# Obsah

Ú	vod		2				
Ι	Oı	otimalizace	3				
1	Základní geometrické pojmy						
	1.1	Přímky a úsečky	4				
	1.2	Afinní prostory	4				
	1.3	Konvexní množiny	5				
	1.4	Kužely	5				
	1.5	Nadroviny a poloprostory	6				
	1.6	Polyedry a polytopy	8				
2	Konvexní optimalizace						
	2.1	Obecná podmíněná úloha	10				
	2.2	Konvexní podmíněná úloha	11				
	2.3	Lagrangeova dualita	11				
		2.3.1 Dolní odhad na $x^*$	11				
			12				
		2.3.3 Slabá dualita	12				
		2.3.4 Silná dualita a Slaterova podmínka	12				
3	Lineární programování 14						
	3.1	Primární úloha	14				
	3.2	Dualita	15				
	3.3	Komplementární skluzovost	17				
4	Semidefinitní programování 19						
	4.1		19				
	4.2	<u> </u>	22				
	4.3	Dualita	23				

OBSAH	1

5	Kombinatorická optimalizace  5.1 Celočíselné lineární programování				
II	Kombinatorické úlohy	26			
6	Shannonova kapacita	27			
	$6.1  \Theta(C_5) = \sqrt{5}  \dots  \dots  \dots  \dots  \dots  \dots  \dots$				
	6.2 Další vlastnosti $\vartheta(G)$				
	6.3 Semidefinitní program pro $\vartheta(G)$				
7	Problém maximálníhu řezu	31			
	7.1 Kvadratické programování	. 3			
	7.2 Relaxace a vektorové programování	. 3			
8	Problém obchodního cestujícího	32			
Zá	Závěr				

# $\acute{\mathbf{U}}\mathbf{vod}$

# Část I Optimalizace

# Základní geometrické pojmy

### 1.1 Přímky a úsečky

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

$$y = \theta x_1 + (1 - \theta)x_2 \tag{1.1}$$

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  $\theta_1x_1+\cdots+\theta_kx_k$  bodů  $x_1,\ldots,x_k$  taková, že  $\theta_1+\cdots+\theta_k=1$ , se nazývá **afinní kombinace** bodů  $x_1,\ldots,x_k$ . Indukcí z definice afinního prostoru lze snadno ukázat, že pokud C je afinní množina,  $x_1,\ldots,x_k\in C$  a  $\theta_1+\cdots+\theta_k=1$ , potom bod  $\theta_1x_1+\cdots+\theta_kx_k\in C$ .

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

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

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

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

Afinní obal je nejmenší afinní prostor, který obsahuje množinu C. Tedy, jestliže S je afinní prostor takový, že  $C \subseteq S$ , potom **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  $x_1, \ldots, x_k$  jako  $\theta_1 x_1 + \cdots + \theta_k x_k$ , kde  $\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$  pro  $i = 1, \ldots, k$ . **Konvexní obal** množiny C je množina všech konvexních kombinací bodů z množiny C, značíme

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

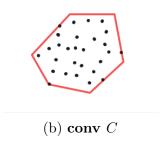
Analogicky, konvexní obal množiny C je nejmenší konvexní množina, která obsahuje množinu C. Pro představu viz obrázek 1.1.

## 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  $\theta_1, \ldots, \theta_k \geq 0$  je **kuželovou kombinací** bodů  $x_1, \ldots, x_k$ . Dále, pokud  $x_i$  leží v konvexním kuželu množiny C, potom libovolná kuželová kombinace bodu  $x_i$  leží rovněž



(a) Množina bodů C



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.

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

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

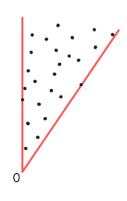
$$\left\{x \mid a^T x = b\right\},\,$$

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

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



#### (a) Množina bodů C



(b) cone C

Obrázek 1.2: Kuželový obal množiny

kde  $x_0$  je libovolný bod této nadroviny a  $\{v \mid a^Tv = 0\}$  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

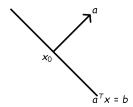
$$\{x \mid a^T x \le b\}$$
, resp.  $\{x \mid a^T x < b\}$ ,

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

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

kde  $a \neq 0$  a  $x_0$  je libovolný bod z nadroviny  $\{x \mid a^Tx = b\}$ . 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).  $P \subseteq \mathbb{R}^n$ .  $Potom P = conv(u_1, \dots, u_r) + cone(v_1, \dots, v_s)$ ,  $kde\ u_i, v_i\ jsou\ extremální\ vrcholy\ P\ právě\ tehdy,\ když\ P = \{x \in \mathbb{R}^n \mid Ax \leq b\}\ , A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ .

