

# Část I

## Optimalizace

# Kapitola 1

## Základní geometrické pojmy

### 1.1 Přímký a úsečky

Mějme dva body  $x_1, x_2 \in \mathbb{R}^n$  takové, že  $x_1 \neq x_2$  a parametr  $\theta \in \mathbb{R}$ . Potom výraz

$$y = \theta x_1 + (1 - \theta)x_2 \quad (1.1)$$

popisuje **přímku** procházející body  $x_1$  a  $x_2$ . Pro  $\theta = 0$  dostáváme bod  $x_2$  a pro  $\theta = 1$  bod  $x_1$ . Omezíme-li  $\theta$  na interval  $\langle 0, 1 \rangle$ , dostaneme **úsečku** s koncovými body  $x_1$  a  $x_2$ . Výraz 1.1 lze přepsat do tvaru

$$y = x_2 + \theta(x_1 - x_2),$$

který můžeme interpretovat jako součet počátečního bodu  $x_2$  a nějakého násobku směrového vektoru  $x_1 - x_2$ .

### 1.2 Afinní prostory

Říkáme, že  $C \subseteq \mathbb{R}^n$  je **afinní prostor**, jestliže přímka procházející libovolnými dvěma různými body z  $C$  leží v  $C$ . Tedy  $C$  obsahuje lineární kombinace libovolných dvou bodů z  $C$ , jestliže součet koeficientů lineární kombinace je roven jedné. To lze zobecnit i pro více než dva body. Lineární kombinace  $\theta_1 x_1 + \dots + \theta_k x_k$  bodů  $x_1, \dots, x_k$  taková, že  $\theta_1 + \dots + \theta_k = 1$ , se nazývá **afinní kombinace** bodů  $x_1, \dots, x_k$ . Indukcí z definice afinního prostoru lze snadno ukázat, že pokud  $C$  je afinní množina,  $x_1, \dots, x_k \in C$  a  $\theta_1 + \dots + \theta_k = 1$ , potom bod  $\theta_1 x_1 + \dots + \theta_k x_k \in C$ .

Nechť  $C$  je afinní prostor a  $x_0 \in C$ , potom množina

$$V = C - x_0 = \{x - x_0 \mid x \in C\}$$

je **vektorový prostor**, tj. množina, která je uzavřená na sčítání a násobení skalárem.

Afinní prostor  $C$  lze vyjádřit jako

$$C = V + x_0 = \{v + x_0 \mid v \in V\},$$

kde  $V$  je vektorový prostor a  $x_0$  je počátek. Poznamenejme, že vektorový prostor  $V$  asociovaný s afinním prostorem  $C$  nezávisí na volbě počátku  $x_0$ .

**Dimenze** afinního prostoru  $C = V + x_0$  je definována jako dimenze vektorového prostoru  $V = C - x_0$ , kde  $x_0$  je libovolný prvek z  $C$ . Množina všech afinních kombinací bodů množiny  $C \subseteq \mathbb{R}^n$  se nazývá **afinní obal** množiny  $C$ . Afinní obal množiny  $C$  budeme značit

$$\mathbf{aff} C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_1, \dots, x_k \in C, \theta_1 + \cdots + \theta_k = 1\}.$$

Afinní obal je nejmenší afinní prostor, který obsahuje množinu  $C$ . Tedy, jestliže  $S$  je afinní prostor takový, že  $C \subseteq S$ , potom  $\mathbf{aff} C \subseteq S$ .

### 1.3 Konvexní množiny

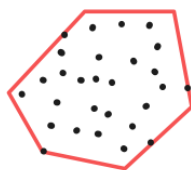
Říkáme, že množina  $C$  je **konvexní**, jestliže úsečka mezi libovolnými dvěma body z  $C$  leží také v  $C$ . Jinak řečeno, jestliže pro libovolné dva body  $x_1, x_2 \in C$  a libovolné  $\theta \in \langle 0, 1 \rangle$  platí, že  $\theta x_1 + (1 - \theta)x_2 \in C$ . Poznamenejme, že každý afinní prostor je zároveň konvexní množinou. Podobně jako afinní kombinaci definujeme **konvexní kombinaci** bodů  $x_1, \dots, x_k$  jako  $\theta_1 x_1 + \cdots + \theta_k x_k$ , kde  $\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$  pro  $i = 1, \dots, k$ . **Konvexní obal** množiny  $C$  je množina všech konvexních kombinací bodů z množiny  $C$ , značíme

$$\mathbf{conv} C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \dots, k, \theta_1 + \cdots + \theta_k = 1\}.$$

Analogicky, konvexní obal množiny  $C$  je nejmenší konvexní množina, která obsahuje množinu  $C$ . Pro představu viz obrázek [1.1](#).

