

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



**Algoritmy metód vnútorného bodu v
lineárnom programovaní**

RIGORÓZNA PRÁCA

2014

Mgr. Marek KABÁT

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Evidenčné číslo:

0f27c61a-9d67-4146-bb95-b14d565c7801

**Algoritmy metód vnútorného bodu v
lineárnom programovaní**

RIGORÓZNA PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	9.1.9. Aplikovaná matematika
Pracovisko:	Katedra aplikovanej matematiky a štatistiky
Konzultant:	doc. RNDr. Margaréta Halická, CSc.

Bratislava 2014

Mgr. Marek KABÁT

Abstrakt

KABÁT, Marek: *Algoritmy metód vnútorného bodu v lineárnom programovaní*. [Rigorózna práca] – Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky. Katedra aplikovanej matematiky a štatistiky. – Konzultant: doc. RNDr. Margaréta Halická, CSc. – Bratislava: FMFI UK, 2014, 55 s.

Rigorózna práca sa zaoberá algoritmi metód vnútorného bodu v lineárnom programovaní. Hlavným cieľom práce je uviesť základnú kategorizáciu primárno-duálnych algoritmov metód vnútorného bodu v lineárnom programovaní, vysvetliť a ilustrovať princíp konkrétnych algoritmov a uviesť a dokázať ich podstatné vlastnosti. Práca vychádza zo základných výsledkov z teórie metód vnútorného bodu v lineárnom programovaní. Ústrednú časť práce tvoria primárno-duálne algoritmy redukcie potenciálu a algoritmy sledovania centrálnej trajektórie. Obsah práce je na záver doplnený o triedu primárno-duálnych neprípustných algoritmov.

Kľúčové slová: lineárne programovanie • metódy vnútorného bodu • centrálna trajektória • primárno-duálny algoritmus • potenciálová funkcia.

Abstract

KABÁT, Marek: *Algorithms of Interior Point Methods in Linear Programming*. [Rigorous Thesis] – Comenius University in Bratislava. Faculty of Mathematics, Physics and Informatics. Department of Applied Mathematics and Statistics. – Supervisor: doc. RNDr. Margaréta Halická, CSc. – Bratislava: FMFI UK, 2014, 55 p.

The rigorous thesis deals with interior point algorithms for linear programming. The main objective of this work is to specify the basic categorization of primal-dual interior point algorithms for linear programming, explain and illustrate the principle of specific algorithms and state and prove their essential characteristics. The work is based on the fundamental results of the theory of interior point methods in linear programming. The main part of thesis consists of primal-dual potential reduction algorithms and path following algorithms. Content of thesis is completed at the end by a class of primal-dual infeasible algorithms.

Keywords: linear programming • interior point methods • central path • primal-dual algorithm • potential function.

Obsah

Úvod	1
1 Metódy vnútorného bodu v lineárnom programovaní	3
1.1 Základné výsledky z teórie lineárneho programovania	3
1.2 Základné predpoklady	4
1.3 Bariérový problém a centrálna trajektória	5
1.4 Primárno-duálne metódy	6
1.5 O zložitosti algoritmov	9
2 Algoritmy redukcie potenciálu	10
2.1 Potenciálová funkcia Φ_ρ	10
2.2 Algoritmus	12
2.3 Redukcia funkcie Φ_ρ a konvergencia	13
2.4 Odhad funkcie Φ_ρ pozdĺž prípustného smeru	14
2.5 Odhad redukcie funkcie Φ_ρ	17
2.6 Polynomiálna zložitosť	19
3 Algoritmy sledovania centrálnej trajektórie	20
3.1 Špecifické okolie centrálnej trajektórie	20
3.2 Polynomiálna zložitosť	25
3.3 Algoritmus s krátkym krokom	26
3.4 Prediktor-korektor algoritmus	31
3.5 Algoritmus s dlhým krokom	35
3.6 Konvergencia	38
3.7 Porovnanie algoritmov	39

4	Neprípustné algoritmy	40
4.1	Neprípustnosť	40
4.2	Algoritmus	40
4.3	Konvergencia a polynomiálna zložitosť	42
4.4	Mehrotrov prediktor-korektor algoritmus	43
5	Dodatok	47
5.1	Dôkaz Lemy 2.1	47
5.2	Dôkaz Lemy 2.4	48
5.3	Odvozenie špecifického okolia pre $n = 2$	48
5.4	Vzťah parametrov θ a σ v algoritme s krátkym krokom	49
	Záver	50
	Zoznam použitej literatúry	52

Zoznam symbolov a značiek

V práci budeme používať nasledovné jednotné značenie.

\mathbf{R}^n	Priestor reálnych n -rozmerných vektorov.
$\mathbf{R}^{m \times n}$	Priestor reálnych matíc rozmeru $m \times n$.
i, j	Indexy označujúce zložky vektorov a matíc.
\emptyset	Prázdna množina.
$[u_i]_{i=1}^n, [u_i]_{i \in \mathcal{M}}$	Vektor so zložkami u_i , kde $i = 1, 2, \dots, n$, resp. $i \in \mathcal{M}$.
u^T	Transpozícia vektora u .
$u^T v$	Skalárny súčin vektorov $u, v \in \mathbf{R}^n$, $u^T v = \sum_{i=1}^n u_i v_i$.
$\ \cdot\ , \ \cdot\ _2$	Euklidova norma. Pre $u \in \mathbf{R}^n$, $\ u\ _2 = \left(\sum_{i=1}^n u_i^2 \right)^{1/2}$.
$\ \cdot\ _1$	Jednotková norma. Pre $u \in \mathbf{R}^n$, $\ u\ _1 = \sum_{i=1}^n u_i $.
$\ \cdot\ _\infty$	Maximová norma. Pre $u \in \mathbf{R}^n$, $\ u\ _\infty = \max_{i=1,2,\dots,n} u_i $.
	Poznámka: pre ľubovoľný vektor $u \in \mathbf{R}^n$ platí $\ u\ _\infty \leq \ u\ _2 \leq \ u\ _1$.
$f(\cdot)$	Funkcia f .
$\nabla f(\cdot)$	Gradient funkcie f .
$\nabla^2 f(\cdot)$	Hessova matica funkcie f .
$\exp(\cdot)$	Exponenciálna funkcia.
$\ln(\cdot)$	Prirodzený logaritmus.
e	Vektor jednotiek príslušného rozmeru, $e = (1, 1, \dots, 1)^T$.
I	Jednotková matica príslušného rozmeru, $I = \text{diag}(e)$.
$h(A)$	Hodnosť matice A .
B^{-1}	Inverzná matica k regulárnej matici B , $BB^{-1} = B^{-1}B = I$.

$(P), (D)$	Štandardná primárna, resp. duálna úloha lineárneho programovania.
x	Vektor premenných primárnej úlohy (P) , $x \in \mathbf{R}^n$.
y	Vektor premenných duálnej úlohy (D) , $y \in \mathbf{R}^m$.
s	Vektor doplnkových premenných duálnej úlohy (D) , $s \in \mathbf{R}^n$.
(A, b, c)	Vstupné údaje lineárneho programu, $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$.
\mathcal{P}, \mathcal{D}	Množina prípustných riešení úlohy (P) , resp. úlohy (D) .
\mathcal{F}	Množina primárno-duálnych prípustných riešení, $\mathcal{F} = \mathcal{P} \times \mathcal{D}$.
$\mathcal{P}^*, \mathcal{D}^*$	Množina optimálnych riešení úlohy (P) , resp. úlohy (D) .
\mathcal{F}^*	Množina primárno-duálnych optimálnych riešení, $\mathcal{F}^* = \mathcal{P}^* \times \mathcal{D}^*$.
$\mathcal{P}^\circ, \mathcal{D}^\circ$	Množina ostro prípustných riešení úlohy (P) , resp. úlohy (D) .
\mathcal{F}°	Množina primárno-duálnych ostro prípustných riešení, $\mathcal{F}^\circ = \mathcal{P}^\circ \times \mathcal{D}^\circ$.
μ	Bariérový parameter, $\mu > 0$, $\mu = x^T s / n$.
(x_μ, y_μ, s_μ)	Parametrizované riešenie systému (1.4)-(1.6).
\mathcal{C}	Centrálna trajektória, $\{(x_\mu, y_\mu, s_\mu) \mid \mu > 0\}$.
k	Iteračný index, $k = 0, 1, 2, \dots$.
(x^k, y^k, s^k)	Iteračný bod primárno-duálneho algoritmu.
$(\Delta x, \Delta y, \Delta s)$	Primárno-duálny Newtonov smer.
σ, σ_k	Centrujúci parameter.
α, α_k	Parameter dĺžky kroku.
$(x(\alpha), y(\alpha), s(\alpha))$	$(x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$
$\mu(\alpha)$	$x(\alpha)^T s(\alpha) / n$
X, X^k	Diagonálna $n \times n$ matica so zložkami vektora x , resp. x^k , $X = \text{diag}(x)$, $X^k = \text{diag}(x^k)$.
S, S^k	Diagonálna $n \times n$ matica so zložkami vektora s , resp. s^k , $S = \text{diag}(s)$, $S^k = \text{diag}(s^k)$.
D	$X^{1/2} S^{-1/2}$
$\Delta X, \Delta S$	Diagonálna $n \times n$ matica so zložkami vektora Δx , resp. Δs , $\Delta X = \text{diag}(\Delta x)$, $\Delta S = \text{diag}(\Delta s)$.
$O(\cdot)$	Označenie zložitosti algoritmov.
$[\cdot]$	Horná celá časť reálneho čísla.
$\Phi_\rho(\cdot)$	Potenciálová funkcia, $\Phi_\rho(x, s) = \rho \ln x^T s - \sum_{i=1}^n \ln x_i s_i$, $\rho \geq n$.
r_P, r_P^k	Primárne rezíduum, $r_P = Ax - b$, $r_P^k = Ax^k - b$.
r_D, r_D^k	Duálne rezíduum, $r_D = A^T y + s - c$, $r_D^k = A^T y^k + s^k - c$.

- $\mathcal{N}_2(\theta)$ Špecifické okolie centrálnej trajektórie pre dané $\theta \in (0, 1)$,
 $\{(x, y, s) \in \mathcal{F}^\circ \mid \|XSe - \mu e\| \leq \theta\mu, \mu = x^T s/n\}$.
- $\mathcal{N}_{-\infty}(\gamma)$ Špecifické okolie centrálnej trajektórie pre dané $\gamma \in (0, 1)$,
 $\{(x, y, s) \in \mathcal{F}^\circ \mid x_i s_i \geq \gamma\mu, i = 1, 2, \dots, n, \mu = x^T s/n\}$.
- $\mathcal{N}_{-\infty}(\gamma, \varrho)$ Špecifické okolie centrálnej trajektórie pre dané $\gamma \in (0, 1), \varrho \geq 1, (x^0, y^0, s^0)$,
 $\{(x, y, s) \mid \|(r_P, r_D)\| \leq [\|(r_P^0, r_D^0)\|/\mu_0]\varrho\mu, (x, s) > 0, x_i s_i \geq \gamma\mu, i = 1, 2, \dots, n\}$.

Úvod

If God did not exist, it would be necessary to invent him. Fortunately, there is no need to solve such a thought for mathematics.

Lineárne programovanie je špeciálnym prípadom matematického konvexného programovania, ktoré rieši problém minimalizácie alebo maximalizácie lineárnej účelovej funkcie na množine prípustných riešení, ktorá je charakterizovaná sústavou lineárnych rovníc alebo neostrých nerovníc. Štandardným algoritmom na riešenie úloh lineárneho programovania je simplexová metóda, ktorú navrhol George Dantzig [2] v roku 1947. Klee a Minty [10] v roku 1972 zistili, že počet iterácií simplexovej metódy môže byť exponenciálny vzhľadom k rozmeru úlohy, čo podnietilo odborníkov ku koštrukcii nových algoritmov s polynomiálnou zložitou. Vznikalo niekoľko nových algoritmov, no prevratným sa stal až výsledok Narendra Karmarkara [8], ktorý v roku 1984 predstavil svoj projektívny algoritmus na riešenie lineárnych úloh v polynomiálnom čase. Ukázalo sa, že Karmarkarov algoritmus úzko súvisí s metódami vnútorného bodu, ktoré sa do toho obdobia aplikovali najmä na riešenie všeobecných úloh nelineárneho programovania. Karmarkarov algoritmus položil základy moderných metód vnútorného bodu, ktoré okrem lineárnych úloh dokážu efektívne riešiť aj širokú triedu štrukturovaných úloh nelineárneho konvexného programovania.

Rigorózna práca sa zaoberá algoritmi metód vnútorného bodu v lineárnom programovaní. V priebehu vývoja metód vnútorného bodu vznikla určitá kategorizácia algoritmov. Cieľom práce je uviesť základnú kategorizáciu primárno-duálnych algoritmov metód vnútorného bodu v lineárnom programovaní, vysvetliť a ilustrovať princíp konkrétnych algoritmov, uviesť ich charakteristické znaky a dokázať ich podstatné vlastnosti. Práca nadväzuje na diplomovú prácu *Metódy vnútorného bodu v lineárnom programovaní a ich aplikácie vo financiách* [7], čím dopĺňa a rozširuje časť venovanú algoritmom. Účelom práce je poskytnúť základný prehľad zaužívaných primárno-duálnych algoritmov metód vnútorného bodu v lineárnom programovaní, pričom sa v práci kladie dôraz na podrobnú analýzu algoritmov, dôkladné vysvetlenie, ilustrácie a jednotný spôsob prezentácie.

Rigorózna práca je rozdelená do piatich kapitol. V prvej kapitole sú uvedené základné výsledky z teórie metód vnútorného bodu v lineárnom programovaní. Druhá kapitola sa zaoberá primárno-duálnymi algoritmi redukcie potenciálu. Ústrednou časťou práce je tretia kapitola, ktorá prezentuje tri základné varianty primárno-duálnych algoritmov sledovania centrálnej trajektórie. Štvrtá kapitola sa venuje neprípustným primárno-duálnym algoritmom, v rámci ktorých je opísaný Mehrotrov prediktor-korektor algoritmus. Obsahom piatej kapitoly sú technické výsledky a postupy, ktoré boli využité v rigoróznej práci.

METÓDY VNÚTORNÉHO BODU V LINEÁRNOM PROGRAMOVANÍ

Prvá kapitola práce je venovaná kľúčovým výsledkom z teórie metód vnútorného bodu v lineárnom programovaní. Kapitola predstavuje širší abstrakt teoretickej časti diplomovej práce *Metódy vnútorného bodu v lineárnom programovaní a ich aplikácie vo financiách* [7]. V tejto kapitole sformulujeme štandardnú úlohu lineárneho programovania a uvedieme základné pojmy a predpoklady. Zároveň vysvetlíme myšlienku metód vnútorného bodu, opíšeme princíp primárno-duálnych metód a vysvetlíme spôsob označenia zložitosti algoritmov.