# Konvexní optimalizace

### 2.1 Obecná podmíněná úloha

min 
$$f(x)$$
  
 $g_i(x) \le 0, i = 1, ..., m$   
 $h_i(x) = 0, i = 1, ..., p$  (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

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

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

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

### 2.2 Konvexní podmíněná úloha

min 
$$f(x)$$
  
 $g_i(x) \le 0, i = 1, ..., m$  (2.2)  
 $a_i^T x = b_i, i = 1, ..., p$ 

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  $\mathcal{D} \neq 0$ . Zobrazení  $L: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$  takové, že

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

se nazývá **Lagrangeova funkce**. Definiční obor **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**  $d: \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$  jako infimum Lagrangeovy funkce L přes všechna  $x \in \mathcal{D}$ . Tedy

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

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

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

Potom pro Lagrangeovu funkci můžeme psát

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

A tedy pro duální funkci platí

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

#### 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.

$$\max_{\lambda > 0} d(\lambda, \mu) \tag{2.5}$$

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:

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

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  $f_1, \ldots, f_k, k \leq m$ , jsou afinní funkce. Potom **modifikovaná Slaterova podmínka** má tvar:

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

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

# 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:

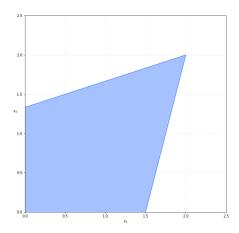
$$\max \left\{ c^T x \mid Ax \le b, x \ge 0 \right\}, \tag{LP-P}$$

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:

$$\max x_1 + x_2 - x_1 + 3x_2 \le 4 4x_1 - x_2 \le 6 x > 0.$$
 (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

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

Pro ní vytvoříme Lagrangeovu funkci

$$L(x,\lambda) = -c^T x + \lambda^T (Ax - b) - \lambda^T x$$
  
=  $-b^T \lambda + (A^T \lambda - c - \lambda)^T x$ .

Z Lagrangeovo funkce přejdeme k duální funkci

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

Tu nakonec použijeme v duální úloze:

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

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

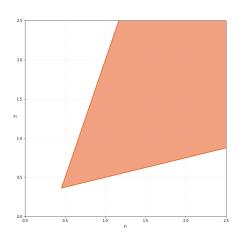
$$\min \{ b^T \lambda \mid A^T \lambda \ge c, \lambda \ge 0 \}$$
 (LP-D)

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

Příklad. Duální úloha k úloze P1 je ve tvaru:

$$\min 4y_1 + 6y_2 
-y_1 + 4y_2 \ge 1 
3y_1 - y_2 \ge 1 
y \ge 0.$$
(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  $c^Tx^*=b^Ty^*$ . 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

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

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:

$$\max\left\{c^{T}x \mid Ax \le b\right\}. \tag{LP-P2}$$

A příslušná duální úloha je ve tvaru:

$$\min\left\{b^T y \mid A^T y = c, y \ge 0\right\}. \tag{LP-D2}$$

Nechť  $\tilde{x}$  je připustné 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:

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

Poslední rovnost přepíšeme maticově:

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

Dostáváme tedy soustavu rovnic  $y_i^*$  ( $b_i - a_i x^*$ ) = 0, kde i = 1, ..., 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, ..., 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.

#### **TODOs**

# 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  $x^TSx>0$  pro všechny nenulové vektory x.

#### Příklad.

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

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

$$x^{T}Sx = 2x_1^2 + 8x_1x_2 + 9x_2^2 = 2(x_1 + 2x_2)^2 + x_2^2.$$

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 sloupečky.

20

Důkaz.

$$x^{T}Sx = x^{T}A^{T}Ax = (Ax)^{T}(Ax) = ||Ax||^{2} \ge 0$$
(4.1)

 $||Ax||^2 > 0$ , jestliže sloupečky matice A jsou lineárně nezávislé  $\Box$ 

Příklad.