### 1.4 Kužely

Množina  $C$  se nazývá **kužel**, jestliže pro každé  $x \in C$  a  $\theta \geq 0$  platí, že  $\theta x \in C$ . Je-li  $C$  navíc konvexní, pak se  $C$  nazývá **konvexní kužel**. Tedy  $C$  je konvexní kužel, jestliže pro libovolné  $x_1, x_2 \in C$  a  $\theta_1, \theta_2 \geq 0$  platí, že  $\theta_1 x_1 + \theta_2 x_2 \in C$ . Říkáme, že bod ve tvaru  $\theta_1 x_1 + \cdots + \theta_k x_k$ , kde  $\theta_1, \dots, \theta_k \geq 0$  je **kuželovou kombinací** bodů  $x_1, \dots, x_k$ . Dále, pokud  $x_i$  leží v konvexním kuželu množiny  $C$ , potom libovolná kuželová kombinace bodu  $x_i$  leží rovněž

(a) Množina bodů  $C$ (b)  $\text{conv } C$ 

Obrázek 1.1: Konvexní obal množiny

v konvexním kuželu množiny  $C$ . Platí, že množina  $C$  je konvexní kužel právě tehdy, když  $C$  obsahuje všechny kuželové kombinace svých bodů. **Kuželový obal** množiny  $C$  je množina, která obsahuje všechny kuželové kombinace množiny  $C$ , tj.

$$\text{cone } C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \dots, k\}.$$

Kuželový obal množiny  $C$  je zároveň nejmenší konvexní kužel, který obsahuje množinu  $C$ . Pro představu viz obrázek [1.2](#).

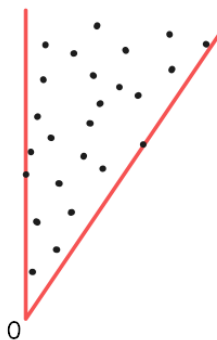
## 1.5 Nadroviny a poloprostory

**Nadrovina** je množina ve tvaru

$$\{x \mid a^T x = b\},$$

kde  $a \in \mathbb{R}^n$ ,  $a \neq 0$  a  $b \in \mathbb{R}$ . Analyticky se na nadrovinu koukáme jako na množinu všech řešení netriviální lineární rovnice. Geometricky zase jako na množinu všech bodů takových, že mají konstantní skalární součin s normálovým vektorem  $a$ . Konstanta  $b$  značí posunutí nadroviny od počátku. Nadrovinu také můžeme vyjádřit jako

$$\{x \mid a^T(x - x_0) = 0\} = x_0 + \{v \mid a^T v = 0\},$$

(a) Množina bodů  $C$ (b) **cone**  $C$ 

Obrázek 1.2: Kuželový obal množiny

kde  $x_0$  je libovolný bod této nadroviny a  $\{v \mid a^T v = 0\}$  je množina všech vektorů, které jsou kolmé k normálovému vektoru  $a$ . Nadrovina je tedy množina, která obsahuje bod  $x_0$  a libovolný bod ve tvaru  $x_0 + v$ , kde  $v$  je vektor, který je kolmý k normálovému vektoru  $a$ . Pro ilustraci v  $\mathbb{R}^2$  viz obrázek [1.3a](#).

Nadrovina dělí  $\mathbb{R}^n$  na dva poloprostory. Množina

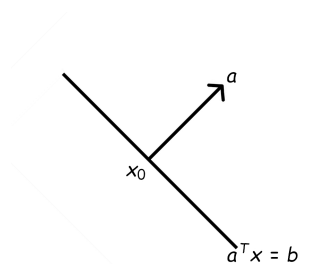
$$\{x \mid a^T x \leq b\}, \text{ resp. } \{x \mid a^T x < b\},$$

kde  $a \neq 0$  se nazývá (uzavřený) **poloprostor**, resp. **otevřený poloprostor**. Je to tedy množina všech řešení netriviální lineární nerovnice. Podobně jako

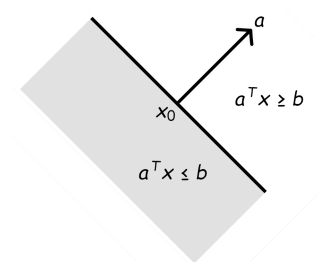
nadrovinu, můžeme poloprostor vyjádřit ve tvaru

$$\{x \mid a^T(x - x_0) \leq 0\}, \text{ resp. } \{x \mid a^T(x - x_0) < 0\},$$

kde  $a \neq 0$  a  $x_0$  je libovolný bod z nadroviny  $\{x \mid a^T x = b\}$ . Poloprostor tedy obsahuje bod  $x_0$  a libovolný bod  $x_0 + v$ , kde  $v$  je vektor, který s vnějším normálovým vektorem svírá tupý nebo pravý úhel. Tato interpretace je v  $\mathbb{R}^2$  ilustrována na obrázku 1.3b. Ještě poznamenejme, že poloprostory jsou konvexní množiny, ale samozřejmě nejsou afinní.



(a) Nadrovina



(b) Poloprostor

Obrázek 1.3: Nadrovina a poloprostor v  $\mathbb{R}^2$ .

## 1.6 Polyedry a polytopy

Polytopy jsou zobecněním konvexních mnohoúhelníků v rovině do více dimenzí. Polytop v  $\mathbb{R}^3$  je konvexní množina, která je ohraničena konečně mnoha konvexními mnohoúhelníky (příkladem polytopů v  $\mathbb{R}^3$  jsou např. Platónská tělesa). Na takovou množinu je možné nahlížet dvěma způsoby. **H-polyedr** je průnik konečně mnoha uzavřených poloprostorů v  $\mathbb{R}^n$ , kde **H-polytop** je omezený H-polyedr. **V-polytop** je konvexní obal konečně mnoha bodů v  $\mathbb{R}^n$ . Následující věta říká, že H-polytop a V-polytop jsou matematicky ekvivalentní množiny.

