 = (\sqrt{5}-1)/(2\sqrt{5})$, takže

$$z_n = F_n = \frac{1}{\sqrt{5}}\left[\tau^{n+1} - \left(-\frac{1}{\tau}\right)^{n+1}\right],$$

z čehož již lze vidět platnost vztahu (7.2.11). Odtud je také vidět, že

$$F_n \sim \frac{\tau^{n+1}}{\sqrt{5}} \qquad \text{pro velká } n.$$

Mimochodem ke stejnému závěru jako v (7.2.11) bychom dospěli, i kdybychom si zvolili posloupnost vytvořenou stejným pravidlem jako v případě Fibonacciho posloupnosti, ale s jinými (libovolnými) počátečními hodnotami.

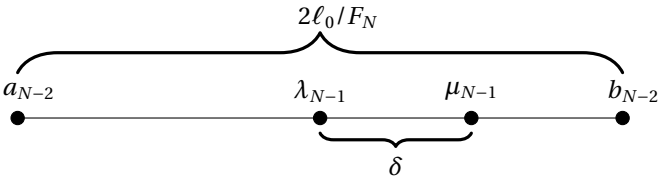
Několik dalších zajímavostí o Fibonacciho číslech:

- F_n vyjadřuje počet způsobů, kterými lze „vyrobit" číslo N jako součet 1 a 2 (např. kolika způsoby lze vyjít schody, můžeme-li chodit po $1-2$ schodech; kolika způsoby lze rozměnit bankovku na 10- a 20-koruny atd.);
- platí $x = F_n$ právě tehdy, když $\sqrt{5x^2+4}$ nebo $\sqrt{5x^2-4}$ je celé číslo;
- pro výpočet (počítačem) si stačí pamatovat pouze poslední dvě hodnoty;
- platí

$$\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}^n = \begin{pmatrix}F_{n+1} & F_n\\ F_n & F_{n-1}\end{pmatrix}$$

- lze vytvořit Fibonacciho čtverec, což je analogie (a aproximace) zlaté spirály z Obrázku 7.30, viz Obrázky 7.35–7.38.

Je až neuvěřitelné, jakou oblibu mají Fibonacciho čísla v přírodě. Kromě nereálného příkladu s králíky a reálného příkladu s rodokmenem včel můžeme Fibonacciho čísla najít např. při počítání okvětních lístků různých květin. Velmi často to je právě některé z Fibonacciho čísel, např. 1 okvětní lístek má ďáblík bahenní, kornoutice africká, lopatkovec neboli toulcovka; 2 okvětní lístky má pryšec zářivý neboli Kristova koruna; 3 okvětní lístky mají podeňka Andersonova, pěknosemenec, lilie, kosatce (ačkoli u posledních dvou by se mohlo zdát, že mají 6 lístků, ve skutečnosti jde pouze o lístky 3 a další tři tvoří kalich sloužící jako vnější ochrana květu, když je v zárodku), 5 okvětních lístků mají blatouchy, divoké růže, karafiáty, azalky, ibišek, pryskyřník, orlíček, hvozdík, 8 okvět-

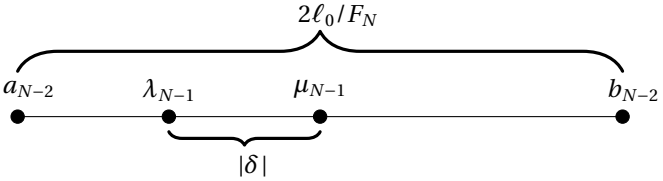


Obrázek 7.44: Poslední krok FM pro $\delta > 0$.

(ii) v případě $\delta \in (-\ell_0/F_N, 0)$ je

$$\lambda_{N-1} = a_{N-2} + \ell_{N-2}/2 + \delta$$

a $\mu_{N-1} = a_{N-2} + \ell_{N-2}/2$ je střed intervalu I_{N-2}, viz Obrázek 7.45.



Obrázek 7.45: Poslední krok FM pro $\delta > 0$.

Následně vyčíslíme a porovnáme funkční hodnoty $f(\lambda_{N-1})$ a $f(\mu_{N-1})$, z čehož s využitím Lemma 7.2.3 získáme poslední ILM I_{N-1} s délkou ℓ_0/F_N nebo $\ell_0/F_N + |\delta|$, jehož střed bereme opět za aproximaci řešení úlohy (7.2.1). Algoritmický popis FM může vypadat následovně.

Algoritmický popis Fibonacciho metody

Zadání: Funkce f, interval $[a,b]$, přesnost $\varepsilon > 0$ nebo číslo $N \geq 2$. Stanovení δ splňujícího $0 < |\delta| < \ell_0/F_N$.

Krok 1: (*Inicializace*) Položíme $a_0 := a$, $b_0 := b$ a $k := 1$. Vypočteme

$$\lambda_1 := a_0 + \frac{F_{N-2}}{F_N}(b_0 - a_0) \quad \text{a} \quad \mu_1 := a_0 + \frac{F_{N-1}}{F_N}(b_0 - a_0).$$

Krok 2: Je-li $k = N-1$, pokračujeme Krokem 7. Jinak následuje Krok 3.

Krok 3: Vyčíslíme $f(\lambda_k)$ a $f(\mu_k)$. Jestliže $f(\lambda_k) > f(\mu_k)$, pokračujeme Krokem 4. V opačném případě následuje Krok 5.

Krok 4: Položíme

$$a_k := \lambda_k, \quad b_k := b_{k-1}, \quad \lambda_{k+1} := \mu_k,$$
$$f(\lambda_{k+1}) := f(\mu_k), \quad \mu_{k+1} := a_k + \frac{F_{N-k-1}}{F_{N-k}}(b_k - a_k)$$

a pokračujeme Krokem 6.

Krok 5: Položíme

$$a_k := a_{k-1}, \quad b_k := \mu_k, \quad \mu_{k+1} := \lambda_k,$$
$$f(\mu_{k+1}) := f(\lambda_k), \quad \lambda_{k+1} := a_k + \frac{F_{N-k-2}}{F_{N-k}}(b_k - a_k)$$

a pokračujeme Krokem 6.

Krok 6: Položíme $k := k+1$ a pokračujeme Krokem 2.

Krok 7: Položíme $\sigma := (b_k + a_k)/2 + \delta$. Vyčíslíme $f\big((b_k+a_k)/2\big)$ a $f(\sigma)$.

(a) Je-li $\sigma < (b_k+a_k)/2$ a $f(\sigma) < f\big((b_k+a_k)/2\big)$, pak $a_{N-1} := a_{N-2}$ a $b_{N-1} := (b_k+a_k)/2$.

(b) Je-li $\sigma < (b_k+a_k)/2$ a $f(\sigma) \geq f\big((b_k+a_k)/2\big)$, pak $a_{N-1} := \sigma$ a $b_{N-1} := b_{N-2}$.

ních lístků najdeme u krásenky, celandine a krvavěnky (krevnice) kanadské, 13 okvětních lístků mají třapatka, heřmánek pravý a některé sedmikrásky, 21 okvětních lístků mají astry, čekanky, kopretina osenní, rmen rolní, zlateň věncová, 34 okvětních lístků má kopretina bílá a 55 nebo 89 okvětních lístků lze napočítat u hvězdnice, jiřina, astra novoanglická, viz Obrázky 7.39–7.42. Pochopitelně se vyskytují i jiné počty okvětních lístků, i když ne tak často a navíc se mnohdy jedná o dvojnásobek některého z Fibonacciho čísel nebo mocniny 2. Někdy se také jedná o čísla ze „spřátelené" *Lucasovy posloupnosti*

$$2, 1, 3, 4, 7, 11, 18, 29, 57$$

vytvořené podle stejného pravidla pouze s jiným začátkem.

I šiška z borovice je v tomto ohledu velmi zajímavá: její šupiny tvoří dvě různě početné skupiny spirál (po a proti směru hodinových ručiček), jejichž počty jsou dvě po sobě jdoucí Fibonacciho čísla. Podobně tomu je u ananasu nebo v semeníku slunečnice, viz Obrázek 7.43. K tomu už existuje skutečně biologické opodstatnění. Velmi úzce to totiž souvisí s tzv. *fylotaxií*, což je oblast zkoumající rozmístění listů na stonku. Když budeme od nějakého listu spirálovitě postupovat po stonku směrem nahoru k nejbližšímu listu a tak dále, než narazíme na list vyrůstající přesně nad výchozím listem, pak poměr otáček a počtu listů, které na „cestě" potkáme, bude často poměrem dvou po sobě jdoucích čísel z Fibonacciho posloupnosti (případně je bude oddělovat pouze jedno další číslo z Fibonacciho posloupnosti)! Např. u slunečnice by to měly být 3/5, u hrušky 3/8, u mandlovníku 5/13 atd. Přesné důvody výskytu Fibonacciho čísel v životě květin jsou zatím předmětem diskuzí, nicméně mnoho faktů již dokázáno bylo. Když na rostlině vyraší nový výhonek, pak dávno před tím, než se začne formovat květ, se na něm vytvoří malé výhonky zvané *primordia*, z nichž později vyrostou semena a další důležité části rostliny. Následná primordia se tvoří ve zlatém úhlu(!). Tuto hodnotu lze matematicky předpovědět, pokud budeme předpokládat, že se primordia rozprostřou po stonku nejefektivnějším způsobem. Na druhé straně je efektivní rozprostření primordií důsledkem elastických vlastností rostoucího stonku, což jsou síly, které ovlivňují vznik primordií. Do hry také vstupuje geometrická „výbava"

(c) Je-li $(b_k + a_k)/2 < \sigma$ a $f\big((b_k + a_k)/2\big) < f(\sigma)$, pak $a_{N-1} := a_{N-2}$ a $b_{N-1} := \sigma$.

(d) Je-li $\sigma < (b_k + a_k)/2$ a $f\big((b_k + a_k)/2\big) \geq f(\sigma)$, pak $a_{N-1} := (b_k + a_k)/2$ a $b_{N-1} := b_{N-2}$.

Posledním ILM je $[a_{k-1}, b_{k-1}]$ a vypočteme $\overline{x} := \frac{a_{k-1}+b_{k-1}}{2}$.

KONEC

Jednotlivé vzdálenosti dělících a krajních bodů a délky odpovídajících ILM v jednotlivých krocích jsou shrnuty v následující tabulce.

i ($i+1$ vyčíslení)	vzdálenost od a_{i-1} (b_{i-1})		délka ILM
	λ_i (μ_i)	μ_i (λ_i)	
1	$\frac{F_{N-2}}{F_N}\ell_0$	$\frac{F_{N-1}}{F_N}\ell_0$	$\frac{F_{N-1}}{F_N}\ell_0$
2	$\frac{F_{N-3}}{F_N}\ell_0$	$\frac{F_{N-2}}{F_N}\ell_0$	$\frac{F_{N-2}}{F_N}\ell_0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$\frac{F_{N-i-1}}{F_N}\ell_0$	$\frac{F_{N-i}}{F_N}\ell_0$	$\frac{F_{N-i}}{F_N}\ell_0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$N-2$	$\frac{1}{F_N}\ell_0$	$\frac{2}{F_N}\ell_0$	$\frac{2}{F_N}\ell_0$
$N-1$ & $\delta>0$	$\frac{1}{F_N}\ell_0$	$\frac{1}{F_N}\ell_{0(-)}^{+}\delta$	$\frac{1}{F_N}\ell_0$ nebo $\frac{1}{F_N}\ell_0+\delta$
$N-1$ & $\delta<0$	$\frac{1}{F_N}\ell_{0(-)}^{+}\delta$	$\frac{1}{F_N}\ell_0$	$\frac{1}{F_N}\ell_0$ nebo $\frac{1}{F_N}\ell_0+\lvert\delta\rvert$

Tabulka 7.2: Průběh FM s výchozím intervalem $I = [a, b]$ s délkou $\ell_0 := b - a$.

Graficky lze FM popsat podobně jako v případě MZŘ. Máme interval $I = [a, b]$ a symetricky ve vzdálenosti $\frac{F_{N-2}}{F_N}\ell_0$ od obou krajních bodů „nasadíme" body λ_1 a μ_1, viz Obrázek 7.34. Porovnáme funkční hodnoty a získáme nový ILM I_1, který již obsahuje jeden z bodů λ_1 nebo μ_1. Překlopíme tento bod podle středu I_1 a získáme druhý bod. Opět porovnáme funkční hodnoty. Celý postup opakujeme až po interval I_{N-2}. Bod obsažený v I_{N-2} je již jeho středem, proto tentokrát pro srovnání použijeme bod získaný posunem o hodnotu δ od středu a získáme poslední ILM I_{N-1}, jehož střed bereme za aproximaci řešení úlohy (7.2.1). Dá se tedy říci, že se MZŘ a FM liší „jen" nasazením bodu λ_1. Je mezi těmito metodami nějaký zásadní rozdíl? Vezmeme-li posloupnost délek výsledných ILM $\{\ell_N - \delta_N\}_{N=1}^{\infty}$, kde $\delta_N \in [0, \ell_0/F_N)$ je buď rovno $|\delta|$ nebo 0, pak

$$\lim_{N\to\infty} \frac{\ell_{N+1} - \delta_{N+1}}{\ell_N - \delta_N} = \lim_{i\to\infty} \frac{\frac{\ell_0}{F_{N+1}}}{\frac{\ell_0}{F_N}} = \lim_{i\to\infty} \frac{F_N}{F_{N+1}} = \frac{1}{\tau}$$

tj. konvergence je stejně jako pro MZŘ lineární s rychlostí $1/\tau$. A v jiných ohledech? Po N vyčísleních bude $\ell^{\mathrm{FM}} = \ell_0/F_N + \delta_N$ a $\ell^{\mathrm{MZŘ}} = \ell_0/\tau^{N-1}$, tedy

$$\frac{\ell^{\mathrm{MZŘ}}}{\ell^{\mathrm{FM}}} = \frac{1}{\sqrt{5}} \frac{\tau^2 + (-1/\tau^2)^N}{1 + \delta_N F_N/\ell_0}.$$

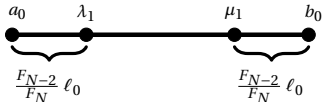
Budeme-li brát δ velmi malé (tj. zanedbatelné), pak

$$\frac{\ell^{\mathrm{MZŘ}}}{\ell^{\mathrm{FM}}} \approx \frac{\tau^2 + (-1/\tau^2)^N}{\sqrt{5}},$$

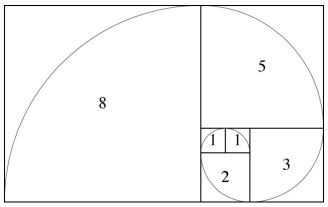
což pro $N = 2$ dává přibližně hodnotu 1,236, pro $N = 3$ to je 1,1458 a s rostoucím N tato hodnota konverguje k $\tau^2/\sqrt{5} \approx 1{,}1708$. Proto apriorní odhad chyby, které se dopustíme při MZŘ, bude vždy větší než v případě FM, což je zcela v souladu s tím, jak jsme se k FM dopracovali.

rostlin. Mnoho rostlin samozřejmě nekopíruje toto matematické schéma zcela přesně. Přesto však matematika a geometrie spojené s Fibonacciho čísly podstatně přispívají k porozumění těmto numerickým jevům ve světě rostlin.

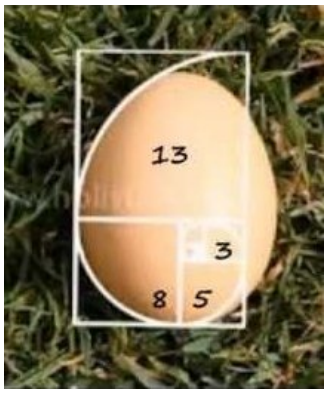
Výskyt Fibonacciho čísel (nejen v matematice) je natolik rozšířený, že existuje dokonce samostatný časopis *Fibonacci Quarterly* zaměřený na jejich studium.



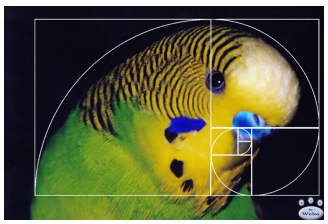
Obrázek 7.34: První krok FM.



Obrázek 7.35: Fibonacciho spirála/čtverec.



Obrázek 7.36: Fibonacciho spirála (podruhé).



Obrázek 7.37: Fibonacciho spirála (potřetí).

Tedy z teoretického hlediska se zdá být FM účinější (přesnější). Jenže z praktického pohledu je MZŘ efektivnější (použitelnější). Lze ji totiž v libovolném kroku přerušit nebo lze několik kroků přidat, aniž bychom vše museli přepočítávat od začátku. Provedeme-li ale FM např. s 10 povolenými vyčísleními a rozhodneme-li se následně pro zvýšení přesnosti tím, že ještě jedno vyčíslení přidáme, pak musíme celý výpočet provést od začátku. MZŘ lze také v jistém smyslu považovat za limitní variantu FM.

Příklad 7.2.16

Uvažme funkci $f(x) = 7x^2 - 6x + 2$ na intervalu $I = [0,1]$. S chybou nejvýše $\varepsilon = 0,14$ určeme aproximaci řešení úlohy (7.2.1) s využitím FM.

Řešení. Maximální požadované chyby dosáhneme při N vyčísleních, pokud $\frac{1}{2}\ell_{N-1} \le 0,14$, tj. $\frac{1}{F_N} + |\delta| \le 0,28$. Zanedbáním δ dostaneme nerovnost

$$\frac{1}{F_N} \le 0,28 \quad \text{neboli} \quad F_N \ge 3,57,$$

což nás přivádí k nejmenší možné hodnotě $F_4 = 5$, tj. $N = 4$. V takovém případě musí $|\delta| \in (0,1/5)$ a současně

$$\frac{1}{F_4} + |\delta| \le 0,28 \quad \text{neboli} \quad |\delta| \le \frac{2}{25}.$$

My si zvolíme $\delta = \frac{1}{20}$.

Začneme s intervalem $I_0 = I = [0,1]$ a v prvním kroku máme body

$$\lambda_1 = 0 + \frac{F_2}{F_4}\ell_0 = \frac{2}{5} \quad \text{a} \quad \mu_1 = 0 + \frac{F_3}{F_4}\ell_0 = \frac{3}{5}$$

s funkčními hodnotami $\frac{18}{25} = f(\lambda_1) < f(\mu_1) = \frac{23}{25}$. Proto novým ILM je $I_1 = [0,3/5]$ s délkou $\ell_1 = \frac{F_3}{F_4}\ell_0 = \frac{3}{5}$.

Ve druhém kroku budeme mít body

$$\lambda_2 = 0 + \frac{F_1}{F_4}\ell_0 = \frac{1}{5} \quad \text{a} \quad \mu_2 = \lambda_1 = \frac{2}{5}$$

s novou funkční hodnotou $f(\lambda_2) = \frac{27}{25}$. Proto novým ILM je $I_2 = [1/5,3/5]$ s délkou $\ell_2 = \frac{F_1}{F_4}\ell_0 = \frac{1}{5}$.

Konečně ve třetím kroku bychom pokračováním v tomto výpočtu dostali

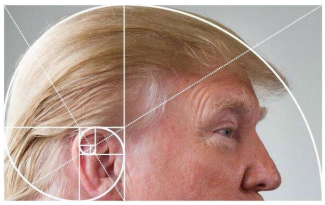
$$\lambda_3 = \mu_2 = \mu_3 = \frac{1}{5} + \frac{F_0}{F_4}\ell_0 = \frac{2}{5} = \mu_3,$$

proto využijeme zvolenou hodnotu δ a budeme mít body

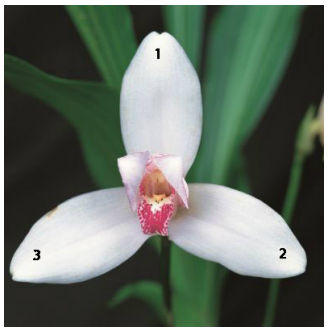
$$\lambda_3 = \frac{2}{5} \quad \text{a} \quad \mu_3 = \frac{2}{5} + \delta = \frac{9}{20}$$

s novou funkční hodnotou $f(\mu_3) = \frac{287}{400}$. Jelikož $f(\lambda_3) > f(\mu_3)$, je posledním ILM $I_3 = [2/5,3/5]$ a aproximací hledaného řešení je $\overline{x} = \frac{1}{2}\left(\frac{2}{5} + \frac{3}{5}\right) = \frac{1}{2}$.

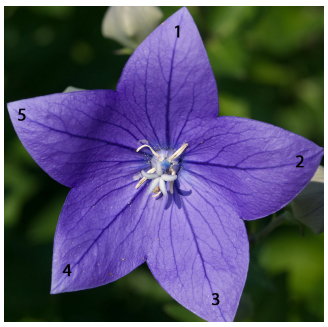
Poznámka 7.2.17. V [510, Kapitola 2] nalezneme velmi zajímavý (alternativní) pohled na některé pojmy, se kterými jsme se setkali v předchozích podkapitolách. Tím prvním je interpretace vyčíslování funkce f v jednotlivých dělících bodech. Funkce f může být v praxi např. výsledkem nějakého experimentu závislého na jednom parametru, takže potřebné vyčíslení f v jednotlivých krocích znamená spuštění celého experimentu znovu s nově nastavenou hodnotou parametru (jednou nebo vícekrát současně v závislosti na zvolené metodě). V takovém případě bude ale samozřejmě složité ověřit unimodální chování celého experimentu, ale v některých případech by mohlo být možné tuto vlastnost přijmout jako základní premisu[48].



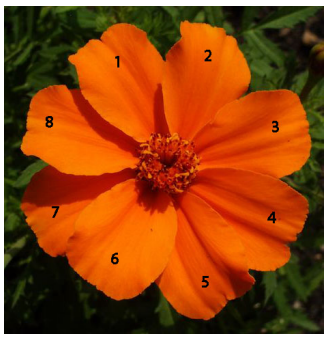
Obrázek 7.38: Fibonacciho spirála (počtvrté – toto už je trochu moc. . .).



Obrázek 7.39: Trillium undulatum neboli trojčetka (3 lístky).



Obrázek 7.40: Platykodon velkokvětý nebo též zvonkovec (5 lístků).



Obrázek 7.41: Cosmos sulphureus neboli krásenka (8 lístků).

MPD bychom v tomto ohledu mohli charakterizovat tak, že všechny experimenty musí proběhnout ve stejnou chvíli, zatímco pro MPI jsou povoleny jen dva (případně tři pro metodu popsanou v Poznámce 7.2.13) experimenty v jednu chvíli a pro MZŘ či FM je to jen jeden experiment (s výjimkou prvního běhu). Předložený výklad je zaměřený na hledání metody, která vede k nejkratšímu z nejdelších ILM v závislosti na volbě dělících bodů. Navzdory odlišnému přístupu v [510, Podkapitoly 2.06 a 2.08] autor dospěje k téže MPD se sudým počtem vyčíslení a MPI jako my. Avšak současně nabízí odlišné pojetí FM, které poskytne ještě kratší poslední ILM. Jak je to možné? Vždyť jsme přeci FM metodu získali minimalizací délky ILM v úloze (7.2.10)! To sice ano, ale pak jsme byli nuceni přidat ještě jeden krok, jehož vliv na délku ILM jsme už nijak více nediskutovali. V tomto ohledu je námi předložena FM *asymetrická*, neboť umístění posledního bodu v této metodě nezachovává symetrii podle středu aktuálního ILM jako ve všech předchozích krocích. Na proti tomu Oliver a Wilde v článku [402] nabízejí *symetrickou* variantu FM, kdy body λ_{N-1} a μ_{N-1} jsou umístěny ve vzdálenosti $\delta/2$ od středu intervalu I_{N-2}, takže délka posledního ILM I_{N-1} bude vždy rovna hodnotě

$$\frac{\ell_0}{F_N} + \frac{F_{N-2}}{F_N}\,\delta.$$

Pak rozdíl délek výsledných ILM v asymetrické a asymetrické variantě bude roven

$$\left(1 - \frac{F_{N-2}}{F_N}\right)\delta = \frac{F_{N-1}}{F_N}\,\delta,$$

což je hodnota, která se s rostoucím N blíží k δ/τ. Tuto asymetrickou metodu získáme umístěním prvních bodů ve vzdálenosti od vzdálenějších krajních bodů

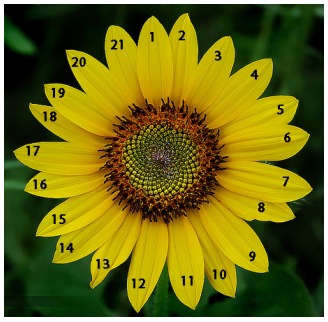
$$\frac{F_{N-1}}{F_N}\,\ell_0 + \frac{(-1)^N\delta}{F_N}$$

od krajních bodů intervalu I_0, takže pro liché N body jsou blíže středu než u asymetrické FM, zatímco pro liché N jsou blíže krajním bodům.

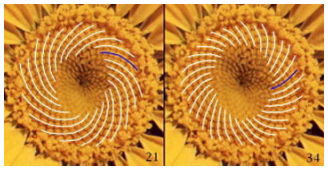
Vyzkoušíme-li symetrickou FM za stejných podmínek jako v Příkladě 7.2.16, dostaneme v prvním kroku body $\lambda_1 = \frac{61}{100}$ a $\mu_1 = \frac{39}{100}$ určující nový ILM $I_1[0,61/100]$. Ve druhém kroku bude $\lambda_2 = \frac{11}{50}$ a $\mu_2 = \lambda_1$ vedoucím k novému ILM $I_2 = [11/50, 61/100]$. Konečně ve třetím kroku získáme body $\lambda_3 = \mu_2$ a $\mu_3 = \frac{11}{25}$ s posledním ILM $I_3 = [39/100, 61/100]$, takže jako aproximace hledaného minima nám opět vyjde $\overline{x} = \frac{1}{2}$. Zdůrazněme, že vzdálenost bodu λ_3 a μ_3 je skutečně rovna hodnotě δ a délka výsledného ILM je $\ell_3 = \frac{11}{50} = \frac{1}{F_4} + \frac{F_2}{F_4}\,\delta$. Stejný výsledek obou variant FM je samozřejmě poměrně sporadická situace. Kdybychom uvážili tutéž funkci i δ a zvolili výchozí interval $I = [-3,2]$, pak by při $N = 7$ nastala úplně jiná situace, která je shrnuta v následující tabulce.

krok	asymetrická FM			symetrická FM		
	λ_i	μ_i	ILM I_i	λ_i	μ_i	ILM
1	$-\frac{23}{21}$	$\frac{2}{21}$	$\left[-\frac{23}{21}, 2\right]$	$-\frac{153}{140}$	$\frac{13}{140}$	$[-\frac{153}{140}, 2]$
2	$\frac{2}{21}$	$\frac{17}{21}$	$\left[-\frac{23}{21}, \frac{17}{21}\right]$	$\frac{13}{140}$	$\frac{57}{70}$	$\left[-\frac{153}{140}, \frac{57}{70}\right]$
3	$-\frac{8}{21}$	$\frac{2}{21}$	$\left[-\frac{8}{21}, \frac{17}{21}\right]$	$-\frac{13}{35}$	$\frac{13}{140}$	$\left[-\frac{13}{35}, \frac{57}{70}\right]$
4	$\frac{2}{21}$	$\frac{1}{3}$	$\left[\frac{2}{21}, \frac{17}{21}\right]$	$\frac{13}{140}$	$\frac{7}{20}$	$\left[\frac{13}{140}, \frac{57}{70}\right]$
5	$\frac{1}{3}$	$\frac{4}{7}$	$\left[\frac{2}{21}, \frac{4}{7}\right]$	$\frac{7}{20}$	$\frac{39}{70}$	$\left[\frac{13}{140}, \frac{39}{70}\right]$
6	$\frac{1}{3}$	$\frac{23}{60}$	$\left[\frac{1}{3}, \frac{4}{7}\right]$	$\frac{3}{10}$	$\frac{7}{20}$	$\left[\frac{3}{10}, \frac{39}{70}\right]$

Tabulka 7.3: Srovnání asymetrické a symetrické FM pro $f(x) = 7x^2 - 6x + 2$ na intervalu $I = [-3,2]$ pro $N = 7$ a $\delta = \frac{1}{20}$.



Obrázek 7.42: Slunečnice (21 lístků)



Obrázek 7.43: Spirály v semeníku slunečnice.

[46] Tohle je bohužel zase trochu *Deus ex machina*...

[47] Tato rovnost je důsledkem d'Ocagneovy identity

$$F_{m+1}F_n - F_m F_{n+1} = (-1)^n F_{m-n-1},$$

pokud zvolíme $m = M$ a $n = M-2$. Tuto rovnost odvodil asi kolem roku 1889 francouzský matematik a inženýr Philbert Maurice d'Ocagne (1862–1938), který byl v roce 1901 prezidentem francouzské matematické společnosti a od roku 1912 profesorem geometrie na pařížské École Polytechnique.

D'Ocagneova rovnost je zobecněním Cassiniho výsledku ukazujícího, že druhá mocnina Fibonacciho čísla se od součinu jeho sousedů vždy liší přesně o jedničku, tj.

$$F_n^2 - F_{n-1}F_{n+1} = (-1)^n$$

z roku 1680 (a znovuobjeveného Simsonem v roce 1753), kterou dostaneme volbou $m = n-1$ a rozšířením rekurentní definice Fibonacciho čísel i na celé $\mathbb{Z}$, což zejména dává $F_{-1} = 0$, $F_{-2} = 1$ a $F_{-3} = -1$ (viz také níže). Cassiniho rovnost hraje klíčovou roli v tzv. Curryho paradoxu z roku 1953, který bývá obvykle ilustrován rozřezáním trojúhelníku ve čtvercové síti (nebo tabulky čokolády) a zpětného složení v jiném pořadí tak, aby jeden dílek chyběl/přebýval (např. $8\times 8 = 64 \neq 65 = 5\times 13$). Lze také odtud získat vzorec pro součet kvadrátů Fibonacciho čísel, tj.

$$\sum_{n=0}^{k} F_n^2 = F_k F_{k+1}.$$

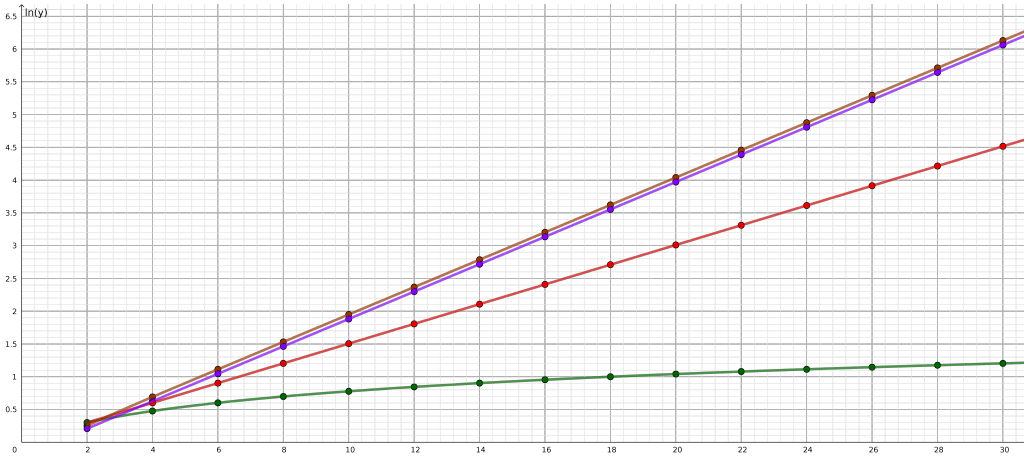
Kromě d'Ocagneovy rovnosti je možné Cassiniho rovnost rozšířit

Z této tabulky vyplývá, že aproximací řešení hledaného pomocí asymetrické FM je $\overline{x} = \frac{19}{42}$, zatímco s využitím symetrické FM získáme $\overline{x} = \frac{3}{7}$, což je dokonce zcela přesné řešení! Náhoda...

Na závěr ještě inspirováni [510, Table 2-1] můžeme čtenáři nabídnout porovnání jednotlivých metod v podobě podílu $\frac{\ell_N}{\ell_0}$ při zanedbání hodnoty δ a sudém počtu vyčíslení, tj. jedná se o faktor zmenšení výchozí intervalu na výsledný ILM po N vyčísleních funkce f, viz Tabulku 7.4 a Obrázek 7.46.

počet vyčíslení	MPD	MPI	MZŘ	FM
2	2	2	1,618	2
4	3	4	4,2358	5
6	4	8	11,089	13
8	5	16	29,030	34
10	6	32	75,999	89
12	7	64	198,96	233
14	8	128	520,86	610
16	9	256	1363,6	1597
18	10	512	3569,7	4181
20	11	1024	9345,3	10946
22	12	2048	24465	28657
24	13	4096	64048	75025
26	14	8192	167673	196418
28	15	16384	438955	514229
30	16	32768	1149151	1346269

Tabulka 7.4: Srovnání poměrného zkrácení výchozího intervalu při daném počtu vyčíslení pro jednotlivé metody.



Obrázek 7.46: Porovnání faktoru zmenšení výchozího intervalu na výsledný ILM po sudém počtu vyčíslení (osa x) pro MPI (zelená barva), MPI (červená barva), MZŘ (fialová barva) a FM (hnědá barva). Hodnoty na ose y jsou v logaritmické stupnici.

▲

také do podoby Catalaniho rovnosti

$$F_{n-k}F_{n+k} - F_n^2 = (-1)^{n+k}F_{k-1}^2$$

z roku 1879 či Vajdovy rovnosti

$$F_n F_{n+i+j} - F_{n+i} Fn + j = \\ = (-1)^n F_{i-1} F_{j-1}$$

publikované v knize z roku 1989, ačkoli ji lze nalézt již v Evermanově úloze zveřejněné v matematickém měsíčníku American Mathematical Monthly z roku 1960. D'Ocagneovu rovnost lze snadno dokázat s pomocí indukce. Začneme však nejprve s důkazem podobné rovnosti

$$F_{m+n} = F_{m+1}F_{n-1} + F_m F_{n-2}. \quad (\star)$$

Pro $n=2$ máme

$$F_{m+2} = F_{m+1}F_1 + F_m F_0 = \\ = F_{m+1} + F_m$$

a pro $n=3$ bude

$$F_{m+3} = F_{m+1}F_2 + F_m F_1 = \\ = 2F_{m+1} + F_m = \\ = F_{m+1} + (F_{m+1} + F_m) = \\ = F_{m+1} + F_{m+2}.$$

Připusťme nyní, že rovnost ($\star$) platí pro libovolné $n \in \mathbb{N}\setminus\{1\}$ a také pro $n+1$. Pak

$$F_{m+n+2} = F_{m+n+1} + F_{m+n} = \\ = F_{m+1}F_n + F_m F_{n-1} + \\ + F_{m+1}F_{n-1} + F_m F_{n-2} = \\ = F_{m+1}(F_n + F_{n-1}) + \\ + F_m(F_{n-1} + F_{n-2}) = \\ = F_{m+1}F_{n+1} + F_m F_n,$$

což ukazuje, že rovnost ($\star$) platí pro všechna $m \in \mathbb{N}$ a $n \in \mathbb{N}\setminus\{1\}$. Rozšíříme-li F_n i pro záporná $n \in \mathbb{Z}$, pak

$$F_{-n} = (-1)^n F_{n-2}$$

pro $n \in \mathbb{N}$. Vskutku, zjevně

$$0 = F_1 - F_0 = F_{-1} = (-1)^1 F_{-1}$$

a

$$1 = F_0 - F_{-1} = F_{-2} = (-1)^2 F_0.$$

Nechť je uvedená rovnost splněna pro libovolné $n \in \mathbb{N}$ a $n+1$, tj. $F_{-n} = (-1)^n F_{n-2}$ a $F_{-n-1} = (-1)^{n+1} F_{n-1}$. Odečtením těchto rovností dostaneme

$$F_{-n-2} = F_{-n} - F_{-n-1} = \\ = (-1)^n (F_{n-2} + F_{n-1}) = \\ = (-1)^n F_n = (-1)^{n+2} F_n.$$

Proto pro $n \in \mathbb{Z}$ máme

$$F_{m-n} = F_{m+1}F_{-n-1} + F_m F_{-n-2} = \\ = (-1)^{n+1}(F_{m+1}F_{n-1} - F_m F_n)$$

neboli

7.2.5 Metoda půlení intervalu II

V předchozích podkapitolách jsme se věnovali metodám nultého řádu, které vyžadují minimum informací o funkci f – pouze funkční hodnoty. Vedle toho jsme požadovali jediné – unimodálnost funkce f na intervalu, ve kterém hledáme minimum. Navzdory těmto slabým předpokladům jsme získali algoritmy s lineární konvergencí, jejíž rychlost *nezávisí* na funkce f. Toto předurčuje tyto metody k širokému uplatnění v praxi. NIcméně naše účelové funkce mají obvykle lepší vlastnosti než unimodální funkce – jsou hladké, tj. třídy $C^1[a,b]$ nebo $C^2[a,b]$ apod. S využitím této skutečnosti můžeme získat další algoritmy, které budou mít navíc některé lepší vlastnosti než metody nultého řádu.

Uvažme úlohu (7.2.1) s funkcí $f \in C^1[a,b]$ takovou, že existuje $x^* \in (a,b)$, ve kterém funkce f' mění znaménko, tj.

$$f'(x) < 0 \quad \text{pro } x \in [a,x^*) \quad \text{a} \quad f'(x) > 0 \quad \text{pro } x \in (x^*,b], \qquad (7.2.18)$$

a tedy f je opět unimodální[49] na $[a,b]$. Kromě tohoto požadavku ještě předpokládejme, že dokážeme určit $f(x)$ a $f'(x)$ v každém bodě $x \in [a,b]$. Potom můžeme modifikovat *metodu půlení intervalu* (viz Podkapitolu 7.2.2) tak, že místo $f\left(\frac{1}{2}(a_k+b_k)\pm\delta\right)$ budeme určovat hodnotu $f'\left(\frac{1}{2}(a_k+b_k)\right)$. V takovém případě mohou pro $x_k := (a_{k-1}+b_{k-1})/2$ nastat 3 možnosti

- $f'(x_k) > 0$, což je díky předpokladu (7.2.18) možné právě tehdy, když $x^* < x_k$, tj. místo intervalu $[a_{k-1}, b_{k-1}]$ bereme za nový ILM $[a_{k-1}, x_k]$, čímž zkrátíme jeho délku na polovinu;
- $f'(x_k) < 0$, takže podobně jako v předchozím případě plyne z (7.2.18), že toto je možné právě tehdy, když $x^* > x_k$, tj. místo intervalu $[a_{k-1}, b_{k-1}]$ bereme za nový ILM $[x_k, b_{k-1}]$, čímž ho opět zkrátíme na polovinu;
- $f'(x_k) = 0$, což je ekvivalentní s $x^* = x_k$, takže jsme našli přesné řešení úlohy (7.2.1).

Tím jsme dospěli k následujícímu algoritmu.

Algoritmický popis Metody půlení intervalu pro diferencovatelné funkce

Zadání: Funkce f, interval $[a,b]$, číslo $N \in \mathbb{N}$, které lze vypočítat i v závislosti na požadované přesnosti $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $a_0 := a$, $b_0 := b$ a $k := 1$. Vypočteme

$$x_k := \frac{a_0 + b_0}{2}.$$

Krok 2: Vyčíslíme $f'(x_k)$. Jestliže $f'(x_k) > 0$, pokračujeme Krokem 3. Jestliže $f'(x_k) < 0$, pokračujeme Krokem 4. Je-li $f'(x_k) = 0$, pokračujeme Krokem 5.

Krok 3: Položíme $a_k := a_{k-1}$, $b_k := x_k$ a pokračujeme Krokem 6.

Krok 4: Položíme $a_k := x_k$, $b_k := b_{k-1}$ a pokračujeme Krokem 6.

Krok 5: Položíme $\overline{x} := x_k$ a máme *přesné* řešení.
KONEC

Krok 6: Je-li $k = N$ pokračujeme Krokem 7. Je-li $k < N$, vypočteme

$$x_{k+1} := \frac{a_k + b_k}{2}.$$

Položíme $k := k+1$ a pokračujeme Krokem 2.

Krok 7: Stanovíme poslední ILM jako $[a_k, b_k]$ a vypočteme $\overline{x} := \frac{a_k+b_k}{2}$.
KONEC

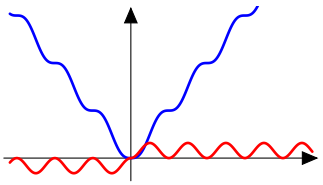
$(-1)^n F_{m-n-1} = F_{m+1}F_n - F_m F_{n+1}$.

[48] Např. z termodynamických zákonů vyplývá unimodalita jisté *funkce volné energie* nebo zákon klesajících výnosů zaručuje „unimodalitu“ *funkce zisku* (zde ale z pohledu maximalizace), viz [510, str. 13].

[49] Avšak nezapomínejme, že unimodálnost a diferencovatelnost implikují pouze, že $f'(x) \le 0$ na $[a, x^*)$ a $f'(x) \ge 0$ na $(x^*, b]$ jako např.

$$f(x) = \begin{cases} -x + \sin x, & x \in [-20, 0], \\ x - \sin x, & x \in [0, 30], \end{cases}$$

viz Obrázek 7.47.



Obrázek 7.47: Funkce f (modrou barvou) a její derivace (červenou barvou).

Jaká je rychlost konvergence? V každém kroku dojde ke zkrácení ILM na polovinu, tudíž pro k vyčísleních máme

$$\ell_k = \frac{\ell_0}{2^k},$$

z čehož plyne

$$\lim_{k\to\infty} \frac{\ell_{k+1}}{\ell_k} = \frac{1}{2},$$

tj. i tato varianta MPI má lineární konvergenci s rychlostí $\frac{1}{2}$, avšak tentokrát v každém kroku využijeme pouze jedno vyčíslení, i když místo f vyčíslujeme f'. Jedná se tedy o metodu, jejíž konvergence je skutečně rychlejší než v případě MZŘ a FM. To ovšem není v rozporu s naší předchozí úvahou o tom, že FM je „nejlepší“. To samozřejmě platí pouze ve srovnání s metodami nultého řádu, ale tentokrát už máme metodu prvního řádu. Tut metodu můžeme považovat za „limitní případ“ MPI s $\delta \to \infty$.

Příklad 7.2.18

Uvažme funkci $f(x) = x^2 + 3\,\mathrm{e}^{-x}$ na intervalu $I = [0,1]$. Pro $N = 4$ určeme aproximaci řešení úlohy (7.2.1) s využitím právě vyložené varianty MPI.

Řešení. Derivace funkce f je $f'(x) = 2x - 3\,\mathrm{e}^{-x}$. Ta vyhovuje podmínce (7.2.18), neboť $f'(0) = -3$ a $f'(1) = 2 - 3\,\mathrm{e}^{-1} > 0$ a současně $f''(x) = 3\,\mathrm{e}^{-x} > 0$, takže první derivace je rostoucí (tj. f je konvexní), a tedy v intervalu $[0,1]$ existuje (jediný) bod x^*, ve kterém f' mění znaménko.

Začneme s intervalem $I_0 = I = [0,1]$ a v prvním kroku máme bod

$$x_1 = \frac{1}{2}$$

s hodnotou $f'(x_1) \approx -0{,}8196 < 0$, takže novým ILM je $I_1 = [1/2, 1]$.

Ve druhém kroku máme bod

$$x_2 = \frac{3}{4}$$

s hodnotou $f'(x_2) \approx 0{,}0829 > 0$, takže novým ILM je $I_2 = [1/2, 3/4]$.

Ve třetím kroku bude

$$x_3 = \frac{5}{8}$$

s hodnotou $f'(x_3) \approx -0{,}3558$, takže novým ILM je $I_3 = [5/8, 3/4]$.

Konečně ve čtvrtém kroku bude

$$x_4 = \frac{11}{16}$$

s hodnotou $f'(x_4) = -0{,}1335$, takže posledním ILM je $I_4 = [11/16, 3/4]$ a aproximací hledaného řešení je bod $\overline{x} = \frac{23}{32} \approx 0{,}71875$.

Poznámka 7.2.19. Samozřejmě je možné oslabit požadavek 7.2.18 např. na situaci, kdy stačí, aby $f'(a) < 0$, $f'(b) > 0$ a f' byla spojitá na (a,b). V takovém případě nám stále MPI poskytne aproximaci stacionárního bodu funkce f v (a,b), tj. $f'(\overline{x}) = 0$. Vskutku, při bližším pohledu na popis MPI vidíme, že generuje intervaly $\ell_0 \supset \ell_1 \supset \ell_2 \supset \dots$ takové, že jejich délky jsou vzájemně poloviční a funkce f' uvnitř mění znaménko ze záporného na kladné. Tento proces skončí pouze ve chvíli, kdy x_k je stacionárním bodem funkce f. Pokud se toto nestane, pak intervaly ℓ_k mají jediný společný bod x^*. Jelikož f' má různá znaménka v libovolném okolí tohoto bodu, nutně platí $f'(x^*) = 0$ díky spojitosti f'. To, jestli se jedná o globální či lokální řešení úlohy (7.2.1), záleží na dalších vlastnostech funkce f (ta může mít extrémů více). Každopádně konvergence k bodu x^* bude lineární s rychlostí 1/2. Odtud můžeme vypozorovat původ této metody – její základ tvoří metoda určená k hledání nulových bodů funkce g na intervalu $[a,b]$ a my jsme ji pouze aplikovali na funkce $g = f'$. Podobně tomu bude i v dalších dvou případech a lze takto implementovat i celou řadu dalších metod určených k hledání řešení nelineární rovnice $g(x) = 0$, viz také [277]. Základní princip původní metody je jistě zřejmý a pro jeho fungování stačí právě pouze, aby $g(a)\,g(b) < 0$ a g byla spojitá. ▲

7.2.6 Interpolační metody

Všechny doposud uvedené metody měly díky požadavku unimodality lineární konvergenci (s různými rychlostmi) a tato konvergence byla globální na intervalu $[a,b]$, tj. bez ohledu na volbě umístění prvního bodu, kterou se jednotlivé metody odlišovali. Můžeme získat i lepší řád konvergence? V následujících dvou metodách uvidíme, že odpověď je kladná. Jejich hlavní myšlenka je založena konstrukci (kvadratického) polynomu, který spolu se svými derivacemi nabývá v daném bodě stejných hodnot jako f', f' a f'' atd. My se v jednom případě omezíme pouze na f, f' a f'', zatímco ve druhém případě to bude pouze f a f'.

Newtonova metoda (NM)[50,51]

Mějme funkci f a předpokládejme, že v bodě x_k je možné určit hodnoty $f(x_k)$, $f'(x_k)$ a $f''(x_k)$. Pak můžeme sestrojit kvadratický (Taylorův) polynom $q(x)$, jehož hodnoty v bodě x_k odpovídají hodnotám funkce f a jejich derivací, tj.

$$q(x) = f(x_k) + f'(x_k)\,(x - x_k) + \frac{f''(x_k)}{2}\,(x - x_k)^2.$$

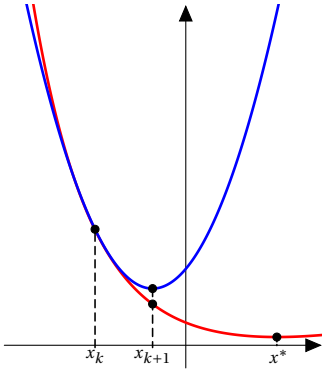
Pak můžeme vypočítat hodnotu x_{k+1} tak, že $q'(x_{k+1}) = 0$, tj. vzhledem k tomu, že

$$q'(x) = f'(x_k) + f''(x_k)\,(x - x_k),$$

bude $q'(x_{k+1}) = 0$ v bodě

$$x_{k+1} := x_k - \frac{f'(x_k)}{f''(x_k)},$$

viz Obrázek 7.48. Tento iterativní proces můžeme dále opakovat v bodě x_{k+1}. Jedná se o aplikaci klasické Newtonovy metody[52] pro řešení rovnice $g(x) = 0$, kde $g = f'$. Všimněme si také, že hodnota x_{k+1} nezávisí na $f(x_k)$. Konverguje tato metoda? Pokud ano, jaký je řád/rychlost konvergence?



Obrázek 7.48: Funkce f (červenou barvou) a odpovídající polynom (červenou barvou) v bodě $x_k = -4$.

Věta 7.2.20 Konvergence NM v $\mathbb{R}$

Nechť $f \in \mathrm{C}^3$ v okolí bodu x^*, který je takový, že $f'(x^*) = 0$ a $f''(x^*) > 0$. Je-li x_0 dostatečně blízko bodu x^*, pak posloupnost $\{x_k\}_{k=0}^{\infty}$ generovaná Newtonovou metodou konverguje k x^* s řádem konvergence (nejméně) $p = 2$.

Důkaz. Nechť $g(x) := f'(x)$. Pak $g(x^*) = 0$ a $g'(x^*) > 0$ a

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}.$$

Jelikož $g = f' \in \mathrm{C}^2$ v okolí x^* a $g'(x^*) > 0$, existují kladná čísla $\delta, k_1, k_2 > 0$ taková, že

$$\left|g''(x)\right| \le k_1 \quad \text{a} \quad g'(x) \ge k_2 \quad \text{pro všechna } x \in \mathcal{O}_\delta(x^*). \tag{7.2.19}$$

Nechť $\rho := \min\{\delta, k_2/k_1\}$ a připusťme, že pro nějaké $k \in \mathbb{N} \cup \{0\}$ je $x_k \in \mathcal{O}_\rho(x^*)$, tj. $x_k \in [x^* - \rho, x^* + \rho]$. Potom $g'(x_k) \ge k_2 > 0$ (neboť $\rho \le \delta$), takže bod x_{k+1} je definován korektně a platí

$$x_{k+1} - x^* = x_k - x^* - \frac{g(x_k)}{g'(x_k)} \overset{g(x^*)=0}{=} x_k - x^* - \frac{g(x_k) - g(x^*)}{g'(x_k)} = \frac{g(x^*) - g(x_k) - g'(x_k)\,(x^* - x_k)}{g'(x_k)}.$$

Čitatel odpovídá Taylorově polynomu prvního řádu funkce g v bodě x_k, tudíž jej můžeme odhadnout zbytkem

$$g(x^*) - g(x_k) - g'(x_k)\,(x^* - x_k) = \frac{g''(\xi)}{2}\,(x^* - x_k)^2 \le \frac{1}{2}\,k_1\,(x^* - x_k)^2$$

dle (7.2.19). Tedy celkem dostáváme, že pro $x_k \in \mathcal{O}_\rho(x^*)$ je

$$\left|x_{k+1} - x^*\right| \le \frac{k_1}{2\,k_2}\left|x_k - x^*\right|^2,$$

což s využitím definice ρ znamená[53], že

[50] Viz také 7.3.2

[51] *Newton's Method*

[52] Pro úlohu $g(x) = 0$ s $g : \mathbb{R} \to \mathbb{R}$ se v Newtonově metodě konstruuje posloupnost $\{x_k\}$ jakožto $x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}$ za předpokladu $g'(x) \neq 0$. Geometricky to znamená, že bod x_{k+1} je průsečíkem tečny ke grafu funkce g v bodě $[x_k, g(x_k)]$ s osou x.

O chování této metody víme následující: (i) Je-li $g \in \mathrm{C}^2[a,b]$, číslo ξ kořenem rovnice $g(x) = 0$ a $g'(\xi) \neq 0$, pak existuje $\delta > 0$ takové, že posloupnost $\{x_k\}$ generovaná Newtonovou metodou konverguje k ξ pro libovolnou počáteční aproximaci $x_0 \in [\xi - \delta, \xi + \delta] \subseteq [a,b]$.

(ii) Je-li $g \in \mathrm{C}^2[a,b]$, má-li rovnice $g(x) = 0$ jediný kořen v $[a,b]$, nemění-li g' a g'' znaménko na $[a,b]$, přičemž $g''(x) \neq 0$ na $[a,b]$, a je-li počáteční aproximace $x_0 \in [a,b]$ zvolena tak, aby $\operatorname{sgn} x_0 = \operatorname{sgn} g''(x)$ na $[a,b]$, pak posloupnost $\{x_k\}$ generovaná Newtonovou metodou konverguje monotónně k onomu jedinému kořenu.

[53] Máme $|x_k - x^*| \le \rho \le \frac{k_2}{k_1}$.

$$\left|x_{k+1}-x^*\right| \le \frac{\left|x_k-x^*\right|}{2}.$$

Proto je-li $x_k \in \mathcal{O}_\rho(x^*)$ pro nějaké $k=\tilde{k}$, pak totéž platí pro všechna $k \ge \tilde{k}$ a posloupnost $\{x_k\}_{k=\tilde{k}}^{\infty}$ konverguje k x^*. Odtud při počáteční podmínce $x_0 \in \mathcal{O}_\rho(x^*)$ plyne[54] tvrzení věty, neboť

$$\frac{\left|x_{k+1}-x^*\right|}{\left|x_k-x^*\right|} \le \frac{k_1}{k_2}$$

implikuje řád konvergence (nejméně) $p=2$. ■

[54] Toto je míněno oním „dostatečně blízko k x^*“.

Poznámka 7.2.21.

(i) Věta 7.2.20 je formulována za požadavku $f''(x^*)>0$, ovšem mohli bychom uvažovat i $f''(x^*)<0$ (či obecněji $f''(x^*)\neq 0$). Pak s jedinou změnou v důkaze v podobě $|g'(x^*)|\ge k_2$ zůstane tvrzení v platnosti a posloupnost $\{x_k\}_{k=0}^{\infty}$ bude konvergovat k bodu x^*, který bude lokálním maximem, tj. NM funguje pro maximalizační i minimalizační problém.

(ii) Oba požadavky $f'(x^*)=0$ a $f''(x^*)>0$ jsou zásadní a nemohou být vynechány. První ukazuje, že x^* je stacionárním bodem f, a jinak to ani být nemůže. Druhá podmínka znamená, že x^* je tzv. *nedegenerovaným* (ostrým) lokálním minimem f, což už ale není obecně nutné pro řešení úlohy (7.2.1). Jenže uvažme např. $f(x)=x^4$, která nabývá minimální hodnoty pro $x^*=0$, avšak $f''(x^*)=0$. V tomto případě

$$x_{k+1}=x_k-\frac{4x_k^3}{12x_k^2}=\frac{2}{3}x_k,$$

což sice konverguje $x_k\to x^*$ pro $k\to\infty$ a libovolnou počáteční aproximaci x_0, ovšem spíše lineárně než kvadraticky, neboť

$$\lim_{k\to\infty}\frac{|x_{k+1}|}{|x_k|}=\lim_{k\to\infty}\frac{\frac{2}{3}|x_k|}{|x_k|}=\frac{2}{3}.$$

Konvergence NM popsaná ve Větě 7.2.20 je poněkud vágní, neboť se hovoří o „x_0 dostatečně blízkém x^*“. Ovšem co to znamená, když ani x^* neznáme? Navíc ani z důkazu není příliš patrné, jak takový bod volit, takže se může snadno stát, že pro některé počáteční aproximace nebude posloupnost generovaná NM vůbec konvergovat. Vskutku, uvažme tentokrát funkci $f(x)=\sqrt{1+x^2}$, která nabývá nejmenší hodnoty pro $x^*=0$. Potom

$$x_{k+1}=\frac{x_k(1+x_k^2)^{3/2}}{\sqrt{1+x_k^2}}=-x_k^3,$$

což ale znamená, že posloupnost $\{x_k\}_{k=0}^{\infty}$ generovaná NM bude konvergovat pouze pro $x_0\in(-1,1)$ s řádem $p=3$, neboť

$$\lim_{k\to\infty}\frac{|x_{k+1}|}{|x_k|^3}=\lim_{k\to\infty}\frac{|-x_k^3|}{|x_k|^3}=1,$$

zatímco pro $x_0\in(-\infty,-1]\cup[1,\infty)$ bude posloupnost divergentní.

Jestliže dokonce $f\in C^3[a,b]$ je taková, že platí tzv. *Fourierovy podmínky*

(i) $f'(a)<0<f'(b)$ a $f''(x)>0$ (případně $f''(x)<0$) pro všechna $x\in[a,b]$,

(ii) f''' nemění na $[a,b]$ znaménko, tj. $f'''(x)\le 0$ pro všechna $x\in[a,b]$ nebo $f'''(x)\le 0$ pro všechna $x\in[a,b]$,

(iii) $|f'(a)/f''(a)|<b-a$ a $|f'(b)/f''(b)|<b-a$,

pak NM konverguje[55] pro libovolnou počáteční aproximaci $x_0\in[a,b]$ ke globálnímu minimu funkce f na $[a,b]$. Nicméně tyto podmínky obvykle nejsou splněny (zejména pro dlouhý interval $[a,b]$) nebo je jejich ověření komplikované. Proto konvergence NM bude na libovolném daném intervalu $[a,b]$ *lokálního* charakteru, tj. za předpokladů Věty 7.2.20 na $[a,b]$ vždy existuje $\delta>0$ takové, že pro každé $x_0\in[x^*-\delta,x^*+\delta]\subseteq[a,b]$ konverguje posloupnost $\{x_k\}_{k=0}^{\infty}$ generovaná NM k bodu x^*.

[55] Konvergence bude dokonce monotónní, zvolíme-li $x_0\in[a,b]$ takové, že $\operatorname{sgn} f(x_0)=\operatorname{sgn} f'''(x_0)$.

(iii) Pro jistotu ještě zdůrazněme, že nemá smysl aplikovat NM na kvadratické funkce – přesněji řečeno s jejím využitím nalezneme hned v prvním kroku nalezneme přesné řešení bez ohledu na počáteční aproximaci. To je pochopitelně způsobeno tím, že v takovém případě je aproximace funkce f Taylorovým polynomem druhého stupně zcela přesná. Např. pro $f(x) = 5x^2 - 4x + 2$ bude $f'(x) = 10x - 4$ a $f''(x) = 10$, takže

$$x_{k+1} = x_k - \frac{10x_k - 4}{10} = \frac{2}{5},$$

což je globální minimum funkce f. ▲

Příklad 7.2.22

Uvažme funkci $f(x) = x^2 + 3\,\mathrm{e}^{-x}$ na intervalu $I = [0,1]$. Pro $N = 3$ určeme aproximaci řešení úlohy (7.2.1) s využitím právě vyložené NM.

Řešení. Jelikož $f'(x) = 2x - 3\,\mathrm{e}^{-x}$, $f''(x) = 2 + 3\,\mathrm{e}^{-x}$ a $f'''(x) = -3\,\mathrm{e}^{-x}$, můžeme snadno ověřit, že na intervalu I nejsou splněny Fourierovy podmínky, neboť $f'(0)/f''(0) = 1$. Nicméně víme, že f' je na intervalu I rostoucí a má v něm nulový bod. Zvolíme např. $x_0 = 1$, pak vzhledem k předpisu

$$x_{k+1} = x_k - \frac{2x - 3\,\mathrm{e}^{-x}}{2 + 3\,\mathrm{e}^{-x}}$$

dostaneme postupně body $x_1 \approx 0,7111900349$, $x_2 \approx 0,7258159305$, $x_3 \approx 0,7258613574$, …, $x_{10} \approx 0,7258613577$, …, $x^* \approx 0,7258613577$, což mimo jiné i ukazuje velmi rychlou konvergenci této metody.

Metoda sečen (MS)[56]

[56] *Secant method*. Někdy chybně označována jako *regula falsi* (*False Position Method*).

Hlavní myšlenka této metody je stejná jako v předchozím případě – aproximace funkce f kvadratickým polynomem, avšak tentokrát místo jednoho bodu s hodnotami f, f' a f'' použijeme dva body s hodnotami f a f', tj. budeme mít metodu prvního řádu a současně tzv. *kvazinewtonovskou*, neboť princip je stejný jako v případě NM, avšak bez využití druhé derivace. Mějme proto dvě předchozí aproximace x_{k-1} a x_k s odpovídajícími hodnotami f a f', pak hodnotu $f''(x_k)$ nahradíme z definice derivace výrazem

$$\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}, \tag{7.2.20}$$

což využijeme v aproximaci f kvadratickým polynomem

$$q(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}\,\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}\,(x - x_k)^2.$$

Čím blíže bude x_{k-1} k x_k, tím více se bude tato aproximace shodovat s tou z NM. Avšak v takové situaci můžeme mít problémy s (numerické) vyčíslením podílu (7.2.20), neboť čitatel i jmenovatel budou mít hodnoty velmi blízké nule. Každopádně nyní máme

$$q'(x) = f'(x_k) + \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}\,(x - x_k)$$

a řešením rovnice $q'(x) = 0$ je bod

$$x_{k+1} := x_k - f'(x_k)\,\frac{x_k - x_{k-1}}{f'(x_k) - f'(x_{k-1})} = x_{k-1} - f'(x_{k-1})\,\frac{x_k - x_{k-1}}{f'(x_k) - f'(x_{k-1})}. \tag{7.2.21}$$

[57] Tečny budou rovnoběžné, ale konstruované v bodech s (obecně) odlišnými funkčními hodnotami.

Z poslední rovnosti je vidět, že ačkoli polynom q nebyl zvolen symetricky vzhledem k[57] x_k a x_{k-1}, tak iterace x_{k+1} (stacionární bod funkce q) je symetrická vzhledem k oběma bodům. Zvolíme-li $g = f'$, pak vlastně hledáme aproximaci nulového bodu této funkce tak, že uděláme sečnu grafu procházející body $[x_{k-1}, g(x_{k-1})]$ a $[x_k, g(x_k)]$ a nalezneme její průsečík s osou x, který označíme jako x_{k+1}.

Věta 7.2.23 Konvergence MS

Nechť $f \in C^3$ v okolí bodu x^* takového, že $f'(x^*) = 0$ a $f''(x^*) > 0$. Je-li počáteční aproximace x_0 a x_1 dostatečně blízko bodu x^*, pak posloupnost $\{x_k\}_{k=0}^{\infty}$ generovaná metodou sečen konverguje k bodu x^* (superlineárně) s řádem konvergence

$$p = \tau = \frac{1+\sqrt{5}}{2} \approx 1{,}618.$$

Důkaz. Nechť $g := f'$ a označme

$$g[a,b] := \frac{g(a)-g(b)}{a-b}.$$

Potom

$$x_{k+1}-x^* = x_k - x^* - g(x_k)\frac{x_k - x_{k-1}}{g(x_k)-g(x_{k-1})} = (x_k-x^*)\left(g[x_k,x_{k-1}] - \frac{g(x_k)}{x_k-x^*}\right)\frac{1}{g[x_k,x_{k-1}]} =$$
$$= (x_k-x^*)\frac{g[x_{k-1},x_k]-g[x_k,x^*]}{g[x_k,x_{k-1}}. \tag{7.2.22}$$

Navíc, pokud označíme

$$g[a,b,c] := \frac{g[a,b]-g[b,c]}{a-c},$$

pak lze (7.2.22) přepsat do tvaru

$$x_{k+1}-x^* = (x_k-x^*)(x_{k-1}-x^*)\frac{g[x_{k-1},x_k,x^*]}{g[x_{k-1},x_k]},$$

a současně z Taylorovy věty plyne, že[58]

$$g[x_{k-1},x_k] = g'(\xi_k) \quad \text{a} \quad g[x_{k-1},x_k,x^*] = \frac{1}{2}g''(\eta_k),$$

kde $\xi_k \in \big(\min\{x_{k-1},x_k\}, \max\{x_{k-1},x_k\}\big)$ a $\eta_k \in \big(\min\{x_{k-1},x_k,x^*\}, \max\{x_{k-1},x_k,x^*\}\big)$ jsou body ležící v intervalech. Tudíž

$$x_{k+1}-x^* = \frac{g''(\eta_k)}{2\,g'(\xi_k)}(x_k-x^*)(x_{k-1}-x^*). \tag{7.2.23}$$

Ze spojitosti g', g'' a $g'(x^*) > 0$ plyne existence čísla $M > 0$ a intervalu $J = [x^*-\varepsilon, x^*+\varepsilon]$ takového, že

$$\left|\frac{g''(\eta_k)}{2g'(\xi_k)}\right| \le M \quad \text{pro všechna } \xi_k, \eta_k \in J.$$

Položíme-li $e_k := M|x_k - x^*|$ a $e_0, e_1 < \min\{1, \varepsilon M\}$, potom platí

$$e_{k+1} \overset{(7.2.23)}{\le} e_k e_{k-1}, \quad \text{a tudíž} \quad e_k \le \min\{1, \varepsilon M\}. \tag{7.2.24}$$

Odtud vyplývá, že

$$e_k \le \kappa^{\tau^k}, \tag{7.2.25}$$

kde[59] $\kappa := \max\{e_0, \sqrt[\tau]{e_1}\}$ a stále $\tau = \frac{1\sqrt{5}}{2}$. Vskutku, pro $k \in \{0,1\}$ je uvedená nerovnost jistě splněna díky volbě čísla κ. Nyní připusťme, že platí pro nějaké $k \in \mathbb{N}$. Pak indukcí a z (7.2.24) dostaneme

$$e_{k+1} \le e_k e_{k-1} \le \kappa^{\tau^k}\kappa^{\tau^{k-1}} = \kappa^{\tau^k+\tau^{k-1}} = \kappa^{\tau^{k-1}(\tau+1)} \overset{\tau^2-\tau-1=0}{=} \kappa^{\tau^{k-1}\tau^2} = \kappa^{\tau^{k+1}}.$$

To znamená, že posloupnost odchylek $\{e_k\}_{k=0}^{\infty}$ se chová alespoň jako $\{\kappa^{\tau^k}\}_{k=0}^{\infty}$, tj. konverguje (superlineárně) s řádem[60] τ. ∎

Poznámka 7.2.24.

(i) Z důkazu je tentokrát patrné, co znamená, „dostatečně blízko“ pro počáteční aproximaci x_0 a x_1, tj. konvergence je opět pouze lokálního charakteru.

[58] Zavedená označení typu $g[x_i,x_j]$ jsou tzv. *poměrné diference*, které se využívají při konstrukci interpolačního polynomu. Postupně platí, že poměrná diference nultého řádu je

$$g[x_i] := g(x_i),$$

prvního řádu

$$g[x_i,x_j] := \frac{g[x_i]-g[x_j]}{x_i-x_j}$$

a obecně m-tého řádu

$$g[x_i,x_{i+1},\dots,x_{i+m}] := \frac{g[x_i,\dots,x_{i+m-1}]-g[x_{i+1},\dots,x_{i+m}]}{x_i-x_{i+m}}.$$

Jsou to symetrické a lineární funkce. Navíc

$$g[x_0,\dots,x_k] = \frac{f^{(k)}(\xi)}{k!},$$

kde

$$\xi \in \big(\min\{x_0,\dots,x_k\}, \max\{x_0,\dots,x_k\}\big).$$

Vskutku, je-li P Lagrangeův interpolační polynom pro funkci g v bodech $x_0,\dots,x_k$, tj.

$$P_k(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \dots + a_k(x-x_0)\cdots(x-x_k) = g[x_0] + \sum_{i=1}^{k} g[x_0,\dots,x_i](x-x_0)\cdots(x-x_i),$$

pak poslední člen je

$$g[x_0,\dots,x_k](x-x_0)\cdots(x-x_k).$$

Označíme-li jako h zbytek po této interpolaci, pak $h := g - P$ a funkce h má $k+1$ nulových bodů $x_0,\dots,x_k$. Postupným aplikováním Rolleovy věty pro funkce $h, h', \dots, h^{(k-1)}$, zjistíme, že v každém z k-tice intervalů (x_{i-1},x_i) existuje y_i splňující $h'(y_i) = 0$, a tedy v každém z $(k-1)$-tice intervalů (y_{i-1},y_i) existuje z_i splňující $h''(z_i) = 0$ atd., až zůstane pouze jeden interval, ve kterém bude existovat číslo ξ splňující $h^{(k)}(\xi) = 0$, tj.

$$0 = h^{(k)}(\xi) = g^{(k)}(\xi) - P_k^{(k)}(\xi)$$

neboli

$$g^{(k)}(\xi) = k!\,g[x_0,\dots,x_k].$$

[59] Neboť

$$\frac{\kappa^{\tau^{k+1}}}{\left(\kappa^{\tau^k}\right)^{\tau}} = 1$$

nebo

$$\frac{\ln\kappa^{\tau^{k+1}}}{\ln\kappa^{\tau^k}} = \frac{\tau^{k+1}}{\tau^k} = \tau.$$

(ii) V jednom kroku metody sečen potřebujeme vyčíslit pouze $f'(x_k)$, neboť $f'(x_{k-1})$ jsme vypočetli již v předchozím kroku, zatímco v Newtonově metodě potřebujeme vždy určit $f'(x_k)$ a $f''(x_k)$. To znamená, že při stejném počtu vyčíslení je[61]

$$e_k = \begin{cases} \leq \kappa^{2^{k+1}} & \text{pro Newtonovu metodu,} \\ \leq \kappa^{\tau^{2k+1}} & \text{pro metodu sečen,} \end{cases}$$

což činí z MS metodu řádu[62] $\tau^2 = \tau + 1 \approx 2,618$, tj. při stejném počtu vyčíslení je (teoreticky) řád MS vyšší než v případě NM.

(iii) Z důkazu Věty 7.2.23 je patrné, že konvergence MS je pouze *lokální* a vyžaduje vhodnou volbu počáteční aproximace. Postačující podmínka je ukryta v (7.2.25) a volbě čísla κ, ovšem její ověření není příliš praktické, neboť vyžaduje znalost x^*. Stanovení praktičtějšího kritéria není úplně snadný úkol, ovšem lze využít např. MPI pro dostatečné přiblížení k x^* a následně v hledání přesnější aproximace pokračovat MS[63].

(iv) NM můžeme považovat za limitní případ MS pro $x_{k-1} \to x_k$, viz (7.2.20). Kombinací MS a MPI dostaneme další kvazinewtonovskou metodu: *metodu regula falsi.* ▲

[59] Toto je míněno oním „dostatečně blízko k x^*".

[60] Jeden krok NM je v tomto ohledu roven dvěma krokům MS

[61] Posloupnosti $\{x_k\}_{k=0}^{\infty}$ generované MS se stejným počtem jako v případě NM můžeme docílit např. tak, že v případě MS budeme uvažovat pouze každý druhý člen.

[62] To, zda jsme již dostatečně blízko x^*, je možné ověřit provedením několika kroků MS a ověřením, zda se již přibližujeme k bodu $f(\overline{x}) \approx 0$.

Příklad 7.2.25

Uvažme opět funkci $f(x) = x^2 + 3\,\mathrm{e}^{-x}$ na intervalu $I = [0,1]$. Pro $N = 5$ určeme aproximaci řešení úlohy (7.2.1) s využitím právě vyložené MS.

Řešení. Zvolíme-li $x_0 = 1$ a $x_1 = 0$, pak podle (7.2.21) máme

$$x_{k+1} = x_k - \frac{(2x_k - 3\,\mathrm{e}^{-x_k})(x_k - x_{k-1})}{2(x_k - x_{k-1}) - 3\,\mathrm{e}^{-x_k} - 3\,\mathrm{e}^{-x_{k-1}}},$$

tj.,

$$x_2 \approx 0,7699490575, \quad \& \quad x_3 \approx 0,7331013907 \quad \&$$
$$x_4 \approx 0,7257946607 \quad \& \quad x_5 \approx 0,7258614592,$$

což je o něco méně přesnější než v případě NM.

Cvičení

7.2.1. Nalezněte přibližné řešení úlohy

$$f(x) = 5x^2 - 7x + 6 \to \min$$

na intervalu $I = [0,1]$ pomocí

a) metody prostého dělení s lichým počtem vyčíslení a chybou nejvýše $\varepsilon = 0,13$;

b) metody prostého dělení se sudým počtem vyčíslení s chybou nejvýše $\varepsilon = 0,18$ a volbou $\delta = 0,1$;

c) metody půlení intervalu s chybou nejvýše $\varepsilon = 0,15$;

d) metody zlatého řezu s chybou nejvýše $\varepsilon = 0,1$ (pracujte s $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$);

e) Fibonacciho metody s chybou nejvýše $\varepsilon = 0,08$, a volbou $\delta = -0,02$.

řešení: a) $N = 7$ & $\overline{x} = 0,75$; b) $N = 6$ & $\overline{x} = 0,625$; c) $N = 6$ & $\overline{x} = 0,65$; d) $N = 5$ & $\overline{x} = 0,69097$; e) $N = 5$ & $\overline{x} = 0,6875$

7.2.2. S využitím metody zlatého řezu určete přibližné řešení úlohy

$$f(x) = 3x^2 - 4x + 2 \to \min$$

na intervalu $I = [0,1]$ s chybou nejvýše $\varepsilon = 0,08$. Při výpočtu pracujte s $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$.

řešení: $\overline{x} = 0,691$

7.2.3. Nalezněte přibližné řešení úlohy

$$f(x) = 7x^2 - 11x + 2 \to \min$$

na intervalu $I = [0,1]$ pomocí

a) metody zlatého řezu s $N = 5$ (pracujte s hodnotou $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0{,}618$). Bez znalosti přesného řešení také odhadněte maximální chybu, které se při této aproximaci dopustíte (odhad vyjádřete pomocí čísla $1/\tau$);

b) Fibonacciho metody s chybou nejvýše $\varepsilon = 3/20$ a volbou $\delta = 1/10$.

řešení: a) $\overline{x} = 0{,}781$ s maximální chybou $\frac{1}{2\tau^4}$; b) $\overline{x} = 3/4$

7.2.4. Nalezněte přibližné řešení úlohy

$$f(x) = 9x^2 - 10x + 16 \to \min$$

na intervalu $I = [0,1]$ pomocí

a) metody zlatého řezu s $N = 5$ (pracujte s $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0{,}618$). Bez znalosti přesného řešení také odhadněte maximální chybu, které se při této aproximaci dopustíte (odhad vyjádřete pomocí čísla $1/\tau$);

b) Fibonacciho metody s chybou nejvýše $\varepsilon = 1/10$ a volbou $\delta = 1/40$.

řešení: a) $\overline{x} = 0{,}545$ s maximální chybou $\frac{1}{2\tau^4}$; b) $\overline{x} = 9/16$

7.2.5. S využitím metody půlení intervalu určete přibližné řešení úlohy

$$f(x) = \mathrm{e}^{x^2 - 5\ln(x+1)} \to \min$$

na intervalu $I = [0,2]$ s chybou nejvýše $\varepsilon = 0{,}25$ a volbou $\delta = 0{,}1$.

řešení: $\overline{x} = 1{,}1125$

7.2.6. S využitím metody půlení intervalu určete přibližné řešení úlohy

$$f(x) = \sqrt{\mathrm{e}^x - 5\ln x + 1} \to \min$$

na intervalu $I = [1/2,2]$ s chybou nejvýše $\varepsilon = 1/5$ a volbou $\delta = 1/10$.

řešení: $\overline{x} = 1{,}33125$

7.2.7. S využitím metody půlení intervalu určete přibližné řešení úlohy

$$f(x) = \mathrm{e}^{4x^2 - 3x + 2} \to \min$$

na intervalu $I = [0,2]$ s chybou nejvýše $\varepsilon = 1/5$ a volbou $\delta = 1/10$.

řešení: $\overline{x} = 0{,}38125$

7.2.8. S využitím Fibonacciho metody určete přibližné řešení úlohy

$$f(x) = -\ln(5\sin x - x^2 + 1) \to \min$$

na intervalu $I = [0,2]$ s chybou nejvýše $\varepsilon = 1/5$ a pro $\delta = 1/10$.

řešení: $\overline{x} = 9/8$

7.2.9. S využitím Fibonacciho metody určete přibližné řešení úlohy

$$f(x) = -(x-1)^4 + \mathrm{e}^{-x} \to \min$$

na intervalu $I = [0,3]$ s chybou nejvýše $\varepsilon = 0{,}21$ a volbou $\delta = 0{,}03$.

řešení: $\overline{x} = 1{,}3275$

7.2.10. S využitím Fibonacciho metody určete přibližné řešení úlohy

$$f(x) = \sin(2x^2 - 6x + 1) \to \min$$

na intervalu $I = [0,1]$ s chybou nejvýše $\varepsilon = 7/100$ a pro $\delta = 1/20$.

řešení: $\overline{x} = 273/520$

7.2.11. S využitím metody půlení intervalu určete přibližné řešení úlohy

$$f(x) = \frac{e^{-x} - x}{e^{x} - x} \to \min$$

na intervalu $I = [-1, 4]$ s volbou $N = 8$ a $\delta = 1/5$.

řešení: $\overline{x} = 1,35625$

7.2.12. S využitím Fibonacciho metody určete přibližné řešení úlohy

$$f(x) = e^{x^2 - \sin x} \to \min$$

na intervalu $I = [-1, 1]$ s chybou nejvýše $\varepsilon = 1/5$ a pro $\delta = 1/20$.

řešení: $\overline{x} = 2/5$

7.2.13. S využitím metody zlatého řezu určete přibližné řešení úlohy

$$f(x) = x \sin(2x - 1) \to \min$$

na intervalu $[0, 1]$ s chybou nejvýše $\varepsilon = 1/10$. Při výpočtu pracujte s hodnotou $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$.

řešení: $\overline{x} = 0,218$

7.2.14. S využitím metody zlatého řezu určete přibližné řešení úlohy

$$f(x) = \operatorname{tg}(x^2 - 3x + 1)$$

na intervalu $[1, 3]$ s chybou nejvýše $\varepsilon = 1/4$. Při výpočtu pracujte s hodnotou $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$.

řešení: $\overline{x} = 1,5277$

7.2.15. S využitím metody zlatého řezu určete přibližné řešení úlohy

$$f(x) = \sin(\sqrt{x} + \ln x + 1)$$

na intervalu $[3, 6]$ s chybou nejvýše $\varepsilon = 1/4$. Při výpočtu pracujte s hodnotou $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$.

řešení: $\overline{x} = 4,6352$

7.2.16. S využitím metody zlatého řezu určete přibližné řešení úlohy

$$f(x) = \sin(x + 1)\cos(x) + x^2$$

na intervalu $[-2, 1]$ s chybou nejvýše $\varepsilon = 0,23$. Při výpočtu pracujte s hodnotou $\frac{1}{\tau} = \frac{2}{1+\sqrt{5}} = 0,618$.

řešení: $\overline{x} = 0,3645$

7.2.17. Vyřešte úlohy dané v Příkladech 7.2.1–7.2.16 také pomocí zbývajících numerických metod (je-li to za daných podmínek možné).

7.2.18. Pomocí různých numerických metod zejména pro $N = 3, 4, 5, 6$ (s vhodným δ, např. $\delta = \pm 1/50$, $\pm 1/20$, $\pm 1/10$) nalezněte přibližné minimum funkce $f(x)$ na intervalu I, kde (hodnota x^* udává přesné řešení)

a) $f(x) = x^2 - 4x + 1$, $I = [1, 4]$, $(x^* = 2)$,

b) $f(x) = x^5 - 5x^3 - 20x + 5$, $I = [0, 3]$, $(x^* = 2)$,

c) $f(x) = 2x^3 - 21x^2 + 60x + 50$, $I = [2, 6]$, $(x^* = 5)$,

d) $f(x) = x^2 + \frac{1}{x}$, $I = [0, 1]$, $(x^* = \frac{1}{\sqrt[3]{2}})$,

e) $f(x) = x^2 + \sqrt{x} - 15$, $I = [0, 4]$, $(x^* = 0)$,

f) $f(x) = e^{x} + e^{-x}$, $I = [-1, 1]$, $(x^* = 0)$,

g) $f(x) = 1 - x\,\mathrm{e}^{-x^2}$, $I = [0,1]$, $(x^* = \frac{1}{\sqrt{2}})$,

h) $f(x) = \frac{x}{\ln x}$, $I = [2,3]$, $(x^* = \mathrm{e})$,

i) $f(x) = -4\sin x\,(1+\cos x)$, $I = [0,\pi/2]$, $(x^* = \pi/3)$,

j) $f(x) = \frac{x^3-54}{x}$, $I = [-4,-2]$, $(x^* = -3)$,

k) $f(x) = \arccos\left(\frac{2x}{1+x^2}\right)$, $I = [0,2]$, $(x^* = 1)$,

l) $f(x) = x - 2\sin x$, $I = [0,2\pi]$, $(x^* = \frac{\pi}{3})$,

m) $f(x) = |2-x| + |5-4x| + |8-9*x|$, $I = [0,3]$, $(x^* = \frac{8}{9})$,

n) $f(x) = \mathrm{e}^{-x} - 1 - \frac{1}{1+x}$, $I = [0,10]$, $(x^* \approx 2,51286)$,

o) $f(x) = \frac{x^2}{\ln(x+1)}$, $I = [-1,0]$, $(x^* \approx -0,71533)$,

p) $f(x) = -\sin^6 x\,\mathrm{tg}(1-x)\,\mathrm{e}^{30x}$, $I = [0,1]$, $(x^* \approx 0,97066)$,

q) $f(x) = \frac{13}{60}(3 - x\,\mathrm{arctg}\,\frac{3}{x}) - \frac{27}{36+4x^2}$, $I = [1,2]$, $(x^* \approx 1,44259)$,

r)
$$f(x) = \begin{cases} -x/2, & x \le 2, \\ x-3, & x > 2, \end{cases} \qquad I = [0,3], \qquad (x^* = 2),$$

s)
$$f(x) = \begin{cases} -x/3, & x \le 2, \\ x-4, & x > 2, \end{cases} \qquad I = [0,3], \qquad (x^* = 2),$$

t)
$$f(x) = \begin{cases} x^2, & x \le -2, \\ x^4 - 3, & x > -2, \end{cases} \qquad I = [-1,2], \qquad (x^* = 0).$$

Určete také maximální chybu, které se touto aproximací dopustíte.

7.2.19. Pomocí různých metod aproximujte minimum funkce f na intervalu I s nejvýše $\varepsilon = 7\%$ chybou vzhledem k délce intervalu I, přičemž funkce f a interval I jsou dány v Příkladech 7.2.1–7.2.18. Uvažte také jiné hodnoty ε.

7.2.20. Pomocí různých metod aproximujte minimum funkce f na intervalu I s chybou nejvýše $\varepsilon = 1/15$, přičemž funkce f a interval I jsou dány v Příkladu 7.2.18. Uvažte také jiné hodnoty ε.

7.3 Numerické metody nepodmíněné optimalizace v $\mathbb{R}^n$

Nyní se budeme věnovat několika základním numerickým metodám pro řešení úlohy

$$f(x) \to \min, \qquad x \in \mathbb{R}^n, \tag{7.3.1}$$

kde $f : \mathbb{R}^n \to \mathbb{R}$ je (jednou/dvakrát/třikrát) spojitě diferencovatelná funkce. Zejména si ukážeme dvě metody 1. řádu

- metoda největšího spádu (MNS),
- metoda sdružených gradientů (MSG),

a jednu metodu 2. řádu

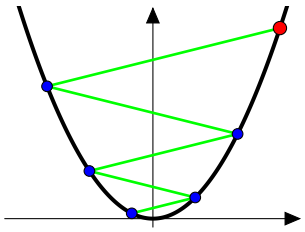
- Newtonova metoda (NM).

[64] To tedy znamená řešení (obecně) nelineární rovnice pro nalezení stacionárního bodu a následné ověření, zda v něm nastává minimum.

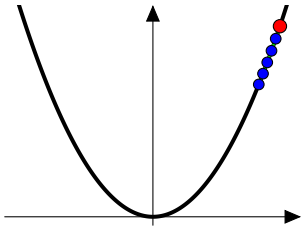
[65] Možná až nekonečné...

[66] Dokonce by body x_k nemuseli ani mířit tím správným směrem, tj. odpovídající funkční hodnoty by mohly přestat být klesající.

[67] První podmínka zabraňuje tomu, abychom byli omezeni konečným součtem délek všech kroků, tj. aby minimalizující posloupnost mohla „dosáhnout" x^* i v případě, kdy je počáteční aproximace $x^{[0]}$ daleko od x^*. Druhá podmínka, která je splněna, pokud posloupnost $\{\alpha_k\}$ konverguje dostatečně rychle k nule, ve většině praktických případů zajišťuje konvergenci $x^{[k]} \to x^*$.



Obrázek 7.49: Zjednodušená ilustrace příliš dlouhé délky kroků při klesající volbě. Při konstantní volbě s příliš dlouhým krokem bychom jenom „skákali" mezi dvěma různými body. Výchozí bod je červenou barvou.



Obrázek 7.50: Zjednodušená ilustrace příliš krátké délky kroků při klesající či konstantní volbě. Výchozí bod je červenou barvou.

[68] *Backtracking line search.*

[69] Tato podmínka je splněna zejména v případě, kdy h_k je *spádovým směrem* funkce f v bodě $x^{[k]}$, tj. existuje $\varepsilon > 0$ takové, že $f(x^{[k]} + \lambda h_k) < f(x^{[k]})$ pro všechna $\lambda \in (0, \varepsilon)$. S využitím Taylorova rozvoje 1. řádu snadno zjistíme, že podmínka $\langle \operatorname{grad} f(x^{[k]}), h_k \rangle < 0$ je dostačující a nerovnost $\langle \operatorname{grad} f(x^{[k]}), h_k \rangle \le 0$ nutná pro to, aby h_k byl spádovým směrem, viz také Podkapitolu 8.1. V takovém případě nalezneme vhodné λ po konečném počtu „drobení".

Tyto metody jsou založeny (přeneseně i NM) na konstrukci tzv. *minimalizující posloupnosti* $\{x^{[k]}\}$, kde

$$x^{[k+1]} = x^{[k]} + \alpha_k h_k,$$

přičemž $\alpha_k \in \mathbb{R}$ se nazývá *délka k-tého kroku* a vektor $h_k \in \mathbb{R}^n$ je *směr k-tého kroku*. Zatímco jednotlivé metody se *zásadně liší volbou vektorů* h_k, tak volba délky kroku α_k může být stejná — jen je nutné zaručit, aby kroky nebyly příliš dlouhé ani příliš krátké. My se budeme držet tzv. *přesné minimalizace*, tj. α_k určíme jako přesné řešení jisté jednorozměrné minimalizační úlohy[64]. Toto samozřejmě není z praktického pohledu příliš efektivní, neboť nalezení přesného řešení může vyžadovat nemalé[65] úsilí. Proto se místo přesné minimalizace používají jiné nástroje, viz mj. [57, str. 29–33]. Nejblíže k přesné minimalizaci má asi *přibližná minimalizace* využívající některou z jednorozměrných numerických metod (např. z těch z Podkapitoly 7.2) nebo *podmíněná minimalizace*, což je přesná minimalizace ale řešení hledáme pouze na nějakém předem daném intervalu, tj. hledáme $\alpha_k \in [0, \tilde{\alpha}]$. Nicméně tyto alternativy jsou stále výpočetně poměrně náročné a/nebo vyžadují „hezké" funkce. Naproti tomu nejjednodušší alternativou je *konstantní volba* $\alpha_k \equiv \alpha$, která však skýtá značné úskalí v tom, že v případě moc dlouhého kroku může vést k velké chybě[66] a v případě moc krátkého kroku k hodně pomalé konvergenci, přičemž rozlišení mezi „dlouhým" a „krátkým" krokem je individuální a závisí především na dané funkci. Drobné vylepšení přináší *klesající volba* $\alpha_k \to 0$ s požadavkem[67] $\sum_{k=1}^{\infty} \alpha_k = \infty$ a případně i $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$, a však u této možnosti přetrvávají problémy s apriorní volbou jednotlivých kroků nezohledňujících aktuální průběh výpočtu, viz Obrázky 7.49–7.50. Proto byla/je již od 60. let 20. století věnována pozornost nalezení nějakého kompromisu mezi všemi těmito možnostmi a výsledkem je tzv. *zpětná minimalizace*[68], která může mít více podob. Základním principem je tzv. *drobení kroku*, kde si nejdříve zvolíme $\beta > 0$ a $\gamma \in (0,1)$ (*koeficient drobení kroku*). Položíme $\lambda = \beta$ a testujeme, zda

$$f(x^{[k]} + \lambda h_k) \le f(x^{[k]}). \tag{7.3.2}$$

Pokud je uvedená nerovnost splněna, volíme $\alpha_k = \beta$, v opačném případě položíme $\lambda = \gamma\beta$ a opět otestujeme nerovnost (7.3.2). V tomto procesu pokračujeme tak dlouho, dokud nenalezneme[69] λ splňující (7.3.2). Jinými slovy, volíme $\alpha_k = \gamma^{j_k}\beta$, kde $j_k \in \mathbb{N} \cup \{0\}$ je nejmenší číslo, pro které platí

$$f(x^{[k]} + \gamma^{j_k}\beta h_k) \le f(x^{[k]}).$$

Tato podmínka je však velmi slabá, neboť bude-li funkce f klesající ve směru h_k, pak bude stačit $j_k = 0$, a tedy „skončíme" opět u konstantní volby délky kroku. Proto i pro tuto metodu existují různé modifikace, ve kterých je popsaným způsobem hledáno λ takové, že

(i) pro dané $\varepsilon_1 \in (0,1)$ platí dokonce *Armijova podmínka*[70]

$$f(x^{[k]} + \lambda h_k) - f(x^{[k]}) \le \varepsilon_1 \lambda \langle \operatorname{grad} f(x^{[k]}), h_k \rangle, \tag{7.3.3}$$

která bývá doplněna o požadavek

$$f(x^{[k]} + \lambda\eta h_k) - f(x^{[k]}) > \varepsilon_1 \lambda\eta \langle \operatorname{grad} f(x^{[k]}), h_k \rangle \tag{7.3.4}$$

pro dané $\eta > 1$ zaručující, že α není příliš malé;

(ii) pro dané $0 < \varepsilon_1 < \varepsilon_2 < 1$ platí *Goldsteinova podmínka*

$$\varepsilon_2 \lambda \langle \operatorname{grad} f(x^{[k]}), h_k \rangle \le f(x^{[k]} + \lambda h_k) - f(x^{[k]}) \le \varepsilon_1 \lambda \langle \operatorname{grad} f(x^{[k]}), h_k \rangle;$$

(iii) pro dané $0 < \varepsilon_1 < \varepsilon_2 < 1$ a $\varepsilon_3 \ge 0$ platí *Wolfeho podmínka* zahrnující (7.3.3) a

$$\varepsilon_2 \langle \operatorname{grad} f(x^{[k]}), h_k \rangle \le \langle \operatorname{grad} f(x^{[k]} + \lambda h_k), h_k \rangle \le \varepsilon_3 \left| \langle \operatorname{grad} f(x^{[k]}), h_k \rangle \right|. \tag{7.3.5}$$

Nerovnosti (7.3.5) představují jakési *gradientní podmínky* a měly by zaručit, že délka kroku není příliš krátká ani příliš dlouhá. Ve *slabé Wolfeho podmínce* se volí $\varepsilon_3 = \infty$, zatímco v *silné Wolfeho podmínce* je $\varepsilon_2 = \varepsilon_3$. Kdybychom navíc zvolili $\varepsilon_3 = \varepsilon_2 = 0$, měli bychom zaručenu ortogonálnost směru h_k a gradientu funkce f v následujícím kroku.

Navzdory různým přístupům k volbě h_k mají všechny metody jednu důležitou vlastnost společnou: jedná se o tzv. *spádové metody*. To znamená, že volbou h_k a α_k dostaneme minimalizující posloupnost $\{x^{[k]}\}$, která generuje klesající posloupnost funkčních hodnot $\{f(x^{[k]})\}$. Jinými slovy, funkční hodnoty $f(x_k)$ jsou klesající s rostoucím $k \to \infty$.

Pro jednotlivé metody si představíme nejenom jejich algoritmický popis, ale podrobněji se zaměříme i na analýzu některých jejich vlastností (zejména na konvergenci a její rychlost). Proč, když v očích většiny lidí věnujících se této oblasti nelineárního programování je budování nějakého algoritmu nebo jeho analýza spojena především s praktickým dopadem na řešení různých problémů[71]? Při velmi extenzivním pohledu, který je odmítavý k roli matematické analýzy, bychom mohli říci, že funkčnost a chování algoritmu musí být pro danou třídu problémů tak jako tak ověřena praktickým experimentem. Takže pokud na základě toho algoritmus vypadá slibně, proč ztrácet čas nějakým analyzováním konvergence? Kromě toho existuje celá řada dalších důležitých otázek, které mohou být hlubší analýzou jen obtížně zodpovězeny. Např. jaký je vliv zaokrouhlovacích chyb, co se bude dít v případě existence několika minim, kdy algoritmus ukončit a jak dobrou aproximaci získáme? To ale asi není úplně správný pohled na věc. Na proti tomu analýza konvergence je čistě matematická záležitost: vlastnosti metod jsou podávány v maximální možné míře jako matematická tvrzení. Obecná a široce platná tvrzení či jednoduché a elegantní důkazy jsou zde na prvním místě. Toto lze ospravedlnit tím, že jednoduchá tvrzení a důkazy jsou snadněji srozumitelné a obecná tvrzení platí nejen pro aktuální problémy ale i pro ty, které se mohou objevit v budoucnu. Pravda bude asi někde mezi těmito dvěma pohledy, jelikož jednoduchost není vždy slučitelná s relevantností a široká platnost je vykoupena jen obtížně ověřitelnými či nedocenitelnými předpoklady. Nicméně matematickým analyzováním často odhalíme (a mnohdy snad i vysvětlíme) zásadní chyby algoritmu, které bychom při pouhém experimentování mohli minout. Tyto nedostatky a jejich analýza nám pak mohou ukázat cestu k modifikacím zlepšujícím chování algoritmu, které jsou uzpůsobeny řešení konkrétních problémů. Současně můžeme analyzováním efektivněji vymezit třídu vhodných problémů, které lze tímto algoritmem řešit, než jak bychom docílili „pouhým“ experimentováním. Náš matematický přístup by v této oblasti ale měl být usměrňován praktickým využitím, přičemž rovnováha mezi teorií a praxí je velmi křehká, subjektivní a závislá na konkrétním problému. Kromě toho, že matematický důkaz přináší nedocenitelný vhled do samotného algoritmu, můžeme čtenáři předložit ještě tři důvody, proč trvat na rigorózní analýze:

(i) vymezení rozsahu použitelnosti různých metod, zejména s ohledem na požadavky na účelovou funkci (diferencovatelnost, konvexnost, singulární či nesingulární minima);

(ii) srovnání kvalitativního chování jednotlivých metod, např. závislost konvergence na volbě počáteční aproximace, konvergence minimalizující posloupnosti či jen odpovídající posloupnosti funkčních hodnot apod., což může být doplněno i různými experimenty;

(iii) poskytnutí návodu pro výběr vhodného algoritmu z řady těch, které jsou pro daný problém použitelné, i s přihlédnutím k rychlosti konvergence (asi není úplně šťastné zvolit algoritmus, který sice konverguje, ale pro daný problém jen velmi pomalu).

Zkušenosti ukazují, že bez dobré analýzy algoritmu může být obtížné vybrat ten správný, aniž bychom kvůli tomu neprovedli mnoho časově/finančně nákladných experimentů. Pro nás bude dostačující, pokud se omezíme pouze na kvadratické funkcím, k čemuž nás vedou především dva důvody. Tím prvním je, že tyto funkce umožňují poměrně hlubokou analýzu a současně je lze považovat za dostatečně reprezentativní. Druhým důvodem je, že nebude-li se uvažovaná metoda chovat „hezky“ pro takovéto funkce, pak je jen velmi malá naděje, že se to změní v případě obecnějších funkcí[72].

Než si ukážeme výše avizované metody, nastíníme si ještě dvě metody nultého řádu, které mohou být užitečné zejména ve chvíli, kdy výpočet derivací je složitý či dokonce nemožný.

(i) *Metoda náhodného hledání*[73,74]: tuto metodu lze považovat za tu nejjednodušší, ale asi také za tu nejméně sofistikovanou. Potřebujeme pouze nějaké kritérium, na jehož zá-

[70] Tato podmínka je splněna např. v situaci, kdy funkce f je diferencovatelná a Lipschitzovská s konstantou $M > 0$, tj. pro každé $x, y \in \mathbb{R}^n$ platí $\|f(x) - f(y)\| \le M\|x - y\|$, jestliže h_k je zvoleno tak, že splňuje $\langle \operatorname{grad} f(x^{[k]}), h_k \rangle < 0$. V takovém případě vyhovuje podmínce (7.3.3) každé $\lambda \in (0, \varepsilon)$ pro

$$\varepsilon = -\frac{(1-\varepsilon_1)\langle \operatorname{grad} f(x^{[k]}), h_k\rangle}{M\|h_k\|^2}.$$

Vskutku, z Lagrangeovy věty o střední hodnotě a Cauchyho-Schwarzovy nerovnosti vyplývá, že pro $\lambda \in (0, \varepsilon)$ a vhodné $\vartheta \in (0,1)$ máme

$$\begin{aligned}
&f(x^{[k]} + \lambda h_k) - f(x^{[k]}) = \\
&= \langle \operatorname{grad} f(x^{[k]} + \lambda\vartheta h_k), \lambda h_k\rangle = \\
&= \langle \operatorname{grad} f(x^{[k]}), \lambda h_k\rangle + \\
&\quad + \langle \operatorname{grad} f(x^{[k]} + \lambda\vartheta h_k) - \\
&\qquad - \operatorname{grad} f(x^{[k]}), \lambda h_k\rangle \le \\
&\le \langle \operatorname{grad} f(x^{[k]}), \lambda h_k\rangle + \\
&\quad \|f(x^{[k]} + \lambda\vartheta h_k) - f(x^{[k]})\| \times \\
&\qquad \times \|\lambda h_k\| \le \\
&\le \lambda\big(\langle \operatorname{grad} f(x^{[k]}), h_k\rangle + \\
&\quad + M\|\lambda\vartheta h_k\| \times \|h_k\|\big) \le \\
&\le \lambda\big(\langle \operatorname{grad} f(x^{[k]}), h_k\rangle + \\
&\qquad + M\lambda\|h_k\|^2\big) \le \\
&\le \lambda\big(\langle \operatorname{grad} f(x^{[k]}), h_k\rangle - \\
&\quad - (1-\varepsilon_1)\langle \operatorname{grad} f(x^{[k]}), h_k\rangle\big) = \\
&= \lambda\varepsilon_1 \langle \operatorname{grad} f(x^{[k]}), h_k\rangle.
\end{aligned}$$

[71] Viz [57, str. 41–43].

[72] Tím samozřejmě netvrdíme, že takové funkce neexistují, ale jsou to spíše singularity než nějaké široké třídy funkcí.

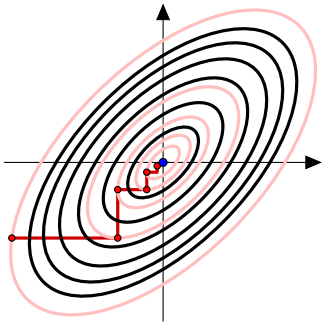
[73] *Random Search Algorithm*

[74] Poněkud sofistikovanější variací na toto náhodné hledání může být *Nelderova–Meadova metoda* (též *metoda klesajících simplexů*), která patří mezi tzv. *komparativní metody* založené na porovnávání funkčních hodnot ve vybraných bodech $\mathbb{R}^n$. V této metodě jsou postupně konstruovány simplexy (tj. pro $n = 2$ trojúhelníky a pro $n = 3$ čtyřstěny), přičemž v každém kroku je na základě porovnání funkčních hodnot v jednotlivých vrcholech aktuálního simplexu sestrojen podle jistých pravidel simplex nový.

kladě (náhodně) vybereme kandidáty pro vyčíslení a jako člena minimalizující posloupnosti vezmeme bod s nejmenší funkční hodnotou. Tato metoda není příliš vhodná pro ruční výpočty, neboť vyžaduje netriviální množství bodů, ale s využitím počítače by již mohla být (alespoň trochu) funkční. Jedná se o tzv. *stochastickou metodu*, u které je nutné ukázat, že s rostoucím počtem bodů/kroků se pravděpodobnost nalezení řešení blíží k jistotě.

(ii) *Metoda souřadnicového spádu*[75]: hlavní myšlenkou této metody je minimalizace funkce f podél souřadných os, tj. ve směru vektorů kanonické báze $e_1,\dots,e_n$, přičemž v každém kroku je postupně řešeno n jednorozměrných minimalizačních úloh. Z výchozího bodu $x^{[0]}$ „vyrazíme“ ve směru e_1 a hledáme bod s nejmenší funkční hodnotou. Z tohoto bodu vyrazíme ve směru e_2, atd. až po e_n, čímž získáme bod $x^{[1]}$. Celý postup pak opakujeme s tímto novým výchozím bodem, viz Obrázek 7.51.

[75] *(Cyclic) Coordinate Descent Method*



Obrázek 7.51: Ilustrace Metody souřadnicového spádu.

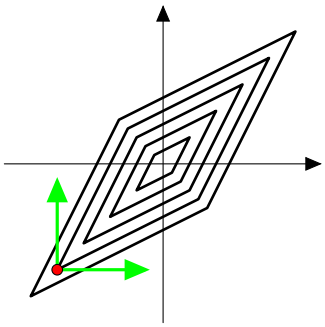
Algoritmický popis Metody souřadnicového spádu

Zadání: Funkce f, výchozí bod $x^{[0]} \in \mathbb{R}^n$, ukončovací pravidlo (např. $\| x^{[k+1]} - x^{[k]} \| \le \varepsilon$) a volba tolerance $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $z^{[1]} := x^{[0]}$, $k := 0$ a $i := 1$.

Krok 2: (*Iterace*) Vypočteme α_i takové, že $f(z^{[i]} + \alpha_i e_i) = \min_{\alpha\in\mathbb{R}} f(z^{[i]} + \alpha e_i)$. Položíme $z^{[i+1]} := z^{[i]} + \alpha_i e_i$. Je-li $i < n$, položíme $i := i+1$ a opakujeme Krok 2. V opačném případě pokračujeme Krokem 3.

Krok 3: (*Ověření ukončovacího pravidla*) Položíme $x^{[k+1]} := z^{[n]}$. Pokud je splněno ukončovací pravidlo KONEC.
V opačném případě položíme $z^{[1]} := x^{[k+1]}$, $k := k+1$, $i := 1$ a pokračujeme Krokem 2.



Obrázek 7.52: Vrstevnice funkce $f(x_1,x_2) = |x_1+x_2|+3|x_2-x_1|$ ilustrující možnou nefunkčnost MSS. Tyto obtíže lze překonat pomocí vhodné modifikace MSS.

Pokud by během celého cyklu v Kroku 2 nedošlo k vylepšení $x^{[k]}$, tj. stále by vycházelo $\alpha_1 = \dots = \alpha_n = 0$, pak bychom našli dokonce přesné řešení. Ale může nastat i situace, kdy metoda nefunguje, viz např. Obrázek 7.52. V případě hladké funkce stačí pro konvergenci k bodu splňujícího $\nabla f(x) = 0$ to, aby funkce f měla jediné minimum podél každého směru $e_1,\dots,e_n$. Říci něco o rychlosti konvergence této metody je celkem složité. Velmi zhruba lze říci, že $n-1$ souřadnicových minimalizací ve druhém kroku je stejně efektivní jako 1 minimalizace ve směru gradientu, které se budeme věnovat za chvíli, viz např. [360, Podkapitola 8.6].

WWW

Všechny níže představené metody jsou dostupné ve formě Mapletu na webu[76]:
☞ https://ln.math.muni.cz/pz/opt/MapletRn

[76] Alternativně lze využít kódy z diplomové práce [245], které jsou určené pro program R. Tyto metody jsou naprogramovány včetně přibližného řešení přidružených jednorozměrných úloh.

7.3.1 Metoda největšího spádu[80]

[80] *Method of Steepest Descent*

Ze základního kurzu matematické analýzy si nepochybně pamatujeme, že vektor $-\operatorname{grad} f(x^{[k]})$ určuje směr, ve kterém funkce f nejrychleji klesá[81] v (nějakém) okolí $x^{[k]}$. Proto nejpřirozenější volbou směru je

$$h_k := -\operatorname{grad} f(x^{[k]}), \tag{7.3.6}$$

což je myšlenka, která byla nejspíše poprvé použita Cauchym v roce 1847, viz [345]. Jedná se proto o tzv. *gradientní metodu*, ve které jsou jako směry voleny tzv. *s gradientem sdružené* vektory, tj. takové, že pro jakoukoli podposloupnost $\{x^{[k]}\}$ konvergující k nestacionárnímu bodu

[81] Přesněji řečeno, vektor

$$\hat{h}_k := -\frac{\operatorname{grad} f(x^{[k]})}{\| \operatorname{grad} f(x^{[k]}) \|}$$

je řešením minimalizační úlohy

$$\langle \operatorname{grad} f(x^{[k]}), h\rangle \to \min \quad \& \quad \|h\| = 1.$$

odpovídající podposloupnost $\{h_k\}$ je ohraničená a splňuje $\limsup_{k\to\infty}\langle \operatorname{grad} f(x^{[k]}), h_k\rangle < 0$, viz [57, str. 35–36]. Zhruba řečeno to znamená, že vektory h_k se nestanou „příliš velké“ nebo „příliš malé“ a současně úhel mezi h_k a $\operatorname{grad} f(x^{[k]})$ nebude „příliš blízko“ $\pi/2$. Tato vlastnost je vztažena k aktuální posloupnosti $\{x^{[k]}\}$, avšak např. Tuto vlastnost budou mít např. vektory směr můžeme získat např. jako $-D_k \operatorname{grad} f(x^{[k]})$

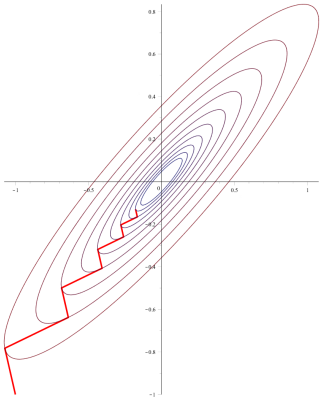
Metoda největšího spádu (MNS) se od ostatních gradientních metod liší právě tím, že směry volíme jako v (7.3.6) a délku kroku určíme jako přesné řešení[82] úlohy

$$f(x^{[k+1]}) = f\big(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})\big) = \min_{\alpha\geq 0} f(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})). \tag{7.3.7}$$

[82] Někdy bývá jako MNS označována pouze metoda s volbou (7.3.6) bez ohledu na způsob určení délky kroku.

Díky této volbě je MNS spádovou metodou, tj. v případě $\operatorname{grad} f(x^{[k]}) \neq 0$ platí $f(x^{[k+1]}) < f(x^{[k]})$. Vskutku, z volby bodů $x^{[k]}$ plyne, že pro funkci

$$\varphi_k(\alpha) := f\big(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})\big)$$



Obrázek 7.53: Několik prvních iterací MNS pro funkci $f(x_1,x_2) = 3x_1^2 - 7x_1x_2 + 5x_2^2$ s $x^{[0]} = [-1,-1]$.

platí $\varphi_k(\alpha_k) \leq \varphi_k(\alpha)$ pro všechna $\alpha \in [0,\infty)$. Potom z pravidla pro derivování složené funkce dostáváme

$$\begin{aligned}\frac{\mathrm{d}\varphi_k}{\mathrm{d}\alpha} &= -\operatorname{grad}^{\top} f\big(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})\big) \frac{\mathrm{d}}{\mathrm{d}\alpha}\Big[x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})\Big] = \\ &= -\Big\langle \operatorname{grad} f\big(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})\big), \operatorname{grad} f(x^{[k]})\Big\rangle,\end{aligned} \tag{7.3.8}$$

což pro $\alpha = 0$ dává

$$\varphi_k'(0) = -\|\operatorname{grad} f(x^{[k]})\|^2.$$

Je-li $\operatorname{grad} f(x^{[k]}) \neq 0$, pak $\varphi_k'(0) < 0$, tj. φ_k je klesající v okolí 0, což znamená, že existuje $\tilde{\alpha} > 0$ takové, že $\varphi_k(0) > \varphi_k(\alpha)$ pro všechna $\alpha \in [0,\tilde{\alpha})$. Proto v tomto případě je nutně $\alpha_k > 0$ a

$$f(x^{[k+1]}) = \varphi_k(\alpha_k) \leq \varphi_k(\tilde{\alpha}) < \varphi_k(0) = f(x^{[k]}).$$

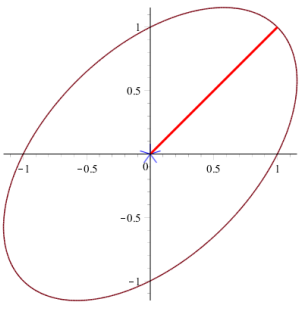
V případě, že $\operatorname{grad} f(x^{[k]}) = 0$, potom α_k může být libovolné a minimalizující posloupnost se stane konstantní, tj. $x^{[k]} = x^{[k+1]}$. Tím jsme dokázali následující tvrzení.

Věta 7.3.1

Jestliže $\{x^{[k]}\}_{k=0}^{\infty}$ je posloupnost generovaná MNS. Pak kdykoli $\operatorname{grad} f(x^{[k]}) \neq 0$ pro nějaké $k \in \mathbb{N}\cup\{0\}$, platí $f(x^{[k+1]}) < f(x^{[k]})$.

Vidíme, že v případě $\operatorname{grad} f(x^{[k]}) = 0$ se iterační proces zastaví a minimalizující posloupnost se od tohoto členu stane konstantní. To je zcela přirozené, neboť jsme nalezli stacionární bod. Avšak z praktického pohledu toto není úplně vhodné ukončovací kritérium a obvykle se volí kritéria typu (absolutní vs. relativní)

$$\|\operatorname{grad} f(x^{[k]})\| < \varepsilon, \quad |f(x^{[k+1]}) - f(x^{[k]})| < \varepsilon, \quad \|x^{[k+1]} - x^{[k]}\| < \varepsilon,$$
$$\frac{|f(x^{[k+1]}) - f(x^{[k]})|}{|f(x^{[k]})|} < \varepsilon, \quad \frac{\|x^{[k+1]} - x^{[k]}\|}{\|x^{[k]}\|} < \varepsilon,$$



Obrázek 7.54: MNS pro funkci $f(x,x_2) = x_1^2 - x_1x_2 + x_2^2$ s $x^{[0]} = [1,1]$. K nalezení řešení stačí jeden krok.

přičemž u posledních dvou podmínek je potřeba dát pozor na dělení nulou. Tomu by šlo zabránit např. tím, že výrazy ve jmenovateli nahradíme za $\max\{1, |f(x^{[k]})|\}$ a $\max\{1, \|x^{[k]}\|\}$.

Také víme, že vektor $\operatorname{grad} f(x^{[k]})$ je tečným vektorem vrstevnice funkce f na úrovni $f(x^{[k]})$. Navíc díky přesné minimalizaci v (7.3.7) dostáváme po dosazení $\alpha = \alpha_k \neq 0$ do (7.3.8), že

$$0 = \frac{\mathrm{d}}{\mathrm{d}\alpha} f(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]}))\Big|_{\alpha=\alpha_k} = -\Big\langle \operatorname{grad} f(x^{[k+1]}), \operatorname{grad} f(x^{[k]})\Big\rangle,$$

tj. směry největšího spádu funkce f v bodech $x^{[k]}$ a $x^{[k+1]}$ reprezentované vektory $\operatorname{grad} f(x^{[k]})$ a $\operatorname{grad} f(x^{[k+1]})$ jsou ortogonální. Proto také platí

$$\langle x^{[k+2]} - x^{[k+1]}, x^{[k+1]} - x^{[k]}\rangle = \alpha_k\alpha_{k+1}\langle \operatorname{grad} f(x^{[k+1]}), \operatorname{grad} f(x^{[k]})\rangle = 0,$$

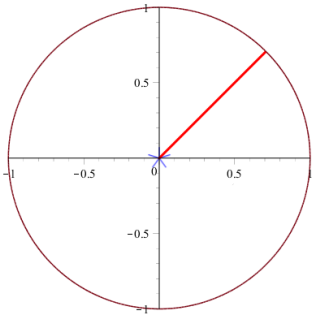
což znamená, že i vektory určené body $x^{[k+1]}$, $x^{[k]}$ a $x^{[k+2]}$, $x^{[k+1]}$ jsou ortogonální. To nás přivádí k typickému obrázku ilustrujícímu MNS a jejímu jednoduchému popisu: „vyjdeme“ z bodu $x^{[k]}$ ve směru vektoru $-\operatorname{grad} f(x^{[k]})$ a hledáme nejbližší vrstevnici, pro kterou bude tato polopřímka tečnou. Toto můžeme vidět na Obrázcích 7.53–7.55.

Avšak v některých případech může být „pohyb“ určený MNS velmi zdlouhavý a vede ke „klikatění“ (tzv. *cik-cak efekt*), kdy „vylepšení“ dané další iterací je zcela nepatrné. Příkladem takové funkce může být *Rosenbrockova funkce*[83] (též *Rosenbrockovo údolí*, *Rosenbrockova banánová funkce*) daná předpisem

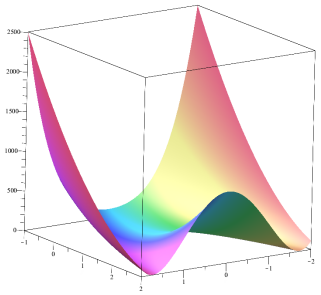
$$f(x_1, x_2) = (a - x_1)^2 + b(x_2 - x_1^2)^2,$$

která má globální minimum v bodě $[a, a^2]$. Nejčastěji se volí hodnoty parametrů $a = 1$ a $b = 100$, viz Obrázky 7.56 a 7.57. Průběh MNS pro tuto funkci s výchozím bodem $x^{[0]} = [0,0]$ je znázorněn na Obrázku 7.58. Další zajímavou testovací funkcí je např. $f(x_1, x_2) = 10(x_2 - \sin x_1)^2 + x_1^2/10$, pro niž je průběh MNS s výchozím bodem $x^{[0]} = [1,1]$ znázorněn na Obrázku 7.59.

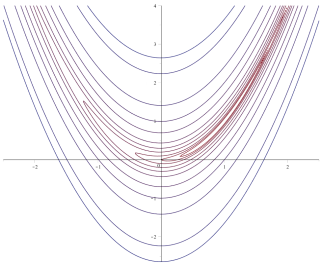
[83] Tato funkce byla zavedena H. Rosenbrockem v roce 1960 jako nekonvexní *testovací funkce* pro testování optimalizačních algoritmů.



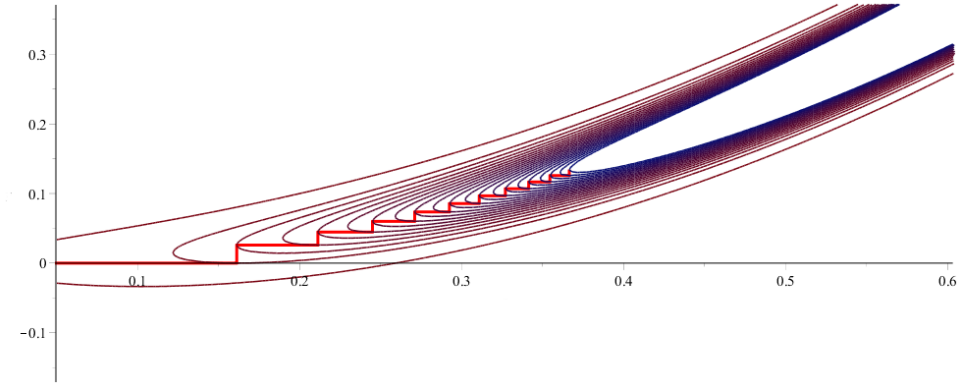
Obrázek 7.55: MNS pro funkci $f(x,x_2) = (x_1^2 + x_2^2)^{3/2}$ s $x^{[0]} = [1/\sqrt{2}, 1/\sqrt{2}]$. K nalezení řešení stačí jeden krok.



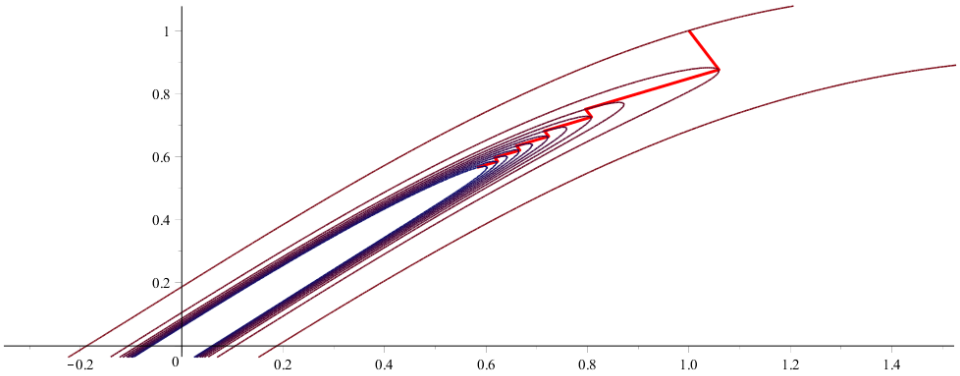
Obrázek 7.56: Graf Rosenbrockovy funkce s volbou $a = 1$ a $b = 100$.



