

# M5170: MATEMATICKÉ PROGRAMOVÁNÍ

PETR ZEMÁNEK (MASARYKOVA UNIVERZITA, BRNO)

## Kapitola 4: Základy matematického programování

(verze: 12. prosince 2022)



### VYMEZENÍ ZÁKLADNÍCH POJMŮ

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

$$f(x) \rightarrow \min, \quad x \in X, \quad (4.1)$$

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

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

tj.

$$X = P \cap \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad i = 1, \dots, k\} \cap \{x \in \mathbb{R}^n \mid g_j(x) = 0, \quad j = k + 1, \dots, m\},$$

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?).

## 4.1 OBECNÁ OPTIMALIZAČNÍ ÚLOHA

## 4.2 NUTNÉ A POSTAČUJÍCÍ PODMÍNKY OPTIMALITY V MP

## 4.3 TEORIE (LAGRANGEOVY) DUALITY

## 4.4 ANALÝZA CITLIVOSTI

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

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

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

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  $x^* \in \text{ri } X$ ?
- (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). Tvzení je asi víceméně zřejmé a důkaz relativně snadný, ale význam tohoto tvzení pro námi budovanou teorii je značný.

**Lemma 4.1.1**

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

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

*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  $f(x^* + th) < f(x^*) \downarrow$  ■

 Obrázek.

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

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

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

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  $f : X \rightarrow \mathbb{R}$  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

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

pro každé  $x \in X$ .

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

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

což je splněno pouze v případě  $\text{grad } f(x^*) = 0$  (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  $f : X \rightarrow \mathbb{R}$  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  $\langle \text{grad } f(x^*), x - x^* \rangle < 0$ . 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

$$0 > \langle \text{grad } f(x^*), x - x^* \rangle = f'_h(x^*) = \lim_{t \rightarrow 0^+} \frac{f(x^* + th) - f(x^*)}{t},$$

tj. existuje  $\alpha_0 > 0$  takové, že  $f(x^* + th) < f(x^*)$  pro každé  $t \in (0, \alpha_0)$ . Tedy celkem  $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) \neq \emptyset$   $\nabla$ .

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

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

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

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

V takovém případě obdržíme z Věty 4.1.3 následující důsledek.

**Důsledek 4.1.4**

Nechť funkce  $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$  je diferencovatelná na nějaké otevřené množině obsahující  $\mathbb{R}_+^n$ . Je-li  $x^* \in \mathbb{R}_+^n$  řešením úlohy

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}_+^n, \tag{4.1.3}$$

pak platí

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

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

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

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.

$$\langle \text{grad } f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in \mathbb{R}_+^n$$

neboli

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

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  $\frac{\partial f}{\partial x_j}(x^*) (\tilde{x}_j - x_j^*) < 0$  pro nějaký index  $j \in \{1, \dots, n\}$  a nějaké  $\tilde{x} \in \mathbb{R}_+^n$ , pak stačí vzít  $\tilde{x} \in \mathbb{R}_+^n$  takové, že  $\tilde{x}_i - x_i^* = 0$  (tj.  $\tilde{x}_i = x_i^*$ ) pro  $i \in \{1, \dots, n\} \setminus \{j\}$  a  $\tilde{x}_j = \tilde{x}_j \neq x_j^*$  (to lze, neboť  $X = \mathbb{R}_+^n$ ). Tím ale dostaneme spor s (\*). Tedy (\*) skutečně platí právě tehdy, když  $\frac{\partial f}{\partial x_i}(x^*) (x_i - x_i^*) \geq 0$  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

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

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

(i) Necht' pro  $i \in \{1, \dots, n\}$  platí levá strana ekvivalence.

(a) Je-li  $x_i^* = 0$ , pak  $x_i - x_i^* \geq 0$  pro všechna  $x \in X$ , takže nutně  $\frac{\partial f}{\partial x_i}(x^*) \geq 0$ , 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, necht' pro  $i \in \{1, \dots, n\}$  platí (4.1.4).

(a) Je-li  $x_i^* = 0$  a  $\frac{\partial f}{\partial x_i}(x^*) \geq 0$ , pak  $\frac{\partial f}{\partial x_i}(x^*) x_i = \frac{\partial f}{\partial x_i}(x^*) (x_i - x_i^*) \geq 0$  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  $\frac{\partial f}{\partial x_i}(x^*) (x_i - x_i^*) \geq 0$ .

Důkaz první části je tímto hotov.