**Věta 1.** *Každý V-polytop je H-polytop. Každý H-polytop je V-polytop.*

Poznamenejme, že V-polytop a H-polytop jsou sice ekvivalentní množiny, ale z algoritmického hlediska je velký rozdíl, zda pracujeme s bodovou množinou, nebo s uzavřenými poloprostory. Pro ilustraci: mějme lineární funkci, kterou chceme minimalizovat na daném polytopu. Pro V-polytop se jedná o triviální problém, protože stačí pro každý bod z množiny  $V$  určit hodnotu dané funkce a vybrat minimum. Na druhou stranu pro H-polytop se jedná o netriviální problém, kterým se zabývá lineární programování. Dále budeme mluvit jen o **polyedrech** a **polytopech**.

Důležitý fakt, že každý polyedr je konečně generovaný, říká Minkowského-Weylova věta.

**Věta 2** (Minkowski-Weyl).  $P \subseteq \mathbb{R}^n$ . Potom  $P = \mathbf{conv}(u_1, \dots, u_r) + \mathbf{cone}(v_1, \dots, v_s)$ , kde  $u_i, v_i$  jsou extrémní vrcholy  $P$  právě tehdy, když  $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ .

# Kapitola 2

## Konvexní optimalizace

### 2.1 Obecná podmíněná úloha

$$\begin{aligned} \min \quad & f(x) \\ & g_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{aligned} \tag{2.1}$$

Hledáme  $x \in \mathbb{R}^n$ , které minimalizuje  $f(x)$ , vzhledem k omezením  $g_i(x)$  a  $h_i(x)$ . Proměnné  $x$  říkáme **optimalizační proměnná**, funkci  $f(x)$  říkáme **cenová** nebo **účelová funkce**. Výrazy  $g_i(x) \leq 0$  jsou **omezení typu nerovnosti** a  $h_i(x) = 0$  jsou **omezení typu rovnosti**. Pokud  $m = p = 0$  problém 2.1 je **neomezený**, jinak je **omezený**.

Definiční obor  $\mathcal{D}$  úlohy 2.1 je

$$\mathcal{D} = \bigcap_{i=1}^m \text{dom } g_i \cap \bigcap_{i=1}^p \text{dom } h_i.$$

Říkáme, že bod  $x \in \mathcal{D}$  je **přípustný**, jestliže splňuje všechna omezení  $g_i(x) \leq 0$  a  $h_i(x) = 0$ . Úloha 2.1 je **přípustná**, jestliže existuje alespoň jeden bod  $x \in \mathcal{D}$ , který je přípustný. Množina všech přípustných bodů  $x \in \mathcal{D}$  se nazývá **přípustná množina**.

**Optimální hodnota**  $x^*$  úlohy 2.1 je definována jako

$$x^* = \{f(x) \mid g_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}.$$



## 2.2 Konvexní podmíněná úloha

$$\begin{aligned} \min \quad & f(x) \\ & g_i(x) \leq 0, i = 1, \dots, m \\ & a_i^T x = b_i, i = 1, \dots, p \end{aligned} \quad (2.2)$$

Oproti obecné úloze 2.1 jsou funkce  $f(x)$ ,  $g_i(x)$  konvexní a funkce  $h_i(x) = a_i^T x - b_i$  jsou afinní. Přípustná množina takové úlohy je konvexní množinou.

## 2.3 Lagrangeova dualita

Mějme úlohu 2.1 s  $\mathcal{D} \neq \emptyset$ . Zobrazení  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$  takové, že

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \mu_i h_i(x) \quad (2.3)$$

se nazývá **Lagrangeova funkce**. Definiční obor  $\text{dom } L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$ . Vektory  $\lambda$  a  $\mu$  nazýváme **duální proměnné** a prvkům těchto vektorů říkáme **Lagrangeovy multiplikátory**. Dále definujeme **duální funkci**  $d : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$  jako infimum Lagrangeovy funkce  $L$  přes všechna  $x \in \mathcal{D}$ . Tedy

$$d(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu) = \inf_{x \in \mathcal{D}} \left( f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \mu_i h_i(x) \right). \quad (2.4)$$

Poznamenejme, že duální funkce je konkávní bez ohledu na to, zda je úloha konvexní a je-li  $L$  zdola neomezená v proměnné  $x$ , potom duální funkce nabývá hodnoty  $-\infty$ .

### 2.3.1 Dolní odhad na $x^*$

Nechť  $\tilde{x}$  je přípustný bod. Pro  $\lambda \geq 0$  je

$$\sum_{i=1}^m \lambda_i g_i(\tilde{x}) + \sum_{i=1}^p \mu_i h_i(\tilde{x}) \leq 0.$$

Potom pro Lagrangeovu funkci můžeme psát

$$L(\tilde{x}, \lambda, \mu) = f(\tilde{x}) + \sum_{i=1}^m \lambda_i g_i(\tilde{x}) + \sum_{i=1}^p \mu_i h_i(\tilde{x}) \leq f(\tilde{x}).$$

A tedy pro duální funkci platí

$$d(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu) \leq L(\tilde{x}, \lambda, \mu) \leq f(\tilde{x}).$$

### 2.3.2 Duální úloha

V části 2.3.1 jsme si ukázali, že duální funkce dává dolní odhad na optimální hodnotu  $x^*$  úlohy 2.1. Stále jsme si ale neřekli, jaký je nejlepší dolní odhad, který pomocí duální funkce jsme schopni dostat. To nás dostává k následující optimalizační úloze.

$$\begin{aligned} \max d(\lambda, \mu) \\ \lambda \geq 0 \end{aligned} \tag{2.5}$$

Úloze 2.5 se říká **Lagrangeova duální úloha** příslušná k úloze 2.1, kterou nazýváme **primární úlohou**.