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

A má lineárně závislé sloupečky, 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  $S=A^TA$  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.

21

#### 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  $S^n_{++}$ .

Věta 10. Množina  $S^n_+$  tvoří konvexní kužel.

 $D\mathring{u}kaz. \ \Theta_1, \Theta_2 \ge 0, \ A, B \in S^n_+$ 

$$x^{T} (\Theta_1 A + \Theta_2 B) x = x^{T} \Theta_1 A x + x^{T} \Theta_2 B x > 0.$$

Množině  $S^n_+$  se říká pozitivně semidefinitní kužel.

??chci ukázat i uzavřenost a další vlastnosti, jak se česky řekne proper cone??

#### Spektraedry

Definujeme tzv. Löwnerovo částečné uspořádání

$$A \succeq B \iff A - B \in S^n_+$$
.

Definice 1. Lineární maticová nerovnost (LMI) je ve tvaru

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

kde  $A_i \in S^n$ .

**Definice 2.** Říkáme, že  $S \subset \mathbb{R}^n$  je spektraedr, jestliže lze napsat ve tvaru

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

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

Spektraedr je tedy množina, která je definována konečným počtem LMI. Můžeme si všimnout analogie s definicí polyedru, který je přípustnou množinu 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 **span**  $\{A_1, \ldots, A_m\}$  posunutého do  $A_0$ .

Spektraedry jsou uzavřené množiny, neboť LMI je ekvivalentní nekonečně mnoha skalárních nerovností 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 vídíme, 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:

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

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

#### Příklad.

$$\min\left\{\left\langle \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \right\rangle \middle| \left\langle \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \right\rangle = 1, \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0 \right\}$$

Po úpravě:

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

Jak vypadá přípustná množina? Použijeme Sylvesterovo kritérium, tj.

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

Z LMI vyjádříme  $x_{22}$ , tj.

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

Dosadíme do přechozího a dostaneme

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

Po úpravě

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

Vidíme tedy, že přípustná množina (zobrazena na obrázku TODO) 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.

TODO: obrázek přípustné množiny s optimálním řešením z mosek implementace

#### 4.3 Dualita

$$\max \left\{ b^T y \; \middle| \; \sum_{i=1}^m A_i y_i \preceq C \right\},\,$$

kde  $b = (b_1, \ldots, b_m)$  a  $y = (y_1, \ldots, y_m)$  jsou duální proměnné.

Vztah mezi primární a duální úlohou je podobně jako u lineárního programování v tom, že řešení jedné úlohy lze použít jako odhad na úlohu druhou. Nechť X je libovolné přípustné řešení primární úlohy a y je libovolné přípustné řešení duální úlohy. Potom

$$\langle C, X \rangle - b^T y = \langle C, X \rangle - \sum_{i=1}^m y_i \langle A_i, X \rangle = \left\langle C - \sum_{i=1}^m A_i y_i, X \right\rangle \ge 0$$

#### WIP

### Odvození duální úlohy

Lagrangeovo funkce

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

primární úloha odpovídá

$$\min_{X \in S^n} \max_{\lambda \in \mathbb{R}^m, Z \succeq 0} L(X, \lambda, Z)$$

z Lagrangeovy duality

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

eliminujeme Z

$$\sup \left\{ b^T \lambda \mid \sum_{i=1}^m A_i \lambda_i \preceq C \right\}$$

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

**Věta 11.** S, T jsou pozitivně definitní  $\implies S + T$  je pozitivně definitní

$$D\mathring{u}kaz. \ x^T(S+T)x = x^TSx + x^TTx > 0$$

Úlohy s racionálními daty nemusí mít racionální optimální řešení. Lagrangovo funkce a dualita konvexního programování (příloha?); z toho dualitu semidefinitního programování; rozdíl oproti lineárnímu programování; příklad

#### Relaxace

vektorové programování

# Kombinatorická optimalizace

# 5.1 Celočíselné lineární programování

složitost úlohy, relaxace, výčet algoritmů, aproximace, ...

# 5.2 Celočíselné semidefinitní programování

složitost úlohy, relaxace, výčet algoritmů, aproximace, ...