Obrázek 7.57: Vrstevnice Rosenbrockovy funkce s volbou $a = 1$ a $b = 100$.



Obrázek 7.58: Několik kroků MNS pro Rosenbrockovu funkce s volbou $a = 1$ a $b = 100$ a $x^{[0]} = [0,0]$. Řešením je $x^*[1,1]$



Obrázek 7.59: Několik kroků MNS pro funkci $f(x_1, x_2) = 10(x_2 - \sin x_1)^2 + x_1^2/10$ s výchozím bodem $x^{[0]} = [1,1]$. Řešením je $x^*[0,0]$.

Algoritmický popis Metody největšího spádu

Zadání: Funkce f, výchozí bod $x^{[0]} \in \mathbb{R}^n$, ukončovací pravidlo (např. $\| x^{[k+1]} - x^{[k]} \| \le \varepsilon$) a volba tolerance $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $k := 0$.

Krok 2: (*Iterace*) Vypočteme $h_k := -\operatorname{grad} f(x^{[k]})$ a α_k takové, že $f(x^{[k]} + \alpha_k h_k) = \min_{\alpha \ge 0} f(x^{[k]} + \alpha h_k)$.

Krok 3: (*Ověření ukončovacího pravidla*) Položíme $x^{[k+1]} := x^{[k]} + \alpha_k h_k$. Pokud je splněno ukončovací pravidlo KONEC.
V opačném případě položíme $k := k + 1$ a pokračujeme Krokem 2.

MNS a kvadratické funkce

Nyní si podrobněji analyzujeme chování MNS pro kvadratické funkce

$$f(x) := \frac{1}{2}\langle Qx, x\rangle - x^\top b, \tag{7.3.9}$$

kde $Q^\top = Q > 0$ je $n \times n$ matice a $b \in \mathbb{R}^n$. Pozitivní definitnost zaručuje Q, že funkce f je ostře (a současně i silně) konvexní, viz Poznámku 6.4.9(ii). Vlastní hodnoty matice Q jsou kladné a můžeme je uspořádat jako

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n.$$

V tomto případě můžeme vyřešit úlohu (7.3.1) „přímo“ — řešením je bod x^* splňující

$$Qx^* = b, \quad \text{tj.} \quad x^* = Q^{-1}b \tag{7.3.10}$$

s odpovídající funkční hodnotou $f(x^*) = -\frac{1}{2}b^\top Q^{-1}b$. Ovšem i v tomto „jednoduchém“ případě může být řešení soustavy nebo výpočet matice Q^{-1} početně velmi náročné.

Gradient funkce f je

$$g(x) := \operatorname{grad} f(x) = Qx - b,$$

což znamená, že pro $g_k := Qx^{[k]} - b$ jsou jednotlivé iterační body tvaru

$$x^{[k+1]} = x^{[k]} - \alpha_k g_k = x^{[k]} - \alpha_k(Qx^{[k]} - b),$$

přičemž α_k je řešením úloh

$$f(x^{[k]} - \alpha g_k) \to \min, \quad \alpha > 0.$$

V tomto speciálním případě můžeme však určit α_k explicitně. Podle (7.3.9) máme

$$\begin{aligned} f(x^{[k]} - \alpha g_k) &= \frac{1}{2}\langle Q(x^{[k]} - \alpha g_k), x^{[k]} - \alpha g_k\rangle - (x^{[k]} - \alpha g_k)^\top b = \\ &= \frac{1}{2}\Big(x^{[k]\top} Q\, x^{[k]} - \alpha\, g_k^\top Q\, x^{[k]} - \alpha\, x^{[k]\top} Q\, g_k + \alpha^2\, g_k^\top Q\, g_k\Big) - x^{[k]\top} b + \alpha\, g_k^\top b, \end{aligned}$$

což po derivování vzhledem k α dává

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} f(x^{[k]} - \alpha\, g_k) = -g_k^\top Q\, x^{[k]} + \alpha\, g_k^\top Q\, g_k + g_k^\top b.$$

Proto

$$\alpha_k = \frac{g_k^\top (Q\, x^{[k]} - b)}{g_k^\top Q g_k} = \frac{g_k^\top g_k}{g_k^\top Q g_k} \quad \text{a} \quad x^{[k+1]} = x^{[k]} - \frac{g_k^\top g_k}{g_k^\top Q g_k}\, g_k. \tag{7.3.11}$$

Jaké hodnoty může α_k nabývat? Jelikož pro libovolné $x \in \mathbb{R}^n$ platí $\lambda_1\, x^\top x \le x^\top Q\, x \le \lambda_n\, x^\top x$, plyne odtud

$$\frac{1}{\lambda_n} = \frac{g_k^\top g_k}{\lambda_n\, g_k^\top g_k} \le \alpha_k \le \frac{g_k^\top g_k}{\lambda_1\, g_k^\top g_k} = \frac{1}{\lambda_1}, \quad \text{tj.} \quad \alpha_k \in [1/\lambda_n, 1/\lambda_1].$$

V tomto případě vrstevnice funkce f jsou n-dimenzionální elipsoidy s osami ve směru vlastních vektorů, které jsou vzájemně ortogonální[84]. Délka poloosy odpovídající i-tému vlastnímu vektoru je násobkem čísla $1/\sqrt{\lambda_i}$, viz opět Obrázek 7.53.

Příklad 7.3.2

Příklad Pomocí MNS určeme první iteraci $x^{[1]}$ pro řešení úlohy

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 2x_2^2 \to \min$$

při počáteční aproximaci $x^{[0]} = [1, 1]$.

Řešení. Platí $\operatorname{grad} f(x) = \binom{2x_1+2x_2}{2x_1+4x_2}$, takže $h_0 = -\operatorname{grad} f(x^{[0]}) = -\binom{4}{6}$. Pak délka kroku α_0 je řešením úlohy[85]

$$f(x^{[0]} + \alpha\, h_0) = (1-4\alpha)^2 + 2(1-4\alpha)(1-6\alpha) + 2(1-6\alpha)^2 \to \min \quad \text{pro } \alpha \ge 0\,.$$

[84] To je díky symetrii matice Q. Vektor b pouze posouvá střed elipsoidů, což komplikuje výpočet poloos atd.

[85] Samozřejmě k výpočtu α_0 a/nebo $x^{[1]}$ je možné využít i vzorec (7.3.11). My se však vydáme přímější cestou. Také zde (míněna celá MNS) lze využít možného „zkrácení“ vektoru h_0, aby ve výpočtu α_0 nevystupovala zbytečně velká čísla, tj. pracovat s $h_0 = -\binom{4}{6} \sim -\binom{2}{3}$.

Tato funkce je konvexní, takže nám stačí najít pouze stacionární bod, tj. derivováním vzhledem k α dostaneme

$$-8(1-4\alpha)+-8(1-6\alpha)-12(1-4\alpha)-24(1-6\alpha)=0 \quad \text{neboli} \quad -52+272\alpha=0,$$

což nás přivádí k $\alpha_0=\frac{52}{272}$. Proto

$$x^{[1]}=x^{[0]}+\alpha_0 h_0=[8/34,-5/34].$$

Analogickým způsobem najdeme $x^{[2]}=[1/170,1/170]$, $x^{[3]}=[2/1445,-1/1156]$ atd.

Jak je to s konvergencí MNS? Tuto otázku si zodpovíme v další části, avšak místo funkce f bude výhodnější se zaměřit na funkci

$$E(x):=\frac{1}{2}(x-x^*)^\top Q(x-x^*)=\frac{1}{2}x^\top Q\,x-\underbrace{x^{*\top}Q}_{\overset{(7.3.10)}{=}b^\top}x+\frac{1}{2}\underbrace{x^{*\top}Q}_{\overset{(7.3.10)}{=}b^\top}\underbrace{x^*}_{Q^{-1}b}=f(x)+\frac{1}{2}b^\top Q^{-1}b=f(x)-f(x^*).$$

Lemma 7.3.3

V MNS pro funkci f z (7.3.9) platí

$$E(x^{[k+1]})=\left\{1-\frac{(g_k^\top g_k)^2}{(g_k^\top Q g_k)(g_k^\top Q^{-1}g_k)}\right\}E(x^{[k]}).$$

Důkaz. Přímým výpočtem dostáváme

$$\begin{aligned}E(x^{[k]})-E(x^{[k+1]})&=f(x^{[k]})-f(x^*)-f(x^{[k+1]})+f(x^*)=\\&=\frac{1}{2}x^{[k]\top}Q\,x^{[k]}-x^{[k]\top}b-\frac{1}{2}x^{[k+1]\top}Q\,x^{[k+1]}+x^{[k+1]\top}b=\\&=\frac{1}{2}x^{[k]\top}Q\,x^{[k]}-x^{[k]\top}b-\frac{1}{2}(x^{[k]}-\alpha_k g_k)^\top Q(x^{[k]}-\alpha_k g_k)+(x^{[k]}-\alpha_k g_k)^\top b=\\&=\alpha_k g_k^\top Q\,x^{[k]}-\frac{1}{2}\alpha_k^2 g_k^\top Q g_k-\alpha_k g_k^\top b=\alpha_k g_k^\top Q(x^{[k]}-\underbrace{Q^{-1}b}_{\overset{(7.3.10)}{=}x^*})-\frac{1}{2}\alpha_k^2 g_k^\top Q g_k,\end{aligned}$$

z čehož plyne

$$\frac{E(x^{[k]})-E(x^{[k+1]})}{E(x^{[k]})}=\frac{2\alpha_k g_k^\top Q\,y_k-\alpha_k^2 g_k^\top Q g_k}{y_k^\top Q\,y_k},$$

kde $y_k:=x^{[k]}-x^*$. Jelikož $g_k=Q\,x^{[k]}-b=Q(x^{[k]}-x^*)=Q\,y_k$, obdržíme odtud s využitím první rovnosti v (7.3.11), že

$$\frac{E(x^{[k]})-E(x^{[k+1]})}{E(x^{[k]})}=\frac{2\frac{g_k^\top g_k}{g_k^\top Q g_k}g_k^\top g_k-\left(\frac{g_k^\top g_k}{g_k^\top Q g_k}\right)^2 g_k^\top Q g_k}{g_k^\top Q^{-1}Q Q^{-1}g_k}=\frac{(g_k^\top g_k)^2}{(g_k^\top Q g_k)(g_k^\top Q^{-1}g_k)},$$

což dává požadované tvrzení. ∎

Právě dokázaná rovnost ukazuje, že $E(x^{[k+1]})=0$ neboli $f(x^{[k+1]})=f(x^*)$ nastane právě tehdy, když platí

$$1=\frac{(g_k^\top g_k)^2}{(g_k^\top Q g_k)(g_k^\top Q^{-1}g_k)}. \tag{7.3.12}$$

Jelikož $f(x)$ je ostře konvexní funkce, platí v takovém případě $x^{[k+1]}=x^*$, a MNS dává přesné řešení po $k+1$ krocích. V jaké situaci nastane (7.3.12)? Musí platit

$$(g_k^\top Q g_k)\times(g_k^\top Q^{-1}g_k)=(g_k^\top g_k)^2 \quad \text{neboli} \quad \langle Q\,g_k,g_k\rangle\times\langle Q^{-1}g_k,g_k\rangle=\langle g_k,g_k\rangle^2,$$

což nastane právě tehdy, když g_k je vlastní vektor[86] matice Q. V $\mathbb{R}^2$ je situace poměrně lehká:

- Má-li matice Q různá vlastní čísla, pak díky ortogonalitě příslušných vlastních vektorů a ortogonalitě jednotlivých gradientů nalezneme řešení buď hned v prvním kroku nebo jich bude potřeba nekonečně mnoho.

[86] Je-li g_k vlastní vektor matice Q, pak je vlastním vektorem i matice Q^{-1}, tj. současně platí $Q\,g_k=\lambda_i g_k$ a $Q^{-1}g_k=\frac{1}{\lambda_i}g_k$ pro nějaké $i\in\{1,\dots,n\}$. Potom

$$\begin{aligned}\langle Q\,g_k,g_k\rangle\langle Q^{-1}g_k,g_k\rangle&=\\&=\langle\lambda_i g_k,g_k\rangle\times\langle\tfrac{1}{\lambda_i}g_k,g_k\rangle=\\&=\langle g_k,g_k\rangle^2.\end{aligned}$$

Naopak, s využitím Cauchyho-Schwarzovy nerovnosti dostaneme

$$\begin{aligned}\langle g_k,g_k\rangle^2=\langle Q^{1/2}g_k,Q^{-1/2}g_k\rangle^2\le\\\le\langle Q^{1/2}g_k,Q^{1/2}g_k\rangle\times\\\times\langle Q^{-1/2}g_k,Q^{-1/2}g_k\rangle=\\=\langle Q\,g_k,g_k\rangle\times\langle Q^{-1}g_k,g_k\rangle.\end{aligned}$$

Jestliže má platit uvedená rovnost, musí se Cauchyho–Schwarzova nerovnost realizovat jako rovnost, což je možné pouze v případě, kdy jsou oba vektory lineárně závislé, tj.

$$Q^{1/2}g_k=\lambda Q^{-1/2}g_k$$

pro nějaké $\lambda\in\mathbb{R}$ (dokonce lze vzít $\lambda>0$). To ale znamená, že $Q g_k=\lambda g_k$, tj. g_k je vlastní vektor matice Q.

- V případě stejných vlastních čísel je jakýkoli vektor vlastním vektorem, a tudíž řešení nalezneme vždy v jednom kroku[87].

V $\mathbb{R}^n$ se situace už komplikuje. Opět platí (i se stejným zdůvodněním jako výše), že v případě jednoho (jediného) vlastního čísla s algebraickou násobností n nalezneme řešení v jednom kroku. Vlastně obecněji lze říci, že jenom jeden krok bude potřeba v situaci, kdy počáteční aproximace bude zvolena jako $x^{[0]} = x^* + \beta v_i$ pro nějaké $\beta \in \mathbb{R}\setminus\{0\}$ a nějaký vlastní vektor[88] v_i. Naopak, je-li $\lambda_1 < \lambda_n$ a současně $\langle x^{[0]} - x^*, v_1\rangle \neq 0$ i $\langle x^{[0]} - x^*, v_n\rangle \neq 0$ pro ortonormalizované vlastní vektory $v_1,\dots,v_n$, pak MNS bude vyžadovat nekonečný počet kroků[89], tj. $x^{[k]} \neq x^*$ pro všechna $k \in \mathbb{N}\cup\{0\}$.

Pokud pro žádné $k \in \mathbb{N}$ nenastane rovnost (7.3.12), tak MNS je nekonečně krokovým iteračním procesem, který pro libovolný výchozí bod $x^{[0]} \in \mathbb{R}^n$ konverguje k x^*: posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná MNS je skutečně minimalizující, a proto je posloupnost $\{f(x^{[k]}) - f(x^*)\}_{k=0}^{\infty}$ klesající a zdola ohraničená, tudíž existuje její limita

$$\lim_{k\to\infty} f(x^{[k]}) = f(\overline{x}).$$

Avšak z ostré konvexnosti f vyplývá, že $\overline{x} = x^*$, neboť v opačném případě by funkce f měla dva stacionární body ↯, viz Větu 6.2.23.

Ještě bychom se měli podívat na rychlost konvergence MNS. K tomu budeme potřebovat následující nerovnost, kterou lze chápat jako zobecnění trojúhelníkové nerovnosti a zároveň jako speciální případ Cauchyho–Schwarzovy nerovnosti.

Lemma 7.3.4 (Kantorovičova nerovnost)

Nechť $A = A^\top > 0$ je $n\times n$ matice. Pak pro každý vektor $x \in \mathbb{R}^n\setminus\{0\}$ platí

$$\frac{(x^\top x)^2}{(x^\top Ax)\times(x^\top A^{-1}x)} \geq \frac{4\sigma_{\min}\sigma_{\max}}{(\sigma_{\min}+\sigma_{\max})^2},$$

kde $\sigma_{\min}$ je nejmenší vlastní číslo matice A a $\sigma_{\max}$ je největší vlastní číslo matice A.

Důkaz. Jelikož $A = A^\top > 0$, lze tuto matici diagonalizovat, tj. existuje ortogonální matice $S \in \mathbb{R}^{n\times n}$ taková, že $A = S^\top \operatorname{diag}\{\sigma_1,\dots,\sigma_n\}\,S$, kde $S^\top = S^{-1}$ a $\sigma_1,\dots,\sigma_n$ jsou vlastní čísla matice A. Proto bez újmy na obecnosti předpokládejme, že $A = \operatorname{diag}\{\sigma_1,\dots,\sigma_n\}$. V takovém případě je

$$\frac{(x^\top x)^2}{(x^\top Ax)\times(x^\top A^{-1}x)} = \frac{\left(\sum_{i=1}^n x_i^2\right)^2}{\left(\sum_{i=1}^n \sigma_i x_i^2\right)\times\left(\sum_{i=1}^n x_i^2/\sigma_i\right)},$$

což můžeme přepsat do tvaru

$$\frac{(x^\top x)^2}{(x^\top Ax)\times(x^\top A^{-1}x)} = \frac{\overbrace{1}^{=:\varphi(\xi)}}{\underbrace{\sum_{i=1}^n \xi_i\sigma_i}_{\sum_{i=1}^n(\xi_i/\sigma_i)}},$$

kde $\xi_i := \frac{x_i^2}{\sum_{i=1}^n x_i^2}$, přičemž jistě platí $0 \leq \xi_i \leq 1$ pro všechna $i \in \{1,\dots,n\}$ a $\sum_{i=1}^n \xi_i = 1$. Tím jsme vlastně výraz na levé straně přepsali do tvaru podílu dvou funkcí, které zahrnují konvexní kombinace čísel σ_i a $1/\sigma_i$. Toto můžeme ilustrovat Obrázkem 7.60, na němž je graf konvexní funkce $f(\sigma) = 1/\sigma$. Bod $\sum_{i=1}^n \xi_i\sigma_i$ leží mezi $\sigma_{\min}$ a $\sigma_{\max}$ a bod

$$\left[\xi, \varphi(\xi)\right] = \left[\sum_{i=1}^n \xi_i\sigma_i,\, 1/\left(\sum_{i=1}^n \xi_i\sigma_i\right)\right]$$

leží na grafu f. Na druhou stranu hodnota $\psi(\xi) = \sum_{i=1}^n \xi_i/\sigma_i = \sum_{i=1}^n \xi_i f(\sigma_i)$ je konvexní kombinací funkčních hodnot, takže „leží“ někde ve vyšrafované oblasti. Body odpovídající témuž ξ musí ležet na stejné svislici. Minimum tohoto podílu dostaneme pro nějaké $\sigma = \xi^*_{\min}\sigma_{\min} +$

[87] Zde jsou vrstevnice kružnice a každá jejich tečna má normálový vektor mířící do středu, což je hledané řešení

[88] Vždyť platí $Qx^* - b = 0$ a $Qv_i = \lambda_i v_i$, tudíž

$$x^{[1]} = x^{[0]} - \alpha_0(Qx^{[0]} - b) = \\ = x^* + \beta(1-\alpha_0\lambda_i)v_i,$$

což nás přivede nutně k volbě $\alpha_0 = 1/\lambda_i$ a k hledanému řešení.

[89] Z vlastností symetrických matic vyplývá, že vlastní vektory $v_1,\dots,v_n$ mohou být zvoleny tak, že tvoří ortonormální bázi $\mathbb{R}^n$. Pak vždy platí

$$x^{[k]} - x^* = \sum_{i=1}^n \beta_i^{[k]} v_i$$

pro vhodná čísla $\beta_1^{[k]},\dots,\beta_n^{[k]} \in \mathbb{R}$. Pro $k=0$ je zjevně $\beta_1^{[0]} \neq 0$ a $\beta_n^{[k]} \neq 0$. Připusťme, že totéž platí pro nějaké $k \in \mathbb{N}\cup\{0\}$. Pak pro $k+1$ je

$$\begin{aligned}
x^{[k+1]} - x^* &= x^{[k]} - x^* - \alpha_k g_k = \\
&= \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i - \alpha_k(Qx^{[k]} - b) = \\
&= \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i - \\
&\quad - \alpha_k[Q(x^* + \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i) - b] = \\
&= \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i - \\
&\quad - \alpha_k[Q(\textstyle\sum_{i=1}^n \beta_i^{[k]} v_i) + Qx^* - b] = \\
&= \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i - \alpha_k \sum_{i=1}^n \beta_i^{[k]} Qv_i = \\
&= \textstyle\sum_{i=1}^n \beta_i^{[k]} v_i - \alpha_k \sum_{i=1}^n \beta_i^{[k]} \lambda_i v_i = \\
&= \textstyle\sum_{i=1}^n (1-\alpha_k\lambda_i)\beta_i^{[k]} v_i,
\end{aligned}$$

tj.

$$\beta_i^{[k+1]} = (1-\alpha_k\lambda_i)\beta_i^{[k]},$$

pro všechna $i \in \{1,\dots,n\}$, přičemž podle první rovnosti v (7.3.11) je

$$\begin{aligned}
\alpha_k &= \frac{(Qx^{[k]}-b)^\top(Qx^{[k]}-b)}{(Qx^{[k]}-b)^\top Q(Qx^{[k]}-b)} = \\
&= \frac{[Q(x^{[k]}-x^*)]^\top Q(x^{[k]}-x^*)}{[Q(x^{[k]}-x^*)]^\top QQ(x^{[k]}-x^*)} = \\
&= \frac{[Q(\sum_{i=1}^n \beta_i^{[k]} v_i)]^\top Q(\sum_{i=1}^n \beta_i^{[k]} v_i)}{(Q\sum_{i=1}^n \beta_i^{[k]} v_i)^\top QQ(\sum_{i=1}^n \beta_i^{[k]} v_i)} = \\
&= \frac{(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)^\top(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)}{(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)^\top Q(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)} = \\
&= \frac{(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)^\top(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)}{(\sum_{i=1}^n \beta_i^{[k]}\lambda_i v_i)^\top(\sum_{i=1}^n \beta_i^{[k]}\lambda_i^2 v_i)} = \\
&= \frac{\sum_{i=1}^n \beta_i^{[k]2}\lambda_i^2}{\sum_{i=1}^n \beta_i^{[k]2}\lambda_i^3}.
\end{aligned}$$

Připusťme, že $\beta_1^{[k+1]} = 0$. Pak $\lambda_1 = \frac{1}{\alpha_k}$ neboli

$$\lambda_1 \sum_{i=1}^{n} \beta_i^{[k]2} \lambda_i^2 - \sum_{i=1}^{n} \beta_i^{[k]2} \lambda_i^3 = 0,$$

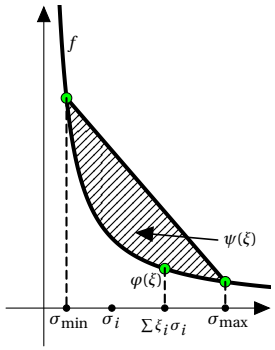
což po úpravách vede na

$$\sum_{i=2}^{n} (\lambda_1 - \lambda_i) \beta_i^{[k]2} \lambda_i^2 = 0,$$

což je ale spor s tím, že $0 < \lambda_1 < \lambda_n$ a $\beta_n^{[k]} \neq 0$. Podobně v případě $\beta_n^{[k+1]} = 0$ obdržíme

$$\sum_{i=1}^{n-1} (\lambda_i - \lambda_n) \beta_i^{[k]2} \lambda_i^2 = 0,$$

což je tentokrát ve sporu s tím, že $0 < \lambda_1 < \lambda_n$ a $\beta_1^{[k]} \neq 0$.



Obrázek 7.60: Ilustrace funkcí φ a ψ.

[90] Stačí uvážit pouze krajní body intervalu $[\sigma_{\min}, \sigma_{\max}]$.

[91] Derivováním funkce

$$\frac{1}{\sigma} \frac{\sigma_{\min}\sigma_{\max}}{\sigma_{\min} + \sigma_{\max} - \sigma}$$

vzhledem k σ dostaneme

$$-\frac{\sigma_{\min}\sigma_{\max}(\sigma_{\min} + \sigma_{\max} - 2\sigma)}{\sigma^2(\sigma_{\min} + \sigma_{\max} - \sigma)^2},$$

což dává jediný stacionární bod $\sigma = (\sigma_{\min} + \sigma_{\max})/2$. Vyšetřením znamének první derivace v okolí tohoto stacionárního bodu pak snadno zjistíme, že se jedná o minimum.

$\xi^*_{\max}\sigma_{\max}$ určené vhodnými[90] $\xi_{\min}, \xi_{\max} \in [0,1]$ splňujícími $\xi_{\min} + \xi_{\max} = 1$. Protože v případě $\xi^* := (\xi^*_{\min}, 0, \ldots, 0, \xi^*_{\max})^\top$ platí

$$\psi(\xi^*) = \frac{\xi^*_{\min}}{\sigma_{\min}} + \frac{\xi^*_{\max}}{\sigma_{\max}} = \frac{\sigma_{\min} + \sigma_{\max} + \xi^*_{\min}\sigma_{\min} + \xi^*_{\max}\sigma_{\max}}{\sigma_{\min}\sigma_{\max}},$$

dostáváme odtud

$$\frac{\varphi(\xi^*)}{\psi(\xi^*)} = \min_{\sigma \in [\sigma_{\min}, \sigma_{\max}]} \frac{1/\sigma}{\frac{\sigma_{\min} + \sigma_{\max} - \sigma}{\sigma_{\min}\sigma_{\max}}}.$$

Minimum na pravé straně je dosaženo[91] pro $\sigma = (\sigma_{\min} + \sigma_{\max})/2$, z čehož plyne

$$\frac{\varphi(\xi)}{\psi(\xi)} \geq \frac{\varphi(\xi^*)}{\psi(\xi^*)} = \frac{\frac{2}{\sigma_{\min} + \sigma_{\max}}}{\frac{2}{\sigma_{\min} + \sigma_{\max}}} = \frac{2}{\sigma_{\min} + \sigma_{\max}} \frac{2\sigma_{\min}\sigma_{\max}}{\sigma_{\min} + \sigma_{\max}} = \frac{4\sigma_{\min}\sigma_{\max}}{(\sigma_{\min} + \sigma_{\max})^2}$$

a důkaz je hotov. ■

Kombinací Lemma 7.3.3 a 7.3.4 dostaneme hlavní výsledek o konvergenci MNS pro kvadratickou funkci (7.3.9).

Věta 7.3.5

Nechť platí (7.3.9). Pak pro libovolné $x^{[0]} \in \mathbb{R}^n$ posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná MNS konverguje k jedinému řešení x^* úlohy (7.3.1). Navíc pro funkci $E(x)$ platí

$$E(x^{[k+1]}) \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 E(x^{[k]}) \quad \text{neboli} \quad \frac{f(x^{[k+1]}) - f^*}{f(x^{[k]}) - f^*} \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2.$$

Důkaz. Z Lemma 7.3.3 a 7.3.4 plyne

$$E(x^{[k+1]}) = \left\{1 - \frac{(g_k^\top g_k)^2}{(g_k^\top Q g_k)(g_k^\top Q^{-1} g_k)}\right\} E(x^{[k]}) \leq \left\{1 - \frac{4\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2}\right\} E(x^{[k]}) = \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 E(x^{[k]}).$$

■

Tedy $E(x^{[k]}) \leq C\beta^k$, kde $k \in \mathbb{N}$, $C = f(x^{[0]}) - f(x^*)$ a $\beta := \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2$. To znamená, že konvergence je *alespoň lineární s rychlostí nejméně* β. Nicméně toto je pouze horní odhad a skutečná rychlost konvergence závisí na volbě $x^{[0]}$. V případě $\lambda_1 = \lambda_n$ (tj. $\lambda_1 = \cdots = \lambda_n$) bude konvergence dokonce alespoň *superlineární*. Ovšem již v případě $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} \neq \lambda_n$ může být konvergence *velmi pomalá*. Dá se také ukázat, že pro „téměř všechny“ výchozí body $x^{[0]}$ bude konvergence lineární s rychlostí velmi blízkou číslu β, viz [51, str. 182]. Hodnota β závisí na rozdílu $\lambda_n - \lambda_1$, což úzce souvisí s tzv. *excentricitou* elipsoidů[92], které jsou vrstevnicemi funkce f. Čím větší excentricita je, tím větší je hodnota β, a tedy pomalejší konvergence. Obzvláště je-li $\lambda_1 \ll \lambda_n$, bude β velmi blízko 1 a konvergence může být velmi pomalá, což povede k „cik-cak efektu“, který se projevuje oscilací čísel α_k v intervalu $[1/\lambda_n, 1/\lambda_1]$.

Číslo β můžeme také zapsat jako

$$\beta = \left(\frac{\kappa - 1}{\kappa + 1}\right)^2,$$

kde $\kappa := \lambda_n/\lambda_1$ vyjadřuje tzv. *podmíněnost matice*[93] Q. Čím větší je hodnota κ, tím horší[94] je podmíněnost matice Q, viz Tabulku 7.5.

Uvažme např. kvadratickou funkci $f(x_1, x_2) = x_1^2 + ax_2^2$. Bude-li $a = 2$, pak $\kappa = 2$ a v podstatě v každém kroku bychom měli snížit hodnotu rozdílu $f(x^{[k]}) - f^*$ desetinásobně oproti předchozímu kroku. Na druhou stranu pro $a = 1/100$ je $\kappa = 100$, což je celkem velká hodnota, a k dosažení téhož desetinásobného snížení by mělo být potřeba cca 58 kroků. Avšak toto jsou pouze horní hranice potřebných počtu kroků, tj. ke kýženému zlepšení může dojít i dříve. A v tomto případě se tomu tak skutečně i stane. S využitím nabídnutých Mapletů se může čtenář snadno přesvědčit, že s prakticky jakoukoli výchozí aproximací dochází k takovémuto zlepšení téměř v každém kroku pro obě hodnoty a. Na proti tomu uvážíme-li funkci

κ	hodnota β	maximální potřebný počet iterací
1,1	0,023	1
3	0,25	2
10	0,67	6
100	0,96	58
200	0,98	116
400	0,99	231

Tabulka 7.5: Nejvyšší potřebný počet iterací k desetinásobnému snížení mezi $f(x^{[k]})$ a f^*, tj. podle Věty 7.3.5 je v pravém sloupci číslo N takové, že $\beta^N \le 0{,}1$

$f(x_1,x_2) = 10x_1^2 + 5x_1x_2 + x_2^2 - 14x_1 - 6x_2 + 10$, pak máme $\kappa = 30{,}234$ a $\beta = 0{,}876$, takže k desetinásobnému snížení je potřeba nejvýše 18 iterací. Opět s využitím Mapletu se lze přesvědčit, že s každou iterací dojde ke snížení rozdílu $f(x^{[k]}) - f^*$ o 0,658 násobek a během prvních 50 iterací dojde cca ke stotisícinásobnému snížení.

Dá se také snadno ukázat, že platí

$$\|x^{[k+1]} - x^*\|_Q \le \frac{\kappa-1}{\kappa+1}\|x^{[k]} - x^*\|_Q \quad \text{a} \quad \|x^{[k+1]} - x^*\| \le \sqrt{\frac{\lambda_n}{\lambda_1}}\left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)\|x^{[k]} - x^*\|, \quad (7.3.13)$$

kde $\|x\|_Q := \sqrt{x^* Q x} = \sqrt{2f(x) + b^\top x}$, viz [467, Theorem 3.1.5].

MNS a nekvadratické funkce

V případě obecné funkce $f : \mathbb{R}^n \to \mathbb{R}$ je potřeba pro konvergenci MNS (k minimu) nějaký dodatečný předpoklad nad rámec diferencovatelnosti funkce f (a pro určení rychlosti konvergence ještě další). Bylo by samozřejmě velmi naivní očekávat, že nás MNS vždy přivede k lokálnímu/globálnímu minimu. Tuto představu lze snadno „rozbít“ jednoduchým příkladem, který lze najít v [513] (viz také [237, NLP Myth 19]):

(i) Funkce $f(x_1,x_2) = x_1^3/3 + x_2^2/2$ je spojitě diferencovatelná a pro $x_1 > 0$ i konvexní. S využitím výchozího bodu $x^{[0]} = [x_1^{[0]}, x_2^{[0]}]$ splňujícího $x_1^{[0]} > 1$ generuje MNS posloupnost $\{[x_1^{[k]}, x_2^{[k]}]\}_{k=0}^\infty$, která nekonverguje (je neohraničená) stejně jako posloupnost funkčních hodnot $\{f(x_1^{[k]}, x_2^{[k]})\}_{k=0}^\infty$, tj. ztrácíme globální konvergenci MNS.

(ii) Vezmeme-li tutéž funkci a výchozí bod splňující $0 < x_1^{[0]} < 1$, pak dostaneme $0 < x_1^{[k]} < 1$ a posloupnost $\{[x_1^{[k]}, x_2^{[k]}]\}_{k=0}^\infty$ konvergující ke stacionárnímu bodu $[0,0]$. To ale není (ani lokální) minimum funkce f, tj. MNS konverguje ke stacionárnímu bodu, který není minimem, viz Obrázek 7.61.

Postačující podmínky pro konvergenci MNS a určení její rychlosti pro nekvadratickou funkci je obsahem následující věty.

Věta 7.3.6 Neasymptotická a asymptotická rychlost globální konvergence MNS

Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ je spojitě diferencovatelná na $\mathbb{R}^n$, zdola ohraničená[95] na $\mathbb{R}^n$ a gradient funkce f je lipschitzovsky spojitý[96], tj. existuje číslo $L > 0$ takové, že

$$\|\operatorname{grad} f(x) - \operatorname{grad} f(y)\| \le L\|x - y\| \quad \text{pro každé } x, y \in \mathbb{R}^n. \quad (7.3.14)$$

Potom pro posloupnost $\{x^{[k]}\}_{k=0}^\infty$ platí následující tvrzení.

(i) Posloupnost $\{f(x^{[k]})\}_{k=0}^\infty$ je nerostoucí a v případě $\operatorname{grad} f(x^{[k]}) \neq 0$ pro $k \in \mathbb{N} \cup \{0\}$ dokonce je $f(x^{[k+1]}) < f(x^{[k]})$, tj. $\{x^{[k]}\}$ je minimalizující posloupnost, přičemž

$$\lim_{k\to\infty} \operatorname{grad} f(x^{[k]}) = 0,$$

[92] Je-li $n = 2$, pak excentrické číslo je podíl vzdálenosti ohnisek od středu a délky hlavní poloosy elipsy $(x-x_0)^2/a^2 + (y-y_0)^2/b^2 = 1$, tj. pro vzdálenost ohniska od středu $e := \sqrt{a^2 - b^2}$ je excentricita rovna $\varepsilon := e/a < 1$. Toto číslo vyjadřuje míru zploštění elipsy, příp. míru odlišnosti od kružnice, pro kterou bychom dostali $e = 0$.

[93] V jistém smyslu toto číslo můžeme chápat jako indikátor citlivosti regulárnosti matice Q na změny. Jinými slovy, číslo $1/\kappa$ vyjadřuje „vzdálenost“ od nejbližší singulární matice (v tomto ohledu je hodnota $\det Q$ nevypovídající, neboť např. pro $Q = \frac{1}{10} I \in \mathbb{R}^{1000\times 1000}$ je $\det Q = 10^{-1000}$, zatímco $\kappa(Q) = 1$). Číslo κ také hraje velmi důležitou roli v citlivosti řešení nehomogenní rovnice $Ax = b$ na změně pravé strany $b + \Delta b$. Je-li hodnota κ malá, pak malá změna Δb způsobí malou změnu v řešení. Avšak bude-li κ velké, pak i malá změna Δb způsobí velkou změnu v řešení soustavy.

[94] Problémy se „špatně“ podmíněnou maticí (tj. vysokou hodnotou κ bývají v angličtině označovány jako *ill-conditioned*.

Věta 7.3.6 Neasymptotická a asymptotická rychlost globální konvergence MNS

tj. každý hromadný bod posloupnosti $\{x^{[k]}\}$ je současně stacionárním bodem funkce f.

(ii) Pro libovolnou počáteční aproximaci $x^{[0]}$ máme

$$\min_{k=0,1,\dots,N} \|\operatorname{grad} f(x^{[k]})\|^2 \le \frac{2L}{N+1}\,(f(x^{[0]}) - f^*),$$

pro libovolné $N \in \mathbb{N}$, což vyjadřuje tzv. *neasymptotickou rychlost konvergence*[97,98].

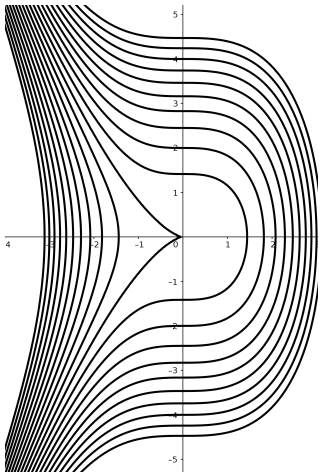
(iii) Je-li funkce f je navíc dvakrát spojitě diferencovatelná na $\mathbb{R}^n$ s

$$aI \le \nabla^2 f(x) \le AI$$

kde $a, A > 0$, pak[99]

$$f(x^{[k+1]}) - f(x^*) \le \left(1 - \frac{a}{A}\right)\big(f(x^{[k]}) - f(x^*)\big), \tag{7.3.15}$$

tj. konvergence je (alespoň) lineární s rychlostí (nejméně) $1 - a/A$.



Obrázek 7.61: Vrstevnice zdola neohraničené funkce $f(x_1, x_2) = x_1^3/3 + x_2^2/2$.

[95] Tj. $f(x) \ge m$ pro nějaké $m \in \mathbb{R}$ a všechna $x \in \mathbb{R}^n$.

[96] Tj. $\operatorname{grad} f \in C_L^{0,1}(\mathbb{R}^n)$ neboli $f \in C_L^{1,1}(\mathbb{R}^n)$. To kvadratická funkce z (7.3.9) jistě splňuje.

[97] *Non-asymptotical rate of convergence.*

[98] Viz [51, Proposition 6.1.1] a/nebo [38, Theorem 4.26], kde jsou analogické výsledky i pro jiné volby délky kroků.

[99] Funkce f je tedy silně konvexní, což pro připomenutí znamená, že (i) dolní obrysová množina $\{x \in X \mid f(x) \le K\}$ je kompaktní pro všechna $K \in \mathbb{R}$; funkce nabývá svého globálního minima na $\mathbb{R}^n$, které je jediné; gradient f je lipschitzovsky spojitý s konstantou A. Současně to znamená, že vlastní čísla matic $\nabla^2 f(x)$ jsou stejnoměrně ohraničená, tj.

$$a \le \lambda_{\min} \nabla^2 f(x) \le \lambda_{\max} \nabla^2 f(x) \le A$$

pro všechna $x \in \mathbb{R}^n$.

Důkaz. Nejdříve dokážeme část (ii). Položme proto $h(t) := f(x + t(y - x))$. Pak $h \in \mathrm{C}^1$ a s využitím Cauchyho–Schwarzovy nerovnosti i (7.3.14) získáme

$$\begin{aligned} |h'(\alpha) - h'(\beta)| &= |\langle y - x, \operatorname{grad} f(x + \alpha(x - y)) - \operatorname{grad} f(x + \beta(x - y))\rangle| \le \\ &\le \|y - x\| \times \|\operatorname{grad} f(x + \alpha(x - y)) - \operatorname{grad} f(x + \beta(x - y))\| \le \\ &\le L|\alpha - \beta| \times \|y - x\|^2 \end{aligned}$$

pro všechna $\alpha, \beta \in \mathbb{R}$. Díky tomu

$$\begin{aligned} f(y) - f(x) - \langle y - x, \operatorname{grad} f(x)\rangle &= h(1) - h(0) - h'(0) = \int_0^1 h'(t)\,\mathrm{d}t - h'(0) = \\ &= \int_0^1 \big[h'(t) - h'(0)\big]\,\mathrm{d}t \le \int_0^1 L\|y - x\|\,t\,\mathrm{d}t = \frac{L}{2}\|y - x\|^2 \end{aligned} \tag{7.3.16}$$

pro všechna $x, y \in \mathbb{R}^n$, z čehož a z konstrukce MNS vyplývá

$$\begin{aligned} f(x^{[k+1]}) &= \min_{\alpha \ge 0} f(x^{[k]} - \alpha \operatorname{grad} f(x^{[k]})) \le \\ &\le \min_{\alpha \ge 0}\left\{f(x^{[k]}) + \langle -\alpha \operatorname{grad} f(x^{[k]}), \operatorname{grad} f(x^{[k]})\rangle + \frac{L}{2}\|\alpha \operatorname{grad} f(x^{[k]})\|^2\right\} = \\ &= f(x^{[k]}) + \|\operatorname{grad} f(x^{[k]})\|^2 \min_{\alpha \ge 0}\left\{-\alpha + \frac{L\alpha^2}{2}\right\} = \\ &= f(x^{[k]}) - \frac{1}{2L}\|\operatorname{grad} f(x^{[k]})\|^2, \end{aligned}$$

tj.

$$f(x^{[k]}) - f(x^{[k+1]}) \ge \frac{1}{2L}\|\operatorname{grad} f(x^{[k]})\|^2, \tag{7.3.17}$$

což ukazuje, že zlepšení účelové funkce v jednom kroku MNS je nejméně přímo úměrné druhé mocnině normy gradientu v předchozím bodě. Vzhledem k monotónnosti MNS nemůže být celkové zlepšení hodnoty účelové funkce v průběhu jistého počtu kroků větší než výchozí rozdíl $f(x^{[0]}) - f^*$. Tudíž při větším počtu iterací se musí objevit krok s malým zlepšením, tj. s malou normou gradientu. Přesněji řečeno, sečtením nerovností (7.3.17) pro $k = 0, \dots, N$ obdržíme

$$f(x^{[0]}) - f^* \ge f(x^{[0]}) - f(x^{[N+1]}) \ge \frac{1}{2L}\sum_{k=0}^{N}\|\operatorname{grad} f(x^{[k]})\|^2 \ge \frac{N+1}{2L}\min_{k=0,\dots,N}\|\operatorname{grad} f(x^{[k]})\|^2.$$

Z (7.3.17) ihned vidíme, že posloupnost $\{f(x^{[k]})\}_{k=0}^{\infty}$ je nerostoucí a současně $f(x^{[k]}) = f(x^{[k+1]})$ může nastat tehdy a jen tehdy, když $\|\operatorname{grad} f(x^{[k]})\| = 0$. Jelikož funkce f je zdola ohraničená, je

tato posloupnost dokonce konvergentní. Proto $f(x^{[k]}) - f(x^{[k+1]}) \to 0$ pro $k \to \infty$, což společně s (7.3.17) ukazuje, že $\lim_{k\to\infty} \operatorname{grad} f(x^{[k]}) = 0$, čímž je dokázána část (i).

Konečně pro důkaz části (iii) nejprve využijeme Taylorův kvadratický polynom

$$f(x^{[k]} - \alpha\, g_k) = f(x^{[k]}) - \alpha\, \|g_k\|^2 + \frac{1}{2}\,\alpha^2\, g_k^{\top} \nabla^2 f(z_k)\, g_k,$$

kde $g_k := \operatorname{grad} f(x^{[k]})$ a $z_k = x^{[k]} + \theta\,(x^{[k]} - \alpha\, g_k)$ pro vhodné $\theta \in (0,1)$. Pak podle Min–Max věty[100] platí

$$f(x^{[k]} - \alpha\, g_k) \le f(x^{[k]}) - \alpha\, \|g_k\|^2 + \frac{1}{2}\, A\alpha^2\, \|g_k\|^2.$$

Minimum na levé straně je $f(x^{[k+1]})$, zatímco minimalizováním pravé strany získáme

$$\begin{aligned} f(x^{[k+1]}) &= \min_{\alpha\ge 0} f(x^{[k]} - \alpha\, g_k) \le \min_{\alpha\ge 0} \Big\{ f(x^{[k]}) - \alpha\, \|g_k\|^2 + \frac{1}{2}\, A\alpha^2\, \|g_k\|^2 \Big\} = \\ &= f(x^{[k]}) + \|g_k\|^2 \min_{\alpha\ge 0} \Big\{ -\alpha + \frac{1}{2}\, A\alpha^2 \Big\} = \\ &= f(x^{[k]}) - \frac{1}{2A} \|g_k\|^2, \end{aligned}$$

srovnej s (7.3.17). Pak

$$f(x^{[k+1]}) - f^* \le f(x^{[k]}) - f^* - \frac{1}{2A} \|g_k\|^2. \tag{7.3.18}$$

Podobně s využitím nerovnosti $aI \le \nabla^2 f(x)$ dostaneme pro libovolné $x \in \mathbb{R}$, že

$$f(x) \ge f(x^{[k]}) + \langle g_k, x - x^{[k]} \rangle + \frac{a}{2} \|x - x_k\|^2.$$

Následně minimalizováním obou stran vzhledem k $x \in \mathbb{R}^n$ obdržíme

$$f^* = \min_{x\in\mathbb{R}^n} f(x) \ge f(x^{[k]}) + \min_{x\in\mathbb{R}^n} \Big\{ \langle g_k, x - x^{[k]} \rangle + \frac{a}{2} \|x - x_k\|^2 \Big\} = f(x^{[k]}) - \frac{1}{2a} \|g_k\|^2,$$

neboť na pravé straně máme minimalizační úlohu s kvadratickou účelovou funkcí, jejíž řešení se realizuje v bodě $\overline{x} = x^{[k]} - \frac{1}{a}\, g_k$. Odtud

$$-\|g_k\|^2 \le 2a \big(f^* - f(x^{[k]}) \big),$$

což po dosazení do (7.3.18) dává (7.3.15), čímž je důkaz hotov. ∎

Odhad efektivnosti ve druhé části tvrzení Věty 7.3.6 dává neasymptotickou horní hranici přesnosti měřenou pomocí $\|\operatorname{grad} f(x^{[k]})\|$, která sublineárně konverguje k 0. Avšak je potřeba si povšimnout, že toto je hranice nepřesnosti té nejlepší (s nejmenším gradientem) z iterací generovaných v průběhu prvních $N+1$ kroků MNS nikoli té poslední $x^{[N+1]}$, protože hodnoty $\|\operatorname{grad} f(x^{[k]})\|$ mohou oscilovat na rozdíl od hodnot $f(x^{[k]})$. Ve třetí části této věty je pak určena asymptotická rychlost konvergence, která pro gradientní metodu s Armijovým pravidlem má podobu

$$f(x^{[k+1]}) - f(x^*) \le \left(1 - \frac{2\varepsilon_1\, a}{\eta\, A}\right) \big(f(x^{[k]}) - f(x^*) \big)$$

pro $\varepsilon_1 < 1/2$, zatímco pro případ $\varepsilon_1 \ge 1/2$ platí

$$\|x^{[k+1]} - x^*\| \le \theta\, \|x^{[k]} - x^*\| \quad \text{a} \quad f(x^{[k]}) - f^* \le \frac{\theta^{2k} A}{a} \big(f(x^{[k]}) - f^* \big),$$

kde

$$\theta := \sqrt{\frac{A\varepsilon_1\eta - (2\varepsilon_1 - 1)(1-\varepsilon_1)\, a}{A\varepsilon_1\eta + (1+\varepsilon_1)\, a}},$$

viz [360, str. 239] a [51, Proposition 6.1.5]. Avšak (7.3.15) dává pouze dolní hranice rychlosti konvergence, která může být ve skutečnosti (mnohem) vyšší, jak jsme mohli vidět v ilustrativním příkladu k Tabulce 7.5. Tato rychlost však ani v případě kvadratické funkce neodpovídá tomu, co jsme odvodili pro kvadratické funkce ve Větě 7.3.5, neboť $\frac{A+a}{A-a} < 1 - a/A$. Tento odhad rychlosti lze snadno snadno získat v případě gradientní metody s konstantní délkou kroku[101] jako např. v [398, Theorem 2.1.15], zatímco pro obecnou funkci už je to mnohem pracnější. Nicméně na [360, str. 239] takové tvrzení nalezneme:

[100] Tj. $\lambda_{\min} \le \frac{h^* A h}{h^* h} \le \lambda_{\max}$, viz str. 499.

[101] **Věta.** *Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je silně konvexní s konstantou silné konvexnosti ϑ a její gradient je lipschitzovsky spojitý s konstantou L. Pak pro posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovanou gradientní metodou s konstantní délkou kroku $\alpha \in (0, 2/(L+\vartheta)]$ platí*

$$\|x^{[k+1]} - x^*\|^2 \le \left(1 - \frac{2\alpha\vartheta L}{L+\vartheta}\right) \|x^{[k]} - x^*\|^2,$$

kde x^* značí (jediné) globální minimum f na $\mathbb{R}^n$. Obzvláště pro $\alpha = 2/(L+\vartheta)$ máme

$$\|x^{[k+1]} - x^*\| \le \left(\frac{L-\vartheta}{L+\vartheta}\right) \|x^{[k]} - x^*\|,$$

a

$$f(x^{[k+1]}) - f(x^*) \le \frac{L}{2} \left(\frac{L-\vartheta}{L+\vartheta}\right)^{2k} \|x^{[0]} - x^*\|^2.$$

První část tohoto tvrzení ukazuje, že pro libovolné $\alpha \in (0, 2/(L+\vartheta)]$ je konvergence posloupnosti $\{\|x^{[k+1]} - x^*\|\}$ lineární (alespoň) s rychlostí $\sqrt{1 - \frac{2\alpha\vartheta L}{L+\vartheta}}$, která je klesající vzhledem k α. Proto nejvyšší rychlosti dosáhneme pro $\alpha = 2/(L+\vartheta)$, a to $\frac{L-\vartheta}{L+\vartheta}$, což je rostoucí vzhledem k hodnotě podílu L/ϑ.

Důkaz. Nejdříve poznamenejme, že pro funkci f za daných předpokladů platí

$$\begin{aligned}\langle \operatorname{grad} f(x) - \operatorname{grad} f(y), x-y\rangle \geq \\ \geq \frac{\vartheta L}{\vartheta+L}\|x-y\|^2 + \\ + \frac{1}{\vartheta+L}\|\operatorname{grad} f(x) - \operatorname{grad} f(y)\|^2.\end{aligned}$$

Označme $r_k := \|x^{[k+1]} - x^*\|$. Pak vzhledem k $\operatorname{grad} f(x^*) = 0$ dostaneme z předchozí nerovnosti

$$\begin{aligned}r_{k+1}^2 &= \|x^{[k+1]} - x^*\|^2 = \\ &= \|x^{[k]} - \alpha \operatorname{grad} f(x^{[k]}) - x^*\|^2 = \\ &= \|x^{[k]} - x^*\|^2 + \\ &\quad + \|\alpha \operatorname{grad} f(x^{[k]})\|^2 + \\ &\quad + 2\langle x^{[k]} - x^*, -\alpha \operatorname{grad} f(x^{[k]})\rangle = \\ &= \|x^{[k]} - x^*\|^2 + \\ &\quad + \alpha^2 \|\operatorname{grad} f(x^{[k]})\|^2 + \\ &\quad - 2\alpha \langle x^{[k]} - x^*, \operatorname{grad} f(x^{[k]}) - \operatorname{grad} f(x^*)\rangle \leq \\ &\leq \|x^{[k]} - x^*\|^2 + \\ &\quad + \alpha^2 \|\operatorname{grad} f(x^{[k]})\|^2 - \\ &\quad - \tfrac{2\alpha\vartheta L}{\vartheta+L}\|x^{[k+1]} - x^*\|^2 - \\ &\quad - \tfrac{2\alpha}{\vartheta+L}\|\operatorname{grad} f(x^{[k]}) - \operatorname{grad} f(x^*)\| = \\ &= \left(1 - \tfrac{2\alpha\vartheta L}{\vartheta+L}\right)\|x^{[k+1]} - x^*\|^2 + \\ &\quad + \left(\alpha^2 - \tfrac{2\alpha}{\vartheta+L}\right)\|\operatorname{grad} f(x^{[k]})\|^2.\end{aligned}$$

Proto pro libovolné $\alpha \in (0, 2/(L+\vartheta)]$ je

$$r_{k+1}^2 \leq \left(1 - \frac{2\alpha\vartheta L}{\vartheta+L}\right) r_k^2,$$

z čehož s využitím (7.3.16) získáme

$$\begin{aligned}f(x^{[k]}) - f(x^*) = \\ = \langle \operatorname{grad} f(x^*), x^{[k]} - x^*\rangle + \\ + \frac{L}{2}\|x^{[k]} - x^*\|^2 = \\ = \frac{L}{2} r_k^2 \leq \frac{L}{2}\left(1 - \frac{2\alpha\vartheta L}{\vartheta+L}\right)^k r_0^2.\end{aligned}$$

Tím je důkaz hotov.

[102] To znamená, že funkce f je v okolí bodu x^* silně konvexní a x^* je její nedegenerované lokální/globální minimum.

Nechť $f \in C^2$, bod x^* je její lokální minimum a[102] $\lambda_{\min}(\nabla^2 f(x^*)) > 0$. Konverguje-li posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná MNS k bodu x^*, pak posloupnost hodnot účelové funkce $\{f(x^{[k]})\}_{k=0}^{\infty}$ konverguje k $f(x^*)$ (alespoň) lineárně s rychlostí (nejméně) $\left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^2$, tj.

$$f(x^{[k+1]}) - f(x^*) \leq \left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^2 \left(f(x^{[k]}) - f(x^*)\right),$$

Tedy i v případě nekvadratické funkce hraje velkou roli podmíněnost matice $\nabla^2 f(x^*)$. Např. pro Rosenbrockovu funkci s parametry $a = 1$ a $b = 100$ mající minimum v $x^* = [1,1]$ jsou vlastní čísla $\lambda_{\max}(\nabla^2 f(x^*)) = 501 + \sqrt{250601}$ a $\lambda_{\min}(\nabla^2 f(x^*)) = 501 - \sqrt{250601}$, takže $\kappa(\nabla^2 f(x^*)) \approx 2508$, viz opět Obrázek 7.57.

Při vhodném oslabení požadavků Věty 7.3.6 můžeme zachovat alespoň poměrně širokou lokální konvergenci MNS. Nezapomínejme, že v případě zdola neohraničené funkce f bude i posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ neohraničená. Opačné tvrzení samozřejmě neplatí, jak se můžeme snadno přesvědčit na Obrázku 7.62.

Věta 7.3.7 Lokální konvergence MNS

Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je spojitě diferencovatelná a množina řešení úlohy (7.3.1) je neprázdná.

(i) Je-li posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná MNS nebo gradientní metodou s využitím Armijovy podmínky pro volbu délky kroku ohraničená, pak každý její hromadný bod je stacionárním bodem funkce f.

(ii) Jestliže $x^{[0]} \in \mathbb{R}^n$ je takové, že množina

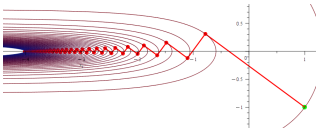
$$S := \{x \in \mathbb{R}^n \mid f(x) \leq f(x^{[0]})\}$$

je ohraničená, pak posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná MNS je také ohraničená, a tudíž všechny hromadné body jsou stacionárními body funkce f.

Důkaz[103]. Část (ii) je přímým důsledkem první části, neboť MNS je spádová metoda, takže v případě $x^{[k+1]} \neq x^{[k]}$ máme $f(x^{[k+1]}) < f(x^{[k]})$. Proto prvky posloupnosti $\{x^{[k]}\}_{k=0}^{\infty}$ musí být obsaženy v dolní obrysové množině S, a tudíž je tato posloupnost ohraničená podle (i) a její hromadné body jsou zároveň stacionárními body f.

Zbývá dokázat část (i). Začneme s gradientní metodou doplněnou o Armijovo pravidlo pro délku kroku. Nechť tedy je posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$, bod $\overline{x}$ jejím nějakým hromadným bodem a připusťme, že $\operatorname{grad} f(\overline{x}) \neq 0$. Hlavní myšlenka celého důkazu je poměrně snadná: jelikož $\operatorname{grad} f(\overline{x}) \neq 0$, učiníme-li krok z $\overline{x}$, pak dojde v souladu s konstrukcí celé metody ke snížení funkční hodnoty o jisté $\delta > 0$. Proto by se intuitivně dalo tušit (a my se o tom i za chvíli přesvědčíme), že musí *existovat malé okolí $\mathcal{O}(\overline{x})$ takové, že při každém kroku MNS nebo gradientní metody z libovolného bodu $x \in \mathcal{O}(\overline{x})$ dojde ke snížení funkční hodnoty funkce f o alespoň nějaké pevné $v > 0$*. Konkrétní hodnota v je pro nás zcela nepodstatná. Potřebujeme pouze zaručit její existenci, kladnost a nezávislost na zvoleném $x \in \mathcal{O}(\overline{x})$. Připusťme, že jsme již toto dokázali. V takovém případě pak velmi snadno docílíme kýženého sporu: jelikož $\overline{x}$ je hromadným bodem posloupnosti $\{x^{[k]}\}_{k=0}^{\infty}$, musí být $x^{[k]} \in \mathcal{O}(\overline{x})$ pro nekonečně mnoho $k \in \mathbb{N}$. Avšak s každým takovým členem musí dojít ke snížení funkční hodnoty o $v > 0$ a současně s žádným krokem nedojde ke zvýšení funkční hodnoty. To ale znamená, že v průběhu výpočtu musí hodnoty účelové funkce $f(x^{[k]})$ klesat bez omezení pro $k \to \infty$, tj. $f(x^{[k]}) \to -\infty$, což je spor s řešitelností úlohy (7.3.1).

Nyní je potřeba dokázat klíčový argument předchozích úvah: existenci $\mathcal{O}(\overline{x})$ a ε. Před tím ale ještě zdůrazněme, že to skutečně jakýsi důkaz vyžaduje navzdory již zmíněné spádovosti uvažovaných metody, tj. že s každým krokem v nestacionárním bodě dojde ke snížení funkční



Obrázek 7.62: Ilustrace několika prvních členů neohraničené posloupnosti $\{x^{[k]}\}_{k=0}^{\infty}$ generované MNS pro ohraničenou funkci $f(x_1,x_2) = e^{x_1}+x_2^2$ s $x^{[0]} = [1,-1]$ a $f^* = 0$.

[103] Důkaz pochází z [51, Theorem 6.1.1]. Alternativní důkaz lze najít v [57, Proposition 1.2.1], kde se však jako Armijova podmínka uplatňuje pouze (7.3.3), což vede k tomu, že posloupnost kroků splňuje $\alpha_k \to 0$. To podmínka (7.3.4) neumožňuje.

hodnoty účelové funkce (a díky spojitosti f je každý bod v dostatečně malé blízkosti nestacionární bodu také nestacionární). Potíž tkví v tom, že snížení hodnoty účelové funkce záleží na aktuálním bodu $x^{[k]}$. Může se tak stát, že v každém bodě v $\mathcal{O}(\overline{x})$ sice dojde ke zlepšení hodnoty účelové funkce, avšak bez jakékoli dolní hranice tohoto snížení. Jenže v našich předchozích úvahách hrála bodová nezávislost tohoto snižování zásadní roli – jinak by totiž kdykoli $x^{[k]} \in \mathcal{O}(\overline{x})$ sice došlo ke snížení, které by však bylo stále menší a menší. Pak by ale součet všech těchto snížení mohl být konečný, což by rozbilo celou naši předchozí argumentaci.

Ve skutečnosti ale takové $\mathcal{O}(\overline{x})$ a v opravdu existují. Stačí nám to dokázat pouze pro gradientní metodu s Armijovým pravidlem, neboť jeden krok MNS poskytuje alespoň stejně tak dobré zlepšení, pokud jej provedeme ze stejného bodu. Jelikož $f \in C^1$ a $\operatorname{grad} f(\overline{x}) \neq 0$, existují kladná čísla r, p, P taková, že

$$p \leq \|\operatorname{grad} f(x)\| \leq P \quad \text{pro všechna } x \in \mathcal{O}_r(\overline{x}).$$

Ze stejného důvodu existuje $r' \in (0,r)$ takové, že pro každé $x, y \in \mathcal{O}_{r'}(\overline{x})$ je

$$\|\operatorname{grad} f(x) - \operatorname{grad} f(y)\| \leq \frac{(1-\varepsilon_1)\,p^2}{P} =: \zeta.$$

Pak pro každé $x \in \mathcal{O}_{r'/2}(\overline{x})$ je délka kroku gradientní metody s Armijovým pravidlem (7.3.3)–(7.3.4) aplikovaným na funkci

$$h_x(\alpha) := f(x - \alpha \operatorname{grad} f(x))$$

nejméně roven $\alpha^* := \frac{r'}{2\eta P}$, což je přesně to, co potřebujeme a dokážeme v následujícím odstavci. Vskutku, snížení hodnoty účelové funkce díky Armijovu pravidlu vedoucímu ke kroku délky α je nejméně $\varepsilon_1 \alpha |h_x'(0)|$, přičemž $h_x'(0) = -\|\operatorname{grad} f(x)\|^2$, viz (7.3.3). Aplikováním tohoto pozorování na jeden krok gradientní metody s Armijovým pravidlem z bodu $x \in \mathcal{O}_{r'/2}(\overline{x})$ dospějeme k závěru, že zlepšení hodnoty účelové funkce f je nejméně $\varepsilon_1 \alpha^* \|\operatorname{grad} f(x)\|^2 \geq \varepsilon_1 \alpha^* p^2$. Hodnota na pravé straně je nezávislá na zvoleném x, a můžeme ji proto vzít jako ono v.

Připusťme, že pro nějaké $x \in \mathcal{O}_{r'/2}(\overline{x})$ je délka kroku určená Armijovým pravidlem $\alpha_x < \alpha^*$. Z (7.3.4) máme

$$h_x(\eta\alpha_x) - h_x(0) > \varepsilon_1 \eta \alpha_x h_x'(0). \tag{7.3.19}$$

Jelikož nyní $\alpha_x < \alpha^*$, úsečka mezi x a $x - \eta\alpha_x \operatorname{grad} f(x)$ má délku nejvýše $\eta\alpha^* P \leq r'/2$, a tudíž vzhledem k tomu, že $x \in \mathcal{O}_{r'/2}(\overline{x})$, musí celá úsečka ležet v $\mathcal{O}_{r'}(\overline{x})$. Proto pro libovolné $\alpha, \beta \in [0, \eta\alpha_x]$ máme

$$\begin{aligned}
&\left|h_x'(\alpha) - h_x'(\beta)\right| = \\
&\quad = \left| -\langle \operatorname{grad} f(x - \alpha \operatorname{grad} f(x)), \operatorname{grad} f(x)\rangle + \langle \operatorname{grad} f(x - \beta \operatorname{grad} f(x)), \operatorname{grad} f(x)\rangle \right| = \\
&\quad = \left|\langle \operatorname{grad} f(x - \beta \operatorname{grad} f(x)) - \operatorname{grad} f(x - \alpha \operatorname{grad} f(x)), \operatorname{grad} f(x)\rangle\right| \leq \\
&\quad \leq \|\operatorname{grad} f(x - \beta \operatorname{grad} f(x)) - \operatorname{grad} f(x - \alpha \operatorname{grad} f(x))\| \times \|\operatorname{grad} f(x)\| \leq \\
&\quad \leq \zeta \|\operatorname{grad} f(x)\| \leq P\zeta = (1-\varepsilon_1)\,p^2,
\end{aligned}$$

neboť se hodnota $\operatorname{grad} f(x)$ podél úsečky mezi x a $x - \eta\alpha_x \operatorname{grad} f(x)$ může změnit nejvýše o ζ, a tudíž derivace h_x na intervalu $[0, \eta\alpha_x]$ se změní nejvýše o $(1-\varepsilon_1)\,p^2$. Toto společně s Větou o střední hodnotě vede k

$$h_x(\eta\alpha_x) - h_x(0) = \eta\alpha_x h_x'(\xi) \leq \eta\alpha_x h_x'(0) + (1-\varepsilon_1)\eta\alpha_x p^2$$

pro vhodné $\xi \in (0, \eta\alpha_x)$. Zkombinováním této nerovnosti s (7.3.19) získáme

$$\begin{aligned}
(1-\varepsilon_1)\eta\alpha_x p^2 &\geq h_x(\eta\alpha_x) - h_x(0) - \eta\alpha_x h_x'(0) > \varepsilon_1 \eta\alpha_x h_x'(0) - \eta\alpha_x h_x'(0) = \\
&= -(1-\varepsilon_1)\eta\alpha_x h_x'(0) = (1-\varepsilon_1)\eta\alpha_x \|\operatorname{grad} f(x)\|^2 \geq (1-\varepsilon_1)\eta\alpha_x p^2,
\end{aligned}$$

což je spor, a důkaz je hotov. ∎

Ještě jednou zdůrazněme to, co jsme již využili v předchozím důkazu: jakákoli volba délky kroku α_k, která dává větší pokles funkčních hodnot v posloupnosti $\{f(x^{[k]})\}$ než Armijovo pra-

vidlo (např. Goldsteinova podmínka), získává tutéž vlastnost, tj. hromadné body $\{x^{[k]}\}$ jsou stacionárními body funkce f. Toto samozřejmě obecně neplatí v případě konstantní volby kroku, ale lze to ukázat pro funkci $f \in \mathrm{C}^1$ s lipschitzovsky spojitým gradientem a volbu délky kroku jako

$$\varepsilon \le \alpha_k \equiv \alpha \le \frac{2-\varepsilon}{L}$$

pro pevně zvolené $\varepsilon > 0$. Důkaz v podobě zahrnující obecnější volbu vektorů h_k, které jsou jakýmsi způsobem přidruženy gradientu grad $f(x^{[k]})$, lze pro konstantní délku kroku najít v [57, Proposition 1.2.3] a pro klesající volbu délky kroku v [57, Proposition 1.2.4].

Souhrn základních vlastností MNS:

(i) 👍 globální konvergence (nezáleží na volbě $x^{[0]}$; pro nekvadratické funkce za dodatečných předpokladů);

(ii) 👎 konvergence je (obvykle) velmi (velmi) pomalá;

(iii) 👎 mnohdy numericky ani nekonverguje;

(iv) 👍 řada jiných metod ale využívá (alespoň jeden krok) MNS ve chvíli, kdy samy neposkytují dostatečné zlepšení nebo selhávají;

(v) 👍 princip je základem pro mnoho velmi efektivních numerických metod, viz např. následující příklad.

Příklad 7.3.8

Gradientní metoda je základem mnoha jiných metod, ačkoli to na první pohled nemusí být patrné. Jedním takovým příkladem může být řešení *Weberova problému* zmíněného již v Příkladu 1.2.9, tj. pro danou m-tici bodů $a_1, \dots, a_m \in \mathbb{R}^n$ (označovaných jako *kotvy*) a m-tici vah $\omega_1, \dots, \omega_m$ nalezněme $x^* \in \mathbb{R}^n$, které je řešení optimalizačního problému

$$\sum_{i=1}^{m} \omega_i \,\|x - a_i\| \to \min, \quad x \in \mathbb{R}^n.$$

Všimněme si, že taková účelová funkce není ani diferencovatelná v bodech $a_1, \dots, a_m$. Jeden z oblíbených přístupů k nalezení řešení pochází z roku 1937 od Weiszfelda[104]. Výchozím bodem je nutná podmínka prvního řádu grad $f(x) = 0$, přičemž implicitně předpokládáme, že $x \neq a_i$ pro všechna $i = 1, \dots, m$. TUto rovnost můžeme explicitně vyjádřit jako

$$\sum_{i=1}^{m} \omega_i \frac{x - a_i}{\|x - a_i\|} = 0,$$

z čehož s trochou námahy získáme

$$\left(\sum_{i=1}^{m} \frac{\omega_i}{\|x - a_i\|}\right) x = \sum_{i=1}^{m} \frac{\omega_i\, a_i}{\|x - a_i\|} \quad \text{neboli} \quad x = \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x-a_i\|}} \sum_{i=1}^{m} \frac{\omega_i\, a_i}{\|x - a_i\|}.$$

Proto podmínku optimality (a vlastně hledání řešení samotného problému) můžeme přeformulovat jako hledání pevného bodu $x = T(x)$ pro operátor

$$T(x) := \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x-a_i\|}} \sum_{i=1}^{m} \frac{\omega_i\, a_i}{\|x - a_i\|}.$$

Proto se zcela přirozeně nabízí využití metody pevného bodu využívající iterace

$$x^{[k+1]} = T(x^{[k]}),$$

což je i Weiszfeldova metoda pro řešení Weberova problému, kde se výchozí $x^{[0]}$ samozřejmě volí různé od $a_1, \dots, a_m$. Pochopitelně tato metoda funguje pouze pokud i všechny následné iterace $x^{[k]}$ budou různá od $a_1, \dots, a_m$. Ačkoli jsme se k této metodě

dopracovali prostřednictvím pevného bodu, jedná se překvapivě současně i o gradientní metodu, neboť

$$x^{[k+1]} = \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x^{[k]} - a_i\|}} \sum_{i=1}^{m} \frac{\omega_i a_i}{\|x^{[k]} - a_i\|} = x^{[k]} - \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x^{[k]} - a_i\|}} \sum_{i=1}^{m} \omega_i \frac{x^{[k]} - a_i}{\|x^{[k]} - a_i\|} =$$
$$= x^{[k]} - \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x^{[k]} - a_i\|}} \operatorname{grad} f(x^{[k]}).$$

V této gradientní metodě je proměnná délka kroku

$$\alpha_k := \frac{1}{\sum_{i=1}^{m} \frac{\omega_i}{\|x^{[k]} - a_i\|}}.$$

Samozřejmě zde zůstává několik nezodpovězených otázek. Je tato metoda korektně definovaná? Jinými slovy, můžeme zaručit, že $x^{[k]}$ nikdy nebude rovno některému z $a_1, \ldots, a_m$? Jsou posloupnost odpovídajících hodnot účelové funkce klesající? Je popsaným postupem získaná posloupnost $\{x^{[k]}\}$ konvergentní? Můžeme sice prozradit, že odpovědi jsou kladné, pokud $f(x^{[0]}) < \min\{f(a_1), \ldots, f(a_m)\}$, ale pro více podrobností raději odkážeme zvídavé čtenáře na [38, str. 70–72].

Nevýhodu MNS v podobě pomalé konvergence při $\beta \approx 1$ lze mírnit různými způsoby. Jednou z možností může být transformace proměnných neboli *škálování* s cílem snížení podmíněnosti Hessovy matice. Uvažme např. funkci

$$f(x_1, x_2) = x_1^2 - 5\,x_1\,x_2 + x_2^4 - 25\,x_1 - 8\,x_2$$

s minimem v bodě $x^* = [20, 3]$ a $f^* = -343$. Podmíněnost Hessovy matice $\nabla^2 f(x^*) = \left(\begin{smallmatrix} 2 & -5 \\ -5 & 108 \end{smallmatrix}\right)$ je cca 61, což naznačuje, že by konvergence MNS mohla být velmi pomalá. O tom se můžeme i snadno přesvědčit ve druhém sloupci Tabulky 7.6.

k-tá iterace	(klasická) MNS	MNS po transformaci $x_2 = y/7$	MNS po transformaci $x_2 = y/2$
1	−230,9958	−334.7231	−266.28517
2	−256,404	−341.2634	−342.9634
3	−279,8423	−342.5611	−342.9983
4	−293,8581	−342.8951	−342.99992
5	−306,19696	−342.9739	−342.9999963
6	−314,2536	−342.9936	−342.9999998
7	−321,14073	−342.9984	−342.999999992
8	−325,8823	−342.9996	−342.99999999963
9	−329,8547	−342.99990	−342.99999999998
10	−332,6878	−342.99997	−342.9999999999992
20	−342,1093		
30	−342,9188		
40	−342,9924		
50	−342,9993		
60	−342,99993		
70	−342,999993		
80	−342.9999994		

Tabulka 7.6: Porovnání funkčních hodnot generovaných klasickou MNS pro funkci $x_1^2 - 5\,x_1\,x_2 + x_2^4 - 25\,x_1 - 8\,x_2$ a pro tutéž funkci po transformaci $y = t\,x_2$.

Jelikož polynomiální funkce f má pouze jeden člen v mocnině vyšší než dva, má Hessova matice $\nabla^2 f(x_1,x_2) = \left(\begin{smallmatrix} 2 & -5 \\ -5 & 12\,x_2^2 \end{smallmatrix}\right)$ jenom jeden nekonstantní člen. Proto se nabízí zkusit transformaci $y = t\,x_2$ pro vhodné t, které by snížilo podmíněnost odpovídající Hessovy matice. Toto nás přivede k matici

$$\nabla^2 f(x_1, y/t) = \begin{pmatrix} 2 & -\frac{5}{t} \\ -\frac{5}{t} & \frac{12\,y^2}{t^4} \end{pmatrix}.$$

Využijeme-li toho, že pro řešení x^* původní úlohy vždy platí $y/t = 3$ bez ohledu na konkrétní volbu t, redukuje se tato matice do podoby

$$\nabla^2 f(x^*) = \nabla^2 f(20,3) = \begin{pmatrix} 2 & -\frac{5}{t} \\ -\frac{5}{t} & \frac{108}{t^2} \end{pmatrix}.$$

s vlastními čísly

$$\lambda_{\min} = \frac{t^2 - \sqrt{t^4 - 83\,t^2 + 2916} + 54}{t^2} \quad \text{a} \quad \lambda_{\max} = \frac{t^2 + \sqrt{t^4 - 83\,t^2 + 2916} + 54}{t^2}.$$

Kdybychom chtěli minimalizovat podmíněnost takovéto matice[105], získali bychom hodnotu $t \approx 7{,}35$ a odpovídající $\kappa \approx 2{,}03$. V [360, Application 3, str. 242–243] volí $t = 7$ a ospravedlňují to přibližnou rovností diagonálních členů. Když tuto hodnotu zvolíme i my, pak funkční hodnoty získané v průběhu výpočtu MNS jsou uvedeny ve třetím sloupci Tabulky 7.6 a je z nich patrné značné urychlení konvergence k číslu -343. Avšak MNS pro nekvadratické funkce je skutečně poněkud mysteriózní. Pro $t = 2$ bude podmíněnost Hessovy matice cca 15, ale konvergence bude extrémně rychlá, jak se můžeme přesvědčit v posledním sloupci Tabulky 7.6.

Ačkoli se konvergence výrazně zrychlila, využili jsme k nalezení vhodného parametru t znalost samotného řešení. To samozřejmě není příliš praktické. Toto můžeme chápat jako ukázku kvazinewtonovských metod, o kterých si řekneme více v závěru příští podkapitoly. Další možností je tzv. *urychlená MNS* pocházející z [396], kde lze volit

$$\lambda_0 := 0 \quad \& \quad \lambda_k := \frac{1 + \sqrt{1 + 4\alpha_{k-1}^2}}{2} \quad \& \quad \alpha_k := \frac{1 - \lambda_k}{\lambda_{k+1}}$$

$$\widetilde{x}^{[k]} := x^{[k]} - \frac{1}{L}\operatorname{grad} f(x^{[k]}) \quad \& \quad x^{[k+1]} := (1-\alpha_k)\,\widetilde{x}^{[k+1]} + \alpha_k\,\widetilde{x}^{[k]}$$

pro funkci f s lipschitzovsky spojitým gradientem s konstantou L, v kterémžto případě získáme posloupnost $\{x^{[k]}\}$ splňující

$$f(\widetilde{x}^{[k+1]}) - f(x^*) \le \frac{2\,L}{k^2}\,\|\,x^{[0]} - x^*\,\|^2$$

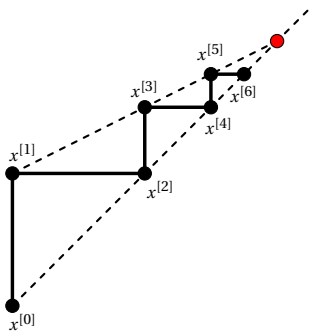
pro všechna $k \in \mathbb{N}$, viz [360, str. 244]. Analogická metoda pro tři různé konstantní volby délek kroků jsou podrobně popsány v [398, str. 78–81].

Za zmínku také rozhodně stojí tzv. *metoda paralelních tečen* (zkráceně PARTAN; též *ovrag-metoda*). V dřívějších numerických experimentech s MNS bylo na dvourozměrných úlohách pozorováno, že společně se značným klikatěním dochází také k tomu, že bod $x^{[k+2]}$ leží poměrně blízko přímky určené body $x^{[k-2]}$ a $x^{[k]}$, viz Obrázky 7.63–7.64. To je samozřejmě způsobeno ortogonalitou jednotlivých směrů a postupným výrazným zkracováním délek jednotlivých kroků, což víme díky dříve popsané konvergenci. Protože každý krok MNS můžeme jako samostatný proces s novým výchozím bodem (není to nijak jinak provázané s předchozími kroky), mohli bychom některé kroky (konkrétně každý třetí) „vynechat“: pro dané $x^{[0]}$ určíme $x^{[1]}$ a $x^{[2]}$ pomocí MNS, ale $x^{[3]} = x^{[2]} - \alpha_2\,h_2$ získáme pomocí směru $h_2 = x^{[2]} - x^{[0]}$ a odpovídajícího α_2, které je řešením jednorozměrné úlohy $\min f(x^{[2]} - \alpha\,h_2)$. Body $x^{[4]}$ a $x^{[5]}$ opět pomocí MNS atd., viz Obrázky 7.65–7.66. Ovšem PARTAN lze formulovat také pouze pomocí $x^{[0]}$ a $x^{[1]}$, avšak s využitím dvou jednorozměrných minimalizací:

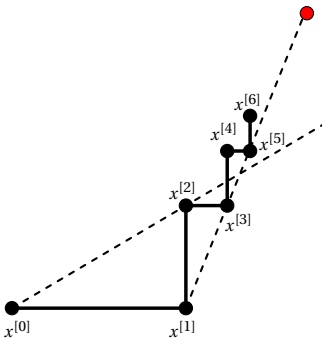
[105] Tj. minimalizovat podíl

$$\frac{\frac{t^2+\sqrt{t^4-83\,t^2+2916}+54}{t^2}}{\frac{t^2-\sqrt{t^4-83\,t^2+2916}+54}{t^2}}$$

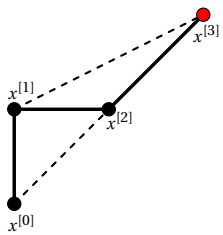
vzhledem k t, což rozhodně není snadný úkol...



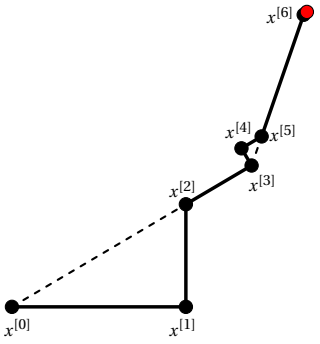
Obrázek 7.63: Vizualizace průběhu MNS pro funkci $f(x_1,x_2) = x_1^2 - 2\,x_1\,x_2 + 2\,x_2^2 - 5\,x_2$ s výchozím bodem $x^{[0]} = [0,0]$. Samotné řešení je červenou barvou.



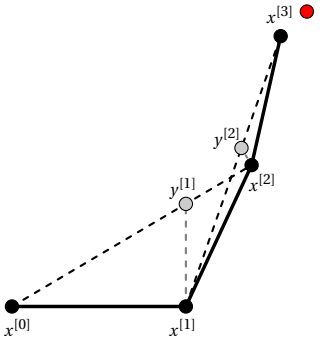
Obrázek 7.64: Vizualizace průběhu MNS pro Rosenbrockovu funkci $f(x_1,x_2) = (x_2 - x_1^2)^2 + (1 - x_1)^2$ s výchozím bodem $x^{[0]} = [0,0]$. Samotné řešení je červenou barvou.



Obrázek 7.65: Vizualizace průběhu „jednoskokové“ PARTAN pro funkci $f(x_1,x_2) = x_1^2 - 2\,x_1\,x_2 + 2\,x_2^2 - 5\,x_2$ s výchozím bodem $x^{[0]} = [0,0]$. Vidíme výrazné urychlení celého procesu. Bod $x^{[3]}$ je spočítán ve směru $h_2 = x^{[2]} - x^{[0]}$ místo $-\operatorname{grad} f(x^{[2]})$ a splývá se samotným řešením, které je červenou barvou!



Obrázek 7.66: Vizualizace průběhu „jednoskokové" PARTAN pro Rosenbrockovu funkci $f(x_1,x_2) = (x_2 - x_1^2)^2 + (1-x_1)^2$ s výchozím bodem $x^{[0]} = [0,0]$. Samotné řešení je červenou barvou.



Obrázek 7.67: Vizualizace průběhu „dvouskokové" PARTAN pro Rosenbrockovu funkci $f(x_1,x_2) = (x_2 - x_1^2)^2 + (1-x_1)^2$ s výchozím bodem $x^{[0]} = [0,0]$. Samotné řešení je červenou barvou.

První minimalizace je provedena (stejně jako u klasické MNS) ve směru $-\operatorname{grad} f(x^{[0]})$, čímž získáme bod $x^{[1]}$. Následně minimalizací ve směru $-\operatorname{grad} f(x^{[1]})$ určíme bod $y^{[1]}$, z něhož pak minimalizací ve směru $h_1 = y^{[1]} - x^{[0]}$ získáme bod $x^{[2]}$. Dále pokračujeme analogickým způsobem, tj. klasickou minimalizací v bodě $x^{[k]}$ ve směru $-\operatorname{grad} f(x^{[k]})$ nalezneme bod $y^{[k]}$, z něhož následnou minimalizací ve směru $h_k = y^{[k]} - x^{[k-1]}$ obdržíme bod $x^{[k+1]}$, viz Obrázek 7.67.

Takto pojatá metoda paralelních tečen vyžaduje pro „podobný" progres v každém kroku (s výjimkou prvního) o jednu jednorozměrnou minimalizaci méně než v první variantě. Podobně jako MNS může být samozřejmě aplikována i na nekvadratické funkce a poprvé byla popsána v článku [451] pro minimalizaci konvexní funkce s cílem nalezení minima kvadratické funkce přesně po n krocích. To je idea, jež stojí i za *Metodou sdružených gradientů* (MSG), které se bude věnovat v přespříští podkapitole, kde si i ukážeme spojení mezi PARTAN a MSG. Velkou výhodou PARTAN je poměrně snadná implementace a zejména charakteristika její globální konvergence. Každý krok je alespoň tak dobrý jako jeden krok MNS: posun z $x^{[k]}$ do $y^{[k]}$ je přesně MNS a dodatečný posun do $x^{[k+1]}$ přináší ještě další zlepšení funkční hodnoty účelové funkce. Proto globální konvergence není vázána na skutečnost, zda je tento proces restartován po n (případně $n+1$) krocích – nicméně „se to" doporučuje, tj. každé $x^{[n+1]}$ je získáno pomocí MNS, čímž docílíme toho, že se tato metoda bude poblíž řešení chovat podobně jako právě MSG. Nevýhodou může být samotná nutnost dvou jednorozměrných minimalizací v každém kroku (kromě prvního), ale tento nedostatek je přirozeně kompenzován tím, že není zcela nutné nalézt přesné řešení (to by mohlo působit potíže i u MSG), tj. i v případě využití jiné volby délky kroku bude PARTAN nejméně stejně „kvalitní" jako MNS.

Algoritmický popis PARTAN:

Zadání: Funkce f, výchozí bod $x^{[0]} \in \mathbb{R}^n$, ukončovací pravidlo (např. $\| x^{[k+1]} - x^{[k]} \| \le \varepsilon$) a volba tolerance $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $k := 0$.

Krok 2: (*První krok*) Určíme $h_k := -\operatorname{grad} f(x^{[k]})$ a α_k takové, že $f(x^{[k]} + \alpha_k h_k) = \min_{\alpha \ge 0} f(x^{[k]} + \alpha h_k)$. Pokračujeme Krokem 5.

Krok 3: (*Iterace y*) Určíme $\tilde{h}_k := -\operatorname{grad} f(x^{[k]})$ a $\tilde{\alpha}_k$ takové, že $f(x^{[k]} + \tilde{\alpha}_k \tilde{h}_k) = \min_{\alpha \ge 0} f(x^{[k]} + \alpha \tilde{h}_k)$. Položíme $y^{[k]} := x^{[k]} + \tilde{\alpha}_k \tilde{h}_k$. Pokračujeme Krokem 4.

Krok 4: (*Iterace x*) Určíme $h_k := y^{[k]} - x^{[k-1]}$ a α_k takové, že $f(y^{[k]} + \alpha_k h_k) = \min_{\alpha \ge 0} f(x^{[k]} + \alpha h_k)$. Pokračujeme Krokem 5.

Krok 5: (*Ověření ukončovacího pravidla*) Položíme $x^{[k+1]} := x^{[k]} + \alpha_k h_k$. Pokud je splněno ukončovací pravidlo KONEC.
V opačném případě položíme $k := k+1$ a pokračujeme Krokem 3.

7.3.2 Newtonova metoda[109]

Viděli jsme, že konvergence MNS může být (a v praxi obvykle je) velmi pomalá. Nyní si naopak ukážeme metodu jejíž konvergence bude velmi rychlá, což ovšem bude vykoupeno celkem silnými požadavky na funkci f. Jedná se o *Newtonovu metodu* (nebo *Newtonovu–Raphsonovu metodu*), jejíž jednorozměrnou variantu jsme si představili již v Podkapitole 7.2.6. Hlavní myšlenka je stále stejná: funkci nahradíme Taylorovým polynomem druhého řádu a hledáme jeho

[109] *Newton's Method*

minimum. Detailněji řečeno, je-li funkce $f \in C^2$, pak ji v $(k+1)$-ním kroku ($k \in \mathbb{N} \cup \{0\}$) aproximujeme Taylorovým polynomem druhého řádu se středem v bodě $x^{[k]}$, tj.

$$T_k(x) := f(x^{[k]}) + \operatorname{grad}^\top f(x^{[k]})\,(x - x^{[k]}) + \frac{1}{2}(x - x^{[k]})^\top \nabla^2 f(x^{[k]})\,(x - x^{[k]}) \approx f(x),$$

a za $x^{[k+1]}$ zvolíme bod, ve kterém nabývá tento polynom nabývá svého minima. Jelikož hledáme řešení úlohy $T_k(x) \to \min$, zderivováním $T_k(x)$ získáme

$$\operatorname{grad} T_k(x) = \operatorname{grad} f(x^{[k]}) + \nabla^2 f(x^{[k]})\,(x - x^{[k]}),$$

což v případě regulární matice $\nabla^2 f(x^{[k]})$ vede díky nutnosti $\operatorname{grad} T_k(x) = 0$ k bodu

$$x^{[k+1]} = x^{[k]} - \left[\nabla^2 f(x^{[k]})\right]^{-1} \operatorname{grad} f(x^{[k]}). \tag{7.3.20}$$

[110] To můžeme chápat tak, že „konvergence je superlineární s řádem ∞".

Opakováním této procedury dostaneme *Newtonovu metodu* (NM) pro řešení úlohy (7.3.1). V případě, že funkce f je dokonce kvadratická, je aproximace Taylorovým polynomem přesná, tj. $f(x) = T_k(x)$. V takovém případě nalezneme řešení úlohy (7.3.1) pomocí (7.3.20) právě v jednom kroku[110]!

Algoritmický popis Newtonovy metody

Zadání: Funkce f, výchozí bod $x^{[0]} \in \mathbb{R}^n$, ukončovací pravidlo (např. $\|x^{[k+1]} - x^{[k]}\| \le \varepsilon$) a volba tolerance $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $k := 0$.

Krok 2: (*Iterace*) Vypočteme $\operatorname{grad} f(x^{[k]})$ a $\nabla^2 f(x^{[k]})$. Je-li $\nabla^2 f(x^{[k]})$ singulární KONEC.
V opačném případě pokračujeme Krokem 3.

Krok 3: (*Ověření ukončovacího pravidla*) Položíme $x^{[k+1]} := x^{[k]} - \left[\nabla^2 f(x^{[k]})\right]^{-1} \operatorname{grad} f(x^{[k]})$. Pokud je splněno ukončovací pravidlo KONEC.
V opačném případě položíme $k := k+1$ a pokračujeme Krokem 2.

[111] Tato nerovnost je splněna díky tomu, že zbytek v Taylorově rozvoji obsahuje třetí parciální derivace funkce f, které jsou spojité, a tudíž i ohraničené na dostatečně malém okolí bodu x^*.

[112] Je-li $A \in \mathbb{R}^{n\times n}$, pak její *spektrální normou* rozumíme číslo $\|A\|_\sigma := \max\{\sqrt{\lambda} \mid \lambda \text{ je vlastní číslo matice } A^*A\}$. Jedná se o maticovou normu indukovanou Eukl(e)idovskou vektorovou normou $\|\cdot\|_2$, tj. platí $\|Au\|_2 \le \|A\|_\sigma \|u\|_2$.

Absence regulárnosti matice $\nabla^2 f(x^{[k]})$ celou situaci velmi výrazně zkomplikuje, takže tento požadavek zatím raději vypouštět nebudeme, i když se jej později pokusíme trochu oslabit. Spokojíme-li se však pouze s touto podmínkou a s existencí řešení úlohy (7.3.1), obdržíme následující tvrzení o lokální konvergenci NM.

Věta 7.3.9 Lokální konvergence NM

Nechť $f \in C^3$ v okolí bodu $x^* \in \mathbb{R}^n$, který je nedegenerovaným minimem, tj. $\operatorname{grad} f(x^*) = 0$ a $\nabla^2 f(x^*) > 0$. Potom pro $x^{[0]} \in \mathbb{R}^n$ dostatečně blízko x^* konverguje posloupnost $\{x^{[k]}\}_{k=0}^\infty$ generovaná NM k bodu x^* superlineárně s řádem (alespoň) $p = 2$ (tj. alespoň kvadraticky).

Důkaz. Připomeňme, že Taylorův rozvoj druhého řádu pro $\operatorname{grad} f(x)$ se středem v $x^{[0]}$ je

$$\operatorname{grad} f(x) - \operatorname{grad} f(x^{[0]}) - \nabla^2 f(x^{[0]})\,(x - x^{[0]}) = \omega(x - x^{[0]}),$$

viz str. 34. Jelikož $f \in C^3$ a $\nabla^2 f(x^*)$ je regulární, existují kladné konstanty ε, c_1, c_2 takové, že pro jakákoli $x^{[0]}, x \in \mathcal{O}_\varepsilon(x^*)$ máme

$$\|\operatorname{grad} f(x) - \operatorname{grad} f(x^{[0]}) - \nabla^2 f(x^{[0]})\,(x - x^{[0]})\| \le c_1 \|x - x^{[0]}\|^2$$

a[111] současně existuje $\left[\nabla^2 f(x)\right]^{-1}$, přičemž její *spektrální norma*[112] splňuje[113]

$$\left\| \left[\nabla^2 f(x)\right]^{-1} \right\|_\sigma \le c_2. \tag{7.3.22}$$

[113] Toto vyplývá z tvrzení: *Nechť $A : \mathbb{R}^r \to \mathbb{R}^{n\times n}$ je maticová funkce, která je spojitá v bodě ξ_0. Pokud existuje $A(\xi_0)^{-1}$, pak existuje i $A(\xi)^{-1}$ pro všechna ξ dostatečně blízko ξ_0 a funkce $A(\cdot)^{-1}$ je spojitá v ξ_0.* Viz [287, Lemma 5.3].

Důkaz. Nejdříve ukážeme existenci $A(\xi)^{-1}$ pro všechna ξ dostatečně blízko ξ_0. Platí

$$A(\xi) = A(\xi_0) - A(\xi_0) + A(\xi) = \\ = A(\xi_0) \times \big(I - A(\xi)\big),$$

kde

$$K(\xi) := A(\xi_0)^{-1}\big(A(\xi_0) - A(\xi)\big).$$

Pak

$$\| K(\xi)\| \le \| A(\xi_0)^{-1}\| \times \| A(\xi_0) - A(\xi)\|$$

a současně

$$\lim_{\|\xi-\xi_0\|\to 0} \| K(\xi)\| = 0. \qquad (7.3.21)$$

Protože A je spojitá v ξ_0, pro všechna ξ dostatečně blízko ξ_0 a vhodné $\theta \in (0,1)$ máme

$$\| A(\xi_0) - A(\xi)\| \le \frac{\theta}{\| A(\xi_0)^{-1}\|}.$$

Tudíž $\| K(\xi)\| \le \theta < 1$ a $\big(I - K(\xi)\big)^{-1}$ existuje. Pak ale

$$\begin{aligned} A(\xi)^{-1} &= \big(A(\xi_0)\big(I - K(\xi)\big)\big)^{-1} = \\ &= \big(I - K(\xi)\big)^{-1} A(\xi_0)^{-1}, \end{aligned}$$

což ukazuje, že $A(\xi)^{-1}$ pro všechna ξ dostatečně blízko ξ_0.

K důkazu spojitosti $A(\cdot)^{-1}$ nejdříve poznamenejme, že

$$\begin{aligned} &\| A(\xi_0)^{-1} - A(\xi)^{-1}\| = \\ &= \| A(\xi)^{-1} - A(\xi_0)^{-1}\| = \\ &= \| \big(\big(I - K(\xi)\big)^{-1} - I\big) A(\xi_0)^{-1}\|. \end{aligned}$$

Avšak protože $\| K(\xi)\| < 1$, máme

$$\begin{aligned} &\big(I - K(\xi)\big)^{-1} - I = \\ &= K(\xi) + K^2(\xi) + \cdots = \\ &= K(\xi)\big(I + K(\xi) + \cdots\big). \end{aligned}$$

Proto

$$\begin{aligned} &\| \big(I - K(\xi)\big)^{-1} - I\| \le \\ &\le \| K(\xi)\| \times \\ &\quad \times \big(1 + \| K(\xi)\| + \| K(\xi)\|^2 + \cdots\big) = \\ &= \frac{\| K(\xi)\|}{1 - \| K(\xi)\|}. \end{aligned}$$

Odtud plyne

$$\begin{aligned} &\| A(\xi)^{-1} - A(\xi_0)^{-1}\| \le \\ &\le \frac{\| K(\xi)\|}{1 - \| K(\xi)\|} \| A(\xi_0)^{-1}\|. \end{aligned}$$

Jelikož platí (7.3.21), plyne odtud

$$\lim_{\|\xi-\xi_0\|\to 0} \| A(\xi)^{-1} - A(\xi_0)^{-1}\| = 0,$$

čímž je důkaz hotov. ■

Věta 7.3.10 „Globální“ konvergence NM

Nechť $x^{[0]} \in \mathcal{O}_\varepsilon(x^*)$. Pak volbou $x = x^*$ v předchozí nerovnosti dostaneme

$$\| \nabla^2 f(x^{[0]})\,(x^{[0]} - x^*) - \operatorname{grad} f(x^{[0]})\| \le c_1 \| x^* - x^{[0]}\|^2. \qquad (7.3.23)$$

Pak pro další krok Newtonovy metody dostaneme

$$\begin{aligned} \| x^{[1]} - x^*\| &= \Big\| x^{[0]} - x^* - \big[\nabla^2 f(x^{[0]})\big]^{-1} \operatorname{grad} f(x^{[0]})\Big\| = \\ &= \Big\| \big[\nabla^2 f(x^{[0]})\big]^{-1} \times \big[\nabla^2 f(x^{[0]})\big(x^{[0]} - x^*\big) - \operatorname{grad} f(x^{[0]})\big]\Big\| \le \\ &\le \big\| \big[\nabla^2 f(x^{[0]})\big]^{-1}\big\|_\sigma \times \big\| \nabla^2 f(x^{[0]})\big(x^{[0]} - x^*\big) - \operatorname{grad} f(x^{[0]})\big\|_\sigma, \end{aligned}$$

což společně s (7.3.22)–(7.3.23) dává

$$\| x^{[1]} - x^*\| \le c_1 c_2 \| x^{[0]} - x^*\|^2.$$

Je-li

$$\| x^{[0]} - x^*\| \le \frac{\alpha}{c_1 c_2}$$

pro nějaké $\alpha \in (0,1)$, pak

$$\| x^{[1]} - x^*\| \le \alpha \| x^{[0]} - x^*\|$$

a indukcí získáme

$$\| x^{[k+1]} - x^*\| \le c_1 c_2 \| x^{[k]} - x^*\|^2 \quad \text{a} \quad \| x^{[k+1]} - x^*\| \le \alpha \| x^{[k]} - x^*\|. \qquad (7.3.24)$$

Proto

$$\lim_{k\to\infty} \| x^{[k]} - x^*\| = 0,$$

u tudíž posloupnost $\{x^{[k]}\}_{k=0}^\infty$ generovaná NM konverguje k x^* a vzhledem k (7.3.24) je řád této konvergence nejméně 2.

Věta 7.3.9 platí dokonce v případě $\det \nabla^2 f(x^*) \neq 0$, kdy např. pro nedegenerované lokální maximum x^* bude při dostatečně blízké počáteční aproximaci $x^{[0]} \in \mathbb{R}^n$ posloupnost $\{x^{[k]}\}$ generovaná NM konvergovat k tomuto maximu x^*.

Co to ale znamená *dostatečně blízko* ve Větě 7.3.9? Z praktického pohledu vlastně nemáme vůbec žádnou informaci, jak zvolit $x^{[0]}$ — ani z důkazu tuto informaci nezískáme, museli bychom totiž znát x^*. Nicméně při zesílení požadavků na funkci f dostaneme následující tvrzení zpřesňující „oblast lokální konvergence“. Za pozornost stojí také skutečnost, že za uvedených předpokladů bude bod $x^{[k+1]}$ dokonce globálním minimem $T_k(x)$.

Nechť $f \in \mathrm{C}^2$ je silně konvexní s konstantou silné konvexnosti $\vartheta > 0$ a funkce $\nabla^2 f(x)$ lipschitzovsky spojitá na $\mathbb{R}^n$ s konstantou M, tj.

$$\| \nabla^2 f(x) - \nabla^2 f(y)\| \le M \| x - y\| \quad \text{pro všechna } x, y \in \mathbb{R}^n. \qquad (7.3.25)$$

Je-li počáteční aproximace $x^{[0]}$ taková, že

$$\| \operatorname{grad} f(x^{[0]})\| < \frac{8\vartheta^2}{M}, \qquad (7.3.26)$$

pak posloupnost $\{x^{[k]}\}_{k=0}^\infty$ generovaná NM konverguje k x^* (což je jediné minimum funkce f na $\mathbb{R}^n$, tj. řešení úlohy (7.3.1)) superlineárně s (alespoň) řádem $p = 2$, tj.

$$\| x^{[k]} - x^*\| \le \frac{4\vartheta}{M} \beta^{2^k}, \quad \text{kde } \beta = \frac{M \| \operatorname{grad} f(x^{[0]})\|}{8\vartheta^2} \in (0,1).$$

Důkaz. Protože funkce f je silně konvexní a spojitá, má podle Věty 6.2.28 jediné globální minimum na $\mathbb{R}^n$, tedy bod x^* je určen jednoznačně. Podle Důsledku 6.4.7 také platí

$$\langle \nabla^2 f(x)\, h, h\rangle \ge 2\vartheta \| h\|^2 \qquad (7.3.27)$$

pro libovolné $x, h \in \mathbb{R}^n$. Protože $\vartheta > 0$, je matice $\nabla^2 f(x)$ pozitivně definitní, a tudíž i regulární. Proto posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ určená vztahem (7.3.20) je definována korektně. Podle Důsledku 6.4.6 máme s využitím Cauchyho–Schwarzovy nerovnosti, že

$$\| x^{[k]} - x^* \|^2 \le \frac{1}{2\vartheta} \langle \operatorname{grad} f(x^{[k]}) - \underbrace{\operatorname{grad} f(x^*)}_{=0}, x^{[k]} - x^* \rangle \overset{\text{C–S}}{\le} \frac{1}{2\vartheta} \| \operatorname{grad} f(x^{[k]}) \| \times \| x^{[k]} - x^* \|$$

neboli

$$\| x^{[k]} - x^* \| \le \frac{1}{2\vartheta} \| \operatorname{grad} f(x^{[k]}) \|. \tag{7.3.28}$$

Nyní odhadneme hodnotu $\| \operatorname{grad} f(x^{[k]}) \|$. Jelikož $f \in \mathrm{C}^2$ platí

$$\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{grad} f\big(x^{[k]} + t(x^{[k+1]} - x^{[k]})\big) = \nabla^2 f\big(x^{[k]} + t(x^{[k+1]} - x^{[k]})\big) \times \big(x^{[k+1]} - x^{[k]}\big),$$

takže

$$\begin{aligned} \operatorname{grad} f(x^{[k+1]}) - \operatorname{grad} f(x^{[k]}) - \nabla^2 f(x^{[k]}) \times \big(x^{[k+1]} - x^{[k]}\big) = \\ = \int_0^1 \Big[\nabla^2 f\big(x^{[k]} + t(x^{[k+1]} - x^{[k]})\big) - \nabla^2 f(x^{[k]})\Big]\big(x^{[k+1]} - x^{[k]}\big)\,\mathrm{d}t, \end{aligned}$$

z čehož společně s (7.3.25) plyne

$$\begin{aligned} \big\| \operatorname{grad} f(x^{[k+1]}) - \operatorname{grad} f(x^{[k]}) - \nabla^2 f(x^{[k]}) \times \big(x^{[k+1]} - x^{[k]}\big) \big\| \le \\ \le \int_0^1 M \big\| x^{[k]} + t(x^{[k+1]} - x^{[k]}) - x^{[k]} \big\| \times \big\| x^{[k+1]} - x^{[k]} \big\|\,\mathrm{d}t = \\ = M \big\| x^{[k+1]} - x^{[k]} \big\|^2 \int_0^1 t\,\mathrm{d}t = \frac{M}{2} \big\| x^{[k+1]} - x^{[k]} \big\|^2. \end{aligned}$$

Pak s využitím rovnosti (7.3.20) dostaneme

$$\begin{aligned} \big\| \operatorname{grad} f(x^{[k+1]}) \big\| = \big\| \operatorname{grad} f(x^{[k+1]}) - \operatorname{grad} f(x^{[k]}) + \nabla^2 f(x^{[k]}) \times \big[\nabla^2 f(x^{[k]})\big]^{-1} \operatorname{grad} f(x^{[k]}) \big\| \le \\ \le \frac{M}{2} \big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} \operatorname{grad} f(x^{[k]}) \big\|^2 \le \\ \le \frac{M}{2} \big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} \big\|_\sigma^2 \times \big\| \operatorname{grad} f(x^{[k]}) \big\|, \end{aligned} \tag{7.3.29}$$

[114] Viz Poznámku 112.

kde $\| \cdot \|_\sigma$ značí *spektrální normu*[114]. Nyní ještě zbývá odhadnout hodnotu $\big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} \big\|_\sigma$. Položme $h := \big[\nabla^2 f(x^{[k]})\big]^{-1} y$ pro $y \in \mathbb{R}^n$. Pak máme

$$\big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} y \big\|^2 \overset{(7.3.27)}{\le} \frac{1}{2\vartheta} \big\langle y, \big[\nabla^2 f(x^{[k]})\big]^{-1} y \big\rangle \overset{\text{C–S}}{\le} \frac{1}{2\vartheta} \| y \| \times \big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} y \big\|,$$

z čehož díky vztahu mezi spektrální a Eukl(e)idovskou normou plyne

$$\big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} \big\|_\sigma = \sup_{y \ne 0} \frac{\big\| \big[\nabla^2 f(x^{[k]})\big]^{-1} y \big\|^2}{\| y \|} \le \frac{1}{2\vartheta}.$$

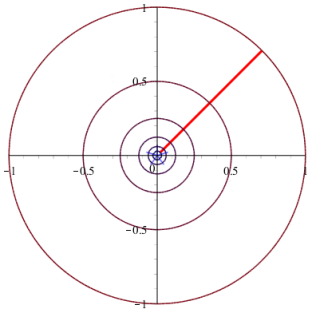
Odtud a z (7.3.29) obdržíme

$$\big\| \operatorname{grad} f(x^{[k+1]}) \big\| \le \frac{M}{2} \Big(\frac{1}{2\vartheta}\Big)^2 \| \operatorname{grad} f(x^{[k]}) \|^2 = \frac{M}{8\vartheta^2} \big\| \operatorname{grad} f(x^{[k]}) \big\|^2.$$

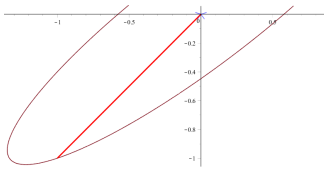
Neboť tento vztah platí pro všechna $k \in \mathbb{N} \cup \{0\}$, indukcí odtud získáme

$$\begin{aligned} \big\| \operatorname{grad} f(x^{[k]}) \big\| \le \frac{M}{8\vartheta^2} \big\| \operatorname{grad} f(x^{[k-1]}) \big\|^2 \le \frac{M}{8\vartheta^2} \Big(\frac{M}{8\vartheta^2}\Big)^2 \big\| \operatorname{grad} f(x^{[k-2]}) \big\|^4 \le \\ \le \cdots \le \Big(\frac{M}{8\vartheta^2}\Big)^{2^k - 1} \big\| \operatorname{grad} f(x^{[k-2]}) \big\|^{2^k} = \frac{8\vartheta^2}{M} \beta^{2^k}, \end{aligned}$$

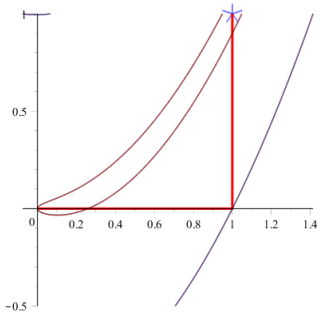
což společně s (7.3.28) dokazuje poslední část tvrzení. ■



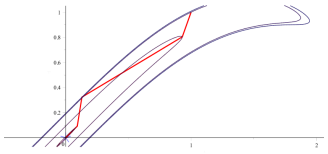
Obrázek 7.68: Ilustrace řešení Příkladu 7.3.11.



Obrázek 7.69: Několik prvních iterací NM pro kvadratickou funkci $f(x_1,x_2)=3x_1^2-7x_1x_2+5x_2^2$ s $x^{[0]}=[-1,-1]$.



Obrázek 7.70: Několik prvních iterací NM pro Rosenbrockovu funkci $f(x_1,x_2)=(1-x_1)^2+100(x_2-x_1^2)^2$ s $x^{[0]}=[0,0]$.



Obrázek 7.71: Několik prvních iterací NM pro testovací funkci $f(x_1,x_2)=10(x_2-\sin x_1)^2+x_1^2/10$ s $x^{[0]}=[1,1]$.

[115] Existují silně konvexní funkce, pro které je při nevhodně zvoleném $x^{[0]}$ posloupnost $\{f(x^{[k]})\}_{k=0}^{\infty}$ generovaná pomocí NM neklesající. Např. pro $f(x_1,x_2)=-1/(x_1^2+x_2^2+1)$ s počáteční aproximací $x^{[0]}=[a,0]$ pro $a>0$ dostatečně velké. Další takovou funkci najdeme v Příkladu 7.3.12.

Příklad 7.3.11

Pomocí NM určeme první iteraci $x^{[1]}$ pro řešení úlohy

$$f(x_1,x_2)=(x_1^2+x_2^2)^{3/2}\to\min$$

při počáteční aproximaci $x^{[0]}=[1/\sqrt{2},1/\sqrt{2}]$, viz Obrázek 7.68 a porovnejme jej s Obrázkem 7.55.

Řešení. Přímým výpočtem získáme

$$\operatorname{grad} f(x^{[0]})=\begin{pmatrix}3/\sqrt{2}\\3/\sqrt{2}\end{pmatrix}\quad\text{a}\quad\nabla^2 f(x^{[0]})=\begin{pmatrix}9/2&3/2\\3/2&9/2\end{pmatrix}\quad\text{a}\quad\left[\nabla^2 f(x^{[0]})\right]^{-1}=\frac{1}{18}\begin{pmatrix}9/2&-3/2\\-3/2&9/2\end{pmatrix},$$

takže

$$x^{[1]}=\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}-\frac{1}{18}\begin{pmatrix}9/2&-3/2\\-3/2&9/2\end{pmatrix}\begin{pmatrix}3/\sqrt{2}\\3/\sqrt{2}\end{pmatrix}=\begin{pmatrix}\frac{1}{2\sqrt{2}}\\\frac{1}{2\sqrt{2}}\end{pmatrix}.$$

Další ilustrace průběhu NM na již dříve zmíněných funkcích můžeme vidět na Obrázcích 7.69–7.71. Z předchozích dvou vět plynou některé výhody a nevýhody NM.

Souhrn základních vlastností NM:

(i) 👍 výhodou je velmi rychlá konvergence;

(ii) 👎 nevýhodou je nutnost dostatečně blízké počáteční aproximace[115];

(iii) 👎 ještě větší nevýhodou je velmi vysoká početní náročnost pro výpočet $\left[\nabla^2 f(x^{[k]})\right]^{-1}$ při velkých hodnotách n (řádově jde o n^3 aritmetických operací při využití Gaussovy–Jordanovy eliminace) a zatížení tohoto výpočtu numerickými chybami v případě, kdy $\det\nabla^2 f(x^{[k]})$ je blízko k nulové hodnotě.

V některých velmi speciálních případech lze hranici „dostatečné blízkosti“ počáteční aproximace určit poměrně snadno, jako tomu je v následujícím příkladě.

Příklad 7.3.12

Pomocí NM určeme první iteraci $x^{[1]}$ pro řešení úlohy

$$f(x_1,x_2)=\sqrt{1+x_1^2+x_2^2}\to\min$$

při počáteční aproximaci $x^{[0]}=[1/2,1/2]$. A co při volbě $\widetilde{x}^{[0]}=[1,1]$? Kdy se „pokazí“ konvergence?

Řešení. Bude-li $x^{[0]}=[1/2,1/2]$, pak přímým výpočtem dostaneme

$$\operatorname{grad} f(x^{[0]})=\begin{pmatrix}1/\sqrt{6}\\1/\sqrt{6}\end{pmatrix}\quad\text{a}\quad\nabla^2 f(x^{[0]})=\frac{\sqrt{6}}{18}\begin{pmatrix}5&-1\\-1&5\end{pmatrix}\quad\text{a}\quad\left[\nabla^2 f(x^{[0]})\right]^{-1}=\frac{\sqrt{6}}{8}\begin{pmatrix}5&1\\1&5\end{pmatrix},$$

takže

$$x^{[1]}=\begin{pmatrix}1/2\\1/2\end{pmatrix}-\frac{\sqrt{6}}{18}\begin{pmatrix}5&-1\\-1&5\end{pmatrix}\begin{pmatrix}1/\sqrt{6}\\1/\sqrt{6}\end{pmatrix}=\begin{pmatrix}-1/4\\-1/4\end{pmatrix},$$

čímž jsme se přiblížili k řešení $x^*=[0,0]$. Čtenář se snadno může přesvědčit, že každá další iterace bude blíže a blíže k x^*, tj. pro tento výchozí bod NM konverguje.

Na druhou stranu pro $\widetilde{x}^{[0]}=[1,1]$ budeme mít

$$\operatorname{grad} f(\widetilde{x}^{[0]})=\begin{pmatrix}\sqrt{3}/3\\\sqrt{3}/3\end{pmatrix}\quad\text{a}\quad\nabla^2 f(\widetilde{x}^{[0]})=\frac{\sqrt{3}}{9}\begin{pmatrix}2&-1\\-1&2\end{pmatrix}\quad\text{a}\quad\left[\nabla^2 f(\widetilde{x}^{[0]})\right]^{-1}=\begin{pmatrix}2\sqrt{3}&\sqrt{3}\\\sqrt{3}&2\sqrt{3}\end{pmatrix},$$

takže

$$\widetilde{x}^{[1]}=\begin{pmatrix}1\\1\end{pmatrix}-\frac{\sqrt{3}}{9}\begin{pmatrix}2&-1\\-1&2\end{pmatrix}\begin{pmatrix}\sqrt{3}/3\\\sqrt{3}/3\end{pmatrix}=\begin{pmatrix}-2\\-2\end{pmatrix},$$

čímž jsme se od bodu x^* vzdálili, tj. pro tento výchozí bod NM nekonverguje.
V tomto poměrně jednoduchém příkladě můžeme určit „oblast konvergence“ NM, tj. množinu vhodných $x^{[0]}$. Položíme-li $x_1 = x_2 = t$, získáme funkci

$$g(t) = f(t,t) = \sqrt{1+2\,t^2} \quad \text{s} \quad g'(t) = \frac{2\,t}{\sqrt{1+2\,t^2}} \quad \text{a} \quad g''(t) = \frac{2}{(1+2\,t^2)^{3/2}}.$$

Vezmeme-li (bez újmy na obecnosti) $t_0 > 0$ libovolné, pak NM dostaneme bod

$$t_1 = t_0 - \frac{2}{(1+2\,t^2)^{3/2}} \frac{2\,t_0}{\sqrt{1+2\,t_0^2}} = t_0 - t_0\,(1+2\,t_0^2).$$

V mezním případě mezi konvergencí a nekonvergencí nastane $t_1 = \pm t_0$. Rovnost $t_1 = t_0$ není možná, zatímco $t_1 = -t_0$ vede na rovnici

$$t_0 = -t_0 + t_0\,(1+2\,t_0^2)$$

s řešením $t_0 = \sqrt{2}/2$. Skutečně zvolíme-li takovéto t_0, pak NM bude generovat posloupnost $\{t_0, -t_0, t_0, \ldots\}$. Proto v případě $\|x^{[0]}\|_2 \geq 1$ NM nekonverguje, zatímco kdykoli $\|x^{[0]}\|_2 < 1$ konverguje[116].

[116] Pro jistotu zdůrazněme, že zde máme na mysli Eukl(e)idovskou normu.

Zmíněné nedostatky NM lze odstranit (alespoň částečně) různými modifikacemi. Nejdříve si můžeme všimnout, že i NM lze brát jako iterační proces podobný MNS[117], tj. postupně generujeme body

$$x^{[k+1]} = x^{[k]} + \alpha_k\,h_k,$$

[117] Nebo spíše podobný gradientní metodě s pevnou délkou kroku.

kde ale tentokrát máme

$$h_k := -\left[\nabla^2 f(x^{[k]})\right]^{-1} \operatorname{grad} f(x^{[k]}) \quad \text{a} \quad \alpha_k \equiv 1.$$

První přirozenou modifikací této metody je proto volit α_k právě jako v MNS, tj.

$$\alpha_k = \operatorname{argmin}_{\alpha \geq 0} \{f(x^{[k]} - \alpha\,h_k)\},$$

čímž dostaneme tzv. *spádovou NM*. Skutečně? A je původní NM také spádovou? Odpověď dává následující tvrzení.

Věta 7.3.13

Nechť $f \in \mathrm{C}^2$ a $\{x^{[k]}\}_{k=0}^{\infty}$ je posloupnost generovaná NM. Jestliže $\nabla^2 f(x^{[k]}) > 0$ a $\operatorname{grad} f(x^{[k]}) \neq 0$, potom směr

$$h_k := -\left[\nabla^2 f(x^{[k]})\right]^{-1} \operatorname{grad} f(x^{[k]}) = x^{[k+1]} - x^{[k]}$$

z bodu $x^{[k]}$ do $x^{[k+1]}$ je spádovým směrem funkce f v tom smyslu, že existuje $\tilde{\alpha} > 0$ takové, že pro každé $\alpha \in (0, \tilde{\alpha})$ platí

$$f(x^{[k]} + \alpha\,h_k) < f(x^{[k]}).$$

Důkaz. Podobně jako v důkazu Věty 7.3.1 dostaneme pro funkci $\varphi_k(\alpha) := f(x^{[k]} + \alpha\,h_k)$ s využitím pravidla pro derivování složené funkce, že

$$\varphi_k'(\alpha) = \left\langle \operatorname{grad} f(x^{[k]} + \alpha\,h_k), h_k \right\rangle.$$

Odtud plyne

$$\varphi_k'(0) = \left\langle \operatorname{grad} f(x^{[k]}), h_k \right\rangle = -\operatorname{grad}^{\top} f(x^{[k]}) \left[\nabla^2 f(x^{[k]})\right]^{-1} \operatorname{grad} f(x^{[k]}) < 0,$$

neboť $\nabla^2 f(x^{[k]}) > 0$ a $\operatorname{grad} f(x^{[k]}) \neq 0$. Tudíž $\varphi_k(\alpha)$ je klesající v okolí $\alpha = 0$, tj. existuje $\tilde{\alpha} > 0$ takové, že pro všechna $\alpha \in (0, \tilde{\alpha})$ je $\varphi_k(\alpha) < \varphi_k(0)$ neboli $f(x^{[k]} + \alpha\,h_k) < f(x^{[k]})$. ∎

Samozřejmě α_k můžeme volit i jiným vhodným způsobem, obzvláště některým z těch popsaných na začátku této podkapitoly, čímž dostaneme další varianty spádové NM. Pro konvergenci takové modifikace je potřebné, aby

$$\text{z nerovnosti } \operatorname{grad} f(x) \neq 0 \quad \text{vždy plynulo} \quad \operatorname{grad}^\top f(x) \left[\nabla^2 f(x)\right]^{-1} \operatorname{grad} f(x) > 0.$$

To je jistě splněno v případě, kdy funkce f je (ostře) konvexní s regulární Hessovou maticí, tj. platí pro ni $\nabla^2 f(x) > 0$ pro všechna $x \in \mathbb{R}^n$. Důkaz následujícího tvrzení je v podstatě stejný jako pro Větu 7.3.7, a tak jej můžeme vynechat. Věta je sice formulována pouze pro klasickou spádovou NM, tj. doplněnou o přesnou minimalizaci, ale její tvrzení zůstane v platnosti i v případě použití Armijovy nebo Goldsteinovy podmínky.

