UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

PROGRAMOVANIE NAD KUŽEĽMI DRUHÉHO RÁDU

Diplomová práca

Bratislava 2013

Bc. Katarína Pokorná

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

PROGRAMOVANIE NAD KUŽEĽMI DRUHÉHO RÁDU

Diplomová práca

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Školiteľ:	RNDr. Mária Trnovská, PhD.
Evidenčné číslo:	89412a36-3b08-47c3-beac-531741e496a8

Bratislava 2013 **Bc. Katarína Pokorná**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Katarína Pokorná
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Programovanie nad kužeľmi druhého rádu

Cieľ: Teoretická práca so zameraním na študovanie teórie a aplikácii v tzv. SOCP/ ice cream cone programovaní, kde sa optimalizuje lineárna funkcia a množinou prípustných riešení je kužel 2. rádu. Cieľom je dôsledné spracovanie teórie duality, popis metódy vnútorného bodu pre takéto úlohy a aplikácie.

Vedúci: RNDr. Mária Trnovská, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Dátum zadania: 25.01.2012

Dátum schválenia: 26.01.2012 prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu

....
študent vedúci práce

Čestné prehlásenie

Čestne prehlasujem, že som diplomovú prácu vypracovala samostatne s využitím svojich vedomostí a s použitím uvedenej literatúry.

V Bratislave, apríl 2013 ....

Katarína Pokorná

Poďakovanie

Touto cestou sa chcem poďakovať svojej vedúcej diplomovej práce RNDr. Márii Trnovskej, PhD. za odborné rady a podnetné pripomienky, ktoré mi pomohli pri písaní tejto práce, ale predovšetkým za ochotu, trpezlivosť a ľudský prístup. Ďakujem aj svojej rodine a priateľom za ich podporu.

Abstrakt

POKORNÁ, Katarína: Programovanie nad kužeľmi druhého rádu [Diplomová práca] - Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky; Katedra aplikovanej matematiky a štatistiky. - Vedúci diplomovej práce: RNDr. Mária Trnovská, PhD.;Katedra aplikovanej matematiky a štatistiky, Univerzita Komenského, Bratislava.- Bratislava 2013 /59s./

Táto diplomová práca sa zaoberá programovaním nad kužeľmi druhého rádu. Ide o úlohy konvexnej optimalizácie, v ktorých množinu prípustných riešení tvorí prienik polyédra a kužeľov druhého rádu. V práci sú stručne uvedené aplikácie ako motivácia pre skúmanie teórie programovania nad kužeľmi druhého rádu. Potom sa práca venuje všeobecne kužeľom a ich vlastnostiam a následne opisuje kužele druhého rádu. Jadro práce tvorí podrobné spracovanie teórie duality, hlavne dôkaz silnej vety o dualite pomocou zovšeobecnenej Farkašovej lemy. Na záver je práca zameraná na metódy vnútorného bodu, kde je aj uvedený algoritmus na riešenie úloh programovania nad kužeľmi druhého rádu.

Kľúčové slová: Programovanie nad kužeľmi druhého rádu (SOCP), kužeľ, dualita, metódy vnútorného bodu

Abstract

POKORNÁ, Katarína: Second-order cone programming [Master's thesis] - Comenius University in Bratislava. Faculty of mathematics, physics and informatics; Department of applied mathematics and statistics. - Supervisor: RNDr. Mária Trnovská, PhD.; Department of applied mathematics and statistics, Comenius University, Bratislava.- Bratislava 2013 /59pp./

This thesis deals with second-order cone programming. There are convex optimization problems where the feasible region is an intersection of polyeder with second-order cones. The work provides a short introduction to the applications as well as motivation to explore theory of second-order cone programming. It includes a basic properties of cones in general and then describes the second-order cones. The main theme of this work is deeper analysis of duality, especially a proof of strong duality theorem using generalized Farkas lemma. In conclusion, this thesis is focused on interior-point methods and there we can find an algorithm to solve the second-order cone programms.

Keywords: Second-order cone programming, cone, duality, interior-point methods

Obsah

Zoznam používaných symbolov

R	reálne čísla
R^n	priestor reálnych n-rozmerných vektorov
R^n_+	nezáporný ortant
R_+, R_{++}	nezáporné, kladné reálne čísla
S^n	priestor symetrických reálnych matíc typu $(n \times n)$
S^n_+, S^n_{++}	priestor symetrických kladne semidefinitných, kladne definitných matíc typu $(n \times n)$
$(x; y)$	$(x; y) = (x^T, y^T)^T$, kde ^{T} označuje transpozíciu matice alebo vektora
$x = (x_0; \bar{x})$	n-rozmerný reálny stĺpcový vektor $(x_0, \bar{x}^T)^T$, kde $x_0 \in R, \bar{x} = (x_1, \dots, x_{n-1})^T \in R^{n-1}$
$\|\|.\|\|$	štandardná Euklidovská norma, pre $x \in R : \|\|x\|\| = (\sum_{i=0}^{n-1} x_i^2)^{1/2} = \sqrt{x^T x}$
$x^T s$	skalárny súčin pre $x, s \in R^n$, t.j. $x^T s = \sum_{i=1}^n x_i s_i$
$\mathcal{A} \times \mathcal{B}$	Kartézsky súčin, $\mathcal{A} \times \mathcal{B} = \{(x; y) : x \in \mathcal{A} \wedge y \in \mathcal{B}\}$, kde $\mathcal{A} \subseteq R^k, \mathcal{B} \subseteq R^l$
$int\mathcal{K}$	vnútro kužeľa $\mathcal{K}$
$bd\mathcal{K}$	hranica kužeľa $\mathcal{K}$
$cl\mathcal{K}$	uzáver kužeľa $\mathcal{K}$
$\mathcal{K}^*$	duálny kužeľ ku kužeľu $\mathcal{K}$, $\mathcal{K}^* = \{z : \forall x \in \mathcal{K}, x^T z \geq 0\}$
$\mathcal{Q}_n$	kužeľ druhého rádu, $\mathcal{Q}_n = \{x = (x_0; \bar{x}) \in R^n : x_0 \geq \\|\bar{x}\\|\}$
$\mathcal{Q}_1$	kužeľ druhého rádu pre $n = 1$: $\mathcal{Q}_1 = \{x_0 \in R : x_0 \geq 0\}$
$B(y, \varepsilon)$	guľa so stredom $y \in R^n$ a polomerom $\varepsilon > 0$, $B(y, \varepsilon) = \{x \in R^n \mid \\|x - y\\| < \varepsilon\}$
S	jednotková guľa (unit sphere), $S = \{x \in R^n \mid \\|x\\| = 1\}$
$conv(A)$	konvexný obal množiny A
$rank(A)$	hodnosť matice A
LP	úloha lineárneho programovania
QP	kvadratické programovanie
LCQP	úloha kvadratického programovania s lineárnymi ohraničeniami
QCQP	úloha kvadratického programovania s kvadratickými ohraničeniami
SOCP	úloha programovania nad kužeľmi druhého rádu
SDP	úloha semidefinitného programovania

Úvod

V dnešnej dobe sa čoraz častejšie stretávame v praxi s optimalizačnými úlohami. Väčšinu z týchto úloh tvoria práve úlohy konvexnej optimalizácie, kde ide o minimalizáciu konvexnej funkcie na konvexnej množine. Nepochybne zaujímavým faktom je, že sa konvexná optimalizácia študuje už viac ako storočie. Napriek tomu sa široká škála aplikácii v oblastiach ako napr. štatistika, financie, komunikácie a siete, či datová analýza alebo spracovanie signálu, objavila až v 90. rokoch minulého storočia. Podnetom pre rozšírenie aplikácii bol vývoj metód vnútorného bodu a v r. 1988 ich zovšeobecnenie na všeobecné konvexné úlohy (Nesterov a Nemorivski). Ukázali sa tak viaceré výhody, prečo je dobré naformulovať optimalizačný problém ako úlohu konvexnej optimalizácie. Jedným z dôvodov je už vyššie spomenutá konvexnosť množiny prípustných riešení. Ďalej sú to práve metódy vnútorného bodu, ktoré zabezpečia efektívne a spoľahlivé riešenie úloh a navyše sa dajú implementovať do rôznych softvérov. A z teoretického hľadiska duálna úloha naformulovaná k pôvodnému problému môže častokrát poskytnúť zaujímavé interpretácie.

Nový pohľad na úlohy konvexnej optimalizácie otvoril možnosti pre rozvoj nových tried konvexnej optimalizácie. Jednou z nich bolo semidefinitné programovanie, kde sa optimalizovala úloha cez prienik afinnej množiny a kužeľa kladne semidefinitných matíc. Uplatnenie našlo napr. v teórii riadenia, štatistike, či spektrálnej analýze. Nevýhodou semidefinitného programovania sa ukázala výpočtová zložitosť potrebná pri jednotlivých iteráciách metód vnútorného bodu pre úlohy veľkého rozsahu. Preto bolo snahou rozvinúť ďalšiu triedu konvexnej optimalizácie, ktorá by riešila nelineárne konvexné úlohy. Tak v roku 2000 vznikla podtrieda semidefinitného programovania, tzv. programovanie nad kužeľmi druhého rádu, ktorá rozšírila možnosti aplikácii konvexnej optimalizácie hlavne v oblasti inžinierstva.

Cieľom tejto práce je spracovanie teórie pre stále pomerne novú podtriedu konvexnej optimalizácie - programovanie nad kužeľmi druhého rádu. Práca je členená na 6 častí. Prvá kapitola je všeobecne venovaná programovaniu nad kužeľmi druhého rádu, kde je uvedený aj motivačný príklad použitia týchto úloh v praxi, ako aj ďalšie možnosti aplikácie. V druhej kapitole oboznámime čitateľa s pojmom kužeľ a spomenieme niektoré základné vlastnosti. Takisto uvedieme dôležité vlastnosti kužeľa druhého rádu, ktoré budeme využívať pri spracovaní teórie. V tretej kapitole naformulujeme základnú úlohu programovania nad kužeľmi druhého rádu a definujeme k nej duálnu úlohu. Potom dokážeme silnú vetu o dualite pomocou zovšeobecnenej Farkašovej lemy a definujeme podmienku komplementarity. Štvrtá kapitola sa zaoberá vzťahmi medzi triedami konvexnej optimalizácie a je v nej uvedený prevod úloh medzi jednotlivými triedami. V piatej kapitole sú opísané metódy vnútorného bodu, kde sa zadefinuje pojem bariérovej funkcie, centrálnej trajektórie a ukáže sa algoritmus vhodný na získanie optimálneho riešenia. Na záver sú v dodatku uvedené niektoré dôkazy, ako aj definovaný tzv. spektrálny rozklad vektora, ktorý je potrebný pri metódach vnútorného bodu.

1 Programovanie nad kužeľmi druhého rádu

Programovanie nad kužeľmi druhého rádu, tzv. *second-order cone programming* (SOCP) alebo ice-cream cone programming, je trieda úloh konvexnej optimalizácie, v ktorých minimalizujeme lineárnu funkciu cez prienik polyedrickej množiny a kužeľov druhého rádu. Základný tvar takejto úlohy si môžeme predstaviť ako

$$\begin{aligned} \min \quad & c^T x, \\ & Ax = b, \\ & \|\bar{x}\| \leq x_0, \end{aligned} \tag{1}$$

kde $c \in R^n, b \in R^m, A \in R^{m \times n}$,$x = (x_0; \bar{x}) \in R^n$, pričom $x_0 \in R, \bar{x} \in R^{n-1}$, a ohraničenie $\|\bar{x}\| \leq x_0$ je ohraničením kužeľa druhého rádu.
SOCP úlohy ponúkajú široký záber aplikácii v inžinierstve,financiách, ako aj štatistike, preto programovanie nad kužeľmi druhého rádu patrí k dôležitým triedam úloh konvexnej optimalizácie. Predtým než sa budeme venovať teórii, ukážeme si motivačný príklad, kde sa dá programovanie nad kužeľmi druhého rádu aplikovať.

1.1 Markowitzov problém

Známy Markowitzov problém sa zaoberá optimalizáciou portfólia zloženého z n aktív. Model je založený na predpoklade, že cieľom investora je maximalizovať funkciu užitočnosti $u(\theta) = \mu_p - \frac{1}{2}\gamma\sigma_p^2$, kde $\mu_p = \mu^T\theta$ je očakávaný výnos portfólia (v eurách), $\sigma_p^2 = \theta^T\Sigma\theta$ je variancia výnosu portfólia a $\gamma > 0$ vyjadruje mieru rizikovej averzie investora. Potom môžeme napísať úlohu v tvare

$$\begin{aligned} \max \quad & \mu^T\theta - \frac{1}{2}\gamma\theta^T\Sigma\theta, \\ & 1^T\theta = c_0, \end{aligned} \tag{2}$$

kde $\theta \in R^n$ je premenná, ktorej zložky predstavujú množstvá v eurách investované do jednotlivých aktív. Ohraničenie, v ktorom $1 = (1, \ldots, 1) \in R^n$, zodpovedá podmienke, že môžeme investovať iba toľko kapitálu, koľko máme práve k dispozícii (c_0). Dá sa ľahko overiť, že úloha (2) je ekvivalentná s úlohou

$$\begin{aligned} \min \quad & \theta^T\Sigma\theta - \frac{2}{\gamma}\mu^T\theta, \\ & 1^T\theta = c_0. \end{aligned} \tag{3}$$

Takto naformulovaná úloha kvadratického programovania sa dá previesť na SOCP úlohu tak, že ohraničením účelovej funkcie dostaneme ohraničenie kužeľa druhého rádu (podrobne v kapitole SOCP úlohy a iné triedy konvexnej optimalizácie). Potom Markowitzov problém ako SOCP úloha bude vyzerať nasledovne

$$\begin{aligned} \min \quad & t \\ 1^T\theta &= c_0 \\ \left\| \begin{matrix} \Sigma^{\frac{1}{2}}\theta \\ \frac{1-\frac{2}{\gamma}\mu^T\theta - t}{2} \end{matrix} \right\| &\leq \frac{1+\frac{2}{\gamma}\mu^T\theta + t}{2}. \end{aligned} \tag{4}$$

Na riešenie SOCP úloh sa vyvíjajú rôzne programy a jedným z nich je CVX, ktoré pracuje v prostredí MATLAB-u [27]. CVX je určené všeobecne pre úlohy konvexnej optimalizácie, ale práve úlohy programovania nad kužeľmi druhého rádu sa tu dajú veľmi jednoducho implementovať. Po definovaní Σ, μ, γ, n by sme mohli Markowitzov problém v tvare SOCP úlohy (4) aplikovať v CVX takto:

```
cvx_begin
  variable theta(n);
  variable t(1);
  minimize(t);
  subject to
    sum(theta) == c0;
    norm([(sigma^0.5)*theta; (1-(2*mu'*theta/gama+t))/2]) <= (1+(2*mu'*theta/gama+t))/2;
cvx_end
```

Poznámka 1.1. Okrem Markowitzovho problému sú známe aj iné prístupy k optimalizácii portfólia a tiež sa dajú preformulovať na SOCP úlohy. Podrobne sa touto témou zaoberá autor v literatúre [12].

1.2 Truss Design

Medzi ďalšie známe aplikácie patrí oblasť tzv. Truss design. Táto problematika sa zaoberá konštrukciami rôznych stavieb. V úlohe máme daný počet uzlov, resp. bodov v priestore a pre každú spojnicu k dvoch uzlov je definovaná jej dĺžka L_k, tzv. Young's modulus E_k a volume t_k. Takisto poznáme vonkajšie sily pôsobiace na uzly. Cieľom je zistiť, ktoré spojnice bodov potrebujeme, aby konštrukcia bola stabilná. Zadefinujeme model na základe [13]

$$\begin{aligned} \min_{f,t,s} \quad & \frac{1}{2}\sum_{k=1}^{m} s_k \\ & Af = -F \\ & Mt \leq d \\ & \frac{L_k^2}{E_k} f_k^2 \leq t_k s_k \quad k = 1, \dots, m \\ & t \geq 0, s \geq 0, \end{aligned} \tag{5}$$

kde $f, t, s \in R^m$. Hlavným problémom pri prevode na úlohu SOCP je ohraničenie

$$\frac{L_k^2}{E_k} f_k^2 \leq t_k s_k,$$

zadefinujeme preto nasledujúce premenné

$$w_k = \frac{1}{2} t_k + \frac{1}{2} s_k,$$

$$y_k = -\frac{1}{2} t_k + \frac{1}{2} s_k.$$

Je zrejme, že

$$s_k = w_k + y_k, \quad t_k = w_k - y_k$$

Môžeme overiť, že

$$t_k s_k = w_k^2 - y_k^2$$

a potom ohraničenie má tvar

$$\sqrt{\frac{L_k^2}{E_k} f_k^2 + y_k^2} \leq w_k.$$

Potom už ľahko prevedeme úlohu (5) na SOCP úlohu

$$\begin{aligned} \min_{f,w,y} \quad & \frac{1}{2} e^T (w + y) \\ & Af = -F \\ & M(w - y) \leq d \\ & \frac{L_k^2}{E_k} f_k^2 \leq t_k s_k \quad k = 1, \ldots, m \\ & \left\| (y_k; \frac{L_k}{\sqrt{E_k}} f_k) \right\| \leq w_k, \quad k = 1, \ldots, m, \end{aligned} \tag{6}$$

kde $w, y, f \in R^m$.

Poznámka 1.2. Iné aplikácie SOCP úloh môžeme nájsť napr. v [18],[8].

2 Kužele a ich vlastnosti

Uvedieme niektoré základné pojmy a vlastnosti kužeľov, ktoré sú dôležité aj pre kužele druhého rádu. Tieto poznatky nám neskôr umožnia pracovať s úlohami programovania nad kužeľmi druhého rádu.

Definícia 2.1. Množinu $\mathcal{K}$ nazveme kužeľom, ak pre každé $x \in \mathcal{K}$ a $\theta \geq 0$ platí $\theta x \in \mathcal{K}$.

Definícia 2.2. Kužeľ $\mathcal{K} \subseteq R^n$ sa nazýva vlastný kužeľ, ak spĺňa nasledujúce vlastnosti:

- $\mathcal{K}$ je uzavretý,
- $\mathcal{K}$ je konvexný, t.j. pre ľubovoľné $x_1, x_2 \in \mathcal{K}$ a $\theta_1, \theta_2 \geq 0$ platí: $\theta_1 x_1 + \theta_2 x_2 \in \mathcal{K}$,
- $\mathcal{K}$ má neprázdne vnútro, t.j $int\mathcal{K} \neq \emptyset$,
- $\mathcal{K}$ je "špicatý"(pointed), čo znamená, že neobsahuje priamku, t.j.

$$x \in \mathcal{K}, -x \in \mathcal{K} \Rightarrow x = 0.$$

Pomocou vlastného kužeľa $\mathcal{K}$ môžeme zadefinovať čiastočné usporiadanie na R^n, tzv. zovšeobecnenú nerovnosť:

$$x \preceq_{\mathcal{K}} y \Leftrightarrow y - x \in \mathcal{K}.$$

Takto definovaná zovšeobecnená nerovnosť spĺňa nasledujúce vlastnosti:

- reflexívnosť: $x \preceq_{\mathcal{K}} x$ (platí, pretože $0 \in \mathcal{K}$),
- antisymetria : $x \preceq_{\mathcal{K}} y, y \preceq_{\mathcal{K}} x \Rightarrow x = y$ (vyplýva zo špicatosti kužeľa $\mathcal{K}$),
- tranzitívnosť: $x \preceq_{\mathcal{K}} y, y \preceq_{\mathcal{K}} z \Rightarrow x \preceq_{\mathcal{K}} z$ ($\mathcal{K}$ je kužeľ a je konvexný),
- zachováva sčítanie: $x \preceq_{\mathcal{K}} y, u \preceq_{\mathcal{K}} v \quad \Rightarrow \quad x + u \preceq_{\mathcal{K}} y + v$.
- zachováva prenásobenie nezáporným skalárom: $x \preceq_{\mathcal{K}} y, \alpha \geq 0 \quad \Rightarrow \quad \alpha x \preceq_{\mathcal{K}} \alpha y$,
- uzavretosť vzhľadom na limitu: ak $x_i \preceq_{\mathcal{K}} y_i \quad \forall i = 1, 2, \ldots$ a $x_i \to x, y_i \to y$ pre $i \to \infty$, potom $x \preceq_{\mathcal{K}} y$.

Podobne je definované aj ostré čiastočné usporiadanie na R^n:

$$x \prec_{\mathcal{K}} y \Leftrightarrow y - x \in int\mathcal{K},$$

pre ktoré platia nasledujúce vlastnosti:

- $x \prec_{\mathcal{K}} y \Rightarrow x \preceq_{\mathcal{K}} y$,

- $x \prec_{\mathcal{K}} y, u \preceq_{\mathcal{K}} v \quad \Rightarrow \quad x + u \prec_{\mathcal{K}} y + v.$
- $x \prec_{\mathcal{K}} y, \alpha > 0 \quad \Rightarrow \quad \alpha x \prec_{\mathcal{K}} \alpha y,$
- $x \not\prec_{\mathcal{K}} x,$
- $x \prec_{\mathcal{K}} y \quad \Rightarrow \quad$ pre u, v dosť malé $x + u \prec_{\mathcal{K}} y + v.$

Opačné nerovnosti $x \succeq_{\mathcal{K}} y$, resp. $x \succ_{\mathcal{K}} y$ sú ekvivalentné zápisom $y \preceq_{\mathcal{K}} x$, resp. $y \prec_{\mathcal{K}} x$. V prípade, že $\mathcal{K} = R_+$, tak zovšeobecnené nerovnosti $\preceq_{\mathcal{K}}$ a $\prec_{\mathcal{K}}$ zodpovedajú štandardným nerovnostiam $\leq$, $<$ v R.

Základné definície kužeľa a vlastnosti zovšeobecnenej nerovnosti sme popísali na základe ([8], kapitola 2) a teraz podrobne v príkladoch overíme vlastnosti vlastného kužeľa pre R^n a S^n_+.

Príklad 2.1. Nezáporný ortant.
Nezáporný ortant je množina bodov s nezápornými zložkami, t.j.

$$R^n_+ = \{x \in R^n | x_i \geq 0, i = 0, \dots, n-1\} = \{x \in R^n | x \succeq 0\}.$$

Ukážeme, že R^n_+ spĺňa vlastnosti vlastného kužeľa.

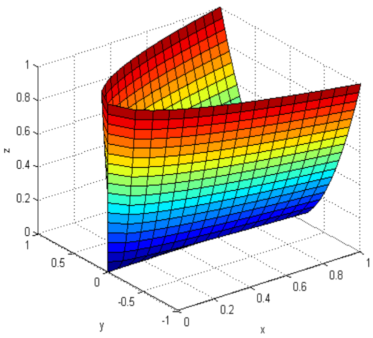
1. Uzavretosť R^n_+ vyplýva z otvorenej komplementárnej množiny $\{x \in R^n | x_i < 0, \forall i\}$ k R^n_+, pretože aj dostatočne malé okolia bodov x_i stále spĺňajú nerovnosti $x_i < 0$.
2. R^n_+ je konvexný kužeľ, ak pre ľubovoľné $x, y \in R^n_+$ a $\theta_1, \theta_2 \geq 0$ platí: $\theta_1 x + \theta_2 y \in R^n_+$, a teda $\forall i : \theta_1 x_i + \theta_2 y_i \geq 0$. To zrejme platí, pretože lineárna kombinácia nezáporných čísel je opäť nezáporné číslo.
3. Je zrejme, že $int R^n_+ \neq \emptyset$, lebo $int R^n_+ = R^n_{++} = \{x \in R^n | x_i > 0, \forall i\} = \{x \in R^n | x \succ 0\}$.
4. A nakoniec R^n_+ je špicatý, pretože $x \in R^n_+, -x \in R^n_+ \Leftrightarrow \forall i : x_i \geq 0 \wedge x_i \leq 0 \Leftrightarrow \forall i : x_i = 0 \Leftrightarrow x = 0$.

