Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského v Bratislave

Katedra aplikovanej matematiky a štatistiky

Metódy vnútorného bodu v lineárnom programovaní

(Diplomová práca)

Milan Sedlák

Vedúci diplomovej práce

Doc. RNDr. Margaréta Halická, CSc.

Bratislava, 2010

Metódy vnútorného bodu v lineárnom programovaní

DIPLOMOVÁ PRÁCA

Milan Sedlák

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA APLIKOVANEJ MATEMATIKY A ŠTATISTIKY

Študijný odbor: 9.1.9 Aplikovaná matematika
Študijný program: Ekonomická a finančná matematika

Vedúci práce: Doc. RNDr. Margaréta Halická, CSc.

BRATISLAVA 2010

Čestné prehlásenie

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

. . . .

Milan Sedlák

Poďakovanie

Ďakujem vedúcej diplomovej práce Doc. RNDr. Margaréte Halickej, CSc., za cenné rady a pripomienky, ktoré mi veľmi pomohli pri vypracovaní tejto práce.

Abstrakt

SEDLÁK, Milan: Metódy vnútorného bodu v lineárnom programovaní /Diplomová práca, 2010/
Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky.
Vedúci práce: Doc. RNDr. Margaréta Halická, CSc.

Hlavnou témou diplomovej práce je optimálne riešenie úlohy lineárneho programovania s najmenšou Euklidovskou normou. Diplomová práca je teoreticky zameraná a pozostáva z dvoch kapitol.
V prvej časti autor predstavuje optimálne riešenie úlohy LP s najmenšou Euklidovskou normou. Popisuje špeciálne riešenie úlohy LP, ktoré je najbližšie k zadanému bodu (GLTN-riešenie), a niektoré jeho vlastnosti. Podáva prehľad algoritmov pre hľadanie takéhoto riešenia.
Druhá kapitola je venovaná regularizovanej centrálnej trajektórii. Táto trajektória je modifikáciou centrálnej trajektórie známej z metód vnútorného bodu. Na rozdiel od nej však existuje pre každú úlohu LP a konverguje k GLTN-riešeniu úlohy LP. Autor popisuje a dokazuje základné vlastnosti regularizovanej centrálnej trajektórie: existenciu a jednoznačnosť, konvergenciu a analytickosť.

Kľúčové slová:
lineárne programovanie, optimálne riešenie s najmenšou normou, regularizovaná centrálna trajektória

Abstract

SEDLÁK, Milan: Interior Point Methods in Linear Programming /Master thesis, 2010/ Comenius University, Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics.
Advisor: Doc. RNDr. Margaréta Halická, CSc.

The main topic of the Master thesis is the least 2-norm solution of linear program. The Master thesis is theoretically oriented and consists of two chapters.
In the first part the author introduces the least 2-norm solution of linear program. He describes special solution of linear program, which is closest to the given point (GLTN-solution), and some of the properties of this solution. He gives an overview of the methods for seeking the GLTN-solution.
The second chapter is devoted to the regularized central path. This trajectory is a modification of the central path, known from the interior-point methods. Unlike the central path, the regularized central path exists for each linear program and it converges to the GLTN-solution of the linear program. The author describes and prooves the basic properties of the regularized central path: existence and uniqueness, convergence and analyticity.

Key words:
linear programming, least 2-norm solution, regularized central path

Obsah

Kapitola 1

Úvod

Lineárne programovanie (LP) je v optimalizácii dobre známou oblasťou. Je to optimalizačný nástroj, ktorého úlohou je minimalizácia (prípadne maximalizácia) lineárnej účelovej funkcie pri lineárnych ohraničeniach. Jeho teoretický aj praktický význam je enormný. Ten praktický spočíva v množstve aplikácií v najrozmanitejších oblastiach. Teoretický význam spočíva v tom, že myšlienky a postupy, ktoré sa aplikujú v LP, je často možné zovšeobecniť a použiť aj na zložitejšie optimalizačné úlohy. Vďaka tomu LP istým spôsobom dláždi cestu k pokrokom v zložitejších oblastiach optimalizácie.
Úlohy LP vieme riešiť už od r. 1947, kedy bola G. B. Dantzigom navrhnutá tzv. simplexová metóda. To bol prvý veľký míľnik v LP. Druhým bol r. 1984, kedy N. Karmarkar prišiel s algoritmom, ktorý je dnes považovaný za začiatok éry tzv. metód vnútorného bodu v LP. Dodnes sa pri riešení úloh LP najčastejšie používajú práve tieto dve triedy metód.
Ak má úloha LP len jedno optimálne riešenie, potom samozrejme každá metóda nájde práve toto riešenie. Môže sa však stať, že úloha LP má viacero riešení. V takom prípade sa jednotlivé metódy líšia typom riešenia, ktoré nájdu. Simplexová metóda (a jej príbuzné metódy) nájdu ako optimálne riešenie krajný bod množiny prípustných riešení. Naopak, metódy vnútorného bodu nájdu ako riešenie taký bod, ktorý je „čo najďalej“ od hraníc množiny prípustných riešení (tzv. analytický stred množiny). Tento poznatok

môže ovplyvniť výber metódy na riešenie konkrétnej úlohy: v niektorých prípadoch totiž charakter úlohy predurčuje istý typ optimálneho riešenia ako výhodnejší.

Hlavným objektom záujmu v tejto práci je špeciálny typ riešenia úlohy LP - optimálne riešenie s najmenšou normou. Práca má teoretické zameranie a pozostáva z dvoch častí (Kapitoly 2 a 3). V tej prvej sa venujeme základným pojmom a prehľadu niektorých doterajších výsledkov v tejto oblasti. V druhej kapitole rozoberáme jednu z novších metód pre hľadanie optimálneho riešenia úlohy LP s najmenšou normou. Ide o metódu, ktorá funguje na podobnom princípe, ako metódy vnútorného bodu v LP. Pri výpočte sleduje istú trajektóriu, ktorá nás povedie k hľadanému riešeniu. Našim cieľom je analyzovať vlastnosti tejto trajektórie, ktoré sú dôležité pri konštrukcii algoritmov na hľadanie optimálneho riešenia úlohy LP s najmenšou normou. Podávame podrobný popis vlastností trajektórie spolu s príslušnými dôkazmi. Pri analýze vychádzame z článku [15], kde je popísaná trajektória pre úlohy LP. Jednotlivé vlastnosti trajektórie však nie sú priamo dokazované - autori len odkazujú na dôkazy vlastností trajektórie pre úlohy lineárnej komplementarity. Druhý článok z ktorého čerpáme je [6], kde sú vlastnosti trajektórie rozoberané podrobnejšie, avšak pre úlohy semidefinitného programovania.

Kapitola 2

Optimálne riešenia s najmenšou normou

Cieľom tejto kapitoly je uviesť čitateľa do problematiky optimálneho riešenia úlohy LP s najmenšou normou. Kapitola má nasledujúcu štruktúru: v prvej časti definujeme úlohu LP a uvádzame vybrané základné výsledky z teórie LP. V druhej časti definujeme optimálne riešenie úlohy LP s najmenšou normou a ďalej sa venujeme základným informáciám o ňom. Záverečná časť kapitoly podáva prehľad a stručný popis niektorých metód, ktoré dokážu nájsť takéto riešenie.

V práci je použité štandardné označenie pre literatúru o metódach vnútorného bodu. Vektory uvažujeme ako stĺpcové. Tu je zoznam niektorých použitých symbolov:

$\mathcal{R}^n$	priestor reálnych n-rozmerných vektorov
$\mathcal{R}^n_+$	priestor nezáporných n-rozmerných vektorov
$\|x\|$	Euklidovská norma vektora x, tiež $\|x\|_2$
e	vektor jednotiek (rozmer je určený kontextom), $e = (1, 1, ..., 1)^T$
I	diagonálna matica s jednotkami na diagonále, $I = diag(e)$
xs	Hadamardov súčin vektorov, $(xs)_i = x_i s_i$

Pre vektor x definujeme:

X matica s vektorom x na diagonále, $X = diag(x)$

x^{-1} $(x^{-1})_i = \frac{1}{x_i}$

X^{-2} $X^{-2} = diag(x^{-2})$

2.1 Úloha lineárneho programovania

Úlohou lineárneho programovania rozumieme hľadanie viazaného extrému lineárnej funkcie pri ohraničeniach v tvare lineárnych rovníc a nerovníc. Existuje niekoľko rôznych formulácii (tvarov) tejto úlohy, ktoré sa navzájom líšia:

-typom ohraničení: len rovnosti, len nerovnosti, zmiešané ohraničenia

-viazanosťou premenných: nezáporné premenné, voľné premenné, zmiešané.

Pre názornosť uveďme všeobecný tvar úlohy LP. Ako je zaužívané v literatúre, m predstavuje počet ohraničení úlohy a n počet premenných. Definujme množinu indexov všetkých premenných $J = \{1, 2, ..., n\}$ a jej podmnožinu $J_{NN} \subseteq J$, ktorá zahŕňa indexy nezáporných premenných. Potom všeobecným tvarom úlohy LP nazývame úlohu

$$
\begin{aligned}
\min\ & c^T x \\
A_E x &= b_E \\
A_I x &\leq b_I \\
x_j &\geq 0, \quad j \in J_{NN},
\end{aligned}
$$

kde $c, x \in \mathcal{R}^n$, $b_E \in \mathcal{R}^{m_1}$, $b_I \in \mathcal{R}^{m-m_1}$, $A_E \in \mathcal{R}^{m_1 \times n}$, $A_I \in \mathcal{R}^{(m-m_1)\times n}$ a m_1 označuje počet rovností spomedzi ohraničení.

Všeobecný tvar úlohy LP predstavuje schému, z ktorej môžeme odvodiť jednotlivé špeciálne tvary. Tieto tvary úlohy LP sú navzájom ekvivalentné: úlohu v jednom tvare vieme previesť na ekvivalentnú úlohu v inom tvare. Preto sa môžeme bez ujmy na všeobecnosti

zaoberať len jedným konkrétnym tvarom úlohy.
Budeme používať tzv. rovnicový tvar úlohy LP (ktorý je typicky preferovaný pri metódach vnútorného bodu) :

$$(P) \qquad \min\{c^T x : Ax = b, x \geq 0\}, \tag{2.1}$$

pričom $A \in \mathcal{R}^{m \times n}$, $x, s, c \in \mathcal{R}^n$, $y, b \in \mathcal{R}^m$. Bez ujmy na všeobecnosti budeme predpokladať, že platí $m \leq n$ a matica A má plnú hodnosť, t.j. $h(A) = m$.
Úlohu (P) budeme nazývať primárnou úlohou. Preusporiadaním údajov v primárnej úlohe dostaneme inú úlohu LP, ktorá sa nazýva duálnou úlohou k úlohe (P):

$$(D) \qquad \max\{b^T y : A^T y + s = c, s \geq 0\}. \tag{2.2}$$

Úlohy (P), (D) spolu úzko súvisia. Vzťahmi medzi nimi sa zaoberá teória duality. Množinou prípustných riešení úlohy (P) nazývame množinu

$$\mathcal{P} = \{x \in \mathcal{R}^n : Ax = b, x \geq 0\}$$

a množinou optimálnych riešení množinu

$$\mathcal{P}^* = \{x^* \in \mathcal{P} : c^T x^* \leq c^T x \;\; \forall x \in \mathcal{P}\}.$$

Analogicky pre úlohu (D) definujeme

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

$$\mathcal{D}^* = \{(y^*, s^*) \in \mathcal{D} : b^T y^* \geq b^T y \;\; \forall (y, s) \in \mathcal{D}\}.$$

O možných výsledkoch, ktoré môžeme dostať riešením úlohy LP, hovorí tzv. základná veta lineárneho programovania:

Veta 1 (Základná veta LP, [9]) *Pre ľubovoľnú úlohu LP nastáva práve jedna z týchto troch možností:*

1. *Neprípustnosť: množina prípustných riešení je prázdna*

2. *Neohraničenosť: účelová funkcia je na množine prípustných riešení v smere optimalízácie neohraničená*

3. *Optimalita: úloha má optimálne riešenie.*

V nasledujúcom texte sa budeme takmer výlučne zaoberať úlohami, pre ktoré platí tretia možnosť z predchádzajúcej vety. Takéto úlohy nazývame riešiteľnými. Z teórie duality v LP vieme, že riešiteľnosť primárnej a duálnej úlohy sú úzko späté: úlohy (P), (D) buď sú obe riešiteľné (majú optimálne riešenia), alebo ani jedna z nich nemá optimálne riešenie. Nutnou a postačujúcou podmienkou pre riešiteľnosť úloh (P), (D) je existencia prípustného riešenia pre obe úlohy. Predpoklad riešiteľnosti úloh (P), (D) preto môžeme formulovať dvoma ekvivalentnými spôsobmi:
- úlohy (P), (D) sú riešiteľné
- obe úlohy (P), (D) majú prípustné riešenia.
Metódy vnútorného bodu využívajú charakterizáciu optimálneho riešenia, ktorá prevádza úlohu LP na riešenie nelineárneho systému rovníc. Ukazuje sa, že riešením tohto systému dostaneme súčasne optimálne riešenie pre primárnu aj pre duálnu úlohu. Uvedenú charakterizáciu možno odvodiť z duálnych vzťahov medzi úlohami (P), (D). Iným spôsobom odvodenia je vyvodiť tento systém ako špeciálny prípad KKT (Karush-Kuhn-Tucker) podmienok optimality. Pre úlohu LP sú totiž KKT podmienky nutnými aj postačujúcimi podmienkami optimality.