1.1 Základné výsledky z teórie lineárneho programovania

V práci budeme vychádzať z úlohy lineárneho programovania v štandardnom tvare

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

kde $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n}$. K úlohe (P) definujeme množinu prípustných riešení \mathcal{P} a množinu optimálnych riešení \mathcal{P}^*

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

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

Duálna úloha k úlohe (P) má tvar

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

kde $s \in \mathbf{R}^n$, $y \in \mathbf{R}^m$. K úlohe (D) analogicky definujeme množinu prípustných riešení \mathcal{D} a množinu optimálnych riešení \mathcal{D}^*

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

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

Dvojicu úloh (P), (D) nazývame primárno-duálnou dvojicou a rozdiel hodnôt ich účelových funkcií $c^T x - b^T y$ nazývame duálnou medzerou. Pre dvojicu úloh (P), (D) zadefinujeme množinu primárno-duálnych prípustných riešení \mathcal{F}

$$\mathcal{F} = \{(x, y, s) \in \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \mid Ax = b, A^T y + s = c, (x, s) \geq 0\}.$$

Veta 1.1 (Slabá dualita). Pre ľubovoľné primárno-duálne prípustné riešenie $(x, y, s) \in \mathcal{F}$ platí $0 \leq x^T s = c^T x - b^T y$. Navyiac, ak pre nejakú trojicu $(x^*, y^*, s^*) \in \mathcal{F}$ platí $c^T x^* = b^T y^*$, potom x^* je optimálnym riešením úlohy (P) a dvojica (y^*, s^*) je optimálnym riešením úlohy (D). [23]

Veta 1.2 (Dualita). Ak obe úlohy (P), (D) majú prípustné riešenie, t.j. $\mathcal{F} \neq \emptyset$, potom obe úlohy (P), (D) majú optimálne riešenie. Trojica (x^*, y^*, s^*) je primárno-duálnym optimálnym riešením úloh (P), (D) práve vtedy, keď

$$Ax^* = b, x^* \geq 0 \tag{1.1}$$

$$A^T y^* + s^* = c, s^* \geq 0 \tag{1.2}$$

$$X^* S^* e = 0. \tag{1.3}$$

Veta 1.2 formuluje nutné a postačujúce podmienky optimality pre primárno-duálnu dvojicu (P), (D). Podmienky (1.1), (1.2) zodpovedajú podmienkam primárnej, resp. duálnej prípustnosti a podmienka (1.3) zodpovedá podmienke komplementarity. To znamená, že za predpokladu $\mathcal{F} \neq \emptyset$ môžeme definovať množinu primárno-duálnych optimálnych riešení \mathcal{F}^*

$$\mathcal{F}^* = \{(x, y, s) \in \mathcal{F} \mid x_i s_i = 0, i = 1, 2, \dots, n\}.$$

Veta 1.3 (Silná dualita). Ak jedna z dvojice úloh (P), (D) má optimálne riešenie, potom má aj druhá a optimálne hodnoty účelových funkcií sa rovnajú. [23]

Veta 1.4 (Ostrá komplementarita). Ak obe úlohy (P), (D) majú prípustné riešenie, t.j. $\mathcal{F} \neq \emptyset$, potom existuje primárno-duálne optimálne riešenie $(x^*, y^*, s^*) \in \mathcal{F}^*$ splňajúce $x^* + s^* > 0$. [23]

1.2 Základné predpoklady

Pri spracovaní teórie metód vnútorného bodu je dôležité a nevyhnutné vychádzať z dvoch základných predpokladov.

Predpoklad 1. Matica $A \in \mathbf{R}^{m \times n}$ má plnú riadkovú hodnosť, t.j. $h(A) = m \leq n$.

Uvedený predpoklad má technický charakter a zabezpečí jednoznačné párovanie medzi vektormi $y \in \mathbf{R}^m, s \in \mathbf{R}^n$ tak, aby $(y, s) \in \mathcal{D}$. Pred formuláciou druhého predpokladu najprv zdefinujeme množinu ostro prípustných primárno-duálnych riešení \mathcal{F}°

$$\mathcal{F}^\circ = \{(x, y, s) \in \mathcal{F} \mid (x, s) > 0\}.$$

Poznamenajme, že množina ostro prípustných riešení \mathcal{F}° tvorí relatívne vnútro množiny \mathcal{F} , a preto ostro prípustné riešenia nazývame aj vnútornými bodmi.

Predpoklad 2. Každá z dvojice úloh (P), (D) má vnútorný bod, t.j. $\mathcal{F}^\circ \neq \emptyset$.

Zatiaľ čo Predpoklad 1 je technický, Predpoklad 2 o existencii vnútorného bodu je nevyhnutný a zaručuje existenciu ústredného objektu metód vnútorného bodu, ktorým je centrálna trajektória. Pre jednoduchosť budeme uvedené predpoklady označovať

$$(P1) \quad h(A) = m \leq n,$$

$$(P2) \quad \mathcal{F}^\circ \neq \emptyset.$$

V ďalších častiach tejto práce budeme vždy predpokladať, že základné predpoklady (P1) a (P2) sú splnené.

1.3 Bariérový problém a centrálna trajektória

Princíp metód vnútorného bodu spočíva v postupnom riešení pomocných bariérových úloh. Ku pôvodnej dvojici úloh (P), (D) priradíme bariérové úlohy (P_μ) , (D_μ)

$$(P_\mu) \quad \min_{x>0} \left\{ c^T x - \mu \sum_{i=1}^n \ln x_i \mid Ax = b \right\}, \quad (D_\mu) \quad \max_{s>0, y} \left\{ b^T y + \mu \sum_{i=1}^n \ln s_i \mid A^T y + s = c \right\},$$

kde $\mu > 0$ je bariérový parameter. Ak predpoklady (P1) a (P2) sú splnené, potom každá z bariérových úloh (P_μ) , (D_μ) má práve jedno optimálne riešenie pre ľubovoľné $\mu > 0$. Nutné a postačujúce podmienky optimality formuluje nasledujúce tvrdenie.

Veta 1.5. *Nech $\mu > 0$. Potom vektor x je optimálnym riešením úlohy (P_μ) a dvojica (y, s) je optimálnym riešením úlohy (D_μ) práve vtedy, keď*

$$Ax = b, \quad x > 0 \tag{1.4}$$

$$A^T y + s = c, \quad s > 0 \tag{1.5}$$

$$XSe = \mu e. \tag{1.6}$$

To znamená, že vyriešiť dvojicu bariérových úloh (P_μ) , (D_μ) znamená vyriešiť systém (1.4)-(1.6). Presné riešenia systému (1.4)-(1.6) definujú centrálnu trajektóriu, ktorá je kľúčovým elementom metód vnútorného bodu. Ak označíme (x_μ, y_μ, s_μ) riešenie systému (1.4)-(1.6) parametrizované parametrom $\mu > 0$, potom pod centrálnou trajektóriou rozumieme množinu

$$\mathcal{C} = \{(x_\mu, y_\mu, s_\mu) \mid \mu > 0\}.$$

Poznamenajme, že centrálna trajektória \mathcal{C} jednoznačne existuje práve vtedy, keď predpoklady (P1) a (P2) sú splnené. Myšlienka metód vnútorného bodu spočíva v postupnom znižovaní parametra $\mu \rightarrow 0$, čím sa duálna medzera $x^T s = n\mu$ znižuje, a tým sa systém (1.4)-(1.6) približuje k nutným a postačujúcim podmienkam optimality z Vety 1.2.

Veta 1.6. *Centrálna trajektória \mathcal{C} konverguje pre $\mu \rightarrow 0$ a jej jediným limitným bodom je primárno-duálne optimálne riešenie $(x^*, y^*, s^*) \in \mathcal{F}^*$ úloh (P), (D). Navyše, dvojica (x^*, s^*) zodpovedá ostro komplementárnemu primárno-duálnemu optimálnemu riešeniu. [24]*

Inými slovami, zmenšovaním parametra μ postupnosť riešení systému (1.4)-(1.6) konverguje k primárno-duálnemu optimálnemu riešeniu $(x^*, y^*, s^*) \in \mathcal{F}^*$. V matematickom kontexte to znamená, že

$$\lim_{\mu \rightarrow 0^+} (x_\mu, y_\mu, s_\mu) = (x^*, y^*, s^*).$$

Na základe postupného riešenia systému (1.4)-(1.6) pre $\mu \rightarrow 0$ je tak možné efektívne vyriešiť dvojicu úloh (P) , (D) súčasne.

1.4 Primárno-duálne metódy

Princíp primárno-duálnych metód spočíva v súčasnom riešení úloh (P) , (D) . Tieto metódy hľadajú primárno-duálne optimálne riešenie postupným riešením systému (1.4)-(1.6) pomocou modifikovanej Newtonovej metódy so skrátenou dĺžkou kroku. Pre $\mu > 0$ definujeme funkciu $F : \mathbf{R}^{2n+m} \rightarrow \mathbf{R}^{2n+m}$

$$F(x, y, s) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ X S e - \mu e \end{bmatrix}, \quad (x, s) \geq 0.$$

Nájsť riešenie systému (1.4)-(1.6) zrejme znamená vyriešiť rovnicu $F(x, y, s) = 0$ s dodatočnou podmienkou $(x, s) \geq 0$. Použitím Taylorovho rozvoja funkcie F v okolí bodu (x^k, y^k, s^k) dostaneme

$$\begin{aligned} F(x, y, s) &\approx F(x^k, y^k, s^k) + \nabla F(x^k, y^k, s^k)(\Delta x, \Delta y, \Delta s)^T = 0, \\ \nabla F(x^k, y^k, s^k)(\Delta x, \Delta y, \Delta s)^T &= -F(x^k, y^k, s^k), \end{aligned} \quad (1.7)$$

kde ∇F je Jakobiho matica funkcie F a $\Delta x \in \mathbf{R}^n$, $\Delta y \in \mathbf{R}^m$, $\Delta s \in \mathbf{R}^n$ sú smerové vektory. Rozpísaním rovnice (1.7) do maticového tvaru dostaneme

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} Ax^k - b \\ A^T y^k + s^k - c \\ X^k S^k e - \mu e \end{bmatrix}. \quad (1.8)$$

Smerové vektory $\Delta x \in \mathbf{R}^n$, $\Delta y \in \mathbf{R}^m$, $\Delta s \in \mathbf{R}^n$, ktoré sú riešením systému (1.8), sa nazývajú primárno-duálne Newtonove smery. Iteračné body Newtonovej metódy vypočítame podľa iteračnej schémy

$$(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + \alpha(\Delta x, \Delta y, \Delta s), \quad (1.9)$$

kde $\alpha \in [0, 1]$ označuje dĺžku kroku pozdĺž smeru $(\Delta x, \Delta y, \Delta s)$. Parameter α volíme v každej iterácii tak, aby boli splnené podmienky $x^k + \alpha \Delta x > 0$, $s^k + \alpha \Delta s > 0$. Týmto spôsobom Newtonova metóda generuje iteračné body (x^k, y^k, s^k) spĺňajúce podmienku $(x^k, s^k) > 0$. Poznamenajme, že ak iteračný bod (x^k, y^k, s^k) je ostro prípustným primárno-

duálnym riešením, potom systém (1.8) môžeme zjednodušiť do tvaru

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \mu e \end{bmatrix}. \quad (1.10)$$

Zároveň platí, že Newtonova metóda zachováva ostrú prípustnosť. Z uvedených vlastností vyplýva, že ak štarovací bod Newtonovej metódy je ostro prípustným riešením, potom v každej iterácii Newtonovej metódy riešime systém (1.10). Týmto spôsobom nás mechanizmus Newtonovej metódy udržuje v každej iterácii v množine vnútorných bodov \mathcal{F}° . Dôležitou súčasťou algoritmu je zabezpečiť správny spôsob zmenšovania parametra μ . V práci sa budeme zaoberať algoritmi s implicitnou voľbou ¹ bariérového parametra

$$\mu = \frac{x^T s}{n}.$$

Tento implicitný spôsob voľby parametra μ je dodatočne upravený centrujúcim parametrom $\sigma \in [0, 1]$, ktorého účelom je zlepšiť priebeh konvergence. Nakoniec získame systém v tvare

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma \mu e \end{bmatrix}. \quad (1.11)$$

V každej iterácii algoritmu zároveň kontrolujeme veľkosť duálnej medzery $x^T s$ pomocou vopred určenej tolerančnej konštanty $\epsilon > 0$. Na základe opísaného princípu môžeme sformulovať schému generického primárno-duálneho algoritmu metód vnútorného bodu na riešenie úloh (P) , (D) .

Schéma generického algoritmu

vstup $(x^0, y^0, s^0) \in \mathcal{F}^\circ, \epsilon > 0$

prirad' $k := 0$

repeat

zvoľ $\sigma_k \in [0, 1]$

prirad' $\mu_k := (x^k)^T s^k / n$

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

¹Poznamenajme, že existujú aj algoritmy s explicitnou voľbou parametra μ . Takéto algoritmy znižujú bariérový parameter podľa schémy $\mu_{k+1} = (1 - \theta)\mu_k$, kde $\theta \in (0, 1)$ je parameter redukcie.

zvoľ $\alpha_k \in [0, 1] : x^k + \alpha_k \Delta x > 0, s^k + \alpha_k \Delta s > 0$
prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k (\Delta x, \Delta y, \Delta s)$
prirad' $k := k + 1$
until $(x^k)^T s^k \leq \epsilon$

Na záver tejto časti uvedieme dve pomocné tvrdenia, ktoré dokazujú, že vektory Δx a Δs sú navzájom ortogonálne a postupnosť $\{\mu_k\}$ generovaná na základe generického algoritmu má nerastúci charakter.

Lema 1.1. Ak trojica $(\Delta x, \Delta y, \Delta s)$ je riešením systému (1.11), potom $\Delta x^T \Delta s = 0$. [24]

Dôkaz. Pre trojicu $(\Delta x, \Delta y, \Delta s)$ zrejme platí $A\Delta x = 0, A^T \Delta y + \Delta s = 0$. Z uvedených vzťahov jednoducho odvodíme, že

$$\Delta x^T \Delta s = \Delta x^T (-A^T \Delta y) = -\underbrace{(A\Delta x)^T}_{=0} y = 0.$$

□

Lema 1.2. Ak $\sigma_k \in [0, 1]$, potom $\mu_{k+1} \leq \mu_k$, pričom rovnosť nastáva práve vtedy, keď $\sigma_k = 1$ alebo $\alpha_k = 0$.