Keďže nezáporný ortant R^n_+ je vlastný kužeľ, tak môžeme definovať zovšeobecnenú nerovnosť $\preceq_{\mathcal{K}}$, ktorá zodpovedá nerovnostiam po zložkách, t.j. $x \preceq y \Leftrightarrow \forall i : x_i \leq y_i$. Analogicky pre ostrú nerovnosť platí $x \prec y \Leftrightarrow \forall i : x_i < y_i$.

Príklad 2.2. Kladne semidefinitný kužeľ.
Pod pojmom kladne semidefinitného kužeľa rozumieme množinu $S^n_+ = \{X \in S^n | X \succeq 0\}$, kde S^n je množina symetrických $(n \times n)$ matíc.
Ukážeme, že S^n_+ spĺňa všetky vlastnosti vlastného kužeľa.

1. S^n_+ je konvexný kužeľ, ak pre ľubovoľné $A, B \in S^n_+$ a $\theta_1, \theta_2 \geq 0$ platí: $\theta_1 A + \theta_2 B \in S^n_+$. Matice A,B sú kladne semidefinitné, ak pre ľubovoľné $x \in R^n$ spĺňajú nerovnosti $x^T A x \geq 0$ a $x^T B x \geq 0$. Potom pre ľubovoľné $x \in R^n$ platí:
$$x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T A x + \theta_2 x^T B x \geq 0 \Leftrightarrow A, B \in S^n_+ \text{ a } \theta_1, \theta_2 \geq 0.$$
2. Aby S^n_+ bol uzavretý, stačí ukázať, že komplement k S^n_+ je otvorená množina. Komplement tvorí množina symetrických matíc, ktoré nie sú kladne semidefinitné. Ak máme maticu M z komplementu, tak $\exists \bar{x} : \bar{x}^T M \bar{x} < 0$. Táto nerovnosť platí pre všetky matice v dostatočnej blízkosti matice M. Preto množina symetrických matíc, ktoré nie sú kladne semidefinitné, je otvorená.

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S^2_+ \Leftrightarrow x \geq 0, z \geq 0, xz \geq y^2$$



Obr. 1: Hranica kladne semidefinitného kužeľa v S^2 zobrazená v R^3 ako (x, y, z).

3. Vnútro S^n_+ je neprázdne, pretože $intS^n_+ = \{X \in S^n | X \succ 0\}$ je množina kladne definitných symetrických matíc S^n_{++}.
4. A S^n_+ je aj špicatý, pretože pre ľubovoľné $x \in R^n$ platí:

$$A \in S^n_+, -A \in S^n_+ \Leftrightarrow x^TAx \geq 0 \wedge x^TAx \leq 0 \Leftrightarrow x^TAx = 0 \Leftrightarrow A = 0.$$

Presvedčili sme sa, že kladne semidefinitný kužeľ je vlastný kužeľ, takže môžeme aj v tomto prípade definovať zovšeobecnenú nerovnosť $\preceq_{\mathcal{K}}$, ktorá predstavuje klasickú maticovú nerovnosť. Teda $X \preceq Y \Leftrightarrow Y - X$ je kladne semidefinitná matica . Ostrá zovšeobecnená nerovnosť zodpovedá ostrej nerovnosti medzi symetrickými maticami, t.j. $X \prec Y \Leftrightarrow Y - X$ je kladne definitná matica .

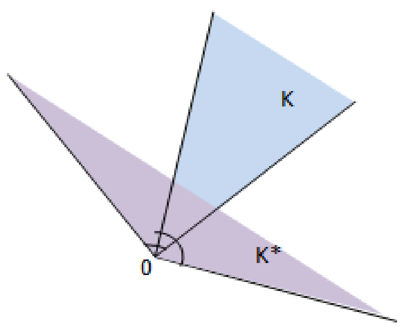
Ďalším štandardným typom kužeľov, ktoré spĺňajú vlastnosti vlastného kužeľa, sú kužele druhého rádu, ktorým sa budeme podrobne venovať neskôr.

2.1 Duálne kužele a ich vlastnosti

Pre každý kužeľ $\mathcal{K}$ definujeme *duálny kužeľ* ako množinu

$$\mathcal{K}^* = \{z : \forall x \in \mathcal{K}, x^Tz \geq 0\},$$

čo geometricky reprezentuje nasledujúci obrázok.



Definícia 2.3. Kužeľ $\mathcal{K}$ sa nazýva samoduálny, ak je zároveň svojím duálnym kužeľom, a teda $\mathcal{K}^* = \mathcal{K}$.

Vlastnosť samoduality (tzv. self-duality) spĺňajú napríklad kužele R^n_+ a S^n_+.

Príklad 2.3. Nezáporný ortant.
Nezáporný ortant R^n_+ je samoduálny, pretože $z^T x \geq 0 \quad \forall x \succeq 0 \Leftrightarrow z \succeq 0$.

Príklad 2.4. Kladne semidefinitný kužeľ.
Ukážeme, že kužeľ S^n_+ je samoduálny, t.j. pre $X, Y \in S^n$ platí:

$$tr(XY) = \sum_{i,j=0}^{n-1} X_{ij} Y_{ij} \geq 0 \quad \forall X \succeq 0 \Leftrightarrow Y \succeq 0.$$

Predpokladajme, že $Y \notin S^n_+$, teda $\exists z \in R^n : z^T Y z < 0$. Z vlastnosti stopy matice vyplýva:

$$z^T Y z = tr(z^T Y z) = tr(z z^T Y) < 0.$$

Matica zz^T je kladne semidefinitná, a teda $\exists X = zz^T : tr(XY) < 0$. Preto $Y \notin (S^n_+)^*$.
Teraz nech $X, Y \in S^n_+$. Použijeme rozklad matice $X = \sum_{i=0}^{n-1} \lambda_i q_i q_i^T$, kde $\lambda_i \geq 0$ sú vlastné čísla matice X. Potom

$$tr(YX) = tr(Y \sum_{i=0}^{n-1} \lambda_i q_i q_i^T) = \sum_{i=0}^{n-1} \lambda_i q_i^T Y q_i \geq 0.$$

To znamená, že $Y \in (S^n_+)^*$, kde $(S^n_+)^*$ označuje duálny kužeľ ku kužeľu S^n_+.

Samoduálnosť oboch kužeľov sa ukázala podobne ako je to uvedené v ([8], kapitola 2), vlastnostiam duálneho kužeľa z nasledujúcej vety sme sa viac venovali v Dodatku - časť Kužele.

Veta 2.1. *Pre duálny kužeľ $\mathcal{K}^*$ platia nasledujúce vlastnosti:*

a) $\mathcal{K}^*$ je konvexný kužeľ.

b) Nech kužele $\mathcal{K}_1, \mathcal{K}_2$ spĺňajú $\mathcal{K}_1 \subseteq \mathcal{K}_2$. Potom pre ich duálne kužele platí $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$.

c) $\mathcal{K}^*$ je uzavretý.

d) Vnútro $\mathcal{K}^*$ je $int\mathcal{K}^* = \{z | z^T x > 0 \quad \forall x \in cl\mathcal{K} \setminus \{0\}\}$.

e) Ak $int\mathcal{K} \neq \emptyset$, potom $\mathcal{K}^*$ je "špicatý".

f) Nech K je konvexný kužeľ. Potom $\mathcal{K}^{**}$ je $cl\mathcal{K}$. Ak $\mathcal{K}$ je aj uzavretý, tak $\mathcal{K}^{**} = \mathcal{K}$.

g) Ak $cl\mathcal{K}$ je "špicatý", tak $int\mathcal{K}^* \neq \emptyset$.

Dôkaz. Dôkaz sa dá nájsť v Dodatku ako Veta(6.2). □

Tieto vlastnosti ukázali, že ak $\mathcal{K}$ je vlastný kužeľ, potom aj k nemu duálny kužeľ $\mathcal{K}^*$ je vlastný, a naviac $\mathcal{K}^{**} = \mathcal{K}$.
Preto podobne ako sme definovali zovšeobecnenú nerovnosť $\preceq_{\mathcal{K}}$ pomocou vlastného kužeľa $\mathcal{K}$, môžeme zadefinovať duálnu zovšeobecnenú nerovnosť $\preceq_{\mathcal{K}^*}$ pomocou duálneho vlastného kužeľa $\mathcal{K}^*$. Keďže $\mathcal{K}^{**} = \mathcal{K}$, tak duálna zovšeobecnená nerovnosť k $\preceq_{\mathcal{K}^*}$ je $\preceq_{\mathcal{K}}$.
Pre samoduálne kužele bude duálna zovšeobecnená nerovnosť $\preceq_{\mathcal{K}^*}$ ekvivalentná s $\preceq_{\mathcal{K}}$.

Tvrdenie 2.1. *Nech* $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$ *je kartézsky súčin kužeľov* $\mathcal{K}_1, \ldots, \mathcal{K}_n$. *Potom platia nasledujúce vlastnosti:*

a) $\mathcal{K}$ *je kužeľ.*

b) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ *konvexné, tak aj ich kartézsky súčin* $\mathcal{K}$ *je konvexný.*

c) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ *samoduálne, potom* $\mathcal{K}$ *je tiež samoduálny kužeľ.*

d) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ *vlastné, potom aj* $\mathcal{K}$ *je vlastný kužeľ.*

Dôkaz. Dôkaz sa dá nájsť v Dodatku ako Tvrdenie(6.2). □

2.2 Kužele druhého rádu

Teraz sa budeme venovať štandardnej triede kužeľov, tzv. kužeľom druhého rádu, ktoré sú základom pri SOCP úlohách.

Definícia 2.4. Kužeľ druhého rádu dimenzie n definujeme ako

$$\mathcal{Q}_n = \{x = (x_0; \bar{x}) \in R^n : x_0 \geq ||\bar{x}||\},$$

kde $\|.\|$ je štandardná Euklidovská norma, n je dimenzia kužeľa $\mathcal{Q}_n$,
$x_0 \in R$, $\bar{x} = (x_1, \ldots, x_{n-1})^T \in R^{n-1}$.

Niekedy sa kužeľ druhého rádu nazýva aj kvadratický kužeľ, pretože sa dá vyjadriť v tvare kvadratických nerovností:

$$Q_n = \left\{ \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \mid \begin{bmatrix} x_0 & \bar{x}^T \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \geq 0, x_0 \geq 0 \right\}^1.$$

Ak poznáme dimenziu n, tak budeme uvádzať označenie kužeľa $\mathcal{Q}$ bez indexu.
Pre $n = 1$ sa kužeľ druhého rádu zdegeneruje na polpriamku ležiacu na osi x_0 a začínajúcu v bode $x_0 = 0$, čo môžeme zapísať ako

$$\mathcal{Q}_1 = \{x_0 \in R : x_0 \geq 0\}.$$

V priestore R^2 a R^3 sú kužele druhého rádu zobrazené na obrázku 2.

Pre kužeľ $\mathcal{Q}$ označíme hranicu kužeľa $\mathcal{Q}$

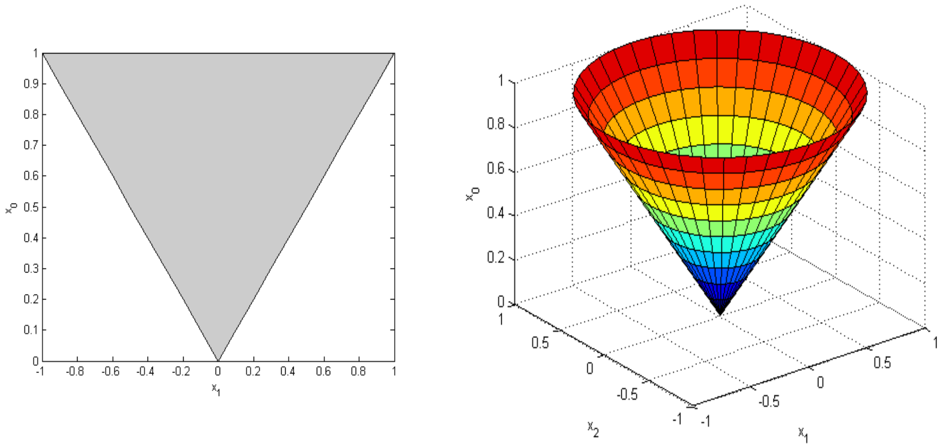
$$bd\mathcal{Q} = \{x \in \mathcal{Q} : x_0 = ||\bar{x}||\}$$

a vnútro kužeľa $\mathcal{Q}$

$$int\mathcal{Q} = \{x \in \mathcal{Q} : x_0 > ||\bar{x}||\}.$$

Potom pod pojmom uzáver kužeľa $\mathcal{Q}$ rozumieme množinu $cl\mathcal{Q} = bd\mathcal{Q} \cup int\mathcal{Q}$.

[1]Kužele 2. rádu sú známe pod viacerými názvami, v angličtine napr. Lorentz cone, ice-cream cone.



Obr. 2: Kužeľ druhého rádu pre $n = 2$ a kužeľ druhého rádu pre $n = 3$

Kužeľ druhého rádu Q_n, podobne ako nezáporný ortant R^n_+ a kladne semidefinitný kužeľ S^n_+, spĺňa vlastnosti vlastného kužeľa.

Tvrdenie 2.2. *Kužeľ druhého rádu Q_n je vlastný kužeľ.*

Dôkaz. Postupne overíme, či pre kužeľ druhého rádu Q_n sú splnené všetky vlastnosti vlastného kužeľa.

1. Najprv dokážeme uzavretosť Q_n. Teda stačí nám ukázať, že komplement $\{x \in R^n : x_0 < ||\bar{x}||\}$ je otvorená množina. Pre ľubovoľné x z tejto množiny platí $x_0 < ||\bar{x}||$ a je vidieť, že táto nerovnosť ostane zachovaná aj pre všetky x' z dostatočne malého okolia x. Takže komplement ku kužeľu $\mathcal{Q}$ je otvorená množina.

2. Ďalej ukážeme, že Q_n je konvexný kužeľ, teda že pre ľubovoľné $x, y \in Q_n$ a $\theta_1, \theta_2 \geq 0$ platí: $\theta_1 x + \theta_2 y \in Q_n$. To znamená, že má byť splnená nerovnosť $\theta_1 x_0 + \theta_2 y_0 \geq ||\theta_1 \bar{x} + \theta_2 \bar{y}||$, čo zrejme platí, pretože $||\theta_1 \bar{x} + \theta_2 \bar{y}|| \leq \theta_1 ||\bar{x}|| + \theta_2 ||\bar{y}||$ a $x_0 \geq ||\bar{x}||, y_0 \geq ||\bar{y}||$.

3. Vnútro Q_n je neprázdne, pretože $(x_0; 0) \in int\mathcal{Q}_n \quad \forall x_0 > 0$.

4. A nakoniec treba ešte overit špicatosť. Tá vyplýva z nasledujúcich ekvivalencii

$$x \in Q_n, -x \in Q_n \quad \Leftrightarrow \quad x_0 \geq ||\bar{x}|| \geq 0 \wedge -x_0 \geq ||\bar{-x}|| \geq 0 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad x_0 \geq 0 \wedge x_0 \leq 0 \quad \Leftrightarrow \quad (x_0 = 0 \Rightarrow ||\bar{x}|| = 0) \quad \Leftrightarrow \quad x = 0.$$

□

Ukázali sme, že kužeľ druhého rádu je vlastný kužeľ, preto môžeme zadefinovať nasledujúce pojmy.

Definícia 2.5. Pre kužeľ druhého rádu definujeme zovšeobecnenú nerovnosť $\preceq_{\mathcal{K}}$ ako:

$$x \preceq_{\mathcal{Q}} y \Leftrightarrow y - x \in \mathcal{Q}.$$

Analogicky definujeme ostrú nerovnosť

$$x \prec_{\mathcal{Q}} y \Leftrightarrow y - x \in int\mathcal{Q}.$$

Poznámka 2.1. V zmysle predchádzajúcej definície budeme pre $x \in \mathcal{Q}$ niekedy písať

$$x \succeq_{\mathcal{Q}} 0.$$

Ďalej dokážeme dôležitú vlastnosť kužeľov druhého rádu, tzv. samoduálnosť.

Tvrdenie 2.3. *Kužeľ druhého rádu $\mathcal{Q}$ je samoduálny.*

Dôkaz. 1.) Najskôr ukážeme, že $\mathcal{Q} \subseteq \mathcal{Q}^*$. Teda chceme ukázať, že $\forall x, z \in \mathcal{Q}$ platí $x^T z \geq 0$. Nech $x = (x_0; \bar{x}), z = (z_0; \bar{z})$ sú ľubovoľné, ale pevne zvolené body kužeľa $\mathcal{Q}$. Potom ak $\bar{x}^T \bar{z} \geq 0$, tak

$$x^T z = x_0 z_0 + \bar{x}^T \bar{z} \geq 0.$$

V opačnom prípade, teda pre $\bar{x}^T \bar{z} \leq 0$ platí $\bar{x}^T \bar{z} = -|\bar{x}^T \bar{z}|$.
Z Cauchyho-Schwarzovej nerovnosti vyplýva

$$|\bar{x}^T \bar{z}| \leq ||\bar{x}||.||\bar{z}|| \leq x_0 z_0.$$

A teda

$$x^T z = x_0 z_0 + \bar{x}^T \bar{z} = x_0 z_0 - |\bar{x}^T \bar{z}| \geq 0.$$

Ukázali sme, že $x^T z \geq 0$, čo znamená, že $z \in \mathcal{Q}^*$.
2.) Teraz ukážeme, že $\mathcal{Q}^* \subseteq \mathcal{Q}$. Predpokladajme, že $z \in \mathcal{Q}^*$, čiže $\forall x \in \mathcal{Q} : x^T z \geq 0$. Ukážeme, že $z \in \mathcal{Q}$ a teda, že $z_0 \geq ||\bar{z}||$. Položme $y_0 = ||\bar{z}||$ a $\bar{y} = -\bar{z}$. Potom, keďže $||\bar{y}|| = ||\bar{z}|| = y_0$, tak zrejme $y = (y_0; \bar{y}) \in \mathcal{Q}$ a z toho vyplýva

$$0 \leq y^T z = y_0 z_0 + \bar{y}^T \bar{z} = ||\bar{z}||z_0 - \bar{z}^T \bar{z} = ||\bar{z}||z_0 - ||\bar{z}||^2.$$

Vidíme, že pre z spĺňajúce $||\bar{z}|| \neq 0$ platí $z_0 \geq ||\bar{z}||$.
Ak $||\bar{z}|| = 0$, t.j. $z_1 = \cdots = z_{n-1} = 0$, tak stačí vziať vektor $y = (y_0, 0, \ldots, 0)^T \in \mathcal{Q}$, kde $y_0 > 0$. Platí $0 \leq y^T z = y_0 z_0$ a teda $z_0 \geq 0$. □

Dôsledok 2.1. *Nech $\mathcal{Q} = \mathcal{Q}_{n_1} \times \mathcal{Q}_{n_2} \times \cdots \times \mathcal{Q}_{n_r}$ je kartézsky súčin kužeľov 2.rádu. Potom $\mathcal{Q}$ je samoduálny.*

Dôkaz. Dôkaz vyplýva z Tvrdenia (2.1). □

Dôsledok 2.2.

a) Ak $\mathcal{Q}$ *je samoduálny kužeľ, tak pre všetky* $x, y \in \mathcal{Q}$ *platí* $x^T y \geq 0$.

b) Ak $\forall x \in \mathcal{Q}$ *platí* $x^T y \geq 0$*, potom* $y \in \mathcal{Q}$.

Dôsledok (2.2) bude veľmi užitočný v teórii duality pre úlohy SOCP, ktorú rozoberieme v nasledujúcej časti.

3 Formulácia SOCP úloh a dualita

Úlohu programovania nad kužeľmi druhého rádu, tzv. SOCP úlohu, môžeme zapísať v štandardnom tvare nasledovne

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{r} (c^i)^T x^i, \\
& \sum_{i=1}^{r} A^i x^i = b, \\
& x^i \succeq_{\mathcal{Q}_{n_i}} 0, \quad i = 1, \ldots, r,
\end{aligned}
\tag{7}
$$

kde $b = (b_1, \ldots, b_m)^T \in R^m, c^i = (c_0^i, \ldots, c_{n_i-1}^i)^T \in R^{n_i}$ a $A^i \in R^{m \times n_i}$ pre $i = 1, \ldots, r$ sú parametre úlohy a $x^i \in R^{n_i}, i = 1, \ldots, r$ sú premenné. Množiny

$$
\mathcal{Q}_{n_i} = \{x^i = (x_0^i; \bar{x}^i) \in R^{n_i} | x_0^i \geq \|\bar{x}^i\|\} \subseteq R^{n_i} \quad \forall i,
$$

sú kužele druhého rádu. Prevedieme úlohu (7) do skráteného tvaru, s ktorým budeme ďalej pracovať.
Nech $n = \sum_{i=1}^{r} n_i$ a nech $x = (x^1; x^2; \ldots; x^r) \in R^n$, kde $x^i = (x_0^i, x_1^i, \ldots, x_{n_i-1}^i)^T \in R^{n_i}$, $i = 1, \ldots, r$ a nech $\mathcal{Q} = \mathcal{Q}_{n_1} \times \mathcal{Q}_{n_2} \times \cdots \times \mathcal{Q}_{n_r}$ je kartézsky súčin kužeľov 2.rádu. Podľa Tvrdenia (2.1 d)) je $\mathcal{Q}$ vlastný kužeľ, a teda možno definovať nerovnosť $\succeq_{\mathcal{Q}}$, pričom

$$
x = (x^1; x^2; \ldots; x^r) \succeq_{\mathcal{Q}} 0 \qquad \Leftrightarrow \qquad x^i \succeq_{\mathcal{Q}_{n_i}} 0, \quad i = 1, \ldots, r.
$$

Označíme $A = [A^1, A^2, \ldots, A^r] \in R^{m \times n}$, kde $A^i \in R^{m \times n_i}$, a $c = (c^1; c^2; \ldots; c^r) \in R^n$, kde $c^i \in R^{n_i}$. Teraz sa dá úloha (7) zapísať nasledovne:

$$
\begin{aligned}
\min_x \quad & c^T x, \\
& Ax = b, \\
& x \succeq_{\mathcal{Q}} 0.
\end{aligned}
\tag{P}
$$

Úloha (P) sa nazýva primárna úloha s primárnou premennou x.
Pre úlohu (P) zavedieme nasledujúce označenia:

- množina prípustných riešení primárnej SOCP úlohy (P)

$$
\mathcal{P} = \{x \in R^n | Ax = b, x \succeq_{\mathcal{Q}} 0\},
$$

- množina ostro prípustných riešení, resp. množina vnútorných bodov primárnej SOCP úlohy (P)

$$
\mathcal{P}^0 = \{x^0 \in R^n | Ax^0 = b, x^0 \succ_{\mathcal{Q}} 0\} = \{x^0 \in \mathcal{P} | x^0 \succ_{\mathcal{Q}} 0\} = \{x^0 \in \mathcal{P} | x^0 \in int\mathcal{Q}\},
$$

- množina optimálnych riešení primárnej SOCP úlohy (P)

$$\mathcal{P}^* = \{x^* \in \mathcal{P} | c^T x^* \le c^T x \quad \forall x \in \mathcal{P}\},$$

- optimálna hodnota účelovej funkcie primárnej SOCP úlohy (P)

$$p^* = \inf_{x \in R^n} \{c^T x | x \in \mathcal{P}\}.$$

Rozšírením účelovej funkcie o ohraničenia v tvare $b - Ax = 0$ a $-x \preceq_{\mathcal{Q}} 0$ (pretože $x \in \mathcal{Q} \Leftrightarrow x^i \succeq_{\mathcal{Q}} 0$) definujeme Lagrangeovu funkciu $L : R^n \times R^m \times R^n \to R$ ako

$$\begin{aligned} L(x, y, s) &= c^T x + y^T (b - Ax) + s^T(-x) \\ &= y^T b + [c - A^T y - s]^T x, \end{aligned}$$

kde $y = (y_1, \ldots, y_m)^T \in R^m$ a $s = (s^1; s^2; \ldots; s^r) \in R^n$ sú Lagrangeove multiplikátory, pričom $s^i = (s_0^i, s_1^i, \ldots, s_{n_i - 1}^i)^T \in R^{n_i}$ pre všetky $i = 1, \ldots, r$.

