Obsah

Úvod

Semidefinitné programovanie je relatívne nová oblasť konvexného programovania. Mnohé úlohy sa dajú formulovať ako úlohy SDP. Pre mnohé ďalšie optimalizačné úlohy sú úlohy SDP dobrým priblížením a teda pomocou SDP ich možno približne riešiť. Využitie SDP je naozaj rozsiahle. Prvé využitie SDP spočívalo vo výpočtoch vlastných čísiel symetrických matíc. Dnes SDP nachádza uplatnenie od teórie riadenia až po kombinatorickú optimalizáciu.

Semidefinitné programovanie je nadstavbou lineárneho programovania. Vektorová nezáporná premenná z lineárneho programovania je v semidefinitnom programovaní nahradená kladne semidefinitnou symetrickou maticou. Podobnosť SDP a LP sa uplatňuje hlavne v teórii duality a v metódach vnútorného bodu. Väčšina výsledkov z teórie duality pre SDP môže byť odvodená analogicky ako v LP. V niektorých prípadoch ale potrebujeme silnejšie predpklady ako boli v LP. SDP môže byť veľmi efektívne riešené metódami vnútorného bodu, ktoré boli pôvodne vyvinuté pre LP.

SDP je stále vo vývoji s množstvom otvorených problémov. Z toho dôvodu je veľmi aktívna publikačná činnosť a mnohé práce k tejto oblasti možno nájsť na internetových adresách:
http://www.zib.de/helmberg/semidef.html
http://new-rutcor.rutgers.edu/ alizadeh/sdp.html
http://www.mcs.anl.gov:80/home/otc/InteriorPoint

Hoci sa už v tejto oblasti objavilo aj niekoľko monografií, tieto majú zvyčajne iba charakter zborníkov a teda zatiaľ v tejto oblasti neexistuje kniha, ktorá by neodborníkovi v tejto oblasti umožňovala ľahko pochopiť základy tejto disciplíny. Preto cieľom tejto diplomovej práce bolo podrobne ale pritom čo najelementárnejšie sformulovať a dokázať niektoré základné výsledky semidefinitného programovania. Pri plnení tohoto cieľa sme vychádzali hlavne z [1] a [2], pričom niektoré postupy boli konfrontované s literatúrou z internetu a to hlavne s prácami [3], [4] a [6].

Okrem teoretických prác sú na internete dostupné aj niektoré programy umožňujúce riešiť konkrétne úlohy semidefinitného programovania pomocou rozličných algoritmov metód vnútorného bodu. Vedľajším cieľom diplomovej práce bolo preskúmať možnosti použitia programu [SeDuMi] na výpočet úloh SDP.

Diplomová práca je rozdelená do piatich častí. V prvej časti je formulovaná úloha SDP. V druhej časti sú rozobraté niektoré špeciálne prípady úloh SDP. Je tam ukázané, za akých podmienok možno úlohu lineárneho programovania stotožniť s úlohou SDP a prevedenie úlohy konvexného-kvadratického programovania a úlohy second order cone programming na úlohu semidefinitného programovania. Tretia časť je o aplikovaní úloh semidefinitného programovania. Úloha minimalizovať maximálne vlastné číslo lineárnej kombinácii matíc je priamo prevedená na úlohu semidefinitného programovania. Problém maximálneho rezu sa vo všeobecnosti nedá priamo previesť na úlohu semidefinitného programovania ale pomocou úlohy SDP možno tento problém približne riešiť. Štvtá časť je o dualite v SDP. V porovnaní s LP je dualita v SDP rozmanitejšia. Piata časť pojednáva o metódach vnútorného bodu. V šiestej kapitole sú zhrnuté poznatky o použití programu [SeDumi]. Siedma kapitola pojednáva o základných vlastnostiach symetrických a (semi)definitných matíc.

1. Úlohy SDP

1.1 Formulácia štandartnej úlohy SDP.

Priestor reálnych symetrických $n \times n$ matíc budeme označovať $S^n \subset R^{nxn}$ a na S^n budeme definovať skalárny súčin pre $M, N \in S^n$ definovaný ako

$$M \bullet N = \text{stopa}(MN) = \sum_{i,j} m_{ij} n_{ij},$$

kde m_{ij} resp. n_{ij} označuje prvok matice M resp. N v i-tom riadku a j-tom stĺpci. Označením $M \succeq 0$ ($M \succ 0$), kde $M \in S^n$, považujeme maticu M za kladne semidefinitnú (definitnú) maticu.

Úlohu semidefinitného programovania v štandartnom tvare môžeme zapísať ako

$$\min_{X \in S^n} \{C \bullet X \mid A_i \bullet X = b_i,\ i = 1, \ldots, m,\ X \succeq 0\}, \tag{SDP}$$

kde $C,\ A_i \in S^n$, $i = 1, \ldots, m$ a $b = (b_1, b_2, \ldots, b_m)^T \in R^m$.

Poznámka 1.1.1. V úlohe (SDP) predpokladáme, že matice $C,\ A_i,\ i = 1, \ldots, m$ sú symetrické. Tento predpoklad nás ale neobmedzuje. Ak by matice $C,\ A_i \notin S^n,\ i = 1, \ldots, m$ ale žiadali by sme aby matica $X \in S^n$ a mali by sme úlohu

$$\min_{X \in S^n} \{stopa(CX) \mid stopa(A_i X) = b_i,\ i = 1, \ldots, m,\ X \succeq 0\}, \tag{$*$}$$

tak na základe vlastností stopy matice dostávame, že

$$stopa(CX) = stopa(C^T X) \tag{1.1.1}$$

a podobne

$$stopa(A_i X) = stopa(A_i^T X). \tag{1.1.2}$$

Z (1.1.2) vidíme, že

$$stopa(A_i X) = b_i, \tag{1.1.3}$$

$$stopa(A_i^T X) = b_i \tag{1.1.4}$$

a sčítaním (1.1.3) a (1.1.4) dostaneme

$$stopa(A_i X + A_i^T X) = 2b_i \tag{1.1.5}$$

$$stopa\left(\frac{A_i + A_i^T}{2} X\right) = b_i. \tag{1.1.6}$$

Použijeme substitúcie

$$\tilde{C} = \frac{C + C^T}{2},\ \tilde{A}_i = \frac{A_i + A_i^T}{2},\ i = 1, \ldots, m \tag{1.1.7}$$

a ľahko nahliadneme, že matice $\tilde{C},\ \tilde{A}_i,\ i = 1,\dots,m$ sú symetrické. Pomocou (1.1.1) a (1.1.7) vyjadríme hodnotu účelovej funkcie

$$\tilde{C} \bullet X = \frac{C + C^T}{2} \bullet X = stopa\left(\frac{CX}{2}\right) + stopa\left(\frac{C^T X}{2}\right) = stopa(CX).$$

Podobne z (1.1.2) a (1.1.7) pre $i = 1,\dots,m$ dostávame

$$\tilde{A}_i \bullet X = \frac{A_i + A_i^T}{2} \bullet X = stopa\left(\frac{A_i X}{2}\right) + stopa\left(\frac{A_i^T X}{2}\right) = stopa(A_i X).$$

Z toho vidíme, že pre úlohu určenú maticami $C,\ A_i \notin S^n,\ i = 1,\dots,m$ ako aj pre úlohu určenú maticami $\tilde{C},\ \tilde{A}_i \in S^n,\ i = 1,\dots,m$ dostaneme tú istú množinu pripustných riešení ako aj tú istú účelovú funkciu. Preto bez obmedzenia môžeme predpokladať, že $C,\ A_i \in S^n,\ i = 1,\dots,m$.

Poznámka 1.1.2. Pre množinu prípustných riešení primárnej úlohy SDP budeme mať označenie

$$\mathcal{P} = \{X \in S^n \mid A_i \bullet X = b_i,\ i = 1,\dots,m,\ X \succeq 0\}.$$

Množinu vnútorných bodov budeme označovať

$$\mathcal{P}^0 = \{X^0 \in S^n \mid A_i \bullet X^0 = b_i,\ i = 1,\dots,m,\ X^0 \succ 0\}.$$

A množinu optimálnych riešení budeme značiť

$$\mathcal{P}^* = \{X^* \in \mathcal{P} \mid C \bullet X^* \geq C \bullet X \text{ pre ľubovoľné } X \in \mathcal{P}\}.$$

1.2 Formulácia duálnej úlohy k SDP.

K primárnej úlohe semidefinitného programovania v štandartnom tvare (SDP) môžeme príslušnú duálnu úlohu napísať ako

$$\max_{y \in R^m, Z \in S^n} \{b^T y \mid \sum_{i=1}^{m} y_i A_i + Z = C,\ Z \succeq 0\}, \tag{SDD}$$

ktorú môžeme prepísať do iného tvaru bez Z

$$\max_{y \in R^m} \{b^T y \mid C - \sum_{i=1}^{m} y_i A_i \succeq 0\}. \tag{1.2.1}$$

To, že táto úloha je naozaj duálna k úlohe (SDP) sa presvedčíme v Kapitole 4 o dualite v SDP.

Poznámka 1.2.1. Pre množinu prípustných riešení duálnej úlohy (SDD) budeme mať označenie

$$\mathcal{D} = \{(y, Z) \in R^m \times S^n \mid \sum_{i=1}^{m} y_i A_i + Z = C,\ Z \succeq 0\}.$$

Množinu vnútorných bodov budeme označovať

$$\mathcal{D}^0 = \{(y^0, Z^0) \in R^m \times S^n \mid \sum_{i=1}^{m} y_i^0 A_i + Z^0 = C, Z^0 \succ 0\}.$$

A množinu optimálnych riešení budeme značiť

$$\mathcal{D}^* = \{(y^*, Z^*) \in D \mid b^T y^* \geq b^T y \text{ pre ľubovoľné } (y, Z) \in \mathcal{D}\}.$$

Predpoklad 1. Matice A_i, $i = 1, ..., m$ sú lineárne nezávislé, čo znamená, že z rovnosti $\sum_{i=1}^m A_i u_i = 0$ vyplýva, že $u_i = 0$, $i = 1, ..., m$.

Poznámka 1.2.2. Predpoklad 1 zaručuje jednoznačnú korešpondenciu medzi y a Z. To znamená, ak máme dané y, tak príslušné Z je jednoznačne určené vzťahom

$$Z = C - \sum_{i=1}^{m} A_i y_i.$$

Opačne, pre každé $Z \succeq 0$ existuje najviac jedno y také, že

$$\sum_{i=1}^{m} a_i y_i + Z = C.$$

Naozaj, ak by existovali také $y^1 \neq y^2$, pre ktoré by platilo

$$\sum_{i=1}^{m} a_i y_i^1 + Z = C, \tag{1.2.2}$$

$$\sum_{i=1}^{m} a_i y_i^2 + Z = C, \tag{1.2.3}$$

tak odčítaním rovnice (1.2.3) od rovnice (1.2.2) dostaneme

$$\sum_{i=1}^{m} A_i (y_i^1 - y_i^2) = 0.$$

Vzhľadom na lineárnu nezávislosť matíc A_i, $i = 1, \ldots, m$, dostávame

$$y_i^1 - y_i^2 = 0, \; i = 1, \ldots, m,$$

čo je spor s predpokladom $y^1 \neq y^2$.

Poznámka 1.2.3. Uvedený Predpoklad 1 nie je obmedzujúci, lebo každá úloha môže byť pretransformovaná na úlohu, ktorá uvedený predpoklad spĺňa.

Definícia 1.2.4. Zadefinujeme vektorovú funkciu maticového argumentu, kde každému $X \in S^n$ priradíme vektor $vec(X) \in R^{n^2}$ nasledovne

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} \rightarrow vec(X) = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \\ x_{21} \\ \vdots \\ x_{2n} \\ \vdots \\ x_{n1} \\ \vdots \\ x_{nn} \end{pmatrix}.$$

Poznámka 1.2.5. Pomocou funkcie vec môžeme definovať maticu A $\in R^{m\times n^2}$ ako

$$A = \begin{pmatrix} vec(A_1)^T \\ vec(A_2)^T \\ \vdots \\ vec(A_m)^T \end{pmatrix}. \tag{1.2.4}$$

Poznámka 1.2.6. Pomocou matice A môžeme Predpoklad 1 formulovať aj ako

$$rank(A) = m.$$

Poznámka 1.2.7. Pomocou funkcie vec možno úlohu (SDP) zapísať v tvare

$$\min_{X\in S^n} \{(vec(C))^T vec(X) \mid Avec(X) = b, X \succeq 0\} \tag{1.2.5}$$

a úlohu (SDD) v tvare

$$\max_{y\in R^n,\ Z\in S^n} \{b^T y \mid A^T y + vec(Z) = vec(C),\ Z \succeq 0\}. \tag{1.2.6}$$

Veta 1.2.8. *Úloha (SDP) je úlohou konvexného programovania.*

Dôkaz. Účelová funkcia v (SDP) je lineárna a teda aj konvexná. Vieme, že množina kladne semidefinitných matíc vytvára konvexný kužel v priestore $R^{n\times n}$, lebo pre ľubovoľné $\alpha \geq 0$ a ľubovoľné $X, Y \in S^n_+$ ľahko nahliadneme, že $\alpha X \in S^n_+$ a $X + Y \in S^n_+$. Ohraničenie $Avec(X) = b$ určuje afinnú rovinu, ktorej prienik s konvexným kuželom vytvára konvexnú množinu. Z toho vidíme že ide o úlohu konvexného programovania. $\square$

Poznámka 1.2.9. Úloha (SDP) sa navonok javí ako lineárna úloha, konvexnosť je skrytá iba v ohraničení $X \succeq 0$.

2. Špeciálne prípady úloh SDP

2.1 Úloha lineárneho programovania.

Uvažujme úlohu lineárneho programovania (LP) v tvare

$$\min_{x\in R^n} \{c^T x \mid {a_i}^T x = b_i,\ i = 1,\dots,m,\ x \geq 0\}, \tag{LP}$$

$$\text{kde } a_i \in R^n,\ i = 1,\dots,m,\ c \in R^n,\ b = (b_1,\dots,b_m)^T \in R^m.$$

Definujme (Definica 7.2)

$$C = Diag(c) = \begin{pmatrix} c_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & c_n \end{pmatrix},\ A_i = Diag(a_i) = \begin{pmatrix} a_i^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_i^n \end{pmatrix},\ i = 1,\dots,m,$$

kde $a_i^j,\ j = 1,\dots,n$ je j-ty prvok vektora a_i kde $i = 1,\dots,m$. K získaným maticiam priradíme úlohu

$$\min_{X\in S^n} \{C \bullet X \mid A_i \bullet X = b_i,\ i = 1,\dots,m;\ X \succeq 0\}. \tag{$\star$}$$

Veta 2.1.1. *Úlohy (LP) a ($\star$) sú ekvivalentné ak každé $x \geq 0$ stotožníme s množinou všetkých kladne semidefinitných matíc takých, že $diag(X) = x$.*

Dôkaz. Ukážeme, že vzhľadom na uvedené stotožnenie obidve úlohy majú rovnakú množinu prípustných aj optimálnych riešení.
(1) Množina prípustných riešení.
Predpokladajme, že x je prípustné riešenie pre úlohu (LP). To ale znamená, že $x \geq 0$ a $a_i^T x = b_i$ pre $i = 1, \ldots, m$. Pre $X = Diag(x)$ platí $A_i \bullet X = a_i^T x = b_i$ pre $i = 1, \ldots, m$. X $\succeq 0$ lebo X je diagonálna matica s nezápornou diagonálou ($x \geq 0$). Z toho vidíme, že ak x je prípustné riešenie pre úlohu (LP), tak $X = Diag(x)$ je prípustné pre úlohu ($\star$).
Keď X je prípustné riešenie pre úlohu ($\star$), tak každej takejto matici priradíme vektor $x = diag(X)$. Potom $x = diag(X)$ je vektor s nezápornými prvkami a $a_i^T x = A_i \bullet X = b_i$ z čoho vidíme, že x je prípustné riešenie pre úlohu (LP).
(2) Množina optimálnych riešení.
Nech x^* je optimálne riešenie úlohy (LP) a nech X je ľubovoľné prípustné riešenie úlohy ($\star$) a teda $diag(X)$ je prípustné riešenie (LP). Pre $X^* = Diag(x^*)$ dostávame $C \bullet X^* = c^T x^* \leq c^T diag(X) = C \bullet X$, teda X^* je optimálne riešenie ($\star$).
Naopak, ak X^* je optimálne riešenie ($\star$), potom $x^* = diag(X^*)$ je optimálne riešenie (LP), lebo pre každé prípustné riešenie x úlohy (LP) platí $c^T x^* = C \bullet X^* \leq C \bullet Diag(x) = c^T x$.

Poznámka 2.1.2. Uvedené stotožnenie je potrebné z dôvodu väčšej množiny prípustných riešení úlohy ($\star$) ako má úloha (LP). Ohraničenia úlohy ($\star$) spĺňajú aj nediagonálne matice, kde však prvky mimo diagonály nemajú vplyv ani na hodnotu účelovej funkcie ani na splnenie lineárnych ohraničení.

2.2 Úloha konvexného-kvadratického programovania.