Dôkaz. Na základe schémy generického algoritmu odvodíme, že

$$\begin{aligned} \mu_{k+1} &= \frac{(x^{k+1})^T s^{k+1}}{n} \\ &= \frac{(x^k + \alpha_k \Delta x)^T (s^k + \alpha_k \Delta s)}{n} \\ &= \frac{(x^k)^T s^k}{n} + \alpha_k \frac{(x^k)^T \Delta s + (s^k)^T \Delta x + \alpha_k \overbrace{(\Delta x)^T \Delta s}^{=0}}{n} \\ &= \frac{(x^k)^T s^k}{n} + \alpha_k \frac{n\sigma_k \mu_k - (x^k)^T s^k}{n} \\ &= \mu_k - \underbrace{\alpha_k(1 - \sigma_k)\mu_k}_{\geq 0} \leq \mu_k. \end{aligned}$$

Keďže $\mu_k > 0$, tak je zrejme, že $\mu_{k+1} = \mu_k$ práve vtedy, keď $\alpha_k = 0$ alebo $\sigma_k = 1$. □

Generický algoritmus je základným konceptom primárno-duálnych algoritmov metód vnútorného bodu v lineárnom programovaní. Existujú dve základné modifikácie, ktoré sa odvíjajú od generického algoritmu – algoritmus redukcie potenciálu a algoritmus sledovania centrálnej trajektórie. Zatiaľ čo algoritmus sledovania centrálnej trajektórie pristupuje k redukcii duálnej medzery priamo, algoritmus redukcie potenciálu znižuje duálnu medzeru nepriamo pomocou potenciálovej funkcie. V nasledujúcich dvoch kapitolách sa zameriame na tieto algoritmy, podrobne opíšeme ich princíp a uvedieme ich základné vlastnosti.

1.5 O zložitosti algoritmov

Teória zložitosti algoritmov sa zaoberá hľadaním horného odhadu počtu aritmetických operácií potrebných na nájdenie dostatočne presného riešenia v závislosti od rozmeru úlohy. V lineárnom programovaní sa stalo zaužívaným považovať n za rozmer úlohy. Na označenie zložitosti algoritmov budeme používať štandardnú $O(\cdot)$ symboliku.

Definícia 1.1. Nech f a g sú nezáporné reálne funkcie definované na množine prirodzených čísel. Píšeme $g(n) = O(f(n))$, ak

$$\exists \beta > 0, \bar{n} \in \mathbf{R} : g(n) \leq \beta f(n), \forall n > \bar{n}.$$

Definícia 1.2. Hovoríme, že algoritmus má zložitost' $O(f(n))$, ak pre horný odhad $\bar{k}(n)$ počtu aritmetických operácií algoritmu pre úlohu rozmeru n platí $\bar{k}(n) = O(f(n))$.

Teória zložitosti zohrala dôležitú úlohu vo vývoji lineárneho programovania a metód vnútorného bodu. V roku 1972 matematici Klee a Minty [10] zistili, že počet operácií simplexovej metódy môže byť exponenciálny vzhľadom k rozmeru úlohy.² Toto zistenie podnietilo matematikov ku konštrukcii nových efektívnych algoritmov. Pod pojmom „efektívne algoritmy“ zvyčajne rozumieme algoritmy s polynomiálnou zložitost'ou. Sú to algoritmy, v ktorých počet aritmetických operácií potrebných na vyriešenie úlohy rozmeru n môžeme odhadnúť zhora polynómom v premennej n . Inými slovami, algoritmus má polynomiálny charakter, ak má zložitost' $O(f(n))$, kde $f(n) \leq n^\kappa$ pre nejakú konštantu κ . V roku 1979 uviedol Leonid Khachiyan [9] eplipsoidný algoritmus na riešenie úloh lineárneho programovania v polynomiálnom čase. Síce Khachiyanov algoritmus mal polynomiálny charakter, no v praxi sa prejavil pomalší ako simplexová metóda, a preto sa neujal. Prevratným sa stal až Karmarkarov projektívny algoritmus [8] z roku 1984, ktorého efektivita sa naplno osvedčila aj praxi.

²Poznamenajme, že simplexová metóda vyžaduje v praxi menej operácií, než uvádzajú teoretické výsledky.

ALGORITMY REDUKCIE POTENCIÁLU

Algoritmy redukcie potenciálu predstavujú pôvodnú triedu moderných algoritmov metód vnútorného bodu. Tieto algoritmy prístupujú k redukcii duálnej medzery nepriamo pomocou potenciálovej funkcie so špecifickými vlastnosťami. V tejto kapitole definujeme primárno-duálnu logaritmickú potenciálovú funkciu a uvedieme jej základné vlastnosti. Opíšeme schému primárno-duálneho algoritmu redukcie potenciálu, vysvetlíme jeho princíp a dokážeme jeho polynomiálnu zložitosť.

2.1 Potenciálová funkcia Φ_ρ

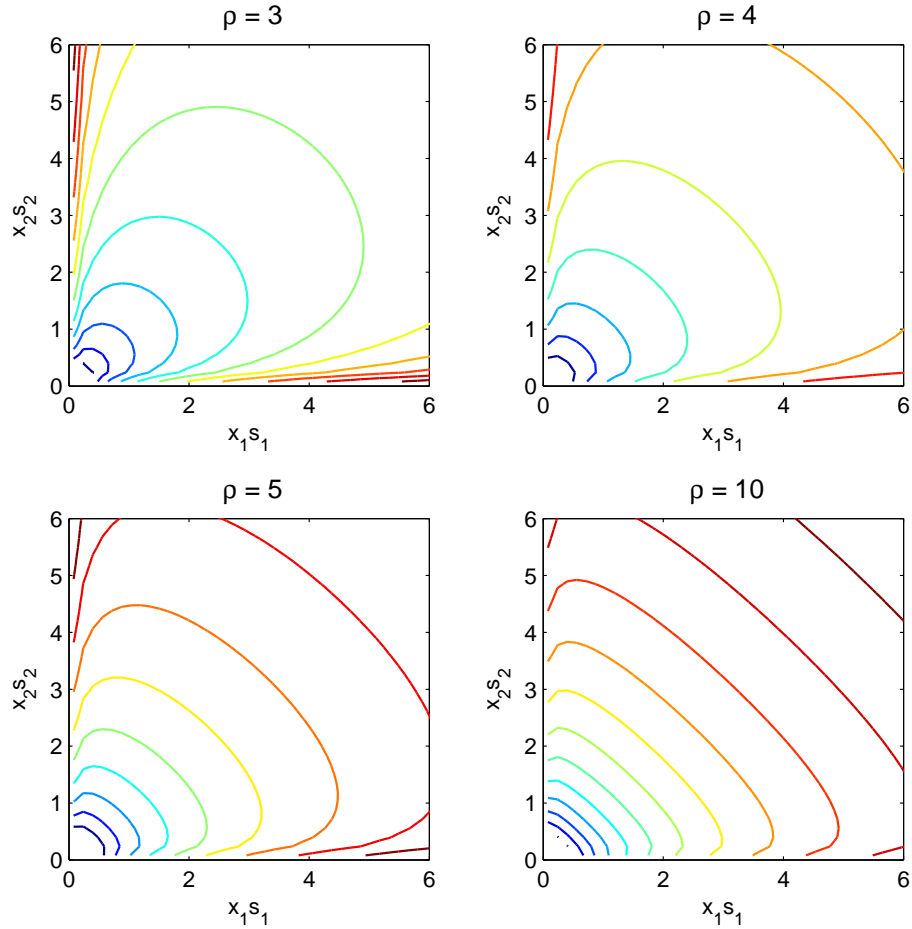
Potenciálové funkcie zohrali vo vývoji metód vnútorného bodu dôležitú rolu. Hlavnou úlohou týchto funkcií je merať kvalitu priebežného riešenia. Poznamenajme, že Karmarkar vo svojom pôvodnom algoritme [8] na riešenie štandardnej úlohy (P) použil logaritmickú potenciálovú funkciu definovanú v tvare

$$\Upsilon_\rho(x) = \rho \ln(c^T x - \lambda) - \sum_{i=1}^n \ln x_i, \quad (2.1)$$

kde $\rho = n + 1$ a λ je dolné ohraňenie účelovej funkcie $c^T x$. Súbežne s Karmarkarovými výsledkami vznikali práce [4], [5], [22], [27], ktoré prezentovali využitie iných potenciálových funkcií, ktoré však záviseli vždy len od vektora primárnych premenných. Vývoj v oblasti primárno-duálnych algoritmov podnietil odborníkov k úvahám o potenciálových funkciách, v ktorých by okrem vektora primárnych premenných vystupoval aj vektor duálnych premenných. Dôležitým výsledkom vo vývoji algoritmov redukcie potenciálu sa stala primárno-duálna logaritmická potenciálová funkcia, ktorú prezentovali Tanabe [26], Todd a Ye [28] v tvare

$$\Phi_\rho(x, s) = \rho \ln x^T s - \sum_{i=1}^n \ln x_i s_i, \quad (2.2)$$

pre nejaký parameter $\rho > n$. Všimnime si vzťah medzi funkciami Υ_ρ a Φ_ρ . Z teórie duality v lineárnom programovaní (Veta 1.1) vieme, že výraz $b^T y$ je pre ľubovoľnú dvojicu $(y, s) \in \mathcal{D}$ dolným ohraňením primárnej účelovej funkcie $c^T x$. Z toho dôvodu môžeme vo funkcii Υ_ρ položiť $\lambda = b^T y$ a následne použiť vzťah $x^T s = c^T x - b^T y$, čím výraz $c^T x - \lambda$ nahradíme výrazom $x^T s$. Druhý člen funkcie Φ_ρ vznikol pridaním sčítanca $-\sum_{i=1}^n \ln s_i$, ktorý reprezentuje bariérový prvok zodpovedajúci ohraňeniu $s \geq 0$.



Obr. 2.1: Vrstevnice funkcie Φ_ρ pre $n = 2$ vykreslené v priestore $x_1s_1 \times x_2s_2$ v závislosti od parametra ρ .

Funkciu Φ_ρ môžeme vyjadriť v tvare

$$\Phi_\rho(x, s) = (\rho - n) \ln x^T s + \Phi_n(x, s) \quad (2.3)$$

$$= (\rho - n) \ln x^T s - \sum_{i=1}^n \ln \left(\frac{x_i s_i}{x^T s / n} \right) + n \ln n. \quad (2.4)$$

Z toho vyplýva, že potenciálová funkcia Φ_ρ má dve dôležité vlastnosti:

1. $\Phi_\rho \rightarrow \infty$ ak $x_i s_i \rightarrow 0$ pre nejaké $i = 1, 2, \dots, n$, ale $\mu = x^T s / n \rightarrow 0$,
2. $\Phi_\rho \rightarrow -\infty$ práve vtedy, keď $(x, y, s) \rightarrow \mathcal{F}^*$.

Prvá vlastnosť hovorí o tom, že funkcia Φ_ρ pôsobí ako bariérová funkcia, ak mechanizmus smeruje k ľubovoľnému bodu (x, y, s) na hranici množiny \mathcal{F}° ($x_i s_i = 0$), ktorý ale nie je optimálnym riešením ($x^T s > 0$). V tomto prípade prvý člen výrazu (2.4) zostáva ohraničený, zatiaľ čo druhý člen nie, a preto $\Phi_\rho \rightarrow \infty$. Druhá vlastnosť formuluje vzťah medzi funkciou Φ_ρ a množinou \mathcal{F}^* . Ak mechanizmus algoritmu spôsobuje $\Phi_\rho \rightarrow -\infty$, potom postupnosť iteračných bodov konverguje k optimálnemu riešeniu. Lema 2.2 do-

kazuje, že člen $\Phi_n(x, s)$ vo výraze (2.3) je zdola ohraňovaný, a preto $\Phi_\rho \rightarrow -\infty$ práve vtedy, keď $(\rho - n) \ln x^T s \rightarrow -\infty$, teda keď $\mu \rightarrow 0$. Toto pozorovanie je podstatou algoritmu redukcie potenciálu, ktorý generuje postupnosť ostro prípustných iteračných bodov (x^k, y^k, s^k) z množiny \mathcal{F}° , pre ktoré $\Phi_\rho(x^k, s^k) \rightarrow -\infty$, čím sa duálna medzera znižuje, a tým mechanizmus algoritmu smeruje k optimalite.

Existuje niekoľko algoritmov [3], [6], [30] založených na potenciálovej funkcii Φ_ρ . V ďalších častiach tejto kapitoly špecifikujeme primárno-duálny algoritmus redukcie potenciálu, ako ho prezentovali Kojima, Mizuno a Yoshise [12]. Uvedieme jeho základné vlastnosti a vysvetlíme dôležitý vzťah medzi funkciou Φ_ρ a parametrom μ , ktorý je kľúčovým elementom konvergenie a polynomiálnej zložitosti.

2.2 Algoritmus

Algoritmus redukcie potenciálu je špeciálnym prípadom generického algoritmu, ktorý pristupuje k zmenšovaniu duálnej medzery nepriamo. Tento algoritmus volí v každej iterácii konštantnú hodnotu centrujúceho parametra $\sigma_k = n/\rho$, kde $\rho > n$. Smer k ďalšiemu iteračnému bodu je štandardným Newtonovým smerom, ktorý vypočítame zo systému (1.11). Algoritmus volí dĺžku kroku α_k tak, aby minimalizovala funkciu Φ_ρ pozdĺž získaného smeru, pričom v každej iterácii je zároveň nutné zachovať ostrú prípustnosť. Práve kvôli podmienke ostrej prípustnosti definujeme hornú hranicu pre dĺžku kroku z bodu $(x^k, y^k, s^k) \in \mathcal{F}^\circ$ pozdĺž smeru $(\Delta x, \Delta y, \Delta s)$

$$\alpha_{\max} = \sup \{ \alpha \in [0, 1] \mid (x, s) + \alpha(\Delta x, \Delta s) > 0 \}.$$

Schéma algoritmu redukcie potenciálu

vstup $\rho > n, (x^0, y^0, s^0) \in \mathcal{F}^\circ, \epsilon > 0$

prirad' $k := 0$

repeat

prirad' $\sigma_k := n/\rho, \mu_k := (x^k)^T s^k / n$

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $\alpha_{\max} := \sup \{ \alpha \in [0, 1] \mid (x^k, s^k) + \alpha(\Delta x, \Delta s) > 0 \}$

prirad' $\alpha_k := \arg \min_{\alpha \in (0, \alpha_{\max})} \Phi_\rho(x^k + \alpha \Delta x, s^k + \alpha \Delta s)$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k(\Delta x, \Delta y, \Delta s)$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \epsilon$

2.3 Redukcia funkcie Φ_ρ a konvergencia

Analýza v tejto časti ukazuje, že hodnota funkcie Φ_ρ je redukovaná aspoň konštantnou mierou v každej iterácii. Dokážeme, že ak $\Phi_\rho \rightarrow -\infty$, potom $\mu \rightarrow 0$. Nasledujúce pomocné tvrdenie je iba technickým výsledkom, a preto dôkaz tohto tvrdenia úvadžame v Dodatku.