Věta 7.3.14

Nechť $f \in C^2$ je konvexní funkce s $\nabla^2 f(x) > 0$ pro všechna $x \in \mathbb{R}^n$ a $x^{[0]}$ je takové, že dolní obrysová množina

$$\left\{x \in \mathbb{R}^n \mid f(x) \leq f(x^{[0]})\right\}$$

přidružená bodu $x^{[0]}$ je ohraničená. Pak posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ generovaná spádovou NM konverguje k jedinému globálnímu minimu funkce f.

Ovšem jenom pouhé zakomponování jednorozměrné minimalizace do NM nemusí být úplně šťastné. Podmínky zaručující její funkčnost jsou poměrně silné a hlavně při takové modifikaci *můžeme přijít o tu největší výhodu NM* — o kvadratickou konvergenci. Navíc se v případě obecnější funkce může stát, že směr h_k nebude spádový, a tudíž nebude možné učinit další krok. Také stále ještě musíme počítat $\left[\nabla^2 f(x^{[k]})\right]^{-1}$. Poslední zmíněný nedostatek odstraňují tzv. *kvazinewtonovské metody*[118] (příp. *škálované verze metody největšího spádu*[119] nebo *metoda proměnné metriky*[120]), které napodobují NM s tím, že matice $\left[\nabla^2 f(x^{[k]})\right]^{-1}$ je nahrazena jistou maticí A_k jako např. $A_k = \nabla^2 f(x^{[k]}) + \mu_k I)^{-1}$ pro vhodné[121] $\mu_k \geq 0$ v Levenbergově–Marquardtově modifikaci, viz [287, Podkapitoly 9.3–9.4]. Tím obecně dostáváme

$$x^{[k+1]} = x^{[k]} - \alpha_k A_k g_k,$$

kde $A_k \in \mathbb{R}^{n\times n}$, $\alpha_k > 0$ a $g_k := \operatorname{grad} f(x^{[k]})$. Volbou $A_k = I$ se vrátíme zpět k MNS, zatímco volba $A_k = \left[\nabla^2 f(x^{[k]})\right]^{-1}$ odpovídá právě studované NM. Tato myšlenka nás (s trochou nadsázky) přivede k následující metodě, kterou lze považovat za rozumný kompromis mezi MNS a NM.

[118] *Quasi-Newton Method*

[119] *Scaled Steepest Descent Method*

[120] *Variable Metric Method*

[121] Jedná se o tzv. *regularizační parametr*.

7.3.3 Metoda sdružených gradientů[125]

[125] *Conjugate Gradient Method*

Vraťme se opět k situaci, kdy v úloze (7.3.1) máme pouze kvadratickou funkci

$$f(x) = \frac{1}{2} x^\top Q x - b^\top x, \tag{7.3.30}$$

přičemž $Q = Q^\top \in \mathbb{R}^{n\times n}$, $Q > 0$ a $b \in \mathbb{R}^n$. V takovém případě je nalezení řešení úlohy (7.3.1) ekvivalentní s řešením soustavy

$$Qx = b. \tag{7.3.31}$$

Úlohu (7.3.31) sice umíme řešit např. Gaussovou eliminací nebo pomocí inverzní matice, jenže Gaussova eliminace je zbytečně zdlouhavá a určení inverzní matice je výpočetně poměrně náročné[126]. V tomto případě, kdy matice Q je reálná, symetrická a s plnou hodností, máme ještě jednu možnost. Můžeme ji pomocí ortogonální matice V složené z normalizovaných vlastních vektorů diagonalizovat, tj. $Q = V^\top Q V = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$. V takovém případě bychom transformací proměnných $y = V^\top x$ neboli $x = Vy$ získali úlohu

$$f(y) = \frac{1}{2} y^\top V^\top Q V y - y^\top V^\top b = \frac{1}{2} y^\top \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\} y - y^\top d$$

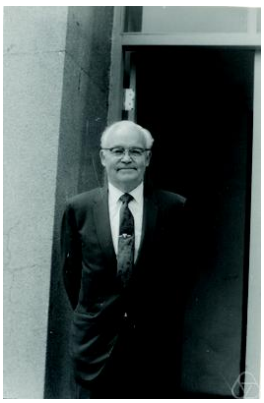
pro $d := V^\top b$, tj. převedli bychom původní n-rozměrnou minimalizační úlohu na n jednorozměrných minimalizačních úloh, neboť $y_i \in \operatorname{argmin}_{t\in\mathbb{R}} \{\frac{1}{2}\lambda_i t^2 - d_i t\}$, tj. stačilo by nám n kroků.

[126] Už dříve jsme se zmínili, že v roce 1945 trval na tehdejších kalkulátorech výpočet inverze matice 10×10 přibližně 15 člověko-dnů. V roce 1949 to údajně bylo (z dnešního pohledu nekonečných) 8 hodin.

Bohužel nalezení takové diagonalizace je (s několika málo výjimkami) mnohem těžší problém než řešení původní úlohy (7.3.1).

Ovšem v roce 1952 publikovali Hestenes[127] a Stiefel[128] přesnou metodu pro řešení nehomogenní soustavy rovnic s pozitivně definitní maticí jakožto alternativu ke zmíněné eliminaci, viz [263]. Tato metoda nalezne řešení soustavy n rovnic o n neznámých v nejvýše n krocích a byla pojmenována jako Metoda sdružených gradientů. Nicméně tato metoda zůstala delší dobu nepovšimnuta, neboť vyžadovala $2n^3 + O(n^2)$ aritmetických operací, zatímco Gaussova eliminace pouze $n^3/3 + O(n^2)$ operací pro $n \to \infty$. Tuto metodu lze použít i v případě pozitivně semidefinitní matice, ovšem tentokrát nás nemusí vždy přivést k výsledku. Ačkoliv se jedná o přímou metodu (řešení získáme po n krocích), lze ji chápat také jako iterativní proces s velmi rychlou konvergencí v případě pozitivně definitní matice. Počátky takovéhoto výzkumu lze vystopovat do roku 1949 a u jeho počátku stál Lanczos[129], na jehož oddělení o něco málo později nastoupil i Hestenes. Výzkum pokračoval napříč padesátými a šedesátými lety a přibližně o 15 let později od zmíněného článku publikovali Fletcher[130] s Reevesem[131] tuto metodu jako nástroj pro řešení velkých nelineárních optimalizačních problémů, viz [179]. Později byla tato metoda revidována a podrobně analyzována i pro úlohu (7.3.1) s funkcí (7.3.30), což si nyní představíme. Pro mnohem více historických podrobností odkazujeme čtenáře na [226].

127



Magnus Rudolph Hestenes (13. února 1906 – 31. května 1991) byl americký matematik, který se věnoval především variačnímu počtu a teorii optimálního řízení. Patří mezi průkopníky matematické informatiky a metody sdružených gradientů.

128

Eduard L. Stiefel (21. dubna 1909 – 25. listopadu 1978) byl švýcarský matematik. Jeho jméno je spojeno především s rozvojem metody sdružených gradientů a s teorií charakteristických tříd.

Nejdříve si nastíníme základní princip této metody. Vrstevnice funkce f jsou (hyper-)elipsoidy se středem v řešení soustavy $Qx = b$. Již víme, jak funguje MNS a MSS. Avšak nejvýhodnější by bylo, kdyby se nám podařilo postupovat třeba tak, jak je zobrazeno na Obrázku 7.72, tj. ve směrech e_1 a e_2 s vhodným výpočtem délky kroku tak, abychom se po dvou krocích „trefili“ do x^*. Jenže toto by fungovalo pouze v případě, že známe bod x^*. MSG samozřejmě funguje mnohem „sofistikovaněji“ a rozhodně nevyžaduje znalost x^*. Zvolíme libovolný bod $x^{[0]}$ a minimalizujeme funkci f ve směru největšího spádu, tj. ve směru $h_0 = b - Qx^{[0]}$, který je kolmý na odpovídající vrstevnici. Tím získáme $x^{[1]}$ a z tohoto bodu pokračujeme ve směru „sdruženého vektoru“, tj. ve směru, který spojuje bod dotyku $x^{[1]}$ a střed elipsy, čímž získáme řešení ve 2 krocích! Viz Obrázek 7.73.

Klíčovým pojmem celé metody je tzv. *Q-sdruženost vektorů.*

Definice 7.3.15

Nechť matice $Q = Q^\top \in \mathbb{R}^{n\times n}$ je pozitivně definitní. Vektory $h_1, h_2 \in \mathbb{R}^n \setminus \{0\}$ se nazývají *Q-sdružené* (též *Q-ortogonální*), jestliže

$$\langle Qh_1, h_2\rangle = h_1^\top Q h_2 = 0.$$

Systém vektorů $\{h_0, \dots, h_{m-1}\}$ v $\mathbb{R}^n \setminus \{0\}$ pro $m \in \{2, \dots, n\}$ se nazývá *Q-sdružený*, jestliže

$$\langle Qh_i, h_j\rangle = 0 \quad \text{pro } i \neq j.$$

Základní vlastností Q-sdružených vektorů je jejich lineární nezávislost.

Věta 7.3.16

Nechť systém vektorů $\{h_0, \dots, h_{m-1}\}$ v $\mathbb{R}^n \setminus \{0\}$ s $m \in \{2, \dots, n\}$ je Q-sdružený. Potom jsou tyto vektory lineárně nezávislé.

129

Cornelius (Cornel) Lanczos (maďarsky Lánczos Kornél, narozen jako Kornél Lőwy; 2. února 1893 – 25. června 1974) byl maďarský, maďarsko-americký a později maďarsko-irský matematik a fyzik

Důkaz. Připusťme, že existují čísla $c_0, \dots, c_{m-1}$ splňující $c_0^2 + \cdots + c_{m-1}^2 \neq 0$ a

$$c_0 h_0 + \cdots + c_{m-1} h_{m-1} = 0.$$

Vynásobením této rovnosti vektorem $h_j^\top Q$ pro libovolné $j \in \{0, \dots, m-1\}$ dostaneme vzhledem ke Q-sdruženosti vektorů

$$0 = c_0 \langle Q h_j, h_0\rangle + \cdots + c_{m-1} \langle Q h_j, h_{m-1}\rangle = c_j \langle Q h_j, h_j\rangle.$$

Současně $Q > 0$, takže $\langle Q h_j, h_j\rangle > 0$, což vede k $c_j = 0$. Avšak index j byl libovolný, tudíž dostáváme spor s tím, že ne všechna čísla $c_0, \dots, c_{m-1}$ jsou nulová. ■

V následujícím tvrzení ukážeme, jak je užitečné (= důležité) míti takový systém Q-sdružených vektorů.

Věta 7.3.17

Nechť $m \in \{2,\dots,n\}$ a mějme systém $\{h_0,\dots,h_{m-1}\}$ Q-sdružených vektorů v $\mathbb{R}^n$. Nechť dále $x^{[0]} \in \mathbb{R}^n$ je dáno a body $x^{[1]},\dots,x^{[m]}$ jsou určeny jako

$$x^{[k+1]} = x^{[k]} + \alpha_k h_k = x^{[0]} + \sum_{i=0}^{k} \alpha_i h_i, \quad k \in \{0,\dots,m-1\}, \tag{7.3.32}$$

kde α_k jsou volena tak, že $f(x^{[k]} + \alpha_k h_k) = \min_{\alpha\in\mathbb{R}} f(x^{[k]} + \alpha h_k)$ pro $k \in \{0,\dots,m-1\}$. Pak pro kvadratickou funkci f definovanou v (7.3.30) platí

$$f(x^{[m]}) = \min_{x\in X_m} f(x),$$

kde $X_m := x^{[0]} + \operatorname{Lin}\{h_0,\dots,h_{m-1}\}$. Zejména pro $m = n$ dostáváme

$$f(x^{[n]}) = \min_{x\in\mathbb{R}^n} f(x),$$

tj. $x^{[n]}$ je řešením úlohy (7.3.1) pro kvadratickou funkci (7.3.30).

Důkaz. Z definice f a (7.3.32) plyne pro libovolné $\lambda_k \in \mathbb{R}$, že

$$\begin{aligned} f(x^{[k]} + \lambda_k h_k) &= \frac{1}{2}\big(x^{[k]} + \lambda_k h_k\big)^\top Q\big(x^{[k]} + \lambda_k h_k\big) - b^\top\big(x^{[k]} + \lambda_k h_k\big) = \\ &= f(x^{[k]}) + \lambda_k \big(Q x^{[k]} - b\big)^\top h_k + \frac{1}{2}\lambda_k^2 h_k^\top Q h_k = \\ &= f(x^{[k]}) + \lambda_k \Big[Q\Big(x^{[0]} + \sum_{i=0}^{k-1} \alpha_i h_i\Big) - b\Big]^\top h_k + \frac{1}{2}\lambda_k^2 h_k^\top Q h_k = \\ &= f(x^{[k]}) + \lambda_k \big(Q x^{[0]} - b\big)^\top h_k + \frac{1}{2}\lambda_k^2 h_k^\top Q h_k. \end{aligned} \tag{7.3.33}$$

Odtud pak dostaneme

$$\begin{aligned} f\Big(x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k\Big) &= \frac{1}{2}\Big(x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k\Big)^\top Q\Big(x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k\Big) - b^\top\Big(x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k\Big) = \\ &= \frac{1}{2} x^{[0]\top} Q x^{[0]} - b^\top x^{[0]} + \frac{1}{2} x^{[0]\top} Q\Big(\sum_{k=0}^{m-1} \lambda_k h_k\Big) + \frac{1}{2}\Big(\sum_{k=0}^{m-1} \lambda_k h_k\Big)^\top Q x^{[0]} + \\ &\quad + \frac{1}{2}\underbrace{\Big(\sum_{k=0}^{m-1} \lambda_k h_k\Big)^\top Q\Big(\sum_{k=0}^{m-1} \lambda_k h_k\Big)}_{= \sum_{k=0}^{m-1} \lambda_k \, h_k^\top Q\left(\sum_{j=0}^{m-1} \lambda_j h_j\right) = \sum_{j=0}^{m-1} \lambda_j h_k^\top Q h_j = \lambda_k h_k^\top Q h_k} - b^\top\Big(\sum_{k=0}^{m-1} \lambda_k h_k\Big) = \\ &= f(x^{[0]}) + \sum_{k=0}^{m-1} \Big[\lambda_k \big(Q x^{[0]} - b\big)^\top h_k + \frac{1}{2}\lambda_k^2 h_k^\top Q h_k\Big] = \\ &\overset{(7.3.33)}{=} f(x^{[0]}) + \sum_{k=0}^{m-1} \big[f(x^{[k]} + \lambda_k h_k) - f(x^{[k]})\big]. \end{aligned}$$

Jelikož každé $x \in X_m$ lze vyjádřit jako $x = x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k$ pro vhodná $\lambda_1,\dots,\lambda_{m-1} \in \mathbb{R}$, platí díky předchozí rovnosti a definici α_k, že

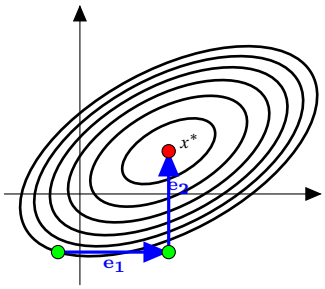
$$\begin{aligned} \min_{x\in X_m} f(x) &= \min_{(\lambda_0,\dots,\lambda_{m-1})^\top \in \mathbb{R}^m} f\Big(x^{[0]} + \sum_{k=0}^{m-1} \lambda_k h_k\Big) = \\ &= f(x^{[0]}) + \sum_{k=0}^{m-1} \Big[\min_{\lambda_k\in\mathbb{R}} f(x^{[k]} + \lambda_k h_k) - f(x^{[k]})\Big] = \\ &= f(x^{[0]}) + \sum_{k=0}^{m-1} \big[f(x^{[k]} + \alpha_k h_k) - f(x^{[k]})\big] = \\ &= f(x^{[0]} + \sum_{k=0}^{m-1} \alpha_k h_k) = f(x^{[m]}). \end{aligned}$$

(židovského původu). Svoji dizertační práci (1921) napsal o teorii relativity a kopii zaslal Albertu Einsteinovi, který mu odpověděl: *Prostudoval jsem váš článek, jak mi to jen moje současné přílišné vytížení dovolovalo. Domnívám se, že mohu říci tolik: něco takového vyžaduje kompetentní a originální intelektuální práci, na jejímž základě by měl být doktorát získat… Rád přijímám čestné věnování.* V roce 1924 objevil přesné řešení Einsteinovy rovnice gravitačního pole. V letech 1928–1929 byl Einsteinovým asistentem. V roce 1939 emigroval se svým synem do USA (jeho žena byla na cestování příliš nemocná a o pár týdnů později zemřela). Na přelomu 40. a 50. let byl však vyšetřován kvůli možnému napojení na komunistickou stranu, a tak v roce 1952 emigroval do Irska, kde na Dublin Institute for Advanced Studies následoval Erwina Schrödingera a kde zůstal až do své smrti v roce 1974.

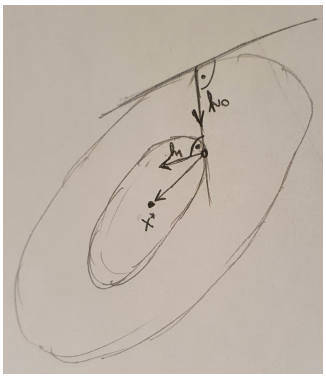
Po roce 1949 pracoval v *Národním úřadu standardů*, kde se věnoval rozvoji *Lanczosova algoritmu* jakožto iteračního procesu pro nalezení vlastních čísel velkých symetrických matic a metodě sdružených gradientů pro řešení systémů lineárních rovnic.

[130] Roger Fletcher (29. ledna 1939 – 15. června 2016) byl britský matematik a profesor na Univerzitě v Dundee. Byl členem Society for Industrial and Applied Mathematics (SIAM) a v roce 2003 byl také zvolem členem Královské společnosti v Londýně. V roce 2006 získal *Lagrangeovu cenu* udílenou SIAMem a v roce obdržel medaili od Královské společnosti v Edinburghu.

[131] C. M. Reeves (Electronic Computing Laboratory, The University, Leeds, UK) – to jsou jediné informace, které se autorovi podařilo najít.



Obrázek 7.72: „Zdokonalení“ MNS a MSS.



Obrázek 7.73: Ilustrace MSG.

Druhá část tvrzení je pouze bezprostředním důsledkem předchozí rovnosti a toho, že vektory $h_0,\dots,h_{n-1}\in\mathbb{R}^n$ jsou lineárně nezávislé podle Věty 7.3.16, tudíž $X_n=\mathbb{R}^n$. ■

Jiný slovy nám Věta 7.3.17 ukazuje, že popsaným způsobem najdeme v k-tém kroku bod, ve kterém se realizuje minimum kvadratické funkce f na k-rozměrném afinním podprostoru určeném bodem $x^{[0]}$ a vektory $h_0,\dots,h_{k-1}$. Díky přesné minimalizaci je $\frac{\mathrm{d}}{\mathrm{d}\alpha}\operatorname{grad} f(x^{[k+1]})=0$, což je ekvivalentní s

$$\operatorname{grad}^\top f(x^{[k+1]})\frac{\mathrm{d}}{\mathrm{d}\alpha}x^{[k+1]}=0 \quad \text{neboli} \quad (Qx^{[k+1]}-b)^\top h_k=0.$$

Avšak současně $b=Qx^*$, takže odtud plyne

$$h_k^\top Q(x^{[k+1]}-x^*)=0,$$

což ukazuje Q-sdruženost vektoru h_k a vektoru mířícího do řešení x^* z bodu $x^{[k+1]}$. Vzhledem k volbě funkce f lze snadno odvodit explicitní předpis pro délku k-tého kroku. Chceme totiž, aby α_k bylo řešením úlohy

$$\frac{1}{2}(x^{[k]}+\alpha h_k)^\top Q(x^{[k]}+\alpha h_k)-b^\top(x^{[k]}+\alpha h_k)\to\min,$$

což po zderivování vzhledem k α vede k rovnici $h_k^\top Q h_k+\alpha h_k^\top Q h_k-b^\top h_k=0$, takže

$$\alpha_k=\frac{h_k^\top(b-Qx^{[k]})}{h_k^\top Q h_k}=-\frac{h_k^\top\operatorname{grad} f(x^{[k]})}{h_k^\top Q h_k}, \tag{7.3.34}$$

kteréžto vyjádření bude v případě $h_k=-\operatorname{grad} f(x^{[k]})$ splývat se vyjádřením pro α_k v MNS, viz (7.3.11). Navíc s využitím rovnosti $x^*=x^{[n]}$ lze na předpis (7.3.32) pro jednotlivé členy iterační posloupnosti také nahlížet jako

$$x^{[k]}-x^*=x^{[0]}-x^*+\sum_{i=0}^{k-1}\alpha_i h_i=x^{[0]}-\Big(x^{[0]}+\sum_{i=0}^{n-1}\alpha_i h_i\Big)+\sum_{i=0}^{k-1}\alpha_i h_i=-\sum_{i=k}^{n-1}\alpha_i h_i,$$

což ukazuje, že v jednotlivých krocích dochází postupně k eliminaci jednoho „zavádějícího“ směru z počáteční aproximace $x^{[0]}$. Odtud také pro $j\in\{0,\dots,k-1\}$ vyplývá ortogonalita

$$0=\sum_{i=k}^{n-1}\alpha_i h_j^\top Q h_i=-h_j^\top Q(x^{[k]}-x^*)=-h_j^\top(Qx^{[k]}-b)=-h_j^\top\operatorname{grad} f(x^{[k]}). \tag{7.3.35}$$

K této rovnosti bychom mohli dospět také jinou cestou: víme, že pokud provedeme jeden krok ve směru h_k, již nikdy se do tohoto směru nevrátíme, tj. směr $x^{[k]}-x^*$ musí být Q-sdružený se všemi předchozími směry hledání. Jenže $\operatorname{grad} f(x^{[k]})=Qxk-b=Q(x^{[k]}-x^*)$, a tudíž je tento gradient nutně ortogonální se všemi $h_0,\dots,h_{k-1}$.

Metoda popsaná ve Větě 7.3.17 se nazývá *Metodou sdružených gradientů* (MSG). Pro její možné použití nám zbývá vhodně zvolit vektory $h_0,\dots,h_{n-1}$. Z Lineární algebry a geometrie známe Gramův–Schmidtův ortogonalizační proces, což je vlastně určení Q-sdružených vektorů pro $Q=I$. V případě libovolné matice $Q>0$ vypadá tento proces následovně. Pro systém lineárně nezávislých vektorů $u_0,\dots,u_{n-1}$ položíme

$$h_0:=u_0 \quad \text{a} \quad h_i:=u_i+\sum_{j=0}^{i-1}\beta_{ij}h_j,$$

přičemž h_1 konstruujeme tak, aby byl Q-sdružený s h_0; h_2 tak, aby byl Q-sdružený s h_0 a h_1 atd. Jinými slovy, v i-tém kroku vezmeme u_i a odečteme ty „složky“, které nejsou Q-sdružené k $h_0,\dots,h_{i-1}$, tj. chceme β_{ij} takové, že rovnost

$$h_i^\top Q h_j=u_i^\top Q h_j+\sum_{k=0}^{i-1}\beta_{ik}h_k^\top Q h_j$$

se pro $i > j$ redukuje na

$$0 = h_i^\top Q\, h_j = u_i^\top Q\, h_j + \beta_{ij}\, h_j^\top Q\, h_j,$$

což nás přivádí k volbě

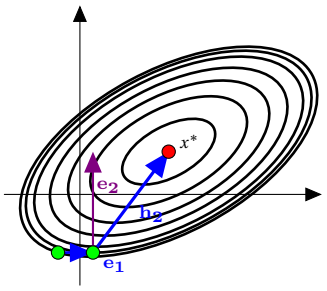
$$\beta_{ij} = -\frac{u_i^\top Q h_j}{h_j^\top Q h_j}, \quad i > j.$$

Nevýhodou tohoto procesu je fakt, že si musíme „pamatovat“ všechny předchozí vektory pro konstrukci dalšího a celkově potřebujeme řádově $O(n^3)$ aritmetických operací. V případě výchozí volby systému vektorů jakožto $e_1, \dots, e_n$ budou konjugované vektory odpovídat tomu, co bychom získali Gaussovou eliminací, viz [475] a Obrázek 7.74. Takže potřebujeme najít výhodnější volbu. My vezmeme vektory $u_k := -\operatorname{grad} f(x^{[k]})$ pro $k \in \{0, \dots, n-1\}$, čímž dostaneme

$$h_0 := -\operatorname{grad} f(x^{[0]}), \quad h_k := -\operatorname{grad} f(x^{[k]}) + \beta_{k-1}\, h_{k-1}, \tag{7.3.36}$$

$$\beta_{k-1} := \frac{\operatorname{grad}^\top f(x^{[k]})\, Q\, h_{k-1}}{h_{k-1}^\top Q\, h_{k-1}}, \tag{7.3.37}$$

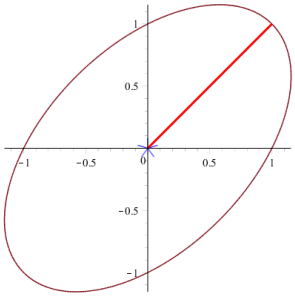
přičemž body $x^{[k]}$ jsou počítány dle (7.3.32). Vlastně volíme $\beta_{ij} = \beta_{k,k-1}$, což zaručuje, že vektory h_{k-1} a h_k jsou Q-sdružené. Je to ale skutečně dobrá volba? Na to nám odpoví následující věta. Viz také Obrázky 7.75–7.76.



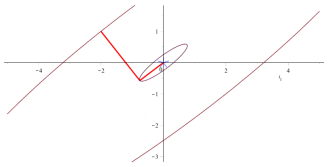
Obrázek 7.74: Ilustrace MSG v případě volby výchozích lineárně nezávislých vektorů jako e_1, e_2 a jejich Q-sdružením. Srovnej s Obrázkem 7.72.

Věta 7.3.18

Nechť $x^{[0]} \in \mathbb{R}^n$ je libovolný a $x^{[1]}, \dots, x^{[n-1]}$, $h_0, \dots, h_{n-1}$ jsou postupně určeny vztahy (7.3.32), (7.3.34), (7.3.36) a (7.3.37). Jsou-li všechny vektory $\operatorname{grad} f(x^{[0]}), \dots, \operatorname{grad} f(x^{[n-1]})$ nenulové[132], potom systém vektorů $\{h_0, \dots, h_{n-1}\}$ je Q-sdružený a libovolná dvojice vektoru $\operatorname{grad} f(x^{[0]}), \dots, \operatorname{grad} f(x^{[n-1]})$ je ortogonální.



Obrázek 7.75: Jeden krok MSG pro funkci $f(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2$ s $x^{[0]} = [1, 1]$.



Obrázek 7.76: Dva kroky MSG pro funkci $f(x_1, x_2) = 3x_1^2 - 7x_1 x_2 + 5x_2^2$ s $x^{[0]} = [-2, 1]$.

Důkaz. Z (7.3.35) víme, že $\operatorname{grad}^\top f(x^{[k+1]})\, h_k$, z[133] čehož pro $k \in \{0, \dots, n-1\}$ s využitím značení $g_k := \operatorname{grad} f(x^{[k]}) = Q x^{[k]} - b$ dostáváme

$$\begin{aligned}
\langle \operatorname{grad} f(x^{[k+1]}), \operatorname{grad} f(x^{[k]}) \rangle &= \langle Q x^{[k+1]} - b, g_k \rangle = \langle Q x^{[k]} + \alpha_k Q h_k - b, g_k \rangle = \\
&= \langle g_k, g_k \rangle + \alpha_k \langle Q h_k, g_k \rangle \overset{(7.3.34)}{=} \langle g_k, g_k \rangle - \frac{\langle h_k, g_k \rangle}{\langle h_k, Q h_k \rangle} \langle Q h_k, g_k \rangle = \\
&\overset{(7.3.36)}{=} \|g_k\|^2 - \frac{\overbrace{\langle g_k, -g_k + \beta_{k-1} h_{k-1} \rangle}^{\langle g_k, \beta_{k-1} h_{k-1} \rangle = 0}}{\underbrace{\langle Q h_k, -g_k + \beta_{k-1} h_{k-1} \rangle}_{\langle Q h_k, \beta_{k-1} h_{k-1} \rangle = 0}} \langle Q h_k, g_k \rangle = \\
&\overset{(7.3.37)}{=} \|g_k\|^2 - \frac{\langle g_k, -g_k \rangle}{\langle Q h_k, -g_k \rangle} \langle Q h_k, g_k \rangle = \|g_k\|^2 - \|g_k\|^2 = 0.
\end{aligned}$$

Nyní ukážeme, že vektory $h_0, \dots, h_{n-1}$ jsou Q-sdružené a také ortogonalitu vektorů $g_0, \dots, g_{n-1}$. Důkaz provedeme indukcí. Vektory h_0 a h_1 jsou Q-sdružené díky volbě (7.3.37). Vektory g_0 a g_1 jsou ortogonální podle předchozího výpočtu. Připusťme, že toto platí pro vektory $h_0, \dots, h_i$ a $g_0, \dots, g_i$ pro nějaké $i \in \{1, \dots, k-1\}$ s $2 \le k \le n-1$. Ukážeme, že totéž je splněno i pro $i = k$. Ortogonalitu g_{k-1} a g_k máme z předchozího výpočtu, zatímco pro $j = 0, \dots, k-2$ vyplývá z

$$\begin{aligned}
\langle g_k, g_j \rangle &= \langle \overbrace{g_{k-1} + \alpha_{k-1} Q h_{k-1}}^{g_k = Q x^{[k]} - b = Q(x^{[k-1]} + \alpha_{k-1} h_{k-1}) - b}, g_j \rangle = \alpha_{k-1} \langle Q h_{k-1}, g_j \rangle = \\
&\overset{(7.3.36)}{=} \alpha_{k-1} \langle Q h_{k-1}, -h_j + \beta_{j-1} h_{j-1} \rangle = 0
\end{aligned}$$

díky indukčnímu předpokladu pro $h_0, \dots, h_i$, přičemž bereme $\beta_{-1} := 0$. Současně platí $h_k \neq 0$, neboť v opačném případě

$$0 = -g_k + \beta_{k-1}\Big(-g_{k-1} + \beta_{k-2}\Big(-g_{k-2} + \beta_{k-3}\big(-g_{k-3} + \dots\big)\Big)\Big),$$

což znamená, že vektory $g_0, \dots, g_k$ jsou lineárně závislé, a to je spor s jejich právě dokázanou ortogonalitou.

[132] Z důkazu by mělo být patrné, co se stane v případě, že $\operatorname{grad} f(x^{[i]}) = 0$ pro nějaké $i \in \{0, \dots, n-1\}$.

[133] Tato rovnost plyne také přímo ze skutečnosti, že v opačném případě by pro směrovou derivaci platilo $f_{h_k}(x^{[k+1]}) = \operatorname{grad}^\top f(x^{[k+1]})\, h_k \neq 0$. V takovém případě by ale funkce f nemohla mít v bodě $x^{[k]}$ extrém ve směru h_k, což je spor ↯, neboť tímto extrémem je $x^{[k+1]}$.

Dále platí

$$\alpha_j = -\frac{\langle g_j, h_j\rangle}{\langle Q h_j, h_j\rangle} \overset{(7.3.36)}{=} -\frac{\langle g_j, -g_j + \beta_{j-1} h_{j-1}\rangle}{\langle Q h_j, h_j\rangle} = \frac{\|g_j\|^2}{\langle Q h_j, h_j\rangle}. \tag{7.3.38}$$

Tudíž $\alpha_j = 0$ právě tehdy, když $g_j = 0$, v kterémžto případě $x^{[j]} = x^*$ je bodem minima funkce f a výpočet vzhledem k její ostré/silné konvexnosti končí. Bude-li $\alpha_j \neq 0$, pak z rovnosti $g_{j+1} = g_j + \alpha_j Q h_j$ dostaneme

$$Q h_j = \frac{1}{\alpha_j}(g_{j+1} - g_j), \tag{7.3.39}$$

což pro $j \in \{0, \ldots, k-2\}$ dává

$$\langle Q h_j, h_k\rangle = \langle Q h_j, -g_k + \beta_{k-1} h_{k-1}\rangle = -\langle Q h_j, g_k\rangle = -\frac{1}{\alpha_j}\langle g_{j+1} - g_j, g_k\rangle = 0$$

díky indukčním předpokladům. Rovnost $\langle Q h_{k-1}, h_k\rangle = 0$ podle volby β_{k-1} v (7.3.36), takže vektory $h_0, \ldots, h_k$ jsou vskutku Q-sdružené. Proto v případě $g_i \neq 0$ pro $i \in \{0, \ldots, n-1\}$ jsou vektory $h_0, \ldots, h_{n-1}$ Q-sdružené a důkaz je hotov. ■

Z předchozího důkazu je mimo jiné patrné, že

$$g_i^\top g_{j+1} = g_i^\top g_j + \alpha_j g_i^\top Q h_j \quad \text{neboli} \quad g_i^\top Q h_j = \begin{cases} \frac{1}{\alpha_{i-1}} g_i^\top g_i, & i = j+1, \\ -\frac{1}{\alpha_i} g_i^\top g_i, & i = j, \\ 0, & i \neq j,\ i \neq j+1, \end{cases}$$

což znamená, že

$$\beta_{k-1} = \frac{g_k^\top Q h_{k-1}}{h_{k-1}^\top Q h_{k-1}} = \frac{1}{\alpha_{k-1}} \frac{g_k^\top g_k}{h_{k-1}^\top Q h_{k-1}} \overset{(7.3.38)}{=} \frac{h_{k-1}^\top Q h_{k-1}}{g_{k-1}^\top g_{k-1}} \frac{g_k^\top g_k}{h_{k-1}^\top Q h_{k-1}} = \frac{g_k^\top g_k}{g_{k-1}^\top g_{k-1}}.$$

Proto místo vzorců (7.3.34), (7.3.36) a (7.3.37) můžeme $x^{[k+1]}$ počítat pomocí

$$\alpha_k = \frac{g_k^\top g_k}{h_k^\top Q h_k} \quad \& \quad \beta_{k-1} = \frac{g_k^\top g_k}{g_{k-1}^\top g_{k-1}} \quad \& \quad h_0 = -g_0 \quad \& \quad h_k = -g_k + \beta_{k-1} h_{k-1}. \tag{7.3.40}$$

Algoritmický popis Metody sdružený gradientů

Zadání: Funkce f, výchozí bod $x^{[0]} \in \mathbb{R}^n$, ukončovací pravidlo (např. $\|x^{[k+1]} - x^{[k]}\| \leq \varepsilon$) a volba tolerance $\varepsilon > 0$.

Krok 1: (*Inicializace*) Položíme $k := 0$.

Krok 2: (*První krok*) Vypočteme $h_k := -\operatorname{grad} f(x^{[k]})$ a α_k takové, že $f(x^{[k]} + \alpha_k h_k) = \min_{\alpha \in \mathbb{R}} f(x^{[k]} + \alpha h_k)$. Pokračujeme Krokem 4.

Krok 3: (*Iterace*) Vypočteme $\beta_{k-1} := \frac{g_k^\top g_k}{g_{k-1}^\top g_{k-1}}$, $h_k := -g_k + \beta_{k-1} h_{k-1}$ a α_k takové, že $f(x^{[k]} + \alpha_k h_k) = \min_{\alpha \in \mathbb{R}} f(x^{[k]} + \alpha h_k)$.

Krok 4: (*Ověření ukončovacího pravidla*) Položíme $x^{[k+1]} := x^{[k]} + \alpha_k h_k$. Pokud je splněno ukončovací pravidlo KONEC.
V opačném případě položíme $k := k+1$ a pokračujeme Krokem 3.

Příklad 7.3.19

Pomocí MSG určeme první iteraci $x^{[1]}$ a směr h_1 další iterace pro řešení úlohy

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$$

při počáteční aproximaci $x^{[0]} = [0, 1]$.

Řešení. První krok je (skoro) stejný jako v případě MNS, tj. určíme $h_0 = -\operatorname{grad} f(x^{[0]}) = -\binom{1}{2}$, takže $x^{[1]} = x^{[0]} + \alpha\, h_0 = [-\alpha, 1-2\alpha]$, což po dosazení do funkce f vede k úloze

$$7\alpha^2 - 5\alpha + 1 \to \min$$

s řešením $\alpha_0 = 5/14$. Proto $x^{[1]} = [-5/14, 2/7]$. Ve druhé kroku nejdříve určíme $\operatorname{grad} f(x^{[1]}) = \binom{-3/7}{3/14}$ a následně podle (7.3.37) získáme

$$\beta_0 = \frac{\operatorname{grad}^\top f(x^{[1]})\, Q\, h_0}{h_0^\top Q\, h_0} = \frac{9}{14^2},$$

přičemž tentýž výsledek bychom dostali i pomocí (7.3.40). Proto

$$h_1 = -\operatorname{grad} f(x^{[1]}) + \beta_0\, h_0 = \frac{15}{14^2}\begin{pmatrix} 5 \\ -4 \end{pmatrix},$$

a tudíž

$$x^{[2]} = x^{[1]} + \alpha\, h_1 = [-5/14 + 75\alpha/14^2, 2/7 - 60\alpha/14^2].$$

Po dosazení tohoto bodu do funkce f dostaneme

$$\frac{675}{5488}\alpha^2 - \frac{45}{196}\alpha + 3/28 \to \min$$

s řešením $\alpha_1 = 14/15$, takže $x^{[2]} = [0,0]$. V souladu s předchozí teorií by měl být bod $x^{[2]}$ řešením dané úlohy. O tom, že tomu tak je, se můžeme snadno přesvědčit např. tím, že $\operatorname{grad} f(x^{[0]}) = \binom{0}{0}$.

Poznámka 7.3.20. V MNS je možné při řešení příkladů „zkracovat“ vektory h_k, čímž si zjednodušíme výpočty při hledání α_k (projeví se to pouze úměrnou změnou α_k). Toto je možné také pro MSG, ale pouze při použití vzorců (7.3.34) a (7.3.36). *Při použití* (7.3.40) *to již fungovat nebude!* Nicméně pomocí obou přístupů musí vyjít tytéž hodnoty $x^{[1]}, x^{[2]}, \ldots$ Např. pro funkci

$$f(x_1, x_2) = 9x_1^2 - 4x_1 x_2 + x_2^2 + 12x_1 - 4x_2 + 4$$

s $x^{[0]} = [0,0]$ vyjde v MSG $x^{[1]} = [-30/47, 10/47]$ a ve druhém kroku

$$\beta_0 = \frac{16}{2209} \quad \text{a} \quad h_1 = \begin{pmatrix} 560/2209 \\ 2320/2209 \end{pmatrix},$$

takže pro výpočet α_1 musíme dosadit do funkce f bod

$$[x_1, x_2] = [-30/47 + 560\alpha/2209, 10/47 + 2320\alpha/2209],$$

čímž následně dostaneme $\alpha_1 = 47/50$. Avšak mnohem jednodušší je pracovat s vektorem $\tilde{h}_1 = (7,\ 29)^\top$, který nás přivede k minimalizaci funkce f v bodě

$$[x_1, x_2] = [-30/47 + 7\alpha, 10/47 + 29\alpha].$$

Každopádně v obou případech dostaneme s použitím správných vzorců $x^{[2]} = [-2/5, 6/5]$. ▲

MSG úzce souvisí s tzv. metodami *Krylovových podprostorů*, které jsou pro dané h_0 a $k \in \mathbb{N}$ (tzv. řád) definovány jako

$$\mathcal{K}(h_0, r) := \operatorname{Lin}\{h_0, Q\,h_0, Q^2 h_0, \ldots, Q^k h_0\}.$$

Tyto metody pocházejí od Hestenese, Stiefela a Lanczose a byly zařazena mezi 10 algoritmů, které nejvíce ovlivnili rozvoj a užití vědy a inženýrství ve 20. století, viz [139]. O tomto propojení se přesvědčíme v následující větě.

Věta 7.3.21

Pokud $x^{[k]}$ získané MSG není řešením úlohy (7.3.1) s kvadratickou funkcí (7.3.30), pak se značením $g_j := \operatorname{grad} f(x^j)$ pro $j = 0,\dots,k$ máme

$$\operatorname{Lin}\{g_0, g_1, \dots, g_k\} = \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^k g_0\} \tag{7.3.41}$$

a

$$\operatorname{Lin}\{h_0, h_1, \dots, h_k\} = \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^k g_0\} \tag{7.3.42}$$

Důkaz. Důkaz provedeme indukcí. Pro $k = 0$ je tvrzení zřejmé. Připusťme, že je pravdivé i pro nějaké k, a ukážeme, že pak bude splněno i pro $k+1$. Z indukčního předpokladu očividně máme

$$g_k \in \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^k g_0\} \quad \text{a} \quad h_k \in \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^k g_0\},$$

přičemž z druhé inkluze plyne

$$Q\,h_k \in \operatorname{Lin}\{Q\,g_0, Q^2\,g_0, \dots, Q^{k+1} g_0\}.$$

Jelikož

$$g_{k+1} = Q\,x^{[k+1]} - b = Q\,(x^{[k]} + \alpha\,h_k) - b = (Q\,x^{[k]} - b) + \alpha\,Q\,h_k = g_k + \alpha_k\,Q\,h_k,$$

plyne odtud

$$g_{k+1} \in \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^{k+1} g_0\},$$

což ukazuje, že

$$\operatorname{Lin}\{g_0, g_1, \dots, g_{k+1}\} \subseteq \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^{k+1} g_0\}.$$

Pro získání opačné inkluzi si všimněme, že z indukčního předpokladu pro (7.3.42) máme

$$Q^{k+1}\,g_0 = Q\,(Q^k\,g_0) \in \operatorname{Lin}\{Q\,h_0, Q\,h_1, \dots, Q\,h_k\}.$$

Zároveň je $Q\,h_i = \frac{1}{\alpha_i}\,(g_{i+1} - g_i)$ pro všechna $i = 0,\dots,k$, takže

$$Q^{k+1}\,g_0 \in \operatorname{Lin}\{g_0, g_1, \dots, g_{k+1}\}.$$

Toto společně s indukčním předpokladem pro (7.3.41) dává požadovanou opačnou inkluzi

$$\operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^{k+1} g_0\} \subseteq \operatorname{Lin}\{g_0, g_1, \dots, g_{k+1}\},$$

čímž je důkaz rovnosti (7.3.41).

Nyní z definice h_k, indukčního předpokladu a s využitím předchozí rovnosti dostaneme

$$\begin{aligned}\operatorname{Lin}\{h_0, h_1, \dots, h_{k+1}\} &= \operatorname{Lin}\{h_0, h_1, \dots, h_k, g_{k+1}\} = \operatorname{Lin}\{g_0, Q\,g_0, \dots, Q^k\,g_0, g_{k+1}\} = \\ &= \operatorname{Lin}\{g_0, g_1, \dots, g_k, g_{k+1}\} = \operatorname{Lin}\{g_0, Q\,g_O, \dots, Q^{k+1}\,g_0\},\end{aligned}$$

čímž je důkaz hotov. ∎

Odtud vyplývá

$$x^{[k+1]} = x^{[0]} + \alpha_0\,h_0 + \cdots + \alpha_k\,h_k = x^{[0]} + \gamma_0\,Q\,g_0 + \cdots + \gamma_k\,Q^k\,g_0 = x^{[0]} + P_k(Q)\,g_0 \tag{7.3.43}$$

pro vhodná čísla $\gamma_0,\dots,\gamma_k$ a $P_k(Q)$ značící odpovídající polynom stupně k. Každý takový polynom $P_k(Q)$ určuje posloupnost $\{x^{[k]}\}$ a platí

$$\begin{aligned}x^{[k+1]} - x^* &= x^{[0]} - x^* + P_k(Q)\,(Q\,x^{[0]} - b) = x^{[0]} - x^* + P_k(Q)\,Q\,(x^{[0]} - x^*) = \\ &= [I + Q\,P_k(Q)]\,(x^{[0]} - x^*),\end{aligned}$$

z čehož pro funkci[134] $E(x) := \frac{1}{2}(x - x^*)^\top Q\,(x - x^*)$ máme

[134] Tutéž funkci jsme využili i v MNS, viz Lemma 7.3.3.

$$E(x^{[k+1]}) = \frac{1}{2}\big(x^{[0]} - x^*\big)^\top Q\,\big[I + Q\,P_k(Q)\big]^2\,\big(x^{[0]} - x^*\big).$$

Nyní bychom se mohli ptát, který polynom $P_k(Q)$ minimalizuje hodnotu $E(x^{[k+1]})$? Ale naše předchozí výsledky ukazují, že to je právě ten polynom, který získáme pomocí MSG. Samozřejmě určení jednotlivých koeficientů $\gamma_0,\dots,\gamma_k$ pomocí $\alpha_0,\dots,\alpha_k$ a $\beta_0,\dots,\beta_{k-1}$ je poněkud komplikované a totéž platí i o jakémsi vztahu mezi polynomem P_k a polynomem P_{k+1}. Do takových podrobností se zde ale pouštět nebudeme. Síla MSG spočívá v tom, že ve svém průběhu postupně řeší takovéto jednotlivé optimalizační polynomické problémy, přičemž se aktualizuje pouze malé množství informací. Tato naše pozorování můžeme shrnout do následující důležité věty.

VĚTA 7.3.22

Bod $x^{[k+1]}$ získaný pomocí MSG splňuje

$$E(x^{[k+1]}) = \min_{P_k} \frac{1}{2}\big(x^{[0]} - x^*\big)^\top Q\,\big[I + Q\,P_k(Q)\big]^2\,\big(x^{[0]} - x^*\big),$$

kde minimum je uvažováno vzhledem ke všem polynomům P_k stupně k.

Jakého zlepšení s každým krokem docílíme? V odpovědi na tuto otázku hrají klíčovou roli vlastní čísla matice Q. Nechť tedy $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ jsou vlastní čísla matice Q a $v_1,\dots,v_n$ odpovídající ortonormální vlastní vektory, takže

$$Q = \sum_{i=1}^{n} \lambda_i\, v_i\, v_i^\top.$$

Jelikož tyto vlastní vektory generují celé $\mathbb{R}^n$, můžeme psát

$$x^{[0]} - x^* = \sum_{i=1}^{n} \xi_i\, v_i$$

pro vhodná čísla $\xi_1,\dots,\xi_n$. Každý vlastní vektor matice Q je také vlastním vektorem matice $P_k(Q)$ pro libovolný polynom P_k. Pro naši matici Q máme

$$P_k(Q)\,v_i = P_k(\lambda_i)\,v_i \quad \text{pro všechna } i = 1,\dots,n,$$

proto pro polynom P_k odpovídající MSG dostaneme

$$x^{[k+1]} - x^* = \sum_{i=1}^{n} \big[1 + \lambda_i\,P_k(\lambda_i)\big]\,\xi_i\,v_i,$$

což nás přivede k rovnosti[135]

$$2\,E(x^{[k+1]}) = \|\,x^{[k+1]} - x^*\,\|_Q^2 = \sum_{i=1}^{n} \lambda_i\,\big[1 + \lambda_i\,P_k(\lambda_i)\big]^2\,\xi_i^2.$$

Vzhledem k optimalitě polynomu P_k získaného pomocí MSG plyne odtud

$$\|\,x^{[k+1]} - x^*\,\|_Q^2 = \min_{P_k} \sum_{i=1}^{n} \lambda_i\,\big[1 + \lambda_i\,P_k(\lambda_i)\big]^2\,\xi_i^2.$$

Vytknutím $\big[1 + \lambda_i\,P_k(\lambda_i)\big]^2$ z tohoto výrazu získáme

$$\begin{aligned}\|\,x^{[k+1]} - x^*\,\|_Q^2 &= \min_{P_k} \max_{1\le i\le n} \big[1 + \lambda_i\,P_k(\lambda_i)\big]^2 \sum_{j=1}^{n} \lambda_j\,\xi_j^2 = \\ &= \min_{P_k} \max_{1\le i\le n} \big[1 + \lambda_i\,P_k(\lambda_i)\big]^2\,\|\,x^{[0]} - x^*\,\|_Q^2, \end{aligned} \tag{7.3.44}$$

neboť $\|\,x^{[0]} - x^*\,\|_Q^2 = \sum_{j=1}^{n} \lambda_j\,\xi_j^2$. Proto platí následující tvrzení, které ukazuje, že každý krok MSG je nejméně tak dobrý jako v případě MNS, neboť pro ni je $x^{[k+1]} = x^{[k]} - \alpha_k\,g_k$, což lze také vyjádřit jako v (7.3.43) pro vhodný polynom P_k.

[135] Pro libovolný vektor z platí $\|\,z\,\|_Q^2 = z^\top Q\,z = \sum_{i=1}^{n} \lambda_i\,(v_i^\top z)^2.$

VĚTA 7.3.23

Pro MSG platí

$$E(x^{[k+1]}) \le \max_{1\le i\le n} \left[1+\lambda_i P_k(\lambda_i)\right]^2 E(x^{[0]})$$

pro libovolný polynom P_k stupně k, kde maximum je bráno vzhledem ke všem vlastním číslům matice Q.

V některých speciálních případech lze explicitně určit polynom P_k a odvodit z něj některé zajímavé vlastnosti MSG jako v následujícím případě, viz také Větu 7.3.28 níže.

VĚTA 7.3.24

Má-li matice Q pouze r různých vlastních čísel, pak MSG nalezne řešení v nejvýše r krocích.

Důkaz. Nechť vlastní čísla $0<\lambda_1\le\lambda_2\le\cdots\le\lambda_n$ matice Q nabývají pouze r různých hodnot $0<\tau_1<\tau_2<\cdots<\tau_r$. Definujme polynom

$$Q_r(\lambda)=\frac{(-1)^r}{\tau_1\cdots\tau_r}(\lambda-\tau_1)\cdots(\lambda-\tau_r).$$

Pak $Q_r(\lambda_i)=0$ pro $i=1,2,\dots,n$ a $Q_r(0)=1$. Proto polynom $Q_r(\lambda)-1$ je polynom stupně r s kořenem $\lambda=0$, můžeme definovat polynom stupně $r-1$ jako

$$R_{r-1}(\lambda):=\frac{Q_r(\lambda)-1}{\lambda}.$$

Pak

$$0\le\min_{P_{r-1}}\max_{1\le i\le n}\left[1+\lambda_i P_{r-1}(\lambda_i)\right]^2\le\max_{1\le i\le n}\left[1+\lambda_i Q_{r-1}(\lambda_i)\right]^2=\max_{1\le i\le n}Q_r^2(\lambda_i)=0,$$

takže na pravé straně (7.3.44) je nulový pro $k=r-1$, a tedy $\|x^{[r]}-x^*\|_Q=0$ neboli $x^{[r]}=x^*$. ■

S podobným zdůvodněním jako lze najít v [360, str. 275], můžeme odvodit i následující tvrzení, viz také nerovnost (7.3.13) pro MNS.

VĚTA 7.3.25

Má-li matice Q vlastní čísla $0<\lambda_1\le\lambda_2\le\cdots\le\lambda_n$, pak pro MSG platí

$$\left\|x^{[k+1]}-x^*\right\|_Q^2\le\left(\frac{\lambda_{n-k}-\lambda_1}{\lambda_{n-k}+\lambda_1}\right)^2\left\|x^{[0]}-x^*\right\|_Q^2. \tag{7.3.45}$$

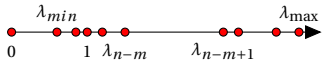
Druhou část můžeme využít pro předpověď chování MSG aplikované na některé specifické problémy. Uvažme situaci, že matice Q má m velkých vlastních hodnot a zbývajících $n-m$ menších „nahromaděných“ kolem 1, viz Obrázek 7.77. Položíme-li $\varepsilon:=\lambda_{n-m}-\lambda_{\min}$, potom po $m+1$ krocích budeme podle (7.3.45) mít[136]

$$\|x^{[m+1]}-x^*\|_Q^2\approx\varepsilon\|x^{[0]}-x^*\|_Q^2,$$

tj. pro velmi malé ε můžeme říci, že $x^{[m+1]}$ získané po pouhých $m+1$ krocích MSG je velmi dobrou aproximací x^*. Odhad (7.3.45) také můžeme vyjádřit pomocí čísla podmíněnosti matice Q, tj. čísla $\kappa(Q):=\lambda_{\max}/\lambda_{\min}$, které dává

$$\|x^{[k]}-x^*\|_Q\le 2\left(\frac{\sqrt{\kappa(Q)}-1}{\sqrt{\kappa(Q)}+1}\right)^k\left\|x^{[0]}-x^*\right\|_Q.$$

[136] Odtud vidíme, že klíčovou roli hraje rozdíl $\lambda_{n-m}-\lambda_{\min}$, tj. mohli bychom mít $\lambda_{\min},\dots,\lambda_{n-m}\in[a,b]$ pro blízká a,b, zatímco $\lambda_{n-m+1},\dots,\lambda_{\max}$ mohou být „libovolná“.



Obrázek 7.77: Ukázka rozložení vlastních čísle matice Q.

Poznámka 7.3.26 (Numerické vlastnosti MSG).

(i) Z teoretického pohledu musí MSG pro úlohu (7.3.1) s funkcí (7.3.30) najít řešení v nejvýše n krocích (dle Věty 7.3.17). V praxi (při numerickém výpočtu) však dochází k zaokrouhlovacím chybám a může se stát, že toto řešení nenalezneme. Nalezneme pouze nějakou jeho aproximaci, přičemž „kvalita“ této aproximace závisí právě na zaokrouhlovacích chybách, a tedy zejména na podmíněnosti matice Q vyjádřené podílem $\lambda_{\max}/\lambda_{\min}$.

V praktických implementacích se proto MSG používá v cyklech délky n, tj. po n krocích se celý proces restartuje a za výchozí bod nového cyklu se bere $x^{[n]}$, tj.

$$\beta_{k-1} := \begin{cases} \beta_{k-1}, & k \neq n, 2n, 3n, \dots, \\ 0, & k = n, 2n, \dots. \end{cases}$$

Volba délky cyklu jako n pochází od Fletchera a Reevese jako přirozená volba s ohledem na teoreticky nejvyšší počet kroků potřebných pro nalezení řešení. Je ale samozřejmě možné zvolit i jinou délku cyklu. V numerických výpočtech je také užitečné zvolit nějaké ukončovací pravidlo, zejména např. $g_k^\top g_k = \|\operatorname{grad} f(x^{[k]})\| < \varepsilon$ pro nějaké $\varepsilon > 0$.

(ii) Hlavní výhodou MSG je její *snadná implementace* (podobně jako pro MNS). Početní náročnost jednoho kroku je[137]

$$O(n) + [\text{počet aritmetických operací pro násobení } h \rightsquigarrow Qh].$$

To znamená, že pro *řídkou* matici Q (tj. počet nenulových prvků $N \ll n^2$) bude počet operací malý[138]. V praxi se mnohdy jedná o systémy, které mají velmi řídké matice (počet nenulových prvků je asi 0,01 %–1 %) nebo jsou součinem dvou řídkých matic (což je početně stejně náročné jako kdybychom měli jednu řídkou matici). Je-li řád matice opravdu velký[139] a není zde žádná „řídkost“ (explicitní či implicitní), tak v podstatě není způsob, jak soustavu $Qx = b$ vyřešit. V řídkém případě lze za dobrých okolností použít Gaussovu eliminaci (nebo např. Choleského rozklad) s vhodnou modifikací pro zachování řídkosti. Jsou-li okolnosti nepříznivé nelze přímé metody z lineární algebry použít a nezbývá nám nic jiného než skutečně použít nějaký iterační proces, jako např. MNS nebo MSG. Nicméně stále je *aritmetická náročnost* relativně malá (závislost je polynomiální) – totéž platí i pro MNS, ale MSG má v jistém smyslu *nejlepší řád konvergence* mezi všemi iteračními procesy založenými na maticovém násobení. Nevýhodou MSG je již zmíněná *citlivost na podmíněnost* matice Q a zaokrouhlování. ▲

[137] Podle (7.3.32)–(7.3.37) potřebujeme 2 maticová násobení ($h_k \rightsquigarrow Qh_k$ pro určení λ_k a $x^{[k]} \rightsquigarrow Qx^{[k]}$ pro určení g_k), ovšem pouze první z nich je nutné, neboť g_k lze zpětně určit z $g_k = g_{k-1} + \alpha_{k-1} Qh_{k-1}$. Všechny ostatní operace jsou „jednoduché“ — výpočet skalárního součinu a lineárních kombinací, které společně řádově „stojí“ $O(n)$ aritmetických operací.

[138] Pro *hustou* matici je počet aritmetických operací při násobení matice–vektor roven $(2n-1)n$, což v případě řídké matice bude menší než $2N \ll (2n-1)n$.

[139] Desítky tisíc, např. v tomografii se jedná o matice velikosti 10^5 až 10^6.

Pamatujete si na PARTAN popsanou v závěru Podkapitoly 7.3.1? Její dvouskoková varianta je pro kvadratické funkce ekvivalentní s MSG!

Věta 7.3.27 PARTAN vs MSG pro kvadratické funkce

Metoda PARTAN je pro kvadratické funkce ekvivalentní s MSG.

Důkaz. Důkaz provedeme pomocí indukce. Pro první krok je tvrzení očividně pravdivé, neboť se v obou případech jedná o MNS. Připusťme, že body $x^{[1]}, x^{[2]}, \dots, x^{[k]}$ byly získány pomocí MSG a bod $x^{[k+1]}$ je vypočten dle PARTAN. Chceme ukázat, že je to tentýž bod, který bychom získali pomocí MSG. Má-li to být pravda, pak tento bod musí minimalizovat funkci f na nadrovině určené h_{k-1} a $g_k := \operatorname{grad} f(x^{[k]})$. Zároveň z popsané MSG víme, že tento bod bude minimalizovat funkci f na afinním podprostoru určeném $x^{[0]}$, vektory $h_0, \dots, h_{k-1}$ a g_k. Jinými slovy, musí se jednat o bod x, pro který je $\operatorname{grad} f(x)$ ortogonální na g_k i na h_{k-1}. Jelikož $y^{[k]}$ v PARTAN minimalizuje f podél g_k, je $\operatorname{grad} f(y^{[k]})$ kolmý na g_k. Současně je $\operatorname{grad} f(x^{[k-1]})$ v podprostoru $\operatorname{Lin}\{h_0, \dots, h_{k-1}\}$ a z (7.3.35) víme, že g_k je ortogonální na všechny tyto vektory (tj. i na celý podprostor), musí být také $\operatorname{grad} f(x^{[k-1]})$ kolmý na g_k. Jelikož $\operatorname{grad} f(x)$ je lineární vzhledem k x, plyne odtud, že v každém bodě x na přímce určené body $x^{[k-1]}$ a $y^{[k]}$ máme $\operatorname{grad} f(x)$ kolmý na g_k. Minimalizováním f podél této přímky získáme bod $x^{[k+1]}$, pro který je $\operatorname{grad} f(x^{[k+1]})$ kolmý na tuto přímku. Proto $\operatorname{grad} f(x^{[k+1]})$ je současně kolmý na g_k i na přímku spojující $x^{[k-1]}$ a $y^{[k]}$, a tedy je kolmý na celou jimi určenu nadrovinu. ■

MSG lze aplikovat také v případě kvadratické funkce s $Q \geq 0$, avšak tentokrát nemáme zaručenu existenci řešení.

Věta 7.3.28

Nechť matice $Q \geq 0$ s hodností $r := \operatorname{rank} Q \leq n$. Pak MSG buď zajistí nalezení minima funkce f dané jako v (7.3.30) v nejvýše r krocích (v takovém případě pro nějaké $k \in$

Věta 7.3.28

$\{0,\dots,r\}$ nastane $\operatorname{grad} f(x^{[k]}) = 0$ a $h_0,\dots,h_{k-1} \neq 0$) nebo signalizuje, že toto minimum neexistuje (v takovém případě $h_k^\top Q h_k = 0$, přičemž $\operatorname{grad} f(x^{[0]}),\dots,\operatorname{grad} f(x^{[k-1]}) \neq 0$ pro nějaké $k \in \{0,\dots,r\}$).

Důkaz. Nejdříve předpokládejme, že odpovídající kvadratická funkce 7.3.30 nabývá svého minima na $\mathbb{R}^n$ a nechť $y \in \mathbb{R}^n$ je takové, že $Qy = 0$. Pak $\langle b, y\rangle = 0$, protože v opačném případě by $f(\alpha y) = -\alpha\langle b, y\rangle \to -\infty$ pro $\alpha \to \pm\infty$ dle znaménka $\langle b, y\rangle$, což je spor. Proto skutečně $\langle b, y\rangle = 0$. Dále pro libovolné $x \in \mathbb{R}^n$ máme

$$\langle \operatorname{grad} f(x), Qx - b\rangle y = \langle Qx, y\rangle = 0.$$

Nepochybně dimenze $\operatorname{Ker} Q$ je $n - r$ a předchozí rovnost ukazuje, že $\operatorname{grad} f(x)$ je pro každé $x \in \mathbb{R}^n$ prvkem ortogonálního doplňku $(\operatorname{Ker} Q)^\perp$ a dle definice $h_0,\dots,h_k$ tamtéž patř také tyto vektory, tj. zejména platí $\langle Qh_k, h_k\rangle = 0$ právě tehdy, když h_k. Ukážeme, že $\operatorname{grad} f(x^{[j]}) = 0$ pro nějaké $j \in \{1,\dots,r\}$. Nechť $h_j \neq 0$, a tedy $\langle Qh_j, h_j\rangle > 0$ pro $j \in \{1,\dots,r-1\}$. I v tomto případě můžeme podobně jako ve Větě 7.3.18 ukázat ortogonalitu vektorů $\operatorname{grad} f(x^{[0]}),\dots,\operatorname{grad} f(x^{[r]})$. Avšak současně musí být tyto vektory lineárně závislé jakožto prvky r dimenzionální podprostoru $\operatorname{Ker} Q$, tudíž alespoň jeden z nich musí být nulový. V odpovídajícím bodě nabývá funkce f svého minima, neboť je konvexní.

Nyní připusťme, že funkce f není zdola ohraničená na $\mathbb{R}^n$. V tomto případě musí nastat rovnost $\langle Qh_k, h_k\rangle = 0$ pro nějaké $k \in \{0,\dots,n-1\}$, což znemožní výpočet konstanty β_k, a výpočet tudíž nemůže pokračovat. Vskutku, kdyby tomu tak nebylo, tj. $\langle Qh_k, h_k\rangle > 0$, pak podobně jako v případě pozitivně definitní matice Q by byly vektory $h_0,\dots,h_{n-1}$ navzájem Q-sdružené, a tedy lineárně nezávislé. Pak by libovolné $x \in \mathbb{R}^n \setminus \{0\}$ bylo možné vyjádřit jako lineární kombinaci $x = \sum_{i=0}^{n-1} \xi_i h_i$, přičemž alespoň jedno z čísel $\xi_0,\dots,\xi_{n-1}$ je nenulové. Pak

$$\langle Qx, x\rangle = \sum i = 0^{n-1} \xi_i^2 \langle Qh_i, h_i\rangle > 0,$$

což ukazuje, že matice Q je pozitivně definitní. V takovém případě je soustava $\operatorname{grad} f(x) = Qx - b = 0$ jednoznačně řešitelná s řešením x^* odpovídajícím minimu funkce f na $\mathbb{R}^n$. To je však spor, a důkaz je hotov. ∎

MSG a nekvadratické funkce

MSG lze využít také pro nekvadratické funkce. Algoritmus funguje stejně jako pro kvadratické funkce pouze se liší volbou β_k, přičemž obvykle dochází k restartu MSG po n krocích. Existuje několik přístupů, zejména (Fletcher–Reeves) β_{k-1}^{FR} stejný jako v (7.3.40) pro $g_k := \operatorname{grad} f(x^{[k]})$ nebo (Polak–Ribiére)

$$\beta_{k-1}^{PR} := \frac{(g_k - g_{k-1})^\top g_k}{g_{k-1}^\top g_{k-1}}$$

nebo (Hestens–Stiefel)

$$\beta_{k-1}^{HS\#1} := \frac{(g_k - g_{k-1})^\top g_k}{h_{k-1}^\top (g_k - g_{k-1})}, \quad \beta_{k-1}^{HS\#2} := \frac{(g_k - g_{k-1})^\top g_k}{g_{k-1}^\top (g_k - g_{k-1})}.$$

Dokonce lze najít strategii

$$\beta_k^{FR-PR} = \begin{cases} -\beta_k^{FR}, & \beta_k^{PR} < -\beta_k^{FR}, \\ \beta_k^{PR}, & |\beta_k^{PR}| \le \beta_k^{FR}, \\ \beta_k^{FR}, & \beta_k^{PR} > \beta_k^{FR}. \end{cases}$$

Obecně platí: různé volby β_k jsou výhodné v různých situacích. Např. je známo, že volba β_k^{PR} dává mnohem lepší výsledky než β_k^{FR} v některých případech – ale již ne v jiných. Existují dokonce případy, kdy g_k jsou odraženy od nuly v případech volby β_k^{PR}, viz [419], ve kterém je na základě analýzy globální konvergence MSG upřednostňována volba β_k^{FR}. Autor téhož článku ale také navrhuje volbu

$$\beta_k = \max\{0, \beta_k^{PR}\}.$$

Následující dvě tvrzení dávají informace o „oblasti“ lokální konvergence a o rychlosti konvergence posloupnosti výsledků MSG s cykly délky[140] n.

Věta 7.3.29 Lokální konvergence MSG ve variantě FR nebo PR

Nechť $f \in C^1$ na $\mathbb{R}^n$ a $x^{[0]} \in \mathbb{R}^n$ je takové, že množina

$$S := \{x \in \mathbb{R}^n \mid f(x) \le f(x^{[0]})\}$$

je ohraničená. Nechť $x^{[1]}$ je výsledek MSG s β^{FR} nebo β^{PR} po n-krocích (tj. cyklu délky n), $x^{[2]}$ výsledek dalšího cyklu s výchozím bodem $x^{[1]}$ atd. Potom posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ je ohraničená a její hromadné body jsou stacionárními body funkce f, tj. $\lim_{j\to\infty} \operatorname{grad} f(x^{[j]}) = 0$ pro každou podposloupnost $\{x^{[j]}\} \subseteq \{x^{[k]}\}$.

Důkaz. Důkaz je v podstatě analogický jako v případě MNS a Věty 7.3.7, a tak můžeme naznačit jen pár detailů[141]. První skutečností je, že hodnota účelové funkce v průběhu výpočtu neroste, jelikož všechna kroky jsou založeny na přesné minimalizaci. Tudíž body $\{x^{[k]}\}$ nikdy neopustí množinu S. Klíčovým faktem je, že první krok k-tého cyklu je shodný s MNS z bodu $x^{[k]}$, a tedy dojde k „podstatnému“ snížení funkční hodnoty, pokud $\operatorname{grad} f(x^{[k]})$ není moc malý[142]. Jelikož následné „mezikroky“ nepovedou ke zvýšení funkční hodnoty, můžeme říci, že součet snížení funkčních hodnot v počátečních krocích splývajících s MNS je shora ohraničený (počáteční odchylkou). V důsledku toho klesá zlepšování funkčních hodnot k nule pro $k \to \infty$, takže malé snížení funkční hodnoty v úvodním kroku MNS z bodu $x^{[k]}$ znamená malou $\|\operatorname{grad} f(x^{[k]})\|$. Proto $\operatorname{grad} f(x^{[k]}) \to 0$ pro $k \to \infty$, a tedy jakýkoli hromadný bod posloupnosti $\{x^{[k]}\}$ je stacionárním bodem funkce f. ■

Naší původní motivací bylo nalezení metody, která je lepší než MNS (rychlejší konvergence), ale méně náročná než NM (funkce z C^2, výpočet inverzní matice atd.). Následující věta ukazuje[143], že výsledky cyklů délky n z MSG pro nekvadratické funkce konvergují k x^* alespoň kvadraticky, tj. jeden cyklus MSG délky n udělá totéž jako jeden krok NM. No, není to skvělé?! Je to skvělé! Viz také Obrázky 7.78–7.84.

Věta 7.3.30 Asymptotická „n-kroková kvadratická“ konvergence MSG pro MSG ve variantě FR nebo PR pro nedegenerovaný nekvadratický problém

Nechť $f \in C^3$ na $\mathbb{R}^n$, $x^{[0]} \in \mathbb{R}^n$ a x^* je nedegenerované lokální minimum, tj. $\operatorname{grad} f(x^*) = 0$ a $\nabla^2 f(x^*) > 0$. Nechť $x^{[k]}$ je výsledek MSG s cyklem délky n a výchozím bode $x^{[k-1]}$ a nechť $x^{[k]} \to x^*$ pro $k \to \infty$. Potom posloupnost $\{x^{[k]}\}$ konverguje k x^* superlineárně s řádem (alespoň) $p = 2$ (tj. kvadraticky), tj.

$$\| x^{[k+1]} - x^* \| \le C \| x^{[k]} - x^* \|^2$$

pro nějaké $C < \infty$ a všechna $k \in \mathbb{N} \cup \{0\}$.

Cvičení

7.3.1. Při dané počáteční aproximaci $x^{[0]}$ vypočtěte alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané pomocí MNS při řešení úlohy

$$f(x) \to \min, \quad x \in \mathbb{R}^n,$$

jestliže

a) $x^{[0]} = [0,3]$ & $f(x) = f(x_1, x_2) = 2x_1^2 + 4x_2^2 - 2x_1 - 8x_2$; $(x^* = [1/2, 1])$

b) $x^{[0]} = [2,3]$ & $f(x) = f(x_1, x_2) = 4x_1^2 - 4x_1 x_2 + 2x_2^2$; $(x^* = [0,0])$

c) $x^{[0]} = [1,0]$ & $f(x) = f(x_1, x_2) = 2x_1^2 - x_1 x_2 + x_2^2$; $(x^* = [1,0])$

d) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1, x_2) = x_1^2 + 2x_1 x_2 + 2x_2^2 + 2x_1$; $(x^* = [-2,1])$

[140] Tj. nikoli o rychlosti konvergence samotné MSG!

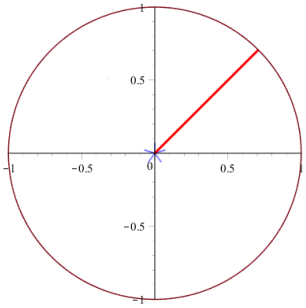
[141] Viz [51, Proposition 7.2.2].

[142] Vše se děje na kompaktní množině S, kde je f spojitě diferencovatelná, takže pro libovolné $x \in S$ máme

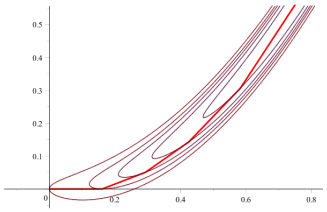
$$f(x+h) \le f(x) + h^\top \operatorname{grad} f(x) + \\ + \omega(\|h\|)\,\|h\|$$

pro libovolné h s $\|h\| \le 1$, přičemž zbytek je nezávislý na $x \in S$ a splňuje $\omega(s) \to 0$ pro $s \to 0$. Tudíž vhodně zvolený krok z bodu $x \in S$ ve směru $-\operatorname{grad} f(x)$ skutečně sníží funkční hodnotu nejméně o $\psi(\|\operatorname{grad} f(x)\|)$ pro nějakou kladnou a neklesající funkcí $\psi(s)$ na polopřímce $s > 0$.

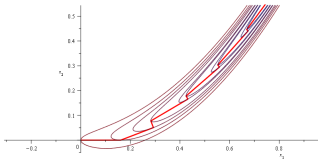
[143] Věta je převzata z [51, Proposition 7.2.3], kde je uvedena bez důkazu.



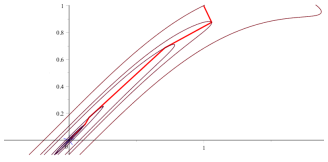
Obrázek 7.78: Několik prvních iterací MSG pro funkci $f(x_1,x_2) = (x_1^2+x_2^2)^{3/2}$ s $x^{[0]} = [1/\sqrt{2}, 1/\sqrt{2}]$.



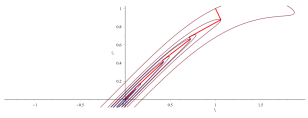
Obrázek 7.79: Několik prvních iterací MSG pro Rosenbrockovu funkci $f(x_1,x_2) = (1-x_1)^2+100(x_2-x_1^2)^2$ s $x^{[0]} = [0,0]$ a bez resetu β.



Obrázek 7.80: Několik prvních iterací MSG pro Rosenbrockovu funkci $f(x_1,x_2) = (1-x_1)^2+100(x_2-x_1^2)^2$ s $x^{[0]} = [0,0]$ a resetem β.



Obrázek 7.81: Několik prvních iterací MSG pro funkci $f(x_1,x_2) = 10(x_2-\sin x_1)^2+x_1^2/10$ s $x^{[0]} = [1,1]$ a bez resetu β.



Obrázek 7.82: Několik prvních iterací MSG pro funkci $f(x_1,x_2) = 10(x_2-\sin x_1)^2+x_1^2/10$ s $x^{[0]} = [1,1]$ a resetem β.

e) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1,x_2) = 4x_1^2+4x_1x_2+2x_2^2-3x_1$; ($x^* = [3/4,-3/4]$)

f) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = x_1-x_2+2x_1^2+2x_1x_2+x_2^2$; ($x^* = [-1,3/2]$)

g) $x^{[0]} = [0,-1]$ & $f(x) = f(x_1,x_2) = (2x_1-x_2-2)^2+x_1x_2+3x_1$; ($x^* = [-2/7,-17/7]$)

h) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = (2x_1-x_2+1)^2+x_1x_2$; ($x^* = [-2/7,4/7]$)

i) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = (3x_1-x_2+2)^2+2x_1x_2$; ($x^* = [-2/5,6/5]$)

j) $x^{[0]} = [2,1]$ & $f(x) = f(x_1,x_2) = (x_1-x_2-1)^2+x_1x_2-x_1$; ($x^* = [4/3,-1/3]$)

k) $x^{[0]} = [1,0]$ & $f(x) = f(x_1,x_2) = (x_1-2x_2-1)^2+x_1x_2+2x_2$; ($x^* = [-2/7,-6/7]$)

l) $x^{[0]} = [4,3]$ & $f(x) = f(x_1,x_2) = (x_1-x_2-2)^2+x_1x_2+x_1$; ($x^* = [2/3,-5/3]$)

m) $x^{[0]} = [1,-2]$ & $f(x) = f(x_1,x_2) = x_1^2+6x_1x_2+9x_2^2$; (nekonečně mnoho řešení tvaru $x_1 = -3x_2$)

n) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = x_1^2+6x_1x_2+9x_2^2+2x_1-x_2$; (úloha nemá řešení)

o) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1,x_2) = (x_1-x_2)^2-x_1+x_2$; (nekonečně mnoho řešení tvaru $x_1 = x_2+1/2$)

p) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1,x_2) = (x_1-x_2)^2+x_1+x_2$; (úloha nemá řešení)

q) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1,x_2) = (2x_1-x_2+1)^2-4x_1+2x_2$; (nekonečně mnoho řešení tvaru $x_2 = 2x_1$)

r) $x^{[0]} = [1,-1]$ & $f(x) = f(x_1,x_2) = (x_1-3x_2+1)^2+2x_1-x_2$; (úloha nemá řešení)

s) $x^{[0]} = [1,-1,0]$ & $f(x) = f(x_1,x_2,x_3) = x_1^2+2x_1x_2+2x_2^2+x_3^2$; ($x^* = [0,0,0]$)

t) $x^{[0]} = [0,-1,1]$ & $f(x) = f(x_1,x_2,x_3) = (2x_1-x_2-1)^2+(x_3-x_2-2)^2+3x_2x_3$; ($x^* = [-3/2,-4,4]$)

u) $x^{[0]} = [0,0,0]$ & $f(x) = f(x_1,x_2,x_3) = (x_1+x_2+1)^2/2+(1-x_3)^2+(x_1-x_3-1)^2$; ($x^* = [2,-3,1]$)

v) $x^{[0]} = [0,0,0]$ & $f(x) = f(x_1,x_2,x_3) = (x_3-2x_2+2)^2/2+(1-x_3)^2+(x_2-3x_1)^2/4$; ($x^* = [1/2,3/2,1]$)

w) $x^{[0]} = [1,0,0]$ & $f(x) = f(x_1,x_2,x_3) = (x_1+x_3+1)^2+(x_1-x_2-1)^2+x_2x_3$; (úloha nemá řešení)

7.3.2. S využitím dané výchozí počáteční aproximace $x^{[0]}$ nalezněte pomocí MSG přesné řešení úloh z Příkladu 7.3.1 (Pozor: v některých příkladech mohou vycházet „komplikovanější“ zlomky, např. v 7.3.1c).

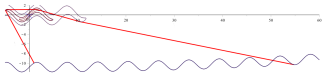
7.3.3. S využitím dané výchozí počáteční aproximace $x^{[0]}$ nalezněte pomocí NM přesné řešení úloh z Příkladu 7.3.1c+o+s+w.

7.3.4. Při dané počáteční aproximaci $x^{[0]}$ vypočtěte alespoň alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané pomocí NM při řešení úlohy

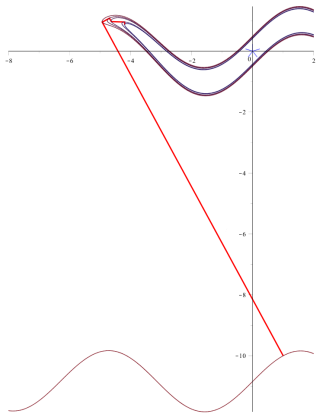
$$f(x) \to \min, \quad x \in \mathbb{R}^n,$$

jestliže

a) $x^{[0]} = [4,2,-1]$ & $f(x) = f(x_1,x_2,x_3) = (x_1-4)^4+x_2^2+\frac{(x_3+5)^2}{4}-3$; ($x^* = [4,0,-5]$)



Obrázek 7.83: Několik prvních iterací MSG pro funkci $f(x_1,x_2) = 10(x_2 - \sin x_1)^2 + x_1^2/10$ s $x^{[0]} = [1,-10]$ a bez resetu β.



Obrázek 7.84: Několik prvních iterací MSG pro funkci $f(x_1,x_2) = 10(x_2 - \sin x_1)^2 + x_1^2/10$ s $x^{[0]} = [1,-10]$ a resetem β.

b) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = 10(x_2 - x_1^2)^2 + (1-x_1)^2$; $(x^* = [1,1])$

c) $x^{[0]} = [1,1]$ & $f(x) = f(x_1,x_2) = \mathrm{e}^{x_1^2+x_2^2}$; $(x^* = [0,0])$

d) $x^{[0]} = [1/2,0]$ a $x^{[0]} = [1/2,1/4]$ & $f(x) = f(x_1,x_2) = \ln(1 + x_1^2 + x_2^2)$; $(x^* = [0,0])$

e) $x^{[0]} = [1,1]$ & $f(x) = f(x_1,x_2) = \mathrm{e}^{1+x_1^2x_2^2}$; $(x^* = [0,0])$

f) $x^{[0]} = [1/4,1/4]$ & $f(x) = f(x_1,x_2) = -\ln(1-x_1-x_2) - \ln x_1 - \ln x_2$; (lokální řešení $x^* = [1/3,1/3]$)

g) $x^{[0]} = [1,1]$ & $f(x) = f(x_1,x_2) = -\mathrm{e}^{2-x_1^2-x_2^2+x_1x_2}$; $(x^* = [0,0])$

h) $x^{[0]} = [0,0]$ & $f(x) = f(x_1,x_2) = (x_1 - x_2 - 1)^4 + \ln(x_1^2 + 2x_2^2 + 1)$; (lokální řešení $x^* \approx [0{,}32125, -0{,}160625]$)

i) $x^{[0]} = [0,1]$ & $f(x) = f(x_1,x_2) = (3x_1 - x_1x_2 + x_2)^4 + \mathrm{e}^{\frac{2x_1^2}{x_2^2+1}}$ $(x^* = [0,0])$

7.3.5. S využitím dané výchozí počáteční aproximace $x^{[0]}$ vypočtěte alespoň alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané pomocí MNS při řešení úloh z Příkladu 7.3.4.

7.3.6. S využitím dané výchozí počáteční aproximace $x^{[0]}$ vypočtěte alespoň alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané pomocí MSG při řešení úloh z Příkladu 7.3.4.

7.3.7. Zvolte si (libovolně – zkusit můžete i různé volby) počáteční aproximaci $x^{[0]}$ a určete alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané metodou největšího spádu a metodou sdružených gradientů při řešení úlohy

$$f(x) \to \min, \quad x \in \mathbb{R}^n,$$

jestliže (x^* značí přesné řešení)

a) libovolná funkce ze zadání Příkladu 7.3.1

b) Matyasova funkce $f(x) = f(x_1,x_2) = \frac{13}{50}(x_1^2 + x_2^2) - \frac{12}{25}x_1x_2$, $(x^* = [0,0])$,

c) $f(x) = f(x_1,x_2) = 1000x_1^2 + 40x_1x_2 + x_2^2$, $(x^* = [0,0])$,

d) $f(x) = f(x_1,x_2) = x_1^2 + 2x_2^2 - 2x_1 - 8x_2$, $(x^* = [1/2,1])$,

e) $f(x) = f(x_1,x_2) = 6x_1^2 + 4x_1x_2 + 10x_2^2 - 11x_1 - 13x_2$, $(x^* = [3/4,1/2])$,

f) Boothova funkce $f(x) = f(x_1,x_2) = (x_1 + 7x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$, $(x^* = [1,3])$,

g) $f(x) = f(x_1,x_2,x_3) = 2x_1^2 - 2x_1x_2 + x_2^2 + x_2x_3 + x_3^2 + 2x_1 + 3x_2 - 4x_3$, $(x^* = [-13/2,-12,8])$,

h) $f(x) = f(x_1,x_2,x_3) = x_1^2 + x_1x_3 + 3x_2^2 + x_2x_3 + 2x_1x_2 + 2x_3^2 + x_1 - 3x_2 - x_3$, $(x^* = [-13/2,-12,8])$,

i) Powellova funkce $f(x) = f(x_1,x_2,x_3,x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^2 + 10(x_1 - x_4)^2$, $(x^* = [0,0,0,0])$,

j) $f(x) = f(x_1,x_2,x_3,x_4) = 3x_1^2 + x_1x_2 + 4x_2^2 + x_2x_3 + 3x_2x_4 + 10x_3^2 + 3x_4^2 + 6x_1 - 3x_2 - x_3 - 5x_4$, $(x^* = [28/39,-27/26,3/13,1/26])$,

7.3.8. Zvolte si (libovolně – zkusit můžete i různé volby) počáteční aproximaci $x^{[0]}$ a spočtěte alespoň dva následující členy v minimalizující posloupnosti $\{x^{[k]}\}$, tj. $x^{[1]}$ a $x^{[2]}$, získané pomocí MNS, NM nebo MSG při řešení úlohy

$$f(x) \to \min, \quad x \in \mathbb{R}^n,$$

jestliže (x^* značí přesné řešení)

a) $f(x) = f(x_1, x_2) = \left(x_1^2 + x_2^2\right)^{3/2}$, ($x^* = [0,0]$),

b) $f(x) = f(x_1, x_2) = 100x_1^4 + \frac{x_2^4}{100}$, ($x^* = [0,0]$),

c) $f(x) = f(x_1, x_2) = \frac{x_1^4}{4} + \frac{x_2^2}{2} - x_1 x_2 + x_1 - x_2$, ($x^* = [-1,0]$ a $x^* = [1,2]$),

d) $f(x) = f(x_1, x_2) = \sqrt{x_1^2 + 1} + \sqrt{x_2^2 + 1}$, ($x^* = [0,0]$),

e) $f(x) = f(x_1, x_2) = -\ln(1 - x_1 - x_2) - \ln x_1 - \ln x_2$, ($x^* = [1/3, 1/3]$),

f) $f(x) = f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2$, ($x^* = [1,0]$),

g) Rosenbrockova funkce $f(x) = f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$, ($x^* = [1,1]$),

h) Baelova funkce ($x^* = [3, 1/2]$)

$$f(x) = f(x_1, x_2) = \left(\frac{3}{2} - x_1 + x_1 x_2\right)^2 + \left(\frac{9}{4} - x_1 + x_1 x_2^2\right)^2 + \left(\frac{453}{200} - x_1 + x_1 x_2^3\right)^2,$$

i) Goldsteinova–Priceho funkce ($x^* = [0,-1]$)

$$f(x) = f(x_1, x_2) = \Big(1 + \left(x_1 + x_2 + 1\right)^2 \left(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1 x_2 + 3x_2^2\right)\Big) \times \\ \times \Big(30 + \left(2x_1 - 3x_2\right)^2 \left(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1 x_2 + 27x_2^2\right)\Big),$$

j) „trojhrbá funkce“ $f(x) = f(x_1, x_2) = 2x_1^2 - \frac{21x_1^4}{20} + \frac{x_1^6}{6} + x_1 x_2 + x_2^2$, ($x^* = [0,0]$),

k) Easomova funkce $f(x_1, x_2) = -\cos x_1 \cos x_2 \exp\left[-(x_1 - \pi)^2 - (x_2 - \pi)^2\right]$, ($x^* = [\pi, \pi]$),

l) nelineární funkce

$$f(x) = f(x_1, x_2, x_3) = \frac{1}{1 + (x_1 - x_2)^2} + \sin(\pi x_2 x_3 / 2) + \exp\left[-\left(\frac{x_1 + x_3}{x_2} - 2\right)^2\right],$$

($x^* = [1,1,1]$), kde $\exp[z] = \mathrm{e}^z$,

m) $f(x) = f(x_1, x_2) = 10(x_2 - \sin x_1)^2 + \frac{x_1^2}{10}$, ($x^* = [0,0]$),

n) $f(x) = f(x_1, x_2) = 10(x_2 - \cos x_1)^2 + \frac{x_1^2}{10}$, ($x^* = [0,1]$),

o) McCormickova funkce $f(x) = f(x_1, x_2) = \sin(x_1 + x_2) + (x_1 - x_2)^2 - \frac{3x_1}{2} + \frac{5y}{2} + 1$, ($x^* \approx [-0.547198, -1.5472]$),

p) Levyho funkce ($x^* = [1,1]$)

$$f(x) = f(x_1, x_2) = \sin^2(3\pi x_1) + (x_1 - 1)^2 \left[1 + \sin^2(3\pi x_2)\right] + (x_2 - 1)^2 \left[1 + \sin^2(2\pi x_2)\right],$$

q) Rosenbrockova funkce#2 ($x^* = [1,1,1]$)

$$f(x) = f(x_1, x_2, x_3) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 100(x_3 - x_2^2)^2 + (1 - x_2)^2,$$

r) Fletcherova/Powellova funkce

$$f(x) = f(x_1, x_2, x_3) = 100\left\{\left[x_3 - 10\varphi(x_1, x_2)\right]^2 + \left(\sqrt{x_1^2 + x_2^2} - 1\right)^2\right\} + x_3^2,$$

($x^* = [1,0,0]$), kde

$$2\pi\varphi(x_1, x_2) = \begin{cases} \operatorname{arctg}\frac{x_2}{x_1}, & x_1 > 0, \\ \pi + \operatorname{arctg}\frac{x_2}{x_1}, & x_1 < 0, \end{cases}$$

s) Rosenbrockova funkce#3 ($x^* = [1,1,1,1]$ a v bodě $[-1,1,1,1]$ je lokální minimum)

$$f(x) = f(x_1, x_2, x_3, x_4) = \sum_{i=1}^{3} 100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2.$$

V případě, že není možné explicitně určit přesné řešení jednorozměrných minimalizačních úloh pro stanovení délky kroku α, aproximujte hodnotu α pomocí některé z metod jednorozměrné minimalizace.

Pojmenované funkce se různě používají jako tzv. testovací funkce pro optimalizační algoritmy, viz např. https://ln.math.muni.cz/pz/opt/funkce1 *nebo* https://ln.math.muni.cz/pz/opt/funkce2.

Kapitola 8

Matematické programování

8.1 *Obecná optimalizační úloha* 619

8.2 *Nutné a postačující podmínky optimality v matematickém programování* 626

8.3 *Teorie (Lagrangeovy) duality* 651

8.4 *Analýza citlivosti* 674

IN THE SUMMER OF 1950 AT THE SECOND BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, HELD IN BERKELEY, CALIFORNIA, A MATHEMATICIAN FROM PRINCETON, ALBERT W. TUCKER, WHO WAS GENERALLY KNOWN AS A TOPOLOGIST, GAVE A TALK WITH THE TITLE „NONLINEAR PROGRAMMING“. IT WAS BASED ON A JOINT WORK OF TUCKER AND A YOUNG MATHEMATICIAN, HAROLD W. KUHN, WHO HAD JUST FINISHED HIS PH.D. STUDY AT PRINCETON UNIVERSITY. THE TALKS WERE PUBLISHED IN A CONFERENCE PROCEEDINGS, AND FOR THE FIRST TIME THE NAME „NONLINEAR PROGRAMMING“ — THE TITLE KUHN AND TUCKER CHOSE FOR THEIR PAPER — APPEARED IN THE MATHEMATICAL LITERATURE. IN THE PAPER KUHN AND TUCKER INTRODUCED A NONLINEAR PROGRAMMING PROBLEM AND PROVED THE MAIN THEOREM OF THE THEORY — THE SO-CALLED „KUHN–TUCKER THEOREM“. THIS THEOREM, WHICH GIVES NECESSARY CONDITIONS FOR THE EXISTENCE OF AN OPTIMAL SOLUTION TO A NONLINEAR PROGRAMMING PROBLEM, LAUNCHED THE MATHEMATICAL THEORY OF NONLINEAR PROGRAMMING.[1]

TINNE HOFF KJELDSEN v úvodu svého článku [322]

[1] *V létě roku 1950 vystoupil na Druhém berkeleyském sympoziu o matematické statistice a pravděpodobnosti, které se konalo v Berkeley (Kalifornie, USA), matematik z Princetonu, Albert W. Tucker, jenž byl obecně známý jako topolog, s přednáškou s názvem „Nelineární programování“. Byla založena na Tuckerově společné práci s mladým matematikem Haroldem W. Kuhnem, který právě získal svůj doktorát na Princetonské univerzitě. Přednášky byly publikovány v konferenčním sborníku, a tak se poprvé v matematické literatuře objevilo označení „nelineární programování“ — název, který Kuhn a Tucker zvolili pro svůj příspěvek. V článku Kuhn a Tucker představili problém nelineárního programování a dokázali hlavní výsledek celé teorie — takzvanou „Kuhnovu–Tuckerovu větu“. Tato věta, která dává nutné podmínky pro existenci optimálního řešení úlohy nelineárního programování, odstartovala budování matematické teorie nelineárního programování.*

Nyní se již konečně dostáváme k obecným *úlohám matematického programování* (a jejich řešení), tj. hlavním objektem této kapitoly je úloha

$$f(x) \to \min, \quad x \in X, \tag{8.1}$$

kde *přípustná množina* X je zadána systémem rovností a nerovností

$$X := \left\{ x \in P \subseteq \mathbb{R}^n \mid g_i(x) \le 0,\ g_j(x) = 0,\ i = 1,\dots,k,\ j = k+1,\dots,m \right\}, \tag{8.2}$$

tj.

$$X = P \bigcap \left\{ x \in \mathbb{R}^n \mid g_i(x) \le 0,\ i = 1,\dots,k \right\} \bigcap \left\{ x \in \mathbb{R}^n \mid g_j(x) = 0,\ j = k+1,\dots,m \right\},$$

přičemž funkce $f, g_1, \dots, g_m$ nemusí být nutně diferencovatelné. Omezení $x \in P \neq \emptyset$ se nazývá *přímé*[2] a omezení určená funkcemi $g_1, \dots, g_m$ se nazývají *funkcionální*. Bez újmy na obecnosti lze místo úlohy se smíšenými funkcionálními omezeními uvažovat úlohy s množinou X vymezenou výhradně rovnostmi/nerovnostmi:

(i) každou rovnost $h(x) = 0$ lze snadno nahradit dvojicí nerovností $h(x) \le 0$ a $-h(x) \le 0$;

(ii) každou nerovnost $h(x) \le 0$ lze s využitím pomocné proměnné z převést na rovnost, tj. $h(x) + z^2 = 0$ nebo případně $h(x) + z = 0$ s doplněním omezení na znaménko $z \ge 0$. Druhá možnost je využívána zejména v lineárním či kvadratickém programování, kdy chceme zachovat „lineární charakter“ jednotlivých omezení a nezápornost všech proměnných. Každopádně pomocná proměnná se neobjevuje v účelové funkci, což je zcela přirozené — reprezentují-li původní proměnné např. spotřebované množství nějakých komodit, pak omezení udávají jejich dostupné množství a pomocné proměnné udávají jejich nevyužité množství (deficitní vs. přebytkové komodity).

V první části se krátce podíváme na některé obecné vlastnosti úlohy (8.1) a ukážeme si několik historických i praktických úloh matematického programování. Aby ale naše pojednání nebylo až příliš abstraktní[3], vrátíme se po chvíli k úloze (8.1) s množinou X popsanou v (8.2) a odvodíme nutné (*Lagrangeův princip* a *Johnovy/Karushovy–Kuhnovy–Tuckerovy podmínky*) a postačující (konvexnost, podmínky 2. řádu) podmínky k tomu, aby bod $x^* \in X$ byl lokálním (příp. globálním) řešením této úlohy. Následně se zaměříme na přidruženou *duální úlohu* a nakonec prozkoumáme *závislost* optimálního řešení *na* různých *parametrech* (reprezentující např. prodejní cenu či stav zásob ve skladu). Pochopitelně v celé kapitole hraje význačnou (a zásadně zjednodušující) roli pojem konvexnosti. Proto jsme si museli nejdříve vybudovat potřebný aparát v Kapitole 6. Ovšem tentokrát na rozdíl od Kapitol 2–4 je naším cílem především vybudování robustního teoretického aparátu pro obecné řešení úloh (8.1) s množinou X popsanou v (8.2) a možné aplikace budou spíše na okraji našeho zájmu.

S využitím nástrojů matematického programování lze vyřešit různé *historické*[4] *optimalizační úlohy*, o kterých jsme se zmínili již v úvodní kapitole, tj. Heronova, Tartagliova či Keplerova úloha na straně 1.3 nebo Steinerův problém popsaný v Příkladu 1.2.9 a případně jeho zobecnění v podobě Weberova problému[5]. Samozřejmě v dnešní době jsou důležité jiné úlohy, zejména z oblasti ekonomie, inženýrství atd. Kromě již dříve zmíněných úloh lineárního (např. optimální výrobní program, směšovací úloha, dopravní problém), celočíselného[6] (např. problém obchodního cestujícího) či kvadratického programovaní (např. teorie portfolia) můžeme uvést několik dalších ukázek.

Příklad 8.1

Ekonomie a její příbuzné obory obsahují řadu rozhodovacích procesů, které můžeme pojmout jako optimalizační úlohy. Avšak oproti našemu dosavadnímu pojetí v nich jde obvykle o maximalizaci (zisku, užitku atd.), takže v nich hrají klíčovou roli konkávní funkce na konvexní množině. Např. základní a současně typický model v *teorii spotřebitele* zahrnuje jednoho spotřebitele, který spotřebovává n různých komodit v nezáporném množství. Označíme-li množství i-komodity jako x_i pro $i = 1, \dots, n$, pak

[2] Typicky bývá P rovno $\mathbb{R}^n$, $\mathbb{R}^n_+$ nebo $\mathbb{R}^n_{++}$.

[3] A vskutku by mohlo být, neboť s využitím striktně geometrického pojetí by se dala budovat jakási analogická teorie i pro mnohem abstraktnější úlohy. I když ani my se jistému geometrickému pojetí nevyhneme. Avšak pozor – nelze hovořit o *geometrickém programování*, neboť toto označení je vyhrazeno pro úlohy matematického programování, kde účelová funkce a nerovnostní omezení jsou dány pomocí *posynomiálů*, což jsou součty *monomiálů* neboli funkcí ve tvaru $a x_1^{a_1} \cdots x_n^{a_n}$, viz např. [34, Podkapitola 11.5].

[4] Tj. převážně úlohy geometrického charakteru s ad hoc řešením.

[5] Za pozornost v tomto ohledu jistě stojí i tzv. *Malfattiho problém* z článku z roku 1803, ve kterém jde o nalezení tří nepřekrývajících se kružnic uvnitř trojúhelníku tak, aby se maximalizoval součet jejich obsahů. Sám Malfatti se domníval že zná správné řešení: nakreslit kruhy tak, aby se každý dotýkal dvou stran trojúhelníku a obou zbývajících kruhů (jedná se o tzv. *Malfattiho kruhy*, jejichž konstrukci popsal v roce 1826 Steiner). Více než sto let se věřilo, že tato ekvivalence mezi konstrukcí Malfattiho kruhů a řešením Malfattiho problému je pravdivá. Avšak v roce 1930 ukázali Lob a Richmond, že pro některé trojúhelníky toto neplatí a že větší obsah lze získat, pokud nejdříve vepíšeme kružnici s největším možným poloměrem, následně vepíšeme kružnici s největším možným poloměrem v některém ze tří vytvořených rohů a konečně vepíšeme kružnici s největším možným poloměrem do některé z pěti částí (to je v podstatě tzv. *hladový algoritmus*, viz později Poznámku 8.4.17). V případě rovno-

stranného trojúhelníku je rozdíl v obsazích nepatrný, cca 1 %, avšak v roce 1946 ukázal Eves, že u rovnoramenných trojúhelníků dlouhého a úzkého tvaru je obsah získaný Malfattiho konstrukcí cca poloviční oproti optimálnímu řešení. Goldberg v roce 1967 dokonce ukázal, že Malfattiho konstrukce není správná pro žádný trojúhelník. Za definitivní řešení tohoto problém bychom mohli považovat výsledek Salgallera a Lose z roku 1992, kdyby jejich zdůvodnění nebylo místy neúplné a v některých pasážích nebylo založena na čistě numerických výpočetních argumentech. První úplný geometricko–analytický se podařilo získat teprve nedávno Lombardimu, který jej zveřejnil v červnu 2022, čímž se tento příběh po více než 200 letech uzavřel – alespoň v případě tří kružnic, ale kružnic samozřejmě můžeme uvažovat i více...

[6] Podmínku $x_i \in \{0,1\}$ lze nahradit za $x_i(1-x_i)=0$ a podmínku $x_i \in \mathbb{Z}$ za $\sin(\pi x_i)=0$.

míra spotřebitelova individuálního užitku ze spotřebování $x_1,\dots,x_n \geq 0$ jednotek všech komodit je dána funkcí $U(x_1,\dots,x_n)$. Spotřebitelův dostupný příjem je I a cena i-té komodity je $p_i \geq 0$, tj. $p := (p_1,\dots,p_n)^\top$ je tzv. vektor cen. Potom rozpočtová množina je

$$X := \{x \in \mathbb{R}^n_+ \mid \langle p, x\rangle \leq I\}$$

a cílem uživatele je maximalizovat míru svého užitku při dodržení rozpočtového omezení, tj. nalézt řešení úlohy

$$U(x_1,\dots,x_n) \to \max, \quad x \in X.$$

Může-li navíc spotřebitel rozdělit svůj čas na práci a volno, přičemž I' je netarifní složka mzdy (tj. pevná částka bez ohledu na odpracovanou dobu), w tarifní složka mzdy, ℓ_0 počet hodin práce a ℓ_1 počet hodin volna, pak je možné maximalizaci spotřebitelova užitku popsat jako

$$U(x_1,\dots,x_n) \to \max, \quad \langle p, x\rangle \leq I' + w\ell_0, \quad \ell_0 + \ell_1 = 24, \quad x \geq 0, \quad \ell_0 \geq 0, \quad \ell_1 \geq 0.$$

Typickou funkcí popisující spotřebitelův užitek je tzv. *Cobbova–Douglasova* funkce

$$U(x_1,\dots,x_n) = x_1^{a_1}\cdots x_n^{a_n}, \quad a_1,\dots,a_n \geq 0,$$

která je konkávní pro $a_1+\cdots+a_n \leq 1$ a ostře konkávní pro $a_1+\cdots+a_n < 1$.
Dalšími významnými ekonomickými úlohami jsou např.

(a) *Minimalizace výdajů*, což je opak předchozí úlohy. Je dán vektor cen $p \in \mathbb{R}^n_+$ a hledá se minimální příjem pro dosažení užitku alespoň o hodnotě $\bar{u}$, je-li číslo $\bar{u}$ předem dáno. Potom přípustná množina je

$$X(\bar{u}) := \{x \in \mathbb{R}^n_+ \mid U(x_1,\dots,x_n) \geq \bar{u}\}$$

a cílem je nalezení řešení úlohy

$$\langle p, x\rangle \to \min, \quad x \in X(\bar{u}).$$

Zde můžeme dokonce $\bar{u}$ chápat jako parametr a analyzovat jeho vliv na hodnotu minima, čímž se dostaneme k tzv. *parametrickému programování*.

(b) V *teorii výrobce* jsou studovány rozhodovací procesy pro firmy. Typickým příkladem v této oblasti je firma, která vyrábí 1 výrobek za použití m komponent (vstupů) s výrobní funkcí $y = g(x_1,\dots,x_n)$, kde $x_1,\dots,x_n$ je použité množství jednotlivých komponent a y je odpovídající vyrobené množství. Je zřejmé, že $g : \mathbb{R}^n_+ \to \mathbb{R}_+$. Jednotková cena i-té komodity je $w_i \geq 0$. Vyrobí-li firma y jednotek výrobku, můžeme jednotkovou cenu získat pomocí tzv. poptávkové funkce $p : \mathbb{R}_+ \to \mathbb{R}_+$ (v případě, že firma podniká v konkurenčním prostředí, obvykle se předpokládá, že $p(y) = \bar{p}$ pro všechny $y \geq 0$ a nějaké pevné $\bar{p} \geq 0$, tj. firma nemá vliv na tržní ceny). Je-li dostupné libovolné množství každé komponenty, tak potom cílová *maximalizace zisku* je řešením úlohy

$$p\big(g(x_1,\dots,x_n)\big)\cdot g(x_1,\dots,x_n) - wx \to \max, \quad x \in \mathbb{R}^n_+.$$

(c) Pro firmy můžeme uvažovat i analogii úlohy s minimalizací výdajů, tj. *minimalizaci nákladů*. V ní je cílem určení množství nakupovaných komodit tak, abychom minimalizovali náklady na výrobu alespoň $\bar{y}$ jednotek výrobku při dané výrobní funkci $g : \mathbb{R}^n_+ \to \mathbb{R}^+$ a ceně vstupů $w_1,\dots,w_n$. Tedy přípustná množina je

$$X(\bar{y}) := \{x \in \mathbb{R}^n_+ \mid g(x) \geq \bar{y}\}$$

a cílem je

$$w_1 x_1 + \cdots + w_n x_n \to \min, \quad x \in X(\bar{y}).$$

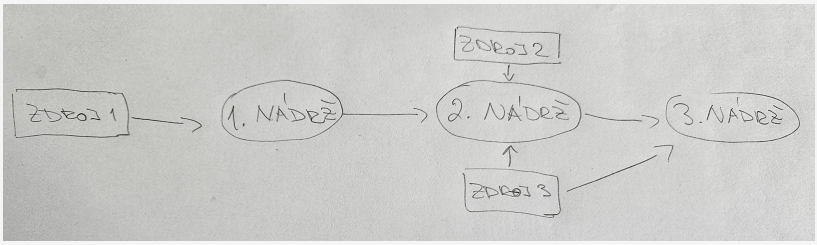
Tato formulace implicitně zahrnuje předpoklad „bezplatné likvidace“, tj. jakékoli nadbytečné vyrobené jednotky mohou být zdarma zlikvidovány (vyřazeny). V opačném případě je nutné tyto výdaje zahrnout do nákladů.

Příklad 8.2

(*Regionální vodní hospodářství I.*) V regionálním vodním hospodářství může být požadováno, aby „producenti“ znečištění vody odstranili toto znečištění z vodního systému. Označme x_j množství biochemické spotřeby kyslíku (BSK, tj. množství kyslíků, které je potřeba k úplné oxidaci biologicky odbouratelných látek obsažených ve zkoumané vodě — vyšší hodnota znamená větší znečištění), které má být odstraněno v j-té zdroji, viz Obrázek 8.1 znázorňující možný systém nádrží a zdrojů znečištění. Označme $f_j(x_j)$ náklady na odstranění BSK v j-tém zdroji, b_i minimální požadované množství vylepšení kvality vody v i-tém bodě systému (tzv. kontrolní místo), a_{ij} efektivita odstranění 1 BSK v j-tém zdroji pro i-tý bod v systému, u_j maximální množství BSK odstranitelné v j-tém místě. Cílem potom je

$$\sum_{j=1}^{n} f_j(x_j) \to \min \quad \& \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_j \quad \& \quad x_j \in [0, u_j]$$

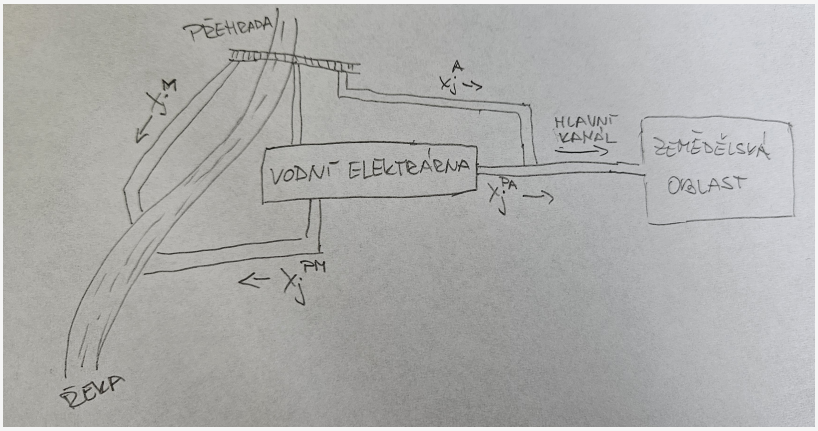
pro $i = 1, \ldots, m$ a $j = 1, \ldots, n$.



Obrázek 8.1: Vodní systém k odstranění znečištění.

Příklad 8.3

(*Regionální vodní hospodářství II.*[7]) Nyní je naším cílem sestavení optimalizačního modelu pro společné využívání vodních zdrojů pro výrobu elektrické energie i pro zemědělské využití. Uvažujme povodí schematicky znázorněné na Obrázku 8.2.



Obrázek 8.2: Schematické znázornění povodí s vodní elektrárnou a zemědělskou oblastí.

[7] Tento příklad pochází z [34, str. 17–20], kde čtenář nalezne navíc i odkazy na literaturu poskytující dodatečné informace k některým funkcím, které se zde vyskytují.

Přehrada napříč řekou poskytuje zásobárnu povrchové vody, která zajišťuje vodu pro výrobu energie a zemědělství. Elektrárna se nachází v blízkosti přehrady a voda pro zemědělství je z přehrady dopravována kanálem přímo nebo skrze vodní elektrárnu po jejím využití k výrobě energie. S problémem jsou spojeny dvě skupiny proměnných.

(a) *Návrhové proměnné:* jaká by měla být optimální kapacita S nádrže, kapacita U kanálu pro zásobování zemědělské oblasti a kapacita E vodní elektrárny?

(b) *Rozhodovací proměnné:* kolik vody by mělo být využíváno pro zemědělství, výrobu elektřiny a další účely?

Z Obrázku 8.2 snadno vyčteme provozní proměnné vztažené k j-tému období:

(a) x_j^A vyjadřující množství vody vypouštěné z přehrady pro zemědělství;

(b) x_j^{PA} vyjadřující množství vody vypouštěné z přehrady pro výrobu elektrické energie a následně pro zemědělské využití;

(c) x_j^{PM} vyjadřující množství vody vypouštěné z přehrady pro výrobu elektrické energie a následně zpět do řeky;

(d) x_j^M vyjadřující množství vody vypouštěné z přehrady přímo do řeky;

Pro účely plánovacího modelu budeme uvažovat horizont N období odpovídající životnosti velkých kapitálových investic, jakými je například přehrada. Cílem je minimalizovat celkové diskontované[8] náklady spojené s nádrží, elektrárnou a kanálem po odečtení příjmů z výroby elektřiny a zemědělství. V souvislosti s elektrárnou jsou náklady

$$C(E) + \sum_{j=1}^{N} \beta_j \widehat{C}_e(E), \tag{8.3}$$

kde $C(E)$ vyjadřuje náklady na vybudování vodní elektrárny s kapacitou E, související stavby a přenosovou soustavu, $\widehat{C}_e(E)$ jsou roční náklady na provoz, údržbu a případnou výměnu zařízení v elektrárně. Číslo β_j je diskontní míra, která udává současnou hodnotu nákladů v j-tém období.

Diskontované výnosy z prodeje elektřiny lze vyjádřit jako

$$\delta \Big\{ \sum_{j=1}^{N} \beta_j \left[p_f F_j + p_d (f_j - F_j) \right] \Big\} + (1-\delta) \Big\{ \sum_{j=1}^{N} \beta_j \left[p_f f_j + p_s (F_j - f_j) \right] \Big\}, \tag{8.4}$$

kde F_j je známá fixní poptávka po elektřině, kterou lze prodat za p_j, a f_j je výroba elektřiny v j-tém období. Přitom $\delta = 1$, jestliže $f_j > F_j$, a přebytečný výkon $f_j - F_j$ lze prodat za dumpingovou cenu p_d. Na druhé straně $\delta = 0$, jestliže $f_j < F_j$, čímž však vzniknou vícenáklady ve výši $p_s(F_j - f_j)$, neboť chybějící výkon se bude muset nakoupit ze sousední elektrické sítě.

Diskontované kapitálové náklady spojené s nádrží a hlavním kanálem jsou dány jako

$$C_r(S) + \alpha\, C_\ell(U), \tag{8.5}$$

kde $C_r(S)$ vyjadřuje investiční náklady na vybudování nádrže s kapacitou S a $C_\ell(U)$ jsou investiční náklady spojené s hlavním kanálem o kapacitě U. Číslo α zohledňuje nižší životnost kanálu ve srovnání s nádrží. Diskontované provozní náklady jsou pak dány vztahem

$$\sum_{j=1}^{N} \beta_j \left[\widehat{C}_r(S) + \widehat{C}_\ell(U) \right]. \tag{8.6}$$

Výnos ze zavlažovaných plodin lze vyjádřit jako funkci R vody použité na zavlažování během j-tého období, takže příjmy ze zemědělství jsou dány vztahem

$$\sum_{j=1}^{N} \beta_j R(x_j^A + x_j^{PA}), \tag{8.7}$$

[8] Diskontování je matematický postup, kdy jsou diskontovány (přepočítány a sečteny) budoucí výnosy (zisky/peníze/peněžní toky) v jednotlivých obdobích na současnou hodnotu investice s použitím diskontní míry (odhadnuté výnosové míry).

přičemž zde pro jednoduchost zanedbáváme vodu získanou prostřednictvím dešťových srážek.
Toto všechno jsou členy vystupující v účelové funkci. Model ale musí také zohledňovat omezení kladená na návrhové a rozhodovací proměnné. Je zřejmé, že vyrobený výkon nesmí překročit energetický potenciál dodávané vody, takže je nutné, aby

$$f_j \leq (x_j^{PM} + x_j^{PA})\Psi(s_j)\,\gamma\,e, \tag{8.8}$$

kde $\Psi(s_j)$ je převýšení vytvořené vodou s_j uloženou v nádrži během j-tého období, γ je konverzní faktor výkonu a e je účinnost energetického systému. Podobně vyrobený výkon nesmí překročit výrobní kapacitu elektrárny, takže

$$f_j \leq \alpha_j\, E\, e\, H_j, \tag{8.9}$$

kde α_j je faktor zatížení definovaný jako poměr průměrné denní výroby k denní špičkové výrobě a H_j je počet provozních hodin. Konečně kapacita elektrárny musí být v rámci známých přijatelných mezích, tj.

$$E' \leq E \leq E''. \tag{8.10}$$

Pokud zanedbáme ztráty způsobené odpařováním, musí se množství vody y_j přitékající do přehrady rovnat změně množství vody uložené v přehradě a vody vypouštěné pro různé účely. To lze vyjádřit jako

$$s_{j+1} - s_j + x_j^A + x_j^M + x_j^{PM} + x_j^{PA} = y_j. \tag{8.11}$$

Druhé omezení v souvislosti s nádrží říká, že zásoby vody v nádrži by měly být přiměřené a v přijatelných mezích, tj.

$$S \geq s_j \quad \& \quad S' \leq S \leq S''. \tag{8.12}$$

Obvykle je také nutné stanovit, že se uvolní určité množství vody M_j, aby byly splněny požadavky na vodu v navazujících tocích, což lze vyjádřit jako

$$x_j^M + x_j^{PM} \geq M_j. \tag{8.13}$$

V neposlední řadě je třeba zaručit, že kapacita kanálu U bude dostatečná pro odvádění vody pro zemědělské účely, tj.

$$x_j^A + x_j^{PA} \leq U. \tag{8.14}$$

Naším cílem je pak minimalizovat čisté náklady představované součtem (8.3), (8.5) a (8.6) pro odečtení příjmů daných v (8.4) a (8.7) za omezení (8.8)–(8.14) a požadavku na nezápornost všech proměnných.

Příklad 8.4

(*Výroba a zásoby.*[9]) Předpokládejme, že společnost vyrábí určitý produkt pro uspokojení známé poptávky. Naším cílem je stanovení výrobního plánu na celkem K-tici po sobě jdoucích období. Poptávku v libovolném období lze uspokojit ze zásob, které jsou k dispozici na začátku období, a z výroby v průběhu daném období. Maximální produkce během libovolného období je pochopitelně omezena kapacitou celého výrobního procesu (např. kvůli použitím strojům apod.) tak, že nemůže překročit b jednotek. Předpokládejme, že v případě potřeby lze najmout přiměřenou dočasnou pracovní sílu a v případě potřeby ji zase propustit. Aby se však zamezilo velké fluktuaci pracovní síly, vede tato změna k nákladům úměrným druhé mocnině rozdílu v pracovní síle během

[9] Tento příklad pochází z [34, str. 5–6]

dvou po sobě jdoucích období. Rovněž vznikají skladovací náklady úměrné zásobám převedeným z jednoho období do druhého. Sestavme úlohu, jejímž řešením bude velikost pracovní síly a zásob během jednotlivých období $1,\dots,K$ tak, aby poptávka byla uspokojena a celkové náklady byly minimalizovány.
V tomto problému existují dvě stavové proměnné, úroveň zásob I_k a pracovní síla L_k na konci k-tého období. Řídicí proměnná u_k je pracovní síla získaná během k-tého období ($u_k < 0$ znamená, že práce je snížena o hodnotu $-u_k$). Problém výroby a zásob lze tedy popsat takto:

$$\sum_{k=1}^{K} c_1\, u_k^2 + c_2\, I_k \to \min$$
$$L_k = L_{k-1} + u_k$$
$$I_k = I_{k-1} + p\, L_{k-1} - d_k$$
$$0 \le L_k \le b/p$$
$$I_k \ge 0$$

pro všechna $k = 0,\dots,K$, kde počáteční zásoby I_0 a počáteční pracovní síla L_0 jsou známé, d_k je známá poptávka během k-tého období a p je počet jednotek vyrobených jedním pracovníkem během jakéhokoli daného období.

