

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



CENTRÁLNA TRAJEKTÓRIA V LINEÁRNOM  
PROGRAMOVANÍ

BAKALÁRSKA PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

**CENTRÁLNA TRAJEKTÓRIA V LINEÁRNOM  
PROGRAMOVANÍ**

**BAKALÁRSKA PRÁCA**

Študijný program: Ekonomická a finančná matematika  
Študijný odbor: 1114 Aplikovaná matematika  
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky  
Vedúci práce: doc. RNDr. Margaréta Halická, CSc.



Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

---

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Matej Poliak  
**Študijný program:** ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)  
**Študijný odbor:** 9.1.9. aplikovaná matematika  
**Typ záverečnej práce:** bakalárska  
**Jazyk záverečnej práce:** slovenský

**Názov:** Centrálna trajektória v lineárnom programovaní.

**Cieľ:** Popísať rôzne spôsoby charakterizácie centrálnej trajektórie, analyzovať vzťahy medzi nimi. V prípade dvojrozmernej úlohy ilustrovať graficky.

**Vedúci:** doc. RNDr. Margaréta Halická, CSc.

**Dátum zadania:** 16.10.2011

**Dátum schválenia:** 27.10.2011

doc. RNDr. Margaréta Halická, CSc.  
garant študijného programu

.....  
študent

.....  
vedúci

**Pod'akovanie** Touto cestou sa chcem poďakovať svojej vedúcej bakalárskej práce Doc. RNDr. Margaréte Halickej, CSc. za ochotu, pomoc, odborné rady a podnetné pripomienky, ktoré boli veľmi cenné písaní tejto práce. Ďakujem aj svojej rodine za neochvejnú podporu.

## Abstrakt v štátnom jazyku

Poliak, Matej: Centrálna trajektória v lineárnom programovaní [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; vedúci: doc. RNDr. Margaréta Halická, CSc., Bratislava, 2012, 55 s.

V práci sa zaoberáme alternatívnymi charakterizáciami centrálnej trajektórie v lineárnom programovaní. Na úvod uvádzame niektoré základné poznatky teórie lineárneho programovania. Nasleduje úvod do metód vnútorného bodu vrátane definície centrálnej trajektórie a jej známa reprezentácia cez  $\mu$ -centrujúci systém. Definujeme analytický stred množiny pre polyédrické množiny a ukážeme existenciu a jednoznačnosť analytického stredy pre množiny vystupujúce v úlohe lineárneho programovania. Na dvojrozmernom príklade pozorujeme, že analytické stredy úrovňových množín duálnej účelovej funkcie sú zhodné s duálnou centrálnou trajektóriou. Dokážeme, že táto vlastnosť platí vo všeobecnosti pre analytické stredy úrovňových množín primárnej aj duálnej účelovej funkcie a tiež podúrovňových množín primárnej a nadúrovňových množín duálnej účelovej funkcie.

**Kľúčové slová:** Lineárne programovanie, Metódy vnútorného bodu, Centrálna trajektória, Analytický stred množiny

## Abstract

Poliak, Matej: Central Path in Linear Optimization [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. RNDr. Margaréta Halická, CSc., Bratislava, 2012, 55 p.

In our work we deal with alternative ways of characterizing the central path in linear optimization. First, we summarize some fundamental results of linear programming theory. Introduction into interior point methods follows along with definition of central path and its well known characterization as the solution of the system of  $\mu$ -centering equations. We define analytic center of a polytope and we show existence and uniqueness of analytic center of sets appearing in linear programming problems. We observe on a two-dimensional example, that analytic centres of level sets of the dual objective function are the same points as the dual central path. We prove, that this is a general property of level sets of dual and primal objective function.

**Keywords:** Linear optimization, Interior point methods, Central path, Analytic center of a set

## Obsah

Zoznam obrázkov	8
Zoznam tabuliek	9
Zoznam použitých symbolov	10
Úvod	11
<b>1 Úloha lineárneho programovania a základy teórie duality</b>	<b>13</b>
<b>2 Centrálna trajektória v LP</b>	<b>19</b>
2.1 Základné vlastnosti . . . . .	19
2.2 Dvojrozmerný príklad . . . . .	24
<b>3 CT ako analytické stredy</b>	<b>27</b>
3.1 Analytický stred množiny . . . . .	27
3.2 CT a úrovňové množiny . . . . .	35
3.3 CT a nad- a podúrovňové množiny . . . . .	48
Záver	54
Zoznam použitej literatúry	55

## Zoznam obrázkov

1	Úrovňové množiny logaritmickéj bariérovej funkcie pre rôzne $\mu$ . . . . .	21
2	Duálna centrálna trajektória v pr. 1 pre $b = (2, 1)^T$ . . . . .	26
3	Ohraničená $\mathcal{D}$ . . . . .	30
4	Neohraničená $\mathcal{D}$ . . . . .	30
5	Vrstevnice $g_S(y_1, y_2)$ . . . . .	36
6	Úrovňové množiny $b^T y = \gamma, y \in \mathcal{D}$ . . . . .	36
7	Hľadanie maxima $g_S$ na $\mathcal{D}_\gamma$ . . . . .	37
8	Maximá $g_S$ na rôznych $\mathcal{D}_\gamma$ . . . . .	37
9	Dôkaz lemy 3.11 . . . . .	41
10	Analytické stredy $\mathcal{D}_\gamma$ v príklade 3 . . . . .	46
11	Hľadanie analytických stredov $\mathcal{D}_{\gamma+}$ pre množinu z príkladu 1 . . . . .	49
12	Analytické stredy $\mathcal{D}_{\gamma+}$ (pr. 1) . . . . .	49
13	Hľadanie analytických stredov $\mathcal{D}_{\gamma+}$ pre množinu z príkladu 3 . . . . .	50
14	Analytické stredy $\mathcal{D}_{\gamma+}$ (pr. 3) . . . . .	50



## Zoznam tabuliek

1	Zhrnutie výsledkov viet 3.12 až 3.15 . . . . .	54
---	--	----

## Zoznam použitých symbolov

- LP lineárne programovanie
- CT centrálna trajektória
- ( $P$ ) primárna úloha lineárneho programovania  $\min \{c^T x \mid Ax = b, x \geq 0_n\}$
- $\mathcal{P}$  primárna prípustná množina  $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0_n\}$
- $\mathcal{P}^0$  množina vnútorných bodov  $\mathcal{P}$ ,  $\mathcal{P}^0 = \{x \in \mathcal{P} \mid x > 0_n\}$
- $\mathcal{P}_\delta$  úrovňová množina primárnej účelovej funkcie,  $\{x \in \mathcal{P} \mid c^T x = \delta\}$
- $\mathcal{P}_{\delta-}$  podúrovňová množina primárnej účelovej funkcie,  $\{x \in \mathcal{P} \mid c^T x \leq \delta\}$
- ( $D$ ) duálna úloha lineárneho programovania  $\max \{b^T y \mid A^T y + s = c, s \geq 0_n\}$
- $\mathcal{D}$  duálna prípustná množina  $\{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^T y + s = c, s \geq 0_n\}$
- $\mathcal{D}^0$  množina vnútorných bodov  $\mathcal{D}$ ,  $\mathcal{D}^0 = \{(y, s) \in \mathcal{D} \mid s > 0_n\}$
- $\mathcal{D}_\gamma$  úrovňová množina duálnej účelovej funkcie,  $\{(y, s) \in \mathcal{D} \mid b^T y = \gamma\}$
- $\mathcal{D}_{\gamma+}$  nadúrovňová množina duálnej účelovej funkcie,  $\{(y, s) \in \mathcal{D} \mid b^T y \geq \gamma\}$

## Úvod

Lineárne programovanie je disciplína v rámci širšieho spektra matematického programovania a zaoberá sa hľadaním optimálnej hodnoty lineárnej účelovej funkcie na množine danej lineárnymi ohraničeniami. Vývoj lineárneho programovania bol podnietený záujmom riešiť praktické úlohy ekonomického plánovania, rozloženia zdrojov [1], [9]. Za prvých významných priekopníkov sú považovaní Leonid Kantorovič (1939) a George Dantzig, známy vypracovaním simplexovej metódy (1947). Simplexová metóda umožňuje rutinne riešiť úlohy lineárneho programovania a spočíva v presúvaní sa po vrcholoch prípustnej množiny tak, že hodnota účelovej funkcie sa zlepšuje, tzn. znižuje pri úlohách minimalizačných a zväčšuje pri úlohách maximalizačných. Klee a Minty [5] ale poukázali na príklad, v ktorom simplexová metóda navštívi každý z  $2n$  vrcholov „zdeformovanej“  $n$ -rozmernej hyperkocky, kým doputuje do posledného, optimálneho vrchola, teda môže vyžadovať až  $2^n - 1$  iterácií. Aj keď takto náročné príklady v praxi pozorované nie sú, pri úlohách veľkých rozmerov sa vyskytujú problémy s trvaním výpočtov a numerickou presnosťou [6]. S prvým návrhom polynomiálneho algoritmu prišiel Chačijan (1979), ale skutočná revolúcia, nielen v lineárnom programovaní ale v celej optimalizácii, prišla s algoritmom Karmarkara (1984) [2]. Na základe Karmarkarovych myšlienok začal prudký vývoj tzv. metód vnútorného bodu. K téme vzniklo veľké množstvo prác a aj učebnice, napríklad [8], [10], ktoré mnohé práce citujú. Na rozdiel od simplexovej metódy prebiehajú iterácie vo vnútri množiny prípustných riešení. To sa zabezpečí transformáciou účelovej funkcie za použitia násobkov špeciálnej, tzv. bariérovej funkcie. Takéto transformované funkcie sa následne dajú minimalizovať, pričom z bodov miním týchto funkcií vznikne krivka nazývaná centrálna trajektória konvergujúca k optimálnemu riešeniu. Algoritmy vnútorného bodu centrálnu trajektóriu sledujú a približujú sa k limitnému bodu centrálnej trajektórie — ostro komplementárnemu optimálnemu riešeniu. Hlavnou výhodou je výpočtová zložitosť, ktorá je pre používané algoritmy z tejto triedy polynomiálna. To teoreticky značí časovú úsporu oproti exponenciálne zložitej simplexovej metóde najmä pre úlohy veľkých rozmerov — napríklad [10, str. 366] uvádza príklad úlohy s rádovo tisíc premennými a ohraničeniami, ktorú simplexová metóda rieši desiatky hodín a algoritmus vnútorného bodu len niekoľko minút. Keďže v súčasnej praxi existuje potreba riešiť úlohy takýchto, aj omnoho väčších

rozmerov, skúmať metódy vnútorného bodu ako efektívne metódy riešenia sa ukazuje ako opodstatnené.

V tejto práci sa zaoberáme predovšetkým už spomínanou centrálnou trajektóriou. Práca môže slúžiť aj ako určité zhrnutie niektorých základných poznatkov o téme, pričom dôraz bude okrem iného kladený na názorné grafické zobrazenie vzorových úloh, čo by mohlo vzbudiť v čitateľovi záujem o tému. Pred tým ako uvedieme definíciu centrálnej trajektórie, zopakujeme základné výsledky teórie lineárneho programovania. Ako východiskovú literaturu tu používame [8]. Na niekoľkých jednoduchých dvojrozmerných príkladoch graficky ilustrujeme, ako môže centrálna trajektória vyzeráť a skúsime odpozorovať nejaké jej vlastnosti. Zaujímať nás budú predovšetkým iné, ekvivalentné definície centrálnej trajektórie, ku ktorým sa dostaneme cez pojem analytického stredu množiny.

# 1 Úloha lineárneho programovania a základy teórie duality

V tejto kapitole uvádzame definíciu úlohy lineárneho programovania (LP) a základné tvrdenia, ktoré sú potrebné pri ďalšom štúdiu subjektu lineárneho programovania.

**Definícia 1.1.** (Primárnou) úlohou lineárneho programovania (P) *rozumieme úlohu*

$$(P) \quad \min \{c^T x \mid Ax = b, x \geq 0_n\},$$

kde  $A$  je matica rozmerov  $m \times n$ , vektory  $x, c \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ . Funkcia  $c^T x : \mathbb{R}^n \rightarrow \mathbb{R}$  ktorú minimalizujeme, sa nazýva (primárna) účelová funkcia a množina  $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  je tzv. množina (primárne) prípustných riešení.

Inak povedané, úlohou lineárneho programovania je nájsť taký vektor  $\hat{x}$  (pokiaľ existuje) z prípustnej množiny  $\mathcal{P}$ , že hodnota účelovej funkcie  $c^T \hat{x}$  je na prípustnej množine minimálna, teda  $c^T \hat{x} \leq c^T x, \quad \forall x \in \mathcal{P}$ .

*Poznámka:* Všeobecný tvar úlohy lineárneho programovania vyzerá nasledovne:

$$\min \{c_1^T x^1 + c_2^T x^2 \mid A_1 x^1 + A_2 x^2 \geq a, B_1 x^1 + B_2 x^2 = b, x^1 \geq 0_{n_1}, x^2 \in \mathbb{R}^{n_2}\}.$$

Podľa [7] sa dá ľubovoľná úloha LP transformovať na úlohu v tvare (P) vhodným zavedením nových premenných, čiže nám sa stačí zaoberať úlohami v tvare (P). Podobne sa nemusíme zvlášť zaoberať úlohami maximalizačnými, keďže vo všeobecnosti úloha  $\max \{f(x) \mid M\}$  je ekvivalentná s úlohou  $\min \{-f(x) \mid M\}$ .

**Definícia 1.2.** Duálnou úlohou (D') k úlohe (P) *rozumieme úlohu:*

$$(D') \quad \max \{b^T y \mid A^T y \leq c\}.$$

Zavedením tzv. doplnkovej premennej  $s$  dostaneme ekvivalentnú úlohu (viac o význame  $s$  je píšeme vo vete 1.14 na konci kapitoly). Vzniknú tým ohraničenia v tvare rovnosti:

**Definícia 1.3.** Duálnou úlohou (D) k úlohe (P) je úloha:

$$(D) \quad \max \{b^T y \mid A^T y + s = c, \quad s \geq 0_n\}.$$

$b^T y$  nazývame duálnou účelovou funkciou a množinu  $\mathcal{D} = \{(y, s) \in \mathbb{R}^{m+n} \mid A^T y + s = c, s \geq 0\}$  duálne prípustnou množinou.

*Poznámka:* Z lineárnej algebry sú známe nasledujúce tvrdenia. Ak má matica  $B$  lineárne nezávislé stĺpce, matica  $B^T B$  je regulárna. Navyše matica  $B^* = (B^T B)^{-1} B^T$  je k  $B$  ľavá inverzná.

V ďalších kapitolách budeme predpokladať, že hodnosť matice  $A$  je  $m \leq n$ , čiže jej riadky sú lineárne nezávislé. K  $A^T$  preto existuje ľavá inverzná matica  $A^{T*}$ . V  $A^T y + s = c$  existuje medzi  $s$  a  $c$  jedno-jednoznačný vzťah:  $s = c - A^T y$ ,  $y = A^{T*}(c - s)$ . Preto budeme pre skrátenie často hovoriť iba o jednom vektore z dvojice  $(y, s) \in \mathcal{D}$ , pričom druhý vektor sa považuje za určený spomenutými vzťahmi (napr. zápisom  $s \in \mathcal{D}$  myslíme  $(y, s) \in \mathcal{D}, y = A^{T*}(c - s)$ ).

**Definícia 1.4.** Ak  $x \in \mathcal{P}$ , hovoríme, že  $x$  je primárne prípustné (riešenie, vektor). Množinu  $\mathcal{P}^0 = \{x \in \mathcal{P} \mid x > 0\}$  nazývame množinou vnútorných bodov primárnej úlohy. Množinu  $\mathcal{P}^* = \{\hat{x} \in \mathcal{P} \mid c^T \hat{x} \leq c^T x, \quad \forall x \in \mathcal{P}\}$  nazývame množinou optimálnych riešení primárnej úlohy.

Ak  $(y, s) \in \mathcal{D}$ , v zmysle predchádzajúcej poznámky stotožňujeme  $(y, s)$ ,  $y$ ,  $s$  a hovoríme, že sú duálne prípustné (riešenia, vektory). Množinu  $\mathcal{D}^0 = \{(y, s) \in \mathcal{D} \mid s > 0\}$  nazývame množinou vnútorných bodov duálnej úlohy. Množinu  $\mathcal{D}^* = \{(\hat{y}, \hat{s}) \in \mathcal{D} \mid b^T \hat{y} \geq b^T y, \quad \forall y \in \mathcal{D}\}$  nazývame množinou optimálnych riešení duálnej úlohy.

Hovoríme, že primárna (duálna) úloha je prípustná, ak  $\mathcal{P} \neq \emptyset$  ( $\mathcal{D} \neq \emptyset$ ). Primárna (duálna) úloha sa ďalej nazýva neohraničená, ak je prípustná, ale nemá optimálne riešenie, teda hodnoty účelovej funkcie nadobúdajú na prípustnej množine ľubovoľne malé (veľké) hodnoty.

**Veta 1.5** (Slabá veta o dualite). Pre každé primárne prípustné  $x$  a každé duálne prípustné  $(y, s)$  platí  $x^T s \geq 0$ , resp.  $c^T x \geq b^T y$ .

*Dôkaz.* Keďže  $x$  je prípustné,  $x \geq 0$ . Keďže  $s$  je prípustné,  $s \geq 0$ . Preto  $x^T s \geq 0$ . Po rozpísaní  $s$  ako  $s = c - A^T y$  dostávame:  $0 \leq x^T s = x^T (c - A^T y) = x^T c - x^T A^T y = x^T c - (Ax)^T y = x^T c - b^T y$ , lebo  $x$  je prípustné a teda  $Ax = b$ .  $\square$

Nasleduje niekoľko viet, ktoré tu uvádzame bez dôkazu. Sú to základné výsledky z teórie duality a majú svoje miesto snáď v každej učebnici či prednáške lineárneho programovania, napr. [8, str. 89], [10, od str. 60], [7].

**Veta 1.6** (Silná veta o dualite). *Ak má jedna z úloh (P), (D) optimálne riešenie, má optimálne riešenie aj druhá, pričom hodnoty účelových funkcií v optimálnych riešeniach sa rovnajú.*

**Veta 1.7.** *Ak  $\mathcal{P} \neq \emptyset$  a  $\mathcal{D} \neq \emptyset$ , potom (P) aj (D) má optimálne riešenie.*

*Pre dvojicu úloh (P) a (D) môže nastať práve jeden z nasledujúcich prípadov:*

- *obidve úlohy majú optimálne riešenie;*
- *jedna z dvojice úloh je neohraničená a druhá neprípustná;*
- *obidve úlohy sú neprípustné.*

**Veta 1.8** (o komplementarite). *Pre ľubovoľnú dvojicu optimálnych riešení  $x^* \in \mathcal{P}^*$ ,  $(y^*, s^*) \in \mathcal{D}^*$  platí  $x^{*T} s^* = 0$  (resp.  $x_i^* s_i^* = 0$ ). Tiež ak pre dve prípustné riešenia  $x \in \mathcal{P}$ ,  $(y, s) \in \mathcal{D}$  platí  $x^T s = 0$ , tieto riešenia sú optimálne.*

**Definícia 1.9** (Ostrá komplementarita). *Optimálne riešenia  $x^* \in \mathcal{P}^*$ ,  $(y^*, s^*) \in \mathcal{D}^*$  sa nazývajú ostro komplementárne, ak pre každé  $i = 1, 2, \dots, n$  platí dvojica implikácií:*

$$x_i^* = 0 \implies s_i^* > 0,$$

$$s_i^* = 0 \implies x_i^* > 0.$$

*Inak zapísané, platí  $x^* + s^* > 0_n$ .*

**Veta 1.10** (o ostrej komplementarite). *Ak majú (P) a (D) optimálne riešenia, existuje dvojica ostro komplementárnych riešení.*

Keďže sa zaoberáme tzv. metódami vnútorného bodu, v nasledovných kapitolách budeme predpokladať, že prípustné množiny majú vnútorné body, teda  $\mathcal{P}^0 \neq \emptyset$ ,  $\mathcal{D}^0 \neq \emptyset$ .