## DŮKAZ DŮSLEDKU 4.1.4 (POKR.)

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

$$\langle \text{grad } f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in \mathbb{R}_+^n,$$

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

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

je (nutná/postačující) podmínka tvaru

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

(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 a  $\alpha_i < \beta_i$ , je (nutná/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

$$X = \{x \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = r\},$$

kde  $r > 0$  je dané číslo (jedná se o zobecnění jednotkového  $(n - 1)$ -simplexu), je (nutná/postačující) podmínka ve tvaru implikace

$$x_i^* > 0 \implies \frac{\partial f}{\partial x_i}(x^*) \leq \frac{\partial f}{\partial x_j}(x^*) \text{ pro } \forall j \in \{1, \dots, n\}.$$

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  $f : X \rightarrow \mathbb{R}$ ,  $f \in C^2$ , na konvexní množině  $X \subseteq \mathbb{R}^n$ , pak

$$(x - x^*)^T \nabla^2 f(x^*) (x - x^*) \geq 0$$

pro všechna  $x \in X$  taková, že  $\text{grad}^T f(x^*) (x - x^*) = 0$ , tj. pro vektory  $(x - x^*) \in \text{Ker grad}^T f(x^*)$ .

**Postačující podmínka pro (4.1)**

Bod  $x^* \in X$  je lokálním minimem funkce  $f : X \rightarrow \mathbb{R}$ ,  $f \in C^2$ , na konvexní množině  $X \subseteq \mathbb{R}^n$ , jestliže

$$\text{grad}^T f(x^*) (x - x^*) \geq 0, \quad \forall x \in X$$

(tj. je to stacionární bod), množina  $X$  je polyedr a platí

$$(x - x^*)^T \nabla^2 f(x^*) (x - x^*) > 0$$

pro všechna  $x \in X$  taková, že  $x \neq x^*$  a  $(x - x^*) \in \text{Ker grad}^T f(x^*)$ .

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

Některé  
historické  
úlohy

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

„Moderní“  
úlohy

- Lineární programování (LP)

$$c^T x \rightarrow \min, \quad Ax \leq b, \quad x \geq 0,$$

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 celočíselným programováním (ZLP), např. problém batohu/zloděje, problém obchodního cestujícího (viz M0160).

- ekonomické úlohy: maximalizace užitku/zisku, minimalizace výdajů/nákladů, tvorba portfolia  $\rightsquigarrow$  kvadratické programování (viz M0160)

## 4.1 OBECNÁ OPTIMALIZAČNÍ ÚLOHA

## 4.2 NUTNÉ A POSTAČUJÍCÍ PODMÍNKY OPTIMALITY V MP

## 4.3 TEORIE (LAGRANGEOVY) DUALITY

## 4.4 ANALÝZA CITLIVOSTI



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.

$$f(x) \rightarrow \min, \quad x \in X, \quad (4.1)$$

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

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*  $L : P \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$  danou předpisem

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

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

Pro výklad ještě potřebujeme následující množiny

$$Q := \{y = (y_1, \dots, y_m)^\top \mid y_1, \dots, y_k \geq 0\},$$

$$I(x^*) := \{i \in \{1, \dots, k\} \mid g_i(x^*) = 0\}, \quad x^* \in X,$$

$$S(x^*) := I(x^*) \cup \{k+1, \dots, m\}, \quad x^* \in X.$$

Množina  $I(x^*)$  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  $f, g_1, \dots, g_k : P \rightarrow \mathbb{R}$  jsou diferencovatelné v bodě  $x^* \in X$  a  $g_{k+1}, \dots, g_m : P \rightarrow \mathbb{R}$  jsou spojitě diferencovatelné v nějakém okolí bodu  $x^*$ . Je-li bod  $x^* \in X$  lokálním řešením úlohy (4.1) & (4.2), pak existují Lagrangeovy multiplikatory  $y_0^*, y_1^*, \dots, y_k^* \geq 0, y_{k+1}^*, \dots, y_m^* \in \mathbb{R}$ , tj.  $y_0^* \geq 0$  a  $y^* \in Q$ , taková, že ne všechna  $y_0^*, \dots, y_m^*$  jsou nulová a platí

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

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

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  $y_1^*, \dots, y_k^* \geq 0$   $\rightsquigarrow$  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). Necht  $x^* \in X$  je lokální řešení úlohy (4.1) & (4.2) a uvažme systém nerovností a rovností na množině  $P$  ve tvaru

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

Předpokládejme, že existuje řešení  $\tilde{x} \in P$  tohoto systému. Označme  $h := \tilde{x} - x^*$ . 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

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

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  $g_i(x^* + th) \leq 0$  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  $g_i(x^* + th) \leq 0$  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.

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

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.  $h \in \mathcal{V}(x^*, X) \cap \mathcal{U}(x^*, f) \neq \emptyset$   $\zeta$  s Lemma 4.1.1.

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

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

tj.

