M5170: Matematické programování

Petr Zemánek (Masarykova Univerzita, Brno)

Kapitola 4: Základy matematického programování

(verze: 10. listopadu 2020)

Nyní se již dostáváme k úlohám matematického programování (a jejich řešení), tj.

$$f(x) \to \min, \quad x \in X, \tag{4.1}$$

kde *přípustná množina* X je zadána systémem rovností a nerovností

$$X := \left\{ x \in P \subseteq \mathbb{R}^n \mid g_i(x) \leqslant 0,\ g_j(x) = 0,\ i = 1, \dots, k,\ j = k+1, \dots, m \right\}, \tag{4.2}$$

tj.

$$X = P \bigcap \left\{ x \in \mathbb{R}^n \mid g_i(x) \leqslant 0,\ i = 1, \dots, k \right\} \bigcap \left\{ x \in \mathbb{R}^n \mid g_j(x) = 0,\ j = k+1, \dots, m \right\},$$

přičemž funkce $f, g_1, \dots, g_m$ nemusí být nutně diferencovatelné. Omezení $x \in P \neq \emptyset$ se nazývá *přímé* (typicky: $\mathbb{R}^n$, $\mathbb{R}^n_+$ nebo $\mathbb{R}^n_{++}$) a omezení určená funkcemi $g_1, \dots, g_m$ se nazývají *funkcionální* (BÚNO lze uvažovat úlohy s omezeními výhradně ve tvaru rovností/nerovností $\rightsquigarrow$ výhodné?).

Nejdříve se podíváme na obecnou úlohu MP, tj. na úlohu (4.1) s (blíže nespecifikovanou) množinou $X \subseteq \mathbb{R}^n$ a funkcí $f : X \to \mathbb{R}$. Zaveďme si dvě množiny

$$\mathcal{V}(x^*, X) := \Big\{h \in \operatorname{Lin} X \;\Big|\; \exists\, \alpha_0 > 0 : \; x^* + th \in X \;\text{ pro }\; \forall t \in (0, \alpha_0)\Big\},$$

$$\mathcal{U}(x^*, f) := \Big\{h \in \operatorname{Lin} X \;\Big|\; \exists\, \alpha_0 > 0 : \; x^* + th \in D(f) \;\& \\ \& \; f(x^* + th) < f(x^*) \;\text{ pro }\; \forall t \in (0, \alpha_0)\Big\}.$$

Význam těchto množin?

(i) $\mathcal{V}(x^*, X)$ je tzv. *množina přípustných vektorů* (směrů pro $\|h\| = 1$) v bodě x^*. Je to kužel? Konvexní? A je-li $x^* \in \operatorname{ri} X$?

(ii) $\mathcal{U}(x^*, f)$ obsahuje tzv. *spádové vektory* a jedná se o tzv. *kužel zlepšujících vektorů* (směrů pro $\|h\| = 1$) funkce f v bodě x^*.

S využitím těchto dvou množin obdržíme první nutnou podmínku pro existenci lokálního řešení úlohy (4.1). Tvrzení je asi víceméně zřejmé a důkaz relativně snadný, ale význam tohoto tvrzení pro námi budovanou teorii je značný.

Lemma 4.1.1

Nechť $X \subseteq \mathbb{R}^n$ a $f : X \to \mathbb{R}$ jsou dány. Je-li bod $x^* \in X$ lokálním řešením úlohy (4.1), potom

$$\mathcal{V}(x^*, X) \bigcap \mathcal{U}(x^*, f) = \emptyset. \tag{4.1.1}$$

Důkaz. Sporem. Nechť $x^* \in X$ je lokálním řešením úlohy (4.1) a existuje $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f)$, tj. $\exists\ \alpha_u, \alpha_v > 0$ taková, že $x^* + th \in X$ pro $t \in (0, \alpha_v)$ a $f(x^* + th) < f(x^*)$ 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 $x^* + th \in X \subset D(f)$ a $f(x^* + th) < f(x^*)$ ↯ ■

Obrázek.

Podmínka (4.1.1) bude jistě splněna, pokud $\mathcal{U}(x^*, f) = \emptyset$. Kdy to nastane? Pro spojitě diferencovatelnou funkci dostaneme požadavek

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

pro každý vektor $h \in \operatorname{Lin} X$ splňující $x^* + th \in D(f)$ pro $t > 0$ dostatečně malá. Zejména v případě $\operatorname{Lin} X = \mathbb{R}^n$ a $x^* \in \operatorname{int} D(f)$ musí platit $\operatorname{grad} f(x^*) = 0$.

Nicméně toto je zbytečně silný požadavek, nám postačuje (4.1.1), což nás přivádí k následující definici (viz Poznámku 2.4.3A).

Definice 4.1.2

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 (4.1) (nebo *stacionárním bodem funkce* f *na množině* X), jestliže

$$\langle \operatorname{grad} f(x^*),\, x - x^*\rangle \geqslant 0 \tag{4.1.2}$$

pro každé $x \in X$.

Je-li $X = \mathbb{R}^n$, pak podmínka (4.1.2) je tvaru

$$\sum_{i=1}^{n} \frac{\partial f(x^*)}{\partial x_i}(x_i - x_i^*) \geqslant 0 \qquad \forall x \in \mathbb{R}^n,$$

což je splněno pouze v případě $\operatorname{grad} f(x^*) = 0$ (viz „klasická“ definice stacionárního bodu).

Následující věta ukazuje, že stacionární bod ve smyslu Definice 4.1.2 má přesně ty vlastnosti, které od něj očekáváme.

Věta 4.1.3 Nechť funkce $f : X \to \mathbb{R}$ je diferencovatelná na (nějaké otevřené množině obsahující konvexní množinu) $X \subseteq \mathbb{R}^n$.

(i) Je-li $x^* \in X$ lokálním extrémem funkce f na X (tj. lokálním řešením úlohy (4.1)), 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 (4.1), tj. (jediným) globálním minimem f na X.

První část Věty 4.1.3 udává nutnou podmínku pro řešení úlohy (4.1), ze které se ovšem ve druhé části (po přidání požadavku konvexnosti funkce f) stala podmínka postačující $\rightsquigarrow$ máme totiž úlohu konvexního programování (minimalizujeme konvexní funkci na konvexní množině $\rightsquigarrow$ stačí „pouze“ najít stacionární body – ty jsou totiž zároveň globálními minimy; pozor na Větu 2.2.6).

✎ Obrázek.

Důkaz. Nechť $x^* \in X$ je lokálním řešením úlohy (4.1) a neplatí (4.1.2), 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. existuje $\alpha_0 > 0$ takové, že $f(x^* + th) < f(x^*)$ pro každé $t \in (0, \alpha_0)$. Tedy celkem $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) \neq \emptyset$ ↯.

Naopak, je-li funkce f (ostře) konvexní a $x^* \in X$ splňuje (4.1.2) pro každé $x \in X$, pak z Věty 2.4.2 plyne

$$f(x) \underset{(>)}{\geq} f(x^*) + \langle \operatorname{grad} f(x^*), x - x^* \rangle \quad \forall x \in X,$$

což vzhledem k (4.1.2) znamená

$$f(x) \underset{(>)}{\geq} f(x^*) \quad \forall x \in X,$$

tj. x^* je (jediným) řešením úlohy (4.1). ■

Z praktického pohledu (viz např. příklady později) je velmi důležitá situace

$$X = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_1, \dots x_n \geqslant 0\}.$$

V takovém případě obdržíme z Věty 4.1.3 následující důsledek.

Důsledek 4.1.4

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

$$f(x) \to \min, \quad x \in \mathbb{R}^n_+, \tag{4.1.3}$$

pak platí

$$\frac{\partial f}{\partial x_i}(x^*) \geqslant 0 \quad \& \quad x_i^* \frac{\partial f}{\partial x_i}(x^*) = 0, \quad i = 1, \dots, n, \tag{4.1.4}$$

neboli pro každé $i \in \{1, \dots, n\}$ máme

$$\frac{\partial f}{\partial x_i}(x^*) \geqslant 0 \quad \text{a navíc} \quad \frac{\partial f}{\partial x_i}(x^*) = 0 \ \text{ v případě } x_i^* > 0$$

Naopak, splňuje-li bod $x^* \in \mathbb{R}^n_+$ podmínku (4.1.4) a funkce f je navíc konvexní na $\mathbb{R}^n_+$, pak x^* je řešením úlohy (4.1.3).

Obrázek.

Důkaz Důsledku 4.1.4

Nechť $x^* \in \mathbb{R}^n$ je řešením (lokálním = globálním) úlohy (4.1.3). Pak podle Věty 4.1.3 je x^* stacionárním bodem, tj.

$$\langle \operatorname{grad} f(x^*), x - x^* \rangle \geqslant 0 \quad \forall x \in \mathbb{R}^n_+$$

neboli

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0 \quad \forall x \in \mathbb{R}^n_+. \tag{$*$}$$

Potom ale nutně $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$ pro každé $i \in \{1, \dots, n\}$ a $x \in \mathbb{R}^n_+$. Vskutku, jestliže $\frac{\partial f}{\partial x_j}(x^*)\,(\widehat{x}_j - x_j^*) < 0$ pro nějaký index $j \in \{1, \dots, n\}$ a nějaké $\widehat{x} \in \mathbb{R}^n_+$, pak stačí vzít $\widetilde{x} \in \mathbb{R}^n_+$ takové, že $\widetilde{x}_i - x_i^* = 0$ (tj. $\widetilde{x}_i = x_i^*$) pro $i \in \{1, \dots, n\} \backslash \{j\}$ a $\widetilde{x}_j = \widehat{x}_j \neq x_j^*$ (to lze, neboť $X = \mathbb{R}^n_+$). Tím ale dostaneme spor s $(*)$. Tedy $(*)$ skutečně platí právě tehdy, když $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$ pro každé $i \in \{1, \dots, n\}$ a $x \in \mathbb{R}^n_+$ (opačná implikace je triviální).

Důkaz Důsledku 4.1.4 (pokr.)

Jenže současně platí

$$\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0 \quad \forall i \in \{1,\dots,n\},\ x \in \mathbb{R}^n_+ \qquad \Longleftrightarrow \qquad \text{platí (4.1.4).}$$

Pro dokončení první části důkazu proto musíme ještě ukázat tuto ekvivalenci.

(i) Nechť pro $i \in \{1,\dots,n\}$ platí levá strana ekvivalence.
   (a) Je-li $x_i^* = 0$, pak $x_i - x_i^* \geqslant 0$ pro všechna $x \in X$, takže nutně $\frac{\partial f}{\partial x_i}(x^*) \geqslant 0$, tj. platí (4.1.4)(i) a podmínka (4.1.4)(ii) je v tomto případě splněna triviálně.
   (b) Je-li $x_i^* > 0$, pak rozdíl $x_i - x_i^*$ může nabývat kladných i záporných hodnot na X, takže nutně $\frac{\partial f}{\partial x_i}(x^*) = 0$, tedy zjevně platí (4.1.4)(i) a (4.1.4)(ii).

(ii) Naopak, nechť pro $i \in \{1,\dots,n\}$ platí (4.1.4).
   (a) Je-li $x_i^* = 0$ a $\frac{\partial f}{\partial x_i}(x^*) \geqslant 0$, pak $\frac{\partial f}{\partial x_i}(x^*)\,x_i = \frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$ pro každé $x \in X$.
   (b) Je-li $x_i^* > 0$, pak $\frac{\partial f}{\partial x_i}(x^*) = 0$ dle (4.1.4), a tudíž $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) = 0$ pro každé $x \in X$.