Veta 2 (KKT podmienky optimality, [14]) *Vektor $x^* \in \mathcal{R}^n$ je optimálnym riešením úlohy (P) a vektor $(y^*, s^*) \in \mathcal{R}^m \times \mathcal{R}^n$ je optimálnym riešením úlohy (D) vtedy a len vtedy, keď vektor (x^*, y^*, s^*) je riešením nasledujúceho systému:*

$$\begin{aligned} A^T y^* + s^* - c &= 0, \\ Ax^* - b &= 0, \\ x_i^* s_i^* &= 0, \ i = 1, 2, ..., n, \\ (x^*, s^*) &> 0. \end{aligned} \tag{2.3}$$

Vzťah $x_i^* s_i^* = 0$, $i = 1, 2, ..., n$, sa nazýva podmienkou komplementarity. Zadefinujeme komplementaritu a ostrú komplementaritu dvojice vektorov:

Definícia 1 *Vektory $a, b \in \mathcal{R}^n$ sa nazývajú komplementárne, ak platí $ab = 0$.*
Vektory $a, b \in \mathcal{R}^n$ sa nazývajú ostrokomplementárne, ak sú komplementárne a navyše platí: $a + b > 0$.

Primárno-duálne optimálne riešenie (x^*, y^*, s^*) nazveme ostrokomplementárnym, ak primárne optimálne riešenie x^* a zložka s^* príslušného duálneho optimálneho riešenia (y^*, s^*) sú ostrokomplementárne vektory.

2.2 O riešení úlohy LP s najmenšou normou

Bude nás zaujímať možný počet riešení úlohy LP. Rozlišujeme tri základné prípady, ktoré môžu nastať:
- 0 optimálnych riešení: neohraničenosť alebo neprípustnosť úlohy. Ide o prípady, ktoré definujú 1. a 2. bod Vety 1
- 1 optimálne riešenie
- viac optimálnych riešení.

Z teórie LP je známe, že množina optimálnych riešení úlohy LP je stenou množiny prípustných riešení. Ak je optimálne riešenie len jedno, ide o stenu najmenšieho rozmeru - jeden bod. Pretože stena množiny prípustných riešení je konvexná množina, s každými dvoma bodmi musí obsahovať celú úsečku, ktorá tieto body spája. Z toho vyplýva, že ak má úloha LP aspoň 2 optimálne riešenia, tak ich musí mať nekonečne veľa.
Práve prípad viacerých optimálnych riešení je zaujímavý. Totiž, ak je riešenie úlohy LP len jedno, potom každá (korektná) metóda musí nájsť práve toto riešenie. Ak je však optimálnych riešení nekonečne veľa, potom je zaujímavé vedieť, aké riešenie nájde konkrétna metóda.

Našu pozornosť teraz upriamime na špeciálne riešenie úlohy LP. Spomedzi všetkých optimálnych riešení chceme vybrať také, ktoré bude mať najmenšiu Euklidovskú normu. Skrátene ho budeme označovať ako LTN-riešenie úlohy LP (označenie je z anglického názvu „Least Two-Norm solution“). Hľadanie takéhoto riešenia už nie je úlohou LP, pretože optimalizujeme nelineárnu účelovú funkciu.

Definícia 2 (LTN-riešenie úlohy LP) *Nech x^* je optimálne riešenie úlohy (P). LTN-riešením úlohy (P) nazývame optimálne riešenie nasledujúcej úlohy:*

$$\min\{\|x\| : Ax = b, c^T x = c^T x^*, x \geq 0\}.$$

Analogicky, nech (y^*, s^*) je optimálne riešenie úlohy (D). LTN-riešením úlohy (D) nazývame optimálne riešenie úlohy:

$$\min\{\|y\| : A^T y + s = c, b^T y = b^T y^*, s \geq 0\}.$$

Z geometrického pohľadu ide o také riešenie, ktoré je spomedzi všetkých optimálnych riešení najbližšie k počiatku súradnicového systému. Prirodzeným zovšeobecnením tohto pojmu je potom také optimálne riešenie, ktoré je najbližšie k ľubovoľnému zadanému bodu. Budeme preň používať označenie GLTN-riešenie („Generalized Least Two-Norm solution“).

Definícia 3 (GLTN-riešenie úlohy LP) *Nech x^* je optimálne riešenie úlohy (P) a nech $a \in \mathcal{R}^n$ je ľubovoľný zadaný vektor. GLTN-riešením úlohy (P) nazývame optimálne riešenie nasledujúcej úlohy:*

$$\min\{\|x - a\| : Ax = b, c^T x = c^T x^*, x \geq 0\}.$$

Analogicky, nech (y^*, s^*) je optimálne riešenie úlohy (D). Pre ľubovoľný zadaný vektor $q \in \mathcal{R}^m$ definujeme GLTN-riešenie úlohy (D) ako optimálne riešenie úlohy:

$$\min\{\|y - q\| : A^T y + s = c, b^T y = b^T y^*, s \geq 0\}.$$

Nasledujúca lema nám ukáže, že GLTN-riešenie primárnej aj duálnej úlohy je jednoznačne určené:

Lema 1 *Nech* (P), (D) *sú riešiteľné úlohy LP a nech* $a \in \mathcal{R}^n$, $q \in \mathcal{R}^m$ *sú ľubovoľné vektory . Potom*
(a) úloha $\min\{\|x - a\| : x \in \mathcal{P}^*\}$ *má práve jedno optimálne riešenie.*
(b) úloha $\min\{\|y - q\| : (y, s) \in \mathcal{D}^*\}$ *má práve jedno optimálne riešenie.*

Dôkaz: (a) Všimnime si, že úloha $\min\{\|x - a\| : x \in \mathcal{P}^*\}$ je ekvivalentná s úlohou $\min\{\sum_{i=1}^n (x_i - a_i)^2 : x \in \mathcal{P}^*\}$. Účelová funkcia $\sum_{i=1}^n (x_i - a_i)^2$ je rýdzokonvexná. Vezmime ľubovoľný bod $x^* \in \mathcal{P}^*$ a definujme

$$\mathcal{P}_{x^*} = \{x \in \mathcal{P}^* : \|x - a\| \leq \|x^* - a\|\}.$$

Ľahko overíme, že množina $\mathcal{P}_{x^*}$ je uzavretá (pretože obsahuje všetky svoje limitné body), neprázdna ($x^* \in \mathcal{P}_{x^*}$), ohraničená a konvexná. Z toho vyplýva, že $\mathcal{P}_{x^*}$ je kompaktná. Spojitá funkcia $\sum_{i=1}^n (x_i - a_i)^2$ nadobúda na kompaktnej množine $\mathcal{P}_{x^*}$ svoje minimum. Z rýdzokonvexnosti funkcie $\sum_{i=1}^n (x_i - a_i)^2$ a z konvexnosti množiny $\mathcal{P}_{x^*}$ zasa vyplýva, že toto minimum je jednoznačné. Z definície množiny $\mathcal{P}_{x^*}$ (obsahuje tie body x z $\mathcal{P}^*$, pre ktoré je výraz $\|x - a\|$ „malý“) vidíme, že jednoznačné minimum funkcie $\sum_{i=1}^n (x_i - a_i)^2$ na $\mathcal{P}_{x^*}$ je zároveň jednoznačným minimom funkcie $\sum_{i=1}^n (x_i - a_i)^2$ na množine $\mathcal{P}^*$.
(b) Úloha $\min\{\|y - q\| : (y, s) \in \mathcal{D}^*\}$ je ekvivalentná s úlohou $\min\{\sum_{i=1}^n (y_i - q_i)^2 : (y, s) \in \mathcal{D}^*\}$. Účelová funkcia $\sum_{i=1}^n (y_i - q_i)^2$ je opäť rýdzokonvexná. Množina $\mathcal{D}^*$ je konvexná, čo nám umožňuje skonštruovať množinu $\mathcal{D}_{y^*} = \{(y, s) \in \mathcal{D}^* : \|y - q\| \leq \|y^* - q\|\}$ s rovnakými vlastnosťami, ako mala množina $\mathcal{P}_{x^*}$. Ďalšia argumentácia a postup dôkazu je rovnaký ako v časti (a). $\square$
Poznámka: Dôkaz sa vzťahuje aj na jednoznačnosť LTN-riešenia, ktoré je špeciálnym prípadom GLTN-riešenia (pre $a = 0$*, resp.* $q = 0$*).*

Krátke vysvetlenie, ktorým chceme zabrániť prípadnému zmäteniu zo striedania pojmov LTN-riešenie a GLTN-riešenie: v tretej kapitole budeme definície a výsledky formulovať pre GLTN-riešenie (s tým, že tieto samozrejme platia aj pre LTN-riešenie ako špeciálny prípad). Len v prípade potreby zdôrazníme niektoré formulácie osobitne aj pre LTN-riešenie. Vo zvyšku tejto kapitoly sa však formulácie pre LTN-riešenie, resp. pre GLTN-riešenie striedajú. Viaceré články k tejto problematike sa totiž zaoberajú len LTN-riešením. Preto sme v týchto častiach textu ponechali formulácie tak, ako boli v originálnych článkoch - t.j. pre LTN-riešenie. Na tieto pasáže čitateľa upozorníme priamo v texte.

GLTN-riešenie úlohy LP môže byť krajným bodom množiny optimálnych riešení, rovnako ako vnútorným bodom tejto množiny. Z toho je zrejmé, že ani simplexová metóda, ani klasické metódy vnútorného bodu vo všeobecnosti nenájdu GLTN-riešenie. Preto je potrebné vyvíjať metódy špeciálne pre úlohu na hľadanie GLTN-riešenia. V ďalšej časti kapitoly sa v krátkosti pozrieme na niektoré z týchto algoritmov.

2.3 Prehľad metód

V tomto prehľade vychádzame z prehľadu v článku [15], ktorý sme doplnili o niektoré novšie výsledky. Upozorňujeme, že prehľad nie je vyčerpávajúci a nemusí zahŕňať všetky v súčasnosti známe metódy a prístupy pre hľadanie LTN-riešenia. Metódy v tomto prehľade budú formulované v tvare, v akom boli formulované v pôvodných článkoch, z ktorých čerpáme (to znamená pre LTN-riešenie. Výnimkou je posledná metóda z podkapitoly o konvexnej reformulácii, ktorá bude uvedená v tvare pre GLTN-riešenie).

2.3.1 Tikhonovova regularizačná metóda

Ide o najstaršiu známu metódu pre hľadanie LTN-riešenia úlohy LP. Je založená na Tikhonovovej regularizácii zle podmienených úloh ([12]). Základnou myšlienkou je riešiť nasledujúcu perturbovanú úlohu:

$$\min\{c^T x + \mu\|x\|^2 : Ax = b, x \geq 0\}, \tag{2.4}$$

kde parameter $\mu > 0$.

Označme $x(\mu) :=$ riešenie úlohy (2.4) pre každé $\mu > 0$. Autori v [12] ukázali, že ak $\mu \to 0$, potom $x(\mu)$ konverguje k LTN-riešeniu úlohy (P). Tento poznatok umožňuje hľadať LTN-riešenie postupným riešením úloh (2.4) pre klesajúce μ. Keďže ale v každom kroku potrebujeme riešiť úlohu kvadratického programovania, takýto postup je časovo dosť náročný.