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

## 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.

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

Podmínka (4.2.4) je splněna triviálně. ■

Podmínka  
(4.2.3) ve  
speciálních  
případech

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

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

(ii) Je-li

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

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

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

(iii) Je-li

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

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

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

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

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

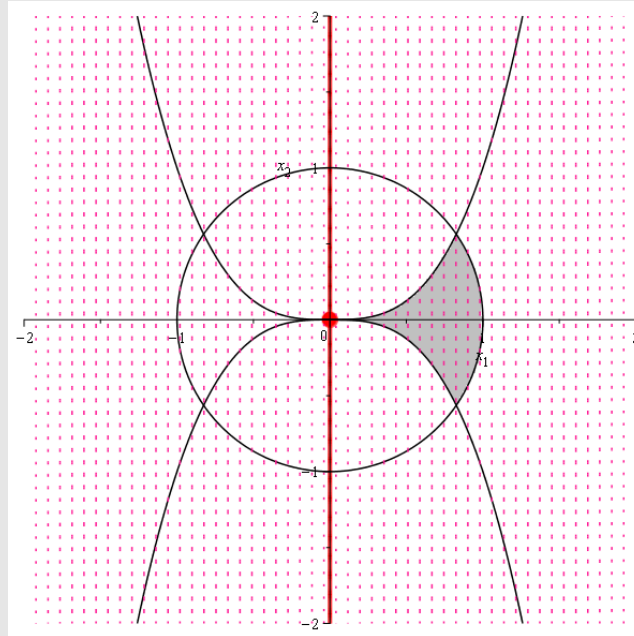
## Příklad

Uvažme úlohu s  $P = \mathbb{R}^2$  a

$$f(x_1, x_2) = x_1 \rightarrow \min,$$

$$g_1(x_1, x_2) = -x_1^3 + x_2 \leq 0, \quad g_2(x_1, x_2) = -x_1^3 - x_2 \leq 0,$$

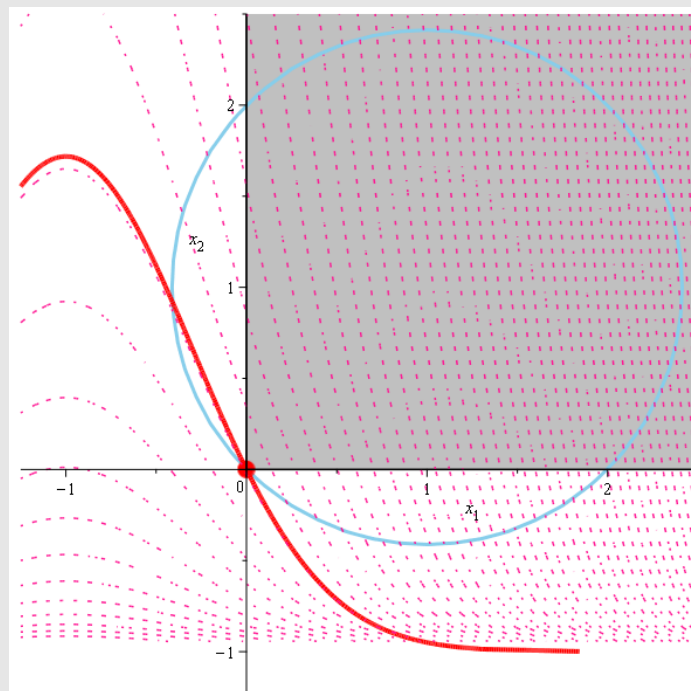
$$g_3(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0.$$

„Divočejší“  
příklad

Uvažme úlohu ( $P = \mathbb{R}^2$ )

$$f(x_1, x_2) = (x_1 + 1)^2 + \ln(x_2 + 1) \rightarrow \min,$$

$$(x_1 - 1)^2 + (x_2 - 1)^2 = 2, \quad x_1 \geq 0, \quad x_2 \geq 0.$$



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  $\text{grad } g_i(x^*)$  pro  $i \in S(x^*)$ .

### Věta 4.2.1A

Nechť jsou splněny předpoklady Věty 4.2.1 a  $x^* \in \text{int } P$ . 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  $x^* \in \text{int } P$  (lokální vs. globální).

### Věta 4.2.1B

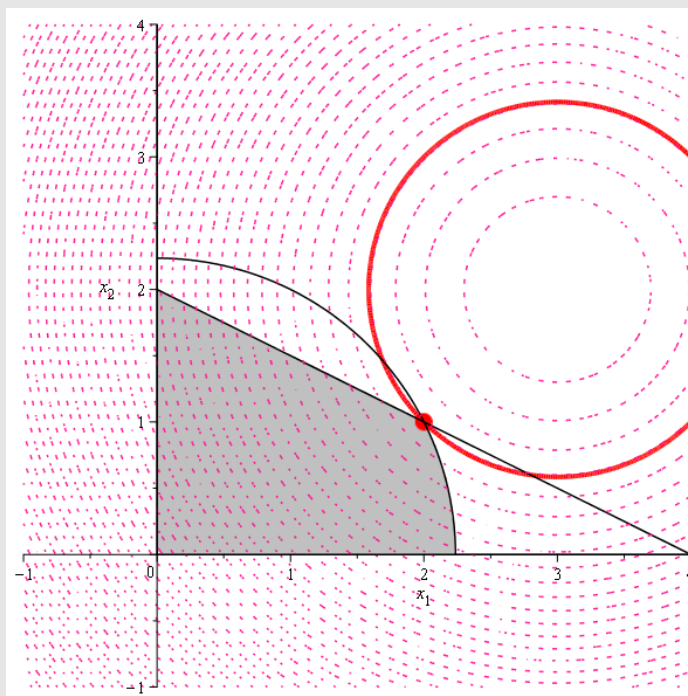
Nechť jsou splněny předpoklady Věty 4.2.1 a  $x^* \in \text{int } P$ . Potom existují (jediné) multiplikátory  $y^* \in Q$  splňující (4.2.3) & (4.2.4) s  $y_0^* = 1$ , jestliže je splněna (alespoň) jedna z následujících podmínek