Lema 2.1.

(i) Pre všetky $\beta > -1$ platí $\ln(1 + \beta) \leq \beta$, pričom rovnosť nastáva práve vtedy, keď $\beta = 0$.

(ii) Pre ľubovoľný vektor $z \in \mathbf{R}^n$, $\|z\|_\infty \leq \tau < 1$ platí

$$-\sum_{i=1}^n \ln(1 + z_i) \leq -e^T z + \frac{\|z\|^2}{2(1 - \tau)}. \quad [29]$$

Nasledujúci výsledok dokazuje, že funkcia Φ_n je zdola ohraničená.

Lema 2.2. Pre $(x, s) > 0$ platí $\Phi_n(x, s) \geq n \ln n$, pričom rovnosť nastáva práve vtedy, keď $XSe = (x^T s/n)e = \mu e$. [29]

Dôkaz. Zo vzťahov (2.3), (2.4) a pomocou Lemy 2.1(i) odvodíme, že

$$\Phi_n(x, s) - n \ln n = -\sum_{i=1}^n \ln\left(\frac{x_i s_i}{x^T s/n}\right) \geq \sum_{i=1}^n \left(\frac{x_i s_i}{\mu} - 1\right) = -(n - n) = 0.$$

Z Lemy 2.1(i) zároveň vyplýva, že rovnosť nastáva práve vtedy, keď $x_i s_i / \mu = 1$, $i = 1, 2, \dots, n$ ($\Leftrightarrow XSe = \mu e$). \square

Ďalšie tvrdenie dokazuje, že funkcia Φ_ρ nie je zdola ohraničená na svojom definičnom obore a zároveň uvádza kľúčový vzťah medzi funkciou Φ_ρ a parametrom μ .

Lema 2.3.

(i) Funkcia Φ_ρ je zdola neohraničená na svojom definičnom obore.

(ii) Pre ľubovoľný bod $(x, y, s) \in \mathcal{F}^\circ$ platí

$$\mu \leq \exp\left(\frac{\Phi_\rho(x, s)}{\rho - n}\right), \text{ kde } \mu = \frac{x^T s}{n}. \quad [29] \quad (2.5)$$

Dôkaz. Z existencie centrálnej trajektórie \mathcal{C} vieme, že pre $\mu > 0$ existuje ostro prípustný bod $(x_\mu, y_\mu, s_\mu) \in \mathcal{F}^\circ$ taký, že $(x_\mu)_i (s_\mu)_i = \mu$ pre $i = 1, 2, \dots, n$. Vyčíslením funkcie Φ_ρ v bode (x_μ, s_μ) dostaneme

$$\begin{aligned} \Phi_\rho(x_\mu, s_\mu) &= (\rho - n) \ln x_\mu^T s_\mu + \Phi_n(x_\mu, s_\mu) \\ &= (\rho - n) \ln(n\mu) + n \ln n, \end{aligned}$$

z čoho vyplýva, že $\Phi_\rho(x_\mu, s_\mu) \rightarrow -\infty$ pre $\mu \rightarrow 0$. Pre dôkaz časti (ii) použijeme vzťahy (2.3), (2.4) a Lemu 2.2. Dostaneme

$$\begin{aligned}
 \Phi_\rho(x, s) &= (\rho - n) \ln x^T s + \Phi_n(x, s) \\
 &\geq (\rho - n) \ln \mu + (\rho - n) \ln n + n \ln n \\
 &\geq (\rho - n) \ln \mu,
 \end{aligned}$$

z čoho priamo vyplýva požadovaný výsledok (2.5). \square

Z uvedených pozorovaní vyplýva, že ak je možné generovať postupnosť iteračných bodov $(x^k, y^k, s^k) \in \mathcal{F}^\circ$, pre ktoré $\Phi_\rho(x^k, s^k) \rightarrow -\infty$, potom $\mu_k \rightarrow 0$. Na záver tejto časti dokážeme, že funkcia Φ_ρ je v každej iterácii redukovaná aspoň konštantnou mierou $\delta > 0$ nezávislou od n . V matematickom kontexte to znamená, že

$$\Phi_\rho(x^{k+1}, s^{k+1}) \leq \Phi_\rho(x^k, s^k) - \delta, \quad k = 0, 1, 2, \dots \quad (2.6)$$

Veta 2.1. *Nech $(x^0, y^0, s^0) \in \mathcal{F}^\circ$ je štartovacím bodom algoritmu. Predpokladajme, že algoritmus generuje postupnosť iteračných bodov $(x^k, y^k, s^k) \in \mathcal{F}^\circ$, ktoré spĺňajú nerovnicu (2.6) pre nejakú kladnú konštantu $\delta > 0$. Potom pre ľubovoľné $\varepsilon \in (0, 1)$ existuje index*

$$\bar{k} = \left\lceil \frac{\Phi_\rho(x^0, s^0)}{\delta} + \frac{\rho - n}{\delta} |\ln \varepsilon| \right\rceil \quad (2.7)$$

taký, že $\mu_k \leq \varepsilon, \forall k \geq \bar{k}$. [29]

Dôkaz. Zlogaritmovaním oboch strán nerovnice (2.5) zistíme, že kritérium $\mu_k \leq \varepsilon$ je priamym dôsledkom nerovnice

$$\Phi_\rho(x^k, s^k) \leq (\rho - n) \ln \varepsilon = -(\rho - n) |\ln \varepsilon|. \quad (2.8)$$

Zo vzťahu (2.6) zároveň vyplýva, že

$$\Phi_\rho(x^k, s^k) \leq \Phi_\rho(x^0, s^0) - k\delta, \quad k = 1, 2, \dots,$$

a preto nerovnica (2.8) platí, ak

$$\Phi_\rho(x^0, s^0) - k\delta \leq -(\rho - n) |\ln \varepsilon|, \quad k = 0, 1, 2, \dots,$$

z čoho je možné elementárnymi úpravami odvodiť explicitný vzorec (2.7) pre \bar{k} . \square

2.4 Odhad funkcie Φ_ρ pozdĺž prípustného smeru

V tejto časti sa zameriame na tú časť algoritmu, ktorá voľbou dĺžky kroku α minimalizuje funkciu Φ_ρ pozdĺž smeru $(\Delta x, \Delta y, \Delta s)$. Pokúsime sa odvodiť kvadratickú funkciu q ako funkciu dĺžky kroku α pozdĺž daného smeru $(\Delta x, \Delta y, \Delta s)$, ktorá zodpovedá hornému ohraničeniu funkcie Φ_ρ . Zdôraznime, že tento odhad je platný iba na intervale $(0, \alpha_\tau] \subset (0, \alpha_{\max})$, kde

$$\alpha_\tau \max(\|X^{-1}\Delta x\|_\infty, \|S^{-1}\Delta s\|_\infty) = \tau, \quad \tau \in (0, 1). \quad (2.9)$$

Poznamenajme, že mechanizmus algoritmu je sústredený v množine ostro prípustných riešení \mathcal{F}° , a preto má zmysel definovať matice X^{-1} , S^{-1} . Kvadratický odhad q funkcie Φ_ρ pozdĺž smeru $(\Delta x, \Delta y, \Delta s)$ získame na základe Lemu 2.1. Z definície funkcie Φ_ρ a zo skutočnosti, že vektory Δx a Δs sú navzájom ortogonálne (Lema 1.1), odvodíme, že

$$\begin{aligned}
 & \Phi_\rho(x + \alpha\Delta x, s + \alpha\Delta s) - \Phi_\rho(x, s) \\
 &= \rho \ln \left[\frac{(x + \alpha\Delta x)^T (s + \alpha\Delta s)}{x^T s} \right] - \sum_{i=1}^n \ln \left(\frac{x_i + \alpha\Delta x_i}{x_i} \right) - \sum_{i=1}^n \ln \left(\frac{s_i + \alpha\Delta s_i}{s_i} \right) \\
 &= \rho \ln \left(1 + \alpha \frac{x^T \Delta s + s^T \Delta x + \alpha \overbrace{(\Delta x)^T \Delta s}^{=0}}{x^T s} \right) - \sum_{i=1}^n \ln \left(1 + \alpha \frac{\Delta x_i}{x_i} \right) - \sum_{i=1}^n \ln \left(1 + \alpha \frac{\Delta s_i}{s_i} \right) \\
 &= \rho \ln \left(1 + \alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} \right) - \sum_{i=1}^n \ln \left(1 + \alpha \frac{\Delta x_i}{x_i} \right) - \sum_{i=1}^n \ln \left(1 + \alpha \frac{\Delta s_i}{s_i} \right). \tag{2.10}
 \end{aligned}$$

Na prvý sčítanec vo výraze (2.10) aplikujeme Lemu 2.1(i), na zvyšné členy aplikujeme Lemu 2.1(ii). Dostaneme, že pre všetky $\alpha \in (0, \alpha_\tau]$ platí

$$\begin{aligned}
 & \Phi_\rho(x + \alpha\Delta x, s + \alpha\Delta s) \\
 & \leq \Phi_\rho(x, s) + \rho\alpha \frac{x^T \Delta s + s^T \Delta x}{x^T s} - \alpha e^T (X^{-1} \Delta x + S^{-1} \Delta s) + \alpha^2 \frac{\|X^{-1} \Delta x\|^2 + \|S^{-1} \Delta s\|^2}{2(1 - \tau)} \\
 & = \Phi_\rho(x, s) + \alpha\xi_1 + \frac{1}{2}\alpha^2\xi_2, \tag{2.11}
 \end{aligned}$$

kde

$$\xi_1 = \rho \frac{x^T \Delta s + s^T \Delta x}{x^T s} - e^T (X^{-1} \Delta x + S^{-1} \Delta s), \tag{2.12}$$

$$\xi_2 = \frac{1}{1 - \tau} (\|X^{-1} \Delta x\|^2 + \|S^{-1} \Delta s\|^2). \tag{2.13}$$

Ak definujeme funkciu $q : \alpha \rightarrow \mathbf{R}$ v tvare

$$q(\alpha) = \Phi_\rho(x, s) + \alpha\xi_1 + \frac{1}{2}\alpha^2\xi_2, \tag{2.14}$$

tak z nerovnice (2.11) dostaneme horný odhad funkcie Φ_ρ pozdĺž smeru $(\Delta x, \Delta y, \Delta s)$

$$\Phi_\rho(x + \alpha\Delta x, s + \alpha\Delta s) \leq q(\alpha), \quad \forall \alpha \in (0, \alpha_\tau].$$

Analyzujme bližšie koeficienty ξ_1 , ξ_2 vo funkcii q . Podľa (2.13) je zrejmé, že $\xi_2 > 0$, a preto q je konvexnou kvadratickou funkciou, a teda a ak má stacionárny bod, tak v ňom funkcia q dosahuje svoje minimum. Ďalej ukážeme, že $\xi_1 < 0$, čo implikuje, že funkcia q klesá na intervale $(0, \hat{\alpha})$, kde $\hat{\alpha}$ je bod, v ktorom funkcia q dosahuje svoje minimum.

Pripomeňme, že $D = X^{1/2}S^{-1/2}$ a zaved' me nové pomocné označenie

$$\begin{aligned}
 V &= (XS)^{1/2}, & v &= Ve = [(x_i s_i)^{1/2}]_{i=1}^n, \\
 v_{\min} &= \min_{i=1,2,\dots,n} v_i, & r &= -v + \frac{n}{\rho} \mu V^{-1} e.
 \end{aligned}$$

Uvedieme pomocné vzťahy, ktoré budú užitočné v ďalších častiach

$$\|v\|^2 = x^T s = n\mu, \quad X = VD, \quad S = VD^{-1}. \quad (2.15)$$

Na základe zavedeného označenia teraz môžeme rovnicu $S\Delta x + X\Delta s = -XSe + \sigma\mu e$, kde $\sigma = n/\rho$, zapísať v troch ekvivalentných tvaroch

$$\begin{aligned}
 S\Delta x + X\Delta s &= Vr \\
 D^{-1}\Delta x + D\Delta s &= r \\
 X^{-1}\Delta x + S^{-1}\Delta s &= V^{-1}r.
 \end{aligned} \quad (2.16)$$

Keďže podľa Lemy 1.1 $(\Delta x)^T \Delta s = 0$, tak z rovnice (2.16) môžeme odvodiť, že

$$\begin{aligned}
 \|r\|^2 &= \|D^{-1}\Delta x + D\Delta s\|^2 \\
 &= \|D^{-1}\Delta x\|^2 + \underbrace{(\Delta x)^T D^{-1} D \Delta s}_{=0} + \|D\Delta s\|^2 \\
 &= \|D^{-1}\Delta x\|^2 + \|D\Delta s\|^2,
 \end{aligned} \quad (2.17)$$

a preto $\|D^{-1}\Delta x\| \leq \|r\|$ a $\|D\Delta s\| \leq \|r\|$. Zo vzťahov (2.15) a (2.17) dostaneme

$$\begin{aligned}
 \|X^{-1}\Delta x\|^2 + \|S^{-1}\Delta s\|^2 &= \|V^{-1}D^{-1}\Delta x\|^2 + \|V^{-1}D\Delta s\|^2 \\
 &\leq \|V^{-1}\|^2 (\|D^{-1}\Delta x\|^2 + \|D\Delta s\|^2) \\
 &\leq \frac{1}{v_{\min}^2} \|r\|^2.
 \end{aligned} \quad (2.18)$$

Z toho vyplýva, že pre koeficient ξ_2 definovaný podľa (2.13) platí

$$\xi_2 \leq \frac{\|r\|^2}{(1-\tau)v_{\min}^2}. \quad (2.19)$$

Pre koeficient ξ_1 definovaný podľa (2.12) dostaneme

$$\begin{aligned}
 \xi_1 &= \frac{\rho}{x^T s} e^T \underbrace{(X\Delta s + S\Delta x)}_{=Vr} - e^T \underbrace{(X^{-1}\Delta x + S^{-1}\Delta s)}_{=V^{-1}r} \\
 &= \frac{\rho}{x^T s} e^T Vr - e^T V^{-1}r \\
 &= -\frac{\rho}{n\mu} \underbrace{\left(-v + \frac{n}{\rho} \mu V^{-1} e\right)^T}_{=r^T} r \\
 &= -\frac{\rho}{n\mu} \|r\|^2.
 \end{aligned} \quad (2.20)$$

Vzťahy (2.19) a (2.20) uvádzajú odhady koeficientov ξ_1, ξ_2 v závislosti od $\|r\|^2$. Keďže $\rho, n, \mu > 0$, tak je zrejmé, že $\xi_1 \leq 0$, pričom $\xi_1 = 0$ práve vtedy, keď $r = 0$. V ďalšej časti dokážeme, že $r \neq 0$, z čoho bezprostredne vyplýva, že $\xi_1 < 0$.