Neskôr bolo dokázané, že na získanie presného LTN-riešenia úlohy (P) stačí vyriešiť jedinú úlohu kvadratického programovania. Presnejšie: existuje taká hodnota $\bar{\mu} > 0$, že pre všetky $\mu \in (0, \bar{\mu}\rangle$ je už príslušné $x(\mu)$ (riešenie úlohy (2.4)) presným LTN-riešením úlohy (P). Tento objav sa pripisuje autorom Mangasarian a Meyer ([8]). Aj tento prístup má však jeden problém - hodnotu $\bar{\mu}$ nevieme určiť. Preto keď vezmeme nejaké $\mu > 0$ a vyriešime úlohu (2.4), je problém zistiť, či μ bolo dostatočne malé nato, aby sme získali presné LTN-riešenie úlohy (P).

Existujú teda dva základné prístupy založené na Tikhonovovej regularizácii:

1. Riešiť postupnosť úloh (2.4) pre zvolenú klesajúcu postupnosť $\mu_k \to 0$. Pritom jednotlivé úlohy (2.4) nie je potrebné riešiť presne, postačujú riešenia s istou toleranciou v presnosti.

2. Určiť kritérium, podľa ktorého sa dá rozhodnúť, či aktuálna hodnota parametra μ je dostatočne malá nato, aby sme vyriešením úlohy (2.4) dostali presné LTN-riešenie úlohy (P).

Oboma prístupmi sa zaoberalo viacero autorov. My sa im však podrobnejšie venovať nebudeme.

2.3.2 Konvexná reformulácia

Táto skupina zahŕňa prístupy, pri ktorých sa úloha na hľadanie LTN-riešenia úlohy LP transformuje na ekvivalentnú úlohu konvexného programovania bez ohraničení. Transformované úlohy sa potom dajú riešiť prostriedkami voľnej optimalizácie. Priekopníkom v tejto oblasti bol Mangasarian ([7]). V súčasnosti je známych viacero konvexných reformulácií, ktoré sú založené na rôznych myšlienkach. Rozoberieme dve novšie:

C. Kanzow, H. Qi a L. Qi ([5]) Základným prvkom reformulácie je výsledok autorov Smith a Wolkowicz ([11]):

Veta 3 *Vektor x^* je LTN-riešením úlohy (2.1) vtedy a len vtedy, keď existuje číslo $R > 0$ také, že pre každé $r \geq R$ má nelineárny systém*

$$A\left[A^T\lambda_r - rc\right]_+ = b$$

riešenie λ_r^* a platí:

$$x^* = \left[A^T\lambda_r - rc\right]_+$$

kde v_+ označuje vektor so zložkami $(v_+)_i = \max\{v_i, 0\}$.
Veta 3 hovorí, že keď zvolíme r a dopočítame k nemu λ_r, vieme vypočítať aj LTN-riešenie. Uvedená charakterizácia LTN-riešenia však závisí od parametra r. Ten musí byť dostatočne veľký, aby sme zo systému z Vety 3 skutočne dostali LTN-riešenie úlohy (P).
Motivovaní Vetou 3, autori v článku [5] zavádzajú funkciu

$$f(\lambda_r) = \frac{1}{2}\|\left[A^T\lambda_r - rc\right]_+\|^2 - b^T\lambda_r,$$

ktorá (rovnako ako λ_r) závisí od parametra r. Táto funkcia má dôležité vlastnosti: je konvexná a spojite diferencovateľná s gradientom

$$\nabla f(\lambda_r) = A[A^T\lambda_r - rc]_+ - b.$$

Keď položíme $\nabla f(\lambda_r) = 0$, vidíme, že stacionárny bod funkcie $f(\lambda_r)$ je riešením nelineárneho systému z Vety 3. Ak navyše parameter r je dostatočne veľký, potom z Vety 3 vyplýva, že vektor $x^* := \left[A^T\lambda_r - rc\right]_+$ je LTN-riešením úlohy (P) práve vtedy, keď λ_r je stacionárny bod funkcie f. Keďže funkcia f je konvexná, každý jej stacionárny bod je jej minimom. Preto x^* je LTN-riešenie úlohy (P) práve vtedy, keď λ_r je minimom funkcie f. Dostávame tak konvexnú reformuláciu pôvodnej úlohy na hľadanie LTN-riešenia:

$$\min\{f(\lambda_r) : \lambda_r \in \mathcal{R}.\} \tag{2.5}$$

Reformulovaná úloha má pekné vlastnosti: je bez ohraničení a účelová funkcia je spojite diferencovateľná. To nám umožňuje použiť na jej riešenie prostriedky voľnej optimalizácie. Treba ešte zdôrazniť prínos takejto reformulácie. Totiž, LTN-riešenie sme mohli dostať aj rovno zo systému, ktorý je uvedený vo Vete 3. Ten však nebol taký „hladký“, ako je účelová funkcia v reformulovanej úlohe (2.5). Preto sa reformulovaná úloha vo všeobecnosti rieši ľahšie ako uvedený nelineárny systém. Autori v článku [5] podávajú aj algoritmus (založený na Newtonovej metóde), ktorý dokáže úlohu (2.5) efektívne riešiť.
Hoci z doterajšieho popisu vyzerá metóda veľmi nádejne, radosť nám kazí hlavný problém: podobne ako pri Tikhonovovej regularizačnej metóde, aj tu riešenie závisí od voľby parametra, ktorého hodnotu vopred nepoznáme. Nevieme overiť, či parameter r je dostatočne veľký (tak ako vyžaduje predpoklad Vety 3). Autori ([5]) nechávajú otázku určovania parametra r nezodpovedanú.

A. I. Golikov, Yu. G. Evtushenko a N. Mollaverdi ([2])

Ďalšia metóda, ktorá prevádza hľadanie LTN-riešenia úlohy LP na hľadanie voľného extrému konvexnej funkcie. Na rozdiel od predchádzajúcich spomínaných metód, autori v článku ([2]) uvádzajú výsledky a formulujú metódu pre GLTN-riešenie. Ich prístup je založený na použití istej pomocnej funkcie, ktorá je odvodená od Lagrangeovej funkcie k úlohe na hľadanie GLTN-riešenia.

Nech je zadaná riešiteľná úloha (P). Zaveďme označenie:

$$f_* = \min_{x \in \mathcal{P}} c^T x.$$

Autori uvažujú úlohu na hľadanie GLTN-riešenia úlohy (P) v nasledujúcom tvare:

$$\min_{x \in \mathcal{P}^*} \frac{1}{2}\|x - a\|^2, \qquad \mathcal{P}^* = \{x \in \mathcal{R}^n : Ax = b, c^T x = f_*, x \geq 0\}, \tag{2.6}$$

kde $a \in \mathcal{R}^n$ je ľubovoľný zadaný vektor.

Lagrangeova funkcia k úlohe (2.6) má tvar:

$$\mathcal{L}(x, p, \beta, a) = \frac{1}{2}\|x - a\|^2 + p^T(b - Ax) + \beta(c^T x - f_*), \tag{2.7}$$

kde $p \in \mathcal{R}^m$ a $\beta \in \mathcal{R}$ sú Lagrangeove multiplikátory. Z KKT podmienok optimality pre úlohu (2.6) vieme vyjadriť

$$x = (a + A^T p - \beta c)_+. \tag{2.8}$$

V ďalšom kroku autori skonštruovali duálnu úlohu k úlohe (2.6), ktorá má tvar:

$$\max_{p \in \mathcal{R}^m} \max_{\beta \in \mathcal{R}} \min_{x \in \mathcal{R}^n_+} \mathcal{L}(x, p, \beta, a). \tag{2.9}$$

Po dosadení výrazu (2.8), ktorý je riešením vnútornej minimalizačnej úlohy v (2.9), sa dostávame k funkcii

$$\tilde{\mathcal{L}}(p, \beta, a) = b^T p - \frac{1}{2}\|(a + A^T p - \beta c)_+\|^2 - \beta f_* + \frac{1}{2}\|a\|^2. \tag{2.10}$$

Maximalizáciou funkcie $\tilde{\mathcal{L}}(p, \beta, a)$ podľa premenných p, β dostaneme optimálne hodnoty Lagrangeových multiplikátorov (ktoré sú zároveň optimálnymi riešeniami duálnej úlohy k úlohe (2.6)). Z nich už potom nie je problém dopočítať optimálne riešenie úlohy (2.6) z výrazu (2.8).

Funkcia $\tilde{\mathcal{L}}(p, \beta, a)$ však obsahuje číslo f_* - optimálnu hodnotu účelovej funkcie úlohy (P). Túto hodnotu na začiatku výpočtu nepoznáme. Autori preto navrhujú nahradiť funkciu $\tilde{\mathcal{L}}(p, \beta, a)$ nasledujúcou funkciou

$$\mathcal{S}(p, \beta, a) = b^T p - \frac{1}{2}\|(a + A^T p - \beta c)_+\|^2. \tag{2.11}$$

Namiesto maximalizácie funkcie $\tilde{\mathcal{L}}(p, \beta, a)$ budeme teda maximalizovať funkciu $\mathcal{S}(p, \beta, a)$. Dostávame sa k úlohe

$$\max_{p \in \mathcal{R}^m} \mathcal{S}(p, \beta, a). \tag{2.12}$$

Všimnime si, že v transformovanej úlohe (2.12) považujeme β za pevnú konštantu a maximalizujeme už len podľa premennej p.

Autori ukázali, že za istých predpokladov je úloha (2.12) ekvivalentná s maximalizáciou funkcie $\tilde{\mathcal{L}}(p, \beta, a)$ a dovedie nás k GLTN-riešeniu úlohy LP. Jedným z týchto predpokladov je ohraničenie na hodnotu parametra β. Tá nesmie klesnúť pod istú konštantu β^*. Hodnota tejto konštanty závisí od nájdeného optimálneho riešenia, čo opäť predstavuje istý problém, keďže túto hodnotu a priori nepoznáme.

Druhým obmedzujúcim predpokladom ekvivalencie úloh (2.9) a (2.12) je podmienka na submaticu matice A, ktorá prislúcha k nájdenému optimálnemu riešeniu (submatica pozostáva z tých stĺpcov matice A, ktoré zodpovedajú nenulovým zložkám nájdeného optimálneho riešenia). Táto submatica musí mať plnú hodnosť m. Len za takéhoto predpokladu sa dá odvodiť formulka pre β^*. Formálne sú podmienky ekvivalencie úloh zhrnuté v nasledujúcej vete:

Veta 4 ([2], Theorem 1) *Nech množina optimálnych riešení úlohy (P) je neprázdna. Nech matica A_l, pozostávajúca z tých stĺpcov matice A ktoré zodpovedajú nenulovým zložkám vektora $\hat{x}^*$, má hodnosť m. Potom pre ľubovoľné číslo $\beta \geq \beta^*$, projekcia $\hat{x}^*$ bodu $a \in \mathcal{R}^n$ na množinu optimálnych riešení úlohy (P) je daná vzťahom*

$$\hat{x}^* = (a + A^T p(\beta) - \beta c)_+,$$

kde $p(\beta)$ je riešením optimalizačnej úlohy (2.12).

Všimnime si, že nutnou podmienkou pre splnenie predpokladu o hodnosti submatice A_l je, aby nájdené optimálne riešenie malo viac nenulových prvkov, ako je počet ohraničení m.

Metóda, ktorú autori navrhujú, je primárne určená na výpočet ľubovoľného optimálneho riešenia (GLTN-riešenie nám nezaručuje, vzhľadom na silné predpoklady Vety 4). Ide o nasledujúci iteračný proces:

$$p_{k+1} \in \arg\max_{p \in \mathcal{R}^m} \{b^T p - \frac{1}{2}\|(x_k + A^T p - \beta c)_+\|^2\}, \tag{2.13}$$

$$x_{k+1} = (x_k + A^T p_{k+1} - \beta c)_+, \tag{2.14}$$

kde za štartovací bod x_0 môžeme zvoliť ľubovoľný vektor.

Na hľadanie voľného extrému v úlohe (2.13) používajú autori zovšeobecnenú Newtonovu metódu, ktorá je stručne popísaná v článku [2]. Uvedený iteračný proces je dobre definovaný. Jeho výsledkom je vždy presné riešenie úlohy (P), ale tiež presné riešenie duálnej úlohy (D).