Tvrdenie 3.1. *Nech úloha (P) má prípustné riešenie. Potom pre ľubovoľné $x \in \mathcal{P}$ a $s \in \mathcal{Q}$ platí*

$$L(x, y, s) \le c^T x.$$

Dôkaz. Nech $x \in \mathcal{P}$ a $s \in \mathcal{Q}$ sú ľubovoľne zvolené.
Z Dôsledku (2.1) vyplýva, že $x^T s \ge 0$, a teda $s^T(-x) \le 0$. Keďže $Ax = b$, tak dostaneme

$$L(x, y, s) = c^T x + s^T(-x) \le c^T x,$$

čo sme chceli ukázať. □

Duálna funkcia $g : R^m \times R^n \to R$ reprezentuje minimálnu hodnotu Lagrangeovej funkcie cez všetky x pre $y \in R^m, s \in R^n$, t.j.

$$g(y, s) = \inf_x L(x, y, s) = y^T b + \inf_x [c - A^T y - s]^T x.$$

Ak je Lagrangeova funkcia zdola neohraničená v premennej x, tak $g(y, s) = -\infty$. Lineárna funkcia je ohraničená zdola iba v prípade, že je identicky rovná 0, preto

$$g(y, s) = \begin{cases} b^T y & \text{ak } c - A^T y - s = 0, \\ -\infty & \text{inak.} \end{cases}$$

Naviac duálna funkcia g je konkávna v (y, s), pretože pre fixné x je Lagrangeova funkcia $L(x, y, s)$ lineárna v (y, s), a teda g je infimum triedy funkcií lineárnych v (y, s) ([8], kapitola 5).
Z Tvrdenia (3.1) vyplýva, že pre $\tilde{x} \in \mathcal{P}$ platí

$$L(\tilde{x}, y, s) \le c^T \tilde{x},$$

a teda zrejme

$$g(y,s) \le c^T\tilde{x} \quad \forall \tilde{x} \in \mathcal{P}. \tag{8}$$

Ak na pravej strane nerovnosti (8) vezmeme infimum cez $\tilde{x} \in \mathcal{P}$, dostaneme

$$g(y,s) \le p^*. \tag{9}$$

To znamená, že pre každé $y \in R^m$ a $s \in \mathcal{Q}$ je duálna funkcia g dolnou hranicou optimálnej hodnoty p^* úlohy (P). Nerovnosť (9) by bola splnená aj pre $g(y,s) = -\infty$, ale nemala by žiaden význam. Preto budeme uvažovať iba (y,s), pre ktoré platí $g(y,s) > -\infty$. Hľadanie takej najlepšej dolnej hranice vedie k úlohe

$$\begin{aligned} \max_{y,s} \quad & g(y,s), \\ & s \succeq_{\mathcal{Q}} 0. \end{aligned} \tag{10}$$

Dosadíme za duálnu funkciu pričom hodnotu $g(y,s) > -\infty$ zabezpečíme podmienkou $c - A^Ty - s = 0$. Takto dostaneme úlohu:

$$\begin{aligned} \max_{y,s} \quad & b^Ty, \\ & A^Ty + s = c, \\ & s \succeq_{\mathcal{Q}} 0. \end{aligned} \tag{D}$$

Najlepšie dolné ohraničenie optimálnej hodnoty p^* úlohy (P) získame vyriešením úlohy (D).

Úloha (D) sa nazýva duálnou úlohou k úlohe (P) s duálnou premennou $y \in R^m$ a doplnkovou premennou $s \in R^n$ [2].

Pre úlohu (D) zavedieme nasledujúce označenia:

- množina prípustných riešení duálnej SOCP úlohy (D)

$$\mathcal{D} = \{(y,s) \in R^m \times R^n | A^Ty + s = c, s \succeq_{\mathcal{Q}} 0\},$$

- množina ostro prípustných riešení, resp. množina vnútorných bodov duálnej SOCP úlohy (D)

$$\mathcal{D}^0 = \{(y^0,s^0) \in R^m \times R^n | A^Ty^0 + s^0 = c, s^0 \succ_{\mathcal{Q}} 0\} = \{(y^0,s^0) \in \mathcal{D} | s^0 \succ_{\mathcal{Q}} 0\},$$

- množina optimálnych riešení duálnej SOCP úlohy (D)

$$\mathcal{D}^* = \{(y^*,s^*) \in \mathcal{P} | b^Ty^* \ge b^Ty \quad \forall (y,s) \in \mathcal{D}\}.$$

[2]Premennú s nájdeme v angličtine pod názvom slack variable.

- optimálna hodnota účelovej funkcie duálnej SOCP úlohy (D)

$$d^* = \sup_{y \in R^m, s \in R^n} \{b^T y | (y,s) \in \mathcal{D}\}.$$

Poznámka 3.1. Z úlohy (D) sa dá odstrániť doplnková premenná s pomocou vyjadrenia z ohraničenia $A^T y + s = c$ a následným dosadením do zovšeobecnenej nerovnosti $\succeq_{\mathcal{Q}}$. Potom môžeme úlohu (D) prepísať do tvaru

$$\begin{aligned} \max_{y} \quad & b^T y \\ & c - A^T y \succeq_{\mathcal{Q}} 0. \end{aligned} \tag{Du}$$

Z predchádzajúcich úvah vyplýva nasledujúca veta.

Veta 3.1 (Slabá veta o dualite)**.** *Nech x je prípustné riešenie primárnej SOCP úlohy (P) a nech (y,s) je duálne prípustné riešenie úlohy (D), t.j. $x \in \mathcal{P}, (y,s) \in \mathcal{D}$.*
Potom

$$p^* \geq d^*.$$

Definícia 3.1. Nech $x \in \mathcal{P}$ a nech $(y,s) \in \mathcal{D}$. Potom výraz $c^T x - b^T y = s^T x$ nazývame duálnou medzerou a výraz $p^* - d^*$ nazývame optimálnou duálnou medzerou.

Dôsledok 3.1.

Ak je primárna úloha (P) zdola neohraničená ($p^* = -\infty$), potom duálna úloha (D) je neprípustná ($d^* := -\infty$), t.j. $\mathcal{D} = \emptyset$,

Ak je duálna úloha (D) zhora neohraničená ($d^* = \infty$), potom primárna úloha (P) je neprípustná ($p^* := \infty$), t.j. $\mathcal{P} = \emptyset$.

3.1 Silná veta o dualite

Skôr než dokážeme silnú vetu o dualite, budeme potrebovať Zovšeobecnenú Farkašovu lemu, resp. jej dôsledok. Preformulovali sme toto tvrdenie z [26] tak, aby sme ho mohli neskôr priamo aplikovať, a následne sme urobili dôkaz.

Tvrdenie 3.2 (Zovšeobecnená Farkašova lema)**.** *Nech $A \in R^{m \times n}, b \in R^m, c \in R^n$ a nech $\mathcal{Q} \subseteq R^n$ je kartézsky súčin kužeľov druhého rádu.*

a) Nech $\mathcal{K} = \{Ax : x \in \mathcal{Q}\}$ je uzavretý konvexný kužeľ. Potom platí práve jedna z nasledujúcich alternatív

   1. $\exists x \in R^n, x \succeq_{\mathcal{Q}} 0 : Ax = b$,

   2. $\exists y \in R^m : A^T y \succeq_{\mathcal{Q}} 0, b^T y < 0$.

b) Nech $\mathcal{K} = \{A^T y + s : s \in \mathcal{Q}, y \in R^m\}$ je uzavretý konvexný kužeľ[3]*. Potom platí práve jedna z nasledujúcich alternatív*

1. $\exists x \in R^n, x \preceq_{\mathcal{Q}} 0 : Ax = 0, c^T x > 0,$

2. $\exists (y,s) \in R^m \times R^n : A^T y + s = c, s \succeq_{\mathcal{Q}} 0.$

Dôkaz. a) Uvažujme dve možnosti, ktoré môžu nastať:

1. Nech $b \in \mathcal{K}$. To znamená, že existuje $x \succeq_{\mathcal{Q}} 0$, pre ktoré $Ax = b$, a teda platí prvá alternatíva.
2. Teraz nech $b \notin \mathcal{K}$. Pretože $\mathcal{K}$ je uzavretý konvexný kužeľ, môžeme použiť Tvrdenie (6.1), ktoré je dôsledkom Vety o oddeľujúcej nadrovine, podľa ktorého

$$\exists y \in R^m, y \neq 0 : \quad y^T b < 0 \quad \text{a} \quad y^T z \geq 0 \quad \forall z \in \mathcal{K}.$$

Všimnime si, že

$$\begin{aligned} y^T z \geq 0 \quad \forall z \in \mathcal{K} \quad &\Leftrightarrow \quad y^T (Az) \geq 0 \quad \forall z \in \mathcal{Q} \quad \Leftrightarrow \\ &\Leftrightarrow \quad (A^T y)^T z \geq 0 \quad \forall z \in \mathcal{Q}, \end{aligned}$$

a teda podľa Dôsledku (2.2 b)) je to ekvivalentné s tým, že $A^T y \in \mathcal{Q}$. Dostali sme, že

$$\exists y \in R^m : \quad y^T b < 0 \quad \text{a} \quad A^T y \succeq_{\mathcal{Q}} 0,$$

čiže platí druhá alternatíva.

Teraz ukážeme, že obe alternatívy nemôžu nastať súčasne. Predpokladajme, že je splnená prvá aj druhá alternatíva. To znamená, že existuje $\tilde{x} \succeq_{\mathcal{Q}} 0, \tilde{y} \in R^m$, pre ktoré platí

$$A\tilde{x} = b, b^T \tilde{y} < 0 \text{ a } A^T \tilde{y} \succeq_{\mathcal{Q}} 0.$$

Dosadíme za b a dostaneme

$$b^T \tilde{y} = (\tilde{x}^T A^T)\tilde{y} = \tilde{x}^T (A^T \tilde{y}) < 0,$$

čo je spor s predpokladom, pretože podľa Dôsledku (2.2 a)) $\tilde{x}^T (A^T \tilde{y}) \geq 0$. Preto platí práve jedna z alternatív $1., 2.$, čo sme chceli ukázať.

b) Podobný postup ako v časti a) zopakujeme aj tu. Uvažujme dve možnosti, ktoré môžu nastať:

[3] Z vlastnosti kužeľa je ľahké overiť, že množiny $\{Ax : x \in \mathcal{Q} \subseteq R^n\}$ a $\{A^T y + s : s \in \mathcal{Q} \subseteq R^n, y \in R^m\}$ sú konvexné kužele.

1. Nech $c \in \mathcal{K}$. To znamená, že existuje $(y, s) \in R^m \times R^n$, pre ktoré $A^T y + s = c, s \succeq_{\mathcal{Q}} 0$, čím je splnená prvá alternatíva.
2. Teraz nech $c \notin \mathcal{K}$. Pretože $\mathcal{K}$ je uzavretý konvexný kužeľ, môžeme opäť použiť Tvrdenie (6.1), podľa ktorého

$$\exists x \in R^n : \quad c^T x < 0 \quad \text{a} \quad x^T z \geq 0 \quad \forall z \in \mathcal{K}.$$

Všimnime si, že

$$\begin{aligned} x^T z \geq 0 \quad \forall z \in \mathcal{K} \quad &\Leftrightarrow \quad x^T(A^T y + s) \geq 0 \quad \forall s \succeq_{\mathcal{Q}} 0, \forall y \in R^m \quad \Leftrightarrow \\ &\Leftrightarrow \quad (Ax)^T y + s^T x \geq 0 \quad \forall s \in \mathcal{Q}, \forall y \in R^m, \end{aligned}$$

a teda podľa Dôsledku (2.2 b)) je to ekvivalentné s tým, že $(Ax = 0 \wedge x \in \mathcal{Q})$. Dostali sme, že

$$\exists x \in R^n : \quad c^T x < 0, Ax = 0 \quad \text{a} \quad x \succeq_{\mathcal{Q}} 0,$$

čo možno upraviť nasledovne

$$\exists x \in R^n : \quad c^T(-x) > 0, A(-x) = 0 \quad \text{a} \quad (-x) \preceq_{\mathcal{Q}} 0,$$

a teda pre $(-x)$ platí druhá alternatíva.

Teraz ukážeme, že obe alternatívy nemôžu nastať súčasne. Predpokladajme, že je splnená prvá aj druhá alternatíva. To znamená, že existuje $\tilde{x} \preceq_{\mathcal{Q}} 0, \tilde{y} \in R^m, \tilde{s} \succeq_{\mathcal{Q}} 0$ pre ktoré platí

$$A\tilde{x} = 0, c^T \tilde{x} > 0 \text{ a } A^T \tilde{y} + \tilde{s} = c.$$

Dosadíme za c a dostaneme

$$c^T \tilde{x} = (\tilde{y}^T A)\tilde{x} + \tilde{s}^T \tilde{x} = \tilde{y}^T (A\tilde{x}) + \tilde{s}^T \tilde{x} > 0.$$

Využijeme $A\tilde{x} = 0$. Potom platí $\tilde{s}^T \tilde{x} > 0$, čo je spor s predpokladom, pretože z Dôsledku (2.2 a)) vyplýva $\tilde{s}^T(-\tilde{x}) \geq 0$. Preto platí práve jedna z alternatív $1., 2.$, čo sme chceli ukázať.

□

Poznámka 3.2. Zovšeobecnená Farkašova lema je založená na uzavretosti kužeľa $\mathcal{K}$. Čo sa stane, keď tento predpoklad nebude splnený, ukážeme v nasledujúcom príklade. Viac o tomto probléme sa dá dozvedieť v literatúre [2] a [14].

Príklad 3.1. Uvažujme kužeľ druhého rádu $\mathcal{Q} = \{x = (x_0; x_1; x_2) : x_0 \geq \|(x_1, x_2)^T\|\}$.
Nech $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Potom

$$\mathcal{K} = \{Ax, x \in \mathcal{Q}\} = \left\{ \begin{pmatrix} x_0 - x_1 \\ x_2 \end{pmatrix}, x_0 \geq \sqrt{x_1^2 + x_2^2} \right\}.$$

Ak by platila prvá alternatíva, tak existuje $x \in \mathcal{Q}$ také, že

$$x_0 - x_1 = 0, x_2 = 1 \quad \Rightarrow \quad x_0 \geq \sqrt{x_0^2 + 1},$$

čo je spor, lebo $0 \not\geq 1$. Ak by platila druhá alternatíva, tak existuje $y = (y_1, y_2)^T$, pre ktoré $A^T y \in \mathcal{Q}$ a $b^T y < 0$. Z toho vyplýva, že

$$b^T y = y_2 < 0, \quad A^T y = \begin{pmatrix} y_1 \\ -y_1 \\ y_2 \end{pmatrix} \in \mathcal{Q} \quad \Rightarrow \quad y_1 \geq \sqrt{(-y_1)^2 + y_2^2},$$

čo je ekvivalentné s $0 \geq |y_2|$, a teda opäť dostaneme spor.
Ani jedna z alternatív zovšeobecnenej Farkašovej lemy neplatí, pozrieme sa preto bližšie na kužeľ $\mathcal{K}$. Aby bolo splnené ohraničenie kužeľa druhého rádu, musí platiť $x_0^2 - x_1^2 \geq x_2^2$ pre $x_0 \geq 0$. Rozoberme nasledujúce prípady:

- Ak $x_0 - x_1 = 0$, tak nutne $x_2 = 0$.
- Možnosť $x_0 - x_1 < 0$ nespĺňa ohraničenie, lebo $0 \leq x_0 < x_1$ implikuje $x_0^2 - x_1^2 < 0$.
- Ak $x_0 - x_1 > 0$, tak ohraničenie je splnené pre $x_0 > |x_1|$. Potom $x_2 \in R$ je ľubovoľne zvolené tak, aby platilo ohraničenie.

Takže kužeľ $\mathcal{K}$ možno vyjadriť aj ako $\{(0;0)\} \cup \{(\tilde{x}_1; \tilde{x}_2) : \tilde{x}_1 > 0, \tilde{x}_2 \in R\}$. Je zrejme, že to nie je uzavretá množina. Preto bol porušený predpoklad zovšeobecnenej Farkašovej lemy a ani jedna alternatíva neplatila.

V nasledujúcej leme uvedieme jednoduchú podmienku, ktorá zabezpečuje, že predpoklad uzavretosti zo Zovšeobecnenej Farkašovej lemy (Tvrdenie 3.2) je splnený.

Lema 3.1. *Nech $A \in R^{m \times n}$ a nech $\mathcal{Q} \subseteq R^n$ je kartézsky súčin kužeľov druhého rádu.*

a) Nech $\exists \tilde{y} \in R^m : A^T \tilde{y} \succ_{\mathcal{Q}} 0$. Potom množina $\mathcal{K} = \{Ax : x \in \mathcal{Q}\}$ je uzavretá.

b) Nech $\exists \tilde{x} \in R^n : A\tilde{x} = 0, \tilde{x} \prec_{\mathcal{Q}} 0$. Potom množina $\mathcal{K} = \{A^T y + s : s \in \mathcal{Q}, y \in R^m\}$ je uzavretá.

Tvrdenie 3.3 (Dôsledok Zovšeobecnenej Farkašovej lemy)**.** *Nech $A \in R^{m \times n}, b \in R^m, c \in R^n$ a nech $\mathcal{Q} \subseteq R^n$ je kartézsky súčin kužeľov druhého rádu.*

a) Nech existuje taký vektor $\tilde{y} \in R^m$, pre ktorý platí $A^T \tilde{y} \succ_{\mathcal{Q}} 0$. Potom platí práve jedna z nasledujúcich alternatív

1. $\exists x \in R^n, x \succeq_{\mathcal{Q}} 0 : Ax = b$,

2. $\exists y \in R^m : A^T y \succeq_{\mathcal{Q}} 0$ a $b^T y < 0$.

b) Nech existuje $\tilde{x} \in R^n$ spĺňajúce podmienky $A\tilde{x} = 0, \tilde{x} \prec_{\mathcal{Q}} 0$. Potom platí práve jedna z nasledujúcich alternatív

1. $\exists x \in R^n, x \preceq_{\mathcal{Q}} 0 : Ax = 0, c^T x > 0$, *2.* $\exists (y,s) \in R^m \times R^n : A^T y + s = c, s \succeq_Q 0.$

Teraz už môžeme prejsť k dôkazu Silnej vety o dualite. Využijeme predchádzajúce tvrdenia a analógiu so semidefinitným programovaním.

Veta 3.2 (Silná veta o dualite)**.** *Uvažujme primárnu a duálnu SOCP úlohu*

$$\begin{aligned} \min_x \quad & c^T x, \\ & Ax = b, \qquad\qquad \text{(P)} \\ & x \succeq_Q 0. \end{aligned} \qquad\qquad \begin{aligned} \max_{y,s} \quad & b^T y, \\ & A^T y + s = c, \qquad\qquad \text{(D)} \\ & s \succeq_Q 0. \end{aligned}$$

a) Nech $\mathcal{P} \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$. Potom primárna úloha (P) má optimálne riešenie a pre optimálne hodnoty účelových funkcii platí $p^* = d^*$.

b) Nech $\mathcal{D} \neq \emptyset$ a $\mathcal{P}^0 \neq \emptyset$. Potom duálna úloha (D) má optimálne riešenie a pre optimálne hodnoty účelových funkcii platí $p^* = d^*$.

c) Nech $\mathcal{P}^0 \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$. Potom primárna úloha (P) a duálna úloha (D) majú optimálne riešenia a pre optimálne hodnoty účelových funkcii platí $p^* = d^*$.

Dôkaz Silnej vety o dualite.

a) Nech $\mathcal{P} \neq \emptyset$ a $\mathcal{D}^0 \neq \emptyset$. Predpokladajme, že nasledujúci systém

$$\{c^T x = d^*, Ax = b, x \succeq_Q 0\} \tag{11}$$

nemá riešenie.
Označme $\tilde{A} = \begin{bmatrix} c^T \\ A \end{bmatrix}, \tilde{b} = \begin{pmatrix} d^* \\ b \end{pmatrix}$ a systém (11) sa dá prepísať do tvaru

$$\{\tilde{A}x = \tilde{b}, x \succeq_Q 0\}. \tag{12}$$

Z predpokladu $\mathcal{D}^0 \neq \emptyset$ vyplýva, že

$$\exists (y^0, s^0) \in R^{m \times n} : A^T y^0 + s^0 = c, s^0 \succ_Q 0,$$

čo je ekvivalentné s

$$\exists y^0 \in R^n : c - A^T y^0 \succ_Q 0. \tag{13}$$

Nech $\tilde{z} = (1; -y^0)$, potom nerovnosť (13) možno vyjadriť ako $\tilde{A}^T \tilde{z} \succ_Q 0$.
Teda vektor $\tilde{z}$ spĺňa predpoklad Tvrdenia (3.3a), a preto ho teraz aplikujeme. Keďže je systém (12) neprípustný, tak musí existovať vektor z, pre ktorý $\tilde{A}^T z \succeq_Q 0$ a $\tilde{b}^T z < 0$, t.j.

$$\exists z = (z_0; \bar{z}) \in R^{m+1} : \quad z_0 c + A^T \bar{z} \succeq_Q 0 \quad \text{a} \quad d^* z_0 + b^T \bar{z} < 0. \tag{14}$$

Budeme uvažovať nasledujúce 3 možnosti:

1. Nech $z_0 = 0$. Potom zo (14) vyplýva $A^T\bar{z} \succeq_{\mathcal{Q}} 0$ a $b^T\bar{z} < 0$. Využijeme nerovnosť (13), vzťah $\alpha A^T\bar{z} \succeq_{\mathcal{Q}} 0$, kde $\alpha \geq 0$ je ľubovoľne zvolený skalár, a dostaneme

$$c - A^T(y^0 - \alpha\bar{z}) \succ_{\mathcal{Q}} 0.$$

To znamená, že vektor $(y^0 - \alpha\bar{z})$ je prípustným riešením duálnej úlohy a hodnota účelovej funkcie je $b^T(y^0 - \alpha\bar{z}) = b^Ty^0 - \alpha b^T\bar{z} \leq d^*$. Keďže $b^T\bar{z} < 0$, tak pre $\alpha \to \infty$ ide hodnota $b^T(y^0 - \alpha\bar{z})$ do nekonečna, a teda aj $d^* \to \infty$. To je však spor s predpokladom $\mathcal{P} \neq \emptyset$.