   Odtud plyne, že vždy platí $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$.

Důkaz první části je tímto hotov.

Důkaz Důsledku 4.1.4 (pokr.)

Naopak, nechť pro $x^* \in X$ platí (4.1.4) a funkce f je konvexní na $\mathbb{R}^n_+$. Potom dle předchozí části $\frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$ pro každé $i \in \{1,\dots,n\}$ a $x \in \mathbb{R}^n_+$, a tudíž také $\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x^*)\,(x_i - x_i^*) \geqslant 0$ neboli

$$\langle \operatorname{grad} f(x^*),\, x - x^* \rangle \geqslant 0 \quad \forall x \in \mathbb{R}^n_+,$$

tj. $x^* \in \mathbb{R}^n$ je stacionární bod funkce f na X. Ovšem vzhledem ke konvexnosti funkce f plyne z Věty 4.1.3, že x^* je řešením úlohy (4.1.3). ■

Podobně jako v Důsledku 4.1.4 můžeme z Věty 4.1.3 získat nutné (a v případě konvexní funkce také postačující) podmínky pro (lokální/globální) řešení úlohy (4.1) pro některé další významné typy množiny X.

(i) Pro množinu

$$X = \left\{x \in \mathbb{R}^n \mid x_i \geqslant 0,\ i \in I\right\}, \quad I \subset \{1,\dots,n\}$$

   je (nutná/postačující) podmínka tvaru

$$\frac{\partial f}{\partial x_i}(x^*) \geqslant 0 \quad \& \quad x_i^*\,\frac{\partial f}{\partial x_i}(x^*) = 0 \ \text{ pro } \ i \in I \quad \& \quad \frac{\partial f}{\partial x_i}(x^*) = 0 \ \text{ pro } \ i \in \{1,\dots,n\} \setminus I.$$

(ii) Pro množinu
$$X = \{x \in \mathbb{R}^n \mid \alpha_i \leqslant x_i \leqslant \beta_i,\ i = 1, \dots, n\},$$
kde $\alpha_i, \beta_i \in \mathbb{R}$ jsou daná čísla a $\alpha_i < \beta_i$, je (nutná/postačující) podmínka tvaru
$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} \geqslant 0, & x_i^* = \alpha_i, \\ \leqslant 0, & x_i^* = \beta_i, \\ = 0, & \alpha_i < x_i^* < \beta_i. \end{cases}$$

(iii) Pro množinu
$$X = \Big\{x \in \mathbb{R}^n \mid x_i \geqslant 0,\ \sum_{i=1}^n x_i = r\Big\},$$
kde $r > 0$ je dané číslo (jedná se o zobecnění jednotkového $(n-1)$-simplexu), je (nutná/postačující) podmínka ve tvaru implikace
$$x_i^* > 0 \implies \frac{\partial f}{\partial x_i}(x^*) \leqslant \frac{\partial f}{\partial x_j}(x^*) \quad \text{pro} \quad \forall j \in \{1, \dots, n\}.$$

V případě, že funkce f není konvexní, potřebujeme k rozhodnutí o „extrémnosti“ stacionárního bodu x^* nějaký další nástroj. Je-li $f \in C^2$, pak máme podmínky druhého řádu:

Nutná podmínka pro (4.1)

Je-li $x^* \in X$ lokálním minimem funkce $f : X \to \mathbb{R}$, $f \in C^2$, na konvexní množině $X \subseteq \mathbb{R}^n$, pak
$$(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) \geqslant 0$$
pro všechna $x \in X$ taková, že $\operatorname{grad}^\top f(x^*)\,(x - x^*) = 0$, tj. pro vektory $(x - x^*) \in \operatorname{Ker} \operatorname{grad}^\top f(x^*)$.

Postačující podmínka pro (4.1)

Bod $x^* \in X$ je lokálním minimem funkce $f : X \to \mathbb{R}$, $f \in C^2$, na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže
$$\operatorname{grad}^\top f(x^*)\,(x - x^*) \geqslant 0, \quad \forall x \in X$$
(tj. je to stacionární bod), množina X je polyedr a platí
$$(x - x^*)^\top \nabla^2 f(x^*)\,(x - x^*) > 0$$
pro všechna $x \in X$ taková, že $x \neq x^*$ a $(x - x^*) \in \operatorname{Ker} \operatorname{grad}^\top f(x^*)$.

Proč polyedr? V takovém případě je totiž množina $\mathcal{V}(x^*, X)$ uzavřená.

Některé historické úlohy

- <u>Keplerův problém</u>: Do dané koule vepište válec s maximálním objemem (planimetrická formulace: kruh $\rightsquigarrow$ obdélník $\rightsquigarrow$ obsah).
- <u>Steinerův problém</u>: V rovinném trojúhelníku najděte takový bod, že součet jeho vzdáleností od vrcholů trojúhelníku je minimální ($\rightsquigarrow$ /Fermatův–/Toricelliho bod).
- <u>Tartagliova úloha</u>: Rozdělte číslo 8 na dvě části tak, aby jejich součin s jejich rozdílem byl maximální.

„Moderní" úlohy

- <u>Lineární programování</u> (LP)

$$c^\top x \to \min, \quad Ax \leqslant b, \quad x \geqslant 0,$$

kde $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ jsou dány (viz M4140 nebo M0160). Např. optimální výrobní program, dopravní problém atd. Toto úzce souvisí s <u>celočíselným programováním</u> (ZLP), např. problém batohu/zloděje, problém obchodního cestujícího (viz M0160).
- <u>ekonomické úlohy</u>: maximalizace užitku/zisku, minimalizace výdajů/nákladů, tvorba portfolia $\rightsquigarrow$ kvadratické programování (viz M0160)

V předchozí části jsme řešili obecnou minimalizační úlohu (4.1) a viděli jsme, že důležitou roli zde hraje geometrie množiny X (vektory $x - x^*$). Abychom se mohli v našich úvahách posunout dále musíme určit množinu X blíže, proto se dále budeme zabývat úlohou (4.1) & (4.2), tj.

$$f(x) \to \min, \quad x \in X, \tag{4.1}$$

$$X := \{x \in P \subseteq \mathbb{R}^n \mid g_i(x) \leqslant 0,\ g_j(x) = 0,\ i = 1, \dots, k,\ j = k+1, \dots, m\}. \tag{4.2}$$

Její speciální případ s $P = \mathbb{R}^n$ a $k = 0$ již řešit umíme za předpokladu regularity Jacobiho matice $DG(x)$ s pomocí stacionárních bodů přidružené (regulární) Lagrangeovy funkce

$$L(x, y) = f(x) + \sum_{i=1}^{m} y_i\, g_i(x).$$

Podobně budeme postupovat i v případě obecné úlohy (4.1) & (4.2). K této úloze přidružíme (obecnou) *Lagrangeovu funkci* $L : P \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ danou předpisem

$$L(x, y_0, y) := y_0\, f(x) + \sum_{i=1}^{m} y_i\, g_i(x), \tag{4.2.1}$$

přičemž v případě $y_0 = 1$ bude $L(x, 1, y) =: L(x, y)$. Čísla $y_0, \dots, y_m$ opět nazýváme *Lagrangeovými multiplikátory*.

Pro výklad ještě potřebujeme následující množiny

$$Q := \{y = (y_1, \dots, y_m)^\top \mid y_1, \dots, y_k \geqslant 0\},$$
$$I(x^*) := \{i \in \{1, \dots, k\} \mid g_i(x^*) = 0\}, \quad x^* \in X,$$
$$S(x^*) := I(x^*) \cup \{k+1, \dots, m\}, \quad x^* \in X.$$

Množina $I(x^*)$ značí množinu *aktivních omezení* v bodě x^*. Množina $S(x^*)$ pak je tvořena indexy všech funkcí určujících množinu X, které se v bodě x^* realizují jako rovnosti.

Věta 4.2.1

(Lagrangeův princip) 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 (4.1) & (4.2), pak existují Lagrangeovy multiplikátory $y_0^*, y_1^*, \dots, y_k^* \geqslant 0$, $y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$, tj. $y_0^* \geqslant 0$ a $y^* \in Q$, taková, že ne všechna $y_0^*, \dots, y_m^*$ jsou nulová a platí

$$\langle \operatorname{grad}_x L(x^*, y_0^*, y^*),\, x - x^* \rangle \geqslant 0 \quad \forall x \in P, \tag{4.2.3}$$

$$y_i^* g_i(x^*) = 0, \quad i = 1, \dots, m. \tag{4.2.4}$$

Podmínka (4.2.3) $\rightsquigarrow$ x^* je stacionárním bodem funkce $L(x, y_0^*, y^*)$ na P.
Podmínka (4.2.4) $\rightsquigarrow$ podmínka komplementarity.
Požadavek $y_1^*, \dots, y_k^* \geqslant 0$ $\rightsquigarrow$ podmínka duality.
(4.2.3)+(4.2.4) $\rightsquigarrow$ Johnovy podmínky a pro $y_0^* = 1$ $\rightsquigarrow$ KKT-podmínky.

Důkaz Věty 4.2.1

Pro jednoduchost se omezíme na situaci, kdy funkce $g_{k+1}, \dots, g_m$ jsou pouze afinní (v opačném případě je potřeba použít Větu o implicitní funkci). Nechť $x^* \in X$ je lokální řešení úlohy (4.1) & (4.2) a uvažme systém nerovností a rovností na množině P ve tvaru

$$\left.\begin{aligned} \langle \operatorname{grad} f(x^*),\, x - x^*\rangle &< 0, \\ \langle \operatorname{grad} g_i(x^*),\, x - x^*\rangle &< 0, \quad i \in I(x^*), \\ \langle \operatorname{grad} g_j(x^*),\, x - x^*\rangle &= 0, \quad j = k+1, \dots, m. \end{aligned}\right\} \tag{4.2.5}$$

Předpokládejme, že existuje řešení $\widetilde{x} \in P$ tohoto systému. Označme $h := \widetilde{x} - x^*$. Potom $x^* + th \in P$ pro libovolné $t \in [0,1]$. 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

$$\frac{f(x^* + th) - f(x^*)}{t} = \langle \operatorname{grad} f(x^*),\, h\rangle + \frac{o(\|th\|)}{t\|h\|}\|h\| < 0 \quad t \to 0^+,$$

takže $h \in \mathcal{U}(x^*, f)$.

Důkaz Věty 4.2.1 (pokr.)

(ii) Podobně pro $i \in I(x^*)$ plyne ze druhé nerovnosti, že funkce g_i jsou klesající ve směru h a současně $g_i(x^*) = 0$, tedy $g_i(x^* + th) \leqslant 0$ pro dostatečně malé $t > 0$, tj. $h \in \mathcal{U}(x^*, g_i)$ pro $i \in I(x^*)$. Pro $i \in \{1, \dots, k\} \setminus I(x^*)$ máme $g_i(x^*) < 0$, tedy v tomto případě $g_i(x^* + th) \leqslant 0$ 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 (4.2.5), tj.

$$g_i(x^* + th) = g_i(x^*) + t \operatorname{grad}^\top g_i(x^*)\, h = g_i(x^*) = 0 \quad \text{pro libovolné } t,$$

a tudíž $h \in \mathcal{U}(x^*, g_i)$ pro všechna $i \in \{1, \dots, m\}$. Jelikož $x^* + th \in P$ plyne odtud, že $h \in \mathcal{V}(x^*, X)$, tj. $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) \neq \emptyset$ ↯ s Lemma 4.1.1.