Veta 5 ([2], Theorem 3) *Nech množina optimálnych riešení úlohy (P) je neprázdna. Potom pre ľubovoľné $\beta > 0$ a ľubovoľný štartovací bod x_0, iteračný proces (2.13), (2.14) konverguje k bodu x^* - optimálnemu riešeniu úlohy (P) - v konečnom počte krokov ω. Navyše, vzťahom $y^* = \frac{p_{\omega+1}}{\beta}$ je určené optimálne riešenie duálnej úlohy (D).*

Autori odporúčajú používať metódu pre riešenie úloh s veľkým počtom premenných (rádovo až v miliónoch) a stredný počet ohraničení (rádovo v tisíckach). Najzložitejším krokom výpočtu je totiž riešenie úlohy na voľný extrém (2.13). Rozmer tejto úlohy je daný počtom ohraničení m, preto časová náročnosť jednotlivých iterácií s rastúcim m stúpa.

Ako sme spomínali, nájdenie GLTN-riešenia uvedenou iteračnou metódou nemáme vo všeobecnosti zaistené. V numerických testoch, ktoré prevádzali autori, bola nastavená konštantná hodnota $\beta = 1$. Vo všetkých testovaných úlohách sa ukázala ako dostatočná - to znamená, že po nájdení riešenia a dopočítaní hodnoty β^* platilo $\beta > \beta^*$. Samozrejme, nemáme teoretickú záruku že by sme GLTN-riešenie našli v ľubovoľnom prípade. Rovnako ako pri predchádzajúcich spomínaných metódach je tu problém odhadnutia neznámeho parametra (β musí byť dostatočne veľké). Po skončení výpočtu však aspoň vieme ľahko overiť, či nájdené optimálne riešenie úlohy (P) je aj GLTN-riešením.

Kapitola 3

Regularizovaná centrálna trajektória

V tejto kapitole sa budeme zaoberať jednou z (relatívne) novších metód pre hľadanie GLTN-riešenia úlohy LP. Bola predstavená v r. 2002 v článku [15] a ide o prvú metódu svojho druhu. Nevyužíva žiadny z prístupov, ktoré sme spomínali v predchádzajúcej kapitole, t.j. nepotrebuje riešiť pomocné úlohy kvadratického programovania ani úlohy voľnej optimalizácie konvexnej funkcie. Vďaka tomu nie sú potrebné žiadne obmedzujúce predpoklady o pôvodnej úlohe. Dokonca sa nepotrebujeme ani vysporiadať s prípadnou neexistenciou vnútorného bodu úlohy (čo je nevyhnutné pri bežnom použití metód vnútorného bodu v LP).

Metóda je založená na podobnom princípe ako metódy vnútorného bodu v LP. Pri výpočte sleduje istú trajektóriu, ktorá je však odlišná od klasickej centrálnej trajektórie, známej z LP. Túto novú trajektóriu nazývame regularizovanou centrálnou trajektóriou. Na rozdiel od klasickej, regularizovaná centrálna trajektória vždy existuje a líši sa od nej aj v niektorých ďalších vlastnostiach. Najdôležitejšou z týchto vlastností je konvergencia regularizovanej centrálnej trajektórie k GLTN-riešeniu úlohy LP.

Kapitola má nasledujúcu štruktúru: najskôr stručne pripomenieme odvodenie klasickej centrálnej trajektórie a niektoré jej vlastnosti. Potom sa budeme zaoberať odvodením regularizovanej centrálnej trajektórie a formuláciou systému rovníc, ktorými je určená. V tretej časti kapitoly rozoberáme jednotlivé vlastnosti regularizovanej centrálnej trajektórie - existenciu, konvergenciu a analytické vlastnosti.

3.1 Klasická centrálna trajektória

Naďalej uvažujeme primárnu aj duálnu úlohu v rovnicovom tvare (úlohy (P), (D) z Kapitoly 2). K úlohe (P) definujme množinu

$$\mathcal{P}^0 = \{x \in \mathcal{R}^n : Ax = b, x > 0\}.$$

Túto množinu nazývame množinou vnútorných bodov úlohy (P) a jej prvky vnútornými bodmi úlohy (P). Analogicky definujeme množinu vnútorných bodov úlohy (D):

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

Pre parameter $\mu > 0$ definujeme tzv. logaritmickú bariérovú úlohu:

$$\min\{c^T x - \mu \sum_{i=1}^{n} \ln x_i : Ax = b, x > 0\}. \tag{3.1}$$

Centrálna trajektória pozostáva z riešení logaritmických bariérových úloh pre jednotlivé hodnoty parametra $\mu > 0$. Centrálnu trajektóriu možno odvodiť len za základného predpokladu:

Predpoklad 1 *Existuje vnútorný bod pre primárnu aj pre duálnu úlohu, t.j.* $\mathcal{P}^0 \neq \emptyset$, $\mathcal{D}^0 \neq \emptyset$.

Za tohto predpokladu, pre každé $\mu > 0$ existuje práve jedno riešenie bariérovej úlohy (3.1) ([13], resp. [3]). Úlohu (3.1) vieme riešiť použitím Lagrangeovej metódy. Bod minima tejto úlohy je stacionárnym bodom Lagrangeovej funkcie

$$\mathcal{L}_\mu(x, y) = c^T x - \mu \sum_{i=1}^{n} \ln x_i - y^T (Ax - b),$$

kde $y \in \mathcal{R}^m$ je vektor Lagrangeových multiplikátorov.

Stacionárny bod Lagrangeovej funkcie nájdeme tak, že položíme parciálne derivácie funkcie $\mathcal{L}_\mu(x,y)$ rovné nule.

$$\begin{aligned} \frac{\partial \mathcal{L}_\mu(x,y)}{\partial x} &= c \quad -A^T y - \mu x^{-1} = 0, \\ \frac{\partial \mathcal{L}_\mu(x,y)}{\partial y} &= \qquad\quad -Ax + b = 0. \end{aligned} \tag{3.2}$$

Ak položíme $s = \mu x^{-1}$, tak spolu s predchádzajúcimi rovnicami (3.2) po úprave dostaneme:

$$\begin{aligned} A^T y + s - c &= 0, \\ Ax - b &= 0, \\ xs &= \mu e, \\ (x,s) &> 0. \end{aligned} \tag{3.3}$$

Systém (3.3) sa nazýva μ-centrujúci systém. Za Predpokladu 1 máme zaručenú existenciu a jednoznačnosť riešenia μ-centrujúceho systému pre ľubovoľné $\mu > 0$ ([13], resp. [3]).

Veta 6 ([13], Corollary 17.3) *Nech množiny vnútorných bodov primárnej aj duálnej úlohy sú neprázdne. Potom pre ľubovoľné $\mu > 0$ existuje práve jedno riešenie systému (3.3).*

Existencia a jednoznačnosť riešenia μ-centrujúceho systému umožňujú pozerať sa na centrálnu trajektóriu ako na funkciu $\mu \mapsto (x(\mu), y(\mu), s(\mu))$.

Centrálna trajektória bola podrobne analyzovaná a boli ukázané mnohé jej zaujímavé vlastnosti. Jednou z kľúčových je jej konvergencia k optimálnemu riešeniu úloh (P), (D). Navyše sa ukazuje, že centrálna trajektória konverguje k význačnému optimálnemu riešeniu - k tzv. analytickému stredu množiny prípustných riešení. Všeobecnú definíciu analytického stredu množiny možno nájsť napr. v [10]. My ju uvedieme v špeciálnom tvare pre množinu $\mathcal{P}^*$:

Definícia 4 (Analytický stred množiny $\mathcal{P}^*$) *Nech (P) je riešiteľná úloha LP a nech existuje vnútorný bod duálnej úlohy (D). Definujme množinu indexov*

$$B = \sigma(\mathcal{P}^*) = \{i : \exists x \in \mathcal{P}^* : x_i > 0\}.$$

Analytickým stredom množiny optimálnych riešení úlohy (P) nazývame optimálne riešenie úlohy

$$\max_{x>0} \left\{ \prod_{i \in B} x_i \| Ax = b \right\}.$$

Voľne povedané, analytický stred množiny optimálnych riešení je taký bod, ktorý je „čo najďalej“ od hraníc tejto množiny. Vieme ho definovať len pre ohraničené množiny (detailne popísané v [10]). Preto v Definícii 4 vystupuje predpoklad existencie vnútorného bodu duálnej úlohy - implikuje totiž ohraničenosť množiny $\mathcal{P}^*$. Ak analytický stred pre danú množinu vieme definovať, potom je jednoznačný - to znamená, že množina môže mať najviac jeden analytický stred.

Centrálna trajektória existuje len za predpokladu existencie vnútorných bodov oboch úloh (P), (D). Tento predpoklad implikuje ohraničenosť množín $\mathcal{P}^*$ aj $\mathcal{D}^*$ a teda aj existenciu (jednoznačného) analytického stredu pre obe množiny $\mathcal{P}^*$, $\mathcal{D}^*$.

Ako sme už spomínali, centrálna trajektória konverguje k analytickému stredu množiny optimálnych riešení úlohy LP:

Veta 7 ([10], [3]) *Ak primárna aj duálna úloha majú vnútorný bod, potom $x(\mu)$ konverguje k analytickému stredu množiny optimálnych riešení primárnej úlohy a $(y(\mu), s(\mu))$ konverguje k optimálnemu riešeniu duálnej úlohy (pre $\mu \to 0$).*

Ďalšou dôležitou vlastnosťou centrálnej trajektórie je jej analytickosť. Analytickosť funkcie je zaujímavá z viacerých dôvodov. Predovšetkým, analytická funkcia je nekonečne diferencovateľná (t.j. existujú pre ňu derivácie všetkých rádov). Navyše, ak je funkcia analytická v nejakom bode μ_0, potom jej rozvoj do Taylorovho radu bude konvergovať k skutočnej hodnote funkcie (na okolí bodu μ_0). Práve táto vlastnosť má veľký význam pri analýze rýchlosti konvergencie algoritmov, ktoré sú založené na sledovaní trajektórie. Analytickosť centrálnej trajektórie pre $\mu > 0$ sa dokazuje použitím vety o implicitnej funkcii. Náročnejší je dôkaz analytickosti v limitnom bode pre $\mu = 0$ (napr. v [4]):

Veta 8 *Centrálna trajektória (definovaná systémom (3.3)) je analytickou funkciou premennej μ na množine $\mu \in \langle 0, \infty)$.*

Centrálna trajektória je osvedčený prostriedok, ktorý umožňuje efektívne hľadať optimálne riešenie úlohy LP. Je však zrejmé, že nie je vhodná pre hľadanie GLTN-riešenia úlohy LP. Jednoducho by sa dalo ukázať, že GLTN-riešenie úlohy LP môže byť aj krajným bodom množiny optimálnych riešení - v takom prípade už nemôže byť analytickým stredom (ak predpokladáme, že úloha má viacero optimálnych riešení). Ukážeme však, že jednoduchá modifikácia Lagrangeovej funkcie k bariérovej úlohe (3.1) povedie k novej trajektórii, pomocou ktorej je možné hľadať GLTN-riešenie. Tzv. regularizovaná centrálna trajektória navyše existuje pre ľubovoľnú dvojicu primárnej a duálnej úlohy - na rozdiel od klasickej centrálnej trajektórie, ktorá vyžaduje existenciu vnútorných bodov primárnej aj duálnej úlohy.

3.2 Odvodenie regularizovanej centrálnej trajektórie

Pri odvodení klasickej centrálnej trajektórie sme hľadali stacionárny bod Lagrangeovej funkcie

$$\mathcal{L}_\mu(x, y) = c^T x - \mu \sum_{i=1}^{n} \ln x_i - y^T(Ax - b).$$

Regularizácia spočíva v pridaní regularizačného člena k tejto funkcii. Dostávame tak modifikovanú Lagrangeovu funkciu:

$$\phi_{\mu,\theta}(x, y) = c^T x - \mu \sum_{i=1}^{n} \ln x_i - y^T(Ax - b) + \frac{1}{2}\theta(\|x - a\|_2^2 - \|y - q\|_2^2),$$

$$\mu > 0, \theta > 0.$$

Ako vidíme, Lagrangeova funkcia bola rozšírená o dve zložky: $\frac{1}{2}\theta\|x-a\|_2^2$ a $-\frac{1}{2}\theta\|y-q\|_2^2$. Význam prvej z nich je intuitívne zrejmý. Je to penalizačný člen, ktorý „pokutuje" tie riešenia, ktoré sú „ďaleko" od bodu a. Inak povedané, „tlačí" nájdené stacionárne body funkcie $\phi_{\mu,\theta}(x,y)$ k takým riešeniam, ktoré sú v Euklidovskej norme najbližšie k zadanému bodu a.