### 2.3.3 Slabá dualita

Optimální řešení Lagrangeovy duální úlohy označíme  $d^*$ , které je už z definice nejlepší dolní odhad na optimální řešení primární úlohy  $p^*$ . Tato nerovnost platí i pokud primární úloha není konvexní. Této nerovnosti říkáme **slabá dualita**. Rozdíl optimálních řešení  $p^* - d^*$  označujeme jako **optimální dualitní rozdíl** primární úlohy. Poznamenejme, že optimální dualitní rozdíl je vždy nezáporný.

### 2.3.4 Silná dualita a Slaterova podmínka

Pokud je optimální dualitní rozdíl  $p^* - d^* = 0$ , pak říkáme, že platí silná dualita. Silná dualita obecně neplatí, ale pro primární úlohu, která splňuje nějaké další podmínky to možné je. Těmto podmínkám se říká **podmínky kvalifikace omezení**. Jednou takovou je **Slaterova podmínka**:

$$\exists x \in \text{relint } \mathcal{D} : f_i(x) < 0, i = 1, \dots, m, Ax = b.$$

Bodu  $x \in \mathcal{D}$ , který splňuje Slaterovu podmínku, říkáme, že je **striktně přípustný**, protože omezení typu nerovnosti jsou ostré. Pokud jsou některé funkce  $f_i$  afinní, můžeme Slaterovu podmínku modifikovat. Nechť tedy  $f_1, \dots, f_k, k \leq m$ , jsou afinní funkce. Potom **modifikovaná Slaterova podmínka** má tvar:

$$\exists x \in \text{relint } \mathcal{D} : f_i(x) \leq 0, i = 1, \dots, k, f_i(x) < 0, i = k + 1, \dots, m, Ax = b.$$

Pro úlohu 2.2 platí následující věta.

**Věta 3** (Slaterova). *Nechť primární úloha je konvexní a platí (modifikovaná) Slaterova podmínka, potom  $p^* = d^*$ .*

## TODO

**Použití duální úlohy:** obecnou primární úlohu je těžké vyřešit, ale duální úloha je vždy konvexní, tak vyřeším tu a mám alespoň dolní odhad na primární úlohu

# Kapitola 3

## Lineární programování

### 3.1 Primární úloha

Úlohou lineárního programování rozumíme minimalizaci nebo maximalizaci lineární **účelové funkce** vzhledem k lineárním **omezením**, kde tato omezení jsou dána soustavou lineární rovnic a nerovnic. Úlohu lineárního programování lze formulovat v několika ekvivalentních tvarech, které se liší zadáním omezení. Úloha v **kanonickém tvaru** má svá omezení dána soustavou lineárních nerovnic  $Ax \leq b$ . Tedy:

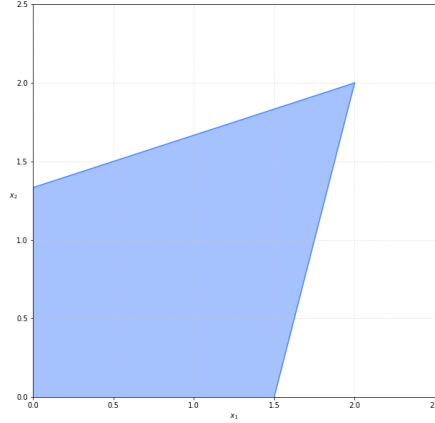
$$\max \{c^T x \mid Ax \leq b, x \geq 0\}, \quad (\text{LP-P})$$

kde  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^n$ ,  $x \in \mathbb{R}^n$  a  $c \in \mathbb{R}^n$ . **Přípustná množina řešení** je průnikem poloprostorů, které jsou definovány soustavou nerovnic  $Ax \leq b$  a **nezáporného ortantu**, tj. množiny  $\{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$ . Obě tyto množiny jsou konvexní a tedy i jejich průnik je rovněž konvexní množina. Dále, protože přípustnou množinu máme popsanou soustavou konečně mnoha lineárních nerovnic, geometricky se na úlohu **LP-P** můžeme koukat jako na maximalizaci lineární funkce přes polyedr, který je definován touto soustavou.

**Příklad.** Mějme následující úlohu:

$$\begin{aligned} \max & x_1 + x_2 \\ & -x_1 + 3x_2 \leq 4 \\ & 4x_1 - x_2 \leq 6 \\ & x \geq 0. \end{aligned} \quad (\text{P1})$$

Přípustná množina řešení je zobrazena na obrázku **3.1**. Řešením úlohy je vektor  $x^* = (2, 2)$  s cenou 4. Implementace v softwaru MOSEK: <https://github.com/c0n73x7/D1PL0MK4/blob/master/mosek/ex1.py>.

Obrázek 3.1: Přípustná množina řešení k úloze P1.