Proto systém (4.2.5) nemá řešení na P a z Věty 2.3.12 plyne existence čísel $y_0^*, y_i^* \geqslant 0$ pro $i \in I(x^*)$ a $y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$ takových, že

$$y_0^* \langle \operatorname{grad} f(x^*),\, x - x^*\rangle + \sum_{i \in S(x^*)} y_i^* \langle \operatorname{grad} g_i(x^*),\, x - x^*\rangle \geqslant 0,$$

tj.

$$\left\langle y_0^* \operatorname{grad} f(x^*) + \sum_{i \in S(x^*)} y_i^* \operatorname{grad} g_i(x^*),\, x - x^* \right\rangle \geqslant 0 \quad \forall x \in P.$$

Důkaz Věty 4.2.1 (pokr.)

Jestliže ještě dodefinujeme $y_i^* := 0$ pro $i \in \{1, \dots, k\} \setminus I(x^*)$, pak z poslední nerovnosti dostáváme (4.2.3), tj.

$$\left\langle y_0^* \operatorname{grad} f(x^*) + \sum_{i=1}^{m} y_i^* \operatorname{grad} g_i(x^*),\ x - x^* \right\rangle \geqslant 0 \quad \forall x \in P.$$

Podmínka (4.2.4) je splněna triviálně. ■

Podmínka (4.2.3) ve speciálních případech

(i) Je-li $x^* \in \operatorname{int} P$ (tedy zejména je-li P otevřená), pak podmínka (4.2.3) se změní na rovnost, tj.

$$\operatorname{grad}_x L(x^*, y_0^*, y^*) = 0.$$

(ii) Je-li

$$P = \left\{x \in \mathbb{R}^n \mid \alpha_i \leqslant x_i \leqslant \beta_i,\ i = 1, \dots, n\right\},$$

kde $-\infty \leqslant \alpha_i < \beta_i \leqslant \infty$ pro $i \in \{1, \dots, n\}$, pak (4.2.3) znamená

$$\frac{\partial L}{\partial x_i}(x^*, y_0^*, y^*) \begin{cases} = 0, & \alpha_i < x_i^* < \beta_i, \\ \geqslant 0, & x_i^* = \alpha_i \neq -\infty, \\ \leqslant 0, & x_i^* = \beta_i \neq \infty. \end{cases}$$

(iii) Je-li

$$P = \left\{x \in \mathbb{R}^n \mid x_i \geqslant 0,\ i = 1, \dots, s\right\},$$

kde $s \in \{1, \dots, n\}$, pak (4.2.3) znamená

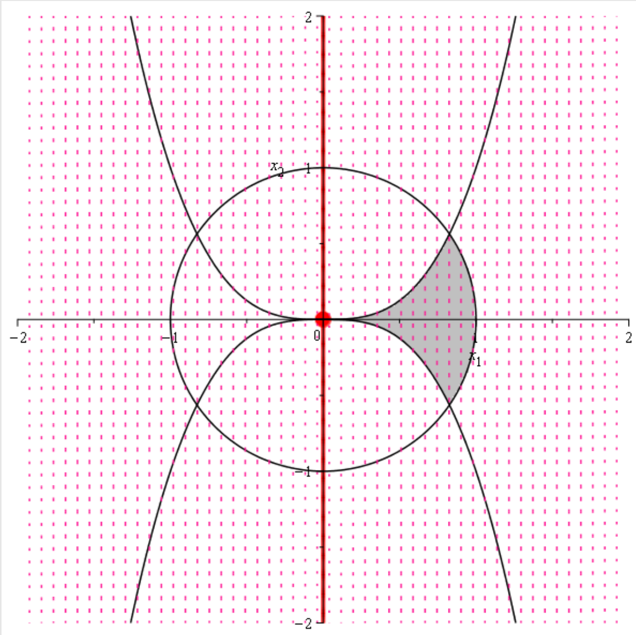
$$\frac{\partial L}{\partial x_i}(x^*, y_0^*, y^*) \geqslant 0 \quad \& \quad x_i^* \frac{\partial L}{\partial x_i}(x^*, y_0^*, y^*) = 0, \quad i \in \{1, \dots, s\},$$
$$\frac{\partial L}{\partial x_i}(x^*, y_0^*, y^*) = 0, \quad i \in \{s+1, \dots, n\}.$$

Je Věta 4.2.1 konzistentní s výkladem pro úlohy s omezeními pouze ve tvaru rovností?

Příklad

Uvažme úlohu s $P = \mathbb{R}^2$ a

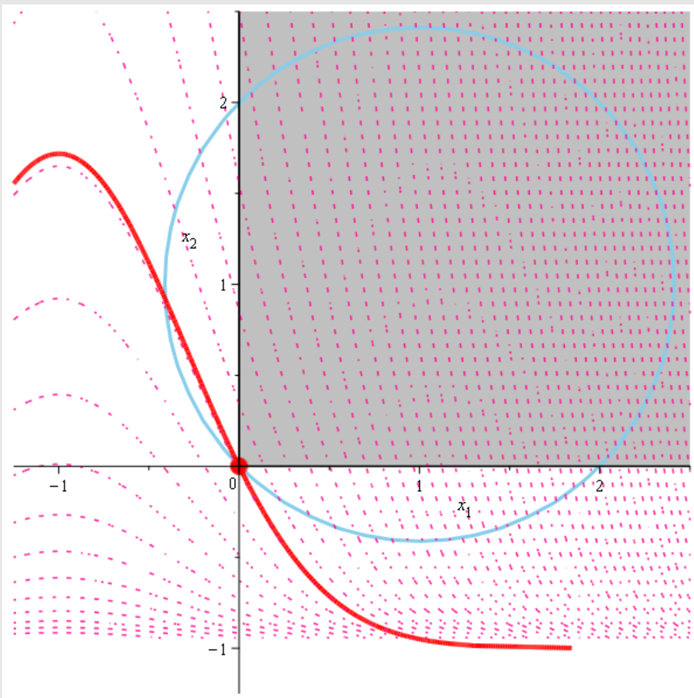
$$f(x_1, x_2) = x_1 \to \min,$$
$$g_1(x_1, x_2) = -x_1^3 + x_2 \leqslant 0, \qquad g_2(x_1, x_2) = -x_1^3 - x_2 \leqslant 0,$$
$$g_3(x_1, x_2) = x_1^2 + x_2^2 - 1 \leqslant 0.$$



„Divočejší“ příklad

Uvažme úlohu ($P = \mathbb{R}^2$)

$$f(x_1, x_2) = (x_1 + 1)^2 + \ln(x_2 + 1) \to \min,$$
$$(x_1 - 1)^2 + (x_2 - 1)^2 = 2, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.$$



Situace s $y_0^* = 0$ je velmi problematická $\rightsquigarrow$ potřebujeme zaručit $y_0^* \neq 0$, což je ekvivalentní s $y_0^* = 1$ (tzv. podmínky kvalifikovaného omezení).

Např. *regulárnost* bodu x^*, tj. lineární nezávislost vektorů $\operatorname{grad} g_i(x^*)$ pro $i \in S(x^*)$.

Věta 4.2.1A

Nechť jsou splněny předpoklady Věty 4.2.1 a $x^* \in \operatorname{int} P$. Jestliže x^* je regulární bod, pak existují (jediné) multiplikátory $y^* \in Q$ takové, že platí (4.2.3) & (4.2.4) s $y_0^* = 1$.

Regulárnost při omezení na znaménko?

V literatuře lze nalézt i několik dalších podmínek zaručující $y_0^* = 1$ při $x^* \in \operatorname{int} P$ (lokální vs. globální).

Věta 4.2.1B

Nechť jsou splněny předpoklady Věty 4.2.1 a $x^* \in \operatorname{int} P$. Potom existují (jediné) multiplikátory $y^* \in Q$ splňující (4.2.3) & (4.2.4) s $y_0^* = 1$, jestliže je splněna (alespoň) jedna z následujících podmínek

(i) (afinní omezení) funkce $g_1, \dots, g_m$ jsou afinní;

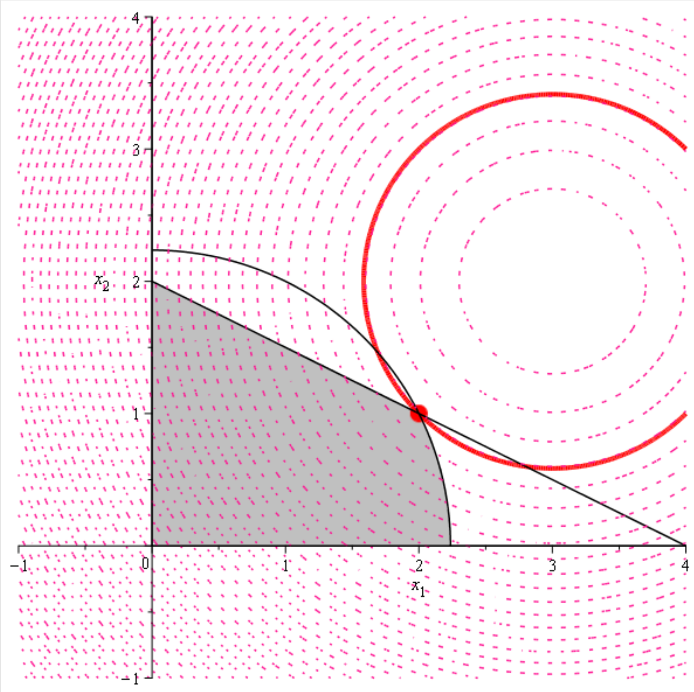
(ii) (Slaterova podmínka) $g_1, \dots, g_k$ jsou konvexní, $g_{k+1}, \dots, g_m$ jsou afinní, konstantní vektory $\operatorname{grad} g_i$ jsou lineárně nezávislé pro $i \in \{k+1, \dots, m\}$ a existuje $\bar{x} \in P$ takový, že $g_i(\bar{x}) < 0$ pro $i \in \{1, \dots, k\}$ a $g_i(\bar{x}) = 0$ pro $i \in \{k+1, \dots, m\}$.

Příklad

Uvažme úlohu

$$(x_1-3)^2+(x_2-2)^2 \to \min,$$

$$x_1^2+x_2^2 \leqslant 5, \qquad x_1+2x_2 \leqslant 4, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.$$



Věta 4.2.1 dává pouze nutnou podmínku, tj. ne každý bod splňující (4.2.3) a (4.2.4) je řešením úlohy (4.1) & (4.2) $\rightsquigarrow$ potřebujeme nějakou postačující podmínku („certifikát optimality“).

Nalezení řešení úlohy (4.1) & (4.2) není ekvivalentní s nalezením extrému Lagrangeovy funkce, a to ani v případě $y_0^*=1$. Nicméně je to postačující.

Věta 4.2.1C

Nechť množina $P\subseteq\mathbb{R}^n$ je konvexní, funkce $f, g_1, \ldots, g_k : P\to\mathbb{R}$ jsou diferencovatelné v bodě $x^*\in X$ a $g_{k+1},\ldots,g_m : P\to\mathbb{R}$ jsou spojitě diferencovatelné v nějakém okolí bodu x^*. Nechť dále pro $x^*\in X$ existují multiplikátory $y^*\in Q$ takové, že platí (4.2.3) a (4.2.4) s $y_0^*=1$. Je-li x^* bodem globálního minima funkce $L(x,y^*)$ na P, pak x^* je globálním řešením úlohy (4.1) & (4.2).