Úlohu konvexného-kvadratického programovania môžeme napísať ako

$$\min_{y \in R^k} \{ f_0(y) \mid f_i(y) \leq 0, \ i = 1, \ldots, L \}, \tag{KKP}$$

$$\text{kde } f_i(y) = (Q_i y + r_i)^T (Q_i y + r_i) - p_i y - o_i \text{ pre } i = 0, \ldots, L,$$

$$Q_i \in R^{n_i \times k}, \ r_i \in R^{n_i}, \ p_i \in R^k, \ o_i \in R.$$

Úloha (KKP) je ekvivalentná s úlohou

$$\min_{\tilde{y} \in R^{k+1}} \{ -y_{k+1} \mid f_0(y) + y_{k+1} \leq 0, \ f_i(y) \leq 0, \ i = 1, \ldots, L \} \tag{2.2.1}$$

kde $\tilde{y} = \begin{pmatrix} y \\ y_{k+1} \end{pmatrix} \in R^{k+1}$ a $y \in R^k$, $y_{k+1} \in R$. Úlohu (2.2.1) môžeme prepísať nasledovne

$$\max_{\tilde{y} \in R^{k+1}} \left\{ y_{k+1} \,\middle|\, \begin{array}{c} (p_0^T y + o_0 - y_{k+1}) - (Q_0 y + r_0)^T (Q_0 y + r_0) \geq 0 \\ (p_i^T y + o_i) - (Q_i y + r_i)^T (Q_i y + r_i) \geq 0, \ i = 1, \ldots, L \end{array} \right\}. \tag{2.2.2}$$

Môžeme si všimnúť, že v ohraničeniach vystupujú výrazy typu

$$p - q^T q \geq 0, \tag{2.2.3}$$

kde $p = (p_0^T y + o_0 - y_{k+1})$ a $q = (Q_0 y + r_0)$, respektíve $p = (p_i^T y + o_i)$ a $q = (Q_i y + r_i)$, pre $i = 1, \ldots, L$. Podľa Vety 7.9 vieme, že matica $\begin{pmatrix} p & q^T \\ q & I \end{pmatrix}$ je kladne semidefinitná práve vtedy, keď $p \geq 0$ a $p - q^T I^{-1} q \geq 0$. Z podmienky (2.2.3) vyplýva, že $p \geq q^T q$, kde $q^T q \geq 0$, tak je automaticky splnené, že $p \geq 0$. Preto podmienka (2.2.3) je ekvivalentná s požiadavkou kladnej semidefinitnosti matice $\begin{pmatrix} p & q^T \\ q & I \end{pmatrix}$. Na základe tejto skutočnosti môžeme úlohu (2.2.2) napísať ako

$$\max_{\tilde{y} \in R^{k+1}} \left\{ y_{k+1} \;\middle|\; \begin{array}{c} \begin{pmatrix} p_i^T y + o_i & (Q_i y + r_i)^T \\ Q_i y + r_i & I_{n_i} \end{pmatrix} \succeq 0, \text{ pre } i = 1, \ldots, L \\ \begin{pmatrix} p_0^T y + o_0 - y_{k+1} & (Q_0 y + r_0)^T \\ Q_0 y + r_0 & I_{n_0} \end{pmatrix} \succeq 0 \end{array} \right\}. \tag{2.2.4}$$

Úlohu (2.2.4) možno ľahko previesť na úlohu semidefinitného programovania, a to na úlohu (SDD). Pritom sa využije fakt (Lema 7.4), že ak máme viacero podmienok typu

$$M_0 \succeq 0, \ldots, M_L \succeq 0$$

tak tieto podmienky môžeme previesť na jednu podmienku, ktorá bude vyzerať

$$\begin{pmatrix} M_0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & M_L \end{pmatrix} \succeq 0.$$

V našom prípade

$$M_0 = \begin{pmatrix} p_0^T y + o_0 - y_{k+1} & (Q_0 y + r_0)^T \\ Q_0 y + r_0 & I_{n_0} \end{pmatrix}, \; M_i = \begin{pmatrix} p_i^T + o_i & (Q_i y + r_i)^T \\ Q_i y + r_i & I_{n_i} \end{pmatrix} \text{ pre } i = 1, \ldots, L.$$

Teraz prepíšeme podmienky $M_i \succeq 0, \; i = 0, \ldots, L$ na tvar podmienky z úlohy (SDD)

$$C - \sum_{i=1}^{k+1} A_i y_i \succeq 0.$$

To znamená, určíme konkrétny tvar matíc C^j, A_i^j tak, aby platilo

$$M_j = C^j - \sum_{i=1}^{k+1} A_i^j \tilde{y}_i \quad \text{pre } j = 0, 1, \ldots, L.$$

Označme

$$C^j = \begin{pmatrix} o_j & r_j^T \\ r_j & I_{n_j} \end{pmatrix} \text{ pre } j = 0, \ldots, L,$$

$$A_i^j = -\begin{pmatrix} p_j^i & (Q_j^i)^T \\ Q_j^i & 0_{n_j} \end{pmatrix} \text{ pre } i = 1, \ldots, k, \quad j = 0, \ldots, L,$$

$$A_{k+1}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 0_{n_0} \end{pmatrix},$$

$$A_{k+1}^j = \begin{pmatrix} 0 & 0 \\ 0 & 0_{n_j} \end{pmatrix}, \text{ pre } j = 1, \ldots, L,$$

$$C = \begin{pmatrix} C^0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C^L \end{pmatrix},$$

$$A_i = \begin{pmatrix} A_i^0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_i^L \end{pmatrix} \text{ pre } i = 1, \ldots, k+1,$$

kde p_j^i predstavuje i-tu zložku vektora p_j a Q_j^i predstavuje i-ty stĺpec matice Q_j.

Nakoniec definovaním vektora $b \in R^{k+1}$, kde $b_i = 0$ pre $i = 1, \ldots, k$ a $b_{k+1} = 1$, kde b_i je i-ty prvok vektora b získame úlohu

$$\max_{\tilde{y} \in R^{k+1}} \{b^T \tilde{y} \mid C - \sum_{i=1}^{k+1} \tilde{y}_i A_i \succeq 0\},$$

ktorá je v tvare (SDD) a ekvivalentná s úlohou (2.2.4).

2.3 Úloha tzv. second order cone programming.

Úlohu second order cone programming môžeme napísať ako

$$\max_x \{b^T x \mid x \in K\}, \tag{2.3.1}$$

$$\text{kde } K = \{x = (x_0, x_1) \in R \times R^{m-1} \mid x_0^2 \geq x_1^T x_1, \ x_0 \geq 0\}.$$

Z toho vidíme, že $x \in K$ práve vtedy, keď $x_0^2 \geq x_1^T x_1$ a $x_0 \geq 0$. Z Vety 7.9 vieme, že matica $\begin{pmatrix} x_0 & x_1^T \\ x_1 & x_0 I \end{pmatrix}$ je kladne semidefinitná práve vtedy, keď $x_0 \geq 0$ a $x_o - x_1^T (x_0 I)^{-1} x_1 \geq 0$ a to je vtedy, keď $x_0 \geq 0$ a $x_0^2 - x_1^T x_1 \geq 0$. Z toho vidíme, že

$$x = (x_0, x_1) \in K \text{ práve vtedy, keď } \begin{pmatrix} x_0 & x_1^T \\ x_1 & x_0 I \end{pmatrix} \succ 0.$$

Keď označíme

$$S = \begin{pmatrix} x_0 & x_1^T \\ x_1 & x_0 I \end{pmatrix} = \begin{pmatrix} x_0 & x_1^1 & x_1^2 & \cdots & x_1^{m-1} \\ x_1^1 & x_0 & 0 & \cdots & 0 \\ x_1^2 & 0 & x_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ x_1^{m-1} & 0 & \cdots & 0 & x_0 \end{pmatrix},$$

kde $x_1^j,\ j = 1, \ldots, m-1$ predstavuje j-ty prvok vektora x_1 tak našou úlohou je nájsť matice C a $A_i,\ i = 0, \ldots, m-1$ spĺňajúce

$$S = C - \sum_{i=0}^{m-1} A_i x_i,$$

kde $x_i,\ i = 0, \ldots, m-1$ predstavuje i-tu zložku vektora $x = (x_0, x_1^1, x_1^2, \ldots, x_1^{m-1})^T$. Z toho vidíme, že

$$\begin{aligned}
C &= 0_{m\times m} \\
A_0 &= -I_m \\
A_1 &= -\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
&\vdots \\
A_{m-1} &= -\begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},
\end{aligned}$$

kde matica $A_k,\ k = 1, \ldots, m-1$ má všetky prvky okrem prvkov $(a_k)_{k+1,1} = (a_k)_{1,k+1} = 1$ nulové. Prvok $(a_k)_{i,j}$ označuje prvok matice A_k v i-tom riadku a j-tom stĺpci. Teda pôvodná úloha (2.3.1) bola prevedná na úlohu (SDD) v tvare

$$\max_{x\in R^m} \{b^T x \mid C - \sum_{i=0}^{m} A_i x_i \succeq 0\}.$$

3. Aplikácie úloh semidefinitného programovania

3.1 Optimalizačné úlohy v spektrálnej analýze.

Vlastné hodnoty symetrických matíc majú významnú úlohu v rôznych oblastiach aplikovanej matematiky. Vo väčšine aplikáciach nepotrebujeme nájsť maticu ako takú, ale potrebujeme poznať jej vlastné čísla. V množstve aplikáciach veľkosť maximálneho vlastného čísla umožňuje posúdiť stabilitu systému. Preto sa môže vyskytnúť úloha minimalizovať maximálne vlastné číslo.

Nech $A_0,\ A_1, \ldots,\ A_m \in S^n$ a našou úlohou je nájsť takú lineárnu kombináciu matíc, t.j.

$$A(y) = A_0 + y_1 A_1 + \cdots + y_m A_m = A_0 + \sum_{i=1}^{m} A_i y_i,$$

aby jej najväčšie vlastné číslo bolo čo najmenšie.

Nech $\lambda_i(A(y)),\ i = 1, \ldots, n$ sú vlastné čísla matice A(y) usporiadané podľa veľkosti, t.j.

$$\lambda_1(A(y)) \geq \lambda_2(A(y)) \geq \cdots \geq \lambda_n(A(y)).$$

Túto úlohu si môžeme napísať ako

$$\min_{y \in R^m} \{\lambda_1(A(y)) \mid A(y) = A_0 + y_1 A_1 + \cdots + y_m A_m\}. \tag{3.1.1}$$

Úlohu (3.1.1) budeme riešiť pomocou úlohy

$$\min_{z \in R,\ y \in R^m} \{z \mid zI - A(y) \succeq 0\}. \tag{3.1.2}$$

Ak zavedieme nasledovné substitúcie

$$\tilde{y} = \begin{pmatrix} y \\ z \end{pmatrix} \in R^{m+1}, \quad \text{kde } \tilde{y}_i = y_i \text{ pre } i = 1, \ldots, m \text{ a } \tilde{y}_{m+1} = z,$$

$$\begin{aligned} \tilde{b} &= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in R^{m+1}, \\ \tilde{C} &= -A_0, \\ \tilde{A}_i &= A_i \text{ pre } i = 1, \ldots, m, \\ \tilde{A}_{m+1} &= -I, \end{aligned}$$

tak z úlohy (3.1.2) dostaneme úlohu

$$\max_{\tilde{y} \in R^{m+1}} \{\tilde{b}^T \tilde{y} \mid \tilde{C} - \sum_{i=1}^{m+1} \tilde{A}_i \tilde{y}_i \succeq 0\}. \tag{3.1.3}$$

Z toho vidíme, že úloha (3.1.2) je úlohou (SDP) v duálnom tvare. Teraz sa pokúsime ukázať, že riešením úlohy (3.1.2) dostaneme riešenie pôvodnej úlohy (3.1.1). Najprv ale budeme musieť dokázať nasledovnú lemu.

Lema 3.1.1. *Nech $A \in S^n$, $z \in R$ a $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$ sú vlastné čísla matice A. Potom*

$$zI - A \succeq 0 \text{ práve vtedy, keď } z \geq \lambda_1(A).$$

Dôkaz. Z Lemy 7.5 vieme, že maticu $A \in S^n$ môžeme napísať ako

$$A = Q^T \Lambda Q = Q^T \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} Q \quad \text{kde} \quad Q^T Q = I$$

a λ_i, $i = 1, \ldots, n$ sú vlastné čísla matice A usporiadané podľa veľkosti $\lambda_1 \geq \cdots \geq \lambda_n$. Postupnou úpravou výrazu $zI - A$ dospejeme k požadovanému výsledku.

$$zI - A = zQ^T Q - Q^T \Lambda Q = Q^T (zI - \Lambda) Q \succeq 0.$$

Platí $Q^T(zI - \Lambda)Q \succeq 0$ práve vtedy, keď $z - \lambda_i \geq 0$ pre ľubovoľné $i \in \{1, 2, \ldots, n\}$ ale pretože λ_1 je najväčšie vlastné číslo tak uvedená podmienka je ekvivalentná s podmienkou $z \geq \lambda_1$. $\square$

Veta 3.1.2. *Nech (z^*, y^*) je optimálne riešenie úlohy (3.1.2). Potom*
(a) $\lambda_1(A(y^*)) = z^*$,
(b) y^* *je optimálnym riešením úlohy (3.1.1).*

Dôkaz. (a) Nech (z^*, y^*) je optimálne riešenie úlohy (3.1.2). Keďže je to optimálne riešenie, tak je to zároveň aj prípustné riešenie, a preto

$$z^*I - A(y^*) \succeq 0$$

a teda podľa Lemy 3.1.1

$$z^* \geq \lambda_1(A(y^*)). \tag{3.1.4}$$

Pretože $\lambda_1(A(y^*))$ je maximálne vlastné číslo matice $A(y^*)$ môžeme napísať

$$\lambda_1(A(y^*))I - A(y^*) \succeq 0.$$

Z toho ale vidíme, že dvojica $(\lambda_1(A(y^*)), y^*)$ je prípustné riešenie pre úlohu (3.1.2) a keďže (z^*, y^*) je optimálne riešenie úlohy (3.1.2) (minimum), tak môžeme napísať

$$z^* \leq \lambda_1(A(y^*)). \tag{3.1.5}$$

Z nerovností (3.1.4) a (3.1.5) dostávame

$$z^* = \lambda_1(A(y^*)).$$

(b) Teraz musíme ukázať, že optimálne riešenie úlohy (3.1.2) y^* je optimálne riešenie aj úlohy (3.1.1). Dá sa ľahko nahliadnuť, že každé $y \in R^m$ je prípustné pre úlohu (3.1.1). Zoberme teda ľubovoľné $y \in R^m$. Potom pre $(\lambda_1(A(y)), y)$ podľa Lemy 3.1.1 platí

$$\lambda_1(A_y(A(y))I - A(y) \succeq 0,$$

z čoho vidieť, že $(\lambda_1(A(y)), y)$ je prípustné pre úlohu (3.1.2), a preto môžeme napísať

$$\lambda_1(A(y)) \geq z^* = \lambda_1(A(y^*))$$

čo vlastne znamená, že pre ľubovoľné $y \in R^m$ platí

$$\lambda_1(A(y^*)) \leq \lambda_1(A(y)),$$

a preto je y^* optimálne riešenie úlohy (3.1.1). □

3.2 Problém maximálneho rezu.

Problém maximálne rezu je jeden z NP-ťažkých problémov v teórií grafov. Nech $G = (V, E)$ označuje vážený, neorientovaný graf bez skokov a viacnásobných hrán. Množina $V = \{1 \ldots, n\}$ predstavuje množinu vrcholov grafu. E bude množina hrán, kde hranu z vrchola i do vrchola j označíme ako $ij \in E$ a a_{ij} bude predstavovať váhu príslušnej hrany $ij \in E$. Uvažujme ľubovoľnú podmnožinu vrcholov $S \subseteq V$, kde rez $\delta(S)$ bude predstavovať množinu tých hrán $ij \in E$, ktoré majú jeden vrchol v množine S a druhý v množine $V \setminus S$. problém maximálneho rezu spočíva v nájdení takého rezu $\delta(S)$, ktorý by maximalizoval sumu váh hrán $ij \in \delta(S)$. Problém maximálneho rezu by sme mohli napísať ako

$$\max_{S \subseteq V} \sum_{ij \in \delta(S)} a_{ij}. \tag{3.2.1}$$

V prípade neúplného grafu definujeme $a_{ij} = 0$ ak $ij \notin E$. Prvky matice $A = (a_{ij}) \in S^n$ predstavujú váhy hrán medzi vrcholmi i a j. Pretože uvažujeme neorientovaný graf, tak váha hrany ij je taká istá, ako váha hrany ji a preto matica A je symetrická.
Zadefinujeme vektor $x \in R^n$, ktorý bude predstavovať príslušný rez. Nech $x_i \in \{-1, 1\}$, $i = 1, \ldots, n$ a $x_i = 1$ znamená, že vrchol $i \in S$ a $x_i = -1$ znamená vrchol $i \in V \setminus S$. Uvažujme nasledovný problém

$$MR^* := \max_{x \in \{-1,1\}} \sum_{0<i<j\leq n} a_{ij} \frac{1 - x_i x_j}{2}. \tag{MR}$$

Ak nastane prípad $x_i = x_j$ čo znamená, že vrcholy i a j sú v tej istej množine, tak hodnota výrazu $\frac{1-x_i x_j}{2}$ je 0. V prípade $x_i \neq x_j$, hodnota výrazu $\frac{1-x_i x_j}{2}$ bude 1.
Využijeme symetriu matice A, $x_i x_i = 1$ a účelovú funkciu prepíšeme nasledovne

$$\begin{aligned}
\frac{1}{2} \sum_{0<i<j\leq n} a_{ij}(1 - x_i x_j) &= \frac{1}{4} \sum_{i,j=1}^{n} a_{ij}(1 - x_i x_j) \\
&= \frac{1}{4} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_i x_i - \sum_{j=1}^{n} a_{ij} x_i x_j \right) \\
&= \frac{1}{4} x^T (Diag(Ae) - A)) x,
\end{aligned}$$

kde vektor $e = (1, \ldots, 1)^T$. Matica

$$L(G) = Diag(Ae) - A \tag{3.2.2}$$

sa nazýva Laplaseova matica grafu G. Keďže matica A je symetrická, ľahko sa dá nahliadnuť, že aj matica $L(G)$ je symetrická. Substitúciou $C = \frac{1}{4} L(G) \in S^n$ dostaneme úlohu

$$\max_{x \in \{-1,1\}^n} x^T C x. \tag{3.2.3}$$

Dá sa ľahko vidieť, že

$$x^T C x = \langle Cx, x \rangle = \langle C, x x^T \rangle = C \bullet x x^T. \tag{3.2.4}$$

Pretože $x \in \{-1, 1\}^n$ matica xx^T má hodnosť 1, je kladne semidefinitná a $diag(xx^T) = e$. Ukážeme, že úloha

$$\begin{aligned}
\max\ & C \bullet X \\
& \mathrm{diag}(X) = e \\
& X \succ 0 \\
& \mathrm{rank}(X) = 1
\end{aligned} \tag{3.2.5}$$

je ekvivalentná s úlohou (3.2.3).