Příklad 8.5

(*Dvouprutová příhradová konstrukce.*[10,11]) Konstrukční inženýři se tradičně snažili vyvinout konstrukce, které by mohly bezpečně unést plánovanou zátěž. Koncept optimality zde byl zahrnut pouze implicitně prostřednictvím standardní praxe a zkušenosti konstruktéra. V poslední době si však návrhy složitějších konstrukcí (jako jsou např. konstrukce v letectví a kosmonautice) vyžádaly mnohem explicitnější zohlednění optimality. Hlavní přístupy používané pro návrh konstrukčních systémů s minimální hmotností jsou založeny na použití matematického programování nebo jiných rigorózních numerických nástrojů kombinované s metodami statické analýzy (šlo především o lineární a nelineární programování a simulace pomocí metody Monte Carlo). V [31] nalezneme následující komentář:

> *Celý proces návrhu složité letecké konstrukce je vícestupňový a sahá od posouzení celkového výkonu systému až po podrobný návrh jednotlivých součástí. Ačkoli všechny úrovně procesu navrhování mají určitou větší či menší míru vzájemné interakce, dosavadní stav techniky v oblasti navrhování vyžadoval předpoklad relativně volné vazby mezi jednotlivými fázemi. Počáteční práce v oblasti optimalizace konstrukcí směřovaly k zachování tohoto rozvrstvení filosofie návrhu, i když tento stav vznikl pravděpodobně spíše v důsledku metodiky použité pro optimalizaci než z touhy zachovat hranice mezi jednotlivými etapami návrhu.*

Následující příklad ilustruje, jak lze s pomocí metod statické analýzy získat problém nelineárního programování zahrnující návrh dvouprutové příhradové konstrukce s minimální hmotností.
Uvažme rovinný příhradový nosník zobrazený na Obrázku 8.3. Tento nosník se skládá ze dvou ocelových trubek, které jsou a jednom konci spojené k sobě a na druhém konci je každá z nich upevněna pomocí čepu. Rozpětí (tj. vzdálenost mezi oběma čepy) je pevně stanovena jako $2r$. Konstruktér pak stojí před otázkou, jak zvolit výšku příhrady a tloušťku a středový průměr ocelových trubek tak, aby vazník unesl zatížení $2T$ při minimalizaci celkové hmotnosti příhrady.

[10] *Příhrada* (angl. *truss*) nebo *příhradová konstrukce* je široce využívaný typ nosné konstrukce, kde namísto masivních a těžkých stěn, desek či bloků jsou využity štíhlé a lehké podélné nosné binární prvky, pruty, nosníky, lana apod. Podstata využití této konstrukce tkví ve vhodném přenesení a rozložení zatížení/mechanického napětí mezi nosnými prvky tak, aby vznikla konstrukce s co nejmenší hmotností, rozměry a finančními náklady.

[11] Tento příklad pochází z [34, str. 9–11].

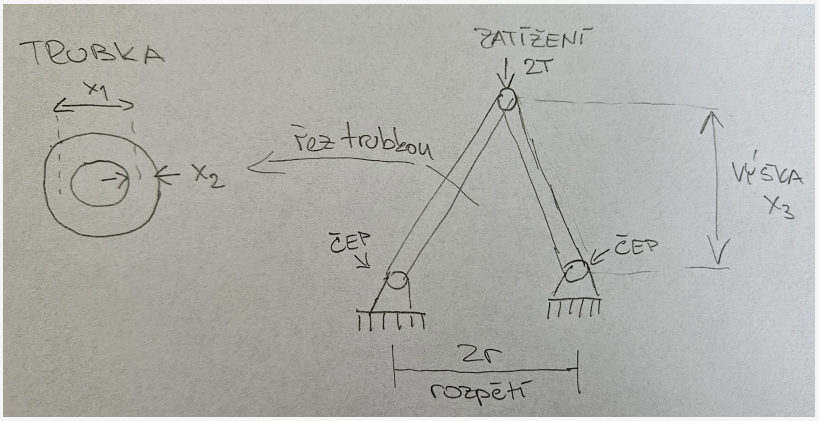
Označme jako x_1, x_2, x_3 postupně (průměrný) průměr trubky, tloušťku trubky a výšku příhrady. Hmotnost celé ocelové konstrukce pak spočteme jako $2\pi\,\rho\,x_1\,x_2\sqrt{s^2+x_3^2}$, kde ρ značí hustotu ocelové trubky. Při tom musíme mít na paměti následující omezení:

(a) z důvodu prostorového omezení nemůže výška příhrady překročit hodnotu b_1, tj. $x_3 \le b_1$;

(b) poměr průměru trubky a její tloušťky nesmí překročit b_2, tj. $x_1 / x_2 \le b_2$;

(c) tlakové napětí v trubkách nesmí překročit mez pružnosti oceli, což vede na omezení $W\sqrt{s^2+x_3^2} \le b_3\,x_1\,x_2\,x_3$;

(d) výška, průměr a tloušťka musí být zvoleny tak, aby se trubky pod zatížením nezkroutily, což lze matematicky vyjádřit s pomocí známého parametru b_4 jako $W\sqrt{(s^2+x_3^2)^3} \le b_4\,x_1\,x_3\,(x_1^2+x_2^2)$.



Obrázek 8.3: Konstrukce dvouprutové příhrady.

Tím získáme následující úlohu nelineárního programování

$$x_1\,x_2\sqrt{s^2+x_3^2} \to \min$$
$$x_3 - b_1 \le 0$$
$$x_1 - b_2\,x_2 \le 0$$
$$W\sqrt{s^2+x_3^2} - b_3\,x_1\,x_2\,x_3 \le 0$$
$$W\sqrt{(s^2+x_3^2)^3} - b_4\,x_1\,x_3\,(x_1^2+x_2^2) \le 0$$
$$x_1 \ge 0 \quad \& \quad x_2 \ge 0 \quad \& \quad x_3 \ge 0.$$

Příklad 8.6

(*Vypuštění rakety s minimální energií.*[12]) Jedná se problém, kde chceme dostat raketu z úrovně země do výšky $\bar{y}$ v čase T. Nechť $y(t)$ označuje výšku rakety od země v čase t a nechť $u(t)$ označuje sílu působící ve vertikálním směru v čase t. Jestliže raketa má hmotnost w, je pohybová rovnice dána vztahem

$$m\,y''(t) + m\,g = u(t) \quad \text{pro } t \in [0,T],$$

kde $y''(t)$ je zrychlení v čase t a g je gravitační zrychlení, které způsobuje zpomalování rakety. Dále předpokládejme, že maximální síla, kterou lze kdykoli vyvinout, nemůže překročit b. Pokud je cílem vynaložit co nejmenší energii, aby raketa dosáhla výšky $\bar{y}$ v čase T, lze tento problém zapsat následovně

$$\int_0^T |u(t)|\,y'(t)\,\mathrm{d}t \to \min\, m\,y''(t) + m\,g = u(t)|u(t)| \le b y(T) = \bar{y}$$

[12] Spojitý problém patří do kategorie úloh *spojitého optimálního řízení* a lze jej nalézt např. v [359, str. 125–126]. Jeho diskrétní analogie v podobě úlohy nelineárního programování pochází z [34, str. 7–8].

pro $t \in [0,T]$, kde $y(0) = 0$. Položíme-li $y_1 := y$ a $y_2 := y'$, pak můžeme tento problém převést na úlohu s pouze první derivací, tj.

$$\int_0^T |u(t)|\, y_2(t)\, \mathrm{d}t \to \min \quad y_1'(t) = y_2(t) \quad m\, y_2'(t) = u(t) - m\, g \quad |u(t)| \le b \quad y_1(T) = \bar{y}$$

pro $t \in [0,T]$, kde $y_1(0) = 0 = y_2(0)$. Nyní rozdělme interval $[0,T]$ na K částí s (pro jednoduchost stejnou) délkou ℓ. Označme sílu, výšku a rychlost pro $k = 1,\dots,K$ postupně jako u_k, $y_{1,k}$ a $y_{2,k}$. Předchozí úlohu pak můžeme aproximovat úlohou nelineárního programování

$$\sum_{k=1}^{K} |u_k|\, y_{2,k} \to \min$$
$$y_{1,k} - y_{1,k-1} = y_{2,k-1}$$
$$m\,(y_{2,k} - y_{2,k-1}) = u_k - m\, g$$
$$|u_k| \le b$$
$$y_{1,K} = |y|$$

pro $k = 1,\dots,K$.

8.1 Obecná optimalizační úloha

Uvažme nyní obecnou úlohu MP, tj. úlohu (8.1) s (blíže nespecifikovanou) množinou $X \subseteq \mathbb{R}^n$ a funkcí $f : X \to \mathbb{R}$. Pro $x^* \in X$ si zaveďme dvě množiny

$$\mathcal{V}(x^*, X) := \left\{ h \in \operatorname{Lin} X \;\middle|\; \text{existuje } \alpha_0 > 0 \text{ takové, že } x^* + th \in X \text{ pro všechna } t \in (0, \alpha_0) \right\}$$

a

$$\mathcal{U}(x^*, f) := \left\{ h \in \operatorname{Lin} X \;\middle|\; \text{existuje } \alpha_0 > 0 \text{ takové, že } x^* + th \in D(f) \right.$$
$$\left. \text{a } f(x^* + th) < f(x^*) \text{ pro všechna } t \in (0, \alpha_0) \right\}.$$

Jaký je význam těchto množin?

(i) Množina $\mathcal{V}(x^*, X)$ je množina všech vektorů „mířících“ z bodu x^* do množiny X. Navíc jistě[13] $0 \in \mathcal{V}(x^*, X)$ a s každým $h \in \mathcal{V}(x^*, X)$ platí i $\beta h \in \mathcal{V}(x^*, X)$ pro libovolné $\beta > 0$, neboť stačí vzít $\tilde{\alpha}_0 := \alpha_0 / \beta$. Proto je $\mathcal{V}(x^*, X)$ tzv. *kužel přípustných vektorů*[14] v bodě x^*. Je-li $x^* \in \operatorname{ri} X$, pak nutně $\mathcal{V}(x^*, X) = \operatorname{Lin} X$. Avšak v obecném případě nemusí tento kužel být konvexní, viz např. Obrázky 8.4–8.5.

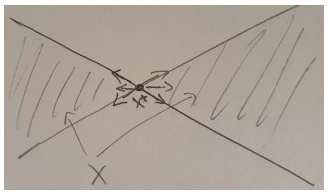
(ii) Množina $\mathcal{U}(x^*, f)$ obsahuje tzv. *spádové vektory*, v jejichž směru je funkce f v bodě x^* klesající. Jedná se o tzv. *kužel*[15] *zlepšujících vektorů* (případně *zlepšujících směrů*) funkce f v bodě x^* (každý směr dává lepší, tj. menší funkční hodnotu f). Je zřejmé, že pokud $x^* + th \notin D(f)$, pak nutně $h \notin \mathcal{U}(x^*, f)$. Na první pohled by se mohlo zdát, že první podmínku $x^* + th \in D(f)$ by šlo nahradit požadavkem $h \in \mathcal{V}(x^*, X)$, jenže to by vyžadovalo $D(f) = X$, zatímco my máme pouze $X \subseteq D(f)$, jako např. pro $\sqrt{x_1 x_2}$ a $X = \mathbb{R}^2_+$. Jinými slovy, funkce může klesat i ve směru nepatřícího do $\mathcal{V}(x^*, X)$.

Odtud vidíme, že množina $\mathcal{V}(x^*, X)$ „charakterizuje“ tvar množiny X a je zcela nezávislá na funkci f, zatímco množina $\mathcal{U}(x^*, f)$ částečně „popisuje“ chování funkce f a na lineární obalu X (tj. konkrétní tvar množiny X je nepodstatný). A právě díky této vzájemné nezávislosti obou množin obdržíme první nutnou podmínku pro existenci lokálního řešení úlohy (8.1), viz Obrázek 8.6. Tvrzení je asi víceméně zřejmé a důkaz relativně snadný, ale význam tohoto tvrzení pro námi budovanou teorii je značný. Všimněme si také, že tato podmínka rozhodně nevyžaduje diferencovatelnost funkce f.

[13] Zas až tak jisté to není. Definice množiny $\mathcal{V}(x^*, X)$ by mohla dávat smysl i pro případ, kdy $x^* \in \overline{X}$, např. v situaci kdy množina X je otevřená, viz [93, Definition 1.20]. My ale po řešení úlohy (8.1) vyžadujeme, aby byly prvkem X, tj. rozšíření tohoto konceptu na množinu $x \in \overline{X} \setminus X$ by nám mohlo dát nějakou dodatečnou informaci, ale pro řešení samotné nám nepomohou.

[14] Omezíme-li se pouze na vektory splňující $\|h\| = 1$, pak se hovoří o *přípustných směrech*.

Obrázek 8.4: Ukázka bodů $x^* \in \operatorname{ri} X = \operatorname{int} X$ a $x^* \in \partial X$ a odpovídajících množin $\mathcal{V}(x^*, X)$.



Obrázek 8.5: Ukázka bodu $x^* \in X$, pro který $\mathcal{V}(x^*, X)$ není konvexní.

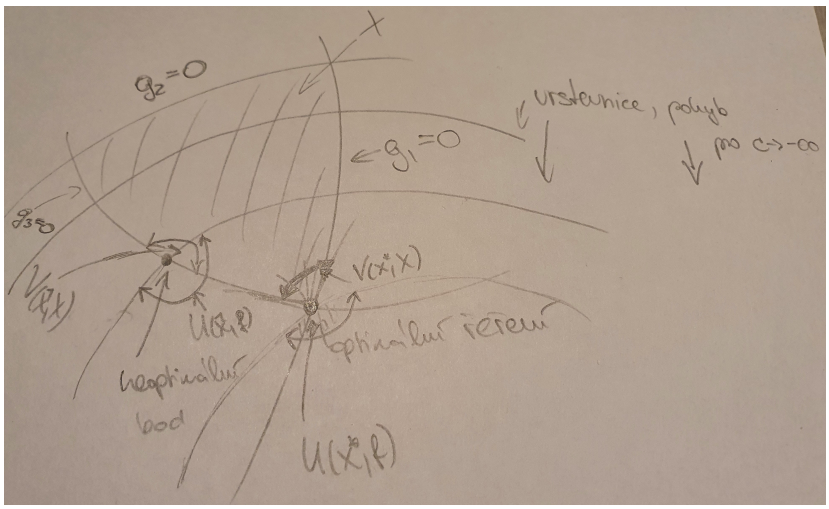
[15] Toto už ale není kužel ve smyslu naší Definice 6.1.5, neboť $0 \notin \mathcal{U}(x^*, f)$.

LEMMA 8.1.1

Nechť $X \subseteq \mathbb{R}^n$ a $f : X \to \mathbb{R}$ jsou dány. Je-li bod $x^* \in X$ lokálním řešením úlohy (8.1), potom

$$\mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) = \emptyset. \tag{8.1.1}$$

Důkaz. Tvrzení dokážeme sporem. Nechť bod $x^* \in X$ je lokálním řešením úlohy (8.1) a připusťme, že existuje vektor $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f)$, tj. existují čísla $\alpha_u, \alpha_v > 0$ taková, že $x^* + th \in X$ pro $t \in (0, \alpha_v)$ a současně $f(x^* + th) < f(x^*)$ pro $t \in (0, \alpha_u)$. Potom tedy pro libovolné $t \in (0, \alpha)$, kde $\alpha \in (0, \min\{\alpha_v, \alpha_u\})$, je současně $x^* + th \in X \subseteq D(f)$ a $f(x^* + th) < f(x^*)$. To je ale zjevný spor s požadavkem lokálního řešení úlohy (8.1) v bodě x^* ↯ a důkaz je hotov. ■



Obrázek 8.6: Bod x^* je řešením úlohy (8.1) a platí v něm (8.1.1). Bod $\hat{x}$ není řešením úlohy (8.1) a podmínka (8.1.1) v něm splněna není.

[16] Nebo pokud $\mathcal{V}(x^*, X) = \{0\}$. Co by to ale znamenalo pro bod x^*? Ten by musel být izolovaný, což lze v souladu s definicí považovat za lokální řešení úlohy (8.1).

Podmínka (8.1.1) bude splněna, zejména ve chvíli, kdy[16] $\mathcal{U}(x^*, f) = \emptyset$, tj. z bodu x^* neexistuje žádný směr v $\operatorname{Lin} X$, v němž by funkce f klesala. Kdy to nastane? Pro každé $h \in \operatorname{Lin} X$ takové, že $x^* + th \in D(f)$ pro $t > 0$ dostatečně malá, musí platit $f(x^* + th) - f(x^*) \geq 0$. Je-li f spojitě diferencovatelná, pak

$$f(x^* + th) = f(x^*) + \langle \operatorname{grad} f(x^*), th \rangle + \omega(th)$$

neboli

$$\frac{f(x^* + th) - f(x^*)}{t} = \langle \operatorname{grad} f(x^*), h \rangle + \frac{\omega(th)}{t},$$

což pro $t \to 0^+$ vede k podmínce

$$\langle \operatorname{grad} f(x^*), h \rangle \geq 0.$$

Zejména v případě $\operatorname{Lin} X = \mathbb{R}^n$ a $x^* \in \operatorname{int} D(f)$ musí platit známý požadavek $\operatorname{grad} f(x^*) = 0$. Nicméně toto je zbytečně silná podmínka a nám postačuje pouze (8.1.1), což nás přivádí k následující vylepšené definici stacionárního bodu, viz také Poznámku 6.4.5(ii).

DEFINICE 8.1.2

Nechť množina $X \subseteq \mathbb{R}^n$ je konvexní a funkce $f : X \to \mathbb{R}$ diferencovatelná na (nějaké otevřené množině obsahující) X. Řekneme, že bod $x^* \in X$ je *stacionárním bodem* úlohy (8.1) (nebo lépe *stacionárním bodem funkce f na množině X*), jestliže

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle \geq 0 \tag{8.1.2}$$

pro každé $x \in X$.

Výraz na levé straně (8.1.2) je roven (oboustranné) směrové derivaci $f_{h(x)}(x^*)$ pro $h(x) := x - x^*$. Máme tedy nerovnost $f_{h(x)}(x^*) \geq 0$ pro všechna $x \in X$, což znamená, že funkce f je neklesající ve všech „přípustných“ směrech z bodu x^*.

Poznámka 8.1.3.

(i) Je-li $X = \mathbb{R}^n$, pak podmínka (8.1.2) je tvaru

$$\sum_{i=1}^{n} \frac{\partial f(x^*)}{\partial x_i}(x_i - x_i^*) \geq 0 \qquad \forall x \in \mathbb{R}^n.$$

Jenže v takovém případě jedinou možností pro splnění této nerovnosti je pouze $\frac{\partial f(x^*)}{\partial x_i} = 0$ pro všechna $i \in \{1, \dots, n\}$, tj. dostáváme klasickou nutnou podmínku prvního řádu pro lokální extrém $\operatorname{grad} f(x^*) = 0$. Vskutku, je-li $\frac{\partial f(x^*)}{\partial x_j} > 0$ pro nějaké j, pak volbou $x \in \mathbb{R}^n$ splňujícího $x_j - x_j^* = -1$ a $x_i - x_i^* = 0$ pro $i \neq j$ dostaneme spor s (8.1.2). Podobně v případě $\frac{\partial f(x^*)}{\partial x_j} < 0$ pro nějaké j, pak volbou $x \in \mathbb{R}^n$ splňujícího $x_j - x_j^* = 1$ a $x_i - x_i^* = 0$ pro $i \neq j$ dostaneme opět spor s (8.1.2).

(ii) V případě $X = [a, b]$ má nerovnost (8.1.2) podobu $f'(x^*)\,(x - x^*) \geq 0$ pro všechna $x \in [a, b]$. Pak podobnou úvahou jako před chvílí můžeme ukázat, že tato nerovnost je ekvivalentní s trojicí podmínek

(a) je-li $x^* \in (a, b)$, pak $f'(x) = 0$, tj. máme „klasický“ stacionární bod;

(b) je-li $x^* = a$, pak $f'(x) \geq 0$, tj. funkce f je neklesající napravo od a;

(c) je-li $x^* = b$, pak $f'(x) \leq 0$, tj. funkce f je nerostoucí nalevo od b,

viz také Poznámku 8.1.7(ii) níže. ▲

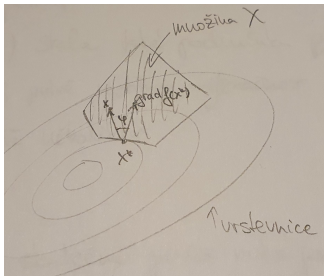
Následující věta ukazuje, že stacionární bod ve smyslu Definice 8.1.2 má přesně ty vlastnosti, které od něj s ohledem na úlohu (8.1) očekáváme, viz Obrázky 8.7–8.8.

Věta 8.1.4

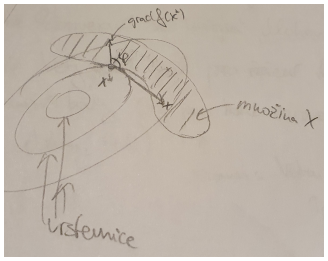
Nechť funkce $f : X \to \mathbb{R}$ je diferencovatelná na (nějaké otevřené množině obsahující) konvexní množině $X \subseteq \mathbb{R}^n$.

(i) Je-li $x^* \in X$ lokálním extrémem funkce f na X (tj. lokálním řešením úlohy (8.1)), pak x^* je stacionárním bodem funkce f na X.

(ii) Naopak, je-li f (ostře) konvexní na X a $x^* \in X$ je stacionárním bodem f na X, pak x^* je (jediným) řešením úlohy (8.1), tj. (jediným) globálním minimem f na X.



Obrázek 8.7: Geometrický význam první části Věty 8.1.4 — v bodě lokálního minima vektor $\operatorname{grad} f(x^*)$ svírá ostrý nebo pravý úhel (měřeno neorientovaně) se všemi možnými směry $x - x^*$, $x \in X$.



Obrázek 8.8: V případě nekonvexní množiny X už první část Věty 8.1.4 nemusí být pravdivá. Zde v bodě x^* sice nastává globální minimum, avšak existuje $x \in X$, pro který vektory $x - x^*$ a $\operatorname{grad} f(x^*)$ svírají tupý úhel, tj. $\langle \operatorname{grad} f(x^*), x - x^* \rangle < 0$.

Poznámka 8.1.5. První část Věty 8.1.4 udává *nutnou*[17] podmínku pro řešení úlohy (8.1), ze které se ovšem ve druhé části (po přidání požadavku konvexnosti funkce f) stala podmínka *postačující*. S tímto jevem, který ukazuje význačnost konvexních funkcí v matematickém programování, se setkáme ještě několikrát.

Úloha (8.1) s konvexní funkcí f na konvexní množině X se nazývá *úlohou konvexního programování*. Pro takovou úlohu podle Věty 8.1.4 stačí „pouze“ najít stacionární body, které jsou lokálními i globálními minimy, tj. řešeními úlohy (8.1), viz[18] Větu 6.2.23(i). ▲

Důkaz Věty 8.1.4. Nechť $x^* \in X$ je lokálním řešením úlohy (8.1) a neplatí nerovnost (8.1.2), tj. existuje $x \in X$ takové, že $\langle \operatorname{grad} f(x^*), x - x^* \rangle < 0$. Potom pro $h := x - x^*$ máme $h \in \mathcal{V}(x^*, X)$, neboť z konvexnosti množiny X plyne

$$x^* + th = x^* + t(x - x^*) = tx + (1 - t)x^* \in X \quad \text{pro každé } t \in [0, 1],$$

a současně $h \in \mathcal{U}(x^*, f)$, neboť platí

$$0 > \langle \operatorname{grad} f(x^*), x - x^* \rangle = f'_h(x^*) = \lim_{t \to 0} \frac{f(x^* + th) - f(x^*)}{t},$$

tj.[19] existuje $\alpha_0 > 0$ takové, že $f(x^* + th) < f(x^*)$ pro každé $t \in (0, \alpha_0)$. Potom ale $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) \neq \emptyset$, což je spor s Lemma 8.1.1 ↯.

Naopak, je-li funkce f (ostře) konvexní a $x^* \in X$ splňuje nerovnost (8.1.2) pro každé $x \in X$, pak z Věty 6.4.3 plyne

$$f(x) \underset{(>)}{\geq} f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle \quad \text{pro všechna } x \in X,$$

[17] Je to skutečně pouze nutná podmínka. Je-li totiž $x^* \in \operatorname{int} X$ stacionární bod, pak nutně $\operatorname{grad} f(x^*) = 0$ a nerovnost (8.1.2) je splněna pro každé $x \in X$. Jenže v takové situaci ztrácíme informaci o růstu funkce f v přípustných směrech a bod x^* již nemusí být lokální minimum.

[18] Pozor na tvrzení Věty 6.2.23(iii)!

[19] V poslední rovnosti zcela záměrně uvažujeme pouze jednostrannou limitu.

což vzhledem k (8.1.2) znamená

$$f(x) \underset{(>)}{\geq} f(x^*) \quad \text{pro všechna } x \in X,$$

tj. bod x^* je (jediným) řešením úlohy (8.1). ■

Z praktického pohledu (viz např. příklady v závěru této podkapitoly) je velmi důležitá zejména situace, kdy

$$X = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_1, \dots x_n \geq 0\}.$$

V takovém případě pak obdržíme z Věty 8.1.4 následující důsledek, jehož geometrickou ilustraci můžeme vidět na Obrázcích 8.9–8.10.

Důsledek 8.1.6

Nechť funkce $f : \mathbb{R}^n_+ \to \mathbb{R}$ je diferencovatelná na nějaké otevřené množině obsahující $\mathbb{R}^n_+$. Je-li $x^* \in \mathbb{R}^n_+$ řešením úlohy

$$f(x) \to \min, \quad x \in \mathbb{R}^n_+, \tag{8.1.3}$$

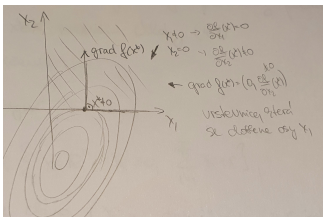
pak platí

$$\frac{\partial f}{\partial x_i}(x^*) \geq 0 \quad \text{a současně} \quad x_i^* \frac{\partial f}{\partial x_i}(x^*) = 0 \quad \text{pro všechna } i = 1, \dots, n, \tag{8.1.4}$$

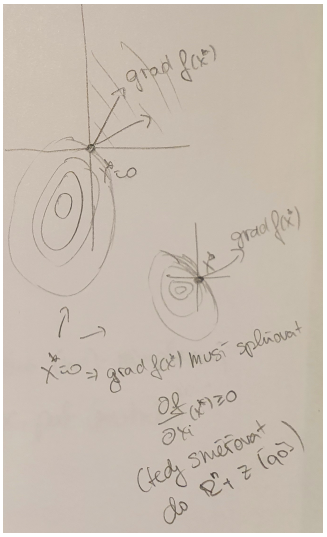
neboli pro každé $i \in \{1, \dots, n\}$ máme

$$\frac{\partial f}{\partial x_i}(x^*) \geq 0 \quad \text{a navíc} \quad \frac{\partial f}{\partial x_i}(x^*) = 0 \text{ v případě } x_i^* > 0$$

Naopak, splňuje-li bod $x^* \in \mathbb{R}^n_+$ podmínku (8.1.4) a funkce f je navíc konvexní na $\mathbb{R}^n_+$, pak x^* je řešením úlohy (8.1.3).



Obrázek 8.9: Geometrický význam Důsledku 8.1.6 — lokální minimum f na X nastává v bodě $x^* \neq 0$ splňujícího $x_1^* \neq 0$ a $x_2^* = 0$. Proto musí být $\frac{\partial f}{\partial x_1}(x^*) = 0$ a zároveň $\frac{\partial f}{\partial x_2}(x^*) \geq 0$. Vidíme však, že x^* není stacionární bod f na $\mathbb{R}^2$, takže nutně $\frac{\partial f}{\partial x_2}(x^*) \neq 0$, tj. $\operatorname{grad} f(x^*) = \left(0, \frac{\partial f}{\partial x_2}(x^*)\right)^\top \neq (0,0)^\top$ je rovnoběžný s osou x_2.



Obrázek 8.10: Geometrický význam Důsledku 8.1.6 — lokální minimum f na X nastává v bodě $x^* = [0,0]$, takže musí být $\frac{\partial f}{\partial x_1}(x^*) \geq 0$ a $\frac{\partial f}{\partial x_2}(x^*) \geq 0$, tj. vektor $\operatorname{grad} f(x^*)$ směřuje z x^* do $\mathbb{R}^2_+$.

Důkaz. Nechť $x^* \in \mathbb{R}^n$ je řešením (lokálním, a tedy také globálním) úlohy (8.1.3). Pak podle Věty 8.1.4 je x^* stacionárním bodem funkce f na X, tj.

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle \geq 0 \quad \text{pro všechna } x \in \mathbb{R}^n_+$$

neboli

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0 \quad \text{pro všechna } x \in \mathbb{R}^n_+. \tag{8.1.5}$$

Potom ale nutně $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$ pro každé $i \in \{1, \dots, n\}$ a $x \in \mathbb{R}^n_+$. Vskutku, jestliže by nastalo $\frac{\partial f}{\partial x_j}(x^*)\,(\widehat{x}_j - x_j^*) < 0$ pro nějaký index $j \in \{1, \dots, n\}$ a nějaké $\widehat{x} = (\widehat{x}_1, \dots, \widehat{x}_n)^\top \in \mathbb{R}^n_+$, pak stačí vzít $\widetilde{x} \in \mathbb{R}^n_+$ takové, že $\widetilde{x}_i - x_i^* = 0$ (tj. $\widetilde{x}_i = x_i^*$) pro $i \in \{1, \dots, n\} \setminus \{j\}$ a $\widetilde{x}_j = \widehat{x}_j \neq x_j^*$, což lze díky tomu, že $X = \mathbb{R}^n_+$. Pak ale v sumě na levé straně (8.1.5) zůstane jediný nenulový člen, kterým bude právě $\frac{\partial f}{\partial x_j}(x^*)\,(\widehat{x}_j - x_j^*) < 0$, což je spor s nerovností (8.1.5) ↯. Jelikož opačné tvrzení je triviální, tj. nezápornost všech $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*)$ implikuje (8.1.5), dostáváme odtud, že (8.1.5) platí právě tehdy, když $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$ pro každé $i \in \{1, \dots, n\}$ a $x \in \mathbb{R}^n_+$.

Zbývá pro ukázat, že

$$\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0 \text{ nastane pro všechna } i \in \{1, \dots, n\} \text{ a } x \in \mathbb{R}^n_+ \text{ právě tehdy, když platí (8.1.4).}$$

O tom se můžeme snadno přesvědčit podobnou úvahou jako před chvílí:

(i) Nechť $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$ pro všechna $i \in \{1, \dots, n\}$ a $x \in \mathbb{R}^n_+$.

  (a) Je-li $x_i^* = 0$, pak $x_i - x_i^* \geq 0$ pro všechna $x \in \mathbb{R}^n_+$, takže nutně $\frac{\partial f}{\partial x_i}(x^*) \geq 0$, tj. platí první nerovnost v (8.1.4), zatímco druhá podmínka je v tomto případě splněna triviálně.

  (b) Je-li $x_i^* > 0$, pak rozdíl $x_i - x_i^*$ může nabývat kladných i záporných hodnot pro $x \in \mathbb{R}^n_+$, takže nutně $\frac{\partial f}{\partial x_i}(x^*) = 0$, v kterémžto případě zjevně platí první i druhá podmínka v (8.1.4).

(ii) Naopak, nechť platí obě podmínky v (8.1.4) a $i \in \{1,\dots,n\}$ je libovolné.

(a) Je-li $x_i^* = 0$ a $\frac{\partial f}{\partial x_i}(x^*) \geq 0$, pak pro každé $x \in \mathbb{R}^n_+$ máme

$$\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) = \frac{\partial f}{\partial x_i}(x^*)\,x_i \geq 0.$$

(b) Je-li $x_i^* > 0$, pak $\frac{\partial f}{\partial x_i}(x^*) = 0$ dle druhé části (8.1.4), a tudíž je $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) = 0$ pro každé $x \in \mathbb{R}^n_+$.

Vidíme, že v obou případech dostáváme $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$, což vzhledem k jeho libovolnosti indexu i znamená, že tato podmínka je splněna pro všechna $i \in \{1,\dots,n\}$. Tím je důkaz první části hotov.

Naopak, nechť pro $x^* \in X$ platí obě podmínky v (8.1.4) a funkce f je konvexní na $\mathbb{R}^n_+$. Potom podle předchozí části máme $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$ pro všechna $i \in \{1,\dots,n\}$ a $x \in \mathbb{R}^n_+$, a tudíž také $\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geq 0$ neboli

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle \geq 0 \quad \text{pro všechna } x \in \mathbb{R}^n_+,$$

tj. $x^* \in \mathbb{R}^n$ je stacionárním bodem funkce f na X. Ovšem v takovém případě plyne vzhledem ke konvexnosti funkce f z Věty 8.1.4, že x^* je řešením úlohy (8.1.3). ■

Poznámka 8.1.7. Podobně jako v Důsledku 8.1.6 můžeme z Věty 8.1.4 získat nutné (a v případě konvexní funkce také postačující) podmínky pro (lokální/globální) řešení úlohy (8.1) v některých dalších význačných případech množiny X.

(i) Pro množinu

$$X = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i \in I\}$$

s libovolnou indexovou množinou[20] $I \subseteq \{1,\dots,n\}$ je nutná (a v případě konvexnosti i postačující) podmínka tvaru[21]

$$\frac{\partial f}{\partial x_i}(x^*) \geq 0 \quad \text{a současně} \quad x_i^* \frac{\partial f}{\partial x_i}(x^*) = 0 \quad \text{pro všechna } i \in I$$
$$\text{a navíc} \quad \frac{\partial f}{\partial x_i}(x^*) = 0 \quad \text{pro všechna } i \in \{1,\dots,n\} \setminus I\,.$$

(ii) Pro množinu

$$X = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1,\dots,n\},$$

kde $\alpha_i, \beta_i \in \mathbb{R}$ jsou daná čísla splňující $\alpha_i < \beta_i$, je[22] nutná (a případně i postačující) podmínka tvaru

$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} \geq 0, & x_i^* = \alpha_i, \\ \leq 0, & x_i^* = \beta_i, \\ = 0, & \alpha_i < x_i^* < \beta_i. \end{cases}$$

(iii) Pro množinu

$$X = \Big\{x \in \mathbb{R}^n \,\Big|\, x_i \geq 0,\ \sum_{i=1}^n x_i = r\Big\},$$

kde $r > 0$ je dané číslo[23], je nutná (a případně i postačující) podmínka ve tvaru implikace

$$x_i^* > 0 \implies \frac{\partial f}{\partial x_i}(x^*) \leq \frac{\partial f}{\partial x_j}(x^*) \quad \text{pro všechna } j \in \{1,\dots,n\}. \qquad ▲$$

V případě, že funkce f není konvexní, však samotná podmínka (8.1.2) nestačí a potřebujeme k rozhodnutí o „extrémnosti“ stacionárního bodu x^* nějaký další nástroj. Tím jsou pro funkci $f \in \mathrm{C}^2$ (na nějaké množině obsahující X) podmínky druhého řádu. Nejdříve si uvedeme nutnou podmínku a několik postačujících podmínek. Jedná se o přirozené analogie tvrzení známých z úloh pro volné extrémy nebo vázané extrémy s rovnostními omezeními, tj. klíčovou roli hraje pozitivní semidefinitnost a definitnost Hessovy matice na podprostoru určeném množinou $\operatorname{Ker} \operatorname{grad}^\top f(x^*)$.

[20] Tj. nezápornost může být požadována pouze po některých proměnných.

[21] První dva požadavky lze formulovat také takto: platí $\frac{\partial f}{\partial x_i}(x^*) \geq 0$ pro všechna $i \in I$ a zároveň $\frac{\partial f}{\partial x_i}(x^*) = 0$ kdykoli $x_i^* > 0$ pro $i \in I$.

[22] V tomto případě je množina X *hyperobdélník* a podmínky by bylo možné upravit tak, že pouze některé souřadnice budou muset patřit do předem daných intervalů.

[23] Množina X je zobecněním jednotkového $(n-1)$-simplexu, viz (2.4.1).

Věta 8.1.8 Nutná podmínka 2. řádu

Je-li $x^* \in X$ lokálním minimem funkce $f : X \to \mathbb{R}$ na konvexní množině $X \subseteq \mathbb{R}^n$ a $f \in C^2(X)$, pak

$$(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) \geq 0 \tag{8.1.6}$$

pro všechna $x \in X$ taková, že $\langle \operatorname{grad} f(x^*), x - x^* \rangle = 0$, tj. pro vektory $(x - x^*) \in \operatorname{Ker} \operatorname{grad}^\top f(x^*)$.

Důkaz. Nechť $x \in X$ je takové, že $\langle \operatorname{grad} f(x^*), x - x^* \rangle = 0$. Pak s využitím Taylorova rozvoje druhého řádu dostaneme pro $\lambda \in [0,1]$ a nějaké $\tilde{\alpha} \in [0,1]$, že

$$f\big(\underbrace{x^* + \lambda(x - x^*)}_{\in X}\big) = f(x^*) + \underbrace{\langle \operatorname{grad} f(x^*), \overbrace{x^* + \lambda(x - x^*)}^{\lambda(x-x^*)} - x^* \rangle}_{=0} + \\ + \frac{1}{2}\lambda^2 (x - x^*)^\top \nabla^2 f\,\underbrace{\big(x^* + \tilde{\alpha}(x - x^*)\big)}_{x^* + \alpha(x-x^*)\ \&\ \alpha \in [0,\lambda]} (x - x^*),$$

tj.

$$f\big(x^* + \lambda(x - x^*)\big) - f(x^*) = \frac{1}{2}\lambda^2 (x - x^*)^\top \nabla^2 f\big(x^* + \alpha(x - x^*)\big)(x - x^*)$$

neboli

$$\frac{f\big(x^* + \lambda(x - x^*)\big) - f(x^*)}{\lambda^2} = \frac{1}{2}(x - x^*)^\top \nabla^2 f\big(x^* + \alpha(x - x^*)\big)(x - x^*).$$

Čitatel na levé straně je nezáporný, neboť x^* je lokálním minimem. Proto limitním přechodem pro $\lambda \to 0^+$ bude i levá strana nezáporná a současně $\alpha \to 0^+$, čímž obdržíme nerovnost (8.1.6). ■

Nerovnost (8.1.6) nabízí první eliminační kritérium pro stacionární body. Pokud se nám podaří najít alespoň jedno $x \in X$, pro které je tato nerovnost porušena, nemůže být uvažovaný stacionární bod x^* lokálním minimem. Avšak ne každý stacionární bod splňující (8.1.6) musí být lokálním minimem. K takovému závěru můžeme dospět až např. pomocí následujícího tvrzení.

Věta 8.1.9 Postačující podmínka 2. řádu

Bod $x^* \in X$ je lokálním minimem funkce $f : X \to \mathbb{R}$ na konvexní množině $X \subseteq \mathbb{R}^n$ pro $f \in C^2(X)$, jestliže

$$\operatorname{grad}^\top f(x^*)\,(x - x^*) \geq 0 \quad \text{pro všechna } x \in X$$

(tj. je to stacionární bod), množina X je *polyedr* a platí

$$(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) > 0 \tag{8.1.7}$$

pro všechna $x \in X \setminus \{x^*\}$ splňující $(x - x^*) \in \operatorname{Ker} \operatorname{grad}^\top f(x^*)$.

Poznámka 8.1.10. Proč nestačí konvexnost množiny X a vyžadujeme, aby byla dokonce polyedrem? V takovém případě je totiž pro libovolné $x^* \in X$ množina všech přípustných směrů[24]

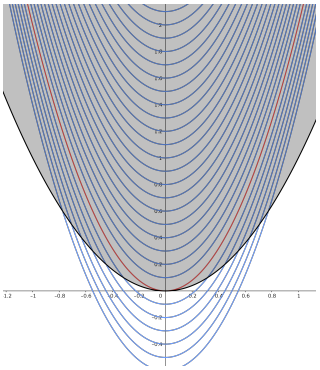
$$\mathcal{V}(x^*, X) = \{h \in \mathbb{R}^n \mid h = \lambda(x - x^*),\ \lambda \geq 0,\ x \in X\} \tag{8.1.8}$$

uzavřená, což je jeden z klíčových bodů celého důkazu. Uvažme např. $X = \{[x_1, x_2] \in \mathbb{R}^2 \mid x_1^2 \leq x_2\}$ a funkci $f(x_1, x_2) = -2x_1^2 + x_2$. Pak X jistě není polyedr a bod $x^* = [0,0] \in X$ splňuje obě podmínky Věty 8.1.9, neboť $\operatorname{grad} f(x) = (-4x_1, 1)^\top$ dává $\operatorname{grad} f(x^*) = (0,1)^\top$ a

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle = \langle (0,1)^\top, (x_1, x_2)^\top \rangle = x_2 > 0$$

pro libovolné $[x_1, x_2] \in X \setminus \{[0,0]\}$, tj. je to stacionární bod a současně podmínka (8.1.7) je splněna triviálně navzdory negativní semidefinitnosti matice $\nabla^2 f(x^*) = \left(\begin{smallmatrix} -4 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, protože neexistuje $x \in X \setminus \{x^*\}$ splňující $\langle \operatorname{grad} f(x^*), x - x^* \rangle = 0$. Ovšem bod x^* není lokálním minimem funkce f

[24] Zřejmé? $x^* + th = x^* + t\lambda(x - x^*) = t\lambda x + (1 - t\lambda)x^* \in X$ pro $t \in [0, 1/\lambda]$.



Obrázek 8.11: Ilustrace řešení příkladu z Poznámky 8.1.10. Vrstevnice procházející počátkem je červenou barvou a z obrázku je patrné, že tento bod nemůže být řešením.

na X. Vskutku, uvažme body $x_n := [1/n, 1/n^2]$. Pak $x_n \in X$ pro všechna $n \in \mathbb{N}$ a $x_n \to x^*$ pro $n \to \infty$, tj. pro libovolné $\varepsilon > 0$ existuje index $n_\varepsilon \in \mathbb{N}$ takový, že $\| x_n - x^* \| < \varepsilon$ pro všechna $n \geq n_\varepsilon$. Jenže $f(x_n) = -1/n^2 < 0 = f(x^*)$, takže v libovolném ε-okolí bodu x^* nalezneme body x_n, pro která má funkce f menší funkční hodnotu, tj. x^* nemůže být lokálním minimem f na X, viz Obrázek 8.11. Důvodem je právě skutečnost, že v tomto případě množina $\mathscr{V}(x^*, X)$ je otevřená, neboť

$$\mathscr{V}(x^*, X) = \left\{ h = (h_1, h_2)^\top \in \mathbb{R}^2 \mid h_2 > 0,\ \| h \| = 1 \right\}.$$

Např. oba vektory $(1,0)^\top$ a $(-1,0)^\top$ patří do $\overline{\mathscr{V}(x^*, X)}$, avšak nepatří do $\mathscr{V}(x^*, X)$, viz také Větu 8.1.11(i). ▲

Důkaz Věty 8.1.9. Tvrzení dokážeme sporem. Připusťme, že jsou splněny podmínky tvrzení, ale bod x^* není lokálním minimem funkce f na X. Pak existuje posloupnost $\{x^{[k]}\}_{k=1}^\infty \subseteq X$ taková, že $x^{[k]} \to x^*$ pro $k \to \infty$ a současně $f(x^{[k]}) < f(x^*)$ pro všechna $k \in \mathbb{N}$. Podle Taylorova rozvoje druhého řádu máme

$$f(x^{[k]}) = f(x^*) + \langle \operatorname{grad} f(x^*), x^{[k]} - x^* \rangle + \frac{1}{2} (x^{[k]} - x^*)^\top \nabla^2 f(x^*) (x^{[k]} - x^*) + \omega(x^{[k]} - x^*), \quad (8.1.9)$$

kde $\lim_{x^{[k]} \to x^*} \frac{\omega(x^{[k]} - x^*)}{\| x^{[k]} - x^* \|^2} = 0$. Pro $k \in \mathbb{N}$ definujme vektory

$$p_k := \frac{x^{[k]} - x^*}{\| x^{[k]} - x^* \|}. \quad (8.1.10)$$

Protože $f(x^{[k]}) < f(x^*)$, dostáváme z (8.1.9) nerovnost

$$0 > \frac{f(x^{[k]}) - f(x^*)}{\| x^{[k]} - x^* \|} = \langle \operatorname{grad} f(x^*), p_k \rangle + \frac{1}{2} p_k^\top \nabla^2 f(x^*) p_k \times \| x^{[k]} - x^* \| + \frac{\omega(x^{[k]} - x^*)}{\| x^{[k]} - x^* \|}. \quad (8.1.11)$$

To dohromady s předpokladem $\langle \operatorname{grad} f(x^*), p_k \rangle \geq 0$ (neboť x^* je stacionární bod) dává

$$\frac{1}{2} p_k^\top \nabla^2 f(x^*) p_k \times \| x^{[k]} - x^* \| + \frac{\omega(x^{[k]} - x^*)}{\| x^{[k]} - x^* \|} < 0 \quad (8.1.12)$$

Posloupnost $\{p_k\}_{k=1}^\infty$ je posloupností přípustných směrů v bodě x^*, tj. $p_k \in \mathscr{V}(x^*, X)$ pro všechna $n \in \mathbb{N}$, neboť $x^* + \alpha p_k = \left(1 - \alpha / \| x^{[k]} - x^* \|\right) x^* + \frac{\alpha}{\| x^{[k]} - x^* \|} x^{[k]} \in X$ pro $\alpha \in \left[0, \| x^{[k]} - x^* \|\right]$ díky konvexnosti X. Současně p_k leží leží na povrchu jednotkové koule (neboť $\| p_k \| = 1$) se středem v x^*. Proto existuje konvergentní podposloupnost $\{p_{k'}\}_{k=1}^\infty$, která konverguje k vektoru $\bar{p}$. Jelikož množina X je polyedr, je množina $\mathscr{V}(x^*, X)$ uzavřená (viz předchozí poznámka), a tudíž $\bar{p} \in \mathscr{V}(x^*, X)$, tj. $\bar{p}$ je také přípustným směrem množiny X v bodě x^*. Potom z předpokladů plyne, že $\langle \operatorname{grad} f(x^*), \bar{p} \rangle \geq 0$ a limitním přechodem pro $k' \to \infty$ dostaneme z (8.1.11) dokonce, že

$$\langle \operatorname{grad} f(x^*), \bar{p} \rangle = 0.$$

V takovém případě ale musí také platit, že

$$\bar{p}^\top \nabla^2 f(x^*) \bar{p} > 0. \quad (8.1.13)$$

Jenže po vydělení číslem $\| x^{[k]} - x^* \| > 0$ obou stran nerovnosti (8.1.12) dostaneme

$$\frac{1}{2} p_k^\top \nabla^2 f(x^*) p_k + \frac{\omega(x^{[k]} - x^*)}{\| x^{[k]} - x^* \|^2} < 0,$$

z čehož následným limitním přechodem pro $k' \to \infty$ získáme

$$\bar{p}^\top \nabla^2 f(x^*) \bar{p} \leq 0,$$

a to je spor s (8.1.13) ↯. ■

Ukážeme si ještě dvě další postačující podmínky, z nichž ta první je de facto zobecněním předchozího tvrzení.

VĚTA 8.1.11 DALŠÍ POSTAČUJÍCÍ PODMÍNKY 2. ŘÁDU

Bod $x^* \in X$ je lokálním minimem funkce $f : X \to \mathbb{R}$ na konvexní množině $X \subseteq \mathbb{R}^n$ pro $f \in C^2(X)$, jestliže

$$\operatorname{grad}^\top f(x^*)\,(x - x^*) \geq 0 \quad \text{pro všechna } x \in X \tag{8.1.14}$$

(tj. je to stacionární bod) a platí alespoň jedna z následujících podmínek:

(i) $h^\top \nabla^2 f(x^*)\, h > 0$ pro všechny nenulové vektory $h \in \overline{\mathcal{V}(x^*, X)} \setminus \{0\}$ splňující $\langle \operatorname{grad} f(x^*), h \rangle = 0$;

(ii) existuje $\gamma > 0$ takové, že

$$(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) \geq \gamma \| x - x^* \|^2 \quad \text{pro všechna } x \in X. \tag{8.1.15}$$

V tomto případě je x^* dokonce ostré lokální minimum.

Důkaz. (i) Nechť jsou splněny předpoklady věty, platí podmínka (i) a připusťme, že bod x^* není lokálním maximem funkce f na X. Pak existuje posloupnost $\{x^{[k]}\}_{k=1}^\infty \subseteq X$ taková, že $x^{[k]} \to x^*$ pro $k \to \infty$ a současně $f(x^{[k]}) < f(x^*)$ pro všechna $k \in \mathbb{N}$. S využitím Taylorova rozvoje druhého řádu v bodě x^* a vektorů p z (8.1.10) obdržíme, že pro všechna $k \in \mathbb{N}$ opět platí (8.1.11) a (8.1.12). Protože stále všechna p_k jsou přípustné směry v x^* ležící na povrchu jednotkové koule, existuje konvergentní podposloupnost $\{p_{k'}\}_{k=1}^\infty$, která konverguje k vektoru $\bar{p} \in \overline{\mathcal{V}(x^*, X)}$. Protože dle předpokladů máme $\langle \operatorname{grad} f(x^*), p_k \rangle \geq 0$ pro všechna $k \in \mathbb{N}$, musí platit také $\langle \operatorname{grad} f(x^*), \bar{p} \rangle \geq 0$. Avšak pro $k' \to \infty$ v (8.1.11) dostaneme také $\langle \operatorname{grad} f(x^*), \bar{p} \rangle \leq 0$, takže $\langle \operatorname{grad} f(x^*), \bar{p} \rangle = 0$. Proto podmínka (i) implikuje $\bar{p}^\top \nabla^2 f(x^*)\, \bar{p} > 0$. Jenže vydělením (8.1.12) výrazem $\| x^{[k']} - x^* \| > 0$ a limitním přechodem pro $k' \to \infty$ dostaneme $\bar{p}^\top \nabla^2 f(x^*)\, \bar{p} \leq 0$, což je spor ↯.

(ii) Z Taylorova rozvoje druhého řádu

$$f(x) = f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) + \omega(x - x^*)$$

pro libovolné $x \in X$ společně s nerovnostmi (8.1.14) a (8.1.15) dostáváme

$$f(x) - f(x^*) \geq \frac{\gamma}{2} \| x - x^* \|^2 + \omega(x - x^*) = \| x - x^* \|^2 \times \left(\frac{\gamma}{2} + \frac{\omega(x - x^*)}{\| x - x^* \|^2} \right).$$

Protože $\lim_{x \to x^*} \frac{\omega(x - x^*)}{\| x - x^* \|^2} = 0$, je výraz na pravé straně nezáporný pro $x \in X$ dostatečně blízko x^*, a je dokonce kladný pro $x \neq x^*$. Proto x^* je ostré lokální minimum f na X. ■

8.2 Nutné a postačující podmínky optimality v matematickém programování

V předchozí části jsme viděli, že při řešení obecné úlohy (8.1) hraje velmi důležitou roli geometrie množiny X (potažmo vektory $x - x^*$). Pro to, abychom mohli být v naší analýze konkrétnější, musíme určit množinu X blíže. Zejména úloze

$$f(x) \to \min, \quad x \in \{x \in \mathbb{R}^n \mid g_i(x) = 0,\ i = 1, \ldots, m\} \tag{8.2.1}$$

jsme se za dodatečného požadavku lineární nezávislosti gradientů omezujících funkcí již věnovali v Kapitole 1, viz úlohu (1.2.5) s nutnými a postačujícími podmínkami uvedenými ve Větách 1.2.16 a 1.2.18. Klíčovou roli zde hrála tzv. (*regulární*) *Lagrangeova funkce*

$$L(x, y) = f(x) + \sum_{i=1}^m y_i\, g_i(x),$$

kde jsme si tentokrát označili *Lagrangeovy multiplikátory* jako $y_1, \ldots, y_m$ místo tradičně užívaných $\lambda_1, \ldots, \lambda_m$. Této změny ve značení se budeme držet v celé této kapitole a výhodné to bude obzvláště v nadcházející podkapitole[25].

[25] Však je to jenom nějaké označení a „nic“ nám nebrání v tom použít jakýkoli symbol uznáme za vhodné. . . V knize [177, str. 17-7] slavného amerického teoretického fyzika Richarda Feynmana najdeme v této souvislosti zajímavý komentář: *We could, of course, use any notation we want; do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations.*
(*Samozřejmě bychom mohli použít jakékoli značení, které chceme; nesmějme se značení; vynalézejme jej, je vlivné. Ve skutečnosti je matematika do značné míry vynalézáním lepšího značení.*)

My se budeme dále zabývat úlohou (8.1) s množinou X popsanou v (8.2) a budeme postupovat podobně jako v předchozím případě. Tentokrát k této úloze přidružíme (obecnou) *Lagrangeovu funkci* $L : P \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ danou předpisem

$$L(x, y_0, y) := y_0 f(x) + \sum_{i=1}^{m} y_i g_i(x), \tag{8.2.2}$$

přičemž v případě $y_0 = 1$ budeme psát pouze $L(x, y)$, tj. $L(x, y) := L(x, 1, y)$. Čísla $y_0, \dots, y_m$ opět nazýváme *Lagrangeovými multiplikátory* a pro další výklad ještě budeme potřebovat následující množiny

$$Q := \{y = (y_1, \dots, y_m)^\top \in \mathbb{R}^m \mid y_1, \dots, y_k \geq 0\},$$
$$I(x^*) := \{i \in \{1, \dots, k\} \mid g_i(x^*) = 0\}, \quad x^* \in X,$$
$$S(x^*) := I(x^*) \cup \{k+1, \dots, m\}, \quad x^* \in X.$$

Množina Q obsahuje všechny možné vektory tvořené Lagrangeovými multiplikátory $y_1, \dots, y_m$ vhodnými pro naši úlohu. Množina $I(x^*)$ značí množinu *aktivních omezení* v bodě x^* a množina $S(x^*)$ pak je tvořena indexy všech funkcí z funkcionálních omezení určujících množinu X, které se v bodě x^* realizují jako rovnosti.

Následující tvrzení je zobecněním Věty 1.2.16 pro případ smíšených omezení, tj. nutnou podmínkou pro lokální řešení úlohy (8.1) a (8.2).

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_k : P \to \mathbb{R}$ jsou diferencovatelné v bodě $x^* \in X$ a $g_{k+1}, \dots, g_m : P \to \mathbb{R}$ jsou spojitě diferencovatelné v nějakém okolí bodu x^*. Je-li bod $x^* \in X$ lokálním řešením úlohy (8.1) a (8.2), pak existují Lagrangeovy multiplikátory $y_0^*, y_1^*, \dots, y_k^* \geq 0$ a $y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$, tj. $y_0^* \geq 0$ a $y^* \in Q$, takové, že ne všechny $y_0^*, \dots, y_m^*$ jsou nulová, tj. $(y_0^*, y^{*\top}) \neq 0$, a platí

$$\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), x - x^* \rangle \geq 0 \quad \text{pro všechna } x \in P, \tag{8.2.3}$$
$$y_i^* g_i(x^*) = 0 \quad \text{pro všechna } i = 1, \dots, m. \tag{8.2.4}$$

Podmínka (8.2.3) vlastně znamená, že x^* je *stacionárním bodem* funkce $L(x, y_0^*, y^*)$ na P (tj. pouze vzhledem k x), viz Definici 8.1.2. Podmínka (8.2.4) se nazývá *podmínkou komplementarity*, neboť nulovost multiplikátoru a aktivnost omezení jsou komplementární (vzájemně se doplňují), tj. buď $y_i^* = 0$ nebo $g_i(x^*) = 0$, přičemž obě rovnosti mohou nastat současně (což vypovídá o jisté *degenerovanosti* bodu x^*). Požadavku $y_1^*, \dots, y_k^* \geq 0$ se říká *podmínka duality*, neboť v případě maximalizace funkce f na množině X se znaménka těchto multiplikátorů změní na opačná[27] . Podmínky (8.2.3)+(8.2.4) bývají v literatuře souhrnně označovány jako *Johnovy podmínky*[28]. Dokážeme-li s pomocí dodatečných podmínek zaručit, že $y_0^* = 1$, pak se jedná o tzv. *Karushovy–Kuhnovy–Tuckerovy podmínky*[29].

Důkaz Věty 8.2.1. Pro začátek se nejdříve omezíme na situaci, kdy funkce $g_{k+1}, \dots, g_m$ jsou afinní. Nechť $x^* \in X$ je lokální řešení úlohy (8.1) a (8.2) a uvažme systém nerovností a rovností na množině P ve tvaru

$$\left.\begin{aligned} \langle \operatorname{grad} f(x^*), x - x^* \rangle &< 0, \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &< 0, \quad i \in I(x^*), \\ \langle \operatorname{grad} g_j(x^*), x - x^* \rangle &= 0, \quad j = k+1, \dots, m. \end{aligned}\right\} \tag{8.2.5}$$

Předpokládejme, že existuje řešení $\tilde{x} \in P$ tohoto systému, a označme $h := \tilde{x} - x^*$. Potom $x^* + th = t\tilde{x} + (1-t)x^* \in P$ pro libovolné $t \in [0, 1]$, neboť P je konvexní. Ukážeme, že $h \in \mathcal{U}(x^*, f)$ a současně $h \in \mathcal{V}(x^*, X)$:

(i) Jelikož funkce f je diferencovatelná v bodě x^*, může využít Taylorovu větu, čímž dostaneme

$$f(x^* + th) = f(x^*) + \langle \operatorname{grad} f(x^*), th \rangle + \omega(th).$$

[26] *Lagrange Multiplier Rule*

[27] Uvážíme-li maximalizační problém

$$F(x) \to \max, \quad x \in X, \qquad (*)$$

kterou volbou $f := -F$ snadno převedeme na úlohu (8.1) a (8.2). S využitím Věty 8.2.1 pak zjistíme, že bod x^* může být lokálním maximem úlohy (∗) pouze v případě, že existují multiplikátory $y_0^* \geq 0$, $y_1^*, \dots, y_k^* \leq 0$ a $y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$ takové, že ne všechny $y_0^*, \dots, y_m^*$ jsou nulové a pro $L(x, y_0, y) := y_0 F(x) + \sum_{i=1}^m y_i g_i(x)$ platí

$$\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), x - x^* \rangle \leq 0 \quad (\square)$$

pro všechna $x \in P$ a zároveň je splněna podmínka (8.2.4). Kdybychom místo úlohy (∗) měli problém

$$F(x) \to \max$$
$$x \in P \ \ \& \ \ h_i(x) \geq 0 \ \ \& \ \ h_j(x) = 0$$
$$i = 1, \dots, k, \ \ j = k+1, \dots, m,$$

pak bychom s využitím $f := -F$, $g_i := -h_i$ a $g_j := -h_j$ pro $i = 1, \dots, k$ a $j = k+1, \dots, m$ a Věty 8.2.1 zjistili, že bod x^* může být lokálním řešením pouze v případě, že existují multiplikátory $y_0^*, \dots, y_m^*$ stejné jako ve Větě 8.2.1 splňující nerovnost $(\square)$ pro všechna $x \in P$ a $L(x, y_0, y) := y_0 F(x) + \sum_{i=1}^m y_i h_i(x)$ a současně

$$y_i^* h_i(x^*) = 0$$

pro všechna $i = 1, \dots, m$.

Věta 8.2.1 Lagrangeův princip[26]

[28]

Fritz John (14. června 1910 – 10. února 1994) byl německý matematik žijící od roku 1935 v USA a věnující se zejména parciálním diferenciálním rovnicí a tzv. „neúplným problémům“. V roce 1984 se stal držitelem ocenění *MacArthur Fellow* (známého neoficiálně také jako „Genius Grant“), čímž se dostal do elitní společnosti čítající v tuto chvíli cca 30 dalších matematiků a stovky dalších skvělých vědců.

Odtud jednoduchou úpravou získáme

$$\underbrace{\frac{f(x^*+th)-f(x^*)}{t}}_{\to f_{h^+}(x^*)\text{ pro }t\to 0^+} = \underbrace{\langle \operatorname{grad} f(x^*), h\rangle}_{\overset{(8.2.5)(i)}{<}0} + \underbrace{\frac{\omega(th)}{t\|h\|}\times\|h\|}_{\to 0\text{ pro }t\to 0^+} < 0,$$

což pro $t\to 0^+$ dává $f_{h^+}(x^*)<0$, tj. funkce f je ve směru vektoru h klesající, tudíž platí $h\in\mathcal{U}(x^*,f)$.

(ii) Podobně pro $i\in I(x^*)$ plyne ze druhé nerovnosti v (8.2.5), že funkce g_i jsou klesající ve směru h. Jelikož současně $g_i(x^*)=0$, plyne odtud $g_i(x^*+th)\le 0$ pro dostatečně malé $t>0$, tj. $h\in\mathcal{U}(x^*,g_i)$ pro každé $i\in I(x^*)$. Poněvadž pro $i\in\{1,\dots,k\}\setminus I(x^*)$ máme $g_i(x^*)<0$, je v tomto případě $g_i(x^*+th)\le 0$ pro $|t|$ dostatečně malé, takže celkem $h\in\mathcal{U}(x^*,g_i)$ pro $i\in\{1,\dots,k\}$. Pro $i\in\{k+1,\dots,m\}$ využijeme předpoklad afinnosti a třetí rovnost v systému (8.2.5), tj.[30]

$$g_i(x^*+th)=g_i(x^*)+t\underbrace{\langle \operatorname{grad} g_i(x^*),h\rangle}_{\overset{(8.2.5)(iii)}{=}0}=g_i(x^*)=0$$

pro libovolné t. Celkově odtud vidíme, že pro dostatečně malé t je $x^*+th\in P$ a současně v tomto bodě platí všechna daná funkcionální omezení, tj. $x^*+th\in X$, a tudíž $h\in\mathcal{V}(x^*,X)$. Nalezli jsme tedy vektor $h\in\mathcal{V}(x^*,X)\cap\mathcal{U}(x^*,f)\neq\emptyset$, což je spor s Lemma 8.1.1 ↯.

Proto systém (8.2.5) nemá řešení na P a z Věty 6.3.21 plyne existence čísel $y_0^*,y_i^*\ge 0$ pro $i\in I(x^*)$ a $y_{k+1}^*,\dots,y_m^*\in\mathbb{R}$ takových, že

$$y_0^*\left\langle \operatorname{grad} f(x^*),x-x^*\right\rangle+\sum_{i\in S(x^*)}y_i^*\left\langle \operatorname{grad} g_i(x^*),x-x^*\right\rangle\ge 0,$$

tj.

$$\left\langle y_0^*\operatorname{grad} f(x^*)+\sum_{i\in S(x^*)}y_i^*\operatorname{grad} g_i(x^*),x-x^*\right\rangle\ge 0\quad\text{pro všechna } x\in P.$$

Jestliže ještě dodefinujeme $y_i^*:=0$ pro $i\in\{1,\dots,k\}\setminus I(x^*)$, pak z poslední nerovnosti dostáváme podmínku (8.2.3), tj.

$$\left\langle y_0^*\operatorname{grad} f(x^*)+\sum_{i=1}^m y_i^*\operatorname{grad} g_i(x^*),x-x^*\right\rangle\ge 0\quad\text{pro všechna } x\in P,$$

zatímco podmínka (8.2.4) je v takovém případě splněna triviálně.

Důkaz tvrzení v obecném případě je založen na Větě o implicitní funkci[31] a tzv. Ljustrenikovo lemma[32]. Nechť $\mathcal{L}:=\operatorname{Lin}P$ a vektory $a_1,\dots,a_s\in\mathbb{R}^n$ tvoří bázi ortogonálního doplňku $\mathcal{L}^\perp$, tj.

$$\mathcal{L}=\{h\in\mathbb{R}^n\mid\langle a_j,h\rangle\text{ pro všechna }j=1,\dots,s\}.\tag{8.2.6}$$

Připusťme, že vektory $a_1,\dots,a_s,\operatorname{grad}g_{k+1}(x^*),\dots,\operatorname{grad}g_m(x^*)$ jsou lineárně závislé, tj.

$$\sum_{j=k+1}^m y_j^*\operatorname{grad}g_j(x^*)+\sum_{i=1}^s\lambda_i a_i=0\tag{8.2.7}$$

pro vhodná čísla $y_{k+1}^*,\dots,y_m^*,\lambda_1,\dots,\lambda_s\in\mathbb{R}$, která nejsou všechna současně nulová. Zároveň však ani $y_{k+1}^*,\dots,y_m^*$ nemohou být všechna nulová, protože by to bylo ve sporu s lineární nezávislostí $a_1,\dots,a_s$. Pak z (8.2.6)–(8.2.7) pro libovolné $h\in\mathcal{L}$ máme

$$\left\langle\sum_{j=k+1}^m y_j^*\operatorname{grad}g_j(x^*),h\right\rangle=-\left\langle\sum_{i=1}^s\lambda_i a_i,h\right\rangle=0.$$

Jelikož $x^*\in\mathcal{L}$, můžeme tento podprostor také vyjádřit jako $\mathcal{L}=\operatorname{aff}P-x^*$. Proto

$$\left\langle\sum_{j=k+1}^m y_j^*\operatorname{grad}g_j(x^*),x-x^*\right\rangle=0$$

[29] Kuhn s Tuckerem publikovali v roce 1951 publikovali tyto podmínky jako nutné pro existenci optimálního řešení v nelineárním programování. Ovšem tato věta byla dokázána již v roce 1939 americkým matematikem Williamem Karushem (1. března 1947 – 22. února 1997) v jeho diplomové práci. John publikoval velmi podobný výsledek v roce 1948 (článek by vyšel dříve, ale rukopis byl odmítnut v prestižním časopise *Duke Mathematical Journal*). Pro více podrobností viz např. [322].

[30] Protože g_i je afinní, můžeme psát $g_i(x)=\langle c,x\rangle+b$ pro nějaké $c\in\mathbb{R}^n$ a $b\in\mathbb{R}$, takže $g_i(x+th)=\langle c,x+th\rangle+b=\langle c,x\rangle+b+t\langle c,h\rangle=g_i(x)+t\langle c,h\rangle$, přičemž $a=\operatorname{grad}g_i(x)$.

[31] Větu o implicitní funkci jsme si připomenuli již v Poznámce 78 na straně 40. Pro naše současné potřeby nám bude postačovat její následující zjednodušená verze: *Nechť funkce* $f_1(r,\alpha),\dots,f_n(r,\alpha):\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ *jsou spojitě diferencovatelné v okolí počátku, přičemž*

$$f_i(0,0)=0$$

pro všechna $i\in\{1,\dots,n\}$ *a gradienty* $\operatorname{grad}_r f_1(0,0),\dots,\operatorname{grad}_r f_n(0,0)$ *jsou lineárně nezávislé a* $\frac{\partial f_i}{\partial\alpha}(0,0)=0$ *pro všechna* $i\in\{1,\dots,n\}$. *Pak existuje funkce* $r:\mathbb{R}\to\mathbb{R}^n$ *taková, že*

$$f_i\big(r(\alpha),\alpha\big)=0$$

pro všechna $i\in\{1,\dots,n\}$ *a pro všechna dostatečně malá* α *s Jacobiho maticí*

$$\mathrm{D}_\alpha r(0)=\lim_{\alpha\to 0}\frac{r(\alpha)}{\alpha}=0\in\mathbb{R}^n.$$

[32] **Ljustrenikovo lemma.** *Nechť funkce $g_1(x),\dots,g_m(x)$ jsou spojitě diferencovatelné v nějakém okolí $x^* \in \mathbb{R}^n$, přičemž*

$$g_i(x^*) = 0$$

pro všechna $i \in \{1,\dots,m\}$ a vektory $\operatorname{grad} g_1(x^*),\dots,\operatorname{grad} g_m(x^*)$ společně s některými vektory $a_{m+1},\dots,a_n \in \mathbb{R}^n$ tvoří bázi $\mathbb{R}^n$. Předpokládejme, že vektor $h \in \mathbb{R}^n$ splňuje podmínku

$$\langle \operatorname{grad} g_i(x^*), h\rangle = 0$$

pro všechna $i \in \{1,\dots,m\}$. Pak existuje funkce $r:\mathbb{R}\to\mathbb{R}^n$ taková, že

$$g_i(x^* + \alpha h + r(\alpha)) = 0$$

pro všechna $i \in \{1,\dots,m\}$ a současně

$$\langle a_i, r(\alpha)\rangle = 0$$

pro všechna $i \in \{m+1,\dots,n\}$ a pro všechna dostatečně malá α s Jacobiho maticí

$$D_\alpha r(0) = \lim_{\alpha\to 0}\frac{r(\alpha)}{\alpha} = 0 \in \mathbb{R}^n.$$

Důkaz. Definujme funkce

$$f_i(r,\alpha) = g_i(x^* + \alpha h + r)$$

pro všechna $i \in \{1,\dots,m\}$ a současně

$$f_i(r,\alpha) = \langle a_i, r\rangle$$

pro všechna $i \in \{m+1,\dots,n\}$. Pak

$$\operatorname{grad}_r f_i(0,0) = \operatorname{grad} g_i(x^*)$$

a

$$\frac{\partial f_i}{\partial \alpha}(0,0) = \langle \operatorname{grad} g_i(x^*), h\rangle$$

pro všechna $i \in \{1,\dots,m\}$ a

$$\operatorname{grad}_r f_i(0,0) = a_i \quad \& \quad \frac{\partial f_i}{\partial \alpha}(0,0) = 0$$

pro všechna $i \in \{m+1,\dots,n\}$. Odtud vidíme, že jsou splněny předpoklady Věty o implicitní funkci, ze které plyne uvedené tvrzení.

[33] Jakožto součet $\widehat{x} \in \operatorname{ri} P \subseteq P$ a nenulového násobku $r(\alpha) \in \operatorname{Lin} P$.

pro libovolné $x \in \operatorname{aff} P$, a tudíž zejména pro všechna $x \in P$. Odtud plyne, že zvolíme-li $y_0^* := 0, \dots, y_k^* := 0$, pak celá m-tice $y_0^*,\dots,y_m^*$ zjevně splňuje obě podmínky (8.2.3)–(8.2.4) v tomto degenerovaném případě.

Konečně uvažme situaci, kdy $a_1,\dots,a_s,\operatorname{grad} g_{k+1}(x^*),\dots,\operatorname{grad} g_m(x^*)$ jsou lineárně nezávislé. Bez újmy na obecnosti můžeme předpokládat, že tvoří bázi $\mathbb{R}^n$, protože v opačném případě bychom k nim mohli chybějící vektory přidat. Připusťme, že systém (8.2.5) má řešení $\widehat{x} \in \operatorname{ri} P$. Položme $h := \widehat{x} - x^*$. Pak třetí část (8.2.5) zaručuje splnění předpokladů Ljustrenikova lemma, a tudíž existuje funkce $r:\mathbb{R}\to\mathbb{R}^n$ taková, že

$$g_i(x^* + \alpha h + r(\alpha)) = 0 \quad \text{pro všechna } i = k+1,\dots,m, \tag{8.2.8}$$

$$\langle a_j, r(\alpha)\rangle = 0 \quad \text{pro všechna } j = 1,\dots,s \tag{8.2.9}$$

pro všechna dostatečně malá α a

$$\lim_{\alpha\to 0}\frac{r(\alpha)}{\alpha} = 0. \tag{8.2.10}$$

Položme $x(\alpha) := x^* + \alpha h + r(\alpha)$. Z diferencovatelnosti funkce f v x^* plyne

$$\begin{aligned} f(x(\alpha)) - f(x^*) &= \langle \operatorname{grad} f(x^*), \alpha h + r(\alpha)\rangle + \omega(\alpha h + r(\alpha)) = \\ &= \alpha\Big(\langle \operatorname{grad} f(x^*), h\rangle + \langle \operatorname{grad} f(x^*), r(\alpha)/\alpha\rangle + \omega(\alpha h + r(\alpha))/\alpha\Big) = \\ &= \alpha\left(\langle \operatorname{grad} f(x^*), h\rangle + \langle \operatorname{grad} f(x^*), r(\alpha)/\alpha\rangle + \|h + r(\alpha)/\alpha\|\,\frac{\omega(\alpha h + r(\alpha))}{\alpha\|h + r(\alpha)/\alpha\|}\right) \end{aligned}$$

pro nějakou funkci $\omega \in o(\|\alpha h + r(\alpha)\|)$, tj. $\lim_{\alpha\to 0}\frac{\omega(\alpha h + r(\alpha))}{\|\alpha h + r(\alpha)\|} = 0$. Pak z první nerovnosti v (8.2.5) a (8.2.10) získáme $f(x(\alpha)) < f(x^*)$ pro všechna dostatečně malá $\alpha > 0$. Podobným způsobem, ale tentokrát za využití druhé části (8.2.5), dostaneme pro všechna $i \in I(x^*)$ a $\alpha > 0$ dostatečně malá, že také $g_i(x(\alpha)) < g(x^*) = 0$. Pro $i \in \{1,\dots,k\}\setminus I(x^*)$ je $g(x^*) < 0$, takže i $g(x(\alpha)) < 0$ pro všechna $\alpha > 0$ dostatečně malá. Z (8.2.6) a (8.2.9) plyne, že $r(\alpha) \in \mathscr{L} = \operatorname{Lin} P$. Nechť $\overline{x}(\alpha) := \widehat{x} + r(\alpha)/\alpha$. Pak[33] $\overline{x}(\alpha) \in \operatorname{aff} P$. Protože $\widehat{x} \in \operatorname{ri} P$ a $\overline{x}(\alpha) \to \widehat{x}$ pro $\alpha \to 0$ podle (8.2.10), máme pro všechna $\alpha > 0$ dostatečně malá $\overline{x}(\alpha) \in P$, viz Větu 6.1.30, a tudíž

$$x(\alpha) = x^* + \alpha(\widehat{x} - x^*) + r(\alpha) = \alpha\overline{x}(\alpha) + (1-\alpha)\,x^* \in P$$

díky konvexnosti P. Tím jsme ale ukázali, že $f(x(\alpha)) < f(x^*)$ a $x(\alpha) \in X$ pro všechna dostatečně malá $\alpha > 0$, přičemž $x(\alpha) \to x^*$ pro $\alpha \to 0$, tj. v každém okolí $\mathcal{O}(x^*)$ leží nějaké $x(\alpha)$. To je však spor s tím, že x^* je lokálním řešením problému (8.2.1). Proto systém (8.2.5) nemůže mít řešení na $\operatorname{ri} P$, a zbytek tvrzení vyplývá z Věty 6.3.21 jako v první části důkazu. ■

Poznámka 8.2.2. Podmínka (8.2.3) ve speciálních případech, viz dřívější Poznámku 8.1.7:

(i) Je-li $x^* \in \operatorname{int} P$ (tedy zejména je-li P otevřená, obzvláště $P = \mathbb{R}^n$), pak podmínka (8.2.3) může být splněna pouze v případě

$$\operatorname{grad}_x L(x^*, y_0^*, y^*) = 0. \tag{8.2.11}$$

(ii) Je-li

$$P = \{x \in \mathbb{R}^n \mid \alpha_i \le x_i \le \beta_i,\ i = 1,\dots,n\},$$

kde $-\infty \le \alpha_i < \beta_i \le \infty$ pro $i \in \{1,\dots,n\}$, pak podmínka (8.2.3) znamená

$$\frac{\partial L}{\partial x_i}(x^*, y_0^*, y^*)\begin{cases} = 0, & \alpha_i < x_i^* < \beta_i, \\ \ge 0, & x_i^* = \alpha_i \ne -\infty, \\ \le 0, & x_i^* = \beta_i \ne \infty. \end{cases}$$

(iii) Je-li

$$P = \{x \in \mathbb{R}^n \mid x_i \ge 0,\ i = 1,\dots,s\},$$

kde $s \in \{1,\dots,n\}$, pak (8.2.3) znamená

$$\frac{\partial L}{\partial x_i}(x^*,y_0^*,y^*) \geq 0 \quad \text{a současně} \quad x_i^* \frac{\partial L}{\partial x_i}(x^*,y_0^*,y^*) = 0 \quad \text{pro } i \in \{1,\dots,s\}$$

$$\text{a navíc} \quad \frac{\partial L}{\partial x_i}(x^*,y_0^*,y^*) = 0 \quad \text{pro } i \in \{s+1,\dots,n\}. \qquad \blacktriangle$$

Zvídavého čtenáře by mohla napadnout otázka, zda je Věta 8.2.1 konzistentní s výkladem pro úlohy s omezeními pouze ve tvaru rovností? Všimněme si, že na rozdíl od úlohy (8.2.1) zde vystupuje také multiplikátor y_0. To je způsobeno tím, že ve Větě 8.2.1 není žádný dodatečný požadavek, který by zaručoval existenci multiplikátorů s $y_0^* \neq 0$, tj. $y_0^* > 0$, což by pak snadno vedlo k multiplikátorům s $y_0^* = 1$ – toto je ve spojení s úlohou (8.2.1) ošetřeno požadavkem na lineární nezávislost vektorů $\operatorname{grad} g_1(x),\dots,\operatorname{grad} g_m(x)$, viz opět Větu 1.2.16. Následující příklad ukazuje, že skutečně pro úlohu (8.1) a (8.2) může nastat situace s $y_0^* = 0$, takže ji nelze „beztrestně“ opomenout.

Příklad 8.2.3

Uvažme úlohu s $P = \mathbb{R}^2$ a

$$f(x_1,x_2) = x_1 \to \min,$$

$$g_1(x_1,x_2) = -x_1^3 + x_2 \leq 0 \quad \& \quad g_2(x_1,x_2) = -x_1^3 - x_2 \leq 0 \quad \& \quad g_3(x_1,x_2) = x_1^2 + x_2^2 - 1 \leq 0.$$

Z Obrázku 8.12 vidíme, že jediným řešením je bod $x^* = [0,0]$. Zde platí

$$\operatorname{grad} f(x^*) = \begin{pmatrix}1\\0\end{pmatrix} \quad \& \quad \operatorname{grad} g_1(x^*) = \begin{pmatrix}0\\1\end{pmatrix} \quad \& \quad \operatorname{grad} g_2(x^*) = \begin{pmatrix}0\\-1\end{pmatrix} \quad \& \quad \operatorname{grad} g_3(x^*) = \begin{pmatrix}0\\0\end{pmatrix},$$

což jsou jistě lineárně závislé vektory. Omezení g_3 je v bodě x^* neaktivní, takže z (8.2.4) vyplývá, že nutně $y_3^* = 0$. Podmínka (8.2.3) pak je tvaru

$$y_0^* \begin{pmatrix}1\\0\end{pmatrix}^\top \begin{pmatrix}x_1-0\\x_2-0\end{pmatrix} + y_1^* \begin{pmatrix}0\\1\end{pmatrix}^\top \begin{pmatrix}x_1-0\\x_2-0\end{pmatrix} + y_2^* \begin{pmatrix}0\\-1\end{pmatrix}^\top \begin{pmatrix}x_1-0\\x_2-0\end{pmatrix} \geq 0$$

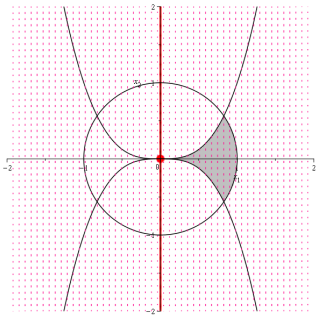
neboli

$$\left[y_0^* \begin{pmatrix}1\\0\end{pmatrix} + y_1^* \begin{pmatrix}0\\1\end{pmatrix} + y_2^* \begin{pmatrix}0\\-1\end{pmatrix} \right]^\top \begin{pmatrix}x_1\\x_2\end{pmatrix} \geq 0$$

pro $y_0^* \geq 0$, $y_1^* \geq 0$ a $y_2^* \geq 0$. Jelikož tato nerovnost má být splněna pro všechna $[x_1,x_2] \in \mathbb{R}^2$, dostáváme odtud rovnici

$$y_0^* \begin{pmatrix}1\\0\end{pmatrix} + y_1^* \begin{pmatrix}0\\1\end{pmatrix} + y_2^* \begin{pmatrix}0\\-1\end{pmatrix} = 0,$$

jejímž řešením jsou multiplikátory $y_0^* = 0$ a $y_1^* = y_2^* = \lambda$ pro libovolné $\lambda > 0$, což je zcela v souladu s tvrzením Věty 8.2.1. Jiná volba ani není možná, neboť vektory $\binom{0}{1}$ a $\binom{0}{-1}$ jsou lineárně závislé, zatímco dvojice vektorů $\binom{1}{0}$, $\binom{0}{1}$ či $\binom{1}{0}$, $\binom{0}{-1}$ jsou lineárně nezávislé[34].



Obrázek 8.12: Grafické znázornění úlohy z Příkladu 8.2.3. Kružnice a dvojice parabol (černá barva) odpovídají jednotlivým omezením, která vymezují přípustnou množinu vyznačenou šedou barvou. Vrstevnice funkce f jsou znázorněny červeně a s rostoucí funkční hodnotou se pohybují zleva doprava.

Tento příklad poukazuje na obecnější fakt, který plyne z Věty 8.2.1 a který jsme již zmínili v Poznámce 8.2.2(i), tj. v případě $x^* \in \operatorname{int} P$ vede (8.2.3) na rovnost

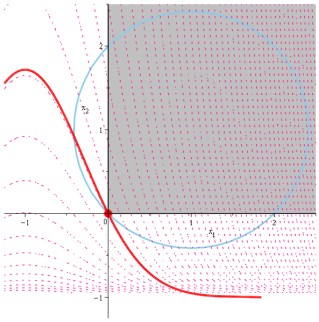
$$\operatorname{grad}_x L(x^*,y_0^*,y^*) = 0 \quad \text{neboli} \quad y_0^* \operatorname{grad} f(x^*) + \sum_{i=1}^m y_i^* \operatorname{grad} g_i(x^*) = 0.$$

Má-li být tato podmínka splněna v bodě x^* pro případ $y_0^* = 0$, musí být

$$0 = \sum_{i=1}^m y_i^* \operatorname{grad} g_i(x^*) = \sum_{i \in S(x^*)} y_i^* \operatorname{grad} g_i(x^*),$$

[34] K témuž výsledku bychom se dopracovali i s využitím Poznámky 8.2.2(i), která společně s podmínkou (8.2.4) vede pro Lagrangeovu funkci $L(x_1,x_2,y_0,y_1,y_2,y_3) = y_0 x_1 + y_1(-x_1^3 + x_2) + y_2(-x_1^3 - x_2) + y_3(x_1^2 + x_2^2 - 1)$ k hodnotě $y_3^* = 0$ a rovnicím $L_{x_1} = y_0 - 3x_1^2(y_1 - y_2) = 0$ a $L_{x_2} = y_1 - y_2 = 0$, jejichž řešení jsou $y_1^* = y_2^*$ a $y_0^* = 0$. Tento postup využijeme později při řešení dalších příkladů.

tj. gradienty funkcí zadávající aktivní omezení (včetně těch rovnostních) v bodě x^* musí být nutně lineárně závislé (jako se to stalo v předchozím příkladě). V takovém případě pak funkce f „nehraje žádnou roli“ v Lagrangeově funkci $L(x^*, 0, y^*)$ a určující je zde pouze geometrie množiny X! V předchozím příkladě byly všechny multiplikátory y_0^*, y_1^*, y_2^* a y_3^* určeny jednoznačně až na násobek kladným číslem, ovšem situace může být i „divočejší“, jak ukazuje následující příklad. Předložený postup řešení je založen jen na podmínkách samotné Věty 8.2.1 místo Poznámky 8.2.2(iii), proto omezení na znaménko považujeme za součást funkcionálních omezení. Alternativní (a současně i praktičtější bez hledání řešení pomocí obrázku) postup založený na zmíněné poznámce si ukážeme v pozdějších příkladech.



Obrázek 8.13: Grafické znázornění úlohy z Příkladu 8.2.4. Kružnice (modrá barva) odpovídají rovnostnímu omezením, takže přípustná množina je část jejího oblouku, která leží v šedě vyznačeném prvním kvadrantu. Vrstevnice funkce f jsou znázorněny červeně a s rostoucí funkční hodnotou se pohybují zdola nahoru.

Příklad 8.2.4

Uvažme úlohu

$$f(x_1, x_2) = (x_1 + 1)^2 + \ln(x_2 + 1) \to \min,$$
$$(x_1 - 1)^2 + (x_2 - 1)^2 = 2 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0.$$

Z Obrázku 8.13 vidíme, že jediným řešením je bod $x^* = [0,0]$. V tuto chvíli si můžeme zvolit, zda omezení na znaménko x_1 a x_2 budeme považovat za přímé, tj. zvolíme $P = \mathbb{R}^2_+$ a funkcionální omezení bude určené pouze funkcí $g_1(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2 - 2$, nebo zda jej zahrneme mezi funkcionální omezení a kromě g_1 budeme mít i $g_2(x_1, x_2) = -x_1$ a $g_3(x_1, x_2) = -x_2$ s tím, že $P = \mathbb{R}^2$. Zvolme si v tuto chvíli druhou možnost. Pak platí

$$\operatorname{grad} f(x^*) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \& \quad \operatorname{grad} g_1(x^*) = \begin{pmatrix} -2 \\ -2 \end{pmatrix} \quad \& \quad \operatorname{grad} g_2(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \& \quad \operatorname{grad} g_3(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Podmínka (8.2.3) má v bodě x^* podobu

$$y_0^* \begin{pmatrix} 2 \\ 1 \end{pmatrix}^\top \begin{pmatrix} x_1 - 0 \\ x_2 - 0 \end{pmatrix} + y_1^* \begin{pmatrix} -2 \\ -2 \end{pmatrix}^\top \begin{pmatrix} x_1 - 0 \\ x_2 - 0 \end{pmatrix} + y_2^* \begin{pmatrix} -1 \\ 0 \end{pmatrix} + y_3^* \begin{pmatrix} 0 \\ -1 \end{pmatrix}^\top \begin{pmatrix} x_1 - 0 \\ x_2 - 0 \end{pmatrix} \geq 0$$

neboli

$$\left[y_0^* \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y_1^* \begin{pmatrix} -2 \\ -2 \end{pmatrix} + y_2^* \begin{pmatrix} -1 \\ 0 \end{pmatrix} + y_3^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right]^\top \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \geq 0,$$

zatímco podmínka (8.2.4) je splněna triviálně, neboť všechna omezení jsou v bodě x^* aktivní. Tato nerovnost má být splněna pro všechna $[x_1, x_2] \in \mathbb{R}^2$, a tudíž nutně

$$y_0^* \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y_1^* \begin{pmatrix} -2 \\ -2 \end{pmatrix} + y_2^* \begin{pmatrix} -1 \\ 0 \end{pmatrix} + y_3^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0$$

pro $y_0^* \geq 0$, $y_1^* \in \mathbb{R}$, $y_2^* \geq 0$ a $y_3^* \geq 0$. Máme soustavu dvou rovnic o čtyřech neznámých, pro kterou snadno najdeme dokonce dvojici lineárně nezávislých řešení (tj. jedna čtveřice není nějakým násobkem té druhé)

$$\begin{pmatrix} y_0^* \\ y_1^* \\ y_2^* \\ y_3^* \end{pmatrix} = \begin{pmatrix} 0 \\ -\lambda \\ 2\lambda \\ 2\lambda \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} y_0^* \\ y_1^* \\ y_2^* \\ y_3^* \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \\ 2\lambda \\ \lambda \end{pmatrix}$$

pro libovolné $\lambda > 0$. To nicméně stále je zcela v souladu s tvrzením Věty 8.2.1.

S ohledem na skutečnost, že v případě $y_0^* = 0$ nehraje funkce f v Lagrangeově funkci žádnou roli, může být následné vyšetřování, zda v nalezeném bodě x^* skutečně nastává extrém, velmi problematické. Proto se nyní zaměříme na hledání podmínek, které zaručí, že $y_0^* \neq 0$ v řešení úlohy (8.1) a (8.2). To lze bez újmy na obecnosti převést na případ s $y_0^* = 1$, čímž z původních

Johnových podmínek (8.2.3) a (8.2.4) získáme Karushovy–Kuhnovy–Tuckerovy podmínky vyžadující existenci Lagrangeových multiplikátorů s $y_0^* = 1$ a $y^* \in Q$. O takových podmínkách se obvykle hovoří jako o těch, které nám zaručují *kvalifikované* (*kvalifikovanost*) *omezení*, případně jako o *podmínkách regularity*[35]. Jednoduchou úvahou zjistíme, že zde hraje důležitou roli např. *regulárnost* bodu x^*, tj. požadavek lineární nezávislosti vektorů $\operatorname{grad} g_i(x^*)$ pro $i \in S(x^*)$. Potom dostáváme následující tvrzení.

[35] *Constraint qualifications* (CQ) nebo *Regularity conditions*

Věta 8.2.5

Nechť jsou splněny předpoklady Věty 8.2.1 a $x^* \in \operatorname{int} P$ je lokálním řešením úlohy (8.1) a (8.2). Jestliže x^* je regulární bod, pak existují (jediné) multiplikátory $y^* \in Q$ takové, že platí (8.2.3) a (8.2.4) s $y_0^* = 1$.

Důkaz. Z Věty 8.2.1 plyne, že

$$y_0^* \operatorname{grad} f(x^*) + \sum_{i=1}^{m} y_i^* \operatorname{grad} g_i(x^*) = 0.$$

Připusťme, že $y_0^* = 0$. Pak díky $y_i^* = 0$ pro $i \in \{1,\dots,m\} \setminus S(x^*)$ musí být

$$\sum_{i=1}^{m} y_i^* \operatorname{grad} g_i(x^*) = \sum_{i \in S(x^*)} y_i^* \operatorname{grad} g_i(x^*) = 0,$$

což vzhledem k lineární nezávislosti vektorů $\operatorname{grad} g_i(x^*)$ pro $i \in S(x^*)$ implikuje, že také $y_i^* = 0$ pro $i \in S(x^*)$. Jenže to znamená, že $y_i^* = 0$ pro všechna $i \in \{0,\dots,m\}$, což je spor s požadavkem $(y_0^*, y^*) \neq 0$ z Věty 8.2.1 ↯. ■

[36] *Hladkou křivkou* rozumíme množinu bodů zadanou rovnicí $h(x) = 0$ pro spojitě diferencovatelnou funkci $h : \mathbb{R}^2 \to \mathbb{R}$, přičemž v každém bodě vyhovujícím této rovnici platí $\operatorname{grad} h(x) \neq 0$.

[37] Příklad pochází z [57, Example 3.1.4].

Příklad 8.2.6

(*Snellův zákon.*[40]) Uvažme hladkou křivku[41] v $\mathbb{R}^2$ určenou rovnicí $h(x) = 0$ pro spojitě diferencovatelnou funkci $h : \mathbb{R}^2 \to \mathbb{R}$, která rozděluje rovinu na dvě části. Nechť v každé z těchto částí je rychlost světa odlišná a body y a z leží na opačných stranách této křivky, jak je znázorněno na Obrázku 8.14. Připusťme, že světlo z bodu y jde do bodu z nejdříve s rychlostí v_y do bodu x^* ležící na dané křivce a odtud s rychlostí v_z do z, přičemž tento pohyb světla splňuje *Fermatův zákon*, tj. potřebný čas je minimální. Pak bod x^* je charakterizován Snellovým zákonem

$$\frac{\sin \vartheta_y}{v_y} = \frac{\sin \vartheta_z}{v_z}, \tag{8.2.12}$$

kde ϑ_y a ϑ_z jsou úhly vyznačené na Obrázku 8.14.
Vskutku, světlo urazí tuto dráhu za čas

$$T(x^*) = \frac{\|y - x^*\|}{v_y} + \frac{\|z - x^*\|}{v_z}$$

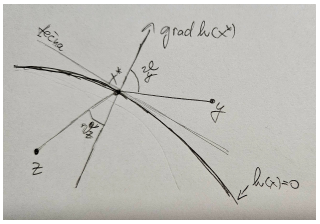
a z Lagrangeova principu obdržíme že gradient

$$\operatorname{grad} T(x^*) = \frac{x^* - y}{v_y \|x^* - y\|} + \frac{x^* - z}{v_z \|x^* - z\|} \tag{8.2.13}$$

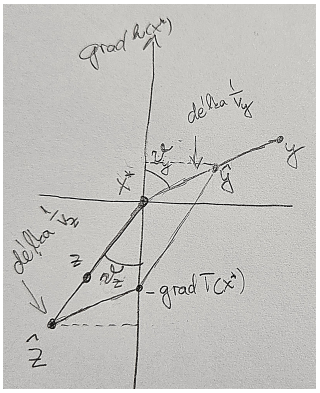
je skalárním násobkem $\operatorname{grad} h(x^*)$, k čemuž jsme využili regulárnost bodu x^*. K interpretaci této podmínky uvažme vektory

$$\widehat{y} := x^* + \frac{y - x^*}{v_y \|y - x^*\|} \quad \& \quad \widehat{z} := x^* + \frac{z - x^*}{v_z \|z - x^*\|}$$

a rovnoběžník se stranami $\widehat{y} - x^*$ a $\widehat{z} - x^*$, který má dle (8.2.13) $-\operatorname{grad} T(x^*)$ jako jednu ze svých úhlopříček. Z Obrázku 8.15 vidíme, že Lagrangeův princip vede k tomu, že tato diagonála je normálou k dané křivce v bodě x^* a současně vzdálenosti $\widehat{y}$ a $\widehat{z}$ od svislé diagonály si musí být rovny. Protože $\|\widehat{y} - x^*\| = 1/v_y$ a $\|\widehat{z} - x^*\| = 1/v_z$, dostáváme odtud Snellův zákon (8.2.12), viz také Cvičení 8.2.43.



Obrázek 8.14: Obrázek k Příkladu 8.2.6. Úhly ϑ_y a ϑ_z jsou určeny normálou ke křivce $h(x)$ v bodě x^*, tj. $\operatorname{grad} h(x^*)$, a vektory $y - x^*$ a $z - x^*$.



Obrázek 8.15: Obrázek k Příkladu 8.2.6. Vzdálenosti bodů $\widehat{y}$ a $\widehat{z}$ od svislé rovnoběžky si musí být rovny jakožto tatáž výška rovnoběžníku.

V situaci $P = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1,\dots,s\}$ pro nějaké $s \in \{0,\dots,n\}$ můžeme omezení z P zahrnout mezi funkcionální omezení jako $g_{m+1}(x) := -x_1, \dots, g_{m+s}(x) := -x_s$, v kterémžto případě regulárnost bodu $x \in X$ znamená, že gradienty $\operatorname{grad} g_1(x), \dots \operatorname{grad} g_{m+s}(x)$ odpovídající aktivním omezením jsou lineárně nezávislé, tj. nejvýše n daných omezení včetně omezení na znaménko je aktivních. Pokud toto dokážeme zaručit pro každé $x \in X$, pak pro řešení úlohy (8.1) a (8.2) musí existovat multiplikátory splňující (8.2.3) a (8.2.4) s $y_0^* = 1$. V takovém případě ale není nutné pracovat s $m+s$ multiplikátory. Vystačíme si s původními m-ticí a Lagrangeovou funkcí $L(x,y) = f(x) + \sum_{i=1}^m y_i\, g_i(x)$, když soustavu (8.2.3) a (8.2.4) upravíme do tvaru

$$\begin{aligned} L_{x_i}(x^*, y^*) &\geq 0 \quad \text{pro všechna } i = 1,\dots,s,\\ L_{x_i}(x^*, y^*) &= 0, \quad \text{je-li } i \in \{1,\dots,s\} \text{ takové, že } x_i^* > 0,\\ L_{x_i}(x^*, y^*) &= 0 \quad \text{pro všechna } i = s+1,\dots,m,\\ y_i^*\, g_i(x^*) &= 0 \quad \text{pro všechna } i = 1,\dots,m. \end{aligned}$$

Bude-li požadavek regulárnosti porušen v nějakém bodě $\widetilde{x} \in X$, pak stále lze použít postup výše s tím, že chování funkce f v bodě $\widetilde{x}$ budeme muset vyšetřit zvlášť.

V literatuře lze nalézt i několik dalších podmínek zaručující $y_0^* = 1$ při $x^* \in \operatorname{int} P$.

VĚTA 8.2.7

Nechť jsou splněny předpoklady Věty 8.2.1 a $x^* \in \operatorname{int} P$ je lokálním řešením úlohy (8.1) a (8.2). Potom existují (jediné) multiplikátory $y^* \in Q$ splňující (8.2.3) a (8.2.4) s $y_0^* = 1$, jestliže je splněna (alespoň) jedna z následujících podmínek

(i) (*afinní omezení*) funkce $g_1,\dots,g_m$ jsou afinní;

(ii) (*Slaterova podmínka*) funkce $g_1,\dots,g_k$ jsou konvexní, funkce $g_{k+1},\dots,g_m$ jsou afinní, konstantní vektory $\operatorname{grad} g_i$ jsou lineárně nezávislé pro $i \in \{k+1,\dots,m\}$ a existuje bod $\overline{x} \in P$ takový, že $g_i(\overline{x}) < 0$ pro $i \in \{1,\dots,k\}$ a $g_i(\overline{x}) = 0$ pro $i \in \{k+1,\dots,m\}$;

(iii) (*Mangasarianova–Fromovitzova podmínka*) vektory $\operatorname{grad} g_i(x^*)$ jsou lineárně nezávislé pro $i \in \{k+1,\dots,m\}$ a existuje vektor $h \in \mathbb{R}^n$ splňující

$$\begin{aligned} \langle \operatorname{grad} g_i(x^*), h\rangle &< 0 \quad \text{pro } i \in I(x^*),\\ \langle \operatorname{grad} g_i(x^*), h\rangle &= 0 \quad \text{pro } i \in \{k+1,\dots,m\}. \end{aligned}$$

Důkaz. Nechť jsou splněny předpoklady věty a platí podmínka (i). Pak místo omezení $g_i(x) \leq 0$ a $g_j(x) = 0$ pro $i = 1,\dots,k$ a $j = k+1,\dots,m$ můžeme psát

$$Ax \leq a \quad \text{a} \quad Bx = b$$

pro vhodné matice $A = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \in \mathbb{R}^{k\times n}$, $B \in \mathbb{R}^{(m-k)\times n}$ a vektory $a \in \mathbb{R}^k$, $b \in \mathbb{R}^{m-k}$. V takovém případě je množina $\mathcal{V}(x^*, X)$ shodná s

$$\{h \in \mathbb{R}^n \mid Bh = 0,\ a_i h \leq 0 \text{ pro } i \in I(x^*)\},$$

neboť opět $A(x^* + th) \leq a$ a $B(x^* + th) = Bx^* = b$ pro libovolné $t \geq 0$. Navíc pro každé $h \in \mathcal{V}(x^*, X)$ musí být $\langle \operatorname{grad} f(x^*), h\rangle \geq 0$, protože v opačném případě bychom dostali spor s lokálním minimem v x^* a Lemma 8.1.1. To znamená, že pro matici $A_{I(x^*)}$ tvořenou řádky matice A s indexy z množiny $I(x^*)$ nemá systém

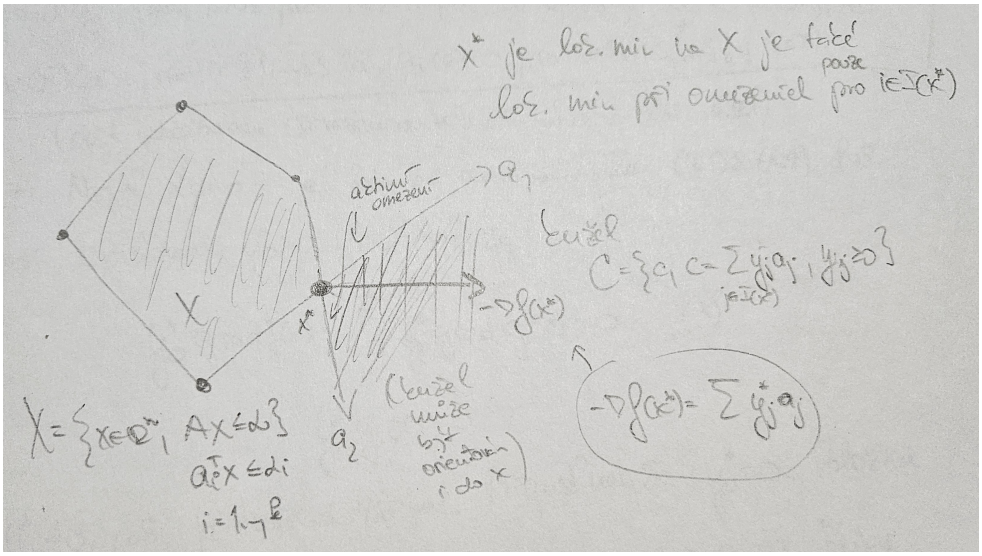
$$\begin{pmatrix} -A_{I(x^*)} \\ B \\ -B \end{pmatrix} h \geq 0 \quad \text{a} \quad \langle \operatorname{grad} f(x^*), h\rangle < 0$$

řešení. Pak z Farkasovy–Minkowského věty (viz Větu 6.3.19) plyne, že existují $\widetilde{y} \in \mathbb{R}^{\#I(x^*)}$, $\widetilde{y}_1, \widetilde{y}_2 \in$

$\mathbb{R}^{m-k}$ taková, že

$$\begin{pmatrix} -A^\top_{I(x^*)} & B^\top & -B^\top \end{pmatrix} \begin{pmatrix} \widetilde{y} \\ \widehat{y}_1 \\ \widehat{y}_2 \end{pmatrix} = \operatorname{grad}^\top f(x^*) \quad \text{a současně} \quad \widetilde{y}, \widehat{y}_1, \widehat{y}_2 \geq 0, \tag{8.2.14}$$

tj. multiplikátory y^* takové, že $y_i^* := \widetilde{y}_i \geq 0$ pro $i \in I(x^*)$, $y_i := 0$ pro $i \in \{1,\dots,k\} \setminus I(x^*)$ a $y_i := (-\widehat{y}_1 + \widehat{y}_2)_{i-k}$ pro $i = k+1,\dots,m$ splňující (8.2.3) a (8.2.4) s $y_0^* = 1$, viz Obrázek 8.16.



Obrázek 8.16: Ilustrace rovnosti 8.2.14 ukazující, že v případě absence rovnostních omezení je vektor $-\operatorname{grad} f(x^*)$ prvkem kužele vymezeného aktivními omezeními.

Nyní uvažme, že je splněna podmínka (ii). Protože $x^* \in \operatorname{int} P$, existují multiplikátory $y_0^* \geq 0$ a $y^* \in Q$ takové, že

$$y_0^* \operatorname{grad} f(x^*) + \sum_{i=1}^m y_i^* \operatorname{grad} g_i(x^*) = 0 \quad \text{a} \quad y_i\, g_i(x^*) = 0 \quad \text{pro } i = 1,\dots,m. \tag{8.2.15}$$

Je-li $y_0^* \neq 0$, pak tvrzení je triviální. Připusťme proto, že $y_0^* = 0$. Položme $h := \overline{x} - x^*$. Pak pro každé $i \in I(x^*)$ platí $0 = g_i(x^*) > g_i(\overline{x})$. Jelikož funkce g_i jsou konvexní pro $i \in \{1,\dots,k\}$, platí podle Věty 6.4.3, že

$$\underbrace{g_i(\overline{x})}_{<0} \geq \underbrace{g_i(x^*)}_{=0} + \langle \operatorname{grad} g_i(x^*), \underbrace{\overline{x} - x^*}_{=h} \rangle \quad \text{neboli} \quad \langle \operatorname{grad} g_i(x^*), h \rangle < 0 \tag{8.2.16}$$

pro $i \in I(x^*)$. Protože g_i jsou afinní pro $i \in \{k+1,\dots,m\}$, tj. $g_i(x) = a_i^\top x + b_i$ pro nějaká $a_i \in \mathbb{R}^n$ a $b_i \in \mathbb{R}$, platí $\langle \operatorname{grad} g_i(x^*), h \rangle = a_i^\top (\overline{x} - x^*) = b_i - b_i = 0$. Celkem tedy máme

$$0 = \underbrace{\left(\sum_{i=1}^m y_i^* \operatorname{grad} g_i(x^*) \right)^\top}_{\overset{(8.2.15)}{=} 0} h = \sum_{i=1}^m \underbrace{y_i^*}_{\substack{=0 \text{ pro} \\ i \in \{1,\dots,k\} \setminus I(x^*)}} \underbrace{\langle \operatorname{grad} g_i(x^*), h \rangle}_{\substack{=0 \text{ pro} \\ i \in \{k+1,\dots,m\}}} = \sum_{i \in I(x^*)} y_i^* \langle \operatorname{grad} g_i(x^*), h \rangle \overset{\substack{(8.2.16)\ \&\\ y_i^* \geq 0 \text{ pro} \\ i \in I(x^*)}}{<} 0,$$

což v případě $I(x^*) \neq \emptyset$ vede ke sporu ↯. Je-li $I(x^*) = \emptyset$, pak (8.2.15) dává

$$\sum_{i=k+1}^m y_i^* \operatorname{grad} g_i(x^*) = 0,$$

což je spor s lineární nezávislostí konstantních vektorů $\operatorname{grad} g_i$ pro $i \in \{k+1,\dots,m\}$ ↯. Tudíž nutně musí být $y_0^* > 0$, což bez újmy na obecnosti vede k $y_0^* = 1$.

Konečně uvažme, že je splněna podmínka (iii). Připusťme, že existují multiplikátory $y_0^* = 0$ a $y^* \in Q$ splňující (8.2.3) a (8.2.4) v bodě x^*, tj.

$$\sum_{i=1}^k y_i^* \operatorname{grad} g_i(x^*) + \sum_{i=k+1}^k y_i^* \operatorname{grad} g_i(x^*) = 0. \tag{8.2.17}$$

Jelikož ne všechny $y_1^*,\dots,y_m^*$ mohou být nulové, nutně $y_i^* > 0$ pro nějaké $i \in I(x^*)$, neboť v opačném případě bychom dostali

$$\sum_{i=k+1}^{k} y_i^* \operatorname{grad} g_i(x^*) = 0,$$

což by vedlo ke sporu s požadavkem lineární nezávislosti vektorů $\operatorname{grad} g_i(x^*)$ pro indexy $i = k+1,\dots,m$ ↯. Jelikož $y_i^* \geq 0$ pro $i = 1,\dots,k$ a $y_i^* = 0$ pro $i \in \{1?\dots,k\} \setminus I(x^*)$ a $y_i^* > 0$ pro alespoň jedno $i \in I(x^*)$, platí pro vektor $h \in \mathbb{R}^n$ z předpokladů tvrzení, že

$$\begin{aligned} 0 &\overset{(8.2.17)}{=} \sum_{i=1}^{k} y_i^* \langle \operatorname{grad} g_i(x^*), h\rangle + \sum_{i=k+1}^{k} y_i^* \langle \operatorname{grad} g_i(x^*), h\rangle = \\ &= \sum_{i\in I(x^*)} y_i^* \underbrace{\langle \operatorname{grad} g_i(x^*), h\rangle}_{<0} + \sum_{i=k+1}^{k} y_i^* \underbrace{\langle \operatorname{grad} g_i(x^*), h\rangle}_{=0} < 0, \end{aligned}$$

což je spor ↯. Proto opět nutně musí být $y_0^* > 0$, což bez újmy na obecnosti vede k $y_0^* = 1$. ■

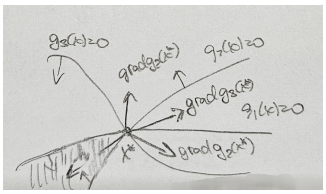
Poznámka 8.2.8. Ukázku vektoru h splňujícího podmínku Věty 8.2.7(iii) můžeme vidět na Obrázku 8.17. Navíc tato podmínka je ekvivalentní se snadněji ověřitelným požadavkem: pro každé $i \in I(x^*)$ existuje vektor $h^{[i]} \in \mathbb{R}^n$ takový, že

$$\left\langle \operatorname{grad} g_j(x^*), h^{[i]} \right\rangle = 0 \quad \text{pro všechna } j \in \{k+1,\dots,m\}$$

a současně

$$\begin{aligned} \left\langle \operatorname{grad} g_i(x^*), h^{[i]} \right\rangle &< 0, \\ \left\langle \operatorname{grad} g_j(x^*), h^{[i]} \right\rangle &\leq 0 \quad \text{pro } j \in I(x^*)\setminus\{i\}. \end{aligned}$$

Pak totiž vektor $h := \sum_{i\in I(x^*)} h^{[i]}$ splňuje obě nerovnosti z Věty 8.2.7(iii). ▲



Obrázek 8.17: Ilustrace vektoru h z Mangasarianovy–Fromovitzovy podmínky.

V následujícím tvrzení si uvedeme další podmínky zaručující $y_0^* = 1$ bez požadavku $x \in \operatorname{int} P$.

Věta 8.2.9

Nechť jsou splněny předpoklady Věty 8.2.1 a x^* je lokálním řešením úlohy (8.1) a (8.2). Jestliže

(i) neexistují $y_{k+1},\dots,y_m$ taková, že

$$\sum_{i=k+1}^{m} y_i \left\langle \operatorname{grad} g_i(x^*), x - x^* \right\rangle \geq 0 \quad \text{pro všechna } x \in P,$$

(ii) existuje vektor $h \in \mathcal{V}(x^*, X)$ takový, že

$$\begin{aligned} \langle \operatorname{grad} g_i(x^*), h\rangle &< 0 \quad \text{pro } i \in I(x^*), \\ \langle \operatorname{grad} g_i(x^*), h\rangle &= 0 \quad \text{pro } i \in \{k+1,\dots,m\}, \end{aligned}$$

potom existují jediné multiplikátory $y^* \in Q$ splňující (8.2.3) a (8.2.4) s $y_0^* = 1$.

Důkaz. Připusťme, že existují multiplikátory $y_0^* = 0$ a $y^* \in Q$ splňující (8.2.3) a (8.2.4), tj.

$$\sum_{i=1}^{m} y_i^* \left\langle \operatorname{grad} g_i(x^*), x - x^* \right\rangle \geq 0 \quad \text{pro všechna } x \in P. \tag{8.2.18}$$

Protože ne všechny y_i^* mohou být nulové, nutně $y_i^* > 0$ pro nějaké $i \in I(x^*)$, protože jinak dostaneme spor s podmínkou (i) ↯. Protože $y_i^* \geq 0$ pro $i \in \{1,\dots,k\}$ a $y_i^* = 0$ pro $i \in \{1,\dots,k\} \setminus I(x^*)$ a současně $y_i^* > 0$ alespoň pro jedno $i \in I(x^*)$, dostáváme pro vektor h z podmínky (ii), že

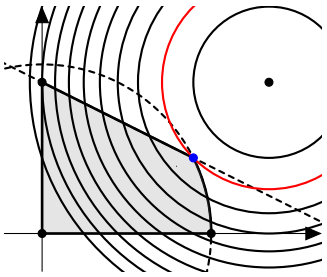
$$\begin{aligned} &\sum_{i=1}^{k} y_i^* \langle \operatorname{grad} g_i(x^*), h\rangle + \sum_{i=k+1}^{m} y_i^* \langle \operatorname{grad} g_i(x^*), h\rangle = \\ &\qquad = \sum_{i\in I(x^*)} y_i^* \underbrace{\langle \operatorname{grad} g_i(x^*), h\rangle}_{<0} + \sum_{i=k+1}^{m} y_i^* \underbrace{\langle \operatorname{grad} g_i(x^*), h\rangle}_{=0} < 0, \end{aligned}$$

což je spor s (8.2.18). Proto nutně musí být $y_0^* > 0$, což bez újmy na obecnosti vede k $y_0^* = 1$. ■

Všimněme si, že zatímco podmínka lineární nezávislosti gradientů aktivních omezení je spíše lokálního charakteru, tak podmínky z Věty 8.2.7(i)–(ii) jsou založeny na globálních vlastnostech daných omezení. V [57, Proposition 3.3.7] lze najít ještě jednu podmínku zobecňující požadavky Věty 8.2.7(i), která dokonce ani nevyžaduje, aby množina X byla konvexní:

> funkce $g_1,\dots,g_m$ jsou konkávní a funkce $g_{k+1},\dots,g_m$ jsou afinní

V následujícím příkladu ilustrujeme platnost tvrzení Věty 8.2.7(ii), tj. existenci multiplikátorů (pouze) s $y_0^*>0$. I tentokrát zůstaneme u hledání multiplikátorů pouze pomocí Věty 8.2.1. Na druhou stranu poznamenejme, že omezení v úloze v Příkladě 8.2.4 nesplňuje žádnou z podmínek ve Větě 8.2.7, ačkoli existují vhodné multiplikátory s $y_0^*=1$. Avšak změníme-li první omezení na $(x_1-1)^2+(x_2-1)^2\le 2$, pak jistě bod $x^*=[0,0]$ zůstane řešením a již bude splněna Slaterova podmínka. V takovém případě už nadále nebude volba s $y_0^*=0$ nebude, neboť $y_1^*=-\lambda<0$ by vedlo ke sporu s nezáporností takového multiplikátoru.



Obrázek 8.18: Grafické znázornění úlohy z Příkladu 8.2.10. Přípustná množina je vyznačena šedou barvou. Vrstevnice funkce f jsou kružnice se zvětšujícím se poloměrem znamenajícím rostoucí funkční hodnotu. Červenou barvou je vyznačena vrstevnice odpovídající nejmenší funkční hodnotě na přípustné množině a řešením je modrý bod se souřadnicemi $[2,1]$.

Příklad 8.2.10

Uvažme úlohu

$$(x_1-3)^2+(x_2-2)^2\to\min$$
$$x_1^2+x_2^2\le 5 \quad\&\quad x_1+2x_2\le 4 \quad\&\quad x_1\ge 0 \quad\&\quad x_2\ge 0.$$

Z Obrázku 8.18 vidíme, že řešením je bod $x^*=[2,1]$. Ukážeme, že v tomto bodě jsou splněny podmínky (8.2.3) a (8.2.4). V bodě x^* jsou aktivní pouze první dvě omezení, tudíž nutně $y_3^*=0=y_4^*$ a podmínka (8.2.4) je splněna pro libovolnou dvojici y_1^*,y_2^*. Dále (funkce g_1 a g_2 odpovídají postupně prvnímu a druhému omezení)

$$\operatorname{grad} f(x^*)=\begin{pmatrix}-2\\-2\end{pmatrix} \quad\&\quad \operatorname{grad} g_1(x^*)=\begin{pmatrix}4\\2\end{pmatrix} \quad\&\quad \operatorname{grad} g_2(x^*)=\begin{pmatrix}1\\2\end{pmatrix}.$$

Podmínka (8.2.3) pak má podobu

$$y_0^*\begin{pmatrix}-2\\-2\end{pmatrix}^{\top}\begin{pmatrix}x_1-2\\x_2-1\end{pmatrix}+y_1^*\begin{pmatrix}4\\2\end{pmatrix}^{\top}\begin{pmatrix}x_1-2\\x_2-1\end{pmatrix}+y_2^*\begin{pmatrix}1\\2\end{pmatrix}^{\top}\begin{pmatrix}x_1-2\\x_2-1\end{pmatrix}\ge 0$$

neboli

$$\left[y_0^*\begin{pmatrix}-2\\-2\end{pmatrix}+y_1^*\begin{pmatrix}4\\2\end{pmatrix}+y_2^*\begin{pmatrix}1\\2\end{pmatrix}\right]^{\top}\begin{pmatrix}x_1-2\\x_2-1\end{pmatrix}\ge 0.$$

Tato podmínka má být splněna pro všechna $[x_1,x_2]\in\mathbb{R}^2=P$, proto musí být

$$y_0^*\begin{pmatrix}-2\\-2\end{pmatrix}+y_1^*\begin{pmatrix}4\\2\end{pmatrix}+y_2^*\begin{pmatrix}1\\2\end{pmatrix}=0$$

pro $y_0^*\ge 0$, $y_1^*\ge 0$ a $y_2^*\ge 0$. Toto vede na soustavu

$$-2y_0^*+4y_1^*+y_2^*=0 \quad\&\quad -2y_0^*+2y_1^*+2y_2^*=0,$$

jejímž řešením je dvojice $y_1^*=y_0^*/3$ a $y_2^*=2y_0^*/3$. Odtud vidíme, že nutně $y_0^*>0$, neboť v opačném případě $y_0^*=0$ by byly všechna multiplikátory nulové. Zejména tedy existují multiplikátory splňující (8.2.3) a (8.2.4) s $y_0^*=1$.

S využitím Farkasovy–Minkowského věty (viz Větu 6.3.19) dostaneme dokonce nutnou a postačující podmínku existence multiplikátorů s $y_0^*=1$, srovnejme s Větou 8.2.7(i).

Věta 8.2.11

Nechť jsou splněny předpoklady Věty 8.2.1 a x^* je lokálním řešením úlohy (8.1) a (8.2). Pak existují multiplikátory $y^*\in Q$ takové, že platí (8.2.3) a (8.2.4) s $y_0^*=1$, právě tehdy,

Věta 8.2.11

když

$$\mathcal{F}(x^*) \cap \mathcal{D}(x^*) = \emptyset, \tag{8.2.19}$$

kde

$$\mathcal{F}(x^*) := \left\{ h \in \mathbb{R}^n \mid \langle \operatorname{grad} g_i(x^*), h \rangle \le 0,\ i \in I(x^*),\ \langle \operatorname{grad} g_j(x^*), h \rangle = 0,\ j \in \{k+1, \dots, m\} \right\}$$

je tzv. *kužel přípustných variací prvního řádu* v bodě x^* a

$$\mathcal{D}(x^*) := \left\{ h \in \mathbb{R}^n \mid \langle \operatorname{grad} f(x^*), h \rangle < 0 \right\}$$

je tzv. *kužel zlepšujících směrů* funkce f v bodě x^*.