Důkaz Věty 4.2.1C

Z předpokladů plyne, že $L(x^*, y^*) \leqslant L(x, y^*)$ pro každé $x \in P$. Navíc z (4.2.4) dostáváme

$$f(x^*) = f(x^*) + \sum_{i=1}^{m} y_i^* g_i(x^*) = L(x^*, y^*).$$

Proto pro každé $x \in X$ platí

$$f(x^*) = L(x^*, y^*) \leqslant L(x, y^*) = f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = f(x) + \sum_{i=1}^{k} y_i^* g_i(x) \leqslant f(x),$$

neboť $g_i(x) = 0$ pro $i \in \{k+1, \dots, m\}$ a $y_i^* g_i(x) \leqslant 0$ pro každé $x \in X$. To znamená, že x^* je globálním minimem funkce f na množině X. ■

Kdy budou splněny podmínky Věty 4.2.1C? Zejména ve chvíli, kdy funkce $L(x, y^*)$ je konvexní $\rightsquigarrow$ úloha konvexního programování.

Důsledek 4.2.2

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 (4.2.3) a (4.2.4) s $y_0^* = 1$. Nechť je dále splněn (alespoň) jeden z následujících předpokladů:

(i) funkce $L(x, y^*)$ je konvexní na množině P,

(ii) úloha (4.1) & (4.2) 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 (4.1) & (4.2).

Důkaz. Jelikož $y^* \in Q$, předpoklad (ii) je speciálním případem (i). Podmínka (4.2.3) 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 4.1.3, že bod x^* je bodem globálního minima funkce $L(x, y^*)$ na P. Proto tvrzení plyne z Věty 4.2.1C. ■

Mohla by úloha konvexního programování vypadat i jinak než je popsáno v (ii)?

Předchozí důsledek ukazuje, že zejména v případě úlohy konvexního programování je existence stacionárního bodu s Lagrangeovým multiplikátorem $y_0^* = 1$ dokonce postačující pro globální řešení.

Kvalifikovanost omezení & konvexnost úlohy $\rightsquigarrow$ základní tvrzení konvexního/nelineárního programování:

Věta 4.2.3

(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) (LNZ) 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 $\bar{x} \in P$ takový, že $g_i(\bar{x}) < 0$ pro $i \in \{1, \dots, k\}$;

(iii) (lineární) množina P je polyedr a funkce $g_1, \dots, g_k$ jsou afinní.

Pak x^* je řešením úlohy (4.1) & (4.2) právě tehdy, když existuje $y^* \in Q$ takové, že platí (4.2.3) a (4.2.4) s $y_0^* = 1$.

Podmínky (ii) a (iii) Věty 4.2.3 se trochu liší od Věty 4.2.1B, což je způsobeno konvexností funkce f, která umožňuje využití regulárních modifikací Věty 2.3.12. V LP a QP vždy platí (iii).

Důkaz Věty 4.2.3

„$\Longleftarrow$“ To plyne přímo z Důsledku 4.2.2 (i bez dodatečných podmínek).

„$\Longrightarrow$“ Musíme dokázat pouze „dostatečnost“ uvedených předpokladů. Nechť tedy $x^* \in X$ řeší úlohu (4.1) & (4.2). V případě podmínky (i) plyne tvrzení přímo z Věty 4.2.1A. Uvažme nyní podmínku (ii). Nechť tedy $k = m$ a uvažme systém nerovností

$$\langle \operatorname{grad} f(x^*),\, x - x^* \rangle < 0, \tag{4.2.6}$$

$$\langle \operatorname{grad} g_i(x^*),\, x - x^* \rangle < 0, \quad i \in I(x^*). \tag{4.2.7}$$

Pak tento systém nemá řešení na P — kdyby totiž $\tilde{x} \in P$ řešilo (4.2.6) a (4.2.7), pak (podobně jako v důkazu Věty 4.2.1) by z diferencovatelnosti funkcí g_i plynulo pro $h = \tilde{x} - x^*$, že

$$g_i(x^* + \alpha h) = g_i(x^*) + \underbrace{\langle \operatorname{grad} g_i(x^*),\, \alpha h \rangle}_{<0 \text{ dle } (4.2.7)} + o(\|\alpha h\|),$$

tj. $g_i(x^* + \alpha h) < g_i(x^*)$ pro $\alpha > 0$ dostatečně malé. Neboť $x^* + \alpha h = \alpha\tilde{x} + (1-\alpha)\, x^* \in P$ pro $\alpha \in [0, 1]$, plyne odtud, že $x^* + \alpha h \in X$ pro $\alpha > 0$ dostatečně malé, a tudíž $h \in \mathcal{V}(x^*, X)$. Současně z (4.2.6) 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. $h \in \mathcal{U}(x^*, f)$. To ale znamená, že $h \in \mathcal{U}(x^*, f) \cap \mathcal{V}(x^*, X) \neq \emptyset$ $\lightning$ s Lemma 4.1.1.

Důkaz Věty 4.2.3 (pokr.)

Nicméně systém (4.2.7) má řešení $\overline{x}$, neboť s využitím Věty 2.4.2 totiž máme

$$\langle \operatorname{grad} g_i(x^*), \overline{x} - x^* \rangle \leqslant g_i(\overline{x}) - g_i(x^*) = g_i(\overline{x}) < 0.$$

Proto z Věty 2.3.13 plyne existence $y_i^* \geqslant 0$ pro $i \in I(x^*)$ takových, že platí

$$\left\langle \operatorname{grad} f(x^*) + \sum_{i \in I(x^*)} y_i^* \operatorname{grad} g_i(x^*), x - x^* \right\rangle \geqslant 0.$$

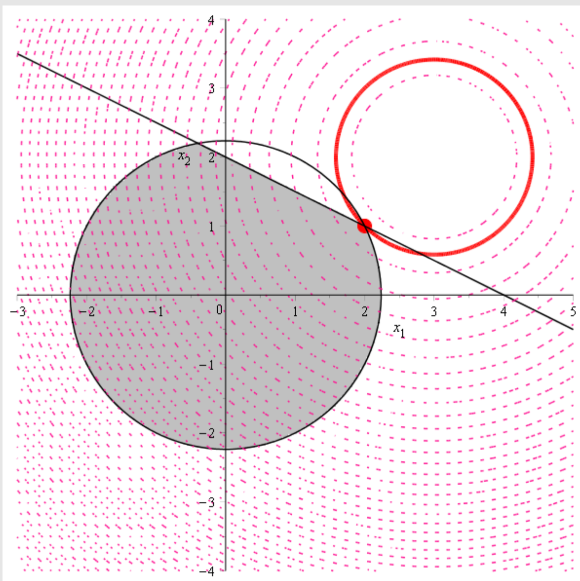
Dodefinujeme-li $y_i^* = 0$ pro $i \in \{1, \dots, k\} \setminus I(x^*)$, pak dostáváme multiplikátory $y_1^*, \dots, y_k^* \geqslant 0$, pro které platí (4.2.3) a (4.2.4) s $y_0^* = 1$.

Důkaz tvrzení za podmínky (iii) plyne z Věty 2.3.14. ■

Příklad

Vyřešme

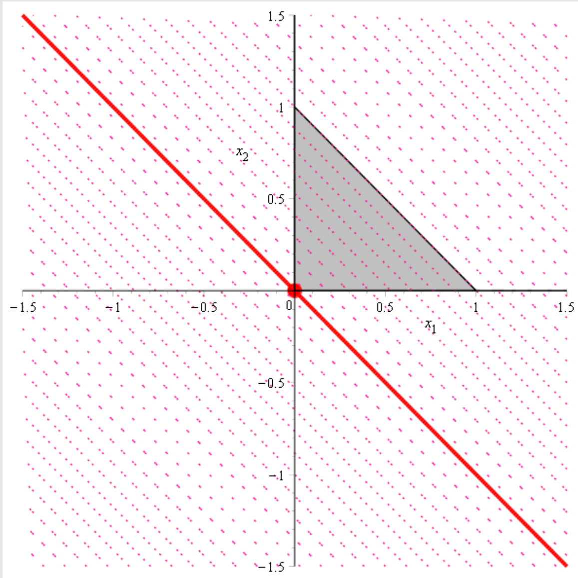
$$f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \to \min,$$
$$x_1^2 + x_2^2 \leqslant 5, \qquad x_1 + 2x_2 \leqslant 4.$$



Příklad

Vyřešme

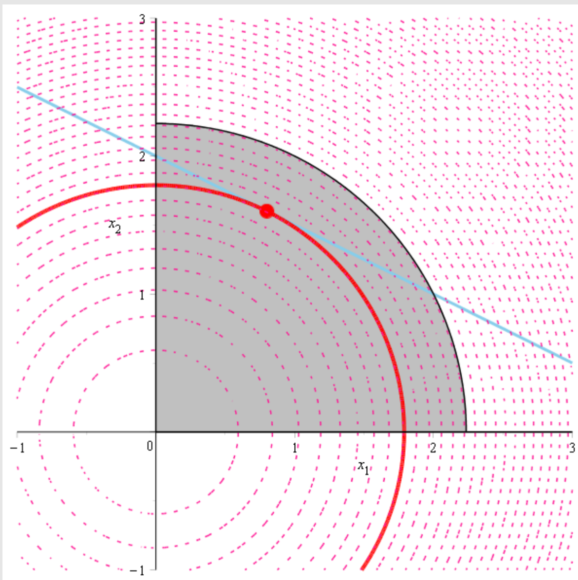
$$f(x_1, x_2) = x_1 + x_2 \to \min,$$
$$x_1 + x_2 \leqslant 1, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.$$



Příklad

Vyřešme

$$f(x_1, x_2) = x_1^2 + x_2^2 \to \min,$$
$$x_1^2 + x_2^2 \leqslant 5, \qquad x_1 + 2x_2 = 4, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.$$



Pro kontrolu vlastních výpočtů je k dispozici maplet s Lagrangeovým principem, viz https://goo.gl/w2JUvL.

Věty 4.2.1C a 4.2.3 dávají postačující podmínky, za kterých stačí pro řešení úlohy (4.1) & (4.2) najít pouze stacionární bod Lagrangeovy funkce.

Bez platnosti těchto podmínek potřebujeme nějaké další kritérium. To je jako vždy založeno na definitnosti matice

$$\nabla_x^2 L(x^*, y_0^*, y^*).$$

Věta 4.2.4

Nechť funkce $f, g_1, \dots, g_m$ jsou dvakrát spojitě diferencovatelné v bodě x^* a $x^* \in \operatorname{int} P$ je takový, že existují multiplikátory $y^* \in Q$ splňující (4.2.3) a (4.2.4) s $y_0^* = 1$ a současně $y_i^* > 0$ pro $i \in I(x^*)$ (tzv. *podmínka ostré komplementarity*), tj.

$$\begin{gathered}\operatorname{grad}_x L(x^*, y^*) = 0, \\ g_i(x^*) \leqslant 0 \quad \text{pro} \quad i \in \{1, \dots, k\}, \qquad g_i(x^*) = 0 \quad \text{pro} \quad i \in \{k+1, \dots, m\}, \\ y_i^* > 0 \quad \text{pro} \quad i \in I(x^*), \qquad y_i^* = 0 \quad \text{pro} \quad i \in \{1, \dots, k\} \setminus I(x^*), \\ y_i^* \in \mathbb{R} \quad \text{pro} \quad i \in \{k+1, \dots, m\}.\end{gathered}$$

Jestliže

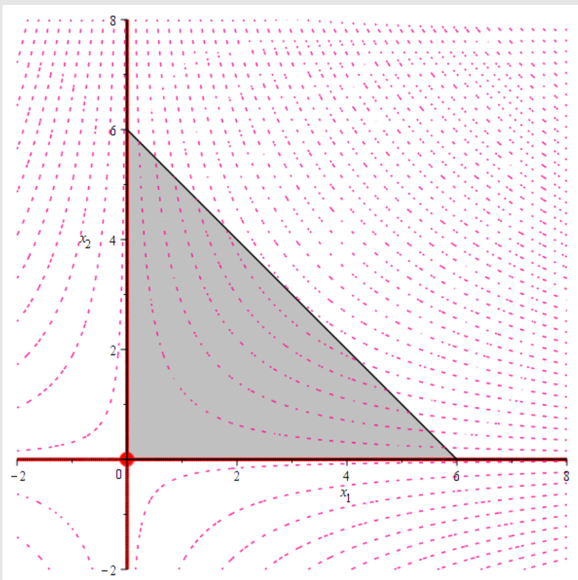
$$\nabla_x^2 L(x^*, y^*) > 0 \quad \text{na} \quad \operatorname{Ker}(\operatorname{grad}^\top g_i(x^*))_{i \in S(x^*)},$$

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.