K týmto podmienkam, ako hovorí nasledovná veta, existujú ekvivalentné podmienky hovoriace o ohraničenosti podmnožín  $\mathcal{P}$ ,  $\mathcal{D}$  definovaných v nasledovne definícii.

**Definícia 1.11.** Pre  $\delta \in \mathbb{R}$  definujeme úrovnové množiny primárnej účelovej funkcie

$$\mathcal{P}_\delta = \{x \in \mathcal{P} \mid c^T x = \delta\} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0_n, c^T x = \delta\}$$

a podúrovnové množiny primárnej účelovej funkcie

$$\mathcal{P}'_{\delta-} = \{x \in \mathcal{P} \mid c^T x \leq \delta\},$$

resp. po zavedení doplnkovej premennej  $\rho$

$$\mathcal{P}_{\delta-} = \{(x, \rho) \in \mathbb{R}^n \times \mathbb{R} \mid Ax = b, x \geq 0_n, c^T x + \rho = \delta, \rho \geq 0\}.$$

Podobne, pre  $\gamma \in \mathbb{R}$  definujeme úrovnové množiny duálnej účelovej funkcie

$$\mathcal{D}_\gamma = \{(y, s) \in \mathcal{D} \mid b^T y = \gamma\} = \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^T y + s = c, s \geq 0_n, b^T y = \gamma\}$$

a nadúrovnové množiny duálnej účelovej funkcie

$$\mathcal{D}'_{\gamma+} = \{y \in \mathcal{D} \mid b^T y \geq \gamma\},$$

čiže po zavedení doplnkovej premennej  $\sigma$  dostávame

$$\mathcal{D}_{\gamma+} = \{(y, s, \sigma) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \mid A^T y + s = c, s \geq 0_n, b^T y - \sigma = \gamma, \sigma \geq 0\}.$$

Horným indexom 0 označujeme podobne ako pre vlastné  $\mathcal{P}$  a  $\mathcal{D}$  množinu vnútorných bodov príslušnej množiny, teda

$$\mathcal{P}_\delta^0 = \{x \in \mathcal{P}_\delta \mid x > 0\},$$

$$\mathcal{P}'_{\delta-}{}^0 = \{(x, \rho) \in \mathcal{P}'_{\delta-} \mid x > 0, \rho > 0\},$$

$$\mathcal{D}_\gamma^0 = \{(y, s) \in \mathcal{D}_\gamma \mid s > 0\},$$

$$\mathcal{D}'_{\gamma+}{}^0 = \{(y, s, \sigma) \in \mathcal{D}'_{\gamma+} \mid s > 0, \sigma > 0\}.$$

*Poznámka:* Hoci by sme mali hovoriť o dvojici  $(x, \sigma) \in \mathcal{P}_{\delta-}$ , často budeme pre skrátenie označovať  $x \in \mathcal{P}_{\delta-}$ , pričom  $\rho$  je určené jednoznačne,  $\rho = \delta - c^T x > 0$  ak  $(x, \rho) \in \mathcal{P}_{\delta-}$ . Analogicky pre  $\mathcal{D}_{\gamma+}$ , kde  $\sigma = b^T y - \gamma$ .

Je zrejmé, že pre niektoré hodnoty  $\gamma$  alebo  $\delta$  sú množiny z definície 1.11 prázdne. Tie považujeme definatoricky za ohraničené.



**Veta 1.12.** *Nech prístupné množiny  $\mathcal{P}$  aj  $\mathcal{D}$  sú neprázdne. Potom sú nasledujúce tvrdenia ekvivalentné:*

1. množina  $\mathcal{D}^0$  je neprázdna;
2. množiny  $\mathcal{D}_{\gamma+}$  sú ohraničené pre každé  $\gamma \in \mathbb{R}$ ;
3. množina  $\mathcal{D}^*$  je ohraničená.

**Veta 1.13.** *Nech prístupné množiny  $\mathcal{P}$  aj  $\mathcal{D}$  sú neprázdne. Potom sú nasledujúce tvrdenia ekvivalentné:*

1. množina  $\mathcal{D}^0$  je neprázdna;
2. množiny  $\mathcal{P}_{\delta-}$  sú ohraničené pre každé  $\delta \in \mathbb{R}$ ;
3. množina  $\mathcal{P}^*$  je ohraničená.

Podrobné dôkazy možno nájsť v [8, str. 100-103].

*Poznámka:* Keďže  $\mathcal{P}_\delta \subseteq \mathcal{P}_{\delta-}$  a  $\mathcal{D}_\gamma \subseteq \mathcal{D}_{\gamma+}$ , z ohraničenosti  $\mathcal{P}_{\delta-}$  vyplýva ohraničenosť  $\mathcal{P}_\delta$  a z ohraničenosti  $\mathcal{D}_{\gamma+}$  vyplýva ohraničenosť  $\mathcal{D}_\gamma$ .

Nasledujúca veta sa zaoberá vektormi typu  $s \geq 0$  z ohraničení duálnej prípustnej množiny. Pohľadom na jednotlivé ohraničenia,  $a_i^T y + s_i = c_i$  možno intuitívne vytušiť, že  $s_i$  predstavuje vzdialenosť od hrany  $a_i^T y = c_i$  ( $a_i$  sme označili  $i$ -ty stĺpec matice  $A$ ). Ako je to presne, objasní nasledujúca veta.

**Veta 1.14** (Interpretácia  $s$ ). *Majme nadrovinu  $N$  danú  $N = \{x \in \mathbb{R}^n \mid a^T x = c\}$ . Vektor  $a$  je pre  $N$  normálový, teda je na nadrovinu  $N$  kolmý. Nadroviny  $a^T x + \sigma = c$ ,  $a^T x - \sigma = c$ ,  $\sigma > 0$  sú s  $N$  rovnobežné a vzdialenosť každej z nich od  $N$  je  $\frac{\sigma}{\|a\|_2}$ . Nadrovina  $a^T x + \sigma = c$  sa nachádza v afínnom polpriestore určenom hranicou  $N$  a smerom  $-a$ , nadrovina  $a^T x - \sigma = c$  zas v polpriestore určenom hranicou  $N$  a vektorom  $a$ .*

*Dôkaz.* Fakt, že vektor  $a$  je pre nadrovinu  $\{x \in \mathbb{R}^n \mid a^T x = c\}$  normálový, je dobre známym tvrdením z analytickej geometrie. Rovnaký normálový vektor majú aj nadroviny

$a^T x = c - \sigma$ ,  $a^T x = c + \sigma$ , sú teda s  $N$  rovnobežné. Majme bod  $\bar{x}$  z nadroviny  $N$ , teda  $a^T \bar{x} = c$ . Zostrojíme priamku  $p$  z bodu  $\bar{x}$  kolmú na  $N$ , čiže so smerovým vektorom  $a$ .

$$p : \bar{x} + \lambda a.$$

Hľadáme prienik  $p$  a nadroviny  $N' : a^T x = c + \sigma'$ ,  $\sigma' \in \mathbb{R}$ :

$$a^T(\bar{x} + \lambda a) = c + \sigma',$$

$$a^T \bar{x} + \lambda a^T a = c + \sigma',$$

$$\lambda = \frac{c + \sigma' - a^T \bar{x}}{a^T a} = \frac{c + \sigma' - c}{a^T a} = \frac{\sigma'}{a^T a} = \frac{\sigma'}{\|a\|_2^2}.$$

Teda priamka  $p$  pretína nadroviny  $N$  a  $N'$  na ňu kolmé v bodoch  $\bar{x}$  a  $\bar{x} + \frac{\sigma'}{\|a\|_2^2} a$ . Dĺžka tejto úsečky, teda vzdialenosť  $N$  a  $N'$  je

$$\left\| \bar{x} - \left( \bar{x} + \frac{\sigma'}{\|a\|_2^2} a \right) \right\| = \left\| \frac{\sigma'}{\|a\|_2^2} a \right\| = \frac{|\sigma'|}{\|a\|_2^2} \|a\| = \frac{|\sigma'|}{\|a\|}.$$

Keď teda máme nadrovinu  $a^T x + \sigma = c$ , čiže  $a^T x = c - \sigma$ ,  $\sigma \geq 0$ , jej vzdialenosť od  $N$  je  $\frac{|\sigma|}{\|a\|} = \frac{\sigma}{\|a\|}$ . Zároveň  $t = \frac{-\sigma}{\|a\|_2^2} < 0$ , čiže od  $N$  k tejto nadrovine smeruje vektor  $-a$ .

Keď máme nadrovinu  $a^T x - \sigma = c$ , čiže  $a^T x = c + \sigma$ ,  $\sigma \geq 0$ , jej vzdialenosť od  $N$  je  $\frac{|\sigma|}{\|a\|} = \frac{\sigma}{\|a\|}$ . Zároveň  $\lambda = \frac{\sigma}{\|a\|_2^2} > 0$ , čiže od  $N$  k tejto nadrovine smeruje vektor  $a$ .  $\square$

## 2 Centrálna trajektória v LP

### 2.1 Základné vlastnosti

V tejto kapitole sa venujeme vlastnému pojmu centrálnej trajektórie v lineárnom programovaní. V prvej časti približujeme ideu metód vnútorného bodu, ktorá sa odvíja od pojmu centrálnej trajektórie, zavádzame základné definície a uvádzame dôležité základné tvrdenia, z ktorých budeme neskôr vychádzať.

Počínajúc touto kapitolou predpokladáme splnenie dvoch predpokladov:

**Predpoklad 1.** Hodnosť matice  $A$  je  $m \leq n$ , čiže matica  $A$  má plnú riadkovú hodnosť a jej riadky sú lineárne nezávislé.

**Predpoklad 2.** Primárna aj duálna prípustná množina majú vnútorné body,  $\mathcal{P}^0 \neq \emptyset$ ,  $\mathcal{D}^0 \neq \emptyset$ .

Zatiaľ čo predpoklad 1 je skôr technického rázu a slúži na zjednodušenie dôkazov, predpoklad 2 je pre fungovanie metód vnútorného bodu nevyhnutný. Na základe predpokladu 2 tiež platia ekvivalentné podmienky uvádzané vo vetách 1.12, 1.13 a platí preto aj tvrdenie z poznámky za vetou 1.13. Taktiež je vzťah medzi  $y$  a  $s$  jednoznačný, ako hovorí poznámka za definíciou 1.3.

Od tejto kapitoly tiež začíname používať do istej miery neštandardné označenie. Ako je zvykom, skalárne premenné spravidla značíme malými gréckymi písmenami. Vektory, ktoré vždy považujeme za stĺpcové vektory, označujeme malými latinskými písmenami. Jednotlivé zložky vektora značíme dolným indexom, teda  $i$ -ta zložka vektora  $v$  je  $v_i$ . Vektor jednotiek vhodného rozmeru označujeme  $e$  keď je jasné aký jeho rozmer je, prípadne pre upresnenie  $e_n (\in \mathbb{R}^n)$ . Podobne vektor núl označujeme jednoducho  $0$  alebo  $0_n$ . Matice značíme veľkými latinskými písmenami. Čo je už neštandardné označenie, vektor, ktorého zložky sú prevrátenými zložkami  $v$  značíme  $v^{-1}$ , teda  $v^{-1} = (\frac{1}{v_1}, \frac{1}{v_2}, \dots, \frac{1}{v_n})^T$ . Pre vektor  $v$  značíme veľkým písmenom  $V$  diagonálnu štvorcovú maticu, ktorá má na diagonále vektor  $v$ .

Príklad. Pre zavedené značenie napríklad platí:

$$\begin{aligned} u &= Ue, \\ v^{-1} &= V^{-1}e, \\ Uv &= Vu = UVe = VUe, \\ Uv = e &\Leftrightarrow v = u^{-1} \Leftrightarrow u = v^{-1}, \\ Vv^{-1} &= VV^{-1}e = e. \end{aligned}$$

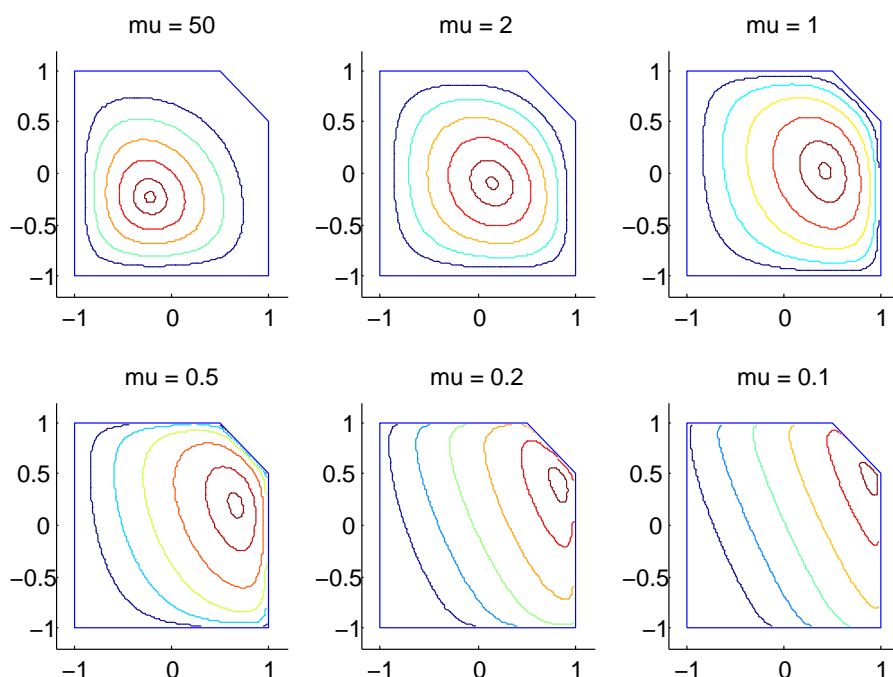
Účelom tejto časti práce je okrem iného priblížiť základné myšlienky za metódami vnútorného bodu. Začneme prvým dôležitým pojmom — pojmom bariérovej funkcie:

**Definícia 2.1.** Pre reálne číslo  $\mu > 0$  rozumieme primárnou logaritmickou bariérovou funkciou funkciu

$$f_\mu(x) = c^T x - \mu \sum_{i=1}^n \ln x_i, \quad f_\mu : \mathbb{R}_{++}^n \rightarrow \mathbb{R}.$$

Úlohou ( $P_\mu$ ) označujeme úlohu  $\min_{x \in \mathbb{R}_{++}^n} \{f_\mu(x) | Ax = b\} = \min_{x \in \mathcal{P}^0} f_\mu(x)$ .

Bariérová funkcia je teda akási funkcia, ktorou nahradíme pôvodnú účelovú funkciu. Táto funkcia už nie je lineárna, onedlho ale ľahko ukážeme, že je (v prípade primárnej funkcie) konvexná. Vďaka jej tvaru ju tiež dokážeme celkom ľahko minimalizovať technikami nelineárnej optimalizácie. Ďalej môžeme pozorovať, že  $f_\mu$  sa od pôvodnej primárnej účelovej  $c^T x$  funkcie až tak nelíši, má iba „navyš“ tzv. *bariérové členy*  $-\mu \sum_{i=1}^n \ln x_i$ . Logaritmy zariadia, že definičný obor účelovej funkcie sa zmenší na kladné vektory. To je zrejma výhoda, pretože ohraničeniami  $x \geq 0$  sa už nemusíme zaoberať — zostáva nám brať do úvahy len ohraničenia v tvare rovnosti  $Ax = b$ . Keďže  $\lim_{t \rightarrow 0^+} \ln t = -\infty$ , keď sa niektoré z  $x_k$  blíži k nule,  $-\mu \ln x_k$  zapríčiní, že hodnota  $f_\mu$  ide do  $+\infty$ . Na kladných číslach má  $f_\mu$  ale konečné hodnoty a túto (konvexnú, ako o chvíľu ukážeme) funkciu môžeme minimalizovať s ohraničením  $Ax = b$  — tvrdenie o existencii a jednoznačnosti tohoto minima je pre teóriu veľmi dôležité. Minimá nájdeme pre rôzne  $\mu > 0$  — tieto tvoria tzv. *centrálnu trajektóriu*. Teraz sa znova pozrieme na tvar bariérovej funkcie:  $c^T x - \mu \sum_{i=1}^n \ln x_i$ . Keď sa  $\mu$  blíži k nule, bariérová funkcia sa čoraz viac začína podobať na primárnu účelovú funkciu  $c^T x$ . Zabezpečili sme teda dodržanie všetkých ohraničení z  $\mathcal{P}$ , pričom na tejto množine máme nájdene minimá funkcií stále viac a viac podobných účelovej funkcie. Očakávame (a ďalšia teória to potvrdí) teda, že v limite sa s  $\mu \rightarrow 0^+$  dostaneme do optimálneho riešenia (P).



Obr. 1: Úrovňové množiny logaritmickej bariérovej funkcie pre rôzne  $\mu$

**Tvrdenie 2.2.** Funkcia  $f_\mu$  je rýdzo konvexná na svojom definičnom obore,  $\mathbb{R}_{++}^n$ .

*Dôkaz.* Počítame:

$$\begin{aligned}\frac{\partial f_\mu}{\partial x_i} &= \frac{\partial}{\partial x_i} \left( \sum_{i=1}^n c_i x_i - \mu \sum_{i=1}^n \ln x_i \right) = c_i - \mu \frac{1}{x_i}, \\ \frac{\partial^2 f_\mu}{\partial x_i^2} &= \frac{\partial}{\partial x_i} \left( c_i - \mu \frac{1}{x_i} \right) = - \left( 0 - \mu \frac{1}{x_i^2} \right) = \mu \frac{1}{x_i^2}, \\ \frac{\partial^2 f_\mu}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_j} \left( c_i - \mu \frac{1}{x_i} \right) = 0, \quad i \neq j.\end{aligned}$$

Teda Hessova matica je pre každé  $x > 0$  diagonálna s kladnými prvkami  $\mu \frac{1}{x_i^2}$ , preto je kladne definitná a  $f_\mu$  rýdzo konvexná.  $\square$

**Definícia 2.3.** Pre  $\mu > 0$  je duálna logaritmická bariérová funkcia

$$g_\mu(y, s) = b^T y + \mu \sum_{i=1}^n \ln s_i, \quad g_\mu : \mathbb{R}^m \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}.$$

Podobne ako pre primárnu bariérovú funkciu je  $(D_\mu)$  úloha

$$\max_{(y,s) \in \mathbb{R}^m \times \mathbb{R}_{++}^n} \{g_\mu(y, s) \mid A^T y + s = c\} = \max_{(y,s) \in \mathcal{D}^0} g_\mu(y, s).$$

Niekedy tiež budeme uvažovať funkciu  $g_\mu$ , do ktorej je dosadené  $s = c - A^T y$ . Vznikne

$$\tilde{g}_\mu(y) = b^T y + \mu \sum_{i=1}^n \ln(c - A^T y)_i,$$

ale vlnovku nad  $g$  vďaka ekvivalencii v zmysle poznámky za definíciou 1.3 často vynechávame.