- (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  $\text{grad } g_i$  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

$$\begin{aligned} & (x_1 - 3)^2 + (x_2 - 2)^2 \rightarrow \min, \\ & x_1^2 + x_2^2 \leq 5, \quad x_1 + 2x_2 \leq 4, \quad x_1 \geq 0, \quad x_2 \geq 0. \end{aligned}$$



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  $f, g_1, \dots, g_k : P \rightarrow \mathbb{R}$  jsou diferencovatelné v bodě  $x^* \in X$  a  $g_{k+1}, \dots, g_m : P \rightarrow \mathbb{R}$  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  $L(x^*, y^*) \leq L(x, y^*)$  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

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

neboť  $g_i(x) = 0$  pro  $i \in \{k+1, \dots, m\}$  a  $y_i^* g_i(x) \leq 0$  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 multiplikaátorem  $y_0^* = 1$  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 necht' platí (alespoň) jedna z následujících podmínek:

- (i) (LNZ) množina  $P$  je otevřená, vektory  $\text{grad } g_i(x)$ ,  $i \in S(x)$ , 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

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

„ $\Rightarrow$ “ Musíme dokázat pouze „dostatečnost“ uvedených předpokladů. Necht' 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). Necht' tedy  $k = m$  a uvažme systém nerovností

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

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

Pak tento systém nemá řešení na  $P$  — kdyby totiž  $\bar{x} \in P$  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  $h = \bar{x} - x^*$ , že

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

tj.  $g_i(x^* + \alpha h) < g_i(x^*)$  pro  $\alpha > 0$  dostatečně malé. Neboť  $x^* + \alpha h = \alpha \bar{x} + (1 - \alpha)x^* \in P$  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  $h \in \mathcal{U}(x^*, f) \cap \mathcal{V}(x^*, X) \neq \emptyset$   $\nmid$  s Lemma 4.1.1.



## DŮKAZ VĚTY 4.2.3 (POKR.)

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

$$\langle \text{grad } g_i(x^*), \bar{x} - x^* \rangle \leq g_i(\bar{x}) - g_i(x^*) = g_i(\bar{x}) < 0.$$

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

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

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 (4.2.3) a (4.2.4) s  $y_0^* = 1$ .

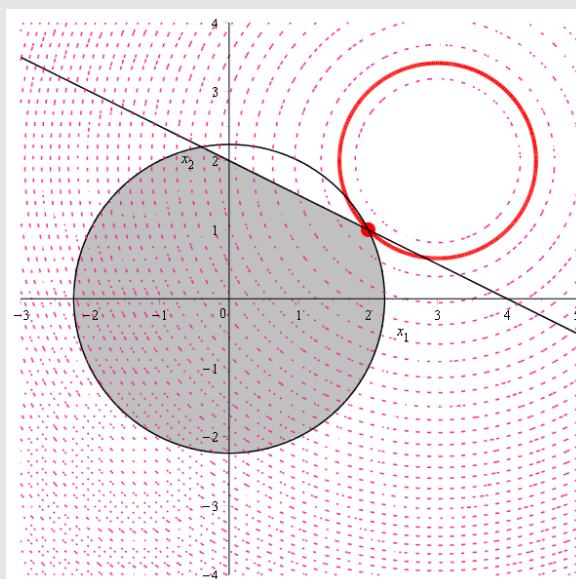
Důkaz tvrzení za podmínky (iii) plyne z Věty 2.3.14. ■