Příklad

Uvažme úlohu

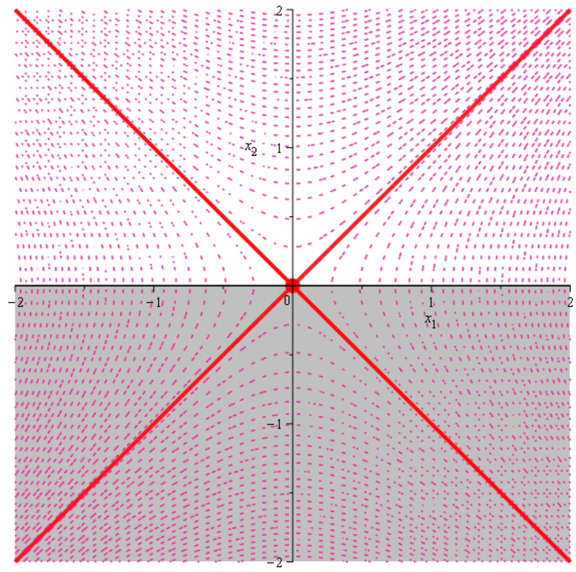
$$\begin{gathered}-x_1\, x_2 \to \min, \\ x_1 + x_2 \leqslant 6, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.\end{gathered}$$



Pozor! Obrázek generován z MAPLE, který dává chybné řešení!

Případ s $y_i^* = 0$ pro nějaké $i \in S(x^*)$, ukazuje jistou „degenerovanost“ tohoto omezení. Např.

$$\frac{x_1^2 - x_2^2}{2} \to \min, \qquad x_2 \leqslant 0.$$



Definice 4.3.1

Vektor $y^* \in Q$ se nazývá *Kuhnovým–Tuckerovým vektorem* (K–T vektorem) úlohy (4.1) & (4.2), jestliže

$$f^* \leqslant f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = L(x, y^*) \quad \forall x \in P, \tag{4.3.0}$$

kde $f^* := \inf_{x \in X} f(x)$ je hodnota úlohy (4.1) & (4.2).

Existuje K–T vektor vždy? Uvažme např. úlohy

$$\begin{aligned} &\text{(i)} \quad -x^2 \to \min, \quad x = 0, \quad P = \mathbb{R}, \\ &\text{(ii)} \quad x - 1 \to \min, \quad x^2 - 1 = 0, \quad P = \mathbb{R}_+, \\ &\Big(\text{(iii)} \quad e^x \to \min, \quad x \leqslant 0, \quad P = \mathbb{R}\Big). \end{aligned}$$

Proto k zaručení existence K–T vektoru potřebujeme některé dodatečné podmínky, které lze získat z regulárních modifikací Věty 2.3.12.

Věta 4.3.2

Nechť úloha (4.1) & (4.2) 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 $\bar{x} \in P$ takové, že $g_i(\bar{x}) < 0$ pro $i = 1, \dots, m$,

(ii) (lineární) množina P je polyedr, funkce $f, g_1, \dots, g_k$ jsou afinní a $X \neq \emptyset$.

Pak existuje K–T vektor úlohy (4.1) & (4.2).

Úloha konvexního programování doplněná o některou z dodatečných podmínek uvedených ve Větě 4.3.2 se nazývá regulární úloha konvexního programování. Podmínka (ii) se ale drobně liší od podmínky (iii) ve Větě 4.2.3 (zde nám LNZ nepomůže; existence vs. neexistence a řešitelnost).

Terminologické „zamyšlení“:

$$\text{omezení} \quad \rightsquigarrow \quad \text{kvalifikované omezení} \quad \rightsquigarrow \quad \text{ostrá komplementarita}$$

úloha MP

$\wr\!\downarrow$

úloha KP

$\wr\!\downarrow$

regulární úloha KP (kvalifikované omezení)

$\wr\!\downarrow$

silně regulární úloha KP (existuje K–T vektor)

Důkaz Věty 4.3.2

Je-li $f^* = -\infty$, pak nerovnost (4.3.1) je splněna triviálně pro každé $y \in Q$. Nechť tedy $f^* > -\infty$ a platí podmínka (i). Uvažujme systém nerovností

$$f(x) - f^* < 0, \qquad g_i(x) < 0, \quad i \in \{1, \dots, m\}.$$

Tento systém nemá řešení na P (kdyby nějaké x splňovalo druhou část, pak $x \in X$ a první část vede ke sporu s definicí f^*). Vynecháme-li první nerovnost, pak systému jistě vyhovuje bod $\bar{x}$ (dle předpokladů). Pak z Věty 2.3.13 plyne existence $y_1^*, \dots, y_m^* \geqslant 0$ takových, že

$$f(x) - f^* + \sum_{i=1}^{m} y_i^* g_i(x) \geqslant 0 \quad \forall x \in P.$$

Tedy $y^* = (y_1, \dots, y_m)^\top$ je K–T vektorem úlohy (4.1) & (4.2). ■

Úloha (4.1) & (4.2) $\rightsquigarrow$ primární $\longleftrightarrow$ duální úloha = úloha konkávního programování.

Motivační příklad 1

Uvažme úlohu

$$4x_1 + 14x_2 \to \min,$$
$$x_1 + 2x_2 \geqslant 3, \qquad x_1 + 4x_2 \geqslant 4,$$
$$3x_1 + x_2 \geqslant 3, \qquad x_1 \geqslant 0, \qquad x_2 \geqslant 0.$$

Motivační příklad 2

Uvažme úlohu

$$x_1^2 + x_2^2 + 2x_1 \to \min,$$
$$x_1 + x_2 = 0.$$

Definice 4.3.3

Nechť $y \in Q$. Definujme funkci

$$\varphi(y) := \inf_{x \in P} L(x, y) = \inf_{x \in P} \Big[f(x) + \sum_{i=1}^{m} y_i\, g_i(x)\Big]$$

a množinu (tzv. efektivní definiční obor)

$$Y := \big\{y \in Q \mid \varphi(y) > -\infty\big\}.$$

Pak úloha

$$\varphi(y) \to \max, \qquad y \in Y, \tag{4.3.1}$$

se nazývá *duální úlohou* k úloze (4.1) & (4.2). Číslo

$$\varphi^* := \sup_{y \in Y} \varphi(y)$$

se nazývá *hodnotou duální úlohy* (4.3.1).

Věta 4.3.4

Úloha (4.3.1) je úlohou konkávního programování, tj. množina Y je konvexní a funkce φ je konkávní na Y.

Důkaz. Nechť $y_1, y_2 \in Y$ a $\lambda \in [0,1]$ jsou libovolná. Potom platí

$$\begin{aligned}\varphi(\lambda y_1 + (1-\lambda)y_2) &= \inf_{x\in P} L(x, \lambda y_1 + (1-\lambda)y_2) = \\ &= \inf_{x\in P}[\lambda L(x,y_1) + (1-\lambda)L(x,y_2)] \geqslant \\ &\geqslant \lambda \inf_{x\in P} L(x,y_1) + (1-\lambda)\inf_{x\in P} L(x,y_2) = \\ &= \lambda\varphi(y_1) + (1-\lambda)\varphi(y_2),\end{aligned}$$

tj. funkce φ je konkávní. Současně odtud také plyne, že pro $y_1, y_2 \in Y$ (tj. $\varphi(y_1) > -\infty$ a $\varphi(y_2) > -\infty$) je také $\lambda y_1 + (1-\lambda)y_2 \in Y$, neboť $\varphi(\lambda y_1 + (1-\lambda)y_2) \geqslant \lambda\varphi(y_1) + (1-\lambda)\varphi(y_2) > -\infty$, tj. množina Y je konvexní. ■

Příklad

Určeme duální úlohu k úloze LP, tj.

$$\langle c, x\rangle \to \min, \qquad x \in X,$$

kde množina X je určena podmínkami

$$\sum_{j=1}^{n} a_{ij}x_j \geqslant b_i, \qquad i \in \{1,\dots,k\}, \quad \text{tj. } \langle a_i, x\rangle \geqslant b_i,$$

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i \in \{k+1,\dots,m\}, \quad \text{tj. } \langle a_i, x\rangle = b_i,$$

$$x_j \geqslant 0, \quad j \in \{1,\dots,s\},$$

kde $c, a_i \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $s \in \{0,\dots,n\}$ jsou dány.

Existuje nějaký vztah mezi f^* a φ^*?

Věta 4.3.5 **(Slabá věta o dualitě)** Pro každé $x \in X$ a každé $y \in Q$ platí

$$f(x) \geqslant \varphi(y).$$

Zejména, pokud $X \neq \emptyset$ a $Y \neq \emptyset$, pak $f^* \geqslant \varphi^*$ (v případě $X = \emptyset$ a/nebo $Y = \emptyset$ je nerovnost splněna triviálně, neboť $\inf \emptyset = \infty$ a $\sup \emptyset = -\infty$).

Důkaz. Pro každé $x \in X$ a $y \in Q$ platí

$$f(x) \geqslant f(x) + \underbrace{\sum_{i=1}^{m} y_i\, g_i(x)}_{\leqslant 0} = L(x,y) \geqslant \inf_{x \in X} L(x,y) \geqslant \inf_{x \in P} L(x,y) \geqslant \varphi(y).$$

Tato nerovnost musí platit také v případě, že na levé straně vezmeme infimum pro $x \in X$ a na pravé straně supremum pro $y \in Y$. ■

Věta 4.3.5 říká, že rozdíl mezi hodnotou účelové funkce primární a duální úlohy je vždy nezáporný, tj. pro libovolná $x \in X \neq \emptyset$ a $y \in Y \neq \emptyset$ platí (v takovém případě jsou obě úlohy nutně řešitelné)

$$g(x,y) := f(x) - \varphi(y) \geqslant 0.$$

Pak číslo $g(x^*, y^*) := f^* - \varphi^*$ udává tzv. „optimální duální rozdíl" (*optimal duality gap*). Věta 4.3.5 také říká, že pro libovolné $y \in Q$ je hodnota $\varphi(y)$ dolní hranici minima účelové funkce úlohy (4.1) & (4.2).

Navíc odtud plyne velmi „jednoduchý" test řešitelnosti úlohy (4.1) & (4.2):

Je-li primární úloha (4.1) & (4.2) neohraničená (tj. $f^* = -\infty$), pak nutně $\varphi^* = -\infty$, tj. $Y = \emptyset \rightsquigarrow$ duální úloha (4.3.1) je nepřípustná.