Pre duálnu bariérovú funkciu môžeme zopakovať podobné pozorovania a úvahy ako pre primárnu. Logaritmy obmedzujú  $s$  na kladné hodnoty. Pre  $\mu \rightarrow 0+$  sa  $g_\mu$  približuje  $b^T y$ . V tejto kapitole používame často práve duálnu úlohu kvôli tomu, že jej ekvivalentný tvar (D') ( $\max \{b^T y \mid A^T y \leq c\}$ ) je veľmi dobre použiteľný na grafické znázornenie ak  $m = 2$ , teda  $y \in \mathbb{R}^2$ . Prípustná množina je teda prienikom  $n$  polrovín. Takto bola zostrojená aj ukážka s úrovňovými množinami (obr. 1). K podobným príkladom sa ešte v tejto kapitole dostaneme.

**Tvrdenie 2.4.** *Funkcia  $g_\mu$  je konkávna na svojom definičnom obore  $\mathbb{R}^m \times \mathbb{R}_{++}^n$ .*

*Za predpokladu 1 je  $\tilde{g}_\mu$  na svojom definičnom obore  $A^T y \leq c$  rýdzo konkávna.*

*Dôkaz.* Derivovanie  $g_\mu$  vykonáme zrýchlene, ľahko môžeme napísať rovno celý gradient a Hessovu maticu.

$$\nabla g_\mu(y, s) = \begin{pmatrix} b \\ \mu s^{-1} \end{pmatrix}, \quad \nabla^2 g_\mu(y, s) = \begin{pmatrix} 0 & 0 \\ 0 & -\mu s^{-2} \end{pmatrix} \preceq 0 \text{ keďže } s > 0.$$

Hessova matica funkcie  $g_S$  je záporne semidefinitná  $\forall s > 0$ , funkcia  $g_S$  je preto konkávna.

Ďalej označíme  $i$ -ty stĺpec matice  $A$  ako  $a_i$ , potom  $s_i = (c - A^T y)_i = c_i - a_i^T y$ . Riadok  $j$  matice  $A$  zase označíme  $A_{j,:}$ . Dostávame:

$$\begin{aligned} \tilde{g}_\mu(y) &= b^T y + \mu \sum_{i=1}^n \ln(c - A^T y)_i, \\ \frac{\partial \tilde{g}_\mu}{\partial y_j}(y) &= b_j + \mu \sum_{i=1}^n \left[ \frac{1}{(c - A^T y)_i} \cdot \frac{\partial}{\partial y_j} (c_i - a_i^T y) \right] = b_j + \mu \sum_{i=1}^n \frac{-(a_i)_j}{s_i} = -A_{j,:} s^{-1}, \\ \nabla \tilde{g}_\mu(y) &= b - \mu A s^{-1}. \end{aligned}$$

Ďalším derivovaním získame Hessovu maticu:

$$\begin{aligned} \nabla^2 \tilde{g}_\mu(y) &= \nabla(b - \mu A s^{-1}) = 0 - \mu A (-S^{-2}) \frac{\partial}{\partial y} s = \mu A S^{-2} \frac{\partial}{\partial y} (c - A^T y) = \\ &= \mu A S^{-2} (-A^T) = -\mu A S^{-2} A^T. \end{aligned}$$

Pre  $s > 0$  platí  $S^{-2} > 0$ . Nasledovne sa dá definovať istá norma určená kladne definitnou maticou  $S^{-2}$ :  $u^T A S^{-2} A^T u = (A^T u)^T S^{-2} (A^T u) = \|A^T u\|_{S^{-2}} > 0$ . Z toho vyplýva, že ak  $A^T$  má lineárne nezávislé stĺpce (teda ak je splnený predpoklad 2),  $u = 0 \Leftrightarrow A^T u = 0 \Leftrightarrow \|A^T u\|_{S^{-2}} = 0$  a Hessova matica  $-A S^{-2} A^T$  je teda záporne definitná. Preto je  $\tilde{g}_\mu$  rýdzo konkávna.  $\square$

**Veta 2.5.** *Pre každé  $\mu > 0$  má  $(P_\mu)$  aj  $(D_\mu)$  práve jedno optimálne riešenie.*

Dôkaz neuvádzame, nachádza sa napríklad v [8, str. 91], [3].

**Definícia 2.6** (Centrálna trajektória). *Pre  $\mu > 0$  definujeme  $x(\mu)$  ako optimálne riešenie  $(P_\mu)$ . Množinu  $\{x(\mu) \mid \mu > 0\}$  nazývame primárna centrálna trajektória.*

*Podobne nech  $(y(\mu), s(\mu))$  je optimálne riešenie  $(D_\mu)$ . Množinu  $\{(y(\mu), s(\mu)) \mid \mu > 0\}$  potom nazývame duálna centrálna trajektória.*

*Množinu  $\{(x(\mu), y(\mu), z(\mu)) \mid \mu > 0\}$ , voláme primárno-duálna centrálna trajektória.*

**Definícia 2.7** ( $\mu$ -centrujúci systém). *Nech  $\mu > 0$ . Potom sa  $\mu$ -centrujúcím systémom nazýva sústava:*

$$(C1) \quad Ax = b, \quad x > 0,$$

$$(C2) \quad A^T y + s = c, \quad s > 0,$$

$$(C3) \quad Xs = \mu e.$$

*Podmienky (C1), (C2) a (C3) nazývame tiež centrujúce podmienky.*

**Veta 2.8** (Ekvivalentná definícia centrálnej trajektórie cez  $\mu$ -centrujúci systém). *Nech  $\mu > 0$ . Potom:*

- $x \in \mathbb{R}^n$  je optimálnym riešením  $(P_\mu) \Leftrightarrow$  existujú  $y \in \mathbb{R}^m, s \in \mathbb{R}^n$  také, že  $(x, y, s)$  je riešením  $\mu$ -centrujúceho systému.
- $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$  je optimálnym riešením  $(D_\mu) \Leftrightarrow$  existuje  $x \in \mathbb{R}^n$  také, že  $(x, y, s)$  je riešením  $\mu$ -centrujúceho systému.

Dôkaz v [8, str. 91], [3]. Idea dôkazu je taká, že centrujúce podmienky sú vlastne nutné a postačujúce podmienky pre minimum  $f_\mu$  resp.  $g_\mu$ .

Na základe vety 2.5 vidíme, že pre každé  $\mu > 0$  existuje práve jeden bod centrálnej trajektórie, čiže práve jedno riešenie  $\mu$ -centrujúceho systému. Teda centrálna trajektória je dobre definovaná.

## 2.2 Dvojrozmerný príklad

**Príklad 1.** Majme nasledovné zadanie úlohy LP:

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Na príklade sa budeme snažiť vypočítať niektoré vlastnosti centrálnej trajektórie, pričom sa k nemu viac krát v priebehu kapitoly vrátíme. Prvým cieľom bude nájsť predpis pre centrálnu trajektóriu.

Nájdeme riešenie  $\mu$ -centrujúceho systému:

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad x_i > 0,$$

$$\begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad s_i > 0,$$

$$\begin{pmatrix} x_1 s_1 \\ x_2 s_2 \\ x_3 s_3 \\ x_4 s_4 \end{pmatrix} = \mu \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$



Matice a vektory roznásobíme:

$$\begin{array}{lll} -x_1 + x_2 = b_1, & -y_1 + s_1 = 1, & s_1 x_1 = \mu, \\ -x_3 + x_4 = b_2, & y_1 + s_2 = 1, & s_2 x_2 = \mu, \\ & -y_2 + s_3 = 1, & s_3 x_3 = \mu, \\ & y_2 + s_4 = 1, & s_4 x_4 = \mu. \end{array}$$

Z rovností zodpovedajúcim  $A^T y + s = c$  vyjadríme  $s_1, \dots, s_4$  a dosadíme do  $Xs = \mu e$ .

Zároveň možno z rovností pre  $Ax = b$  vyjadriť  $x_2$  pomocou  $x_1$  a  $x_4$  pomocou  $x_3$ :

$$\begin{array}{ll} x_1(1 + y_1) = \mu, & (1) \quad x_1(1 + y_1) = \mu, \\ x_2(1 - y_1) = \mu, & x_2 = b_1 + x_1, \quad (2) \quad (b_1 + x_1)(1 - y_1) = \mu, \\ x_3(1 + y_2) = \mu, & (3) \quad x_3(1 + y_2) = \mu, \\ x_4(1 - y_2) = \mu, & x_4 = b_2 + x_3, \quad (4) \quad (b_2 + x_3)(1 - y_2) = \mu. \end{array}$$

V treťom stĺpci sme dosadili  $x_2$  a  $x_4$ . Z (1) vyjadríme  $x_1$  a dosadíme do (2), čím získame jedinú rovnicu pre  $y_1$ :

$$z(1): \quad x_1 = \frac{\mu}{1+y_1}, \text{ pričom } 1 + y_1 = s_1 > 0.$$

$$(2): \quad (b_1 + \frac{\mu}{1+y_1})(1 - y_1) = \mu \quad \text{roznásobíme a } \mu \text{ presunieme na ľavú stranu:}$$

$$b_1 - b_1 y_1 + \frac{\mu}{1+y_1} - \frac{\mu}{1+y_1} y_1 - \mu = 0 \quad / \cdot (1 + y_1)$$

$$b_1 + b_1 y_1 - b_1 y_1 - b_1 y_1^2 + \mu - \mu y_1 - \mu - \mu y_1 = 0$$

$$-b_1 y_1^2 - 2\mu y_1 + b_1 = 0.$$

Dostali sme kvadratickú rovnicu, z ktorej vyjadríme  $y_1$ :

$$y_1 = \frac{2\mu \pm \sqrt{4\mu^2 + 4b_1^2}}{-2b_1}$$

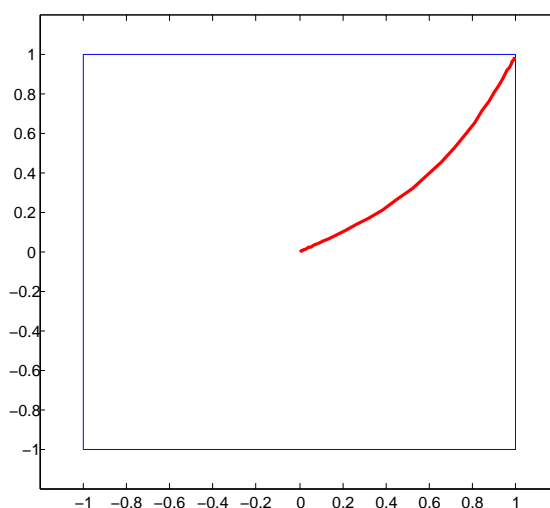
Pre každé  $\mu > 0$  musí byť splnené ohraňenie  $s_1 = 1 + y_1 > 0$ , teda  $y_1 > -1$  a  $s_2 = 1 - y_1 > 0$ , čiže  $y_1 < 1$ . Vo vetve „+“ pre  $\mu \rightarrow \infty$  ide čitateľ  $2\mu + \sqrt{4\mu^2 + 4b_1^2} \rightarrow +\infty$ . Podľa znamienka  $b_1$  preto ide celý zlomok do  $\pm\infty$  a preto nie je splnené jedno z ohraňení  $y_1 > -1, y_1 < 1$ . Vyberieme teda vetvu „-“:

$$\begin{aligned} y_1 &= \frac{2\mu - \sqrt{4\mu^2 + 4b_1^2}}{-2b_1} = \frac{\mu - \sqrt{\mu^2 + b_1^2}}{-b_1} \cdot \frac{\mu + \sqrt{\mu^2 + b_1^2}}{\mu + \sqrt{\mu^2 + b_1^2}} = \frac{\mu^2 - \mu^2 - b_1^2}{-b_1(\mu + \sqrt{\mu^2 + b_1^2})} \\ &= \frac{-b_1^2}{-b_1(\mu + \sqrt{\mu^2 + b_1^2})} = \frac{b_1}{\mu + \sqrt{\mu^2 + b_1^2}}. \end{aligned}$$

Rovnice pre  $y_2$ , teda (3), (4) sú zhodné s (1), (2), iba v nich namiesto  $b_1$  vystupuje  $b_2$  (a  $x_3$  namiesto  $x_1$ ). Preto analogicky  $y_2 = \frac{b_2}{\mu + \sqrt{\mu^2 + b_2^2}}$ .

Platí  $b_i^2 < \mu^2 + b_i^2$ , preto vidíme, že  $|b_i| < |\mu + \sqrt{\mu^2 + b_i^2}|$  pre všetky  $b_i \in \mathbb{R}, \mu > 0$ . Preto podiel týchto čísel,  $\left| \frac{b_i}{\mu + \sqrt{\mu^2 + b_i^2}} \right| = |y_i| < 1$ , z čoho  $-1 < y_i < 1$  a nakoniec skutočne  $s > 0$ , keďže po zložkách  $s_1 = 1 + y_1, s_2 = 1 - y_1, s_3 = 1 + y_2, s_4 = 1 - y_2$ .

Zvyšné prvky centrálnej trajektórie už dopočítame ľahko. Pre každé  $\mu > 0$  už máme  $y(\mu)$ , pomocou ktorého postupne vyjadríme  $s(\mu)$  a  $x(\mu)$ :  $s(\mu) = c - A^T y(\mu), x(\mu) = \mu s^{-1}(\mu)$ . Nás však bude viac zaujímať  $y(\mu) \in \mathbb{R}^2$ , čiže duálna centrálna trajektória v úlohe (D'), ktorú ako dvojrozmernú vieme zobrazit' v rovine.



**Obr. 2:** Duálna centrálna trajektória v pr. 1 pre  $b = (2, 1)^T$

Pre  $b = (2, 1)^T$  je centrálna trajektória zobrazená na obrázku 2. Pre takéto  $b$  je optimálne riešenie príslušnej úlohy (D) bod  $(1, 1)^T$ . Vidíme, že centrálna trajektória k tomuto bodu konverguje. Toto je všeobecná vlastnosť centrálnej trajektórie.  $\square$

**Veta 2.9.** *Centrálna trajektória  $(x(\mu), y(\mu), s(\mu))$  konverguje pre  $\mu \rightarrow 0+$  k optimálnemu riešeniu primárnej a duálnej úlohy.*

Dôkaz v [3]. Do istej miery bude tiež zrejmy z pozorovaní v nasledujúcej podkapitole.

### 3 Centrálna trajektória ako analytické stredy

#### 3.1 Analytický stred množiny

Na príklade 1 skúmame ďalšie vlastnosti centrálnej trajektórie. Podľa obr. 2 sa zdá, že vychádza z bodu  $(0, 0)$ , teda „stred“ prípustnej množiny. Toto pozorovanie overíme:

$$\lim_{\mu \rightarrow \infty} y_i = \lim_{\mu \rightarrow \infty} \frac{b_i}{\mu + \sqrt{\mu^2 + b_i^2}} = 0.$$

Tento výsledok sme dostali pri duálnej centrálnej trajektórii bez ohľadu na vektor  $b$  — bod  $y(\infty) := \lim_{\mu \rightarrow \infty} y(\mu)$  teda závisí iba od duálnej prípustnej množiny. Ako sme pozorovali, bod  $y(\infty)$  je tiež „stredom“ duálne prípustnej množiny a práve pojem stredy je predmetom nasledujúcej definície.

**Definícia 3.1.** *Majme množinu*

$$M = \{(u, v) \in \mathbb{R}^k \times \mathbb{R}^l \mid N_1 u + N_2 v = d, u \geq 0\},$$

kde  $N_1$  je matrica rozmerov  $r \times k$ , matrica  $N_2$  má rozmery  $r \times l$ ,  $d$  je vektor z  $\mathbb{R}^r$  a celé čísla  $k > 0$ ,  $l \geq 0$ ,  $r > 0$ . Množina vnútorných bodov  $M$  je  $M^0 = \{(u, v) \in M \mid u > 0\}$ . Ak  $M^0 \neq \emptyset$ , analytickým stredom množiny  $M$  nazývame bod

$$\operatorname{argmax} \{h(u) := \sum_{i=1}^k \ln u_i \mid (u, v) \in M, u > 0\}.$$

Ďalej označíme  $M_u = \{u \mid (u, v) \in M\}$ ,  $M_u^0 = \{u \in M_u \mid u > 0\} = \{u \mid (u, v) \in M^0\}$ . Keďže funkcia  $h(u)$  je funkciou iba premennej  $u$  a nie  $v$ ,  $M_u^0$  je definičný obor funkcie  $h(u)$ .

*Poznámka:* Pre duálnu prípustnú množinu  $\mathcal{D}$  je analytickým stredom riešenie úlohy  $(S_D) : \max \left\{ \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0 \right\}$ , označujeme ako bod  $(y_S, s_S)$ . Účelovú funkciu tejto úlohy označujeme  $\sum_{i=1}^n \ln s_i =: g_S(s)$ , resp. ak  $s$  je vyjadrené pomocou  $y$ ,  $\tilde{g}_S(y) = \sum_{i=1}^n \ln(c - a^T y)_i$ . Podobne ako pre  $\tilde{g}_\mu$  budeme vlnovku pre jednoduchosť vynechávať na základe jedno-jednoznačného vzťahu medzi  $y$  a  $s$  (poznámka za definíciou 1.3). Keď si bude potrebné rozdiel medzi  $\tilde{g}_S(y)$  a  $g_S(s)$  uvedomiť, upozorníme na to.

Analytickým stredom  $\mathcal{P}$  je riešenie  $(S_P) : \max \left\{ \sum_{i=1}^n \ln x_i \mid Ax = b, x > 0 \right\}$ , bod  $x_S$ , pričom podobne tu účelovú funkciu značíme  $\sum_{i=1}^n \ln x_i =: f_S(x)$ .

Nasleduje niekoľko tvrdení o existencii a jednoznačnosti analytických stredov najprv všeobecných množín a neskôr množín  $\mathcal{P}$  a  $\mathcal{D}$ . Tieto tvrdenia sú dôležité, keďže analytické stredy hrajú centrálnu úlohu v ďalších pozorovaniach v práci.

**Lema 3.2.** *Účelová funkcia v úlohe hľadania analytického stredy množiny  $h(u) = \sum_{i=1}^n \ln u_i$  je rýdzo konkávna na svojom definičnom obore  $M_u^0$ .*

*Dôkaz.* Nájdeme druhú deriváciu funkcie  $h(u)$ .

$$\begin{aligned}\nabla h(u) &= \nabla \sum_{i=1}^n \ln u_i = u^{-1}, \\ \nabla^2 h(u) &= -1 \cdot U^{-2} = -U^{-2}.\end{aligned}$$

Hessova matica funkcie  $h(u)$  je  $-U^{-2}$ , diagonálna matica so zápornými členmi, keďže na  $M_u^0$  platí  $u > 0$ . Je preto záporne definitná a  $h(u)$  je rýdzo konkávna.  $\square$

**Lema 3.3.** *Konvexná množina  $K$  je neohraničená  $\Leftrightarrow K$  obsahuje polpriamku.*

*Ak konvexná množina  $K$  obsahuje polpriamku, obsahuje aj polpriamku s ňou rovnobežnú vedenú z ľubovoľného bodu  $K$ .*

Ide o známe tvrdenia o konvexných množinách, dôkaz napríklad v [7].