Zložka $-\frac{1}{2}\theta\|y-q\|_2^2$ je o čosi menej intuitívna - možno by sme ju čakali skôr s kladným znamienkom. Ukazuje sa však, že táto zložka je nevyhnutná - má rozhodujúcu regularizačnú funkciu. Zabezpečuje existenciu regularizovanej centrálnej trajektórie a niektoré jej ďalšie vlastnosti (to bude lepšie vidno pri jednotlivých dôkazoch v ďalšom texte).

Váhu kladenú na regularizačný člen meníme pomocou parametra θ. Základná myšlienka je podobná ako pri klasickej centrálnej trajektórii: keď parametre μ a θ zmenšujeme smerom k nule, nájdené riešenia (stacionárne body funkcie $\phi_{\mu,\theta}(x,y)$) budú čoraz presnejšie aproximovať skutočné optimálne riešenia úloh (P), (D).

Parametre μ a θ sa môžu meniť nezávisle od seba. My však prijmeme zjednodušenie (podľa vzoru článku [15]), ktoré nám umožní zbaviť sa jedného parametra: μ a θ zviažeme vzťahom $\theta = \mu^p$, kde $p \in (0,1)$ je konštanta. Takáto voľba θ zabezpečí teoretické vlastnosti trajektórie, navyše sa ukazuje výhodná pri konštrukcii algoritmov. Odteraz budeme uvažovať Lagrangeovu funkciu s jedným parametrom:

$$\phi_\mu(x,y) = c^T x - \mu \sum_{i=1}^{n} \ln x_i - y^T(Ax-b) + \frac{1}{2}\mu^p(\|x-a\|_2^2 - \|y-q\|_2^2),$$

$$\mu > 0, x > 0, y \in \mathcal{R}^m.$$

Regularizovaná centrálna trajektória pozostáva zo stacionárnych bodov funkcie $\phi_\mu(x,y)$ pre jednotlivé hodnoty $\mu > 0$. Pre nájdenie stacionárneho bodu, položme prvé parciálne derivácie rovné nule:

$$\begin{aligned} \frac{\partial \phi_\mu(x,y)}{\partial x} &= c \quad -A^T y - \mu x^{-1} + \mu^p (x-a) = 0, \\ \frac{\partial \phi_\mu(x,y)}{\partial y} &= \qquad\qquad -Ax + b - \mu^p (y-q) = 0. \end{aligned} \tag{3.4}$$

Označme $s = \mu x^{-1}$, potom po úpravách dostaneme:

$$\begin{aligned} A^T y + s - c &= \mu^p (x-a), \\ Ax - b &= -\mu^p (y-q), \\ xs &= \mu e, \\ (x,s) &> 0. \end{aligned} \tag{3.5}$$

Systém (3.5) definuje stacionárne body funkcie $\phi_\mu(x,y)$. Pre každé $\mu > 0$ označme $(x(\mu), y(\mu), s(\mu))$ riešenie systému (3.5). Funkciu $\mu \mapsto (x(\mu), y(\mu), s(\mu))$, resp. množinu bodov $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$ nazveme regularizovanou centrálnou trajektóriou.

3.3 Vlastnosti regularizovanej centrálnej trajektórie

Tak ako pri klasickej centrálnej trajektórii, aj pri regularizovanej nás budú zaujímať niektoré jej vlastnosti. Tieto nám umožnia konštruovať algoritmy, ktoré umožňujú nájsť optimálne riešenia úloh (P), (D) sledovaním trajektórie. Navyše ukážeme, že regularizovaná centrálna trajektória nás povedie k špeciálnym optimálnym riešeniam týchto úloh (k optimálnym riešeniam, ktoré sú v Euklidovskej norme najbližšie k vopred zvolenému bodu). Jednotlivé vlastnosti budeme študovať v takomto poradí:

1. **Existencia a jednoznačnosť**: základná vlastnosť, ktorá nám umožní študovať regularizovanú centrálnu trajektóriu ako funkciu.

2. **Konvergencia**: limitný bod regularizovanej centrálnej trajektórie predstavuje špeciálne optimálne riešenia úloh (P), (D) - GLTN-riešenia.

3. **Analytickosť**: umožňuje odhadovať rýchlosť konvergencie pri sledovaní regularizovanej centrálnej trajektórie.

Pri analýze uvedených vlastností vychádzame z článku [6], kde sú rovnaké vlastnosti rozoberané pre regularizovanú centrálnu trajektóriu k úlohám semidefinitného programovania.

3.3.1 Existencia

Regularizovanú centrálnu trajektóriu sme definovali ako funkciu, ktorá každému $\mu > 0$ priradí stacionárny bod Lagrangeovej funkcie $\phi_\mu(x, y)$ (a teda riešenie systému (3.5)). Cieľom tejto podkapitoly je ukázať, že príslušné stacionárne body sú jednoznačne určené a regularizovaná centrálna trajektória je tak dobre definovaná.
Upozorňujeme na jeden podstatný rozdiel medzi klasickou a regularizovanou centrálnou trajektóriou: zatiaľčo klasická centrálna trajektória existuje len v prípade existencie vnútorného bodu oboch úloh (P), (D), regularizovaná centrálna trajektória existuje aj v prípade nesplnenia tohto predpokladu.

Veta 9 *Pre každé $\mu > 0$ má funkcia $\phi_\mu(x, y)$ práve jeden stacionárny bod, t.j. systém (3.5) má pre ľubovoľné $\mu > 0$ práve jedno riešenie.*

Dôkaz: Definujme funkciu $g(x) = \sup_{y \in \mathcal{R}^m} \phi_\mu(x, y)$. Keďže pre pevné $x > 0$ je $\phi_\mu(x, y)$ konkávna kvadratická funkcia premennej y, suprémum je jednoznačne určené. V bode, v ktorom sa nadobúda, musí platiť

$$0 = \frac{\partial \phi_\mu(x, y)}{\partial y} = -Ax + b - \mu^p(y - q).$$

Z toho dostávame vzťah pre y:

$$y(x) = \frac{b - Ax}{\mu^p} + q. \tag{3.6}$$

Dosadením (3.6) do funkcie $g(x)$ dostaneme:

$$
\begin{aligned}
g(x) &= \phi_\mu(x, y(x)) \\
&= c^T x - \mu \sum_{i=1}^{n} \ln x_i + \left(\frac{b - Ax}{\mu^p} + q \right)^T (b - Ax) + \frac{1}{2}\mu^p \sum_{i=1}^{n} (x_i - a_i)^2 - \frac{1}{2}\mu^p \left\| \frac{(b - Ax)}{\mu^p} \right\|_2^2 \\
&= c^T x - \mu \sum_{i=1}^{n} \ln x_i + \frac{1}{\mu^p}(b - Ax)^T (b - Ax) + q^T (b - Ax) + \frac{1}{2}\mu^p \sum_{i=1}^{n} (x_i - a_i)^2 \\
&- \frac{1}{2\mu^p}(b - Ax)^T (b - Ax) \\
&= c^T x - \mu \sum_{i=1}^{n} \ln x_i + \frac{1}{2\mu^p}(b - Ax)^T (b - Ax) + q^T (b - Ax) + \frac{1}{2}\mu^p \sum_{i=1}^{n} (x_i - a_i)^2
\end{aligned} \tag{3.7}
$$

Spočítajme teraz gradient a Hessián funkcie $g(x)$ v ľubovoľnom bode $x > 0$:

$$
\begin{aligned}
\nabla g(x) &= c - \mu x^{-1} - \frac{1}{\mu^p} A^T (b - Ax) - A^T q + \mu^p (x - a), \\
\nabla^2 g(x) &= \mu X^{-2} + \frac{1}{\mu^p} A^T A + \mu^p I.
\end{aligned} \tag{3.8}
$$

Všimnime si bližšie Hessovu maticu $\nabla^2 g(x)$. Matice μX^{-2} a $\mu^p I$ sú diagonálne matice s kladnými diagonálami, preto sú obe kladne definitné. Pre ľubovoľnú maticu A je matica $A^T A$ kladne semidefinitná. Napokon, súčet kladne definitnej a kladne semidefinitnej matice dáva kladne definitnú maticu. Z toho je zrejmé, že $\nabla^2 g(x)$ je kladne definitná a preto funkcia $g(x)$ je rýdzokonvexná.

Vezmime ľubovoľný bod $\bar{x} > 0$ a definujme množinu $\bar{P} = \{x > 0 : g(x) \leq g(\bar{x})\}$. Táto množina má nasledujúce vlastnosti:

1. $\bar{P}$ *je neprázdna*: pretože $\bar{x} \in \bar{P}$.

2. $\bar{P}$ *je konvexná*: ľahko vidíme, že pre ľubovoľné body $x, y \in \bar{P}$, celá úsečka xy leží v $\bar{P}$.

3. $\bar{P}$ *je uzavretá*: stačí ukázať, že $\bar{P}$ obsahuje všetky svoje hromadné body. Vezmime ľubovoľnú postupnosť $\{x_k\}_{k=1}^{\infty} \subset \bar{P}$ takú, že $\lim_{k\to\infty} x_k = x$. Z platnosti $g(x_k) \leq g(\bar{x})$

pre $\forall k$ vyplýva $g(x) \leq g(\bar{x})$. Z platnosti $x_k > 0$ vyplýva po limitnom prechode „len“ platnosť $x \geq 0$. Predpokladajme však, že $x_i = 0$ pre nejaké i (i-ta zložka bodu x). Potom by platilo

$$g(x) = \infty = \lim_{k\to\infty} g(x_k). \tag{3.9}$$

Zároveň však, pretože $\{x_k\}_{k=1}^{\infty}$ leží v množine $\bar{P}$, postupnosť $\{g(x_k)\}_{k=1}^{\infty}$ je zhora ohraničená. Tým sa dostávame do sporu s vlastnosťou (3.9). Predpoklad $x_i = 0$ bol nesprávny, pre všetky zložky x musí platiť $x_i > 0$. Pre x - ľubovoľný hromadný bod množiny $\bar{P}$ teda platí: $g(x) \leq g(\bar{x})$ a súčasne $x > 0$, t.j. $x \in \bar{P}$.

4. $\bar{P}$ *je ohraničená*: Najskôr si všimnime limitnú vlastnosť funkcie $g(x)$ pri rastúcej norme argumentu x. O limite pre $\|x\|_2 \to \infty$ budú rozhodovať členy najvyššieho stupňa: $\frac{1}{2}\mu^p \sum_{i=1}^{n}(x_i - a_i)^2$ a $\frac{1}{2\mu^p}(b - Ax)^T(b - Ax)$. Platí:

$$\|x\|_2 \to \infty \;\Rightarrow\; \frac{1}{2}\mu^p \sum_{i=1}^{n}(x_i - a_i)^2 \to \infty,$$

$$\forall x : \frac{1}{2\mu^p}(b - Ax)^T(b - Ax) \geq 0.$$

Z toho už môžeme nahliadnuť platnosť vzťahu:

$$\lim_{\|x\|_2\to\infty} g(x) = \infty. \tag{3.10}$$

Výraz (3.10) prepíšeme podľa definície nevlastnej limity funkcie:

$$\lim_{\|x\|_2\to\infty} g(x) = \infty \;\Leftrightarrow\; \forall K \in \mathcal{R}\; \exists c \in \mathcal{R}\; \forall x; \|x\|_2 > c : g(x) > K. \tag{3.11}$$

Pre ľubovoľné číslo $k \in \mathcal{R}$ definujme množinu $P_k = \{x > 0 : g(x) \leq k\}$. O tejto množine chceme ukázať, že je ohraničená.

Sporom: predpokladajme, že by P_k bola neohraničená. To by znamenalo, že nech vezmeme ľubovoľné číslo $\gamma \in \mathcal{R}$, vždy bude k nemu existovať x_γ také, že:

$$x_\gamma \in P_k\;,\; \|x_\gamma\|_2 > \gamma. \tag{3.12}$$

Takže: máme dané číslo $k \in \mathcal{R}$, ktoré určuje množinu P_k. Podľa (3.11), k nášmu číslu k existuje taká konštanta $c_k \in \mathcal{R}$, že pre všetky x platí:

$$\|x\|_2 > c_k \;\Rightarrow\; \|g(x)\|_2 > k. \tag{3.13}$$

Povedané voľnejšie, ak s normou $\|x\|_2$ prekročíme istú hranicu c_k, funkčná hodnota $g(x)$ nám (v bode x) presiahne hranicu k.