Lema 3.2.1. *Úlohy (3.2.3) a (3.2.5) sú ekvivalentné.*

Dôkaz. Nech X je prípustné pre úlohu (3.2.5). Pretože $X \succeq 0$ a $rank(X) = 1$ existuje taký vektor $x \in R^n$, že $X = xx^T$. Podľa podmienky $diag(X) = e$ platí $x_{ii} = 1 = x_i x_i,\ i = 1, \ldots, n$ a preto $x \in \{-1, 1\}^n$. Naopak vidíme, že pre ľubovoľné $x \in \{-1, 1\}^n$ je matica xx^T prípustná pre úlohu (3.2.5). Preto úlohy (3.2.3) a (3.2.5) optimalizujú tí istú funkciu (3.2.4) s rovnakou množinou prípustných riešení. $\square$

V úlohe (3.2.5) zostáva problémom podmienka $rank(X) = 1$, ktorá nie je podmienkou semidefinitného programovania. Vypustením podmienky $rank(X) = 1$ dostaneme úlohu semidefinitného programovania

$$\begin{aligned} MR^*_{SDP} := \max\ & C \bullet X \\ & diag(X) = e \\ & X \succeq 0, \end{aligned} \tag{MR_{SDP}}$$

ktorá je horným ohraničením úlohy (MR) a platí

$$MR^*_{SDP} \geq MR^*.$$

Ak definujeme maticu $A_i,\ i = 1, \ldots, n$ tak, že okrem prvka na diagonále $a_{ii} = 1$ bude mať všetky ostatné prvky nulové a keď $b = (1, \ldots, 1)^T \in R^n$ tak potom z podmienky $A_i \bullet X = b_i,\ i = 1, \ldots, m$ vyplýva, že $diag(X) = e$, lebo $A_i \bullet X = x_{ii} = 1 = b_i,\ i = 1, \ldots, n$ kde x_{ii} je prvok matice X v i-tom riadku a i-tom stĺpci.

Hoci sa nám nepodarilo pôvodnú úlohu (MR) pomocou úlohy (MR_{SDP}) exaktne vypočítať, získali sme určitú aproximáciu problému (MR), ktorý je NP-ťažký.

4. Dualita v SDP

4.1 Vety o dualite.

Definícia 4.1.1. Nech $X \in \mathcal{P},\ (y, Z) \in \mathcal{D}$. Potom rozdiel $C \bullet X - b^T y$ nazývame duálnou medzerou.

Lema 4.1.2. *Nech $X \in \mathcal{P}, (y, Z) \in \mathcal{D}$. Potom pre duálnu medzeru platí*

$$C \bullet X - b^T y = X \bullet Z. \tag{4.1.1}$$

Dôkaz. Postupnými úpravami duálnej medzeri dostaneme požadovaný výsledok

$$\begin{aligned} C \bullet X - b^T y &= C \bullet X - \sum_{i=1}^{m} b_i y_i \\ &= C \bullet X - \sum_{i=1}^{m} (A_i \bullet X) y_i \\ &= X \bullet C - \sum_{i=1}^{m} (X \bullet A_i) y_i \\ &= X \bullet (C - \sum_{i=1}^{m} A_i y_i) \\ &= X \bullet Z \quad \square \end{aligned}$$

Slabá veta o dualite. *Nech $X \in \mathcal{P}$, $(y, Z) \in \mathcal{D}$. Potom platí*

$$C \bullet X \geq b^T y. \tag{4.1.2}$$

Dôkaz. Vzhľadom na skutočnosť, že matice X a Z sú kladne semidefinitné a na základe vlastnosti skalárneho súčinu matíc môžeme napísať

$$C \bullet X - b^T y = X \bullet Z \geq 0 \quad \square$$

Teraz zaveďme nasledujúce označenia

$$p^* = inf\{C \bullet X \mid X \in \mathcal{P}\}, \tag{4.1.3}$$

$$d^* = sup\{b^T y \mid (y, Z) \in \mathcal{D}\}. \tag{4.1.4}$$

Definícia 4.1.3. Budeme hovoriť, že

(a) úloha (SDP) je prípustná ak $\mathcal{P} \neq \emptyset$. Inak je neprípustná. Definitoricky pre neprípustnú úlohu položíme $p^* = \infty$.
(b) úloha (SDD) je prípustná ak $\mathcal{D} \neq \emptyset$. Inak je neprípustná. Definitoricky pre neprípustnú úlohu položíme $d^* = -\infty$.
(c) úlohy (SDP), (SDD) sú neohraničené, ak $p^* = -\infty$ respektíve $d^* = \infty$.

Poznámka 4.1.4. Množiny $\mathcal{P}^*$ a $\mathcal{D}^*$ môžeme zapísať nasledovne

$$\mathcal{P}^* = \{X \in \mathcal{P} \mid C \bullet X = p^*\}, \tag{4.1.5}$$

$$\mathcal{D}^* = \{(y, Z) \in \mathcal{D} \mid b^T y = d^*\}. \tag{4.1.6}$$

Veta o dualite.

(a) *Ak $\mathcal{D}^0 \neq \emptyset$, tak potom $p^* = d^*$ a ak navyše p^* je konečné číslo, tak primárna úloha má optimálne riešenie $X^* \in \mathcal{P}^*$.*
(b) *Ak $\mathcal{P}^0 \neq \emptyset$, tak potom $p^* = d^*$ a ak navyše d^* je konečné číslo, tak duálna úloha má optimálne riešenie $(y^*, Z^*) \in \mathcal{D}^*$.*
(c) *Ak $\mathcal{P}^0 \neq \emptyset$ a súčasne $\mathcal{D}^0 \neq \emptyset$, tak potom p^* a d^* sú konečné čísla a primárna aj duálna úloha majú optimálne riešenia X^* a (y^*, Z^*), pre ktoré platí $C \bullet X^* = p^* = d^* = b^T y^*$.*

Poznámka 4.1.5. Na rozdiel od lineárneho programovania v prípade semidefinitného programovania môže nastať situácia kedy

(a) primárna aj duálna úloha majú optimálne riešenia, ale $p^* \neq d^*$,
(b) $p^* = d^*$ je konečné, ale primárna (resp. duálna) úloha nemá riešenie.

Príklad 4.1.6. V lineárnom programovaní mohol nastal prípad $p^* \neq d^*$ iba vtedy, keď primárna aj duálna úloha boli neprípustné a teda $d^* = -\infty$, $p^* = \infty$. V semidefinitnom programovaní môže nastať prípad $p^* > d^*$ aj v prípade konečných hodnôt p^* a d^*.

Majme úlohu (SDP)

$$\min_{X\in S^3}\left\{\begin{pmatrix}0 & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\bullet X \,\middle|\, \begin{pmatrix}0 & -\frac{1}{2} & 0\\ -\frac{1}{2} & 0 & 0\\ 0 & 0 & 1\end{pmatrix}\bullet X = 1,\ \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\bullet X = 0,\right.$$

$$\left.\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}\bullet X = 0,\ \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\bullet X = 0,\ X\succeq 0\right\}.$$

Označme

$$X=\begin{pmatrix}x_{11} & x_{12} & x_{13}\\ x_{12} & x_{22} & x_{23}\\ x_{13} & x_{23} & x_{33}\end{pmatrix}.$$

Ľahko nahliadneme, že pôvodnú úlohu môžeme napísať ako

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

Z podmienky kladnej semidefinitnosti dostávame

$$\begin{pmatrix}0 & x_{12}\\ x_{12} & x_{22}\end{pmatrix}\succeq 0 \text{ práve vtedy, keď } -x_{12}^2\geq 0 \text{ a z toho vidíme, že } x_{12}=0.$$

Takže množina prípustných riešení bude vyzerať

$$\mathcal{P}=\left\{\begin{pmatrix}0 & 0 & 0\\ 0 & x_{22} & 0\\ 0 & 0 & 1\end{pmatrix}|x_{22}\geq 0\right\}.$$

Z toho vidíme, že množina vnútorných bodov je prázdna, $\mathcal{P}^0=\emptyset$. Pretože $x_{12}=0$ tak $p^*=0$ a $\mathcal{P}=\mathcal{P}^*$.
Teraz vypočítame duálnu úlohu. Duálna úloha je v tvare

$$\max_{y\in R^4,\ Z\in S^3}\left\{y_1\,\middle|\,y_1\begin{pmatrix}0 & -\frac{1}{2} & 0\\ -\frac{1}{2} & 0 & 0\\ 0 & 0 & 1\end{pmatrix}+y_2\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}+\right.$$

$$\left.+y_3\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}+y_4\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}+Z=\begin{pmatrix}0 & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\ Z\succeq 0\right\}$$

a ekvivalentne ju môžeme zapísať ako

$$\max_{y\in R^4}\left\{y_1\,\middle|\,\begin{pmatrix}-y_2 & \frac{y_1+1}{2} & -y_3\\ \frac{y_1+1}{2} & 0 & -y_4\\ -y_3 & -y_4 & -y_1\end{pmatrix}\succeq 0\right\}.$$

Z podmienky kladnej semidefinitnosti postupne dostávame, že $y_1 = -1,\ y_2 \le 0,\ y_4 = 0,\ -y_2 \ge y_3^2$. Množinu prípustných riešení duálnej úlohy môžeme zapísať ako

$$\mathcal{D} = \left\{ \begin{pmatrix} -1 \\ y_2 \\ y_3 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} -y_2 & 0 & -y_3 \\ 0 & 0 & 0 \\ -y_3 & 0 & 1 \end{pmatrix} \succeq 0,\ y_2 \le 0,\ -y_2 \ge y_3^2 \right\}$$

Pretože $z_{22} = 0$, tak vidíme, že množina vnútorných bodov duálnej úlohy je prázdna, $\mathcal{D}^0 = \emptyset$. Pretože maximalizujeme $y_1 = -1$ tak ľahko nahliadneme, že $\mathcal{D}^* = \mathcal{D}$ a $d^* = -1$. Pre tento príklad sme ukázali, že

$$-1 = d^* < p^* = 0,$$

kde množiny vnútorných bodov oboch úloh $\mathcal{P}^0$ a $\mathcal{D}^0$ boli prázdne.

Príklad 4.1.7. V tomto príklade ukážeme, že ak množina vnútorných bodov duálnej úlohy je prázdna $\mathcal{D}^0 = \emptyset$, tak môže nastať situácia, že množina optimálnych riešení primárnej úlohy je prázdna $\mathcal{P}^* = \emptyset$, aj keď $p^* = d^*$ sú konečné čísla.
Majme úlohu (SDP)

$$\min_{X \in S^2} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet X \middle| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X = 2,\ X \succeq 0 \right\}.$$

Označme

$$X = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}.$$

Z podmienky máme

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X = 2x_3 = 2$$

preto $x_3 = 1$. Preto množina prípustných riešení bude vyzerať

$$\mathcal{P} = \left\{ \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \middle| \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \succeq 0 \right\},$$

z čoho dostávam, aby bola splnená podmienka kladnej semidefinitnosti, tak musí platiť $x_1 \ge 0$ a zároveň $x_1x_2 - 1 \ge 0$. Preto môžeme úlohu preformulovať ako

$$\min_{x_1,\ x_2 \in R} \{x_1 \mid x_1x_2 \ge 1,\ x_1 \ge 0\}.$$

Optimálnu hodnotu môžeme vyjadriť ako

$$p^* = \inf_{x_1,\ x_2 \in R} \{x_1 \mid x_1x_2 \ge 1,\ x_1 \ge 0\} = 0.$$

Pretože žiadne riešenie nenadobúda optimálnu hodnotu, tak množina optimálnych riešení je prázdna. Množina vnútorných vyzerá nasledovne

$$\mathcal{P}^0 = \left\{ \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \middle| x_1 > 0,\ x_1x_2 > 1 \right\}.$$

Ukázali sme, že $p^* = 0$, $P^* = \emptyset$ a $\mathcal{P}^0 \neq \emptyset$.
Duálnu úlohu môžeme napísať ako

$$\max_{y \in R,\ Z \in S^2} \left\{ 2y \,\middle|\, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} y + Z = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ Z \succeq 0 \right\},$$

ktorú ale môžeme preformulovať na úlohu

$$\max_{y \in R} \left\{ 2y \,\middle|\, \begin{pmatrix} 1 & -y \\ -y & 0 \end{pmatrix} \succeq 0 \right\}.$$

Matica $\begin{pmatrix} 1 & -y \\ -y & 0 \end{pmatrix}$ je kladne semidefinitná práve vtedy, keď $0 - y^2 \geq 0$. Z toho ale je vidieť, že $y = 0$ a preto množina prípustných riešení duálnej úlohy obshaju práve jeden prvok $\left(0, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right)$, ktorý je aj oprimálnym riešením. Zrejme $D^0 = \emptyset$, $D^* \neq \emptyset$ a $d^* = 0$. Teda v tomto príklade sme mali situáciu. kde $p^* = d^* = 0$ a zároveň $P^* = \emptyset$ a $D^* \neq \emptyset$

V lineárnom programovaní dôkaz silnej vety o dualite spočíval vo využití Farkašovej lemy.

Farkašova Lema. *Nech $G \in R^{k \times n}$, $g \in R^k$. Potom nasledujúce dva výroky sú ekvivalentné*

(a) *neexistuje vektor $x \geq 0$ taký, že $Gx = g$,*
(b) *existuje $u \in R^k$ také, že $G^T u \geq 0$ a zároveň $g^T u < 0$.*

Farkašova lema je založená na vlastnosti, že množina

$$K_{LP} = \{h \in R^k \mid \exists x \geq 0 \ : \ Gx = h\}$$

je uzavretá a konvexná. Bohužial, semidefinitný analóg takejto množiny, t.j.

$$K_{SDP} = \{h \in R^m \mid \exists X \succeq 0 \ : \ G_i \bullet X = h_i,\ i = 1, \ldots, m\}$$

už nemusí byť uzavretou množinou (Príklad 4.1.8). Preto analógia Farkašovej lemy pre semidefinitné matice si vyžaduje dodatočný predpoklad o uzavretosti množiny K_{SDP}. Pretože takýto predpoklad sa ťažko overuje, uvedieme jednu postačujúcu podmienku pre splnenie tohoto predpokladu.

Príklad 4.1.8. Ukážeme, že obraz uzavretého kužela pri lineárnom zobrazení už nemusí byť uzavretý. Nech

$$K_G = \left\{ \begin{pmatrix} r \\ s \end{pmatrix} \in R^2 \,\middle|\, \exists X \succeq 0 : \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \bullet X = r,\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet X = s \right\}.$$

Ľahko sa ukáže, že

$$K_G = \left\{ \begin{pmatrix} r \\ s \end{pmatrix} \in R^2 \,\middle|\, \begin{pmatrix} s & r \\ r & t \end{pmatrix} \succeq 0 \right\}.$$

Matica $\begin{pmatrix} s & r \\ r & t \end{pmatrix}$ je kladne semidefinitná práve vtedy, keď $s \geq 0$ a zároveň $st - r^2 \geq 0$. Ak $s = 0$ tak $r = 0$. V prípade $s > 0$ tak vhodnou voľbou t vieme docieliť, aby $r \in R$ bolo ľubovoľné. Dostávame množinu

$$K_G = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} r \\ s \end{pmatrix} \ s > 0,\ r \in R \right\},$$

ktorá ale neobsahuje všetky svoje hromadné body. Množina hromadných bodov množiny K_G, ktoré nepatria do K_G vyzerá nasledovne $\left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in R^2 \middle| \alpha \in R,\ \alpha \neq 0,\ \beta \in R \right\}$. Preto množina K_G nie je uzavretá.