Naopak, je-li duální úloha (4.3.1) neohraničená (shora, tj. $\varphi^* = \infty$), pak nutně $f^* = \infty$, tj. $X = \emptyset \rightsquigarrow$ primární úloha (4.1) & (4.2) je nepřípustná.

Celkem může nastat 9 možností ohledně přípustnosti, ohraničenosti a řešitelnosti primární a duální úlohy. Věta 4.3.5 vylučuje 3 možnosti a 2 možnosti potvrzuje (více zatím není možné rozhodnout):

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✔	✖	✖
PaO	?	?	✖
NP ($\varphi^* = -\infty$)	?	?	✔

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^*)

Věta 4.3.5 také dává velmi „jednoduchou“ postačující podmínku pro ověření optimality.

„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 (ekvivalence? viz Větu 4.3.8).

Duální rozdíl je také úzce spojen s existencí K–T vektoru. Jestliže duální rozdíl je nenulový, tj. $f^* > \varphi^*$, pak množina K–T vektorů musí být prázdná. Vskutku, připusťme, že existuje K–T vektor $y^* \in Q$, tj.

$$f^* \leqslant \inf_{x \in P} L(x, y^*) = \varphi(y^*) \leqslant \varphi^* < f^*$$

(v případě $X = \emptyset$ je $f^* = \infty$, a tudíž také $\varphi^* = \infty$ ↯, podobně pro $Y = \emptyset$ je $\varphi^* = -\infty$, a tudíž $f^* = -\infty$ ↯).

To ale znamená, že existence K–T vektoru zaručuje nulový duální rozdíl. Připomeňme, že situaci $f^* = -\infty$ jsme již vyřešili.

Věta 4.3.6

(Silná věta o dualitě) Nechť úloha (4.1) & (4.2) je regulární úlohou konvexního programování (viz Věta 4.3.2). 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 (4.3.1) je neprázdná a shodná s množinou všech K–T vektorů úlohy (4.1) & (4.2).

Samozřejmě nulového duálního rozdílu lze dosáhnout i za jiných předpokladů (např. P konvexní a uzavřená, $f, g_1, \dots, g_k$ konvexní a spojité na P, g_{k+1}, g_m afinní a množina řešení primární úlohy je neprázdná a ohraničená, pak $Y \neq \emptyset$ a $f^* = \varphi^*$).

Praktický význam tvrzení Věty 4.3.6 nemusí být na první pohled zřejmý, ale v některých případech může být duální úloha jednodušší (za daných předpokladů jsou obě úlohami KP) – některé algoritmy LP a QP jsou založeny právě na řešení duální úlohy.

Záměna maxima (suprema) a minima (infima) ⇝ von Neumannova věta o „minimaxu“ ⇝ sedlový bod (viz později) ⇝ teorie her.

Důkaz Věty 4.3.6

Jelikož máme regulární úlohu KP, podle Věty 4.3.2 existuje K–T vektor $y^* \in Q$ úlohy (4.1) & (4.2), a tudíž (opět)

$$f^* \leqslant \inf_{x \in P} L(x, y^*) = \varphi(y^*) \leqslant \varphi^*.$$

Jelikož $f^* > -\infty$, je také $\varphi^* > -\infty$, a tedy $y^* \in Y \neq \emptyset$. Současně předpoklady zaručují $X \neq \emptyset$. Proto z Věty 4.3.5 máme $f^* \geqslant \varphi^*$, takže celkem $f^* = \varphi^*$.

Nechť nyní $y^* \in Y$ je řešením duální úlohy (4.3.1). Potom

$$f^* = \varphi^* = \varphi(y^*) = \inf_{x \in P} L(x, y^*) \leqslant f(x) + \sum_{i=1}^{m} y_i^* g_i(x), \quad \forall x \in P,$$

tj. y^* je K–T vektorem úlohy (4.1) & (4.2).

Konečně, nechť $y^* \in Q$ je K–T vektorem úlohy (4.1) & (4.2). Pak platí

$$\varphi^* = f^* \leqslant f(x) + \sum_{i=1}^{m} y_i^* g_i(x) = L(x, y^*), \quad \forall x \in P.$$

Odtud plyne $\varphi^* \leqslant \inf_{x \in P} L(x, y^*) = \varphi(y^*)$, a tudíž z definice φ^* dostáváme rovnost $\varphi^* = \varphi(y^*)$. To ale znamená, že y^* je řešením úlohy (4.3.1). ■

Teorie (Lagrangeovy) duality

Z Věty 4.3.6 bezprostředně vyplývají dvě důležité implikace.

Důsledek 4.3.7

Nechť úloha (4.1) & (4.2) je regulární úlohou konvexního programování (viz Věta 4.3.2).

(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 4.3.5 plyne $f^* \geqslant \varphi^* > -\infty$, takže úloha (4.1) & (4.2) je řešitelná podle Věty 4.3.6.

(ii) Jestliže platí $Y = \emptyset$, pak nutně $f^* = -\infty$ (v případě $f^* > -\infty$ totiž dostaneme spor s Větou 4.3.6). ■

Z Vět 4.3.5 a 4.3.6 a z Důsledku 4.3.7 vyplývá, že v případě regulární úlohy KP mohou nastat pouze dvě možnosti (regularita $\rightsquigarrow X \neq \emptyset$):

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✖	✖	✖
PaO	✖	✔	✖
NP ($\varphi^* = -\infty$)	✖	✖	✔

V případě, kdy primární úloha může být i nepřípustná máme 4+1 možnost:

DÚ \ PÚ	NP ($f^* = \infty$)	PaO	NO ($f^* = -\infty$)
NO ($\varphi^* = \infty$)	✔	✖	✖
PaO	?✔?	✔	✖
NP ($\varphi^* = -\infty$)	✔	✖	✔

Pozor: v úlohách LP a QP varianta PÚ=NP & DÚ=PaO není možná, neboť duální úloha je téhož typu ($\rightsquigarrow$ záměna role primární a duální úlohy). Nicméně obecně tato situace může nastat (např. $x \to \min$ & $x \leqslant 0$ & $P = \mathbb{R}_{++}$).

S využitím teorie duality můžeme získat nutné a postačující podmínky pro řešení regulární úlohy KP bez předpokladu diferencovatelnosti.

Věta 4.3.8

(Kuhnova–Tuckerova v nediferenciálním tvaru) Nechť úloha (4.1) & (4.2) je regulární úlohou konvexního programování (viz Věta 4.3.2). 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

$$L(x^*, y^*) = \min_{x \in P} L(x, y^*), \tag{4.3.2}$$

$$y_i^* g_i(x^*) = 0, \quad i \in \{1, \dots, m\}. \tag{4.3.3}$$

Navíc množina takovýchto vektorů $y^* \in Q$ splývá s množinou řešení duální úlohy (4.3.1) (a podle Věty 4.3.6 tedy také s množinou K–T vektorů úlohy (4.1) & (4.2)).

Jsou-li navíc funkce $f, g_1, \dots, g_m$ diferencovatelné v bodě x^*, pak podmínka (4.3.2) je ekvivalentní s (4.2.3) pro $y_0^* = 1$, zatímco (4.3.3) odpovídá (4.2.4). Tudíž Věta 4.3.8 je skutečně zobecněním KKT věty (Věty 4.2.3) pro případ nediferencovatelných funkcí. Koncept K–T vektoru je zobecněním Lagrangeových multiplikátorů (splývá za podmínek Věty 4.2.3).

Důkaz Věty 4.3.8

(i) „$\Longrightarrow$“ Nechť x^* je řešením úlohy (4.1) & (4.2), tj. $f(x^*) = f^*$, a $y^* \in Q$ je K–T vektor (existuje díky regulárnosti úlohy). Pak podle druhé části Věty 4.3.6 je y^* řešením duální úlohy (4.3.1), tj. $\varphi(y^*) = \varphi^*$. Současně podle první části Věty 4.3.6 dostáváme $f(x^*) = f^* = \varphi^* = \varphi(y^*)$.

„$\Longleftarrow$“ Nechť $f(x^*) = \varphi(y^*)$ pro nějaké $x^* \in X$ a $y^* \in Q$. Jelikož platí $f(x^*) \leqslant f(x^*) = \varphi(y^*) \leqslant \varphi(y^*)$, plyne z Věty 4.3.5 $f^* = f(x^*)$ a $y^* = \varphi(y^*)$, tj. x^* je řešením úlohy (4.1) & (4.2) a y^* je řešením duální úlohy (4.3.1).

Důkaz Věty 4.3.8 (pokr.)

(ii) „$\Longrightarrow$“ Nechť x^* je řešením úlohy (4.1) & (4.2). Pak podle předchozí části existuje $y^* \in Q$ takové, že $f(x^*) = \varphi(y^*)$, tj.

$$f(x^*) = \varphi(y^*) \overset{\mathrm{D}.4.3.3}{=} \inf_{x \in P} L(x, y^*) \leqslant f(x) + \sum_{i=1}^{m} y_i^* g_i(x) \quad \forall x \in P.$$

Odtud volbou $x = x^*$ dostaneme $\sum_{i=1}^{m} y_i^* g_i(x^*) \geqslant 0$. Jelikož ale $g_i(x^*) \leqslant 0$ a $y_i^* \geqslant 0$ pro $i = 1, \dots, k$, musí platit (4.3.3), z čehož dále vyplývá

$$L(x^*, y^*) = f(x^*) + \underbrace{\sum_{i=1}^{m} y_i^* g_i(x^*)}_{=0} = f(x^*) = \varphi(y^*) \leqslant L(x, y^*) \quad \forall x \in P$$

neboli

$$L(x^*, y^*) = \min_{x \in P} L(x, y^*)$$

„$\Longleftarrow$“ Nechť pro nějaké $x^* \in X$ a $y^* \in Q$ platí (4.3.2) a (4.3.3). Potom

$$f(x^*) = f(x^*) + \underbrace{\sum_{i=1}^{m} y_i^* g_i(x)}_{=0 \ \text{ dle (4.3.3)}} = L(x^*, y^*) \overset{(4.3.2)}{=} \min_{x \in P} L(x, y^*) = \inf_{x \in P} L(x, y^*) = \varphi(y^*)$$

Pak z části (i) vyplývá, že x^* je řešením úlohy (4.1) & (4.2). ∎

Příklad

Určeme duální úlohu pro

$$f(x) = e^{-x} \to \min, \qquad x \leqslant 1,$$

a ověřme platnost vztahu duality $f^* = \varphi^*$.

?? Co je duální úloha k duální úloze ??

Tuto část zakončíme ještě jednou alternativou k původní KKT větě, s jejíž pomocí snadno získáme odpověď na otázku:

jak s pomocí řešení duální úlohy získáme řešení primární úlohy?

Definice 4.3.9

Bod $[x^*, y^*] \in P \times Q$ se nazývá *sedlovým bodem* Lagrangeovy funkce $L(x, y)$ úlohy (4.1) & (4.2) na $P \times Q$, jestliže

$$L(x^*, y) \leqslant L(x^*, y^*) \leqslant L(x, y^*) \quad \forall x \in P,\ y \in Q,$$

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 (4.1) & (4.2)?

Věta 4.3.10

(Kuhnova–Tuckerova pro sedlový bod) Nechť úloha (4.1) & (4.2) je regulární úlohou konvexního programování (viz Věta 4.3.2). Pak bod $x^* \in P$ je řešením úlohy (4.1) & (4.2) právě tehdy, když existuje $y^* \in Q$ takové, že $[x^*, y^*]$ je sedlovým bodem Lagrangeovy funkce $L(x, y)$ úlohy (4.1) & (4.2) na $P \times Q$.