Podľa (3.12) máme zaručenú existenciu bodu x_{c_k} so zaujímavými vlastnosťami:

$$x_{c_k} \in P_k,$$

$$\|x_{c_k}\|_2 > c_k.$$

Lenže podľa (3.13), z platnosti $\|x_{c_k}\|_2 > c_k$ vyplýva $\|g(x_{c_k})\|_2 > k$. To je spor s tým, že $x_{c_k} \in P_k$. Predpoklad o neohraničenosti P_k bol nesprávny - množina P_k musí byť ohraničená.

Napokon, podľa definície množiny $\bar{P}$ platí:

$$\bar{P} = P_{g(\bar{x})},$$

a $g(\bar{x}) \in \mathcal{R}$. Množina $\bar{P}$ je ohraničená.

Z vlastností $1., 3.$ a 4. vyplýva, že $\bar{P}$ je kompaktná množina. Dobre známy výsledok z matematickej analýzy (Weierstrassova veta) hovorí, že spojitá funkcia na kompaktnej množine nadobúda maximum aj minimum. Navyše, rýdzokonvexná funkcia $g(x)$ môže na konvexnej množine $\bar{P}$ nadobúdať najviac jedno minimum. Teda $g(x)$ nadobúda na $\bar{P}$ práve jedno minimum.
A keďže $\bar{P} \subset P = \{x > 0\}$, pričom funkcia $g(x)$ nadobúda svoje „najmenšie hodnoty“ práve na množine $\bar{P}$ (stačí si všimnúť definíciu $\bar{P}$), tak $g(x)$ nadobúda minimum na množine P, a toto minimum je jednoznačné.

Označme x^* bod minima funkcie $g(x)$ na $P = \{x > 0\}$. Musí ísť o stacionárny bod funkcie $g(x)$, preto platí:

$$0 = \nabla g(x^*) = c - \mu x^{*-1} - \frac{1}{\mu^p} A^T(b - Ax^*) - A^T q + \mu^p (x^* - a). \tag{3.14}$$

Definujme bod

$$y^* := \frac{b - Ax^*}{\mu^p} + q. \tag{3.15}$$

Dosadením (3.15) do (3.14) dostaneme:

$$c - \mu x^{*-1} - A^T y^* + \mu^p (x^* - a) = 0. \tag{3.16}$$

Označme ešte bod

$$s^* := \mu x^{*-1}. \tag{3.17}$$

Spojením vzťahov (3.15), (3.16) a (3.17) s použitím jednoduchých úprav dostaneme:

$$\begin{aligned} A^T y^* + s^* - c &= \mu^p (x^* - a), \\ Ax^* - b &= -\mu^p (y^* - q), \\ x^* s^* &= \mu e, \\ (x^*, s^*) &> 0. \end{aligned} \tag{3.18}$$

Takže bod (x^*, y^*, s^*) vyhovuje systému (3.5). Tým sme ukázali, že systém (3.5) má aspoň jedno riešenie.

Ostáva ukázať jednoznačnosť nájdeného riešenia. Nech má systém (3.5) ľubovoľné dve riešenia, označme ich (x_1, y_1, s_1) a (x_2, y_2, s_2). Pre bod (x_1, y_1, s_1) platí:

$$A^T y_1 + s_1 - c = \mu^p (x_1 - a), \tag{3.19}$$

$$Ax_1 - b = -\mu^p (y_1 - q), \tag{3.20}$$

$$x_1 s_1 = \mu e. \tag{3.21}$$

Z (3.20), resp. (3.21) vyjadrime:

$$y_1 = \frac{b - Ax_1}{\mu^p} + q, \tag{3.22}$$

$$s_1 = \mu(x_1)^{-1}. \tag{3.23}$$

Dosadením (3.22), (3.23) do vzťahu (3.19) dostávame:

$$A^T \left(\frac{b - Ax_1}{\mu^p} + q \right) + \mu(x_1)^{-1} - c = \mu^p(x_1 - a)$$

a po úprave

$$c - \mu(x_1)^{-1} - \frac{1}{\mu^p} A^T(b - Ax_1) - A^T q + \mu^p(x_1 - a) = 0.$$

To ale znamená, že $\nabla g(x_1) = 0$ (ľahko vidíme napr. z (3.14)), a teda x_1 je stacionárnym bodom funkcie $g(x)$. Pripomeňme, že (x_1, y_1, s_1) je riešením systému (3.5) - preto platí $x_1 > 0$ a x_1 je tak z definičného oboru funkcie $g(x)$. Napokon, keďže pre $\forall x$ je $\nabla^2 g(x) \succ 0$, x_1 musí byť lokálnym minimom funkcie $g(x)$. Úplne analogickým postupom by sme ukázali, že aj bod x_2 je lokálnym minimom $g(x)$.

Pretože $g(x)$ je rýdzokonvexná funkcia, nemôže nadobúdať lokálne minimá v dvoch rôznych bodoch. Z toho nevyhnutne vyplýva $x_1 = x_2$. Z vyjadrení (3.20), (3.21) (a analogických pre bod (x_2, y_2, s_2)) potom dostaneme rovnosť $(x_1, y_1, s_1) = (x_2, y_2, s_2)$. □

3.3.2 Konvergencia

Jednou z kľúčových vlastností centrálnej trajektórie je jej konvergencia k optimálnemu riešeniu úlohy LP. Konvergencia klasickej centrálnej trajektórie je dobre známym výsledkom teórie metód vnútorného bodu. Bolo ukázané, že táto trajektória konverguje k špeciálnemu bodu - k analytickému stredu množiny prípustných riešení. Presnú definíciu analytického stredu množiny $\mathcal{P}^*$ sme uviedli v časti 3.1. Zjednodušene, analytický stred je taký bod danej množiny, ktorý je „najviac vzdialený“ od hraníc tejto množiny. Množina prípustných riešení má práve jeden analytický stred.

Ukazuje sa, že aj regularizovaná centrálna trajektória konverguje k istému význačnému bodu množiny optimálnych riešení úlohy LP. Je to taký bod množiny $\mathcal{P}^*$, ktorý je spomedzi všetkých optimálnych riešení najbližšie (v Euklidovskej norme) k zadanému bodu (GLTN-riešenie). K takémuto riešeniu konverguje primárna aj duálna regularizovaná centrálna trajektória.

Ešte raz pripomeňme, že existenciu regularizovanej centrálnej trajektórie sme dokázali pre akúkoľvek dvojicu úloh (P), (D). Pri konvergencii to už možné nebude - potrebujeme predpoklad riešiteľnosti úloh (P), (D). Uvedieme ho preto hneď na úvod s tým, že sa vzťahuje na všetky vety a lemy v tejto podkapitole:

Predpoklad 2 *Existuje optimálne riešenie pre primárnu aj pre duálnu úlohu, t.j.* $\mathcal{P}^* \neq \emptyset, \mathcal{D}^* \neq \emptyset$.

Aby nedošlo k nedorozumeniu, upresnime čo budeme rozumieť pod primárno-duálnym GLTN-riešením úloh (P), (D):

x^* := GLTN-riešenie úlohy (P) (prislúchajúce k bodu a),

y^* := GLTN-riešenie úlohy (D) (prislúchajúce k bodu q),

s^* := duálny doplnok prislúchajúci k y^*, t.j. $(y^*, s^*) \in \mathcal{D}$. Primárno-duálnym GLTN-riešením potom rozumieme bod (x^*, y^*, s^*).

Základným stavebným prvkom dôkazu existencie limitného bodu regularizovanej centrálnej trajektórie je (podobne ako pri klasickej centrálnej trajektórii) ohraničenosť množiny bodov $\{(x(\mu), y(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\}$. K odvodeniu ohraničenosti potrebujeme nasledujúci technický výsledok:

Lema 2 *Pre ľubovoľný bod* $(x(\mu), y(\mu), s(\mu))$ *na regularizovanej centrálnej trajektórii a pre ľubovoľné primárno-duálne optimálne riešenie* (x^*, y^*, s^*) *platí:*

$$\|y(\mu) - q\| \left(\|y(\mu) - q\| - \|y^* - q\|\right) + \|x(\mu) - a\| \left(\|x(\mu) - a\| - \|x^* - a\|\right) \leq n\mu^{1-p}.$$

Dôkaz: Budeme odhadovať výraz $(s(\mu) - s^*)^T(x(\mu) - x^*)$. Najskôr nájdeme horný odhad:

$$(s(\mu) - s^*)^T(x(\mu) - x^*) = s(\mu)^T x(\mu) - (s^*)^T x(\mu) - s(\mu)^T x^* + (s^*)^T x^*.$$

Bod (x^*, s^*) je optimálnym riešením úloh (P), (D), preto duálna medzera $(s^*)^T x^* = 0$. Pre duálnu medzeru pozdĺž regularizovanej centrálnej trajektórie platí $s(\mu)^T x(\mu) = n\mu$. Napokon, výrazy $(s^*)^T x(\mu)$ a $s(\mu)^T x^*$ sú nezáporné (keďže $x(\mu), s(\mu), x^*, s^* \geq 0$). Vďaka tomu dostávame odhad:

$$(s(\mu) - s^*)^T (x(\mu) - x^*) \ \leq n\mu. \tag{3.24}$$

Dolný odhad dostaneme nasledovne:

$$\begin{aligned}
&(s(\mu) - s^*)^T (x(\mu) - x^*) = \\
&= (c - A^T y(\mu) + \mu^p (x(\mu) - a) - c + A^T y^*)^T (x(\mu) - x^*) \\
&= (-A^T y(\mu) + A^T y^* + \mu^p (x(\mu) - a))^T (x(\mu) - x^*) \\
&= (A^T (y^* - y(\mu)))^T (x(\mu) - x^*) + \mu^p (x(\mu) - a)^T (x(\mu) - x^*) \\
&= -(A(x(\mu) - x^*))^T (y(\mu) - y^*) + \mu^p (x(\mu) - a)^T (x(\mu) - x^*) \\
&= -(Ax(\mu) - b)^T (y(\mu) - y^*) + \mu^p (x(\mu) - a)^T (x(\mu) - x^*) \\
&= \mu^p (y(\mu) - q)^T (y(\mu) - y^*) + \mu^p (x(\mu) - a)^T (x(\mu) - x^*) \\
&= \mu^p (y(\mu) - q)^T (y(\mu) - q + q - y^*) + \mu^p (x(\mu) - a)^T (x(\mu) - a + a - x^*) \\
&= \mu^p (y(\mu) - q)^T (y(\mu) - q) - \mu^p (y(\mu) - q)^T (y^* - q) \\
&\quad + \mu^p (x(\mu) - a)^T (x(\mu) - a) - \mu^p (x(\mu) - a)^T (x^* - a).
\end{aligned} \tag{3.25}$$

Teraz rozpíšeme výsledný výraz v (3.25). Prvú a tretiu zložku vieme vyjadriť (s využitím definície skalárneho súčinu) takto:

$$\begin{aligned}
\mu^p (y(\mu) - q)^T (y(\mu) - q) &= \mu^p \, \|y(\mu) - q\|^2 \\
\mu^p (x(\mu) - a)^T (x(\mu) - a) &= \mu^p \, \|x(\mu) - a\|^2
\end{aligned} \tag{3.26}$$

Na druhú a štvrtú zložku využijeme Cauchy-Schwarz-Bunyakovského nerovnosť:

$$\begin{aligned}
\mu^p (y(\mu) - q)^T (y^* - q) &\leq \mu^p \, \|y(\mu) - q\| \, \|y^* - q\| \\
\mu^p (x(\mu) - a)^T (x^* - a) &\leq \mu^p \, \|x(\mu) - a\| \, \|x^* - a\|
\end{aligned} \tag{3.27}$$

Keď dosadíme (3.26), (3.27) do posledného vzťahu v (3.25), dostaneme hľadaný dolný odhad:

$$\begin{aligned}(s(\mu) - s^*)^T (x(\mu) - x^*) \ \geq& \\ \mu^p(\|y(\mu) - q\|^2 - \|y(\mu) - q\|\,\|y^* - q\|) &+ \mu^p(\|x(\mu) - a\|^2 - \|x(\mu) - a\|\,\|x^* - a\|).\end{aligned} \tag{3.28}$$

Nakoniec porovnáme odhady (3.24) a (3.28):

$$\begin{aligned}\mu^p(\|y(\mu) - q\|^2 - \|y(\mu) - q\|\,\|y^* - q\| + \|x(\mu) - a\|^2 - \|x(\mu) - a\|\,\|x^* - a\|) &\leq n\mu \\ (\|y(\mu) - q\|^2 - \|y(\mu) - q\|\,\|y^* - q\|) + (\|x(\mu) - a\|^2 - \|x(\mu) - a\|\,\|x^* - a\|) &\leq n\mu^{1-p} \\ \|y(\mu) - q\|\,(\|y(\mu) - q\| - \|y^* - q\|) + \|x(\mu) - a\|\,(\|x(\mu) - a\| - \|x^* - a\|) &\leq n\mu^{1-p},\end{aligned} \tag{3.29}$$

čím je dôkaz dokončený. □

Ak chceme dokázať konvergenciu, potrebujeme v prvom rade ukázať existenciu nejakého limitného bodu regularizovanej centrálnej trajektórie. Tú nám zaručí nasledujúca lema o ohraničenosti a jej dôsledok.