Důkaz. Nejdříve uvažme, že nejsou žádná omezení ve tvaru rovností. Podmínka (8.2.19) je ekvivalentní s $\langle \operatorname{grad} f(x^*), h \rangle \ge 0$ pro všechny $h \in \mathbb{R}^n$ taková, že $\langle \operatorname{grad} g_i(x^*), h \rangle \le 0$ pro $i \in I(x^*)$, což je dle Farkasovy–Minkowského věty ekvivalentní s existencí multiplikátorů $y^* \in Q$ splňujících (8.2.3) a (8.2.4) s $y_0^* = 1$. V případě dalšího rovnostního omezení, jej můžeme nahradit dvojicí nerovností, takže tvrzení v takovém případě plyne z první části. ■

Nicméně Věta 8.2.1 dává pouze nutnou podmínku, tj. ne každý bod splňující (8.2.3) a (8.2.4) může být řešením úlohy (8.1) a (8.2), viz Příklad 8.2.22. K definitivnímu rozhodnutí o tom, zda ve stacionárním bodě nastává hledaný extrém či nikoliv, potřebujeme nějakou postačující podmínku – jakýsi *certifikát optimality*. Měli bychom zdůraznit, že nalezení řešení úlohy (8.1) a (8.2) rozhodně není ekvivalentní s nalezením extrému Lagrangeovy funkce, a to ani v případě s $y_0^* = 1$! Avšak je to podmínka postačující, tj. pokud se nám podaří najít minimum funkce $L(x, y^*)$, pak máme i hledané řešení dané úlohy.

Věta 8.2.12 Obecná postačující podmínka

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_k : P \to \mathbb{R}$ jsou diferencovatelné v bodě $x^* \in X$ a $g_{k+1}, \dots, g_m : P \to \mathbb{R}$ jsou spojitě diferencovatelné v nějakém okolí bodu x^*. Nechť dále pro $x^* \in X$ existují multiplikátory $y^* \in Q$ takové, že platí (8.2.3) a (8.2.4) s $y_0^* = 1$. Je-li x^* bodem globálního minima funkce $L(x, y^*)$ na P, pak x^* je globálním řešením úlohy (8.1) a (8.2).

Důkaz. Z předpokladů plyne, že $L(x^*, y^*) \le L(x, y^*)$ pro každé $x \in P$. Navíc z (8.2.4) dostáváme

$$f(x^*) = f(x^*) + \sum_{i=1}^{m} y_i^* g_i(x^*) = L(x^*, y^*).$$

Proto pro libovolné $x \in X$ platí

$$f(x^*) = L(x^*, y^*) \le L(x, y^*) = f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = f(x) + \sum_{i=1}^{k} y_i^* g_i(x) \le f(x),$$

neboť $g_i(x) = 0$ pro $i \in \{k+1, \dots, m\}$ a $y_i^* g_i(x) \le 0$ pro každé $x \in X$. To znamená[42], že x^* je globálním minimem funkce f na množině X. ■

Kdy ale budou splněny podmínky Věty 8.2.12? Zejména, když se zde objeví konvexnost! Přesněji řečeno ve chvíli, když funkce Lagrangeova funkce $L(x, y^*)$ bude konvexní na P, což se stane obzvláště v případě, kdy úloha (8.1) a (8.2) je úlohou konvexního programování. To je obsahem následujícího důsledku.

Důsledek 8.2.13

Nechť $P \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f, g_1, \dots, g_m$ jsou diferencovatelné na (nějaké otevřené množině obsahující) P a pro $x^* \in X$ existují multiplikátory $y^* \in Q$ takové, že platí (8.2.3) a (8.2.4) s $y_0^* = 1$. Nechť je dále splněn (alespoň) jeden z následujících předpokladů:

(i) funkce $L(x, y^*)$ je konvexní na množině P;

[42] Proč požadujeme, aby x^* bylo globálním minimem funkce $L(x, y^*)$ na P? Vždyť předchozí výpočet bude platit i v případě, že x^* je globálním minimem funkce $L(x, y^*)$ pouze na X. Jenže rozhodnout o tom, zda takový extrém nastává na X je vlastně stejně obtížné jako vyřešit původní úlohu. Chceme-li poskytnout nějakou snadněji ověřitelnou podmínku, pak globální minimum $L(x, y^*)$ na P je asi tou nejjednodušší možností. Množina P by totiž měla být mnohem jednodušší než X (typicky $P = \mathbb{R}^n$), takže ověření globálního minima na P by mohlo být výrazně snadnější. Navíc ještě poznamenejme, že řešení úlohy (8.1) a (8.2) není ani totéž jako nalezení globálního minima funkce $L(x, y^*)$ na X.

Důsledek 8.2.13

(ii) úloha (8.1) a (8.2) je úlohou konvexního programování, tj. na konvexní množině P jsou funkce $f, g_1, \dots, g_k$ konvexní a funkce $g_{k+1}, \dots, g_m$ afinní.

Pak bod x^* je globálním řešením úlohy (8.1) & (8.2).

Důkaz. Jelikož $y^* \in Q$, předpoklad (ii) implikuje část (i). Podmínka (8.2.3) znamená, že bod x^* je stacionárním bodem funkce $L(x, y^*)$ na P. Jelikož funkce $L(x, y^*)$ je konvexní a množina P je také konvexní, plyne z Věty 8.1.4, že bod x^* je bodem globálního minima funkce $L(x, y^*)$ na P. Proto tvrzení plyne z Věty 8.2.12. ■

Poznámka 8.2.14. Mohl by problém (8.1) a (8.2) být úlohou konvexního programování i za jiných podmínek než těch, které jsou popsány v části (ii)? V případě rovnostních omezení je nutné, aby funkce $g_{k+1}, \dots, g_m$ byly afinní, jinak by totiž X nemohla být konvexní množina. V případě nerovnostních omezení zaručí konvexnost X to, že nerovností $g_i(x) \le 0$ určující vlastně dolní vrstevnicovou množinu V_0 bude vymezena konvexní množina. Tuto vlastnost mají nepochybně konvexní funkce (viz Větu 6.2.10), ovšem my již víme, že právě tato vlastnost definuje *kvazikonvexní* funkce, viz Poznámku 6.2.11. ▲

Z předchozího důsledku vidíme, že zejména v případě úlohy konvexního programování je existence stacionárního bodu s Lagrangeovým multiplikátorem $y_0^* = 1$ dokonce postačující pro globální řešení úlohy (8.1) a (8.2). V souladu s předchozí částí tak dostaneme následující *základní tvrzení pro úlohy konvexního programování.*

Věta 8.2.15 Karushova–Kuhnova–Tuckerova v diferenciálním tvaru

Nechť $P \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f, g_1, \dots, g_k$ konvexní na P a diferencovatelné na (nějaké otevřené množině obsahující) P, funkce $g_{k+1}, \dots, g_m$ afinní na P a nechť platí (alespoň) jedna z následujících podmínek:

(i) (*lineární nezávislost*) množina P je otevřená, vektory $\operatorname{grad} g_i(x)$, $i \in S(x)$, jsou lineárně nezávislé pro každé $x \in X$;

(ii) (*Slaterova*) funkcionální omezení jsou pouze ve tvaru nerovností, tj. $k = m$, a existuje bod $\overline{x} \in P$ takový, že $g_i(\overline{x}) < 0$ pro $i \in \{1, \dots, k\}$;

(iii) (*lineární*) množina P je polyedr a funkce $g_1, \dots, g_k$ jsou afinní.

(iv) (*zobecněná lineární*) množina P je polyedr a existuje $\ell \in \{1, \dots, k\}$ takové, že funkce $g_{\ell+1}, \dots, g_k$ jsou afinní, a současně existuje $\overline{x} \in X$ takový, že $g_i(\overline{x}) < 0$ pro $i \in \{1, \dots, \ell\}$;

(v) (*zobecněná Slaterova*) funkce $g_{\ell+1}, \dots, g_k$ jsou afinní pro nějaké $\ell \in \{0, \dots, k\}$ a existuje $\overline{x} \in \operatorname{ri} P \cap X$ takové, že $g_i(\overline{x}) < 0$ pro $i \in \{1, \dots, \ell\}$;

Pak x^* je řešením úlohy (8.1) & (8.2) právě tehdy, když existuje $y^* \in Q$ takové, že platí (8.2.3) a (8.2.4) s $y_0^* = 1$.

Poznámka 8.2.16. Předchozí podmínky (ii) a (iii) jsou mírně odlišné od podmínek (ii) a (i) uvedených ve Větě 8.2.7. To je způsobeno požadavkem konvexností funkce f, která umožňuje využití regulárních modifikací Věty 6.3.21. Bod $\overline{x}$ vyžadovaný v podmínce (ii) musí splňovat $\overline{x} \in \operatorname{int} X$. Volba $\ell = 0$ by v podmínce (iv) vedla přímo na (iii). Volbou jakéhokoli ℓ v (iv) a (v) dostaneme rozšíření/kombinaci Slaterovy podmínky (ii) a podmínky (iii), zejména je-li $\ell = k$. Rozdíl spočívá pouze v pozici bodu $\overline{x}$. V podmínce (iv) to musí být bod z přípustné množiny X, zatímco v (v) musí navíc patřit do relativního vnitřku množiny udávající přímé omezení, tj. podmínka „P je mnohostěn“ je nahrazena za „$x \in \operatorname{ri} P$“. Všimněme si také, že v případě $P = \mathbb{R}^n$ jsou podmínky (iv) a (v) shodné. Nakonec ještě dodejme, že úlohy lineárního a kvadratického programování jsou význačné tím, že vždy platí podmínka (iii). ▲

Důkaz Věty 4.2.3. Implikace „$\Longleftarrow$“ (dostatečnost) plyne přímo z Důsledku 8.2.13 (i bez dodatečných podmínek).

Musíme dokázat především implikaci „$\Longrightarrow$“, tj. dostatečnost uvedených předpokladů pro existenci multiplikátorů s $y_0^* = 1$. Nechť tedy $x^* \in X$ řeší úlohu (8.1) a (8.2). V případě podmínky (i)

plyne tvrzení přímo z Věty 8.2.5. Uvažme nyní podmínku (ii). Nechť tedy $k = m$ a uvažme systém nerovností

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle < 0, \tag{8.2.20}$$

$$\langle \operatorname{grad} g_i(x^*), x - x^* \rangle < 0, \quad i \in I(x^*). \tag{8.2.21}$$

Pak tento systém nemá řešení na P — kdyby totiž $\widetilde{x} \in P$ řešilo (8.2.20) a (8.2.21), pak (podobně jako v důkazu Věty 8.2.1) by z diferencovatelnosti funkcí g_i plynulo pro $h = \widetilde{x} - x^*$, že

$$g_i(x^* + \alpha h) = g_i(x^*) + \underbrace{\langle \operatorname{grad} g_i(x^*), \alpha h \rangle}_{<0 \text{ dle } (8.2.21)} + \omega(\alpha h),$$

tj. $g_i(x^* + \alpha h) < g_i(x^*)$ pro $\alpha > 0$ dostatečně malé. Jelikož $x^* + \alpha h = \alpha \widetilde{x} + (1-\alpha) x^* \in P$ pro všechna $\alpha \in [0,1]$, plyne odtud, že $x^* + \alpha h \in X$ pro $\alpha > 0$ dostatečně malé, a tudíž $h \in \mathscr{V}(x^*, X)$. Současně z (8.2.20) plyne, že směrová derivace funkce f v bodě x^* a směru h je záporná, tj.

$$f(x^* + \alpha h) < f(x^*)$$

pro $\alpha > 0$ dostatečně malé, tj. $h \in \mathscr{U}(x^*, f)$. To ale znamená, že $h \in \mathscr{U}(x^*, f) \cap \mathscr{V}(x^*, X) \neq \emptyset$, což je spor s Lemma 8.1.1 ↯.

Nicméně systém (8.2.21) má řešení $\overline{x}$, neboť s využitím daných předpokladů a Věty 6.4.3 totiž máme

$$\langle \operatorname{grad} g_i(x^*), \overline{x} - x^* \rangle \leq g_i(\overline{x}) - g_i(x^*) = g_i(\overline{x}) < 0. \tag{8.2.22}$$

Proto z Důsledku 6.3.22 plyne existence $y_i^* \geq 0$ pro $i \in I(x^*)$ takových, že platí

$$\Big\langle \operatorname{grad} f(x^*) + \sum_{i \in I(x^*)} y_i^* \operatorname{grad} g_i(x^*), x - x^* \Big\rangle \geq 0.$$

Dodefinujeme-li $y_i^* := 0$ pro $i \in \{1, \dots, k\} \setminus I(x^*)$, pak dostáváme multiplikátory $y_1^*, \dots, y_k^* \geq 0$, pro které platí (8.2.3) a (8.2.4) s $y_0^* = 1$.

Nyní nechť je splněna podmínka (iii) a uvažme lineární systém

$$\begin{aligned} \langle \operatorname{grad} f(x^*), x - x^* \rangle &< 0, \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &\leq 0 \quad \text{pro } i \in I(x^*), \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &= 0 \quad \text{pro } i = k+1, \dots, m, \end{aligned}$$

Pokud x splňuje (8.2), pak pro $h = x - x^*$ a libovolné $i \in I(x^*)$ z afinnosti funkcí g_i dostaneme

$$g_i(x^* + \alpha h) = g_i(x^*) + \alpha \langle \operatorname{grad} g_i(x^*), h \rangle \leq g_i(x^*) = 0$$

pro všechna $\alpha \geq 0$. Nyní stejně jako výše můžeme ukázat, že systém (8.2)–(8.2) nemá řešení na P, přitom ale samotný bod x^* je řešením soustavy (8.2)–(8.2). Tvrzení pak plyne z Důsledku 6.3.24.

Je-li splněna podmínka (iv), pak uvažme systém

$$\begin{aligned} \langle \operatorname{grad} f(x^*), x - x^* \rangle &< 0, \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &< 0 \quad \text{pro } i \in I(x^*) \cap \{1, \dots, \ell\}, \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &\leq 0 \quad \text{pro } i \in I(x^*) \cap \{\ell+1, \dots, k\}, \\ \langle \operatorname{grad} g_i(x^*), x - x^* \rangle &= 0 \quad \text{pro } i = k+1, \dots, m, \end{aligned}$$

Podobně jako výše můžeme ukázat, že tato soustava nemá řešení na P a zároveň je bod $\overline{x}$ řešením podsystému (8.2)–(8.2) na P. Vskutku, pro jakékoli $i \in I(x^*) \cap \{1, \dots, \ell\}$ platí (8.2.22), zatímco pro jakékoli $i \in I(x^*) \cap \{\ell+1, \dots, k\}$ máme díky afinnosti jednotlivých funkcí

$$\langle \operatorname{grad} g_i(x^*), \overline{x} - x^* \rangle = g_i(\overline{x}) - g_i(x^*) = g_i(\overline{x}) \leq 0$$

[43] **Věta.** *Nechť $X \subseteq \mathbb{R}^n$ je polyedr, funkce $f_0, \dots, f_\ell$ konvexní na relativně otevřené množině $U \supseteq X$ a funkce $f_{\ell+1}, \dots, f_k, f_{k+1}, \dots, f_m$ afinní. Předpokládejme, že systém*

$$f_0(x) < 0 \tag{△}$$

$$\left.\begin{aligned} f_i(x) &< 0 \quad \text{pro } i = 1, \dots, \ell \\ f_i(x) &\leq 0 \quad \text{pro } i = \ell+1, \dots, k \\ f_i(x) &= 0 \quad \text{pro } i = k+1, \dots, m \end{aligned}\right\} \tag{∗}$$

nemá řešení na X, ale podsystém (∗) řešení má. Pak existují čísla $y_1, \dots, y_k \geq 0$ a $y_{k+1}, \dots, y_m \in \mathbb{R}$ taková, že

$$f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \geq 0$$

pro všechna $x \in X$.

Důkaz. Definujme množinu

$$V := \{x \in X \mid f_i(x) \leq 0 \text{ pro } i = \ell+1, \dots, k, \; f_j(x) = 0 \text{ pro } j = k+1, \dots, m\}.$$

Z předpokladů víme, že systém (△) společně s první částí (∗) nemá řešení na V, zatímco pouze první část (∗) řešení na V má. Z Důsledku 6.3.22 plyne existence $y_1, \dots, y_\ell \geq 0$ takových, že

$$h(x) = f_0(x) + \sum_{i=1}^{\ell} y_i f_i(x) \geq 0$$

pro všechna $x \in V$. Funkce h je očividně konvexní na U. Navíc množina V je konvexní (jakožto průnik několika dolních obrysových množin nebo poloprostorů, viz Větu 6.2.10), platí $V \subseteq U = \operatorname{ri} U$ a z Věty 6.1.29 dostaneme, že $\emptyset \neq \operatorname{ri} V = \operatorname{ri} U \cap \operatorname{ri} V$. Potom podle Lemma 6.5.5 s $X = U$, $Y = V$, $f = h$ a $g = 0$ existuje afinní funkce ℓ taková, že $h(x) \geq \ell(x)$ pro všechna $x \in U$ a současně $\ell(x) \geq 0$ pro všechna $x \in V$. Z této poslední nerovnosti a definice množiny V vyplývá, že systém složený z nerovnosti $\ell(x) < 0$ a ze druhé a třetí části (∗) nemá řešení na X. Proto podle Věty 6.3.24 existují čísla $y_{\ell+1}, \dots, y_k \geq 0$ a $y_{k+1}, \dots, y_m \in \mathbb{R}$ taková, že

$$\ell(x) + \sum_{i=\ell+1}^{m} y_i f_i(x) \geq 0$$

pro všechna $x \in X$. Jelikož $X \subseteq U$ plyne odtud požadovaná nerovnost. ∎

V právě dokázané větě se nelze omezit pouze na to, aby $f_0, \dots, f_l$ byly konvexní jen na X. Příkladem toho může být systém $f_0(x) = -\sqrt{x_1 x_2} < 0$ a $f_1(x) = x_1 \leq 0$ na množině $\mathbb{R}^2_+$.

a pro $i = k+1, \dots, m$ platí analogický výpočet akorát s rovností na konci. Tvrzení pak plyne z podobného tvrzení[43] jako výše.

Konečně je-li splněna podmínka (v), pak systém (8.2)–(8.2) stále nemá řešení na P, zatímco bod $\bar{x}$ je řešením podsystému (8.2)–(8.2) na ri$P$. Samotné tvrzení pak díky tomu opět plyne z dalšího podobného tvrzení[44] jako výše. ■

V následujících příkladech si ilustrujeme využití Věty 8.2.15 při řešení konkrétních úloh.

WWW

Ke kontrole výpočtů lze využít Maplet na webu:
☞ https://ln.math.muni.cz/pz/opt/MapletLagrange

Příklad 8.2.17

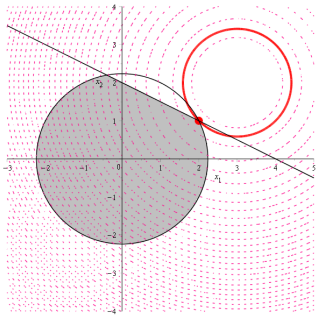
Vyřešme úlohu

$$f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \to \min,$$
$$x_1^2 + x_2^2 \le 5 \quad \& \quad x_1 + 2x_2 \le 4,$$

viz Obrázek 8.19.

Video ☞ https://ln.math.muni.cz/pz/opt/mp1_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp1_zapisnik

⚠ Rovnice v čase 09:51 má být $q/2 + 2q = 4$ (ale $5q/2 = 4$ a tedy $q = 8/5$ je správně).



Obrázek 8.19: Zadání a „řešení“ Příkladu 8.2.17.

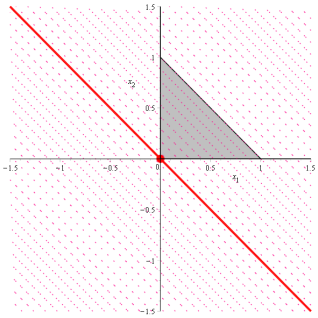
Příklad 8.2.18

Vyřešme úlohu

$$f(x_1, x_2) = x_1 + x_2 \to \min,$$
$$x_1 + x_2 \le 1 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0,$$

viz Obrázek 8.20.

Video ☞ https://ln.math.muni.cz/pz/opt/mp2_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp2_zapisnik



Obrázek 8.20: Zadání a „řešení“ Příkladu 8.2.18.

Příklad 8.2.19

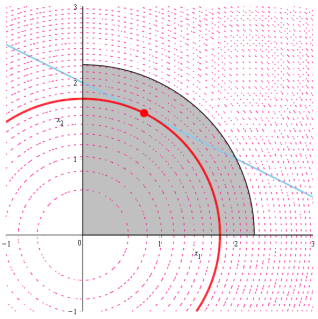
Vyřešme úlohu

$$f(x_1, x_2) = x_1^2 + x_2^2 \to \min,$$
$$x_1^2 + x_2^2 \le 5 \quad \& \quad x_1 + 2x_2 = 4 \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0,$$

viz Obrázek 8.21.

Video ☞ https://ln.math.muni.cz/pz/opt/mp3_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp3_zapisnik

⚠ Rovnice v čase 21:49 má být $2x_2 + 2y_1 x_2 + 2y_2 = 0$. Dále je ale počítáno správně, zejména v čase 28:27 je použita správná rovnice.



Obrázek 8.21: Zadání a „řešení“ Příkladu 8.2.19.

Věty 8.2.12 a 8.2.15 dávají postačující podmínky, za kterých stačí pro řešení úlohy (8.1) a (8.2) najít pouze stacionární bod Lagrangeovy funkce. Bez platnosti těchto podmínek potřebujeme

nějaké další kritérium. To je jako vždy založeno na definitnosti matice

$$\nabla_x^2 L(x^*, y_0^*, y^*).$$

na jistém podprostoru, což nutně vyžaduje $f, g_1, \dots, g_m \in \mathrm{C}^2$.

Věta 8.2.20 Postačující podmínka 2. řádu při ostré komplementaritě

Nechť funkce $f, g_1, \dots, g_m$ jsou dvakrát spojitě diferencovatelné v bodě $x^* \in X \cap \operatorname{int} P$ takovém, že existují multiplikátory $y^* \in Q$ splňující (8.2.3) a (8.2.4) s $y_0^* = 1$ a současně $y_i^* > 0$ pro $i \in I(x^*)$, tj.

$$\begin{gathered}
\operatorname{grad}_x L(x^*, y^*) = 0, \\
g_i(x^*) \le 0 \quad \text{pro } i \in \{1, \dots, k\}, \qquad g_i(x^*) = 0 \quad \text{pro } i \in \{k+1, \dots, m\}, \\
y_i^* > 0 \quad \text{pro } i \in I(x^*), \qquad y_i^* = 0 \quad \text{pro } i \in \{1, \dots, k\} \setminus I(x^*), \\
y_i^* \in \mathbb{R} \quad \text{pro } i \in \{k+1, \dots, m\}.
\end{gathered} \tag{8.2.23}$$

Jestliže

$$\nabla_x^2 L(x^*, y^*) > 0 \quad \text{na } \operatorname{Ker}(\operatorname{grad}^\top g_i(x^*))_{i \in S(x^*)}, \tag{8.2.24}$$

tj. $h^\top \nabla_x^2 L(x^*, y^*)\, h > 0$ pro všechna $h \in \mathbb{R}^n \setminus \{0\}$ taková, že $\langle \operatorname{grad} g_i(x^*), h \rangle = 0$ pro $i \in S(x^*)$, pak bod x^* je ostré lokální minimum funkce f na množině X.

Důkaz. Tvrzení dokážeme pomocí transformace úlohy (8.1) a (8.2) do tvaru výhradně s rovnostními omezeními a následným využitím Věty 1.2.18. Místo úlohy (8.1) a (8.2) tedy uvažme úlohu

$$\left.\begin{aligned}
f(x) &\to \min \\
g_i(x) + z_i^2 &= 0 \quad \text{pro } i \in \{1, \dots, k\}, \\
g_j(x) &= 0 \quad \text{pro } j \in \{k+1, \dots, m\}.
\end{aligned}\right\} \tag{8.2.25}$$

Uvažme bod $[x^*, z^*]$, kde $z^* := [z_1^*, \dots, z_k^*]$ se složkami

$$z_i^* := \sqrt{-g_i(x^*)} \quad \text{pro } i \in \{1, \dots, k\}.$$

Ukážeme, že $[x^*, z^*]$ a y^* splňují předpoklady Věty 1.2.18, z čehož vyplyne, že tento bod je ostrým lokálním (vázaným) minimem úlohy 8.2.25, a tedy x^* řešením původní úlohy (8.1) a (8.2). Označme jako $\overline{L}(x, z, y)$ Lagrangeovu funkci úlohy (8.2.25), tj.

$$\overline{L}(x, z, y) := f(x) + \sum_{i=1}^{k} y_i \left[g_i(x) + z_i^2 \right] + \sum_{i=k+1}^{m} y_i\, g_i(x).$$

Pak

$$\begin{aligned}
\operatorname{grad}_{(x,z)} \overline{L}(x^*, z^*, y^*) &= \left(\operatorname{grad}_x^\top \overline{L}(x^*, z^*, y^*), \operatorname{grad}_z^\top \overline{L}(x^*, z^*, y^*) \right)^\top = \\
&= \left(\operatorname{grad}_x^\top L(x^*, y^*), 2\, y_1^* z_1^*, \dots, 2\, y_k^* z_k^* \right)^\top = (0, 0)^\top,
\end{aligned}$$

kde poslední rovnost plyne z předpokladů $\operatorname{grad}_x^\top L(x^*, y^*) = 0$, $y_i^* = 0$ pro $i \in \{1, \dots, k\} \setminus I(x^*)$ a $z_i^* = 0$ pro $i \in I(x^*)$. Současně z (8.2.25) vyplývá, že platí

$$\operatorname{grad}_y \overline{L}(x^*, z^*, y^*) = \left(g_1(x^*) + (z_1^*)^2, \dots, g_k(x^*) + (z_k^*)^2, g_{k+1}(x^*), \dots, g_m(x^*) \right)^\top = (0, 0)^\top.$$

Nyní ukážeme, že pro každé $(h^\top, u^\top)^\top \in \mathbb{R}^{n+k} \setminus \{0\}$ splňující

$$\left.\begin{aligned}
\langle \operatorname{grad} g_i(x^*), h \rangle + 2\, z_i^*\, u_i &= 0 \quad \text{pro } i \in \{1, \dots, k\}, \\
\langle \operatorname{grad} g_j(x^*), h \rangle &= 0 \quad \text{pro } j \in \{k+1, \dots, m\}
\end{aligned}\right\} \tag{8.2.26}$$

platí

$$\begin{pmatrix} h^\top & u^\top \end{pmatrix} \begin{pmatrix}
\nabla_x^2 L(x^*, y^*) & 0 & \cdots & \cdots & 0 \\
0 & 2\, y_1^* & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \ddots & 2\, y_{k-1}^* & 0 \\
0 & 0 & \cdots & 0 & 2\, y_k^*
\end{pmatrix} \begin{pmatrix} h \\ u \end{pmatrix} > 0, \tag{8.2.27}$$

[44] **Věta.** *Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina, funkce $f_0, \dots, f_\ell$ konvexní na X a $f_{\ell+1}, \dots, f_m$ afinní funkce. Jestliže systém (△) & (∗) z předchozí poznámky nemá řešení na X a podsystém (∗) jej má na $\operatorname{ri} X$. Pak existují čísla $y_1, \dots, y_k \ge 0$ a $y_{k+1}, \dots, y_m \in \mathbb{R}$ taková, že*

$$f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \ge 0$$

pro všechna $x \in X$.

Důkaz. Společně s množinou V z předchozí poznámky uvažme také množinu

$$\begin{aligned}
\widehat{V} := \{ x \in \operatorname{aff} X \mid \\
f_i(x) \le 0 \text{ pro } i = \ell+1, \dots, k, \\
f_j(x) = 0 \text{ pro } j = k+1, \dots, m \}.
\end{aligned}$$

Je jasné, že $V = X \cap \widehat{V}$. Pak podobně jako v předchozí poznámce odvodíme s využitím Důsledku 6.3.22, že existují čísla $y_1, \dots, y_\ell \ge 0$ taková, že

$$h(x) = f_0(x) + \sum_{i=1}^{\ell} y_i f_i(x) \ge 0$$

pro všechna $x \in V$. Funkce h je konvexní na X a $\operatorname{ri} X \cap \widehat{V} \ne \emptyset$, $\widehat{V} \subseteq \operatorname{aff} X$, a tudíž $\operatorname{ri} X \cap \operatorname{ri} \widehat{V} \ne \emptyset$ (toto je obecně platné tvrzení: jsou-li množiny $X_1, X_2 \subseteq \mathbb{R}^n$ konvexní, $\operatorname{ri} X_1 \cap X_2 \ne \emptyset$ a $X_2 \subseteq \operatorname{aff} X_1$, pak $\operatorname{ri} X_1 \cap \operatorname{ri} X_2 \ne \emptyset$. Jeho důkaz už ponecháme čtenáři k samostatnému objevení.) Potom podle Lemma 6.5.5 s X, $Y = \widehat{V}$, $f = h$ a $g = 0$ existuje afinní funkce ℓ taková, že $h(x) \ge \ell(x)$ pro všechna $x \in X$ a současně $\ell(x) \ge 0$ pro všechna $x \in \widehat{V}$. Z této poslední nerovnosti a definice množiny $\widehat{V}$ vyplývá, že systém složený z nerovnosti $\ell(x) < 0$ a ze druhé a třetí části (∗) nemá řešení na $\operatorname{aff} X$. Současně množina $\operatorname{aff} X$ je polyedr, takže podle Věty 6.3.24 existují čísla $y_{\ell+1}, \dots, y_k \ge 0$ a $y_{k+1}, \dots, y_m \in \mathbb{R}$ taková, že

$$\ell(x) + \sum_{i=l+1}^{m} y_i f_i(x) \geq 0$$

pro všechna $x \in \operatorname{aff} X$, čímž je důkaz hotov. ■

V uvedené větě se nelze omezit pouze na to, že podsystém (∗) má řešení na X. Příkladem toho může být volba $f_0(x) = x_1 - 1 < 0$ a $f_1(x) = x_1 + x_2 - 2 \leq 0$ na množině $X = \{[x_1, x_2] \in \mathbb{R}^2_+ \mid x_1\, x_2 \geq 1\}$.

přičemž levá strana (8.2.27) může být zapsána jako

$$h^\top \nabla^2 L(x^*, y^*)\, h + 2 \sum_{i=1}^{k} y_i^*\, u_i^2. \tag{8.2.28}$$

Nechť tedy $(h^\top, u^\top)^\top \in \mathbb{R}^{n+k} \setminus \{0\}$ je libovolný vektor splňující (8.2.26). Pro $i \in I(x^*)$ platí $z_i^* = 0$, což dle (8.2.26) dává

$$\langle \operatorname{grad} g_i(x^*), h \rangle \quad \text{pro } i \in S(x^*).$$

Je-li $h \neq 0$, pak z předpokladů plyne

$$h^\top \nabla^2 L(x^*, y^*)\, h > 0,$$

což díky (8.2.28) a faktu $y_i^* \geq 0$ pro $i \in \{1, \ldots, k\}$ implikuje, že $(h^\top, u^\top)^\top$ splňuje (8.2.27).

Je-li $h = 0$, pak nutně $u_j \neq 0$ pro nějaké $j \in \{1, \ldots, k\}$. V takovém případě díky (8.2.26) máme

$$2\, z_i^*\, u_i = 0 \quad \text{pro } i \in \{1, \ldots, k\},$$

z čehož plyne, že z_j^* musí být rovno 0, a tudíž $j \in I(x^*)$. Z předpokladů máme $y_j^* > 0$, tj. $2\, y_j^*\, u_j^2 > 0$, a tudíž

$$2 \sum_{i=1}^{k} y_i^*\, u_i^2 > 0,$$

což opět potvrzuje platnost (8.2.27). Tím je důkaz hotov. ■

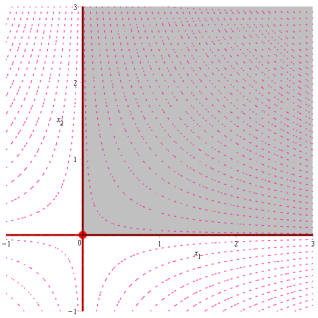
Příklad 8.2.21

Vyřešme úlohu

$$-x_1\, x_2 \to \min,$$
$$x_1 + x_2 \leq 6 \quad \& \quad x_1 \geq 0 \quad \& \quad x_2 \geq 0,$$

viz Obrázek 8.22.

Video ☞ https://ln.math.muni.cz/pz/opt/mp4_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp4_zapisnik



Obrázek 8.22: Zadání a „řešení“ Příkladu 8.2.21. Pozor! Obrázek je generován z MAPLE, který chybně označuje bod $[0,0]$ za řešení!

Platnost (8.2.24) lze ověřit buď přímo pomocí vektorů z $\{h \in \mathbb{R}^n \mid \mathrm{D}G(x^*)\, h = 0\}$, kde $\mathrm{D}G(x^*)$ značí Jacobiho matici vektorové funkce

$$G := \big(g_i(x)\big)_{i \in S(x^*)} \in \mathbb{R}^{\#I(x^*)+m-k},$$

nebo pomocí rozšířené Hessovy matice

$$H := \begin{pmatrix} 0 & \mathrm{D}G(x^*) \\ \mathrm{D}^\top G(x^*) & \nabla^2 L(x^*, y^*) \end{pmatrix},$$

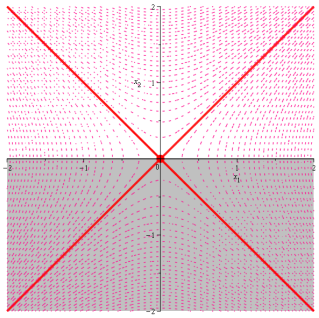
pro kterou posledních $n - (\#S(x^*))$ vedoucích hlavní minorů musí mít stejné znaménko jako číslo $(-1)^{\#S(x^*)}$.

Podmínka (8.2.23) se nazývá *podmínkou ostré komplementarity*. Její nesplnění, tj. $y_i^* = 0$ pro nějaké $i \in S(x^*)$, ukazuje jistou „degenerovanost“ příslušného omezení, jak můžeme vidět v následujícím příkladu. Tento příklad je také ukázkou toho, že podmínky (8.2.3) a (8.2.4) jsou pouze nutné (včetně situace s $y_0^* = 1$) nikoli postačující pro nalezení řešení úlohy (8.1) a (8.2).

Příklad 8.2.22

Uvažme úlohu

$$\frac{x_1^2 - x_2^2}{2} \to \min \quad \& \quad x_2 \leq 0,$$



Obrázek 8.23: Zadání a „řešení“ Příkladu 8.2.22.

viz Obrázek 8.23. Snadno můžeme ověřit, že bod $x^* = [0,0]$ splňuje (8.2.3) a (8.2.4) s $y_0^* = 1$ a $y_1^* = 0$. Poněvadž účelová funkce není konvexní, potřebujeme nějaké další kritérium. Ovšem není splněna podmínka ostré komplementarity (v x^* je dané omezení aktivní a zároveň je příslušný Lagrangeův multiplikátor nulový), takže nelze využít postačující podmínku z Věty 8.2.20. Kdybychom přeci jen zkusili tuto větu aplikovat, zjistili bychom, že Hessova matice $\nabla_x^2 L(x^*, y^*)$ je sice indefinitní, ale vzhledem aktivnosti daného omezení v x^* je potřeba otestovat definitnost této matice pouze na $\ker \operatorname{grad}^\top g(x^*) = \ker(0,1)$. Protože v takovém případě máme $h^\top \nabla_x^2 L(x^*, y^*)\, h > 0$ pro libovolné $h \in \ker \operatorname{grad}^\top g(x^*) \setminus \{0\}$, vyplývalo by z Věty 8.2.20, že x^* je ostré lokální minimum. Nicméně je téměř zřejmé, že to není pravda (tj. v bodě $[0,0]$ nemůže být lokální minimum), neboť pro libovolný přípustný bod $[x_1, x_2] = [0, x_2]$ s $x_2 < 0$ je $f(0, x_2) < 0$, zatímco $f(0,0) = 0$.

Není-li tedy splněna podmínka ostré komplementarity a/nebo $x \notin \operatorname{int} P$ je formulace postačující podmínky pro ověření lokálního minima komplikovanější. Nabídneme dvě z nich[45], které se opírají o následující dvě množiny

$$\mathscr{V}(x^*) := \Big\{ h \in \mathbb{R}^n \mid h = \lambda (x - x^*),\ \lambda > 0 \ \ \&\ \ x \in P \Big\},$$
$$\mathscr{H}(x^*) := \Big\{ h \in \mathbb{R}^n \mid \langle \operatorname{grad} f(x^*), h \rangle \le 0,\ \langle \operatorname{grad} g_i(x^*), h \rangle \le 0,$$
$$\langle \operatorname{grad} g_j(x^*), h \rangle = 0,\ i \in I(x^*),\ j \in \{k+1, \dots, m\} \Big\}.$$

definované pro $x^* \in P$. Všimněme si, že pro $x^* \in \operatorname{int} P$ bude $\mathscr{V}(x^*) = \mathbb{R}^n$. Začneme se zcela obecným tvrzení nevyžadujícím konvexnost či kvalifikovanost omezení.

[45] Viz [466, Věta 2.6 a 2.7 na str. 161–163] a jejich důsledky.

Věta 8.2.23

Nechť funkce $f, g_1, \dots, g_m$ jsou dvakrát spojitě diferencovatelné v bodě $x^* \in X$, pro který existují multiplikátory $y_0^* \ge 0$ a $y^* \in Q$ splňující (8.2.3) a (8.2.4). Jestliže

$$h^\top \nabla_x^2 L(x^*, y_0^*, y^*)\, h > 0 \quad \text{pro všechna } h \in \overline{\mathscr{V}(x^*)} \cap \mathscr{H}(x^*) \setminus \{0\}, \tag{8.2.29}$$

pak x^* je ostré lokální (vázané) minimum funkce f na X, tj. ostré lokální řešení úlohy (8.1) a (8.2).

Důkaz. Je-li x^* izolovaným bodem množiny X, pak je tvrzení triviální. Nechť tedy dále je x^* hromadným bodem množiny X a připusťme, že není ostrým lokálním (vázaným) minimem funkce f na X. Pak existuje posloupnost $\{x^{[k]}\}_{k=0}^{\infty}$ taková, že

$$x^{[k]} \in X \quad \& \quad x^{[k]} \ne x^* \quad \& \quad x^{[k]} \to x^* \quad \& \quad f(x^{[k]}) \le f(x^*). \tag{8.2.30}$$

Díky tomu můžeme jednotlivé členy $x^{[k]}$ vyjádřit jako

$$x^{[k]} = x^* + \alpha_k h_k,$$

kde $\alpha_k := \| x^{[k]} - x^* \| > 0$ a $h_k := (x^{[k]} - x^*)/\alpha_k \in \mathscr{V}(x^*) \setminus \{0\}$. Protože $\| h_k \| = 1$, můžeme bez újmy na obecnosti připustit, že samotná posloupnost $\{h_k\}$ je konvergentní s[46] $h_k \to h \ne 0$, a tudíž $h \in \overline{\mathscr{V}(x^*)}$. Navíc díky (8.2.30) máme

$$0 \ge f(x^{[k]}) - f(x^*) = \langle \operatorname{grad} f(x^*), \alpha_k h_k \rangle + \omega(\alpha_k h_k),$$
$$0 \ge g_i(x^{[k]}) - g_i(x^*) = \langle \operatorname{grad} g_i(x^*), \alpha_k h_k \rangle + \omega(\alpha_k h_k) \quad \text{pro všechna } i \in I(x^*),$$
$$0 = g(x^{[k]}) - g(x^*) = \langle \operatorname{grad} g_i(x^*), \alpha_k h_k \rangle + \omega(\alpha_k h_k) \quad \text{pro všechna } i = \{k+1, \dots, m\}.$$

Podělením těchto výrazů α_k a následným limitním přechodem dostaneme, že $h \in \mathscr{H}(x^*)$, tj. $h \in \overline{\mathscr{V}(x^*)} \cap \mathscr{H}(x^*) \setminus \{0\}$.

Dále z podmínky (8.2.3) vyplývá, že

$$\big\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), h_k \big\rangle \ge 0, \tag{8.2.31}$$

[46] Však i $\| h \| = 1$, proto nutně $h \ne 0$.

a s ohledem na (8.2.4) a (8.2.30) platí

$$L(x^{[k]}, y_0^*, y^*) = y_0^* f(x^{[k]}) + \sum_{i=1}^{m} y_i^* g_i(x^{[k]}) \le y_0^* f(x^{[k]}) \le$$
$$\le y_0^* f(x^*) = y_0^* f(x^*) + \sum_{i=1}^{m} y_i^* g_i(x^*) = L(x^*, y_0^*, y^*). \tag{8.2.32}$$

Vzhledem k předpokladům věty je funkce $L(x, y_0^*, y^*)$ dvakrát spojitě diferencovatelná v okolí x^*, takže z Taylorova rozvoje[47] získáme

$$L(x^{[k]}, y_0^*, y^*) = L(x^*, y_0^*, y^*) + \langle \operatorname{grad}_x L(x^*, y_0^*, y^*), \alpha_k h_k \rangle +$$
$$+ \frac{\alpha_k^2}{2} \langle \nabla_x^2 L(x^*, y_0^*, y^*) h_k, h_k \rangle + \omega(\alpha_k h_k)$$

Odtud a z (8.2.31)–(8.2.32) plyne

$$\frac{\alpha_k^2}{2} \langle \nabla_x^2 L(x^*, y_0^*, y^*) h_k, h_k \rangle + \omega(\alpha_k h_k) \le 0,$$

z čehož podělením α_k^2 a limitním přechodem obdržíme (8.2.29). ■

[47] Viz Poznámku 62 zmíněnou na straně 62.

Všimněme si, že pro libovolné $y_0^* \ge 0$ a $y^* \in Q$ splňující (8.2.3) a (8.2.4) z podmínky $h \in \overline{\mathcal{V}(x^*)} \cap \mathcal{H}(x^*) \setminus \{0\}$ vyplývá

$$\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), h \rangle = 0, \tag{8.2.33}$$
$$y_0^* \langle \operatorname{grad} f(x^*), h \rangle = 0, \tag{8.2.34}$$
$$y_i^* \langle \operatorname{grad} g_i(x^*), h \rangle = 0 \quad \text{pro všechna } i \in I(x^*). \tag{8.2.35}$$

Tyto rovnosti jsou ve skutečnosti důsledkem toho, že z podmínky (8.2.3) a volby $h \in \overline{\mathcal{V}(x^*)}$ plyne $\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), h \rangle \ge 0$, zatímco z (8.2.4) dostaneme pro $h \in \mathcal{H}(x^*)$ opačnou nerovnost

$$\langle \operatorname{grad}_x L(x^*, y_0^*, y^*), h \rangle = \langle y_0^* \operatorname{grad} f(x^*) + \sum_{i \in S(x^*)} y_i^* \operatorname{grad} g_i(x^*), h \rangle =$$
$$= \langle y_0^* \operatorname{grad} f(x^*), h \rangle + \langle \sum_{i \in I(x^*)} y_i^* \operatorname{grad} g_i(x^*), h \rangle \le 0.$$

Odtud a z důkazu Věty 8.2.23 plyne, že předpoklady samotného tvrzení lze mírně modifikovat s cílem jeho snadnější aplikace. První z takových modifikací nevyžaduje na rozdíl od množiny $\mathcal{H}(x^*)$ explicitně „žádnou“ informaci o $\operatorname{grad} f(x^*)$.

Důsledek 8.2.24

Nechť jsou splněny předpoklady Věty 8.2.23. Jestliže platí

$$h^\top \nabla_x^2 L(x^*, y_0^*, y^*) h > 0$$

pro každé $h \in \overline{\mathcal{V}(x^*)} \setminus \{0\}$ splňující

$$\langle \operatorname{grad} g_i(x^*), h \rangle \le 0 \quad \text{a}^{48} \quad y_i^* \langle \operatorname{grad} g_i(x^*), h \rangle = 0 \quad \text{pro všechna } i \in I(x^*),$$
$$\langle \operatorname{grad} g_j(x^*), h \rangle \le 0 \quad \text{pro všechna } j \in \{k+1, \dots, m\},$$

pak x^* je ostré lokální (vázané) minimum funkce f na X, tj. ostré lokální řešení úlohy (8.1) a (8.2).

[48] První dvě podmínky vlastně znamenají, že buď $y_i^* = 0$ a $\langle \operatorname{grad} g_i(x^*), h \rangle \le 0$ nebo $y_i^* > 0$ a $\langle \operatorname{grad} g_i(x^*), h \rangle = 0$.

Všimněme si, že v případě $k = 0$ (tj. máme-li pouze rovnostní omezení), $x^* \in \operatorname{int} P$ a $y_0^* = 1$ dostaneme tvrzení Věty 1.2.18. Tvrzení Věty 8.2.23 lze samozřejmě upravit také tak, aby nevyžadovalo žádnou informaci o Hessově matici $\nabla_x^2 L(x^*, y_0^*, y^*)$, i když to je za cenu poměrně silného požadavku.

DŮSLEDEK 8.2.25

Nechť jsou splněny předpoklady Věty 8.2.23. Je-li $\overline{\mathscr{V}(x^*)} \cap \mathscr{H}(x^*) = \{0\}$, pak x^* je ostré lokální (vázané) minimum funkce f na X, tj. ostré lokální řešení úlohy (8.1) a (8.2).

V případě úlohy konvexního programování jsou podmínky Věty 8.2.23 a jejich důsledků postačující pro jednoznačnost (globálního) řešení úlohy (8.2.1). Na závěr si ještě uvedeme obecnou nutnou podmínku druhého řádu pro případ $x^* \in \operatorname{int} P$. Zdůrazněme, že její předpoklady vylučují množnost $y_0^* = 0$. Navíc v případě $k = m$ (tj. omezení pouze ve tvaru rovností) se toto tvrzení redukuje na druhou část Věty 1.2.16.

VĚTA 8.2.26

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_m$ dvakrát spojitě diferencovatelné v bodě $x^* \in \operatorname{int} P \cap X$. Nechť navíc funkce g_i jsou spojitě diferencovatelné v nějakém okolí x^* pro $i \in S(x^*)$ a vektory $\operatorname{grad} g_i(x^*)$ jsou lineárně nezávislé pro $i \in S(x^*)$. Je-li x^* lokální řešením úlohy (8.1) a (8.2), pak nutně

$$h^\top \nabla_x^2 L(x^*, y_0^*, y^*)\, h \geq 0 \tag{8.2.36}$$

pro libovolné $h \in \mathscr{H}(x^*)$ a libovolné multiplikátory $y_0^* \geq 0$ a $y^* \in Q$ splňující (8.2.3) a (8.2.4).

Důkaz. Pro libovolné $h \in \mathscr{H}(x^*)$ definujme množiny

$$I(x^*, h) := \{i \in I(x^*) \mid \langle \operatorname{grad} g_i(x^*), h\rangle\} = \{i \in \{1, \dots, k\} \mid g_i(x^*) = 0 \ \&\ \langle \operatorname{grad} g_i(x^*), h\rangle\}$$
$$S(x^*, h) := I(x^*, h) \cup \{k+1, \dots, m\} = \{i \in \{1, \dots, m\} \mid g_i(x^*) = 0 \ \&\ \langle \operatorname{grad} g_i(x^*), h\rangle\}.$$

Podle Ljustrenikova lemmatu[49] existuje funkce $r : \mathbb{R} \to \mathbb{R}^n$ taková, že

$$g_i(x^* + \alpha h + r(\alpha)) = 0 \quad \text{pro všechna } i \in S(x^*, h) \tag{8.2.37}$$

pro všechna dostatečně malá α s Jacobiho maticí

$$D_\alpha r(0) = \lim_{\alpha \to 0} \frac{r(\alpha)}{\alpha} = 0 \in \mathbb{R}^n. \tag{8.2.38}$$

Nechť $x(\alpha) := x^* + \alpha h + r(\alpha)$. Pak pro libovolné $i \in \{1, \dots, k\} \setminus I(x^*, h)$ je splněna některá z následujících dvou nerovností:

$$\text{buď} \quad g_i(x^*) < 0 \quad \text{nebo} \quad \langle \operatorname{grad} g_i(x^*), h\rangle < 0 \tag{8.2.39}$$

Proto v prvním případě společně s (8.2.4) a ve druhém případě s ohledem na (8.2.35) obdržíme pro tyto indexy, že $y_i^* = 0$. Odtud a z (8.2.37) máme $y_i^*\, g_i(x(\alpha)) = 0$ pro všechna $i \in \{1, \dots, m\}$ a všechna dostatečně malá $\alpha > 0$. Díky tomu je

$$L(x(\alpha), y_0^*, y^*) = y_0^*\, f(x(\alpha)) + \sum_{i=1}^{m} y_i^*\, g_i(x(\alpha)) = y_0^*\, f(x(\alpha)) \tag{8.2.40}$$

pro všechna dostatečně malá α. Z (8.2.37)–(8.2.39) a předpokladu $x^* \in \operatorname{int} P$ lze snadno odvodit, že $x(\alpha) \in X$ pro všechna dostatečně malá $\alpha > 0$, takže $f(x^*) \leq f(x(\alpha))$ pro všechna dostatečně malá $\alpha > 0$, protože x^* je lokálním řešením úlohy (8.1) a (8.2). Potom z (8.2.4), (8.2.11), (8.2.40) a dvojnásobné diferencovatelnosti funkce $L(x, y_0^*, y^*)$ v x^* dostaneme s využitím Taylorova rozvoje[50]

$$0 \leq y_0^*\, f(x(\alpha)) - y_0^*\, f(x^*) = L(x(\alpha), y_0^*, y^*) - L(x^*, y_0^*, y^*) =$$
$$= \frac{1}{2}\big\langle \nabla^2 L(x^*, y_0^*, y^*)\,(\alpha h + r(\alpha)), \alpha h + r(\alpha)\big\rangle + \omega(\alpha h + r(\alpha)),$$

a tudíž

$$\frac{1}{2}\big\langle \nabla^2 L(x^*, y_0^*, y^*)\,(h + r(\alpha)/\alpha), h + r(\alpha)/\alpha\big\rangle + \frac{\omega(\alpha h + r(\alpha))}{\alpha^2} \geq 0,$$

z čehož limitním přechodem získáme (8.2.36). ∎

[49] Viz Poznámku 32 zmíněnou na straně 629.

[50] Viz Poznámku 62 zmíněnou na straně 62.

V případě úlohy konvexního programování je tato věta (na rozdíl od předchozí věty) nezajímavá, neboť podmínka (8.2.36) je splněna automaticky pro všechna $h \in \mathbb{R}^n$ díky konvexnosti funkce $L(x, y_0^*, y^*)$ a Důsledku 6.4.7.

Cvičení

8.2.1. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

(a) $f(x) = \sqrt{1+x^2} \to \min, \quad x \le 1.$

(b) $f(x) = \sqrt{1+x^2} \to \min, \quad x+1 \le 0.$

řešení: 0, −1

8.2.2. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

(a) $f(x) = |x| \to \min, \quad x \le 1.$

(b) $f(x) = |x| \to \min, \quad x+1 \le 0.$

řešení: 0, −1

8.2.3. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = x_1^2 + x_2^2 - 14x_1 - 6x_2 \to \min, \quad x_1 + x_2 \le 2, \quad x_1 + 2x_2 \le 3.$$

řešení: $[3, -1]$

8.2.4. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + x_2^2 - 20x_1 - 14x_2 \to \min, \quad x_1 + 3x_2 \le 5, \quad 2x_1 - x_2 \le 4.$$

řešení: $[17/7, 6/7]$

8.2.5. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = x_1^2 + x_2 \to \min, \quad x_1^2 + x_2^2 \le 9, \quad x_1 + x_2 \le 1.$$

řešení: $[0, -3]$

8.2.6. Pomocí Lagrangeova principu vyřešte úlohu

$$f(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 + 1 \to \min, \quad x_1 + x_2 \le 1, \quad x_1^2 + x_2^2 \le 4.$$

řešení: $[4/3, -1/3]$

8.2.7. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = 9x_1^2 + 12x_2^2 + 18x_1x_2 + 3x_1 - 3x_2 + 1 \to \min,$$
$$x_1 + x_2 = 3, \quad x_1 - x_2 \ge 1, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[2, 1]$

8.2.8. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = 3x_1^2 - x_1x_2 + x_2^2 + x_1 + x_2 + 2 \to \min,$$
$$2x_1 + x_2 = 1, \quad x_1 - 3x_2 \ge -1, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[1/3, 1/3]$

8.2.9. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = x_1^2 - x_1x_2 + 3x_2^2 - 3x_1 + x_2 + 2 \to \min,$$
$$x_1 + 2x_2 = 1, \quad 2x_1 - x_2 \ge 1, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[1, 0]$

8.2.10. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1, x_2) = x_1^2/2 - x_1x_2 + x_2^2 - 2x_1 + x_2 \to \min,$$
$$x_1 + x_2 \le 3, \quad 2x_1 - x_2 \le 4, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: [7/3, 2/3]

8.2.11. Pomocí Lagrangeova principu vyřešte úlohu lineárního programování

$$f(x) = \langle c, x \rangle \to \min$$

na množině

$$X = \{x \in \mathbb{R}^n \mid Bx \le b,\ x \ge 0\},$$

je-li $c = (2,2)^\top$, $B = \begin{pmatrix} -3 & 2 \\ -3 & -1 \\ 1 & 0 \end{pmatrix}$ a $b = (6,-3,3)^\top$.

řešení: [1, 0]

8.2.12. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1^2 + x_2 \to \min, \quad x_1^2 + x_2^2 \le 9, \quad x_1 + x_2 \le 1, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: [0, 0]

8.2.13. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1^2 + 2x_2^2 - 4x_1 + 1 \to \min, \quad x_1 + x_2 \le 1, \quad x_1^2 + x_2^2 \le 4, \quad x_1 \ge 0.$$

řešení: [4/3, −1/3]

8.2.14. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 2x_1^2 + x_1x_2 + x_2^2 - 4x_1 - 2x_2, \quad x_1^2 + x_2 \le 1, \quad x_1 + x_2 \le 1, \quad x_1 \ge 0.$$

(**Nápověda:** polynom $4z^3 - 3z^2 + 4z - 3$ má jediný reálný kořen $z = 3/4$.)

řešení: [3/4, 1/4]

8.2.15. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 9x_2 \to \min,$$
$$4x_1^2 - x_2 \le 2, \quad 4x_1 + 3x_2 \le 10, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: [28/37, 86/37]

8.2.16. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2 \to \min, \quad x_1 + x_2 + x_3 = 3, \quad 2x_1 - x_2 + x_3 \le 5.$$

řešení: [1, 1, 1]

8.2.17. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = 4x_1^2 + 4x_1x_2 + 3x_2^2 - 4x_2x_3 + 2x_3^2 + x_1 + 4x_2 - 2x_3 \to \min,$$
$$x_1 - x_2 + 2 = 0, \quad x_1 + 2x_2 - x_3 \le 1, \quad x_1 \ge 0.$$

řešení: [0, 2, 3]

8.2.18. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 8x_1 - 6x_2 - 4x_3 + 9,$$
$$x_1 + x_2 + x_3 = 4, \quad x_1 - 2x_2 - 2x_3 \ge 1, \quad x_3 \ge 0.$$

řešení: [3, 1, 0]

8.2.19. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = 3x_1^2 + x_1x_2 + x_2^2 - 2x_2x_3 + 2x_3^2 \to \min,$$
$$x_1 + 2x_3 = 1, \quad x_1 - 2x_2 - x_3 \le 2, \quad x_2 \ge 0, \quad x_3 \ge 0.$$

řešení: [0, 1/2, 1/2]

8.2.20. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = 3x_1^2 + x_1x_2 + x_2^2 - 2x_2x_3 + 2x_3^2 \to \min,$$
$$x_1 + 2x_3 = 2, \quad x_1^2 - 2x_2 - x_3 \le 2, \quad x_2 \ge 0, \quad x_3 \ge 0.$$

(**Nápověda:** polynom $-32z^3 + 156z^2 - 241z + 92$ má jediný reálný kořen $z \approx 0,56356$.)

řešení: $[0,1,1]$

8.2.21. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2,x_3) = 4x_1^2 + 4x_1x_2 + 3x_2^2 - 4x_2x_3 + 2x_3^2 + x_1 + 4x_2 - 2x_3 \to \min,$$
$$x_1 - x_2 + 2 \le 0, \quad 2x_1 + x_2 - x_3 \le 2, \quad x_1 \ge 0.$$

řešení: $[0,2,5/2]$

8.2.22. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x) = \frac{1}{2}\|x\|^2 - \langle c,x\rangle \to \min$$

na množině $\mathbb{R}^n_+$, kde $c \in \mathbb{R}^n$.

řešení: $[\max\{0,c_1\},\max\{0,c_2\},\ldots,\max\{0,c_n\}]$

8.2.23. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x) = \|x\| - \langle c,x\rangle \to \min$$

na množině $\mathbb{R}^n_+ \setminus \{0\}$, kde $c \in \mathbb{R}^n$.

8.2.24. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1 + x_2^2 + 1 \to \min, \quad x_1^3 - x_2^3 \le 1, \quad x_1 \ge 0.$$

řešení: $[0,0]$

8.2.25. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1^3 - 3x_1x_2 + x_2^3 \to \min, \quad x_2 - 2x_1^2 \le 0, \quad 0 \le x_1 \le 3, \quad x_2 \ge 0.$$

řešení: $[1,1]$

8.2.26. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1^2 + 2x_1x_2 - 4x_1 - 2x_2 \to \min,$$
$$x_1^2 - 2x_2 \le 0, \quad 2x_1 + x_2 \le 6, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[0,6]$

8.2.27. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = x_1^2 + (x_2-1)^2 \to \min, \quad x_1^2 + 4x_2^2 \le 4, \quad 2x_1^2 + x_2 \ge 2, \quad x_1 \ge 2x_2.$$

řešení: $[\sqrt{65}/8 - 1/8, \sqrt{65}/16 - 1/16]$

8.2.28. Uvažte úlohu matematického programování

$$\left.\begin{array}{c} f(x_1,x_2,x_3) = -x_1^3 + x_2^3 - 2x_1x_3^2 \to \min, \\ 2x_1 + x_2^2 + x_3 = 5, \quad 5x_1^2 - x_2^2 - x_3 \ge 2, \quad x_1 \ge 0, \quad x_3 \ge 0. \end{array}\right\} \tag{8.2.41}$$

a) Pomocí Lagrangeova principu určete nutné podmínky pro stacionární body úlohy (8.2.41).

b) Pomocí podmínek druhého řádu ukažte, že bod $[1,0,3]$ je bodem lokálního vázaného minima úlohy (8.2.41).

8.2.29. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 4x_1 + 2x_2 - 2x_1x_2 - x_1^2 \to \min,$$
$$x_1^2 - 2x_2 \le 0, \quad 2x_1 + x_2 \le 6, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[0,0]$ a $[2,2]$

8.2.30. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 4x_1 + 2x_2^2 - 2x_1x_2 - x_1^2 \to \min,$$
$$x_1^2 - 2x_2 \le 0, \quad 2x_1 + x_2 \le 6, \quad x_1 \ge 0, \quad x_2 \ge 0.$$

řešení: $[0,0]$

8.2.31. Nechť $\rho \in \mathbb{R}$ a uvažte úlohu matematického programování

$$f(x_1,x_2) = (x_1-1)^2 + x_2^2 \to \min, \quad x_1 - \rho\, x_2^2 \le 0. \tag{8.2.42}$$

Pro jaké hodnoty ρ je bod $x^* = [0,0]$ řešením úlohy (8.2.42)?

řešení: $\rho \le 1/2$

8.2.32. Firma vyrábějící letadla je může provozovat ve dvou zemích. Její náklady v zemi A závisí na počtu provozovaných letadel $x_1 \ge 0$ a jsou dány předpisem $C_A(x_1) = \ln(1+3x_1/100)$. Pro zemi B jsou $C_B(x_2) = \ln(1+3x_1/100)$. Firma chce rozdělit své kapacity mezi počty x_1 a x_2 tak, aby minimalizovala své celkové náklady při celkovém provozování alespoň $q > 0$ kusů.

a) Ukažte, že optimální volba x_1 a x_2 musí být řešením jistého optimalizačního problému s omezením na znaménko.

b) Ukažte, že v závislosti na hodnotě q existují jeden, dva nebo tři kandidáti na řešení.

c) Ukažte, že v optimálním případě provozuje firma letadla pouze v zemi A, je-li q pod kritickou hodnotou q^*, a pouze v zemi B, je-li $q > q^*$.

d) Určete optimální náklady v závislosti na q a ukažte, že tato funkce není diferencovatelná v bodě q^*. Je v tomto bodě subdiferencovatelná? Pokud ano, určete příslušný subdiferenciál.

řešení: a) $\ln(1+3x_1/100) + 2\ln(1+x_2/100) \to \min$ při omezení $x_1 + x_2 \ge q$, $x_1, x_2 \ge 0$

b) pro $q \in (0, 50/3)$ je jediný kandidát $[0,q]$, pro $q = 50/3$ jsou dva kandidáti $[0,q]$ a $[q,0]$,
pro $q > 50/3$ jsou tři kandidáti $[0,q]$, $[q,0]$ a $[q/3+100/9, 2q/3-100/9]$

c) pro $q \le 100$ je řešením $[0,q]$, pro $q \ge 100$ je řešením $[q,0]$, tj. $q^* = 100$

d) $\partial C^*(q^*) = \emptyset$

8.2.33. *Dokažte:* Nechť

$$X = \{x \in \mathbb{R}^n \mid x_i \ge 0,\ i \in I\},$$

kde $I \subseteq \{1,\dots,n\}$ je indexová množina, a nechť $f : X \to \mathbb{R}$ je diferencovatelná na (nějaké otevřené množině obsahující) X. Je-li $x^* \in X$ lokálním minimem funkce f na X, potom platí

$$\frac{\partial f}{\partial x_i}(x^*) \ge 0 \quad \& \quad x_i^* \cdot \frac{\partial f}{\partial x_i}(x^*) = 0 \quad \text{pro všechna } i \in I \tag{8.2.43}$$

a současně

$$\frac{\partial f}{\partial x_i}(x^*) = 0 \quad \text{pro všechna } i \in \{1,\dots,n\} \setminus I \tag{8.2.44}$$

Je-li funkce navíc f konvexní na X, pak bod $x^* \in X$ splňující dvojici podmínek (8.2.43) a (8.2.44) je dokonce globálním minimem funkce f na X.

8.2.34. *Dokažte:* Nechť

$$X = \{x \in \mathbb{R}^n \mid \alpha_1 \le x_i \le \beta_i,\ i = 1,\dots,n\},$$

kde $\alpha_i, \beta_i \in \mathbb{R}$ jsou daná čísla, a nechť $f : X \to \mathbb{R}$ je diferencovatelná na (nějaké otevřené množině obsahující) X. Je-li $x^* \in X$ lokálním minimem funkce f na X, potom platí

$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} \ge 0, & \text{pokud } x_i^* = \alpha_i, \\ \le 0 & \text{pokud } x_i^* = \beta_i, \\ = 0 & \text{pokud } \alpha_i < x_i^* < \beta_i. \end{cases} \tag{8.2.45}$$

Je-li funkce f navíc konvexní na X, pak bod $x^* \in X$ splňující podmínku (8.2.45) je dokonce globálním minimem funkce f na X.

8.2.35. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 4x_1^2 - x_1\,x_2 + 2x_2^2 \to \min$$

na obdélníku

$$X = \{[x_1,x_2] \in \mathbb{R}^2 \mid 4 \le x_1 \le 8,\ -1 \le x_2 \le 2\}.$$

řešení: $[4,1]$

8.2.36. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = 2x_1^2 + x_1x_2 + x_2^2 \to \min$$

na „obdélníku“

$$X = \{[x_1,x_2] \in \mathbb{R}^2 \mid -1 \le x_1 \le 1,\ x_2 \ge 2\}.$$

řešení: $[-1/2, 2]$

8.2.37. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,x_2) = \mathrm{e}^{a^2x_1} + \mathrm{e}^{a^2x_2} + 2ax_1 - x_2 \to \min$$

na čtverci

$$X = \{[x_1,x_2] \in \mathbb{R}^2 \mid 0 \le x_1 \le 1,\ -1 \le x_2 \le 0\}.$$

8.2.38. *Dokažte:* Nechť

$$X = \left\{x \in \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^n x_i = r\right\},$$

kde $r \in \mathbb{R}$ je dané číslo, a nechť $f : X \to \mathbb{R}$ je diferencovatelná na (nějaké otevřené množině obsahující) X. Je-li $x^* \in X$ lokálním minimem funkce f na X, potom platí implikace

$$x_i^* > 0 \implies \frac{\partial f}{\partial x_i}(x^*) \le \frac{\partial f}{\partial x_j}(x^*) \quad \text{pro všechna } j \in \{1,\dots,n\}. \tag{8.2.46}$$

Je-li funkce f navíc konvexní na X, pak bod $x^* \in X$ splňující podmínku (8.2.46) je dokonce globálním minimem funkce f na X.

8.2.39. Nechť $a_1,\dots,a_n > 0$ jsou daná čísla. Pomocí Lagrangeova principu vyřešte úlohu matematického programování

$$f(x_1,\dots,x_n) = x_1^{a_1}\cdots x_n^{a_n} \to \max$$

na množině (jednotkovém simplexu)

$$X = \left\{x \in \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^n x_i = 1\right\}.$$

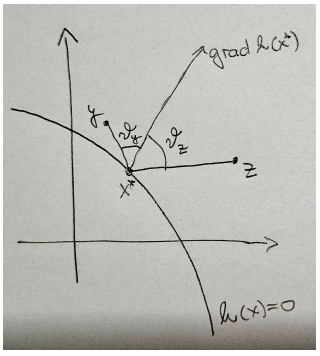
8.2.40. Vyřešte *Tartagliovu úlohu*:

Rozdělte číslo 8 na dvě části tak, aby součin těchto čísel vynásobený jejich rozdílem byl maximální.

řešení: $4 \pm 4/\sqrt{3}$

8.2.41. Ukažte, že mezi všemi pyramidami, které mají tutéž trojúhelníkovou základnu a stejnou výšku, ta s minimálním povrchem (tj. součtem obsahů bočních trojúhelníků) má následující vlastnost:

projekcí vrcholu pyramidy ležícího mimo základnu do trojúhelníku tvořícího základny je bod, jenž má stejnou vzdálenost od všech stran trojúhelníku.



Obrázek 8.24: Obrázek k *zákonu odrazu*, kdy chceme ukázat, že úhly ϑ_y a ϑ_z jsou shodné.

8.2.42. Vyřešte *Steinerův problém*:

V rovinném trojúhelníku najděte takový bod, že součet jeho vzdáleností od vrcholů trojúhelníku je minimální.

řešení: Fermatův(–Torricelliho) bod

8.2.43. Podobně jako v Příkladě 8.2.6 uvažte hladkou křivku v $\mathbb{R}^2$ určenou rovnicí $h(x) = 0$ pro spojitě diferencovatelnou funkci $h : \mathbb{R}^2 \to \mathbb{R}$. Nechť y a z jsou body, které leží na téže straně určené danou křivkou, viz Obrázek 8.24. Ukažte, že pokud bod x^* minimalizuje součet vzdáleností

$$\|y - x\| + \|z - x\|$$

mezi všemi body ležícími na dané křivce, pak úhly ϑ_y a ϑ_z vyznačené na Obrázku 8.24 musí být shodné. Tento fakt je v optice znám jako *zákon odrazu*.

8.3 Teorie (Lagrangeovy) duality

Při řešení úloh matematického programování hraje velmi důležitou roli tzv. *teorie duality*, které jsme se věnovali již v Podkapitole 2.6 a mnohem stručněji i v Podkapitole 4.3. Naši výchozí úlohu (8.1) a (8.2) můžeme označit jako *primární* a v této podkapitole si ukážeme, jak získat duální úlohu a jaké jsou vlastnosti. Její zásadní výhoda tkví v tom, že se jedná o úlohu konvexního programování, takže pak už stačí „jenom“ najít stacionární bod odpovídající Lagrangeovy funkce. Jenže jak řešení duální úlohy souvisí s řešením primární úlohy? A jak jej z duální úlohy získat? Odpovědi na tyto otázky jsou hlavním cílem této podkapitoly. Začneme však nejdříve s následujícím pojmem, který je úzce spojen s duální úlohou.

DEFINICE 8.3.1

Vektor $y^* \in Q$ se nazývá *Kuhnovým–Tuckerovým vektorem* úlohy (8.1) a (8.2), jestliže

$$f^* \le f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = L(x, y^*) \quad \text{pro všechna } x \in P, \tag{8.3.1}$$

kde $f^* := \inf_{x \in X} f(x)$ je hodnota úlohy (8.1) a (8.2).

Existuje Kuhnův–Tuckerův vektor vždy? Pochopitelně ne, jak bude patrné z následujících několika jednoduchých příkladů.

Příklad 8.3.2

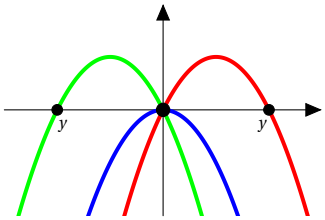
(i) Uvažme nejdříve úlohu

$$-x^2 \to \min \quad \& \quad x = 0 \quad \& \quad P = \mathbb{R}^2.$$

Pak zjevně jediným přípustným bodem a zároveň řešením je $x^* = 0$. Proto nerovnost (8.3.1) vede k hledání $y \in \mathbb{R}$ takového, že

$$0 = f^* \le -x^2 + yx = x(y - x) \quad \text{pro všechna } x \in \mathbb{R}. \tag{8.3.2}$$

Jenže pro $y \neq 0$ nabývá kvadratická funkce $g(x) := x(y-x)$ kladných i záporných hodnot a pro $y = 0$ nabývá pouze nekladných hodnot, viz Obrázek 8.25, tj. nerovnost (8.3.2) není splněna pro žádné $y \in \mathbb{R}$, a tudíž pro danou úlohu neexistuje Kuhnův–Tuckerův vektor.



Obrázek 8.25: Graf kvadratické funkce $g(x) := x(y-x)$ z Příkladu 8.3.2(i) pro $y > 0$ (červená barva), $y < 0$ (zelená) a $y = 0$ (modrá). Její kořeny jsou 0 a y.

(ii) Podívejme se nyní na úlohu

$$x - 1 \to \min \quad \& \quad x^2 - 1 = 0 \quad \& \quad P = \mathbb{R}_+.$$

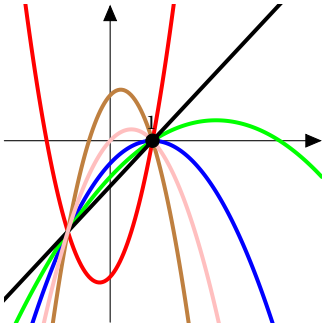
Jediným přípustným bodem a zároveň řešením je $x^* = 1$. Proto nerovnost (8.3.1) vede k hledání $y \in \mathbb{R}$ takového, že

$$0 = f^* \le x - 1 + y(x^2 - 1) = yx^2 + x - 1 - y \quad \text{pro všechna } x > 0.$$

Je-li $y = 0$, dostáváme nerovnost $0 \le x - 1$ neboli $1 \le x$, což je spor s požadavkem $x > 0$. Pro $y \neq$ má kvadratická funkce $h(x) := yx^2 + x - 1 - y$ kořeny

$$x_{1,2} = \frac{-1 \pm \sqrt{1 + 4y(y+1)}}{2y} = \frac{-1 \pm \sqrt{(2y+1)^2}}{2y} = \frac{-1 \pm |2y+1|}{2y}$$

a vrchol v $x = -1/(2y)$. Je-li $y \in \mathbb{R} \setminus \{0, -1/2\}$, pak máme dva reálné kořeny 1 a $-(1+y)/y$, přičemž $-(1+y)/y < 0$ pro $y \in (-\infty, -1) \cup (0, \infty)$, zatímco $-(1+y)/y > 0$ pro $y \in (-1, 0) \setminus \{-1/2\}$ a pro $y = -1$ je druhý kořen nulový. Každopádně funkce $h(x)$ nabývá na intervalu $(0, \infty)$ kladných i záporných hodnot. Pro $y = -1/2$ máme dvojnásobný kořen a $h(x) < 0$ pro všechna $x \in (0,1) \cup (1, \infty)$, viz Obrázek 8.26. To znamená, že ani pro tuto úlohu Kuhnův–Tuckerův vektor neexistuje.



Obrázek 8.26: Graf kvadratické funkce $h(x) := yx^2 + x - 1 - y$ z Příkladu 8.3.2(ii) pro $y > 0$ (červená barva), $y \in (-1, 0) \setminus \{-1/2\}$ (zelená), $y < -1$ (hnědá), $y = -1$ (růžová), $y = -1/2$ (modrá) a $y = 0$ (černá).

(iii) Uvažme ještě také úlohu

$$e^x \to \min \quad \& \quad x \le 0 \quad \& \quad P = \mathbb{R}.$$

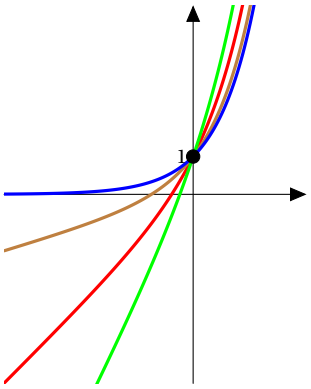
Tato úloha sice nemá řešení, avšak $f^* = \inf_{x\le 0} e^x = 0$. Nerovnost (8.3.1) proto vede k hledání $y \ge 0$ takového, že

$$0 = f^* \le e^x + y\,x \quad \text{pro všechna } x \in \mathbb{R}.$$

Tato nerovnost je jistě splněna pro $y = 0$. Je-li ovšem $y > 0$, pak funkce $\ell(x) := e^x + y\,x$ nabývá kladných i záporných hodnot, neboť např. pro libovolné $y > 0$ máme

$$\lim_{x\to-\infty} e^x + y\,x = -\infty \quad \text{a} \quad \lim_{x\to\infty} e^x + y\,x = \infty,$$

viz Obrázek 8.27. Proto $y = 0$ je jediný Kuhnův–Tuckerův „vektor“ příslušný dané úloze.



Obrázek 8.27: Graf funkce $\ell(x) := e^x + y\,x$ z Příkladu 8.3.2(iii) pro $y = 1$ (červená barva), $y = 2$ (zelená), $y = 3/10$ (hnědá) a $y = 0$ (modrá).

Ačkoli každý z uvedených příkladů je trochu jiný, ukazují nám, že ne pro každou úlohu matematického programování existuje Kuhnův–Tuckerův vektor a že jeho existence nemusí být závislá na řešitelnosti dané úlohy. Všimněme si, že první dva příklady nejsou úlohami konvexního programování, neboť v (i) není účelová funkce konvexní a v (ii) máme rovnostní omezení dáno funkcí, která není afinní. Třetí příklad již je úlohou konvexního programování, ale řešení neexistuje, zatímco Kuhnův–Tuckerův vektor ano. Kdybychom chtěli zabránit „problémům“ s neexistencí Kuhnova–Tuckerova vektoru, museli bychom místo regulární Lagrangeovy funkce $L(x, y^*)$ v (8.3.1) uvážit $L(x^*, y_0^*, y^*)$, v kterémžto případě bychom z Věty 6.3.21 dostali následující tvrzení.

Věta 8.3.3

Nechť úloha (8.1) a (8.2) je úlohou konvexního programování, tj. množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_k$ konvexní na P a $g_{k+1}, \dots, g_m$ afinní. Je-li množina X neprázdná, pak existuje $y_0^* \ge 0$ a $y^* \in Q$ takové, že $(y_0, y^\top) \ne 0$ a platí

$$y_0^* f^* \le y_0^* f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = L(x, y_0^*, y^*) \quad \text{pro všechna } x \in P. \tag{8.3.3}$$

Důkaz. Jestliže $f^* = -\infty$, pak nerovnost (8.3.3) platí pro libovolné $y_0^* > 0$ a $y^* \in Q$. Nechť tedy dále je $f^* > -\infty$. Uvažme systém

$$\begin{aligned} f(x) - f^* &< 0, \\ g_i(x) &< 0 \quad \text{pro } i = 1, \dots, k, \\ g_j(x) &= 0 \quad \text{pro } j = k+1, \dots, m. \end{aligned}$$

Podle definice f^* tento systém nemůže mít řešení na P. Proto z Věty 6.3.21 vyplývá existence m-tice konstant $y_0^*, y_1^*, \dots, y_k^* \ge 0$ a $y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$ takových, že ne všechny jsou současně nulové a platí

$$y_0^* \left[f(x) - f^* \right] + \sum_{i=1}^{m} y_i^* g_i(x) \ge 0 \quad \text{pro všechna } x \in P,$$

tj. platí nerovnost (8.3.3). ∎

V [58, Podkapitole 6.1] se můžeme setkat s podobným konceptem *geometrického vektoru multiplikátorů* $y^* \in Q$, který splňuje

$$f^* = \inf_{x\in P} L(x, y^*)$$

Tato drobná změna nám trochu usnadní geometrickou interpretaci tohoto konceptu prostřednictvím množiny $\{(g(x), f(x)) \mid x \in P\}$ a opěrné nadroviny s normálovým vektorem $(y, 1)$, která protíná svislou osu v bodě $L(x, y)$. To je ale totéž jako v případě geometrické interpretace duální úlohy z diskutované v Poznámce 8.3.8.

Chceme-li, aby mohlo být $y_0^* = 1$, jsou k zaručení existence Kuhnova–Tuckerova vektoru potřeba některé dodatečné podmínky, které lze získat z regulárních modifikací Věty 6.3.21.

VĚTA 8.3.4

Nechť úloha (8.1) a (8.2) je úlohou konvexního programování, tj. množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_k$ konvexní na P a $g_{k+1}, \dots, g_m$ afinní, a nechť dále platí (alespoň) jedna z podmínek regularity:

(i) (*Slaterova*) $k = m$ a existuje $\overline{x} \in P$ takové, že $g_i(\overline{x}) < 0$ pro $i = 1, \dots, m$;

(ii) (*lineární*) množina P je polyedr, funkce $f, g_1, \dots, g_k$ jsou afinní a navíc $X \neq \emptyset$;

(iii) (*modifikovaná lineární*) množina P je polyedr a existuje $\ell \in \{0, \dots, k\}$ takové, že funkce $f, g_1, \dots, g_\ell$ jsou konvexní na relativně otevřené (vzhledem k $\operatorname{aff} P$) konvexní množině $\mathcal{U} \supseteq P$, funkce $g_{\ell+1}, \dots, g_m$ jsou afinní a existuje $\overline{x} \in X$ splňující $g_i(\overline{x}) < 0$ pro všechna $i \in \{1, \dots, \ell\}$;

(iv) (*modifikovaná Slaterova*) funkce $g_{\ell+1}, \dots, g_k$ jsou afinní pro nějaké $\ell \in \{0, \dots, k\}$ a existuje $\overline{x} \in \operatorname{ri} P \cap X$ takové, že $g_i(\overline{x}) < 0$ pro všechna $i \in \{1, \dots, \ell\}$.

Pak existuje Kuhnův–Tuckerův vektor úlohy (8.1) a (8.2).

Důkaz. Jestliže $f^* = -\infty$, pak nerovnost (8.3.1) platí pro libovolné $y^* \in Q$. Nechť tedy dále je $f^* > -\infty$ a připusťme, že je splněna podmínka (i). Uvažme systém nerovností

$$\begin{aligned} f(x) - f^* &< 0, \\ g_i(x) &< 0 \quad \text{pro } i = 1, \dots, m. \end{aligned}$$

Vzhledem k definici f^* nemůže tento systém mít řešení na P. Podle předpokladů existuje $\overline{x} \in P$ splňující dnou soustavu bez první nerovnosti. Proto z Věty 6.3.22 plyne existence m-tice konstant $y_1^*, \dots, y_m^* \geq 0$ takových, že platí

$$f(x) - f^* + \sum_{i=1}^{m} y_i^* g_i(x) \geq 0 \quad \text{pro všechna } x \in P,$$

tj. platí nerovnost (8.3.1).

V případě splnění podmínky (ii) plyne tvrzení z Důsledku 6.3.24 aplikovaného na systém

$$\begin{aligned} f(x) - f^* &< 0 \\ g_i(x) &\leq 0 \quad \text{pro } i = 1, \dots, k \\ g_i(x) &= 0 \quad \text{pro } i = k+1, \dots, m \end{aligned}$$

na množině P.

Je-li splněna podmínka (iii) nebo (iv) plyne tvrzení z vět uvedených v poznámkách 43 a 44 na stranách 639 a 641 aplikovaných na systém

$$\begin{aligned} f(x) - f^* &< 0 \\ g_i(x) &< 0 \quad \text{pro } i = 1, \dots, \ell \\ g_i(x) &\leq 0 \quad \text{pro } i = \ell+1, \dots, k \\ g_i(x) &= 0 \quad \text{pro } i = k+1, \dots, m \end{aligned}$$

na množině P. ∎

Úloha konvexního programování doplněná o některou z dodatečných podmínek uvedených ve Větě 8.3.4 se nazývá *regulární úlohou konvexního programování*[51]. Všimněme si, že podmínky (ii)–(iii) jsou silnější než analogické podmínky ve Větě 8.2.15 a požadavek lineární nezávislosti gradientů aktivních omezení zde nevystupuje vůbec. To je způsobeno tím, že nyní je potřeba skutečně zajistit existenci Kuhnova–Tuckerova vektoru, zatímco případná neexistence vhodného $y^* \in Q$ ve Větě 8.2.15 znamenala, že daná úloha nemá řešení. Např. pro úlohu konvexního programování

$$-\sqrt{x_1 x_2} \to \min \quad \& \quad x_1 \leq 0 \quad \& \quad [x_1, x_2] \in P = \mathbb{R}^2_+$$

[51] (*Terminologické „zamyšlení“.*) V předchozí podkapitole jsme se v úloze matematického programování zaměřili především na „kvalitu“ daného omezení:

omezení
⟆
kvalifikované omezení
⟆
ostrá komplementarita.

Pohled skrze Kuhnův–Tuckerův vektor nabízí trochu odlišný pohled:

úloha matematického programování
⟆
úloha konvexního programování
⟆
regulární úloha KP (kvalifikované omezení)
⟆
silně regulární úloha KP (existuje K–T vektor)

je splněna lineární podmínka z Věty 8.2.15, avšak nerovnost (8.3.3) je splněna pouze pro $y_0^* = 0$. Pro tuto úlohu neplatí žádná z podmínek Věty 8.3.4, jelikož množina X neobsahuje žádný bod s $x_1 < 0$, funkce f není lineární a současně je konvexní pouze na P. Na druhou stranu však zdůrazněme, že ani samotnou Větu 8.2.15 nemůžeme použít, neboť účelová funkce není diferencovatelná v bodě $x^* = [0,0]$, který je řešením dané úlohy. Dodejme ještě, že v literatuře lze najít i další podmínky zaručující „regularitu“ úlohy (8.1) a (8.2), tj. existenci Kuhnova–Tuckerova vektoru, např. P je konvexní a uzavřená množina, funkce $f, g_1, \dots, g_k$ konvexní a spojité na P, $g_{k+1}, \dots, g_m$ afinní a množina řešení úlohy (8.1) a (8.2) neprázdná a ohraničená.

Nyní si ukážeme jednoduchý motivační příklad pro duální úlohu z pohledu obecného matematického programování, srovnej s Příkladem 2.6.1.

Příklad 8.3.5

Uvažme jednoduchou úlohu matematického programování

$$x_1^2 + x_2^2 + 2x_1 \to \min \quad \& \quad x_1 + x_2 = 0. \tag{8.3.4}$$

Její řešení lze nalézt snadno např. díky vztahu $x_1 = -x_2$, který úlohu redukuje na hledání volného extrému pro funkci $2x_1^2 + 2x_1$. Tím získáme jediné řešení $x_1^* = -1/2$ a $x_2^* = 1/2$ s hodnotou $f^* = -1/2$.
My se ale na tuto úlohu podíváme z „duální strany“. Nejdříve bude naším cílem nalezení dolního odhadu hodnoty f^* s pomocí řešení úlohy na volné extrémy, tj. po vynechání daného omezení dostaneme úlohu

$$x_1^2 + x_2^2 + 2x_1 \to \min, \tag{8.3.5}$$

pročež nepochybně platí

$$\text{hodnota úlohy (8.3.5)} \le \text{hodnota úlohy (8.3.4)}.$$

Řešením úlohy (8.3.5) je $x_1 = -1$ a $x_2 = 0$ s hodnotou -1, což je skutečně menší než f^*. K nalezení jiného dolního odhadu použijeme následující „trik“: vezmeme libovolné reálné číslo y a uvážíme úlohu

$$x_1^2 + x_2^2 + 2x_1 + y(x_1 + x_2) \to \min \quad \& \quad x_1 + x_2 = 0,$$

která je zjevně ekvivalentní s (8.3.4). Nyní vynecháním daného omezení získáme

$$x_1^2 + x_2^2 + 2x_1 + y(x_1 + x_2) \to \min \tag{8.3.6}$$

a pro libovolné $y \in \mathbb{R}$ platí podobně jako výše

$$\text{hodnota úlohy (8.3.6)} \le \text{hodnota úlohy (8.3.4)}.$$

Snadno zjistíme, že řešením úlohy (8.3.6) je pro libovolné $y \in \mathbb{R}$ dvojice $x_1 = -1 - y/2$ a $x_2 = -y/2$, takže hodnota úlohy (8.3.6) je

$$-\frac{1}{2} - \frac{(y+1)^2}{2},$$

která je menší nebo rovna než $-1/2$. Např. pro $y = 0$ dostaneme -1, což je zcela v souladu s hodnotou úlohy (8.3.5). Jaký ale nejlepší odhad můžeme takto získat? Tato otázka není nic jiného než nalezení řešení úlohy

$$\max\{\text{hodnota úlohy (8.3.6)} \mid y \in \mathbb{R}\} = \max\left\{-\frac{1}{2} - \frac{(y+1)^2}{2} \;\middle|\; y \in \mathbb{R}\right\},$$

což je právě ona duální úloha. Jejím řešením je $y = -1$ s odpovídající funkční hodnotou $-1/2$, což se shoduje s hodnotou f^*. Navíc dosazením $y = -1$ do vyjádření x_1 a x_2 při řešení úlohy (8.3.6) dostaneme právě výše zmíněné $x_1^* = -1/2$ a $x_2^* = 1/2$.

V předchozím příkladu nám vyšlo f^* rovné hodnotě duální úlohy. Byla to náhoda? Odpověď zjistíme za chvíli. Teď už totiž máme vše potřebné k tomu, abychom si definovali obecnou duální úlohu a odvodili některé její základní vlastnosti včetně podmínek zaručující rovnost hodnot primární a duální úlohy.

DEFINICE 8.3.6

Nechť $y \in Q$. Definujme funkci

$$\varphi(y) := \inf_{x \in P} L(x,y) = \inf_{x \in P} \left\{ f(x) + \sum_{i=1}^{m} y_i\, g_i(x) \right\}$$

a označme její *efektivní definiční obor* jako Y, tj.

$$Y := \{ y \in Q \mid \varphi(y) > -\infty \}.$$

Pak úloha

$$\varphi(y) \to \max, \quad y \in Y, \tag{8.3.7}$$

se nazývá *duální úlohou* k úloze (8.1) a (8.2). Číslo

$$\varphi^* := \sup_{y \in Y} \varphi(y)$$

se nazývá *hodnotou duální úlohy* (8.3.7).

V Kapitolách 2 a 4 jsme se zmínili o tom, že duální úloha k úlohám lineárního a kvadratického programování je téhož typu. Tento fakt si nyní potvrdíme alespoň v případě úloh lineárního programování.

Příklad 8.3.7

Určeme duální úlohu pro obecnou úlohu lineárního programování

$$c^\top x \to \min, \quad x \in X, \tag{8.3.8}$$

kde vektor $c \in \mathbb{R}^n$ je dán a množina X je určena podmínkami

$$\sum_{j=1}^{n} a_{ij}\, x_j = \langle a_i, x \rangle \geq b_i, \quad i = 1, \ldots, k$$

$$\sum_{j=1}^{n} a_{ij}\, x_j = \langle a_i, x \rangle = b_i, \quad i = k+1, \ldots, m$$

$$x_j \geq 0, \quad j = 1, \ldots, s$$

pro dané vektory $a_1, \ldots, a_m \in \mathbb{R}^n$, čísla $b_1, \ldots, b_m \in \mathbb{R}$ a $s \in \{0, \ldots, n\}$.
Řešení. Jestliže řádky matice A jsou tvořené postupně vektory $a_1, \ldots, a_m$, tj.

$$A := \begin{pmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{pmatrix} \in \mathbb{R}^{m \times n},$$

pak příslušná regulární Lagrangeova funkce má podobu

$$L(x,y) = \langle c, x \rangle + \langle y, b - Ax \rangle = \langle y, b \rangle + \langle c - A^\top y, x \rangle, \quad x \in P,\ y \in Q,$$

přičemž podmínku nezápornosti některých proměnných bereme jako přímá omezení definující množinu P, tj. $P := \{ x \in \mathbb{R}^n \mid x_j \geq 0,\ j \in \{1, \ldots, s\} \}$. Pak pro každé $y \in Q$ máme

$$\varphi(y) = \inf_{x_1, \ldots, x_s \geq 0} L(x,y), \tag{8.3.9}$$

přičemž platí

$$\operatorname{grad}_x L(x,y) = c - A^\top y.$$

Budeme-li $y \in Q$ považovat v (8.3.9) za parametr, pak máme úlohu konvexního programování bez funkcionálních omezení, takže stačí najít $\widetilde{x} \in P$ splňující[52]

$$\langle c - A^\top y, x - \widetilde{x} \rangle \geq 0 \quad \text{pro všechna } x \in P\,.$$

To v souladu s Poznámkou 8.1.7(i) po rozepsání do jednotlivých souřadnic znamená, že hledáme $\widetilde{x}$ splňující

$$\left.\begin{aligned} (c - A^\top y)_i \geq 0 \quad \text{a} \quad \widetilde{x}_i\,(c - A^\top y)_i = 0, \quad \text{pro } i = 1,\dots,s, \\ (c - A^\top y)_i = 0 \quad \text{pro } i = s+1,\dots,n. \end{aligned}\right\} \tag{8.3.10}$$

Pokud pro dané $y \in Q$ takové $\widetilde{x} \in P$ nalezneme, pak se v něm realizuje infimum $L(x,y)$ a platí $\langle c - A^\top y, \widetilde{x} \rangle = 0$ neboli

$$\varphi(y) = \langle y, b \rangle + \inf_{x_1,\dots,x_s \geq 0} \langle c - A^\top y, x \rangle = \langle y, b \rangle.$$

Je-li některá z podmínek v (8.3.10) porušena, pak hodnota infima[53] je $-\infty$. Proto je přípustná množina Y duální úlohy dána systém rovností a nerovností z (8.3.10), tj.

$$Y = \Big\{ y \in Q \;\Big|\; \sum_{i=1}^m a_{ij}\, y_i \leq c_j,\ \sum_{i=1}^m a_{ir}\, y_i = c_r,\ j \in \{1,\dots,s\},\ r \in \{s+1,\dots,n\} \Big\}.$$

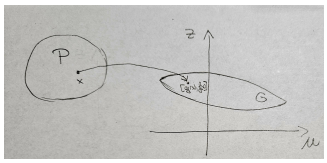
Duální úlohou tedy celkem je

$$\langle y, b \rangle \to \max, \quad y \in Y,$$

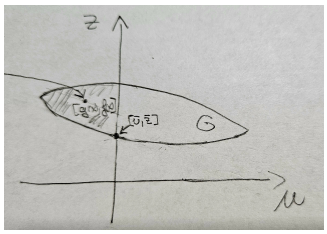
což je skutečně opět úloha lineárního programování.

[52] Zde místo x^* píšeme raději $\widetilde{x}$, neboť symbol x^* používáme pro řešení primární úlohy (8.3.8), zatímco nyní $\widetilde{x}$ je bodem, ve kterém se realizuje infimum Lagrangeovy funkce z (8.3.9).

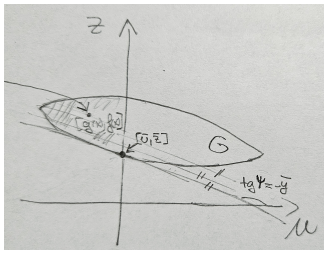
[53] V případě $(c - A^\top y)_i < 0$ pro nějaké $i \in \{1,\dots,s\}$ můžeme vzít x^* s $x^*_j = 0$ pro $j \neq i$ a $x_i \to \infty$. Podobně pro $(c - A^\top y)_i \neq 0$ můžeme vzít x^* s $x^*_j = 0$ a $x^*_i \to \pm\infty$ dle znaménka $(c - A^\top y)_i$. V obou případech bude $\langle c - A^\top y, x^* \rangle \to \infty$.



Obrázek 8.28: Geometrická vizualizace duální úlohy – množina G.



Obrázek 8.29: Geometrická vizualizace duální úlohy – řešení primární úlohy $[\overline{u},\overline{z}]$.



Obrázek 8.30: Geometrická vizualizace duální úlohy – hledání funkce $\varphi(y)$ je totéž jako hledání „nejnižší“ přímky se směrnicí $-y$ a mající neprázdný průnik s G. V případě bodu $[\overline{u},\overline{z}]$ to bude přímka se směrnicí $-\overline{y}$.

Poznámka 8.3.8. Můžeme také poskytnout geometrický pohled na řešení duální úlohy, který ukazuje, že tento problém je úzce spojen s opěrnými nadrovinami. Uvažme pro jednoduchost úlohu

$$f(x) \to \min \quad \& \quad g(x) \leq 0 \quad \& \quad x \in P$$

a označme jako G obraz množiny P skrze zobrazení zprostředkované vektorovou funkcí $\binom{g}{f}$, tj.

$$G := \big\{ [u,z] \in \mathbb{R}^2 \mid u = g(x) \ \&\ z = f(x) \ \&\ x \in P \big\},$$

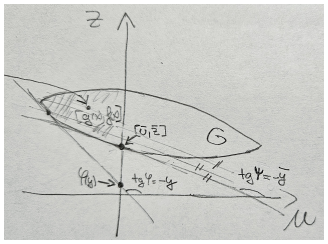
viz Obrázek 8.28. Primární úloha pak je vlastně nalezení bodu v množině G takového, že $u \leq 0$ a souřadnice z má nejmenší možnou hodnotu, na Obrázku 8.29 se jedná o bod $[\overline{u},\overline{z}]$. Nechť je nyní dáno $y \geq 0$. K určení $\varphi(y)$ musíme minimalizovat $f(x) + y\,g(x)$ na množině P. To je ale totéž jako minimalizace $z + y\,u$ na G. Současně $z + y\,u = \alpha$ je rovnice přímky se směrnicí $-y$ a průsečíkem s osou z v bodě $[0,\alpha]$, viz Obrázek 8.30. Proto k minimalizaci $z + y\,u$ na G potřebujeme posouvat tuto přímku směrem dolů tak dlouho dokud bude mít neprázdný průnik s množinou G. Jinými slovy, posouváme tuto přímku dolů tak dlouho, dokud se z ní nestane opěrná „nadrovina“ množiny G, tj. „dotkneme“ se množiny G zdola a odpovídající hodnota α dává $\varphi(y)$, viz Obrázek 8.31. Duální úloha pak je ekvivalentní s nalezením směrnice přímky (opěrné „nadroviny“ ke G) takové, že průsečík na ose z má největší možnou hodnotu. Na Obrázku 8.30 to je právě $-\overline{y}$ a tato odpovídající přímka prochází právě bodem $[\overline{u},\overline{z}]$, tj. v takovém případě je řešením duální úlohy $\overline{y}$ a $\varphi^* = \overline{z} = f^*$ (tj. platí vztah duality, viz později Větu 8.3.16). Nicméně rovnost $\varphi^* = f^*$ není splněna automaticky, jak můžeme vidět na Obrázku 8.32. ▲

Následující tvrzení ukazuje hlavní výhodu duální úlohy – bez ohledu na primární úlohu (8.1) a (8.2) to je vždy úloha konvexního programování, tudíž při jejím řešení nám postačí najít jen stacionární bod. Pouze místo minima konvexní funkce hledáme maximum konkávní funkce, o čemž ale již velmi dobře víme, že je vzájemně ekvivalentní. Nicméně abychom nenabyli mylného dojmu – přechod od primární k duální úloze není úplně triviální a skrývá v sobě nemalé[54]

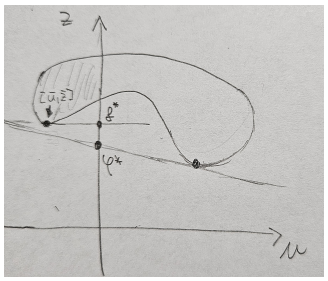
úsilí s ohledem na definici funkce $\varphi(\cdot)$. Výhodné to je zejména ve chvíli, kdy umíme snadno určit odpovídající duální úlohu (jako v případě lineárního či kvadratického programování), avšak i v takovém případě potřebujeme ještě navíc zaručit spojení mezi řešením primární a duální úlohy, čemuž se budeme věnovat v další části.

Věta 8.3.9

Úloha (8.3.7) je úlohou konkávního programování, tj. množina Y je konvexní a funkce φ je konkávní na Y.



Obrázek 8.31: Geometrická vizualizace duální úlohy – další přímka určující $\varphi(y)$. Řešení duální úlohy pak znamená nalezení takové přímky s nejvýše položeným průsečíkem na ose z. V tomto případě to je přímka se směrnicí $-\bar{y}$, tj. platí $\varphi^* = f^*$.



Obrázek 8.32: Geometrická vizualizace duální úlohy – tentokrát neplatí rovnost $\varphi^* = f^*$. Ovšem nezapomínejme, že f^* hledáme pouze v té části množiny G, kde $u \le 0$.

Důkaz. Nechť $y_1, y_2 \in Y$ a $\lambda \in [0,1]$ jsou libovolná. Potom platí

$$\varphi\big(\lambda y_1 + (1-\lambda) y_2\big) = \inf_{x\in P} L\big(x, \lambda y_1 + (1-\lambda) y_2\big) = \inf_{x\in P} \big[\lambda L(x,y_1) + (1-\lambda) L(x,y_2)\big] \ge$$
$$\ge \lambda \inf_{x\in P} L(x,y_1) + (1-\lambda) \inf_{x\in P} L(x,y_2) = \lambda\varphi(y_1) + (1-\lambda)\varphi(y_2),$$

tj. funkce φ je skutečně konkávní. Současně odtud také plyne, že pro libovolná $y_1, y_2 \in Y$ je také[55] $\lambda y_1 + (1-\lambda) y_2 \in Y$, neboť $\varphi\big(\lambda y_1 + (1-\lambda) y_2\big) \ge \lambda\varphi(y_1) + (1-\lambda)\varphi(y_2) > -\infty$, tj. množina Y je konvexní. ■

Jak již bylo naznačeno, míříme k jednomu z hlavních výsledků v teorii duality, kterým jsou postačující podmínky pro rovnost $f^* = \varphi^*$, v kterémžto případě stačí následně najít „pouze“ řešení rovnice $f(x) = \varphi^*$ pro $x \in X$, čímž dostaneme řešení úlohy 8.1 a 8.2. Později uvidíme, že to spojení bude ještě sofistikovanější.

Jinou motivaci pro rovnost $f^* = \varphi^*$ můžeme najít v teorii her. Definujme na P funkci[56]

$$\psi(x) := \sup_{y\in Q} L(x,y) = \begin{cases} f(x), & x \in X, \\ \infty, & x \in P \setminus X. \end{cases}$$

Je-li $X = \emptyset$, pak formálně položíme $f^* = \infty$, tj. $\sup_{y\in Q} L(x,y) = \infty$ pro každé $x \in P$. Podobně, pro $Y = \emptyset$ položíme $\varphi^* = -\infty$ pro každé $y \in Q$. Vždy pak platí $\inf_{x\in X} f(x) = \inf_{x\in P} \psi(x)$, z čehož plyne

$$f^* = \inf_{x\in P} \psi(x) = \inf_{x\in P} \sup_{y\in Q} L(x,y) \quad \text{a zároveň dle definice} \quad \varphi^* = \sup_{y\in Q} \inf_{x\in P} L(x,y).$$

Tedy rovnost $f^* = \varphi^*$ zaručuje možnost záměny pořadí suprema a infima v Lagrangeově funkci, což je důležité např. ve zmíněné teorii her[57].

Věta 8.3.10 von Neumannova o „minimaxu“

Nechť $X \subseteq \mathbb{R}^n$ a $Y \subseteq \mathbb{R}^n$ jsou kompaktní a konvexní množiny. Nechť $f : X \times Y \to \mathbb{R}$ je spojitá funkce na $X \times Y$ taková, že pro každé $x \in X$ je $f(x,\cdot) : Y \to \mathbb{R}$ konvexní a pro každé $y \in Y$ je $f(\cdot, y) : X \to \mathbb{R}$ konkávní. Potom

$$\max_{x\in X} \min_{y\in Y} f(x,y) = \min_{y\in Y} \max_{x\in X} f(x,y) \tag{8.3.11}$$

a existuje dvojice $[x^*, y^*] \in X \times Y$ taková, že pro každé $x \in X$ a $y \in Y$ platí

$$f(x,y^*) \le f(x^*,y^*) \le f(x^*,y). \tag{8.3.12}$$

Navíc pokud označíme společnou hodnotu v (8.3.11) jako v, pak $f(x^*,y^*) = v$.

Bod $[x^*, y^*]$ s vlastností (8.3.12) se nazývá *sedlový bod* funkce f, viz také Definici 8.3.25. Navíc požadovaná spojitost funkce f a kompaktnost množin X, Y zaručují existenci uvedených minim a maxim. Pozor – rovnost (8.3.11) ale rozhodně není splněna pro libovolnou spojitou funkci f definovanou na $X \times Y$, kde X a Y jsou konvexní a kompaktní množiny. Uvažme např. $X = Y = [0,1] \subseteq \mathbb{R}$ a funkci $f(x,y) = (x-y)^2$. Pak pro pevné $x \in X$ je $\min_{y\in Y} f(x,y) = f(x,x) = 0$, tj. levá strana (8.3.11) má nulovou hodnotu. Na druhou stranu pro pevné $y \in Y$ je

$$\max_{x\in X} f(x,y) = \begin{cases} (1-y)^2, & y \in [0,1/2], \\ y^2, & y \in [1/2,1], \end{cases}$$

[54] Toto může být problém i leckterý výpočetní software, neboť zde vystupuje y jako parametr. Výhodu zde bude mít software, který umí počítat symbolicky namísto numerických výpočtů. Avšak ani ten si nemusí umět poradit se všemi úskalími.

[55] Nezapomeňme, že $\varphi(y) > -\infty$ pro všechna $y \in Y$ a současně $\varphi(y) < \infty$ pro všechna $y \in Y$. V jakém případě by mohla být druhá nerovnost porušena? Jinými slovy, v jakém případě by se mohlo stát $\inf_{x\in P} L(x,y) = \infty$?

tj. levá strana (8.3.11) má hodnotu 1/4. V obecném případě se zcela libovolnou funkcí na jakékoli množině $X \times Y$ je potřeba pracovat pouze se supremem a infimem a rovnost v (8.3.11) nahradit obecně platnou nerovností

$$\sup_{x\in X}\inf_{y\in Y} f(x,y) \le \inf_{y\in Y}\sup_{x\in X} f(x,y), \tag{8.3.13}$$

což je známo jako tzv. *Max-min nerovnost*[58]. Ještě poznamenejme, že požadavek konvexnosti funkce $f(\cdot,y)$ na X a konkávnosti $f(x,\cdot)$ na Y je pouze postačující, což můžeme snadno ilustrovat např. volbou $X = Y = [0,1]$ a funkce $f(x,y) = \mathrm{e}^x\,\mathrm{e}^y$, která je konvexní v obou proměnných, v kterémžto případě bude na obou stranách rovnosti (8.3.11) číslo e.

Poznámka 8.3.11 (dlouhá o teorii her). Nejjednodušším případem tzv. *her dvou hráčů s nulovým součtem*[59] jsou maticové hry. Tyto hry hrají dva hráči – Červený a Modrý. Modrý volí přirozené číslo i z konečné množiny $I := \{1,\dots,m\} \subset \mathbb{N}$. Červený současně a bez ohledu na Modrého volí přirozené číslo j z podobné konečné množiny $J := \{1,\dots,n\} \subset \mathbb{N}$. *Výplata* (*odměna*) pro Modrého v případě, že zahrál i a Červený zahrál j, je předem určena jako reálné číslo $a(i,j)$. V této situaci je výplata pro Červeného[60] $-a(i,j)$. *Strategie* Modrého/Červeného je pravidlo, které jim říká, jaké $i \in I$ nebo $j \in J$ mohou zvolit[61]. Cílem Modrého je zvolit $i \in I$ tak, aby maximalizoval $a(i,j)$. Podobně cílem Červeného je zvolit $j \in J$ tak, aby minimalizoval $a(i,j)$, tj. maximalizoval $-a(i,j)$. Množina strategií se zatím skládá jen z *čistých strategií* pro hrače, tj. např. pro Modrého to je strategie *hrát* $i_0 \in I$.

Takové hry mohou být také popsány následovně pomocí matic. Nechť $A \in \mathbb{R}^{m\times n}$ je daná matice s prvky $a(i,j)$ pro $i \in I$ a $j \in J$. Modrý pak volí i-tý řádek a Červený j-tý sloupec. Volba je učiněna současně a bez znalosti soupeřova rozhodnutí. V případě hry kámen, nůžky, papír vypadá tato „matice“ následovně.

	kámen	nůžky	papír
kámen	0	+1	−1
nůžky	−1	0	+1
papír	+1	−1	0

Tabulka 8.1: Tabulka výplat pro hru kámen-nůžky-papír.

My se budeme věnovat námořní bitvě zvané „Plukovník Blotto“[62]. Dvě červené nákladní lodě jsou chráněny třemi doprovodnými fregatami pod vedením kapitána Pluta. Celá flotila je pod útokem skupiny čtyř modrých ponorek pod velením kapitána Nema. Lodě převáží nebezpečný náklad, a proto si udržují rozestup tak, že jedna ponorka nemůže zaútočit na obě lodě. Problém kapitána Nema je určení rozmístění ponorek pro útok na obě lodě. Může je rozmístit tak, že všechny zaútočí pouze na jednu loď a druhou nechají, nebo tak, že tři ponorky zaútočí na jednu loď a čtvrtá ponorka na druhou loď, nebo 2 ponorky na první loď a zbývající dvě ponorky na druhou loď. Tedy možnosti (strategie) kapitána Nema jsou dvojice $(4,0)$, $(0,4)$, $(3,1)$, $(1,3)$ a $(2,2)$, kde (α,β) značí, že α ponorek zaútočí na první loď a β ponorek na druhou loď. Problém kapitána Pluta je naopak rozmístění fregat k obraně lodí. K tomu má čtyři možnosti (strategie), z nichž každá může být zapsána jako uspořádaná dvojice (γ,δ) s $\gamma,\delta \in \{0,1,2,3\}$ a $\gamma+\delta = 3$, kde γ fregat chrání první loď a δ fregat ochrání druhou loď. Předpokládejme, že okolnosti na moři jsou takového rázu, že kapitán Nemo nemůže při rozmisťování ponorek určit polohu fregat a ani kapitán Pluto nemůže při rozmisťování fregat určit polohu ponorek.

Pokud daný počet útočících ponorek překročí počet bránících fregat, pak výplata kapitána Nema je rovna počtu fregat plus 1, kde hodnota +1 je odměna za (předpokládané) potopení nákladní lodi. Je-li ale počet útočících ponorek menší než počet bránicích fregat, pak jeho výplata je minus (počet útočících ponorek +1), kde +1 je odečteno za to, že nákladní loď může pokračovat do cíle. Je-li počet ponorek a fregat stejný, pak jde o remízu a výplata je 0. Celková

[56] Pro jistotu zdůrazněme, že pro pevně zvolené $x \in X$ je buď $g_i(x) < 0$ nebo $g_i(x) = 0$, takže v takovém případě je $L(x,y) \le f(x)$ pro všechna $y \in Q$ a $\sup_{y\in Q} L(x,y) = f(x)$.

[57] Autorem Věty 8.3.10 je von Neumann, který ji v roce 1928 zveřejnil v článku [501]. Toto tvrzení později okomentoval slovy *As far as I can see, there could be no theory of games … without that theorem … I thought there was nothing worth publishing until the Minimax Theorem was proved.* (*Podle toho, co vím, nemohla by existovat žádná teorie her … bez této věty … Myslel jsem si, že neexistuje nic, co by stálo za zveřejnění, dokud nebude dokázána Věta o minimaxu.*), viz [95, str. 19].

[58] Její důkaz je celkem snadný. Označme pro každé $x \in X$ označme $g(x) := \inf_{y\in Y} f(x,y)$. Pak na $X \times Y$ platí $g(x) \le f(x,y)$. Vezmeme-li nyní libovolné $y \in Y$, pak $\sup_{x\in X} g(x) \le \sup_{x\in X} f(x,y)$. Avšak tato nerovnost má platit pro všechna $y \in Y$, tudíž nutně máme také pro nejmenší možnou hodnotu pravé strany, tj.

$$\sup_{x\in X} g(x) \le \inf_{y\in Y}\sup_{x\in X} f(x,y),$$

což je právě nerovnost (8.3.13).

[59] Tj. co jeden získá, druhý ztratí, takže spolupráce nemá smysl. Např. kámen, nůžky, papír o 1 Kč (každý učiní tuto sázku, tj. výhra není z externího zdroje a nic se nikam nepřevádí), kdy tuto výhru získá buď jeden nebo druhý (v případě remízy se hra opakuje, ale sázka se nenavyšuje, nebo se sázky vrací).

[60] Skutečně je to hra s nulovým součtem.

[61] Množiny I a J představují *prostor strategií*, např. kámen, nůžky, papír.

[62] *Colonel Blotto game*

výplata kapitánu Nemovi je součet výplat při obou soubojích. Výplata kapitánu Plutovi pak je minus výplata Nemovi. Tato hra může být „maticově“ zapsána následující způsobem.

kapitán Pluto / kapitán Nemo	(3,0)	(0,3)	(2,1)	(1,2)
(4,0)	4	0	2	1
(0,4)	0	4	1	2
(3,1)	1	−1	3	0
(1,3)	−1	1	0	3
(2,2)	−2	−2	2	2

Tabulka 8.2: Tabulka výplat pro hru Plukovník Blotto.

Zvolí-li kapitán Nemo řádek i, pak má zaručenu výplatu nejméně

$$\min\{a(i,j) \mid j = 1,2,3,4\}.$$

Pokud zvolí řádek, který mu dá výplatu[63]

$$v^- := \max_i \min_j a(i,j),$$

pak má zaručeno, že získá nejméně v^-. Kapitán Nemo tak může zvolit první nebo druhý řádek, aby měl zaručenu výplatu alespoň 0, tj. $v^- = 0$. Na druhou stranu, zvolí-li kapitán Pluto j-tý sloupec, ztratí nejvýše

$$\max\{a(i,j) \mid i = 1,2,3,4,5\}.$$

Pokud kapitán Pluto zvolí řádek, který kapitánu Nemovi dává výplatu[64]

$$v^+ := \min_j \max_i a(i,j),$$

má zaručeno, že neztratí více než v^+. Proto kapitán Pluto může zvolit třetí nebo čtvrtý sloupec, což mu zaručí, že neztratí více než 3.

Pokud by kapitán Nemo věděl (třeba „telepaticky“), že kapitán Pluto zvolí třetí sloupec, pak by vybral třetí řádek, což by mu přineslo výplatu $3 > v^- = 0$. Na druhou stranu, kdyby kapitán Pluto věděl, že kapitán Nemo zvolí druhý řádek, zvolil by první sloupec, čímž by ztratil 0 jednotek, což je méně než ztráta $3 = v^+$ jednotek. Při pohledu na matici výplat je vidět, že pokud jeden velitel zná volbu druhého, pak *může* vylepšit svoji volbu k dosažení většího zisku nebo menší ztráty. Tedy v této námořní bitvě je práce v tajnosti naprosto klíčová. Co je ale nejlepší strategie pro každého hráče?

K zodpovězení této otázky budeme muset nejdříve odhlédnout od této konkrétní námořní bitvy a uvažme obecnou maticovou hru s maticí výplat

$$\begin{pmatrix} 3 & 7 & 4 & 3 \\ 13 & 10 & 7 & 8 \\ 10 & 4 & 1 & 9 \\ 3 & 5 & 6 & 7 \end{pmatrix}.$$

Zde je

$$\max_i \min_j a(i,j) = \min_j \max_i a(i,j) = 7.$$

[63] V každém řádu určíme nejmenší hodnotu (minimální možná výplata) a z těchto hodnot následně vybereme tu největší (maximalizujeme minimální výplatu).

[64] V každém sloupci nejdříve určíme největší hodnotu (maximální možná ztráta) a z těchto hodnot následně vyberemu tu nejmenší (minimalizujeme maximální ztrátu).

V tomto případě má strategie $(i,j) = (i^*, j^*) = (2,3)$ tu vlastnost, že $a(i^*, j^*) = 7$ a

$$a(i, j^*) \le a(i^*, j^*) \le a(i^*, j)$$

pro všechna $i, j = 1,2,3,4$. To znamená, že *hodnota hry* je 7 a $i^* = 2$ je *optimální* (čistá) *strategie* pro Modrého (volí řádek) a $j^* = 3$ je *optimální* (čistá) *strategie* pro Červeného (volí sloupec), neboť zvolí-li Modrý optimálně $i^* = 2$, pak nejlepší volba Červeného je $j^* = 3$ a naopak (pokud Červený zvolí $j^* = 3$, pak nejlepší volba pro Modrého je $i^* = 2$). Tedy pokud hráč oznámí svůj tah dopředu a zvolí-li přitom optimální strategii, soupeř nemůže využít této informace k vlastnímu prospěchu jinak, než že i on také zvolí optimální strategii. V obecné maticové hře s $A \in \mathbb{R}^{m\times n}$ platí[65]

$$v^- = \max_i \min_j a(i,j) \le \min_j \max_i a(i,j) = v^+.$$

Pokud je splněna i opačná nerovnost $v^+ \ge v^-$, pak definujeme

$$v := \max_i \min_j a(i,j) = \min_j \max_i a(i,j)$$

a říkáme, že maticová hra *má cenu* v v čistých strategiích. V takovém případě existuje dvojice (i^*, j^*) taková, že

$$a(i^*, j^*) = v \tag{8.3.14}$$

a pro všechna $i = 1,\dots,m$ a $j = 1,\dots,n$ platí

$$a(i, j^*) \le a(i^*, j^*) \le a(i^*, j). \tag{8.3.15}$$

Strategie i^*, j^* se nazývají *optimální* (čisté) *strategie*[66], přičemž nemusí být jediné. Dvojice strategií (i^*, j^*) se nazývá *sedlový bod* v čistých strategiích. Jestliže má maticová hra sedlový bod v čistých strategiích, pak má hra cenu v a platí (8.3.14). Má-li maticová hra cenu v, pak Modrý (jeho cílem je maximalizovat výplatu) má zaručenu výplatu alespoň v a Červený (jeho cílem je minimalizovat výplatu) má zaručeno, že ztratí nejvýše $-v$. Význam nerovností (8.3.15) tkví v tom, že pokud Modrý zvolí optimální strategii a Červený nikoli, pak Modrý může získat získat více, než kdyby Červený zahrál optimální strategii (a naopak).

Ovšem jak hrát, pokud hra nemá sedlový bod jako naše námořní bitva? *Smíšenou strategií*[67] pro Modrého v maticové hře s $A \in \mathbb{R}^{m\times n}$ rozumíme rozdělení pravděpodobnosti mezi jednotlivé řádky A, tj. jedná se o vektor $x = (x_1,\dots,x_m)^\top$, kde $x_i \ge 0$ pro všechna $i = 1,\dots,m$ a $\sum_{i=1}^m x_i = 1$. Množinu všech těchto smíšených strategií označíme pro další použití jako P_m. Podobně pro Červeného je smíšenou strategií rozdělení pravděpodobnosti mezi jednotlivé sloupce A, tj. vektor $y = (y_1,\dots,y_n)^\top$ z množiny P_n, kde $y_i \ge 0$ pro všechna $i = 1,\dots,n$ a $\sum_{i=1}^n y_i = 1$. Množina čistých strategií je zjevně podmnožinou smíšených strategií, neboť čistá strategie pro Modrého odpovídá volbě $x = e_1 = (0,\dots,0,1,0,\dots,0)^\top$ s hodnotou 1 na i-tém řádku, což znamená, že zvolíme i-tý řádek s pravděpodobností 1.