2.5 Odhad redukcie funkcie Φ_ρ

Nasledujúce tvrdenie uvádza dolné ohraňenie pre $\|r\|$.

Lema 2.4. *Pre ľubovoľný bod $(x, y, s) \in \mathcal{F}^\circ$ a pre $\rho > n + \sqrt{n}$ platí*

$$\|r\| \geq \frac{\sqrt{3n\mu}}{2\rho v_{\min}}. \quad [29] \quad (2.21)$$

Dôkaz. Na základe zavedeného označenia dostaneme

$$\begin{aligned} \frac{\rho^2}{n^2\mu^2}\|r\|^2 &= \left\| \frac{\rho}{n\mu}v - V^{-1}e \right\|^2 \\ &= (V^{-1}e)^T(V^{-1}e) - 2\frac{\rho}{n\mu}v^T V^{-1}e + \frac{\rho^2}{n^2\mu^2}v^T v \\ &= (V^{-1}e)^T(V^{-1}e) - 2\frac{\rho}{\mu} + \frac{\rho^2}{n\mu} \\ &= (V^{-1}e)^T(V^{-1}e) + \frac{\rho^2 - 2n\rho + n^2 - n}{n\mu} - \frac{n^2 - n}{n\mu}. \end{aligned} \quad (2.22)$$

Keďže $\rho^2 - 2n\rho + n^2 - n = (\rho - n - \sqrt{n})^2 + 2\sqrt{n}(\rho - n - \sqrt{n})$, tak je zrejmé, že výraz $\frac{\rho^2 - 2n\rho + n^2 - n}{n\mu}$ v rovnici (2.22) je nezáporný pre $\rho \geq n + \sqrt{n}$ a nulový pre $\rho = n + \sqrt{n}$. Z toho dôvodu platí

$$\frac{\rho^2}{n^2\mu^2}\|r\|^2 \geq (V^{-1}e)^T(V^{-1}e) - \frac{n^2 - n}{n\mu}, \quad (2.23)$$

pričom rovnosť nastáva, ak $\rho = n + \sqrt{n}$. V ďalšej časti dôkazu využijeme technický výsledok, že vektory v a $V^{-1}e - v/\mu$ sú navzájom ortogonálne, čo znamená, že $v^T(V^{-1}e - v/\mu) = 0$. Dôkaz tohto výsledku je uvedený v Dodatku. Dosadením $\rho = n + \sqrt{n}$ do nerovnice (2.23) a použitím definície vektora $r = -v + n\mu V^{-1}e/\rho$ dostaneme

$$\begin{aligned} (V^{-1}e)^T(V^{-1}e) - \frac{n^2 - n}{n\mu} &= \frac{\rho^2}{n^2\mu^2}\|r\|^2 \\ &= \frac{(n + \sqrt{n})}{n^2\mu^2} \left\| \frac{n\mu}{n + \sqrt{n}}V^{-1}e - v \right\|^2 \\ &= \left\| V^{-1}e - \frac{n + \sqrt{n}}{n\mu}v \right\|^2 \\ &= \left\| V^{-1}e - \frac{1}{\mu}v \right\|^2 - 2\underbrace{\frac{\sqrt{n}}{n\mu}v^T \left(V^{-1}e - \frac{1}{\mu}v \right)}_{=0} + \left\| \frac{\sqrt{n}}{n\mu}v \right\|^2 \\ &= \left\| V^{-1}e - \frac{1}{\mu}v \right\|^2 + \left\| \frac{\sqrt{n}}{n\mu}v \right\|^2. \end{aligned} \quad (2.24)$$

Prvý sčítanec v rovnici (2.24) môže byť zdola ohaničený štvorcom ľubovoľnej zložky vektora $V^{-1}e - v/\mu$, a preto

$$\left\| V^{-1}e - \frac{1}{\mu}v \right\|^2 \geq \left(\frac{1}{v_{\min}} - \frac{v_{\min}}{\mu} \right)^2.$$

Pomocou tohto ohraničenia môžeme rovnicu (2.24) ďalej rozpísať do tvaru

$$\begin{aligned} \frac{\rho^2}{n^2\mu^2} \|r\|^2 &\geq \left(\frac{1}{v_{\min}} - \frac{v_{\min}}{\mu} \right)^2 + \frac{1}{n\mu^2} v^T v \\ &= \frac{1}{\mu^2} \left[\left(\frac{\mu}{v_{\min}} - v_{\min} \right)^2 + \mu \right] \\ &= \frac{1}{\mu^2} \left[\left(\frac{\mu}{2v_{\min}} - v_{\min} \right)^2 + \frac{3\mu^2}{4v_{\min}^2} \right] \\ &\geq \frac{3}{4v_{\min}^2}, \end{aligned}$$

z čoho priamo vyplýva výsledok (2.21). \square

Na záver tejto časti ukážeme, že existuje taká dĺžka kroku $\bar{\alpha}$, ktorá vedie ku konštantnej (fixnej) redukcii funkcie q a Φ_ρ . Pre jednoduchosť a bez ujmy na všeobecnosti použijeme konkrétnu hodnotu $\tau = 0.5$, pričom analogicky je možné dokázať podobný výsledok pre ľubovoľnú hodnotu parametra τ z intervalu $(0, 1)$.

Veta 2.2. *Nech $\tau = 0.5$. Definujme $\bar{\alpha} = \tau v_{\min} / \|r\|$. Potom $q(\bar{\alpha}) - q(0) \leq -0.15$, a tak spôsob redukcie (2.6) platí s konštantou $\delta = 0.15$. [29]*

Dôkaz. Najprv overíme, že dĺžka kroku $\bar{\alpha}$ vyhovuje podmienke $0 < \bar{\alpha} \leq \alpha_\tau$, kde α_τ je dĺžka kroku definovaná podľa (2.9). Keďže $\tau, v_{\min}, \|r\| > 0$, tak zrejme platí $\bar{\alpha} > 0$. Predpokladajme, že $\|X^{-1}\Delta x\|_\infty \geq \|S^{-1}\Delta s\|_\infty$. Z nerovnice (2.18) vieme, že

$$\|X^{-1}\Delta x\|_2 \leq \frac{\|r\|}{v_{\min}}, \quad \|S^{-1}\Delta s\|_2 \leq \frac{\|r\|}{v_{\min}},$$

a keďže $\|\cdot\|_\infty \leq \|\cdot\|_2$, tak

$$\bar{\alpha} \|X^{-1}\Delta x\|_\infty \leq \bar{\alpha} \|X^{-1}\Delta x\|_2 \leq \frac{\tau v_{\min}}{\|r\|} \frac{\|r\|}{v_{\min}} = \tau = \alpha_\tau \|X^{-1}\Delta x\|_\infty.$$

Z toho priamo vyplýva, že $\bar{\alpha} \leq \alpha_\tau$. Analogicky môžeme odvodiť rovnaký výsledok, ak $\|X^{-1}\Delta x\|_\infty \leq \|S^{-1}\Delta s\|_\infty$. Na základe odhadov (2.19) a (2.20) koeficientov ξ_1, ξ_2 odvodíme, že

$$\begin{aligned} q(\bar{\alpha}) - q(0) &= \bar{\alpha}\xi_1 + \frac{1}{2}\bar{\alpha}^2\xi_2 \\ &\leq -\frac{\rho}{n\mu}\bar{\alpha}\|r\|^2 + \frac{1}{2}\bar{\alpha}^2\frac{\|r\|^2}{v_{\min}^2(1-\tau)} \end{aligned}$$

$$\begin{aligned}
 &\leq -\frac{\rho}{n\mu} \frac{\tau v_{\min}}{\|r\|^2} \|r\|^2 + \frac{1}{2} \left(\frac{\tau v_{\min}}{\|r\|^2} \right)^2 \frac{\|r\|^2}{v_{\min}^2 (1-\tau)} \\
 &= -\frac{\rho}{n\mu} \tau v_{\min} \|r\| + \frac{\tau^2}{2(1-\tau)}.
 \end{aligned}$$

Použitím dolného ohraničenia (2.21) pre $\|r\|$ a dosadením $\tau = 0.5$ dostaneme

$$\begin{aligned}
 q(\bar{\alpha}) - q(0) &\leq -\frac{\rho}{n\mu} \tau v_{\min} \left(\frac{\sqrt{3}n\mu}{2\rho v_{\min}} \right) + \frac{\tau^2}{2(1-\tau)} \\
 &= -\frac{\sqrt{3}\tau}{2} + \frac{\tau^2}{2(1-\tau)} \\
 &= -\frac{1-\sqrt{3}}{4} \leq -0.15.
 \end{aligned}$$

Podľa definície funkcie q a zo vzťahu (2.11) nakoniec odvodíme, že

$$\begin{aligned}
 \min_{\alpha \in (0, \alpha_\tau)} \Phi_\rho(x + \alpha \Delta x, s + \alpha \Delta s) &\leq \Phi_\rho(x + \bar{\alpha} \Delta x, s + \bar{\alpha} \Delta s) \\
 &\leq q(\bar{\alpha}) \leq q(0) - 0.15 = \Phi_\rho(x, s) - 0.15.
 \end{aligned}$$

Z toho vyplýva, že spôsob redukcie (2.6) platí s $\delta = 0.15$. □

2.6 Polynomiálna zložitosť

Na záver tejto kapitoly sformulujeme tvrdenie, ktoré dokazuje polynomiálnu zložitosť algoritmu redukcie potenciálu.

Veta 2.3. *Nech $\rho \geq n + \sqrt{n} a \varepsilon > 0$. Predpokladajme, že štartovací bod $(x^0, y^0, s^0) \in \mathcal{F}^\circ$ spĺňa okrem podmienky (2.6) aj podmienku $\Phi_\rho(x^0, s^0) \leq \kappa(\rho - n) |\ln \varepsilon|$ pre nejakú kladnú konštantu $\kappa > 0$ nezávislú od n . Potom existuje index*

$$\bar{k} = \left\lceil \frac{\kappa + 1}{\delta} (\rho - n) |\ln \varepsilon| \right\rceil = O((\rho - n) |\ln \varepsilon|)$$

taký, že $(x^k, y^k, s^k) \in \mathcal{F}^\circ$, $\mu_k \leq \varepsilon$, $\forall k \geq \bar{k}$. [29]

Dôkaz. Polynomiálna zložitosť je priamym dôsledkom Vety 2.1, ak v explicitnom vzorci (2.7) použijeme ohraničenie $\Phi_\rho(x^0, s^0) \leq \kappa(\rho - n) |\ln \varepsilon|$. □

Predpoklady polynomiálnej zložitosti z Vety 2.3 vedú prirodzene k otázke vhodnej voľby parametra ρ . Kojima, Mizuno a Yoshise [12] vo svojom pôvodnom algoritme zvolili $\rho = n + \sqrt{n}$, čo podľa Vety 2.3 vedie k zložitosti $O(\sqrt{n} |\ln \varepsilon|)$. Na druhej strane, voľbou $\rho = n + \sqrt{n}$ pre $n \gg 1$ získame hodnoty centrujúceho parametra σ blízke 1, čo podľa Lemy 1.2 spôsobuje menšiu redukciu duálnej medzery. Menšie hodnoty parametra σ môžeme získať voľbou $\rho = 10n$ alebo $\rho = n + n^{3/2}$, čo síce znižuje odhad polynomiálnej zložitosti na $O(n |\ln \varepsilon|)$, resp. $O(n^{3/2} |\ln \varepsilon|)$, no v praxi sa táto voľba parametra ρ osvedčila a vykazuje lepšie numerické výsledky. [29]

ALGORITMY SLEDOVANIA CENTRÁLNEJ TRAJEKTÓRIE

V úvodnej kapitole sme uviedli, že ak základné predpoklady (P1) a (P2) sú splnené, potom systém (1.4)-(1.6) definujúci centrálnu trajektóriu \mathcal{C} má jediné riešenie (x_μ, y_μ, s_μ) pre ľubovoľné $\mu > 0$. Zároveň sme uviedli výsledok, podľa ktorého riešenia (x_μ, y_μ, s_μ) konvergujú pre $\mu \rightarrow 0$ k primárno-duálnemu optimálnemu riešeniu. Analytické riešenie nelineárneho systému (1.4)-(1.6) nie je možné, a preto je tento prístup nevhodný pre praktické implementácie. Keďže presné riešenia (x_μ, y_μ, s_μ) konvergujú k optimálnemu riešeniu, tak sa môžeme domnievať, že aj postupnosť približných riešení konverguje. Ak by sme pre nejaké $\mu > 0$ dokázali identifikovať bod v dobre definovanom okolí presného riešenia (x_μ, y_μ, s_μ) na centrálnej trajektórii, potom by sme zrejme vedeli nájsť ďalší bod blízko presného riešenia $(x_{\bar{\mu}}, y_{\bar{\mu}}, s_{\bar{\mu}})$ zodpovedajúce parametru $\bar{\mu} < \mu$. Túto myšlienku by sme mohli použiť opakovane, čím by sme získali približné riešenia systému (1.4)-(1.6), ktoré by pozdĺž centrálnej trajektórie \mathcal{C} v smere znižovania parametra μ smerovali do množiny \mathcal{F}^* . Táto myšlienka je podstatou algoritmov sledovania centrálnej trajektórie. Ich spoločným charakteristickým znakom je princíp sledovania centrálnej trajektórie v jej dobre definovanom špecifickom okolí, v ktorom iteračné body algoritmu konvergujú k primárno-duálnemu optimálnemu riešeniu. K tomu, aby iteračné body ležali v špecifickom okolí, je nevyhnutné merať ich vzdialenosť od centrálnej trajektórie. To znamená, že na vyriešenie pôvodnej dvojice úloh (P), (D) je postačujúce nájsť len približné riešenia systému (1.4)-(1.6). Z toho dôvodu môžeme tieto algoritmy interpretovať ako voľné sledovanie centrálnej trajektórie, kde centrálna trajektória slúži ako navigátor určujúci smer optimalizácie z množiny \mathcal{F}° do množiny \mathcal{F}^* .