## Příklad

Vyřešme

$$f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \rightarrow \min,$$

$$x_1^2 + x_2^2 \leq 5, \quad x_1 + 2x_2 \leq 4.$$

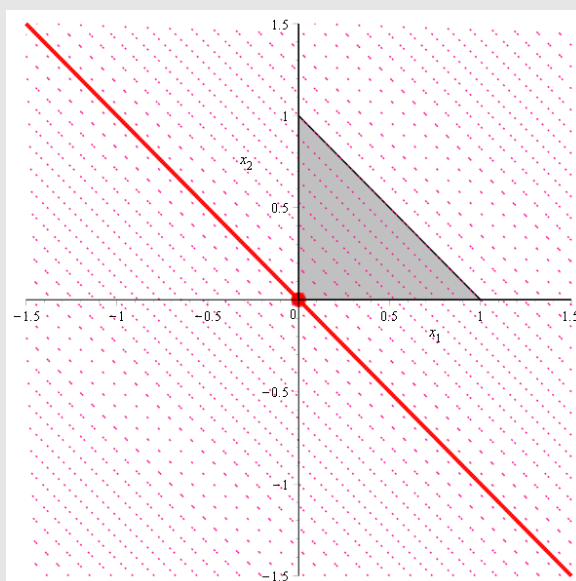


## Příklad

Vyřešme

$$f(x_1, x_2) = x_1 + x_2 \rightarrow \min,$$

$$x_1 + x_2 \leq 1, \quad x_1 \geq 0, \quad x_2 \geq 0.$$

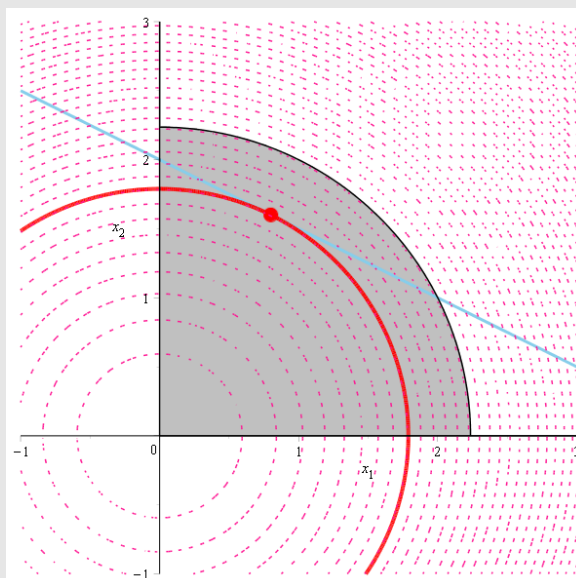


## Příklad

Vyřešme

$$f(x_1, x_2) = x_1^2 + x_2^2 \rightarrow \min,$$

$$x_1^2 + x_2^2 \leq 5, \quad x_1 + 2x_2 = 4, \quad x_1 \geq 0, \quad x_2 \geq 0.$$



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  $x^* \in \text{int } P$  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.

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

$$g_i(x^*) \leq 0 \quad \text{pro } i \in \{1, \dots, k\}, \quad g_i(x^*) = 0 \quad \text{pro } i \in \{k+1, \dots, m\},$$

$$y_i^* > 0 \quad \text{pro } i \in I(x^*), \quad y_i^* = 0 \quad \text{pro } i \in \{1, \dots, k\} \setminus I(x^*),$$

$$y_i^* \in \mathbb{R} \quad \text{pro } i \in \{k+1, \dots, m\}.$$

Jestliže

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

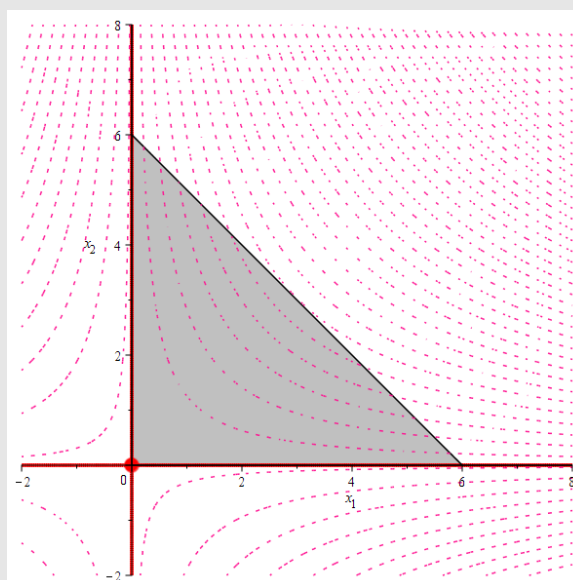
tj.  $h^\top \nabla_x^2 L(x^*, y^*) h > 0$  pro všechna  $h \in \mathbb{R}^n \setminus \{0\}$  taková, že  $\langle \text{grad } g_i(x^*), h \rangle = 0$  pro  $i \in S(x^*)$ , pak bod  $x^*$  je ostré lokální minimum funkce  $f$  na množině  $X$ .