Uvažme, že v naší hře Modrý použije k výběru řádku náhodný generátor, který dá hodnotu i s pravděpodobností x_i, přičemž v takovém případě Modrý skutečně zvolí ve hře i-tý řádek. Červený postupuje obdobně. Je-li použita smíšená strategie, žádný hráč nezkouší odhadnout tah protihráče a oba věří ve štěstí. Navíc pokud Modrý/Červený užije smíšenou strategii, pak soupeř nemůže tuto informaci ani nikterak využít ve svůj prospěch. Je-li povolena smíšená strategie, pak musíme definovat výplatu maticové hry jako *očekávanou hodnotu* výplaty, tj. pokud označíme výplatu odpovídající volbě strategií (x,y) jako $P(x,y)$, pak

$$P(x,y) := x^\top A y = \langle x, Ay\rangle.$$

V maticové hře se smíšenými strategiemi, kde Modrý volí $x \in P_m$ a Červený $y \in P_n$, je výplata Modrému daná jako $P(x,y)$ a Červenému jako $-P(x,y)$. Řekneme, že maticová hra se smíšenými strategiemi *má cenu*, pokud

$$\max_{x\in P_m} \min_{y\in P_n} P(x,y) = \min_{y\in P_n} \max_{x\in P_m} P(x,y).$$

[65] Tzv. *Max–Min inquality.*

[66] Toto mimo jiné úzce souvisí s tzv. *Nashovou rovnováhou*. Tu si intuitivně můžeme představit následovně. Připusťme, že hráči navzájem znají svoje strategie. Pak se každý hráč může zeptat sám sebe: „Pokud znám strategie ostatních a budu je brát za pevně dané, mohu změnou mé strategie získat nějaký prospěch?“. Pokud bude odpovědí „ano“, pak takováto situace není Nashovou rovnováhou. Pokud ale všichni hráči odpoví „ne“, pak je toto řešení stabilním a označujeme ho za Nashovu rovnováhu. Nashovu rovnováhu tedy můžeme označit jako nejlepší možnou reakci na strategie ostatních.

[67] Koncept smíšených strategií byl zaveden Johnem von Neumannem a Oscarem Morgensternem v jejich knize *Theory of Games and Economic Behavior* z roku 1944.

Toto číslo se nazývá *cenou hry* a opět se značí jako v. Řekneme, že hra má *sedlový bod* (x^*, y^*), kde x^* je optimální smíšená strategie pro Modrého a y^* pro Červeného, pokud

$$P(x, y^*) \le P(x^*, y^*) \le P(x^*, y)$$

pro všechna $x \in P_m$ a $y \in P_n$. Bez dalších podrobností zdůrazněme, že sedlový bod pro takové hry existuje vždy.

Věta. Pro libovolnou maticovou hru vždy existuje sedlový bod (x^*, y^*) ve smíšených strategiích a hra má cenu

$$v = P(x^*, y^*) = \max_{x\in P_m} \min_{y\in P_n} P(x, y) = \min_{y\in P_n} \max_{x\in P_m} P(x, y).$$

Výše zmíněná námořní bitva má cenu $v = 14/9$ s optimálními smíšenými strategiemi $x^* = (4/9, 4/9, 0, 0, 1/9)^\top$ a $y^* = (1/18, 1/18, 4/9, 4/9)^\top$. K nalezení tohoto řešení je možné využít lineárního programování[68] , zatímco hry s nenulovým součtem by nás přivedly ke kvadratickému programování. ▲

Nerovnost (8.3.13) by nás mohla inspirovat k hledání podobného výsledku pro primární a duální úlohy matematického programování. To je obsahem následující věty, která ukazuje, že hodnota primární úlohy nikdy není menší než hodnota duální úlohy.

Pro každé $x \in X$ a každé $y \in Q$ platí

$$f(x) \ge \varphi(y).$$

Zejména, pokud $X \ne \emptyset$ a $Y \ne \emptyset$, pak[69] $f^* \ge \varphi^*$.

Věta 8.3.12 Slabá věta o dualitě

Důkaz. Pro každé $x \in X$ a $y \in Q$ platí

$$f(x) \ge f(x) + \underbrace{\sum_{i=1}^{m} y_i\, g_i(x)}_{\le 0} = L(x, y) \ge \inf_{x\in X} L(x, y) \overset{X\subseteq P}{\ge} \inf_{x\in P} L(x, y) \ge \varphi(y).$$

Tato nerovnost musí platit také v případě, že na levé straně vezmeme infimum pro $x \in X$ a na pravé straně supremum pro $y \in Y$, což dává $f^* \ge \varphi^*$. ■

Věta 8.3.12 ukazuje, že je-li $y \in Q$ libovolné, pak hodnota $\varphi(y)$ vždy udává dolní hranici minima (infima) účelové funkce f úlohy (8.1) a (8.2). V případě $X \ne \emptyset$ a $Y \ne \emptyset$ jsou dokonce obě úlohy ohraničené zdola/shora, a tudíž i řešitelné alespoň ve smyslu infima a suprema, tj. určení f^* a φ^*. Navíc odtud vidíme, že rozdíl mezi hodnotou účelové funkce primární a duální úlohy je vždy nezáporný, tj. pro libovolná $x \in X \ne \emptyset$ a $y \in Y \ne \emptyset$ platí

$$g(x, y) := f(x) - \varphi(y) \ge 0.$$

Číslo $g(x^*, y^*) := f^* - \varphi^*$ udává tzv. (*optimální*) *duální rozdíl*[70]. Tím získáváme „jednoduchý“ test přípustnosti/řešitelnosti primární a duální úlohy (viz také Důsledek 8.3.17):

(i) je-li *primární úloha* (8.1) *a* (8.2) (zdola) *neohraničená* (tj. $f^* = -\infty$), pak nutně $\varphi^* = -\infty$, tj. $Y = \emptyset$ neboli *duální úloha* (8.3.7) *je nepřípustná*;

(ii) je-li *duální úloha* (8.3.7) (shora) *neohraničená* (tj. $\varphi^* = \infty$), pak nutně $f^* = \infty$, tj. $X = \emptyset$ neboli *primární úloha* (8.1) *a* (8.2) *je nepřípustná*.

Toto pozorování si ilustrujeme na jednoduché úloze v následujícím příkladu. Zároveň zdůrazněme, že uvedené implikace platí pouze jedním směrem, tj. neohraničenost jedné úlohy implikuje nepřípustnost druhé úlohy, zatímco samotná nepřípustnost jedné úlohy může nastat i v případě nepřípustné či řešitelné druhé úlohy, viz Příklady 8.3.15(i)–(iii).

[68] Výpočet hodnoty

$$\max_{x\in P_m} \min_{y\in P_n} P(x, y)$$

je ekvivalentní s úlohou

$$\min_{y} e_i^\top P\, x \to \max$$
$$\sum_{j=1}^{m} x_j = 1$$
$$x_1 \ge 0 \quad \& \quad \dots \quad \& \quad x_m \ge 0,$$

což můžeme s pomocí nové proměnné $v \in \mathbb{R}$ přepsat jako úlohu lineárního programování

$$v \to \max$$
$$v \le e_1^\top P\, x$$
$$\vdots$$
$$v \le e_n^\top P\, x$$
$$\sum_{j=1}^{m} x_j = 1$$
$$x_1 \ge 0 \quad \& \quad \cdots \quad \& \quad x_m \ge 0$$

neboli s pomocí vektoru $e = (1, \dots, 1)^\top \in \mathbb{R}^m$ jako

$$v \to \max$$
$$v\, e - P\, x \le 0$$
$$e^\top x = 1$$
$$x \ge 0.$$

Analogicky, hodnotu

$$\min_{y\in P_n} \max_{x\in P_m} P(x, y)$$

můžeme najít pomocí úlohy lineárního programování

$$w \to \max$$
$$w\, e - P^\top y \le 0$$
$$e^\top y = 1$$
$$y \ge 0,$$

kde tentokrát $e = (1, \dots, 1)^\top \in \mathbb{R}^n$.

[69] V případě $X = \emptyset$ a/nebo $Y = \emptyset$ je nerovnost $f^* \ge \varphi^*$ splněna triviálně, neboť $\inf \emptyset = \infty$ a $\sup \emptyset = -\infty$.

[70] *(Optimal) duality gap*

Příklad 8.3.13

Uvažme úlohu

$$x \to \min \quad \& \quad g(x) \equiv \pi \le 0 \quad \& \quad P = \mathbb{R}_+,$$

která je zjevně nepřípustná, takže $f^* = \infty$. To je v souladu s tím, že účelová funkce duální úlohy je

$$\varphi(y) = \inf_{x \ge 0} [x + \pi y] = \pi y,$$

takže $Y = \mathbb{R}_+$ a $\varphi^* = \infty$, tj. duální úloha je neohraničená.

Bezprostředně z Věty 8.3.12 také obdržíme následující velmi „jednoduché“ kritérium (v podobě postačující podmínky) pro ověření optimality. Jeho praktické využití sice může být poněkud diskutabilní, ale jako „certifikát optimality“ funguje dokonale, viz také Větu 8.3.20.

Důsledek 8.3.14 Certifikát optimality

Jsou-li $x^* \in X$ a $y^* \in Q$ taková, že platí

$$f(x^*) = \varphi(y^*),$$

pak x^* a y^* jsou optimálními řešeními svých příslušných úloh.

Platí i opačné tvrzení? Věta 8.3.12 naznačuje, že odpověď je bez dalších požadavků záporná, tj. ne vždy je duální rozdíl nulový, což si ilustrujeme v následujícím příkladu. Pro úplnost však s předstihem dodejme, že v případě regulární úlohy konvexního programování již ekvivalence bude pravdivá, viz Větu 8.3.20.

Příklad 8.3.15

Uvažme úlohu z Příkladu 8.3.2(ii), tj.

$$x - 1 \to \min, \quad x^2 - 1 = 0, \quad x \ge 0.$$

Pak $X = \{1\}$, takže $f^* = 0$. Účelová funkce duální úlohy je[71]

$$\varphi(y) = \inf_{x \ge 0} [x - 1 + y(x^2 - 1)] = \begin{cases} -1 - y, & y \ge 0, \\ -\infty, & y < 0. \end{cases}$$

Duální úloha tedy je

$$-1 - y \to \max, \quad y \ge 0$$

a není těžké zjistit, že její hodnota je $\varphi^* = -1$, tj. máme ostrou nerovnost

$$0 = f^* > \varphi^* = -1.$$

Mohou ale nastat i „extrémnější“ situace, které si ilustrujeme na úlohách tvaru

$$f(x) \to \min \quad \& \quad g(x) \le 0 \quad \& \quad x \in P.$$

(i) V případě $f(x) = -x$, $g(x) \equiv \pi$ a $P = \mathbb{R}_+$ máme zjevně nepřípustnou úlohu, takže $f^* = \infty$. Účelová funkce duální úlohy je

$$\varphi(y) = \inf_{x \ge 0} [-x + \pi y] \equiv -\infty,$$

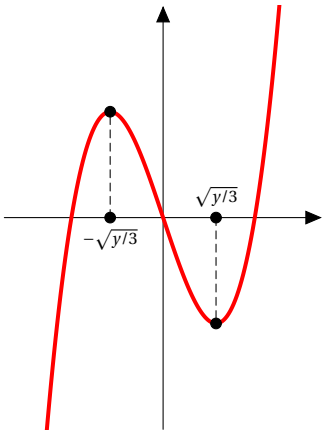
takže $Y = \emptyset$, tj. duální úloha je také nepřípustná, a tudíž $f^* = \infty > -\infty = \varphi^*$.

(ii) V případě $f(x) = x$, $g(x) = x$ a $P = \mathbb{R}_{++}$ máme opět zjevně nepřípustnou úlohu, takže $f^* = \infty$. Účelová funkce duální úlohy je

$$\varphi(y) = \inf_{x > 0} [x + xy] = \inf_{x > 0} [(1 + y)x] = \begin{cases} 0 & \text{pro } y \ge -1, \\ -\infty & \text{pro } y < -1, \end{cases}$$

takže $Y = \mathbb{R}_+$ a $\varphi^* = 0$, tj. duální úloha je řešitelná, ale $f^* = \infty > 0 = \varphi^*$.

[71] Toto plyne z faktu, že pro $h(x) = x - 1 + y(x^2 - 1)$ je $h'(x) = 1 + 2xy > 0$ pro $x, y \ge 0$, takže minimum nastává pro $x = 0$, zatímco v případě $x \ge 0$ a $y < 0$ je tato funkce klesající a zdola neohraničená na intervalu $(-1/(2y), \infty)$.



Obrázek 8.33: Graf funkce $(x^2 - y)\,x$ z Příkladu 8.3.15(iii).

(iii) V případě $f(x) = x^3$, $g(x) = -x$ a $P = \mathbb{R}$ máme $X = \mathbb{R}_+$ a $f^* = 0$. Účelová funkce duální úlohy je

$$\varphi(y) = \inf_{x\in\mathbb{R}} \left[x^3 - x\,y\right] = \inf_{x\in\mathbb{R}} \left[(x^2 - y)\,x\right] \equiv -\infty,$$

neboť stacionární body $x_{1,2} = \pm\sqrt{y/3}$ funkce $h(x) = (x^2 - y)\,x$ určují pouze lokální extrémy, viz Obrázek 8.33. Proto $Y = \emptyset$, a tudíž $\varphi^* = -\infty$, tj. nalezli jsme řešitelnou primární úlohu s nepřípustnou duální úlohou, takže $f^* = 0 > -\infty = \varphi^*$.

Ohledně přípustnosti, ohraničenosti a řešitelnosti primární a duální úlohy může nastat celkem 9 možností. Věta 8.3.12 vylučuje 3 z nich a naopak 2 z nich potvrzuje, přičemž o zbývajících 4 zatím není možné rozhodnout, viz Tabulku 8.3.

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✔	✖	✖
PaO	?	?	✖
NP ($\varphi^* = -\infty$)	?	?	✔

Tabulka 8.3: Vztah primární a duální úlohy dle Věty 8.3.12. **PÚ**: primární úloha; **DÚ**: duální úloha; **NP**: nepřípustná úloha (tj. $X = \emptyset$ nebo $Y = \emptyset$); **NO**: neohraničená úloha (tj. $f^* = -\infty$ nebo $\varphi^* = \infty$); **PaO**: přípustná a ohraničená úloha (tj. existuje konečné f^* nebo φ^*).

Duální rozdíl je také úzce spojen s existencí Kuhnova–Tuckerova vektoru. Jestliže duální rozdíl je nenulový, tj. $f^* > \varphi^*$, pak[72] $X \neq \emptyset$ a $Y \neq \emptyset$ a *množina Kuhnových–Tuckerových vektorů musí být prázdná*[73]. Vskutku, připusťme, že existuje Kuhnův–Tuckerův vektor $y^* \in Q$, tj.

$$f^* \le \inf_{x\in P} L(x, y^*) = \varphi(y^*) \le \varphi^* < f^* \quad \lightning.$$

To ale znamená, že existence Kuhnova–Tuckerova vektoru zaručuje *nulový duální rozdíl*, což je obsahem následujícího tvrzení, která podtrhuje význam regulárních úloh konvexního programování. Připomeňme, že situaci $f^* = -\infty$ jsme již vyřešili.

Věta 8.3.16 Silná věta o dualitě

Nechť úloha (8.1) a (8.2) je regulární úlohou konvexního programování (viz Větu 8.3.4). Pokud $f^* > -\infty$, pak platí tzv. *vztah duality*

$$f^* = \varphi^*, \quad \text{tj.} \quad \inf_{x\in P} \sup_{y\in Q} L(x, y) = \sup_{y\in Q} \inf_{x\in P} L(x, y),$$

přičemž množina řešení duální úlohy (8.3.7) je neprázdná a shodná s množinou všech Kuhnových–Tuckerových vektorů úlohy (8.1) a (8.2).

Důkaz. Jelikož máme regulární úlohu konvexního programování, Věta 8.3.4 zaručuje existenci Kuhnova–Tuckerova vektoru $y^* \in Q$ úlohy (8.1) a (8.2), a tudíž (opět)

$$f^* \le \inf_{x\in P} L(x, y^*) = \varphi(y^*) \le \varphi^*.$$

Jelikož $f^* > -\infty$, je také $\varphi^* > -\infty$, a tedy $y^* \in Y \neq \emptyset$. Současně dané předpoklady zaručují $X \neq \emptyset$. Proto z Věty 8.3.12 máme $f^* \ge \varphi^*$, takže celkem $f^* = \varphi^*$.

Nechť nyní $y^* \in Y$ je řešením duální úlohy (8.3.7). Potom

$$f^* = \varphi^* = \varphi(y^*) = \inf_{x\in P} L(x, y^*) \le f(x) + \sum_{i=1}^{m} y_i^*\, g_i(x) \quad \text{pro všechna } x \in P,$$

tj. y^* je Kuhnovým–Tuckerovým vektorem úlohy (8.1) a (8.2).

[72] Je-li $X = \emptyset$, pak $f^* = \infty$, a tudíž také $\varphi^* = \infty$ $\lightning$. Podobně pro $Y = \emptyset$ je $\varphi^* = -\infty$, a tudíž $f^* = -\infty$ $\lightning$.

[73] Např. v první úloze v Příkladu 8.3.15 je existence Kuhnova–Tuckerova vektoru ekvivalentní existenci čísla $y \in \mathbb{R}$ takového, že

$$0 \le (x-1)\,(1 + y\,(x+1))$$

pro všechna $x \ge 0$. To je ale v případě $x - 1 \ge 0$ splněno pouze pro $1 + y\,(x+1) \ge 0$ neboli

$$y \ge -1/(x+1) \ge 0,$$

zatímco pro $x - 1 < 0$ musí být $0 \le x < 1$ a $1 + y(x+1) < 0$ neboli

$$y < -1/(x+1) < -1,$$

což potvrzuje neexistenci vhodného y.

Konečně, nechť $y^* \in Q$ je Kuhnovým–Tuckerovým vektorem úlohy (8.1) a (8.2). Pak platí

$$\varphi^* = f^* \le f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = L(x, y^*) \quad \text{pro všechna } x \in P$$

z čehož plyne $\varphi^* \le \inf_{x \in P} L(x, y^*) = \varphi(y^*)$, a tudíž z definice φ^* dostáváme rovnost $\varphi^* = \varphi(y^*)$. To ale znamená, že y^* je řešením úlohy (8.3.7). ■

Praktický význam tvrzení Věty 8.3.16 nemusí být na první pohled zřejmý (za daných předpokladů jsou totiž obě úlohami konvexního programování), ale v některých případech může být duální úloha jednodušší. Navíc v Kapitolách 2–4 jsme viděli, že některé algoritmy pro řešení úloh lineárního, celočíselného či kvadratického programování jsou založeny právě na řešení duální úlohy. Z této věty navíc bezprostředně vyplývají dvě důležité implikace.

DŮSLEDEK 8.3.17

Nechť úloha (8.1) a (8.2) je regulární úlohou konvexního programování (viz Větu 8.3.4).

(i) Jestliže $Y \neq \emptyset$, pak duální úloha je řešitelná a $f^* > -\infty$.

(ii) Jestliže $Y = \emptyset$, pak $f^* = -\infty$.

Důkaz. (i) Jestliže platí $Y \neq \emptyset$, z definice φ^* vyplývá $\varphi^* > -\infty$. Potom z Věty 8.3.12 plyne $f^* \ge \varphi^* > -\infty$, takže úloha (8.1) & (8.2) je řešitelná podle Věty 8.3.16.

(ii) Jestliže platí $Y = \emptyset$, pak nutně $f^* = -\infty$. V případě $f^* > -\infty$ bychom totiž dostali spor s Větou 8.3.16. ■

Z Vět 8.3.12 a 8.3.16 a z Důsledku 8.3.17 vyplývá, že v případě regulární úlohy konvexního programování mohou ve skutečnosti nastat pouze dvě možnosti[74] vyznačené v Tabulce 8.4.

[74] Nezapomeňme, že regularita primární úlohy zaručuje $X \neq \emptyset$, tj. primární úlohy nemůže být nepřípustná.

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✖	✖	✖
PaO	✖	✔	✖
NP ($\varphi^* = -\infty$)	✖	✖	✔

Tabulka 8.4: Vztah regulární primární úlohy a příslušné duální úlohy dle Vět 8.3.12 a 8.3.16.

V případě, kdy buď máme regulární úlohu konvexního programování nebo nepřípustnou úlohu (tj. $X = \emptyset$), může nastat celkem 4+1 možnost, které jsou uvedené v Tabulce 8.5.

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✔	✖	✖
PaO	?✔?	✔	✖
NP ($\varphi^* = -\infty$)	✔	✖	✔

Tabulka 8.5: Vztah primární duální úlohy v případě, kdy primární úloha je buď nepřípustná nebo regulární.

Oproti Tabulce 8.4 máme navíc 3 možnosti: obě úlohy mohou být nepřípustné (to není v rozporu s Větou 8.3.12, neboť $-\infty = \varphi^* \le f^* = \infty$); nepřípustná duální úloha a neohraničená primární úloha dle Tabulky 8.3; primární úloha nepřípustná a duální úloha řešitelná. Jednoduché příklady ilustrující tyto možnosti jsme viděli již v Příkladě 8.3.15. Avšak měli bychom ještě zmínit, že poslední uvedená možnost (označená v Tabulce 8.5 otazníky) není např. v úlohách lineárního a kvadratického programování možná. To je způsobeno tím, že duální úlohy

jsou stejného typu jako primární úlohy, takže záměnou role primární a duální úlohy bychom dostali spor s tím, že není možné, aby primární úloha byla řešitelná (a tedy nutně i regulární) a duální úloha byla nepřípustná.

Ovšem předpoklady Věty 8.3.16 dávají pouze postačující podmínky pro platnost vztahu duality. Jinými slovy, rovnost $f^* = \varphi^*$ může být splněna i pro „neregulární“ úlohy, jak můžeme vidět v následujícím příkladě.

Příklad 8.3.18

Uvažme např.

$$x \to \min, \quad x^2 \le 0.$$

Pak $X = \{0\}$, a tudíž $f^* = 0$. Účelová funkce duální úlohy je

$$\varphi(y) = \inf_{x\in\mathbb{R}} \left[x + y x^2\right] = \begin{cases} -\frac{1}{4y}, & y > 0, \\ -\infty, & y \le 0, \end{cases}$$

neboť funkce $x + y x^2 = x(1 + y x)$ má jediný stacionární bod $x = -1/2y$ pro $y > 0$ určující její globální minimum na $\mathbb{R}$. Proto $Y = (0,\infty)$ a duální úloha je

$$-\frac{1}{4y} \to \max, \quad y > 0.$$

Tato úloha sice nemá řešení, avšak platí $\varphi^* = \sup_{y>0}\left[-1/(4y)\right] = 0 = f^*$.

Tento příklad je jen ukázkou úlohy splňující požadavky následujícího tvrzení.

Nechť v úloze (8.1) a (8.2) je množina P uzavřená a konvexní, funkce $f, g_1, \dots, g_k$ spojité a konvexní na P, funkce $g_{k+1}, \dots, g_m$ afinní a množina řešení této úlohy je neprázdná a ohraničená. Potom $Y \ne \emptyset$ a platí $f^* = \varphi^*$.

Důkaz. Omezme se na situaci, kdy $k = m$, tj. nejsou žádná rovnostní omezení. Pro libovolné $\varepsilon > 0$ uvažme úlohu

$$f(x) \to \min \quad \& \quad g_1(x) \le \varepsilon \quad \& \quad \cdots \quad \& \quad g_m(x) \le \varepsilon \quad \& \quad x \in P, \tag{8.3.16}$$

přičemž pochopitelně pro $\varepsilon = 0$ dostaneme původní problém. Označme jako X_ε přípustnou množinu úlohy (8.3.16) a f_ε^* hodnotu této úlohy.

Protože $X \subset X_\varepsilon$, je $f_\varepsilon^* \le f^*$, což společně s předpoklady věty zaručuje, že úloha (8.3.16) má řešení[75] x_ε^*, tj. $x_\varepsilon^* \in X_\varepsilon$ a $f(x_\varepsilon^*) = f_\varepsilon^*$. Všimněme si, že pro jakékoli $\varepsilon \in (0,1]$ bod x_ε^* vyhovuje soustavě nerovností

$$f(x) \le f^*$$
$$g_1(x) \le 1 \quad \& \quad \cdots \quad \& \quad g_m(x) \le 1.$$

Každopádně[76] množina řešení tohoto systému je ohraničená a bez újmy na obecnosti můžeme předpokládat[77], že existuje $x^* := \lim_{\varepsilon\to 0} x_\varepsilon^*$. Pak $x^* \in X$, a protože $f(x_\varepsilon^*) \le f^*$, máme také $f(x^*) \le f^*$, takže z definice f^* plyne $f(x^*) = f^*$, tj. x^* je řešením úlohy (8.1) a (8.2). Tím pádem ze spojitosti f dostaneme, že

$$\lim_{\varepsilon\to 0} f_\varepsilon^* = \lim_{\varepsilon\to 0} f(x_\varepsilon^*) = f(x^*) = f^*. \tag{8.3.17}$$

Nyní se podívejme na duální úlohu pro (8.3.16). Tou je

$$\varphi_\varepsilon \to \max \quad \& \quad y \in Y_\varepsilon, \tag{8.3.18}$$

kde

$$\varphi_\varepsilon(y) := \inf_{x\in P}\left\{f(x) + \sum_{i=1}^m y_i\,(g_i(x) - \varepsilon)\right\} \quad \text{a} \quad Y_\varepsilon := \{y \in Q \mid \varphi_\varepsilon(y) > -\infty\}.$$

[75] **Věta.** *Uvažme problém*

$$f(x) \to \min$$
$$g_1(x) \le b_1$$
$$\vdots$$
$$g_m(x) \le b_m$$
$$x \in P,$$

kde $P \subseteq \mathbb{R}^n$ je konvexní uzavřená množina a $f, g_1, \dots, g_m$ spojité a konvexní funkce na P. Pro libovolné $b = (b_1, \dots, b_m)^\top \in \mathbb{R}^m$ označme jako $X(b)$ přípustnou množinu dané úlohy a jako $X^*(b)$ množinu jejich řešení. Je-li pro nějaké $\tilde b$ množina $X^*(\tilde b)$ neprázdná a ohraničená, pak pro libovolné $b \in \mathbb{R}^m$ s $X(b) \ne \emptyset$, je $X^*(b)$ neprázdná a ohraničená.

Důkaz. Pro $b \in \mathbb{R}^m$ a $\beta \in \mathbb{R}$ uvažme množiny

$$Y(b,\beta) := \{x \in X(b) \mid f(x) \le \beta\}.$$

Nechť $\tilde\beta := f(x)$ pro nějaké $x \in X^*(\tilde b)$. Množina $Y(\tilde b, \tilde\beta) = X^*(\tilde b)$ je podle předpokladů věty neprázdná a ohraničená. Uvažme libovolný vektor $b \in \mathbb{R}^m$ takový, že existuje $\tilde x \in X(b)$. Podle Důsledku C níže je také neprázdná množina $Y(b, f(\tilde x))$ ohraničená. Pak je množina $X^*(b)$ neprázdná díky Větě A níže a současně je ohraničená, neboť $X^*(b) \subseteq Y(b, f(\tilde x))$, čímž je důkaz hotov. ■

Věta A. *Nechť $X \subseteq \mathbb{R}^n$ je uzavřená množina a funkce f spojitá na X. Jestliže pro nějaké $\tilde x \in X$ je množina*

$$N(\tilde x) := \{x \in X \mid f(x) \le f(\tilde x)\}$$

ohraničená, pak existuje globální minimum funkce f na X.

Důkaz. Z uzavřenosti X a spojitosti f plyne, že množina $N(\tilde x)$ je uzavřená, a tedy i kompaktní. Podle Weierstrassovy věty pak nutně f nabývá své nejmenší hodnoty na $N(\tilde x)$, což vzhledem k definici $N(\tilde x)$ znamená, že to bude také bod minima f na X. ■

Věta B. *Nechť f je spojitá a konvexní funkce na uzavřené konvexní množině X. Pokud je množina $X_\beta := \{x \in X \mid f(x) \le \beta\}$ neprázdná a ohraničená pro nějaké $\beta \in \mathbb{R}$, pak je X_β ohraničená pro každé $\beta \in \mathbb{R}$.*

Důkaz. Nechť $\tilde\beta$ je právě takové číslo, že množina X_β je neprázdná a ohraničená pro $\beta = \tilde\beta$. Jestliže $\beta < \tilde\beta$, pak $X_\beta \subseteq X_{\tilde\beta}$ a ohraničenost X_β plyne ihned z ohraničenost $X_{\tilde\beta}$.

Nechť proto dále $\beta > \tilde\beta$ a připusťme, že X_β není ohraničená. Uvažme libovolný bod $x^* \in X_{\tilde\beta} \subseteq X_\beta$. Protože X_β je konvexní a uzavřená, existuje polopřímka vycházející z bodu x^* a ležící celá v X_β, tj. existuje vektor $h \neq 0$ takový, že

$$x^* + \alpha h \in X_\beta$$

pro všechna $\alpha \geq 0$, tj.

$$x^* + \alpha h \in X$$

a současně

$$f(x^* + \alpha h) \leq \beta, \qquad (*)$$

viz [466, Věta 1.15, str. 72]. Defi-

Věta 8.3.20 Karushova–Kuhnova–Tuckerova v nediferenciálním tvaru

nujme funkci $g(\alpha) := f(x^* + \alpha h)$ pro $\alpha \geq 0$. Pak není těžké přímým výpočtem ověřit, že taková funkce je také konvexní na $\mathbb{R}_+$. Připusťme, že $g(\tilde\alpha) > g(0)$ pro nějaké $\tilde\alpha > 0$. Pak pro jakékoli $\alpha > \tilde\alpha$ máme

$$g(\tilde\alpha) = g\big((1 - \tilde\alpha/\alpha)\cdot 0 + (\tilde\alpha/\alpha)\cdot\alpha\big) \leq$$
$$\leq (1 - \tilde\alpha/\alpha)\, g(0) + (\tilde\alpha/\alpha)\, g(\alpha),$$

z čehož dostáváme

$$g(\alpha) \geq (\alpha/\tilde\alpha)\big(g(\tilde\alpha) - g(0)\big) + g(0).$$

Protože $g(\tilde\alpha) > g(0)$, je $g(\alpha) \to \infty$ pro $\alpha \to \infty$, což je ve sporu s $(*)$ ↯. Tudíž pro jakékoli $\alpha > 0$ je

$$f(x^* + \alpha h) = g(\alpha) \leq$$
$$\leq g(0) = f(x^*) \leq \tilde\beta,$$

tj. $x^* + \alpha h \in X_{\tilde\beta}$, což je ale ve sporu s ohraničeností množiny $X_{\tilde\beta}$ ↯ a důkaz je hotov. ■

Důsledek C. *Nechť $f_1, \ldots, f_m$ jsou spojité konvexní funkce na uzavřené konvexní množině X. Pro $b = (b_1, \ldots, b_m)^\top \in \mathbb{R}^m$ označme množinu*

$$X(b) := \{x \in X \mid f_i(x) \leq b_i,\ i = 1, \ldots, m\}.$$

Je-li $X(\tilde b)$ neprázdná a ohraničená pro nějaké $\tilde b \in \mathbb{R}^m$, pak je $X(b)$ ohraničená pro všechna $b \in \mathbb{R}^m$.

Důkaz. Položme $f(x) := \max_{i=1,\ldots,m}\{f_i(x) - b_i\}$ a uvažme příslušné množiny X_β definované v předchozí Větě B. Předpoklady tvrzení nám zaručují, že $X_0 = X(\tilde b)$ je neprázdná a ohraničená. Navíc funkce f je konvexní a spojitá na X. Pro jakékoli b existuje β takové, že $X(b) \subseteq X_\beta$. Podle Věty B je množina X_β ohraničená, a tudíž také množina $X(b)$ je ohraničená. ■

Pro každé $y \in Q$ samozřejmě máme $\varphi_\varepsilon(y) \leq \varphi(y)$, v důsledku čehož je $Y_\varepsilon \subseteq Y$ a

$$\varphi_\varepsilon^* := \sup_{y \in Y_\varepsilon} \varphi_\varepsilon(y) \leq \sup_{y \in Y} \varphi(y) = \varphi^*. \qquad (8.3.19)$$

Je zřejmé, že úloha (8.3.16) splňuje Slaterovu podmínku pro libovolné $x \in X$. Z Věty 8.3.16 pak jednak vyplývá, že úloha (8.3.18) má řešení, a tudíž $Y_\varepsilon \neq \emptyset$, což implikuje $Y \neq \emptyset$, a jednak také $f_\varepsilon^* = \varphi_\varepsilon^*$. Odtud s ohledem na (8.3.17) a (8.3.19) obdržíme $f^* \leq \varphi^*$, což společně s Větou 8.3.12 konečně vede ke kýžené rovnosti $f^* = \varphi^*$. ■

Každopádně s využitím předchozích výsledků můžeme rozšířit tvrzení Karushovy–Kuhnovy-Tuckerovy věty (viz Větu 8.2.15) na případ regulární úlohy konvexního programování bez explicitního předpokladu diferencovatelnosti.

Nechť úloha (8.1) a (8.2) je regulární úlohou konvexního programování (viz Větu 8.3.4). Pak $x^* \in X$ je řešením této úlohy právě tehdy, když platí (alespoň) jedna z podmínek:

(i) existuje $y^* \in Q$ takové, že $f(x^*) = \varphi(y^*)$,

(ii) existuje $y^* \in Q$ takové, že

$$L(x^*, y^*) = \min_{x \in P} L(x, y^*), \qquad (8.3.20)$$

$$y_i^*\, g_i(x^*) = 0, \quad i \in \{1, \ldots, m\}. \qquad (8.3.21)$$

Navíc množina takovýchto vektorů $y^* \in Q$ splývá s množinou řešení duální úlohy[78] (8.3.7).

Důkaz.

(i) „$\Longrightarrow$“ Nechť x^* je řešením úlohy (8.1) & (8.2), tj. $f(x^*) = f^*$, a $y^* \in Q$ je Kuhnův–Tuckerův vektor (ten existuje díky regulárnosti úlohy). Pak podle druhé části Věty 8.3.16 je y^* řešením duální úlohy (8.3.7), tj. $\varphi(y^*) = \varphi^*$. Současně podle první části Věty 8.3.16 dostáváme $f(x^*) = f^* = \varphi^* = \varphi(y^*)$.

„$\Longleftarrow$“ Nechť $f(x^*) = \varphi(y^*)$ pro nějaké $x^* \in X$ a $y^* \in Q$. Jelikož platí $f(x^*) \leq f(x^*) = \varphi(y^*) \leq \varphi(y^*)$, plyne z Věty 4.3.5 $f^* = f(x^*)$ a $y^* = \varphi(y^*)$, tj. x^* je řešením úlohy (8.1) a (8.2) a y^* je řešením duální úlohy (8.3.7).

(ii) „$\Longrightarrow$“ Nechť x^* je řešením úlohy (8.1) a (8.2). Pak podle předchozí části existuje $y^* \in Q$ takové, že $f(x^*) = \varphi(y^*)$, tj.

$$f(x^*) = \varphi(y^*) \overset{\text{Definice } 8.3.6}{=} \inf_{x \in P} L(x, y^*) \leq f(x) + \sum_{i=1}^m y_i^*\, g_i(x) \quad \text{pro všechna } x \in P.$$

Odtud volbou $x = x^*$ dostaneme $\sum_{i=1}^m y_i^*\, g_i(x^*) \geq 0$. Jelikož ale $g_i(x^*) \leq 0$ a $y_i^* \geq 0$ pro $i = 1, \ldots, k$, musí platit (8.3.21), z čehož dále vyplývá

$$L(x^*, y^*) = f(x^*) + \underbrace{\sum_{i=1}^m y_i^*\, g_i(x^*)}_{=0} = f(x^*) = \varphi(y^*) \leq L(x, y^*) \quad \text{pro všechna } x \in P.$$

neboli

$$L(x^*, y^*) = \min_{x \in P} L(x, y^*)$$

„$\Longleftarrow$“ Nechť pro nějaké $x^* \in X$ a $y^* \in Q$ platí (8.3.20) a (8.3.21). Potom

$$f(x^*) = f(x^*) + \underbrace{\sum_{i=1}^m y_i^*\, g_i(x)}_{=0 \text{ dle } (8.3.21)} = L(x^*, y^*) \overset{(8.3.20)}{=} \min_{x \in P} L(x, y^*) = \inf_{x \in P} L(x, y^*) = \varphi(y^*)$$

Pak z části (i) vyplývá, že x^* je řešením úlohy (8.1) a (8.2). ■

[76] Podle Důsledku C v předchozí poznámce, přičemž pokud nahradíme 1 za 0 dostaneme soustavu určující množinu řešení úlohy (8.1) a (8.2), která je dle předpokladů neprázdná a ohraničená.

[77] „Posloupnost“ $\{x_\varepsilon\}$ je ohraničená, takže díky Bolzanově–Weierstrassově větě z ní lze vybrat konvergentní podposloupnost, a tudíž může vzít, že konvergentní je dokonce samotná posloupnost.

[78] A podle Věty 8.3.16 také s množinou Kuhnových–Tuckerových vektorů úlohy (8.1) a (8.2).

Poznámka 8.3.21. (i) Jsou-li navíc k předpokladům Věty 8.3.20 funkce $f, g_1, \dots, g_m$ diferencovatelné v bodě x^*, pak podmínka (8.3.20) je ekvivalentní s (8.2.3) pro $y_0^* = 1$, zatímco (8.3.21) odpovídá (8.2.4). Tudíž Věta 8.3.20 je skutečně zobecněním Karushovy-Kuhnovy–Tuckerovy věty pro případ nediferencovatelných funkcí. Při tomto pohledu bychom měli zdůraznit, že koncept Kuhnova–Tucekrova vektoru můžeme brát jako zobecnění Lagrangeových multiplikátorů (tj. vektorů $y^* \in Q$ splňujících (8.2.3) a (8.2.4) s $y_0^* = 1$), neboť tyto dva pojmy splývají při splnění podmínek Věty 8.2.15. Samotná Věta 8.3.20 je užitečná zejména ve chvíli, kdy je možné „snadno“ určit řešení duální úlohy. Pak se nalezení řešení primární úlohy redukuje na řešení rovnice $\varphi(y^*) = f(x^*)$ nebo systému (8.3.20) a (8.3.21) na X.

(ii) Kdy můžeme požadavek $x^* \in X$ při hledání řešení systému (8.3.20) a (8.3.21) „ignorovat“, tj. kdy bude tato podmínka splněna „automaticky“? Např. máme-li regulární úlohu konvexního programování, vektor y^* je řešením (8.3.7) a víme, že úloha (8.1) a (8.2) má řešení. Je-li $x^* \in \mathbb{R}^n$ jediný bod splňující $\varphi(y^*) = f(x^*)$ nebo podmínky (8.3.20) a (8.3.21), pak x^* je jediným řešením úlohy (8.1) a (8.2).

(iii) Z poslední části Věty 8.3.20 také plyne, jak s využitím řešení duální úlohy nalézt řešení primární úlohy. V případě regulární úlohy konvexního programování máme pro řešitelnou duální úlohu zaručenu rovnost $\varphi^* = f^*$, a tedy potřebujeme najít x^* takové, aby platilo (8.3.20). Jenže to již známe z výpočtu $\varphi(y)$, kde jsme určili $\inf_{x\in P} L(x, y)$ pro libovolné $y \in Q$. Stačí tedy tento výpočet využít a vyčíslit $\inf_{x\in P} L(x, y^*)$, viz též Důsledek 8.3.28 později.

(iv) Co je duální úloha k duální úloze? Přepíšeme-li duální úlohu do tvaru

$$-\varphi(y) \to \min, \quad y \in Y,$$

tak při hledání duální–duální úlohy máme dvě možnosti. Budeme-li podmínku $y \in Y$ považovat za přímé omezení, pak nemáme žádné funkcionální omezení a účelová funkce příslušné duální–duální úlohy bude

$$\psi(z) = \inf_{y\in Y} -\varphi(y) = -\sup_{y\in Y} \varphi(y) = -\sup_{y\in Y} \inf_{x\in P} L(x, y),$$

což je konstantní funkce a její určení je vlastně totéž jako samotné vyřešení duální úlohy. Druhou možností je považovat $y \in Y$ za funkcionální omezení – pokud tedy $Y \neq \mathbb{R}^m$, tj. pokud v primární úloze nejsou pouze rovnostní omezení a/nebo vyvstane nějaké omezení z požadavku $\varphi(y) > -\infty$, protože v opačném případě bychom v duální úloze hledali volný extrém, což není příliš zajímavé, viz např. Cvičení 8.3.26. Uvažme proto ještě situaci, kdy primární úloha má $0 < k < m$ nerovnostních omezení. Jestliže $\varphi(y) > -\infty$ pro všechna $y \in Q$, pak účelová funkce duální úlohy k duální úloze je

$$\psi(z) = \inf_{y\in\mathbb{R}^m} \{-\varphi(y) - z_1 y_1 - \dots - z_k y_k\},$$

pro kterou stačí najít (klasický) stacionární bod v závislosti na z. V případě, kdy primární úloha má dokonce pouze nerovnostní omezení, tj. $k = m$, dostaneme odtud

$$\psi(z) = \inf_{y\in\mathbb{R}^m} \{-\varphi(y) - z_1 y_1 - \dots - z_m y_m\} = -\sup_{y\in\mathbb{R}^m} \{\varphi(y) + \langle z, y\rangle\},$$

což je totožné s Fenchelovou transformací konvexní funkce $-\varphi(y)$, tj. $\psi(z) = -(-\varphi)^*(y)$. Obzvláště v situaci, kdy již v samotné primární úloze jsou jen a pouze omezení na znaménko jednotlivých proměnných $x \geq 0$, obdržíme

$$\varphi(y) = \inf_{x\in\mathbb{R}^n} \left\{f(x) - \langle y, x\rangle\right\} = -\sup_{x\in\mathbb{R}^n} \left\{\langle y, x\rangle - f(x)\right\} = -f^*(y),$$

což nás přivede k duální–duální účelové funkci $\psi(z) = -f^{**}(z)$, a tudíž otázka shody mezi primární a duální–duální úlohou je úzce spojena s případnou rovností $f = f^{**}$, viz Větu 6.6.13 a také Poznámku 8.4.16.

Nicméně spojení mezi primární a duální-duální úlohou je poněkud delikátnější. Stále totiž musíme mít na paměti, že v duální úloze hledáme maximum jenom na množině Y určené mimo jiné podmínkou $\varphi(y) > -\infty$. Ta s sebou může přinést i některá další omezení, jako je tomu např. v úloze lineárního nebo kvadratického programování, viz Příklad 8.3.7 a Podkapitoly 2.6 a 4.3 nebo Cvičení 8.3.27–8.3.30. ▲

Příklad 8.3.22

Pomocí vztahu duality vyřešme úlohu

$$\sum_{j=1}^{m} |x_j - a_j| \to \min, \quad \sum_{j=1}^{m} x_j = 0$$

pro daný vektor $a = (a_1, \dots, a_n)^\top \in \mathbb{R}^n$.
Řešení. Lagrangeova funkce této úlohy je

$$L(x, y) = \sum_{j=1}^{m} |x_j - a_j| + y \sum_{j=1}^{m} x_j$$

pro $y \in \mathbb{R}$. Protože

$$\inf_{x_j \in \mathbb{R}} \left[|x_j - a_j| + y x_j \right] = \begin{cases} y a_j, & |y| \le 1, \\ -\infty, & |y| > 1, \end{cases}$$

viz Obrázek 8.34, máme účelovou funkci duální úlohy

$$\varphi(y) = \begin{cases} A y, & |y| \le 1, \\ -\infty, & |y| > 1, \end{cases}$$

kde $A := \sum_{j=1}^{n} a_j$. Duální úloha tedy je

$$A y \to \max, \quad y \in Y = [-1, 1].$$

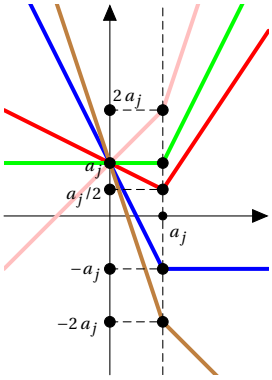
Jejím řešením je $y^* = \operatorname{sgn} A$ a hodnota duální úlohy je $\varphi(y^*) = |A|$. Ze vztahu duality pak plyne, že řešení x^* primární úlohy je řešením rovnice

$$\sum_{j=1}^{n} |x_j - a_j| = |A|, \quad \sum_{j=1}^{n} x_j = 0.$$

Je-li $A = 0$, pak musí být $x_j = a_j$ pro každé $j \in \{1, \dots, n\}$, tj. $x^* = a$. Je-li $A \neq 0$, můžeme zkusit hledat řešení ve tvaru $x_j^* = a_j - u_j A$ pro $j \in \{1, \dots, n\}$ a nějaká čísla $u_1, \dots, u_n \in \mathbb{R}$. Potom rovnice $\sum_{j=1}^{n} |x_j - a_j| = |A|$ přejde do tvaru $\sum_{j=1}^{n} |u_j A| = |A|$, tj. $\sum_{j=1}^{n} |u_j| = 1$ a současně podmínka $\sum_{j=1}^{n} x_j = 0$ dává

$$0 = \sum_{j=1}^{n} (a_j - u_j A) = \sum_{j=1}^{n} a_j - A \sum_{j=1}^{n} u_j = A(1 - \sum_{j=1}^{n} u_j),$$

tj. $\sum_{j=1}^{n} u_j = 1$. Proto nutně $u_j \ge 0$, v opačném případě bychom totiž dostali spor s rovnostmi $\sum_{j=1}^{n} |u_j| = 1 = \sum_{j=1}^{n} u_j$. Celkem proto vidíme, že hledaným řešením primární úlohy je $x^* = a - A u$ pro libovolné $u \in \mathbb{R}^n_+$ splňující $\sum_{j=1}^{n} u_j = 1$.



Obrázek 8.34: Graf funkce $|x_j - a_j| + y x_j$ z Příkladu 8.3.22. pro $a_j > 0$ a $y = 2$ (růžová barva), $y = 1/2$ (červená), $y = 1$ (zelená), $y = -1$ (modrá) a $y = -2$ (hnědá).

Příklad 8.3.23

Určeme duální úlohu pro

$$f(x) = \mathrm{e}^{-x} \to \min, \quad x \le 1. \tag{8.3.22}$$

Vyřešme ji a zjistěme, jaké informace nám toto řešení dává o primární úloze (8.3.22).

Video ☞ https://ln.math.muni.cz/pz/opt/mp5_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp5_zapisnik

Příklad 8.3.24

Určeme duální úlohu pro

$$f(x_1,x_2) = 2x_1^2 + x_1x_2 + x_2^2 - 4x_1 + 2x_2 \to \min, \quad 3x_1 + x_2 \le 2, \quad x_1 \ge 0. \tag{8.3.23}$$

Vyřešme ji a zjistěme, jaké informace nám toto řešení dává o primární úloze (8.3.23).

Video ☞ https://ln.math.muni.cz/pz/opt/mp6_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp6_zapisnik

Tuto část zakončíme ještě jinou alternativou k původní Karushově–Kuhnově–Tuckerově větě, k níž nás může přivést poznámka o von Neumannově větě o „minimaxu“ (viz Větu 8.3.10) a nerovnost (8.3.13). S pomocí této reformulace Věty 8.2.15 již dokážeme dát definitivní odpověď na otázku: *jak s pomocí řešení duální úlohy získáme řešení primární úlohy?* Toto tvrzení je založeno na pojmu sedlového bodu Lagrangeovy funkce.

DEFINICE 8.3.25

Bod $[x^*, y^*] \in P \times Q$ se nazývá *sedlovým bodem* Lagrangeovy funkce $L(x,y)$ úlohy (8.1) a (8.2) na $P \times Q$, jestliže

$$L(x^*, y) \le L(x^*, y^*) \le L(x, y^*) \quad \text{pro všechny } x \in P \text{ a } y \in Q\,,$$

tj. platí

$$L(x^*, y^*) = \min_{x\in P} L(x, y^*) = \max_{y\in Q} L(x^*, y).$$

A jak to souvisí s úlohou (8.1) a (8.2)?

VĚTA 8.3.26 KARUSHOVA–KUHNOVA–TUCKEROVA PRO SEDLOVÝ BOD

Nechť úloha (8.1) a (8.2) je regulární úlohou konvexního programování (viz Větu 8.3.4). Pak bod $x^* \in P$ je řešením úlohy (8.1) a (8.2) právě tehdy, když existuje $y^* \in Q$ takové, že $[x^*, y^*]$ je sedlovým bodem Lagrangeovy funkce $L(x,y)$ úlohy (8.1) a (8.2) na $P \times Q$.

Důkaz. Důkaz je založen na Větě 8.3.20(ii). „$\Longrightarrow$“ Nechť $x^* \in P$ je řešením úlohy (8.1) & (8.2). Pak $x^* \in X$ a podle Věty8.3.20(ii) existuje $y^* \in Q$ takový, že

$$L(x^*, y^*) = \min_{x\in P} L(x, y^*) \quad \text{a} \quad y_i^*\, g_i(x^*) = 0, \quad i \in \{1,\dots,k\},$$

tj. z první části máme $L(x^*, y^*) \le L(x, y^*)$ pro každé $x \in P$. Současně s využitím druhé části dostaneme

$$L(x^*, y^*) = f(x^*) + \underbrace{\sum_{i=1}^m y_i^*\, g_i(x^*)}_{=0} = f(x^*) \ge f(x^*) + \underbrace{\sum_{i=1}^m y_i\, g_i(x^*)}_{\le 0} = L(x^*, y) \quad \text{pro všechna } y \in Q.$$

To dohromady s první částí ukazuje, že bod $[x^*, y^*]$ je sedlovým bodem Lagrangeovy funkce úlohy (8.1) a (8.2).

„$\Longleftarrow$“ Naopak, nechť nyní $[x^*, y^*]$ je sedlovým bodem Lagrangeovy funkce. Potom $L(x^*, y^*) = \min_{x\in P} L(x, y^*)$ a zároveň $L(x^*, y^*) \ge L(x^*, y)$ pro každé $y \in Q$, tj.

$$f(x^*) + \sum_{i=1}^m y_i^*\, g_i(x^*) \ge f(x^*) + \sum_{i=1}^m y_i\, g_i(x^*) \quad \text{pro všechna } y \in Q$$

neboli

$$\sum_{i=1}^m y_i^*\, g_i(x^*) \ge \sum_{i=1}^m y_i\, g_i(x^*) \quad \text{pro všechna } y \in Q\,.$$

Zejména pro $y = 0 \in Q$ odtud dostaneme $\sum_{i=1}^{m} y_i^* g_i(x^*) \geq 0$. Současně volba $y = y^* + e_i \in Q$ pro $i \in \{1,\dots,k\}$ dává $0 \geq g_i(x^*)$ a $y = y^* \pm e_i \in Q$ pro $i \in \{k+1,\dots,m\}$ dává dokonce dvojici nerovností $0 \geq g_i(x^*)$ a $0 \leq g_i(x^*)$, tj. $g_i(x^*) = 0$. Proto $x^* \in X$ a $y_i^* g_i(x^*) \leq 0$ podle definice množiny Q, z čehož plyne $y_i^* g_i(x^*) = 0$ pro všechna $i \in \{1,\dots,m\}$. Tedy jsou splněny podmínky Věty 8.3.20(ii), což znamená, že bod $x^* \in P$ je řešením úlohy (8.1) a (8.2). ■

Poznámka 8.3.27. (i) Z předchozího důkazu je patrné, že implikace „$\Longleftarrow$“ ve Větě 8.3.26 platí i bez předpokladu[79] regularity úlohy (8.1) a (8.2). Bez tohoto předpokladu dokonce platí i to, že sedlový bod splňuje (8.3.20) a (8.3.21).

(ii) Analogické tvrzení jako ve Větě 8.3.26 platí také pro duální úlohu[80]:

> Bod $y^* \in Y$ je řešením duální úlohy (8.3.7) právě tehdy, když existuje $x^* \in P$ takový, že $[x^*, y^*] \in P \times Q$ je sedlovým bodem Lagrangeovy funkce regulární úlohy (8.1) a (8.2).

▲

V literatuře bývají mnohdy Věty 8.3.20 a 8.3.26 uváděny společně.

> Nechť úloha (8.1) a (8.2) je regulární úlohou konvexního programování (viz Větu 8.3.4).
> (i) Body $x^* \in X$ a $y^* \in Y$ řeší úlohy (8.1) a (8.2) a (8.3.7) právě tehdy, když platí podmínka (i) nebo (ii) z Věty 8.3.20.
> (ii) Body $x^* \in X$ a $y^* \in Y$ řeší úlohy (8.1) a (8.2) a (8.3.7) právě tehdy, když $[x^*, y^*]$ je sedlový bod Lagrangeovy funkce úlohy (8.1) a (8.2).

A co to znamená pro řešení primární úlohy? Známe-li řešení duální úlohy, stačí dopočítat x^* tak, aby $[x^*, y^*]$ byl sedlovým bodem Lagrangeovy funkce. To ale z Definice 8.3.1 znamená, že musíme určit $\min_{x \in P} L(x, y^*)$, což jsme již při výpočtu $\varphi(y)$ učinili dokonce pro libovolné $y \in Q$, tj. dosazením hodnoty $y = y^*$ do tohoto výpočtu získáme odpovídající x^*, které je řešením úlohy (8.1) a (8.2), viz také komentář k Větě 8.3.20 v Poznámce 8.3.21.

Cvičení

8.3.1. Určete duální úlohu k úloze matematického programování

(a) $f(x) = x^2 + 2x + 2 \to \min \quad \& \quad x \leq 1$
(b) $f(x) = x^2 + 2x + 2 \to \min \quad \& \quad x + 2 \leq 0$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: (a) $\varphi(y) = -y^2/4 - 2y + 1$ a $y^* = 0$,
(b) $\varphi(y) = -y^2/4 + y + 1$ a $y^* = 2$

8.3.2. Určete duální úlohu k úloze matematického programování

(a) $f(x) = \sqrt{1 + x^2} \to \min \quad \& \quad x \leq 1$
(b) $f(x) = \sqrt{1 + x^2} \to \min \quad \& \quad x + 1 \leq 0$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: (a) $\varphi(y) = \sqrt{1 - y^2} - y$ pro $y \in [0,1)$, $\varphi(y) = -\infty$ pro $y \geq 1$, $y^* = 0$,
(b) $\varphi(y) = \sqrt{1 - y^2} + y$ pro $y \in [0,1)$, $\varphi(y) = -\infty$ pro $y \geq 1$, $y^* = 1/\sqrt{2}$

8.3.3. Určete duální úlohu k úloze matematického programování

(a) $f(x) = |x| \to \min \quad \& \quad x \leq 1$
(b) $f(x) = |x| \to \min \quad \& \quad x + 1 \leq 0$
(c) $f(x) = |x| \to \min \quad \& \quad x \leq 0$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

[79] Absence předpokladu regularity úlohy (8.1) a (8.2) nám zamezuje využít Větu 8.3.20(ii) v závěru důkazu. Nechť $x \in P$ a $y \in Q$ jsou libovolné. Nicméně z rovnosti $y_i^* g_i(x^*)$ plyne pro libovolné $x \in X$, že

$$f(x^*) = L(x^*, y^*) \leq L(x, y^*) = \\ = f(x) + \underbrace{\sum_{i=1}^{m} y_i^* g_i(x)}_{\leq 0} \leq f(x),$$

tj. x^* je globálním řešením úlohy (8.1) a (8.2).

Důsledek 8.3.28

[80] Nechť y^* je řešením duální úlohy. Pak podle Věty 8.3.16 také existuje řešení primární úlohy, takže dvojice x^*, y^* splňuje podmínky (8.3.20) a (8.3.21) ve Větě 8.3.20(ii). Zbytek důkazu je pak totožný s první částí důkazu Věty 8.3.26.

Naopak, nechť $[x^*, y^*]$ je sedlový bod Lagrangeovy funkce. Pak jsou splněny podmínky Věty 8.3.20(ii) podle důkazu druhé části Věty 8.3.26, z čehož plyne, že y^* je řešením duální úlohy.

řešení: (a) $\varphi(y) = -y$ pro $y \in [0,1]$, $\varphi(y) = -\infty$ pro $y > 1$, $y^* = 0$,
(b) $\varphi(y) = y$ pro $y \in [0,1]$, $\varphi(y) = -\infty$ pro $y > 1$, $y^* = 1$,
(c) $\varphi(y) = 0$ pro $y \in [0,1]$, $\varphi(y) = -\infty$ pro $y > 1$, y^* je libovolný bod z $[0,1]$

8.3.4. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1 \to \min \quad \& \quad x_2 \le x_1$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2/2 - y - 1$ pro $y \ge 0$ a $y^* = 0$

8.3.5. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = x_1^2 + x_2^2 + 2x_1 \to \min \quad \& \quad x_2 + 1 \le x_1$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2/2 + 2y - 1$ pro $y \ge 0$ a $y^* = 2$

8.3.6. Určete duální úlohu k úloze

$$f(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 \to \min \quad \& \quad x_1 + x_2 \ge 1$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2 + y$ pro $y \ge 0$ a $y^* = 1/2$

8.3.7. Určete duální úlohu k úloze

$$f(x_1, x_2) = 3x_1^2 + 2x_1 x_2 + x_2^2 - 2x_1 + x_2 \to \min \quad \& \quad x_1 + 2x_2 + 2 \le 0$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -9y^2/8 + y/4 - 11/8$ pro $y \ge 0$ a $y^* = 1/9$

8.3.8. Určete duální úlohu k úloze

$$f(x_1, x_2) = 3x_1^2 + x_2^2 + 4x_2 - 1 \to \min \quad \& \quad x_1^2 + 2x_2 \le 1$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2 - 5y - 5$ pro $y \ge 0$ a $y^* = 0$

8.3.9. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = x_1^2 - x_1 x_2 + 2x_2^2 - 3x_1 + 2x_2 \to \min \quad \& \quad x_1 + x_2 \le 1 \quad \& \quad x_1 \ge 0$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = (-28y^2 + 14y - 112)/49$ pro $y \in [0,2]$,
$\varphi(y) = -(y^2 + 12y + 4)/8$ pro $y \in [2,\infty)$ a $y^* = 1/4$

8.3.10. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = x_1^2 + 6x_2^2 + 2x_1 - 8x_2 \to \min \quad \& \quad 2x_2 - x_1 \le 1 \quad \& \quad x_2 \ge 0$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2/4 - 1$ pro $y \in [0,4]$,
$\varphi(y) = -5y^2/12 + 4y/3 - 11/3$ pro $y \in [4,\infty)$ a $y^* = 8/5$

8.3.11. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = x_1^2 + 6x_2^2 + 2x_1 - 8x_2 \to \min \quad \& \quad 2x_2 - x_1 + 2 \le 0 \quad \& \quad x_2 \ge 0$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = -y^2/4 + 3y - 1$ pro $y \in [0,4]$,
$\varphi(y) = -5y^2/12 + 13y/3 - 11/3$ pro $y \in [4,\infty)$ a $y^* = 6$

8.3.12. Určete duální úlohu k úloze matematického programování

$$f(x_1, x_2) = 4x_1^2 + 3x_1 x_2 + x_2^2 + 2x_1 - 5x_2 \to \min \quad \& \quad 2x_1 - 3x_2 - 5 \le 0 \quad \& \quad x_2 \ge 0$$

a ověřte platnost vztahu duality $f^* = \varphi^*$.

řešení: $\varphi(y) = (-58y^2 - 141y - 134)/7$ pro $y \ge 0$ a $y^* = 0$

8.3.8.3.12.A) Určete duální úlohu k úloze matematického programování

$$f(x_1,x_2)=4x_1^2+4x_1x_2+x_2^2-2x_1+3x_2\to\min \quad \& \quad x_1-3x_2\le 4 \quad \& \quad x_1\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-9y^2/4+y/2-9/4$ pro $y\ge 8/7$, $\varphi(y)=-\infty$ pro $y\in[0,8/7)$
$y^*=8/7$

8.3.13. Určete duální úlohu k úloze

$$f(x_1,x_2)=x_1^2+x_2^2\to\min \quad \& \quad x_1+x_2\ge 4 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-y^2/2+4y$ pro $y\ge 0$ a $y^*=4$

8.3.14. Určete duální úlohu k úloze matematického programování

$$f(x_1,x_2)=x_1^2+x_2^2+3x_1\to\min \quad \& \quad x_2^2-x_1+x_2\le 1 \quad \& \quad x_1\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

Nápověda: polynom $2z^3+3z^2-2$ má jediný reálný kořen $z\approx 0,67765$.

řešení: $\varphi(y)=-(5y^2+4y)/(4+4y)$ pro $y\in[0,3]$,
$\varphi(y)=-(y^3+7y+9)/(4+4y)$ pro $y\ge 3$ a $y^*=0$

8.3.15. Určete duální úlohu k úloze matematického programování

$$f(x_1,x_2)=2x_1^2+x_1x_2+x_2^2-x_2\to\min \quad \& \quad x_1^2-2x_2+x_1\le 2 \quad \& \quad x_1\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-y^2-3y-1/4$ pro $y\ge 0$ a $y^*=0$

8.3.15.A) Určete duální úlohu k úloze matematického programování

$$f(x_1,x_2)=2x_1^2+x_1x_2+x_2^2+x_2\to\min \quad \& \quad x_1^2-2x_2+x_1\le 2 \quad \& \quad x_1\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

(**Nápověda:** platí $16z^4+53z^3+17z^2-48z+14=(7+4z)(4z^3+7z^2-8z+2)$ a polynom $16z^3+72z^2+105z+17$ má jediný reálný kořen $z\approx -0,18422$.)

řešení: $\varphi(y)=(-4y^3-15y^2-6y-2)/(7+4y)$ pro $y\in[0,1/4]$,
$\varphi(y)=-y^2-y-1/4$ pro $y\ge 1/4$ a $y^*=0$

8.3.16. Určete duální úlohu k úloze

$$f(x_1,x_2)=8x_1+18x_2\to\min \quad \& \quad x_1+3x_2\ge 2 \quad \& \quad 2x_1+4x_2\ge 1 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y_1,y_2)=2y_1+y_2$ pro $8-y_1-2y_2\ge 0$, $18-3y_1-4y_2\ge 0$ a $y_1,y_2\ge 0$
$[y_1^*,y_2^*]=[6,0]$

8.3.17. Určete duální úlohu k úloze

$$f(x_1,x_2)=x_1+x_2\to\min \quad \& \quad x_1+2x_2\ge 3 \quad \& \quad 2x_1+x_2\ge 3 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $[y_1^*,y_2^*]=[1,1]$

8.3.18. Určete duální úlohu k úloze

$$f(x_1,x_2)=x_1^2+x_2^2\to\min \quad \& \quad x_1+x_2\le 1 \quad \& \quad x_1-x_2+2\le 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y_1,y_2)=-(y_1+1)^2/2-(y_2-2)^2/2+5/2$ pro $[y_1,y_2]\in\mathbb{R}^2_+$
$[y_1^*,y_2^*]=[0,2]$

8.3.19. Určete duální úlohu k úloze

$$f(x_1,x_2,x_3)=4x_3^2+2x_2-6x_1\to\min \quad \& \quad 2x_1+2x_2+x_3\le 0 \quad \& \quad -2x_1+4x_2+x_3^2=0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y_1,y_2)=-y_1^2/(16+4y_2)$ pro $y_1\ge 0$, $y_1-y_2=3$ a $2+2y_1+4y_2\ge 0$
$[y_1^*,y_2^*]=[5/3,-4/3]$

8.3.20. Určete duální úlohu k úloze

$$f(x_1,x_2)=\sqrt{1+x_1^2+x_2^2}\to\min \quad \& \quad x_1+x_2\le 1$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=\sqrt{1-2y^2}-y$ pro $y\in[0,1/\sqrt{2}]$ a $y^*=0$

8.3.21. Určete duální úlohu k úloze

$$f(x_1,x_2)=x_1+\mathrm{e}^{x_2}\to\min \quad \& \quad 3x_1-2\,\mathrm{e}^{x_2}\ge 10 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y_1,y_2)=y_2-y_2\ln[y_2/(1+2y_1)]+10y_1$ pro $y_1=1/3$ a $y_2\ge 0$
$[y_1^*,y_2^*]=[1/3,5/3]$

8.3.22. Určete duální úlohu k úloze

$$f(x_1,x_2)=-\sqrt{x_1}-\sqrt{3x_2}\to\min \quad \& \quad x_1+x_2\le 4 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-1/y-4y$ pro $y>0$ a $\varphi(y)=-\infty$ pro $y=0$
$y^*=1/2$

8.3.23. Určete duální úlohu k úloze

$$f(x_1,x_2,x_3)=x_1-4x_2+x_3^4\to\min \quad \& \quad x_1+x_2+x_3^2\le 2 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-2y$ pro $y\ge 4$ a $\varphi(y)=-\infty$ pro $y<4$
$y^*=4$

8.3.24. Určete duální úlohu k úloze

$$f(x_1,x_2,x_3)=\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}\to\min \quad \& \quad x_1+x_2+x_3\le 2 \quad \& \quad x_1>0 \quad \& \quad x_2>0 \quad \& \quad x_3>0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$.

řešení: $\varphi(y)=-2y+6\sqrt{y}$ pro $y\ge 0$ a $y^*=9/4$

8.3.25. Určete duální úlohu k úloze

$$f(x_1,x_2,x_3,x_4)=x_1+2x_2+3x_3+4x_4+\sum_{i=1}^{4}x_i\ln x_i\to\min,$$
$$x_1+x_2+x_3+x_4=1 \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0 \quad \& \quad x_3\ge 0 \quad \& \quad x_4\ge 0$$

a ověřte platnost vztahu duality $f^*=\varphi^*$. (Pro jednoduchost užíváme konvenci $0\ln 0=0$.)

řešení: $\varphi(y)=-\mathrm{e}^{-2-y}-\mathrm{e}^{-3-y}-\mathrm{e}^{-4-y}-\mathrm{e}^{-5-y}-y$ pro $y\in\mathbb{R}$ a $y^*=\ln(\mathrm{e}^3+\mathrm{e}^2+\mathrm{e}+1)-5$

8.3.26. Určete duální úlohu k úloze

$$f(x)=\langle x,x\rangle\to\min \quad \& \quad Ax=b,$$

kde $A\in\mathbb{R}^{m\times n}$ a $b\in\mathbb{R}^m$.

řešení: $\varphi(y)=-\frac{1}{4}\,y^\top AA^\top y-y^\top b$ pro $y\in\mathbb{R}^m$

8.3.27. Určete duální úlohu k úloze kvadratického programování

$$f(x)=\tfrac{1}{2}\,\langle Qx,x\rangle+\langle c,x\rangle\to\min \quad \& \quad Ax\le b,$$

kde $Q\in\mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní matice, $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$, $c\in\mathbb{R}^n$.

řešení: $\varphi(y) = -\frac{1}{2}\, y^\top AQ^{-1}A^\top y - y^\top (AQ^{-1}c + b) - c^\top Q^{-1}c/2$ pro $y \geq 0$

8.3.28. Určete duální úlohu k úloze kvadratického programování

$$f(x) = \tfrac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle \to \min \quad \& \quad Ax = b \quad \& \quad x \geq 0,$$

kde $Q \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní matice, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

řešení: $\varphi(y) = -\frac{1}{2}\langle Qx, x\rangle - \langle y, b\rangle$ pro $Qx + A^\top y \geq -d$

8.3.29. Určete duální úlohu k úloze kvadratického programování

$$f(x) = \tfrac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle \to \min \quad \& \quad Ax \leq b \quad \& \quad x \geq 0,$$

kde $Q \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní matice, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

řešení: $\varphi(y) = -\frac{1}{2}\langle Qx, x\rangle - \langle y, b\rangle \to \max$ pro $Qx + A^\top y \geq -d$ a $y \geq 0$

8.3.30. Určete duální úlohu k úloze kvadratického programování

$$f(x) = \tfrac{1}{2}\langle Qx, x\rangle + \tfrac{1}{2}\langle Pz, z\rangle + \langle c, x\rangle \to \min \quad \& \quad Ax + Pz \geq -b \quad \& \quad x \geq 0$$

kde $P, Q \in \mathbb{R}^{n\times n}$ jsou symetrické a pozitivně definitní matice, $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

řešení: $\varphi(y) = -\frac{1}{2}\langle Qx, x\rangle - \frac{1}{2}\langle Py, y\rangle - \langle y, b\rangle$ pro $Qx - A^\top y \geq -d$ a $y \geq 0$

8.4 Analýza citlivosti

Nyní se budeme věnovat *závislosti řešení* úloh matematického programování na (*vstupních*) *parametrech*, tj. situaci kdy úloha (8.1) a (8.2) obsahuje i nějaké parametry, které mohou vystupovat v účelové funkci a/nebo ve funkcionálních omezeních:

(i) parametry v účelové funkci mohou zastupovat prodejní ceny výrobků získaných v jistém výrobním procesu, nákupní ceny, míru užitku, ...

(ii) parametry v omezeních mohou zastupovat maximální investice v různých segmentech výrobního procesu, dostupnost skladových zásob, skladovací/výrobní kapacitu, přijatelnou míra rizika, ...

Této oblasti se říká *parametrické programování*, o čemž jsme se zmiňovali již při motivačních úlohách z oblasti kvadratického programování. Pro jednoduchost se však nebudeme věnovat zcela obecné úloze matematického programování, ale uvážíme nejdříve úlohu s omezeními pouze ve tvaru rovností s „obecnou“ závislostí na parametrech (tím se dostaneme k tzv. *Větě o obálce*) a poté se zaměříme na úlohu s omezeními ve tvaru nerovností, kde ovšem parametry budou vystupovat pouze v podobě absolutních členů funkcí zadávajících jednotlivá omezení.

Uvidíme, že i v této oblasti hrají zásadní roli Lagrangeovy multiplikátory. TO je vlastně i původní motivace pro studium zmíněné závislosti, neboť v některých úlohách matematického programování (zejména v lineárním programování) lze Lagrangeovy multiplikátory (které jsou úzce spojeny s řešením duální úlohy, jak víme z předchozí části) velmi zajímavým způsobem interpretovat – v ekonomických úlohách jde o nějaké ceny (to si i ukážeme), zatímco např. ve fyzice představují veličiny s konkrétním fyzikálním významem. Uvidíme také, že z matematického pohledu mohou být chápány jako ukazatel míry změny optimální hodnoty účelové funkce při změně parametrů.

Začněme proto s jednoduchou úlohu pro dvě proměnné a jedno omezení ve tvaru rovnosti, tj.

$$f(x_1, x_2) \to \min, \qquad g(x_1, x_2) = c, \tag{8.4.1}$$

o které pro jednoduchost předpokládáme, že je jednoznačně řešitelná pro každé[81] $c \in \mathbb{R}$. Potom řešením (8.4.1) je dvojice $x_1^*(c)$ a $x_2^*(c)$ závislá na parametru c a připusťme, že tyto funkce $x_1^*(\cdot)$ a $x_2^*(\cdot)$ jsou diferencovatelné vzhledem k c. Hodnota úlohy (8.4.1) také závisí na c, tj.

$$f^*(c) = f\big(x_1^*(c), x_2^*(c)\big),$$

[81] Stačilo by na nějakém okolí bodu $\bar{c}$.

stejně jako odpovídající Lagrangeův multiplikátor $y^* = y^*(c)$. Je-li navíc i funkce $f^*(c)$ diferencovatelná vzhledem k c, pak platí

$$\frac{\mathrm{d}}{\mathrm{d}c} f^*(c) = \frac{\mathrm{d}}{\mathrm{d}c} f\big(x_1^*(c), x_2^*(c)\big) = f_{x_1}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_1^*(c)}{\mathrm{d}c} + f_{x_2}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_2^*(c)}{\mathrm{d}c}. \tag{8.4.2}$$

Protože $x_1^*(c)$ a $x_2^*(c)$ řeší úlohu (8.4.1), plyne z Lagrangeova principu (Věta 8.2.1) že pro funkci $L(x_1, x_2, y) = f(x_1, x_2) + y\, g(x_1, x_2)$ musí platit $L_{x_1}(x_1^*, x_2^*) = 0$, $L_{x_2}(x_1^*, x_2^*) = 0$ a $g(x_1^*, x_2^*) = c$, tj.

$$\left.\begin{aligned} f_{x_1}(x_1^*(c), x_2^*(c)) + y^*(c)\, g_{x_1}(x_1^*(c), x_2^*(c)) = 0, \\ f_{x_2}(x_1^*(c), x_2^*(c)) + y^*(c)\, g_{x_2}(x_1^*(c), x_2^*(c)) = 0. \end{aligned}\right\} \tag{8.4.3}$$

Současně derivováním rovnosti $g(x_1^*(c), x_2^*(c)) = c$ vzhledem k c obdržíme

$$g_{x_1}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_1^*(c)}{\mathrm{d}c} + g_{x_2}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_2^*(c)}{\mathrm{d}c} = 1. \tag{8.4.4}$$

Zkombinováním rovností (8.4.2), (8.4.3) a (8.4.4) dostaneme

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}c} f^*(c) &\overset{(8.4.2)}{\overset{(8.4.3)}{=}} -y^*(c)\, g_{x_1}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_1^*(c)}{\mathrm{d}c} - y^*(c)\, g_{x_2}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_2^*(c)}{\mathrm{d}c} = \\ &= -y^*(c) \left[g_{x_1}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_1^*(c)}{\mathrm{d}c} + g_{x_2}(x_1^*(c), x_2^*(c)) \frac{\mathrm{d}x_2^*(c)}{\mathrm{d}c} \right] = \\ &\overset{(8.4.4)}{=} -y^*(c). \end{aligned}$$

Odtud je vidět, že hodnota Lagrangeova multiplikátoru udává míru (rychlost) změny hodnoty dané úlohy (tj. čísla f^*) vzhledem ke změně c. Jinými slovy, je-li Δc (dostatečně malá) změna hodnoty c, pak

$$f^*(c + \Delta c) - f^*(c) \approx -y^*(c)\, \Delta c.$$

V ekonomii mnohdy číslo c na pravé straně omezení vyjadřuje dostupné zásoby nějakého zdroje (včetně času, peněz apod.) a funkce f vyjadřuje užitek nebo zisk dosažený spotřebováváním tohoto zdroje. Potom hodnota $-y^*(c)\, \Delta c$ měří přibližnou míru změny užitku/zisku, pokud přidáme/ubereme Δc jednotek zdroje. Jelikož v našem případě $y^*(c) \in \mathbb{R}$, mohou se změny dít oběma směry. Ekonomové nazývají multiplikátor $-y(c)$ jako *stínovou cenu*[82] zdroje, což má původ a hlubší souvislosti v jeho roli v duální úloze lineárního programování, viz Podkapitolu 2.6. Slovo „stínová“ je zde proto, že nemusí jednat o skutečnou tržní cenu, např. zdroj může být náš čas věnovaný jisté aktivitě a pro takový zdroj asi nedokážeme určit jeho skutečnou tržní cenu – vždyť pro takovou „komoditu“ asi neexistuje ani vhodný trh. Speciálně při volbě $\Delta c = 1$ získáme

$$f^*(c + 1) - f^*(c) \approx -y^*(c), \tag{8.4.5}$$

tj. $-y^*(c)$ přibližně udává, o kolik vzroste zisk/užitek $f^*(c)$, pokud navýšíme zdroj o 1 jednotku.

[82] *Shadow price*

Příklad 8.4.1

Produkční funkce firmy je $f(x_p, x_m) = 50\, x_p^{1/2}\, x_m^2$, kde x_p udává cenu práce a x_m cenu materiálu v tisících Kč. Firma má dostupný rozpočet 79 tisíc Kč. Jak se změní hodnota optimální produkce, pokud firma investuje 80 tisíc Kč?

Máme dánu úlohu

$$-50 x_p^{1/2}\, x_m^2 \to \min, \quad x_p + x_m = 79,$$

tj. $L(x_p, x_m, y) = -50\, x_p^{1/2}\, x_m^2 + y\,(x_p + x_m - 79)$ s parciálními derivacemi

$$L_{x_p}(x_p, x_m, y) = -25\, x_p^{-1/2}\, x_m^2 + y \quad \text{a} \quad L_{x_m}(x_p, x_m, y) = -100\, x_p^{1/2}\, x_m + y.$$

Řešením soustavy

$$L_{x_p}(x_p, x_m, y) = 0, \quad L_{x_m}(x_p, x_m, y) = 0 \quad \text{a} \quad x_p + x_m = 79$$

je trojice $x_m^*(79) = 63{,}2$, $x_p^*(79) = 15{,}8$ a $y^*(79) \approx 25121{,}5$ s hodnotou $f^*(79) = -793839{,}5$, tj. maximální produkce má hodnotu 793839,5 tisíc Kč. V souladu s (8.4.5) dostaneme[83] při navýšení rozpočtu o 1 tisíc Kč přibližně

$$f^*(80) - f^*(79) \approx -y^*(79) \approx -25121{,}5.$$

To znamená, že hodnota f^* klesne zhruba o 25121,5, tj. hodnota maximální produkce vzroste o 25121,5 tisíc Kč. Při podobném výpočtu řešení pro $c = 80$ bychom získali $x_m^*(80) = 64$, $x_p^*(80) = 16$ a $y^*(80) = 25600$ s hodnotou $f^*(80) = -819200$, tj. přesný nárůst maximální produkce je 25360,5 tisíc Kč.
Ještě bychom měli zmínit, že tento přibližný výpočet nefunguje recipročně. Hodnota $y^*(80) \approx 25600$ totiž znamená, že při snížení rozpočtu z 80 tisíc Kč na 79 tisíc Kč dojde k poklesu hodnoty maximální produkce přibližně o 25600 tisíc Kč, což je mírně odlišné odhad než při nárůstu ze 79 tisíc na 80 tisíc

[83] Neměli bychom zapomenout na to, že odhad (8.4.5) byl odvozen za předpokladu diferencovatelnosti řešení a hodnoty účelové funkce vzhledem k parametru c. To jsme tentokrát neověřili - ono to totiž vyžaduje znalost řešení a Lagrangeových multiplikátorů pro libovolné c (a v takovém případě už není potřeba odhadovat přibližný nárůst či pokles f^* při změnách c, když to známe zcela přesně). Nicméně nemělo by být příliš těžké zjistit, že pro libovolné $c > 0$ budeme mít $x_m^*(c) = 4c/5$, $x_p^*(c) = c/5$ a $y^*(c) = 80\,c.\sqrt{c/5}$ s odpovídající hodnotou $f^*(c) = -32\,c^2\sqrt{c/5}$, přičemž diferencovatelnost všech těchto funkcí vzhledem k c je zřejmá.

Analogický výsledek můžeme dostat při větším počtu omezení a bez ohledu na počet proměnných, tj. pro $c = (c_1, \dots, c_m)^\top$ máme

$$f^*(c + \Delta c) \approx f^*(c) - y_1^*(c)\,\Delta c_1 - \cdots - y_m^*(c)\,\Delta c_m. \tag{8.4.6}$$

Např. řešení úlohy

$$f(x,y,z) = -x - 2z \to \min, \quad x + y + z = 1, \quad x^2 + y^2 + z = 7/4$$

pomocí Lagrangeových multiplikátorů dává $[x^*, y^*, z^*] = [0, -1/2, 1/3]$ a $y_1 = y_2 = 1$ s hodnotou $f^* = -3$. Změníme-li daná omezení na

$$x + y + z = 98/100 \quad \text{a} \quad x^2 + y^2 + z = 9/5,$$

pak $\Delta c_1 = -2/100$ a $\Delta c_2 = 5/100$, takže

$$f^*(98/100, 9/5) - f^*(1, 7/4) = -1 \cdot (-2)/100 - 1 \cdot 5/100 = -0{,}03,$$

tj. optimální hodnota klesna přibližně o 3/100 (ve skutečnosti to je o 29/1000).

Odhady (8.4.5) a (8.4.6) jsou speciálním případem následujícího tvrzení.

Věta 8.4.2 O obálce

Mějme úlohu

$$f(x,r) \to \min, \quad g_1(x,r) = 0, \ \dots, \ g_m(x,r) = 0,$$

kde $x \in \mathbb{R}^n$, $r \in \mathbb{R}^k$, $f, g_1, \dots, g_m \in \mathrm{C}^1$. Připusťme, že pro každou hodnotu parametru r má úloha (8.4) jediné řešení, které označíme $x^*(r)$. Potom hodnota úlohy (8.4) je

$$f^*(r) = f(x^*(r), r).$$

Je-li $x^*(r)$ diferencovatelná vzhledem k r a Jacobiho matice $\mathrm{D}G(x^*(r), r) \in \mathbb{R}^{m \times n}$ má plnou hodnost m, pak platí

$$\frac{\partial}{\partial r_i} f^*(r) = \frac{\partial f}{\partial r_i}(x^*(r), r) + \sum_{j=1}^{m} y_j^*(r)\,\frac{\partial g_j}{\partial r_i}(x^*(r), r). \tag{8.4.7}$$

Důkaz. Jsou-li splněny předpoklady věty, pak podle řetězového pravidla platí

$$\frac{\partial}{\partial r_i} f^*(r) = \frac{\partial}{\partial x_1} f(x^*(r), r) \cdot \frac{\partial}{\partial r_i} x_1^*(r) + \cdots + \frac{\partial}{\partial x_n} f(x^*(r), r) \cdot \frac{\partial}{\partial r_i} x_n^*(r) + \frac{\partial}{\partial r_i} f(x^*(r), r).$$

Protože Jacobiho matice má plnou hodnost, existují čísla $y_1(r), \dots, y_m(r)$ taková, že

$$f_{x_i}(x^*(r), r) + \sum_{j=1}^{m} y_j(r) \cdot \frac{\partial}{\partial x_i} g_j(x^*(r), r) = 0 \quad \text{pro } i = 1, \dots, n.$$

Tudíž odtud máme

$$\frac{\partial}{\partial r_i} f^*(r) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j} f(x^*(r), r) \cdot \frac{\partial}{\partial r_i} x_j^*(r) + \frac{\partial}{\partial r_i} f(x^*(r), r) =$$
$$= -\sum_{j=1}^{n} \left(\sum_{\ell=1}^{m} y_\ell(r) \cdot \frac{\partial}{\partial x_j} g_\ell(x^*(r), r) \right) \frac{\partial}{\partial r_i} x_j^*(r) + \frac{\partial}{\partial r_i} f(x^*(r), r).$$

Záměnou pořadí sum obdržíme

$$\frac{\partial}{\partial r_i} f^*(r) = -\sum_{\ell=1}^{m} y_\ell(r) \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j} g_\ell(x^*(r), r) \cdot \frac{\partial}{\partial r_i} x_j^*(r) \right) + \frac{\partial}{\partial r_i} f(x^*(r), r).$$

Současně platí $g_\ell(x^*(r), r) = 0$, což zderivováním vzhledem k r_i dává

$$\sum_{j=1}^{n} \frac{\partial}{\partial x_j} g_\ell(x^*(r), r) \cdot \frac{\partial}{\partial r_i} x_j^*(r) + \frac{\partial}{\partial r_i} g_\ell(x^*(r), r) = 0,$$

tudíž

$$\frac{\partial}{\partial r_i} f^*(r) = \sum_{\ell=1}^{m} y_\ell(r) \frac{\partial}{\partial r_i} g_\ell(x^*(r), r) + \frac{\partial}{\partial r_i} f(x^*(r), r),$$

čímž je důkaz hotov. ■

Všimněme si, že vyjádření $\frac{\partial}{\partial r_i} f^*(r)$ na pravé straně (8.4.7) sna trochu překvapivě neobsahuje derivace vzhledem k původním proměnným $x_1, \ldots, x_n$ a je vlastně derivací Lagrangeovy funkce $L(x, y, r) = f(x, r) + \sum_{\ell=1}^{m} y_\ell g_\ell(x, r)$ vzhledem k r_i v bodě $[x^*(r), y(r), r]$, tj. vzhledem k $(n + m + i)$-té proměnné funkce $L(x, y, r)$, neboť $g_\ell(x^*(r), r)$. Proč se tomuto tvrzení říká „Věta o obálce“? Pro jednoduchost uvažme, že $g_i \equiv 0$ pro všechna $i \in \{1, \ldots, m\}$, tj. hledáme pouze

$$f(x) \to \min \quad \text{pro } x \in \mathbb{R}^n.$$

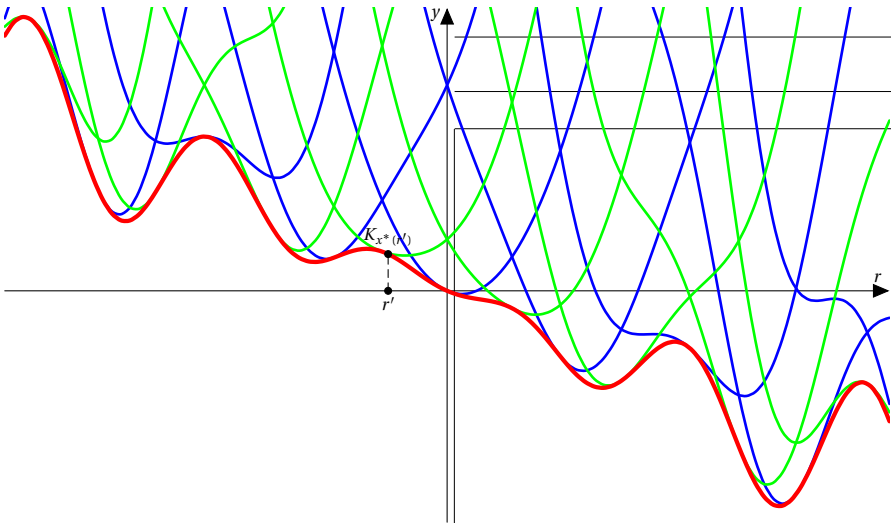
Potom rovnost (8.4.7) se redukuje na

$$\frac{\partial}{\partial r_i} f^*(r) = \frac{\partial f}{\partial r_i}(x^*(r), r). \tag{8.4.8}$$

Obálkou se obvykle rozumí křivka (nebo obecně spíše nadplocha), která je tečnou ke každé křivce (nadploše) z dané množiny. Označíme-li v našem případě symbolem K_x funkci $y = f(x, r)$ pro pevné $x \in \mathbb{R}^n$, pak žádná z nadploch K_x neleží pod grafem funkce $f^*(r)$, neboť pro každé x a r platí

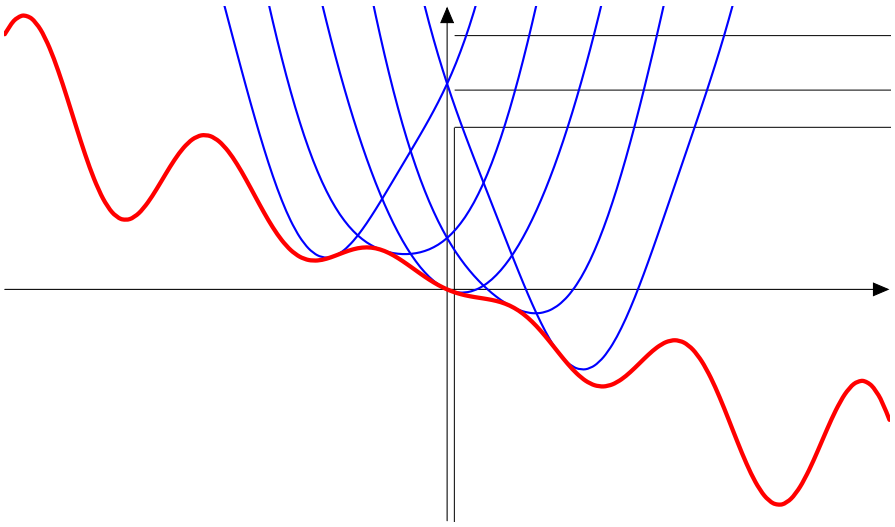
$$f(x, r) \geq \min_{x \in \mathbb{R}^n} f(x, r) = f^*(r).$$

Současně, vzhledem k předpokladům Věty 8.4.2, pro každé r existuje právě jedno x takové, že $f(x, r) = f^*(r)$ — tím je právě $x = x^*(r)$, tj. nadplocha $K_{x^*(r)}$ se dotkne grafu funkce $f^*(r)$ v bodě $[x^*(r), f^*(r)] = [x^*, f(x^*, r)]$ a podle (8.4.8) mají obě tyto nadplochy tutéž směrnici tečny. Na Obrázku 8.35 je tato situace znázorněna pro $n = 1$.



Obrázek 8.35: Ilustrace Věty o obálce (Věta 8.4.2).

Některé křivky K_x jsou vyznačeny modrou a zelenou barvou, graf funkce $f^*(r)$ je vykreslen červenou barvou. Vidíme, že každá z křivek K_x má právě jeden společný bod s $f^*(r)$. Tím je je nejnižší bod na všech K_x pro odpovídající hodnotu r, viz také Obrázek 8.36 s cca dvěma sty křivek K_x.



Obrázek 8.36: Další ilustrace Věty o obálce (Věta 8.4.2).

Možné využití Věty o obálce si ukážeme v několika následujících ilustrativních příkladech.

Příklad 8.4.3

Uvažme úlohu bez omezujících podmínek

$$\pi(q,p) = -p\,q + q^2/2 \to \min,$$

kde $p$je parametr. Pomocí Věty o obálce (Věta 8.4.2) vyřešme následující úkoly:

(i) Bez nalezení explicitního řešení dané úlohy určeme $\frac{\partial}{\partial p}\pi^*(p)$;

(ii) Určeme řešení $q^*(p)$. S pomocí předchozí části určeme znaménko $\frac{\partial}{\partial p}\pi^*(p)$ a interpretujme tento výsledek;

(iii) Určeme explicitně $\pi^*(p)$;

(iv) Načrtněme $\pi^*(p)$ a $\pi(q,p)$ pro $q \in \{1,2\}$, tj. $\pi(1,p) = -p + 1/2$, $\pi(3/2,p) = -3p/2 + 9/8$ a $\pi(2,p) = -2p + 2$;

(v) Rozhodněme o konvexnosti/konkávnosti funkce $\pi^*(p)$.

Řešení. Podle (8.4.8) platí

$$\frac{\partial}{\partial p}\pi^*(p) = \frac{\partial}{\partial p}\pi(q^*(p),p) = -q^*(p).$$

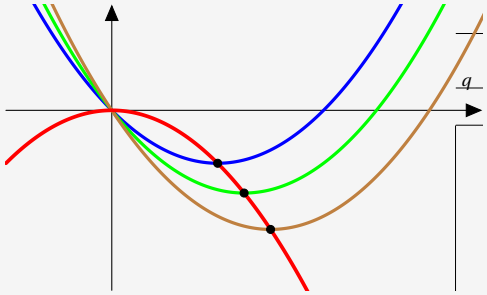
Funkce $\pi(\cdot,p)$ je konvexní pro libovolné $p \in \mathbb{R}$, takže pro vyřešení dané úlohy stačí nalézt stacionární bod této funkce, tj.

$$0 = \frac{\partial}{\partial q}\pi(q,p) = -p + q \quad \text{neboli} \quad q^*(p) = p.$$

Je-li $p > 0$, pak $\frac{\partial}{\partial p}\pi^*(p) = q^*(p) = p$ znamená, že funkce π^* je klesající, zatímco pro $p < 0$ je tato funkce rostoucí. Na Obrázku 8.37 jsou grafy funkcí $\pi(q,p)$ pro $p = 1$, $p = 5/4$ a $p = 3/2$ a také funkce

$$\pi^*(p) = \pi(q^*(p),p) = -p^2/2.$$

Průsečíky jednotlivých grafů s $\pi^*(p)$ jsou v bodech, kde nastává minimum $\pi(q,p)$ pro dané hodnoty p, tj. pro $q = p$ s odpovídajícími hodnotami minima $\pi^*(1) = -1/2$, $\pi^*(5/4) = -25/32$ a $\pi^*(3/2) = -9/8$.



Obrázek 8.37: Grafy funkcí $\pi(q,p)$ pro $p=1$ (modrá barva), $p=5/4$ (zelená barva) a $p=3/2$ (hnědá barva) a také funkce $\pi^*(p)$ (červená barva). Pozor — grafu funkce $\pi^*(p)$ tentokrát *není* obálkou ostatních křivek jako na Obrázcích 8.35–8.36. Nyní jsou totiž grafy vykresleny pro proměnnou q s pevně zvolenými hodnotami parametru p.

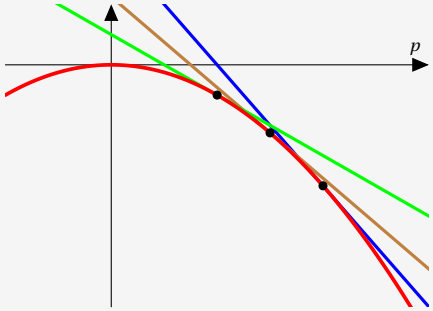
Graf funkce $\pi^*(p)$ je obálkou grafů funkcí $\pi(q,p)$. Toto můžeme vidět na Obrázku 8.38 s grafy $\pi(q,p)$ pro $q=1$, $q=3/2$ a $q=2$. Jaké jsou souřadnice průsečíků? To jsou pro dané hodnoty q body odpovídající hodnotě p takové, že $\pi(q,p)=\pi^*(p)$, tj. vzhledem k předchozím výpočtům dostáváme rovnost $q=p$, která nás přivede k bodům $[p,\pi^*(p)]$ neboli $[1,-1/2]$, $[3/2,-9/8]$ a $[2,-2]$. Navíc tečna ke grafu funkce $\pi^*(p)$ v bodě $[p_0,\pi^*(p_0)]$ je přímka o rovnici

$$t:\ y-\pi^*(p_0)=\frac{\mathrm{d}}{\mathrm{d}p}\pi^*(p_0)\,(p-p_0)\quad\text{neboli}\quad y=-p_0\,(p-p_0)-p_0^2/2$$

neboli

$$t:\ y=-p_0\,p+p_0^2/2=\pi(p_0,p),$$

tj. jednotlivé grafy funkcí $\pi(p_0,p)$ jsou dokonce tečnami $\pi^*(p)$ v bodech $[p_0,\pi^*(p_0)]$.



Obrázek 8.38: Grafy funkcí $\pi(q,p)$ pro $q=1$ (modrá barva), $q=3/2$ (hnědá barva) a $q=2$ (zelená barva) a jejich obálka jakožto graf funkce $\pi^*(p)$ (červená barva).

Konečně, funkce $\pi^*(p)$ je konvexní, neboť

$$\frac{\mathrm{d}}{\mathrm{d}p^2}\pi^*(p)=-1<0.$$

Příklad 8.4.4

Uvažme produkční funkci

$$F(x_1,\ldots,x_n),$$

která je diferencovatelná a konkávní. Nechť p je prodejní cena výrobku a $q_1,\ldots,q_n$ jednotkové ceny jednotlivých vstupů $x_1,\ldots,x_n$. Potom zisk je

$$\pi(x,p,q)=p\,F(x)-q_1\,x_1-\cdots-q_n\,x_n.$$

Podmínka prvního řádu pro maximalizaci zisku (tj. $-\pi(x,p,q)\to\min$) dává

$$\frac{\partial}{\partial x_i}\pi(x,p,q) = p\,\frac{\partial}{\partial x_i}F(x) - q_i = 0 \quad \text{pro všechna } i \in \{1,\dots,n\}.$$

Nechť $x^* = [x_1^*,\dots,x_n^*]$ je řešením této soustavy, přičemž předpokládejme, že platí $x_1^* > 0, \dots, x_n^* > 0$. Je jasné, že x^* je řešením také uvažované maximalizační úlohy (protože funkce $\pi(x,p,q)$ je konkávní) a současně $x^* = x^*(p,q)$. Označme funkci optimální hodnoty $\pi^*(p,q) \coloneqq \pi(x^*(p,q),p,q)$. Pak podle Věty o obálce (Věta 8.4.2) a potažmo rovnosti (8.4.8) platí

$$\frac{\partial}{\partial p}\pi^*(p,q) = \frac{\partial}{\partial p}\pi(x^*,p,q) = F(x^*) \quad \text{a} \quad \frac{\partial}{\partial q_i}\pi^*(p,q) = \frac{\partial}{\partial p}\pi(x^*,p,q) = -x_i^*.$$

Těmto rovnostem se říká *Hotellingovo lemma*[84] a jejich interpretace nás (snad) přivede k intuitivně zřejmému pozorování: vzroste-li prodejní cena o Δp, pak maximální zisk vzroste zhruba o $\Delta p\,F(x^*)$; zatímco vzroste-li nákupní cena i-tého vstupu o Δq_i, pak maximální zisk klesne zhruba o $\Delta q_i\,x_i^*$.

[84] Jen pro jistotu zdůrazněme, že ačkoli tento výsledek je skutečně pojmenován po Haraldu Hotellingovi zmíněném v Poznámce 60 na straně 80, autor se již setkal i s označením typu „hotelové lemma“...

Příklad 8.4.5

Uvažme užitkovou funkci

$$U(x) \to \max \quad \text{při omezení } \langle p,x\rangle = w,$$

kde x je vektor zboží/komodit, p vektor cen a w je zákazníkův majetek/příjem. Označme řešení jako $x^*(p,w)$ a odpovídající hodnotu užitku jako

$$U^*(p,q) \coloneqq U(x^*(p,w)).$$

Tato funkce se nazývá *nepřímou užitkovou funkcí*. Aplikováním Věty o obálce (viz Větu 8.4.2) na úlohu $-U(x) \to \min$ a $\langle p,x\rangle = w$ obdržíme

$$\frac{\partial}{\partial p_i}\mathscr{U}^*(p,w) = -y^*(p,w)\,x_i^*(p,w) \quad \text{a} \quad \frac{\partial}{\partial w}\mathscr{U}^*(p,w) = y^*(p,w)\,x_i^*(p,w), \qquad (8.4.9)$$

jelikož užitková funkce $U(x)$ nezávisí na p ani w a $\frac{\partial}{\partial w}\big(\langle p,x\rangle - w\big) = -1$. Druhá rovnost v (8.4.9) ukazuje, že $y^*(p,w)$ udává míru růstu maximálního užitku, pokud roste příjem w, tj. $y^*(p,w)$ vyjadřuje tzv. *mezní užitek příjmu*. První rovnost v (8.4.9) je tzv. *Royova identita*, která vyjadřuje to, že „mezní neužitečnost růstu j-té ceny je rovna součinu mezního užitku $(-y^*)$ a velikosti i-té poptávky (x_i^*)“. Navíc z obou těchto rovností vyplývá, že

$$\frac{\frac{\partial}{\partial p_i}\mathscr{U}^*(p,w)}{\frac{\partial}{\partial w}\mathscr{U}^*(p,w)} = -x_i^*(p,w),$$

tj. známe-li nepřímou užitkovou funkci můžeme snadno určit optimální poptávkovou funkci.

Vezmeme-li konkrétně za $U(x)$ Cobbovu–Douglasovu funkci

$$U(x_1,\dots,x_n) = A\,x_1^{a_1}\cdots x_n^{a_n}$$

pro nějaká pevně zvolená čísla $A, a_1, \dots, a_n > 0$. Pak řešením je[85]

$$x_i^*(p,w) = \frac{a_i\,w}{(a_1+\cdots+a_n)\,p_i},$$

takže nepřímá užitková funkce je

$$U^*(p,w) = B\,\frac{w^a}{P(p_1,\dots,p_n)},$$

[85] Vskutku, pro úlohu

$$-A\,x_1^{a_1}\cdots x_n^{a_n} \to \min$$

za podmínky $\langle p,x\rangle = w$ máme Lagrangeovu funkci $L(x,p,w) = -A\,x_1^{a_1}\cdots x_n^{a_n} + y\big(\langle p,x\rangle - w\big)$, která je konvexní. Musíme tedy najít řešení soustavy

$$L_{x_1} = -A\,a_1\,x_1^{a_1-1}\cdots x_n^{a_n} + y\,p_1 = 0$$
$$\vdots$$
$$L_{x_n} = -A\,a_n\,x_1^{a_1}\cdots x_n^{a_n-1} + y\,p_n = 0$$
$$p_1\,x_1 + \cdots + p_n\,x_n = w,$$

přičemž je dobré mít na paměti, že hledané minimum se jistě nerealizuje v bodě s $x_i = 0$ pro jakékoli $i \in \{1,\dots,n\}$. Sečtením první n-tice rovnic dostaneme

$$-A(a_1+\cdots+a_n)\,x_1^{a_1}\cdots x_n^{a_n} + y\,(p_1\,x_1+\cdots+p_n\,x_n) = 0,$$

z čehož s využitím poslední rovnosti dostaneme

$$y = \frac{A\,a\,x_1^{a_1}\cdots x_n^{a_n}}{w}$$

pro $a \coloneqq a_1+\cdots+a_n$. To zpětným dosazením do těchto rovnic dává

$$x_i^*(p,w) = \frac{a_i\,w}{a\,p_i}$$

pro všechna $i \in \{1,\dots,n\}$, a tudíž

$$y^*(p,w) = \frac{B\,a\,w^{a-1}}{P(p_1,\dots,p_n)}$$

pro číslo $B \coloneqq \frac{A\,a_1^{a_1}\cdots a_n^{a_n}}{a^a}$ a funkci $P(p_1,\dots,p_n) \coloneqq p_1^{a_1}\cdots p_n^{a_n}$.

kde $a := a_1 + \cdots + a_n$ je tzv. *index cen*, $B := \frac{A\, a_1^{a_1} \cdots a_n^{a_n}}{a^a}$ a $P(p_1, \ldots, p_n) := p_1^{a_1} \cdots p_n^{a_n}$, což je vlastně opět nějaká Cobbova–Douglasova funkce. Potom přímým výpočtem snadno získáme

$$\frac{\partial}{\partial w} U^*(p, w) = \frac{B\, a\, w^{a-1}}{P(p_1, \ldots, p_n)} = y^*(p, w)$$

a

$$\frac{\partial}{\partial p_i} U^*(p, w) = \frac{B\, w^a}{p_1^{a_1} \cdots p_{i-1}^{a_{i-1}}\, p_{i+1}^{a_{i+1}} \cdots p_n^{a_n}} = -\frac{B\, w^a\, a_i}{p_i\, P(p_1, \ldots, p_n)} =$$
$$= \frac{\partial}{\partial w} U^*(p, w)\, x_i^*(p, w) = -y^*(p, w)\, x_i^*(p, w),$$

což přesně odpovídá rovnostem z (8.4.9).

Jenže předpoklady uvedené Věty 8.4.2 jsou celkem silné (předpokládáme existenci řešení pro každé r). V jistém smyslu analogické tvrzení lze získat i za slabších předpokladů (alespoň některých).

VĚTA 8.4.6

Nechť $f, g_1, \ldots, g_m \in \mathrm{C}^2$ a x^* je lokálním řešením úlohy

$$f(x) \to \min, \quad g_1(x) = 0,\ \ldots,\ g_m(x) = 0$$

s odpovídající Lagrangeovým multiplikátorem $y^* \in \mathbb{R}^m$. Nechť dále tato dvojice splňuje postačující podmínku druhého řádu (tj. $\nabla_x^2 L(x^*, y^*) > 0$ na $\operatorname{Ker} DG(x^*)$), přičemž současně x^* je regulárním bodem, tj. Jacobiho matice $\mathrm{D}G(x^*) \in \mathbb{R}^{m \times n}$ má plnou hodnost m. Uvažme úlohu parametrického programování

$$f(x) \to \min, \quad G(x) = u$$

pro vektor parametrů $u \in \mathbb{R}^m$. Pak existuje otevřená koule S se středem v počátku ($u = 0$) taková, že pro každé $u \in S$ existuje lokální řešení $x^*(u) \in \mathbb{R}^n$ úlohy (8.4) a odpovídající $y^*(u) \in \mathbb{R}^m$. Navíc $x^*(\cdot)$ a $y^*(\cdot)$ jsou spojitě diferencovatelné funkce na S a platí $x^*(0) = x^*$, $y^*(0) = y^*$ a pro každé $u \in S$ máme

$$\operatorname{grad} f^*(u) = -y^*(u),$$

kde $f^*(u)$ značí optimální hodnota úlohy (8.4) vzhledem k u, tj. klademe $f^*(u) := f(x^*(u))$.

Důkaz. Uvažme systém rovnic

$$\operatorname{grad} f(x) + y_1 \operatorname{grad} g_1(x) + \cdots + y_m \operatorname{grad} g_m(x) = 0 \quad \text{a} \quad G(x) := \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} = u. \tag{8.4.10}$$

Pro každé pevné u tento systém představuje $n + m$ rovnic o $n + m$ neznámých, což jsou složky vektorů x a y. Pro $u = 0$ má systém řešení x^* a y^*. Odpovídající $(n + m) \times (n + m)$ Jacobiho matice Lagrangeovy funkce $L(x, y)$ vzhledem k x a y je

$$\mathrm{J} := \begin{pmatrix} \nabla_{xx}^2 L(x^*, y^*) & \mathrm{D}^\top G(x^*) \\ \mathrm{D}G(x^*) & 0 \end{pmatrix}.$$

Ukážeme, že matice $\mathcal{J}$ je regulární. Připusťme opak, tj. existenci nenulového vektoru $\binom{\xi}{\eta} \in \operatorname{Ker} \mathcal{J}$, tj.

$$\nabla_{xx}^2 L(x^*, y^*)\, \xi + \mathrm{D}^\top G(x^*)\, \eta = 0 \quad \text{a současně} \quad DG(x^*)\, \xi = 0. \tag{8.4.11}$$

Odtud po vynásobení první rovnice $\xi^\top$ zleva a druhé $\eta^\top$ taktéž zleva obdržíme

$$\xi^\top \nabla_{xx}^2 L(x^*, y^*)\, \xi = 0,$$

což vzhledem k předpokladu platnosti postačující podmínky druhého řádu dává $\xi = 0$, neboť druhá rovnost v (8.4.11) znamená, že $\xi \in \operatorname{Ker} \mathrm{D}G(x^*)$. Pak ale $\mathrm{D}^\top G(x^*)\eta = 0$ dle první rovnosti v (8.4.11), což vzhledem k předpokladu regulárnosti bodu x^* je možné pouze v případě, že $\eta = 0$. Celkem tedy $\binom{\xi}{\eta} = \binom{0}{0}$, což je spor ↯, tj. matice $\mathcal{J}$ je regulární.

[86] Viz Poznámku 78 zmíněnou na straně 40.

Z regularity matice $\mathcal{J}$ a z Věty o implicitní funkci[86] plyne, že existuje otevřená koule S se středem v počátku taková, že pro každé $u \in S$ existují spojitě diferencovatelné funkce $x^*(u)$ a $y^*(u)$ splňující $x^*(0) = x^*$, $y^*(0) = y^*$, a

$$\operatorname{grad}_x f(x^*(u)) + \underbrace{y_1 \operatorname{grad}_x g_1(x^*(u)) + \cdots + y_m \operatorname{grad}_x g_m(x^*(u))}_{=\mathrm{D}^\top G(x^*(u))\, y \in \mathbb{R}^n} = 0 \quad \text{a} \quad h(x^*(u)) = u. \tag{8.4.12}$$

Pro u dostatečně blízko počátku vektory $x^*(u)$ a $y^*(u)$ splňují postačující podmínku druhého řádu pro danou úlohu z tvrzení věty, jelikož ta je splněna zejména pro $u = 0$. Toto lze ověřit přímo díky požadavkům spojitosti funkcí $f, g_1, \ldots, g_m$. Kdyby to totiž nebyla pravda, existovala by posloupnost $\{u^{[k]}\}_{k=1}^\infty$ splňující $u^{[k]} \to \infty$ pro $k \to \infty$ a také posloupnost $\{\xi^{[k]}\}_{k=1}^\infty$ s normou $\|\xi^{[k]}\| = 1$ tak, že $\mathrm{D}G(x(u^{[k]}))\,\xi^{[k]} = 0$ a

$$\xi^{[k]\top} \nabla_{xx} L(x(u^{[k]}), y(u^{[k]}))\, \xi^{[k]} \le 0$$

pro každé $k \in \mathbb{N}$. Avšak odtud limitním přechodem pro $k \to \infty$ a libovolnou konvergentní podposloupnost z $\{\xi^{[k]}\}_{k=1}^\infty$, jejíž existence je zaručena kompaktností jednotkové koule, na jejímž povrchu leží všechna $\xi^{[k]}$, dostaneme spor s podmínkou druhého řádu ↯. Proto $x^*(u)$ je lokálním řešením uvažované úlohy a $y^*(u)$ jsou odpovídající multiplikátory.

Zbývá dokázat, že $\operatorname{grad} f^*(u) = -y^*(u)$. Z (8.4.12) máme po vynásobení maticí $\mathrm{D}^\top x(u) \in \mathbb{R}^{m\times n}$, že

$$\underbrace{\mathrm{D}^\top x^*(u)}_{\in\mathbb{R}^{m\times n}} \underbrace{\operatorname{grad} f(x^*(u))}_{\in\mathbb{R}^n} + \underbrace{\mathrm{D}^\top x^*(u)}_{\in\mathbb{R}^{m\times n}} \underbrace{\mathrm{D}^\top G(x^*(u))}_{\in\mathbb{R}^{n\times m}} \underbrace{y^*(u)}_{\in\mathbb{R}^m} = 0.$$

Derivováním rovnosti $G(x(u)) = u$ dostaneme

$$I = \begin{pmatrix} \frac{\partial}{\partial x_1} g_1(x^*(u)) \frac{\partial}{\partial u_1} x_1^*(u) + \frac{\partial}{\partial x_2} g_1(x^*(u)) \frac{\partial}{\partial u_1} x_2^*(u) + \cdots + \frac{\partial}{\partial x_n} g_1(x^*(u)) \frac{\partial}{\partial u_1} x_n^*(u) & \cdots \\ \vdots & \vdots \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial}{\partial u_1} x_1^*(u) & \cdots & \frac{\partial}{\partial u_1} x_n(u) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial u_m} x_1^*(u) & \cdots & \frac{\partial}{\partial u_m} x_n(u) \end{pmatrix} \times \begin{pmatrix} \frac{\partial}{\partial x_1} g_1(x^*(u)) & \cdots & \frac{\partial}{\partial x_n} g_m(x^*(u)) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n} g_1(x^*(u)) & \cdots & \frac{\partial}{\partial x_n} g_m(x^*(u)) \end{pmatrix} =$$

$$= \mathrm{D}^\top x^*(u) \times \mathrm{D}^\top G(x^*(u)),$$

což po dosazení do předchozí rovnosti dává

$$\mathrm{D}^\top x^*(u) \operatorname{grad}_x f(x^*(u)) + y^*(u) = 0 \quad \text{neboli} \quad \operatorname{grad} f^*(u) = -y^*(u),$$

neboť

$$\operatorname{grad} f^*(u) = \operatorname{grad}_u f(x^*(u)) = \begin{pmatrix} \frac{\partial}{\partial u_1} f(x^*(u)) \\ \vdots \\ \frac{\partial}{\partial u_m} f(x^*(u)) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1} f(x^*(u)) \frac{\partial}{\partial u_1} x_1^*(u) + \frac{\partial}{\partial x_2} f(x^*(u)) \frac{\partial}{\partial u_1} x_2^*(u) + \cdots + \frac{\partial}{\partial x_n} f(x^*(u)) \frac{\partial}{\partial u_1} x_n^*(u) \\ \vdots \\ \frac{\partial}{\partial x_1} f(x^*(u)) \frac{\partial}{\partial u_m} x_1^*(u) + \frac{\partial}{\partial x_2} f(x^*(u)) \frac{\partial}{\partial u_m} x_2^*(u) + \cdots + \frac{\partial}{\partial x_n} f(x^*(u)) \frac{\partial}{\partial u_m} x_n^*(u) \end{pmatrix} \in \mathbb{R}^m$$

a současně

$$\mathrm{D}^\top x^*(u) \operatorname{grad}_x f(x^*(u)) = \underbrace{\begin{pmatrix} \frac{\partial}{\partial u_1} x_1^*(u) & \cdots & \frac{\partial}{\partial u_1} x_n^*(u) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial u_m} x_1^*(u) & \cdots & \frac{\partial}{\partial u_m} x_n^*(u) \end{pmatrix}}_{\in\mathbb{R}^{m\times n}} \times \underbrace{\begin{pmatrix} \frac{\partial}{\partial x_1} f(x^*(u)) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x^*(u)) \end{pmatrix}}_{\in\mathbb{R}^n} \in \mathbb{R}^m.$$

■

Ve druhé části této podkapitoly se zaměříme na úlohu matematického programování, kde budou vystupovat jen omezení ve tvaru nerovností a parametry budou pro jednoduchost pouze absolutními členy v těchto nerovnostech. To je z praktického pohledu přeci jen užitečnější. Navíc každou rovnost lze zaměnit za dvojici nerovností.

Budeme mít proto úlohu závislou na m-tici parametrů $b = (b_1, \dots, b_m)^\top \in \mathbb{R}^m$, tj.

$$f(x) \to \min, \qquad x \in X(b) := \{x \in P \subseteq \mathbb{R}^n \mid g_i(x) \le b_i,\ i \in 1, \dots, m\}. \tag{8.4.13}$$

K našemu výsledku budeme ještě potřebovat následující označení

$$G(x) := \big(g_1(x), \dots, g_m(x)\big)^\top, \qquad X(b) := \Big\{x \in P \mid G(x) \le b\Big\},$$

$$B := \Big\{b \in \mathbb{R}^m \mid X(b) \neq \emptyset\Big\}, \qquad F(b) := \inf_{x \in X(b)} f(x),\ b \in B,$$

$$Y(b) := \Big\{y \in \mathbb{R}^m \mid y \ge 0,\ F(b) \le f(x) + \langle y, G(x) - b\rangle\ \forall x \in P\Big\},$$

$$\partial F(b) := \Big\{a \in \mathbb{R}^m \mid F(b') - F(b) \ge \langle a, b' - b\rangle\ \forall b' \in B\Big\}.$$

V předchozích podkapitolách jsme pracovali s $X = X(0)$ a $f^* = F(0)$. Z definice množiny B je patrné, že připouštíme i situaci, kdy $X(b)$ je pro některé hodnoty parametru b prázdná. Množina $Y(b)$ je množinou Kuhnových–Tuckerových vektorů úlohy (8.4.13) a množina $\partial F(b)$ je subdiferenciálem funkce $F(b)$, přičemž i tyto množiny mohou být pro některé hodnoty b prázdné.