Lema 4.1.9. *Nech existuje vektor $y^0 \in R^k$ taký, že $\sum_{i=1}^k G_i y_i^0 \succ 0$. Potom množina $K = \{h \in R^k \mid$ existuje $X \succeq 0$ také, že $G_i \bullet X = h_i,\ i = 1, \ldots, k\}$ je uzavretá.*

Ukážeme, že Príklad 4.1.8 nespĺňa podmienku Lemy 4.1.9. Chceme ukázať, že neexistuje $y \in R^2$ pre ktoré by platilo, že matica

$$\sum_{i=1}^{2} G_i y_i = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} y_1 + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y_2 = \begin{pmatrix} y_2 & \frac{y_1}{2} \\ \frac{y_1}{2} & 0 \end{pmatrix}$$

by bola kladne definitná. Zrejme $0 - \frac{1}{4} y_1^2 \leq 0$ pre ľubovoľné y_1 čo je ale spor s kladnou definitnosťou.

Teraz môžeme sformulovať tzv. Zovšeobecnenú Farkašovu lemu pre semidefinitné programovanie.

Zovšeobecnená Farkašova Lema. *Nech matice $G_i \in S^n$ pre $i = 1, \ldots, k$ a vektor $g \in R^k$. Ďalej nech existuje taký vektor $u^0 \in R^k$, pre ktorý platí $\sum_{i=1}^k G_i u_i^0 \succ 0$. Potom nasledujúce dva výroky sú ekvivalentné*

(a) *neexistuje matica $X \succeq 0$ taká, že $G_i \bullet X = g_i$ pre $i = 1, \ldots, k$,*

(b) *existuje vektor $u \in R^k$ taký, že $\sum_{i=1}^k G_i u_i \succeq 0$ a zároveň $g^T u < 0$.*

Dôkaz Vety o dualite. Dokážeme tvrdenie (a) vety. Tvrdenie (b) vety sa dokazuje analogicky.

Zo slabej vety o dualite máme $p^* \geq d^*$. Z predpokladov vety vieme, že množina vnútorných bodov duálnej úlohy je neprázdna ($\mathcal{D}^0 \neq \emptyset$) teda je neprázdna aj množina prípustných riešení duálnej úlohy ($\mathcal{D} \neq \emptyset$) z čoho vidíme, že $d^* > -\infty$. Preto môžeme napísať

$$p^* \geq d^* > -\infty.$$

Keby nastal prípad $d^* = \infty$, tak to by znamenalo $p^* = d^* = \infty$ a veta by bola tým pádom dokázaná.

Predpokladajme preto, že d^* je konečné číslo. Chceme ukázať rovnosť $p^* = d^*$. Sporom. Budeme predpokladať ostrú nerovnosť $p^* > d^*$ a budeme sa snažiť dospieť k sporu.

Predpoklad $p^* > d^*$ znamená, že

$$\text{neexistuje } X \succeq 0 \text{ také, že } C \bullet X = d^* \text{ a zároveň } A_i \bullet X = b_i,\ i = 1, \ldots, m \tag{4.1.7}$$

lebo pre každé $X \succeq 0$ spĺňajúce podmienku $A_i \bullet X = b_i,\ i = 1, \ldots, m$ platí $p^* \leq C \bullet X$ a keďže $d^* < p^*$ tak dostávame $d^* < C \bullet X$.

Spravíme nasledujúce substitúcie

$$G_i = A_i,\ i = 1, \ldots, m, \tag{4.1.8}$$
$$G_{m+1} = C, \tag{4.1.9}$$
$$g_i = b_i,\ i = 1, \ldots, m, \tag{4.1.10}$$
$$g_{m+1} = d^*, \tag{4.1.11}$$
$$k = m + 1.$$

Teraz budeme aplikovať Zovšeobecnenú Farkašovu lemu. V označení (4.1.8)-(4.1.11) budú vzťahy (4.1.7) mať tvar výroku (a) Zovšeobecnenej Farkašovej lemy. Overíme či je splnený predpoklad lemy. Vzhľadom na skutočnosť, že $A_i \in S^n$ pre $i = 1, \ldots, m$ a $C \in S^n$, tak aj $G_i \in S^n$ pre $i = 1, \ldots, k$. Keďže $g = (b_1, \ldots, b_m, d^*)^T$, tak $g \in R^k$. Pretože množina vnútorných bodov $\mathcal{D}^0$ je neprázdna vieme, že

$$\text{existuje vektor } y^0 \in R^m \text{ taký, že } C - \sum_{i=1}^{m} y_i^0 A_i \succ 0. \tag{4.1.12}$$

Z toho vidíme, že substitúciami

$$\begin{aligned} u_i^0 &= -y_i^0, \text{ pre } i = 1, \ldots, m, \\ u_{m+1}^0 &= 1, \end{aligned}$$

získame vektor $u^0 = (u_1^0, \ldots, u_k^0)^T \in R^k$, pre ktorý vďaka (4.1.8),(4.1.9) a (4.1.12) platí

$$\sum_{i=1}^{k} G_i u_i^0 = G_{m+1} u_{m+1}^0 + \sum_{i=1}^{m} G_i u_i^0 = C + \sum_{i=1}^{m} A_i(-y_i^0) \succ 0.$$

Z toho vidíme, že sme splnili predpoklad a použitím Zovšeobecnenej Farkašovej lemy dostávame, že existuje vektor $\begin{pmatrix} u \\ u_0 \end{pmatrix} \in R^k$, kde $u \in R^m$ a $u_0 \in R$ taký, že

$$u_0 C + \sum_{i=1}^{m} A_i u_i \succeq 0, \tag{4.1.13}$$

$$u_0 d^* + u^T b < 0. \tag{4.1.14}$$

(4.1.13) a (4.1.14) napíšeme ako

$$u_0 C - \sum_{i=1}^{m} A_i(-u_i) \succeq 0 \tag{4.1.15}$$

$$u_0 d^* - b^T(-u) < 0. \tag{4.1.16}$$

Teraz preberieme tri prípady $u_0 = 0$, $u_0 > 0$ a $u_0 < 0$, a ukážeme, že ani jeden z týchto prípadov nemôže nastať.
Ak by $u_0 = 0$, tak po dosadení do (4.1.15) a (4.1.16) dostávame

$$-\sum_{i=1}^{m} A_i(-u_i) \succeq 0, \tag{4.1.17}$$

$$-b^T(-u) < 0. \tag{4.1.18}$$

Pretože $\mathcal{D}^0 \neq \emptyset$, tak

$$\text{existuje } y^0 \in R^m \text{ také, že } C - \sum_{i=1}^{m} y_i^0 A_i \succ 0. \tag{4.1.19}$$

Označme

$$y_\alpha = y^0 + \alpha(-u) \text{ pre ľubovoľné } \alpha \geq 0.$$

Využitím (4.1.17) a (4.1.19) dostávame

$$C - \sum_{i=1}^{m}(y_i^0 + \alpha(-u_i))A_i = C - \sum_{i=1}^{m} A_i y_i^0 - \sum_{i=1}^{m} A_i(-u_i) \succ 0$$

a teda y_α je prípustné riešenie pre úlohu (SDD) pre ľubovoľné $\alpha \geq 0$ a pre hodnotu účelovej funkcie platí

$$\begin{aligned} b^T y_\alpha &= b^T(y^0 - \alpha u) \\ &= b^T y^0 - \alpha b^T u. \end{aligned} \tag{4.1.20}$$

Pretože z (4.1.18) máme $b^T u < 0$, tak potom $-b^T u > 0$ a preto z (4.1.20) platí

$$\lim_{\alpha \to \infty} b^T y_\alpha = +\infty$$

čo je spor s predpokladom, že $d^* = sup\{b^T y \mid (y, Z) \in \mathcal{D}\} < +\infty$.
Prípad $u_0 > 0$ by po predelení (4.1.15) a (4.1.16) číslom u_0 znamenal

$$d^* - \left(-\frac{u}{u_0}\right)^T b < 0, \tag{4.1.21}$$

$$C - \sum_{i=1}^{m}\left(-\frac{u_i}{u_0}\right) A_i \succeq 0, \tag{4.1.22}$$

z čoho vidíme, že

$$\left(-\frac{u}{u_0}\right) \in \mathcal{D}$$

a zároveň

$$b^T \left(-\frac{u}{u_0}\right) > d^*,$$

čo je ale spor s definíciou d^*, lebo $d^* = sup\{b^T y \mid (y, Z) \in \mathcal{D}\}$.
Pre $u_0 < 0$, z (4.1.15) a (4.1.16) dostávame

$$-C + \sum_{i=1}^{m}\left(-\frac{u_i}{u_0}\right) A_i \succeq 0, \tag{4.1.23}$$

$$-d^* + b^T\left(-\frac{u}{u_0}\right) < 0. \tag{4.1.24}$$

Ľahko je vidieť, že existuje $\varepsilon_1 > 0$ také, že

$$-d^* + b^T\left(-\frac{u}{u_0}\right) < -\varepsilon_1 < 0. \tag{4.1.25}$$

Zvoľme ε_2 také, že $0 < \varepsilon_2 < \varepsilon_1$. Označme

$$\tilde{y} = -\frac{u}{u_0}. \tag{4.1.26}$$

Dosadením (4.1.26) do (4.1.23) a (4.1.25) dostávame

$$-C + \sum_{i=1}^{m} A\tilde{y}_i \geq 0, \tag{4.1.27}$$

$$-d^* + b^T\tilde{y} < -\varepsilon_1. \tag{4.1.28}$$

Pretože $d^* = sup\{b^T y \mid (y, Z) \in \mathcal{D}\}$, tak z definície suprema vieme, že existuje také $\overline{y}$, že

$$d^* < b^T\overline{y} + \varepsilon_2, \tag{4.1.29}$$

$$C - \sum_{i=1}^{m} A_i\overline{y}_i \succeq 0. \tag{4.1.30}$$

Sčítaním (4.1.28) a (4.1.29) dostávame

$$b^T(\tilde{y} - \overline{y}) < \varepsilon_2 - \varepsilon_1 < 0 \tag{4.1.31}$$

a podobne sčítaním (4.1.27) a (4.1.30) dostaneme

$$\sum_{i=1}^{m} A_i(\tilde{y}_i - y_i) \succeq 0. \tag{4.1.32}$$

Substitúciou $\tilde{u} = \tilde{y} - \overline{y}$ dostávame

$$b^T\tilde{u} < 0, \tag{4.1.33}$$

$$\sum_{i=1}^{m} A_i(\tilde{y}_i - \overline{y}_i) \succeq 0, \tag{4.1.34}$$

čo je ale tá istá situácia ako (4.1.17) a (4.1.18) v prípade $u_0 = 0$. Preto aj v prípade $u_0 < 0$ nastáva spor. Predpoklad $p^* > d^*$ vedie k sporu a keďže $p^* \geq d^*$, tak platí $p^* = d^*$ Ukážeme, že bod (c) je len dôsledkom bodov (a) a (b). Ak množina vnútorných bodov oboch úloh je neprázdna ($\mathcal{D}^0 \neq \emptyset$, $\mathcal{P}^0 \neq \emptyset$), tak

$$-\infty < b^T y^0 \leq d^* = p^* \leq C \bullet X^0 < \infty,$$

$$X^0 \in \mathcal{P}^0,\ y^0 \in \mathcal{D}^0.$$

Takže $p^* = d^*$ sú konečné a teda podľa (a) a (b) primárna aj duálna úloha majú optimálne riešenia.

Veta 4.1.10. *Predpokladajme, že množiny vnútorných bodov úloh (SDP) a (SDD) sú neprázdne, teda $\mathcal{P}^0 \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$. Potom $X^* \in \mathcal{P}$, $(y^*, Z^*) \in D$ sú optimálne riešenia úloh (SDP) a (SDD) práve vtedy, keď*

$$X^* \bullet Z^* = 0.$$

Dôkaz. Nech množiny vnútorných bodov $\mathcal{P}^0$ a $\mathcal{D}^0$ sú neprázdne. Potom z Vety o dualite vieme, že existujú optimálne riešenie X^* a (y^*, Z^*), pre ktoré platí

$$C \bullet X^* = b^T y^*. \tag{4.1.35}$$

Z Lemy 4.1.1 vieme, že pre duálnu medzeru platí

$$C \bullet X - b^T y = X \bullet Z. \tag{4.1.36}$$

Z (4.1.35) a (4.1.36) teda dostávame

$$X^* \bullet Z^* = 0.$$

Naopak, nech $X^* \in \mathcal{P}$, $(y^*, Z^*) \in \mathcal{D}$ a $X^* \bullet Z^* = 0$. Keďže $\mathcal{P}^0 \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$, tak zo silnej vety o dualite vieme, že existujú optimálne riešenia $\tilde{X}$ a $(\tilde{y}, \tilde{Z})$, pre ktoré platí

$$b^T y^* \leq b^T \tilde{y} = C \bullet \tilde{X} \leq C \bullet X^*. \tag{4.1.37}$$

Z Lemy 4.1.1 vieme, že pre duálnu medzeru platí

$$C \bullet X^* - b^T y^* = X^* \bullet Z^*$$

a pretože podľa predpokladu $X^* \bullet Z^* = 0$ tak dostávame

$$C \bullet X^* = b^T y^*. \tag{4.1.38}$$

Z toho vidíme, že využitím (4.1.37) a (4.1.38) dostávame

$$C \bullet \tilde{X} = C \bullet X^*$$
$$b^T \tilde{y} = b^T y^*$$

a preto aj X^* a (y^*, Z^*) sú optimálne riešenia. $\square$

Dôsledok 4.1.11. *Predpokladajme, že množina vnútorných bodov úloh (SDP) a (SDD) sú neprázdne, teda $\mathcal{P}^0 \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$. Potom $X^* \in \mathcal{P}$, $(y^*, Z^*) \in \mathcal{D}$ sú optimálne riešenia úloh (SDP) a (SDD) práve vtedy, keď*

$$X^* Z^* = Z^* X^* = 0.$$

Dôkaz. Z Vety 4.1.10 vieme, že X^* a (y^*, Z^*) sú optimálne riešenia práve vtedy, keď $X^* \bullet Z^* = 0$. Z Lemy 7.11 máme, že $X^* \bullet Z^* = 0$ práve vtedy, keď $X^* Z^* = Z^* X^* = 0$. $\square$

4.2 Komplementarita v SDP.

Zo slabej vety o dualite vieme, že ak $X \in \mathcal{P}$, $(y, Z) \in \mathcal{D}$, tak potom $C \bullet X \geq b^T y$. Ak navyše predpokladáme, že $\mathcal{P}^0 \neq \emptyset$, $\mathcal{D}^0 \neq \emptyset$, tak $X^* \bullet Z^* = 0$. Ale vzhľadom na skutočnosť, že $X^* \succeq 0$, $Z^* \succeq 0$ dostávame podmienku, ktorá sa interpretuje ako podmienka komplementarity

$$X^* \bullet Z^* = 0 \Leftrightarrow X^* Z^* = Z^* X^* = 0_{n \times n}.$$

Pretože $X^* Z^*$ komutujú tak z [9] vieme, že majú spoločný systém vlastných vektorov, teda

$$\exists Q : \; QQ^T = I_n$$

$$X^* = Q \Lambda Q^T, \text{ kde } \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix},$$

$$Z^* = Q \Omega Q^T, \text{ kde } \Omega = \begin{pmatrix} \omega_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \omega_n \end{pmatrix}.$$

Dosadíme a upravujeme

$$0 = X^* Z^* = Q \Lambda Q^T Q \Omega Q^T$$

$$0 = \Lambda \Omega$$

$$\lambda_i \omega_i = 0 \; \forall i = 1, \ldots, n$$

Uvedený vzťah sa nazýva aj komplementarita vlastných čísiel.

Definícia 4.2.1. Ostro komplementárne optimálne riešenia sú také, pre ktoré platí

$$rank(X^*) + rank(Z^*) = n.$$

Poznámka 4.2.2. Zatiaľ čo v lineárnom programovaní z existencie optimálneho riešenia, vyplýva existencia ostrokomplementárneho optimálneho riešenia, v semidefinitnom programovaní to tak nemusí byť. Existujú také úlohy SDP, ktoré (aj pri splnení predpokladov $\mathcal{P}^0 \neq \emptyset$, $\mathcal{D}^0 \neq \emptyset$) nemajú ostrokomplementárne riešenia. Neexistencia tejto vlastnosti komplikuje analýzu algoritmov metód vnútorného bodu.

5. Metódy vnútorného bodu

5.1 Bariérové úlohy.

K Predpokladu 1 pribudne ďalší predpoklad.

Predpoklad 2. Obidve úlohy (SDP) aj (SDD) majú ostro prípustné riešenia, čo znamená že

$$\text{existuje } X^0 \in S^n \text{ také, že } A_i \bullet X^0 = b_i, \text{ pre } i = 1, ..., m, \quad X^0 \succ 0,$$

$$\text{existuje } (y^0, Z^0) \in R^m \times S^n \text{ také, že } \sum_{i=1}^m A_i y_i^0 + Z^0 = C, \quad Z^0 \succ 0.$$

Inými slovami $\mathcal{P}^0 \neq \emptyset$, $\mathcal{D}^0 \neq \emptyset$.