#### Příklad

Uvažme úlohu

$$-x_1 x_2 \rightarrow \min,$$

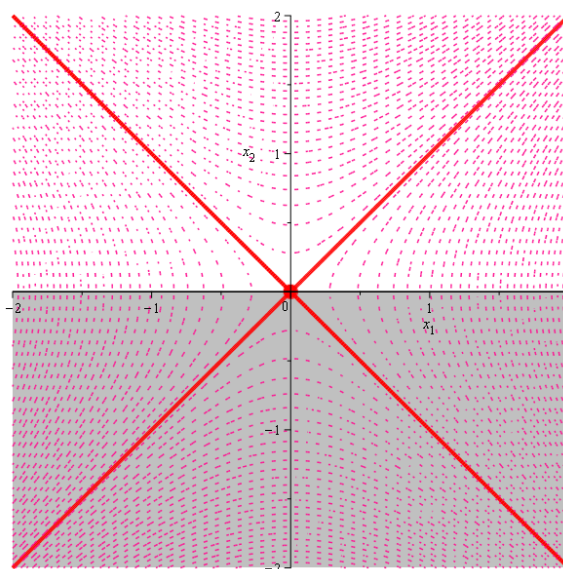
$$x_1 + x_2 \leq 6, \quad x_1 \geq 0, \quad x_2 \geq 0.$$



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.

$$\frac{x_1^2 - x_2^2}{2} \rightarrow \min, \quad x_2 \leq 0.$$



- 4.1 OBECNÁ OPTIMALIZAČNÍ ÚLOHA
- 4.2 NUTNÉ A POSTAČUJÍCÍ PODMÍNKY OPTIMALITY V MP
- 4.3 TEORIE (LAGRANGEOVY) DUALITY
- 4.4 ANALÝZA CITLIVOSTI

## Definice 4.3.1

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

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

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

Existuje K–T vektor vždy? Uvažme např. úlohy

- (i)  $-x^2 \rightarrow \min, \quad x = 0, \quad P = \mathbb{R},$
- (ii)  $x - 1 \rightarrow \min, \quad x^2 - 1 = 0, \quad P = \mathbb{R}_+,$
- (iii)  $e^x \rightarrow \min, \quad x \leq 0, \quad P = \mathbb{R}.$

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 necht' 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 LNŽ nepomůže; existence vs. neexistence a řešitelnost).

Terminologické „zamyšlení“:

omezení  $\rightsquigarrow$  kvalifikované omezení  $\rightsquigarrow$  ostrá komplementarita



## DŮKAZ VĚTY 4.3.2

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

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

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  $y_1^*, \dots, y_m^* \geq 0$  takových, že

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

Tedy  $y^* = (y_1, \dots, y_m)^T$  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

$$\begin{aligned} 4x_1 + 14x_2 &\rightarrow \min, \\ x_1 + 2x_2 &\geq 3, & x_1 + 4x_2 &\geq 4, \\ 3x_1 + x_2 &\geq 3, & x_1 &\geq 0, & x_2 &\geq 0. \end{aligned}$$

### Motivační příklad 2

Uvažme úlohu

$$\begin{aligned} x_1^2 + x_2^2 + 2x_1 &\rightarrow \min, \\ x_1 + x_2 &= 0. \end{aligned}$$

### Definice 4.3.3

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 množinu (tzv. efektivní definiční obor)

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

Pak úloha

$$\varphi(y) \rightarrow \max, \quad y \in Y, \tag{4.3.1}$$

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  $\varphi$  je konkávní na  $Y$ .

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

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

tj. funkce  $\varphi$  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  $\varphi(\lambda y_1 + (1 - \lambda)y_2) \geq \lambda \varphi(y_1) + (1 - \lambda)\varphi(y_2) > -\infty$ , tj. množina  $Y$  je konvexní. ■

## Příklad

Určeme duální úlohu k úloze LP, tj.

$$\langle c, x \rangle \rightarrow \min, \quad x \in X,$$

kde množina  $X$  je určena podmínkami

$$\begin{aligned} \sum_{j=1}^n a_{ij} x_j &\geq b_i, \quad i \in \{1, \dots, k\}, \quad \text{tj. } \langle a_i, x \rangle \geq b_i, \\ \sum_{j=1}^n a_{ij} x_j &= b_i, \quad i \in \{k+1, \dots, m\}, \quad \text{tj. } \langle a_i, x \rangle = b_i, \\ x_j &\geq 0, \quad j \in \{1, \dots, s\}, \end{aligned}$$

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



## Věta 4.3.5

Existuje nějaký vztah mezi  $f^*$  a  $\varphi^*$ ?

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

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

Zejména, pokud  $X \neq \emptyset$  a  $Y \neq \emptyset$ , pak  $f^* \geq \varphi^*$  (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

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

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)

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

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 nepřipustná.

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 nepřipustná.

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: nepřipustná 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  $\varphi^*$ )

„Certifikát  
optimality“

Věta 4.3.5 také dává velmi „jednoduchou“ postačující podmínku pro ověření 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, připustíme, že existuje K–T vektor  $y^* \in Q$ , tj.

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

(v případě  $X = \emptyset$  je  $f^* = \infty$ , a tudíž také  $\varphi^* = \infty \not\leq$ , podobně pro  $Y = \emptyset$  je  $\varphi^* = -\infty$ , a tudíž  $f^* = -\infty \not\geq$ ).

