Obsah

Ú	vod	2				
Ι	Konvexní optimalizace	3				
1	Základní geometrické pojmy	4				
	1.1 Přímky a úsečky	4				
	1.2 Affiní prostory	4				
	1.3 Konvexní množiny	5				
	1.4 Kužely	5				
	1.5 Nadroviny a poloprostory	6				
	1.6 Polyedry a polytopy	8				
2	Lineární programování					
	2.1 Formulace úlohy	10				
	2.2 Dualita	11				
	2.3 Komplementární skluzovost	12				
3	Semidefinitní programování	14				
4	Kuželové programování					
II	Kombinatorické úlohy	16				
5	Shannonova kapacita	17				
	5.1 $\Theta(C_5) = \sqrt{5}$	18				
	5.2 Semidefinitní program pro $\vartheta(G)$					
6	Problém maximálníhu řezu					
7	Problém obchodního cestujícího					

OBSAH			1

23

Závěr

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

Část I Konvexní 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 θ 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 Affiní 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 affiních kombinací bodů množiny $C \subseteq \mathbb{R}^n$ se nazývá **affiní obal** množiny C. Affiní obal množiny C budeme značit

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

Affiní obal je nejmenší affiní prostor, který obsahuje množinu C. Tedy, jestliže S je affiní 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 affiní 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



(b) 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.

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



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 vektorů a. Nadrovina je tedy množina, která obsahuje bod x_0 a libovolný bod ve tvarů $x_0 + v$, kde v je vektor, který je kolmý k normálovému vektorů 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 affiní.



(b) Poloprostor

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

1.6 Polyedry a polytopy

Mějmě konečně mnoho uzavřených poloprostorů v \mathbb{R}^n . Množina, která vznikne jejich průnikem se nazývá **polytop**. Je-li navíc polytop omezený, potom ho nazýváme **polyedr**. Polyedr lze také ekvivalentně definovat jako konvexní obal konečně mnoha bodů v \mathbb{R}^n . Příkladem polyedrů v \mathbb{R}^3 jsou např. platónská tělesa, viz obrázek 1.4. Důležitý fakt říká Minkowského-Weyleova věta: každý polytop P je konečně generovaný a můžeme ho vyjadřit jako

$$P =$$
conv $(u_1, ..., u_r) +$ **cone** $(v_1, ..., v_s),$

kde u_i, v_i jsou extremální vrcholy P.



(a) Dodecahedron.



(b) Icosahedron.



(c) Octahedron.

Obrázek 1.4: Platónská tělesa.

Lineární programování

2.1 Formulace úlohy

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 2.1. 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 2.1: Přípustná množina řešení k úloze P1.

2.2 Dualita

Úloha LP-P se nazývá **primární úloha**. Ke každé primární úloze můžeme přiřadit příslušnou **duální úlohu**. Je to opět úloha lineárního programování, která pro případ LP-P je ve tvaru:

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

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 2.2. 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.

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, kterou dokázala skupina kolem Alberta W. Tuckera v roce 1948. Začneme slabou větou o dualitě.

Věta 1 (Slabá o dualitě). Nechť \tilde{x} je přípustné řešení LP-P a \tilde{y} je přípustné řešení LP-D. Potom $c^T \tilde{x} \leq b^T \tilde{y}$.



Obrázek 2.2: Přípustná množina řešení k úloze P2.

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 2.3. 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 2.3: Slabá věta o dualitě.

Věta 2 (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 2.3 upravit na obrázek 2.4.

2.3 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



Obrázek 2.4: Ceny přípustných řešení primární a příslušné duální úlohy.

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:

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

Dostáváme tedy soustavu rovnic $y_i^* (b_i - a_i x^*) = 0$, kde i = 1, ..., 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 3. 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 3 Semidefinitní programování

Kapitola 4 Kuželové programování

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.

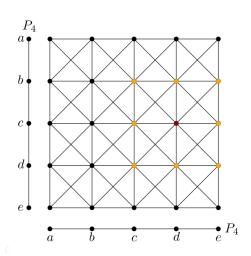
Maximální počet nezaměnitelných zpráv délky 1 je roven $\alpha(G)$, kde $\alpha(G)$ značí velikost největší nezávislé množiny v grafu G. Pro popis delších zpráv definujeme **silný součin** $G \cdot H$ grafů G a H následovně:

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

Příklad. Pro graf $P_4 = a - b - c - d - e$ je silný součin $P_4 \cdot P_4$ zobrazen na obrázku 5.1. 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 5.1: $P_4 \cdot P_4$

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 $\Theta(G)$.

5.1
$$\Theta(C_5) = \sqrt{5}$$

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

$$(x \circ y)^{T} (u \circ v) = (x^{T}u) (y^{T}v).$$
(5.1)

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:

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

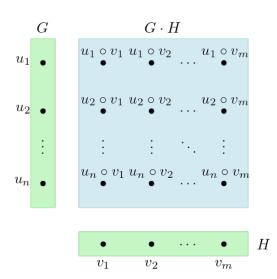
Mějme graf G = (V, E), kde $V = \{1, \dots, n\}$. Systém $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ jednotkových vektorů v Euklidovském prostoru takový, že

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

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 5.1. $(u_i \circ v_j)^T (u_k \circ v_l) = (u_i^T u_k) (v_j^T v_l) = 0 \iff ik \notin E(G) \lor jl \notin E(H)$.



Obrázek 5.2: 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 "rukojeť" (anglicky 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:

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

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

Důkaz. TODO (máš to někde na papíře) □

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

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

Věta 4. $\Theta(C_5) = \sqrt{5}$

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

5.2 Semidefinitní program pro $\vartheta(G)$

formulace semidefinitních programů (jsou dva ekvivalentní), subgradientní aproximační metoda (zkusit naimplementovat???)

Problém maximálníhu řezu

formulace úlohy, approximační algoritmy, porovnání semidefinitních programů..

Problém obchodního cestujícího

Závěr