Poznámka 5.1.1. Predpoklad 2 umožňuje použiť metódy vnútorného bodu.

Za týchto predpokladov nutné a postačujúce podmienky optimality pre dvojicu úloh (SDP) a (SDD) sú

$$A_i \bullet X = b_i,\ i = 1, \dots, m,\ X \succeq 0 \tag{5.1.1}$$

$$\sum_{i=1}^{m} A_i y_i + Z = C,\ Z \succeq 0 \tag{5.1.2}$$

$$XZ = 0_{n\times n} \tag{5.1.3}$$

kde podmienka (5.1.1) predstavuje primárnu prípustnosť, (5.1.2) duálnu prípustnosť a (5.1.3) podmienku komplementarity.

Pre úlohu (SDP)

$$\min_{X\in S^n} \{C \bullet X \mid A_i \bullet X = b_i, i = 1, \dots, m,\ X \succeq 0\}, \tag{SDP}$$

definujeme transformačnú bariérovú funkciu

$$\phi_\mu(X) = C \bullet X - \mu \ln(det(X)),\ X \succ 0 \tag{5.1.4}$$

a pre pevné $\mu > 0$ vytvoríme bariérovú úlohu

$$\min_{X\in S^n_{++}} \{\phi_\mu(X) |\ A_i \bullet X = b_i,\ i = 1, \dots, m\}. \tag{SDPμ}$$

Poznámka 5.1.2. V porovnaní s úlohou (SDP) je problém (SDPμ) omnoho jednoduchší, lebo sa podarilo zbaviť nepríjemného ohraničenia $X \succeq 0$. Navyše účelová funkcia je rýdzo konvexná. Analogické zjednodušenie získame aj úlohou (SDDμ).

Analogicky pre úlohu (SDD)

$$\max_{y\in R^m, Z\in S^n} \{b^T y \mid \sum_{i=1}^{m} y_i A_i + Z = C,\ Z \succeq 0\}, \tag{SDD}$$

definujeme transformačnú bariérovú funkciu

$$\Psi_\mu(y, Z) = b^T y + \mu \ln(det(Z)),\ Z \succ 0 \tag{5.1.5}$$

a tiež pre pevné $\mu > 0$ vytvoríme bariérovú úlohu

$$\max_{y\in R^m, Z\in S^n_{++}} \{\Psi_\mu(y, Z) \mid \sum_{i=1}^{m} y_i A_i + Z = C\}. \tag{SDDμ}$$

Lema 5.1.3. *Funkcia* $\ln(det(V)) : S^n_{++} \to R$ *je rýdzokonkávna.*

Dôkaz. Uvažujme ľubovoľné $V, W \in S^n_{++}$ ale navzájom rôzne $V \neq W$. Chceme ukázať, že pre ľubovoľné $\alpha \in (0, 1)$ platí

$$\ln(det(\alpha V + (1-\alpha)W)) > \alpha \ln(det(V)) + (1-\alpha)\ln(det(W)). \tag{5.1.6}$$

Z [9] vieme, že

$$W = P\Lambda P^T, \tag{5.1.7}$$
$$V = PP^T. \tag{5.1.8}$$

$$\begin{aligned}
\ln(det(\alpha V + (1-\alpha)W)) &= \ln(det(\alpha PP^T + (1-\alpha)P\Lambda P^T)) \\
&= \ln(det(P(\alpha I + (1-\alpha)\Lambda)P^T)) \\
\text{pretože } det(AB) = det(A)det(B) \text{ tak} & \\
&= ln(det(P).det(\alpha I + (1-\alpha)\Lambda)det(P^T)) \\
&= \ln(det(PP^T)) + \ln(det(\alpha I + (1-\alpha)\Lambda)) \\
\text{matica } \alpha I + (1-\alpha)\Lambda \text{ je diagonálna, preto} & \\
&= \ln(det(V)) + ln\left(\prod_{i=1}^{n}(\alpha + (1-\alpha)\lambda_i)\right) \\
&= \ln(det(V)) + \sum_{i=1}^{n}\ln(\alpha + (1-\alpha)\lambda_i) \\
\text{funkcia ln je konkávna, takže} & \\
&> \ln(det(V)) + \sum_{i=1}^{n}(\alpha\ln(1) + (1-\alpha)\ln(\lambda_i)) \\
&= \ln(det(V)) + (1-\alpha)\sum_{i=1}^{n}\ln(\lambda_i) \\
&= \ln(det(V)) + (1-\alpha)ln\left(\prod_{i=1}^{n}\lambda_i\right) \\
&= \ln(det(V)) + (1-\alpha)\ln(det(\Lambda)) \\
\text{pričítame nulu v tvare } & + \alpha\ln(det(V)) - \alpha\ln(det(V)) \text{ a dostaneme} \\
&= \alpha\ln(det(V)) + (1-\alpha)(\ln(det(V)) + \ln(det(\Lambda))) \\
&= \alpha\ln(det(V)) + (1-\alpha)(\ln(det(PP^T)) + \ln(det(\Lambda))) \\
&= \alpha\ln(det(V)) + (1-\alpha)\ln(det(P\Lambda P^T)) \\
&= \alpha\ln(det(V)) + (1-\alpha)\ln(det(W)) \quad \square
\end{aligned}$$

Veta 5.1.4.

(a) *Funkcia* $\phi_\mu(X) = C \bullet X - \mu ln(det(X))$ *je rýdzokonvexná na* $\mathcal{P}^0$.

(b) *Funkcia* $\psi_\mu(y, Z) = b^T y + \mu ln(det(Z))$ *je rýdzokonkávna na* $\mathcal{D}^0$.

Dôkaz. (a) Ľahko sa dá nahliadnuť, že funkcia $C \bullet X$ je lineárna a funkcia $-\mu ln(det(X))$ je rýdzokonvexná na množine kladne definitných matíc. Z toho vidíme, že funkcia $\phi_\mu(X)$ je rýdzokonvexná na množine kladne definitných matíc a tým pádom je rýdzokonvexná aj na množime $\mathcal{P}^0 \subset S^n_{++}$.
(b) Vidíme, že funkcia $b^T y$ je lineárna na R^m a funkcia $\mu ln(det(Z))$ je rýdzokonkávna

na množine kladne definitných matíc. Priamo z toho ale nevyplýva rýdzokonkávnosť funkcie $\psi_\mu(y, Z)$ na množine $R^m \times S^n_{++}$. Môže sa stať, že síce $(y_1, Z_1) \neq (y_2, Z_2)$ ale $Z_1 = Z_2$ a teda nemáme rýdzokonkávnosť, ale len konkávmosť. Na základe Predpokladu 1, ale vieme, že medzi y a Z existuje jednoznačná korešpondencia, a preto nemôže nastať prípad, kde $(y_1, Z_1) \neq (y_2, Z_2)$ a zároveň $Z_1 = Z_2$ alebo $y_1 = y_2$ a teda funkcia $\psi_\mu(y, Z)$ je na množine $\mathcal{D}^0$ rýdzokonkávna. $\square$

Dôsledok 5.1.5. *Úlohy* (SDP_μ) *a* $(SDDD_\mu)$ *majú najviac jedno riešenie.*

Dôkaz. V prípade úlohy (SDP_μ) hľadáme minimum rýdzokonvexnej funkcie, a preto ak existuje, tak je najviac jedno. Podobne je to pri úlohe (SDD_μ), kde hľadáme maximum rýdzokonkávnej funkcie na $\mathcal{D}^0$. $\square$

Vidíme, že sa nám podarilo ukázať, že úlohy (SDP_μ) a (SDD_μ) majú najviac jedno riešenie. Teraz ukážeme existenciu. Budeme k tomu potrebovať nasledovnú lemu.

Lema 5.1.6. *Nech K je konvexná, neprázdna a neohraničená podmnožina množiny S^n_{++}. Potom pre ľubovoľné $V \in K$ existuje $W \in K,\ V \neq W$ také, že*

$$p = \{V_t \mid V_t = V + t(W - V),\ t \geq 0\} \subset K$$

$$\textit{a navyše}\ \ W - V \succ 0.$$

Poznámka 5.1.7. Táto Lema je dôsledkom dobre známej vety z konvexnej analýzy o existencii neohraničeného lúča pre neohraničenú konvexnú množinu.

Veta 5.1.8. *Pre každé pevné $\mu > 0$ má úloha (SDP_μ) jediné optimálne riešenie.*

Dôkaz. Zvoľme pevné $\mu > 0$. Z Predpokladu 2 vieme, že množina vnútorných bodov je neprázdna, teda $\mathcal{P}^0 \neq \emptyset$. Preto nech $X^0 \in \mathcal{P}^0$. Označme

$$\mathcal{P}(X^0) = \{X \in \mathcal{P}^0 \mid \phi_\mu(X) \leq \phi_\mu(X^0)\}. \tag{5.1.9}$$

Ľahko sa dá nahliadnuť, že množina $\mathcal{P}(X^0)$ je neprázdna, lebo $X^0 \in \mathcal{P}(X^0)$. Priamo z definície množiny $\mathcal{P}(X^0)$ vidíme, že pre každé $X \in \mathcal{P}^0 - \mathcal{P}(X^0)$ platí $\phi_\mu(X) > \phi_\mu(X^0)$ a teda

$$\min_{X \in S^n} \{\phi_\mu(X) \mid X \in \mathcal{P}^0\} = \min_{X \in S^n} \{\phi_\mu(X) \mid X \in \mathcal{P}(X^0)\}. \tag{5.1.10}$$

Stačí teda ukázať, že úloha

$$\min_{X \in S^n} \{\phi_\mu(X) \mid X \in \mathcal{P}(X^0)\} \tag{5.1.11}$$

má jediné optimálne riešenie. Ukážeme, že množina $\mathcal{P}(X^0)$ je uzavretá a ohraničená a teda kompaktná.

(a) Uzavretosť množiny $\mathcal{P}(X^0)$. Majme ľubovoľnú postupnosť $\{X^n\}_{n=1}^\infty \subset \mathcal{P}(X^0)$ a nech $\lim_{n\to\infty} X^n = X$. Ukážeme, že $X \in \mathcal{P}(X^0)$. Zo spojitosti funkcie ϕ_μ vieme, že

$$\lim_{n\to\infty} \phi_\mu(X^n) = \phi_\mu(X). \tag{5.1.12}$$

Z definície množiny $\mathcal{P}(X^0)$ (5.1.9) vieme, že pre ľubovoľné n je $\phi_\mu(X^n) \leq \phi_\mu(X^0)$ a preto

$$\lim_{n\to\infty} \phi_\mu(X^n) \leq \phi_\mu(X^0), \tag{5.1.13}$$

$$\phi_\mu(X) \leq \phi_\mu(X^0). \tag{5.1.14}$$

Z toho vidíme, že $X \in \mathcal{P}(X^0)$ a keďže množina $\mathcal{P}(X^0)$ obsahuje všetky svoje hromadné body, tak je uzavretá.
(b) Ohraničenosť množiny $\mathcal{P}(X^0)$. Sporom. Pretože množina $\mathcal{P}(X^0)$ je úrovňovou množinou konvexnej funkcie ϕ_μ, tak množina $\mathcal{P}(X^0)$ je konvexná. Teraz budeme predpokladať neohraničenosť množiny $\mathcal{P}(X^0)$ a dospejeme k sporu. Ak $\mathcal{P}(X^0)$ je neohraničená množina, tak na základe Lemy 5.1.6 vieme, že existujú matice $X_1, X_2,\ X_1 \neq X_2$ a $X_2 - X_1 \succ 0$ také, že

$$P = \{X_t \mid X_t = X_1 + t(X_2 - X_1),\ t \geq 0\} \subset \mathcal{P}(X^0). \tag{5.1.15}$$

Teda pre ľubovolné $t \geq 0$ platí

$$\phi_\mu(X_t) = \phi_\mu(X_1 + t(X_2 - X_1)) \leq \phi_\mu(X_0). \tag{5.1.16}$$

Počítajme

$$\begin{aligned}
\lim_{t\to\infty} \phi_\mu(X_t) &= \lim_{t\to\infty} \phi_\mu(X_1 + t(X_2 - X_1)) \\
&= \lim_{t\to\infty} (C \bullet X_1 + tC \bullet (X_2 - X_1) - \mu ln(det(X_1 + t(X_2 - X_1))) \\
&= \lim_{t\to\infty} \left(C \bullet X_1 + t\left(C \bullet (X_2 - X_1) - \frac{\mu \ln(det(X_1 + t(X_2 - X_1)))}{t}\right)\right).
\end{aligned}$$

Výraz $p_k(t) = det(X_1 + t(X_2 - X_1))$ predstavuje polynóm stupňa $k > 0$ v premennej t. Deriváciou polynómu $p_k(t)$ získame polynóm $q_{k-1}(t) = p_k'(t)$ stupňa $k-1$. Preto

$$\lim_{t\to\infty} \mu \frac{\ln(det(X_1 + t(X_2 - X_1)))}{t} = \mu \lim_{t\to\infty} \frac{\ln(p_k(t))}{t} =$$

použitím L'Hospitalovho pravidla dostávame

$$= \mu \lim_{t\to\infty} \frac{d}{dt}(\ln(p_k(t))) = \mu \lim_{t\to\infty} \frac{q_{k-1}(t)}{p_k(t)} = 0.$$

Teda ukázali sme

$$\lim_{t\to\infty} \mu \frac{\ln(det(X_1 + t(X_2 - X_1)))}{t} = 0. \tag{5.1.17}$$

Pretože $\mathcal{D}^0 \neq \emptyset$, tak existuje $(y^0, Z^0) \in \mathcal{D}^0$. Vieme, že $Z^0 \succ 0$ a $C = \sum_{i=1}^m A_i y_i^0 + Z^0$. Počítajme

$$C \bullet (X_2 - X_1) = (\sum_{i=1}^m A_i y_i^0 + Z^0) \bullet (X_2 - X_1) = (\sum_{i=1}^m A_i y_i^0) \bullet (X_2 - X_1) =$$

$$= \sum_{i=1}^m y_i^0 (A_i \bullet X_2) - \sum_{i=1}^m y_i^0 (A_i \bullet X_1) + Z^0 \bullet (X_2 - X_1) = Z^0 \bullet (X_2 - X_1) > 0$$

pretože $Z^0 \succ 0$ a zároveň $X_2 - X_1 \succ 0$. Preto platí

$$C \bullet (X_2 - X_1) > 0. \tag{5.1.18}$$

Z (5.1.17) a (5.1.18) dostávame, že existuje $\varepsilon > 0$ a t_0 také, že pre ľubovoľné $t > t_0$ platí

$$C \bullet (X_2 - X_1) - \frac{\mu \ln(det(X_1 + t(X_2 - X_1)))}{t} > \varepsilon_0. \tag{5.1.19}$$

Z toho vidíme, že

$$\lim_{t\to\infty} \phi_\mu(X_t) = \infty. \tag{5.1.20}$$

To je ale spor s predpokladom $\phi_\mu(X_t) \leq \phi_\mu(X_0)$ pre ľubovoľné $t \geq 0$. Preto množina $\mathcal{P}(X^0)$ je ohraničená.
Ukázali sme kompaktnosť množiny $\mathcal{P}(X^0)$. Na základe Weierstrassovej vety má spojitá funkcia $\phi_\mu(X)$ na kompakte minimum a keďže funkcia $\phi_\mu(X)$ je navyše rýdzokonvexná, tak to minimum bude jediné.