## Věta 4.3.6

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

**(Silná věta o dualitě)** Necht' 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 spojitě 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)  $\rightsquigarrow$  von Neumannova věta o „minimaxu“  $\rightsquigarrow$  sedlový bod (viz později)  $\rightsquigarrow$  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)

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

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  $f^* \geq \varphi^*$ , takže celkem  $f^* = \varphi^*$ .

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

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

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

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

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

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  $\varphi^*$  vyplývá  $\varphi^* > -\infty$ . Potom z Věty 4.3.5 plyne  $f^* \geq \varphi^* > -\infty$ , 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 nepřipustná 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.  $x \rightarrow \min$  &  $x \leq 0$  &  $P = \mathbb{R}_{++}$ ).

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

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

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

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) „ $\implies$ “ Necht  $x^*$  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  $f(x^*) = f^* = \varphi^* = \varphi(y^*)$ .

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

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

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

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

neboli

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

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

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

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

## Příklad

Určeme duální úlohu pro

$$f(x) = e^{-x} \rightarrow \min, \quad x \leq 1,$$

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

?? Co je duální úloha k duální úloze ??

## Příklad

Určeme duální úlohu pro

$$f(x_1, x_2) = 2x_1^2 + x_1x_2 + x_2^2 - 4x_1 + 2x_2 \rightarrow \min$$

$$3x_1 + x_2 \leq 2 \quad \& \quad x_1 \geq 0$$

Duální úlohu vyřešme. Jaké informace nám toto řešení dává o primární 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

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

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).

„ $\implies$ “ Necht'  $x^* \in P$  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  $L(x^*, y^*) \leq L(x, y^*)$  pro každé  $x \in P$ . Současně s využitím druhé části dostaneme

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

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

## DŮKAZ VĚTY 4.3.10 (POKR.)

„ $\impliedby$ “ Necht' nyní  $[x^*, y^*]$  je sedlovým bodem. Pak  $L(x^*, y^*) = \min_{x \in P} L(x, y^*)$  a zároveň  $L(x^*, y^*) \geq L(x^*, y)$  pro každé  $y \in Q$ , tj.

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

neboli

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

Protože tato nerovnost platí pro všechna  $y \in Q$ , nutně také pro  $y = 0 \in Q$ , což dává  $\sum_{i=1}^m y_i^* g_i(x^*) \geq 0$ . 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  $0 \geq g_i(x^*)$  &  $0 \leq g_i(x^*)$ , 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 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  $\Leftarrow$  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).

### 4.1 OBECNÁ OPTIMALIZAČNÍ ÚLOHA

### 4.2 NUTNÉ A POSTAČUJÍCÍ PODMÍNKY OPTIMALITY V MP

### 4.3 TEORIE (LAGRANGEOVY) DUALITY

### 4.4 ANALÝZA CITLIVOSTI



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 užítka, ...
- (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.

Začneme s jednoduchou úlohou pro 2 proměnné a 1 omezení ve tvaru rovnosti, tj.

$$f(x_1, x_2) \rightarrow \min, \quad g(x_1, x_2) = c, \quad (4.4.1)$$

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

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

Současně z Lagrangeova principu máme

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

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

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

Zkombinováním (\*), (□) a (△) dostaneme

$$\begin{aligned} \frac{d}{dc} f^*(c) &\stackrel{(*),(\square)}{=} -y^*(c) g_{x_1}(x_1^*(c), x_2^*(c)) \frac{dx_1^*(c)}{dc} - y^*(c) g_{x_2}(x_1^*(c), x_2^*(c)) \frac{dx_2^*(c)}{dc} = \\ &= -y^*(c) \left[ g_{x_1}(x_1^*(c), x_2^*(c)) \frac{dx_1^*(c)}{dc} + g_{x_2}(x_1^*(c), x_2^*(c)) \frac{dx_2^*(c)}{dc} \right] = \\ &\stackrel{(\triangle)}{=} -y^*(c). \end{aligned}$$

(Tvzení 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  $\Delta c \rightsquigarrow$  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

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

### Příklad

Produkční funkce firmy je  $f(x_p, x_m) = 50x_p^{1/2}x_m^2$ , 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

$$f(x, r) \rightarrow \min, \quad g_1(x, r) = 0, \dots, g_m(x, r) = 0, \quad (*)$$

kde  $x \in \mathbb{R}^n$ ,  $r \in \mathbb{R}^k$ ,  $f, g_1, \dots, g_m \in C^1$ . Pripustme, 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

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

s odpovídající Lagrangeovým multiplikátorem  $y^*$ . Nechť dále tato dvojice splňuje postačující podmínku druhého řádu (tj.  $\nabla_x^2 L(x^*, y^*) > 0$  na  $\text{Ker DG}(x^*)$ ), přičemž současně  $x^*$  je regulárním bodem, tj.  $D_x G(x^*) \in \mathbb{R}^{m \times n}$  má plnou hodnotu  $m$ . Uvažme úlohu parametrického programování

$$f(x) \rightarrow \min, \quad G(x) = u \quad (*)$$

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

$$\text{grad } f^*(u) = -y^*(u),$$

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  $b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$ , tj.