## 3.2 Dualita

Pomocí Lagrangeovy duality odvodíme duální úlohu k primární úloze LP-P. Máme tedy optimalizační úlohu

$$\min \{ -c^T x \mid Ax \leq b, x \geq 0 \}.$$

Pro ní vytvoříme Lagrangeovu funkci

$$\begin{aligned} L(x, \lambda) &= -c^T x + \lambda^T (Ax - b) - \lambda^T x \\ &= -b^T \lambda + (A^T \lambda - c - \lambda)^T x. \end{aligned}$$

Z Lagrangeovy funkce přejdeme k duální funkci

$$\begin{aligned} d(\lambda) &= \inf_x L(x, \lambda) \\ &= \inf_x -b^T \lambda + (A^T \lambda - c - \lambda)^T x \\ &= \begin{cases} -b^T \lambda & \text{pokud } A^T \lambda - c - \lambda = 0, \\ -\infty & \text{jinak.} \end{cases} \end{aligned}$$

Tu nakonec použijeme v duální úloze:

$$\max \{ -b^T \lambda \mid A^T \lambda - c - \lambda = 0 \}$$

$$\max \{ -b^T \lambda \mid A^T \lambda \geq c, \lambda \geq 0 \}$$

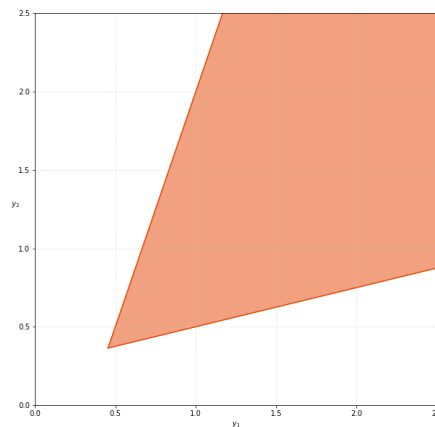
$$\min \{b^T \lambda \mid A^T \lambda \geq c, \lambda \geq 0\} \quad (\text{LP-D})$$

Dostáváme tedy duální úlohu **LP-D** k primární úloze **LP-P**.

**Příklad.** Duální úloha k úloze **P1** je ve tvaru:

$$\begin{aligned} \min & 4y_1 + 6y_2 \\ & -y_1 + 4y_2 \geq 1 \\ & 3y_1 - y_2 \geq 1 \\ & y \geq 0. \end{aligned} \quad (\text{P2})$$

Přípustná množina řešení je zobrazena na obrázku 3.2. Řešením úlohy je vektor  $y^* \approx (0.4546, 0.3636)$  s cenou 4. Implementace v softwaru MOSEK: <https://github.com/c0n73x7/D1PL0MK4/blob/master/mosek/ex2.py>.



Obrázek 3.2: Přípustná množina řešení k úloze **P2**.

Všimněme si, že v příkladech **P1** a **P2** mají řešení  $x^*$  i  $y^*$  stejnou cenu. To není náhoda a tento fakt je obsahem silné věty o dualitě lineárního programování, kterou dokázala skupina kolem Alberta W. Tuckera v roce 1948. Začneme slabou větou o dualitě lineárního programování.

**Věta 4** (Slabá o dualitě). *Nechť  $\tilde{x}$  je přípustné řešení **LP-P** a  $\tilde{y}$  je přípustné řešení **LP-D**. Potom  $c^T \tilde{x} \leq b^T \tilde{y}$ .*

Tedy každé přípustné řešení  $\tilde{y}$  duální úlohy **LP-D** nám dává horní odhad na maximum účelové funkce primární úlohy **LP-P**. Graficky můžeme slabou

větu o dualitě interpretovat jako na obrázku 3.3. Zatím tedy nevíme, zda vždy existují přípustná (optimální) řešení  $x^*$  pro úlohu LP-P a  $y^*$  pro úlohu LP-D, pro která platí  $c^T x^* = b^T y^*$ . Kladnou odpověď dostaneme z již zmíněné silné věty od dualitě.



Obrázek 3.3: Slabá věta o dualitě.

**Věta 5** (Silná o dualitě). *Jestliže úlohy LP-P a LP-D mají přípustná řešení. Potom*

$$\max \{c^T x \mid Ax \leq b, x \geq 0\} = \min \{b^T y \mid A^T y \geq c, y \geq 0\}.$$

Se znalostí silné věty o dualitě můžeme obrázek 3.3 upravit na obrázek 3.4.



Obrázek 3.4: Ceny přípustných řešení primární a příslušné duální úlohy.

### 3.3 Komplementární skluzovost

Pro odvození tzv. podmínky komplementární skluzovosti nejprve převedeme úlohy LP-P a LP-D do jiných tvarů. V primární úloze povolíme  $x \in \mathbb{R}^n$ . Tedy primární úloha je ve tvaru:

$$\max \{c^T x \mid Ax \leq b\}. \quad (\text{LP-P2})$$

A příslušná duální úloha je ve tvaru:

$$\min \{b^T y \mid A^T y = c, y \geq 0\}. \quad (\text{LP-D2})$$

Nechť  $\tilde{x}$  je přípustné řešení a  $x^*$  je optimální řešení úlohy LP-P2,  $\tilde{y}$  je přípustné řešení a  $y^*$  je optimální řešení úlohy LP-D2. **Dualitní rozdíl**  $\tilde{x}$  a  $\tilde{y}$  je číslo  $b^T \tilde{y} - c^T \tilde{x} \geq 0$ . Ze silné věty o dualitě samozřejmě plyne, že pro optimální řešení  $x^*$  a  $y^*$  je dualitní rozdíl roven 0. Vyjdeme z dualitního rozdílu optimálních řešení:

$$b^T y^* - c^T x^* = y^{*T} b - y^{*T} A x^* = y^{*T} (b - A x^*) = 0.$$

Poslední rovnost přepíšeme maticově:

$$[y_1^*, \dots, y_m^*] \left( \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} - \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix} \right) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Dostáváme tedy soustavu rovnic  $y_i^* (b_i - a_{i-} x^*) = 0$ , kde  $i = 1, \dots, m$ . Tedy buď  $y_i^* = 0$  nebo  $b_i - a_{i-} x^* = 0$ . **Podmínka komplementární skluzovosti** je splněna, jestliže pro přípustná řešení  $\tilde{x}, \tilde{y}$  platí buď  $\tilde{y}_i = 0$  nebo  $b_i - a_{i-} \tilde{x} = 0$ ,  $i = 1, \dots, m$ . Pokud nastane  $b_i - a_{i-} \tilde{x} = 0$ , potom říkáme, že **vazba**  $a_{i-} \tilde{x} \leq b_i$  je **aktivní**.

**Věta 6.** *Nechť  $\tilde{x}$  je přípustné řešení LP-P2 a  $\tilde{y}$  je přípustné řešení LP-D2. Potom  $\tilde{x}, \tilde{y}$  jsou optimální právě tehdy, když platí podmínka komplementární skluzovosti.*



# Kapitola 4

## Semidefinitní programování

Na semidefinitní programování se můžeme koukat jako na zobecnění lineárního programování, kde proměnné jsou symetrické matice. Jedná se tedy o optimalizaci lineární funkce vzhledem k tzv. lineárním maticovým nerovnostem.

### 4.1 Vsuvka z lineární algebry

#### Pozitivně definitní matice

Pracujeme s reálnými symetrickými maticemi  $S = S^T$ . Ty mají všechna vlastní čísla reálná a některé z nich mají zajímavou vlastnost, že všechna jejich vlastní čísla jsou kladná. Takovým maticím říkáme, že jsou pozitivně definitní. Alternativní definicí je, že matice  $S$  je pozitivně definitní, jestliže  $x^T S x > 0$  pro všechny nenulové vektory  $x$ .

**Příklad.**

$$x^T S x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 8x_1x_2 + 9x_2^2$$

Je pro všechny  $x$  nenulové  $x^T S x > 0$ ? Ano, protože můžeme výraz přepsat na součet čtverců:

$$x^T S x = 2x_1^2 + 8x_1x_2 + 9x_2^2 = 2(x_1 + 2x_2)^2 + x_2^2.$$

TODO: obrázek kyblíčku

Ukážeme si několik kritérií, jak otestovat pozitivní definitnost dané matice.

**Věta 7.**  $S = S^T$  je pozitivně definitní, jestliže lze napsat jako  $S = A^T A$  pro nějakou matici  $A$ , která má lineárně nezávislé sloupcečky.

*Důkaz.*

$$x^T S x = x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0 \quad (4.1)$$

$\|Ax\|^2 > 0$ , jestliže sloupčky matice  $A$  jsou lineárně nezávislé  $\square$

**Příklad.**

$$S = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 7 \\ 4 & 7 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} = A A^T$$

$A$  má lineárně závislé sloupčky, tj.  $S$  není pozitivně definitní

Dalším testem je tzv. Sylvesterovo kritérium.

**Věta 8.**  $S = S^T$  je pozitivně definitní, jestliže všechny hlavní minory  $S$  jsou kladné.

**Příklad.**

$$S = \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}, D_1 = 3, D_2 = 3 \cdot 6 - 4 \cdot 4 = 2$$

hlavní minory  $D_1, D_2 > 0$ ; matice  $S$  je pozitivně definitní

A poslední, které si uvedeme souvisí s Gaussovou eliminací.

**Věta 9.**  $S = S^T$  je pozitivně definitní, jestliže jsou všechny pivoty při Gaussově eliminaci kladné.

**Příklad.**

$$S = \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 3 & 4 \\ 0 & \frac{2}{3} \end{bmatrix}, p_1 = 3, p_2 = \frac{2}{3}$$

pivoty  $p_1, p_2 > 0$ ; matice  $S$  je pozitivně definitní

## Pozitivně semidefinitní matice

Pro pozitivní semidefinitnost modifikujeme předcházejí definice a tvrzení pro pozitivně definitní matice následovně:

1.  $S = S^T$  je pozitivně semidefinitní, jestliže všechna její čísla jsou nezáporná.
2.  $S = S^T$  je pozitivně semidefinitní, jestliže  $x^T S x \geq 0$  pro všechny nenulové vektory  $x$ .
3.  $S = S^T$  je pozitivně semidefinitní, jestliže lze napsat jako  $S = A^T A$  pro nějakou matici  $A$ .
4.  $S = S^T$  je pozitivně semidefinitní, jestliže všechny hlavní minory  $S$  jsou nezáporné.
5.  $S = S^T$  je pozitivně semidefinitní, jestliže jsou všechny pivoty při eliminaci nezáporné.