Příklad 8.4.7

Uvažme úlohu

$$\frac{1}{2}k x^2 \to \min \quad \& \quad x \ge b \tag{8.4.14}$$

pro libovolné $k > 0$. Tato úloha má pro $b > 0$ jednoduchou fyzikální interpretaci: chceme natáhnout pružinu alespoň o b tak, aby se minimalizovala její potenciální energie. V takovém případě je zřejmé, že řešením je nejmenší možné natažení pružiny, tj. $x^* = b$, k čemuž potřebujeme sílu rovnu $k x^*$. Je-li $b \le 0$, pak je dané omezení splněno i volným globálním minimem funkce f, kterým je bod $x^* = 0$. Samozřejmě toto je velmi jednoduchý optimalizační problém, k jehož řešení nejsou Lagrangeovy multiplikátory potřeba. Nicméně my si na něm ilustrujeme závislost řešení na parametru b. Pro jednoduchost budeme uvažovat $k = 2$ a dodejme, že volba jiného k nemá na popsané chování podstatný vliv. Lagrangeova funkce dané úlohy je

$$L(x, y) = x^2 + y(b - x).$$

a můžeme si všimnout, že pro $x = b$ nezávisí na y. Jelikož řešením soustavy

$$\frac{\partial L(x,y)}{\partial x} = 2x - y = 0 \qquad \& \qquad \frac{\partial L(x,y)}{\partial y} = b - x = 0$$

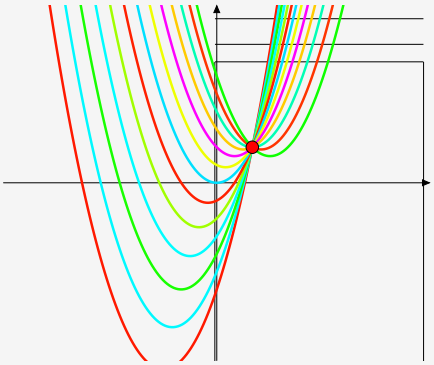
je $x = b$ a $y = 2b$ a $L(x, y)$ je konvexní, realizuje se v tomto bodě globální minimum Lagrangeovy funkce. Je-li $b \ge 0$, pak je tento bod $x^* = b$ $y^* = 2b$ i sedlovým bodem Lagrangeovy funkce, neboť

$$\underbrace{b^2}_{L(b,y)} \le \underbrace{b^2}_{L(b,2b)} \le \underbrace{x^2 + 2b(b - x)}_{L(x,2b)} = (x - b)^2 + b^2$$

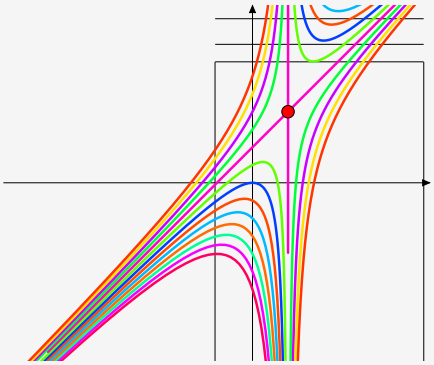
pro libovolné $x \in \mathbb{R}$ a $y \ge 0$ a také

$$\min_{x \in \mathbb{R}} \underbrace{x^2 + 2b(b - x)}_{L(x,2b)} = \underbrace{b^2}_{L(b,2b)} = \max_{y \ge 0} \underbrace{b^2}_{L(b,y)},$$

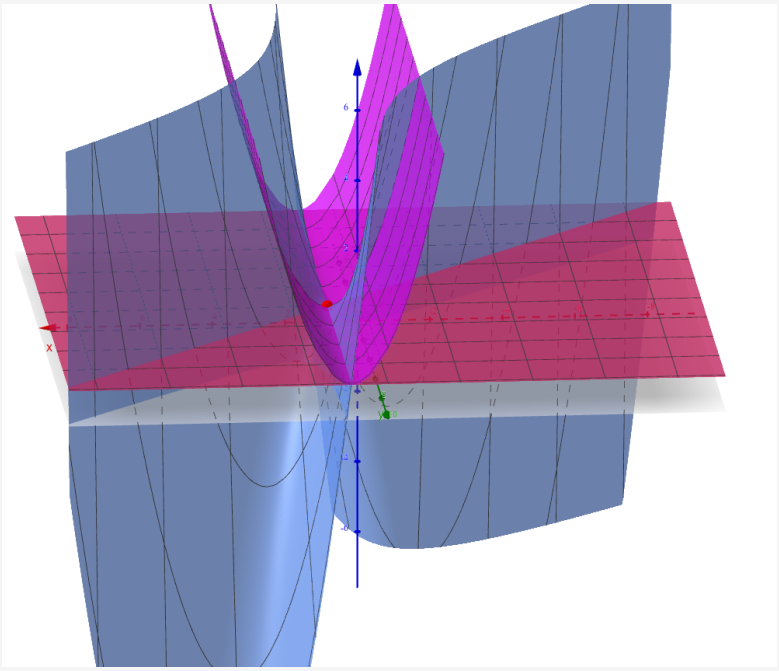
viz Obrázky 8.39.a, 8.39.b a 8.40. V tomto případě odpovídá hodnota Lagrangeova multiplikátoru y^* právě síle potřebné pro dosažení minimální potenciální energie $f^*(b) = b^2/2$.



Obrázek 8.39.a: Lagrangeova funkce úlohy (8.4.14) pro $b = 1$ pro některé hodnoty $y \in [-3,3]$. Všechny paraboly prochází bodem $[1,1]$.



Obrázek 8.39.b: Několik vrstevnic Lagrangeovy funkce úlohy (8.4.14) pro $b = 1$. Bod $[1,2]$ je jejím sedlovým bodem.



Obrázek 8.40: Graf Lagrangeovy funkce úlohy (8.4.14) pro $b = 1$ (modrou barvou) společně s grafem konstantní funkce $z = L(b,y)$ (červenou barvou) a grafem $z = L(x,2b)$ (fialovou barvou). Odtud je opět patrné, že $L(b,y) \le L(x,2b)$ a všechny tyto grafy se protnou v sedlovém bodě $[1,2]$.

Je-li $b < 0$, pak nalezené globální minimum již nesplňuje podmínku $y^* \ge 0$ a sedlovým bodem Lagrangeovy funkce je bez ohledu na konkrétní hodnotu $b < 0$ vždy dvojice $x^* = 0$ a $y^* = 0$, neboť

$$\underbrace{y\,b}_{L(0,y)} \le \underbrace{0}_{L(0,0)} \le \underbrace{x^2}_{L(x,0)}$$

pro libovolné $x \in \mathbb{R}$ a $y \ge 0$ a také

$$\min_{x\in\mathbb{R}} \underbrace{x^2}_{L(x,0)} = \underbrace{0}_{L(0,0)} = \max_{y\ge 0} \underbrace{y\,b}_{L(0,y)},$$

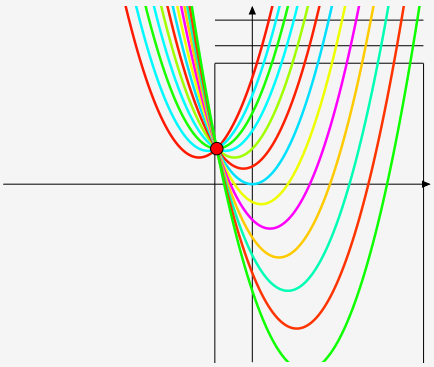
viz Obrázky 8.41.a, 8.41.b a 8.42. Tentokrát v sedlovém bodě platí

$$\left.\frac{\partial L(x,y)}{\partial x}\right|_{\substack{x=0\\y=0}} = 0 \qquad \& \qquad \left.\frac{\partial L(x,y)}{\partial y}\right|_{\substack{x=0\\y=0}} = b < 0$$

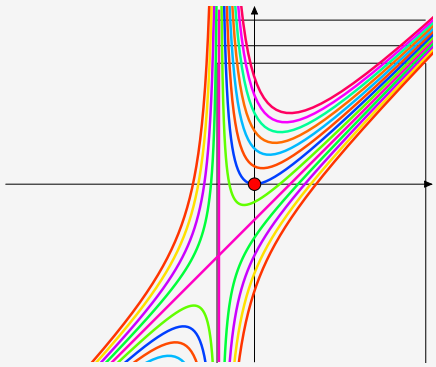
I zde $y^* = 0$ odpovídá síle „potřebné“ pro dosažení minimální potenciální energie $f^*(b) = 0$. Celkem tedy máme

$$[x^*(b), y^*(b)] = \begin{cases} [b, 2b], & b \geq 0 \\ [0,0], & b < 0 \end{cases} \qquad \& \qquad f^*(b) = \begin{cases} b^2/2 & b \geq 0 \\ 0 & b < 0. \end{cases}$$

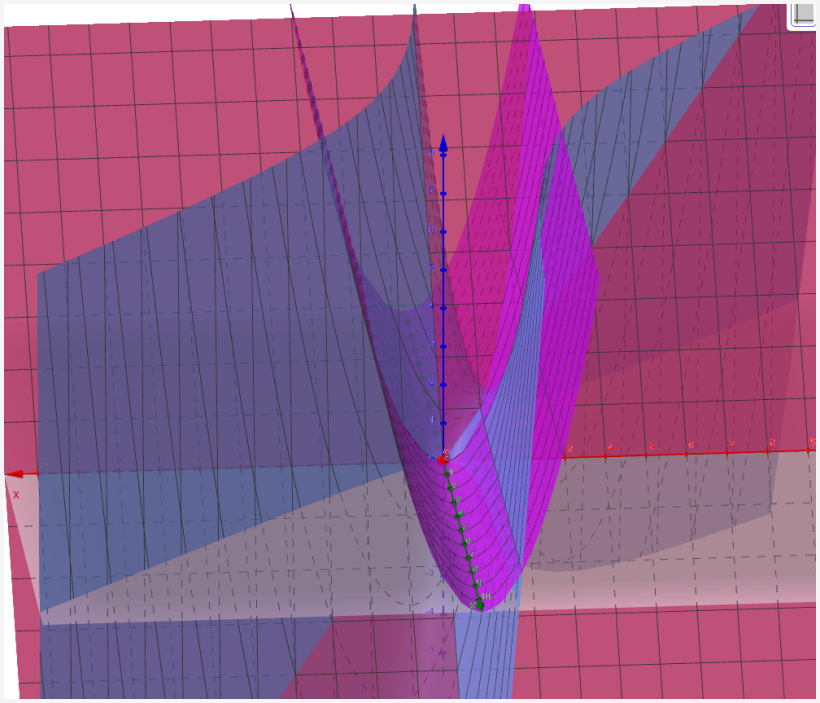
Za povšimnutí stojí, že tentokrát pro hodnoty $[x^*(b), y^*(b)]$ a $f^*(b)$ jsou spojité i pro $b = 0$ a $f^*(b)$ je diferencovatelná pro všechna b.



Obrázek 8.41.a: Lagrangeova funkce úlohy (8.4.14) pro $b = -1$ pro některé hodnoty $y \in [-3,3]$. Všechny paraboly prochází bodem $[-1,-1]$.



Obrázek 8.41.b: Několik vrstevnic Lagrangeovy funkce úlohy (8.4.14) pro $b = -1$. Bod $[0,0]$ je jejím sedlovým bodem.



Obrázek 8.42: Graf Lagrangeovy funkce úlohy (8.4.14) pro $b = -1$ (modrou barvou) společně s grafem funkce $z = L(0,y)$ (červenou barvou) a $z = L(x,0)$ (fialovou barvou). Odtud je opět patrné, že $L(0,y) \leq L(x,0)$ pro $x \in \mathbb{R}$ a $y \geq 0$. Sedlovým bodem Lagrangeovy funkce je tentokrát $[0,0]$.

V následující větě je popsána závislost řešení úlohy (8.4.13) na parametru b.

Věta 8.4.8

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \ldots, g_m$ jsou konvexní na P a platí $0 \in B$, $F(0) > -\infty$ a $Y(0) \neq \emptyset$. Potom

(i) množina B je konvexní;

(ii) funkce $F(b)$ je konečná, konvexní a nerostoucí[87] na B;

(iii) platí $\partial F(b) = -Y(b)$ pro všechna $b \in B$.

[87] Pro funkci více proměnných tímto rozumíme to, že pro libovolná $\widetilde{b}, \widehat{b} \in B$, jejichž odpovídající složky splňují $\widetilde{b}_i \leq \widehat{b}_i$ pro všechna $i \in \{1, \ldots, m\}$, platí $F(\widetilde{b}) \geq F(\widehat{b})$.

Poznámka 8.4.9.

(i) Předpoklady Věty 8.4.8 říkají, že
   (a) $0 \in B$ neboli $X(0) \neq \emptyset$, tj. úloha (8.4.13) s $b = 0$ je přípustná;
   (b) $F(0) > -\infty$, tj. úloha (8.4.13) s $b = 0$ má konečné řešení,
   (c) $Y(0) \neq \emptyset$, tj. množina Kuhnových–Tuckerových vektorů úlohy (8.4.13) s $b = 0$ je neprázdná. Toto není splněno automaticky, neboť současné požadavky nezaručují, že úloha (8.4.13) s $b = 0$ je regulární, viz Větu 8.3.4.

(ii) Je-li funkce F dokonce diferencovatelná v b, pak množina $\partial F(b)$ je jednoprvková, obsahuje pouze $\operatorname{grad}^\top F(b)$ a tento vektor je roven (-1)-násobku Kuhnova–Tuckerova vektoru. To je analogie Věty o obálce (Věta 8.4.6).

(iii) Z důkazu poslední části Věty 8.4.8(iii) uvidíme, že množiny $\partial F(b)$ a $Y(b)$ musí být současně prázdné/neprázdné. Z Věty 6.5.6 víme, že jistě $\partial F(b) \neq \emptyset$ pro každé $b \in \operatorname{ri} B$, ale pro ostatní hodnoty již „nevíme“ nic. Současně je možné, že $Y(b) = \emptyset$, neboť explicitně nepožadujeme regulární úlohu konvexního programování.

(iv) Ve Větě 8.3.16 jsme charakterizovali Kuhnovy–Tucekrovy vektory úlohy (8.1) a (8.2) pomocí řešení duální úlohy (8.3.7). Část (iii) právě zformulované věty jim dává ještě jinou charakteristiku — pomocí subgradientu hodnoty úlohy parametrického programování (8.4.13): v případě regulární úlohy konvexního programování dostáváme zkombinováním těchto dvou výsledků $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy. ▲

Důkaz Věty 8.4.8.

(i) Je-li množina B jednoprvková ($B \neq \emptyset$ neboť $0 \in B$), pak je tvrzení zřejmé. Nechť tedy $\widetilde{b}, \widehat{b} \in B$ a $\lambda \in [0,1]$ jsou libovolná, tj. existují $\widetilde{x} \in X(\widetilde{b})$ a $\widehat{x} \in X(\widehat{b})$. Položme $x := \lambda\widetilde{x} + (1-\lambda)\widehat{x}$. Potom

$$G(x) = G(\lambda\widetilde{x} + (1-\lambda)\widehat{x}) \overset{\substack{G \text{ je} \\ \text{konvexní}}}{\leq} \lambda G(\widetilde{x}) + (1-\lambda) G(\widehat{x}) \leq \lambda\widetilde{b} + (1-\lambda)\widehat{b}. \tag{8.4.15}$$

Tedy $x \in X(\lambda\widetilde{b} + (1-\lambda)\widehat{b})$, tj. $X(\lambda\widetilde{b} + (1-\lambda)\widehat{b}) \neq \emptyset$ a $\lambda\widetilde{b} + (1-\lambda)\widehat{b} \in B$, což dokazuje konvexnost množiny B.

(ii) Nechť $y^* \in Y(0)$. Pak $y^* \geq 0$ a $F(0) \leq f(x) + \langle y^*, G(x)\rangle$ pro všechna $x \in P$. Protože $y^* \geq 0$, platí $F(0) \leq f(x) + \langle y^*, b\rangle$ pro libovolné $b \in B$ a $x \in X(b)$, a tudíž také

$$F(0) \leq F(b) + \langle y^*, b\rangle.$$

Protože $F(0) > -\infty$, je zřejmé, že také $F(b) > -\infty$, a tedy funkce $F(b)$ je konečná[88].

Nechť nyní $\widetilde{b}, \widehat{b} \in B$ a $\lambda \in [0,1]$ jsou libovolná. Položme $b := \lambda\widetilde{b} + (1-\lambda)\widehat{b}$. Pak pro libovolná $\widetilde{x} \in X(\widetilde{b})$ a $\widehat{x} \in X(\widehat{b})$ označme $x := \lambda\widetilde{x} + (1-\lambda)\widehat{x}$. Podle části (i) je $x \in X(b)$, viz (8.4.15), a tedy

$$F(b) \leq f(x) \leq \lambda f(\widetilde{x}) + (1-\lambda) f(\widehat{x}).$$

Protože $\widetilde{x}, \widehat{x}$ byla zvolena libovolně, musí tato nerovnost platit i pro nejmenší hodnoty funkce f na množinách $X(\widetilde{b})$ a $X(\widehat{b})$, z čehož plyne

$$F(b) \leq \lambda F(\widetilde{b}) + (1-\lambda) F(\widehat{b}),$$

tj. funkce F je konvexní.

Konečně, jsou-li $\widetilde{b}, \widehat{b} \in B$ a $\widetilde{b} \leq \widehat{b}$ (po složkách), pak zřejmě $X(\widetilde{b}) \subseteq X(\widehat{b})$. Proto $F(\widetilde{b}) \geq F(\widehat{b})$, a tudíž funkce F je nerostoucí na B.

(iii) Nechť $b \in B$ splňující $\partial F(b) \neq \emptyset$ je libovolné a nechť $y^* \in \partial F(b)$, tj.

$$F(b') - F(b) \geq \langle y^*, b' - b\rangle \quad \text{pro všechna } b' \in B. \tag{8.4.16}$$

Jelikož F je nerostoucí dle (ii), plyne odtud[89] $y^* \leq 0$. Nechť dále $x \in P$ je libovolné a

[88] Hodnota $F(b) = \infty$ by byla možná pouze pro $f \equiv \infty$ ↯.

[89] Je-li $b' - b \geq 0$ (po složkách), pak $F(b') \leq F(b)$ neboli $F(b') - F(b) \leq 0$, takže máme

$$0 \geq F(b') - F(b) \geq \langle y^*, \underbrace{b' - b}_{\geq 0}\rangle,$$

což dává $y^* \leq 0$. Tentýž výsledek získáme analogickým postupem i pro $b' - b \leq 0$ (po složkách).

položme $b' := G(x)$. Pak $b' \in B$, $x \in X(b')$, a tudíž $F(b') \le f(x)$ dle definice $F(b)$. Proto z (8.4.16) máme

$$F(b) \le F(b') - \langle y^*, b' - b\rangle \le f(x) + \langle -y^*, b' - b\rangle = f(x) + \langle -y^*, G(x) - b\rangle. \qquad (8.4.17)$$

Jelikož $x \in P$ bylo zvoleno libovolně, platí nerovnost (8.4.17) pro každé $x \in P$, a tudíž $-y^* \in Y(b)$ dle definice $Y(b)$ a z faktu $-y^* \ge 0$ výše, tj. $\partial F(b) \subseteq -Y(b)$.

Naopak, nechť $b \in B$ splňující $Y(b) \ne \emptyset$ je libovolné a nechť $-y^* \in Y(b)$, tj. $-y^* \ge 0$ a

$$F(b) \le f(x) + \langle -y^*, G(x) - b\rangle \quad \text{pro každé } x \in P.$$

Potom pro každé $b' \in B$ a $x \in X(b')$ dostáváme

$$F(b) \le f(x) + \langle \underbrace{-y^*}_{\ge 0}, G(x) - b\rangle \overset{G(x)\le b'}{\le} f(x) + \langle -y^*, b' - b\rangle.$$

Jelikož tato nerovnost platí pro všechna $x \in X(b')$, musí být splněna i pro nejmenší funkční hodnotu na $X(b')$, tj.

$$F(b) \le F(b') + \langle -y^*, b' - b\rangle \quad \text{pro všechna } b' \in B\,,$$

a tudíž $y^* \in \partial F(b)$, tj. tentokrát $-Y(b) \subseteq \partial F(b)$. Celkem tedy dostáváme $\partial F(b) = -Y(b)$. ■

Právě provedený důkaz Věty 8.4.8 nebyl příliš náročný – „stačila“ především znalost základních definic. Nicméně s pomocí této věty a vlastností konvexních funkcí dokážeme odvodit několik důležitých vlastností původní úlohy matematického programování, tj. úlohy (8.4.13) s $b = 0$. My si toto tvrzení snadno ukážeme při zesílení některých požadavků Věty 8.4.8, viz Slaterovu podmínku ve Větě 8.3.4.

DŮSLEDEK 8.4.10

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_m$ jsou konvexní na P, platí $F(0) > -\infty$ a existuje $\overline{x} \in P$ takové, že $G(\overline{x}) < 0$. Potom $0 \in \operatorname{int} B$ a

(i) funkce $F(\cdot)$ je spojitá v bodě $b = 0$;

(ii) pro libovolné $h \in \mathbb{R}^m$ existuje jednostranná směrová derivace

$$F_{h^+}(0) = \max_{y^* \in Y(0)} \langle -y^*, h\rangle;$$

(iii) funkce F je diferencovatelná v bodě $b = 0$ právě tehdy, když $Y(0)$ je jednoprvková, tj. $Y(0) = \{y^*\}$. Navíc platí $\operatorname{grad}^\top F(0) = -y^*$.

Z části (iii) ihned vyplývá, že pro úlohu (8.4.13) s $b = 0$ (při splnění uvedených předpokladů) existuje více Kuhnových–Tuckerových vektorů právě tehdy, když funkce F není v bodě $b = 0$ diferencovatelná.

Důkaz Důsledku 8.4.10. Protože existuje bod $\overline{x} \in P$ takový, že $G(\overline{x}) < 0 = b$, je zřejmé, že $b = 0 \in \operatorname{int} B$ (neboť $G(\overline{x}) \le b$ pro všechna dostatečně malá b). Tvrzení pak plyne z Věty 8.4.8 a Věty 6.4.1 pro část (i), Věty 6.5.8 pro (ii) a konečně Věty 6.5.9 pro (iii). ■

Příklad 8.4.11

V závislosti na parametru b určeme hodnotu $F(b)$ parametrické úlohy

$$f(x) = x^2 + 2x + 2 \to \min, \quad x \le b.$$

Určeme dále duální úlohu k této úloze, vyřešme ji (bez použití předchozí části) a ověřme platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

Video ☞ https://ln.math.muni.cz/pz/opt/mp7_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp7_zapisnik

Příklad 8.4.12

V závislosti na parametru b určeme hodnotu $F(b)$ parametrické úlohy

$$f(x) = x_1^2 + x_2^2 + 3x_1 \to \min, \quad x_2^2 - x_1 \le b, \quad x_1 \ge 0.$$

Určeme dále duální úlohu k této úloze, vyřešme ji (bez použití předchozí části) a ověřme platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

Video ☞ https://ln.math.muni.cz/pz/opt/mp8_video
Zápisník ☞ https://ln.math.muni.cz/pz/opt/mp8_zapisnik

Na začátku této podkapitoly jsme si ukázali ekonomický význam Lagrangeových multiplikátorů v případě omezení ve tvaru rovností. O totéž se pokusíme i v případě nerovnostních omezení (tentokrát však pro Kuhnův–Tuckerův vektor), k čemuž ale budeme potřebovat ještě následující důsledek.

DŮSLEDEK 8.4.13

(i) Nechť jsou splněny předpoklady Věty 8.4.8 a $y_i^* = 0$ pro všechna $y^* \in Y(0)$ a dané $i \in \{1,\dots,m\}$. Potom $F(\alpha e_i) = F(0)$ pro každé $\alpha \ge 0$.

(ii) Nechť jsou splněny předpoklady Důsledku 8.4.10 a $y_i^* > 0$ pro všechna $y^* \in Y(0)$ a dané $i \in \{1,\dots,m\}$. Potom $F(\alpha e_i) < F(0)$ pro každé $\alpha > 0$.

Důkaz.

(i) V důkazu Věty 8.4.8(ii) jsme ukázali, že $F(0) \le F(b) + \langle y^*, b\rangle$ pro všechna $y^* \in Y(0)$. Volbou $b = \alpha e_i$ dostaneme $F(\alpha e_i) \ge F(0)$. Protože funkce F je nerostoucí, musí současně platit $F(0) \ge F(\alpha e_i)$ pro každé $\alpha \ge 0$. Tedy celkem máme $F(0) = F(\alpha e_i)$.

(ii) Z Důsledku 8.4.10(ii) máme

$$F_{e_i}^+(0) = \lim_{\alpha\to 0^+} \frac{F(\alpha e_i) - F(0)}{\alpha} = \max_{y^*\in Y(0)} \langle -y^*, e_i\rangle = \max_{y^*\in Y(0)} (-y_i^*) < 0,$$

tj. $F(\alpha e_i) - F(0) < 0$ pro $\alpha > 0$ dostatečně malé. Z monotonie funkce F ale plyne, že tato nerovnost musí platit pro všechna $\alpha > 0$. ■

V následujících dvou příkladech si ukážeme slibované ekonomické využití těchto výsledků.

Příklad 8.4.14

Nechť je výrobní proces charakterizován n-rozměrným vektorem $x \in \mathbb{R}^n$, který popisuje výsledek tohoto procesu, přičemž k realizaci tohoto procesu je potřeba m zdrojů, jejichž spotřeba při výrobě charakterizované vektorem x je popsána m-ticí funkcí $G(x) := (g_1(x),\dots,g_m(x))^\top$ splňující $G(x) \ge 0$. Označme dále jako $b := (b_1,\dots,b_m)^\top \ge 0$ vektor zásob zdrojů (nebo výrobní kapacitu), funkci $f(x)$ určující zisk při výrobě na úrovni x a $P \subseteq \mathbb{R}^n$ udávající množinu technologicky možných výrobních procesů[90]. Předpokládejme, že funkce f je konkávní, množina P konvexní a funkce G konvexní (tyto požadavky mají jistou přirozenou ekonomickou interpretaci), $0 \in P$ (tj. podnik může „nevyrábět“) a $G(0) = 0$ (při nulové výrobě je nulová spotřeba materiálu), viz také (6.2.6). Uvažme proto úlohu

$$f(x) \to \max \quad \& \quad G(x) \le b \quad \& \quad x \in P \tag{8.4.18}$$

a označme jako $F(b)$ hodnotu této úlohy. Pak jsou splněny požadavky Věty 8.4.8. Jestliže $x^* = x^*(b)$ je řešením úlohy (8.4.18), tj. $x^*(b)$ je skladba optimálního výrob-

[90] Např. hodnoty $x_1,\dots,x_n$ určující počty zaměstnanců/strojů v různých kategoriích; omezení $g_i(x_1,\dots,x_n) \le b_i$ popisující dostupné kapacity (finance, prostory, poměry mezi kategoriemi; $f(x_1,\dots,x_n)$ udávající zisk z výroby produkované těmito zaměstnanci/stroji. Nebo $x_1,\dots,x_n$ určující komponenty/polotovary (intermediate goods); omezení $g_i(x_1,\dots,x_n)$ vyjadřující potřebný materiál (např. šrouby, výrobní kapacity, ...) a b_i vyjadřující dostupné množství nebo $g_i(x_1,\dots,x_n)$ vyjadřující objem výroby i-tého výrobku z $x_1,\dots,x_n$ (dáváme polotovarům nějakou přidanou hodnotu) a b_i vyjadřující požadované množství i-tého výrobku; $f(x_1,\dots,x_n)$ udávající zisk z tohoto výrobního procesu (materiál ⇝ kompletace ⇝ zisk z finálního výrobku).

ního procesu při omezení zdrojů daném vektorem b, pak hodnota této maximalizační úlohy je rovna

$$\Phi(b) := f(x^*(b)),$$

Připusťme, že ekvivalentní úloha k (8.4.18), tj.

$$-f(x) \to \min \quad \& \quad G(x) - b \leq 0 \quad \& \quad x \in P$$

má jediný Kuhnův–Tuckerův vektor $y^*(b)$, přičemž nezapomeňme, že v tomto případě je $F(b) = -\Phi(b)$. Jestliže při optimální skladbě výroby není i-tý zdroj zcela spotřebován, tj. $g_i(x^*(b)) < b_i$, pak z podmínek komplementarity plyne $y_i^*(b) = 0$ a při splnění požadavků Důsledku 8.4.13(i) (neboli Věty 8.4.8) plyne, že zvětšení zásob i-tého zdroje nevede ke zvýšení zisku, neboť $\Phi(b + \alpha e_i) = \Phi(b)$ pro libovolné $\alpha \geq 0$. V takovém případě říkáme, že i-tý zdroj *není deficitní*. Na druhou stranu je-li $y_i^*(b) > 0$, pak při splnění požadavků Důsledku 8.4.13(ii) (neboli Věty 8.4.8, neboť $G(0) < b$ pro $b > 0$) je i-tý zdroj *deficitní* a zvětšení jeho zásob vede ke zvýšení zisku, neboť $\Phi(b+\alpha e_i) > \Phi(b)$ pro každé $\alpha > 0$, přičemž podle Důsledku 8.4.10

$$\Phi_{h^+}(b) = \langle y^*(b), h \rangle \quad \text{neboli} \quad \operatorname{grad} \Phi(b) = y^*(b).$$

Odtud plyne, že největší „lokální“ nárůst zisku zajistí zvětšení zásob toho zdroje, pro který platí $y_i^*(b) \geq y_j^*(b)$ pro všechna $j \in \{1, \ldots, m\}$, tj. toho, který je „nejvíce deficitní“. Jestliže je možné zvyšovat zásoby všech zdrojů, je vhodné je zvyšovat přímo úměrně (proporcionálně) ke komponentám vektoru $y^*(b) = \left(y_1^*(b), \ldots, y_m^*(b)\right)^\top$.
Kuhnův–Tuckerův vektor se tedy ukazuje jako důležitý indikátor nedostatku používaných zdrojů. My jsme však výše nezohlednili jeden důležitý fakt – pořízení dalších zdrojů bude nepochybně vyžadovat některé náklady na straně podniku. Proto nyní vezmeme v úvahu i tuto okolnost. Nechť $p = (p_1, \ldots, p_m)^\top$ je vektor (kladných) cen jednotlivých zdrojů a připusťme, že podnik může nakupovat potřebné zdroje do svých zásob nebo prodávat „nepotřebné“ zdroje za účelem maximalizace celkové tržby, která zohledňuje i náklady/příjmy spojené s obchodními operacemi se zdroji. Pak úloha 8.4.18 se změní na

$$f(x) - \langle p, h \rangle \to \max \quad \& \quad G(x) \leq b + h \quad \& \quad x \in P \quad \& \quad h \geq -b,$$

kde vektor $h = (h_1, \ldots, h_m)^\top$ popisuje prodej a nákup jednotlivých zdrojů, přičemž $h_i > 0$ znamená, že i-tý zdroj je nakoupen, zatímco v případě $h_i < 0$ je i-tý zdroj prodán. Podmínka $h \geq -b$ zohledňuje to, že podnik nemůže prodat větší množství zdrojů, něž kolik má aktuálně k dispozici. Nechť $[x^*, h^*]$ je řešením této úlohy a y^* je Kuhnův–Tuckerův vektor odpovídající funkcionálnímu omezení $G(x) \leq b + h$. Pak podle Věty 8.3.20 máme

$$L(x^*, h^*, y^*) = \min_{\substack{x \in P \\ h \geq -b}} L(x, h, y^*),$$

kde

$$L(x, h, y) = -f(x) + \langle p, h \rangle + \langle y, G(x) - b - h \rangle.$$

Přirozeně můžeme očekávat, že podnik bude chtít nerentabilní zdroje prodat, tj. bude spíše platit $h > -b$. Pak bude část gradientu Lagrangeovy funkce $L(x, h, y^*)$ vzhledem k h nulová v bodě $[x^*, h^*]$, tj.

$$\operatorname{grad}_h L(x^*, h^*, y^*) = p - y^* = 0,$$

a tedy Kuhnův–Tuckerův vektor y^* ve skutečnosti nebude nic jiného než vektor aktuálních cen jednotlivých zdrojů.

Příklad 8.4.15

Uvažme opět úlohu (8.4.13), kterou můžeme interpretovat takto: „optimální hodnota udává celkovou ztrátu při vymáhání dluhů různých kategorií (např. známe úspěšnost v různých kategoriích) s omezením maximálních nákladů na lidi, soudy vybavení atd.“ Předpokládejme, že stávající rozhodovací proces (jeho omezení) je charakterizován vektorem $b = 0$, tedy minimum účelové funkce na přípustné množině $X(0)$ je $F(0)$. Připusťme, že ale nejsme spokojeni s touto optimální hodnotou $F(0)$ a chceme dosáhnout lepšího výsledku $F(b) < F(0)$ (tj. nižších ztrát při změně investovaných nákladů) vhodnou změnou rozhodovacích procesů, tj. hodnoty omezení b. Nechť náklady na jednotkovou změnu omezení b jsou charakterizovány jistým vektorem $y \in \mathbb{R}^m$, tedy změna omezení z $b = 0$ na hodnotu b znamená náklady $\langle y, b\rangle$. Při důkazu Věty 8.4.8(ii) jsme odvodili nerovnost

$$F(b) + \langle y^*, b\rangle \geq F(0) \quad \text{tj.} \quad F(0) - F(b) \leq \langle y^*, b\rangle.$$

Tuto nerovnost lze interpretovat takto: rozdíl ve ztrátě $F(0) - F(b)$ při přechodu z 0 na b bude nejvýše $\langle y^*, b\rangle$. Je-li vektor y charakterizující jednotkové náklady na změnu omezení právě Kuhnovým–Tuckerovým vektorem úlohy 8.4.13 pro $b = 0$, tj. $y = y^*(0)$, pak tato nerovnost znamená, že náklady vynaložené na změnu omezení jsou společně s hodnotou $F(b)$ (tj. nová ztrátovost plus náklady) vyšší nebo stejné ve srovnání s původní ztrátou. V takovém, případě jsou tedy náklady související se změnou rozhodovacích procesů větší než efekt, kterého touto změnou dosáhneme. Neboli pokud vektor charakterizující náklady na změnu omezení je Kuhnovým–Tuckerovým vektorem původní úlohy, pak snaha o tuto změnu není účelná, neboť náklady vynaložené na tuto změnu jsou vyšší než snížení ztrát, kterého tímto dosáhneme.

Poznámka 8.4.16. Náš výklad v posledních dvou podkapitolách byl založen na Kuhnově–Tuckerově vektoru. V literatuře lze ovšem nalézt také přístupy založené na Fenchelově transformaci. Ukážeme si proto na závěr, jak lze tímto odvodit např. vztah duality $f^* = \varphi^*$ pro úlohy (8.1) a (8.2) a (8.3.7). Uvažme úlohu

$$f(x) \to \min, \quad g_i(x) + b_i \leq 0, \quad i \in \{1, \dots, m\} \quad x \in P \subseteq \mathbb{R}^n$$

(zde přechod od $g_i(x) \leq b_i$ k $g_i(x) + b_i \leq 0$ je pouze z formálních důvodů) a nechť $F(b)$ je hodnota této úlohy. Pak pro konjugovanou funkci platí

$$\begin{aligned}
F^*(y) &= \sup_{b\in B}\{\langle y, b\rangle - F(b)\} = \sup_{b\in\mathbb{R}^m}\{\langle y, b\rangle - \inf_{\substack{x\in P\\ G(x)+b\leq 0}} f(x)\} = \\
&= \sup_{b\in\mathbb{R}^m}\{-\inf_{\substack{x\in P\\ G(x)\leq -b}}\big(f(x) - \langle y, b\rangle\big)\} = \sup_{b\in\mathbb{R}^m}\sup_{\substack{x\in P\\ G(x)\leq -b}}\{\langle y, b\rangle - f(x)\} = \\
&= \sup_{x\in P}\sup_{b\leq -G(x)}\{\langle y, b\rangle - f(x)\} = \\
&= \begin{cases} \sup_{x\in P}\{-\langle y, G(x)\rangle - f(x)\} = -\inf_{x\in P}\{f(x) + \langle y, G(x)\rangle\}, & y \geq 0, \\ \infty, & y < 0. \end{cases}
\end{aligned}$$

Tedy $F^*(y) = -\varphi(y)$ a duální úlohu (8.3.7) lze psát jako

$$-F^*(y) \to \max, \quad y \geq 0.$$

Navíc platí

$$F^{**}(0) = \sup_{y\geq 0}\{\langle 0, y\rangle - F^*(y)\} = \sup_{y\geq 0}\{-F^*(y)\} = \sup_{y\geq 0}\{\varphi(y)\} = \varphi^*.$$

Pokud je funkce F spojitá v 0, platí podle Věty 6.6.13 rovnost $F(0) = F^{**}(0)$, tedy platí vztah duality $f^* = \varphi^*$. To znamená, že každá podmínka, která zajistí spojitost funkce F v 0 je zároveň postačující podmínkou pro platnost vztahu duality, např. Slaterova podmínka a požadavek $F(0) > -\infty$ implikují $0 \in \operatorname{int} B$, a tedy F je spojitá podle Vět 6.4.1 a 8.4.8(ii). ▲

Poznámka 8.4.17. V celé této kapitole a částečně i v té předchozí hrála ústřední roli konvexnost. Zejména v případě numerických metod je důvodem je to, že nám tato vlastnost poskytuje dostatek informací o chování funkce v okolí aktuálního odhadu a umožňuje nám se přesunout do jiného bodu, který nám přinese podstatné zlepšení. Algoritmus, který je založen na této myšlence „lokálně optimální volby“ v každém kroku, se ve výpočtové vědě (či matematické informatice) nazývá *hladovým algoritmem*[91]. Proto dokud je náš problém konvexní můžeme se držet motta Gordona Gekka[92]: „Chamtivost je správná. Chamtivost funguje.“[93]

Proto je užitečné učinit konvexními, co nejvíce problémů můžeme. Např. máme-li v úloze (8.1) a (8.2) rovnostní omezení $g(x) = 0$ pro funkci $g : \mathbb{R}^n \to \mathbb{R}$, která je sice konvexní, ale není afinní, pak máme nekonvexní problém. Tento nedostatek se můžeme pokusit odstranit tím, že tuto rovnost nahradíme nerovností $g(x) \leq 0$, která již požadavek konvexnosti splňovat bude. Tím však současně zvětšíme přípustnou množinu a může se objevit řešení, které není přípustné pro původní úlohu. Předejít této situaci můžeme obvykle dvěma způsoby: (i) ukázat, že jakékoli řešení zvětšeného problému bude splňovat $g(x) = 0$; (ii) nebo můžeme vyřešit zvětšený problém a následně ověřit, zda v nalezeném řešení platí $g(x) = 0$. I když ani jedna z možností nezafunguje, může nám tato analýza poskytnout užitečné informace k nalezení řešení původního problému. Tato technika je známa jako *konvexní rozvolnění* (nebo *relaxace*)[94,95].

Nyní předpokládejme, že máme minimalizační úlohu s účelovou funkcí $f(x) = \max\{r_i(x) \mid i \in I\}$, kde I je nějaká indexová množina a každá funkce r_i je konvexní a diferencovatelná. Funkce f je také konvexní[96], ale rozhodně už nemusí být diferencovatelná. Existuje ale jednoduchý trik, jak z tohoto učinit úlohu konvexního programování s diferencovatelnou (dokonce lineární) účelovou funkcí. Nejdříve zavedeme pomocnou pomocnou proměnnou t a budeme minimalizovat novou účelovou funkce $\tilde{f}(x,t) := t$ vzhledem k původním omezením doplněním o podmínku $f(x) - t \leq 0$. Tomuto přechodu se říká tzv. *nadgrafová transformace*[97], neboť dodatečnému omezení pro dané t vyhovují všechny body z dolní obrysové množiny funkce f na úrovni t, tj. přesunuli jsme se z definičního oboru funkce f do jejího nadgrafu. Zatím jsme si však moc nepomohli, neboť teď máme omezení určené nediferencovatelnou funkcí f. To lze však snadno napravit, neboť $f(x) \leq t$ je ekvivaletní s požadavkem $r_i(x) \leq t$ pro všechna $i \in I$, takže přidáním jednoho takového omezení pro každou funkci $r_1, r_2, \ldots$ již budeme mít diferencovatelnou úlohu konvexního programování. ▲

[91] *greedy algorithm*

[92] Gordon Gekko je fiktivní postava z filmů Wall Street a Wall Street: Peníze nikdy nespí, kterou ztvárnil Michael Douglas (za svůj výkon získal Oscara).

[93] Greed is right. Greed works.

[94] *convex relaxation*

[95] Podobnou myšlenku jsme využili při řešení úloh celočíselného programování, kdy jsme „ignorovali“ požadavek celočíselnosti a řešili přidruženou úlohu lineárního programování.

[96] Viz také Cvičení 6.2.14.

[97] *epigraph transformation*

Cvičení

8.4.1. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\sqrt{1+x^2} \to \min \quad \& \quad x \leq b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 1$ pro $b \geq 0$ a $F(b) = \sqrt{1+b^2}$ pro $b \leq 0$

8.4.2. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x \ln x \to \min \quad \& \quad 0 < x \leq b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1/\mathrm{e}$ pro $b \geq 1/\mathrm{e}$ a $F(b) = b \ln b$ pro $b \in (0, 1/\mathrm{e}]$

8.4.3. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$|x| \to \min \quad \& \quad x \leq b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \geq 0$ a $F(b) = -b$ pro $b < 0$

8.4.4. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$e^{|x|} \to \min \quad \& \quad x \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 1$ pro $b \ge 1$ a $F(b) = \mathrm{e}^{-b}$ pro $b \le 0$

8.4.5. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\cosh(x) := \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \to \min \quad \& \quad x \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = \cosh(b)$ pro $b \le 0$ a $F(b) = 1$ pro $b \ge 0$

8.4.6. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 4x_2^2 \to \min \quad \& \quad x_1 + 2x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge 0$ a $F(b) = b^2/2$ pro $b \le 0$

8.4.7. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 \to \min \quad \& \quad x_1 - x_2 - 1 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge -1$ a $F(b) = (b+1)^2/2$ pro $b \le -1$

8.4.8. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 + 2x_1 \to \min \quad \& \quad -x_1 + x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1$ pro $b \ge 1$ a $F(b) = (b-1)^2/2 - 1$ pro $b \le -1$

8.4.9. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 2x_2^2 - 2x_1 + 1 \to \min \quad \& \quad x_1 + x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge 1$ a $F(b) = 2(b-1)/3$ pro $b \le 1$

8.4.10. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 + 2x_2 \to \min \quad \& \quad x_2 - x_1 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1$ pro $b \ge -1$ a $F(b) = (b+1)^2/4 - 1$ pro $b \le -1$

8.4.11. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 + 2x_1 \to \min \quad \& \quad x_1 + x_2 + 1 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1$ pro $b \ge 0$ a $F(b) = b^2/2 - 1$ pro $b \le 0$

8.4.11.A) V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 - 2x_2 \to \min \quad \& \quad x_1 + x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1$ pro $b \ge 1$ a $F(b) = (b-1)^2/2 - 1$ pro $b \le 1$

8.4.12. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 2x_1x_2 + 2x_2^2 \to \min \quad \& \quad x_1 - x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge 0$ a $F(b) = b^2/5$ pro $b \le 0$

8.4.13. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$3x_1^2 + 2x_1x_2 + x_2^2 - 2x_1 + x_2 \to \min \quad \& \quad x_1 + 2x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -11/8$ pro $b \ge -7/4$ a $F(b) = 2b^2/9 + 7b/9 - 25/36$ pro $b \le -7/4$

8.4.14. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\sqrt{1 + x_1^2 + x_2^2} \to \min \quad \& \quad x_1 + x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 1$ pro $b \ge 0$ a $F(b) = \sqrt{1 + b^2/2}$ pro $b \le 0$

8.4.15. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\sqrt{1 + 4x_1^2 + 4x_2^2} \to \min \quad \& \quad x_2 - x_1 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 1$ pro $b \ge 0$ a $F(b) = \sqrt{1 + 2b^2}$ pro $b \le 0$

8.4.16. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1 + x_2 \to \min \quad \& \quad x_1^2 + x_2^2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -\sqrt{2b}$ pro $b \ge 0$

8.4.17. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$3x_1^2 + x_2^2 + 4x_2 - 1 \to \min \quad \& \quad x_1^2 + 2x_2 \le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -5$ pro $b \ge -4$ a $F(b) = b^2/4 + 2b - 1$ pro $b \le -4$

8.4.18. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$4x_1^2 + x_2^2 - 2x_2 \to \min \quad \& \quad x_1^2 + 5x_2^2 \le b + 1.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = -1$ pro $b \ge 4$ a $F(b) = (b+1)/5 - 2\sqrt{(b+1)/5}$ pro $b \in [-1, 4]$

8.4.19. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$2x_1^2 + 2x_1x_2 + x_2^2 \to \min \quad \& \quad x_1 + x_2 \le b \quad \& \quad x_1 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge 0$ a $F(b) = b^2$ pro $b \le 0$

8.4.20. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 6x_2^2 + 2x_1 - 8x_2 \to \min \quad \& \quad 2x_2 - x_1 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = b^2 - 2b$ pro $b \le -1$, $F(b) = (3b^2 - 14b - 2)/5$ pro $b \in [-1, 7/3]$ a $F(b) = -11/3$ pro $b \ge 7/3$

8.4.21. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 6x_2^2 + 2x_1 - 8x_2 \to \min \quad \& \quad x_1 - 2x_2 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.22. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 6x_2^2 - 2x_1 + 8x_2 \to \min \quad \& \quad x_1 - 2x_2 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.23. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 - 3x_1x_2 + 4x_2^2 - 2x_1 - x_2 \to \min \quad \& \quad 2x_2 - x_1 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.24. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 3x_1x_2 + 4x_2^2 - 2x_1 + x_2 \to \min \quad \& \quad x_1 + 2x_2 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.25. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 + x_2 \to \min \quad \& \quad 3x_1 + 2x_2 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = b^2/9 - 2b/3$ pro $b \le 21/10$, $F(b) = b^2/34 - 11b/34 - 49/136$ pro $b \in [21/10, 11/2]$ a $F(b) = -5/4$ pro $b \ge 11/2$

8.4.26. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + 4x_2^2 + 2x_2 \to \min \quad \& \quad 3x_1^2 - x_2 \le b \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b \ge 0$ a $F(b) = 4b^2 - 2b$ pro $b \le 0$

8.4.27. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$2x_1^2 + 2x_1x_2 + x_2^2 \to \min \quad \& \quad x_1 + x_2 \le b \quad \& \quad x_1 \ge 0 \quad \& \quad x_2 \ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.28. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$-\sqrt{x_1}-\sqrt{3x_2}\to\min \quad \& \quad x_1+x_2\le b \quad \& \quad x_1\ge 0 \quad \& \quad x_2\ge 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=-2\sqrt{b}$ pro $b\ge 0$

8.4.29. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{x_1}+\frac{1}{x_2}\to\min \quad \& \quad x_1+x_2\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=4/b$ pro $b>0$

8.4.30. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{x_1+x_2}\to\min \quad \& \quad x_1^2+x_2^2\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=1/\sqrt{2b}$ pro $b>0$

8.4.31. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{\sqrt{x_1x_2}}\to\min \quad \& \quad x_1+x_2\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=2/b$ pro $b>0$

8.4.32. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{\sqrt{x_1}}+\frac{1}{\sqrt{x_2}}\to\min \quad \& \quad x_1+x_2\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=2\sqrt{2/b}$ pro $b>0$

8.4.33. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2+x_2^2+x_3^2\to\min \quad \& \quad x_1+x_2+x_3\le b.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=0$ pro $b\ge 0$ a $F(b)=b^2/3$ pro $b\le 0$

8.4.34. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}\to\min \quad \& \quad x_1+x_2+x_3\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0 \quad \& \quad x_3>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b)=9/b$ pro $b>0$

8.4.35. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{x_1^2+x_2^2+x_3^2}\to\min \quad \& \quad x_1+x_2+x_3\le b \quad \& \quad x_1>0 \quad \& \quad x_2>0 \quad \& \quad x_3>0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b)=-Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 3/b^2$ pro $b > 0$

8.4.36. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\frac{1}{x_1 x_2 x_3} \to \min \quad \& \quad x_1 + x_2 + x_3 \le b \quad \& \quad x_1 > 0 \quad \& \quad x_2 > 0 \quad \& \quad x_3 > 0.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 27/b^3$ pro $b > 0$

8.4.37. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 \to \min \quad \& \quad x_1 \le b_1 \quad \& \quad x_2 \le b_2.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b_1, b_2 \ge 0$,

$F(b) = b_2^2$ pro $b_1 \ge 0$ & $b_2 \le 0$,

$F(b) = b_1^2$ pro $b_1 \le 0$ & $b_2 \ge 0$,

$F(b) = b_1^2 + b_2^2$ pro $b_1 \le 0$ & $b_2 \le 0$

8.4.38. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$x_1^2 + x_2^2 \to \min \quad \& \quad x_1 + x_2 \le b_1 \quad \& \quad x_1 - x_2 \le b_2.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = 0$ pro $b_1, b_2 \ge 0$,

$F(b) = b_2^2/2$ pro $b_1 \ge 0$ & $b_2 \le 0$,

$F(b) = b_1^2/2$ pro $b_1 \le 0$ & $b_2 \ge 0$,

$F(b) = b_1^2/2 + b_2^2/2$ pro $b_1 \le 0$ & $b_2 \le 0$

8.4.39. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$\mathrm{e}^{x_1+x_2} \to \min \quad \& \quad x_2 - 2x_1 \le b_1 \quad \& \quad x_1 - 2x_2 \le b_2.$$

Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = \mathrm{e}^{-b_1-b_2}$

8.4.40. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$f(x) = f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{\alpha_i}{x_i} \to \min \quad \& \quad \sum_{i=1}^{n} x_i \le b \quad \& \quad x_1 > 0 \quad \& \quad \ldots \quad \& \quad x_n > 0,$$

kde $\alpha_i > 0$ jsou reálné konstanty. Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

řešení: $F(b) = \alpha^2/b$ pro $\alpha := \sum_{i=1}^{n} \sqrt{\alpha_i}$

8.4.41. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$f(x) = f(x_1, \ldots, x_n) = -\sum_{i=1}^{n} \alpha_i \sqrt{x_i} \to \min \quad \& \quad \sum_{i=1}^{n} x_i \le b \quad \& \quad x_1 \ge 0 \quad \& \quad \ldots \quad \& \quad x_n \ge 0,$$

kde $\alpha_1 > 0, \ldots, \alpha_n > 0$ jsou reálné konstanty. Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

8.4.42. V závislosti na parametru b určete hodnotu $F(b)$ parametrické úlohy matematického programování

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle c, x\rangle \to \min, \quad \langle a, x\rangle \le b,$$

kde $C \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní matice, $x, c, a \in \mathbb{R}^n$. Určete dále duální úlohu k této úloze, vyřešte ji (bez využití předchozí části) a ověřte platnost vztahu $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy.

Kapitola 9
Variační počet

9.1 *Není to jen o nejrychlejší klouzačce* 700
9.2 *Variace bez kombinace* 700
9.3 *Základní úloha variačního počtu* 705
9.4 *Úloha o brachistochroně* 710

Tua solutio problematis isoperimetrici continet, ut video, quidquid in hac questione desiderari potest; et ego maxime gaudeo hoc argumentum, quod fere solus post primos conatus tractaveram, a te potissimum ad summum perfectionis fastigium esse evectum. Rei dignitas me excitavit ut tuis luminibus adjutus, ipse solutionem analyticam conscripserim quam tamen celare statui, donec ipse tuas meditationes publici juris feceris, ne ullam partem gloriae tibi debitae praeripiam.[1]

Leonhard Euler

It seems to me that Berlin would not be at all suitable for me while M Euler is there.[2]

If I had been rich, I probably would not have devoted myself to mathematics.[3]

Eulerův výrok pochází z dopisu Lagrangeovi datovaném 2. října 1759, viz soupis jejich vzájemné korespondence https://ln.math.muni.cz/pz/opt/EulerLagrange. Další dva výroky pochází z Lagrangeovy biografie https://ln.math.muni.cz/pz/opt/lagrange1, přičemž ten první je úryvkem z Lagrangeova druhého odmítnutí pozice na Královské Pruské akademii věd v Berlíně z roku 1766 (poprvé mu byla tato pozice nabídnuta již v roce 1756 – kdepak asi v té době byl pan Euler?).

[1] *Vaše řešení isoperimetrického problému neponechává nic, co by se mělo ještě objasnit, a já jsem nesmírně rád, že šťastný los připadl na Vás, abyste řešení problému dotáhl na nejvyšší stupeň dokonalosti a vytvořil teorii, kterou jsem od samého počátku rozvíjel jen já. Důležitost tohoto tématu mě podnítila za pomoci Vašich k vybudování analytického řešení, které budu držet v tajnosti tak dlouho, dokud Vaše rozvahy nebudou publikovány, abych Vám neubral část slávy, kterou si zasloužíte.*
(Your solution of the isoperimetric problem contains, insofar as I can see, all that may be desired in this area, and I am extremely happy that this theory, which I hardly touched upon after my first attempts, has been brought to such very great perfection by you. The importance of that subject has stimulated me to develop, aided by your lights, an analytical solution which I will keep secret as long as your own meditations are not published, lest I take away from you a part of the glory which you deserve.)

[2] *Mně se zdá, že Berlín, dokud tam je pan Euler, by pro mě nebyl vůbec vhodný.*

[3] *Kdybych byly bohatý, pravděpodobně bych se nevěnoval matematice.*

9.1 Není to jen o nejrychlejší klouzačce

9.2 Variace bez kombinace

Problémy variačního počtu zmíněné v úvodní kapitole patří mezi tzv. *základní úlohy*, ve kterých má funkcionál $\mathcal{J}[\cdot]$ velmi specifický tvar. Než se ale pustíme do jejich řešení, musíme si vybudovat alespoň některé základní teoretické nástroje, které vlastně budou analogií tvrzení známých z diferenciálního počtu více proměnných. Připomeňme, že funkci $f:\mathbb{R}^n\to\mathbb{R}$ definovanou v nějakém okolí bodu x_0 nazveme v tomto bodě *diferencovatelnou*, jestliže pro funkci $\omega:\mathbb{R}^n\to\mathbb{R}$ danou vztahem

$$\omega(h) := f(x_0+h) - f(x_0) - (a_1 h_1 + \cdots + a_n h_n) = f(x_0+h) - f(x_0) - \langle a, h\rangle$$

existuje n-tice $a := (a_1,\dots,a_n)^\top \in \mathbb{R}^n$ taková, že

$$\lim_{h\to 0} \frac{\omega(h)}{\|h\|} = 0. \tag{9.2.1}$$

Takový vektor a pak je nutně určen jednoznačně. Nazýváme jej *gradientem* funkce f v bodě x_0 a pro jeho složky platí $a_1 = f_{x_1}(x_0)$, …, $a_n = f_{x_n}(x_0)$, zatímco odpovídající skalární součin jakožto funkce $\langle a, h\rangle : \mathbb{R}^n \to \mathbb{R}$ se nazývá (*totálním*, *silným* nebo *Fréchetovým*) *diferenciálem*. To znamená, že v okolí bodu x_0 platí

$$\Delta f(x_0) := f(x_0+h) - f(x_0) = \langle a, h\rangle + \omega(h), \tag{9.2.2}$$

tj. máme *přírůstek závisle proměnné* f na levé straně vyjádřený pomocí lineární části $\langle a, h\rangle$ a nelineární části $\omega(h)$, která je v jistém smyslu „malá“.

Naším cílem nyní bude zobecnit vyjádření (9.2.2) pro funkcionály a s jeho pomocí odvodit některá tvrzení důležitá pro řešení optimalizačních úloh. Ovšem naše výsledky budou muset být formulovány trochu jinak, než jsme to dělali v základních kurzech matematické analýzy. V nich jsme totiž podmínky pro lokální/globální extrémy budovali pomocí (parciálních) derivací nikoli pomocí diferenciálu, který jsme si zavedli „tak trochu bokem“. Budete-li mít ale stále na paměti spojení mezi (totálním) diferenciálem a (parciálními) derivacemi, tak analogie snad bude víceméně zřejmá (alespoň na začátku). Také bude potřeba zohlednit to, že už obecně nebude možné perturbovat nezávislou proměnnou libovolným „vektorem“, jak jsme to dělali v případě $x_0 + h$, tj. budeme se moci pohybovat pouze v jisté přípustné množině.

Mějme nyní dán funkcionál $\mathcal{J}[\cdot] : \mathcal{V} \to \mathbb{R}$ na normovaném vektorovém prostoru $\mathcal{V}$ s normou $\|\cdot\|$. Naším úkolem je najít minimum/maximum tohoto funkcionálu na *přípustné* množině $\mathcal{N} \subseteq \mathcal{V}$. Taková množina ovšem samozřejmě nemusí být vektorovým podprostorem, jak lze snadno vidět na příkladu $\mathcal{N} = \{y \in \mathcal{V} \mid y(a) = \alpha,\ \ y(b) = \beta\}$ s volbou $\alpha \neq 0$ nebo $\beta \neq 0$. Pak totiž pro $y_1, y_2 \in \mathcal{N}$ a $c_1, c_2 \in \mathbb{R}$ máme $c_1\, y_1(a) + c_2\, y_2(a) = (c_1 + c_2)\,\alpha$ a $c_1\, y_1(b) + c_2\, y_2(b) = (c_1 + c_2)\,\beta$, z čehož lze snadno vyvodit, že vyjma několika velmi speciálních voleb platí $c_1\, y_1 + c_2\, y_2 \notin \mathcal{N}$. Společně s množinou $\mathcal{N}$ budeme potřebovat ještě také tzv. přípustné *variace* nezávisle proměnné, jejichž přičtením k $y \in \mathcal{N}$ se neporuší přípustnost, tj. funkce $\eta \in \mathcal{V}$ takové, že $y + \eta \in \mathcal{N}$ pro libovolné $y \in \mathcal{N}$. Tuto množinu označíme jako $\mathcal{N}_0$. Např. pro výše zmíněnou množinu $\mathcal{N}$ budou přípustné variace tvořit vektorový podprostor $\mathcal{N}_0 = \{\eta \in \mathcal{V} \mid \eta(a) = 0 = \eta(b)\}$.

Před následující definicí ještě bude užitečné připomenout, že (*spojitým*) *lineárním funkcionálem* na normovaném vektorovém prostoru $\mathcal{N}_0$ rozumíme zobrazení $\varphi : \mathcal{N}_0 \to \mathbb{R}$ splňující

$$\varphi(c_1\,\eta_1 + c_2\,\eta_2) = c_1\,\varphi(\eta_1) + c_2\,\varphi(\eta_2) \quad \text{pro libovolná } c_1, c_2 \in \mathbb{R} \text{ a } \eta_1, \eta_2 \in \mathcal{N}_0$$

a spojité pro všechna $\eta \in \mathcal{N}_0$, tj. pro každé $\eta \in \mathcal{N}_0$ a libovolné $\varepsilon > 0$ existuje $\delta > 0$ takové, že $|\varphi(\eta) - \varphi(\tilde\eta)| < \varepsilon$ pro všechna $\tilde\eta \in \mathcal{N}_0$ splňující $\|\eta - \tilde\eta\| < \delta$. Takovými funkcionály jsou např.

(i) $\varphi(\eta) := \eta(x_0)$ pro nějaké pevně dané $x_0 \in [a, b]$,

(ii) $\varphi(\eta) := \int_a^b \eta(x)\,\mathrm{d}x$ pro $\eta \in \mathrm{C}[a,b]$,

(iii) $\varphi(\eta) := \int_a^b \alpha(x)\eta(x)\,\mathrm{d}x$ pro $\eta \in \mathrm{C}[a,b]$ a pevně zvolenou funkci $\alpha \in \mathrm{C}[a,b]$,

(iv) $\varphi(\eta) := \int_a^b \left[\alpha_0(x)\eta(x) + \alpha_1(x)\eta'(x) + \cdots + \alpha_n(x)\eta^{(n)}(x)\right]\mathrm{d}x$ pro $\eta \in \mathrm{C}^n[a,b]$ a pevně zvolené funkce $\alpha_0, \ldots, \alpha_n \in \mathrm{C}[a,b]$.

DEFINICE 9.2.1

Nechť $y \in \mathcal{N}$ je dáno. Přírůstek funkcionálu $\mathcal{J}[y]$ odpovídající přírůstku $\eta \in \mathcal{N}_0$ v „nezávislé proměnné“ y označíme jako

$$\Delta\mathcal{J}_y[\eta] = \Delta\mathcal{J}[y;\eta] := \mathcal{J}[y+\eta] - \mathcal{J}[y].$$

Funkcionál $\mathcal{J}[\cdot]$ nazveme *diferencovatelným* v y, jestliže

$$\Delta\mathcal{J}_y[\eta] = \varphi(\eta) + \omega(\eta)\,\|\eta\|, \tag{9.2.3}$$

pro všechny přípustné variace $\eta \in \mathcal{N}_0$, nějaký lineární funkcionál $\varphi(\eta)$ a nelineární funkcionál $\omega(\cdot)$ splňující $\omega(\eta) \to 0$ pro $\|\eta\| \to 0$. Funkcionál φ se nazývá *variace* (nebo *diferenciál*) *funkcionálu* $\mathcal{J}[\cdot]$ v y a značí se $\mathcal{J}'[y;\eta]$, příp. jen $\mathcal{J}'[y]$ nebo $\delta\mathcal{J}[y]$.

VĚTA 9.2.2

Variace $\mathcal{J}'[y;\eta]$ diferencovatelného funkcionálu $\mathcal{J}[\cdot]$ je určena jednoznačně.

Důkaz. Nejdříve ukážeme, že jediným lineárním funkcionálem φ splňujícím $\varphi(\eta)/\|\eta\| \to 0$ pro $\|\eta\| \to 0$ je nulový funkcionál, tj. $\varphi \equiv 0$. Připusťme proto, že $\varphi(\eta)/\|\eta\| \to 0$ a současně $\varphi(\eta_0) \neq 0$ pro nějaké $\eta_0 \in \mathcal{N}_0 \setminus \{0\}$. Pak $\lambda := \varphi(\eta_0)/\|\eta_0\| \neq 0$ a pro nenulovou posloupnost $\eta_n := \frac{1}{n}\eta_0$ platí

$$\|\eta_n\| = \tfrac{1}{n}\|\eta_0\| \to 0 \quad \text{pro } n \to \infty,$$

takže vzhledem k linearitě funkcionálu φ máme

$$\frac{\varphi(\eta_n)}{\|\eta_n\|} = \frac{\frac{1}{n}\varphi(\eta_0)}{\frac{1}{n}\|\eta_0\|} = \lambda \neq 0,$$

což je ve sporu s předpokladem $\varphi(\eta)/\|\eta\| \to 0$. ↯

Nechť nyní platí

$$\Delta\mathcal{J}_y[\eta] = \varphi_1(\eta) + \omega_1(\eta)\,\|\eta\|,$$
$$\Delta\mathcal{J}_y[\eta] = \varphi_2(\eta) + \omega_2(\eta)\,\|\eta\|$$

pro nějaké funkcionály $\varphi_1, \varphi_2, \omega_1, \omega_2$ splňující požadavky Definice 9.2.1. Pak

$$0 = \varphi_1(\eta) - \varphi_2(\eta) + [\omega_1(\eta) - \omega_2(\eta)]\,\|\eta\|,$$

takže $[\varphi_1(\eta) - \varphi_2(\eta)]/\|\eta\| = \omega_1(\eta) - \omega_2(\eta) \to 0$ pro $\|\eta\| \to 0$. Z první části důkazu proto vyplývá, že $\varphi_1(\eta) - \varphi_2(\eta) \equiv 0$, tj. $\varphi_1(\eta) = \varphi_2(\eta)$, což dokazuje uvedenou jednoznačnost. ■

Pro funkci $f : \mathbb{R}^n \to \mathbb{R}$ definujeme lokální extrémy v bodě x_0 pomocí nerovností $f(x_0) \leq f(x)$ a $f(x_0) \geq f(x)$ platné pro všechna x z nějakého okolí $\mathcal{O}(x_0)$. Zcela analogicky to učiníme i pro funkcionály.

DEFINICE 9.2.3

Funkcionál $\mathcal{J}[\cdot]$ nabývá na množině $\mathcal{N}$ *lokálního extrému* pro funkci $\widehat{y} \in \mathcal{N}$, pokud výraz $\mathcal{J}[y] - \mathcal{J}[\widehat{y}]$ nemění znaménko v nějakém okolí $\widehat{y}$, tj. pro všechny funkce $y \in \mathcal{N}$ splňující $\|y - \widehat{y}\| < \varepsilon$ pro nějaké $\varepsilon > 0$.

Ačkoliv po formální stránce jsou definice lokálních extrémů pro funkci f a funkcionál $\mathcal{J}$ „totožné“, je mezi nimi jeden zásadní rozdíl. Ten tkví v pojmu okolí bodu $x_0 \in \mathbb{R}^n$ a funkce $y \in \mathcal{N} \subseteq \mathcal{V}$, které je vždy určeno nějakou metrikou či normou. V $\mathbb{R}^n$ to je typicky *eukl(e)idovská* norma, avšak lokální extrém se nezmění, pokud zvolíme jinou normu, neboť v konečně dimenzionálním prostoru jsou všechny normy ekvivalentní. Jenže vektorový prostor $\mathcal{V}$ je nekonečně dimenzionální, takže volba konkrétní metriky zde již může hrát zcela zásadní roli. Toto si ilustrujeme v následujícím velmi důležitém příkladě s prostorem spojitě diferencovatelných funkcí $\mathrm{C}^1[a,b]$, se kterým budeme pracovat v dalších částech této kapitoly.

Příklad 9.2.4

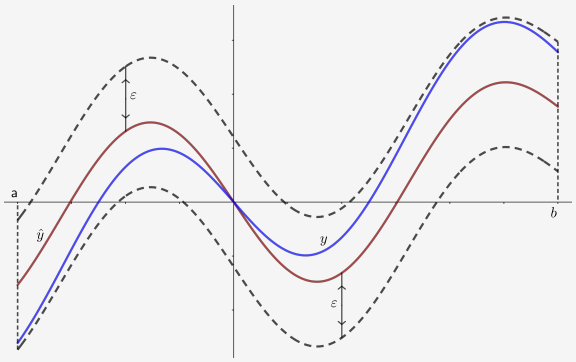
Na prostoru spojitých funkcí $\mathrm{C}[a,b]$ můžeme zavést *maximální normu*

$$\|y\|_0 := \max_{x\in[a,b]} |y(x)|,$$

takže funkce y leží ve vzdálenosti ε od funkce $\widehat{y}$, jestliže

$$|y(x) - \widehat{y}(x)| < \varepsilon \quad \text{pro všechna } x \in [a,b],$$

tj. „funkční hodnoty na $[a,b]$ leží dostatečně blízko“, viz Obrázek 9.1.



Obrázek 9.1: Funkce $\widehat{y}$ a její ε-okolí zahrnující i funkci y.

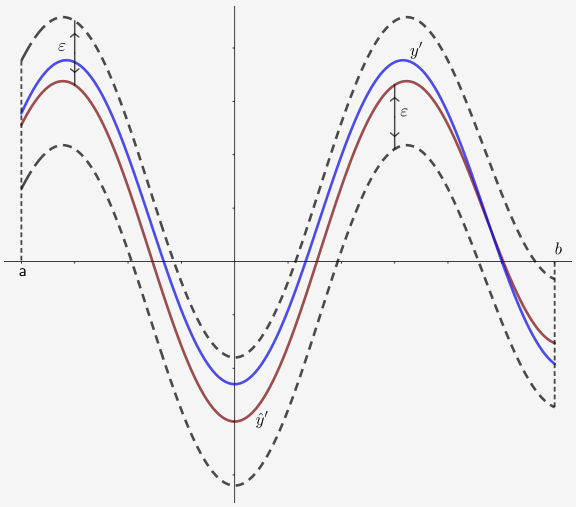
Na prostoru spojitě diferencovatelných funkcí $\mathrm{C}^1[a,b]$ můžeme také zavést tuto normu, avšak mnohem výhodnější je využít v normě i derivaci funkce y, tj.

$$\|y\|_1 := \max_{x\in[a,b]} |y(x)| + \max_{x\in[a,b]} |y'(x)|.$$

Pak funkce y leží ve vzdálenosti ε od $\widehat{y}$, jestliže

$$|y(x) - \widehat{y}(x)| < \varepsilon/2 \quad \text{a} \quad |y'(x) - \widehat{y}'(x)| < \varepsilon/2 \quad \text{pro všechna } x \in [a,b],$$

tj. „funkční hodnoty a hodnoty derivace na $[a,b]$ leží dostatečně blízko“[4], viz také Obrázek 9.2.



Obrázek 9.2: Funkce $\tilde{y}'$ a její ε-okolí zahrnující i derivaci funkce y.

Např. pro vzdálenosti funkcí $y_1(x) = x$ a $y_2(x) = \sin x$ na intervalu $[0, \pi/4]$ platí[5]

$$\| y_1 - y_2 \|_0 = \max_{x \in [0,\pi/4]} |x - \sin x| = \pi/4 - \sqrt{2}/2,$$

zatímco derivací dostaneme vzdálenost

$$\| y_1 - y_2 \|_1 = \max_{x \in [0,\pi/4]} |x - \sin x| + \max_{x \in [0,\pi/4]} |1 - \cos x| = \pi/4 + 1 - \sqrt{2}/2.$$

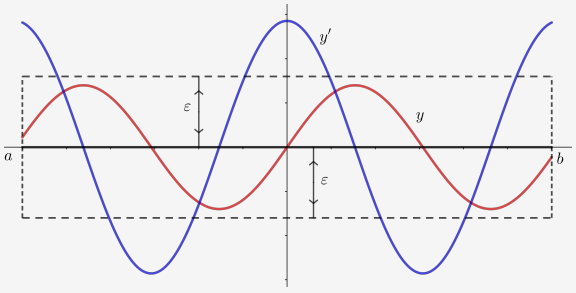
Avšak situace může být i výrazně odlišná. Uvažme např. funkci

$$y(x) = c \sin(x/c^2),$$

na intervalu $[a, b]$ jistě patří do ε-okolí nulové funkce $y \equiv 0$ pro libovolné $0 < c < \varepsilon$ vzhledem k normě $\|\cdot\|_0$, neboť $\| y - 0 \|_0 = \max_{x \in [a,b]} |y(x)| \leq c$. Avšak již nepatří do ε-okolí nulové funkce $y \equiv 0$ určeného normou $\|\cdot\|_1$, neboť z rovnosti $y'(x) = \frac{1}{c} \cos(x/c^2)$ plyne

$$\| y - 0 \|_1 = \max_{x \in [a,b]} |y(x)| + \max_{x \in [a,b]} |y'(x)| \leq c + 1/c.$$

Zejména je-li $k\pi c^2 \in [a, b]$ a $(\pi/2 + m\pi)\, c^2 \in [a, b]$ pro nějaké $k, m \in \mathbb{Z}$, pak dokonce $\| y - 0 \|_1 = c + 1/c$. Má-li ale být splněno $c + 1/c < \varepsilon$ neboli $c^2 - c\varepsilon + 1 < 0$, pak nutně $\varepsilon^2 - 4 > 0$, tj.[6] $\varepsilon > 2$.



Obrázek 9.3: Okolí nulové funkce vs. y a y' pro $c = 0{,}7$ a $\varepsilon = 0{,}8$.

Funkce y ukazuje, že okolí určené normou $\|\cdot\|_0$ obsahuje více funkcí než v případě normy $\|\cdot\|_1$, tj. $\mathcal{O}_{\|\cdot\|_0}(y) \supseteq \mathcal{O}_{\|\cdot\|_1}(y)$. Budeme-li pak chtít ověřit, zda $\hat{y}$ je lokálním extrémem, budeme mít pro $\|\cdot\|_0$ více funkcí k testování neměnnosti znaménka výrazu $\mathcal{J}[y] - \mathcal{J}[\hat{y}]$, a tudíž i více možností v nichž může být tato podmínka porušena. Proto

[4] V literatuře se lze setkat i s definicí okolí pomocí hodnot funkce a její derivace nikoli jako součtu obou maxim, ale požadavku aby obě maxima byla menší než ε, jak jsme to před chvíli částečně učinili my. Snad není příliš těžké se přesvědčit o tom, že v kontextu lokálních extrémů jsou tyto přístupy ekvivalentní.

[5] Na intervalu $(0, \pi/4]$ máme $(x - \sin x)' = 1 - \cos x > 0$, tj. funkce $x - \sin x$ je na daném intervalu rostoucí. Podobně $(1 - \cos x)' = \sin x > 0$ pro $x \in (0, \pi/4]$.

[6] Však také $\min_{c>0}\{c + 1/c\} = 2$.

se v této souvislosti hovoří o $\|\cdot\|_0$ jako o *silné* normě a odpovídajícím $\widehat{y}$ jako o *silném* lokálním extrému, zatímco v případě $\|\cdot\|_1$ se hovoří o *slabé* normě a *slabém* lokálním extrému. Jaká je motivace pro právě tuto terminologii[7]? Odpovědí je vztah mezi silnými a slabými lokálními extrémy. Je-li $\widehat{y}$ silným lokálním extrémem, pak je nutně i slabým lokálním extrémem díky výše popsané inkluzi okolí $\widehat{y}$, tj. pokud $\mathcal{J}[y]-\mathcal{J}[\widehat{y}]$ nemění znaménko pro všechna $y \in \mathrm{C}^1[a,b]$ splňující $\|y-\widehat{y}\|_0 < \varepsilon$, pak totéž jistě platí i pro funkce vyhovující dokonce nerovnosti $\|y-\widehat{y}\|_1 < \varepsilon$. Opačná implikace už samozřejmě neplatí, ovšem kontrapozicí alespoň dostaneme, že v případě absence slabého extrému nebude existovat ani silný extrém. Proto každá postačující podmínka pro silný extrém je i postačující podmínkou pro slabý extrém, zatímco každá nutná podmínka pro slabý extrém je nutnou podmínkou i pro silný extrém. Vzhledem k výše uvedenému snad nepřekvapí, že nalezení slabých extrémů je jednodušší a že se právě na ně zaměříme[8].

Pro úplnost ještě dodejme, že analogicky můžeme pro funkce $y \in \mathrm{C}^n[a,b]$ definovat *maximální normu k-tého řádu* pro $k \in \{0,\dots,n\}$ jako

$$\|y\|_k := \sum_{j=0}^{k} \max_{x\in[a,b]} |y^{(j)}(x)|.$$

[7] Méně vhodných kandidátů ≈ „slabá terminologie“.

[8] Také Lagrange se ve své práci věnoval právě slabým extrémům. Počátky studia silných extrémů jsou spojeny až s prací Weierstrasse.

Bez ohledu na konkrétní volbu prostoru $\mathcal{V}$ a jeho normy musí každý lokální extrém nutně splňovat následující podmínku.

VĚTA 9.2.5

Nechť funkcionál $\mathcal{J}[\cdot]$ je diferencovatelný a má na $\mathcal{N}$ lokální extrém v $\widehat{y}$. Pak $\mathcal{J}'[\widehat{y};\eta]=0$ pro všechny přípustné variace $\eta \in \mathcal{N}_0$.

Důkaz. Nechť pro jednoduchost $\mathcal{J}[\cdot]$ má na $\mathcal{N}$ lokální minimum v $\widehat{y}$. Pak z Definic 9.2.1 a 9.2.3 vyplývá, že

$$0 \le \mathcal{J}[\widehat{y}+\eta]-\mathcal{J}[\widehat{y}] = \Delta\mathcal{J}_{\widehat{y}}[\eta] = \mathcal{J}'[\widehat{y};\eta]+\omega(\eta)\|\eta\| \tag{9.2.4}$$

pro všechny přípustné variace $\eta \in \mathcal{N}_0$ vyhovující podmínce $\|\eta\| < \varepsilon$ pro nějaké $\varepsilon > 0$. Připusťme, že $\mathcal{J}'[\widehat{y},\eta_0] \neq 0$ pro nějaké $\eta_0 \in \mathcal{N}_0 \setminus \{0\}$. Pak pro libovolné $\delta \in (-\varepsilon/\|\eta_0\|, \varepsilon/\|\eta_0\|)$ bude $\delta\eta_0 \in \mathcal{N}_0$ splňovat $\|\delta\eta_0\| < \varepsilon$ a z rovnosti (9.2.4) díky linearitě $\varphi(\cdot) = \mathcal{J}'[\widehat{y};\cdot]$ dostaneme

$$0 \le \Delta\mathcal{J}_{\widehat{y}}[\delta\eta_0] = \delta\varphi(\eta_0)+\omega(\delta\eta_0)|\delta|\|\eta_0\| \quad \text{neboli} \quad 0 \le \operatorname{sgn}(\delta)\varphi(\eta_0)+\omega(\delta\eta_0)\|\eta_0\|,$$

kde výraz $\operatorname{sgn}(\delta)\varphi(\eta_0)$ je konstantní, zatímco hodnota $\omega(\delta\eta_0)\|\eta_0\|$ se pro $\delta \to 0$ blíží k nule. Pak ale jistě lze nalézt δ dostatečně malé a s opačným znaménkem než $\varphi(\eta_0)$ tak, že poslední nerovnost nebude splněna. ↯ ■

Podobně jako jsme v případě funkcí $f:\mathbb{R}^n \to \mathbb{R}$ pojmenovali body splňující nutnou podmínku prvního řádu *stacionárními*, bude i nyní velmi užitečné zavést si pro funkce splňující nutnou podmínku z Věty 9.2.5 „zvláštní“ terminologii.

DEFINICE 9.2.6

Funkce $y \in \mathcal{N}$ splňující $\mathcal{J}'[y;\eta]=0$ pro všechny přípustné variace $\eta \in \mathcal{N}_0$ se nazývá *extremálou*.

Jak ale ověřit, zda nalezená extremála je lokálním extrémem? K tomu budeme potřebovat pojem *kvadratického funkcionálu*. Zobrazení $\varphi : \mathcal{N}_0 \times \mathcal{N}_0 \to \mathbb{R}$, které je lineární v obou složkách, se nazývá *bilineární funkcionál*, tj. $\varphi(\eta,\zeta)$ pro pevné $\zeta \in \mathcal{N}_0$ a libovolné $c_1,c_2 \in \mathbb{R}$, $\eta_1,\eta_2 \in \mathcal{N}_0$ splňuje

$$\varphi(c_1\eta_1+c_2\eta_2,\zeta) = c_1\varphi(\eta_1,\zeta)+c_2\varphi(\eta_2,\zeta)$$

a podobně pro pevné $\eta \in \mathcal{N}_0$ a libovolné $c_1, c_2 \in \mathbb{R}$, $\zeta_1, \zeta_2 \in \mathcal{N}_0$ platí

$$\varphi(\eta, c_1\zeta_1 + c_2\zeta_2) = c_1\varphi(\eta, \zeta_1) + c_2\varphi(\eta, \zeta_2).$$

Kvadratický funkcionál pak získáme volbou $\eta = \zeta$, tj. jako $\varphi(\eta, \eta)$. S tímto jsme se již setkali v konečně dimenzionálním případě (matice), kdy jsme hovořili o *kvadratických formách*.

Je-li navíc k (9.2.3) možné získat dokonce vyjádření

$$\Delta\mathcal{J}[y;\eta] = \varphi_1(\eta) + \varphi_2(\eta) + \omega(\eta)\,\|\eta\|^2$$

pro nějaký lineární funkcionál $\varphi_1(\eta)$, kvadratický funkcionál $\varphi_2(\eta)$ a nelineární funkcionál $\omega(\cdot)$ splňující $\omega(\eta) \to 0$ pro $\|\eta\| \to 0$, pak o funkcionálu $\mathcal{J}[\cdot]$ řekneme, že je *dvakrát diferencovatelný*, a kvadratický funkcionál $\varphi_2(\eta)$ nazveme jeho *druhou variací* v y a označíme ji $\mathcal{J}''[y;\eta]$ příp. $\delta^2\mathcal{J}[y]$. Také $\mathcal{J}''[y;\eta]$ je určena jednoznačně a s její pomocí získáme nutnou i postačující podmínku pro lokální extrém. Je-li totiž $\widehat{y}$ lokální minimum funkcionálu $\mathcal{J}[\cdot]$, pak nutně

$$\mathcal{J}''[\widehat{y};\eta] \geq 0 \quad \text{pro všechna } \eta \in \mathcal{N}_0.$$

[9] To znamená, že existuje $k > 0$ takové, že $\mathcal{J}''[\widehat{y};\eta] \geq k\|\eta\|^2$ pro všechny přípustné variace $\eta \in \mathcal{N}_0$.

Na druhou stranu, je-li $\widehat{y}$ extremála a $\mathcal{J}''[\widehat{y};\eta]$ je pozitivně definitní[9], pak $\widehat{y}$ je lokální minimum. Pochopitelně analogické podmínky platí pro lokální maximum.

9.3 Základní úloha variačního počtu

Nyní s využitím předchozích výsledků odvodíme nutné a postačující podmínky pro řešení úlohy o brachistochroně a dalších podobných úloh, ve kterých budeme mezi všemi funkcemi $y \in \mathrm{C}^1[a,b]$ vyhovujícím podmínkám

$$y(a) = \alpha \qquad y(b) = \beta \tag{9.3.1}$$

hledat takovou, že

$$\mathrm{J}[y] := \int_a^b F\big(x, y(x), y'(x)\big)\,\mathrm{d}x \to \mathrm{opt} \tag{9.3.2}$$

pro danou spojitou funkci $F : \mathbb{R}^3 \to \mathbb{R}$ se spojitými parciálními derivacemi až do druhého řádu vzhledem ke všem třem proměnným. Funkce F se nazývá lagrangián a právě tomuto typu úloh se Lagrange ve své práci věnoval. Naše řešení má proto mnoho společného s Lagrangeovým čistě analytickým přístupem nevyžadujícím žádné geometrické nástroje, který poprvé představil Eulerovi v dopise z 12. srpna 1755. V tomto dopise ilustroval sílu své metody na řešení třech úloh z Eulerovy knihy [159] a následně s její pomocí v dopise z 20. listopadu 1755 vyřešil i úlohu o brachistochroně s tzv. *volným koncem*, viz [185] a případně korespondenci mezi Eulerem a Lagrangem dostupnou na adrese https://ln.math.muni.cz/pz/opt/EulerLagrange. Toto je sice z dnešního pohledu značné zjednodušení, neboť máme velmi specifické počáteční podmínky, řád derivací i tvar lagrangiánu, ale v praxi se řada fyzikálních nebo geometrických úloh redukuje právě do této podoby. Připomeňme, že řešení úlohy (9.3.1)–(9.3.2) závisí na volbě normy v prostoru $\mathrm{C}^1[a,b]$. Tomuto problému jsme se podrobně věnovali již v Příkladu 9.2.4 a my se v našich následujících úvahách omezíme na hledání slabých extrémů.

Abychom mohli využít tvrzení Věty 9.2.5 budeme potřebovat pár pomocných výsledků týkajících se „nulovosti některých integrálů“.

Lemma 9.3.1

Nechť $\alpha \in \mathrm{C}[a,b]$. Pokud $\int_a^b \alpha(x)\,\eta(x)\,\mathrm{d}x = 0$ pro všechna $\eta \in \mathrm{C}[a,b]$ splňující $\eta(a) = 0 = \eta(b)$, pak $\alpha(x) = 0$ pro všechna $x \in [a,b]$.

Důkaz. Připusťme, že $\alpha \not\equiv 0$ na $[a,b]$, tj. existuje $x_0 \in [a,b]$ takové, že $\alpha(x_0) \neq 0$. V takovém případě ale ze spojitosti funkce $\alpha(\cdot)$ vyplývá, že musí být $\alpha(x) \neq 0$ na nějakém netriviálním podintervalu $[x_1,x_2] \subseteq [a,b]$. Bez újmy na obecnosti vezměme $\alpha(x) > 0$ na $[x_1,x_2]$. Položíme-li

$$\eta(x) := \begin{cases} (x-x_1)(x_2-x), & x \in [x_1,x_2], \\ 0, & \text{jinak}, \end{cases}$$

pak jistě $\eta \in \mathrm{C}[a,b]$ a $\eta(x_1) = 0 = \eta(x_2)$, takže podle předpokladů by měl být $\int_a^b \alpha(x)\eta(x)\,\mathrm{d}x = 0$. Jenže $\alpha(x) > 0$ a $\eta(x) > 0$ na (x_1,x_2), což znamená, že dostáváme spor, neboť

$$\int_a^b \alpha(x)\eta(x)\,\mathrm{d}x = \int_{x_1}^{x_2} \alpha(x)\eta(x)\,\mathrm{d}x > 0. \lightning$$

Proto nutně $\alpha \equiv 0$ na $[a,b]$. ∎

Tvrzení předchozího lemma zůstane v platnosti i v případě, že uvažujeme $\eta \in \mathrm{C}^n[a,b]$. V takovém případě ale nelze v důkazu vzít totéž $\eta(\cdot)$, protože $\eta'(x) = -2x + x_1 + x_2 \neq 0$ pro $x \in \{x_1,x_2\}$ implikuje $\eta \notin \mathrm{C}^1[a,b]$. Místo toho je potřeba zvolit

$$\eta(x) := \begin{cases} (x-x_1)^{n+1}(x_2-x)^{n+1}, & x \in [x_1,x_2], \\ 0, & \text{jinak}, \end{cases}$$

což již vyhovuje požadavku $\eta \in \mathrm{C}^n[a,b]$.

Následující tvrzení bývá v literatuře označováno jako *fundamentální lemma variačního počtu.* Avšak pro nás bude mnohem důležitější jeho důsledek.

LEMMA 9.3.2

Nechť $\alpha \in \mathrm{C}[a,b]$. Pokud $\int_a^b \alpha(x)\eta'(x)\,\mathrm{d}x = 0$ pro všechna $\eta \in \mathrm{C}^1[a,b]$ splňující $\eta(a) = 0 = \eta(b)$, pak $\alpha \equiv c$ na $[a,b]$ pro vhodné $c \in \mathbb{R}$.

Důkaz. Nechť c je střední hodnota funkce $\alpha(\cdot)$ na $[a,b]$, tj. $c := \frac{1}{b-a}\int_a^b \alpha(x)\,\mathrm{d}x$ a vzhledem ke spojitosti $\alpha(\cdot)$ dokonce máme $c = \alpha(x_0)$ pro nějaké $x_0 \in [a,b]$, viz [144, Důsledek 2.36]. Položíme-li $\eta(x) := \int_a^x [\alpha(t)-c]\,\mathrm{d}t$, pak ze spojitosti $\alpha(x)-c$ plyne $\eta \in \mathrm{C}^1[a,b]$, viz [144, Věta 2.39]. Navíc zjevně máme $\eta(a) = 0$ a také

$$\eta(b) = \int_a^b [\alpha(x)-c]\,\mathrm{d}x = \int_a^b \alpha(x)\,\mathrm{d}x - c(b-a) = 0.$$

Odtud s využitím předpokladů lemma získáme přímým výpočtem

$$\int_a^b [\alpha(x)-c]^2\,\mathrm{d}x = \int_a^b [\alpha(x)-c]\,\eta'(x)\,\mathrm{d}x = \underbrace{\int_a^b \alpha(x)\eta'(x)\,\mathrm{d}x}_{=0} - c\int_a^b \eta'(x)\,\mathrm{d}x = -c\big[\eta(x)\big]_a^b = 0,$$

jenže $[\alpha(x)-c]^2$ je spojitá a nezáporná funkce na $[a,b]$, tudíž nutně $[\alpha(x)-c]^2 \equiv 0$ na $[a,b]$ neboli $\alpha(x) \equiv c$. ∎

DŮSLEDEK 9.3.3

Nechť $\alpha,\beta \in \mathrm{C}[a,b]$. Pokud

$$\int_a^b \big[\alpha(x)\eta(x) + \beta(x)\eta'(x)\big]\,\mathrm{d}x = 0$$

pro všechna $\eta \in \mathrm{C}^1[a,b]$ splňující $\eta(a) = 0 = \eta(b)$, pak funkce β je diferencovatelná (dokonce $\beta \in \mathrm{C}^1[a,b]$) a platí $\beta'(x) = \alpha(x)$ pro všechna $x \in [a,b]$.

Důkaz. Nechť $A(x)$ je primitivní funkce k $\alpha(x)$, tj.

$$A(x) = \int_a^x \alpha(t)\,\mathrm{d}t, \quad x \in [a,b],$$

viz [144, Důsledek 2.40]. Pak s využitím integrace per-partes získáme

$$\int_a^b \alpha(x)\,\eta(x)\,\mathrm{d}x \left| \begin{matrix} u = \eta(x) & u' = \eta'(x) \\ v = A(x) & v' = \alpha(x) \end{matrix} \right| = \underbrace{\big[\tau(x)\,A(x)\big]_a^b}_{=0} - \int_a^b A(x)\,\eta'(x)\,\mathrm{d}x = -\int_a^b A(x)\,\eta'(x)\,\mathrm{d}x,$$

díky čemuž máme

$$0 = \int_a^b \big[\alpha(x)\,\eta(x) + \beta(x)\,\eta'(x)\big]\,\mathrm{d}x = \int_a^b \big[-A(x) + \beta(x)\big]\,\eta'(x)\,\mathrm{d}x.$$

Jelikož poslední rovnost platí pro libovolnou funkci η vyhovující daným předpokladům, plyne z Lemma 9.3.2, že $-A(x) + \beta(x) \equiv c$ na $[a,b]$. Tudíž $\beta \in \mathrm{C}^1[a,b]$ a $\beta'(x) = A'(x) = \alpha(x)$ pro všechna $x \in [a,b]$. ■

Všechna tři uvedená tvrzení jsou platná dokonce i jako ekvivalence, přičemž opačný se snadno ověří přímým výpočtem. Zejména v Lemma 9.3.2 využijeme rovnosti $\int_a^b \eta'(x) = \big[\eta(x)\big]_a^b$ a v Důsledku 9.3.3 pro volbou $\alpha = \beta'$ dostaneme

$$\int_a^b \big[\beta'(x)\,\eta(x) + \beta(x)\,\eta'(x)\big]\,\mathrm{d}x = \int_a^b \big[\beta(x)\,\eta(x)\big]'\,\mathrm{d}x = \big[\beta(x)\,\eta(x)\big]_a^b.$$

Nyní určíme první variaci funkcionálu z úlohy (9.3.2). Připomeňme, že $\mathscr{V} = \mathrm{C}^1[a,b]$, $\mathscr{N} = \{y \in \mathscr{V} \mid y(a) = \alpha, y(b) = \beta\}$ a $\mathscr{N}_0 = \{\eta \in \mathscr{V} \mid \eta(a) = 0 = \eta(b)\}$. Pak

$$\Delta \mathrm{J}[y;\eta] = \mathrm{J}[y+\eta] - \mathrm{J}[y] = \int_a^b \big[F(x, y+\eta, y'+\eta') - F(x,y,y')\big]\,\mathrm{d}x,$$

na což můžeme díky spojitosti parciálních derivací aplikovat Taylorovu větu, čímž obdržíme[10]

$$\begin{aligned}\Delta \mathscr{J}[y;\eta] = \int_a^b \big[&\underbrace{F_x(x,y,y')\,(x-x)}_{=0} + F_y(x,y,y')\,\underbrace{(y+\eta-y)}_{=\eta} \\ &+ F_z(x,y,y')\,\underbrace{(y'+\eta'-y')}_{=\eta'} + R(\eta^2,\eta'^2)\big]\,\mathrm{d}x \\ = \int_a^b &\big[F_y(x,y,y')\,\eta + F_z(x,y,y')\,\eta'\big]\,\mathrm{d}x + \int_a^b R(\eta^2,\eta'^2)\,\mathrm{d}x.\end{aligned}$$

[10] Abychom předešli případným nejasnostem, měli bychom zdůraznit, že výrazem $F_z(x,y,y')$ rozumíme parciální derivaci funkce $F(\cdot,\cdot,\cdot)$ vzhledem ke třetí proměnné vyčíslenou v „bodě“ $[x,y,y']$.

Toto vyjádření odpovídá rovnosti (9.2.3), přičemž první integrál na pravé straně je lineární vzhledem k η, zatímco druhý integrál obsahuje kvadratické a smíšené členy. Proto variace daného funkcionálu je rovna

$$\mathrm{J}'[y;\eta] = \int_a^b \big[F_y(x,y,y')\,\eta + F_z(x,y,y')\,\eta'\big]\,\mathrm{d}x.$$

Má-li tedy $\widehat{y} \in \mathrm{C}^1[a,b]$ být lokálním řešením úlohy (9.3.1)–(9.3.2) musí dle Věty 9.2.5 platit

$$0 = \mathrm{J}'[\widehat{y};\eta] = \int_a^b \big[F_y(x,\widehat{y},\widehat{y}')\,\eta + F_z(x,\widehat{y},\widehat{y}')\,\eta'\big]\,\mathrm{d}x \tag{9.3.3}$$

pro všechna $\eta \in \mathscr{N}_0$. Volbou $\alpha(x) = F_y(x,\widehat{y},\widehat{y}')$ a $\beta(x) = F_z(x,\widehat{y},\widehat{y}')$ v Důsledku 9.3.3 pak vyplývá, že rovnost (9.3.3) je splněna právě tehdy, když β je diferencovatelná a $\beta'(x) = \alpha(x)$ pro všechna $x \in [a,b]$, tj. funkce $\widehat{y}$ musí být řešením *Eulerovy–Lagrangeovy rovnice*

$$\frac{\mathrm{d}}{\mathrm{d}x} F_z(x,\widehat{y},\widehat{y}') = F_y(x,\widehat{y},\widehat{y}') \quad \text{na } [a,b]. \tag{9.3.4}$$

Řešením této rovnice obdržíme všechny (silné i slabé) extremály úlohy (9.3.1)–(9.3.2). Obecně se jedná o nelineární diferenciální rovnici 2. řádu. Podaří-li se nám najít její řešení, pak bude

záviset na dvou konstantách, k jejichž určení využijeme okrajové podmínky. Eulerova–Lagrangeova rovnice tak převádí úlohu variačního počtu (9.3.1)–(9.3.1) na úlohu nalezení partikulárního řešení příslušné diferenciální rovnice (případně zjištění, zda vůbec takové řešení existuje). To ale v obecném případě může být dosti obtížné. V úloze (9.3.1)–(9.3.1) není nic vyžadováno pro $\widehat{y}''$, ovšem pokud budeme mít zaručenu jeho existenci, můžeme levou stranu rovnice (9.3.4) rozpočítat pomocí derivace složené funkce, čímž získáme explicitní tvar zmíněné nelineární diferenciální rovnice 2. řádu

$$F_y(x,\widehat{y},\widehat{y}') - F_{zx}(x,\widehat{y},\widehat{y}') - F_{zy}(x,\widehat{y},\widehat{y}')\,\widehat{y}' - F_{zz}(x,\widehat{y},\widehat{y}')\,\widehat{y}'' = 0. \tag{9.3.5}$$

Ať už je to s existencí $\widehat{y}''$ jakkoli, skýtají některé speciální typy funkce F výrazné zjednodušení Eulerovy–Lagrangeovy rovnice:

(i) Pokud F explicitně nezávisí na y, tj.

$$\mathrm{J}[y] = \int_a^b \widehat{F}(x,y')\,\mathrm{d}x,$$

pak $F_y(x,\widehat{y},\widehat{y}') \equiv 0$ a Eulerova–Lagrangeova je tvaru

$$\frac{\mathrm{d}}{\mathrm{d}x} F_z(x,\widehat{y},\widehat{y}') = 0 \quad \text{na } [a,b],$$

tj. $F_z(x,\widehat{y},\widehat{y}') \equiv \text{konst.}$, což je diferenciální rovnice 1. řádu.

(ii) Pokud F explicitně nezávisí na y', tj.

$$\mathrm{J}[y] = \int_a^b \widehat{F}(x,y)\,\mathrm{d}x,$$

pak $F_y(x,\widehat{y},\widehat{y}') \equiv 0$ a a Eulerova–Lagrangeova je tvaru

$$F_y(x,\widehat{y},\widehat{y}') = 0 \quad \text{na } [a,b],$$

neboli $F(x,\widehat{y},\widehat{y}') \equiv \text{konst.}$, což je dokonce „pouze“ algebraická rovnice.

(iii) Pokud F explicitně nezávisí na x, tj.

$$\mathrm{J}[y] = \int_a^b \widehat{F}(y,y')\,\mathrm{d}x,$$

pak lze s využitím podobných úprav jako v (9.3.5) přepsat Eulerovu–Lagrangeovu rovnici do tvaru

$$\begin{aligned} 0 &= y'\left[F_y(x,\widehat{y},\widehat{y}') - \frac{\mathrm{d}}{\mathrm{d}x}F_z(x,\widehat{y},\widehat{y}')\right] = \\ &= y'\left[F_y(x,\widehat{y},\widehat{y}') - \widehat{y}'\,F_{zy}(x,\widehat{y},\widehat{y}') - \widehat{y}''\,F_{zz}(x,\widehat{y},\widehat{y}')\right] + \widehat{y}''\,F_z(x,\widehat{y},\widehat{y}') - \widehat{y}''\,F_z(x,\widehat{y},\widehat{y}') = \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[F(x,\widehat{y},\widehat{y}') - \widehat{y}'\,F_z(x,\widehat{y},\widehat{y}')\right] \end{aligned}$$

z čehož ihned plyne tzv. *Beltramiho rovnost*

$$F(x,\widehat{y},\widehat{y}') - \widehat{y}'\,F_z(x,\widehat{y},\widehat{y}') \equiv \text{konst.} \quad \text{na } [a,b].$$

(iv) Je-li F lineární vzhledem k y, tj. $F(x,y,y') = A(x,y)\,y' + B(x,y)$, pak Eulerova–Lagrangeova rovnice má podobu

$$A_x(x,\widehat{y}) + A_y(x,\widehat{y})\,\widehat{y}' = A_y(x,\widehat{y})\,y' + B_y(x,\widehat{y}) \quad \text{neboli} \quad A_x(x,\widehat{y}) = B_y(x,\widehat{y}).$$

To vlastně není diferenciální rovnice, ale spíše implicitní vyjádření, které určuje funkci y proměnné x.

Příklad 9.3.4

Určeme extremály variačního problému

$$\mathcal{J}[y] = \int_1^2 y'(1+x^2\,y')\,\mathrm{d}x \to \mathrm{opt}, \quad y \in \mathrm{C}^1[1,2], \quad y(1)=4, \quad y(2)=3.$$

Řešení. Máme $F(x,y,y') = \widehat{F}(x,y') = y'(1+x^2\,y')$, tudíž $F_y(x,y,y') = 0$ a $F_z(x,y,y') = 1 + x^2\,y' + y'\,x^2$. Využijeme-li přímo Eulerovu–Lagrangeovu rovnici, musí každá extremála $y \in \mathrm{C}^1[1,2]$ dané úlohy vyhovovat diferenciální rovnici

$$\frac{\mathrm{d}}{\mathrm{d}x}F_z(x,y,y') = 4x\,y' + 2x^2\,y'' = 0 = F_y(x,y,y') \quad \text{neboli} \quad x^2\,y'' + 2x\,y' = 0.$$

To je tzv. *Eulerova diferenciální rovnice*, k jejímuž řešení využijeme substituci $t = \ln|x|$. Jelikož $x \in [1,2]$, je $x = \mathrm{e}^t$ a platí

$$y' = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = \dot{y}\,\frac{1}{x} \quad \text{a} \quad y'' = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\dot{y}\,\frac{1}{x}\Big) = \frac{1}{x}\,\frac{\dot{y}}{\mathrm{d}x} + \dot{y}\,\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{x}\Big) = \frac{1}{x^2}\,(\ddot{y} - \dot{y}).$$

Dosazením do Eulerova diferenciální rovnice dostaneme lineární diferenciální rovnici s konstantními koeficienty

$$0 = x^2\,y'' + 2x\,y' = \ddot{y} + \dot{y},$$

jejíž charakteristická rovnice $\lambda^2 + \lambda = 0$ má kořeny $\lambda = 0$ a $\lambda = -1$, takže její obecné řešení je

$$y(t) = c_1\,\mathrm{e}^{-t} + c_2, \quad c_1, c_2 \in \mathbb{R}.$$

Zpětným dosazením a využitím daných okrajových podmínek obdržíme

$$y(x) = \frac{c_1}{x} + c_2, \quad 4 = y(1) = c_1 + c_2 \quad \& \quad 3 = y(2) = c_1/2 + c_2,$$

čemuž vyhovuje pouze funkce

$$\widehat{y}(x) = \frac{2}{x} + 2.$$

Z předcházející poznámky víme, že funkcionál v dané úloze nabízí výrazné zjednodušení Eulerovy–Lagrangeovy rovnice, které jsme ale v tuto chvíli nevyužili. Kdybychom tak učinili, hledali bychom pouze řešení diferenciální rovnice 1. řádu $F_z(x,y,y') = c$ neboli

$$1 + 2x^2\,y' = c, \quad \text{tj.} \quad y' = \frac{c-1}{x^2},$$

což přímým integrováním vede k funkci

$$y(x) = -\frac{c-1}{2x} + d,$$

z níž opět s využitím daných okrajových podmínek získáme hledané řešení odpovídající volbě $c = -3$ a $d = 2$.

Pro daný funkcionál z (9.3.2) je jeho druhá variace

$$\mathrm{J}''[y;\eta] = \int_a^b \big[P(x)\,\eta'^2 + Q(x)\,\eta^2\big]\,\mathrm{d}x$$

pro $P(x) := \frac{1}{2}F_{zz}(x,y,y')$ a $Q(x) := \frac{1}{2}[F_{yy}(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x}F_{yz}(x,y,y')]$. Potom nutná podmínka pro to, aby extremála $\widehat{y}$ byla slabým lokálním minimem je tzv. *Legendreova podmínka*

$$F_{zz}(y,\widehat{y},\widehat{y}') \geq 0 \quad \text{pro všechna } x \in [a,b].$$

V případě ostré nerovnosti hovoříme o *zesílené Legendreově podmínce*. Legendre se (neúspěšně) snažil ukázat, že tato zesílená podmínka je postačující pro slabé lokální minimum. Jenže to není pravda a je potřeba k ní ještě jeden požadavek doplnit.

VĚTA 9.3.5

Nechť $\widehat{y} \in \mathcal{N}$ je přípustná funkce taková, že

(i) je splněna Eulerova–Lagrangeova rovnice (9.3.4),

(ii) $F_{zz}(x, \widehat{y}, \widehat{y}') = 2P(x) > 0$ pro všechna $x \in [a, b]$,

(iii) existuje řešení *Jacobiho diferenciální rovnice*

$$Q(x)\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left[P(x)\eta'\right] = 0$$

vyhovující počátečním podmínkám $\eta(a) = 0$ a $\eta'(a) = 1$, pro které je $x = a$ jediným nulovým bodem, tj. $\eta(x) \neq 0$ pro $x \in (a, b]$.

Potom $\mathcal{J}''[\widehat{y}; \eta] > 0$ pro všechny přípustné variace $\eta \in \mathcal{N}_0$, tj. $\widehat{y}$ je slabým lokálním minimem funkcionálu $\mathcal{J}[\cdot]$ neboli slabým lokálním řešením minimalizační úlohy (9.3.1)–(9.3.2).

Poslední podmínku v předchozím tvrzení lze nahradit existencí řešení tzv. *Riccatiho rovnice*

$$w^2 = P(x)\,[Q(x) + w'] \quad \text{nebo} \quad w' = F_{yy}(x, \widehat{y}, \widehat{y}') - \frac{\left[F_{yz}(x, \widehat{y}, \widehat{y}') - w\right]^2}{F_{zz}(x, \widehat{y}, \widehat{y}')}$$

definovaného na celém intervalu $[a, b]$.

Příklad 9.3.6

Ověřme, že extremála nalezená v Příkladě 9.3.4 je slabým lokálním extrémem a určete jeho charakter.

Řešení. Jedinou nalezenou extremálou je funkce $\widehat{y} = 2/x + 2$. Pak

$$F_{zz}(x, \widehat{y}, \widehat{y}') = 2x^2 > 0, \quad x \in [1, 2],$$

tj. je splněna zesílená Legendreova podmínka. Jelikož $P(x) = F_{zz}(x, \widehat{y}, \widehat{y}')/2 = x^2$ a $Q(x) = \frac{1}{2}[F_{yy}(x, \widehat{y}, \widehat{y}') - \frac{\mathrm{d}}{\mathrm{d}x}F_{yz}(x, \widehat{y}, \widehat{y}')] = 0$, máme Jacobiho rovnici

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\eta'\right) = 0 \quad \text{neboli} \quad x^2\eta' = c,$$

jejímž obecným řešením je $\eta(x) = -c/x + d$. S využitím počátečních podmínek pak nalezneme její řešení $\eta(x) = -1/x + 1$, které jistě na intervalu $(1, 2]$ nemá žádný nulový bod. Poněvadž $F_{zz}(x, \widehat{y}, \widehat{y}') > 0$ jedná se o slabé lokální minimum.

Kdybychom místo Jacobiho rovnice chtěli využít Riccatiho rovnici dostaneme

$$w^2 = x^2\,w' \quad \text{neboli} \quad \frac{w'}{w^2} = x^2.$$

Jejím obecným řešením je funkce $w(x) = x/(1 - cx)$, která je pro libovolnou volbu integrační konstanty $c \in (-\infty, 1/2) \cup (1, \infty)$ definovaná na celém intervalu $[1, 2]$.

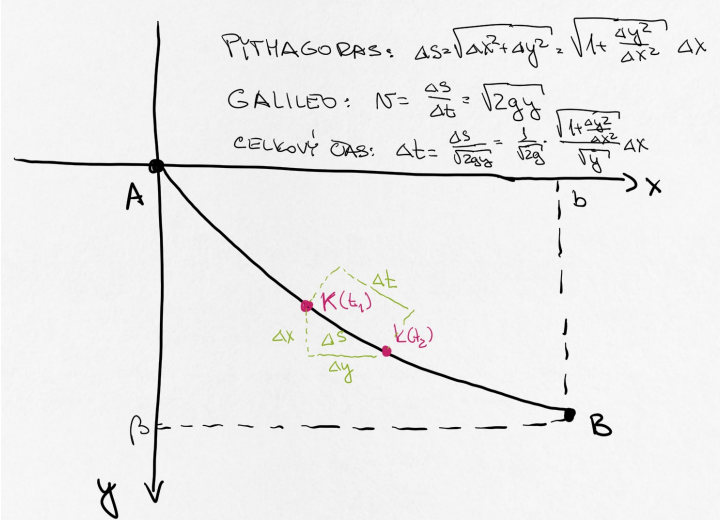
9.4 Úloha o brachistochroně

Nyní už je na čase pustit se do řešení slovutné úlohy o brachistochroně.

> *Uvažme proto kuličku (tedy spíše hmotný bod) o hmotnosti m v tíhovém poli Země startující v bodě $A = [0, 0]$, která se má při zanedbání tření a s nulovou počáteční rychlostí dostat v nejkratším čase do bodu $B = [b, \beta]$, kde $b > 0$ a $\beta \geq 0$[11]. Po jaké dráze toho docílíme?*

Celou situaci máme zakreslenu na Obrázku 9.4 níže, kde jsme pro přehlednost otočili orientaci osy y.

[11] Pro určitost volíme $b > 0$, abychom se pohybovali směrem doprava od A (mohli bychom ale volit i $b < 0$). V našem řešení si otočíme orientaci osy y, takže podmínka $\beta \geq 0$ zaručuje, že nechceme, aby kulička dojela výše než byl výchozí bod, což totiž vzhledem k předpokladu nulové počáteční rychlosti není možné.



Obrázek 9.4: Situační obrázek pro úlohu o brachistochroně.

Podle definice rychlosti platí

$$v(t) = \frac{\mathrm{d}s}{\mathrm{d}t} \quad \text{neboli} \quad \mathrm{d}t = \frac{\mathrm{d}s}{v},$$

kde $\mathrm{d}s$ značí délku dráhy uražené za čas $\mathrm{d}t$. Také víme, že pro délku segmentu křivky určené částí grafu funkce $y = y(x)$ v rozsahu $\mathrm{d}x$ platí

$$\mathrm{d}s = \sqrt{1 + y'^2(x)}\,\mathrm{d}x.$$

Rychlost v pohybujícího se hmotného bodu v tíhovém poli Země dokážeme stanovit pomocí *zákona zachování energie* pro izolovaný systém. Protože součet kinetické energie $E_k = mv^2/2$ a potenciální energie $E_p = mgy$ se nemění, musí vzhledem k otočené orientaci osy y platit[12]

$$0 = mv^2 - mgy \quad \text{neboli} \quad v = \sqrt{2gy},$$

kde $g \approx 9{,}8\,m/s^2$ představuje tíhové zrychlení na Zemi. Proto celková doba pohybu kuličky po dráze je rovna

$$t = \int_0^b \sqrt{\frac{1 + y'^2(x)}{2g\,y(x)}}\,\mathrm{d}x.$$

Naším cílem je tedy nalézt mezi všemi funkcemi $y \in \mathrm{C}^1[0,b]$ splňujícími $y(0) = 0$ a $y(b) = \beta$ minimum funkcionálu

$$\mathcal{J}[y] = \sqrt{\frac{1 + y'^2}{2gy}},$$

který explicitně nezávisí na x. V takové případě se Eulerova–Lagrangeova rovnice redukuje na Beltramiho rovnost

$$\sqrt{\frac{1 + y'^2}{2gy}} - \frac{y'^2}{\sqrt{2gy(1 + y'^2)}} = c \quad \text{neboli} \quad \sqrt{y(1 + y'^2)} = K$$

pro nějaké $c > 0$ a $K := (C\sqrt{2g})^{-1}$. Nahradíme-li $y' = \mathrm{tg}\,\varphi$, pak z předchozí rovnosti získáme $\sqrt{y(1 + \mathrm{tg}^2\,\varphi)} = K$, což dává parametrické vyjádření y-ové souřadnice hledaného řešení jako[13]

$$y = \frac{K^2}{1 + \mathrm{tg}^2\,\varphi} = K^2 \cos^2\varphi = L(1 + \cos 2\varphi)/2$$

pro $L := K^2$. Derivováním podle x odtud dostaneme diferenciální rovnici

$$\mathrm{tg}\,\varphi = y' = -L\varphi' \sin 2\varphi = -L \sin 2\varphi \frac{\mathrm{d}\varphi}{\mathrm{d}x}, \quad \text{tj.} \quad \mathrm{d}x = -L\frac{2\sin\varphi\cos\varphi}{\mathrm{tg}\,\varphi}\,\mathrm{d}\varphi = -2L\cos^2\varphi\,\mathrm{d}\varphi,$$

[12] Součet těchto energií musí být konstantní. Jenže vzhledem k počátečním podmínkám $y(0) = 0$ a $v(0) = 0$ máme $E_k(0) = 0 = E_p(0)$, takže tento součet je dokonce nulový.

[13] Zde se pochopitelně sluší připomenout některé užitečné rovnosti pro goniometrické funkce: $\frac{1}{1+\mathrm{tg}^2\,\varphi} = \cos^2\varphi = (1 + \cos 2\varphi)/2$, $\sin 2\varphi = 2\sin\varphi\cos\varphi$, $\int \cos^2\varphi\,\mathrm{d}\varphi = (2\varphi + \sin 2\varphi)/4$, $\sin(\pi - t) = \sin t$ a $\cos(\pi - t) = -\cos t$.

[14] Toto je jiné „t“ než na začátku, tj. není to čas.

z čehož přímým integrováním získáme parametrické vyjádření x-ové osy ve tvaru

$$x = -2L[\varphi/2 + (\sin 2\varphi)/4] + C = -L(2\varphi + \sin 2\varphi)/2 + C.$$

Zavedeme si nyní dva nové parametry[14] t a r tak, že položíme $t := \pi - 2\varphi$ (proměnná) a $r := L/2$ (konstanta). Potom

$$x = -r[\pi - t + \sin(\pi - t)] + C = r(t - \sin t) + C - r\pi \quad \text{a} \quad y = r(1 - \cos t).$$

Jelikož $x(0) = 0$, platí $C = r\pi$, takže máme parametrické vyjádření řešení

$$x = -r[\pi - t + \sin(\pi - t)] + C = r(t - \sin t) \quad \text{a} \quad y = r(1 - \cos t),$$

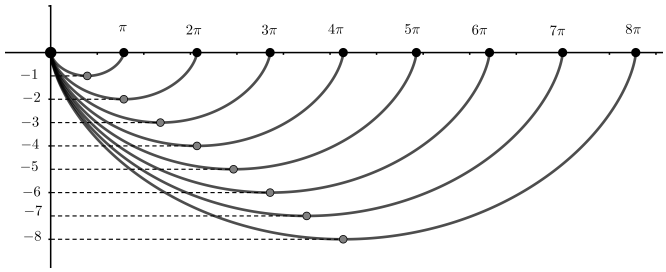
což je rovnice *cykloidy* určené kružnicí o poloměru r.

Podmínka $y(0) = 0$ znamená, že výchozí hodnota parametru t je $2k\pi$ pro nějaké $k \in \mathbb{Z}$. Pro jednoduchost můžeme zvolit $k = 0$ a jakákoli jiná volba $k \neq 0$ vše pouze posune o $2k\pi$. Maximální hodnotu parametru t pak získáme řešením rovnice

$$\frac{y(t_{\max})}{x(t_{\max})} = \frac{\beta}{b} = \frac{1 - \cos t_{\max}}{t_{\max} - \sin t_{\max}},$$

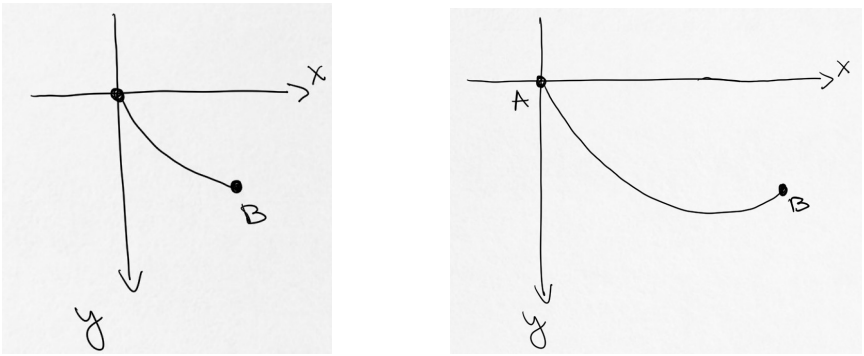
které bude $t_{\max} \in (0, 2\pi)$. Odtud také dostaneme poloměr vymezující kružnice

$$r = \frac{\beta}{1 - \cos t_{\max}} = \frac{b}{t_{\max} - \cos t_{\max}}.$$



Obrázek 9.5: Několik cykloid procházejících počátkem a body na ose x tvaru $[k\pi, 0]$ s vyznačeným nejnižším bodem. Všimněte si, že každým bodem v tomto kvadrantu prochází právě jedna křivka – ta je jednoznačně určena poloměrem r, který je roven y-ové souřadnici nejnižšího bodu. Omezením parametru $t \in [0, t_{\max}]$ pak získáme hledanou brachistochronu.

Jestliže $t_{\max} < \pi$ neboli $\frac{\beta}{b} > \frac{2}{\pi}$, tak kulička podél celé dráhy pouze klesá, a to až do nejnižšího bodu B. Je-li $t_{\max} \geq \pi$ neboli $\frac{\beta}{b} < \frac{2}{\pi}$, tak kulička podél brachistochrony klesá do nejnižšího bodu $[r\pi, 2r]$ odpovídajícího $t = \pi$ a pak stoupá až do bodu B.



Obrázek 9.6: Ilustrace situací $t_{\max} \lessgtr \pi$.

Např. pro $B = [1,1]$ je $t_{\max} \approx 2,412011$ a $r \approx 0,5729$. Čas potřebný pro cestu z A do B pak je roven $0,5831927828\,\mathrm{s}$. Podobně pro bod $B = [\pi,2]$ dostaneme $t_{\max} = \pi$ a $r = 1$ s potřebným časem $0,7705\,\mathrm{s}$.

Je nalezená extremála i hledaným řešením. Intuitivně můžeme říci, že ano, protože nějaké řešení „musí“ existovat a zároveň jsme nalezli jediného možného kandidáta. Ovšem toto matematicky nestačí

LITERATURA

[1] S. Aaronson, $\mathbf{P} \stackrel{?}{=} \mathbf{NP}$, Electronic Colloquium on Computational Complexity **TR17-004**, Report No. 4 (2017), str. 1–120. (Citováno na straně 309.)

[2] U. A. Acar, G. E. Blelloch, *Algorithms: Parallel and Sequential*, 2019. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/acar1. (Citováno na straně 258.)

[3] F. S. Acton, *Numerical Methods That Work*, Mathematical Association of America, Washington, 1990. (Citováno na straně 77.)

[4] J. Adda, R. W. Cooper, *Dynamic Economics: Quantitative Methods and Applications*, The MIT Press, Cambridge, 2003. (Citováno na straně 386.)

[5] I. Adler, G. C. Resende, G. Veiga, N. Karmarkar, *An implementation of Karmarkar's algorithm for linear programming*, Math. Programming **44** (1989), č. 3, str. 297–335. (Citováno na straně 249.)

[6] D. J. Albers, C. Reid, G. B. Dantzig, *An interview with George B. Dantzig: The father of linear programming*, College Math. J. **17** (1986), č. 4, str. 292–314. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/dantzig1. (Citováno na stranách 61 a 71.)

[7] D. Alevras, M. W. Padberg, *Linear Optimization and Extensions: Problems and Solutions*, Springer, New York, 2001. (Citováno na stranách 211 a 233.)

[8] O. Y. M. Al-Rawi, T. Mukherjee, *Application of linear programming in optimizing labour scheduling*, Journal of Mathematical Finance **9** (2019), str. 272–285. (Citováno na straně 216.)

[9] N. Amenta, G. M. Ziegler, *Deformed products and maximal shadows of polytopes*, str. 57–90 v „Advances in Discrete and Computational Geometry", B. Chazelle, J. E. Goodman a R. Pollack (editoři), Contemp. Math. **223**, Amer. Math. Soc., Providence, 1999. (Citováno na straně 239.)

[10] D. R. Anderson, D. J. Sweeney, T. A. Williams, J. D. Camm, M. Kipp, *An Introduction to Management Science: Quantitative Approaches to Decision Making*, třinácté vydání, South-Western Cengage Learning, Mason, 2012. (Citováno na straně 199.)

[11] N. Angier, *Folding the perfect corner*, Time Magazine **124** (1984), č 55. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/timemagazine. (Citováno na straně 250.)

[12] D. L. Applegate, R. E. Bixby, V. Chvátal, W. J. Cook, *The Traveling Salesman Problem: A Computational Study*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, 2006. (Citováno na stranách 258 a 268.)

[13] V. I. Arnold, *Yesterday and Long Ago*, z ruského originálu „Istorii davnie i nedavnie" (PHASIS, Moskva, 2006) přeloženo L. P. Kotovou a O. L. deLangem , Springer-Verlag, Berlin, 2007. ISBN 978-3-540-28734-6. (Citováno na straně 68.)