Veta 5.1.9. *Pre každé pevné $\mu > 0$ má úloha (D_μ) jediné optimálne riešenie.*

Dôkaz. Dôkaz je variáciou predchádzajúceho dôkazu. Z Predpokladu 2 vieme, že množina $\mathcal{D}^0$ je neprázdna, preto existuje $(y^0, Z^0) \in \mathcal{D}^0$. Označme

$$\mathcal{D}(y^0, Z^0) = \{(y, Z) \in \mathcal{D}^0 \mid \psi_\mu(y, Z) \geq \psi_\mu(y^0, Z^0)\}. \tag{5.1.21}$$

Množina $\mathcal{D}(y^0, Z^0)$ je neprázdna, lebo $(y^0, Z^0) \in \mathcal{D}(y^0, Z^0)$. Vidíme, že pre ľubovoľné $(y, Z) \in \mathcal{D}^0 \setminus \mathcal{D}(y^0, Z^0)$ platí $\phi_\mu(y, Z) < \phi_\mu(y^0, Z^0)$. Stačí teda ukázať, že úloha

$$\max_{y\in R^m,\ Z\in S^n} \{\psi_\mu(y, Z) \mid (y, Z) \in \mathcal{D}(y^0, Z^0) \tag{5.1.22}$$

má jediné optimálne riešenie. Opäť ukážeme, že množina $\mathcal{D}(y^0, Z^0)$ je uzavretá a ohraničená a tým pádom kompaktná.
(a) Uzavretosť množiny $\mathcal{D}(y^0, Z^0)$. Uvažujme ľubovoľnú postupnosť $\{(y^n, Z^n)\}_{n=1}^\infty \subset \mathcal{D}(y^0, Z^0)$ a nech $\lim\limits_{n\to\infty}(y^n, Z^n) = (y, Z)$. Ukážeme, že $(y, Z) \in \mathcal{D}(y^0, Z^0)$. Zo spojitosti funkcie ψ_μ vieme, že

$$\lim_{n\to\infty} \psi_\mu(y^n, Z^n) = \psi_\mu((y, Z)). \tag{5.1.23}$$

Z definície množiny $\mathcal{D}(y^0, Z^0)$ vieme, že $\psi_\mu(y^n, Z^n) \geq \psi_\mu(y^0, Z^0)$ a preto

$$\lim_{n\to\infty} \psi_\mu(y^n, Z^n) \geq \psi_\mu(y^0, Z^0), \tag{5.1.24}$$

$$\psi_\mu(y, Z) \geq \psi_\mu(y^0, Z^0). \tag{5.1.25}$$

Z toho vidíme, že $(y, Z) \in \mathcal{D}(y^0, Z^0)$ a pretože množina $\mathcal{D}(y^0, Z^0)$ obsahuje všetky svoje hromadné body, tak je uzavretá.
(b) Ohraničenosť množiny $\mathcal{D}(y^0, Z^0)$. Sporom. Predpokladajme neohraničenosť množiny $\mathcal{D}(y^0, Z^0)$. Pretože množina $\mathcal{D}(y^0, Z^0)$ je konvexná a neohraničená tak na základe Lemy 5.1.6 vieme, že existuje polpriamka

$$p = \{(y_t, Z_t) \mid (y_t, Z_t) = (y_1, Z_1) + t[(y_2, Z_2) - (y_1, Z_1)],\ t \geq 0\} \subset \mathcal{D}(y^0, Z^0)$$

kde $Z_2 - Z_1 \succ 0$. Vypočítame limitu hodnôt funkcie ψ_μ na polpriamke

$$\begin{aligned}\lim_{t\to\infty} \psi_\mu(y_t, Z_t) &= \lim_{t\to\infty} \psi_\mu((y_1, Z_1) + t[(y_2, Z_2) - (y_1, Z_1)]) \\ &= \lim_{t\to\infty} \left(b^T y_1 + t\left(b^T(y_2 - y_1) - \mu\frac{ln(det(Z_1 + t(Z_2 - Z_1)))}{t}\right)\right).\end{aligned}$$

Podobne ako v dôkaze Vety 5.1.8 sa ukáže, že

$$\lim_{t\to\infty} \frac{\ln\det(Z_1 + t(Z_2 - Z_1))}{t} = 0. \tag{5.1.26}$$

Pretože $\mathcal{P}^0 \neq \emptyset$, tak existuje $X^0 \in \mathcal{P}^0$. Vieme, že $X^0 \succ 0$ a $A_i \bullet X = b_i,\ i = 1, \ldots, m$ preto

$$b^T(y_2 - y_1) = \sum_{i=1}^{m} b_i(y_{2_i} - y_{1_i}) = \sum_{i=1}^{m}(A_i \bullet X^0)(y_{2_i} - y_{1_i}) =$$

$$= \sum_{i=1}^{m} A_i(y_{2_i} - y_{1_i}) \bullet X^0 = -(Z_2 - Z_1) \bullet X^0 < 0.$$

Takže analogicky ako v dôkaze Vety 5.1.8 platí

$$\lim_{t\to\infty} \psi_\mu((y_t, Z_t)) = -\infty, \tag{5.1.27}$$

čo je ale spor s predpokladom $\psi_\mu(y_t, Z_t) \geq \psi_\mu(y^0, Z^0)$ a preto množina $\mathcal{D}(y^0, Z^0)$ je ohraničená.
Ukázali sme kompaktnosť množiny $\mathcal{D}(y^0, Z^0)$ a spojitá rýdzokonkávna funkcia $\psi_\mu(y, Z)$ tu má svoje jediné maximum. □

Dôsledok 5.1.10. *Pre každé pevné $\mu > 0$ majú úlohy (SDP_μ) a (SDD_μ) jediné riešenie $X(\mu),\ y(\mu),\ Z(\mu)$.*

5.2 Centrálna trajektória.

Definícia 5.2.1. Množina bodov

$$\{(X(\mu), y(\mu), Z(\mu));\ \mu > 0\}$$

sa nazýva centrálna trajektória.

Lema 5.2.2. *Nech $f(X) = ln(det(X))$, pre $X \in S^n_{++}$. Potom*

$$\nabla_X f(X) = X^{-1}. \tag{5.2.1}$$

Dôkaz. Urobíme Laplaceov rozvoj $det(X)$ podľa j-teho stĺpca

$$det(X) = \sum_{k=1}^{n}(-1)^{k+j} x_{kj} det(X_{kj}), \tag{5.2.2}$$

kde x_{kj} je prvok matice X nachádzajúci sa v k-tom riadku a j-tom stĺpci a matica $X_{kj} \in R^{(n-1)\times(n-1)}$ vznikne z matice X vynechaním k-teho riadku a j-teho stĺpca. Derivovaním (5.2.2) podľa prvku $x_{kj},\ k, j = 1, \ldots, n$ dostaneme

$$\frac{\partial ln(det(X))}{\partial x_{kj}} = \frac{1}{det(X)}(-1)^{k+j} det(X_{kj}). \tag{5.2.3}$$

Prvok matice X^{-1} v k-tom riadku a j-tom stĺpci môžeme vyčísliť ako

$$(X^{-1})_{kj} = \frac{1}{det(X)}(-1)^{k+j} det(X_{kj}) \tag{5.2.4}$$

a preto

$$\nabla_X ln(det(X)) = X^{-1}. \quad \square \tag{5.2.5}$$

Veta 5.2.3. *Nech platí Predpoklad 1 a Predpoklad 2. Potom* $X \in \mathcal{P}^0$, $(y, Z) \in \mathcal{D}^0$ *je optimálne riešenie* (SDP_μ), (SDD_μ) *práve vtedy, keď*

$$A_i \bullet X = b_i,\ i = 1, 2, \ldots, m, \quad X \succ 0, \tag{1}$$

$$\sum_{i=1}^{m} A_i y_i + Z = C, \quad Z \succ 0, \tag{2}$$

$$XZ = \mu I. \tag{3}$$

Dôkaz. Úloha (SDPμ) je konvexná úloha s Lagrangeovou funkciou

$$L_{P_\mu}(X, y) = C \bullet X - \mu \ln(det(X)) - \sum_{i=1}^{m} y_i(A_i \bullet X - b_i).$$

Preto $X \in S^n_{++}$ je optimálnym riešením tejto úlohy vtedy a len vtedy ak existuje také y, že v bode (X,y) je $\nabla_X L_{P_\mu} = 0$ a zároveň $\nabla_y L_{P_\mu} = 0$.

Pomocou Lemy 5.2.2 dostávame

$$\nabla_X L_{P_\mu} = C - \mu X^{-1} - \sum_{i=1}^{m} y_i A_i = 0, \quad X \succ 0, \tag{$\tilde{1}$}$$

$$\nabla_{y_i} L_{P_\mu} = A_i \bullet X - b_i = 0 \text{ pre } i = 1, \ldots, m. \tag{$\tilde{2}$}$$

$$\text{Označme} \quad Z = \mu X^{-1}. \tag{$\tilde{3}$}$$

Zrejme $Z \in S^n_{++}$ a podmienky $(\tilde{1}) - (\tilde{3})$ možno prepísať do tvaru (1)-(3).

Podobne pre úlohu $(SDD\mu)$ je

$$L_{D_\mu}(X, y, Z) = b^T y + \mu \ln(det(Z)) - X \bullet (\sum_{i=1}^{m} y_i A_i + Z - C)$$

a (y, Z) je optimálnym riešením vtedy a len vtedy, ak existuje také X, že platí

$$\nabla_X L_{D_\mu} = \sum_{i=1}^{m} y_i A_i + Z - C = 0, \quad Z \succ 0, \tag{$\tilde{4}$}$$

$$\nabla_{y_i} L_{D_\mu} = b_i - X \bullet A_i = 0 \text{ pre } i = 1, \ldots, m, \tag{$\tilde{5}$}$$

$$\nabla_Z L_{D_\mu} = \mu Z^{-1} - X = 0. \tag{$\tilde{6}$}$$

Z podmienky $(\tilde{6})$ je vidieť, že $X \in S^n_{++}$ a podmienky $(\tilde{4})$-$(\tilde{6})$ možno prepísať do tvaru (1)-(3).

Z toho vyplýva, že úlohy (SDPμ) a (SDDμ) majú spoločný systém nutných a postačujúcich podmienok. Z jednoznačnosti optimálnych riešení úloh (SDPμ) a (SDDμ) vyplýva aj jednoznačnosť (pre pevne zvolené μ) riešenia systému (1)-(3). X-ová časť je optimálnym riešením úlohy (SDPμ) a zároveň Lagrangeovým multiplikátorom pre úlohu (SDDμ).Časť (y, Z) je zase optimálnym riešením úlohy (SDDμ) a zároveň aj Lagrangeovým multiplikátorom pre úlohu (SDPμ).

$$A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \quad X \succ 0, \tag{1}$$

$$\sum_{i=1}^{m} A_i y_i + Z = C, \quad Z \succ 0, \tag{2}$$

$$XZ = \mu I. \tag{3}$$

kde (1) predstavuje primárnu prípustnosť, (2) duálnu prípustnosť a (3) perturbovanú podmienku komplementarity. □

Máme teda charakterizáciu bodov na centrálnej trajektórii. Bod $(X(\mu), y(\mu), Z(\mu))$ je bodom centrálnej trajektórie práve vtedy, keď je riešením systému (1-3). Inými slovami vnútorný bod úloh $(SDP\mu)$ a $(SDD\mu)$) je bodom centrálnej trajektórie práve vtedy, keď spĺňa perturbovanú podmienku komplementarity (3). Teraz môžeme formulovať nasledovné dôsledky.

Dôsledok 5.2.4. *Bod $(X(\mu), y(\mu), Z(\mu))$ pre $\mu > 0$ je bodom centrálnej trajektórie práve vtedy, keď je riešením systému (1-3).*

Dôsledok 5.2.5. *Ak bod $(X(\mu), y(\mu), Z(\mu))$ pre $\mu > 0$ je bodom centrálnej trajektórie, tak pre duálnu medzeru (4.1.1) platí*

$$C \bullet X(\mu) - b^T y(\mu) = n\mu.$$

Dôkaz. Podľa Lemy 4.1.2 dostávame

$$C \bullet X(\mu) - b^T y(\mu) = X \bullet Z = stopa(XZ) = stopa(\mu I_n) = n\mu. \quad \square$$

Lema 5.2.6. *Nech* $0 < \mu_1 < \mu_2$. *Potom*

$$C \bullet X(\mu_1) \leq C \bullet X(\mu_2).$$

Dôkaz. $X(\mu_1)$ je optimálne riešenie pre úlohu $SDP(\mu_1)$ a $X(\mu_2)$ pre $SDP(\mu_2)$ a platí

$$\phi_{\mu_2}(X(\mu_2)) = C \bullet X(\mu_2) - \mu_2 ln(det(X(\mu_2))) \tag{5.2.6}$$

$$\phi_{\mu_2}(X(\mu_1)) = C \bullet X(\mu_1) - \mu_2 ln(det(X(\mu_1))). \tag{5.2.7}$$

Keďže

$$\phi_{\mu_2}(X(\mu_2)) \leq \phi_{\mu_2}(X(\mu_1)) \tag{5.2.8}$$

tak môžeme napísať

$$C \bullet X(\mu_2) - \mu_2 ln(det(X(\mu_2))) \leq C \bullet X(\mu_1) - \mu_2 ln(det(X(\mu_1))) \tag{5.2.9}$$

a po úprave dostaneme

$$\mu_2(ln(det(X(\mu_1))) - ln(det(X(\mu_2)))) \leq C \bullet X(\mu_1) - C \bullet X(\mu_2). \tag{5.2.10}$$

Analogicky z nerovnosti

$$\phi_{\mu_1}(X(\mu_1)) \leq \phi_{\mu_1}(X(\mu_2)) \tag{5.2.11}$$

dostaneme

$$\mu_1(ln(det(X(\mu_1))) - ln(det(X(\mu_2)))) \geq C \bullet X(\mu_1) - C \bullet X(\mu_2). \tag{5.2.12}$$

Z nerovníc (5.1.10) a (5.2.12) dostávame

$$\frac{C \bullet X(\mu_1) - C \bullet X(\mu_2)}{\mu_1} \leq \frac{C \bullet X(\mu_1) - C \bullet X(\mu_2)}{\mu_2}. \tag{5.2.13}$$

Nakoniec postupnými úpravami dospejeme k požadovanému výsledku

$$C \bullet X(\mu_1) - C \bullet X(\mu_2) \leq \frac{\mu_1}{\mu_2} C \bullet X(\mu_1) - \frac{\mu_1}{\mu_2} C \bullet X(\mu_2),$$

$$C \bullet X(\mu_1)(1 - \frac{\mu_1}{\mu_2}) \leq C \bullet X(\mu_2)(1 - \frac{\mu_1}{\mu_2}),$$

keďže $0 < \mu_1 < \mu_2$ tak $1 - \frac{\mu_1}{\mu_2} > 0$ a preto

$$C \bullet X(\mu_1) \leq C \bullet X(\mu_2) \quad \square$$

Analogická monotónnosť platí aj pre účelovú funkciu duálnej úlohy.

Lema 5.2.7. *Nech* $0 < \mu_1 < \mu_2$. *Potom*

$$b^T y(\mu_1) \geq b^T y(\mu_2).$$

Dôkaz. Nech (y_{μ_2}, Z_{μ_2}) je optimálne riešenie pre úlohu $SDD(\mu_2)$ a (y_{μ_1}, Z_{μ_1}) je optimálne riešenie pre úlohu $SDD(\mu_1)$ a platí

$$\psi_{\mu_2}(y(\mu_2), Z_(\mu_2)) = b^T y(\mu_2) + \mu_2 ln(det(Z(\mu_2))) \tag{5.2.14}$$

$$\psi_{\mu_2}(y(\mu_1), Z(\mu_1)) = b^T y(\mu_1) + \mu_2 ln(det(Z(\mu_1))). \tag{5.2.15}$$

Keďže

$$\psi_{\mu_2}(y(\mu_2), Z(\mu_2)) \geq \psi_{\mu_2}(y(\mu_1), Z(\mu_1)) \tag{5.2.16}$$

tak môžeme napísať

$$b^T y(\mu_2) + \mu_2 ln(det(Z(\mu_2))) \geq b^T y(\mu_1) + \mu_2 ln(det(Z(\mu_1))) \tag{5.2.17}$$

a po úprave dostaneme

$$\mu_2[ln(det(Z(\mu_2))) - ln(det(Z(\mu_1)))] \geq b^T y(\mu_1) - b^T y(\mu_2). \tag{5.2.18}$$

Analogicky z nerovnosti

$$\psi_{\mu}(y(\mu_1), Z(\mu_1)) \geq \psi_{\mu_1}(y(\mu_2), Z(\mu_2)) \tag{5.2.19}$$

dostaneme

$$\mu_1[ln(det(Z(\mu_2))) - ln(det(Z(\mu_1)))] \leq b^T y(\mu_1) - b^T y(\mu_2). \tag{5.2.20}$$

Z nerovníc (5.2.18) a (5.2.20) dostávame

$$b^T y(\mu_1) - b^T y(\mu_2) \geq \frac{\mu_1}{\mu_2}(b^T y(\mu_1) - b^T y(\mu_2)) \tag{5.2.21}$$

a postupnými úpravami dospejeme k požadovanému výsledku.

$$(1 - \frac{\mu_1}{\mu_2} b^T y(\mu_1) \geq (1 - \frac{\mu_1}{\mu_2}) b^T y(\mu_2)$$

a keďže $0 < \mu_1 < \mu_2$ tak $1 - \frac{\mu_1}{\mu_2} > 0$ a preto

$$b^T y(\mu_1) \geq b^T y(\mu_2). \quad \square$$

Lema 5.2.8. (O ohraničenosti množín).
Množiny

(a) $\mathcal{P}^1 = \{X \in \mathcal{P} \mid C \bullet X \leq C \bullet X(1)\}$
(b) $\mathcal{D}^1 = \{(y, Z) \in \mathcal{D} \mid b^T y \geq b^T y(1)\}$

sú ohraničené.