## Pozitivně semidefinitní kužel

Množinu všech symetrických matic značíme  $S^n$ , množinu všech pozitivně semidefinitních matic značíme  $S_+^n$  a množinu všech pozitivně definitních matic značíme  $S_{++}^n$ .

**Lemma 1.** *Množina  $S_+^n$  je uzavřená*

*Důkaz.* TODO □

**Lemma 2.** *Množina  $S_+^n$  tvoří konvexní kužel.*

*Důkaz.*  $\Theta_1, \Theta_2 \geq 0, A, B \in S_+^n$

$$x^T (\Theta_1 A + \Theta_2 B) x = x^T \Theta_1 A x + x^T \Theta_2 B x \geq 0.$$

□

**Lemma 3.** *Kužel  $S_+^n$  je pointed.*

*Důkaz.* TODO □

**Lemma 4.** *Kužel  $S_+^n$  je samoduální.*

*Důkaz.* TODO □

Shrneme předchozí lemmata do následující věty.

**Věta 10.** *Množina  $S_+^n$  tvoří konvexní, pointed a uzavřený kužel, který je samoduální.*

Říkáme, že  $S_+^n$  je **pozitivně semidefinitní kužel**.

## Spektraedry

Definujeme tzv. **Löwnerovo částečné uspořádání**:

$$A \succeq B \iff A - B \in S_+^n,$$

tj. matice  $A - B$  je pozitivně semidefinitní. **Lineární maticová nerovnost** (LMI) je ve tvaru

$$A_0 + \sum_{i=1}^n A_i x_i \succeq 0,$$

kde  $A_i \in S^n$ .

Množina  $S \subset \mathbb{R}^n$ , která je definována pomocí konečně mnoha LMI, se nazývá **spektraedr**. Tedy:

$$S = \left\{ (x_1, \dots, x_m) \in \mathbb{R}^m \mid A_0 + \sum_{i=1}^m A_i x_i \succeq 0 \right\}$$

pro nějaké symetrické matice  $A_0, \dots, A_m \in S^n$ .

Můžeme si všimnout analogie s definicí polyedru, který je přípustnou množinou pro lineární program. Podobně spektraedr je přípustnou množinou pro semidefinitní program.

Geometricky je spektraedr definován jako průnik pozitivně semidefinitního kuželu  $S_+^n$  a afinního podprostoru  $\text{span}\{A_1, \dots, A_m\}$  posunutého do  $A_0$ .

Spektraedry jsou uzavřené množiny, neboť LMI je ekvivalentní nekonečně mnoha skalárním nerovnostem ve tvaru  $v^T(A_0 + \sum_{i=1}^m A_i x_i)v \geq 0$ , jednu pro každou hodnotu  $v \in \mathbb{R}^n$ .

Vždy můžeme několik LMI „scucnout“ do jedné. Stačí zvolit matice  $A_i$  blokově diagonální. Odtud snadno víme, že polyedr je speciálním případem spektraedru. Polyedr bude mít všechny matice  $A_i$  diagonální.

**Příklad.**

$$\left\{ (x, y) \in \mathbb{R}^2 \mid A(x, y) = \begin{bmatrix} x+1 & 0 & y \\ 0 & 2 & -x-1 \\ y & -x-1 & 2 \end{bmatrix} \succeq 0 \right\}$$

## 4.2 Primární úloha

Semidefinitní program je lineární optimalizační problém přes spektraedr. Primární úlohu ve standardním tvaru můžeme napsat jako:

$$\inf \{ \langle C, X \rangle \mid \langle A_i, X \rangle = b_i, i = 1, \dots, m; X \succeq 0 \}, \quad (\text{SDP-P})$$

kde  $C, A_i \in S^n$ ,  $\langle X, Y \rangle = \text{Tr}(X^T Y) = \sum_{ij} X_{ij} Y_{ij}$  a  $X \in S^n$  je proměnná, nad kterou provádíme minimalizaci.

**Příklad.**

$$\inf \left\{ \left\langle \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \right\rangle \mid \left\langle \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \right\rangle = 1, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0 \right\} \quad (\text{P3})$$

Po úpravě:

$$\inf \left\{ 2x_{11} + 2x_{12} \mid x_{11} + x_{22} = 1, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0 \right\}.$$

Jak vypadá přípustná množina? Použijeme Sylvesterovo kritérium, tj.

$$x_{11} \geq 0, x_{11}x_{22} - x_{12}^2 \geq 0.$$

Z LMI vyjádříme  $x_{22}$ , tj.

$$x_{22} = 1 - x_{11}$$

Dosadíme do přechozího a dostaneme

$$x_{11} \geq 0, x_{11}(1 - x_{11}) - x_{12}^2 \geq 0$$

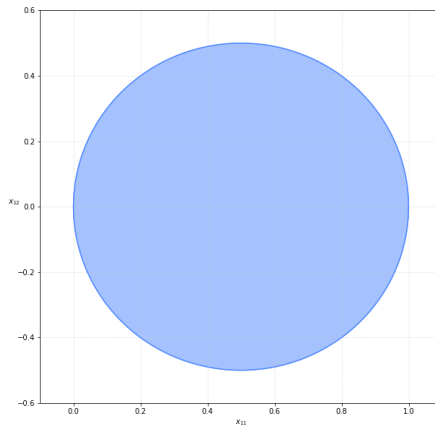
Po úpravě

$$x_{11} \geq 0, \left(x_{11} - \frac{1}{2}\right)^2 + x_{12}^2 \leq \frac{1}{4}$$

Vidíme tedy, že přípustná množina (zobrazena na obrázku 4.1) je kruh s poloměrem  $\frac{1}{2}$  a se středem v bodě  $(x_{11}, x_{12}) = (\frac{1}{2}, 0)$ . Řešením úlohy je matice

$$X^* \approx \begin{bmatrix} 0.1464 & -0.3536 \\ -0.3536 & 0.8536 \end{bmatrix}$$

s cenou  $\approx -0.4142$ . Implementace v softwaru MOSEK: <https://github.com/c0n73x7/D1PL0MK4/blob/master/mosek/ex3.py>.