Lema 3 *Pre ľubovoľné číslo* $\bar{\mu} > 0$*, množina* $\{(x(\mu), y(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\}$ *je ohraničená.*

Dôkaz: Podľa dokázaného vzťahu z Lemy 2, pre $\mu \in (0, \bar{\mu}\rangle$ platí takýto odhad:

$$\|y(\mu) - q\|\,(\|y(\mu) - q\| - \|y^* - q\|) + \|x(\mu) - a\|\,(\|x(\mu) - a\| - \|x^* - a\|) \leq n\mu^{1-p}. \tag{3.30}$$

Dokážeme najskôr ohraničenosť množiny $\{\|y(\mu) - q\| : \mu \in (0, \bar{\mu}\rangle\}$. Sporom: keby táto množina bola neohraničená, musela by existovať postupnosť $\{\mu_k\}_{k=1}^{\infty} \subset (0, \bar{\mu}\rangle$ taká, že $\|y(\mu_k) - q\| \to \infty$ pre $k \to \infty$. Keďže výraz $\|y(\mu) - q\|\,(\|y(\mu) - q\| - \|y^* - q\|)$ je konvexnou kvadratickou funkciou premennej $\|y(\mu) - q\|$, platilo by:

$$\|y(\mu_k) - q\|\,(\|y(\mu_k) - q\| - \|y^* - q\|) \to \infty \,(\textit{pre}\, k \to \infty). \tag{3.31}$$

Výraz $\|x(\mu) - a\| \, (\|x(\mu) - a\| - \|x^* - a\|)$ je konvexnou kvadratickou funkciou premennej $\|x(\mu) - a\|$, preto musí nadobúdať konečné minimum (na svojom definičnom obore). Označme $M := \min\{\|x(\mu) - a\| \, (\|x(\mu) - a\| - \|x^* - a\|) : \|x(\mu) - a\| \geq 0\}$. Z (3.30) potom dostaneme:

$$\|y(\mu_k) - q\| \, (\|y(\mu_k) - q\| - \|y^* - q\|) \leq n\bar{\mu}^{1-p} - M. \tag{3.32}$$

Podľa (3.31) však ľavá strana v (3.32) pre $k \to \infty$ neobmedzene rastie, čo je spor. Preto množina $\{\|y(\mu) - q\| : \mu \in (0, \bar{\mu}\rangle \, \|\}$ musí byť ohraničená. Analogicky by sme ukázali aj ohraničenosť množiny $\{\|x(\mu) - a\| : \mu \in (0, \bar{\mu}\rangle \, \|\}$.

Z ohraničenosti množiny $\{\|y(\mu) - q\| : 0 < \mu \leq \bar{\mu}\}$ vyplýva $\|y(\mu) - q\| < r$ (pre nejaké kladné r). Potom môžeme s využitím trojuholníkovej nerovnosti ohraničiť aj $\|y(\mu)\|$:

$$\begin{aligned} &\|y(\mu)\| = \|(y(\mu) - q) + q\| \leq \|(y(\mu) - q\| + \|q\|, \\ &\|y(\mu)\| < r + \|q\|. \end{aligned} \tag{3.33}$$

To znamená, že $\{\|y(\mu)\| : 0 < \mu \leq \bar{\mu}\}$ je ohraničená. Analogicky by sme ukázali aj ohraničenosť množiny $\{\|x(\mu)\| : 0 < \mu \leq \bar{\mu}\}$. Napokon, ohraničenosť množiny $\{\|s(\mu)\| : 0 < \mu \leq \bar{\mu}\}$ vyplýva zo vzťahu pre s zo systému (3.5):

$$s(\mu) = \mu^p(x(\mu) - a) - A^T y(\mu) + c$$

a z ohraničenosti $\|x(\mu) - a)\|$ a $\|y(\mu)\|$. $\square$

Dôsledok 1 *Regularizovaná centrálna trajektória má aspoň jeden limitný bod pre $\mu \to 0$.*

Dôkaz: Vezmime ľubovoľnú klesajúcu postupnosť $\{\mu_k\}_{k=1}^{\infty}$ takú, že $\mu_k \to 0$ pre $k \to \infty$. Podľa Lemy 3 je potom postupnosť $\{(x(\mu_k), y(\mu_k), s(\mu_k))\}_{k=1}^{\infty}$ ohraničená. Každá ohraničená postupnosť obsahuje konvergentnú podpostupnosť. Limita tejto podpostupnosti je zároveň limitným bodom regularizovanej centrálnej trajektórie pre $\mu \to 0$. $\square$

Existenciu limitného bodu máme teda zaručenú. Teraz si všimnime dôležitú vlastnosť limitných bodov:

Lema 4 *Každý limitný bod regularizovanej centrálnej trajektórie je primárno-duálnym optimálnym riešením úloh* (P), (D).

Dôkaz: Stačí si všimnúť, ako sa správa systém (3.5) pre $\mu \to 0$. Po limitnom prechode dostaneme, že limitný bod vyhovuje systému (2.3).
Tento systém predstavuje dobre známe tzv. KKT-podmienky pre úlohy (P), (D). Ako sme spomínali, tieto podmienky sú nutnými aj postačujúcimi podmienkami optimality pre úlohu LP. Keďže limitný bod regularizovanej centrálnej trajektórie tieto podmienky spĺňa, je primárno-duálnym optimálnym riešením pre úlohy (P), (D). $\square$

V nasledujúcej vete ukážeme hlavný výsledok tejto podkapitoly:

Veta 10 *Ak* $(\bar{x}, \bar{y}, \bar{s})$ *je limitným bodom regularizovanej centrálnej trajektórie, potom* $\bar{x}$ *je GLTN-riešením úlohy* (P) *a* $\bar{y}$ *je GLTN-riešením úlohy* (D).

Dôkaz: Už sme ukázali, že limitný bod trajektórie je optimálnym primárno-duálnym riešením. Stačí teda ukázať, že ide o GLTN-riešenie. Označme primárno-duálne GLTN-riešenie k úlohám (P), (D) ako (x^*, y^*, s^*). Kľúčovým pre nás bude vzťah ukázaný v Leme 2. Limitným prechodom pre $\mu \to 0$ z tohto vzťahu dostaneme:

$$\|\bar{y} - q\| \left(\|\bar{y} - q\| - \|y^* - q\|\right) + \|\bar{x} - a\| \left(\|\bar{x} - a\| - \|x^* - a\|\right) \leq 0. \tag{3.34}$$

Všimnime si, že nerovnosť (3.34) je splnená vtedy a len vtedy, keď budú platiť rovnosti:

$$\|\bar{y} - q\| \left(\|\bar{y} - q\| - \|y^* - q\|\right) = 0, \tag{3.35}$$

$$\|\bar{x} - a\| \left(\|\bar{x} - a\| - \|x^* - a\|\right) = 0. \tag{3.36}$$

Dokážeme prvú rovnosť. Môžu nastať dva prípady:

1. $\|\bar{y} - q\| = 0$: z toho hneď dostávame, že $\bar{y} = q$. Žiadny iný bod nemôže byť bližšie k bodu q ako samotný bod q. Preto triviálne platí $\bar{y} = y^*$.

2. $\|\bar{y}-q\| > 0$: vtedy jediná možnosť ako splniť vzťah (3.35) je položiť $\|\bar{y}-q\| = \|y^*-q\|$. Podľa Lemy 1 je GLTN-riešenie úlohy LP jednoznačné. Preto:

$$\|\bar{y} - q\| = \|y^* - q\| \Rightarrow \bar{y} = y^*.$$

Takže $\bar{y}$ je GLTN-riešením úlohy (D). Zo vzťahu (3.36) je možné úplne analogicky ukázať, že $\bar{x}$ je GLTN-riešením úlohy (P). $\square$

Napokon si všimnime, že jednoznačnosť GLTN-riešenia úlohy LP spolu s existenciou limitného bodu regularizovanej centrálnej trajektórie (Dôsledok 1) a práve dokázanou Vetou 10 zaručujú konvergenciu regularizovanej centrálnej trajektórie.

Dôsledok 2 *Regularizovaná centrálna trajektória $(x(\mu), y(\mu), s(\mu))$ konverguje (t.j. má práve jeden limitný bod) pre $\mu \to 0$.*

3.3.3 Analytické vlastnosti

V tejto časti ukážeme, že regularizovaná centrálna trajektória je analytickou funkciou premennej μ.
Analytickosť regularizovanej centrálnej trajektórie dokážeme len pre $\mu > 0$. Dôkaz vychádza z vety o implicitnej funkcii, presnejšie z analytickej verzie tejto vety:

Veta 11 ([1]) *Nech $F(f, \mu) : \mathcal{R}^n \times \mathcal{R} \to \mathcal{R}^n$ a nech $F(f_0, \mu_0) = 0$. Ďalej, nech $F(f, \mu)$ je analytická funkcia na $U \times I$ okolí bodu (f_0, μ_0). Nech Jacobyho matica $\frac{\partial F}{\partial f}(f_0, \mu_0)$ je regulárna. Potom existuje také okolie $\mathcal{O}(\mu_0) \subset I$ bodu μ_0 a taká analytická funkcia $f(\mu) : \mathcal{O}(\mu_0) \to U$, že $f(\mu_0) = f_0$ a pre každé $\mu \in \mathcal{O}(\mu_0)$ platí $F(f(\mu), \mu) = 0$.*

Takto vyzerá tvrdenie, ktoré zhŕňa analytickosť regularizovanej centrálnej trajektórie:

Veta 12 *Regularizovaná centrálna trajektória je analytickou funkciou na $\mu \in (0, \infty)$.*

Dôkaz: Skúsme aplikovať Vetu 11 na regularizovanú centrálnu trajektóriu. Funkciu F dostaneme tak, že prepíšeme systém (3.5):

$$F(x(\mu), y(\mu), s(\mu), \mu) = \begin{pmatrix} A^T y + s - c - \mu^p(x-a) \\ Ax - b + \mu^p(y-q) \\ xs - \mu e \end{pmatrix} \tag{3.37}$$

V zmysle označenia z Vety 11 máme $f = (x, y, s)$. Všimnime si, že $F : \mathcal{R}^{2n+m} \times \mathcal{R} \to \mathcal{R}^{2n+m}(x, s \in \mathcal{R}^n, y \in \mathcal{R}^m)$.

Zvoľme teraz ľubovoľné $\mu_0 > 0$. Z Vety 9 vyplýva, že existuje práve jedno riešenie systému (3.5), označme ho $f_0 = (x(\mu_0), y(\mu_0), s(\mu_0))$. V tomto bode zjavne platí $F(x(\mu_0), y(\mu_0), s(\mu_0), \mu_0) = 0$. Ďalší predpoklad Vety 11 - analytickosť funkcie $F(f, \mu)$ na nejakom okolí bodu (f_0, μ_0) - je splnený. $F(f, \mu)$ je totiž analytickou funkciou všetkých svojich premenných x, y, s, μ (ľahko vidíme z (3.37)).