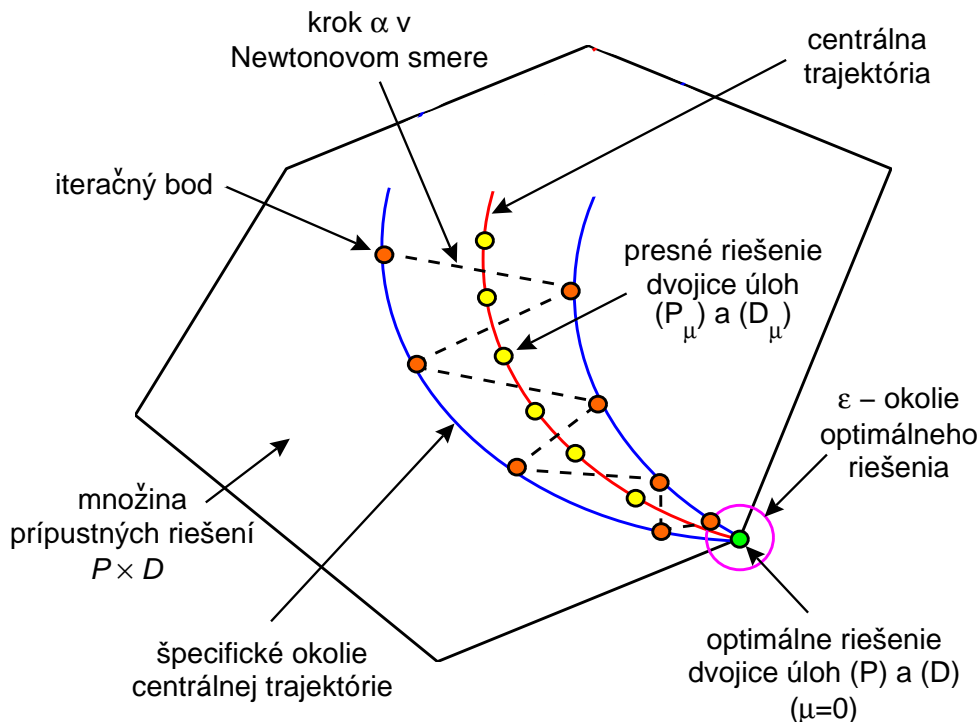
3.1 Špecifické okolie centrálnej trajektórie

V súvislosti s generovaním približných riešení systému (1.4)-(1.6) by sme mali byť schopní definovať a merať ich presnosť. Ak bod (x, y, s) aproximuje presné riešenie (x_μ, y_μ, s_μ) pre nejaké $\mu > 0$, potom presnosť aproximácie vieme sledovať štandardne pomocou nejakej normy

$$\|(x, y, s) - (x_\mu, y_\mu, s_\mu)\|,$$

na základe ktorej môžeme definovať množinu približných riešení

$$\{(x, y, s) \in \mathcal{F}^\circ \mid \|(x, y, s) - (x_\mu, y_\mu, s_\mu)\| \leq \omega\}, \text{ pre nejaké } \omega \geq 0.$$



Obr. 3.1: Ilustrácia primárno-duálneho algoritmu sledovania centrálnej trajektórie.

Avšak ako sme už uviedli, nájsť analytické riešenie (x_μ, y_μ, s_μ) nie je možné, a preto by sme museli riešiť systém (1.4)-(1.6) numericky. Namiesto toho využijeme implicitný popis riešenia cez systém (1.4)-(1.6), čím získame množiny, ktorých charakteristika implikuje presnosť aproximácie. Takéto množiny nazývame špecifickým okolím centrálnej trajektórie. V moderných metódach vnútorného bodu sa štandardne používa dvojica špecifických okolí

$$\mathcal{N}_2(\theta) = \{(x, y, s) \in \mathcal{F}^\circ \mid \|XSe - \mu e\|_2 \leq \theta\mu, \mu = x^T s/n\}, \text{ pre dané } \theta \in (0, 1),$$

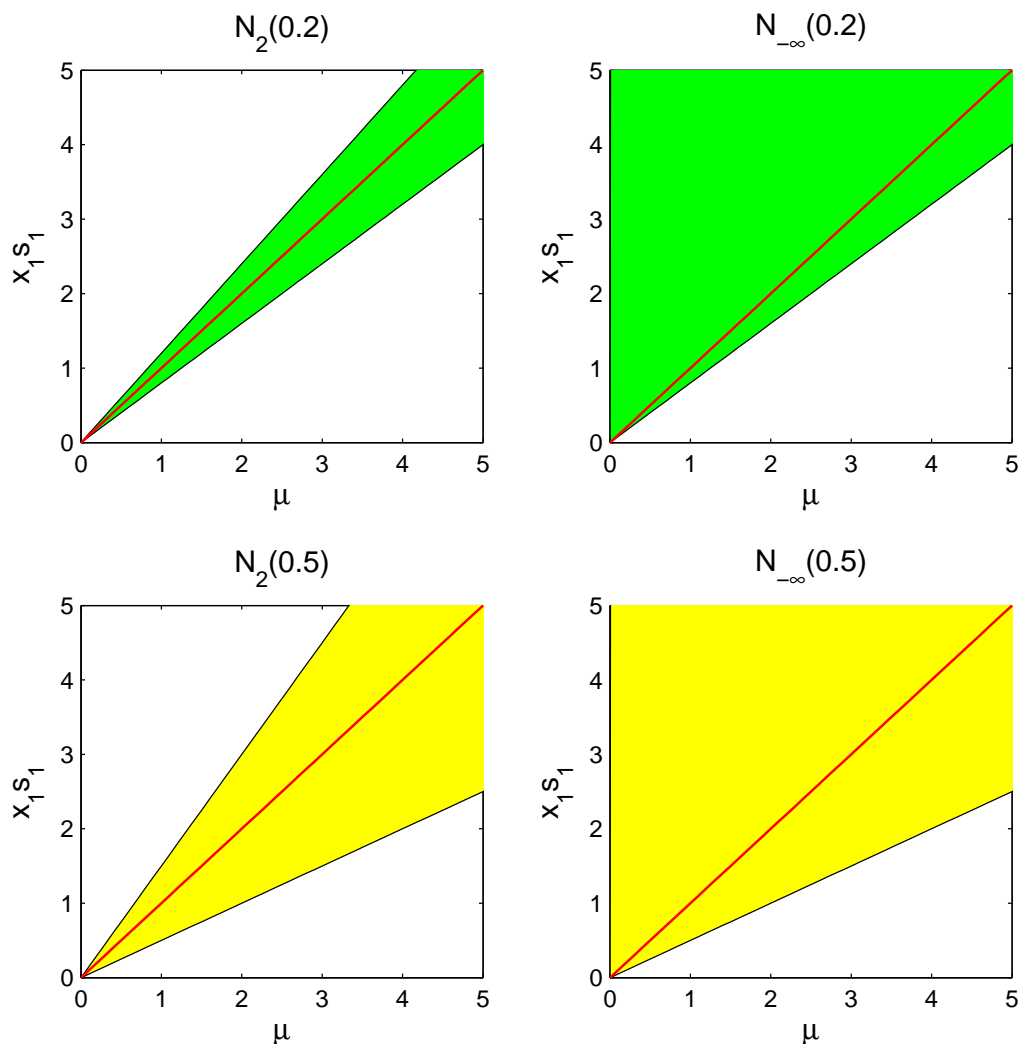
$$\mathcal{N}_{-\infty}(\gamma) = \{(x, y, s) \in \mathcal{F}^\circ \mid x_i s_i \geq \gamma\mu, i = 1, 2, \dots, n, \mu = x^T s/n\}, \text{ pre dané } \gamma \in (0, 1).$$

Ak pre bod (x, y, s) platí $\|XSe - \mu e\|_2 \leq \theta\mu$ a $\mu = x^T s/n$, potom hovoríme, že bod (x, y, s) spĺňa podmienku presnosti pre okolie $\mathcal{N}_2(\theta)$. Analogicky hovoríme o podmienke presnosti pre okolie $\mathcal{N}_{-\infty}(\gamma)$. Analyzujme bližšie obe špecifické okolia. V prípade špecifického okolia $\mathcal{N}_2(\theta)$ pre podmienku presnosti zrejme platí

$$\|XSe - \mu e\|_2 \leq \theta\mu \Leftrightarrow \sum_{i=1}^n \left(\frac{x_i s_i}{\mu} - 1 \right)^2 \leq \theta^2.$$

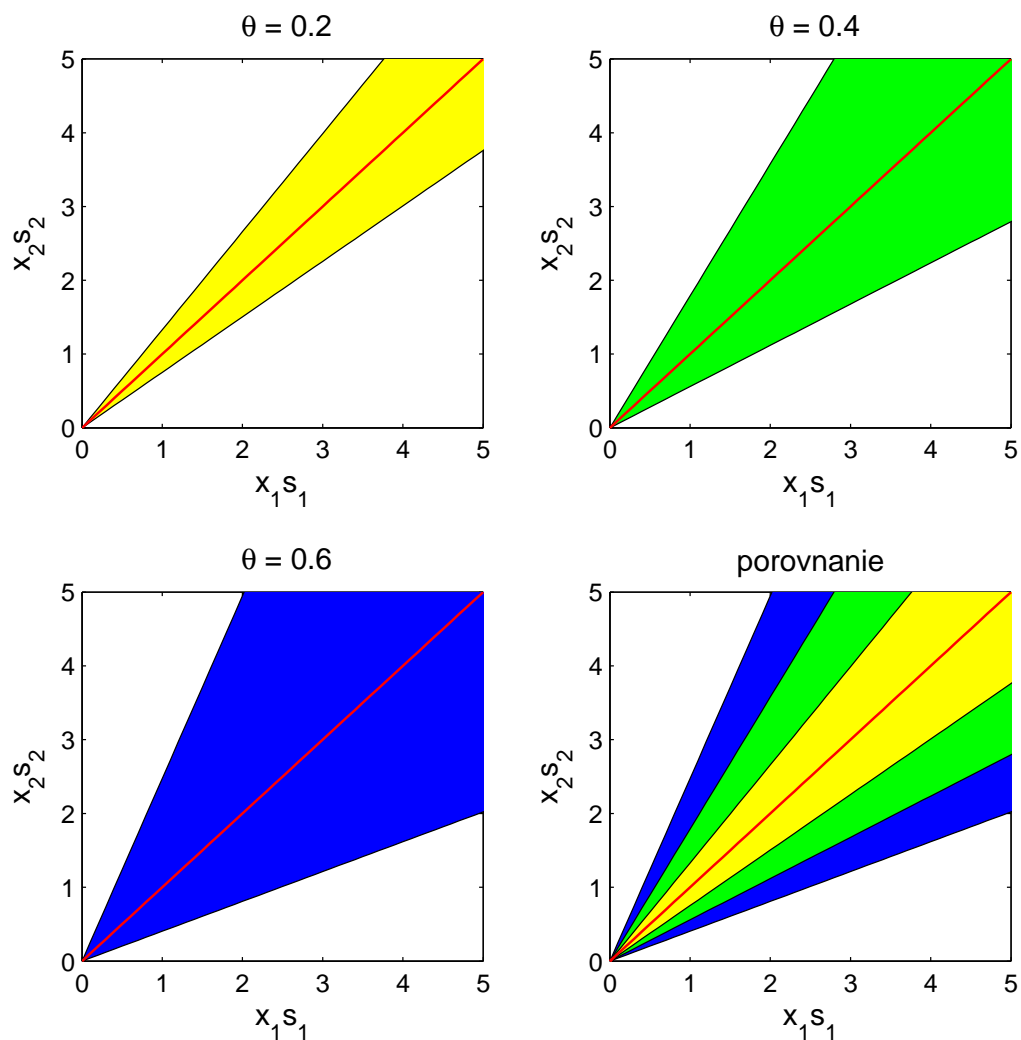
To znamená, že súčet štvorcov relatívnych odchýliek $x_i s_i$ od ich priemeru μ nesmie presiahnuť $\theta^2 < 1$. Z toho dôvodu okolie $\mathcal{N}_2(\theta)$ obsahuje iba malú časť ostro prípustných bodov z množiny \mathcal{F}° , a preto algoritmy využívajúce toto špecifické okolie majú me-

nej priestoru na manévrovanie pozdĺž centrálnej trajektórie. Na druhej strane, špecifické okolie $\mathcal{N}_{-\infty}(\gamma)$ môže pre malé hodnoty parametra γ zabrať skoro celú množinu \mathcal{F}° . Jedinou podmienkou špecifického okolia $\mathcal{N}_{-\infty}(\gamma)$ je, aby súčiny $x_i s_i$ neboli výrazne menšie ako ich priemer μ . Táto podmienka zabráni, aby sa dvojica (x, s) priblížila k hranici nezáporného ortantu $(x, s) \geq 0$. Z uvedených pozorovaní vyplýva, že špecifické okolia $\mathcal{N}_2(\theta)$ a $\mathcal{N}_{-\infty}(\gamma)$ sa kvalitatívne líšia, čomu nasvedčuje aj nasledujúca ilustrácia, ktorá porovnáva obe špecifické okolia v jednorozmernom prípade.

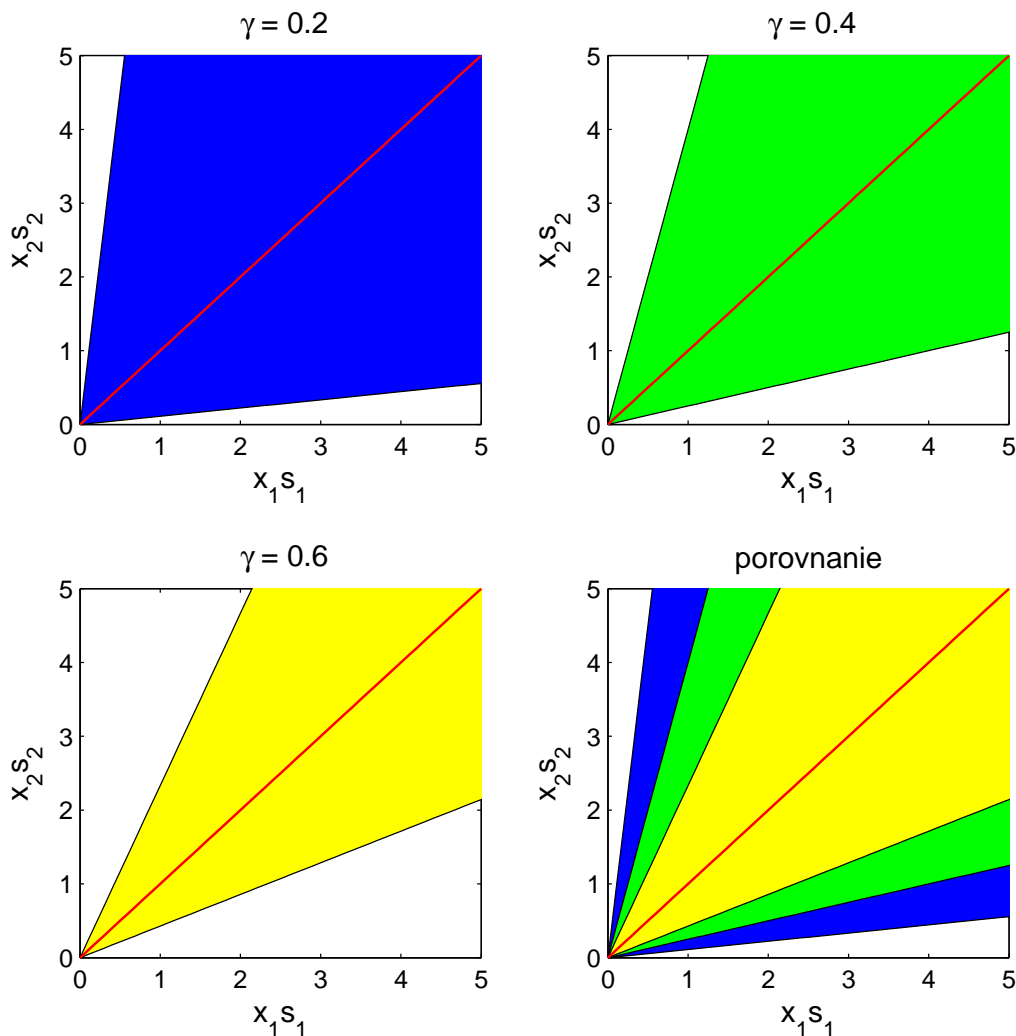


Obr. 3.2: Kvalitatívny rozdiel špecifických okolí $\mathcal{N}_2(\theta)$ a $\mathcal{N}_{-\infty}(\gamma)$ pre $n = 1$ ilustrovaný v priestore $\mu \times x_1 s_1$. Červená polpriamka predstavuje centrálnu trajektóriu.