**Veta 3.4** (Existencia a jednoznačnosť analytického stredy). *Majme množinu  $M$  a z nej vytvorené  $M^0 \neq \emptyset$ ,  $M_u$  a  $M_u^0$  podľa definície 3.1. Ak  $M$  je ohraničená, jej analytický stred existuje. Naopak, ak  $M_u$  je neohraničená, analytický stred  $M$  neexistuje.*

*Ak hodnosť matice  $N_2$  je  $l$  (vrátane prípadu keď  $l = 0$  a maticu  $N_2$  ani vektor  $v$  nemáme), analytický stred  $M$  je jednoznačný.*

*Dôkaz.* Prvá časť dôkazu podľa [4]. Predpokladajme, že  $M$  je ohraničená. Keďže  $M^0 \neq \emptyset$ , obsahuje aspoň jeden prvok a ten označíme  $(u^0, v^0) \in M^0$ . Vytvoríme množinu  $\tilde{M} = \{(u, v) \in M^0 \mid h(u) \geq h(u^0)\}$ . Je zrejmé, že  $\tilde{M} \subset M$  a maximum  $h(u)$  na  $M$  (ak existuje) leží v  $\tilde{M}$ . Keďže  $M$  je ohraničená, je ohraničená aj jej podmnožina  $\tilde{M}$ . Keď ukážeme, že  $\tilde{M}$  je uzavretá, budeme vedieť, že  $\tilde{M}$  je kompaktná, čiže funkcia  $h$  nadobúda svoje maximum na  $\tilde{M}$  a teda aj na  $M$ . Majme teda postupnosť  $\{(u^k, v^k)\} \subset \tilde{M}$  konvergujúcu k  $(\hat{u}, \hat{v})$ . Ukážeme, že  $(\hat{u}, \hat{v}) \in \tilde{M}$ , čo je definícia uzavretosti množiny.

$$N_1 u^k + N_2 v^k = d \quad \forall k \quad \Rightarrow \quad N_1 \hat{u} + N_2 \hat{v} = d,$$

$$u^k > 0 \quad \forall k \quad \Rightarrow \quad \hat{u} \geq 0,$$

$$h(u^k) \geq h(u^0) \quad \forall k \quad \Rightarrow \quad h(\hat{u}) \geq h(u^0) \quad \Rightarrow \quad \hat{u} > 0.$$

Limitným prechodom sa rovnosť aj neostrá nerovnosť zachovali, pričom keďže súčet logaritmov, funkcia  $h(\hat{u})$  má konečnú hodnotu, platí  $\hat{u} > 0$ . Teda  $(\hat{u}, \hat{v}) \in \tilde{M}$ ,  $\tilde{M}$  je uzavretá a podľa predchádzajúcich úvah maximum  $h(u)$  na  $M$  existuje.

Ďalej predpokladajme, že  $M_u$  je neohraničená. Ukážeme, že analytický stred  $M$  neexistuje. Podľa lemy 3.3 obsahuje  $M_u$  polpriamku  $p' : \{u^1 + \xi u^S \mid \xi \geq 0\}$ . Platí  $u^S \neq 0$  aby definícia polpriamky dávala zmysel a tiež  $u^S \geq 0$ . Keby totiž pre nejaký index  $i_0$  bol  $u_{i_0}^S < 0$ , existovalo by také  $\xi_0$ , že  $\xi > \xi_0 \Rightarrow u_{i_0}^1 + \xi u_{i_0}^S < 0$ , teda  $u^1 + \xi u^S \not\geq 0$ ,  $p' \notin M_u$ . Keďže  $M^0 \neq \emptyset$ , tiež platí, že  $M_u^0 \neq \emptyset$ , teda existuje  $u^0 \in M_u^0$ . Podľa lemy 3.3 leží aj polpriamka  $p : \{u^0 + \xi u^S \mid \xi \geq 0\}$  v množine  $M_u$ . Platí  $u^S \geq 0$ , teda  $u^0 + \xi u^S \geq u^0 > 0$ , teda  $p \subset M_u^0$ . Ďalej  $u^S \neq 0$ , teda pre nejaký index  $i_v$  je  $u_{i_v}^S > 0$ . Preto pre  $\xi \rightarrow +\infty$  ide  $u_{i_v}^0 + \xi u_{i_v}^S \rightarrow +\infty$ , čiže aj  $\ln(u_{i_v}^0 + \xi u_{i_v}^S) \rightarrow +\infty$ . Pre ostatné indexy platí  $u_i^S \geq 0$ , teda  $u_i^0 + \xi u_i^S \geq u_i^0$ , teda  $\ln(u_i^0 + \xi u_i^S) \geq \ln(u_i^0)$ . Preto súčet logaritmov, funkcia  $h(u^0 + \xi u^S) \rightarrow +\infty$  pre  $\xi \rightarrow +\infty$ . Účelová funkcia nadobúda ľubovoľne veľké hodnoty na  $M_u^0$ , preto jej maximum neexistuje.

Zostáva ukázať jednoznačnosť analytického stredy za predpokladu, že hodnosť matice  $N_2$  je  $l$ . Podľa lemy 3.2 je funkcia  $h(u)$  na svojom definičnom obore, konvexnej množine  $M_u^0$  rýdzo konkávna, preto ak existuje  $u_S$ , maximum  $h(u)$ , je jediné. Ak  $l = 0$ , vektor  $v$  nemáme,  $M_u = M$  a  $u_S$  je analytický stred  $M$ . Nech  $l > 0$ , pričom matica  $N_2$  má plnú stĺpcovú hodnosť  $l$ . Nulový priestor matice  $N_2$  je potom  $\{0_l\}$ . Podľa poznámky ze definíciou 1.3 k  $N_2$  existuje ľavá inverzná matica,  $N_2^*$ . Potom

$$N_1 u_S + N_2 v = d,$$

$$N_2 v = d - N_1 u_S,$$

$$v = v_S = N_2^*(d - N_1 u_S).$$

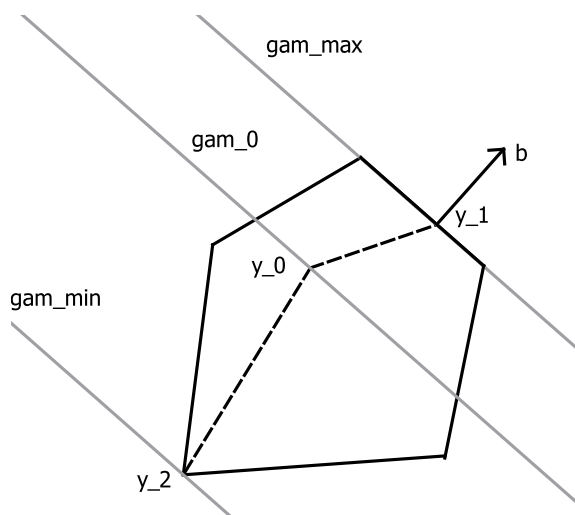
Bod  $v_S$  je určený bodom  $u_S$  jednoznačne, preto je  $(u_S, v_S)$  jednoznačný analytický stred množiny  $M$ . □

**Definícia 3.5.** *Označíme:*

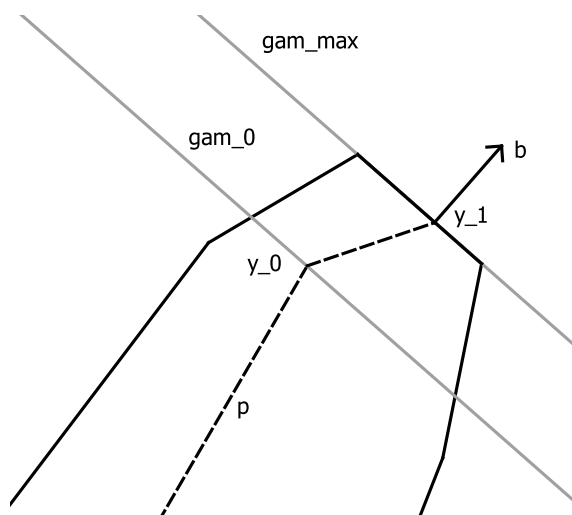
- $\delta_{max}$  hodnotu v maxime primárnej účelovej funkcie na  $\mathcal{P}$ , resp.  $\delta_{max} := +\infty$  ak maximum neexistuje (čiže  $\mathcal{P}$  je neohraničená);
- $\delta_{min}$  hodnotu v minime primárnej účelovej funkcie na  $\mathcal{P}$ ;
- $\delta_S$  hodnotu primárnej účelovej funkcie v  $x_S$ , analytickom strede  $\mathcal{P}$ , resp.  $\delta_S := +\infty$  ak analytický stred neexistuje (čiže  $\mathcal{P}$  je neohraničená);
- $\gamma_{max}$  hodnotu v maxime duálnej účelovej funkcie na  $\mathcal{D}$ ;
- $\gamma_{min}$  hodnotu v minime duálnej účelovej funkcie na  $\mathcal{D}$ , resp.  $\gamma_{min} := -\infty$  ak minimum neexistuje (čiže  $\mathcal{D}$  je neohraničená);
- $\gamma_S$  hodnotu duálnej účelovej funkcie v  $y_S$ , analytickom strede  $\mathcal{D}$ , resp.  $\gamma_S := -\infty$  ak analytický stred neexistuje (čiže  $\mathcal{D}$  je neohraničená).

*Poznámka:* Za predpokladu 2 sú primárna aj duálna úloha prípustné, preto podľa vety 1.7 existujú optimálne riešenia oboch úloh,  $\gamma_{max}$  a  $\delta_{min}$  existujú a sú konečné.

**Lema 3.6.** Nech  $c \neq 0_m$  a  $b \neq 0_m$ . Množina  $\mathcal{P}_\delta$  obsahuje vnútorné body pre  $\delta \in (\delta_{min}, \delta_{max})$ . Množina  $\mathcal{P}_{\delta-}$  obsahuje vnútorné body pre  $\delta > \delta_{min}$ . Množina  $\mathcal{D}_\gamma$  obsahuje vnútorné body pre  $\gamma \in (\gamma_{min}, \gamma_{max})$ . Množina  $\mathcal{D}_{\gamma+}$  obsahuje vnútorné body pre  $\gamma < \gamma_{max}$ .



Obr. 3: Ohraničená  $\mathcal{D}$



Obr. 4: Neohraničená  $\mathcal{D}$

*Dôkaz.* Duálny prípad. Ak  $b \neq 0_m$ , duálna účelová funkcia môže svoje minimum alebo maximum nadobúdať iba na hranici prípustnej množiny, inak sa vždy dá posunúť v

smere vektora  $b$  do vyšších hodnôt  $b^T y$ , resp, v smere  $-b$  do nižších hodnôt  $b^T y$ . Teda množiny  $\mathcal{D}_{\delta_{max}}$  a  $\mathcal{D}_{\delta_{min}}$  (ak je  $\delta_{min}$  konečné) obsahujú iba hraničné body  $\mathcal{D}$ , čiže nijaký vnútorný bod.

Podľa predpokladu 2 existuje vnútorný bod  $\mathcal{D}$ , ktorý označíme  $y^0$ . K nemu prislúchajúce  $s^0 > 0$ , a keďže na úrovni  $b^T y = \gamma_{max}$  vnútorné body neležia, platí  $\gamma_0 := b^T y^0 < \gamma_{max}$ . Vyberieme jedno optimálne riešenie (D), bod  $y^1$ . Platí preto  $b^T y^1 = \gamma_{max}$ ,  $s^1 \geq 0$ . Úsečka bez krajných bodov spájajúca  $y^0$  a  $y^1$  pretína každú úroveň množiny  $b^T y = \gamma$ ,  $\gamma \in (\gamma_0, \gamma_{max})$  práve v jednom bode. Ukážeme, že tieto body sú vnútorné body množiny  $\mathcal{D}$ . Majme teda bod  $y_\lambda = \lambda y^1 + (1 - \lambda)y^0$ ,  $\lambda \in (0, 1)$ . Potom

$$\begin{aligned} s_\lambda &= c - A^T y_\lambda = (\lambda + (1 - \lambda))c - A^T(\lambda y^1 + (1 - \lambda)y^0) = \\ &= \lambda(c - A^T y^1) + (1 - \lambda)(c - A^T y^0) = \lambda s^1 + (1 - \lambda)s^0 \end{aligned}$$

Platí  $s^1 \geq 0$ , preto  $\lambda s^1 \geq 0$ . Tiež vieme, že  $s^0 > 0$ , preto aj  $(1 - \lambda)s^0 > 0$ , čiže  $s_\lambda > 0$ ,  $y_\lambda$  je vnútorný bod  $\mathcal{D}$ , čiže vnútorný bod úrovňovej množiny  $\mathcal{D}_\gamma$  pre  $\gamma = b^T y_\lambda$ .

Ak je  $\mathcal{D}$  ohraničená, analogicky by sa ukázalo, že vnútro úsečky spájajúcej  $y^0$  a  $y^2$ , bod minima  $b^T y$  na  $\mathcal{D}$ , má jeden spoločný bod s každou z množín  $\mathcal{D}_\gamma$ ,  $\gamma \in (\gamma_{min}, \gamma_0)$ . Tieto body sú zároveň vnútorné body  $\mathcal{D}_\gamma$ .

Teda ak je  $\mathcal{D}$  ohraničená,  $\mathcal{D}_\gamma$  obsahuje vnútorné body pre  $\gamma \in (\gamma_{min}, \gamma_{max})$ .

Ak je  $\mathcal{D}$  neohraničená, podľa lemy 3.3 existuje polpriamka  $p$  idúca z bodu  $x^0$ . Analogicky s druhou časťou dôkazu vety 3.4 sú všetky body polpriamky  $p$  vnútorné body  $\mathcal{D}$ . Zároveň  $p$  smeruje ku menším hodnotám  $b^T y$  ako  $\gamma_0$ . Keby smerovala k väčším hodnotám, potom by  $b^T y$  nemala na  $\mathcal{D}$  maximum a  $p$  tiež nemôže ležať v  $\mathcal{D}_{\gamma_0}$ , keďže táto množina je podľa vety 1.12 ohraničená. Teda každá množina  $\mathcal{D}_\gamma$  pre  $\gamma < \gamma_0$  má jeden spoločný bod s  $p$ , čo je vnútorný bod  $\mathcal{D}_\gamma$ . Spojením s prvou časťou dôkazu dostávame, že  $\mathcal{D}_\gamma$  obsahuje vnútorné body práve pre  $\gamma \in (-\infty, \gamma_{max}) = (\gamma_{min}, \gamma_{max})$  keďže pre neohraničenú množinu sme definovali  $\gamma_{min} = -\infty$ .

Množina  $\mathcal{D}_{\gamma_{max}+} = \mathcal{D}_{\gamma_{max}}$ , keďže väčšie hodnoty ako  $\gamma_{max}$  účelová funkcia nenadobúda, táto množina teda vnútorné body  $\mathcal{D}$  neobsahuje. Pre  $\gamma \in (\gamma_{min}, \gamma_{max})$  množina  $\mathcal{D}_{\gamma+}$  obsahuje množinu  $\mathcal{D}_{\gamma_1}$ ,  $\gamma_1 = \frac{\gamma_{max} + \gamma}{2} \in (\gamma_{min}, \gamma_{max})$ , ktorá obsahuje vnútorný bod, teda aj  $\mathcal{D}_{\gamma+}$  obsahuje vnútorný bod. Pre  $\gamma \leq \gamma_{min}$  je už  $\mathcal{D}_{\gamma+}$  celá  $\mathcal{D}$ , ktorá obsahuje vnútorný bod.

Dôkaz pre primárny prípad je analogický. □

**Dôsledok 3.7.** *Analytický stred množín  $\mathcal{P}$  a  $\mathcal{D}$  existuje práve vtedy, keď je príslušná množina ohraničená a ak existuje, je jednoznačný. Pre hodnoty  $\gamma$  a  $\delta$  v rozsahu podľa lemy 3.6 analytické stredy množín  $\mathcal{P}_\delta$ ,  $\mathcal{P}_{\delta-}$ ,  $\mathcal{D}_\gamma$ ,  $\mathcal{D}_{\gamma+}$  existujú a sú jednoznačné.*

*Dôkaz.* Dôkaz využíva značenie zavedené v definícii 3.1.

Nech je splnený predpoklad 2, preto majú  $\mathcal{P}$  aj  $\mathcal{D}$  vnútorné body. Ak je potom množina  $\mathcal{P}$  ( $\mathcal{D}$ ) ohraničená, existuje analytický stred podľa vety 3.4. V množine  $\mathcal{P} = \{x \mid Ax = b, x \geq 0\}$  sa vyskytujú iba nezáporné premenné, preto  $l = 0$  a podľa vety 3.4 je analytický stred jednoznačný. Pre množinu  $\mathcal{D} = \{(y, s) \mid A^T y + s = c, s \geq 0\}$  zas platí  $N_2 = A^T$ , podľa predpokladu 1 má matica  $A^T$  plnú stĺpcovú hodnotu  $m = l$ , preto znovu podľa 3.4 je analytický stred jednoznačný.

Ak je  $M = \mathcal{P}$  neohraničená,  $M_u$  k množine  $\mathcal{P}$  je vlastná  $\mathcal{P}$ , pretože  $l = 0$ .  $M_u$  je teda neohraničená, podľa vety 3.4 analytický stred neexistuje. Nech je množina  $\mathcal{D}$  neohraničená. Ak je neohraničená v premenných  $y$ , existujú body ľubovoľne ďaleko od niektorých hraničných nadrovín  $A^T y = c$ . Podľa vety 1.14 je práve vektor  $s$  priamo úmerný týmto vzdialenostiam, teda ak sú vzdialenosti neohraničené,  $\mathcal{D}$  je neohraničená aj v premennej  $s$ . Práve množina  $\{s \mid (y, s) \in \mathcal{D}\}$  je pre  $\mathcal{D}$  množina  $M_u$ , preto podľa vety 3.4 analytický stred  $\mathcal{D}$  neexistuje.