[14] K. J. Arrow, T. Harris, J. Marschak, *Optimal inventory policy*, Econometrica **19** (1951), str. 250–272. (Citováno na straně 368.)

[15] K. J. Arrow, D. Blackwell, M. A. Girshick, *Bayes and minimax solutions of sequentials decision problems*, Econometrica **17** (1949), str. 213–244. (Citováno na straně 368.)

[16] M. Ashbaugh, R. Benguria, *The problem of queen Didó: Overview of the Subject of Isoperimetry*, 2010. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/dido1. (Citováno na straně 52.)

[17] A. A. Assad, S. I. Gass, *Profiles in Operations Research: Pioneers and Innovators*, International Series in Operations Research & Management Science **147**, Springer, New York, 2011. ISBN 978-1-4419-6280-5. (Citováno na stranách 79, 80 a 585.)

[18] J.–P. Aubin, *Mathematical Methods of Game and Economic Theory*, Stud. Math. Appl. **7**, North-Holland Publishing, Amsterdam–New York, 1979. (Citováno na straně 520.)

[19] D. Avis, *Postscript to „What is the worst case behavior of the simplex algorithm?“*, str. 145–147 v „Polyhedral Computation“, D. Avis, D. Bremner a A. Deza (editoři), CRM Proc. Lecture Notes **48**, Amer. Math. Soc., Providence, 2009. (Citováno na straně 240.)

[20] D. Avis, V. Chvátal, *Notes on Bland's pivoting rule*, str. 24–34 v „Polyhedral Combinatorics: Dedicated to the Memory of D. R. Fulkerson“, M. L. Balinski a A. J. Hoffman (editoři), Math. Programming Stud. **8** (1978). (Citováno na straně 239.)

[21] D. Avis, B. Kaluzny and D. Titley-Péloquin, *Visualizing and constructing cycles in the simplex method*, Operations Research **56** (2008), č. 2, str. 512–518. (Citováno na straně 155.)

[22] M. Avriel, W. E. Diewert, S. Schaible, I. Zang, *Generalized Concavity*, Classics In Applied Mathematics **63**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2010. (Citováno na straně 540.)

[23] E. Balas, S. Ceria, G. Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, Math. Programming **58** (1993), č. 3, str. 295–324. (Citováno na straně 303.)

[24] E. Balas, S. Ceria, G. Cornuéjols, N. Natraj, *Gomory cuts revisited*, Oper. Res. Lett. **19** (1996), č. 1, str. 1–9. (Citováno na straně 303.)

[25] E. Balas, M. W. Padberg, *Set partitioning: a survey*, SIAM Review **18** (1976), č. 4, str. 710–760. (Citováno na straně 272.)

[26] I. Bárány, *A generalization of Carathéodory's theorem*, Discrete Math. **40** (1982), č. 2-3, str. 141–152. (Citováno na straně 437.)

[27] I. Bárány, S. Onn, *Carathéodory's theorem, colourful and applicable*, str. 11–21 v „Intuitive Geometry (Budapest, 1995)“, I. Bárány a K. Böröczky (editoři), Bolyai Soc. Math. Stud. **6**, Budapest, 1997. (Citováno na straně 437.)

[28] I. Bárány, P. Soberón, *Tverberg's theorem is 50 years old: a survey*, Bull. Amer. Math. Soc. **55** (2018), č. 4, str. 459–492. (Citováno na straně 440.)

[29] E. R. Barnes, *A variation on Karmarkar's algorithm for solving linear programming problems*, Math. Programming **36** (1986), č. 2, str. 174–182. (Citováno na straně 251.)

[30] A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics, Vol. 54, American Mathematical Society, Providence, 2002. (Citováno na straně 429.)

[31] J. R. Batt, R. A. Gellatly, *A Discretized Program for the Optimal Design of Complex Structures*, AGARD Lecture Series M70, NATO, 1974. (Citováno na straně 617.)

[32] H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2011. (Citováno na straně 472.)

[33] M. S. Bazaraa, J. J. Jarvis, H. D. Sherali, *Linear Programming and Network Flows*, čtvrté vydání, John Wiley & Sons, Hoboken, 2010. (Citováno na stranách 162, 187 a 190.)

[34] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, třetí vydání, John Wiley & Sons, Hoboken, 2006. (Citováno na stranách 612, 614, 616, 617 a 618.)

[35] E. M. L. Beale, *An alternative method for linear programming*, Proc. Cambridge Philos. Soc. **50** (1954), str. 513–523. (Citováno na straně 175.)

[36] E. M. L. Beale, *Cycling in the dual simplex algorithm*, Naval Res. Logist. Quart. **2** (1955), č. 4, str. 269–275. (Citováno na stranách 156 a 178.)

[37] E. M. L. Beale, R. Small, *Mixed integer programming by a branch and bound technique*, str. 450–451 v „Proceedings of the Third IFIP Congress“, 1965. (Citováno na straně 281.)

[38] A. Beck, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*, MOS-SIAM Series on Optimization **19**, Society for Industrial and Applied Mathematics (SIAM) & Mathematical Optimization Society, Philadelphia, 2014. (Citováno na stranách 472, 580 a 585.)

[39] E. F. Beckenbach, *Convex functions*, Bull. Amer. Math. Soc. **54** (1948), č. 5, str. 439–460. (Citováno na straně 452.)

[40] M. J. Beckmann, *Dynamic programming and the secretary problem*, Comput. Math. Appl. **19** (1990), č. 11, str. 25–28. (Citováno na straně 390.)

[41] E. Behrends, *On Bárány's theorems of Carathéodory and Helly type*, Studia Math. **141** (2000), č. 3, str. 235–250. (Citováno na straně 437.)

[42] E. T. Bell, *Men of Mathematics*. Penguin Books, Melbourne, 1953. (Citováno na straně 268.)

[43] D. R. Bellhouse, *The problem of Waldegrave*, J. Électron. Hist. Probab. Stat. **3** (2007), č. 2, 12 str. (elektronicky). (Citováno na straně 56.)

[44] D. R. Bellhouse, N. Fillion, *Le her and other problems in probability discussed by Bernoulli, Montmort and Waldegrave*, Statist. Sci. **30** (2015), č. 1, str. 26–39. (Citováno na straně 56.)

[45] R. E. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, 1957. (Citováno na stranách 367, 380, 390 a 399.)

[46] R. E. Bellman, S. E. Dreyfus *Applied Dynamic Programming*, Princeton University Press, Princeton, 1962. (Citováno na straně 390.)

[47] R. E. Bellman, *Eye of the Hurricane: An Autobiography*, World Scientific Publishing, Singapore, 1984. ISBN 9971-966-01-8. (Citováno na stranách 367 a 369.)

[48] M. Ben–Daya, C. M. Shetty, *Polynomial barrier function algorithms for convex quadratic programming*, Arabian J. Sci. Engrg. **15** (1990), č. 4, str. 657–670. (Citováno na straně 359.)

[49] A. Ben-Israel, *On Some Problems of Mathematical Programming*, disertační práce, Northwestern University, 1962. (Citováno na straně 303.)

[50] A. Ben-Israel, A. Charnes, *On some problems of diophantine programming*, Cahiers Centre Études Rech. Opér. **4** (1962), str. 215–280. (Citováno na straně 303.)

[51] A. Ben-Tal, A. Nemirovski, *Lecture Notes Optimization III: Convex Analysis, Nonlinear Programming Theory, Nonlinear Programming Algorithms*, 2015. Dostupné online na adrese https://www2.isye.gatech.edu/~nemirovs/OPTIII_LectureNotes2015.pdf. (Citováno na stranách 10, 578, 580, 581, 583 a 605.)

[52] J. Van den Bergh, J. Beliën; P. De Bruecker, E. Demeulemeester, L. De Boeck, *Personnel scheduling: a literature review*, European J. Oper. Res. **226** (2013), č. 3, str. 367–385. (Citováno na straně 216.)

[53] D. Bergman, *An exact algorithm for the quadratic multiknapsack problem with an application to event seating*, INFORMS J. Comput. **31** (2019), č. 3, str. 477–492. (Citováno na straně 328.)

[54] D. P. Bertsekas, *Dynamic Programming and Optimal Control I*, Athena Scientific, Belmont, 1995. (Citováno na straně 390.)

[55] D. P. Bertsekas, *Dynamic Programming and Optimal Control II*, Athena Scientific, Belmont, 1995. (Citováno na straně 390.)

[56] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Athena Scientific, Belmont, 1996. (Citováno na straně 535.)

[57] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, 1999. (Citováno na stranách 10, 570, 571, 573, 583, 584, 632 a 636.)

[58] D. P. Bertsekas, *Convex Optimization Theory*, Athena Scientific, Belmont, 2009. (Citováno na stranách 10 a 652.)

[59] D. P. Bertsekas, A. Nedic, A. Ozdeglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, 2003. (Citováno na straně 10.)

[60] D. Bertsimas, J. N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, 1997. (Citováno na stranách 141 a 233.)

[61] A. K. Bhunia, L. Sahoo, A. A. Shaikh, *Advanced Optimization and Operations Research*, Springer Optimization and Its Applications **153**, Springer, Singapore, 2019. (Citováno na straně 187.)

[62] L. Biacino, *A geometrical solution by Fermat to a problem of maximum*, Appl. Math. Sci. (Ruse) **8** (2014), no. 136, str. 6827–6834. (Necitováno.)

[63] N. L. Biggs, *T. P. Kirkman, mathematician*, Bull. London Math. Soc. **13** (1981), č. 2, str. 97–120. (Citováno na straně 265.)

[64] N. L. Biggs, E. K. Lloyd, R. J. Wilson, *Graph Theory 1736–1936*, druhé vydání, The Clarendon Press, Oxford, 1986. (Citováno na stranách 258 a 262.)

[65] V. Bissonnette, *Operational Research and the British Bombing Offensive against Germany, 1941–1945: Calculating Choices in War*, War in History **25** (2018), č. 4, str. 534–554. (Citováno na straně 56.)

[66] R. E. Bixby, *A brief history of linear and mixed-integer programming computation*, str. 107–121 v „Optimization Stories (21st International Symposium on Mathematical Programming, Berlin, Germany, 2012)“, M. Grötschel (editor), Doc. Math. **2012** (2012). (Citováno na straně 306.)

[67] C. Blair Jr., *Passing of a great mind*, LIFE **42** (25. února 1957), č. 8, str. 89–104. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/vonneumann1. (Citováno na straně 65.)

[68] R. G. Bland, *New finite pivoting rules for the simplex method*, Math. Oper. Res. **2** (1977), č. 2, str. 103–107. (Citováno na straně 156.)

[69] R. G. Bland, D. Goldfarb, M. J. Todd, *The ellipsoid method: a survey*, Oper. Res. **29** (1981), č. 6, str. 1039–1091. (Citováno na straně 241.)

[70] F. Bock, *Mathematical programming solution of traveling salesman examples*, str. 339–341 v „Recent Advances in Mathematical Programming“, R. L. Graves a P. S. Wolfe (editoři), McGraw–Hill, New York, 1963. (Necitováno.)

[71] I. Boldyrev, T. Düppe, *Programming the USSR: Leonid V. Kantorovich in context*, British J. Hist. Sci. **53** (2020), č. 2, str. 255–278. (Citováno na stranách 66 a 68.)

[72] A. Bollard, *Economists at War: How a Handful of Economists Helped Win and Lose the World Wars*, Oxford University Press, Oxford, 2020. ISBN 978-019884600-0. (Citováno na straně 68.)

[73] T. Bonnesen, W. Fenchel, *Theorie der konvexen Körper* [německy], opravený dotisk původního vydání z roku 1934, Springer–Verlag, Berlin–New York, 1974. (Citováno na stranách 436 a 439.)

[74] B. T. Borchers, *Improved Branch and Bound Algorithms for Integer Programming*, disertační práce, Rensselaer Polytechnic Institute, 1992. (Citováno na straně 281.)

[75] B. T. Borchers, J. E. Mitchell, *An improved branch and bound algorithm for mixed integer nonlinear programs*, Comput. Oper. Res. **21** (1994), č. 4, str. 359–367. (Citováno na straně 281.)

[76] K. Borsuk, *Drei Sätze über die n-dimensionale euklidische Sphäre* [německy], Fundam. Math. **20** (1933), str. 177–190. (Citováno na straně 442.)

[77] J. M. Borwein, A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, druhé vydání, CMS Books in Mathematics, Springer, 2006. (Citováno na straně 10.)

[78] V. J. Bowman Jr., G. L. Nemhauser, *A finiteness proof for modified Dantzig cuts in integer programming*, Naval Res. Logist. Quart. **17** (1970), č. 3, str. 309–313. (Citováno na straně 292.)

[79] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004. (Citováno na straně 10.)

[80] S. Boyd, L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*, Cambridge University Press, Cambridge, 2018. ISBN 978-1-316-51896-0. (Citováno na straně 325.)

[81] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, New York, 2011. (Citováno na straně 483.)

[82] G. Bruno, A. Genovese, G. Improta, *Routing problems: a historical perspective*, BSHM Bull. **26** (2011), č. 2, str. 118–127. (Citováno na straně 258.)

[83] M. W. Browne, *A Soviet Discovery Rocks World of Mathematics*, New York Times (November 2, 1979). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/NYT. (Citováno na straně 241.)

[84] M. W. Browne, *An Approach to Difficult Problems*, New York Times (November 27, 1979). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/NYT2. (Citováno na straně 243.)

[85] J. J. Burckhardt, E. A. Fellmann, W. Habicht (editoři), *Leonhard Euler 1707–1783: Beiträge zu Leben und Werk* Birkhäuser Verlag, Basel, 1983. (Necitováno.)

[86] P. Bussotti, *On the genesis of the Lagrange multipliers*, J. Optim. Theory Appl. **117** (2003), č. 3, str. 453–459. (Citováno na straně 35.)

[87] E. R. Butchers, P. R. Day, A. P. Goldie, S. Miller, J. A. Meyer, D. M. Ryan, A. C. Scott, Ch. A. Wallace, *Optimized crew scheduling at Air New Zealand*, INFORMS Journal on Applied Analytics **31** (2001), str. 30–56. (Citováno na straně 306.)

[88] P. Camion, *Characterization of totally unimodular matrices*, Proc. Amer. Math. Soc. **16** (1965), str. 1068–1073. (Citováno na straně 275.)

[89] J. D. Camm, T. E. Chorman, F. A. Dill, J. R. Evans, D. J. Sweeney, G. W. Wegryn, *Blending OR/MS, judgment, and GIS: restructuring P&G's supply chain*, Interfaces **27** (1997), č. 1, str. 128–142. (Citováno na straně 278.)

[90] C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen* [německy], Math. Ann. **64** (1907), str. 95–115. (Citováno na straně 436.)

[91] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen* [německy], Palermo Rend. **32** (1911), str. 193–217. (Citováno na straně 436.)

[92] F. Caro, J. Gallien, M. Díaz, J. García, J. M. Corredoira, M. Montes, J. A. Ramos, J. Correa, *Zara uses operations research to reengineer its global distribution process*, INFORMS Journal on Applied Analytics **40** (2010), č. 1, str. 71–84. (Citováno na straně 306.)

[93] B. Chachuat, *Nonlinear and Dynamic Optimization: From Theory to Practice*, École Polytechnique Fédérale de Lausanne, Lausanne, 2007. (Citováno na straně 619.)

[94] G. Cornuéjols, *Revival of the Gomory cuts in the 1990's*, Ann. Oper. Res. **149** (2007), str. 63–66. (Citováno na straně 303.)

[95] J. Casti, *Five Golden Rules: Great Theories of 20th-Century Mathematics — and Why They Matter*, John Wiley & Sons, 1996. (Citováno na straně 658.)

[96] A. Cobham, *The intrinsic computational difficulty of functions*, str. 24–30 v „Logic, Methodology and Philosophy of Science: Proceedings of the 1964 International Congress", Y. Bar-Hillel (editor), Studies in Logic and the Foundations of Mathematics, North-Holland Publishing Company, Amsterdam, 1965. (Citováno na straně 308.)

[97] C. W. Combettes, S. Pokutta, *Revisiting the approximate Carathéodory problem via the Frank–Wolfe algorithm*, Math. Program. (2021), v tisku. (Citováno na straně 437.)

[98] M. Conforti, G. Cornuéjols, G. Zambelli, *Integer Programming*, Graduate Texts in Mathematics 271, Springer, Cham, 2014. (Citováno na stranách 300 a 303.)

[99] S. A. Cook, *The complexity of theorem-proving procedures*, str. 151–158 v „STOC '71: Proceedings of the third annual ACM symposium on Theory of computing", Association for Computing Machinery, 1971. (Citováno na straně 308.)

[100] W. J. Cook, *Markowitz and Manne + Eastman + Land and Doig = branch and bound*, str. 227–238 v „Optimization Stories (21st International Symposium on Mathematical Programming, Berlin, Germany, 2012)", M. Grötschel (editor), Doc. Math. **2012** (2012). (Citováno na straně 303.)

[101] W. J. Cook, *In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation*, Princeton University Press, Princeton, 2012. Vyšlo také česky jako „Po stopách obchodního cestujícího: Matematika na hranicích možností", Argo, Praha, 2012. (Citováno na stranách 258 a 268.)

[102] W. W. Cooper, *Abraham Charnes and W. W. Cooper (et al.): A brief history of a long collaboration in developing industrial uses of linear programming*, 50th Anniversary Issue of Operations Research, Oper. Res. **50** (2002), č. 1, str. 35–41. (Citováno na straně 162.)

[103] R. W. Cottle, *Symmetric dual quadratic programs*, Research Report **19**, Operations Research Center, University of California, Berkeley, 1962. (Citováno na straně 334.)

[104] R. W. Cottle, *George B. Dantzig: operations research icon*, Oper. Res. **53** (2005), č. 6, str. 892–898. (Citováno na straně 80.)

[105] R. W. Cottle, *George B. Dantzig: a legendary life in mathematical programming*, Math. Program. **105** (2006), č. 1, str. 1–8. (Citováno na straně 80.)

[106] R. W. Cottle, B. C. Eaves, M. N. Thapa, *Dantzig, George B. (1914–2005)*, v „The New Palgrave Dictionary of Economics", M. Palgrave (editor), Palgrave Macmillan, London, 2008. (Citováno na straně 80.)

[107] R. W. Cottle, E. Johnson, R. Wets, *George B. Dantzig (1914–2005)*, Notices Amer. Math. Soc. **54** (2007), č. 3, str. 344–362. (Citováno na straně 80.)

[108] W. H. Cunningham, *Theoretical properties of the network simplex method*, Math. Oper. Res. **4** (1979), č. 2, str. 196–208. (Citováno na straně 161.)

[109] R. J. Dakin, *A tree-search algorithm for mixed integer programming problems*, Comput. J. **8** (1965), str. 250–255. (Citováno na straně 281.)

[110] R. C. Daniel, M. Jeffreys, *Unboundedness in integer and discrete programming L.P. relaxations*, J. Oper. Res. Soc. **30** (1979), str. 1119–1128. (Citováno na straně 282.)

[111] G. B. Dantzig, *On the non-existence of tests of „Student's" hypothesis having power functions independent of* σ, Ann. Math. Statistics **11** (1940), str. 186–192. (Citováno na straně 72.)

[112] G. B. Dantzig, *Programming of interdependent activities II: Mathematical model*, Econometrica **17** (1949), č. 3-4, str. 200–211. (Citováno na straně 78.)

[113] G. B. Dantzig, *Maximization of a linear function of variables subject to linear inequalities*, str. 339–347 v „Activity Analysis of Production and Allocation", T. C. Koopmans (editor), Cowles Commission Monograph No. 13, John Wiley & Sons, New York; Chapman & Hall, London, 1951. (Citováno na stranách 140, 149, 150 a 167.)

[114] G. B. Dantzig, *Linear programming under uncertainty*, Management Sci. **1** (1955), č. 3-4, str. 197-206. (Citováno na straně 59.)

[115] G. B. Dantzig, *Discrete-variable extremum problems*, Oper. Res. **5** (1957), č. 2, str. 266–288. (Citováno na straně 285.)

[116] G. B. Dantzig, *Note on solving linear programs in integers*, Naval Res. Logist. Quart. **6**, č. 1, str. 75–76 (Citováno na stranách 289 a 292.)

[117] G. B. Dantzig, *Linear Programming and Extensions*, Princeton University Press, Princeton, 1963. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/dantzig2. (Citováno na stranách 10, 63, 71, 78, 154, 155, 163 a 238.)

[118] G. B. Dantzig, *Reminiscences about the origins of linear programming*, Oper. Res. Lett. **1** (1982), č. 2, str. 43–48. (Citováno na stranách 78, 80 a 164.)

[119] G. B. Dantzig, *Origins of the simplex method*, Technical Report SOL 87-5, Department of Operations Research, Stanford University, Stanford, 1987. Publikováno také na str. 141–151 v „A History of Scientific Computing", Association for Computing Machinery, New York, 1990. (Citováno na straně 71.)

[120] G. B. Dantzig, *The diet problem*, Interfaces **20** (1990), č. 4, str. 43–47. (Citováno na straně 77.)

[121] G. B. Dantzig, *Linear programming*, v „The New Palgrave Dictionary of Economics", M. Palgrave (editor), Palgrave Macmillan, London, 2008. (Necitováno.)

[122] G. B. Dantzig, R. Fulkerson, S. Johnson, *Solution of a large-scale traveling-salesman problem*, J. Operations Res. Soc. Amer. **2** (1954), str. 393–410. (Citováno na stranách 263 a 264.)

[123] G. B. Dantzig, R. Fulkerson, S. Johnson, *On a linear-programming, combinatorial approach to the traveling-salesman problem*, Operations Res. **7** (1959), str. 58–66. (Citováno na straně 263.)

[124] G. B. Dantzig, A. Orden, P. S. Wolfe, *The generalized simplex method for minimizing a linear form under linear inequality restraints*, Pacific J. Math. **5** (1955), str. 183–195. (Citováno na stranách 156 a 341.)

[125] G. B. Dantzig, M. N. Thapa, *Linear Programming I: Introduction*, Springer, New York, 1997. (Citováno na stranách 167, 190, 198, 233 a 251.)

[126] G. B. Dantzig, M. N. Thapa, *Linear Programming II: Theory and Extensions*, Springer, New York, 2003. (Citováno na stranách 155 a 190.)

[127] G. B. Dantzig, A. Wald, *On the fundamental lemma of Neyman and Pearson*, Ann. Math. Statistics **22** (1951), str. 87–93. (Citováno na straně 72.)

[128] L. Danzer, B. Grünbaum, V. Klee, *Helly's theorem and its relatives*, Proc. Sympos. Pure Math. **7** (1963), str. 101–180. (Citováno na stranách 437, 440 a 441.)

[129] M. Dascal (editor), *Gottfried Wilhelm Leibniz: The Art of Controversies*, New Synthese Historical Library 60, Springer, Dordrecht, 2006. (Citováno na straně 258.)

[130] M. Davis, *Engines of Logic: Mathematicians and the Origin of the Computer*, opravené první vydaní, W. W. Norton, New York, 2001. (Citováno na straně 306.)

[131] P. Davis, R. Hersh, *The Mathematical Experience*, Houghton Mifflin, Boston, 1981. (Necitováno.)

[132] D. A. d'Esopo, *A convex programming procedure*, Naval Res. Logist. Quart. **6** (1959), č. 1, str. 33–42. (Citováno na straně 335.)

[133] J. C. Dickson, F. P. Frederick, *A decision rule for improved efficiency in solving linear programming problems with the simplex algorithm*, Comm. ACM **3** (1960), str. 509–512. (Citováno na straně 150.)

[134] I. I. Dikin, *Iterative solution of problems of linear and quadratic programming* [rusky], Dokl. Akad. Nauk SSSR **174** (1967), č. 4, str. 747–748. Pro anglický překlad viz Soviet Math. Dokl. **8** (1967), str. 674–675. (Citováno na straně 251.)

[135] I. I. Dikin, *On the speed of an iterative process* [rusky], Upravlyaemye Sistemi **12**, str. 54–60. (Citováno na straně 251.)

[136] I. I. Dikin, *A study of optimal programming problems by the interior points method* [rusky], str. 72–108 a 169 v „Optimization Methods, Applied Mathematics" [rusky], Sibirsk. Ènerget. Inst. Sibirsk. Otdel. Akad. Nauk SSSR Irkutsk, 1974. (Citováno na straně 251.)

[137] I. I. Dikin, *Letter to the editor: „A variation on Karmarkar's algorithm for solving linear programming problems" [Math. Programming 36 (1986), no. 2, 174–182] by E. R. Barnes*, Math. Programming **41** (1988), č. 3, str. 393–394. (Citováno na straně 251.)

[138] M. H. Dodani, A. J. G. Babu, *Karmarkar's projective method for linear programming: a computational survey*, Internat. J. Math. Ed. Sci. Tech. **21** (1990), č. 2, str. 191–212. (Citováno na straně 249.)

[139] J. Dongarra, F. Sullivan, *Guest editors' introduction: The top 10 algorithms*, Computing in Science and Engg. **2** (2000), str. 22–23. (Citováno na stranách 78 a 599.)

[140] R. Dorfman, *Mathematical, or "linear," programming: A nonmathematical exposition*, The American Economic Review **43** (1953), č. 5, str. 797–825. (Citováno na straně 58.)

[141] Z. Došlá, O. Došlý, *Metrické prostory*, Masarykova univerzita, Brno, 2000. (Citováno na stranách 438 a 482.)

[142] Z. Došlá, J. Kuben, *Diferenciální počet funkcí jedné proměnné*, Masarykova univerzita, Brno, 2004. (Citováno na straně 452.)

[143] O. Došlý, *Základy konvexní analýzy a optimalizace v* $\mathbb{R}^n$, Masarykova univerzita, Brno, 2005. (Citováno na stranách 10, 29 a 340.)

[144] O. Došlý, P. Zemánek, *Integrální počet v* $\mathbb{R}$, Masarykova univerzita, Brno, 2011. (Citováno na stranách 706 a 707.)

[145] N. J. Driebeek, *An algorithm for the solution of mixed integer programming problems*, Management Sci. **12** (1966), str. 576–587. (Citováno na straně 281.)

[146] D. Dugošija, A. Savić, Z. Maksimović, *A new integer linear programming formulation for the problem of political districting*, Ann. Oper. Res. **288** (2020), str. 247–263. (Citováno na straně 272.)

[147] W. Dunham, *Journey through Genius: The Great Theorems of Mathematics*, John Wiley & Sons, New York, 1990. (Citováno na straně 51.)

[148] G. W. Dunnington, *Carl Friedrich Gauss: Titan of Science*, The Mathematical Association of America, Washington, 2004. (Citováno na straně 19.)

[149] F. Dyson, *Prof. Hermann Weyl, For.Mem.R.S.*, Nature **177** (1956), str. 457–458. (Citováno na straně 24.)

[150] J. Eckhoff, *Radon's theorem revisited*, str. 164–185 v „Contributions to Geometry: Proceedings of the Geometry-Symposium held in Singen, 1978", J. Tölke a M. Wills (editoři), Birkhäuser, Basel, 1979. (Citováno na stranách 439 a 441.)

[151] J. Eckhoff, *Helly, Radon, and Carathéodory type theorems*, str. 389–448 v „Handbook of Convex Geometry, Volume A", P. M. Gruber a J. M. Wills (editoři), North-Holland, Amsterdam, 1993. (Citováno na straně 437.)

[152] J. Edmonds, *Paths, trees, and flowers*, Canadian Journal of Mathematics **17** (1965), str. 449–467. (Citováno na straně 308.)

[153] H. A. Eiselt, C. L. Sandblom, *Integer Programming and Network Models*, Springer-Verlag, Berlin, 2000. (Citováno na stranách 285 a 289.)

[154] H. A. Eiselt, C. L. Sandblom, *Linear Programming and its Applications*, Springer, New York, 2007. (Citováno na stranách 187, 220, 224, 233 a 251.)

[155] H. A. Eiselt, C. L. Sandblom, *Nonlinear Optimization: Methods and Applications*, International Series in Operations Research & Management Science **282**, Springer, Cham, 2019. ISBN 978-3-030-19461-1; 978-3-030-19462-8. (Citováno na stranách 341 a 360.)

[156] A. Emmett, *Karmarkar's algorithm: A threat to simplex? An elegant new method of solving linear equations has yet to be thoroughly tested*, IEEE Spectrum **22** (1985), č. 12, str. 54–55. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/ieee. (Citováno na straně 250.)

[157] L. Euler, *De linea brevissima in superficie quacunque duo quaelibet puncta jungente* [O nejkratší čáře spojující dva libovolné body jakékoli plochy (On the shortest line joining two points on a surface)], Comm. Ac. Petrop. **3** (1732), str. 110–124. Publikováno také v Opera Omnia: Series 1, Volume 25, Orell Füssli, Turici, 1952. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler7. (Citováno na straně 52.)

[158] L. Euler, *Solutio problematis ad geometriam situs pertinentis* [Řešení úlohy týkající se geometrie polohy (The solution of a problem relating to the geometry of position)], Comm. Ac. Petrop. **8** (1741), str. 128–140. Publikováno také v Opera Omnia: Series 1, Volume 7, Orell Füssli, Turici, 1923. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler11. (Citováno na straně 258.)

[159] L. Euler, *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti* [Metoda nalézání křivek majících vlastnost maxima nebo minima neboli řešení isoperimetrického problému pojatého v nejširším smyslu (A Method for Finding Curved Lines Enjoying Properties of Maximum or Minimum, or Solution of Isoperimetric Problems in the Broadest Accepted Sense)], Apud Marcum-Michaelem Bousquet & Socios, Lausannae & Genevae, 1744. Publikováno také v Opera Omnia: Series 1, Volume 24, Orell Füssli, Turici, 1952. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler2. (Citováno na stranách 15, 53 a 705.)

[160] L. Euler, *Elementa calculi variationum* [Základy variačního počtu (The Elements of the Calculus of Variations)], N. Comm. Ac. Petrop. **10** (1766), str. 51–93. Publikováno také v Opera Omnia: Series 1, Volume 25, Orell Füssli, Turici, 1952. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler8. (Citováno na straně 53.)

[161] L. Euler, *Dopis Ch. Goldbachovi z 26. dubna 1757*. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler9. (Citováno na straně 261.)

[162] L. Euler, *Solution d'une question curieuse que ne paroit soumise à aucune analyse* [Řešení kuriózní otázky, která zřejmě nebyla podrobena žádné analýze (Solution of a curious question which does not seem to have been subjected to any analysis)], Mémoires de l'académie des sciences de Berlin **15** (1766), str. 310–337. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/euler10. (Citováno na straně 261.)

[163] U. Faigle, W. Kern, G. Still, *Algorithmic Principles of Mathematical Programming*, Kluwer Texts in the Mathematical Sciences **24**, Kluwer Academic Publishers, Dordrecht, 2002. ISBN 1-4020-0852-X. (Citováno na straně 162.)

[164] K. Fan, I. Glicksberg, A. J. Hoffman, *Systems of inequalities involving convex functions*, Proc. Amer. Math. Soc. **8** (1957), č. 3, str. 617–622. (Citováno na straně 490.)

[165] S.–Ch. Fang, S. C. Puthenpura, *Linear Optimization and Extensions: Theory and Algorithms*, Prentice Hall, Englewood Cliffs, 1993. (Citováno na stranách 245, 249 a 251.)

[166] J. Farkas, *Theorie der einfachen Ungleichungen* [německy], J. Reine Angew. Math. **124** (1902), str. 1–27. (Citováno na straně 486.)

[167] Y. Fathi, *Computational complexity of LCPs associated with positive definite symmetric matrices*, Math. Programming **17** (1979), č. 3, str. 335–344. (Citováno na straně 239.)

[168] Y. Fathi, C. Tovey, *Affirmative action algorithms*, Math. Programming **34** (1986), č. 3, str. 292–301. (Citováno na straně 161.)

[169] M. W. Fenchel, *On conjugate convex functions*, Can. J. Math. **1** (1949), str. 73–77. (Citováno na straně 513.)

[170] M. W. Fenchel, *Convex Cones, Sets and Functions*, Princeton University, Princeton, 1953. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/fenchel1. „Modernější" verze tohoto textu je k dispozici online na adrese https://ln.math.muni.cz/pz/opt/fenchel2. (Citováno na straně 58.)

[171] R. W. Farebrother, *Fitting Linear Relationships: A History of the Calculus of Observations 1750–1900*, Springer Series in Statistics, Springer-Verlag, New York, 1999. (Citováno na straně 69.)

[172] R. W. Farebrother, *A linear programming procedure based on de la Vallée Poussin's minimax estimation procedure*, Comput. Statist. Data Anal. **51** (2006), č. 2, str. 453–456. (Citováno na straně 69.)

[173] T. S. Ferguson, *Who solved the secretary problem?*, Statist. Sci. **4** (1989), č. 3, str. 282–296. (Citováno na straně 390.)

[174] P. de Fermat, *Methodus ad disquirendam maximam et minimam* [latinsky], vyšlo v „Varia opera mathematica D. Petri de Fermat", Tolosæ (1679). Text byl také publikován v souhrnném díle od P. Tanneryho a Ch. Henryho, „Œuvres de Fermat I", Gauthier–Villars, Paříž (1891) na str. 133–135, viz online https://ln.math.muni.cz/pz/opt/fermat1. Francouzský překlad vyšel v dalším díle „Œuvres de Fermat III" (1896) na str. 121–123, viz online https://ln.math.muni.cz/pz/opt/fermat2. Tato verze byla následně přeložena J. Rossem jako *Method for the study of maxima and minima* (2008), viz https://ln.math.muni.cz/pz/opt/fermat3. (Citováno na stranách 22 a 26.)

[175] P. de Fermat, *Methodus de maxima et minima* [latinsky], vyšlo v „Varia opera mathematica D. Petri de Fermat", Tolosæ (1679). Text byl také publikován v souhrnném díle od P. Tanneryho a Ch. Henryho, „Œuvres de Fermat I", Gauthier–Villars, Paříž (1891) na str. 147–153, viz online https://ln.math.muni.cz/pz/opt/fermat1. Francouzský překlad vyšel v dalším díle „Œuvres de Fermat III" (1896) na str. 131–136, viz online https://ln.math.muni.cz/pz/opt/fermat2. Tato verze byla následně přeložena J. Rossem jako *Method for the study of maxima and minima* (2008), viz https://ln.math.muni.cz/pz/opt/fermat3. (Citováno na stranách 25 a 26.)

[176] M. C. Ferris, A. B. Philpott, *On the performance of Karmarkar's algorithm*, The Journal of the Operational Research Society **39** (1988), č. 3, str. 257–270. (Citováno na straně 249.)

[177] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics I: Mainly Mechanics, Radiation, and Heat*, Addison–Wesley Publishing, Reading, 1963. (Citováno na straně 626.)

[178] A. V. Fiacco, G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York-London-Sydney, 1968. (Citováno na straně 250.)

[179] R. Fletcher, C. M. Reeves, *Function minimization by conjugate gradients*, Computer J. **7** (1964), str. 149–154. (Citováno na straně 594.)

[180] Ch. A. Floudas, P. M. Pardalos (editoři), *Encyclopedia of Optimization*, Springer, New York, 2009. (Citováno na straně 487.)

[181] Ch. A. Floudas, V. Visweswaran, *Quadratic optimization*, str. 217–269 v „Handbook of Global Optimization“, R. Horst P. M. Pardalos (editoři), Nonconvex Optim. Appl. 2, Kluwer Acad. Publ., Dordrecht, 1995. (Citováno na straně 360.)

[182] J. J. Forrest, D. Goldfarb, *Steepest-edge simplex algorithms for linear programming*, Math. Programming **57** (1992), č. 3, str. 341–374. (Citováno na straně 150.)

[183] J. B. Fourier, [referováno v] *Analyse des travaux de l'Academic Royale des Sciences, pendant l'annee 1823*, Partie mathèmatique, Histoire de l'Academie Royale des Sciences de l'Institut de France **6** [1823] (1826), str. xxix–xli. Částečně přetištěno jako *Premier extrait* v „Oeuvres de Fourier II“, G. Darboux (editor), Gauthier–Villars, Paříž, 1890, str. 321–324. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/fourier1. (Citováno na straně 69.)

[184] M. Frank, P. S. Wolfe, *An algorithm for quadratic programming*, Naval Res. Logist. Quart. **3** (1956), č. 1-2, str. 95–110. (Citováno na straně 341.)

[185] C. Fraser, *J. L. Lagrange's changing approach to the foundations of the calculus of variations*, Arch. Hist. Exact Sci. **32** (1985), č. 2, str. 151–191. (Citováno na straně 705.)

[186] C. Fraser, *Joseph Louis Lagrange, Théorie des fonctions analytiques, first edition (1797)*, Kapitola 19 v „Landmark Writings in Western Mathematics 1640–1940“, I. Grattan–Guinness (editor), Elsevier, Amsterdam, 2005. (Citováno na straně 35.)

[187] P. R. Freeman, *The secretary problem and its extensions: a review*, Internat. Statist. Rev. **51** (1983), č. 2, str. 189–206. (Citováno na straně 390.)

[188] O. Friedmann, *A subexponential lower bound for Zadeh's pivoting rule for solving linear programs and games*, str. 192–206 v „Integer programming and combinatorial optimization“, O. Günlük a G. J. Woeginger (editoři), Lecture Notes in Comput. Sci. **6655**, Springer, Heidelberg, 2011. (Citováno na straně 240.)

[189] R. Frisch, *La résolution des problèmes de programme linéaire par la méthode du potentiel logarithmique* [francouzsky; The logarithmic potential method for solving linear programming problems], Cahiers du Séminaire d'Économétrie **1956** (1956), č. 4, str. 7–23. (Citováno na straně 250.)

[190] R. Frisch, *Programme convexe et planification pour le développement national* [francouzsky; The logarithmic potential method of convex programming. With particular application to the dynamics of planning for national developments], Colloques internationaux du Centre National de la recherche scientifique LXII, Les modèles dynamiques, Conférence à Paris mai 1955, 1956, str. 47–80. (Citováno na straně 250.)

[191] R. Frisch, *Macroeconomics and linear programming*, str. 38–67 v „25 Economic Essays in Honour of Erik Lindahl 21 November 1956“, Ekonomisk Tidskrift, 1956. (Citováno na straně 251.)

[192] G. Frobenius, *Theorie der linearen Formen mit ganzen Coefficienten* [německy; Teorie lineárních forem s celočíselnými koeficienty], J. Reine Angew. Math. **86** (1879), str. 146–208. (Citováno na straně 279.)

[193] G. Frobenius, *Theorie der linearen Formen mit ganzen Coefficienten (Fortsetzung der Abhandlung im 86. Bande dieses Journals)* [německy; Teorie lineárních forem s celočíselnými koeficienty (pokračování pojednání z 86. svazku tohoto časopisu)], J. Reine Angew. Math. **88** (1880), str. 96–116. (Citováno na straně 279.)

set triple

[194] E. Fuchs, I. Netuka, *Johann Radon*, Pokroky Mat. Fyz. Astronom. **33** (1988), č. 5, str. 282–285. (Citováno na straně 443.)

[195] F. Furini, E. Malaguti, D. Thomopulos, *Modeling two-dimensional guillotine cutting problems via integer programming*, Informs Journal On Computing **28** (2016), str. 736–751. (Citováno na straně 224.)

[196] D. Gale, The Theory of Linear Economic Models, McGraw–Hill, New York, 1960. (Citováno na straně 110.)

[197] D. Gale, *How to solve linear inequalities*, Amer. Math. Monthly **76** (1969), str. 589–599. (Citováno na straně 239.)

[198] G. Galilei, *Il Saggiatore*, Giacomo Mascardi, Rome, 1623. (Citováno na straně 26.)

[199] R. Gardner, *L. V. Kantorovich: The Price Implications of Optimal Planning*, Journal of Economic Literature **28** (1990), č. 2, str. 638–648. (Citováno na straně 68.)

[200] M. R. Garey, D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, A Series of Books in the Mathematical Sciences, W. H. Freeman, San Francisco, 1979. ISBN 0-7167-1045-5. (Citováno na straně 360.)

[201] R. S. Garfinkel, G. L. Nemhauser, *Optimal political districting by implicit enumeration techniques*, Management Sci. **16**, č. 8, Application Series, str. B495–B508. (Citováno na straně 272.)

[202] R. S. Garfinkel, G. L. Nemhauser, *Integer Programming*, John Wiley & Sons, Canada, 1972. (Citováno na straně 297.)

[203] S. I. Gass, *Linear Programming: Methods and Applications*, čtvrté vydání, McGraw–Hill, New York, 1975. (Citováno na stranách 190 a 206.)

[204] S. I. Gass, *Letter to the editor – Comments on the possibility of cycling with the simplex method*, Oper. Res. **27** (1979), č. 4, str. 848–852. (Citováno na straně 162.)

[205] S. I. Gass, *On artificial variables and the simplex method*, J. Inform. Optim. Sci. **8** (1987), č. 1, str. 77–79. (Citováno na straně 167.)

[206] S. I. Gass, *Encounters with degeneracy: a personal view*, Ann. Oper. Res. **46** (1993), č. 2, str. 335–342. (Citováno na straně 156.)

[207] S. I. Gass, *The first linear-programming shoppe*, v „50th anniversary issue of Operations Research“, L. M. Wein (editor), Oper. Res. **50** (2002), č. 1, str. 61–68. (Citováno na straně 77.)

[208] S. I. Gass, A. A. Assad, *An Annotated Timeline of Operations Research: An Informal History*, International Series in Operations Research & Management Science **75**, Kluwer Academic Publishers, Boston, 2005. ISBN 1-4020-8112-X. (Citováno na straně 54.)

[209] S. I. Gass, S. Vinjamuri, *Cycling in linear programming problems*, Comput. Oper. Res. **31** (2004), č. 2, str. 303–311. (Citováno na stranách 155, 156 a 162.)

[210] S. I. Gass, S. Vinjamuri, *Erratum to „Cycling in linear programming problems“ [Computers and Operations Research 31 (2002) 303–311]*, Comput. Oper. Res. **33** (2006), č. 8, str. 2445. (Citováno na straně 155.)

[211] C. F. Gauss, F. Bolyai, *Briefwechsel zwischen Carl Friedrich Gauss und Wolfgang Bolyai. Mit Unterstützung der Ungarischen Akademie der Wissenschaften herausgegeben von Franz Schmidt und Paul Stäckel*, B. G. Teubner, Leipzig, 1899. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/gauss1. (Citováno na straně 19.)

[212] B. Gendron, *Scheduling employees in Quebec's liquor stores with integer programming*, INFORMS Journal on Applied Analytics **35** (2005), č. 5, str. 402–410. (Citováno na straně 306.)

[213] A. M. Geoffrion, R. E. Marsten, *Integer programming algorithms: A framework and state-of-the-art survey*, Management Sci. **18** (1972), str. 465–491. (Citováno na straně 281.)

[214] A. Ghouila-Houri, *Caractérisation des matrices totalement unimodulaires* [francouzsky], C. R. Acad. Sci. Paris **254** (1962), str. 1192–1194. (Citováno na straně 275.)

[215] P. E. Gill, W. Murray, M. A. Saunders, M. H. Wright, *Practical Anti-Cycling Procedure for Linear and Nonlinear Programming*, Technical Report SOL 88-4, Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, 1988. (Citováno na straně 162.)

[216] Ch. C. Gillispie (editor), Dictionary of Scientific Biography 14, Charles Scribner's Sons, New York, 1976. (Citováno na straně 65.)

[217] J. Gleick, *Breakthrough in Problem Solving*, The New York Times (November 19, 1984). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/NYT4. (Citováno na straně 250.)

[218] F. Glover, *Generalized cuts in diophantine programming*, Management Sci. 13 (1966), str. 254–268. (Citováno na straně 297.)

[219] F. Glover, *A new foundation for a simplified primal integer programming algorithm*, Operations Res. **16** (1968), str. 727–740. (Citováno na straně 303.)

[220] P. Gočev, *Spor o ekonomickou kalkulaci*, Marathon, 2002, č. 4. (Citováno na straně 64.)

[221] T. E. Goldberg, L. M. Wilson, *Limits of golden constructions*, Math. Mag. **96** (2023), č. 4, str. 399–412. (Citováno na straně 553.)

[222] D. Goldfarb, S. Liu, *An $O(n^3L)$ primal interior point algorithm for convex quadratic programming*, Math. Programming **49** (1990/91), č. 3, str. 325–340. (Citováno na straně 359.)

[223] D. Goldfarb, J. K. Reid, *A practicable steepest-edge simplex algorithm*, Math. Programming **12** (1977), č. 3, str. 361–371. (Citováno na straně 150.)

[224] D. Goldfarb, W. Y. Sit, *Worst case behavior of the steepest edge simplex method*, Discrete Appl. Math. **1** (1979), č. 4, str. 277–285. (Citováno na straně 239.)

[225] D. Goldfarb, M. J. Todd, *Linear programming*, str. 73–170 v „Optimization", Handbooks Oper. Res. Management Sci. bf 1, North-Holland, Amsterdam, 1989. (Citováno na straně 249.)

[226] G. H. Golub, D. P. O'Leary, *Some history of the conjugate gradient and Lanczos algorithms: 1948–1976*, SIAM Review **31** (1989), č. 1, str. 50–102. (Citováno na straně 594.)

[227] R. E. Gomory, *Outline of an algorithm for integer solutions to linear programs*, Bull. Amer. Math. Soc. **64** (1958), str. 275–278. (Citováno na stranách 289 a 300.)

[228] R. E. Gomory, *An algorithm for the mixed integer problem*, Research Memoranda **RM-2597-PR**, The RAND Corporation, Santa Monica, 1960. (Citováno na stranách 289 a 304.)

[229] R. E. Gomory, *An all-integer programming algorithm*, IBM Research Report RC-189, New York, 1960. (Citováno na straně 297.)

[230] R. E. Gomory, *An algorithm for integer solutions to linear programs*, str. 269–302 v „Recent Advances in Mathematical Programming", R. L. Graves a P. Wolfe (editoři), McGraw–Hill, New York, 1963. (Citováno na stranách 289 a 300.)

[231] R. E. Gomory, *An all-integer programming algorithm*, str. 193–206 v „Industrial Scheduling", J. F. Muth a G. L. Thompson (editoři), Prentice Hall, Englewood Cliffs, 1963. (Citováno na straně 303.)

[232] R. E. Gomory, *Early integer programming*, Oper. Res. **50** (2002), č. 1, str. 78–81. (Citováno na straně 291.)

[233] R. E. Gomory, A. J. Hoffman, T. J. Watson *On the convergence of an integer-programming process*, Naval Res. Logist. Quart. **10** (1963), č. 1, str. 121–123. (Citováno na straně 292.)

[234] P. Gordan, *Über die Auflösungen linearer Gleichungen mit reelen coeffienten* [německy], Math. Ann. **6** (1873), str. 23–28. (Citováno na straně 486.)

[235] N. I. M. Gould, P. L. Toint, *A quadratic programming bibliography*, Numerical Analysis Group Internal Report 2000–1 Rutherford Appleton Laboratory, Chilton, England. Aktualizovaná verze je dostupná na webu ftp://ftp.numerical.rl.ac.uk/pub/qpbook/qp.pdf. (Citováno na straně 328.)

[236] F. Granot, J. Skorin–Kapov, *Towards a strongly polynomial algorithm for strictly convex quadratic programs: An extension of Tardos' algorithm*, Math. Programming **46** (1990), č. 2, str. 225–236. (Citováno na straně 359.)

[237] H. J. Greenberg, *Myths and Counterexamples in Mathematical Programming*, autorovy poznámky publikované na webu INFORMS Computing Society v rámci Mathematical Programming Glossary, verze 2010-10-20. (Citováno na straně 579.)

[238] M. Grötschel, L. Lovász, A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, druhé vydání, Algorithms and Combinatorics **2**, Springer-Verlag, Berlin, 1993. (Citováno na straně 245.)

[239] M. Grötschel, Y. Yuan, *Euler, Mei-Ko Kwan, Königsberg, and a Chinese postman*, Doc. Math. **2012** (2012), Extra vol.: Optimization stories, str. 43–50. (Citováno na straně 260.)

[240] P. M. Gruber and J. M. Wills (editoři), *Handbook of Convex Geometry: Volume A*, North-Holland Publishing, Amsterdam, 1993. (Citováno na straně 442.)

[241] P. M. Gruber and J. M. Wills (editoři), *Handbook of Convex Geometry: Volume B*, North-Holland Publishing, Amsterdam, 1993. (Citováno na straně 442.)

[242] P. Guerrero-García, Á. Santos-Palomo, *On Hoffman's celebrated cycling LP example*, Comput. Oper. Res. **34** (2007), č. 9, str. 2709–2717. (Citováno na straně 155.)

[243] O. K. Gupta, *Applications of quadratic programming*, J. Inform. Optim. Sci. **16** (1995), č. 1, str. 177–194. (Citováno na straně 328.)

[244] O. K. Gupta, A. Ravindran, *Branch and bound experiments in convex nonlinear integer programming*, Management Sci. **31** (1985), č. 12, str. 1533–1546. (Citováno na straně 281.)

[245] M. Habrovec, *Numerické metody nepodmíněné optimalizace*, diplomová práce, Masarykova univerzita, 2017. Dostupné online na adrese https://is.muni.cz/th/lb4jf. (Citováno na stranách 11 a 572.)

[246] G. Hadley, *Linear Programming*, Addison–Wesley Series in Industrial Management, Addison–Wesley Publishing, Reading, 1962. (Citováno na straně 167.)

[247] H. Hadwiger, H. Debrunner, *Ausgewählte Einzelprobleme der kombinatorischen Geometrie in der Ebene* [německy], Enseign. Math. **1** (1955), str. 56–89. (Citováno na straně 442.)

[248] H. Hadwiger, H. Debrunner, *Combinatorial Geometry in the Plane*, s novou kapitolou a dalšími přidanými pasážemi od překladatele, přeloženo V. Kleem, Holt, Rinehart and Winston, New York, 1964. (Citováno na straně 442.)

[249] J. A. J. Hall, K. I. M. McKinnon, *The simplest examples where the simplex method cycles and conditions where EXPAND fails to prevent cycling*, Math. Program. B **100** (2004), č. 1, str. 133–150. (Citováno na straně 162.)

[250] M. Hamala, *Nelineárne programovanie*, druhé vydání, Alfa, Bratislava, 1976. (Citováno na stranách 10, 341 a 359.)

[251] T. L. Hankins, *Sir William Rowan Hamilton*, Johns Hopkins University Press, Baltimore, 1980. (Citováno na straně 262.)

[252] G. Hardy, *A Mathematician's Apology*, Cambridge University Press, Cambridge, 2012. (Citováno na straně 26.)

[253] P. M. J. Harris, *Pivot selection methods of the Devex LP code*, Math. Program.**5** (1973), str. 1–28. (Citováno na straně 150.)

[254] J. Hartmanis, *Gödel, von Neumann and the P=?NP problem*, Bull. EATCS **38** (1989), str. 101–107. (Citováno na straně 307.)

[255] P. Hasil, P. Zemánek, *Sbírka řešených příkladů z matematické analýzy II*, Masarykova univerzita, Brno, 2016. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/SRPzMA2. (Citováno na stranách 32 a 44.)

[256] M. Heidegger, *Die Frage nach dem Ding: Zu Kants Lehre von den transzendentalen Grundsätzen* [německy], V. Klostermann, Frankfurt am Main, 1984. (Citováno na stranách 18 a 19.)

[257] M. Heidegger, *What is a Thing?*, přeloženo W. B. Bartonem Jr. a V. Deutsch s analýzou od E. T. Gendlina, Gateway Editions, South Bend, 1967. (Citováno na stranách 18 a 19.)

[258] R. Heinlein, *Time Enough For Love*, ACE Books, New York, 1973. (Citováno na straně 25.)

[259] I. Heller, C. B. Tompkins, *An extension of a theorem of Dantzig's*, str. 247–254 v „Linear Inequalities and Related Systems", H. W. Kuhn a A. W. Tucker (editoři), Annals of Mathematics Studies 38, Princeton University Press, Princeton, 1956. (Citováno na straně 276.)

[260] E. Helly, *Über lineare Funktionaloperationen* [německy], Wien. Ber. **121** (1912), str. 265–297. (Citováno na straně 440.)

[261] E. Helly, *Über Mengen konvexer Körper mit gemeinschaftlichen Punkten* [německy], Deutsche Math.-Ver. **32** (1923), str. 175–176. (Citováno na straně 440.)

[262] R. Herz-Fischler, *A Mathematical History of the Golden Number*, Dover Publications, 1998. (Citováno na straně 550.)

[263] M. R.. Hestenes, E. Stiefel, *Methods of conjugate gradients for solving linear systems*. J. Res. NBS **49** (1952), str. 409–436. (Citováno na straně 594.)

[264] D. Hilbert, *Mathematische Probleme – ortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900*, Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, Band Heft 3 (1900), str. 253–297. (Citováno na stranách 17 a 20.)

[265] C. G. Hildreth, *Point estimates of ordinates of concave functions*, J. Amer. Statist. Assoc. **49** (1954), str. 598–619. (Citováno na straně 335.)

[266] C. G. Hildreth, *A quadratic programming procedure*, Naval Res. Logist. Quart. **4** (1957), č. 1, str. 79–85. (Citováno na straně 335.)

[267] F. L. Hitchcock, *The distribution of a product from several sources to numerous localities*, J. Math. Phys. Mass. Inst. Tech. **20** (1941), str. 224–230. (Citováno na straně 68.)

[268] A. J. Hoffman, *Selected Papers of Alan Hoffman: With Commentary*, editováno Ch. A. Micchellim, World Scientific Publishing, River Edge, 2003. (Citováno na stranách 154, 155 a 276.)

[269] A. J. Hoffman, *What the transportation problem did for me*, Ann. Oper. Res. **149** (2007), str. 117–120. (Citováno na straně 277.)

[270] A. J. Hoffman, J. B. Kruskal, *Integral boundary points of convex polyhedra*, str. 223–246 v „Linear Inequalities and Related Systems", H. W. Kuhn a A. W. Tucker (editoři), Annals of Mathematics Studies 38, Princeton University Press, Princeton, 1956. (Citováno na stranách 276 a 278.)

[271] A. J. Hoffman, P. S. Wolfe, *History*, str. 1–15 v „The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization", E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan a D. B. Shmoys (editoři), Wiley-Interscience Series in Discrete Mathematics and Optimization, A Wiley-Interscience Publication, John Wiley & Sons, Chichester, 1990. (Citováno na straně 258.)

[272] M. Hojati, *Optimal political districting*, Computers Ops Res. **23** (1996), č. 12, str. 1147–1161. (Citováno na straně 272.)

[273] O. Hölder, *Über einen Mittelwertsatz* [německy], Gött. Nachr. **1889** (1889), str. 38–47. (Citováno na stranách 460, 463 a 477.)

[274] N. Hooker, *Karmarkar's linear programming algorithm*, Interfaces **16** (1986), č. 4, str. 75–90. (Citováno na straně 249.)

[275] B. Hopkins, R. J. Wilson, *The truth about Königsberg*, College Math. J.**35** (2004), č. 3, str. 198–207. (Citováno na straně 258.)

[276] R. A. Horn, Ch. R. Johnson, *Matrix Analysis*, druhé vydání, Cambdridge University Press, Cambridge, 2013. (Citováno na straně 27.)

[277] I. Horová, J. Zelinka, *Numerické metody*, 2. rozšířené vydání, Masarykova univerzita, Brno, 2004. (Citováno na straně 561.)

[278] L. G. Čačijan, *A polynomial algorithm in linear programming* [rusky], Doklady Akad. Nauk SSSR **244** (1979), str. 1093–1096. Pro anglický překlad viz Soviet Math. Dokl. **20** (1979), č. 1, str. 191–194. (Citováno na straně 241.)

[279] L. G. Čačijan, *Polynomial algorithms in linear programming*, Zh. Vychisl. Mat. i Mat. Fiz. **20** (1980), č. 1, str. 51–68, 260. Pro anglický překlad viz U.S.S.R. Comput. Math. and Math. Phys. **20**, str. 53–72. (Citováno na straně 241.)

[280] A. Chakraborty, V. Chandru, M. R. Rao, *A linear programming primer: from Fourier to Karmarkar*, Ann. Oper. Res. **287** (2020), č. 2, str. 593–616. (Citováno na straně 69.)

[281] R. Chandrasekaran, *Total unimodularity of matrices*, SIAM J. Appl. Math. **17** (1969), str. 1032–1034. (Citováno na straně 275.)

[282] A. Charnes, *Optimality and degeneracy in linear programming*, Econometrica **20** (1952), č. 2, str. 160–170. (Citováno na straně 156.)

[283] A. Charnes, W. W. Cooper, *Management Models and Industrial Applications of Linear Programming*, John Wiley & Sons, New York–London, 1961. (Necitováno.)

[284] A. Charnes, W. W. Cooper, *Handling artificial variables in a simplex method: a commentary*, J. Inform. Optim. Sci. **4** (1983), č. 1, str. 41–42. (Citováno na straně 167.)

[285] A. Charnes, W. W. Cooper, R. Mellon, *Blending aviation gasolines – a study in programming interdependent activities in an integrated oil company*, Econometrica **20** (1952), č. 2, str. 135–159. (Citováno na straně 161.)

[286] D.-S. Chen, R. G. Batson, Y. Dang, *Applied Integer Programming: Modeling and solution*, John Wiley & Sons, Hoboken, 2010. (Citováno na straně 303.)

[287] E. K. P. Chong, S. H. Zak, *An Introduction to Optimization*, třetí vydání, Wiley–Interscience, Hoboken, 2008. (Citováno na stranách 233, 588 a 593.)

[288] P. Choypeng, P. Puakpong, R. E. Rosenthal, *Optimal ship routing and personnel assignment for naval recruitment in Thailand*, INFORMS Journal on Applied Analytics **16** (1986), č. 4, str. 47–52. (Citováno na straně 306.)

[289] S. J. Chung, K. G. Murty, *Polynomially bounded ellipsoid algorithms for convex quadratic programming*, str. 439–485 v „Nonlinear Programming 4: Proceedings of the Fourth Nonlinear Programming Symposium held at the University of Wisconsin 1980“, O. L. Mangasarian, R. R. Meyer a S. M. Robinson (editoři), Academic Press, New York-London, 1981. (Citováno na straně 359.)

[290] V. Chvátal, *A combinatorial theorem in plane geometry*. J. Combinatorial Theory Ser. B **18** (1975), str. 39–41. (Citováno na straně 445.)

[291] V. Chvátal, *Hard knapsack problems*, Oper. Res. **28** (1980), č. 6, str. 1402–1411. (Citováno na stranách 283 a 308.)

[292] V. Chvátal, *Linear Programming*, W. H. Freeman, New York, 1983. (Citováno na stranách 155, 157, 158, 167, 220 a 233.)

[293] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, druhé vydání, Cambdridge Texts in Applied Mathematics, Vol. 44, Cambdridge University Press, Cambridge, 2009. (Citováno na straně 533.)

[294] A. Iványi (editor), *Algorithms of Informatics 3*, AnTonCom, Budapest, 2011. (Citováno na straně 280.)

[295] R. G. Jeroslow, *The simplex algorithm with the pivot rule of maximizing criterion improvement*, Discrete Math. **4** (1973), str. 367–377. (Citováno na straně 239.)

[296] R. G. Jeroslow, *Trivial integer programs unsolvable by branch-and-bound*, Math. Programming **6** (1974), str. 105–109. (Citováno na straně 283.)

[297] P. Jež, *Kolik mám denně přijmout bílkovin, tuků a sacharidů, obecně energie?* [elektronicky], Faktor plus, Benešov, 2017. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/LPjidelnicek1. (Citováno na straně 208.)

[298] P. Jolly, *Handling artificial variables in the simplex method*, J. Inform. Optim. Sci. **2** (1981), č. 3, str. 241–248. (Citováno na straně 167.)

[299] H. W. E. Jung, *Über die kleinste Kugel, die eine räumliche Figur einschliesst* [německy], J. Reine Angew. Math. **123** (1901), str. 241–257. (Citováno na straně 441.)

[300] H. W. E. Jung, *Über den kleinsten Kreis, der eine ebene Figur einschliesst* [německy], J. Reine Angew. Math. **137** (1910), str. 310–313. (Citováno na straně 441.)

[301] M. Jünger, T. Liebling, D. Naddef, G. L. Nemhauser, W. Pulleyblank, G. Reinelt, G. Rinaldi, L. Wolsey (editoři) *50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art*, Springer-Verlag, Berlin, 2010. (Citováno na stranách 258 a 276.)

[302] S. N. Kabadi, A. P. Punnen, *Anti-stalling pivot rule for linear programs with totally unimodular coefficient matrix*, str. 15–20 v „Mathematical Programming and Game Theory for Decision Making“, S. K. Neogy, R. B. Bapat, A. K. Das a T. Parthasarathy (editoři), Stat. Sci. Interdiscip. Res. 1, World Sci. Publ., Hackensack, 2008. (Citováno na straně 161.)

[303] Z. Kadelburg, D. Đukić, M. Lukić, I. Matić, *Inequalities of Karamata, Schur and Muirhead, and some applications*, Teach. Math. **8** (2005), č. 1, str. 31–45. (Citováno na straně 467.)

[304] G. Kalai, *A subexponential randomized simplex algorithm (extended abstract)*, str. 475–482 v „STOC '92: Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing“, Association for Computing Machinery, 1992. (Citováno na straně 240.)

[305] G. Kalai, *Some old and new problems in combinatorial geometry I: around Borsuk's problem*, str. 147–174 v „Surveys in Combinatorics 2015“, A. Czumaj, A. Georgakopoulos, D. Král, V. Lozin a O. Pikhurko (editoři), London Math. Soc. Lecture Note Ser. **424**, Cambridge Univ. Press, Cambridge, 2015. (Citováno na straně 442.)

[306] S. A. Kaliadina, *W. W. Leontief and the Repressions of the 1920s: an Interview*, Economic Systems Research **18** (2006), č. 4, str. 347–355. (Citováno na straně 68.)

[307] S. A. Kaliadina, N. I. Pavlova, *The Family of W. W. Leontief in Russia*, Economic Systems Research **18** (2006), č. 4, str. 335–345. (Citováno na straně 68.)

[308] D. Kalman, *Leveling with Lagrange: an alternate view of constrained optimization*, Math. Mag. **82** (2009), č. 3, str. 186–196. (Citováno na stranách 35 a 36.)

[309] L. V. Kantorovič, *On the translocation of masses*, Management Sci. **5** (1958), str. 1–4. Přeloženo z ruského originálu: C. R. (Doklady) Acad. Sci. URSS (N.S.) **37** (1942), str. 199–201. (Citováno na stranách 65 a 68.)

[310] L. V. Kantorovič, *Mathematical methods of organizing and planning production*, Management Sci. **6** (1960), č. 4, str. 366–422. (Citováno na straně 67.)

[311] L. V. Kantorovič, *The Best Use of Economic Resources*, Pergamon Press, 1965. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/kantorovich1. (Citováno na straně 66.)

[312] H. Karloff, *Linear Programming*, Progress in Theoretical Computer Science, Birkhäuser, Boston, 1991. (Citováno na stranách 245 a 249.)

[313] J. Karamata, *Sur une inégalité rélative aux fonctions convexes* [francouzsky], Publ. Math. Univ. Belgrade **1** (1932), str. 145–148. (Citováno na straně 467.)

[314] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, str. 302–311 v „STOC '84: Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing", Association for Computing Machinery, 1984. (Citováno na stranách 247 a 249.)

[315] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica **4** (1984), č. 4, str. 373–395. (Citováno na stranách 247 a 249.)

[316] E. Katok, W. Tarantino, R. Tiedeman, *Improving Performance and Flexibility at Jeppesen: The World's Leading Aviation-Information Company*, INFORMS Journal on Applied Analytics **31** (2001), str. 7–29. (Citováno na straně 306.)

[317] M. Katz, *Jung's theorem in complex projective geometry*, Quart. J. Math. Oxford **36** (1985), č. 144, str. 451–466. (Citováno na straně 441.)

[318] V. J. Katz, *A History of Mathematics*, třetí vydání, Pearson, Boston, 2009. (Citováno na straně 33.)

[319] M. G. Katz, D. M. Schaps, S. Shnider, *Almost equal: the method of adequality from Diophantus to Fermat and beyond*, Perspectives on Science **21** (2013), č. 3, str. 283–324. (Citováno na straně 26.)

[320] P. Kirchberger, *Über Tchebychefsche Annäherungsmethoden* [německy], Math. Ann. **57** (1903), no. č, str. 509–540. (Citováno na straně 441.)

[321] T. P. Kirkman, *On the representation of polyedra* [sic!], Phil. Trans. Royal Soc. London **146** (1856), str. 413–418. (Citováno na stranách 260 a 267.)

[322] T. H. Kjeldsen, *A contextualized historical analysis of the Kuhn–Tucker Theorem in nonlinear programming: The impact of World War II*, Hist. Math. **27** (2000), č. 4, str. 331–361. (Citováno na stranách 611 a 628.)

[323] T. H. Kjeldsen, *Different motivations and goals in the historical development of the theory of systems of linear inequalities*, Arch. Hist. Exact Sci. **56** (2002), č. 6, str. 469–538. (Citováno na straně 486.)

[324] V. Klee, *A class of linear programming problems requiring a large number of iterations*, Numer. Math. **7** (1965), str. 313–321. (Citováno na straně 237.)

[325] V. Klee, G. J. Minty, *How good is the simplex algorithm?*, str. 159–175 v „Inequalities III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)", Academic Press, New York, 1972. (Citováno na straně 238.)

[326] A. Koberstein, *Progress in the dual simplex algorithm for solving large scale LP problems: techniques for a fast and stable implementation*, Comput. Optim. Appl. **41** (2008), č. 2, str. 185–204. (Citováno na straně 175.)

[327] D. A. Kohler, *Projections of Convex Polyhedral Sets*, Operational Research Center Report, ORC 67-29, University of California, Berkeley, 1967. (Citováno na straně 87.)

[328] G. B. Kolata, *Mathematicians Amazed by Russian's Discovery: An unknown Russian discovers a new solution to linear programming problems. But his result goes unnoticed for months*, Science **206** (1979), č. 4418, str. 545–546. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/science. (Citováno na straně 241.)

[329] G. B. Kolata, *A fast way to solve hard problems: A new algorithm to solve linear programming problems is so fast that experts in the field are taken aback*, Science **225** (1984), str. 1379–1380. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/science2. (Citováno na straně 250.)

[330] B. Kolman, R. E. Beck, *Elementary Linear Programming with Applications*, 2. vydání, Academic Press, San Diego, 1995. (Citováno na stranách 95, 187 a 233.)

[331] T. J. Koopmans, *Activity Analysis of Production and Allocation*, Cowles Commission Monograph **13**, John Wiley & Sons, New York, London, 1951. (Citováno na straně 63.)

[332] T. W. Körner, *The Pleasure of Counting*, Cambridge University Press, Cambdridge, 1996. (Citováno na straně 57.)

[333] M. K. Kozlov, S. P. Tarasov, L. G. Chačijan, *Polynomial solvability of convex quadratic programming* [rusky], Dokl. Akad. Nauk SSSR **248** (1979), č. 5, str. 1049–1051. (Citováno na straně 359.)

[334] E. Kročilová, *Funkcionální rovnice a jejich aplikace*, diplomová práce, Masarykova univerzita, 2024. Dostupné online na adrese https://is.muni.cz/th/bnxxw/. (Citováno na straně 391.)

[335] Z. Křístková, *Simulace dopadu nástrojů zemědělské politiky na českou ekonomiku v kontextu obecné rovnováhy*, disertační práce, Česká zemědělská univerzita v Praze, Praha, 2009. (Citováno na straně 64.)

[336] H. W. Kuhn, R. E. Quandt, *An experimental study of the simplex method*, str. 107–124 v „Proc. Sympos. Appl. Math. Vol. XV", Amer. Math. Soc., Providence, 1963. (Citováno na straně 150.)

[337] H. P. Künzi, W. Krelle, R. von Randow, *Nichtlineare Programmierung*, ve spolupráci s W. Oettlim, druhé přepracované a rozšířené vydání, Hochschultext, Springer-Verlag, Berlin–New York, 1979. (Citováno na stranách 10, 340 a 359.)

[338] S. S. Kutateladze, *Fundamentals of Functional Analysis*, Kluwer Academic Publishers Group, Dordrecht, 1996. (Citováno na straně 482.)

[339] M. Kwan, *Graphic programming using odd or even points* [v čínštině], Acta Math. Sinica **10**, str. 263–266; anglický překlad vyšel Chinese Math. **1** (1960), str. 273–277. (Citováno na straně 259.)

[340] B. Labatut, *Maniac*, Paseka, 2024. (Citováno na straně 65.)

[341] J.-L. Lagrange, *Recherches sur la libration de la Lune, dans lesquelles on tâche de résoudre la question proposéee par l'Académie royale des sciences pour le Prix de l'année 1764*, str. 1–50 v „Recueil des pieces qui ont remporté le Prix de l'Académie royale des sciences", tomme neuvieme, Quí contient les Pieces de 1764, 1765, 1766, 1770 & 1772, Académie royale des sciences, Paříž, 1777. Původní vydání je dostupné online na adrese https://ln.math.muni.cz/pz/opt/lagrange2 a jeho pozdější přetisk na adrese https://ln.math.muni.cz/pz/opt/lagrange3. (Citováno na straně 34.)

[342] J.-L. Lagrange, *Théorie de la libration de la Lune et des autres phénomènes qui dépendent de la figure non sphérique de cette planète*, str. 203–309 v „Nouveaux Mémoires de l'Académie royale des sciences et Belles–Lettres de Berlin", Berlín, 1780. Pozdější přetisk je dostupný online na adrese https://ln.math.muni.cz/pz/opt/lagrange4. (Citováno na straně 34.)

[343] J.-L. Lagrange, *Mécanique Analytique I*, anglický překlad druhého vydání z roku 1811, přeloženo A. Bissonnadem a V. N. Vaglientem, Boston studies in the philosophy of science, Vol. 191, Springer Science+Business Media, Dordrecht, 1997. Originál prvního vydání z roku 1788 je dostupný online na adrese https://ln.math.muni.cz/pz/opt/lagrange5. Druhé vydání z roku 1811 je pak k dispozici na adrese https://ln.math.muni.cz/pz/opt/lagrange6. (Citováno na stranách 34, 35 a 39.)

[344] J.-L. Lagrange, *Théorie des fonctions analytiques, contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits ou d'évanouissans, de limites ou de fluxions, et réduits a l'analyse algébrique des quantités finies*, De l'imprimerie de la République, Paris, 1797. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/lagrange7. (Citováno na straně 35.)

[345] C. Lemaréchal, *Cauchy and the gradient method*, str. 251–254 v „Optimization Stories (21st International Symposium on Mathematical Programming, Berlin, Germany, 2012)", M. Grötschel (editor), Doc. Math. **2012** (2012). (Citováno na straně 572.)

[346] A. H. Land, A. G. Doig, *An automatic method of solving discrete programming problems*, Econometrica **28** (1960), č. 3, str. 497–520. (Citováno na straně 281.)

[347] A. H. Land, A. G. Doig, *A survey of the operational use of ILP models*, Ann. Oper. Res. **149** (2007), str. 147–156. (Citováno na straně 303.)

[348] E. L. Lawler, *The Great Mathematical Sputnik of 1979*, Math. Intelligencer **2** (1980), č. 4, str. 191–198. Pro český překlad viz Pokroky Mat. Fyz. Astronom. **27** (1982), č. 1, str. 39–47. (Citováno na straně 242.)

[349] E. L. Lawler, D. E. Wood, *Branch-and-bound methods: A survey*, Operations Res. **14** (1966), 699–719. (Citováno na straně 281.)

[350] S. R. Lay, *Convex Sets and Their Applications*, Pure & Applied Mathematics, A Wiley–Interscience Series of Texts, Monographs & Tracts, John Wiley & Sons, New York, 1982. (Citováno na straně 441.)

[351] J. Lee, *Classroom Note: Hoffman's circle untangled*, SIAM Rev. **39** (1997), č. 1, str. 98–105. (Citováno na straně 155.)

[352] J. Lee, *A First Course in Combinatorial Optimization*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2004. (Citováno na straně 300.)

[353] C. E. Lemke, *The dual method of solving the linear programming problem*, Naval Res. Logist. Quart. **1** (1954), č. 1, str. 36–47. (Citováno na straně 175.)

[354] L. A. Levin, *Universal enumeration problems* [rusky], Problemy Peredači Informacii **9** (1973), č. 3, str. 115–116. (Citováno na straně 308.)

[355] D. V. Lindley, *Dynamic programming and decision theory*, Appl. Statist. **10** (1961), str. 39–51. (Citováno na straně 390.)

[356] J. D. C. Little, K. G. Murty, D. W. Sweeney, C. Karel, *An algorithm for the traveling salesman problem*, Operations Research **11** (1963), č. 6, str. 972–989. Viz také https://ln.math.muni.cz/pz/opt/BaB1. (Citováno na stranách 266 a 281.)

[357] M. Livio, *Zlatý řez: Příběh fí, nejpodivuhodnějšího čísla na světě*, Argo & Dokořán, 2006. (Citováno na straně 550.)

[358] L. Lovász, *A new linear programming algorithm — better or worse than the simplex method?* Math. Intelligencer **2** (1980), č. 3, str. 141–146. Pro český překlad viz Pokroky Mat. Fyz. Astronom. **26** (1981), č. 4, str. 193–202. (Citováno na straně 242.)

[359] D. G. Luenberger, *Optimization by Vector Space Methods*, Series in Decision & Control, John Wiley & Sons, New York, 1969. (Citováno na straně 618.)

[360] D. G. Luenberger, Y. Ye, *Linear and Nonlinear Programming*, čtvrté vydání, International Series in Operations Research & Management Science Vol. 228, Springer, New York, 2016. (Citováno na stranách 10, 572, 581, 586 a 602.)

[361] M. S. Mahoney, *The Mathematical Career of Pierre de Fermat: 1601-1665*, druhé vydání, Princeton University Press, Princeton, 1994. (Citováno na stranách 26 a 31.)

[362] H. M. Markowitz, *Portfolio selection*, J. Finance **7** (1952), č. 1, str. 77–91. (Citováno na stranách 320 a 325.)

[363] H. M. Markowitz, *The utility of wealth*, The Journal of Political Economy **60** (1952), č. 2, str. 151–158. (Citováno na stranách 320 a 325.)

[364] H. M. Markowitz, *The optimization of a quadratic function subject to linear constraints*, Naval Res. Logist. Quart. **3** (1956), č. 1-2, str. 111–133. (Citováno na straně 341.)

[365] H. M. Markowitz, *The elimination form of the inverse and its application to linear programming*, Management Sci. **3** (1957), č. 3, str. 255–269. (Citováno na straně 320.)

[366] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, Cowles Foundation for Research in Economics at Yale University, John Wiley & Sons, New York, 1959. (Citováno na straně 320.)

[367] H. M. Markowitz, B. Hausner, H. W. Karr, *SIMSCRIPT: A Simulation Programming Language*, Research Memoranda **RM-3310-PR**, The RAND Corporation, Santa Monica, 1962. (Citováno na straně 320.)

[368] I. Maros, *A general Phase-I method in linear programming*, European J. Oper. Res. **23** (1986), č. 1, str. 64–77. (Citováno na straně 167.)

[369] K. T. Marshall, J. W. Suurballe, *A note on cycling in the simplex method*, Naval Res. Logist. Quart. **16** (1969), č. 1, str. 121–137. (Citováno na straně 156.)

[370] S. Martello, P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Chichester, 1990. (Citováno na straně 286.)

[371] C. H. Martin, *Ohio University's College of Business uses integer programming to schedule classes*, INFORMS Journal on Applied Analytics **34** (2004), č. 6, str. 460–465. (Citováno na straně 306.)

[372] P. Massé, *Sur un cas particulier remarquable de la régulation d'un débit aléatoire par un réservoir* [francouzsky], C. R. Acad. Sci. Paris **219** (1944), str. 173–175. (Citováno na straně 368.)

[373] J. Matoušek, *Lineární programování*, Univerzita Karlova, 2006 (elektronicky). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/matousek. (Citováno na straně 150.)

[374] J. Matoušek, B. Gärtner, *Understanding and Using Linear Programming*, Universitext, Springer, 2007. (Citováno na stranách 10, 98, 102, 149, 150 a 220.)

[375] J. Matoušek, M. Sharir, E. Welzl, *A subexponential bound for linear programming*, Algorithmica **16** (1996), č. 4-5, 498–516. (Citováno na straně 240.)

[376] B. A. McCarl, H. Moskowitz, H. Furtan, *Quadratic programming applications*, Omega **5** (1977), č. 1, str. 43–55. (Citováno na straně 328.)

[377] G. McGuire, B. Tugemann, G. Civario, *There is no 16-clue Sudoku: solving the Sudoku minimum number of clues problem via hitting set enumeration*, Exp. Math. **23** (2014), č. 2, str. 190–217. (Citováno na straně 275.)

[378] K. Menger, *Das Botenproblem*, str. 11–12 v „ErgebnisseEines Mathematischen Kolloquiums 2“, D. Egbert a S. Karl (editoři), Teubner, Leipzig, Germany, 1932. (Citováno na straně 262.)

[379] R. E. Miller, P. D. Blair, *Input-Output Analysis: Foundations and Extensions*, druhé vydání, Cambridge University Press, Cambridge, 2009. (Citováno na straně 77.)

[380] H. Minkowski, *Geometrie der Zahlen* [německy], Teubner, Leipzig, 1896. (Citováno na straně 486.)

[381] L. G. Mitten, *Composition principles for synthesis of optimal multistage processes*, Oper. Res. **12** (1964), str. 610–619. (Necitováno.)

[382] R. D. C. Monteiro, I. Adler, *Interior path following primal-dual algorithms II: Convex quadratic programming*, Math. Programming **44** (1989), č. 1, str. 43–66. (Citováno na straně 359.)

[383] R. D. C. Monteiro, I. Adler, M. G. C. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Math. Oper. Res. **15** (1990), č. 2, str. 191–214. (Citováno na straně 359.)

[384] J. J. Moreau, *Fonctionnelles convexes*, Séminaire Jean Leray **2** (1966-1967), str. 1–108. (Citováno na straně 513.)

[385] T. S. Motzkin, *Beiträge zur Theorie der Linearen Ungleichungen*, Bazilejská univerzita, 1933. (Citováno na stranách 69 a 486.)

[386] J. R. Munkres, *Topology*, druhé vydání, Upper Saddle River, Prentice Hall, 2000. (Citováno na straně 440.)

[387] F. H. Murphy, V. Panchanadam, *Understanding linear programming modeling through an examination of the early papers on model formulation*, Oper. Res. **45** (1997), č. 3, str. 341–356. (Citováno na straně 208.)

[388] K. G. Murty, *Computational complexity of complementary pivot methods*, str. 61–73 v „Complementarity and Fixed Point Problems“, M. L. Balinski a R. W. Cottle (editoři), Math. Programming Stud. **7** (1978). (Citováno na straně 239.)

[389] K. G. Murty, *Linear Complementarity, Linear and Nonlinear Programming*, Heldermann Verlag, Berlin, 1988. (Citováno na straně 359.)

[390] K. G. Murty, S. N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Math. Programming **39** (1987), č. 2, str. 117–129. (Citováno na straně 360.)

[391] P. J. Nahin, *When Least Is Best*, Princeton University Text, 2004. ISBN 978-0-691-13052-1. (Citováno na straně 10.)

[392] J. F. Nash Jr., korespondence s Národní bezpečností agenturou v USA z roku 1955. Korespondence byla v roce 2012 odtajněna a originál je dostupný na webu https://ln.math.muni.cz/pz/opt/nash, viz také https://ln.math.muni.cz/pz/opt/nash2. (Citováno na straně 308.)

[393] T. Negishi, *History of Economic Theory*, Advanced Textbooks in Economics **26**, North-Holland, Amsterdam, 1989. (Citováno na straně 62.)

[394] G. L. Nemhauser, *Einführung in die Praxis der dynamischen Programmierung*, R. Oldenbourg Verlag, Mnichov, 1969. (Citováno na straně 390.)

[395] G. L. Nemhauser, L. Wolsey, *Integer and Combinatorial Optimization*, reprint originálu z roku 1988, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, 1999. (Citováno na stranách 285, 300 a 303.)

[396] Y. E. Nesterov, *A method for solving the convex programming problem with convergence rate* $\mathcal{O}(1/k^2)$ [rusky], Dokl. Akad. Nauk SSSR **269** (1983), č. 3, str. 543–547. (Citováno na straně 586.)

[397] Y. E. Nesterov, *Polynomial methods in the linear and quadratic programming* [rusky], Izv. Akad. Nauk SSSR Tekhn. Kibernet. **1988**(1988), č. 3, str. 3–6 a 218. Pro anglický překlad viz Soviet J. Comput. Systems Sci. 26 (1989), č. 5, str. 98–101. (Citováno na straně 359.)

[398] Y. E. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Appl. Optim. **87**, Kluwer Academic Publishers, Boston, 2004. (Citováno na stranách 581 a 586.)

[399] I. Netuka, J. Veselý, *Eduard Helly, konvexita a funkcionální analýza*, Pokroky Mat. Fyz. Astronom. **29** (1984), č. 6, str. 301–312. (Citováno na stranách 440, 441 a 444.)

[400] F. Neuman, *Funkcionální rovnice*, Matematický seminář SNTL, Státní nakladatelství technické literatury, Praha, 1986. (Citováno na straně 391.)

[401] M. Ohm, *Die reine Elementar-Mathematik II: Die ebene Raumgrössenlehre*, druhé vydání, Jonas Berlags–Buchhandlung, Berlin, 1835. (Citováno na straně 549.)

[402] L. T. Oliver, D. J. Wilde, *Symmetric sequential minimax search for a maximum*, Fibonacci Quart. **2** (1964), č. 3, str. 169–175. (Citováno na straně 558.)

[403] J. B. Orlin, *On the simplex algorithm for networks and generalized networks*, str. 166–178 v „Mathematical Programming Essays in Honor of George B. Dantzig, Part I“, R. W. Cottle (editor), Math. Programming Stud. **24** (1985). (Citováno na straně 161.)

[404] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Classics In Applied Mathematics **30**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2000. (Citováno na straně 535.)

[405] B. G. Pachpatte, *Mathematical Inequalities*, North–Holland Mathematical Library **67**, Elsevier, Amsterdam, 2005. (Citováno na stranách 466 a 467.)

[406] M. Padberg, G. Rinaldi, *Optimization of a 532-city symmetric traveling salesman problem by branch and cut*, Oper. Res. Lett. **6** (1987), str. 1–7. (Citováno na straně 303.)

[407] M. Padberg, G. Rinaldi, *A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems*, SIAM Rev. **33** (1991), č. 1, str. 60–100. (Citováno na straně 303.)

[408] Ch. H. Papadimitriou, K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, druhé upravené vydání, Dover Publications, Mineola, 1998. (Citováno na stranách 245 a 303.)

[409] P. M. Pardalos, *Global optimization algorithms for linearly constrained indefinite quadratic problems*, Comput. Math. Appl. **21** (1991), č. 6-7, str. 87–97. (Citováno na straně 360.)

[410] P. M. Pardalos, G. Schnitger, *Checking local optimality in constrained quadratic programming is NP-hard*, Oper. Res. Lett. **7** (1988), č. 1, str. 33–35. (Citováno na straně 360.)

[411] R. G. Parker, R. L. Rardin, *Discrete Optimization*, Computer Science and Scientific Computing, Academic Press, Boston, 1988. (Citováno na stranách 300 a 304.)

[412] A. Phillips, *The Tableau Économique as a Simple Leontief Model*, The Quarterly Journal of Economics **69** (1955), str. 137–144. (Citováno na straně 62.)

[413] A. B. Philpott, *Linear programming made simpler than simplex*, New Scientist **1433** (1984), str. 20. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/newscientist. (Citováno na straně 250.)

[414] J. Plesník, J. Dupačová, M. Vlach, *Lineárne programovanie*, Alfa, vydavateľstvo technickej a ekonomickej literatúry, Bratislava, 1990. (Citováno na stranách 187, 198 a 206.)

[415] H. Poincaré, *Second complement a l'analysis situs*, Proc. Lond. Math. Soc. **32** (1900), str. 277–308. Text byl také publikován v souhrnném díle od A. Appella a G. Darbouxe, „Œuvres de Henri Poincaré VI“, Gauthier–Villars, Paříž, 1916, str. 338–370, viz online https://ln.math.muni.cz/pz/opt/poincare1. (Citováno na straně 276.)

[416] T. Popoviciu, *Les Fonctions Convexes* [francouzsky], Actualités Scientifiques et Industrielles **992**, Hermann & Cie, Paříž, 1944. (Citováno na straně 452.)

[417] T. Popoviciu, *Sur certaines inégalités qui caractérisent les fonctions convexes* [rumunsky], An. Şti. Univ. „Al. I. Cuza“ Iaşi Secţ. I a Mat. (N.S.) **11B** (1965), str. 155–164. (Citováno na straně 467.)

[418] V. Pourová, A. Jakešová, *O výživě*, Pointa, Praha, 2019. (Citováno na straně 208.)

[419] M. J. D. Powell, *Convergence properties of algorithms for nonlinear optimization*, SIAM Review **28** (1986), str. 487–500. (Citováno na straně 604.)

[420] M. J. D. Powell, *Karmarkar's algorithm: a view from nonlinear programming*, Bull. Inst. Math. Appl. **26** (1990), č. 8-9, str. 165–181. (Citováno na straně 249.)

[421] M. J. D. Powell, *On the number of iterations of Karmarkar's algorithm for linear programming*, Math. Programming **62** (1993), č. 1, str. 153–197. (Citováno na straně 249.)

[422] S. D. Promislow, *A First Course in Functional Analysis*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, A Wiley–Interscience Publication, John Wiley & Sons, Hoboken, 2008. (Citováno na straně 483.)

[423] M. Propato, J. G. Uber, *Booster system design using mixed–integer quadratic programming*, Journal of Water Resources Planning and Management **130** (2004), č. 4, str. 348–352. (Citováno na straně 328.)

[424] A. P. Punnen, *The traveling salesman problem: applications, formulations and variations*, str. 1–28 v „The Traveling Salesman Problem and its Variations", G. Gutin a A. P. Punnen (editoři), Combinatorial Optimization **12**, Kluwer Academic Publishers, Dordrecht, 2002. (Citováno na straně 258.)

[425] J. Radon, *Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten* [německy], Math. Ann. **83**, č. 1-2, str. 113–115. (Citováno na stranách 438 a 440.)

[426] A. M. Raĭgorodskiĭ, *The Borsuk partition problem: the seventieth anniversary*, Math. Intelligencer **26** (2004), č. 3, str. 4–12. (Citováno na straně 442.)

[427] S. S. Rao, *Engineering Optimization: Theory and Practice*, čtvrté vydání, John Wiley & Sons, New Jersey, 2009. (Citováno na straně 10.)

[428] N. V. Reinfeld, W. R. Vogel, *Mathematical Programming*, Prentice Hall, Englewood Cliffs, 1958. (Citováno na straně 196.)

[429] F. Ricca, A. Scozzari, B. Simeone, *Political districting: from classical models to recent approaches*, 4OR **9** (2011), č. 3, str. 223–254. (Citováno na straně 272.)

[430] N. Robertson, P. Seymour, *Graph Minors XX: Wagner's conjecture*, Journal of Combinatorial Theory Series B **92** (2004), č. 2, str. 325–357. (Citováno na straně 308.)

[431] J. Robinson, *On the Hamiltonian game (a traveling salesman problem)*, Research Memoranda **RM-303**, The RAND Corporation, Santa Monica, 1949. (Citováno na straně 263.)

[432] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, New Jersey, 1970. (Citováno na stranách 58, 446 a 513.)

[433] R. T. Rockafellar, *Lagrange multipliers and optimality*, SIAM Review **35** (1993), č. 2, str. 183–238. (Citováno na straně 59.)

[434] L. J. Rogers, *An extension of a certain theorem in inequalities*, Mess. **XVII** (1888), č. 10, str. 145–150. (Citováno na straně 463.)

[435] N. Rose, *Mathematical Maxims and Minims*, Rome Press, Raleigh, 1988. (Citováno na stranách 19 a 24.)

[436] J. Rosenhouse, L. Taalman, *Taking Sudoku Seriously: The Math Behind the World's Most Popular Pencil Puzzle*, Oxford University Press, Oxford, 2011. (Citováno na straně 275.)

[437] B. Russell, *Mysticism and Logic: And Other Essays*, The Mayflower Press, London, 1949. (Citováno na straně 25.)

[438] J. Rust, *Dynamic programming*, v „The New Palgrave Dictionary of Economics", M. Palgrave (editor), Palgrave Macmillan, London, 2008. (Citováno na straně 368.)

[439] S. Sahni, *Computationally related problems*, SIAM J. Comput. **3** (1974), str. 262–279. (Citováno na straně 360.)

[440] H. Sachs, M. Stiebitz, R. J. Wilson, *An historical note: Euler's Königsberg letters*, J. Graph Theory **12** (1988), č. 1, str. 133–139. (Citováno na stranách 258 a 259.)

[441] K. S. Sarkaria, *Tverberg's theorem via number fields*, Israel J. Math. **79** (1992), č. 2-3, str. 317–320. (Citováno na straně 439.)

[442] F. J. Scott, *The Correspondence of Isaac Newton IV (1694–1709)*, The Syndics of the Cambdridge University Press, Cambridge, 1967. (Citováno na straně 50.)

[443] P. D. Seymour, *Decomposition of regular matroids*, J. Combin. Theory Ser. B **28** (1980), č. 3, str. 305–359. (Citováno na straně 276.)

[444] M. Shaw, *Review of computational experience in solving large mixed integer programming problems*, str. 406–412 v „Applications of Mathematical Programming Techniques", E. M. L. Beale (editor), English Universities Press, London, 1970. (Citováno na straně 306.)

[445] C. M. Shetty, *A simplified procedure for quadratic programming*, Operations Res. **11** (1963), 248–260. (Citováno na straně 348.)

[446] T. C. Schelling, *The Strategy of Conflict*, Harvard University Press, Cambridge, 1980. (Citováno na straně 307.)

[447] A. Schrijver, *The new linear programming method of Karmarkar*, CWI Newslett. **8** (1985), str. 2–14. (Citováno na stranách 249 a 250.)

[448] A. Schrijver, *Theory of Linear and Integer Programming*. Wiley–Interscience Series in Discrete Mathematics, A Wiley–Interscience Publication, John Wiley & Sons, Chichester, 1986. (Citováno na stranách 300 a 303.)

[449] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Algorithms and Combinatorics **24**, Springer-Verlag, Berlin Heidelberg, 2004. (Citováno na stranách 258 a 263.)

[450] A. Schrijver, *On the history of combinatorial optimization (till 1960)*, str. 1–68 v „Discrete Optimization", K. Aardal, G. L. Nemhauser a R. Weismantel (editoři), Handbooks in Operations Research and Management Science **12**, Elsevier Science, Amsterdam, 2005. (Citováno na straně 258.)

[451] B. V. Shah, R. J. Buehler, O. Kempthorne, *Some algorithms for minimizing a function of several variables*, J. Soc. Indust. Appl. Math. **12** (1964), str. 74–92. (Citováno na straně 587.)

[452] S. M. Sinha, *Mathematical Programming – Theory and Methods*, Elsevier, Delhi, 2006. (Citováno na stranách 10 a 359.)

[453] R. Shamir, *The efficiency of the simplex method: a survey*, Management Sci. **33** (1987), č. 3, str. 301–334. (Citováno na straně 149.)

[454] H. J. S. Smith, *On systems of linear indeterminate equations and congruences*, Philos. Trans. Roy. Soc. London Ser. A **151** (1861), str. 293–326. (Citováno na straně 279.)

[455] D. E. Smith, *A Source Book in Mathematics II*, Dover Publications, New York, 1959. (Citováno na straně 50.)

[456] M. Sniedovich, *Dynamic Programming*, Pure and Applied Mathematics **154**, Marcel Dekker, New York, 1992. (Citováno na straně 390.)

[457] D. A. Spielman, S.-H. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, J. ACM **51** (2004), č. 3, str. 385–463. (Citováno na stranách 239 a 240.)

[458] A. E. Steenge, R. van den Berg, *Transcribing the Tableau Économique: input-output analysis à la Quesnay*, Journal of the History of Economic Thought **29** (2007), str. 331–358. (Citováno na straně 62.)

[459] E. Steinitz, *Bedingt konvergente Reihen und konvexe Systeme* [německy], J. Reine Angew. Math. **143** (1913), str. 128–176. (Citováno na straně 436.)

[460] E. Steinitz, *Bedingt konvergente Reihen und konvexe Systeme* [německy], J. Reine Angew. Math. **144** (1914), str. 1–40. (Citováno na straně 436.)

[461] E. Steinitz, *Bedingt konvergente Reihen und konvexe Systeme* [německy], J. Reine Angew. Math. **146** (1916), str. 1–52. (Citováno na straně 436.)

[462] E. Stiemke, *Über positive Lösungen homogener linearer Gleichungen* [německy], Math. Ann. **76** (1915), str. 340–342. (Citováno na straně 486.)

[463] G. J. Stigler, *The cost of subsistence*, Journal of Farm Economics **27** (1945), č. 2, str. 303–314. (Citováno na straně 71.)

[464] O. Stolz, *Grundzüge der Differential- und Integralrechnung. Erster Teil: Reelle Veränderliche und Functionen* [německy], B. G. Teubner, Leipzig, 1893. (Citováno na straně 452.)

[465] P. Strømholm, *Fermat's methods of maxima and minima and of tangents: A reconstruction*, Arch. History Exact Sci. **5** (1968), no. 1, str. 47–69. (Citováno na straně 26.)

[466] A. Sucharev, A. Timochov, V. Fedorov, *Kurs metodov optimizacii* [rusky], Nauka, Moskva, 1986. (Citováno na stranách 10, 643 a 666.)

[467] W. Sun, Y.-X. Yuan, *Optimization Theory and Methods – Nonlinear Programming*, Springer, New York, 2006. (Citováno na stranách 10 a 579.)

[468] G. Sierksma, Y. Zwols, *Linear and Integer Optimization: Theory and Practice*, třetí vydání, CRC Press, Boca Raton, 2015. (Citováno na straně 163.)

[469] N. Z. Šor, *Cut-off method with space extension in convex programming problems* [rusky], Kibernetika (Kiev) 1977, č. 1, str. 94–95. Pro anglický překlad viz Cybernetics **13** (1977), č. 1, str. 94–96. (Citováno na straně 241.)

[470] K. Sydsæter, P. Hammond, A. Seierstad, A. Strøm, *Further Mathematics for Economic Analysis*, druhé vydání, Prentice Hall, Harlow, 2008. (Citováno na straně 466.)

[471] L. Taalman, *Taking Sudoku seriously*, Math Horiz. **15** (2007), str. 5–9. (Citováno na straně 275.)

[472] H. A. Taha, *Integer Programming: Theory, Applications and Computations*, Operations Research and Industrial Engineering, Academic Press, New York, 1975. (Citováno na straně 303.)

[473] B. W. Taylor, *Introduction to Management Science*, 10. vydání, Prentice Hall, Upper Saddle River, 2010. (Citováno na straně 187.)

[474] R. L. Teixeira, *Quadratic Programming as an Extension of Linear Programming*, diplomová práce, University of Cape Town, Kapské Město, 1968. (Citováno na stranách 346 a 359.)

[475] J. Tenne, S. W. Armfield, *A note on the relation of Gaussian elimination to the conjugate directions algorithm*, ANZIAM J. **46** (2005), str. C971–C986. (Citováno na straně 597.)

[476] T. Terlaky, *A convergent criss-cross method*, Optimization **16** (1985), č. 5, str. 683–690. (Citováno na straně 161.)

[477] T. Terlaky, S. Zhang, *Pivot rules for linear programming: A survey on recent theoretical developments*, Ann. Oper. Res. **46** (1993), str. 203–233. (Citováno na straně 149.)

[478] P. R. Thie, G. E. Keough, *An Introduction to Linear Programming and Game Theory*, třetí vydání, Wiley, 2008. (Citováno na straně 233.)

[479] W. Thomson, *Popular Lectures and Addresses II*, Macmillan & Co., London, 1894. Úryvek o isoperimetrických problémech na str. 571–592 je dostupný také online na adrese https://ln.math.muni.cz/pz/opt/kelvin1. (Citováno na straně 52.)

[480] M. J. Todd, *Polynomial algorithms for linear programming*, str. 49–66 v „Advances in Optimization and Control (Proceedings of the Conference „Optimization Days 86“ Held at Montreal, Canada, 1986)“, H. A. Eiselt a G. Pederzoli (editoři), Lecture Notes in Econom. and Math. Systems **302**, Springer, Berlin, 1988. (Citováno na straně 249.)

[481] M. J. Todd, *The many facets of linear programming*, Math. Program. Ser. B. **91** (2002), č. 3, str. 417–436. (Citováno na straně 251.)

[482] A. N. Tolstoĭ, *Metody nakhozhdeniya naimen'shego summovogo kilometrazha pri planirovanii perevozok v prostranstve* [rusky; Metody pro nalezení nejmenšího počtu najetých kilometrů při plánování nákladní přepravy v prostoru (Methods of finding the minimal total kilometrage in cargo-transportation planning in space)], str. 23–55 v „Planirovanie Perevozok, Sbornik pervyĭ“ [rusky; Plánování dopravy, Díl I (Transportation Planning, Volume I)], Transpechat' NKPS [TransPress of the National Commissariat of Transportation], Moskva, 1930. (Citováno na straně 68.)

[483] A. N. Tolstoĭ, *Metody ustraneniya neratsional'nykh perevozok pri planirovanii* [rusky; Metody eliminace iracionální dopravy při plánování (Methods of removing irrational transportation in planning)], Sotsialisticheskiĭ Transport **9** (1939), str. 28–51. Také publikováno jako „pamflet“: *Metody ustraneniya neratsional'nykh perevozok pri sostavlenii operativnykh planov* [rusky; Metody eliminace iracionální dopravy při konstrukci operativních plánů (Methods of Removing Irrational Transportation in the Construction of Operational Plans)], Transzheldorizdat, Moskva, 1941. (Citováno na straně 68.)

[484] J. A. Tomlin, *An improved branch-and-bound method for integer programming*, Operations Res. **19** (1971), str. 1070–1075. (Citováno na straně 281.)

[485] K. Truemper, *Algebraic characterizations of unimodular matrices*, SIAM J. Appl. Math. **35** (1978), č. 2, str. 328–332. (Citováno na straně 275.)

[486] H. Tverberg, *A generalization of Radon's theorem*, J. London Math. Soc. **41** (1966), str. 123–128. (Citováno na straně 439.)

[487] S. Ulam, *Marian Smoluchowski and the theory of probabilities in physics*, American Journal of Physics **25** (1957), str. 475–481. (Citováno na straně 275.)

[488] A. F. Veinott Jr., G. B. Dantzig, *Integral extreme points*, SIAM Rev. **10** (1968), str. 371–372. (Citováno na straně 278.)

[489] C. J. de la Vallée Poussin, *Sur la méthode de l'approximation minimum*, Ann. Soc. Sci. Bruxelles **3**5 (1911), str. 1–16. (Citováno na straně 69.)

[490] R. Vandebril, M. Van Barel, N. Mastronardi, *Matrix Computations and Semiseparable Matrices I: Linear Systems*, Johns Hopkins University Press, Baltimore, 2008. (Citováno na straně 275.)

[491] R. J. Vanderbei, *Linear Programming: Foundations and Extensions*, třetí vydání, Springer, Boston, 2008. (Citováno na stranách 152, 158, 229 a 233.)

[492] R. J. Vanderbei, J. C. Lagarias, *I. I. Dikin's convergence result for the affine-scaling algorithm*, str. 109–119 v „Mathematical Developments Arising from Linear Programming (Proceedings of the Joint Summer Research Conference held at Bowdoin College, Brunswick, 1988)“, J. C. Lagarias a M. J. Todd (editoři), Contemp. Math. **114**, Amer. Math. Soc., Providence, 1990. (Citováno na straně 251.)

[493] R. J. Vanderbei, M. S. Meketon, B. A. Freedman, *A modification of Karmarkar's linear programming algorithm*, Algorithmica **1** (1986), č. 4, str. 395–407. (Citováno na straně 251.)

[494] S. A. Vavasis, *Quadratic programming is in NP*, Inform. Process. Lett. **36** (1990), č. 2, str. 73–77. (Citováno na straně 360.)

[495] S. A. Vavasis, *Nonlinear Optimization: Complexity Issues*, International Series of Monographs on Computer Science **8**, The Clarendon Press, Oxford University Press, New York, 1991. ISBN 0-19-507208-1. (Citováno na straně 360.)

[496] O. Veblen, P. Franklin, *On matrices whose elements are integers*, Ann. of Math. (2) **23** (1921), č. 1, str. 1–15. (Citováno na straně 279.)

[497] A. Vershik, *L. V. Kantorovich and linear programming*, online, arXiv:0707.0491, 2007. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/vershik1. (Citováno na straně 67.)

[498] J. Ville, *Sur la théorie générale des jeux où intervient l'habileté des joueurs* [francouzsky], str. 105–113 v „Traité du calcul des probabilités et de ses applications IV, 2“, E. Borel (editor), Gauthier–Villars, Paříž, 1938. (Citováno na straně 486.)

[499] S. Vojíř, *Leonid Vitalievič Kantorovich*, dostupné online na adrese https://ln.math.muni.cz/pz/opt/kantorovich2. (Citováno na straně 66.)

[500] S. Vojíř, *Tjalling Charles Koopmans*, dostupné online na adrese https://ln.math.muni.cz/pz/opt/koopmans1. (Citováno na straně 69.)

[501] J. von Neumann, *Zur Theorie der Gesellschaftsspiele* [německy], Math. Ann. **100** (1928), str. 295–320. (Citováno na straně 658.)

[502] J. von Neumann, *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes* [německy], v „Ergebnisse eines mathematischen Kolloquiums 1935–36“, K. Menger (editor), 1937. Pro anglický překlad s názvem „A model of general economic equilibrium“ viz Review of Economic Studies **13** (1945), str. 1–9. (Citováno na straně 62.)

[503] J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1944. (Citováno na stranách 368 a 486.)

[504] A. Wald, *Sequential Analysis*, Wiley & Sons, New York, 1947. (Citováno na straně 368.)

[505] S. Walukiewicz, *Integer Programming*, Mathematics and its Applications (East European Series) **46**, Kluwer Academic Publishers Group, Dordrecht; PWN–Polish Scientific Publishers, Varšava, 1991. (Citováno na straně 285.)

[506] D. Watson, *A refinement of theorems of Kirchberger and Carathéodory*, J. Austral. Math. Soc. **15** (1973), str. 190–192. (Citováno na straně 436.)

[507] D. de Werra, *On some characterisations of totally unimodular matrices*, Math. Programming **20** (1981), č. 1, str. 14–21. (Citováno na straně 275.)

[508] R. S. Westfall, *Never at Rest: A Biography of Isaac Newton*, Cambridge University Press, Cambridge, 2010. ISBN 978-0-521-23143-5. (Citováno na straně 51.)

[509] A. N. Whitehead, *Science and the Modern World*, The New American Library of World Literature, New York, 1948. (Citováno na straně 25.)

[510] D. J. Wilde, *Optimum Seeking Methods*, Prentice-Hall, Englewood Cliffs, 1965. (Citováno na stranách 557, 558, 559 a 561.)

[511] H. P. Williams, *Fourier's method of linear programming and its dual*, Amer. Math. Monthly **93** (1986), č. 9, str. 681–695. (Citováno na straně 69.)

[512] P. S. Wolfe, *The simplex method for quadratic programming*, Econometrica **27** (1959), str. 382–398. (Citováno na stranách 340, 341 a 342.)

[513] P. S. Wolfe, *Convergence conditions for ascent methods II: Some corrections*, SIAM Rev. **13** (1971), str. 185–188. (Citováno na straně 579.)

[514] R. D. Young, *A primal (all-integer) integer programming algorithm*, J. Res. Nat. Bur. Standards Sect. B **69B** (1965), str. 213–250. (Citováno na straně 303.)

[515] R. D. Young, *A simplified primal (all-integer) integer programming algorithm*, Operations Res. **16** (1968), str. 750–782. (Citováno na straně 303.)

[516] D. B. Yudin, A. S. Nemirovski, *Informational complexity and effective methods for the solution of convex extremal problems* [rusky], Èkonom. i Mat. Metody **12** (1976), č. 2, str. 357–369. Pro anglický překlad viz Matekon: Translations of Russian and East European Math. Economics **13** (1977), str. 25–45. (Citováno na straně 241.)

[517] Y. Yang, *A collection of cycling problems in linear programming*, online, arXiv:2101.01805v3, 2021. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/yang. (Citováno na straně 155.)

[518] Y. Ye, *An extension of Karmarkar's algorithm and the trust region method for quadratic programming*, str. 49–63 v „Progress in Mathematical Programming: Interior-Point and Related Methods" (sborník z konference konané v roce 1987 v Pacific Grove), N. Megiddo (editor), Springer, New York, 1989. (Citováno na straně 359.)

[519] N. Zadeh, *A bad network problem for the simplex method and other minimum cost flow algorithms*, Math. Programming **5** (1973), str. 255–266. (Citováno na straně 239.)

[520] N. Zadeh, *What is the worst case behavior of the simplex algorithm?*, str. 131–143 v „Polyhedral Computation", D. Avis, D. Bremner a A. Deza (editoři), CRM Proc. Lecture Notes **48**, Amer. Math. Soc., Providence, 2009. (Citováno na stranách 239 a 240.)

[521] G. M. Ziegler, *Typical and extremal linear programs*, str. 217–230 v „The Sharpest Cut: The Impact of Manfred Padberg and His Work", MPS/SIAM Ser. Optim., SIAM, Philadelphia, 2004. (Citováno na straně 240.)

[522] G. M. Ziegler, *Who solved the Hirsch conjecture?*, str. 75–85 v „Optimization Stories (21st International Symposium on Mathematical Programming, Berlin, Germany, 2012)", M. Grötschel (editor), Doc. Math. **2012** (2012). (Citováno na straně 69.)

[523] S. Zhang, *On anti-cycling pivoting rules for the simplex method*, Oper. Res. Lett. **10** (1991), č. 4, str. 189–192. (Citováno na straně 161.)

[524] S. Zhang, *New variants of finite criss-cross pivot algorithms for linear programming*, Eur. J. Oper. Res. **116** (1999), č. 3, str. 607–614. (Citováno na straně 161.)

[525] P. Zörnig, *Systematic construction of examples for cycling in the simplex method*, Comput. Oper. Res. **33** (2006), č. 8, str. 2247–2262. (Citováno na straně 155.)

[526] P. Zörnig, *A note on cycling LP examples with permutation structure*, Comput. Oper. Res. **35** (2008), č. 3, str. 994–1002. (Citováno na straně 155.)

[527] M. Zoubková, J. Zuzaňáková, P. Zemánek, *Sbírka řešených příkladů z matematického programování*, dostupné online na adrese https://ln.math.muni.cz/pz/opt/sbirkaMP. (Citováno na stranách 11 a 44.)

[528] J. Zuzaňáková, *Lineární programování v praxi*, diplomová práce, Masarykova univerzita, 2020. Dostupné online na adrese https://is.muni.cz/th/i23nz. (Citováno na stranách 10, 208 a 211.)

[529] P. Ženčák, *Lineární programování*, Univerzita Palackého v Olomouci, Olomouc, 2013. (Citováno na stranách 81 a 211.)

[530] *Acta Eruditorum*, ročník 1696. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/ae96. (Citováno na straně 48.)

[531] *Acta Eruditorum*, ročník 1697. Dostupné online na adrese https://ln.math.muni.cz/pz/opt/ae97. (Citováno na stranách 48 a 50.)

[532] *A RUSSIAN'S SOLUTION IN MATH QUESTIONED; Americans Who Studied Khachian Linear Programming Method Express Doubt on Scope Effect Would Be Widespread 'Emphatically Not Solved' Critique Explains Significance What Khachian Actually Showed*, The New York Times (March 21, 1980). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/NYT3. (Citováno na straně 243.)

[533] *Bible*, kniha *Genesis* **3**, Český ekumenický překlad, dostupné online na adrese https://ln.math.muni.cz/pz/opt/bible1. (Citováno na straně 54.)

[534] *Bible*, kniha *Genesis* **41**, Český ekumenický překlad, dostupné online na adrese https://ln.math.muni.cz/pz/opt/bible2. (Citováno na straně 54.)

[535] *Doporučené denní dávky vitamínů a minerálů* [elektronicky], 2017. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/LPjidelnicek2. (Citováno na straně 208.)

[536] *Fit Výživa Rob* [elektronicky], 2015. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/LPjidelnicek4. (Citováno na straně 209.)

[537] *From the archive, 29 October 1979: Russian Way with the Mathematical Travelling Salesman*, The Guardian. Dostupné na webu https://ln.math.muni.cz/pz/opt/guardian. (Citováno na straně 241.)

[538] *Kalorické tabulky* [elektronicky], 2020. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/LPjidelnicek3. (Citováno na straně 209.)

[539] *Lady's and gentleman's diary for the year of our Lord 850*, The Company of Stationers, London, 1850. Dostupné na webu https://ln.math.muni.cz/pz/opt/diary1. (Citováno na straně 265.)

[540] *Leonid Vitaliyevich Kantorovich – Autobiography* [elektronicky], NobelPrize.org, Nobel Prize Outreach, 2022. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/kantorovich3 (Citováno na straně 68.)

[541] *Obchod Tesco online* [elektronicky], 2020. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/LPjidelnicek5. (Citováno na straně 209.)

[542] *Optimal Road Trips: History of the TSP with road distances*, Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/TSP1. (Citováno na straně 258.)

[543] *Rozhovor George B. Dantziga s Irvem Lustigem z 5. března 2001*, INFORMS History & Traditions, New York. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/dantzig3. (Citováno na straně 61.)

[544] *Report on the Twenty-Seventh Meeting of the British Association for the Advancement of Science; Held at Dublin in August and September 1857*, John Murray, London, 1858. (Citováno na straně 261.)

[545] *Rozhovor Ralpha E. Gomoryho s Williamem Thomasem z 19. června 2010*, Niels Bohr Library & Archives, American Institute of Physics, College Park. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/gomory1. (Citováno na straně 255.)

[546] *Rozhovor Ralpha E. Gomoryho s Irvem Lustigem z 10. května 2017*, INFORMS History & Traditions, New York. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/gomory2. (Citováno na stranách 255 a 291.)

[547] *Russian Assails Soviet Economics*, The New York Times (May 29, 1959). Dostupné online na adrese https://ln.math.muni.cz/pz/opt/NYT5. (Citováno na straně 70.)

[548] http://mathshistory.st-andrews.ac.uk/ (Citováno na straně 10.)

[549] *Dynamic Programming*, ACME. Citováno 4. června 2025. Dostupné na webu https://ln.math.muni.cz/pz/opt/DPmarriage (Citováno na straně 386.)

Zdroje převzatých obrázků

Strana 15 ..... Zdroj: https://www.sciencephoto.com/media/82984/view

Strana 19 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Carl_Friedrich_Gauss.jpg

Strana 20 . Zdroj: https://mathshistory.st-andrews.ac.uk/Biographies/Titchmarsh/pictdisplay/

Strana 20 ..... nedohledáno

Strana 21 ..... Zdroj: https://upload.wikimedia.org/wikipedia/commons/e/ee/Emil_du_Bois-Reymond.jpg

Strana 22 ..... Zdroj: https://albert.ias.edu/entities/archivalmaterial/a265ccb2-76a3-477e-830d-8e9f2bdad06b

Strana 22 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Niccol%C3%B2_Tartaglia.jpg

Strana 22 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Gerolamo_Cardano_(colour).jpg

Strana 23 ..... Zdroj: https://mathshistory.st-andrews.ac.uk/Biographies/Fermat/

Strana 23 ..... Zdroj: https://greatestgreeks.wordpress.com/2017/04/04/diophantus/

Strana 24 ..... Zdroj: https://sarah-kaiser.com/artwork/1297713_Hermann_Weyl.html

Strana 25 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Galileo_facing_the_Roman_Inquisition.jpg

Strana 29 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Marin_mersenne.jpg

Strana 29 ..... Zdroj: https://de.wikipedia.org/wiki/Fermat-Punkt

Strana 30 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Descartes2.jpg

Strana 30 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Linear_regression.svg

Strana 35 Zdroj: https://commons.wikimedia.org/wiki/File:Evangelista_Torricelli_by_Lorenzo_Lippi_(circa_1647,_Galleria_Silvano_Lodi_%26_Due).jpg

Strana 35 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Cavalieri%27s_Principle_in_Coins.JPG

Strana 36 ..... Zdroj: https://de.wikipedia.org/wiki/Datei:Karl_Weierstrass_1.jpg

Strana 37 ..... Zdroj: https://mathshistory.st-andrews.ac.uk/Miller/images/lagrange.jpg

Strana 48 ..... Zdroj: https://commons.wikimedia.org/wiki/File:Johann_Bernoulli2.jpg

Strana 51 ..... Zdroj: http://www.sothebys.com/en/auctions/ecatalogue/2014/old-master-british-paintings-day-l14037/lot.193.html

Strana 52 Zdroj: https://commons.wikimedia.org/wiki/File:Plan_de_Paris_1223_BNF07710747.png

Strana 52 . . . . . Zdroj: https://cs.m.wikipedia.org/wiki/Soubor:1575_Belleforest_x.jpg

Strana 52 . . . . . Zdroj: https://commons.wikimedia.org/wiki/File:1574_Mediolanum.jpg

Strana 52 . . . . . Zdroj: https://pages.uoregon.edu/dluebke/Reformations441/Frankfurt1574.jpg

Strana 58 . . . . . Zdroj: http://mathshistory.st-andrews.ac.uk/Biographies/Motzkin.html

Strana 58 . . . . . Zdroj: https://commons.wikimedia.org/wiki/File:Werner_Fenchel.jpeg

Strana 59 . . . . . Zdroj: https://www.convexoptimization.com/wikimization/index.php/Rockafellar

Strana 63 . . . . . Zdroj: https://gallica.bnf.fr/ark:/12148/bpt6k106140h/

Strana 63 . . . . . Zdroj: https://gallica.bnf.fr/ark:/12148/bpt6k106140h/

Strana 69 . . . . . Zdroj: https://www.econlib.org/library/Enc/bios/Koopmans.html

Strana 78 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Albert_W._Tucker.gif

Strana 79 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Harold_W._Kuhn.jpg

Strana 79 . . . . . Zdroj: https://miro.medium.com/max/840/1*GYw6bp17svYihnJhkLUchg.jpeg

Strana 80 . . . . . Zdroj: https://www.math.spbu.ru/user/jvr/LVK_html/70_dkk.html

Strana 80 . . . . . nedohledáno

Strana 104 . . . . . Zdroj: https://www.hpmuseum.org/adverts/41nasaad.jpg

Strana 156 . . Zdroj: https://www.engineering.cornell.edu/faculty-directory/robert-gary-bland

Strana 159 . . . . . Zdroj: https://www.af.mil/News/Article-Display/Article/123803/analysts-recall-pioneering-scientist/

Strana 162 . . . . . Zdroj: https://dynaprog.wordpress.com/2014/09/05/abraham-charnes/

Strana 164 . . . . . Zdroj: https://commons.wikimedia.org/wiki/File:Chvatal-kyoto2007-costume-head.png

Strana 176 . . . . . nedohledáno

Strana 237 . . . . . Zdroj: https://old.maa.org/press/periodicals/convergence/whos-that-mathematician-paul-r-halmos-collection-page-35

Strana 340 . . . . . Zdroj: https://pubsonline.informs.org/do/10.1287/orms.2017.01.21in/full/

Strana 367 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Richard_Ernest_Bellman.jpg

Strana 442 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Johann_Radon_cr.jpg

Strana 444 . . . . . Zdroj: https://mathshistory.st-andrews.ac.uk/Biographies/Helly/poster/lived/

Strana 488 . . . . Zdroj: https://en.wikipedia.org/wiki/Gyula_Farkas_(natural_scientist)#/media/File:Farkas_Gyula_(1847-1930).JPG

Strana 488 . . . . . Zdroj: https://en.wikipedia.org/wiki/Hermann_Minkowski#/media/File:De_Raum_zeit_Minkowski_Bild_(cropped).jpg

Strana 594 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Magnus_Hestenes.jpg

Strana 594 Zdroj: https://en.wikipedia.org/wiki/File:Eduard_Stiefel_ETH-Bib_Portr_00817.jpg

Strana 594 . . . . . Zdroj: https://en.wikipedia.org/wiki/File:Lanczos_Kornel_photo_in_1947.jpg

Strana 627 . . . . . Zdroj: https://commons.wikimedia.org/wiki/File:Fritz_John.jpg