$$f(x) \rightarrow \min, \quad x \in X(b) := \{x \in P \subseteq \mathbb{R}^n \mid g_i(x) \leq b_i, i \in 1, \dots, m\}. \quad (4.4.2)$$

Ještě budeme potřebovat označení

$$G(x) := (g_1(x), \dots, g_m(x))^T, \quad X(b) := \{x \in P \mid G(x) \leq b\},$$

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

$$Y(b) := \{y \in \mathbb{R}^m \mid y \geq 0, F(b) \leq f(x) + \langle y, G(x) - b \rangle \quad \forall x \in P\},$$

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

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  $\text{grad}^\top F(b)$  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  $\tilde{b}, \hat{b} \in B$  a  $\lambda \in [0, 1]$  a jsou libovolná, tj. existují  $\tilde{x} \in X(\tilde{b})$  a  $\hat{x} \in X(\hat{b})$ . Položme  $x := \lambda \tilde{x} + (1 - \lambda)\hat{x}$ . Potom

$$G(x) = G(\lambda \tilde{x} + (1 - \lambda)\hat{x}) \stackrel{G..konvx}{\leq} \lambda G(\tilde{x}) + (1 - \lambda) G(\hat{x}) \leq \lambda \tilde{b} + (1 - \lambda)\hat{b}. \quad (*)$$

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

## DŮKAZ VĚTY 4.4.3(II)

Nechť  $y^* \in Y(0)$ . Pak  $y^* \geq 0$  a  $F(0) \leq f(x) + \langle y^*, G(x) \rangle$  pro  $\forall x \in P$ . Protože  $y^* \geq 0$ , platí  $F(0) \leq f(x) + \langle y^*, b \rangle$  pro libovolné  $b \in B$  a  $x \in X(b)$ , 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 ( $F(b) = \infty$  pouze pro  $f \equiv \infty$   $\downarrow$ ).

Nechť nyní  $\tilde{b}, \hat{b} \in B$  a  $\lambda \in [0, 1]$  jsou libovolná. Položme  $b := \lambda \tilde{b} + (1 - \lambda)\hat{b}$ . Pak pro libovolná  $\tilde{x} \in X(\tilde{b})$  a  $\hat{x} \in X(\hat{b})$  označme  $x := \lambda \tilde{x} + (1 - \lambda)\hat{x}$ . Podle části (i) je  $x \in X(b)$ , viz (\*), tedy

$$F(b) \leq f(x) \leq \lambda f(\tilde{x}) + (1 - \lambda)f(\hat{x}).$$

Protože  $\tilde{x}, \hat{x}$  byla zvolena libovolně, plyne odtud

$$F(b) \leq \lambda F(\tilde{b}) + (1 - \lambda)F(\hat{b}),$$

tedy funkce  $F$  je konvexní.

Je-li  $\tilde{b}, \hat{b} \in B$  a  $\tilde{b} \leq \hat{b}$  (po složkách), pak zřejmě  $X(\tilde{b}) \subseteq X(\hat{b})$ . Proto  $F(\tilde{b}) \geq F(\hat{b})$ , 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.

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

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

$$F(b) \leq F(b') - \langle y^*, b' - b \rangle \leq f(x) + \langle -y^*, b' - b \rangle = f(x) + \langle -y^*, G(x) - b \rangle. \quad (*)$$

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  $-y^* \geq 0$ ).

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

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

Vzhledem k libovlnosti  $b' \in B$ , 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  $b \in \text{ri } B$ , 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  $0 \in \text{int } B$  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  $\text{grad}^T F(0) = -y^*$ .

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  $b = 0 \in \text{int } B$  (neboť  $G(\bar{x}) \leq b$  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  $b$  určíme hodnotu  $F(b)$  parametrické úlohy

$$f(x) = x^2 + 2x + 2 \rightarrow \min, \quad x \leq b.$$

Určíme dále duální úlohu k této úloze, vyřešíme ji (bez použití předchozí části) a ověříme platnost vztahu  $\partial F(b) = -Y^*(b)$ , kde  $Y^*(b)$  je množina řešení duální úlohy.

### Příklad

V závislosti na parametru  $b$  určíme hodnotu  $F(b)$  parametrické úlohy

$$f(x_1, x_2) = x_1^2 + x_2^2 + 3x_1 \rightarrow \min, \quad x_2^2 - x_1 \leq b, \quad x_1 \geq 0.$$

Určíme dále duální úlohu k této úloze, vyřešíme ji (bez použití předchozí části) a ověříme 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  $\alpha \geq 0$ .
- (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

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

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

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

## FENCHEL VS. K–T (POKR.)

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

$$-F^*(y) \rightarrow \max, \quad y \geq 0.$$

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 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  $0 \in \text{int } B$ , a tedy  $F$  je spojitá podle Vět 2.4.1 a 4.4.3(ii).

Konec.