Za splnenia predpokladu 2 sú podľa viet 1.12 a 1.13 a poznámky za nimi množiny  $\mathcal{P}_\delta$ ,  $\mathcal{P}_{\delta-}$ ,  $\mathcal{D}_\gamma$ ,  $\mathcal{D}_{\gamma+}$  ohraničené. Ich analytické stredy teda existujú ak majú množiny vnútorné body, čiže ak  $\gamma$  a  $\delta$  sú v rozsahu podľa lemy 3.6. Zostáva ukázať jednoznačnosť. Množiny  $\mathcal{P}_\delta$ ,  $\mathcal{P}_{\delta-}$  obsahujú iba nezáporné premenné, v nich je preto  $l = 0$ , analytický stred je jednoznačný. Množina  $\mathcal{D}_\gamma$  má ohraničenia:

$$\begin{pmatrix} I_n \\ 0_n^T \end{pmatrix} s + \begin{pmatrix} A^T \\ b^T \end{pmatrix} y = \begin{pmatrix} c \\ \gamma \end{pmatrix}.$$

Ak sú stĺpce  $A^T$  lineárne nezávislé, lineárne nezávislé sú aj stĺpce  $\begin{pmatrix} A^T \\ b^T \end{pmatrix}$ , teda táto matica má plnú stĺpcovú hodnotu, preto je analytický stred jednoznačný. Podobne množina  $\mathcal{D}_{\gamma+}$  má ohraničenia:

$$\begin{pmatrix} I_n & 0_n \\ 0_n^T & -1 \end{pmatrix} \begin{pmatrix} s \\ \sigma \end{pmatrix} + \begin{pmatrix} A^T \\ b^T \end{pmatrix} y = \begin{pmatrix} c \\ \gamma \end{pmatrix}.$$

Maticou  $N_2$  je tu rovnaká matica  $\begin{pmatrix} A^T \\ b^T \end{pmatrix}$  s lineárne nezávislými stĺpcami, preto je analytický stred jednoznačný. □



Neskôr využijeme gradienty funkcií  $g_S$  a  $f_S$ , preto ich vypočítame v leme:

**Lema 3.8.** Gradient funkcie  $\tilde{g}_S(y) = \sum_{i=1}^n \ln(c - A^T y)_i$ , teda účelovej funkcie pri hľadaní analytických stredov  $\mathcal{D}$  je  $\nabla \tilde{g}_S(y) = -As^{-1}$

Gradient  $f_S(x) = \sum_{i=1}^n \ln x_i$  je  $\nabla f_S(x) = x^{-1}$ .

*Dôkaz.* Pre  $\tilde{g}_S$  je výpočet zhodný s dôkazom tvrdenia 2.4, ibaže bez člena  $b^T y$ . Platí teda

$$\nabla \tilde{g}_S(y) = -As^{-1}.$$

Pre  $f_S(x)$ :

$$\nabla f_S(x) = \nabla \sum_{i=1}^n \ln x_i = x^{-1}.$$

□

Vráťme sa k bodu  $y(\infty) = \lim_{\mu \rightarrow \infty} y(\mu)$ . Pozrieme sa na takýto bod vo všeobecnom prípade. Z definície centrálnej trajektórie je takýto bod optimálnym riešením  $(D_\mu)$  pre  $\mu \rightarrow \infty$  (ak také maximum existuje). Pripomenieme úlohu  $(D_\mu) : \max_{(y,s) \in \mathcal{D}^0} g_\mu(y, s)$ , čiže maximalizujeme funkciu  $g_\mu(y, s)$  na akejsi množine (nezávislej na  $\mu$ ). Keď účelovú funkciu násobíme kladnou konštantou, maximum sa nezmení: ak  $g_\mu(\hat{y}, \hat{s}) > g_\mu(y, s) \Rightarrow \alpha g_\mu(\hat{y}, \hat{s}) > \alpha g_\mu(y, s)$  pre  $\alpha > 0$ . Preto môžeme namiesto  $g_\mu(y, s)$  maximalizovať funkciu  $\frac{1}{\mu} g_\mu(y, s) = \frac{b^T y}{\mu} + \sum_{i=1}^n \ln s_i$ . Pre  $\mu \rightarrow \infty$  sa hodnota  $\frac{b^T y}{\mu}$  blíži k nulovej a zostane  $(D_\infty) : \max \left\{ \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0 \right\}$ , čo je presne definícia analytického stredy  $\mathcal{D}$ . Z predchádzajúcich viet tiež vieme, kedy limita existuje — vtedy, keď existuje analytický stred  $\mathcal{D}$ , čiže práve vtedy, keď je množina  $\mathcal{D}$  ohraničená.

Pojem analytického stred vystupuje aj v ďalšom pozorovaní o centrálnej trajektórii:

**Veta 3.9.** Ak  $b = 0_m$ , duálna centrálna trajektória je jediný bod nezávislý na  $\mu$ , a to  $y_S$ , analytický stred  $\mathcal{D}$ . Ak  $c = 0_n$ , primárna centrálna trajektória je jediný bod nezávislý na  $\mu$ ,  $x_S$ , analytický stred  $\mathcal{P}$ .

*Dôkaz.* Nech  $b = 0_m$ . Zisťujeme, ako vyzerá centrálna trajektória.  $\mu$ -centrujúci systém

má tvar:

$$\begin{aligned} Ax &= 0, & x &> 0, \\ A^T y + s &= c, & s &> 0, \\ Xs &= \mu e. \end{aligned}$$

Z tretej rovnosti vyjadríme  $x = \mu s^{-1}$  a dosadíme do prvej:

$$\mu A s^{-1} = 0.$$

Keďže  $\mu > 0$ , môžeme vydeliť a dostávame systém:

$$\begin{aligned} A s^{-1} &= 0, \\ A^T y + s &= c, & s &> 0. \end{aligned}$$

Okamžite je očividné, že systém od  $\mu$  nezávisí. Navyše, podľa lemy 3.8 je gradient funkcie  $g_S$  rovný  $-A s^{-1}$ . Náš systém je teda systémom

$$\nabla g_S = 0, \quad A^T y + s = c, \quad s > 0,$$

čiže je to úloha hľadania voľného maxima funkcie  $g_S(y)$  (presnejšie funkcie  $\tilde{g}_S(y) = g_S(c - A^T y)$ ) na  $\mathcal{D}$ , čiže úloha hľadania analytického stredu  $\mathcal{D}$ . Zistili sme teda, že ak  $b = 0$ , centrálna trajektória je jeden bod, a to analytický stred duálne prípustnej množiny (ak ten existuje). To nie je veľmi prekvapujúce — duálna účelová funkcia je konštantná, nulová, preto všetky prípustné riešenia sú aj optimálne. Jednobodová centrálna trajektória v tomto prípade zaujímavá nie je, preto budeme v budúcich tvrdeniach väčšinou predpokladať, že  $b \neq 0$ .

Podobný výsledok očakávame pre primárnu úlohu, v prípade keď zas  $c = 0_n$ . V  $\mu$ -centrujúcom systéme vyjadríme  $s = \mu x^{-1}$  a dostaneme systém:

$$\begin{aligned} A^T y + \mu x^{-1} &= 0, & Ax &= b, & x &> 0, \\ A^T \frac{y}{\mu} + x^{-1} &= 0, & Ax &= b, & x &> 0. \end{aligned} \tag{1}$$

Hľadanie analytického stredu  $\mathcal{P}$  ale nie je hľadanie voľného extrémumu. V primárnom prípade sa rieši úloha:  $\max \left\{ \sum_{i=1}^n \ln x_i \mid Ax = b, x > 0 \right\}$ . Lagrangeova funkcia je:

$$L = \sum_{i=1}^n \ln x_i + u(Ax - b), \quad u \in \mathbb{R}^m, x > 0.$$

Podmienky optimality sú:

$$\nabla_x L = x^{-1} + A^T u = 0, \quad Ax = b, \quad x > 0. \quad (2)$$

Dostávame rovnaký systém ako (1), pričom vektor Lagrangeových multiplikátorov je  $u = \frac{y}{\mu}$ . Systém (2) zostáva rovnaký, preto sa pre rôzne  $\mu$  síce môže meniť  $y$  v (1),  $u$  aj  $x$  zostávajú rovnaké. Ak  $c = 0$ , primárna centrálna trajektória je preto jediný bod, analytický stred  $\mathcal{P}$ .  $\square$

### 3.2 Centrálna trajektória a úrovnňové množiny

Vráťme sa späť k príkladu 1. Pri hľadaní analytického streda duálne prípustnej množiny  $\mathcal{D}$  riešime úlohu ( $S_D$ ):  $\max \left\{ \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0 \right\}$ , teda konkrétne pre náš príklad:

$$\max \left\{ \sum_{i=1}^4 \ln s_i \mid -y_1 + s_1 = 1, y_1 + s_2 = 1, -y_2 + s_3 = 1, y_2 + s_4 = 1, s > 0 \right\}.$$

Vektor  $s$  vyjadríme pomocou  $y$ , pričom pamätáme na to, že  $s > 0$  a preto musíme nájsť vhodné hranice pre  $y$ :

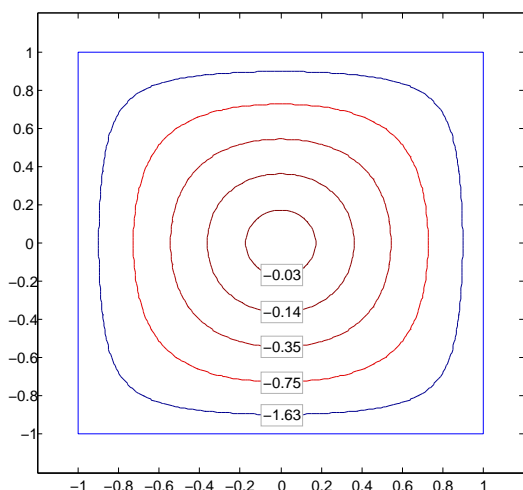
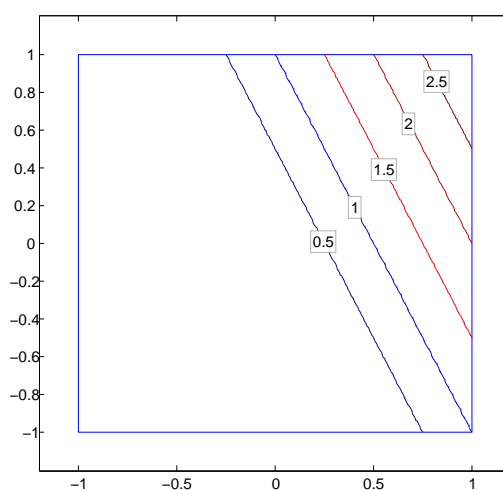
$$\begin{aligned} s_1 &= 1 + y_1, & y_1 &= s_1 - 1 > 0 - 1 = -1, \\ s_2 &= 1 - y_1, & y_1 &= 1 - s_2 < 1 - 0 = 1, \\ s_3 &= 1 + y_2, & y_2 &= s_3 - 1 > 0 - 1 = -1, \\ s_4 &= 1 - y_2, & y_2 &= 1 - s_4 < 1 - 0 = 1. \end{aligned}$$

Po dosadení vznikne:

$$\max \{ g_S(y) = \ln(1 + y_1) + \ln(1 - y_1) + \ln(1 + y_2) + \ln(1 - y_2) \mid -1 < y_1 < 1, -1 < y_2 < 1 \}.$$

V rovine môžeme zakresliť vrstevnice účelovej funkcie  $\tilde{g}_S$  (obr. 5).

Limitný bod centrálnej trajektórie,  $y(\infty)$ , sme už vedeli vyjadriť ako analytický stred nejakej množiny — konkrétne  $\mathcal{D}$ . Naším cieľom je teraz vyjadriť vlastné body centrálnej trajektórie ako analytické stredy akýchsi množín. Dáva zmysel pozrieť sa na

Obr. 5: Vrstevnice  $g_S(y_1, y_2)$ Obr. 6: Úrovnňové množiny  $b^T y = \gamma, y \in \mathcal{D}$ 

úrovnňové množiny duálnej účelovej funkcie  $b^T y$ . Zatiaľ čo pri výpočte  $y(\infty)$  sa nepoužila informácia o  $b$ , samotná centrálna trajektória už od  $b$  závisí a práve túto informáciu dodáme v duálnej účelovej funkcii  $b^T y$ . Spomínané úrovnňové množiny sú už v prvej kapitole označené množiny:

$$\mathcal{D}_\gamma : \{(y, s) \mid A^T y + s = c, s \geq 0, b^T y = \gamma\}$$

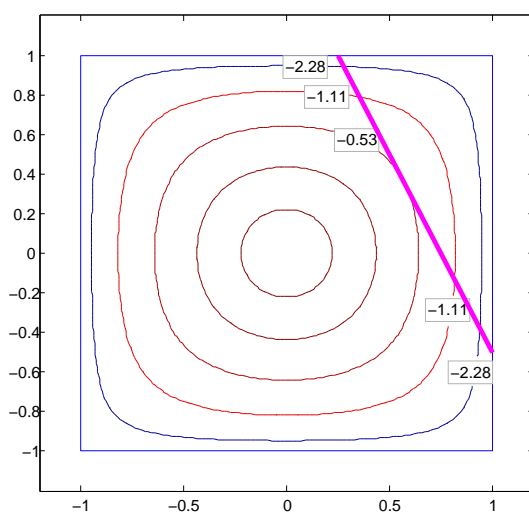
(obr. 6 pre  $b = (2, 1)^T$ ; vidíme, že v tomto prípade ide o úsečky). Podľa dôsledku 3.7 má zmysel hľadať analytické stredy pre vhodné hodnoty  $\gamma$ . Analytickým stredom takejto množiny pre nejaké  $\gamma$  je podľa definície riešenie úlohy

$$\max \left\{ \sum_{i=1}^4 \ln s_i \mid A^T y + s = c, s > 0, b^T y = \gamma \right\}.$$

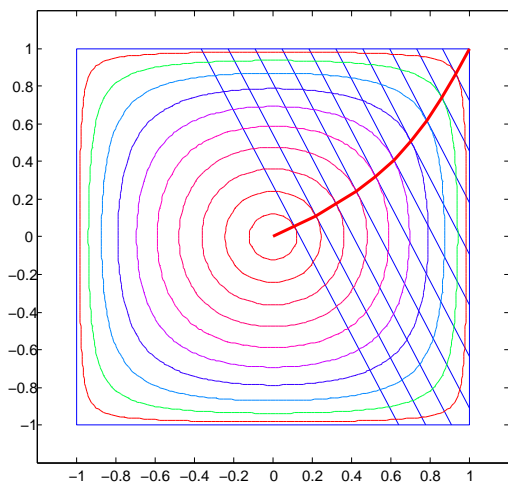
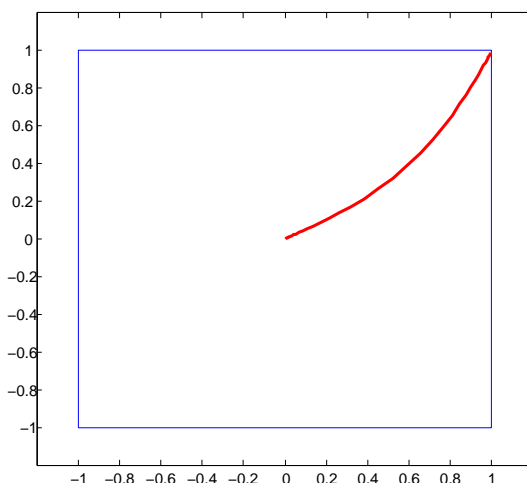
Teda k úlohe hľadania analytického stredy  $\mathcal{D}$  sme pridali dodatočné ohraničenie  $b^T y = \gamma$ . Vyjadríme  $s$  pomocou  $y$  a dostávame:

$$\max \{ \ln(1+y_1) + \ln(1-y_1) + \ln(1+y_2) + \ln(1-y_2) \mid -1 < y_1 < 1, -1 < y_2 < 1, b^T y = \gamma \}.$$

Majme teraz  $\gamma$  pevné. Maximalizujme funkciu  $g_S(y)$  na množine  $\mathcal{D}_\gamma = \{y \in \mathcal{D} \mid b^T y = \gamma\}$ . Z Obr. 5 máme predstavu o tom, ako  $g_S(y)$  vyzerá. Na „kružniciach“ s menším polomerom má  $g_S(y)$  postupne väčšie hodnoty. Najväčšiu hodnotu na množine  $\mathcal{D}_\gamma$  dosiahne  $g_S(y)$  v takom bode z  $\mathcal{D}_\gamma$ , ktorý leží na vrstevnici  $g_S(y)$  s čo najmenším polomerom — takej, ktorá sa úsečky  $\mathcal{D}_\gamma$  dotýka (obr. 7).

Obr. 7: Hľadanie maxima  $g_S$  na  $\mathcal{D}_\gamma$ 

Takéto maximá nájdeme pre rôzne  $\gamma$ , čím získame akúsi krivku (obr. 8 vs. obr. 1). V skutočnosti ide o duálnu centrálnu trajektóriu, čo dokážeme vo všeobecnosti v nasledovnej vete. Najprv však sformulujeme a dokážeme dve pomocné tvrdenia.

Obr. 8: Maximá  $g_S$  na rôznych  $\mathcal{D}_\gamma$ 

Obr. 1: Duálna centrálna trajektória

**Lema 3.10** (Podmienky na analytický stred  $\mathcal{D}_\gamma$  a  $\mathcal{P}_\delta$ ). a) Nech  $\gamma \in (\gamma_{min}, \gamma_{max})$ . Bod  $y \in \mathcal{D}_\gamma$  je analytickým stredom množiny  $\mathcal{D}_\gamma$  práve vtedy, keď existuje také  $\tau \in \mathbb{R}$ , že platí

$$\nabla \tilde{g}_S(y) = \tau b.$$

b) Nech  $\delta \in (\delta_{min}, \delta_{max})$ . Bod  $x \in \mathcal{P}_\delta$  je analytickým stredom množiny  $\mathcal{P}_\delta$  práve vtedy,

keď existujú také  $\tau \in \mathbb{R}$  a  $u \in \mathbb{R}^m$ , že platí

$$\nabla f_S(x) - A^T u = \tau c.$$

*Dôkaz.* a) Nech  $\gamma \in (\gamma_{min}, \gamma_{max})$ . Z dôsledku 3.7 vieme, že množina  $\mathcal{D}_\gamma$  má analytický stred práve pre  $\gamma$  v tomto intervale. Úloha hľadania analytického stredu je úlohou konvexného programovania, preto vieme napísať nutné a postačujúce podmienky na existenciu jej riešenia. Riešime teda úlohu:

$$\max \left\{ \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0, b^T y = \gamma \right\}.$$

Napíšeme Lagrangeovu funkciu (s mínusovou účelovou funkciou, keďže ide o maximalizačnú úlohu):

$$L(y, s, u, \tau) = - \sum_{i=1}^n \ln s_i + u^T (A^T y + s - c) + \tau (b^T y - \gamma), \quad L : \mathbb{R}^m \times \mathbb{R}_{++}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}.$$

Nutné a postačujúce podmienky sú:

1.  $y$  je voľná premenná, preto  $\nabla_y L = 0$ ;
2.  $s > 0$  je otvorená množina, preto hľadáme vzhľadom na  $s$  voľný extrém,  $\nabla_s L = 0$ ;
3. Pri  $u$  sú ohraničenia v tvare rovnosti, preto  $\nabla_u L = 0$ ;
4. Pri  $\tau$  sú ohraničenia v tvare rovnosti, preto  $\nabla_\tau L = 0$ .