Špecifické okolia $\mathcal{N}_2(\theta)$ a $\mathcal{N}_{-\infty}(\gamma)$ môžeme veľmi dobre ilustrovať aj v dvojrozmernom prípade, kedy okolia majú kvalitatívne podobný charakter. Tieto ilustrácie uvádzame na obrázkoch Obr. 3.3 a Obr. 3.4, pričom postup odvodenia nerovnic generujúcich tieto okolia uvádzame v Dodatku.



Obr. 3.3: Špecifické okolie $\mathcal{N}_2(\theta)$ pre $n = 2$ vykreslené v priestore $x_1s_1 \times x_2s_2$ v závislosti od parametra $\theta \in (0, 1)$. Červená polpriamka predstavuje centrálnu trajektóriu.



Obr. 3.4: Špecifické okolie $\mathcal{N}_{-\infty}(\gamma)$ pre $n = 2$ vykreslené v priestore $x_1 s_1 \times x_2 s_2$ v závislosti od parametra $\gamma \in (0, 1)$. Červená polpriamka predstavuje centrálnu trajektóriu.

Predchádzajúce ilustrácie vedú prirodzene k domnienke, ktorú uvádzame v nasledujúcom tvrdení.

Lema 3.1. Pre dvojicu okolí $\mathcal{N}_2(\theta)$, $\mathcal{N}_{-\infty}(\gamma)$ platí:

- (i) ak $0 < \theta_1 < \theta_2 < 1$, potom $\mathcal{N}_2(\theta_1) \subset \mathcal{N}_2(\theta_2)$;
- (ii) ak $0 < \gamma_1 < \gamma_2 < 1$, potom $\mathcal{N}_{-\infty}(\gamma_2) \subset \mathcal{N}_{-\infty}(\gamma_1)$;
- (iii) ak $\gamma \leq 1 - \theta$, potom $\mathcal{N}_2(\theta) \subset \mathcal{N}_{-\infty}(\gamma)$. [29]

Dôkaz. Dôkazy tvrdení (i) a (ii) sú zjavné. Dokážeme tvrdenie (iii), v ktorom treba dokázať implikáciu $\forall (x, y, s) \in \mathcal{N}_2(\theta) : (x, y, s) \in \mathcal{N}_2(\theta) \Rightarrow (x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$, $\gamma \leq 1 - \theta$. Nech (x, y, s) je ľubovoľný bod z okolia $\mathcal{N}_2(\theta)$. Keďže $(x, y, s) \in \mathcal{N}_2(\theta)$, tak $\|XSe - \mu e\|_2 \leq \theta\mu$. Pomocou nerovnosti $\|\cdot\|_\infty \leq \|\cdot\|_2$ odvodíme, že

$$\theta\mu \geq \|XSe - \mu e\|_2 = \|\mu e - XSe\|_2 \geq \|\mu e - XSe\|_\infty = \max_{i=1,2,\dots,n} |\mu - x_i s_i|.$$

Keďže $\mu > 0$ a $x_i s_i > 0$ pre $i = 1, 2, \dots, n$, tak zrejme platí

$$\max_{i=1,2,\dots,n} |\mu - x_i s_i| = \mu - \min_{i=1,2,\dots,n} x_i s_i.$$

Z odvodených vzťahov za predpokladu $\gamma \leq 1 - \theta$ dostaneme

$$\mu - \min_{i=1,2,\dots,n} x_i s_i = \max_{i=1,2,\dots,n} |\mu - x_i s_i| \leq \theta \mu \leq (1 - \gamma) \mu.$$

Elementárnymi úpravami nerovnice $\mu - \min_i x_i s_i \leq (1 - \gamma) \mu$ nakoniec odvodíme, že $\min_i x_i s_i \geq \gamma \mu$. Z toho bezprostredne vyplýva, že $x_i s_i \geq \gamma \mu$ pre $i = 1, 2, \dots, n$. Tým sme dokázali, že ľubovoľný bod (x, y, s) z okolia $\mathcal{N}_2(\theta)$ spĺňa podmienku presnosti pre okolie $\mathcal{N}_{-\infty}(\gamma)$, a preto $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$. \square

Algoritmus sledovania centrálnej trajektórie sleduje schému generického algoritmu, pričom vždy volí jedno z dvojice okolí $\mathcal{N}_2(\theta), \mathcal{N}_{-\infty}(\gamma)$. Zároveň volí centrujúci parameter σ a dĺžku kroku α tak, aby každý iteračný bod (x^k, y^k, s^k) ležal vo vybranom špecifickom okolí. V závislosti od voľby parametrov σ a α existujú tri základné varianty algoritmu sledovania centrálnej trajektórie – algoritmus s krátkym krokom [13], [20], perdiktor-korektor algoritmus [19] a algoritmus s dlhým krokom [14]. Spoločným znakom týchto algoritmov je polynomiálna zložitosť. V nasledujúcich častiach opíšeme princíp týchto algoritmov, uvedieme ich podstatné vlastnosti a na základe Vety 3.1 dokážeme ich polynomiálny charakter.

3.2 Polynomiálna zložitosť

V tejto časti uvedieme všeobecné tvrdenie o polynomiálnej zložitosti pre algoritmy sledovania centrálnej trajektórie. Veta hovorí o tom, že ak znižovanie parametra μ v každej iterácii závisí určitým spôsobom od rozmeru n , a ak počiatočná duálna medzera nie je „príliš veľká“, potom algoritmus sledovania centrálnej trajektórie má polynomiálnu zložitosť.

Veta 3.1. *Nech $\varepsilon \in (0, 1)$. Predpokladajme, že algoritmus na riešenie systému (1.11) generuje postupnosť iteračných bodov (x^k, y^k, s^k) , ktoré spĺňajú*

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n^\omega}\right) \mu_k, \quad k = 0, 1, 2, \dots \quad (3.1)$$

pre nejaké kladné konštanty $\delta, \omega > 0$. Zároveň predpokladajme, že štartovací bod (x^0, y^0, s^0) spĺňa podmienku $\mu_0 \leq 1/\varepsilon^\kappa$ pre nejakú kladnú konštantu $\kappa > 0$. Potom existuje index $\bar{k} = O(n^\omega |\ln \varepsilon|)$ taký, že $\mu_k \leq \varepsilon, \forall k \geq \bar{k}$. [29]

Dôkaz. Zlogaritmovaním oboch strán nerovnice (3.1) dostaneme

$$\ln \mu_{k+1} \leq \ln \left(1 - \frac{\delta}{n^\omega} \right) + \ln \mu_k. \quad (3.2)$$

Na základe rekurentného vzťahu (3.2) so začiatočnou podmienkou $\mu_0 \leq 1/\varepsilon^\kappa$ dostaneme

$$\begin{aligned} \ln \mu_k &\leq k \ln \left(1 - \frac{\delta}{n^\omega} \right) + \ln \mu_0 \\ &\leq k \ln \left(1 - \frac{\delta}{n^\omega} \right) + \kappa \ln \left(\frac{1}{\varepsilon} \right). \end{aligned} \quad (3.3)$$

Aplikovaním Lemy 2.1(i) na prvý člen výrazu (3.3) dostaneme

$$\ln \mu_k \leq k \ln \left(-\frac{\delta}{n^\omega} \right) + \kappa \ln \left(\frac{1}{\varepsilon} \right). \quad (3.4)$$

Z nerovnice (3.4) vyplýva, že konvergenčné kritérium $\mu_k \leq \varepsilon$ je splnené, ak

$$k \ln \left(-\frac{\delta}{n^\omega} \right) + \kappa \ln \left(\frac{1}{\varepsilon} \right) \leq \ln \varepsilon. \quad (3.5)$$

Z nerovnice (3.5) nakoniec odvodíme, že

$$\begin{aligned} k &\geq \left(\frac{n^\omega}{\delta} \right) \left[\kappa \ln \left(\frac{1}{\varepsilon} \right) - \ln \varepsilon \right] \\ &= \left(\frac{n^\omega}{\delta} \right) (\kappa \ln 1 - \kappa \ln \varepsilon - \ln \varepsilon) \\ &= \left(\frac{n^\omega}{\delta} \right) (-\kappa - 1) \ln \varepsilon \\ &= \left(\frac{n^\omega}{\delta} \right) (\kappa + 1) |\ln \varepsilon|. \end{aligned}$$

Tým sme dokázali, že existuje index $\bar{k} = (n^\omega/\delta)(\kappa + 1)|\ln \varepsilon|$ taký, že konvergenčné kritérium $\mu_k \leq \varepsilon$ je splnené pre $\forall k \geq \bar{k}$, pričom $\bar{k} = O(n^\omega |\ln \varepsilon|)$. \square

3.3 Algoritmus s krátkym krokom

Najjednoduchším variantom algoritmu sledovania centrálnej trajektórie sú algoritmy s krátkym krokom. Tieto algoritmy sledujú centrálnu trajektóriu v špecifickom okolí $\mathcal{N}_2(\theta)$, pričom v každej iterácii volia jednotkovú dĺžku kroku $\alpha_k = 1$ a konštantnú hodnotu centrujúceho parametra $\sigma_k = \sigma$, kde $\theta \in (0, 1)$ a $\sigma \in (0, 1)$ spĺňajú nižšie uvedený vzťah (3.12). Pre jednoduchosť uvedieme schému tohto algoritmu s konkrétnymi parametrami $\theta = 0.4$ a $\sigma = 1 - 0.4/\sqrt{n}$.

Schéma algoritmu s krátkym krokom

vstup $\theta = 0.4, \sigma = 1 - 0.4/\sqrt{n}, (x^0, y^0, s^0) \in \mathcal{N}_2(\theta), \epsilon > 0$

prirad' $k := 0$

repeat

prirad' $\sigma_k := \sigma, \mu_k := (x^k)^T s^k / n$

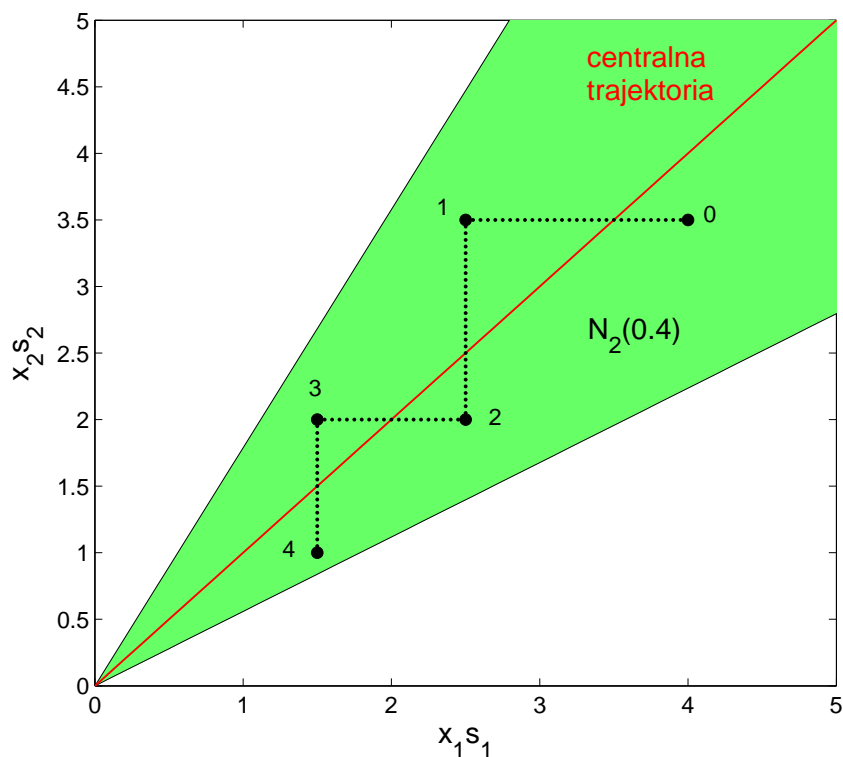
vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + (\Delta x, \Delta y, \Delta s)$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \epsilon$



Obr. 3.5: Iterácie algoritmu s krátkym krokom v špecifickom okolí $\mathcal{N}_2(0.4)$ ilustrované v priestore $x_1 s_1 \times x_2 s_2$.