Důkaz Věty 4.3.10

Důkaz je založen na Větě 4.3.8(ii).

„$\Longrightarrow$" Nechť $x^* \in P$ je řešením úlohy (4.1) & (4.2). Pak $x^* \in X$ a podle Věty 4.3.8(ii) existuje $y^* \in Q$ takový, že

$$L(x^*, y^*) = \min_{x \in P} L(x, y^*) \quad \& \quad y_i^* g_i(x^*) = 0, \quad i \in \{1, \dots, k\},$$

tj. z první části máme $L(x^*, y^*) \leqslant L(x, y^*)$ pro každé $x \in P$. Současně s využitím druhé části dostaneme

$$L(x^*, y^*) = f(x^*) + \underbrace{\sum_{i=1}^{m} y_i^* g_i(x^*)}_{=0} = f(x^*) \geqslant f(x^*) + \underbrace{\sum_{i=1}^{m} y_i\, g_i(x^*)}_{\leqslant 0} = L(x^*, y) \quad \forall y \in Q.$$

To dohromady s první částí ukazuje, že bod $[x^*, y^*]$ je sedlovým bodem.

Důkaz Věty 4.3.10 (pokr.)

„$\Longleftarrow$" Nechť nyní $[x^*, y^*]$ je sedlovým bodem. Pak $L(x^*, y^*) = \min_{x \in P} L(x, y^*)$ a zároveň $L(x^*, y^*) \geqslant L(x^*, y)$ pro každé $y \in Q$, tj.

$$f(x^*) + \sum_{i=1}^{m} y_i^* g_i(x^*) \geqslant f(x^*) + \sum_{i=1}^{m} y_i\, g_i(x^*) \quad \forall y \in Q$$

neboli

$$\sum_{i=1}^{m} y_i^* g_i(x^*) \geqslant \sum_{i=1}^{m} y_i\, g_i(x^*) \quad \forall y \in Q.$$

Protože tato nerovnost platí pro všechna $y \in Q$, nutně také pro $y = 0 \in Q$, což dává $\sum_{i=1}^{m} y_i^* g_i(x^*) \geqslant 0$. Současně volba $y = y^* + e_i \in Q$ pro $i \in \{1, \dots, k\}$ dává $0 \geqslant g_i(x^*)$ a $y = y^* \pm e_i \in Q$ pro $i \in \{k+1, \dots, m\}$ dává dokonce $0 \geqslant g_i(x^*)$ & $0 \leqslant g_i(x^*)$, tj. $g_i(x^*) = 0$. Proto $x^* \in X$ a $y_i^* g_i(x^*) \leqslant 0$ 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 4.3.8(ii), což znamená, že bod $x^* \in P$ je řešením úlohy (4.1) & (4.2). ■

Regularita ve Větě 4.3.10

Implikace $\Longleftarrow$ ve Větě 4.3.10 platí i bez předpokladu regularity úlohy (4.1) & (4.2). Bez tohoto předpokladu dokonce platí i to, že sedlový bod splňuje (4.3.2) a (4.3.3).

Sedlový bod a duální úloha

Analogické tvrzení jako ve Větě 4.3.10 platí také pro duální úlohu:

> Bod $y^* \in Y$ je řešením duální úlohy (4.3.1) právě tehdy, když existuje $x^* \in P$ takový, že $[x^*, y^*] \in P \times Q$ je sedlovým bodem Lagrangeovy funkce regulární úlohy (4.1) & (4.2).

V literatuře bývají Věty 4.3.8 a 4.3.10 uváděny společně.

Důsledek 4.3.11

Nechť úloha (4.1) & (4.2) je regulární úlohou konvexního programování (viz Věta 4.3.2).

(i) Body $x^* \in X$ a $y^* \in Y$ řeší úlohy (4.1) & (4.2) a (4.3.1) právě tehdy, když platí (i) nebo (ii) z Věty 4.3.8.

(ii) Body $x^* \in X$ a $y^* \in Y$ řeší úlohy (4.1) & (4.2) a (4.3.1) právě tehdy, když $[x^*, y^*]$ je sedlový bod Lagrangeovy funkce úlohy (4.1) & (4.2).

Nyní se budeme věnovat závislosti řešení úloh MP na parametrech, tj. situaci kdy úloha (4.1) & (4.2) bude obsahovat i nějaké parametry $\rightsquigarrow$ parametrické programování (viz QP). Parametry mohou vystupovat v účelové funkci a/nebo ve funkcionálních omezeních:

(i) v účelové funkci $\rightsquigarrow$ prodejní ceny, nákupní ceny, míru užitku, ...

(ii) v omezeních $\rightsquigarrow$ dostupnost skladových zásob, výrobní kapacita, maximální investice, přijatelná míra rizika, ...

Nebudeme se věnovat obecné úloze MP. Nejdříve uvážíme úlohu MP s omezeními pouze ve tvaru rovností ale s obecnou závislostí na parametrech. Poté se podíváme na úlohu s omezeními ve tvaru nerovností ale s parametry pouze v podobě absolutních členů ve funkcích zadávajících jednotlivá omezení.

Motivace: interpretace řešení duální úlohy v LP $\rightsquigarrow$ K–T vektor $\rightsquigarrow$ Lagrangeovy multiplikátory.

Začněme s jednoduchou úlohu pro 2 proměnné a 1 omezení ve tvaru rovnosti, tj.

$$f(x_1, x_2) \to \min, \qquad g(x_1, x_2) = c, \tag{4.4.1}$$

která je jednoznačně řešitelná pro každé $c \in \mathbb{R}$ (pro jednoduchost). Potom řešením (4.4.1) je dvojice $x_1^*(c)$ a $x_2^*(c)$ a předpokládejme, že $x_1^*(\cdot)$ a $x_2^*(\cdot)$ jsou diferencovatelné vzhledem k c. Hodnota úlohy (4.4.1) také závisí na c, tj.

$$f^*(c) = f(x_1^*(c), x_2^*(c)),$$

stejně jako odpovídající Lagrangeův multiplikátor $y^*(c)$. Je-li navíc $f^*(c)$ také diferencovatelná vzhledem k c, pak platí

$$\frac{d}{dc} f^*(c) = f_{x_1}(x_1^*(c), x_2^*(c))\, \frac{dx_1^*(c)}{dc} + f_{x_2}(x_1^*(c), x_2^*(c))\, \frac{dx_2^*(c)}{dc}. \tag{$*$}$$

Současně z Lagrangeova principu máme

$$\left.\begin{aligned} f_{x_1}(x_1^*(c), x_2^*(c)) + y^*(c)\, g_{x_1}(x_1^*(c), x_2^*(c)) = 0, \\ f_{x_2}(x_1^*(c), x_2^*(c)) + y^*(c)\, g_{x_2}(x_1^*(c), x_2^*(c)) = 0. \end{aligned}\right\} \tag{$\Box$}$$

Konečně derivováním rovnosti $g(x_1^*(c), x_2^*(c)) = c$ obdržíme

$$g_{x_1}(x_1^*(c), x_2^*(c))\, \frac{dx_1^*(c)}{dc} + g_{x_2}(x_1^*(c), x_2^*(c))\, \frac{dx_2^*(c)}{dc} = 1. \tag{$\triangle$}$$

Zkombinováním ($*$), ($\square$) a ($\triangle$) dostaneme

$$\frac{d}{dc} f^*(c) \overset{(*),(\square)}{=} -y^*(c)\, g_{x_1}(x_1^*(c), x_2^*(c))\, \frac{dx_1^*(c)}{dc} - y^*(c)\, g_{x_2}(x_1^*(c), x_2^*(c))\, \frac{dx_2^*(c)}{dc} =$$

$$= -y^*(c) \left[g_{x_1}(x_1^*(c), x_2^*(c))\, \frac{dx_1^*(c)}{dc} + g_{x_2}(x_1^*(c), x_2^*(c))\, \frac{dx_2^*(c)}{dc} \right] =$$

$$\overset{(\triangle)}{=} -y^*(c).$$

(Tvrzení zůstane v platnosti i pokud se omezíme pouze na nějaké okolí daného $c_0 \in \mathbb{R}$.)

Takže hodnota $-y^*(c)$ udává míru změny hodnoty $f^*(c)$, neboť platí

$$f^*(c + \Delta c) \approx f^*(c) - y^*(c)\, \Delta c.$$

V ekonomii mnohdy c vyjadřuje dostupné zásoby nějakého zdroje (vč. času, peněz apod.) a funkce f užitek nebo zisk. Potom hodnota $-y^*(c)\,\Delta c$ měří přibližnou míru změny užitku/zisku, pokud změníme zásoby o Δc $\rightsquigarrow$ stínová cena (tržní cena? čas, živiny apod.). Takže zejména při volbě $\Delta c = 1$ zjistíme, že $-y^*(c)$ určuje změnu optimálního zisku/užitku $f^*(c)$ při navýšení zásob o 1 jednotku.

Analogicky dostaneme při větším počtu omezení

$$f^*(c + \Delta c) \approx f^*(c) - y_1^*(c)\, \Delta c_1 - \cdots - y_m^*(c)\, \Delta c_m.$$

Příklad

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?

Věta 4.4.1

(Věta o obálce) Mějme úlohu

$$f(x, r) \to \min, \quad g_1(x, r) = 0, \ \ldots, \ g_m(x, r) = 0, \qquad (*)$$

kde $x \in \mathbb{R}^n$, $r \in \mathbb{R}^k$, $f, g_1, \ldots, g_m \in C^1$. Připusťme, že pro každou hodnotu parametru r má úloha ($*$) jediné řešení, které označíme $x^*(r)$. Potom hodnota úlohy ($*$) je

$$f^*(r) = f(x^*(r), r).$$

Je-li $x^*(r)$ diferencovatelná vzhledem k r a Jacobiho matice $D_x G(x^*(r), r) \in \mathbb{R}^{m \times n}$ má plnou hodnost m, pak platí

$$\frac{\partial}{\partial r_i} f^*(r) = \frac{\partial f}{\partial r_i}(x^*(r), r) + \sum_{j=1}^{m} y_j^*(r)\, \frac{\partial g_j}{\partial r_i}(x^*(r), r).$$

Proč „Věta o obálce“? Obrázek.

Jenže předpoklady Věty 4.4.1 jsou celkem silné ($\forall r$ existuje jediné řešení). Zjednodušení úlohy $\rightsquigarrow$ oslabení požadavků.

Věta 4.4.2

Nechť $f, g_1, \dots, g_m \in C^2$ a x^* je lokálním řešením úlohy

$$f(x) \to \min, \quad g_1(x) = 0, \; \dots, \; g_m(x) = 0$$

s odpovídající Lagrangeovým multiplikátorem y^*. 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. $D_x G(x^*) \in \mathbb{R}^{m \times n}$ má plnou hodnost m. Uvažme úlohu parametrického programování

$$f(x) \to \min, \quad G(x) = u \tag{$*$}$$

pro parametr $u \in \mathbb{R}^m$. 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 ($*$) 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 ($*$) vzhledem k u, tj. klademe $f^*(u) := f(x^*(u))$.

Nyní se zaměříme na úlohu s omezeními ve tvaru nerovností (to je z praktického pohledu přeci jen užitečnější; navíc 1 rovnost=2 nerovnosti), kde parametry budou vystupovat pouze ve formě absolutních členů (viz Věta 4.4.2).