Ostáva overiť posledný predpoklad Vety 11: regularitu Jakobiánu $\frac{\partial F}{\partial f}$ v bode (f_0, μ_0). Deriváciou F podľa premenných x, y, s dostaneme:

$$\frac{\partial F}{\partial(x, y, s)} = \begin{pmatrix} -\mu^p I & A^T & I \\ A & \mu^p I & 0 \\ S & 0 & X \end{pmatrix} \tag{3.38}$$

Regularitu ľubovoľnej matice M môžeme popísať takto: M je regulárna vtedy a len vtedy, keď pre ľubovoľný vektor h (príslušného rozmeru) platí:

$$Mh = 0 \;\Rightarrow\; h = 0, \tag{3.39}$$

t.j. homogénny systém rovníc určený maticou M má jediné riešenie, ktorým je 0-vektor. Označme $h := (\Delta x, \Delta y, \Delta s)^T$. Našim cieľom bude vyšetriť riešenia systému

$$\begin{pmatrix} -\mu^p I & A^T & I \\ A & \mu^p I & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = 0.$$

Po rozpísaní rovníc dostaneme

$$\begin{array}{llll} -\mu^p \Delta x & +A^T \Delta y & +\Delta s & = 0, \\ A\Delta x & +\mu^p \Delta y & & = 0, \\ S\Delta x & & +X\Delta s & = 0. \end{array} \tag{3.40}$$

Teraz vyjadríme Δy z druhého vzťahu v (3.40):

$$\Delta y = -\frac{1}{\mu^p} A\Delta x, \tag{3.41}$$

a z tretieho vzťahu v (3.40) zasa Δs:

$$\Delta s = -X^{-1} S\Delta x. \tag{3.42}$$

Dosadením vyjadrení (3.41) a (3.42) do prvého vzťahu v (3.40) máme:

$$(-\mu^p I - \frac{1}{\mu^p} A^T A - X^{-1} S)\Delta x = 0. \tag{3.43}$$

Označme $P := -\mu^p I - \frac{1}{\mu^p} A^T A - X^{-1} S$ a všimnime si matice, ktorých súčet vytvára P:
Matica $-\mu^p I$ je diagonálna, pričom všetky jej diagonálne prvky sú záporné. Z toho vyplýva, že je záporne definitná.
Z teórie definitnosti matíc je známe, že pre ľubovoľnú maticu C je matica $C^T C$ kladne semidefinitná. Preto $A^T A$ je kladne definitná matica. Prenásobením zápornou konštantou dostaneme zápornú semidefinitnosť matice $-\frac{1}{\mu^p} A^T A$.
Napokon, X^{-1} a S sú diagonálne matice s kladnými diagonálami. Ich súčin je tiež diagonálna matica s kladnou diagonálou (a teda kladne definitná matica). Preto matica $-X^{-1}S$ je záporne definitná.
Súčet dvoch záporne definitných matíc je záporne definitná matica. Súčet záporne definitnej a záporne semidefinitnej matice je takisto záporne definitná matica. Matica P je tvorená súčtom dvoch záporne definitných a jednej záporne semidefinitnej matice , preto je záporne definitná. Zo zápornej definitnosti zas vyplýva, že je regulárna.
Vráťme sa k vzťahu (3.43), podľa ktorého platí $P\Delta x = 0$. Z tohto tvrdenia a z regularity matice P vyplýva $\Delta x = 0$. Z (3.41) a (3.42) jasne vidíme, že aj $\Delta y = 0$ a $\Delta s = 0$. Tým

sme ukázali, že systém (3.40) má jediné riešenie: vektor $(\Delta x, \Delta y, \Delta s) = 0$. Podľa (3.39) to znamená, že Jakobián $\frac{\partial F}{\partial(x,y,s)}$ v bode (f_0, μ_0) je regulárny.

Overili sme teda všetky predpoklady z Vety 11. Vďaka tomu máme zaručenú existenciu analytickej funkcie $f(\mu)$ na nejakom okolí $\mathcal{O}(\mu_0)$ bodu μ_0. Vieme tiež, že táto analytická funkcia sa v bode μ_0 zhoduje s regularizovanou centrálnou trajektóriou, pretože bod $f_0 = (x(\mu_0), y(\mu_0), s(\mu_0)$ sme definovali ako riešenie systému (3.5).
Potrebujeme ešte ukázať, že analytická funkcia $f(\mu)$ sa zhoduje s regularizovanou centrálnou trajektóriou na celom okolí $\mathcal{O}(\mu_0)$. V prvom rade musíme obmedziť okolie bodu μ_0 na také okolie I_1 (pozri predpoklady Vety 11), aby platilo

$$\forall \mu \in I_1 : \mu > 0.$$

Z Vety 11 vieme, že na okolí $\mathcal{O}(\mu_0) \subset I_1$ platí $F(f(\mu), \mu) = 0$. Po prepísaní z toho dostávame platnosť prvých troch vzťahov systému (3.5). Nato aby sme zaručili aj posledný vzťah zo systému (3.5) - nezápornosť premenných x, s - potrebujeme presnejšie definovať okolie U bodu f_0 z predpokladov Vety 11.
Pre bod $x(\mu_0), y(\mu_0), s(\mu_0)$ (ako riešenie systému (3.5)) platilo

$$x(\mu_0) > 0, s(\mu_0) > 0.$$

Existuje teda také okolie U_1 bodu $(x(\mu_0), y(\mu_0), s(\mu_0))$, v ktorom bude splnené

$$\forall (x, y, s) \in U_1 : x > 0, s > 0.$$

Vo Vete 11 teda zvolíme U_1 ako okolie bodu f_0. To nám zaručí, že nájdená analytická funkcia f bude vracať hodnoty x a s kladné. Vďaka tomu body $f(\mu) = (x(\mu), y(\mu), s(\mu))$ budú vyhovovať všetkým vzťahom zo systému (3.5), čo znamená že funkcia f nie je nič iné ako naša regularizovaná centrálna trajektória.
Analytickosť regularizovanej centrálnej trajektórie sme tak ukázali v nejakom okolí zvoleného bodu $\mu_0 > 0$. Aplikáciou uvedeného dôkazu na všetky body $\mu_0 > 0$ dostávame

hľadaný výsledok. □

Ešte pár slov k analytickosti regularizovanej centrálnej trajektórie v limitnom bode $\mu = 0$. V článku [6] o regularizovanej centrálnej trajektórii k úlohám semidefinitného programovania je dokazovaná aj analytickosť v bode $\mu = 0$. Tento dôkaz je však spravený za silného predpokladu: nájdené GLTN-riešenie musí byť ostrokomplementárne.

Z teórie LP je dobre známe, že ak úloha LP má optimálne riešenie, potom existuje aj ostrokomplementárne optimálne riešenie. Z teórie metód vnútorného bodu je zasa známe, že limitným bodom klasickej centrálnej trajektórie je vždy ostrokomplementárne optimálne riešenie ([3], [10]). Na tomto poznatku je založený aj dôkaz analytickosti klasickej centrálnej trajektórie v limitnom bode $\mu = 0$ ([4]).

GLTN-riešenie úlohy LP však môže byť aj krajným bodom množiny prípustných riešení. V takom prípade už nemôže byť ostrokomplementárnym optimálnym riešením. Preto dôkaz analytickosti regularizovanej centrálnej trajektórie v bode $\mu = 0$ z článku [6] nie je všeobecný. Z nám dostupných článkov sa nám nepodarilo nájsť ani jeden, v ktorom by bol uvedený úplný dôkaz analytickosti regularizovanej centrálnej trajektórie v limitnom bode.

Zdôraznime, že pri odhade rýchlosti lokálnej konvergencie algoritmov na sledovanie centrálnej trajektórie sa využíva práve jej analytickosť v limitnom bode $\mu = 0$.

Kapitola 4

Záver

V práci sme sa zaoberali špeciálnym typom optimálneho riešenia úlohy LP - optimálnym riešením, ktoré je najbližšie k zadanému bodu.
V prvej časti práce bolo našim cieľom motivovať čitateľa a oboznámiť ho s problematikou. Poskytnutý prehľad metód má za cieľ ilustrovať súčasný stav v tejto oblasti. Ako si môžeme všimnúť, ani jedna z predstavených metód nie je „dokonalá“ - neumožňuje rutinne riešiť úlohy na hľadanie GLTN-riešenia. Väčšinou narážame na problém odhadovania vopred neznámeho parametra.
Regularizovaná centrálna trajektória, ktorú predstavujeme v druhej časti práce, má potenciál odstrániť tento problém. Je to pomerne nová myšlienka a preto nie je ešte detailne rozpracovaná. Snažili sme sa čitateľovi popísať základné vlastnosti trajektórie tak, aby ich mohol nielen prijať, ale tiež uvidieť, z čoho vyplývajú. Viaceré myšlienky dôkazov sme prevzali z článku [6] pre úlohy semidefinitného programovania. Podávame ich však pre úlohy LP, navyše ucelenejšie a podrobnejšie. Cieľom je, aby čitateľ nemusel strácať čas s „lúskaním“ dôkazov, aby mohol rýchlejšie napredovať. Navyše je takto myšlienka regularizovanej centrálnej trajektórie lepšie dostupná tým čitateľom, ktorí nemajú základy semidefinitného programovania.
Ohľadom regularizovanej centrálnej trajektórie ostáva veľa otvorených otázok. Predovšet-

kým, stále nie je zrejmé, či trajektória je analytická v limitnom bode $\mu = 0$. S tým potom súvisí otázka konštrukcie efektívneho algoritmu na hľadanie GLTN-riešenia, založeného na sledovaní trajektórie. Takisto nie je jasné, v akom najkratšom čase dokážeme nájsť GLTN-riešenie.

Pre viaceré úlohy z oblasti ekonomickej a finančnej matematiky môže mať GLTN-riešenie zaujímavú interpretáciu. Len samotnú základnú myšlienku (GLTN-riešenie je spomedzi optimálnych riešení najbližšie k zadanému bodu) možno aplikovať v situáciách, keď optimalizujeme už zabehnutý systém. V takom prípade býva často výhodné, ak sa nové optimálne riešenie čo najmenej líši od predchádzajúceho - napríklad z dôvodu dodatočných transakčných nákladov. Ak by sme chceli transakčné náklady modelovať presne, často dostaneme oveľa zložitejšiu úlohu. Môžeme ich však zohľadniť aspoň tým, že namiesto riešenia obyčajnej úlohy LP hľadáme GLTN-riešenie.

Aplikáciám GLTN-riešenia sme sa v práci nevenovali. Myslíme si však, že môže byť užitočné v rôznych úlohách a že si zaslúži pozornosť. Naša práca si kladie za cieľ pomôcť prípadným záujemcom o štúdium v tejto oblasti.

Literatúra

[1] J. Dieudonne , *Foundations of Modern Analysis*, Vol. I., Academic Press, New York, 1969.

[2] A. I. Golikov, Yu. G. Evtushenko, N. Mollaverdi , *Application of Newton's Method for Solving Large Linear Programming Problems*, Computational Mathematics and Mathematical Physics, Vol. 44, No. 9, 2004, pp. 1484-1493.

[3] M. Halická , *Prednášky z predmetu Lineárne programovanie 2* v akademickom roku 2007/2008.

[4] M. Halická , *Two simple proofs for analyticity of the central path in linear programming*, Operations Research Letters, 28, No. 1, 2001, pp. 9-19.

[5] C. Kanzow, H. Qi, L. Qi , *On the Minimum Norm Solution of Linear Programs*, Journal of Optimization Theory and Applications, Vol 116, No. 2 (2004), pp. 333-345.

[6] A. Lin , *On a special class of regularized central paths for semidefinite programs (second revised version)*, článok prijatý do Mathematical Programming.

[7] O. L. Mangasarian , *Least-norm linear programming solution as an unconstrained minimization problem*, J. Math. Analy. and Applic. 92, 1983 ,pp. 240-251.

[8] O. L. Mangasarian, R. R. Meyer , *Nonlinear perturbation of linear programs*, SIAM Journal on Control and Optimization 17, 1979, pp. 745-757.

[9] J. Plesník, J. Dupačová, M. Vlach , *Lineárne programovanie*, Alfa, Bratislava, 1990.

[10] C. Roos, T. Terlaky, J-Ph. Vial , *Interior Point Methods for Linear Optimization*, Springer (Second Edition), 2006.

[11] P. W. Smith, H. Wolkowicz , *A nonlinear equation for linear programming*, Mathematical Programming 34, 1986, pp. 235-238.

[12] A. N. Tikhonov, V. Y. Arsenin , *Solutions of ill-posed Problems*, Halsted Press, Wiley, New York, 1977.

[13] R.J. Vanderbei , *Linear Programming: Foundations and Extensions*, Kluwer Academic Publishers: Boston, MA, 1996.

[14] S. J. Wright , *Primal-dual interior-point methods*, SIAM, Philadelphia, PA, 1997.

[15] Y.-B. Zhao, D. Li , *Locating the Least 2-Norm Solution of Linear Program*, SIAM Journal of Optimization, Vol.12, No.4, (2002), pp. 893-912.