Dôkaz. (a) Sporom. Nech $\mathcal{P}^1$ je neohraničená množina. $\mathcal{P}^1$ je konvexná množina čo vyplýva z konvexnosti množiny $\mathcal{P}$ a z linearity funkcie $C \bullet X$. Potom z Lemy 5.1.6 vieme, že musí existovať polpriamka

$$p = \{X_t \mid X_t = X_1 + t(X_2 - X_1) \mid t \geq 0\},$$

$$X_1, X_2 \in \mathcal{P}^1,\ X_2 - X_1 \succ 0,$$

ktorá celá leží v množine $\mathcal{P}^1$ teda $p \subset \mathcal{P}^1$. Označme

$$W_t = \frac{X_t}{\| X_t \|}. \tag{5.2.22}$$

Platí $\| W_t \| = 1$ a $W_t \succ 0$. Množina $\{W_t \mid t > 0\}$ je ohraničená, a preto z nej vieme vybrať konvergujúcu podpostupnosť $\{W_{t_k}\}_{k=1}^{\infty}$. Príslušnú limitu označíme

$$W = \lim_{k\to\infty} W_{t_k} \tag{5.2.23}$$

pričom zrejme $W \succeq 0$ a $\| W \| = 1$. Z definície W a W_t dostávame

$$A_i \bullet W = \lim_{k\to\infty} A_i \bullet W_{t_k} = \lim_{k\to\infty} \frac{A_i \bullet X_{t_k}}{\| X_{t_k} \|}. \tag{5.2.24}$$

Keďže $A_i \bullet X_{t_k} = b_i$ a pre $k \to \infty$ ide $\| X_{t_k} \| \to \infty$ dostávame

$$A_i \bullet W = \lim_{k\to\infty} \frac{b_i}{\| X_{t_k} \|} = 0. \tag{5.2.25}$$

Podobne

$$C \bullet W = \lim_{k\to\infty} C \bullet W_{t_k} = \lim_{k\to\infty} \frac{C \bullet X_{t_k}}{\| X_{t_k} \|} \leq \lim_{k\to\infty} \frac{C \bullet X(1)}{\| X_{t_k} \|} = 0. \tag{5.2.26}$$

Podľa Predpokladu 2 existuje $(y^0, Z^0) \in \mathcal{D}^0,\ Z^0 \succ 0$ také, že

$$\sum_{i=1}^{m} A_i y_i^0 = C - Z^0. \tag{5.2.27}$$

Z rovnosti (5.2.25) dostávame

$$(\sum_{i=1}^{m} A_i y_i^0) \bullet W = 0 \tag{5.2.28}$$

a po dosadení (5.2.27) do (5.2.28) máme

$$(C - Z^0) \bullet W = 0 \tag{5.2.29}$$

$$C \bullet W - Z^0 \bullet W = 0. \tag{5.2.30}$$

Pretože podľa (5.2.26) je $C \bullet W \leq 0$, musí platiť

$$Z^0 \bullet W \leq 0. \tag{5.2.31}$$

Ale keďže $Z^0 \succ 0$, $W \succeq 0$ tak $W = 0$ čo je ale spor s predpokladom $\| W \| = 1$.
(b) Sporom. Nech $\mathcal{D}^1$ je neohraničená množina. Vieme, že množina $\mathcal{D}^1$ je konvexná. Z Lemy 5.1.6 vieme, že musí existovať polpriamka

$$d = \{(y_t, Z_t) \mid (y_t, Z_t) = (y_1, Z_1) + t[(y_2, Z_2) - (y_1, Z_1)] \mid t \geq 0\},$$

$$(y_1, Z_1),\ (y_2, Z_2) \in \mathcal{D}^1,\ Z_1 - Z_2 \succ 0,$$

ktorá celá leží v množine $\mathcal{D}^1$. Označme

$$(v_t, W_t) = \frac{(y_1, Z_1) + t[(y_2, Z_2) - (y_1, Z_1)]}{\| (y_1, Z_1) + t[(y_2, Z_2) - (y_1, Z_1)] \|}. \tag{5.2.32}$$

Platí $\| (v_t, W_t) \| = 1$ a $W_t \succeq 0$. Na normu v menovateli sa môžeme pozerať ako na vektorovú normu v $n^2 + m$ rozmernom vektorovom priestore. Množina $\{(v_t, W_t) \mid t > 0\}$ je ohraničená, a preto z nej vieme vybrať konvergujúcu podpostupnosť $\{(v_{t_k}, W_{t_k})\}_{k=1}^{\infty}$. Príslušnú limitu označíme

$$(v, W) = \lim_{k\to\infty} (v_{t_k}, W_{t_k}). \tag{5.2.33}$$

Zrejme z vlastnosti limity $\| (v, W) \| = 1,\ W \succeq 0$. Ak podmienku z duálnej úlohy

$$\sum_{i=1}^{m} A_i y_i + Z = C, \text{ kde } Z \succeq 0 \tag{5.2.34}$$

podelíme výrazom $\| (y_1, Z_1) + t_k[(y_2, Z_2) - (y_1, Z_1)] \|$ a pošleme $k \to \infty$ tak dostaneme

$$\sum_{i=1}^{m} v_i A_i + W = \lim_{k\to\infty} (\sum_{i=1}^{m} (v_{t_k})_i A_i + W_{t_k}) = \tag{5.2.35}$$

$$= \lim_{k\to\infty} \frac{\sum_{i=1}^{m} (y_{1i} + t_k(y_{2i} - (y_{1i}) A_i + (Z_1 + t_k(Z_2 - Z_1))}{\| (y_1, Z_1) + t_k((y_2, Z_2) + (y_1, Z_1)) \|} =$$

$$= \lim_{k\to\infty} \frac{C}{\| (y_1, Z_1) + t_k((y_2, Z_2) + (y_1, Z_1)) \|} = 0$$

Podobne

$$b^T v = \lim_{k\to\infty} b^T v_{t_k} \geq \lim_{k\to\infty} \frac{b^T y(1)}{\| (y_1, Z_1) + t_k((y_2, Z_2) + (y_1, Z_1)) \|} = 0. \tag{5.2.36}$$

Podľa Predpokladu 2 primárna úloha má vnútorný bod, teda existuje $X^0 \succ 0$, ktoré spĺňa

$$A_i \bullet X^0 = b_i \text{ pre } i = 1, 2, \ldots, m. \tag{5.2.37}$$

Úpravou dostaneme

$$(\sum_{i=1}^{m} v_i A_i + W) \bullet X^0 = 0$$

$$\sum_{i=1}^{n} v_i A_i \bullet X^0 + W \bullet X^0 = 0$$

$$\sum_{i=1}^{n} v_i b_i + W \bullet X^0 = 0$$

$$b^T v + W \bullet X^0 = 0$$

$$W \bullet X^0 = -b^T v \leq 0$$

$$0 \geq W \bullet X^0 = tr(W X^0) = tr(W^{\frac{1}{2}} X^0 W^{\frac{1}{2}})$$

Z podmienok máme $X^0 \succ 0$ a preto dostávame $W = 0$ čo je spor s predpokladmi. $\square$

Dôsledok 5.2.9. *Množiny*

$$M = \{X(\mu) \mid 0 < \mu \leq 1\},$$
$$N = \{y(\mu), Z(\mu) \mid 0 < \mu \leq 1\}$$

sú ohraničené.

Dôkaz. Pretože $C \bullet X(\mu) \leq C \bullet X(1)$ pre ľubovoľné $\mu \in (0, 1\rangle$ tak $M \subset \mathcal{P}^1$. Podobne $b^T y(\mu) \geq b^T y(1)$ pre ľubovolné $\mu \in (0, 1\rangle$ a preto $N \subset \mathcal{D}^1$. Vzhľadom na to, že množiny $\mathcal{P}^1$ a $\mathcal{D}^1$ sú ohraničené, tak aj množiny M a N sú ohraničené.

Veta 5.2.10. *Pre ľubovoľnú postupnosť $\{\mu_k\}_{k=1}^{\infty}$, ktorá zhora konverguje k 0, existuje taká z nej vybraná podpostupnosť $\{\mu_{k_i}\}_{i=1}^{\infty}$, že postupnosti*

$$\{X(\mu_{k_i})\}_{i=1}^{\infty} \ a \ \{(y(\mu_{k_i}),\ Z(\mu_{k_i}))\}_{i=1}^{\infty}$$

konvergujú a limitný bod každej konvergentnej postupnosti je optimálne riešenie primárno-duálnej dvojice čo môžeme napísať ako

$$\lim_{i\to\infty} X(\mu_{k_i}) \in \mathcal{P}^*,$$
$$\lim_{i\to\infty} (y(\mu_{k_i}), Z(\mu_{k_i})) \in \mathcal{D}^*.$$

Dôkaz. Z Dôsledku 5.2.9 vieme, že každá postupnosť $\{X(\mu_k)y(\mu_k), Z(\mu_k)\}_{k=1}^{\infty}$ je ohraničená a teda z nej možno vybrať konvergentnú podpostupnosť takú, že

$$\lim_{j\to\infty} X(\mu_{k_j}) \to \tilde{X},$$
$$\lim_{j\to\infty} y(\mu_{k_j}) \to \tilde{y},$$
$$\lim_{j\to\infty} Z(\mu_{k_j}) \to \tilde{Z}.$$

Pretože

$$X(\mu_{k_j})Z(\mu_{k_j}) = \mu_{k_j} I,$$

tak prechodom k limite ($j \to \infty$) dostávame

$$\tilde{X}\tilde{Z} = 0$$

a teda konverhujú k optimálnemu riešeniu. $\square$

5.3 Generický algoritmus.

Veľa metód vnútorného bodu pre semidefinitné programovanie bolo odvodených z lineárneho programovania.

Budeme riešiť úlohu (SDPμ) resp. (SDDμ) a pre zmenšujúce μ budeme sledovať centrálnu trajektóriu. Vieme, že pre ľubovoľné $\mu > 0$ existuje jednoznačné riešenie $X(\mu), y(\mu), Z(\mu)$ úloh (SDPμ) a (SDDμ), ktoré spĺňa centrujúci systém

$$A_i \bullet X = b_i,\ i = 1, \ldots, m,\ X \succ 0, \tag{5.3.1}$$

$$\sum_{i=1}^{m} A_i y_i + Z = C,\ Z \succ 0. \tag{5.3.2}$$

$$XZ = \mu I. \tag{5.3.3}$$

Majme danú iteráciu (X^k, y^k, Z^k), $X^k \succ 0$, $Z^k \succ 0$. Ďalšia iterácia bude $(X^{k+1}, y^{k+1}, Z^{k+1}) = (X^k + \triangle X, y^k + \triangle y, Z^k + \triangle Z)$, ktorá tiež musí spĺňať centrujúci systém, teda

$$A_i \bullet (X^k + \triangle X) = b_i,\ i = 1, \ldots, m, \tag{5.3.4}$$

$$\sum_{i=1}^{m} A_i (y_i^k + \triangle y_i) + (Z^k + \triangle Z) = C, \tag{5.3.5}$$

$$(X^k + \triangle X)(Z^k + \triangle Z) = \mu I. \tag{5.3.6}$$

Linearizáciou (5.3.4)-(5.3.6) získame

$$A_i \triangle X = -(A_i \bullet X^k - b_i),\ i = 1, \ldots, m, \tag{5.3.7}$$

$$\sum_{i=1}^{m} A_i \triangle y_i + \triangle Z = -(\sum_{i=1}^{m} A_i y_i^k + Z^k - C), \tag{5.3.8}$$

$$\triangle X Z^k + X^k \triangle Z \approx -(X^k Z^k - \mu I). \tag{5.3.9}$$

Vo výraze (5.3.9) sme v dôsledku linearizácie vynechali člen $\triangle X \triangle Z$. Z toho vidíme, že v každej iterácii budeme riešiť systém

$$A_i \triangle X = -(A_i \bullet X - b_i),\ i = 1, \ldots, m, \tag{5.3.10}$$

$$\sum_{i=1}^{m} A_i \triangle y_i + \triangle Z = -(\sum_{i=1}^{m} A_i y_i + Z - C), \tag{5.3.11}$$

$$\triangle X Z + X \triangle Z = -(XZ - \mu I) \tag{5.3.12}$$

vzhľadom na smer $(\triangle X, \triangle y, \triangle Z)$. Riešením rovníc (5.3.10)-(5.3.12) získame síce symetrickú maticu $\triangle Z$ (5.3.11), ale matica $\triangle X$ vo všeobecnosti nemusí byť symetrická. Preto ju potrebujeme dodatočne symetrizovať. Jedna z možností je definovať $\triangle \tilde{X}$ nasledovne

$$\triangle \tilde{X} = \frac{\triangle X + \triangle X^T}{2}.$$

Ľahko nahliadneme, že matica $\triangle \tilde{X}$ je symetrická. Potom položíme

$$(X_N, y_N, Z_N) = (X, y, Z) + \alpha(\triangle \tilde{X}, \triangle y, \triangle Z),$$

kde parameter α vyberieme tak, aby $X_N \succ 0$ a $Z_N \succ 0$ čo je vlastne Newtonova metóda s voľbou dĺžky kroku.

Teraz môžeme formulovať

Generický algoritmus.

Vstup: $A_i,\ i = 1, \ldots, m,\ b,\ C,\ \sigma > \varepsilon,\ X^0 \succ 0, y^0, Z^0 \succ 0$, položíme $k = 0$.

1. Zvolíme μ_k.
2. Riešime systém (5.3.10)-(5.3.12) a vypočítame $(\triangle\tilde{X}, \triangle y, \triangle Z)$ v danom bode (X^k, y^k, Z^k).
3. Zvolíme $\alpha > 0$ tak, aby
$$X^k + \alpha\triangle\tilde{X} \succ 0,$$
$$Z^k + \alpha\triangle Z \succ 0.$$
4. Položíme $(X^{k+1}, y^{k+1}, Z^{k+1}) = (X^k, y^k, Z^k) + \alpha(\triangle\tilde{X}, \triangle y, \triangle Z)$.
5. Ak pre všetky $j \in R^m$ platí $\| A_j(X^{k+1} - b) \| + \| \sum_{i=1}^m A_i y_i^{k+1} + Z^{k+1} - C \| < \sigma$ a zároveň $X^{k+1}\tilde{Z}^{k+1} < \varepsilon$ tak STOP. V opačnom prípade položíme $k := k + 1$ a ideme na bod 1.

Znižujúc hodnotu μ v bode 1. a idúc s μ k nule dostávame primárno-duálnu metódu pre SDP. Generický algoritmus predstavuje všeobecnú štruktúru metód vnútorného bodu pre SDP. Preto existujú rozličné variácie tohoto algoritmu. Napríklad výber parametra μ v kroku 1. Jedna z možností je nemeniť parameter μ v každom kroku, ide o úlohy s tzv. dlhým krokom. Parameter μ sa mení až po splnení určitého kritéria, napríklad $\| X \bullet Z - \mu n \| < \gamma$. Potom nastane vôčšia zmena μ, preto úlohy s dlhým krokom. Iná možnosť je meniť parameter μ v každom kroku. V tomto prípade ale potrebujeme mať zaruené, že od jednej iterácii k druhej sa zmenší duálna medzera, ktorá je určená rovnicou (5.3.3), lebo

$$\begin{aligned} XZ &= \mu I \\ stopa(XZ) &= stopa(\mu I) \\ X \bullet Z &= \mu n. \end{aligned}$$

Ak sa nachádzame v k-tej iterácii, tak príslušná duálna medzera je $X^k \bullet Z^k$. Chceme zabezpečiť, aby pri prechode k ďalšej iterácii $k+1$ bola duálna medzera menšia, teda $X^{k+1} \bullet Z^{k+1} < X^k \bullet Z^k$. Vyjadríme duálnu medzeru v $k+1$ iterácii

$$\begin{aligned} X^{k+1} \bullet Z^{k+1} &= (X^k + \alpha\triangle X) \bullet (Z^k + \alpha\triangle Z) \\ &= X^k \bullet Z^k + \alpha(X^k \bullet \triangle Z + Z^k \bullet \triangle X) + \alpha^2 \triangle X \bullet \triangle Z \\ &= X^k \bullet Z^k + \alpha(X^k \bullet \triangle Z + Z^k \bullet \triangle X) \\ &= X^k \bullet Z^k + \alpha stopa(X^k \triangle Z + Z^k \triangle X) \\ &= X^k \bullet Z^k + \alpha stopa(\mu I - X^k Z^k) \\ &= X^k \bullet Z^k + \alpha\mu n - \alpha X^k \bullet Z^k \\ &= (1-\alpha)(X^k \bullet Z^k) + \alpha\mu n. \end{aligned}$$

Teda parameter μ musí spĺňať

$$\begin{aligned} (1-\alpha)(X^k \bullet Z^k) + \alpha\mu n &< X^k \bullet Z^k \\ -\alpha X^k \bullet Z^k + \alpha\mu n &< 0 \\ \mu n &< X^k \bullet Z^k \\ \mu &< \frac{X^k \bullet Z^k}{n}. \end{aligned}$$

Ak budeme voliť $\mu < \frac{X^k \bullet Z^k}{n}$ tak potom v ďalšej iteráciii príde k polkesu duálnej medzery. Voľbou parametra $\mu = 0$ získame najväčší pokles duálnej medzery, naopak voľbou $\mu = \frac{X^k \bullet Z^k}{n}$ nenastane žiadny pokles duálnej medzery.

6. SEDUMI

SEDUMI je prídavný modul pre program MATLAB, pomocou ktorého sa dajú riešiť optimalizačné problémy s lineárnymi, kvadratickými a semidefinitnými podmienkami a umožňuje aj prácu s komplexnými dátami a premennými. SEDUMI je voľne prístupné na internetovej adrese http://fewcal.kub.nl/sturm/software/sedumi/html.

Na implementovanie modulu SEDUMI potrebujeme program MATLAB verziu 5 alebo novšiu. Postup inštalácie je nasledovný:

(1) Z uvedenej internetovej adresy stiahneme súbor SeDuMi104.zip, ktorý rozpakujeme do adresára určeného pre prídavné moduly (Toolbox), ktorý je najčastejšie v tvare C:\MATLAB\TOOLBOX.

(2) Teraz musíme v MATLAB-e nastaviť cestu k modulu SEDUMI. Nájdeme súbor MATLABRC.M, ktorý sa najčastejšie nachádza v adresáry C:\MATLAB\TOOLBOX\LOCAL a na koniec súbora pridáme text:

```
MATLABROOT = 'C:\MATLAB\';
addpath([MATLABROOT 'toolbox\SeDuMi104']);
```

samozrejme pokiaľ uvedená cesta k súboru platí.