Budeme mít úlohu závislou na m-tici parametrů $b = (b_1, \dots, b_m)^\top \in \mathbb{R}^m$, tj.

$$f(x) \to \min, \qquad x \in X(b) := \left\{x \in P \subseteq \mathbb{R}^n \mid g_i(x) \leqslant b_i, \; i \in 1, \dots, m\right\}. \tag{4.4.2}$$

Ještě budeme potřebovat označení

$$G(x) := \big(g_1(x), \dots, g_m(x)\big)^\top, \qquad X(b) := \left\{x \in P \mid G(x) \leqslant b\right\},$$

$$B := \left\{b \in \mathbb{R}^m \mid X(b) \neq \emptyset\right\}, \qquad F(b) := \inf_{x \in X(b)} f(x), \; b \in B,$$

$$Y(b) := \left\{y \in \mathbb{R}^m \mid y \geqslant 0, \; F(b) \leqslant f(x) + \langle y, \, G(x) - b\rangle \;\; \forall x \in P\right\},$$

$$\partial F(b) := \left\{a \in \mathbb{R}^m \mid F(b') - F(b) \geqslant \langle a, \, b' - b\rangle \;\; \forall b' \in B\right\}.$$

Souvislost s dřívějším značením: $X = X(0)$ a $f^* = F(0)$. Množina $Y(b)$ je množinou K–T vektorů úlohy (4.4.2). Množina $\partial F(b)$ je subdiferenciálem funkce $F(b)$.

A jak tedy závisí úloha (4.4.2) na parametru b?

Věta 4.4.3

Nechť množina $P \subseteq \mathbb{R}^n$ je konvexní, funkce $f, g_1, \dots, g_m$ 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 na B,

(iii) platí $\partial F(b) = -Y(b)$ pro všechna $b \in B$.

Poznámky k Větě 4.4.3

(i) Předpoklady Věty 4.4.3 říkají, že
- $0 \in B \iff X(0) \neq \emptyset$: úloha (4.4.2) s $b = 0$ je přípustná,
- $F(0) > -\infty$: úloha (4.4.2) s $b = 0$ má konečné řešení,
- $Y(0) \neq \emptyset$: množina K–T vektorů úlohy (4.4.2) s $b = 0$ je neprázdná (pozor – toto není splněno automaticky, viz Věta 4.3.2).

(ii) Je-li F dokonce diferencovatelná v b, pak $\partial F(b)$ je jednoprvková, obsahuje pouze $\operatorname{grad}^{\top} F(b)$ a tento vektor je roven $(-1)\times$ K–T vektoru, což je analogie Věty o obálce (Věta 4.4.2).

(iii) Ve Větě 4.3.6 jsme charakterizovali K–T vektory úlohy (4.1) & (4.2) pomocí řešení duální úlohy (4.3.1). Část (iii) předchozí věty jim dává ještě jinou charakteristiku — pomocí subgradientu hodnoty úlohy parametrického programování (4.4.2): v případě regulární úlohy KP dostáváme zkombinováním těchto dvou výsledků $\partial F(b) = -Y^*(b)$, kde $Y^*(b)$ je množina řešení duální úlohy (viz Věta 4.3.6 a LP).

Důkaz Věty 4.4.3(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 $\lambda \in [0,1]$ a 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

$$G(x) = G(\lambda\widetilde{x} + (1-\lambda)\widehat{x}) \overset{G..\text{konvx}}{\leqslant} \lambda G(\widetilde{x}) + (1-\lambda)\,G(\widehat{x}) \leqslant \lambda\widetilde{b} + (1-\lambda)\widehat{b}. \qquad (*)$$

Tedy $x \in X(\lambda\widetilde{b} + (1-\lambda)\widehat{b})$, tj. $X(\lambda\widetilde{b} + (1-\lambda)\widehat{b}) \neq \emptyset$, tj. $\lambda\widetilde{b} + (1-\lambda)\widehat{b} \in B$.

Důkaz Věty 4.4.3(ii)

Nechť $y^* \in Y(0)$. Pak $y^* \geqslant 0$ a $F(0) \leqslant f(x) + \langle y^*, G(x) \rangle$ pro $\forall x \in P$. Protože $y^* \geqslant 0$, platí $F(0) \leqslant f(x) + \langle y^*, b \rangle$ pro libovolné $b \in B$ a $x \in X(b)$, tudíž také

$$F(0) \leqslant F(b) + \langle y^*, b \rangle .$$

Protože $F(0) > -\infty$, je zřejmé, že také $F(b) > -\infty$, a tedy funkce $F(b)$ je konečná ($F(b) = \infty$ pouze pro $f \equiv \infty$ $\lightning$).

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 $(*)$, tedy

$$F(b) \leqslant f(x) \leqslant \lambda f(\widetilde{x}) + (1-\lambda) f(\widehat{x}).$$

Protože $\widetilde{x}, \widehat{x}$ byla zvolena libovolně, plyne odtud

$$F(b) \leqslant \lambda F(\widetilde{b}) + (1-\lambda) F(\widehat{b}),$$

tedy funkce F je konvexní.

Je-li $\widetilde{b}, \widehat{b} \in B$ a $\widetilde{b} \leqslant \widehat{b}$ (po složkách), pak zřejmě $X(\widetilde{b}) \subseteq X(\widehat{b})$. Proto $F(\widetilde{b}) \geqslant F(\widehat{b})$, a tedy funkce F je nerostoucí.

Důkaz Věty 4.4.3(iii)

Nechť $b \in B$ splňující $\partial F(b) \neq \emptyset$ je libovolné a nechť $y^* \in \partial F(b)$, tj.

$$F(b') - F(b) \geqslant \langle y^*, b' - b \rangle \quad \forall b' \in B. \tag{4.4.3}$$

Jelikož F je nerostoucí dle (ii), plyne odtud $y^* \leqslant 0$. Nechť dále $x \in P$ je libovolné a položme $b' := G(x)$. Pak $b' \in B$, $x \in X(b')$, a tudíž $F(b') \leqslant f(x)$ dle definice $F(b)$. Proto z (4.4.3) máme

$$F(b) \leqslant F(b') - \langle y^*, b' - b \rangle \leqslant f(x) + \langle -y^*, b' - b \rangle = f(x) + \langle -y^*, G(x) - b \rangle . \tag{$*$}$$

Jelikož $x \in P$ bylo zvoleno libovolně, platí nerovnost $(*)$ pro každé $x \in P$, a tudíž $-y^* \in Y(b)$ (dle definice $Y(b)$ a z faktu $-y^* \geqslant 0$).

Naopak, nechť $b \in B$ splňující $Y(b) \neq \emptyset$ je libovolné a nechť $-y^* \in Y(b)$, tj. $-y^* \geqslant 0$ a $F(b) \leqslant f(x) + \langle -y^*, G(x) - b \rangle$ pro každé $x \in P$. Potom pro každé $b' \in B$ a $x \in X(b')$ dostáváme

$$F(b) \leqslant f(x) + \langle -y^*, G(x) - b \rangle \leqslant f(x) + \langle -y^*, b' - b \rangle \implies F(b) \leqslant F(b') + \langle -y^*, b' - b \rangle .$$

Vzhledem k libovolnosti $b' \in B$, platí tato nerovnost pro každé $b' \in B$, a tudíž $y^* \in \partial F(b)$.

Jelikož z těchto dvou „implikací“ také vyplývá, že množiny $\partial F(b)$ a $Y(b)$ musí být současně prázdné/neprázdné (jistě $\partial F(b) \neq \emptyset$ pro každé $b \in \operatorname{ri} B$, ale jinak? – je možné $Y(b) = \emptyset$, neboť explicitně nepožadujeme regulární úlohu KP), platí $\partial F(b) = -Y(b)$. ∎

S pomocí Věty 4.4.3 můžeme získat několik důležitých vlastností původní úlohy MP, tj. úlohy (4.4.2) s $b = 0$.

Důsledek 4.4.4

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 $\bar{x} \in P$ takové, že $G(\bar{x}) < 0$ (viz Slaterovu podmínku ve Větě 4.3.2). 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 (4.4.2) s $b = 0$ (při splnění uvedených předpokladů) existuje více K–T vektorů $\iff$ funkce F není v bodě $b = 0$ diferencovatelná.

Důkaz. Protože existuje $\bar{x} \in P$ takové, že $G(\bar{x}) < 0 = b$, je zřejmé, že $b = 0 \in \operatorname{int} B$ (neboť $G(\bar{x}) \leqslant b$ pro všechna dostatečně malá b). Tvrzení pak plyne z Věty 4.4.3 a (i) Věty 2.4.1, (ii) Věty 2.5.6 a (iii) Věty 2.5.7. ■

Příklad

V závislosti na parametru b určeme hodnotu $F(b)$ parametrické úlohy
$$f(x) = x^2 + 2x + 2 \to \min, \quad x \leqslant b.$$
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.

Ukážeme si (alespoň) jeden ekonomicky motivovaný příklad založený na předchozích výsledcích. K tomu bude potřeba ještě následující důsledek.

Důsledek 4.4.5

(i) Nechť jsou splněny předpoklady Věty 4.4.3 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 $\alpha \geqslant 0$.

(ii) Nechť jsou splněny předpoklady Důsledku 4.4.4 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 $\alpha > 0$.

Příklad

Ekonomická interpretace K–T vektoru (viz Lagrangeovy multiplikátory $\rightsquigarrow$ stínová cena).

Fenchel vs. K–T

Celý výklad byl založen na K–T vektoru. Nyní si ukážeme, jak lze odvodit např. vztah duality $f^* = \varphi^*$ pomocí Fenchelovy transformace. Uvažme úlohu

$$f(x) \to \min, \quad g_i(x) + b_i \leqslant 0, \quad i \in \{1, \dots, m\} \quad x \in P \subseteq \mathbb{R}^n$$

a nechť $F(b)$ je hodnota této úlohy. Pak pro konjugovanou funkci platí

$$\begin{aligned} F^*(y) &= \sup_{b \in B} \left\{ \langle y, b \rangle - F(b) \right\} = \sup_{b \in \mathbb{R}^m} \Big\{ \langle y, b \rangle - \inf_{\substack{x \in P \\ G(x)+b \leqslant 0}} f(x) \Big\} = \\ &= \sup_{b \in \mathbb{R}^m} \Big\{ - \inf_{\substack{x \in P \\ G(x) \leqslant -b}} \big(f(x) - \langle y, b \rangle \big) \Big\} = \sup_{b \in \mathbb{R}^m} \sup_{\substack{x \in P \\ G(x) \leqslant -b}} \left\{ \langle y, b \rangle - f(x) \right\} = \\ &= \sup_{x \in P} \sup_{b \leqslant -G(x)} \left\{ \langle y, b \rangle - f(x) \right\} = \\ &= \begin{cases} \sup_{x \in P} \left\{ - \langle y, G(x) \rangle - f(x) \right\} = -\inf_{x \in P} \left\{ f(x) + \langle y, G(x) \rangle \right\}, & y \geqslant 0, \\ \infty, & y < 0. \end{cases} \end{aligned}$$

Fenchel vs. K–T (pokr.)

Tedy $F^*(y) = -\varphi(y)$ a duální úlohu (4.3.1) lze psát jako

$$-F^*(y) \to \max, \quad y \geqslant 0.$$

Navíc platí

$$F^{**}(0) = \sup_{y \geqslant 0} \left\{ \langle 0, y \rangle - F^*(y) \right\} = \sup_{y \geqslant 0} \left\{ - F^*(y) \right\} = \sup_{y \geqslant 0} \left\{ \varphi(y) \right\} = \varphi^*.$$

Pokud je funkce F spojitá v 0, platí podle Věty 2.6.6 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 2.4.1 a 4.4.3(ii).

Konec.