2. Nech $z_0 > 0$. Potom predelíme (14) skalárom z_0 a dostaneme

$$c + A^T\left(\frac{\bar{z}}{z_0}\right) \succeq_{\mathcal{Q}} 0 \quad \text{a} \quad d^* + b^T\left(\frac{\bar{z}}{z_0}\right) < 0.$$

Úpravou získame

$$c - A^T\left(-\frac{\bar{z}}{z_0}\right) \succeq_{\mathcal{Q}} 0 \quad \text{a} \quad d^* - b^T\left(-\frac{\bar{z}}{z_0}\right) < 0.$$

Z prvej nerovnosti vyplýva, že vektor $\left(-\frac{\bar{z}}{z_0}\right)$ je prípustným riešením duálnej úlohy a z druhej nerovnosti $d^* < b^T\left(-\frac{\bar{z}}{z_0}\right)$, čo je spor s predpokladom, že d^* je optimálna hodnota účelovej funkcie duálnej úlohy.

3. Nech $z_0 < 0$. Opäť predelíme (14) skalárom z_0 a dostaneme

$$c + A^T\left(\frac{\bar{z}}{z_0}\right) \preceq_{\mathcal{Q}} 0 \quad \text{a} \quad d^* + b^T\left(\frac{\bar{z}}{z_0}\right) > 0. \tag{15}$$

Ostrá nerovnosť v (15) sa zachová, aj keď zmenšíme hodnotu $d^* + b^T\left(\frac{\bar{z}}{z_0}\right)$ o malé $\varepsilon_1 > 0$, teda

$$\varepsilon_1 > 0 : d^* + b^T\left(\frac{\bar{z}}{z_0}\right) - \varepsilon_1 > 0.$$

Na druhej strane pre

$$\varepsilon_2 : 0 < \varepsilon_2 < \varepsilon_1$$

musí podľa definície d^* ako suprema existovať také $\tilde{y}$, pre ktoré platí

$$d^* - \varepsilon_2 < b^T\tilde{y} \quad \text{a} \quad c - A^T\tilde{y} \succeq_{\mathcal{Q}} 0. \tag{16}$$

Sčítaním upravenej (15) a (16) dostaneme

$$A^T\left(-\tilde{y} - \frac{\bar{z}}{z_0}\right) \succeq_{\mathcal{Q}} 0 \quad \text{a} \quad b^T\left(-\tilde{y} - \frac{\bar{z}}{z_0}\right) < \varepsilon_2 - \varepsilon_1 < 0,$$

čo vedie k rovnakému problému ako sme riešili v prípade $z_0 = 0$. Podobným postupom by sme sa tiež dopracovali k sporu.

Všetky 3 prípady ukázali, že predpoklad vedie k sporu, a teda musí existovať riešenie $x^* \in R^n$ systému (11), t.j.

$$\{c^T x^* = d^*, Ax^* = b, x^* \succeq_{\mathcal{Q}} 0\}.$$

Vidíme, že $x^* \in \mathcal{P}$. Zo Slabej vety o dualite (Veta 3.1) vyplýva $d^* \leq p^*$, takže x^* je optimálnym riešením úlohy (P), a naviac $p^* = d^*$, čo sme chceli ukázať.

b) Budeme postupovať analogicky ako v prípade a).
Nech $\mathcal{D} \neq \emptyset$ a $\mathcal{P}^0 \neq \emptyset$. Predpokladajme, že nasledujúci systém

$$\{b^T y = p^*, A^T y + s = c, s \succeq_{\mathcal{Q}} 0\} \tag{17}$$

nemá riešenie.
Označme $\tilde{A} = \begin{bmatrix} b & A \end{bmatrix}, \tilde{c} = \begin{pmatrix} p^* \\ c \end{pmatrix}$ a systém (17) sa dá prepísať do tvaru

$$\{\tilde{A}^T y + \tilde{s} = \tilde{c}, \tilde{s} \succeq_{\tilde{\mathcal{Q}}} 0\}, \tag{18}$$

kde $\tilde{s} = (0; s) \in R^{n+1}$ a $\tilde{\mathcal{Q}} = (-\mathcal{Q}_1) \times \mathcal{Q}$ je kartézsky súčin kužeľa druhého rádu dimenzie 1 a pôvodného kužeľa $\mathcal{Q}$.
Z predpokladu $\mathcal{P}^0 \neq \emptyset$ vyplýva, že

$$\exists x^0 \in R^n : Ax^0 = b, x^0 \succ_{\mathcal{Q}} 0. \tag{19}$$

Nech $\tilde{z} = (1; -x^0)$, potom (19) možno vyjadriť ako $\tilde{A}^T \tilde{z} = 0, \tilde{z} \prec_{\tilde{\mathcal{Q}}} 0$.
Teda vektor $\tilde{z}$ spĺňa predpoklad Tvrdenia (3.3b), a preto ho teraz aplikujeme. Keďže je systém (18) neprípustný, tak musí existovať vektor z, pre ktorý $\tilde{A}^T z = 0, z \preceq_{\tilde{\mathcal{Q}}} 0$ a $\tilde{c}^T z > 0$, t.j.

$$\exists z = (z_0; \bar{z}) \in R^{n+1} : \quad z_0 b + A\bar{z} = 0, \quad z_0 \geq 0, \bar{z} \preceq_{\mathcal{Q}} 0 \quad \text{a} \quad p^* z_0 + c^T \bar{z} > 0. \tag{20}$$

Budeme uvažovať nasledujúce 2 možnosti, ktoré môžu nastať:

1. Nech $z_0 = 0$. Potom z (20) vyplýva $A\bar{z} = 0, \bar{z} \preceq_{\mathcal{Q}} 0$ a $c^T \bar{z} > 0$.
   Využijeme (19) a vzťahy $\alpha A\bar{z} = 0, \alpha\bar{z} \preceq_{\mathcal{Q}} 0$, kde $\alpha \geq 0$ je ľubovoľne zvolený skalár, a dostaneme

   $$A(x^0 - \alpha\bar{z}) = b \quad \text{a} \quad (x^0 - \alpha\bar{z}) \succ_{\mathcal{Q}} 0.$$

   To znamená, že vektor $(x^0 - \alpha\bar{z})$ je prípustným riešením primárnej úlohy a hodnota účelovej funkcie je $c^T(x^0 - \alpha\bar{z}) = c^T x^0 - \alpha c^T \bar{z} \geq p^*$. Keďže $c^T \bar{z} > 0$, tak pre $\alpha \to \infty$ ide hodnota $c^T(x^0 - \alpha\bar{z})$ k hodnote $-\infty$, a teda aj $p^* \to -\infty$. To je však spor s predpokladom $\mathcal{D} \neq \emptyset$.

2. Nech $z_0 > 0$. Potom predelíme (20) skalárom z_0 a dostaneme

$$b + A\left(\frac{\bar{z}}{z_0}\right) = 0, \quad \left(\frac{\bar{z}}{z_0}\right) \preceq_{\mathcal{Q}} 0 \quad \text{a} \quad p^* + c^T\left(\frac{\bar{z}}{z_0}\right) > 0.$$

Úpravou získame

$$b - A\left(-\frac{\bar{z}}{z_0}\right) = 0, \quad \left(-\frac{\bar{z}}{z_0}\right) \succeq_{\mathcal{Q}} 0 \quad \text{a} \quad p^* - c^T\left(-\frac{\bar{z}}{z_0}\right) > 0.$$

Z toho vyplýva, že vektor $\left(-\frac{\bar{z}}{z_0}\right)$ je prípustným riešením primárnej úlohy, ale z poslednej nerovnosti $p^* > c^T\left(-\frac{\bar{z}}{z_0}\right)$, čo je v spore s predpokladom, že p^* je optimálna hodnota účelovej funkcie primárnej úlohy.

Oba prípady ukázali, že predpoklad vedie k sporu, a teda musí existovať riešenie $(y^*, s^*) \in R^m \times R^n$ systému (17), t.j.

$$\{b^T y^* = p^*, A^T y^* + s^* = c, s^* \succeq_{\mathcal{Q}} 0\}.$$

Vidíme, že $(y^*, s^*) \in \mathcal{D}$. Zo Slabej vety o dualite (Veta 3.1) vyplýva $d^* \leq p^*$, takže (y^*, s^*) je optimálnym riešením úlohy (D), a naviac $d^* = p^*$, čo sme chceli ukázať.

c) Tento bod je dôsledkom bodov a), b).
Z predpokladov $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$ vyplýva, že $\mathcal{P} \neq \emptyset, \mathcal{D} \neq \emptyset$ a existujú vnútorné body, pre ktoré platí:

podľa a) $\quad \exists x^0 \in \mathcal{P}^0 : d^* = p^* \leq c^T x^0 < \infty,$

podľa b) $\quad \exists (y^0, s^0) \in \mathcal{D}^0 : -\infty < b^T y^0 \leq d^* = p^*.$

Spojením dostaneme, že existujú optimálne riešenia primárnej aj duálnej úlohy (P),(D) a optimálne hodnoty ich účelových funkcii sa rovnajú, čo sme chceli dokázať.

□

Nasledujúci príklad ukazuje, že ak sú splnené predpoklady tvrdenia b) Silnej vety o dualite, optimálne riešenie úlohy (P) sa nutne nemusí nadobúdať.

Príklad 3.2. Skúmajme primárnu a duálnu úlohu

$$\begin{aligned} \min_{x_0,x_1,x_2} \quad & x_0 + x_2 \\ & x_1 = 1 \\ & x_0 \geq \sqrt{x_1^2 + x_2^2}. \end{aligned} \tag{21}$$

$$\begin{aligned} \max_{y,s_0,s_1,s_2} \quad & y \\ & s_0 = 1 \\ & y + s_1 = 0 \\ & s_2 = 1 \\ & s_0 \geq \sqrt{s_1^2 + s_2^2}. \end{aligned} \tag{22}$$

Ide o SOCP úlohy v tvare (P),(D), kde $A = (0,1,0)^T, b = 1, c = (1,0,1)^T$ a $x_0 \geq \sqrt{x_1^2 + x_2^2}$, $s_0 \geq \sqrt{s_1^2 + s_2^2}$ sú ohraničenia kužeľa druhého rádu.
Je zrejme, že optimálne riešenie úlohy (22) je

$$y^* = 0, s_0^* = 1, s_1^* = 0, s_2^* = 1,$$

a zároveň je to jej jediné prípustné riešenie. Optimálna hodnota účelovej funkcie je $d^* = 0$.
Ďalej vidíme, že $x_1^* = 1$. Potom musí platiť

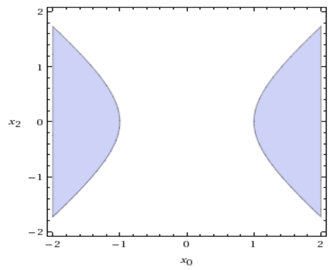
$$x_0 \geq \sqrt{1 + x_2^2} \quad \Leftrightarrow \quad (x_0 + x_2)(x_0 - x_2) \geq 1, \quad x_0 \geq 0 \tag{23}$$

aby bolo splnené ohraničenie kužeľa druhého rádu v úlohe (21).

Nerovnosť (23) predstavuje hyperbolu zobrazenú na obrázku. Aby sme minimalizovali $x_0 + x_2$, tak optimálne riešenie sa bude nachádzať v oblasti $x_0 > 1 \wedge x_2 < 0$. Ďalej z (23) vyplýva, že

$$p^* = \inf_{x_0, x_1, x_2} x_0 + x_2 = 0,$$

ale je zrejme, že sa táto hodnota nikdy nenadobudne, t.j. $x_0 + x_2 \neq 0$.



Optimálne hodnoty účelových funkcií sa síce rovnajú, t.j. $p^* = d^*$, ale v úlohe (21) nevieme nájsť optimálne riešenie s optimálnou hodnotou p^*.

V ďalšom príklade si ukážeme, že v úlohách programovania nad kužeľmi druhého rádu môže nastať prípad, kedy jedna z úloh je neprípustná a druhá úloha má optimálne riešenie.

Príklad 3.3. Skúmajme primárnu a duálnu úlohu

$$\begin{aligned} &\min_{x_0,x_1,x_2} \quad 0 \\ &x_0 + x_2 = 0 \\ &x_1 = 1 \\ &x_0 \geq \sqrt{x_1^2 + x_2^2}. \end{aligned} \tag{24}$$

$$\begin{aligned} &\max_{y,s_0,s_1,s_2} \quad y_2 \\ &y_1 + s_0 = 0 \\ &y_2 + s_1 = 0 \\ &y_1 + s_2 = 0 \\ &s_0 \geq \sqrt{s_1^2 + s_2^2}. \end{aligned} \tag{25}$$

Ide o SOCP úlohy v tvare (P),(D), kde

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

a $x_0 \geq \sqrt{x_1^2 + x_2^2}, s_0 \geq \sqrt{s_1^2 + s_2^2}$ sú ohraničenia kužeľa druhého rádu.
Z ohraničení úlohy (25) vyplýva, že $s_0 = -y_1, s_1 = -y_2, s_2 = -y_1$, a teda

$$-y_1 \geq \sqrt{y_2^2 + y_1^2} \quad \Leftrightarrow \quad y_1^2 \geq y_2^2 + y_1^2, \quad y_1 \leq 0 \quad \Leftrightarrow \quad 0 \geq y_2^2.$$

Teda pre $y \in \mathcal{D}$ nutne musí platiť $y_2 = 0$. Preto $d^* = 0$.
Z ohraničení úlohy (24) dostaneme

$$x_0 \geq \sqrt{1^2 + (-x_0)^2} \quad \Leftrightarrow \quad x_0^2 \geq 1 + x_0^2, \quad x_0 \geq 0 \quad \Leftrightarrow \quad 0 \geq 1,$$

čo znamená, že úloha (24) je neprípustná, t.j. $p^* = \infty$. Ukázalo sa, že aj napriek tomu, že je duálna úloha (25) ohraničená a má optimálne riešenie, tak primárna úloha (24) je neprípustná.

Poznámka 3.3. V lineárnom programovaní silná veta o dualite vyžadovala buď prípustné riešenia v primárnej aj duálnej úlohe, alebo prípustné riešenie v primárnej (resp. duálnej) úlohe s ohraničenou účelovou funkciou zdola (resp. zhora). Tieto podmienky zaručili existenciu optimálnych riešení a rovnosť optimálnych hodnôt účelových funkcii, t.j. $p^* = d^*$. Predchádzajúce príklady ukázali, že v úlohách programovania nad kužeľmi druhého rádu je potrebná silnejšia podmienka, ako sme aj ukázali vo Vete (3.2), inak nie je zaručená silná dualita.

3.2 Komplementarita a podmienky optimality

Zo Silnej vety o dualite (3.2) vieme, že ak existujú vnútorné body úloh (P) a (D), tak v optimálnom riešení $x^*, (y^*, s^*)$ platí $p^* = d^*$. To znamená, že pre optimálnu duálnu medzeru dostaneme

$$p^* - d^* = (x^*)^T s^* = 0. \tag{26}$$

Lema 3.2. *Nech* $x = (x_0; \bar{x}) \in R^p, s = (s_0; \bar{s}) \in R^p$ *a nech* $x \in \mathcal{Q}, s \in \mathcal{Q}$, *kde* $\mathcal{Q} \subseteq R^p$ *je kužeľ druhého rádu. Potom*

$$x^T s = 0 \qquad \Leftrightarrow \qquad x_0 \bar{s} + s_0 \bar{x} = 0.$$

Dôkaz. Nech $x = (x_0; \bar{x}) \in \mathcal{Q}, s = (s_0; \bar{s}) \in \mathcal{Q}$. Potom

$$x_0^2 \geq \|\bar{x}\|^2, \quad s_0^2 \geq \|\bar{s}\|^2. \tag{27}$$

Ak $x_0 = 0$, tak $\bar{x} = 0$. Podobne ak $s_0 = 0$, tak $\bar{s} = 0$. V takom prípade je dôkaz triviálny, preto uvažujme prípad, keď $x_0 > 0, s_0 > 0$. Urobíme sled ekvivalentných úprav. Z nerovnosti (27) predelením získame

$$1 \geq \frac{\|\bar{s}\|^2}{s_0^2}$$

a prenásobením výrazom x_0^2 dostaneme

$$x_0^2 \geq \frac{\|\bar{s}\|^2 x_0^2}{s_0^2}. \tag{28}$$

Teraz nech platí $x^T s = 0$. Vyjadríme

$$-x_0 s_0 = \bar{x}^T \bar{s}$$

a úpravou dostaneme

$$-2x_0^2 = \frac{2\bar{x}^T \bar{s} x_0}{s_0}. \tag{29}$$

Spojením (27) a (28) vznikla nerovnosť

$$0 \geq \|\bar{x}\|^2 - x_0^2 + \frac{\|\bar{s}\|^2 x_0^2}{s_0^2} - x_0^2 \overset{(29)}{=}$$
$$\overset{(29)}{=} \|\bar{x}\|^2 + \frac{\|\bar{s}\|^2 x_0^2}{s_0^2} + \frac{2\bar{x}^T \bar{s} x_0}{s_0} = \left(\bar{x} + \frac{\bar{s}x_0}{s_0}\right)^T \left(\bar{x} + \frac{\bar{s}x_0}{s_0}\right) = \left\|\bar{x} + \tfrac{\bar{s}x_0}{s_0}\right\|^2.$$

A teda platí

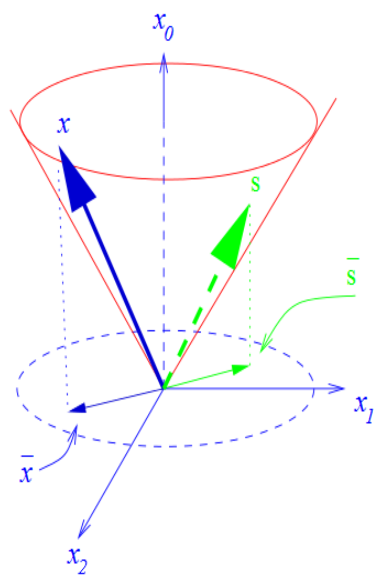
$$\bar{x} + \frac{\bar{s}x_0}{s_0} = 0, \tag{30}$$

čo je to isté ako $x_0\bar{s} + s_0\bar{x} = 0$. □

Podmienka z lemy (3.2)

$$x_0\bar{s} + s_0\bar{x} = 0. \tag{31}$$

sa nazýva *podmienkou komplementarity* a geometricky ju možno interpretovať nasledovne:
$x \in \mathcal{Q}$ a $s \in \mathcal{Q}$ sú ortogonálne práve vtedy, keď x alebo s je nulové, alebo obe ležia na hranici kužeľa $\mathcal{Q}$ takým spôsobom, že ich projekcie na rovinu $x_1, \ldots, x_{p-1}$ sú kolineárne, t.j. ležiace na jednej priamke, ako je znázornené na nasledujúcom obrázku ([3]).



Zavedieme označenie $\circ : R^p \times R^p \to R^p$

$$x \circ s = \begin{pmatrix} x^T s \\ x_0\bar{s} + s_0\bar{x} \end{pmatrix},$$

kde $x = (x_0; \bar{x}) \in R^p$ a $s = (s_0; \bar{s}) \in R^p$ (Pozri Dodatok - operátor $\circ$).

Poznámka 3.4. Všimnime si, že

$$x \circ s \quad \Leftrightarrow \quad x^T s = 0 \quad \Leftrightarrow \quad x_0\bar{s} + s_0\bar{x} = 0.$$

Pre $x = (x^1; \ldots; x^r) \in R^n, s = (s^1; \ldots; s^r) \in R^n$, kde $x^i, s^i \in R^{n_i}$, definujeme

$$x \bullet s = (x^1 \circ s^1, \ldots, x^r \circ s^r).$$

Veta 3.3 (Podmienky optimality)**.** *Nech úlohy (P) a (D) majú ostro prípustné riešenie, t.j. $\mathcal{P} \neq \emptyset, \mathcal{D} \neq \emptyset$. Potom $x^* \in R^n$ a $(y^*, s^*) \in R^m \times R^n$ sú optimálne riešenia v (P) a (D) práve vtedy, keď spĺňajú nasledujúce podmienky*

$$\begin{aligned} Ax = b, &\quad x \in \mathcal{Q}, \\ A^T y + s = c, &\quad s \in \mathcal{Q}, \\ x \bullet s = 0. & \end{aligned} \tag{32}$$

Dôkaz. Predpokladajme, že $\mathcal{P} \neq \emptyset, \mathcal{D} \neq \emptyset$.
Je zrejme, že prvá podmienka

$$Ax = b, \quad x \in \mathcal{Q}$$

zaručuje primárnu prípustnosť, t.j. $x \in \mathcal{P}$. A tiež ju bude spĺňať ľubovoľné optimálne riešenie úlohy (P). Podobne z druhej podmienky

$$A^T y + s = c, \quad s \in \mathcal{Q}$$

vyplýva $(y, s) \in \mathcal{D}$, takže podmienka bude platiť aj pre $(y^*, s^*) \in \mathcal{D}^*$.
Stačí ukázať, že

$$p^* = d^* \quad \Leftrightarrow \quad x^* \bullet s^* = 0.$$

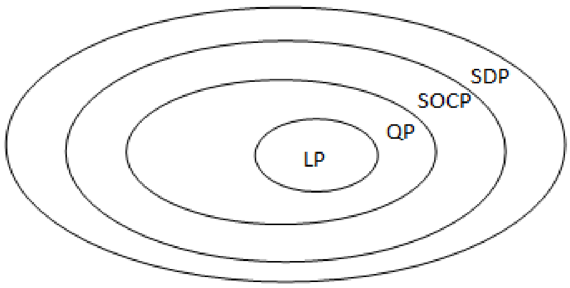
Zo silnej vety o dualite (3.2 c)) vyplýva, že

$$p^* = d^* \quad \Leftrightarrow \quad p^* - d^* = 0 \quad \overset{(26)}{\Leftrightarrow} \quad (x^*)^T s^* = 0 \quad \overset{\text{Lema (3.2)}}{\Leftrightarrow} \quad x^* \bullet s^* = 0,$$

čo sme chceli ukázať. $\square$

4 SOCP úlohy a iné triedy konvexnej optimalizácie

V tejto časti ukážeme, že úlohy lineárneho a kvadratického programovania sa vždy dajú previesť na SOCP úlohu, ale na druhej strane je programovanie nad kužeľmi druhého rádu podtriedou semidefinitného programovania. Vzťahy medzi triedami konvexnej optimalizácie, ktoré bližšie uvedieme v tejto kapitole, sú graficky zobrazené na nasledujúcom obrázku.