# Část II Kombinatorické úlohy

# Shannonova kapacita

Představme si zašuměný komunikační kanál, kterým posíláme zprávy, které jsou složeny ze symbolů (písmen) nějaké konečné abecedy. Vlivem šumu mohou být některé symboly špatně interpretovány a naším cílem je vybrat co největší počet slov délky k tak, aby žádná dvě slova nebyla vlivem šumu zaměnitelná.

Problém si formalizujeme v řeči teorie grafů. Mějme neorientovaný graf G=(V,E), kde množina vrcholů představuje symboly z konečné abecedy a dva vrcholy x,y jsou spojeny hranou, pokud vrchol x může být vlivem šumu zaměněn za y.

Maximální počet nezaměnitelných zpráv délky 1 je roven  $\alpha(G)$ , kde  $\alpha(G)$  značí velikost největší nezávislé množiny v grafu G. Pro popis delších zpráv definujeme **silný součin**  $G \cdot H$  grafů G a H následovně:

$$\begin{split} V(G \cdot H) &= V(G) \times V(H), \\ E(G \cdot H) &= \{(i, u)(j, v) \mid ij \in E(G) \wedge uv \in E(H)\} \cup \\ \{(i, u)(j, v) \mid ij \in E(G) \wedge u = v\} \cup \\ \{(i, u)(j, v) \mid i = j \wedge uv \in E(H)\} \,. \end{split}$$

**Příklad.** Pro graf  $P_4 = a - b - c - d - e$  je silný součin  $P_4 \cdot P_4$  zobrazen na obrázku 6.1. Z obrázku je hezky vidět, že např. zpráva cd (na obrázku červeně) může být zaměněna s bc, bd, be, cc, ce, dc, dd a de (na obrázku oranžově). Podobně pro další zprávy.

Pro jednoduchost budeme silný součin k kopií grafu G značit  $G^k$ . Tedy  $\alpha(G^k)$  je maximální počet nezaměnitelných zpráv délky k. Shannonova kapacita grafu G je definována jako

$$\Theta(G) = \sup \{ \alpha(G^k)^{1/k} \mid k = 1, 2, \dots \}.$$



Obrázek 6.1:  $P_4 \cdot P_4$ 

Neví se, zda pro libovolný graf G existuje vůběc nějaký algoritmus, kterým bychom určili hodnotu  $\Theta(G)$ . Přesto je alespoň něco známo. Pro perfektní grafy Claude E. Shannon ukázal, že  $\Theta(G) = \alpha(G)$ . To také znamená, že pro perfektní grafy lze  $\Theta(G)$  určit v polynomiálním čase. Dalším kdo se problémem zabýval byl László Lovász, který velmi hezkým způsobem ukázal, že kružnice délky 5 má kapacitu  $\sqrt{5}$ . Na Lovászův postup se dále podíváme, protože vede k obecnému hornímu odhadu na  $\Theta(G)$ .

**6.1** 
$$\Theta(C_5) = \sqrt{5}$$

Tenzorový součin vektorů  $\mathbf{u} = (u_1, \dots, u_n)$  a  $\mathbf{v} = (v_1, \dots, v_m)$  je

$$\mathbf{u} \circ \mathbf{v} = (u_1 v_1, \dots, u_1 v_m, u_2 v_1, \dots, u_n v_m).$$

Užitečné bude následující pozorování, které dává do souvisloti skalární a tenzorový součin.

Pozorování. Nechť **x**, **u** jsou vektory délky n a **y**, **v** jsou vektory délky m. Potom platí

$$(x \circ y)^{T} (u \circ v) = (x^{T}u) (y^{T}v).$$
(6.1)

Důkaz. Levá strana:

$$(x_1y_1, x_1y_2, \dots, x_1y_m, \dots, x_ny_m)^T (u_1v_1, u_1v_2, \dots, u_1v_m, \dots, u_nv_m) = x_1y_1u_1v_1 + x_1y_2u_1v_2 + \dots + x_1y_mu_1v_m + \dots + x_my_mu_nv_m$$

Pravá strana:

$$(x_1u_1 + \dots + x_nu_n) \cdot (y_1v_1 + \dots + y_nv_m) = x_1y_1u_1v_1 + x_1y_2u_1v_2 + \dots + x_1y_mu_1v_m + \dots + x_my_mu_nv_m$$

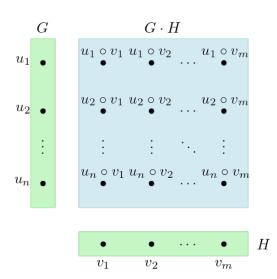
Mějme graf G = (V, E), kde  $V = \{1, \dots, n\}$ . Systém  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  jednotkových vektorů v Euklidovském prostoru takový, že

$$\forall ij \notin E \implies \mathbf{v}_i \perp \mathbf{v}_i$$

nazýváme **ortonormální reprezentace** grafu G. Poznamenejme, že každý graf má nějakou ortonormální reprezentaci, např.  $1 \mapsto \mathbf{e}_1, \dots, n \mapsto \mathbf{e}_n$ .