Obrázek 4.1: Přípustná množina řešení k úloze P3.

### 4.3 Dualita

#### Duální úloha

Podobně jako u lineárního programování použijeme Lagrangeovu dualitu k odvození duální úlohy k úloze **SDP-P**. Lagrangeova funkce je ve tvaru:

$$L(X, \lambda, Z) = \langle C, X \rangle - \sum_{i=1}^m \lambda_i (\langle A_i, X \rangle - b_i) - \langle Z, X \rangle.$$

K ní duální funkce:

$$d(\lambda, Z) = \inf_{X \in S^n} L(X, \lambda, Z) = \begin{cases} \lambda^T b & \dots \ C - \sum_{i=1}^m \lambda_i A_i - Z = 0, \\ -\infty & \dots \text{jinak.} \end{cases}$$

Duální funkci použijeme v duální úloze:

$$\sup \left\{ \lambda^T b \mid C - \sum_{i=1}^m \lambda_i A_i \succeq 0 \right\}, \quad (\text{SDP-D})$$

kde  $\lambda = (\lambda_1, \dots, \lambda_m)$  je duální proměnná.

Dostali jsme duální úlohu **SDP-D** k úloze **SDP-P**.

#### Slabá dualita semidefinitního programování

Vztah mezi primární a duální úlohou je stejně jako u lineárního programování takový, že řešení jedné úlohy lze použít jako odhad na úlohu druhou. Necht  $X$  je libovolné přípustné řešení primární úlohy a  $y$  je libovolné přípustné řešení duální úlohy. Potom

$$\langle C, X \rangle - b^T y = \langle C, X \rangle - \sum_{i=1}^m y_i \langle A_i, X \rangle = \left\langle C - \sum_{i=1}^m A_i y_i, X \right\rangle \geq 0. \quad (4.2)$$

Za pozornost stojí poslední nerovnost, která plyne z toho, že skalární součin dvou pozitivně semidefinitních matic je nezáporný. Odvození je následující: mějme dvě matice  $S, T \succeq 0$ . Matici  $S$  můžeme napsat jako součet „rank one“ matic. Označme  $r_S = \text{rank}(S)$ , tj.

$$S = \sum_{i=1}^{r_S} \lambda_i s_i s_i^T,$$

kde  $s_i$  je  $i$ -tý sloupeček matice  $S$ . Dále se podíváme na součin  $T \cdot S_i$ . Tedy pro  $i = 1, \dots, r_S$  platí

$$T \cdot S_i = \lambda_i s_i^T T s_i \geq 0,$$

kde nerovnost plyne z toho, že matice  $T$  je pozitivně semidefinitní.

O nerovnosti 4.2 se mluví jako o slabé dualitě semidefinitního programování.

### Silná dualita semidefinitního programování

**Věta 11** (podmínka optimality). *Nechť  $X$  je přípustné řešení úlohy  $\text{SDP-P}$  a  $y$  je přípustné řešení úlohy  $\text{SDP-D}$  taková, že splňují podmínku (komplementární skluzovosti)*

$$\left( C - \sum_{i=1}^m A_i y_i \right) X = 0.$$

*Potom  $X$  je optimální řešení úlohy  $\text{SDP-P}$  a  $y$  je optimální řešení úlohy  $\text{SDP-D}$ .*

Obracená implikace sama o sobě neplatí, což znamená, že obecně dualitní rozdíl u semidefinitního programování není nulový. Musíme přidat podmínku kvalifikace omezení, kterou je například již zmíněná Slaterova podmínka. Ta je pro úlohu  $\text{SDP-P}$  ve tvaru  $X \succ 0$  a pro úlohu  $\text{SDP-D}$  ve tvaru  $C - \sum_i A_i y_i \succ 0$ .

**Věta 12** (silná dualita semidefinitního programování). *Nechť úloha  $\text{SDP-P}$  a úloha  $\text{SDP-D}$  jsou striktně přípustné. Potom dualitní rozdíl jejich optimálních řešení je nulový.*

## BACKLOG

Stopa matice  $A \in R^{n \times n}$  je  $\text{Tr}(A) = \sum_{i=1}^n A_{ii}$ .

**Věta 13.**  $S, T$  jsou pozitivně definitní  $\implies S + T$  je pozitivně definitní

*Důkaz.*  $x^T(S + T)x = x^T S x + x^T T x > 0$  □

Úlohy s racionálními daty nemusí mít racionální optimální řešení. příklad  
Jen 3 kužely s vlastnostmi jako PSD cone. Zmrzlina, a ještě jeden. Vnitřek  $S_+^n$  je  $S_{++}^n$ .

Fejérová věta