Obr. 3: Vzťahy medzi LP, QP, SOCP a SDP úlohami

4.1 Úloha lineárneho programovania

Uvažujme štandardný tvar úlohy lineárneho programovania

$$\begin{aligned} \min_x \quad & c^T x, \\ & Ax = b, \\ & x \succeq 0, \end{aligned} \tag{LP}$$

kde $x \in R^n, c \in R^n, A \in R^{m \times n}$ a $b \in R^m$. Vidíme,že ide o minimalizáciu lineárnej funkcie cez prienik afinneho podpriestoru s nezáporným ortantom. Rozdiel oproti SOCP úlohe (P) je práve v ohraničení $x \succeq 0$, ktoré reprezentuje nerovnosti po zložkách, t.j. $x_i \geq 0$ pre všetky $i = 1, \dots, n$. Avšak nerovnosť $x_i \geq 0$ je ohraničením kužeľa druhého rádu dimenzie 1. To znamená, že ohraničenie $x \succeq 0$ je ekvivalentné ohraničeniu $x \succeq_{\mathcal{Q}} 0$, kde $\mathcal{Q}$ je kartézsky súčin kužeľov druhého rádu dimenzie 1, a teda každú úlohu (LP) možno transformovať na SOCP úlohu (P). Úlohy lineárneho programovania sú špeciálnym prípadom programovania nad kužeľmi druhého rádu.

4.2 Úlohy kvadratického programovania

Úloha kvadratického programovania s lineárnymi ohraničeniami je definovaná v tvare

$$\begin{aligned} \min_x \quad & \frac{1}{2} x^T P x + q^T x + r, \\ & Ax = b, \\ & Cx \leq d, \end{aligned} \tag{LCQP}$$

kde P je symetrická, kladne semidefinitná matica, t.j. $P \in S^n_+, x \in R^n, q \in R^n, r \in R$, $A \in R^{m\times n}$, $b \in R^m, C \in R^{p\times n}$ a $d \in R^p$.

Úloha LCQP sa dá aj viac zovšeobecniť, keď okrem lineárnych ohraničení bude mať naviac aj kvadratické ohraničenia. Potom ju môžeme zapísať v tvare

$$\begin{aligned} \min_x \quad & \frac{1}{2}x^T P_0 x + q_0^T x + r_0, \\ & \frac{1}{2}x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1,\dots,m \\ & Ax = b, \end{aligned} \tag{QCQP}$$

kde pre všetky $i : P_i \in S^n_+$. Ďalej $q_i \in R^n, r_i \in R$ pre všetky $i = 0,\dots,m, x \in R^n$, $A \in R^{m\times n}$ a $b \in R^m$.

Ukážeme, že obe uvedené úlohy kvadratického programovania sa dajú previesť na SOCP úlohu. Vidíme, že ohraničenia úlohy LCQP sú lineárne, problémom je teda iba kvadratická účelová funkcia. V úlohe QCQP bude treba previesť nielen účelovú funkciu, ale aj kvadratické ohraničenia na ohraničenia kužeľov druhého rádu.

Je zrejme, že

$$\min_x \quad \frac{1}{2}x^T P x + q^T x + r, \qquad \Leftrightarrow \qquad \begin{aligned} \min_x \quad & v, \\ & \frac{1}{2}x^T P x + q^T x + r \leq v. \end{aligned}$$

Keďže $P \in S^n_+$, tak existuje taká štvorcová matica B typu $n\times n$, pre ktorú platí $P = B^T B$. Potom

$$\frac{1}{2}x^T P x + q^T x + r - v = \frac{1}{2}x^T B^T B x + q^T x + r - v \leq 0. \tag{33}$$

Využijeme vzťah $z = \left(\frac{1+z}{2}\right)^2 - \left(\frac{1-z}{2}\right)^2$ a dostaneme

$$\frac{1}{2}x^T B^T B x + q^T x + r - v = \frac{1}{2}x^T B^T B x + \left(\frac{1 + (q^T x + r - v)}{2}\right)^2 - \left(\frac{1 - (q^T x + r - v)}{2}\right)^2 \leq 0.$$

Teda

$$\frac{1}{2}x^T B^T B x + \left(\frac{1 + (q^T x + r - v)}{2}\right)^2 = \left\|\frac{1}{\sqrt{2}}Bx\right\|^2 + \left(\frac{1 + (q^T x + r - v)}{2}\right)^2 \leq \left(\frac{1 - (q^T x + r - v)}{2}\right)^2.$$

Z toho už vidieť, že ohraničenie (33) je ekvivalentné s nasledujúcim ohraničením kužeľa druhého rádu

$$\left\|\begin{matrix} \frac{1}{\sqrt{2}}Bx \\ \frac{1+q^T x + r - v}{2} \end{matrix}\right\| \leq \frac{1 - (q^T x + r - v)}{2}. \tag{34}$$

Potom úlohu LCQP môžeme previesť na SOCP úlohu

$$
\begin{aligned}
\min_{x,v} \quad & v, \\
& Ax = b, \\
& Cx \leq d, \\
& \left\| \begin{matrix} \frac{1}{\sqrt{2}} Bx \\ \frac{1+q^T x + r - v}{2} \end{matrix} \right\| \leq \frac{1 - q^T x - r + v}{2},
\end{aligned}
$$

ktorá bude mať rovnakú množinu optimálnych riešení ako úloha LCQP.
Podobným spôsobom sa prevedie účelová funkcia v úlohe QCQP. Všimnime si, že keď v ohraničení (33) zvolíme $v = 0$, tak dostaneme kvadratické ohraničenia úlohy QCQP, a teda celú úlohu QCQP môžeme transformovať na SOCP úlohu

$$
\begin{aligned}
\min_{x,v} \quad & v, \\
& Ax = b, \\
& \left\| \begin{matrix} \frac{1}{\sqrt{2}} B_0 x \\ \frac{1+q_0^T x + r_0 - v}{2} \end{matrix} \right\| \leq \frac{1 - q_0^T x - r_0 + v}{2}, \\
& \left\| \begin{matrix} \frac{1}{\sqrt{2}} B_i x \\ \frac{1+q_i^T x + r_i}{2} \end{matrix} \right\| \leq \frac{1 - q_i^T x - r_i}{2}, \quad i = 1, \ldots, m.
\end{aligned}
$$

Ukázali sme spôsob, akým sa dajú previesť kvadratické ohraničenia a kvadratická účelová funkcia na ohraničenia kužeľov druhého rádu, a následne sme previedli obe úlohy LCQP, QCQP na SOCP úlohy. To znamená, že kvadratické programovanie je podtriedou programovania nad kužeľmi druhého rádu.

4.3 Úloha semidefinitného programovania

Uvažujme duálnu SOCP úlohu (D) pre $r = 1$, t.j. $\mathcal{Q}$ je iba jeden kužeľ druhého rádu.
K vektoru $s = (s_0; \bar{s}) \in R^n$ definujeme tzv. *arrow-shaped* maticu ako

$$
Arw(s) = \begin{pmatrix} s_0 & \bar{s}^T \\ \bar{s} & s_0 I \end{pmatrix}
$$

(Pozri Dodatok o vektore x). Chceme ukázať, že každá duálna SOCP úloha sa dá previesť na duálnu úlohu semidefinitného programovania v tvare

$$
\begin{aligned}
\max_{y,S} \quad & b^T y, \\
& \sum_{i=1}^{m} y_i A_i + S = C \\
& S \succeq 0,
\end{aligned}
\tag{SDP}
$$

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

Lema 4.1. *Nech*

$$S = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

je symetrická bloková matica, kde matice A, C sú symetrické a štvorcové. Predpokladajme, že C je kladne definitná matica. Potom nasledujúce vlastnosti sú ekvivalentné:

- *S je kladne semidefinitná matica.*
- *Schurov doplnok k matici C v matici S, definovaný ako $A - BC^{-1}B^T$, je kladne semidefinitná matica.*

Dôkaz. Dôkaz sa dá nájsť napr. v ([16], kapitola 14.8). □

Teraz pomocou lemy (4.1) ukážeme, že ohraničenie kužeľa druhého rádu je ekvivalentné s kladnou semidefinitnosťou priradenej symetrickej arrow-shaped matice, t.j.

$$s \succeq_{\mathcal{Q}} 0 \Leftrightarrow Arw(s) \succeq 0. \tag{35}$$

Ak $s_0 = 0$, je zrejme, že $s = 0$. Potom vektor s spĺňa ohraničenie kužeľa druhého rádu a platí aj $Arw(s) \succeq 0$.
V druhom prípade, t.j. ak $s_0 > 0$, je $s_0 I \succ 0$ a Schurov doplnok k matici $s_0 I$ v matici $Arw(s)$ je $s_0 - \bar{s}^T(s_0 I)^{-1}\bar{s}$. Potom z lemy 4.1 vyplýva

$$Arw(s) \succeq 0 \Leftrightarrow s_0 > 0 \text{ a } s_0 - \bar{s}^T(s_0 I)^{-1}\bar{s} \geq 0.$$

Všimnime si, že pre $s_0 > 0$ je

$$s \succeq_{\mathcal{Q}} 0 \quad \Leftrightarrow \quad s_0^2 \geq \bar{s}^T\bar{s} \quad \Leftrightarrow \quad s^0 \geq \frac{\bar{s}^T\bar{s}}{s^0} \quad \Leftrightarrow \quad s_0 - \bar{s}^T(s_0 I)^{-1}\bar{s} \geq 0.$$

To znamená, že platí (35), čo sme chceli ukázať.
Nech $c = (c_0; \bar{c})$ a $A = \begin{bmatrix} a & \bar{A} \end{bmatrix}$. Potom $s = c - A^T y \succeq_{\mathcal{Q}} 0$ sa dá vyjadriť ako

$$\begin{pmatrix} c_0 - a^T y \\ \bar{c} - \bar{A}^T y \end{pmatrix} \succeq_{\mathcal{Q}} 0,$$

a potom podľa (35)

$$\begin{pmatrix} c_0 - a^T y & (\bar{c} - \bar{A}^T y)^T \\ \bar{c} - \bar{A}^T y & (c_0 - a^T y) I \end{pmatrix} \succeq 0. \tag{36}$$

Z (36) dostaneme ohraničenie úlohy SDP, ak položíme $S = Arw(s), C = Arw(c)$ a

$$\begin{pmatrix} a^T y & y^T \bar{A} \\ \bar{A}^T y & a^T y I \end{pmatrix} = \sum_{i=1}^{m} y_i \begin{pmatrix} a_i & \bar{A}_i \\ (\bar{A}_i)^T & a_i I \end{pmatrix},$$

kde $\bar{A}_i$ je i-ty riadok matice $\bar{A}$, a_i je i-ta zložka vektora a.
Teraz sa už ľahko prepíše SOCP úloha (D) na úlohu SDP:

$$\begin{aligned} \max_{y,S} \quad & b^T y, \\ & \sum_{i=1}^{m} y_i \begin{pmatrix} a_i & \bar{A}_i \\ (\bar{A}_i)^T & a_i I \end{pmatrix} + \begin{pmatrix} s_0 & \bar{s}^T \\ \bar{s} & s_0 I \end{pmatrix} = \begin{pmatrix} c_0 & \bar{c}^T \\ \bar{c} & c_0 I \end{pmatrix}, \\ & \begin{pmatrix} s_0 & \bar{s}^T \\ \bar{s} & s_0 I \end{pmatrix} \succeq 0. \end{aligned} \tag{37}$$

Teda programovanie nad kužeľmi druhého rádu je špeciálnou podtriedou semidefinitného programovania.

Poznámka 4.1. Nech $x = (x^1; \ldots; x^r) \in R^n$. Potom

$$Arw(x) = Arw(x^1) \oplus \cdots \oplus Arw(x^r),$$

kde operátor $\oplus$ je definovaný pre r matíc $A_1, A_2, \ldots, A_r$ ako

$$A_1 \oplus A_2 \oplus \cdots \oplus A_r = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & A_r \end{pmatrix}.$$

Potom SDP úloha by vyzerala analogicky ako v (37), len matice by bolo blokovo diagonálne.

Poznámka 4.2. Prevod duálnych úloh bol priamočiary, ale prevod medzi primárnou SOCP úlohou a primárnou SDP úlohou nie je taký jednoduchý. Viac sa dá o tom dozvedieť v literatúre [21].

5 Metódy vnútorného bodu

V tejto kapitole sa budeme zaoberať metódami vnútorného bodu pre úlohy programavania nad kužeľmi druhého rádu. Hlavnou myšlienkou týchto metód je riešiť postupnosť tzv. bariérových úloh alebo alternatívne postupnosť perturbovaných systémov podmienok optimality, pričom optimálne riešenia týchto úloh slúžia ako štartovací bod pre riešenie nasledujúcej úlohy a konvergujú k optimálnemu riešeniu (P),(D). Najprv však zavedieme nasledujúce predpoklady.

Predpoklad 1. Matica $A = [A^1, A^2, \dots, A^r] \in R^{m\times n}$ zo SOCP úloh (P) a (D) má m lineárne nezávislých riadkov, teda plnú hodnosť $rank(A) = m$.

Poznámka 5.1. Predpoklad 1 zabezpečuje jednoznačnosť medzi y a s pre duálne riešenie (y, s):
1.) Ak máme dané y, tak definujeme $s = c - A^T y$.
2.) Predpokladajme teraz, že máme dané s a k nemu prípustné $y_1, y_2 : y_1 \neq y_2$. To znamená, že platí $A^T y_1 + s = c, A^T y_2 + s = c$. Po odčítaní dostaneme $A^T(y_1 - y_2) = 0$. Z lineárnej nezávislosti riadkov matice A vyplýva $y_1 - y_2 = 0$, čo je spor s predpokladom $y_1 \neq y_2$. Takže ak poznáme s, vieme k nemu jednoznačne určiť prípustné y.
Predpoklad 1 je technického charakteru a nie je vôbec obmedzujúci, pretože ak by boli riadky matice A lineárne závisle, tak nadbytočné by sme mohli odstrániť.

Predpoklad 2. Obe SOCP úlohy, primárna (P) aj duálna (D) úloha, majú ostro prípustné riešenie, takže

- existuje $x^0 \in R^n$, pre ktoré platí $Ax^0 = b, x^0 \in int\mathcal{Q}$,
- existuje $(y^0, s^0) \in R^m \times R^n$ také, že $A^T y^0 + s^0 = c, s^0 \in int\mathcal{Q}$,

t.j. $\mathcal{P}^0 \neq \emptyset, \mathcal{D}^0 \neq \emptyset$.

Poznámka 5.2. Predpoklad 2 predstavuje Slaterovú podmienku, ktorá zaručuje existenciu optimálnych riešení úloh (P),(D) (Silná veta o dualite (3.2)), a preto je to dôležitý predpoklad pre aplikáciu metód vnútorného bodu.

Vhodná bariérová funkcia $b(x) : int\mathcal{Q} \to R$ by mala byť konvexná na množine $int\mathcal{Q}$ a pre ľubovoľnú postupnosť bodov $\{x_i\}$ blížiacich sa k $bd\mathcal{Q}$ musí platiť $b(x_i) \to \infty$.
V semidefinitnom programovaní je bariérová funkcia definovaná ako

$$b(X) = -\ln\det X, \text{ kde } X \in S^n_{++}.$$

V snahe vytvoriť analógiu metód vnútorného bodu pre úlohy programovania nad kužeľmi druhého rádu s metódami vnútorného bodu v semidefinitnom programovaní, definuje sa

podobná štruktúra pre vektor x (Pozri Dodatok - Spektrálny rozklad vektora x).
Bariérová funkcia bude mať tvar

$$b(x) = -\ln\det x,$$

kde $\det x = \prod_{i=1}^{r}((x_0^i)^2 - \|\bar{x}^i\|^2)$ pre $x \in \mathcal{Q} = \mathcal{Q}_{n_1} \times \cdots \times \mathcal{Q}_{n_r}$. Definičným oborom bariérovej funkcie b je $int\mathcal{Q}$, t.j. $x \succ_{\mathcal{Q}} 0$.
Z Lemy (5.2) a z vlastnosti

$$\{x_i\} \to bd\mathcal{Q} \quad \Leftrightarrow \quad \det x_i \to 0$$

vyplýva, že $b(x)$ je vhodne zvolená bariérová funkcia, pretože

$$\lim_{y \to 0+} (\ln y) = -\infty.$$

Potom môžeme definovať pomocné bariérové úlohy v tvare

$$\begin{aligned} \min_x \quad & c^T x - \mu \ln\det x, \\ & Ax = b, \end{aligned} \qquad (P_\mu) \qquad\qquad \begin{aligned} \max_{y,s} \quad & b^T y + \mu \ln\det s, \\ & A^T y + s = c, \end{aligned} \qquad (D_\mu)$$

kde $\mu > 0, x \in R^n, s \in R^n, y \in R^m, A \in R^{m \times n}, b \in R^m, c \in R^n$.
Ukážeme, že pomocné bariérové úlohy (P_μ),(D_μ) majú jednoznačné riešenie.

Lema 5.1. *Pre $z = (z_0; \bar{z}) \in R^p$, kde $z_0 \in R, \bar{z} \in R^{p-1}$, platí*

$$\nabla_z(-\ln(z_0^2 - \|\bar{z}\|^2)) = -\frac{2}{z_0^2 - \|\bar{z}\|^2}\begin{pmatrix} z_0 \\ -\bar{z} \end{pmatrix}.$$

Dôkaz. Počítajme

$$\begin{aligned} \nabla_z(-\ln(z_0^2 - \|\bar{z}\|^2)) &= -\frac{1}{z_0^2 - \|\bar{z}\|^2}(2z_0, -2z_1, \ldots, -2z_{n-1})^T = \\ &= -\frac{2}{z_0^2 - \|\bar{z}\|^2}\begin{pmatrix} z_0 \\ -\bar{z} \end{pmatrix}. \end{aligned}$$

□

Lema 5.2. *Bariérová funkcia $b(x) = -\ln\det x$ je rýdzokonvexná na $int\mathcal{Q} \subseteq R^n$.*

Dôkaz. Funkcia $b(x)$ bude rýdzokonvexná na $int\mathcal{Q}$ práve vtedy, keď

$$\nabla_x^2(-\ln\det x) \succ 0.$$

Všimnime si, že

$$b(x) = -\ln\det x = -\ln\prod_{i=1}^{r}((x_0^i)^2 - \|\bar{x}^i\|^2) = -\sum_{i=1}^{r}\ln((x_0^i)^2 - \|\bar{x}^i\|^2).$$

Keďže súčet rýdzokonvexných funkcii je rýdzokonvexná funkcia, tak stačí ukázať, že funkcia $-\ln((x_0^i)^2 - \|\bar{x}^i\|^2)$ pre pevne zvolené i je rýdzokonvexná na $int\mathcal{Q}_{n_i} \subseteq R^{n_i}$.
Funkcia $b(x)$ je definovaná pre $\det x \neq 0$, preto platí aj $\det x^i \neq 0, i = 1, \ldots, r$.
Z lemy (5.1) vyplýva, že

$$\nabla_{x^i}(-\ln((x_0^i)^2 - \|\bar{x}^i\|^2)) = -\frac{2}{(x_0^i)^2 - \|\bar{x}^i\|^2}\begin{pmatrix} x_0^i \\ -\bar{x}^i \end{pmatrix}.$$

A teraz postupne vypočítame parciálne derivácie

$$\frac{\partial}{\partial x_0^i}\left(-\frac{2x_0^i}{(x_0^i)^2 - \|\bar{x}^i\|^2}\right) = \frac{-2((x_0^i)^2 - \|\bar{x}^i\|^2) + (2x_0^i)^2}{((x_0^i)^2 - \|\bar{x}^i\|^2)^2} = \frac{2\|x^i\|^2}{((x_0^i)^2 - \|\bar{x}^i\|^2)^2}.$$

Pre $j = 1, \ldots, n-1$ dostávame

$$\frac{\partial}{\partial x_j^i}\left(-\frac{2x_0^i}{(x_0^i)^2 - \|\bar{x}^i\|^2}\right) = \frac{4x_0^i x_j^i}{((x_0^i)^2 - \|\bar{x}^i\|^2)^2}$$

a ďalej

$$\frac{\partial}{\partial x_j^i}\left(\frac{2x_j^i}{(x_0^i)^2 - \|\bar{x}^i\|^2}\right) = \frac{2((x_0^i)^2 - \|\bar{x}^i\|^2) + (2x_j^i)^2}{((x_0^i)^2 - \|\bar{x}^i\|^2)^2}.$$

Nakoniec pre $k \neq j$, kde $k, j \in \{1, \ldots, n-1\}$, platí

$$\frac{\partial}{\partial x_k^i}\left(\frac{2x_j^i}{(x_0^i)^2 - \|\bar{x}^i\|^2}\right) = \frac{4x_k^i x_j^i}{((x_0^i)^2 - \|\bar{x}^i\|^2)^2}.$$

Takže

$$\nabla_{x^i}^2(-\ln((x_0^i)^2 - \|\bar{x}^i\|^2)) = \frac{2}{(\det x^i)^2}\begin{pmatrix} \|x^i\|^2 & -2x_0^i(\bar{x}^i)^T \\ -2x_0^i\bar{x}^i & (\det x^i)I + 2\bar{x}^i(\bar{x}^i)^T \end{pmatrix}.$$

Stačí teda ukázať, že $\nabla_{x^i}^2(-\ln((x_0^i)^2 - \|\bar{x}^i\|^2)) \succ 0$. Využijeme analógiu Lemy (4.1) o Schurovom doplnku. Keďže $\|x^i\|^2 > 0$, tak

$$\nabla_{x^i}^2(-\ln((x_0^i)^2 - \|\bar{x}^i\|^2)) \succ 0 \quad \Leftrightarrow \quad M \succ 0,$$

kde

$$M = (\det x^i)I + 2\bar{x}^i(\bar{x}^i)^T - \frac{4(x_0^i)^2}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T$$

je Schurov doplnok k $\|x^i\|^2$. Maticu M môžeme upraviť do tvaru

$$\begin{aligned} M &= (\det x^i)I + 2\bar{x}^i(\bar{x}^i)^T - \frac{4(x_0^i)^2}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T = \\ &= (\det x^i)I + \frac{2(x_0^i)^2 + 2\|\bar{x}^i\|^2 - 4(x_0^i)^2}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T = \\ &= (\det x^i)I - \frac{2\det x^i}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T = (\det x^i)\left[I - \frac{2}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T\right]. \end{aligned}$$

Ukážeme, že pre ľubovoľné $u \in R^{n_i-1}$ platí $u^T M u > 0$, potom bude matica M kladne definitná. Je zrejme, že $\det x^i > 0$, lebo $(x_0^i)^2 > \|\bar{x}^i\|^2$. Ekvivalentnými úpravami s využitím Cauchy-Schwarzovej nerovnosti

$$\|u\|^2\|\bar{x}^i\|^2 \geq (u^T\bar{x}^i)^2$$

dostaneme

$$\|x^i\|^2\|u\|^2 = (x_0^i)^2\|u\|^2 + \|u\|^2\|\bar{x}^i\| > 2\|\bar{x}^i\|^2\|u\|^2 \geq 2(u^T\bar{x}^i)^2 \geq 0,$$

a teda

$$\|x^i\|^2\|u\|^2 > 2(u^T\bar{x}^i)^2.$$

Potom môžeme vidieť, že

$$u^T M u = u^T(\det x^i)\left[I - \frac{2}{\|x^i\|^2}\bar{x}^i(\bar{x}^i)^T\right]u = (\det x^i)\left[\|u\|^2 - \frac{2}{\|x^i\|^2}(u^T\bar{x}^i)^2\right] > 0,$$

čo sme chceli ukázať. □

Veta 5.1. *Nech $\mu > 0$. Potom*

a) $f_\mu(x) = c^T x - \mu \ln\det x$ *je rýdzokonvexná na* $\mathcal{P}^0$,

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

Dôkaz. Nech $\mu > 0$.

a) Z lemy (5.2) vyplýva, že $-\mu \ln\det x$ je rýdzokonvexná funkcia v x na množine $int\mathcal{Q}$ a pričítanie lineárnej funkcie $c^T x$ túto rýdzokonvexnosť nezmení. Teda funkcia $f_\mu(x)$ je rýdzokonvexná na $\mathcal{P}^0$.

b) Z lemy (5.2) vyplýva, že $\mu \ln\det s$ je rýdzokonkávna funkcia v premennej s na množine $int\mathcal{Q}$. Funkcia $g_\mu(y, s)$ je v premennej y iba lineárna, preto ešte nie je zaručená rýdzokonkávnosť $g_\mu(y, s)$ na $\mathcal{D}^0$. Predpoklad 1 však zabezpečuje jednoznačnosť medzi y a s, a teda funkcia $g_\mu(y, s)$ bude rýdzokonkávna na $\mathcal{D}^0$.