Dostávame teda rovnice:

$$\nabla_y L = Au + \tau b = 0 \quad \Rightarrow \quad -Au = \tau b, \quad (3)$$

$$\nabla_s L = -s^{-1} + u = 0 \quad \Rightarrow \quad u = s^{-1}, \quad (4)$$

$$\nabla_u L = A^T y + s - c = 0 \quad \Rightarrow \quad A^T y + s = c, \quad (5)$$

$$\nabla_\tau L = b^T y - \gamma = 0 \quad \Rightarrow \quad b^T y = \gamma. \quad (6)$$

Dosadením  $u$  z (4) do (3) dostávame:

$$-As^{-1} = \tau b. \quad (7)$$

Lema 3.8 hovorí, že  $\nabla \tilde{g}_S(y) = -As^{-1}$ . Tiež  $b = \nabla b^T y$ . Preto dostávame nasledovnú podmienku:

$$\nabla \tilde{g}_S(y) = \tau \nabla b^T y = \tau b. \quad (8)$$

Zostáva ešte zahrnúť nerovnosť  $s > 0$ , ktorá vystupuje už v definícii úlohy. Podmienky (5) a (6) spolu s  $s > 0$  dávajú  $y \in \mathcal{D}_\gamma$ . Dodatočná podmienka je podmienka (7), resp. (8).

b) Primárny prípad. Nech  $\delta \in (\delta_{min}, \delta_{max})$ . Podľa dôsledku 3.7 má  $\mathcal{P}_\delta$  jednoznačný analytický stred. Vieme vyjadriť podmienky na riešenie úlohy:

$$\max \{f_s(x) = \sum_{i=1}^n \ln x_i \mid Ax = b, x > 0, c^T x = \delta\}.$$

Lagrangeova funkcia:

$$L(x, u, \tau) = - \sum_{i=1}^n \ln x_i + u^T (Ax - b) + \tau (c^T x - \delta) \quad L : \mathbb{R}_{++}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}.$$

Podmienky optimality:

- $x$  je definované na otvorenej množine  $x > 0$ , preto  $\nabla_x L = 0$ ;
- ohraničenia pri  $u$  sú v tvare rovností, preto  $\nabla_u L = 0$ ;
- ohraničenie pri  $\tau$  je v tvare rovnosti, preto  $\nabla_\tau L = 0$ .

To sú teda rovnice:

$$\nabla_x L = -x^{-1} + A^T u + \tau c = 0 \quad \Rightarrow \quad \tau c = x^{-1} - A^T u, \quad (9)$$

$$\nabla_u L = Ax - b = 0 \quad \Rightarrow \quad Ax = b, \quad (10)$$

$$\nabla_\tau L = c^T x - \delta = 0 \quad \Rightarrow \quad c^T x = \delta. \quad (11)$$

Pozrieme sa na (9). Podľa lemy 3.8 platí  $\nabla f_S(x) = x^{-1}$ , teda z (9) dostávame:

$$\tau c = \nabla f_S(x) - A^T u. \quad (12)$$

Dostali sme teda podmienky  $x \in \mathcal{P}_\delta$  s dodatočnou podmienkou (12).  $\square$

Na základe vlastnosti (8) vieme ukázať vzťah medzi  $\gamma$  a  $\tau$ . Podobný vzťah existuje aj v primárnom prípade.

**Lema 3.11.** a) *Majme bod  $y_\gamma$ ,  $\gamma \in (\gamma_{min}, \gamma_{max})$  v ktorom platí podmienka z lemy 3.10,  $\nabla \tilde{g}_S(y_\gamma) = \tau b$  a  $b^T y_\gamma = \gamma$ ,  $b \neq 0$ . Potom platí:*

$$\tau > 0 \quad \Leftrightarrow \quad \gamma \in (\gamma_{min}, \gamma_S),$$

$$\tau = 0 \quad \Leftrightarrow \quad \gamma = \gamma_S,$$

$$\tau < 0 \quad \Leftrightarrow \quad \gamma \in (\gamma_S, \gamma_{max}).$$

(Ak je  $\mathcal{D}$  neohraničená, označili sme  $\gamma_{min} = -\infty$ ,  $\gamma_S = -\infty$ , zostáva preto iba tretia podmienka,  $\tau < 0 \Leftrightarrow \gamma < \gamma_{max}$ )

b) Nech  $\delta \in (\delta_{min}, \delta_{max})$  a pre bod  $x_\delta$  platí  $c^T x_\delta = \delta$  ( $c \neq 0$ ) a existuje  $u$ , že  $\nabla f_S(x) - A^T u = \tau c$ . Potom:

$$\tau > 0 \Leftrightarrow \delta \in (\delta_{min}, \delta_S),$$

$$\tau = 0 \Leftrightarrow \delta = \delta_S,$$

$$\tau < 0 \Leftrightarrow \delta \in (\delta_S, \delta_{max}).$$

*Dôkaz.* Úplný dôkaz uvádzame pre duálny prípad.

Máme konkávnou funkciu  $g_S$  definovanú na otvorenej množine  $\mathcal{D}^0$ . Uvažujeme body  $y_\gamma$ , pre ktoré platí  $\nabla g_S(y_\gamma) = \tau b$ , pričom predpokladáme, že  $b \neq 0$ . Zisťujeme, aké znamienko má  $\tau$  v závislosti od  $\gamma$ .

Nech je množina  $\mathcal{D}$  neohraničená. Podľa dôsledku 3.7 analytický stred  $\mathcal{D}$  neexistuje. Vektor  $b$  smeruje k existujúcemu optimálnemu riešeniu, teda smerom k hranici  $\mathcal{D}$ . Funkcia  $g_S$  rastie smerom dovnútra množiny — je totiž rastúcou funkciou zložiek vektora  $s$  ktoré je podľa vety 1.14 vzdialenosťami od jednotlivých hrán  $\mathcal{D}$  a tieto vzdialenosti sa smerom dovnútra množiny zväčšujú. Gradient  $g_S$  je smer najväčšieho rastu funkcie, teda smeruje dovnútra  $\mathcal{D}$ , opačným smerom ako  $b$ . Preto  $\tau < 0$  pre každé  $\gamma < \gamma_{max}$ .

Nech  $\mathcal{D}$  je ohraničená množina, jej analytický stred preto existuje a označíme ho  $y_S$ , pričom hodnotu duálnej účelovej funkcie v ňom označíme  $\gamma_S = b^T y_S$ .

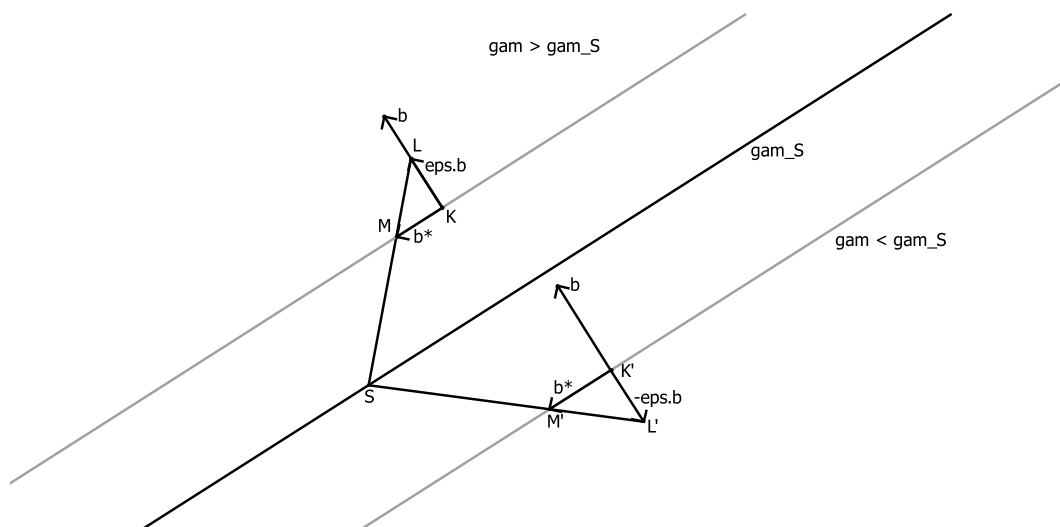
Nech  $\gamma = \gamma_S$ . Bod  $y_\gamma$  je analytickým stredom  $\mathcal{D}_{\gamma_S}$ , ktorý vieme, že je jediný. Na úrovni  $b^T y = \gamma_S$  leží voľný extrém funkcie  $g_S(y)$ , bod  $y_S$ . Bod  $y_S$  je teda aj viazaným extrémom funkcie  $g_S(y)$  na úrovni  $b^T y = \gamma_S$ , teda je analytickým stredom množiny  $\mathcal{D}_{\gamma_S}$ . Platí  $\nabla g_S(y_\gamma) = 0$  a z  $g_S(y_\gamma) = 0_m = \tau b$  vyplýva  $\tau = 0$ .

Naopak, nech  $\tau = 0$ . Potom  $g_S(y_\gamma) = \tau b = 0_m$ ,  $y_\gamma$  je voľným extrémom  $g_S$ , čiže analytickým stredom  $\mathcal{D}$ . Vieme, že analytický stred je jediný, platí preto  $y_\gamma = y_S$  a teda  $\gamma = \gamma_S$ .

Týmto ostávame ekvivalenciu  $\tau = 0 \Leftrightarrow \gamma = \gamma_S$ .

Nech  $\gamma > \gamma_S$ . Bod  $y_S$  budeme skrátene označovať  $S$  a bod  $y_\gamma$  označíme  $K$ . Dokážeme, že  $\tau < 0$ . Už sme vylúčili možnosť, že  $\tau = 0$ . Sporom predpokladajme, že  $\tau > 0$ , čiže  $\nabla g_S(K)$  má rovnaký smer ako  $b$ . Ako bod  $L_\varepsilon$  označíme bod  $K + \varepsilon b$ ,  $\varepsilon > 0$ .  $\varepsilon$  môžeme voliť dostatočne malé, tak, aby bod  $L_\varepsilon = K + \varepsilon b$  stále ležal v otvorenej





Obr. 9: Dôkaz lemy 3.11

množine  $\mathcal{D}^0$ . Keďže  $\gamma > \gamma_S$  a vektor  $b$  smeruje k ešte väčším hodnotám funkcie  $b^T y$  ako  $\gamma$ , úsečka spájajúca body  $S$  a  $L_\varepsilon$  má priesečník s nadrovinou  $b^T y = \gamma$  — tento priesečník označíme  $M_\varepsilon$ . Vektor idúci z bodu  $K$  do bodu  $M_\varepsilon$  označíme  $b_\varepsilon^*$ . Keďže body  $K$  a  $M_\varepsilon$  ležia na nadrovine  $b^T y = \gamma$ , vektor ich spájajúci,  $b_\varepsilon^*$ , je kolmý na normálový vektor nadroviny,  $b$ . Symbolicky zapísané:  $b^T b_\varepsilon^* = 0$ . Derivácia funkcie  $g_S$  v bode  $K$  v smere  $b$  je:  $(\nabla g_S(K))^T b = \tau b^T b = \tau \|b\|^2$  — smer gradientu poznáme z matematickej analýzy ako smer najväčšieho rastu funkcie. Derivácia funkcie  $g_S$  v bode  $K$  v smere  $b_\varepsilon^*$  je:  $(\nabla g_S(K))^T b_\varepsilon^* = \tau b^T b_\varepsilon^* = \tau 0 = 0$ . Existuje dostatočne malé  $\varphi$ -okolie bodu  $K$ , na ktorom je funkcia  $g_S$  dostatočne dobre aproximovaná lineárne pomocou svojej prvej derivácie, teda na tomto okolí v smere  $b$  rastie približne lineárne a v smere  $b^*$  je na  $\varphi$ -okolí približne konštantná. Zvolíme  $\varepsilon$  tak, aby aj  $L_\varepsilon$  aj  $M_\varepsilon$  ležali vnútri spomínaného  $\varphi$ -okolia. Platí preto  $g_S(M_\varepsilon) \approx \gamma < g_S(L_\varepsilon)$ . Zároveň je bod  $S$  maximom funkcie  $g_S$ , preto  $g_S(M_\varepsilon) < g_S(L_\varepsilon) \leq g_S(S)$ . Označíme  $\alpha_L = g_S(L_\varepsilon)$ . Funkcia  $g_S$  je konkávna a pre každú konkávnu funkciu platí, že jej nadúrovnňové množiny  $\{y \mid g_S(y) \geq \alpha\}$  sú konvexné. Množina  $\{y \in \mathcal{D}^0 \mid g_S(y) \geq \alpha_L\}$  ale konvexná nie je — obsahuje body  $S$  a  $L_\varepsilon$  ale neobsahuje bod na úsečke medzi nimi,  $M_\varepsilon$ , lebo  $g_S(M_\varepsilon) < \alpha_L = g_S(L_\varepsilon)$ . Teda dostávame spor a platí  $\tau < 0$ .

Nech  $\gamma < \gamma_S$ . Dokážeme, že  $\tau > 0$ . Postupujeme analogicky ako v prípade  $\gamma > \gamma_S$ . Bod  $y_S$  označíme  $S$  a bod  $y_\gamma$  označíme  $K'$ . Sporom predpokladáme, že  $\tau < 0$ , čiže  $\nabla g_S(K')$  má opačný smer ako  $b$ . Označíme  $L'_\varepsilon = K' - \varepsilon b, \varepsilon > 0$ . Vektor  $-b$

smeruje k hodnotám  $b^T y < \gamma < \gamma_S$ , preto úsečka spájajúca  $L'_\varepsilon$  a  $S$  má priesečník s nadrovinou  $b^T y = \gamma$  — tento priesečník nech je  $M'_\varepsilon$ .  $\varepsilon$  zvolíme dostatočne malé na to, aby body  $L'_\varepsilon$  aj  $M'_\varepsilon$  ležali na  $\varphi$ -okolí bodu  $K'$ , na ktorom funkciu  $g_S$  dostatočne presne aproximuje lineárna funkcia určená deriváciou  $g_S(K')$ . Keďže predpokladáme  $\tau < 0$ , gradient  $\nabla g_S(K')$  má rovnaký smer ako vektor  $-b$ , preto  $g_S(S) \geq \alpha_{L'} := g_S(L') > g_S(K') \approx g_S(M')$ ,  $S$  je maximum funkcie  $g_S$ . Nadúrovnňová množina  $\{y \in \mathcal{D}^0 \mid g_S(y) \geq \alpha_{L'}\}$  konkávnej funkcie  $g_S$  nie je konvexná, pretože neobsahuje bod  $M'$  na úsečke spájajúcej  $S$  a  $L'$ , čím dostávam spor. Platí preto  $\tau > 0$ .

Dokázali sme zatiaľ ekvivalenciu  $\tau = 0 \Leftrightarrow \gamma = \gamma_S$  a implikácie

$$\tau < 0 \Leftrightarrow \gamma > \gamma_S, \quad \tau > 0 \Leftrightarrow \gamma < \gamma_S. \quad (13)$$

Obmenami implikácii (13) vzniknú ekvivalentné tvrdenia

$$\tau \geq 0 \Rightarrow \gamma \leq \gamma_S, \quad \tau \leq 0 \Rightarrow \gamma \geq \gamma_S. \quad (14)$$

Použitím ekvivalencie  $\gamma = \gamma_S \Leftrightarrow \tau = 0$  na (14) dostávame implikácie

$$\tau > 0 \Rightarrow \gamma < \gamma_S, \quad \tau < 0 \Rightarrow \gamma > \gamma_S. \quad (15)$$

Dokazované ekvivalencie vzniknú spojením (13) a (15).

Pre primárny prípad ukážeme iba ekvivalenciu  $\tau = 0 \Leftrightarrow \delta = \delta_S$ . Zvyšné tvrdenia platia analogicky s duálnym prípadom, iba by dôkaz bol zložitejší kvôli  $A^T u$ , členu navyše v podmienke  $\nabla f_S(x) - A^T u = \tau c$ . Keď  $\delta = \delta_S$ , maximalizujeme funkciu  $f_S$  na úrovni  $c^T x = \delta_S$ , teda úrovni na ktorej leží  $x_S$ , maximum  $f_S$  na množine  $\mathcal{P}$ . Analytický stred  $\mathcal{P}_\delta$  je teda analytický stred  $\mathcal{P}$ , v ktorom, ako sme ukázali v dôkaze vety 3.9 platí:

$$x^{-1} + A^T u' = 0, \quad Ax = b, \quad x > 0.$$

Porovnaním s podmienkou  $x^{-1} - A^T u = \tau c$  dostávame z jednoznačnosti Langrangeovho multiplikátora v úlohe konvexného programovania  $u' = -u$  a teda  $\tau = 0$ .

Naopak, ak predpokladáme  $\tau = 0$ , dostaneme z (9)  $x^{-1} - A^T u = 0$ . Z jednoznačnosti existencie analytického stredy dostávame, že riešenie  $x^{-1} - A^T u = 0$  je  $x_S$  a teda  $\delta = \delta_S$ .  $\square$

**Veta 3.12** (Reprezentácia centrálnej trajektórie cez analytické stredy úrovnňových množín duálnej účelovej funkcie). *Majme množiny  $\mathcal{D}_\gamma = \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^T y + s =$*

$c, s \geq 0, b^T y = \gamma$ , pričom  $b \neq 0$  (prípád  $b = 0$  je rozoberaný vo vete 3.9). Analytické stredy množín  $\mathcal{D}_\gamma$  pre  $\gamma \in (\gamma_S, \gamma_{max})$  sú práve duálna centrálna trajektória.

*Dôkaz.* Pripomenieme podmienky pre analytický stred  $\mathcal{D}_\gamma$  z lemy 3.10 (ľavý stĺpec) a  $\mu$ -centrujúci systém (pravý stĺpec):

$$(7) \quad -As^{-1} = \tau b \qquad Ax = b, \quad x > 0, \quad (C1)$$

$$(5) \quad A^T y + s = c, \quad s > 0 \qquad A^T y + s = c, \quad s > 0, \quad (C2)$$

$$(6) \quad b^T y = \gamma \qquad x = \mu s^{-1}. \quad (C3)$$

Dokážeme najprv, že bod spĺňajúci podmienky na analytický stred  $\mathcal{D}_\gamma, \gamma \in (\gamma_S, \gamma_{max})$  spĺňa aj podmienky  $\mu$ -centrujúceho systému a následne opačné tvrdenie, že pre bod spĺňajúci centrujúce podmienky spĺňa podmienky (7), (5), (6).

Nech  $\gamma \in (\gamma_S, \gamma_{max})$ . Podľa lemy 3.11 platí  $\tau < 0$ . Aby sme zo (7) dostali  $Ax = b$ , zavedieme ďalšiu rovnosť:

$$\begin{aligned} s^{-1} = -\tau x &\Leftrightarrow x = -\frac{1}{\tau} s^{-1}, & (16) \\ -As^{-1} = \tau b, & \\ \tau Ax = \tau b, & \\ Ax = b. & \end{aligned} \quad (17)$$

Keďže  $\tau \neq 0$ , môžeme bez problémov deliť. Aby sme z (16) dostali (C3), zostáva substituovať  $\mu = -\frac{1}{\tau}$ . Keďže  $\tau < 0$ , platí  $\mu = -\frac{1}{\tau} > 0$  a potom keďže aj  $s > 0$ , z (16) dostávame  $x > 0$ . Nerovnosti z  $\mu$ -centrujúceho systému,  $x > 0$  a  $s > 0$  sú teda splnené. Rovnosti z  $\mu$ -centrujúceho systému tvoria (17), (5), (16), pričom máme dodatočnú podmienku (6), čiže  $b^T y = \gamma$ . Táto podmienka je potrebná, pretože  $\mu$  je v úlohe hľadania analytických stredov neznáma navyše oproti  $\mu$ -centrujúcemu systému. K podmienke (6) vieme vyjadriť ekvivalentnú podmienku:

$$\begin{aligned} \mu e &= Xs \quad / \text{súčet po zložkách,} \\ n\mu &= x^T s = x^T (c - A^T y) = x^T c - x^T A^T y = c^T x - b^T y, \\ \mu &= \frac{c^T x - b^T y}{n} = \frac{c^T x - \gamma}{n}. \end{aligned} \quad (18)$$

Teraz ukážeme, že platí aj opačné tvrdenie. Z  $\mu$ -centrujúceho systému pre dané  $\mu > 0$  dostaneme podmienky na riešenie (7), (5) a (6), pričom platí  $\gamma \in (\gamma_S, \gamma_{max})$ .