Po spustení programu MATLAB môžeme napísať
≫ help SeDuMi104
alebo
≫ help sedumi
a budeme oboznámení s funkciami a ich popisom.

Teraz zapíšeme úlohu

$$\min_{X \in S^n} \{C \bullet X \mid A_i \bullet X = b_i,\ i = 1, \ldots, m,\ X \succeq 0\} \tag{SDP}$$

v tvare potrebnom pre modul SEDUMI. V (1.1.12) bola úloha (SDP) zapísaná pomocou funkcie vec nasledovne

$$\min_{X \in S^n} \{vec(C)^T vec(X) \mid Avec(X) = b,\ X \succeq 0\}. \tag{6.1}$$

Pre modul SEDUMI budú parametre vstupovať v tvare:

$$At = \begin{pmatrix} vec(A_1)^T \\ \vdots \\ vec(A_m)^T \end{pmatrix},\ At \in R^{m \times n^2},$$

$$c = vec(C),\ c \in R^{n^2},$$

$$b = b,\ b \in R^m,$$

$$K.s = [n],\ \text{hľadáme semidefinitné matice } n \times n$$

a úlohu zapíšeme nasledovne
≫ At=[vec(A1);vec(A2);...;vec(Am)];
≫ c=[vec(C)];
≫ b=[b];
≫ K.[s]=n;
Po zadaní vstupných dát výpočet spustíme príkazom sedumi(At,b,c,K).
≫ [x,y,info] = sedumi(At,b,c,K);
V premenných x, y a info dostaneme výsledok optimalizácie. V x a y budú nájdené riešenia primárnej a duálnej úlohy. V premennej info bude základná charakteristika nájdených riešení ako prípustnosť, počet iterácií a čas konvergencie.

Príklad 6.1. Majme úlohu (SDP), kde

$$A_1 = \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ide vlastne o Príklad 4.1.6. Úlohu zadáme nasledovne.
≫ At=[0,-1/2,0,-1/2,0,0,0,0,1;1,0,0,0,0,0,0,0,0;0,0,1,0,0,0,1,0,0;0,0,0,0,0,1,0,1,0];
≫ c=[0,1/2,0,1/2,0,0,0,0,0];
≫ K.s=[3];
≫ b=[1,0,0,0];
Teraz môžeme spustiť samotné riešenie úlohy.
≫ [x,y,info]=sedumi(At,b,c,K);
Po spustení úlohy sa na obrazovke objaví nasledujúci text.

```
SeDuMi 1.04 Pre-Test by Jos F. Sturm, 1998, 2000. 15/9/00
Alg = 1: v-corrector, theta = 0.250, beta = 0.500
eqs m = 4, order n = 4, dim = 10, blocks = 2
nnz(A) = 5 + 0, nnz(ADA) = 16, nnz(L) = 10
it :   b*y       gap     delta  rate  t/maxt   feas
00 :           4.00E+000 0.000
01 :  9.93E-002 3.22E-001 0.000 0.0806  0.9900  1.00
02 : -6.43E-002 9.45E-002 0.000 0.2931 0.9000   0.20
03 : -1.29E-001 2.52E-002 0.000 0.2666 0.9000   0.11
04 : -1.70E-001 5.74E-003 0.000 0.2278 0.9000   0.05
05 : -2.07E-001 5.73E-004 0.000 0.0998 0.9900   0.03
06 : -2.13E-001 2.06E-004 0.000 0.3597 0.9000   0.01
07 : -2.18E-001 6.89E-005 0.000 0.3340 0.9000   0.00
08 : -2.23E-001 2.08E-005 0.000 0.3016 0.9000   0.00
09 : -2.27E-001 5.44E-006 0.000 0.2618 0.9000   0.00
10 : -2.30E-001 1.17E-006 0.000 0.2155 0.9000   0.00
11 : -2.34E-001 9.87E-008 0.000 0.0842 0.9900   0.00
12 : -2.34E-001 3.12E-008 0.000 0.3161 0.9000   0.00
13 : -2.35E-001 8.71E-009 0.000 0.2792 0.9000   0.00
14 : -2.35E-001 2.05E-009 0.000 0.2352 0.9000   0.00
```

15 : -2.36E-001 3.83E-010 0.000 0.1869 0.9000 0.00
16 : -2.36E-001 2.19E-011 0.000 0.0572 0.9900 0.00
17 : -2.36E-001 5.08E-012 0.000 0.2322 0.9000 0.00
18 : -2.36E-001 9.33E-013 0.000 0.1838 0.9000 0.00
19 : -2.36E-001 5.09E-014 0.000 0.0545 0.9900 0.00
20 : -2.36E-001 1.14E-014 0.000 0.2234 0.9000 0.00
21 : -2.36E-001 1.05E-015 0.000 0.0925 0.9900 -0.00
22 : -2.36E-001 3.57E-016 0.000 0.3392 0.9000 0.00
23 : -2.36E-001 1.10E-016 0.000 0.3079 0.9000 -0.00
24 : -2.36E-001 2.96E-017 0.000 0.2693 0.9000 0.00
25 : -2.37E-001 6.63E-018 0.000 0.2242 0.9000 -0.00
26 : -2.32E-001 6.24E-019 0.000 0.0941 0.9900 0.00
iter seconds digits c*x b*y
26 1.8 Inf -3.7622234692e-001 -2.3166954389e-001
—Ax-b— = 1.1e-010, —x—= 1.4e+009, —y—= 1.4e+009
Teraz môžeme vypísať riešenia, ktoré modul Sedumi našiel.
≫ x
x =
1.0e+009 *

0.0000
-0.0000
0
-0.0000
1.3998
0
0
0
0.0000

Matica X z primárnej úlohy bude vyzerať nasledovne
≫ mat(x)
1.0e+009 *

0.0000	-0.0000	0
-0.0000	1.3998	0
0	0	0.0000

≫ y
y =
1.0e+009 *

-0.0000
-1.4245
0
0

Keď máme riešenia nájdené modulom Sedumi, tak vlastnosti týchto riešení sa nachádzajú v premmennej info. Ak pinf=dinf=0 tak x a y sú približne optimálne riešenia. Ak pinf=1 tak primárna úloha je neprípustná, podobne ak dinf=1 tak duálne úloha je neprípustná. V tomto príklade vidíme, že pinf=dinf=0.

$\gg$ info

info =

cpusec: 1.7500

iter: 26

feasratio: 0.0015

numerr: 0

pinf: 0

dinf: 0

7. Dodatok

Definícia 7.1. Matica $A \in S^n$ práve vtedy, keď $A^T = A$.

Definícia 7.2.

(a) $A \in S^n$ je kladne semidefinitná ($A \in S^n_+,\ A \succeq 0$) ak pre $\forall x \in R^n\ :\ x^T Ax \geq 0$.

(b) $A \in S^n$ je kladne definitná ($A \in S^n_{++},\ A \succ 0$) ak pre $\forall x \in R^n,\ x \neq 0\ :\ x^T Ax > 0$.

Definícia 7.2. Zadefinujeme vektorovú funkciu maticového argumentu, kde každej matici $X \in R^{n\times n}$ priradíme vektor $diag(X) \in R^n$ nasledovne

$$x = diag(X) = \begin{pmatrix} x_{11} \\ x_{22} \\ \vdots \\ x_{nn} \end{pmatrix},$$

kde x_{ii} je prvok matice X v i-tom riadku a i-tom stĺpci. Podobne môžeme definovať maticovú funkciu vektorového argumentu, kde každému vektoru $x \in R^n$ priradíme maticu $Diag(x) \in R^{n\times n}$ nasledovne

$$Diag(x) = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_n \end{pmatrix},$$

kde x_i je i-ty prvok vektora x.

Poznámka 7.3. Nech matica $X \in R^{n\times n}$. Potom sa dá ľahko nahliadnuť, že

$$Diag(diag(X)) = \begin{pmatrix} x_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_{nn} \end{pmatrix},$$

kde x_{ii} je prvok matice X v i-tom riadku a i-tom stĺpci.

Lema 7.4. *Nech $A_i \in S^{n_i}$ pre $i = 1, \dots, m$. Potom matica*

$$A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_m \end{pmatrix}$$

je kladne definitná resp. semidefinitná práve vtedy, keď každá matica A_i, $i = 1, \ldots, m$ je kladne definitná resp. semidefinitná.

Dôkaz. Nech $x_{n_i} \in R^{n_i}$, $i = 1, \ldots, m$ sú ľubovoľné vektory a nech $n = \sum_{i=1}^{m} n_i$. Definujme vektor $x = (x_{n_1}^T, \ldots, x_{n_m}^T)^T \in R^n$. Potom platí

$$x^T A x = x_{n_1}^T A_1 x_{n_1} + \cdots + x_{n_m}^T A_m x_{n_m}.$$

Ak sú matice A_i, $i = 1, \ldots, m$ kladne semidefinitné tak ľahko nahliadneme, že aj matica A je kladne semidefinitná. Naopak nech je kladne semidefinitná matica A. Sporom. Ak by existoval vektor x_{n_j} pre nejaké $j \in \{1, \ldots, m\}$ taký, že $x_{n_j}^T A_{n_j} x_{n_j} < 0$ tak položením $x_{n_i} = 0$ pre $i \in \{1, \ldots, m\} \setminus j$ získame spor $x^T A x = x_{n_j}^T A_j x_{n_j} < 0$.

Lema 7.5. *Pre ľubovoňú maticu $A \in S^n$ existuje matica C ortonormálna ($CC^T = I$) a diagonálna matica D, kde $A = CDC^T$. Matica D má na diagonále vlastné čísla matice A a k-ty stĺpec C je vlastný vektor prislúchajúci ku k-temu vlastnému číslu A.*

Dôkaz. Nech v_i je vlastný vektor A prislúchajúci k vlastnému číslu λ_i a nech v_j je vlastný vektor patriaci k λ_j, kde $i \neq j$. Počítajme

$$\begin{aligned} Av_i &= \lambda_i v_i \\ v_i^T A &= \lambda_i v_i^T \qquad (/v_j) \\ v_i^T A v_j &= \lambda_i v_i^T v_j \\ v_i^T \lambda_j v_j &= \lambda_i v_i^T v_j \\ v_i^T v_j &= \frac{\lambda_i}{\lambda_j} v_i^T v_j. \end{aligned}$$

Z toho vidíme, že $v_i v_j = 0$ pre $i \neq j$. Normovaním vieme docieliť aby $v_i v_i = 1$. □

Veta 7.6. *Nech $A \in S^n$. Potom nasledujúce výroky sú ekvivalentné*

(a) *A je kladne definitná,*
(b) *vlastné čísla $\lambda_i(A) > 0$ pre $i = 1, \ldots, n$,*
(c) *existuje matica $C \in R^{n \times n}$ taká, že $rank(C) = n$ a $A = C^T C$.*

Dôkaz. Ukážeme, že $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$.
$(a) \Rightarrow (b)$ Nech v je ľubovoľný vlastný vektor matice A a nech λ je prislušné vlastné číslo. Potom

$$\begin{aligned} \lambda v &= Av \quad \text{/zľava } v^T \\ \lambda v^T v &= v^T A v > 0. \end{aligned}$$

Z toho vidíme,že $\lambda > 0$.
$(b) \Rightarrow (c)$ Nech $A = V \Lambda V^T$ kde stĺpce V sú vlastné vektory a Λ má na diagonále príslušné vlastné čísla $\lambda_i(A)$, $i = 1, \ldots, n$. Keďže $\lambda_i(A) > 0$ tak $\Lambda = \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}}$ a preto $C = \Lambda^{\frac{1}{2}} V^T$.
$(c) \Rightarrow (a)$ Pre ľubovoľné $v \in R^n$, $v \neq 0$ platí

$$v^T A v = v^T C^T C v = (Cv)^T C v > 0$$

lebo $Cv \neq 0$. □

Veta 7.7. *Nech $A \in S^n$. Potom nasledujúce výroky sú ekvivalentné*

(a) *A je kladne semidefinitná,*
(b) *vlastné čísla $\lambda_i(A) \geq 0$ pre $i = 1, \ldots, n$,*
(c) *existuje matica $C \in R^{m\times n}$ taká, že $rank(C) = rank(A)$ a $A = C^TC$.*

Dôkaz. Analogický ako pri predchádzajúcej vete.

Lema 7.8. *Nech $B \in R^{n\times n}$ také, že $rank(B) = n$. Potom*

(a) *$A \in S^n_+$ práve vtedy, keď $B^TAB \in S^n_+$,*
(b) *$A \in S^n_{++}$ práve vtedy, keď $B^TAB \in S^n_{++}$.*

Dôkaz. Pre ľubovoľné $x \in R^n$ a $y = B^{-1}x$ platí

$$x^TAx = x^T(B^{-1})^TB^TABB^{-1}X = y^TB^TABy. \quad \square$$

Veta 7.9. (Shurov doplnok). *Nech $A \in S^m_{++}$, $C \in S^n_+$ a $B \in R^{m\times n}$. Potom*

(a) $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succ 0$ *práve vtedy, keď $C - B^TA^{-1}B \succ 0$,*
(b) $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0$ *práve vtedy, keď $C - B^TA^{-1}B \succeq 0$.*

Dôkaz. Nech

$$X = \begin{pmatrix} I_m & -A^{-1}B \\ 0 & I_n \end{pmatrix}.$$

Vidíme, že X je regulárna matica a platí

$$X^T \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} X = \begin{pmatrix} A & 0 \\ 0 & C - B^TA^{-1}B \end{pmatrix}.$$

Lema 7.10. *Nech $A, B \in S^n_+$. Potom $A \bullet B \geq 0$.*

Dôkaz. Pretože $A \in S^n_+$ tak $A = A^{\frac{1}{2}}A^{\frac{1}{2}}$. Potom

$$A \bullet B = stopa(AB) = stopa(A^{\frac{1}{2}}BA^{\frac{1}{2}}) \geq 0$$

lebo $A^{\frac{1}{2}}BA^{\frac{1}{2}} \succeq 0$.

Lema 7.11. *Nech $X,\ Z \in S^n$ a $X \succeq 0,\ Z \succeq 0$. Potom*

$$X \bullet Z = 0 \text{ práve vtedy, keď } XZ = ZX = 0_{n\times n}.$$

Dôkaz. Máme predpoklad, že $X \bullet Z = 0$. Počítajme

$$0 = X \bullet Z = stopa(XZ) = stopa(X^{\frac{1}{2}}ZX^{\frac{1}{2}}).$$

Z toho vidíme, že súčet prvkov na diagonále matice $X^{\frac{1}{2}}ZX^{\frac{1}{2}}$ je nulový a pretože $X \succeq 0$, $Z \succeq 0$ tak aj $X^{\frac{1}{2}}ZX^{\frac{1}{2}} \succeq 0$ a preto matica $X^{\frac{1}{2}}ZX^{\frac{1}{2}}$ má na diagonále nuly. Navyše matica $X^{\frac{1}{2}}ZX^{\frac{1}{2}}$ je aj symetrická a preto je celá nulová. Takže

$$\begin{aligned} 0 &= X^{\frac{1}{2}}ZX^{\frac{1}{2}} \\ &= X^{\frac{1}{2}}Z^{\frac{1}{2}}Z^{\frac{1}{2}}X^{\frac{1}{2}} \\ &= (X^{\frac{1}{2}}Z^{\frac{1}{2}})(X^{\frac{1}{2}}Z^{\frac{1}{2}})^T. \end{aligned}$$

Z toho vidíme, že matica $X^{\frac{1}{2}}Z^{\frac{1}{2}}$ musí byť nulová. Preto

$$XZ = 0_{n\times n}.$$

Pretože $X \bullet Z = Z \bullet X$ podobne sa úkáže, že aj $ZX = 0_{n\times n}$.
Opačne nech $VW = 0_{n\times n}$. Potom

$$0 = Stopa(0_{n\times n}) = Stopa(VW) = V \bullet W. \quad \square$$

Literatúra

[1] M.Halická, *Semidefinitné programovanie*, prednášky, zimný semester 2000/01.

[2] Z.Křížková, *Semidefinitné programovanie. Teória a metódy vnútorného bodu,*, Písomná časť dizertačnej skúšky, (1999).

[3] C.Helmberg, *Semidefinite Programming for Combinatorial Optimization*, Konrad-Zuse-Zentrum fur Informationstechnik Berlin, (2000).

[4] F.Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J.Opt. 5, 13-51, (1997).

[5] JOS F.Sturm, *Using SEDUMI 1.0x, a MATLAB toolbox for optimalization over symmetric cones*, Department of Quantitative Economics, Maastricht University, (1999).

[6] P.Dreyer, *Semidefinite programming*, SDP lecture notes, (1995).

[7] M.Nayakkankuppam, H.Wolkowocz, *http://www-fp.mcs.anl.gov/otc/Guide/OptWeb/continuous/constrained/sdp/, (1997).*

[8] Renato D.C Monteiro, *Primal-Dual Path-Following Algorithms for Semidefinite Programming*, School of Industrial and Systems Engineering, Georgia Tech, (1996).

[9] R.A.Horn, Ch.R.Johnson, *Matrix Analysis*, Cambridge University Press, (1986).