□

Z predchádzajúcej vety vyplýva, že ak budú mať úlohy (P_μ),(D_μ) riešenie, tak bude jediné. Nasledujúca lema je dôsledkom známej vety pre neohraničenú množinu ([10],veta 1.2.9).

Lema 5.3. *Nech K je neprázdna konvexná a neohraničená podmnožina $int\mathcal{Q}$, kde $\mathcal{Q} \subseteq R^n$ je kartézsky súčin kužeľov druhého rádu. Potom*

$$\forall u \in K\ \exists v \in K, u \neq v : u + t(v - u) \in K \quad \forall t \geq 0,$$
$$v - u \succeq_{\mathcal{Q}} 0.$$

Veta 5.2.

a) *Pre každé $\mu > 0$ existuje optimálne riešenie úlohy (P_μ).*

b) *Pre každé $\mu > 0$ existuje optimálne riešenie úlohy (D_μ).*

Dôkaz.

a) Chceme ukázať, že $\forall \mu > 0$ má úloha (P_μ) optimálne riešenie.
Nech $\mu > 0$ je pevne zvolené. Z predpokladu 2 vieme, že $\exists x^0 \in \mathcal{P}^0$.
Uvažujme množinu

$$\mathcal{P}(x^0) = \{x \in \mathcal{P}^0 | f_\mu(x) \le f_\mu(x^0)\},$$

kde $f_\mu(x) = c^T x - \mu \ln \det x$ je účelová funkcia úlohy (P_μ).
Je zrejme, že množina optimálnych riešení úlohy

$$\min_x \{f_\mu(x) | x \in \mathcal{P}(x^0)\} \qquad (\tilde{P}_\mu)$$

je rovnaká ako množina optimálnych riešení úlohy (P_μ), preto stačí ukázať, že existuje optimálne riešenie úlohy $(\tilde{P}_\mu)$.
Množina $\mathcal{P}(x^0)$ je úrovňovou množinou konvexnej funkcie, teda je to konvexná množina. Funkcia f_μ je spojitá a $\mathcal{P}(x^0) \neq \emptyset$, lebo $x^0 \in \mathcal{P}(x^0)$. Preto ak ukážeme, že $\mathcal{P}(x^0)$ je uzavretá a ohraničená, t.j. kompaktná, tak aplikovaním Weierstrassovej vety dostaneme, že existuje optimálne riešenie úlohy $(\tilde{P}_\mu)$.
1.) Najprv teda ukážeme uzavretosť množiny $\mathcal{P}(x^0)$, t.j. limitný bod ľubovoľnej konvergentnej postupnosti bodov $\{x^n\}_{n=1}^\infty$ množiny $\mathcal{P}(x^0)$ patrí tiež do tejto množiny. Predpokladajme, že $x^n \in \mathcal{P}(x^0)\ \forall n$ a $x^n \to \tilde{x}$ pre $n \to \infty$. Je zrejme, že potom

$$Ax^n = b \quad \forall n = 1, 2, \dots \text{ implikuje } A\tilde{x} = b.$$

Keďže funkcia f_μ je spojitá, tak platí

$$\lim_{n\to\infty} f_\mu(x^n) = f_\mu(\tilde{x})$$

a z nerovnosti $f_\mu(x^n) \le f_\mu(x^0)$ vyplýva nerovnosť

$$f_\mu(\tilde{x}) \le f_\mu(x^0). \qquad (38)$$

Problémom je, že body $x^n \in int\mathcal{Q}$ môžu konvergovať ku hranici kužeľa $\mathcal{Q}$. Ak by $\tilde{x}$ patrilo hranici kužeľa $\mathcal{Q}$, tak potom $\det \tilde{x} = 0$ a to by bol spor s ohraničením (38), pretože

$$-\ln \det \tilde{x} = \infty.$$

To znamená, že $\tilde{x} \in int\mathcal{Q}$, takže bod $\tilde{x}$ patrí do množiny $\mathcal{P}(x^0)$, čo sme chceli dokázať.

2.) Teraz ukážeme, že množina $\mathcal{P}(x^0)$ je ohraničená. Predpokladajme, že je $\mathcal{P}(x^0)$ neohraničená. Už sme spomínali, že je neprázdna a konvexná, preto podľa Lemy (5.3) existuje také $x_1 \in \mathcal{P}(x^0), x_2 \in \mathcal{P}(x^0), x_1 \neq x_2, x_2 - x_1 \succeq_{\mathcal{Q}} 0$, pre ktoré platí

$$\{x_t | x_t = x_1 - t(x_2 - x_1), t \geq 0\} \subset \mathcal{P}(x^0).$$

Z toho vyplýva, že

$$f_\mu(x_t) \leq f_\mu(x^0).$$

Pozrieme sa na hodnotu $f_\mu(x_t)$ pre $t \to \infty$, t.j.

$$\lim_{t\to\infty} f_\mu(x_t) = \lim_{t\to\infty} c^T x_1 + t\left[c^T(x_2 - x_1) - \frac{\mu \ln\det(x_1 + t(x_2 - x_1))}{t}\right]. \qquad (39)$$

Označme polynóm $p(t) = \det(x_1 + t(x_2 - x_1))$ a využijeme L'Hospitalovo pravidlo na výpočet limity

$$\lim_{t\to\infty} \frac{-\mu \ln p(t)}{t} = \lim_{t\to\infty} \frac{-\mu p'(t)}{p(t)} = 0.$$

Z predpokladu 2 existuje $(y^0, s^0) \in \mathcal{D}^0$ a z Dôsledku (2.2 a)) vyplýva

$$(s^0)^T(x_2 - x_1) \geq 0.$$

Potom

$$c^T(x_2 - x_1) = (y^0)^T A x_2 - (y^0)^T A x_1 + (s^0)^T(x_2 - x_1) = (y^0)^T b - (y^0)^T b + (s^0)^T(x_2 - x_1) \geq 0.$$

To znamená, že

$$\lim_{t\to\infty} f_\mu(x_t) = \infty,$$

čo je spor s ohraničením $f_\mu(x_t) \leq f_\mu(x^0)$. Preto je množina $\mathcal{P}(x^0)$ ohraničená.

b) Teraz chceme ukázať, že $\forall \mu > 0$ má úloha (D_μ) optimálne riešenie.
Nech $\mu > 0$ je pevne zvolené. Z predpokladu 2 vieme, že $\exists (y^0, s^0) \in \mathcal{D}^0$.
Uvažujme množinu

$$\mathcal{D}(y^0, s^0) = \{(y, s) \in \mathcal{D}^0 | g_\mu(y, s) \geq g_\mu(y^0, s^0)\},$$

kde $g_\mu(y, s) = b^T y + \mu \ln\det s$ je účelová funkcia úlohy (D_μ).
Je zrejme, že množina optimálnych riešení úlohy

$$\min_{y,s}\{g_\mu(y, s) | (y, s) \in \mathcal{D}(y^0, s^0)\} \qquad (\tilde{D_\mu})$$

je rovnaká ako množina optimálnych riešení úlohy (D_μ), preto stačí ukázať, že existuje optimálne riešenie úlohy $(\tilde{D_\mu})$.
Množina $\mathcal{D}(x^0)$ je úrovňovou množinou konkávnej funkcie, teda je to konvexná

množina. Funkcia g_μ je spojitá a $\mathcal{D}(y^0, s^0) \neq \emptyset$, lebo $(y^0, s^0) \in \mathcal{D}(y^0, s^0)$. Preto ak ukážeme, že $\mathcal{D}(y^0, s^0)$ je uzavretá a ohraničená, t.j. kompaktná, tak aplikovaním Weierstrassovej vety dostaneme, že existuje optimálne riešenie úlohy $(\tilde{D}_\mu)$.

1.) Najprv teda ukážeme uzavretosť množiny $\mathcal{D}(y^0, s^0)$, t.j. limitný bod ľubovoľnej konvergentnej postupnosti bodov $\{y^n\}_{n=1}^\infty, \{s^n\}_{n=1}^\infty$ množiny $\mathcal{D}(y^0, s^0)$ patrí tiež do tejto množiny. Predpokladajme, že $(y^n, s^n) \in \mathcal{D}(y^0, s^0)$ $\forall n$ a $y^n \to \tilde{y}, s^n \to \tilde{s}$ pre $n \to \infty$. Je zrejme, že potom

$$A^T y^n + s^n = c \quad \forall n = 1, 2, \dots \text{ implikuje } A\tilde{y} + \tilde{s} = c.$$

Keďže funkcia g_μ je spojitá, tak platí

$$\lim_{n\to\infty} g_\mu(y^n, s^n) = g_\mu(\tilde{y}, \tilde{s})$$

a z nerovnosti $g_\mu(y^n, s^n) \leq g_\mu(y^0, s^0)$ vyplýva nerovnosť

$$g_\mu(\tilde{y}, \tilde{s}) \leq g_\mu(y^0, s^0). \tag{40}$$

Problémom je, že body $s^n \in int\mathcal{Q}$ môžu konvergovať ku hranici kužeľa $\mathcal{Q}$. Ak by $\tilde{s}$ patrilo hranici kužeľa $\mathcal{Q}$, tak potom $\det \tilde{s} = 0$ a to by bol spor s ohraničením (40), pretože

$$\ln \det \tilde{s} = -\infty.$$

To znamená, že $\tilde{s} \in int\mathcal{Q}$, takže bod $\tilde{s}$ patrí do množiny $\mathcal{D}(y^0, s^0)$, čo sme chceli dokázať.

2.) Teraz ukážeme, že množina $\mathcal{D}(y^0, s^0)$ je ohraničená. Predpokladajme, že $\mathcal{D}(y^0, s^0)$ je neohraničená. Už sme spomínali, že je neprázdna a konvexná, preto podľa Lemy (5.3) existuje také $(y_1, s_1) \in \mathcal{D}(y^0, s^0), (y_2, s_2) \in \mathcal{D}(y^0, s^0), (y_1, s_1) \neq (y_2, s_2)$, $s_2 - s_1 \succeq_\mathcal{Q} 0$ (z predpokladu 1 $s_1 \neq s_2$), pre ktoré platí

$$\{(y_t, s_t) | (y_t, s_t) = (y_1, s_1) - t[(y_2, s_2) - (y_1, s_1)], t \geq 0\} \subset \mathcal{D}(y^0, s^0).$$

Z toho vyplýva, že

$$g_\mu(y_t, s_t) \geq g_\mu(y^0, s^0).$$

Pozrieme sa na hodnotu $g_\mu(y_t, s_t)$ pre $t \to \infty$, t.j.

$$\lim_{t\to\infty} g_\mu(y_t, s_t) = \lim_{t\to\infty} b^T y_1 + t\left[b^T(y_2 - y_1) + \frac{\mu \ln \det(s_1 + t(s_2 - s_1))}{t}\right]. \tag{41}$$

Označme polynóm $p(t) = \det(s_1 + t(s_2 - s_1))$ a využijeme L'Hospitalovo pravidlo na výpočet limity

$$\lim_{t\to\infty} \frac{\mu \ln p(t)}{t} = \lim_{t\to\infty} \frac{\mu p'(t)}{p(t)} = 0.$$

Z predpokladu 2 existuje $x^0 \in \mathcal{P}^0$ a z Dôsledku (2.2 a)) vyplýva

$$(x^0)^T(s_2 - s_1) \geq 0.$$

Potom

$$b^T(y_2 - y_1) = (x^0)^T A^T y_2 - (x^0)^T A^T y_1 + (x^0)^T(c - c) = -(x^0)^T(s_2 - s_1) \leq 0.$$

To znamená, že

$$\lim_{t\to\infty} g_\mu(y_t, s_t) = -\infty,$$

čo je spor s ohraničením $g_\mu(y_t, s_t) \geq g_\mu(y^0, s^0)$. Preto je množina $\mathcal{D}(y^0, s^0)$ ohraničená.

□

V dôkaze Vety (5.2) sme využili analógiu tejto vety zo semidefinitného programovania (napr. [17]).

Dôsledok 5.1. *Pre každé $\mu > 0$ existuje práve jedno optimálne riešenie úloh (P_μ) a (D_μ).*

Dôkaz. Dôsledok vyplýva z Vety (5.1) a Vety (5.2). □

Veta 5.3. *Nech $\mu > 0$. Potom $(\hat{x}, \hat{y}, \hat{s})$ je optimálne riešenie úloh (P_μ),(D_μ) práve vtedy, keď*

$$\begin{aligned} A\hat{x} = b, &\quad \hat{x} \succ_{\mathcal{Q}} 0, \\ A^T\hat{y} + \hat{s} = c, &\quad \hat{s} \succ_{\mathcal{Q}} 0, \\ \hat{x} \bullet \hat{s} = 2\mu e, & \end{aligned} \tag{42}$$

kde $e = (e^1; \ldots; e^r)$, pričom $e^i = (1; 0) \in R^{n_i}$.

Dôkaz. Nech $\mu > 0$. K úlohe (P_μ) definujeme Lagrangeovu funkciu v tvare

$$L(x, y) = c^T x - \mu \ln \det x + y^T(b - Ax),$$

kde $x \in int\mathcal{Q}, y \in R^m$.

Optimálne riešenie $\hat{x} \in int\mathcal{Q}$ úlohy (P_μ) musí spĺňať podmienky

$$\bigtriangledown_x L(\hat{x}, \hat{y}) = 0, \quad \bigtriangledown_y L(\hat{x}, \hat{y}) = 0$$

pre ľubovoľné $\hat{y} \in R^m$.

Najprv si všimnime, že z lemy (5.1) a z definície x^{-1} pre $r > 1$ (v Dodatku) vyplýva, že

$$\bigtriangledown_x(-\ln \det x) = -2x^{-1}.$$

Potom

$$\begin{aligned} \nabla_x L(\hat{x},\hat{y}) &= c - 2\mu\hat{x}^{-1} - A^T\hat{y} = 0, \\ \nabla_y L(\hat{x},\hat{y}) &= b - A\hat{x} = 0. \end{aligned} \tag{43}$$

Označme $\hat{s} = c - A^T\hat{y}$. Potom z (43) dostaneme

$$\begin{aligned} \hat{s} &= 2\mu\hat{x}^{-1}, \\ A\hat{x} &= b, \\ A^T\hat{y} + \hat{s} &= c. \end{aligned} \tag{44}$$

Z vlastnosti $\|-z\| = \|z\|$ je zrejme, že $\hat{s} \in int\mathcal{Q}$. Ďalej z (44) vyplýva, že

$$\hat{s}^i = 2\mu(\hat{x}^i)^{-1} \quad \forall i,$$

a teda

$$\hat{s}_0^i = \frac{2\mu\hat{x}_0^i}{\det\hat{x}^i}, \quad \hat{\bar{s}}_0^i = \frac{-2\mu}{\det\hat{x}^i}\hat{\bar{x}}_0^i,$$

z čoho sa ľahko overí ekvivalencia

$$\hat{s}^i = 2\mu(\hat{x}^i)^{-1} \quad \forall i \quad \Leftrightarrow \quad \hat{x}^i \circ \hat{s}^i = 2\mu e^i \quad \forall i \quad \Leftrightarrow \quad \hat{x} \bullet \hat{s} = 2\mu e.$$

Podobným spôsobom by sme z Lagrangeovej funkcie k úlohe (D_μ)

$$L(y,s,x) = b^T y + \mu \ln\det s + x^T(c - A^T y - s),$$

kde $s \in int\mathcal{Q}$, dostali podmienky pre optimálne riešenie $(\hat{y},\hat{s})$ úlohy (D_μ), t.j.

$$\nabla_y L(\hat{y},\hat{s},\hat{x}) = 0, \quad \nabla_s L(\hat{y},\hat{s},\hat{x}) = 0, \quad \nabla_x L(\hat{y},\hat{s},\hat{x}) = 0,$$

a z toho opäť systém (42), čo sme chceli dokázať. □

Ukázalo sa, že pomocné bariérové úlohy (P_μ),(D_μ) majú pre každé $\mu > 0$ jednoznačné riešenie a to získame riešením tzv. pertubovaného systému

$$\begin{aligned} Ax &= b, \quad x \succ_{\mathcal{Q}} 0, \\ A^T y + s &= c, \quad s \succ_{\mathcal{Q}} 0, \\ x \bullet s &= 2\mu e. \end{aligned} \tag{45}$$

Zavedieme pojem centrálnej trajektórie.

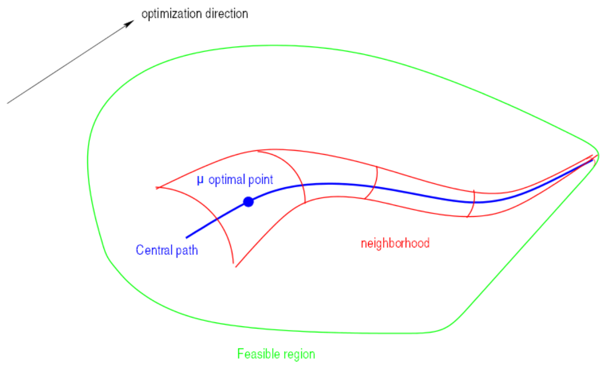
Definícia 5.1. Množina

$$\{(x(\mu), y(\mu), s(\mu)); \mu > 0\}$$

riešení úloh (P_μ),(D_μ) sa nazýva centrálna trajektória.

Dôsledok 5.2. *Nech $\mu > 0$. Potom bod $(x(\mu), y(\mu), s(\mu))$ je bodom centrálnej trajektórie práve vtedy, ak je riešením systému (45).*

Na základe analógie zo semidefinitného programovania sa dá ukázať, že pre $\mu \to 0$ existuje limita centrálnej trajektórie a konverguje k optimálnemu riešeniu. Túto vlastnosť využívajú primárno-duálne metódy vnútorného bodu. Ich myšlienkou je začať v blízkom okolí alebo v bode priamo na centrálnej trajektórii a postupne sa posúvať k optimálnemu riešeniu redukovaním duálnej medzery. Smer posunu $(\triangle x, \triangle y, \triangle s)$ získame z Newtonovej metódy aplikovanej na systém (45) a konvergenciu riešenia zabezpečíme, keď ostaneme v dostatočnej blízkosti centrálnej trajektórie (obrázok - oblasť "neighborhood", [3]).



Predpokladajme, že máme začiatočný bod (x, y, s) pre $\mu > 0$. Aplikujeme Newtonovu metódu pre systém (45):

$$\begin{aligned} A(x + \triangle x) &= b, \\ A^T(y + \triangle y) + (s + \triangle s) &= c, \\ (x + \triangle x) \bullet (s + \triangle s) &= 2\mu e. \end{aligned} \tag{46}$$

Zanedbáme člen $\triangle x \bullet \triangle s$ a po úprave dostaneme

$$\begin{aligned} A\triangle x &= b - Ax, \\ A^T \triangle y + \triangle s &= c - A^T y - s, \\ x \bullet \triangle s + \triangle x \bullet s &= 2\mu e - x \bullet s. \end{aligned} \tag{47}$$

Dá sa ľahko overiť, že platí nasledujúca ekvivalencia

$$x \bullet \triangle s + \triangle x \bullet s = Arw(s)\triangle x + Arw(x)\triangle s,$$

kde Arw(x) je definované ako v Poznámke (4.1). Potom systém (47) môžeme prepísať do maticového tvaru

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Arw(s) & 0 & Arw(x) \end{pmatrix} \begin{pmatrix} \triangle x \\ \triangle y \\ \triangle s \end{pmatrix} = \begin{pmatrix} b - Ax \\ c - A^T y - s \\ 2\mu e - x \bullet s \end{pmatrix}. \tag{48}$$

Označíme $r_P = b - Ax, r_D = c - A^T y - s, r_C = 2\mu e - x \bullet s$. Všimnime si, že $r_P = 0$, ak x je primárne prípustné a $r_D = 0$, ak (y, s) je duálne prípustné. Zo systému (47) vyjadríme smer $(\triangle x; \triangle y; \triangle s)$ a dostaneme

$$\begin{aligned}
\triangle x &= -Arw^{-1}(s)(Arw(s)\triangle s - r_C),\\
\triangle y &= [AArw^{-1}(s)Arw(x)A^T]^{-1}[AArw^{-1}(s)(Arw(x)r_D - r_C) + r_P],\\
\triangle s &= r_D - A^T\triangle y.
\end{aligned} \tag{49}$$

Následne zvolíme dĺžku kroku a ak bod $(x+\triangle x; y+\triangle y; s+\triangle s)$ je prípustný, tak bude novým štartovacím bodom do ďalšej iterácie. Takýto postup sa bude opakovať, až kým sa dostatočne nepriblížime k optimálnemu riešeniu.
Pre vhodne zvolený začiatočný bod, dĺžku kroku a redukciu duálnej medzery sa dá konvergencia k optimálnemu riešeniu dosiahnuť polynomiálnym počtom iterácii. Výhodou metód vnútorného bodu pre SOCP úlohy oproti úlohám semidefinitného programovania je, že výpočtová zložitosť na interáciu je menšia než pre úlohy SDP približne rovnakej veľkosti a štruktúry. Podľa literatúry ([1], časť 7) definujeme tzv. proximity measure $d_f(x,s) = \|Ms - \mu e\|_F$, kde $M = M_1 \oplus \cdots \oplus M_r$ a

$$M_i = \begin{pmatrix} (x_0^i)^2 & (\bar{x}^i)^T \\ \bar{x}^i & (\det x^i)^{\frac{1}{2}} I + \frac{\bar{x}^i(\bar{x}^i)^T}{(\det x^i)^{\frac{1}{2}} + x_0^i} \end{pmatrix}.$$

Pomocou tejto miery môžeme zadefinovať okolie centrálnej trajektórie pre $\gamma \in (0,1)$ ako

$$\mathcal{N}_F(\gamma) = \{(x;y;s) | Ax = b, A^Ty + s = c, x \succ_Q 0, s \succ_Q 0, d_F(x,s) \le \gamma\mu\},$$

čo nám poslúži v podmienke pre redukciu μ. Potom algoritmus pre riešenie úloh programovania nad kužeľmi druhého rádu vyzerá nasledovne:

Primárno-duálny algoritmus

1. Zvolíme presnosť duálnej medzery $\varepsilon > 0$, $\gamma \in (0,1), \delta \in (0,1)$. Ďalej zvolíme štartovací bod $(x_0; y_0; s_0)$ a $\mu_0 > 0$ tak, aby $(x_0; y_0; s_0) \in \mathcal{N}_F(\gamma)$.