Dosadením (C3) do (C1) dostávame

$$\mu As^{-1} = b, \quad \text{resp.} \quad -As^{-1} = -\frac{1}{\mu}b. \quad (19)$$

Vidíme, že podmienka (19) je (7) pre  $\tau = -\frac{1}{\mu}$ . Podmienka (5) je zhodná s (C2). Podmienka (6) je  $b^T y = \gamma$ . Riešenie  $\mu$ -centrujúceho systému, čiže bod centrálnej trajektórie čo je vnútorný bod množiny  $\mathcal{D}$ , preto  $\gamma = b^T y < \gamma_{max}$ . Nakoniec môžeme vďaka (19) použiť lemu 3.11 pre  $\tau = -\frac{1}{\mu} < 0$  a teda platí  $\gamma > \gamma_S$ .  $\square$

**Príklad 2.** Hľadáme analytický stred  $\mathcal{D}_\gamma$  pre náš príklad 1, pričom  $b_1, b_2 \neq 0$ :

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Riešime teda systém:

$$Ax = b, \quad A^T y + s = c, \quad Xs = \mu e, \quad \gamma = b^T y$$

pre nami zvolené  $\gamma$ . Z prvých troch rovností sme pri hľadaní centrálnej trajektórie (príklad 1) dostali:

$$b_1 y_1^2 + 2\mu y_1 - b_1 = 0, \quad (20)$$

$$b_2 y_2^2 + 2\mu y_2 - b_2 = 0. \quad (21)$$

Z podmienky  $\gamma = b^T y$  dostávame:

$$b_1 y_1 + b_2 y_2 = \gamma. \quad (22)$$

Vyjadríme  $y_2$  z (22):  $y_2 = \frac{\gamma - b_1 y_1}{b_2}$ , čo dosadíme do (21):

$$b_2 \frac{\gamma^2 - 2\gamma b_1 y_1 + b_1^2 y_1^2}{b_2^2} + 2\mu \frac{\gamma - b_1 y_1}{b_2} - b_2 = 0 \quad / \cdot b_2,$$

$$\gamma^2 - 2\gamma b_1 y_1 + b_1^2 y_1^2 + 2\mu\gamma - 2\mu b_1 y_1 - b_2^2 = 0 \quad / - b_1 \cdot (20),$$

$$\gamma^2 - 2\gamma b_1 y_1 + 2\mu\gamma - 4\mu b_1 y_1 - b_2^2 + b_1^2 = 0.$$

Z (20) vyjadríme  $2\mu$  a dosadíme:

$$2\mu = \frac{b_1 - b_1 y_1^2}{y_1} = b_1 \frac{1 - y_1^2}{y_1} \quad (23)$$

$$\begin{aligned} \gamma^2 - 2\gamma b_1 y_1 + \gamma b_1 \frac{1 - y_1^2}{y_1} - 2b_1 y_1 b_1 \frac{1 - y_1^2}{y_1} - b_2^2 + b_1^2 &= 0 \quad / \cdot y_1, \\ \gamma^2 y_1 - 2\gamma b_1 y_1^2 + \gamma b_1 (1 - y_1^2) - 2b_1^2 y_1 (1 - y_1^2) - b_2^2 y_1 + b_1^2 y_1 &= 0, \\ \gamma^2 y_1 - 2\gamma b_1 y_1^2 + \gamma b_1 - \gamma b_1 y_1^2 - 2b_1^2 y_1 + 2b_1^2 y_1^3 - b_2^2 y_1 + b_1^2 y_1 &= 0, \\ 2b_1^2 y_1^3 - 3\gamma b_1 y_1^2 + (\gamma^2 - b_2^2 - b_1^2) y_1 + \gamma b_1 &= 0. \end{aligned}$$

Pre  $y_1$  sme dostali kubickú rovnicu, ktorá môže mať tri korene. Ku každému z koreňov dopočítame  $y_2 = \frac{\gamma - b_1 y_1}{b_2}$  a za vyhovujúce riešenie vyberieme takú dvojicu, pre ktorú je  $s > 0$ , teda, ako sme už ukázali,  $-1 < y_1 < 1$  a  $-1 < y_2 < 1$ .

Pre konkrétne čísla  $b_1 = 2, b_2 = 1, \gamma = 1$  dostávame kubickú rovnicu:

$$\begin{aligned} 2 \cdot 2^2 y_1^3 - 3 \cdot 1 \cdot 2 y_1^2 + y_1 (1^2 - 2^2 - 1^2) + 1 \cdot 2 &= 0, \\ 8y_1^3 - 6y_1^2 - 4y_1 + 2 &= 0. \end{aligned}$$

MATLAB nájde približné korene polynómu ako  $y_1 : -0.6404, 1.0000, 0.3904$ . K nim sú príslušné  $y_2 = \frac{\gamma - b_1 y_1}{b_2} : 2.2808, -1.0000, 0.2192$ . V hraniciach  $(-1, 1)$  je tretia dvojica, teda

$$y = \begin{pmatrix} 0.3904 \\ 0.2192 \end{pmatrix}.$$

$\mu$  môžeme vyjadriť podľa (23):  $\mu = b_1 \frac{1 - y_1^2}{2y_1} = 2 \cdot \frac{1 - 0.3904^2}{2 \cdot 0.3904} = 2.1712$ .

Podľa lemy 3.8 vypočítame

$$\nabla g_S(y) = -As^{-1} = -A(c - A^T y)^{-1} = \begin{pmatrix} -0.9212 \\ -0.4606 \end{pmatrix}.$$

Pozorujeme, že skutočne platí z dôkazu vety 3.12.

$$-\mu \nabla g_S(y) = -2.1712 \begin{pmatrix} -0.9212 \\ -0.4606 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = b.$$

**Príklad 3.** Podobne ako v príklade 1 chceme zobrazit analytické stredy úrovnňových množín, tentokrát ale pre úlohu s neohraničenou prípustnou množinou. Zadanie je:

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad c = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Dostávame  $\mu$ -centrujúci systém:

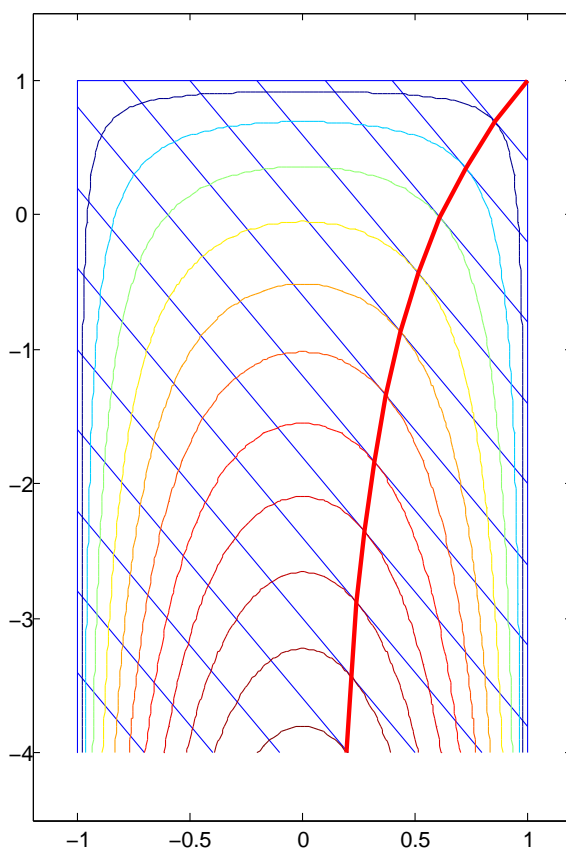
$$\begin{array}{lll} -x_1 + x_2 = 2, & -y_1 + s_1 = 1, & s_1 x_1 = \mu, \\ x_3 = 1, & y_1 + s_2 = 1, & s_2 x_2 = \mu, \\ & y_2 + s_3 = 1, & s_3 x_3 = \mu. \end{array}$$

Rovnice pre  $y_1$  sú zhodné s príkladom 1, dostávame teda  $y_1 = \frac{b_1}{\mu + \sqrt{\mu^2 + b_1^2}} = \frac{2}{\mu + \sqrt{\mu^2 + 4}}$ .

Už ľahšie sa vyjadrí  $y_2$ :

$$\begin{aligned} s_3 &= \frac{\mu}{x_3} = \frac{\mu}{1} = \mu \\ y_2 &= 1 - s_3 = 1 - \mu \end{aligned}$$

Vidíme, že centrálna trajektória je skutočne neohraničená, keďže  $\lim_{\mu \rightarrow +\infty} y_2 = \lim_{\mu \rightarrow +\infty} 1 - \mu = -\infty$ . Podobne ako na obrázkoch 7 a 8 nájdeme graficky analytické stredy rôznych  $\mathcal{D}_\gamma$ , čím dostaneme duálnu centrálnu trajektóriu pre príklad 3 (obr. 10).



Obr. 10: Analytické stredy  $\mathcal{D}_\gamma$  v príklade 3

□

Centrálnu trajektóriu teda vieme reprezentovať ako analytické stredy úrovnňových množín duálnej účelovej funkcie. Očakávame, že podobné tvrdenie platí aj pre analytické stredy primárnej účelovej funkcie. Znovu využijeme lemy 3.10 a 3.11.

**Veta 3.13** (Reprezentácia centrálnej trajektórie cez analytické stredy úrovnňových množín primárnej účelovej funkcie). *Majme množiny  $\mathcal{P}_\delta = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0, c^T x = \delta\}$ ,  $c \neq 0$ . Analytické stredy množín  $\mathcal{P}_\delta$  pre  $\delta \in (\delta_{min}, \delta_S)$  sú práve primárna centrálna trajektória.*

*Dôkaz.* Chceme ukázať, že z podmienok existencie analytického stredy  $\mathcal{P}_\delta$  pre  $\delta \in (\delta_{min}, \delta_S)$  dostaneme  $\mu$ -centrujúci systém teda že keď je bod  $x_\delta$  analytickým stredom  $\mathcal{P}_\delta$  pre  $\delta$  zo spomínaného intervalu, je to bod centrálnej trajektórie. Pripomenieme podmienky na analytický stred  $\mathcal{P}_\delta$  (ľavý stĺpec) a  $\mu$ -centrujúci systém (pravý stĺpec):

$$(9) \quad x^{-1} - A^T u = \tau c \quad Ax = b, \quad x > 0, \quad (C1)$$

$$(10) \quad Ax = b, \quad x > 0 \quad A^T y + s = c, \quad s > 0, \quad (C2)$$

$$(11) \quad c^T x = \delta \quad x = \mu s^{-1}. \quad (C3)$$

Upravíme (9). Predpokladáme  $\gamma \in (\gamma_{min}, \gamma_S)$ , preto podľa lemy ?? platí  $\tau > 0$  a môžeme  $\tau$  aj deliť.

$$\begin{aligned} x^{-1} - A^T u = \tau c & \quad / -\tau c; \quad \cdot(-1), \\ -x^{-1} + A^T u + \tau c = 0 & \quad / \cdot X, \\ -e + X(A^T u + \tau c) = 0 & \quad / + e; \quad \cdot \frac{1}{\tau}, \\ X(A^T \frac{u}{\tau} + c) = \frac{1}{\tau} e. & \end{aligned} \quad (24)$$

Označíme  $s = c + A^T \frac{u}{\tau}$  a substituujeme  $u = -\tau y$  aby  $s = c - A^T y$  Nakoniec označíme  $\mu = \frac{1}{\tau}$ . Dostávame:

$$(10) \quad Ax = b, \quad A^T y + s = c, \quad (24) \quad Xs = \mu e, \quad (11) \quad c^T x = \delta.$$

Dostali sme teda  $\mu$ -centrujúci systém s dodatočnou podmienkou  $c^T x = \delta$ . Ako už vieme, z rovníc  $\mu$ -centrujúceho systému platí  $\mu = \frac{c^T x - b^T y}{n}$ , preto ekvivalentná podmienka ku  $c^T x = \delta$  je  $\mu = \frac{\delta - b^T y}{n}$ .

Zostáva určiť, či platia nerovnosti z  $\mu$ -centrujúceho systému. Máme už  $x > 0$  a vieme, že  $\tau > 0$ , preto  $\mu = \frac{1}{\tau} > 0$ . Z  $Xs = \mu e$  potom platí  $s > 0$ .

Ukážeme, že platí opačné tvrdenie, teda budeme vychádzať z bodu, ktorý je riešením  $\mu$ -centrujúceho systému a ukážeme, že tento bod je analytickým stredom množiny  $\mathcal{P}_\delta$ , pričom  $\delta \in (\delta_{min}, \delta_S)$ .

Majme teda bod, pre ktorý platia centrujúce podmienky (C1), (C2), (C3). Dosadením  $s$  z (C2) do (C3) dostávame:

$$\begin{aligned} X(c - A^T y) &= \mu e & / \cdot X^{-1}; \cdot \frac{1}{\mu}, \\ \frac{1}{\mu} c - A^T \frac{y}{\mu} &= x^{-1}, \\ \tau c &= x^{-1} + A^T(\tau y). \end{aligned} \tag{25}$$

Podľa (25) môžeme použiť lemu 3.11. Platí  $\tau = \frac{1}{\mu} > 0$ , preto  $b^T y = \delta \in (\delta_{min}, \delta_S)$ , splnená je podmienka (11). Rovnosť (25) je po substitúcii  $u = -\tau y$  zhodná s (9) a (C1) nám dáva (10). Rovnosti z podmienok na analytický stred  $\mathcal{P}_\delta$  sú splnené. Nerovnosť  $x > 0$  je tiež splnená, a to už z podmienok  $\mu$ -centrujúceho systému.  $\square$

### 3.3 Centrálna trajektória a nadúrovňové a podúrovňové množiny

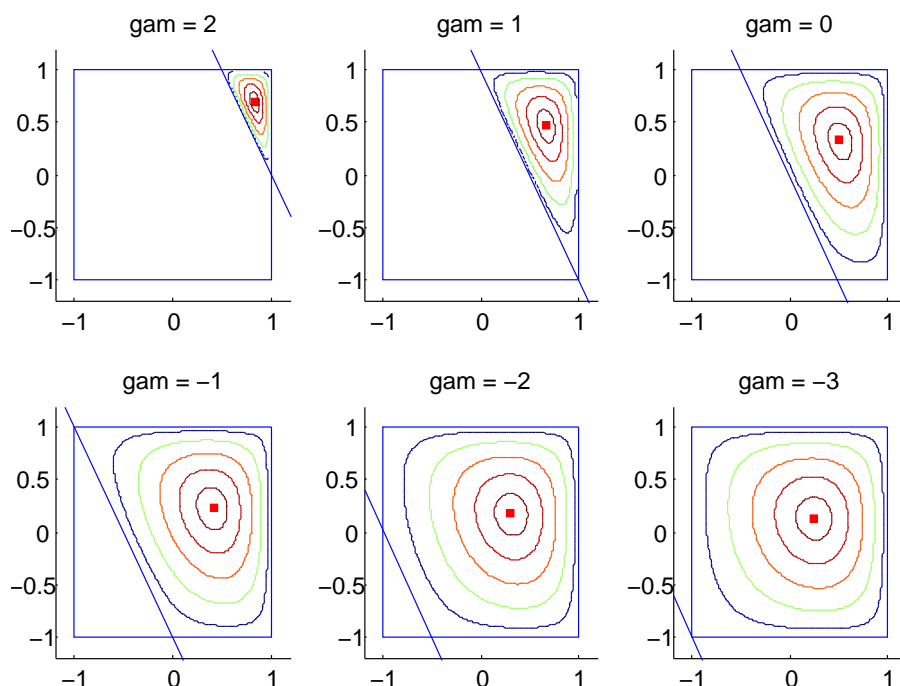
Ukázali sme, že analytické stredy úrovňových množín tvoria centrálnu trajektóriu. Podľa dôsledku 3.7 aj podúrovňové množiny primárnej účelovej funkcie  $\mathcal{P}_{\delta-} = \{x \in \mathcal{P} \mid c^T x \leq \delta\}$  a nadúrovňové množiny duálnej účelovej funkcie  $\mathcal{D}_{\gamma+} = \{y \in \mathcal{D} \mid b^T y \geq \gamma\}$  majú analytické stredy pre vhodné  $\gamma, \delta$ . Pozrieme sa na tieto analytické stredy najprv pre príklad 1.

Máme nájsť analytický stred množiny  $\mathcal{D}'_{\gamma+} = \{(y, s) \mid A^T y + s = c, s \geq 0, b^T y \geq \gamma\}$ . V definícii pre analytický stred sa však vyskytujú iba ohraničenia v tvare rovnosti, preto  $b^T y \geq \gamma$  prepíšeme ako  $b^T y - \sigma = \gamma, \sigma \geq 0$ . Nová doplnková premenná  $\sigma$  je nezáporná, preto podľa definície 3.1 pri hľadaní analytického stredy  $\mathcal{D}_{\gamma+}$  vystupuje v účelovej funkcii aj  $\ln \sigma$ . Analytický stred  $\mathcal{D}_{\gamma+}$  je teda riešenie úlohy:

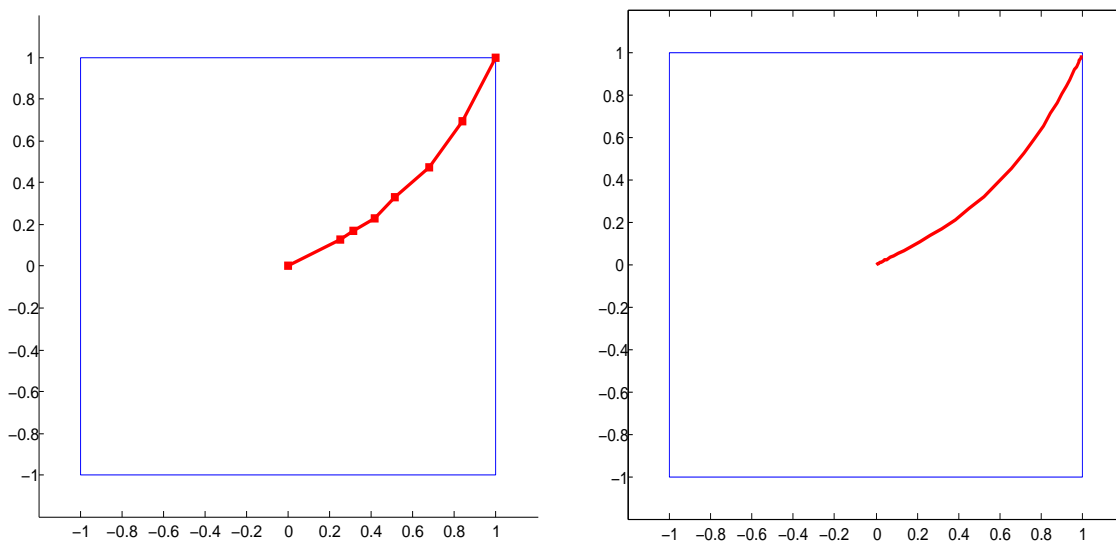
$$\max \left\{ \ln \sigma + \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0, b^T y - \sigma = \gamma, \sigma > 0 \right\}.$$

Pre niekoľko hodnôt  $\gamma$  je približné maximum spolu s vrstevnicami účelovej funkcie zobrazené na obrázku 11.