**Lemma 1.** Nechť  $(\mathbf{u}_1, \dots, \mathbf{u}_n)$  je ortonormální reprezentace grafu G a  $(\mathbf{v}_1, \dots, \mathbf{v}_m)$  je ortonormální reprezentace grafu H. Potom  $\mathbf{u}_i \circ \mathbf{v}_j$  je ortonormální reprezentace grafu  $G \cdot H$ .

 $D\mathring{u}kaz$ . Použijeme vztah 6.1.  $(u_i \circ v_j)^T (u_k \circ v_l) = (u_i^T u_k) (v_j^T v_l) = 0 \iff ik \notin E(G) \lor jl \notin E(H)$ .



Obrázek 6.2: Ortornomální reprezentace  $G \cdot H$ .

**Hodnotu** ortonormální reprezentace  $(u_+, \ldots, u_n)$  definujeme jako:

$$\min_{c} \max_{i=1,\dots,n} \frac{1}{\left(c^T u_i\right)^2}.$$

Vektoru c, pro který nastává minimum říkáme **handle** dané ortonormální reprezentace.

Dále definujeme funkci  $\vartheta(G)$  jako minimální hodnotu přes všechny ortonormální reprezentace grafu G. Ortonormální reprezentaci, pro kterou nastává minumum nazýváme **optimální**.

Funkci  $\vartheta(G)$  se říká **Lovászova theta funkce** a ona je právě již zmíněným horním odhadem na  $\Theta(G)$ . Podívejme se na některé její vlastnosti.

Lemma 2. 
$$\vartheta(G \cdot H) \leq \vartheta(G)\vartheta(H)$$

 $D\mathring{u}kaz$ . Nechť  $(u_1,\ldots,u_n)$  je optimální ortonormální reprezentace grafu G s "rukojetí" c a  $(v_1,\ldots,v_m)$  je optimální ortonormální reprezentace grafu H s "rukojetí" d. Pak  $c \circ d$  je jednotkový vektor a platí:

$$\vartheta(G \cdot H) \le \max_{i,j} \frac{1}{\left(\left(c \circ d\right)^T \left(u_i \circ v_j\right)\right)^2} = \max_i \frac{1}{\left(c^T u_i\right)^2} \cdot \max_j \frac{1}{\left(d^T v_j\right)^2} = \vartheta(G)\vartheta(H).$$

Lemma 3.  $\alpha(G) \leq \vartheta(G)$ 

 $D\mathring{u}kaz$ . TODO (máš to někde na papíře)

Lemma 4.  $\Theta(G) < \vartheta(G)$ 

 $D\mathring{u}kaz$ . TODO (máš to někde na papíře)

Věta 12.  $\Theta(C_5) = \sqrt{5}$ 

Důkaz. TODO (obě nerovnosti, obrázek, spherical cosine theorem)

## **6.2** Další vlastnosti $\vartheta(G)$

vztah k barvení  $(\overline{G})$ , ...

### 6.3 Semidefinitní program pro $\vartheta(G)$

formulace semidefinitních programů (jsou dva ekvivalentní – to asi nenaimplementuješ, je to docela pekelný, nic nespočítáš), subgradientní aproximační metoda (zkusit naimplementovat???)

# Problém maximálníhu řezu

# 7.1 Kvadratické programování

kvadratický program, striktní kvadratický program

# 7.2 Relaxace a vektorové programování

ekvivalence se semidefinitním programováním

### **TODO**

formulace úlohy, approximační algoritmy, porovnání semidefinitních programů,  $\dots$  komplexní semidefinitní programování,  $\dots$ 

# Kapitola 8 Problém obchodního cestujícího

# Závěr