2. Nech $x := x_0, y := y_0, s := s_0, \mu := \mu_0$. Opakujeme nasledujúci postup:

    Pokiaľ je $x^Ts \ge \varepsilon$, tak

    $\mu = \sigma\mu$, kde $\sigma = 1 - \frac{\delta}{\sqrt{r}}$,

    Pokiaľ platí $(x;y;s) \notin \mathcal{N}_F(\gamma)$, tak

    Vypočítame $(\triangle x; \triangle y; \triangle s)$ riešením systému (47) v bode $(x;y;s)$.

    Zvolíme α tak, aby $x + \alpha\triangle x \succ_Q 0, s + \alpha\triangle s \succ_Q 0$, potom

    $$(x;y;s) := (x;y;s) + \alpha(\triangle x; \triangle y; \triangle s).$$

    inak STOP,

    inak STOP.

V tomto pr'ipade je iteračná zložitosť algoritmu $\mathcal{O}(\sqrt{r})$. Iné možnosti voľby napr. okolia centrálnej trajektórie alebo redukcie μ sú podrobnejšie uvedené v ([1] časť 7, resp.[5]).

6 Dodatok

6.1 Kužele

Nasledujúcu vetu využijeme v niektorých dôkazoch vlastností duálneho kužeľa.

Veta 6.1 (Veta o oddeľujúcej nadrovine)**.** *Nech $C \subseteq R^n$ je uzavretá konvexná množina a nech $x_0 \notin C$. Potom existuje také $a \in R^n, a \neq 0, b \in R$, že $a^T x \geq b \quad \forall x \in C$ a $a^T x_0 < b$.*

Dôkaz. Dôkaz sa dá nájsť v literatúre ([19], str. 3).[4] □

Jej analógia pre kužele sa využíva pri dôkazoch Zovšeobecnenej Farkašovej lemy (Tvrdenie 3.2). Formulácie oboch viet o oddeľujúcej nadrovine sa dajú nájsť v ([2], Lec03).

Tvrdenie 6.1 (Veta o oddeľujúcej nadrovine pre kužele)**.** *Nech $\mathcal{K} \subset R^n$ je uzavretý konvexný kužeľ a nech $z \notin \mathcal{K}$. Potom existuje také nenulové $a \in R^n$, že*

$$a^T z < 0 \qquad a \qquad a^T x \geq 0 \quad \forall x \in \mathcal{K}.$$

Poznámka 6.1. Veta o oddeľujúcej nadrovine pre kužele je dôsledkom Vety (6.1).

Veta 6.2. *Pre duálny kužeľ $\mathcal{K}^*$ platia nasledujúce vlastnosti:*

a) $\mathcal{K}^*$ je konvexný kužeľ.

b) Nech kužele $\mathcal{K}_1, \mathcal{K}_2$ spĺňajú $\mathcal{K}_1 \subseteq \mathcal{K}_2$. Potom pre ich duálne kužele platí $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$.

c) $\mathcal{K}^*$ je uzavretý.

d) Vnútro $\mathcal{K}^*$ je $int\mathcal{K}^* = \{z | z^T x > 0 \quad \forall x \in cl\mathcal{K} \setminus \{0\}\}$.

e) Ak $int\mathcal{K} \neq \emptyset$, potom $\mathcal{K}^*$ je "špicatý".

f) Nech K je konvexný kužeľ. Potom $\mathcal{K}^{**}$ je $cl\mathcal{K}$. Ak $\mathcal{K}$ je aj uzavretý, tak $\mathcal{K}^{**} = \mathcal{K}$.

g) Ak $cl\mathcal{K}$ je "špicatý", tak $int\mathcal{K}^* \neq \emptyset$.

Dôkaz. a) $\mathcal{K}^*$ je konvexný, ak pre ľubovoľné $y, z \in \mathcal{K}^*$ a $\theta_1, \theta_2 \geq 0$ platí: $\theta_1 y + \theta_2 z \in \mathcal{K}^*$. Nech $y, z \in \mathcal{K}^*$ a $\theta_1, \theta_2 \geq 0$. Teda $x^T y \geq 0 \quad \forall x \in \mathcal{K}$ a tiež $x^T z \geq 0 \quad \forall x \in \mathcal{K}$. Potom je zrejme, že

$$x^T(\theta_1 y + \theta_2 z) = \theta_1 x^T y + \theta_2 x^T z \geq 0 \quad \forall x \in \mathcal{K},$$

pretože lineárna kombinácia kladných čísel je kladná. Z toho vyplýva $\theta_1 y + \theta_2 z \in \mathcal{K}^*$.

[4] Táto veta sa dá v angličtine nájsť pod názvom Separating hyperplane theorem a má rôzne alternatívne varianty, napr. ([8], str. 46)

b) Nech $\mathcal{K}_1 \subseteq \mathcal{K}_2$ a nech $z \in \mathcal{K}_2^*$ je ľubovoľné. To znamená, že $x^T z \geq 0 \quad \forall x \in \mathcal{K}_2$.
Keďže $\mathcal{K}_2 \supseteq \mathcal{K}_1$, tak platí aj $x^T z \geq 0 \quad \forall x \in \mathcal{K}_1$, a teda $z \in \mathcal{K}_1^*$.
Ukázali sme, že ľubovoľný prvok z $\mathcal{K}_2^*$ patrí aj do $\mathcal{K}_1^*$, t.j. $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$.

c) Ukážeme, že limitný bod ľubovoľnej konvergentnej postupnosti bodov y_n kužeľa $\mathcal{K}^*$ patrí tiež do tohto kužeľa. Predpokladajme, že $y_n \in \mathcal{K}^*$ pre všetky n a $y_n \to y$ pre $n \to \infty$. Potom platí $\forall x \in \mathcal{K} : y_n^T x \geq 0$. Vidíme, že

$$\lim_{n\to\infty} x^T y_n = x^T y, \text{ a teda } x^T y \geq 0 \quad \forall x \in \mathcal{K}.$$

Takže $y \in \mathcal{K}^*$, z čoho vyplýva uzavretosť $\mathcal{K}^*$.

d) Označme $V := \{z | z^T x > 0 \quad \forall x \in cl\mathcal{K} \setminus \{0\}\}$.
1.) Najskôr ukážeme, že $int\mathcal{K}^* \subseteq V$.
Budeme dokazovať obmenu implikácie, t.j. $y \in \mathcal{K}^* \setminus V \Rightarrow y \notin int\mathcal{K}^*$.
Nech $y \in \mathcal{K}^* \setminus V$. Potom pre nejaké nenulové $x \in \mathcal{K}$ platí $y^T x = 0$. Uvažujme guľu $B(y, \varepsilon)$ so stredom y a ľubovoľným polomerom $\varepsilon > 0$. Zvolíme $u \in R^n$ tak, aby $u^T x < 0$ a zvolíme dostatočne malé číslo $\alpha > 0$ spĺňajúce $\|\alpha u\| = \alpha \|u\| < \varepsilon$.
Potom $y + \alpha u \in B(y, \varepsilon)$, ale

$$(y + \alpha u)^T x = y^T x + \alpha u^T x = 0 + \alpha u^T x < 0,$$

čo znamená, že $y + \alpha u \notin \mathcal{K}^*$, a teda $B(y, \varepsilon) \not\subseteq \mathcal{K}^*$. Toto platí pre všetky $\varepsilon > 0$, preto $y \notin int\mathcal{K}^*$.
2.) Teraz chceme ukázať, že $V \subseteq int\mathcal{K}^*$, t.j. ľubovoľný bod z V patrí aj do $int\mathcal{K}^*$.
Nech $y \in V$, t.j. $y^T x > 0 \quad \forall x \in cl\mathcal{K} \setminus \{0\}$. Z definície kužeľa platí

$$y^T x > 0 \text{ aj pre všetky } x \in (cl\mathcal{K} \setminus \{0\}) \cap S,$$

kde $S = \{x \in R^n | \|x\| = 1\}$ je jednotková guľa. Dá sa ukázať, že S je kompaktná a $(cl\mathcal{K} \setminus \{0\})$ je uzavretá množina. To znamená, že aj prienik $(cl\mathcal{K} \setminus \{0\}) \cap S$ je kompaktná množina. Potom z kompaktnosti vyplýva, že

$$\exists \varepsilon > 0 : (y + u)^T x > 0 \quad \forall x \in (cl\mathcal{K} \setminus \{0\}) \cap S \quad \forall u : \|u\| < \varepsilon.$$

Teda $B(y, \varepsilon) \subset V$, takže y je vnútorným bodom V. Z konštrukcie vyplýva, že $y, y + u \in \mathcal{K}^*$, t.j. $B(y, \varepsilon) \subset \mathcal{K}^*$. A teda $y \in int\mathcal{K}^*$, čo sme chceli ukázať.

e) Ukážeme, že ak má kužeľ $\mathcal{K}$ neprázdne vnútro, tak $\mathcal{K}^* \cap (-\mathcal{K}^*) = \{0\}$.
Nech $y = \mathcal{K}^* \cap (-\mathcal{K}^*)$. To znamená, že

$$x^T y \geq 0 \wedge x^T y \leq 0 \quad \forall x \in \mathcal{K},$$

teda $x^T y = 0 \quad \forall x \in \mathcal{K}$. Ďalej nech $int\mathcal{K} \neq \emptyset$, potom existuje neprázdna otvorená guľa $\mathcal{B} \subseteq \mathcal{K}$, pre ktorú $y^T x = 0 \quad \forall x \in \mathcal{B}$.
Predpokladajme, že $y \neq 0$. Potom existuje také $u \in R^n, u \neq 0$, že $y^T u > 0$.
Zvolíme ľubovoľné $x \in \mathcal{B}$ a keďže $\mathcal{B}$ je otvorená, existuje dostatočne malé $\alpha > 0$ tak, aby $x + \alpha u \in \mathcal{B}$. Dostávame teda

$$0 = y^T(x + \alpha u) = y^T x + \alpha y^T u = 0 + \alpha y^T u > 0,$$

čo je spor. Preto $y = 0$. Z toho vyplýva, že $0 \in \mathcal{K}^* \cap (-\mathcal{K}^*)$, takže $\mathcal{K}^* \cap (-\mathcal{K}^*) = \{0\}$, čo sme chceli ukázať.

f) 1.) Najprv ukážeme, že $cl\mathcal{K} \subseteq \mathcal{K}^{**}$.
Pre ľubovoľné $x \in \mathcal{K}$ platí $x^T y \geq 0 \quad \forall y \in \mathcal{K}^*$, čo vyplýva z definície duálneho kužeľa. Kužeľ $\mathcal{K}^{**}$ je duálnym kužeľom kužeľa $\mathcal{K}^*$, a teda

$$z \in \mathcal{K}^{**} \Leftrightarrow z^T y \geq 0 \quad \forall y \in \mathcal{K}^*.$$

Túto podmienku spĺňa aj naše ľubovoľne zvolené x, takže $x \in \mathcal{K}^{**}$. Ukázali sme, že $\mathcal{K} \subseteq \mathcal{K}^{**}$, a preto aj $cl\mathcal{K} \subseteq \mathcal{K}^{**}$.
2.) Teraz ukážeme, že $\mathcal{K}^{**} \subseteq cl\mathcal{K}$, t.j. ľubovoľne zvolené $z \in \mathcal{K}^{**}$ platí $z \in cl\mathcal{K}$.
Nech $z \in \mathcal{K}^{**}$. Predpokladajme, že $z \notin cl\mathcal{K}$. Potom podľa vety 6.1 existuje nenulové $c \in R^n$, pre ktoré platí

$$c^T z < 0 \quad \text{a} \quad c^T x \geq 0 \quad \forall x \in cl\mathcal{K}.$$

Z druhej nerovnosti vyplýva, že $c \in \mathcal{K}^*$. To je však v spore s prvou podmienkou $c^T z < 0$, pretože $z \in \mathcal{K}^{**}$, a teda $c^T z \geq 0 \quad \forall c \in \mathcal{K}^*$. Takže $z \in cl\mathcal{K}$, a teda sme ukázali $\mathcal{K}^{**} \subseteq cl\mathcal{K}$.
Spojením dvoch inklúzii dostaneme $\mathcal{K}^{**} = cl\mathcal{K}$. Ak je kužeľ $\mathcal{K}$ uzavretý, tak $\mathcal{K} = cl\mathcal{K}$, a preto $\mathcal{K}^{**} = K$.

g) Nech $cl\mathcal{K}$ je špicatý, t.j. $cl\mathcal{K} \cap (-cl\mathcal{K}) = \{0\}$. Ukážeme, že $0 \notin conv(cl\mathcal{K} \cap S)$, t.j. bod 0 nemôžeme napísať pomocou konvexnej kombinácie bodov z $cl\mathcal{K} \cap S$, kde S je jednotková guľa v R^n.
Predpokladajme, že $0 \in conv(cl\mathcal{K} \cap S)$. Potom

$$0 = \sum_{i=1}^{m} \theta_i x_i, \text{ kde pre všetky } i \text{ platí } \theta_i \geq 0, \sum_{i=1}^{m} \theta_i = 1 \text{ a } x_i \in cl\mathcal{K} \cap S.$$

Vyberieme index j, pre ktorý $\theta_j \neq 0$, potom

$$\theta_j x_j \in cl\mathcal{K} \text{ a } -\theta_j x_j = \sum_{i:i\neq j} \theta_i x_i \in cl\mathcal{K}.$$

To znamená, že $0 \neq \theta_j x_j \in cl\mathcal{K} \cap (cl\mathcal{K})$, čo je v spore s vlastnosťou špicatosti kužeľa. Preto $0 \notin conv(cl\mathcal{K} \cap S)$. Keďže $cl\mathcal{K} \cap S$ je kompaktná množina, tak aj jej konvexný obal je kompaktná množina, a preto môžeme použiť vetu 6.1.
Teda existuje nenulové $a \in R^n$, pre ktoré $a^T x \geq 0 \quad \forall x \in cl\mathcal{K} \cap S$ a $a^T 0 < b$, z čoho vyplýva $b > 0$. Zvolíme ľubovoľné nenulové $y \in cl\mathcal{K}$. Potom $\frac{y}{\|y\|} \in S$ a

$$a^T y = \|y\| a^T \frac{y}{\|y\|} \geq \|y\| b > 0,$$

takže $a \in \mathcal{K}^*$. Teraz ukážeme, že $B(a, b) \subset \mathcal{K}^*$. Nech $u \in R^n : \|u\| < b$ je ľubovoľné. Potom s použitím Cauchy-Schwarzovej nerovnosti platí

$$u^T y \geq -|u^T y| \geq -\|u\|\|y\|,$$

t.j. $\frac{u^T y}{\|y\|} \geq -\|u\|$, a teda

$$(a + u)^T y \geq a^T y + u^T y \geq b\|y\| - \|u\|\|y\| = (b - \|u\|)\|y\| \geq 0.$$

To znamená, že $B(a, b) \subset \mathcal{K}^*$, takže $a \in int\mathcal{K}^*$, a teda $int\mathcal{K}^* \neq \emptyset$.

$\square$

Poznámka 6.2. Myšlienka dôkazu Vety (6.2 d),e)) bola prevzatá z [25], iný spôsob overenia niektorých vlastností sa dá nájsť napr. v [9],[22].

Tvrdenie 6.2. *Nech $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$ je kartézsky súčin kužeľov $\mathcal{K}_1, \ldots, \mathcal{K}_n$. Potom platia nasledujúce vlastnosti:*

a) $\mathcal{K}$ je kužeľ.

b) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ konvexné, tak aj ich kartézsky súčin $\mathcal{K}$ je konvexný.

c) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ samoduálne, potom $\mathcal{K}$ je tiež samoduálny kužeľ.

d) Ak sú kužele $\mathcal{K}_1, \ldots, \mathcal{K}_n$ vlastné, potom aj $\mathcal{K}$ je vlastný kužeľ.

Dôkaz. a) Nech $x \in \mathcal{K}$ a nech $\alpha \geq 0$. Potom $x = (x_1; x_2; \ldots; x_n)$ pre ľubovoľné $x_1 \in \mathcal{K}_1, \ldots, x_n \in \mathcal{K}_n$. Keďže $\mathcal{K}_1, \ldots, \mathcal{K}_n$ sú kužele, tak $\alpha x_1 \in \mathcal{K}_1, \ldots, \alpha x_n \in \mathcal{K}_n$. Z toho vyplýva, že $\alpha x = (\alpha x_1; \alpha x_2; \ldots; \alpha x_n) \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$, a preto $\mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_n$ je kužeľ, čo sme chceli ukázať.

b) Ukážeme, že pre ľubovoľné $x, y \in \mathcal{K}$ a $\theta_1, \theta_2 \geq 0$ platí $\theta_1 x + \theta_2 y \in \mathcal{K}$.
Nech $x = (x_1; x_2; \ldots; x_n), y = (y_1; y_2; \ldots; y_n) \in \mathcal{K}$ sú ľubovoľne zvolené a $\theta_1, \theta_2 \geq 0$. Potom $x_1, y_1 \in \mathcal{K}_1; \ldots; x_n, y_n \in \mathcal{K}_n$. Z konvexnosti kužeľov $\mathcal{K}_1, \ldots, \mathcal{K}_n$ vyplýva

$$\theta_1 x_1 + \theta_2 y_1 \in \mathcal{K}_1, \ldots, \theta_1 x_n + \theta_2 y_n \in \mathcal{K}_n.$$

Takže $\theta_1 x + \theta_2 y = (\theta_1 x_1 + \theta_2 y_1; \ldots; \theta_1 x_n + \theta_2 y_n) \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$.

c) Podľa a) je kartézsky súčin kužeľov $\mathcal{K}$ kužeľ.
Označme $\mathcal{K}^* = (\mathcal{K}_1 \times \cdots \times \mathcal{K}_n)^*$ duálny kužeľ kužeľa $\mathcal{K}$.
Keďže $\mathcal{K}_1, \ldots, \mathcal{K}_n$ sú samoduálne kužele, tak platí $\mathcal{K}_1^* = \mathcal{K}_1, \ldots, \mathcal{K}_n^* = \mathcal{K}_n$.
Preto stačí ukázať, že $\mathcal{K}^* = \mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^*$. Potom po dosadení za $\mathcal{K}_i^*, i = 1, \ldots, n$, dostaneme $\mathcal{K}^* = \mathcal{K}$.
1.) Najprv ukážeme, že $\mathcal{K}^* \subseteq \mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^*$.
Z definície duálneho kužeľa vyplýva

$$\begin{aligned}\mathcal{K}^* &= \{x = (x_1; x_2; \ldots; x_n) | x^T z \geq 0 \quad \forall z = (z_1; z_2; \ldots; z_n) \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_n\} \\ &= \{x = (x_1; x_2; \ldots; x_n) | x_1^T z_1 + \cdots + x_n^T z_n \geq 0 \quad \forall z_1 \in \mathcal{K}_1, \ldots, z_n \in \mathcal{K}_n\}\end{aligned}$$

Zvolíme také j, že $z_i = 0 \quad \forall i \neq j$. Potom $x_j \in \mathcal{K}_j^*$. Takto môžeme postupovať pre $j = 1, \ldots, n$ a postupne dostaneme $x_1 \in \mathcal{K}_1^*, \ldots, x_n \in \mathcal{K}_n^*$.
Takže $\mathcal{K}^* \subseteq \mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^*$.
2.) Teraz ukážeme opačnú inklúziu, t.j. $\mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^* \subseteq \mathcal{K}^*$.
Nech $x_1 \in \mathcal{K}_1^*, \ldots, x_n \in \mathcal{K}_n^*$. Potom pre ľubovoľné $(z_1; z_2; \ldots; z_n) \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$ platí

$$(x_1; x_2; \ldots; x_n)^T (z_1; z_2; \ldots; z_n) = x_1^T z_1 + \cdots + x_n^T z_n \geq 0.$$

Z toho vyplýva, že $\mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^* \subseteq \mathcal{K}^*$.
Spojením oboch inklúzii dostaneme $\mathcal{K}^* = \mathcal{K}_1^* \times \cdots \times \mathcal{K}_n^*$, z čoho je už ľahko vidieť samoduálnosť kartézskeho súčinu samoduálnych kužeľov.

d) Ukážeme, že $\mathcal{K}$ spĺňa vlastnosti vlastného kužeľa, t.j. je uzavretý, konvexný, špicatý a $int\mathcal{K} \neq \emptyset$. Keďže každý z kužeľov $\mathcal{K}_1, \ldots, \mathcal{K}_n$ spĺňa tieto vlastnosti, tak je zrejme, že konvexnosť kužeľa $\mathcal{K}$ vyplýva z b). Ďalej $\exists x_i^0 \in int\mathcal{K}_i \quad \forall i = 1, \ldots, n$. Potom $x^0 = (x_1^0; \ldots; x_n^0) \in int\mathcal{K}$, a teda vnútro kužeľa $\mathcal{K}$ je neprázdne. Teraz ukážeme, že

$$z \in \mathcal{K} \cap (-\mathcal{K}) \text{ implikuje } z = 0.$$

Nech $z = (z_1; \ldots; z_n) \in \mathcal{K}$ a $-z \in \mathcal{K}$. To znamená, že $z_i \in \mathcal{K}_i \cap (-\mathcal{K}_i)$ pre $i = 1, \ldots, n$. Kužele $\mathcal{K}_i, i = 1, \ldots, n$ sú špicaté, preto $z_i = 0 \quad \forall i$, čo je ekvivalentné s $z = 0$. A nakoniec ostáva ukázať uzavretosť kužeľa $\mathcal{K}$.
Nech $\{p^m = (p_1^m; \ldots, p_n^m)\}_{m=1}^{\infty}$ je postupnosť bodov z $\mathcal{K}$ konvergujúca k $\tilde{p} = (\tilde{p}_1; \ldots, \tilde{p}_n)$. Pre všetky $i = 1, \ldots, n$ platí $p_i^m \in \mathcal{K}_i \to \tilde{p}_i$. Pretože $\mathcal{K}_i$ je uzavretý kužeľ, tak $\tilde{p}_i \in \mathcal{K}_i \ \forall i$, z čoho vyplýva $\tilde{p} \in \mathcal{K}$. A teda kužeľ $\mathcal{K}$ je uzavretý.
Podarilo sa nám ukázať, že $\mathcal{K}$ spĺňa všetky vlastnosti vlastného kužeľa.