Obr. 11: Hľadanie analytických stredov  $\mathcal{D}_{\gamma+}$  pre množinu z príkladu 1

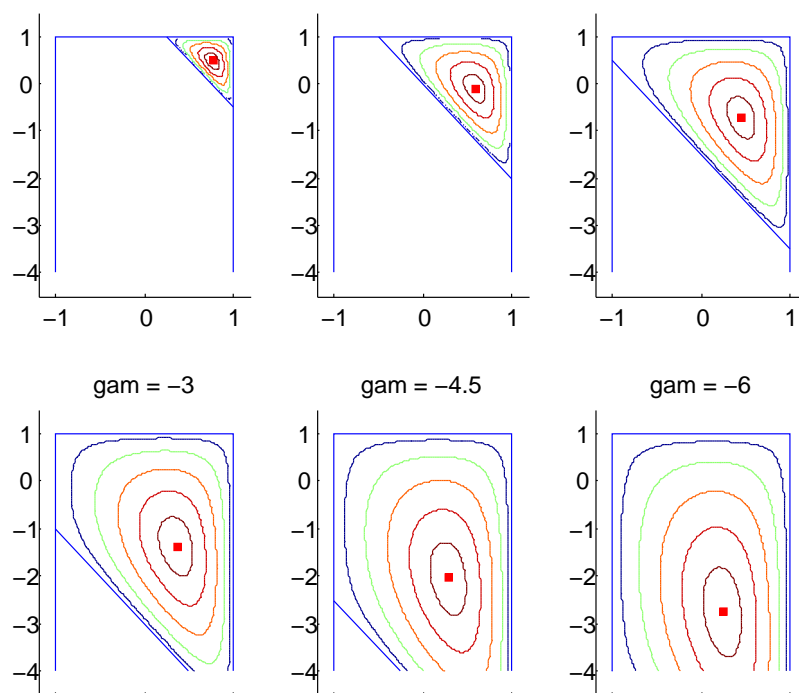


Obr. 12: Analytické stredy  $\mathcal{D}_{\gamma+}$  (pr. 1)

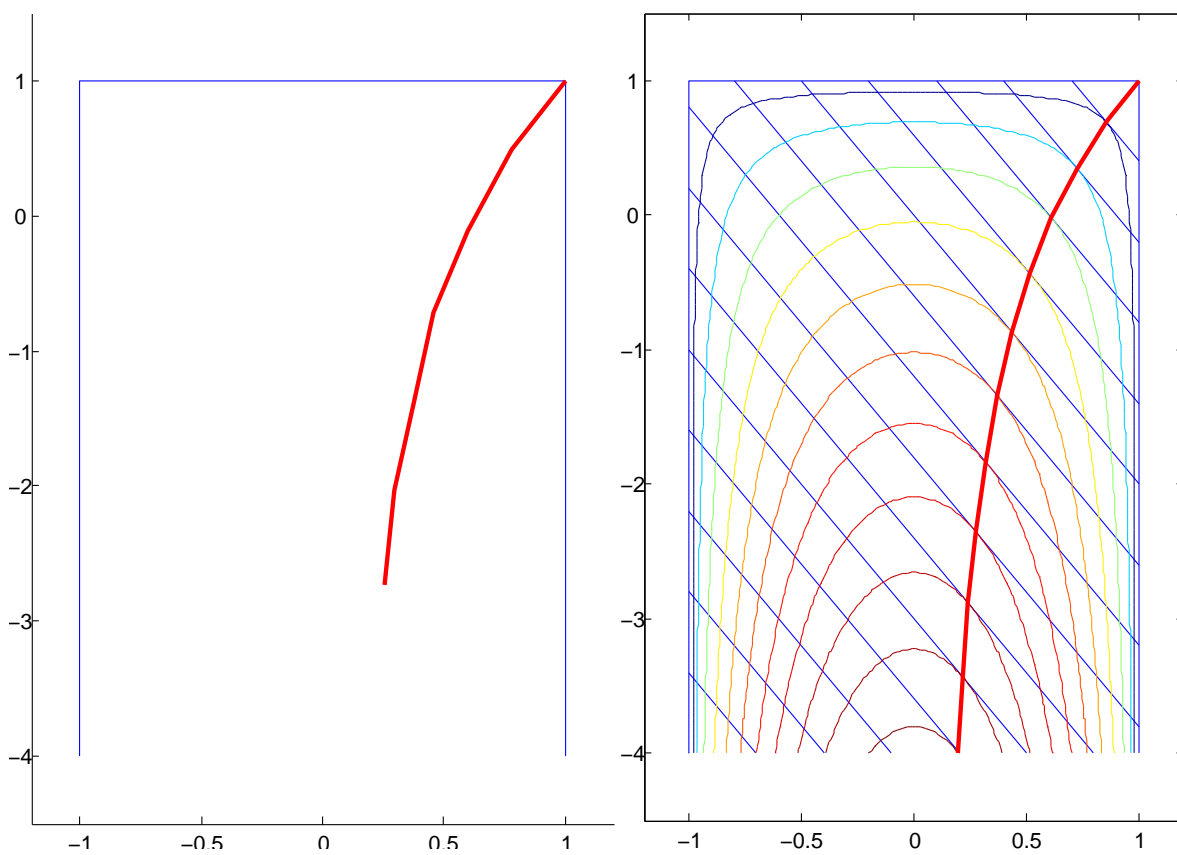
Obr. 2: Duálna centrálna trajektória (pr. 1)

Pospájané body maxim sú na obr. 12. Rovnakým postupom zhotovíme obrázky pre množinu z príkladu 3. V oboch prípadoch sme dostali krivku, ktorá znovu vyzerá ako centrálna trajektória. Že to tak naozaj je, ukážeme vo všeobecnosti v nasledujúcej vete.

**Veta 3.14** (Reprezentácia centrálnej trajektórie cez analytické stredy nadúrovňových



Obr. 13: Hľadanie analytických stredov  $\mathcal{D}_{\gamma+}$  pre množinu z príkladu 3



Obr. 14: Analytické stredy  $\mathcal{D}_{\gamma+}$  (pr. 3)

Obr. 10: Duálna centrálna trajektória (pr. 3)

množín duálnej účelovej funkcie). Majme množinu  $\mathcal{D}_{\gamma+} = \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^T y + s = c, s \geq 0, b^T y - \sigma = \gamma, \sigma \geq 0\}$ , pričom  $b \neq 0$ . Analytické stredy množín  $\mathcal{D}_{\gamma+}$  pre  $\gamma < \gamma_{max}$  sú duálna centrálna trajektória.

*Dôkaz.* Chceme získať nutné a postačujúce podmienky pre bod analytického stredy  $\mathcal{D}_{\gamma+}$ . Podľa dôsledku 3.7 riešenie existuje a je jediné práve pre  $\gamma < \gamma_{max}$ . Vieme teda napísať nutné a postačujúce podmienky na riešenie:

$$\max \left\{ \ln \sigma + \sum_{i=1}^n \ln s_i \mid A^T y + s = c, s > 0, b^T y - \sigma = \gamma, \sigma > 0 \right\}.$$

Lagrangeova funkcia je:

$$L(y, s, \sigma, u, \tau) = -\ln \sigma - \sum_{i=1}^n \ln s_i + u^T (A^T y + s - c) + \tau (b^T y - \sigma - \gamma),$$

$$L : \mathbb{R}^m \times \mathbb{R}_{++}^n \times \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}.$$

Nutné aj postačujúce podmienky:

$$\nabla_y L = Au + \tau b = 0 \quad \Rightarrow \quad -A \frac{u}{\tau} = b, \quad (26)$$

$$\nabla_s L = -s^{-1} + u = 0 \quad \Rightarrow \quad Su = e, \quad (27)$$

$$\nabla_\sigma L = -\sigma^{-1} - \tau = 0 \quad \Rightarrow \quad \sigma = -\frac{1}{\tau}, \quad (28)$$

$$\nabla_u L = A^T y + s - c = 0 \quad \Rightarrow \quad A^T y + s = c, \quad (29)$$

$$\nabla_\tau L = b^T y - \sigma - \gamma = 0 \quad \Rightarrow \quad b^T y - \sigma = \gamma. \quad (30)$$

Všimneme si, že  $\tau = -\frac{1}{\sigma} < 0$ , preto  $\tau \neq 0$  a môžeme  $\tau$  deliť. Snažíme sa upraviť podmienky na rovnice  $\mu$ -centrujúceho systému. Substituujeme  $u = -\tau x$  aby sa z (26) stalo (C1):  $Ax = b$ . (27) sa tým zmení na  $Sx = -\frac{1}{\tau} e = \sigma e$  podľa (28). Nahradením  $\mu = \sigma$  dostávame (C3):  $Sx = \mu e$ . Pridaním (29)~(C2) dostávame rovnosti  $\mu$ -centrujúceho systému. Dodatočná podmienka je (30), čiže  $b^T y - \sigma = b^T y - \mu = \gamma$ .

Aby boli splnené aj nerovnosti z  $\mu$ -centrujúceho systému, zostáva ukázať, že  $x > 0$ . Platí  $\mu = \sigma > 0$ , preto z  $Xs = \mu e$  vyplýva  $x > 0$ .

K podmienke  $b^T y - \mu = \gamma$  vieme znova vyjadriť ekvivalentnú podmienku dosadením

$b^T y$  do  $\mu = \frac{c^T x - b^T y}{n}$ .

$$\begin{aligned}\mu &= \frac{c^T x - b^T y}{n}, \\ \mu &= \frac{c^T x - \gamma - \mu}{n} \quad / \cdot n; +\mu, \\ n\mu + \mu &= c^T x - \gamma \quad / : (n+1), \\ \mu &= \frac{c^T x - \gamma}{n+1}.\end{aligned}$$

Ukážeme opačné tvrdenie, že bod spĺňajúci  $\mu$ -centrujúci systém je analytickým stredom nejakej  $\mathcal{D}_{\gamma+}$ ,  $\gamma < \gamma_{max}$ . Majme teda splnené pre zadané  $\mu > 0$ :

$$(C1) Ax = b, \quad x > 0, \quad (C2) A^T y + s = c, \quad s > 0, \quad (C3) Xs = \mu e.$$

Označíme  $\mu = \sigma = -\frac{1}{\tau}$ , čím je splnené (28). Označíme  $\gamma = b^T y - \sigma = b^T y - \mu$  čo je (30), pričom keďže  $\mu > 0$ , dostávame  $\gamma = b^T y - \mu < b^T y < \gamma_{max}$ . Zavedením  $x = -\frac{u}{\tau} = \mu u$  dostávame (26) z (C1) a zároveň z (C3) máme  $\mu S u = \mu e$ , čo po vydelení  $\mu$  dáva (27). Zostávajúca rovnosť, (29) je priamo (C2). Jediná nerovnosť z podmienok na analytický stred  $\mathcal{D}_{\gamma+}$ ,  $s > 0$ , už je splnená v  $\mu$ -centrujúcom systéme.  $\square$

*Poznámka:* Pri hľadaní analytického stredy  $\mathcal{D}_{\gamma+}$  vieme interpretovať  $\mu$  vďaka podmienke  $\mu = b^T y - \gamma$ . Keď podmienku prepíšeme na  $b^T y - \mu = \gamma$ , podľa vety 1.14 je vzdialenosť nadroviny  $b^T y = \gamma$  a analytického stredy  $\{y \in \mathcal{D} \mid b^T y \geq \gamma\}$  rovná  $\frac{\mu}{\|b\|}$ .

Už skrátene sa zameriame na primárny prípad. Keďže úloha je minimalizačná, má zmysel sa zaoberať podúrovňovými množinami  $\mathcal{P}_{\delta-} = \{x \in \mathcal{P} \mid c^T x \leq \delta\}$ , ktoré sú podľa vety 1.13 ohraničené.

**Veta 3.15** (Reprezentácia centrálnej trajektórie cez analytické stredy podúrovňových množín primárnej účelovej funkcie). *Majme množinu  $\mathcal{P}_{\delta-} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0, c^T x + \rho = \delta, \rho \geq 0\}$ . Analytické stredy množín  $\mathcal{P}_{\delta-}$  pre  $\delta > \delta_{min}$  sú primárna centrálna trajektória.*

*Dôkaz.* Podrobne dokazovať nebudeme, dôkaz je analógiou dôkazov viet 3.14 a 3.13. Množina  $\mathcal{P}_{\delta-}$  má analytický stred pre  $\delta > \delta_{min}$ . Z podmienok optimality pre riešenie úlohy

$$\max \left\{ \ln \rho + \sum_{i=1}^n \ln x_i \mid Ax = b, x > 0, c^T x + \rho = \delta, \rho > 0 \right\}$$

dostaneme systém

$$Ax = b, x > 0 \quad A^T y + s = c \quad Sx = \mu e \quad \mu = \delta - c^T x \quad \left( \mu = \frac{\delta - b^T y}{n+1} \right). \quad (31)$$

To je  $\mu$ -centrujúci systém s dodatočnou podmienkou  $\mu = \delta - c^T x$ , resp. podmienkou k nej ekvivalentnou  $\mu = \frac{\delta - b^T y}{n+1}$ . Podobne ako v dôkaze vety 3.14 platí  $\mu = \rho > 0$ , preto platí  $s = \mu x^{-1} > 0$ .

Rovnako platí opačné tvrdenie, zo systému (31) sa dajú získať centrujúce podmienky, pričom  $\delta > \delta_{min}$ . □

## Záver

V práci sme ukázali, že centrálna trajektória v lineárnom programovaní sa dá interpretovať ako množina analytických stredov rôznych množín. Značnú časť práce sme venovali definícii analytického stredy množiny (definícia 3.1) a niektorým dôležitým základným vlastnostiam ako existencii a jednoznačnosti, ako vo všeobecnom prípade (veta 3.4), tak aj špeciálne pre množiny nášho záujmu,  $\mathcal{P}$ ,  $\mathcal{P}_\delta$ ,  $\mathcal{P}_{\delta-}$ ,  $\mathcal{D}$ ,  $\mathcal{D}_\gamma$ ,  $\mathcal{D}_{\gamma+}$  (dôsledok 3.7). Dôsledok 3.7 bol často citovaný pri dokazovaní viet o interpretácii 3.12, 3.13, 3.14 a 3.15. Zhrnutie výsledkov týchto štyroch viet sa nachádza v tabuľke 1. Stĺpce tabuľky sú: množina M, čiže množina, ktorej analytický stred v danej vete hľadáme; označenie množiny M používané v texte; dodatočná podmienka v úlohe hľadania analytického stredy k podmienkam  $\mu$ -centrujúceho systému; k nej ekvivalentná alternatívna podmienka; reprezentácia  $\mu$ , kde  $\tau$  je  $g_S(y) = \tau b$ ,  $d$  je vzdialenosť nadroviny  $b^T y = \gamma$  ( $c^T x = \delta$ ) a bodu analytického stredy M.

Množina M	Ozn. M	Úrovne	Dod. podm.	Alt. podm.	Interp. $\mu$
$b^T y = \gamma, y \in \mathcal{D}$	$\mathcal{D}_\gamma$	$(\gamma_S, \gamma_{max})$	$b^T y = \gamma$	$\mu = \frac{c^T x - \gamma}{n}$	$\mu = -\frac{1}{\tau}$
$c^T x = \delta, x \in \mathcal{P}$	$\mathcal{P}_\delta$	$(\delta_{min}, \delta_S)$	$c^T x = \delta$	$\mu = \frac{\delta - b^T y}{n}$	
$b^T y \geq \gamma, y \in \mathcal{D}$	$\mathcal{D}_{\gamma+}$	$\gamma < \gamma_{max}$	$b^T y - \mu = \gamma$	$\mu = \frac{c^T x - \gamma}{n+1}$	$d = \frac{\mu}{\ b\ }$
$c^T x \leq \delta, x \in \mathcal{P}$	$\mathcal{P}_{\delta-}$	$\delta > \delta_{min}$	$c^T x + \mu = \delta$	$\mu = \frac{\delta - b^T y}{n+1}$	$d = \frac{\mu}{\ c\ }$

**Tabuľka 1:** Zhrnutie výsledkov viet 3.12 až 3.15

Výsledky môžu mať teoretický význam — mohli by sme vďaka nim odvodzovať ďalšie vlastnosti centrálnej trajektórie. Navyše je prístup geometricky názorný. Tvrdenia ktorých sa práca dotkla môžu mať aj praktický význam — hoci dnešné praktické algoritmy skôr nami skúmané interpretácie nevyužívajú, toto sa môže v budúcnosti zmeniť. Text môže poslúžiť aj pre záujemcu o problematiku metód vnútorného bodu ako zhrnutie istých poznatkov a možno trochu odlišný pohľad na problematiku, ako používa väčšina učebníc. Vytváranie práce tiež dodalo autorovi množstvo skúseností so samostatnou prácou, so samostatným riešením problémom a taktiež s dodržiavaním formálnej stránky práce.

## Zoznam použitej literatúry

- [1] Dantzig, G. B.: *Linear Programming*, prístupné na internete: [http://www2.informs.org/History/dantzig/LinearProgramming\\_article.pdf](http://www2.informs.org/History/dantzig/LinearProgramming_article.pdf) (3. 12. 2011)
- [2] Halická, M.: *Dvadsať rokov moderných metód vnútorného bodu*, Pokroky matematiky, fyziky a astronomie 49 (2004), 234-244, prístupné na internete: <http://www.iam.fmph.uniba.sk/institute/halicka/clanoklatex.pdf>
- [3] Halická, M.: *Metódy vnútorného bodu v lineárnom programovaní*, prednášky, FMFI UK, Bratislava, 2011
- [4] Halická, M.: *osobná komunikácia*, FMFI UK, Bratislava, 2011, 2012
- [5] Klee, V., Minty, G. J.: *How good is the simplex algorithm?* Inequalities III, Academic Press New York-London (1972) 159–175
- [6] Lesaja, G.: *Introducing Interior-Point Methods for Introductory Operations Research Courses and/or Linear Programming Courses*, The Open Operational Research Journal 3 (2009), 1-12,, prístupné na internete: <http://www.benthamscience.com/open/toorj/V003-GenArt.htm> (3. 12. 2011)
- [7] Plesník, J.: *Lineárne programovanie*, prednášky, FMFI UK, Bratislava, 2011
- [8] Roos, C., Terlaky, T., Vial, J. P.: *Interior point methods for linear optimization*, Springer, New York, 2006.
- [9] Todd, M. J.: *The many facets of linear programming*, Springer-Verlag 2001, prístupné na internete: <http://cowles.econ.yale.edu/lec-lun/2001/todd-011024.pdf>
- [10] Vanderbei, R. J.: *Linear programming: Foundations and Extensions*, Kluwer, Boston, 1997