□

Poznámka 6.3. Body *a*) a *c*) z predchádzajúceho tvrdenia sú rozšírením [7] a ([22],s.334) pre $n > 2$. Najdôležitejšou vlastnosťou je vlastnosť *d*), ktorá sa využije už pri formulácii SOCP úloh.

6.2 Operátor $\circ$ a geometria vektora x

Nech $x = (x_0; \bar{x}) \in R^p$ a $s = (s_0; \bar{s}) \in R^p$.
Zavedieme označenie

$$x \circ s = \begin{pmatrix} x^T s \\ x_0 \bar{s} + s_0 \bar{x} \end{pmatrix}. \tag{50}$$

Ďalej k vektoru x definujeme tzv. *arrow-shaped* maticu ako

$$Arw(x) = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix}.$$

Všimnime si, že potom (50) môžeme zapísať pomocou arrow-shaped matíc nasledovne

$$x \circ s = Arw(x)s = Arw(x)Arw(s)e,$$

kde $e = (1, 0, \ldots, 0)^T \in R^p$.
V nasledujúcom tvrdení spomenieme a overíme základné vlastnosti operátora $\circ$.

Tvrdenie 6.3. *Operátor $\circ$ spĺňa nasledujúce vlastnosti:*

1. *Nech $z = x \circ y$. Potom $\forall \alpha, \beta \in R$ platí*
$$x \circ (\alpha y + \beta z) = \alpha x \circ y + \beta x \circ z,$$
$$(\alpha y + \beta z) \circ x = \alpha y \circ x + \beta z \circ x.$$
2. $x \circ s = s \circ x$
3. *Nech $e = (1; 0) \in R^p$, kde $0 \in R^{p-1}$. Potom $x \circ e = x$.*
4. *Nech $x^2 = x \circ x$. Potom matice $Arw(x)$ a $Arw(x^2)$ komutujú. Naviac $\forall y \in R^p$ platí:*
$$x \circ (x^2 \circ y) = x^2 \circ (x \circ y).$$

Dôkaz. 1. Nech $z = x \circ y$. Potom úpravami na základe vzťahu (50) dostaneme

$$\begin{aligned} x \circ (\alpha y + \beta z) &= x \circ (\alpha y + \beta Arw(x)y) = Arw(x)(\alpha y + \beta Arw(x)y) = \\ &= \alpha Arw(x)y + \beta Arw(x)Arw(x)y = \alpha x \circ y + \beta Arw(x)z = \\ &= \alpha x \circ y + \beta x \circ z. \end{aligned}$$

Podobne ukážeme aj druhú rovnosť. Najprv ľavá strana rovnice:

$$\begin{aligned} (\alpha y + \beta z) \circ x &= \begin{pmatrix} \alpha y_0 + \beta x^T y \\ \alpha \bar{y} + \beta x_0 \bar{y} + \beta y_0 \bar{x} \end{pmatrix} \circ x = \\ &= \begin{pmatrix} \alpha y_0 + \beta x^T y & (\alpha \bar{y} + \beta x_0 \bar{y} + \beta y_0 \bar{x})^T \\ \alpha \bar{y} + \beta x_0 \bar{y} + \beta y_0 \bar{x} & (\alpha y_0 + \beta x^T y) I \end{pmatrix} \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix} = \\ &= \begin{pmatrix} \alpha x_0 y_0 + \beta x_0 x^T y + \alpha \bar{y}^T \bar{x} + \beta x_0 \bar{y}^T \bar{x} + \beta y_0 \bar{x}^T \bar{x} \\ \alpha x_0 \bar{y} + \beta x_0^2 \bar{y} + \beta y_0 x_0 \bar{x} + \alpha y_0 \bar{x} + \beta x^T y \bar{x} \end{pmatrix} \end{aligned}$$

Teraz vyjadríme pravú stranu rovnice:

$$\alpha y \circ x + \beta z \circ x = \alpha \begin{pmatrix} y^T x \\ y_0 \bar{x} + x_0 \bar{y} \end{pmatrix} + \beta \begin{pmatrix} x^T y & (x_0 \bar{y} + y_0 \bar{x})^T \\ x_0 \bar{y} + y_0 \bar{x} & x^T y I \end{pmatrix} \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha y^T x + \beta x_0 x^T y + \beta x_0 \bar{y}^T \bar{x} + \beta y_0 \bar{x}^T \bar{x} \\ \alpha y_0 \bar{x} + \alpha x_0 \bar{y} + \beta x_0^2 \bar{y} + \beta x_0 y_0 \bar{x} + \beta x^T y \bar{x} \end{pmatrix}$$

Vidíme, že ľavá strana=pravá strana, čo sme chceli ukázať.

2. Ľavá strana:

$$x \circ y = Arw(x) y = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix} \begin{pmatrix} y_0 \\ \bar{y} \end{pmatrix} = \begin{pmatrix} x_0 y_0 + \bar{x}^T \bar{y} \\ y_0 \bar{x} + x_0 \bar{y} \end{pmatrix}.$$

Pravá strana:

$$y \circ x = Arw(y) x = \begin{pmatrix} y_0 & \bar{y}^T \\ \bar{y} & y_0 I \end{pmatrix} \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix} = \begin{pmatrix} y_0 x_0 + \bar{y}^T \bar{x} \\ x_0 \bar{y} + y_0 \bar{x} \end{pmatrix}.$$

Vidíme, že ľavá strana=pravá strana, čo sme chceli ukázať.

3. Nech $e = (1; 0) \in R^p$, kde $0 \in R^{p-1}$. Potom

$$x \circ e = Arw(x) e = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix} = x.$$

4. Matice $Arw(x)$ a $Arw(x^2)$ komutujú, ak $Arw(x)Arw(x^2) = Arw(x^2)Arw(x)$. Pozrieme sa najprv na ľavú a potom na pravú stranu rovnice a dostaneme:

$$Arw(x)Arw(x^2) = \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix} \begin{pmatrix} x^T x & 2x_0 \bar{x}^T \\ 2x_0 \bar{x} & x^T x I \end{pmatrix} = \begin{pmatrix} x_0 x^T x + 2x_0 \bar{x}^T \bar{x} & 2x_0^2 \bar{x}^T + x^T x \bar{x}^T \\ x^T x \bar{x} + 2x_0^2 \bar{x} & 2x_0 \bar{x} \bar{x}^T + x_0 x^T x I \end{pmatrix}$$

$$Arw(x^2)Arw(x) = \begin{pmatrix} x^T x & 2x_0 \bar{x}^T \\ 2x_0 \bar{x} & x^T x I \end{pmatrix} \begin{pmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{pmatrix} = \begin{pmatrix} x_0 x^T x + 2x_0 \bar{x}^T \bar{x} & x^T x \bar{x}^T + 2x_0^2 \bar{x}^T \\ 2x_0^2 \bar{x} + x^T x \bar{x} & 2x_0 \bar{x} \bar{x}^T + x_0 x^T x I \end{pmatrix}$$

Opäť nastala rovnosť a to sme chceli ukázať.

Zvoľme ľubovoľné $y \in R^p$. Potom na základe predchádzajúceho výsledku platí

$$x \circ (x^2 \circ y) = x \circ (Arw(x^2) y) = Arw(x) Arw(x^2) y = Arw(x^2) Arw(x) y =$$
$$= Arw(x^2)(x \circ y) = x^2 \circ (x \circ y).$$

$\square$

6.2.1 Spektrálny rozklad vektora x

V snahe vytvoriť analógiu metód vnútorného bodu pre úlohy programovania nad kužeľmi druhého rádu s metódami vnútorného bodu v semidefinitnom programovaní, definuje sa podobná štruktúra pre vektor x ako poznáme pre matice. Teda definujeme pojmy vlastných čísel, spektrálneho rozkladu a podobne.
Uvažujme nasledujúcu rovnosť:

$$x^2 - 2x_0 x + (x_0^2 - \|\bar{x}\|^2)e = 0.$$

Ľahko overíme, že platí, pretože

$$\begin{pmatrix} x^T x \\ 2x_0\bar{x} \end{pmatrix} - 2x_0 \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix} + \begin{pmatrix} x_0^2 - \|\bar{x}\|^2 \\ 0 \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} \|x\|^2 - 2x_0^2 + x_0^2 - \|\bar{x}\|^2 \\ 2x_0\bar{x} - 2x_0\bar{x} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Potom môžeme definovať *charakteristický polynóm* $p(\lambda, x)$ ako

$$p(\lambda, x) = \lambda^2 - 2x_0\lambda + (x_0^2 - \|\bar{x}\|^2) = (\lambda - (x_0 + \|\bar{x}\|))(\lambda - (x_0 - \|\bar{x}\|)).$$

Korene rovnice $p(\lambda, x) = 0$ sú

$$\lambda_1 = x_0 + \|\bar{x}\| \quad \text{a} \quad \lambda_2 = x_0 - \|\bar{x}\|$$

a nazývajú sa *vlastné hodnoty* vektora x. To znamená, že každý prvok x má iba 2 reálne vlastné hodnoty.
Potom môžeme definovať *stopu* vektora x ako

$$tr(x) = \lambda_1 + \lambda_2 = x_0 + \|\bar{x}\| + x_0 - \|\bar{x}\| = 2x_0$$

a *determinant* vektora x

$$\det x = \lambda_1\lambda_2 = (x_0 + \|\bar{x}\|)(x_0 - \|\bar{x}\|) = x_0^2 - \|\bar{x}\|^2.$$

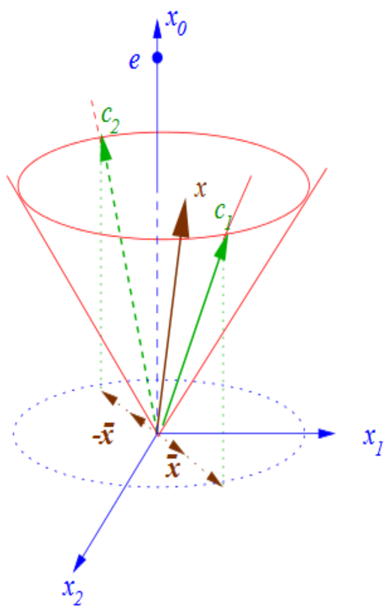
Spektrálny rozklad vektora x predstavuje rovnosť

$$x = \lambda_1 c_1 + \lambda_2 c_2,$$

kde

$$c_1 = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} \quad \text{a} \quad c_2 = \frac{1}{2}\begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix},$$

čo geometricky reprezentuje nasledujúci obrázok.



Obr. 4: Spektrálny rozklad x: Nech $x = (x_0; \bar{x}) \in R^p$. Projekciou x do roviny $x_1, \ldots, x_{p-1}$, normalizáciou $\bar{x}, -\bar{x}$ a následným zdvihnutím normalizovaných vektorov na hranicu kužeľa získame c_1, c_2([1],s20).

Dokážeme nasledujúce tvrdenie.

Tvrdenie 6.4. *Vektory* $c_1 = \frac{1}{2}\begin{pmatrix} 1 & \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}^T, c_2 = \frac{1}{2}\begin{pmatrix} 1 & -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}^T$ *spĺňajú nasledujúce vlastnosti:*

1. $c_1 \circ c_2 = 0,$
2. $c_1^2 = c_1, c_2^2 = c_2,$
3. $c_1 + c_2 = e,$
4. $c_1 = Rc_2, c_2 = Rc_1,$ *kde*
$$R = \begin{pmatrix} 1 & 0^T \\ 0 & -I \end{pmatrix},$$
5. $c_1, c_2 \in bd\mathcal{Q}.$

Dôkaz. 1.
$$c_1 \circ c_2 = \begin{pmatrix} \frac{1}{4}(1 - \frac{\bar{x}^T\bar{x}}{\|\bar{x}\|^2}) \\ -\frac{\bar{x}}{4\|\bar{x}\|} + \frac{\bar{x}}{4\|\bar{x}\|} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

2. Ukážeme $c_1^2 = c_1$ a analogicky sa spraví $c_2^2 = c_2$. Teda
$$c_1 \circ c_1 = \begin{pmatrix} \frac{1}{4}(1 + \frac{\bar{x}^T\bar{x}}{\|\bar{x}\|^2}) \\ \frac{\bar{x}}{4\|\bar{x}\|} + \frac{\bar{x}}{4\|\bar{x}\|} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} = c_1.$$

3. Ďalej
$$c_1 + c_2 = \frac{1}{2}\left[\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} + \begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e.$$

4. Počítajme
$$Rc_1 = \frac{1}{2}\begin{pmatrix} 1 & 0^T \\ 0 & -I \end{pmatrix}\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} = c_2,$$
$$Rc_2 = \frac{1}{2}\begin{pmatrix} 1 & 0^T \\ 0 & -I \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} = c_1,$$
čo sme chceli ukázať.

5. Nech $\bar{c} = \frac{\bar{x}}{2\|\bar{x}\|}$. Potom $c_1 = (\frac{1}{2}; \bar{c}), c_2 = (\frac{1}{2}; -\bar{c})$. Chceme ukázať, že $\|\bar{c}\| = \|-\bar{c}\| = \frac{1}{2}$.
Teda
$$\|\bar{c}\| = \left\|\frac{\bar{x}}{2\|\bar{x}\|}\right\| = \frac{\|\bar{x}\|}{2\|\bar{x}\|} = \frac{1}{2},$$
a preto $c_1, c_2 \in bd\mathcal{Q}$.

□

Ak $\det x \neq 0$, môžeme definovať
$$x^{-1} = \lambda_1^{-1} c_1 + \lambda_2^{-1} c_2 = \frac{1}{x_0^2 - \|\bar{x}\|^2}\begin{pmatrix} x_0 \\ -\bar{x} \end{pmatrix}.$$
Všimnime si, že $x \circ x^{-1} = e$.
Pre nezáporné číslo p ďalej definujeme p-tu mocninu vektora x ako
$$x^p = \lambda_1^p c_1 + \lambda_2^p c_2.$$
A nakoniec zavedieme normy:
$$\|x\|_F = \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{2}\|x\|,$$
$$\|x\|_2 = \max\{|\lambda_1|, |\lambda_2|\} = |x_0| + \|\bar{x}\|, \text{ak } x \in \mathcal{Q}.$$

Zatiaľ sme uvažovali, že $r = 1$, teda vektor $x = (x_0; \bar{x}) \in R^p$ je ako jeden blok.
Teraz nech $x = (x^1; \ldots; x^r) \in R^n$, kde $x^i \in R^{n_i}$.
Potom charakteristický polynóm definujeme ako
$$p(x) = \prod_{i=1}^{r} p_i(x^i).$$
To znamená, že vektor x bude mať $2r$ vlastných čísel
$$\lambda_1^i = x_0^i + \|\bar{x}^i\|, \quad \lambda_2^i = x_0^i - \|\bar{x}^i\|, \quad i = 1, \ldots, r,$$

takže

$$\det x = \prod_{i=1}^{r}((x_0^i)^2 - \|\bar{x}^i\|^2)$$

a

$$tr(x) = 2\sum_{i=1}^{r} x_0^i.$$

Vektor x v spektrálnom rozklade zapíšeme analogicky

$$x = \begin{pmatrix} \lambda_1^1 c_1^1 + \lambda_2^1 c_2^1 \\ \vdots \\ \lambda_1^r c_1^r + \lambda_2^r c_2^r \end{pmatrix}$$

a potom ak $\det x \neq 0$, tak definujeme

$$x^{-1} = ((x^1)^{-1}; \dots; (x^r)^{-1}),$$

a tiež pre nezáporné celé číslo p mocninu vektora x ako

$$x^p = ((x^1)^p; \dots; (x^r)^p).$$

A nakoniec zavedieme normy:

$$\|x\|_F^2 = \sum_{i=1}^{r} \|x^i\|_F^2,$$

$$\|x\|_2 = \max_i \|x^i\|.$$

Poznámka 6.4. Ďalšie vlastnosti operátora $\circ$ a rozšírenie geometrie vektora x sa dá nájsť v ([1], časť 4), pre naše účely to však nie je potrebné.

Záver

Cieľom tejto práce bolo spracovanie teórie pre úlohy programovania nad kužeľmi druhého rádu. Motiváciou pre analýzu tejto teórie boli možné aplikácie programovania nad kužeľmi druhého rádu v praxi. V prvom rade sme sa postupne oboznámili so všeobecnými kužeľmi a ich vlastnosťami, ako aj s kužeľmi druhého rádu. Zistili sme, že kužeľ druhého rádu je vlastný kužeľ a jeho základnou vlastnosťou je tzv. samoduálnosť. Vďaka tomu sa môže zadefinovať zovšeobecnená nerovnosť, ktorá tak vytvára analógiu s ohraničeniami v lineárnom a semidefinitnom programovaní.
Následne sme naformulovali úlohu programovania nad kužeľmi druhého rádu a k nej duálnu úlohu. Pri skúmaní duality sme zistili, že slabá veta o dualite platí podobne ako v lineárnom programovaní. Silná veta o dualite v úlohách programovania nad kužeľmi druhého rádu si na rozdiel od lineárneho programovania vyžaduje prísnejšiu podmienku, tzv. ostrú prípustnosť riešenia. Ukázali sme na príklade, čo sa môže stať, ak táto podmienka nie je splnená. A vetu sme dokázali pomocou zovšeobecnenej Farkašovej lemy, ktorá je analógiou zovšeobecnenej Farkašovej lemy pre semidefinitné programovanie. Pri podmienke komplementarity sme zaviedli nový operátor a vysvetlili interpretáciu tejto podmienky aj geometricky.
V nasledujúcej kapitole sme ukázali súvis medzi úlohami programovania nad kužeľmi druhého rádu a inými triedami konvexnej optimalizácie. Lineárne programovanie a kvadratické programovanie je podtriedou programovania nad kužeľmi druhého rádu, takže všetky úlohy lineárneho a kvadratického programovania možno previesť na úlohy programovania nad kužeľmi druhého rádu. Na druhej strane je programovanie nad kužeľmi druhého rádu podtriedou semidefinitného programovania ako to bolo spomenuté aj v úvode tejto práce.
Na záver sme sa zaoberali metódami vnútorného bodu. Ich myšlienkou bolo vytvoriť analógiu bariérovej funkcie ako v semidefinitnom programovaní, preto sa zaviedol nový pojem spektrálneho rozkladu vektora, ktorý sme podrobnejšie uviedli v Dodatku. Následne sme uviedli potrebné predpoklady, definovali sme pomocné bariérové úlohy a ukázali sme existenciu optimálneho riešenia týchto úloh. Nakoniec sme v algoritme opísali postup, akým sa priblížime k optimálnemu riešeniu pôvodnej primárnej a duálnej úlohy.
Jednou z možnosti rozšírenia tejto práce je venovať sa podrobnejšie vlastnostiam centrálnej limitnej trajektórie alebo skúmať rôzne voľby dĺžky kroku, resp. redukcie duálnej medzery v primárno-duálnom algoritme. Druhou alternatívou je venovať sa viac aplikáciám programovania nad kužeľmi druhého rádu, ale tým by sme prekročili rozsah práce.

Literatúra

[1] Alizadeh F., Goldfarb D.: *Second-order cone programming*, Springer, 2002

[2] Alizadeh F.: *Semidefinite and Second Order Cone Programming Seminar*, Fall 2012, Rutgers University, New Jersey, <http://rutcor.rutgers.edu/ alizadeh/CLASSES/12fallSDP/index.html>

[3] Alizadeh F.: *Semidefinite and Second Order Cone Programming*, Institute for Mathematics and its Applications - Workshop, University of Minnesota, 2003 <http://www.ima.umn.edu/talks/workshops/3-11.2003/alizadeh/tutorialProsper.pdf>

[4] Antoniou A., Wu-Sheng Lu: *Practical Optimization: Algorithms and Engineering Applications*, Chapter 14, Springer, 2007

[5] Bai Y.Q., Wang G.Q., Roos C.: *A new Primal-Dual Interior-Point Algorithm for Second-Order Cone Optimization*, Department of Mathematics, College Science, Shanghai University, Shanghai, 2004

[6] Ben-Tal A., Nemirovski A.: *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Series on Optimization, 2001

[7] Bertsekas D. P.: *Convex Analysis and Optimization*, Supplementary material: Chapter 1, Athena Scientific, 2003, <http://www.athenasc.com/convexity.html>

[8] Boyd S., Vandenberghe L.: *Convex optimization*, Cambridge University Press, 2004

[9] Dattorro J.: *Convex Optimization & Euclidean Distance Geometry*, Chapter 2, M$\varepsilon\beta$oo Publishing, 2005

[10] Cambini A., Martein L.: *Generalized Convexity and Optimization: Theory and Applications*, Springer, 2009

[11] Drewes S.: *Mixed Integer Second Order Cone Programming*, Dissertation 2009, Technischen Universität, Darmstadt

[12] Engels M.: *Portfolio Optimization: Beyond Markowitz*, Master's Thesis, 2004, Universiteit Leiden, Netherlands

[13] Freund R. M.: *Truss Design and Convex Optimization*, Massachusetts Institute of Technology, 2004

[14] Gärtner B., Matoušek J.: *Approximation Algorithms and Semidefinite Programmming*, Springer, 2012

[15] Halická M.: *Semidefinitné programovanie*, Prednášky 2011/2012, MFF UK, Bratislava

[16] Harville D.A.: *Matrix Algebra From a Statistician's Perspective*, Springer, 1997

[17] Hladík Ľ.: *Semidefinitné programovanie*, Diplomová práca 2001, MFF UK, Bratislava

[18] Lobo M. S., Vandenberghe L., Boyd S., Lebret H.: *Applications of second-order cone programming*, Linear Algebra and its Applications, 1998, vol.284, p.193-228

[19] Mishra D.: *Introduction To Convex Sets with Applications to Economics*, 2011 <http://www.isid.ac.in/ dmishra/mpdoc/lecconv11.pdf>

[20] Renegar J.: *A Mathematical View of Interior-Point Methods in convex Optimization*, MPS-SIAM Series on Optimization, 2001

[21] Sim Ch-K., Zhao G.: *A Note on Treating Second Order Cone Problem as a Special Case of Semide nite Problem*, JournalMathematical Programming: Series A and B, Volume 102 Issue 3, April 2005, Pages 609 - 613, Springer

[22] Stewart D.E.: *Dynamics with Inequalities: Impacts and Hard Constraints*, SIAM, 2011

[23] Terlaky T., Wang Z.: *On Identification of the Optimal Partition of Second Order Cone Optimization Problem*, Lehigh University

[24] Trnovská M.: *Centrálna trajektória v semidefinitnom programovaní*,Písomná práca k dizertačnej skúške, MFF UK, Bratislava, 2005

[25] Wolkowicz H.: *Convex Optimization and Analysis*, CO463/663, Fall 2012, University of Waterloo, Canada <http://orion.math.uwaterloo.ca/ hwolkowi/henry/teaching/f12/663.f12/663priv.d/ass312.pdf>

[26] Ye Y.: *Conic Linear Optimization*, Course Winter 2012-2013, Standford University, California, <http://www.stanford.edu/ yyye/>

[27] <http://cvxr.com/cvx/>