Pre algoritmus s krátkym krokom najprv dokážeme jeho polynomiálnu zložitosť. Zaved' me pomocné označenie

$$\begin{aligned}(x(\alpha), y(\alpha), s(\alpha)) &= (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s), \\ \mu(\alpha) &= x(\alpha)^T s(\alpha)/n.\end{aligned}$$

Lema 3.2. Ak $(\Delta x, \Delta y, \Delta s)$ je riešením systému (1.11), potom $\mu(\alpha) = [1 - \alpha(1 - \sigma)]\mu$. [29]

Dôkaz. Tvrdenie dokážeme pomocou rovnice $X\Delta s + S\Delta x = -XSe + \sigma\mu e$. Sčítaním n zložiek na oboch stranách rovnice dostaneme

$$x^T \Delta s + s^T \Delta x = -x^T s + n\sigma\mu = -x^T s + \sigma x^T s = -(1 - \sigma)x^T s.$$

Rozpísaním výrazu $x(\alpha)^T s(\alpha)$ dostaneme

$$\begin{aligned}x(\alpha)^T s(\alpha) &= x^T s + \alpha(x^T \Delta s + s^T \Delta x) + \alpha^2 \underbrace{\Delta x^T \Delta s}_{=0} \\ &= x^T s + \alpha[-(1 - \sigma)x^T s] \\ &= [1 - \alpha(1 - \sigma)]x^T s.\end{aligned}$$

Prenásobením oboch strán odvodenej rovnice $x(\alpha)^T s(\alpha) = [1 - \alpha(1 - \sigma)]x^T s$ výrazom $1/n$ nakoniec dostaneme požadovaný vzťah. \square

Keďže $\alpha_k = 1$ a $\sigma = 1 - 0.4/\sqrt{n}$, tak podľa Lemy 3.9 pre algoritmus s krátkym krokom platí

$$\mu_{k+1} = [1 - \alpha_k(1 - \sigma)]\mu_k = \sigma\mu_k = \left(1 - \frac{0.4}{\sqrt{n}}\right)\mu_k, \quad k = 0, 1, 2, \dots, \quad (3.6)$$

a preto polynomiálna zložitosť $O(\sqrt{n} \ln 1/\varepsilon)$ tohto algoritmu je priamym dôsledkom vzťahu (3.6) a Vety 3.1. V ďalšej časti tejto kapitoly postupne dokážeme, že všetky iteračné body algoritmu s krátkym krokom ležia v špecifickom okolí $\mathcal{N}_2(\theta)$.

Lema 3.3. Pre ľubovoľné čísla $\beta, \gamma \in \mathbf{R}$ splňajúce podmienku $\beta\gamma \geq 0$ platí

$$\sqrt{|\beta\gamma|} \leq \frac{1}{2}|\beta + \gamma|. \quad (3.7)$$

Lema 3.4. Nech $u, v \in \mathbf{R}^n$ sú ľubovoľné vektory splňajúce podmienku $u^T v \geq 0$. Potom

$$\|UVe\| \leq 2^{-3/2}\|u + v\|^2,$$

kde $U = \text{diag}(u)$, $V = \text{diag}(v)$. [29]

Dôkaz. Na základe predpokladu $u^T v \geq 0$ môžeme odvodiť, že

$$0 \leq u^T v = \sum_{u_i v_i \geq 0} u_i v_i + \sum_{u_i v_i < 0} u_i v_i = \sum_{i \in \mathcal{V}} |u_i v_i| - \sum_{i \in \mathcal{W}} |u_i v_i|, \quad (3.8)$$

kde $\mathcal{V} = \{i \mid u_i v_i \geq 0\}$, $\mathcal{W} = \{i \mid u_i v_i < 0\}$. Postupnými úpravami dostaneme

$$\begin{aligned}
 \|UVe\| &= (\|[u_i v_i]_{i \in \mathcal{V}}\|^2 + \|[u_i v_i]_{i \in \mathcal{W}}\|^2)^{1/2} \\
 &\leq (\|[u_i v_i]_{i \in \mathcal{V}}\|_1^2 + \|[u_i v_i]_{i \in \mathcal{W}}\|_1^2)^{1/2} && \text{keďže } \|\cdot\|_2 \leq \|\cdot\|_1 \\
 &\leq (2\|[u_i v_i]_{i \in \mathcal{V}}\|_1^2)^{1/2} && \text{podľa (3.8)} \\
 &\leq \sqrt{2} \left\| \left[\frac{1}{4}(u_i + v_i)^2 \right]_{i \in \mathcal{V}} \right\|_1 && \text{podľa (3.7)} \\
 &= 2^{-3/2} \sum_{i \in \mathcal{V}} (u_i + v_i)^2 \\
 &\leq 2^{-3/2} \sum_{i=1}^n (u_i + v_i)^2 \\
 &= 2^{-3/2} \|u + v\|^2
 \end{aligned}$$

□

Lema 3.5. Ak $(x, y, s) \in \mathcal{N}_2(\theta)$, potom

$$\|\Delta X \Delta S e\| \leq \frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} \mu. \quad [29]$$

Dôkaz. V úvode dôkazu pripomeňme, že $D = X^{1/2} S^{-1/2}$. Prenásobením oboch strán rovnice $X \Delta s + S \Delta x = -X S e + \sigma \mu e$ maticou $(X S)^{-1/2}$ dostaneme $D^{-1} \Delta x + D \Delta s = (X S)^{-1/2} (-X S e + \sigma \mu e)$. Teraz priradíme $u := D^{-1} \Delta x$, $v := D \Delta s$ a na takto definované vektory $u, v \in \mathbf{R}^n$ aplikujeme Lemu 3.4. Dostaneme

$$\begin{aligned}
 \|\Delta X \Delta S e\| &= \|(D^{-1} \Delta X)(D \Delta S) e\| \\
 &\leq 2^{-3/2} \|D^{-1} \Delta x + D \Delta s\|^2 \\
 &= 2^{-3/2} \|(X S)^{-1/2} (-X S e + \sigma \mu e)\|^2 \\
 &= 2^{-3/2} \sum_{i=1}^n \frac{(-x_i s_i + \sigma \mu)^2}{x_i s_i} \\
 &\leq 2^{-3/2} \frac{\|X S e - \sigma \mu e\|^2}{\min_i x_i s_i}. \tag{3.9}
 \end{aligned}$$

Z predpokladu $(x, y, s) \in \mathcal{N}_2(\theta)$ môžeme podobne ako v dôkaze Lemy 3.1 odvodiť, že $\min_i x_i s_i \geq (1 - \theta) \mu$. Využitím tohto vzťahu v nerovnici (3.9) dostaneme

$$\|\Delta X \Delta S e\| \leq 2^{-3/2} \frac{\|X S e - \sigma \mu e\|^2}{(1 - \theta) \mu}. \tag{3.10}$$

Rozpísaním výrazu $\|X S e - \sigma \mu e\|^2$ na pravej strane nerovnice (3.10) dostaneme

$$\begin{aligned}
 \|X S e - \sigma \mu e\|^2 &= \|(X S - \mu e) + (1 - \sigma) \mu e\|^2 \\
 &= \|X S - \mu e\|^2 + 2(1 - \sigma) \underbrace{\mu e^T (X S e - \mu e)}_{=0} + (1 - \sigma)^2 \mu^2 e^T e \\
 &\leq \theta^2 \mu^2 + (1 - \sigma)^2 \mu^2 n.
 \end{aligned}$$

Z toho vyplýva, že pre nerovnicu (3.10) ďalej platí

$$\begin{aligned}\|\Delta X \Delta S e\| &\leq 2^{-3/2} \frac{\theta^2 \mu^2 + (1 - \sigma)^2 \mu^2 n}{(1 - \theta) \mu} \\ &= \frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} \mu.\end{aligned}$$

□

Dôsledok 3.1. Ak $(x, y, s) \in \mathcal{N}_2(\theta)$, potom

$$\begin{aligned}\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\leq |1 - \alpha| \|XSe - \mu e\| + \alpha^2 \|\Delta X \Delta Se\| \\ &\leq |1 - \alpha| \theta \mu + \alpha^2 \left[\frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} \right] \mu. \quad [29] \quad (3.11)\end{aligned}$$

Dôkaz. Využitím Lemy 3.2 dostaneme, že

$$\begin{aligned}x_i(\alpha)s_i(\alpha) - \mu(\alpha) &= x_i s_i + \alpha(x_i \Delta s_i + s_i \Delta x_i) + \alpha^2 \Delta x_i \Delta s_i - [1 - \alpha(1 - \sigma)] \mu \\ &= x_i s_i (1 - \alpha) + \alpha \sigma \mu + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha + \alpha \sigma) \mu \\ &= x_i s_i (1 - \alpha) + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha) \mu.\end{aligned}$$

Zoradením prvkov odvodenej rovnice do vektora dostaneme

$$\begin{aligned}\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &= \|[x_i s_i (1 - \alpha) - (1 - \alpha) \mu + \alpha^2 \Delta x_i \Delta s_i]_{i=1}^n\| \\ &\leq |1 - \alpha| \|XSe - \mu e\| + \alpha^2 \|\Delta X \Delta Se\| \\ &\leq |1 - \alpha| \theta \mu + \alpha^2 \frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} \mu.\end{aligned}$$

□

Nasledujúce tvrdenie formuluje vzťah medzi parametrami θ a σ , ktorý zabezpečí, že jednotlivé iteračné body algoritmu zotrývajú v špecifickom okolí $\mathcal{N}_2(\theta)$. Tvrdenie zároveň hovorí o tom, že ani voľbou plného kroku $\alpha = 1$ pozdĺž Newtonovho smeru mechanismus optimalizácie neopustí špecifické okolie $\mathcal{N}_2(\theta)$.

Veta 3.2. Nech parametre $\theta \in (0, 1)$, $\sigma \in (0, 1)$ spĺňajú vzťah

$$\frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} \leq \sigma \theta. \quad (3.12)$$

Potom, ak $(x, y, s) \in \mathcal{N}_2(\theta)$, tak $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_2(\theta)$, $\forall \alpha \in [0, 1]$. [29]

Dôkaz. Z nerovnice (3.11) a pomocou predpokladu (3.12) jednoducho odvodíme, že

$$\begin{aligned}\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\leq |1 - \alpha| \theta \mu + \alpha^2 \sigma \theta \mu \\ &\leq (1 - \alpha + \sigma \alpha) \theta \mu && \text{keďže } \alpha \in [0, 1] \\ &= \theta \mu(\alpha). && \text{podľa Lemy 3.2}\end{aligned}$$

To znamená, že bod $(x(\alpha), y(\alpha), s(\alpha)), \forall \alpha \in [0, 1]$ spĺňa podmienku presnosti pre špeci-
fické okolie $\mathcal{N}_2(\theta)$. Nakoniec dokážeme, že $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{F}^\circ, \forall \alpha \in [0, 1]$. Na základe
schémy algoritmu s krátkym krokom odvodíme, že

$$Ax(\alpha) = A(x + \alpha\Delta x) = \underbrace{Ax}_{=b} + \alpha \underbrace{A\Delta x}_{=0} = b,$$

$$A^T y(\alpha) + s(\alpha) = A^T(y + \alpha\Delta y) + s + \alpha\Delta s = \underbrace{A^T y + s}_{=c} + \alpha \underbrace{(A^T \Delta y + \Delta s)}_{=0} = c.$$

Ak $\alpha = 0$, potom $(x(\alpha), s(\alpha)) = (x, s) > 0$. Keďže $\theta, \sigma \in (0, 1)$, tak pre $\alpha \in (0, 1]$ platí

$$x_i(\alpha)s_i(\alpha) \geq (1 - \theta)\mu(\alpha) = (1 - \theta)[1 - \alpha(1 - \sigma)]\mu > 0.$$

Z toho vyplýva, že ak $\alpha \in (0, 1]$, potom pre žiadny index $i = 1, 2, \dots, n$ nemôže platiť
 $x_i(\alpha) = 0$, ani $s_i(\alpha) = 0$. Z toho dôvodu platí $(x(\alpha), s(\alpha)) > 0, \forall \alpha \in [0, 1]$. Tým sme
dokázali, že $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{F}^\circ, \forall \alpha \in [0, 1]$. \square

Na záver zostáva overiť, že parametre $\theta = 0.4, \sigma = 1 - 0.4/n$ spĺňajú podmienku
(3.12) pre $n \geq 1$. Túto časť uvádzame v Dodatku.

3.4 Prediktor-korektor algoritmus

V algoritme s krátkym krokom sme centrujúci parameter volili v každej iterácii koštantne
 $\sigma_k = \sigma$ z intervalu $(0, 1)$. Takto zvolený cetrujúci parameter v každej iterácii zlepšil cen-
trovanie smerom k centrálnej trajektórii a zároveň zmenšil duálnu medzeru. Prediktor-
korektor algoritmus rieši tieto dva problémy postupne, a to tak, že iterácie algoritmu
alternujú medzi dvomi krokmi.

- *Prediktor krok*, ktorý zodpovedná voľbe $\sigma_k = 0$. Touto voľbou získame *afinno-škálovací*
smer $(\Delta x, \Delta y, \Delta s)$, ktorý smeruje k hranici prípustných riešení, a pozdĺž ktorého
duálna medzera poklesne najviac.
- *Korektor krok*, ktorý zodpovedá voľbe $\sigma_k = 1$. Touto voľbou získame *centrujúci smer*
 $(\Delta x, \Delta y, \Delta s)$, ktorý zlepšuje centrovanie smerom k centrálnej trajektórii, a pozdĺž
ktorého sa duálna medzera nezmení.

Dôležitou súčasťou prediktor-korektor algoritmu je dvojica \mathcal{N}_2 -okolí, kde jedno okolie
(vnútorné) tvorí podmnožinu druhého (vonkajšieho). Párne iterácie prediktor-korektor
algoritmu sú sústredené vo vnútornom okolí, zatiaľ čo nepárne iterácie ležia vo vonkaj-
šom okolí. Stručne opíšeme prvé dve iterácie prediktor-korektor algoritmu a na základe
nich sformulujeme schému tohto algoritmu. Začínajúc štartovacím bodom (x^0, y^0, s^0) z
vnútorného okolia vypočítame prediktor krok voľbou $\sigma_0 = 0$. Pozdĺž získaného afinno-
škálovacieho smeru prejdeme až k hranici vonkajšieho okolia, kde definujeme nový ite-
račný bod (x^1, y^1, s^1) . Korektor krok vypočítame voľbou $\sigma_1 = 1$. Voľbou jednotkovej

dĺžky kroku $\alpha = 1$ pozdĺž centrujúceho smeru získame ďalšiu iteráciu (x^2, y^2, s^2) , ktorá opäť leží vo vnútornom okolí. Takto opísaný dvojkrokový cyklus sa opakuje, a tým generuje postupnosť iteračných bodov s párnou iteráciou vo vnútornom okolí a s nepárnou iteráciou na hranici vonkajšieho okolia centrálnej trajektórie. Pre jednoduchosť a bez ujmy na všeobecnosti sformulujeme schému prediktor-korektor algoritmu s konkrétnymi okoliami $\mathcal{N}_2(0.25)$ a $\mathcal{N}_2(0.5)$. Poznamenajme, že podľa Lemy 3.1 vnútorné okolie $\mathcal{N}_2(0.25)$ naozaj tvorí podmnožinu vonkajšieho okolia $\mathcal{N}_2(0.5)$.

Schéma prediktor-korektor algoritmu

vstup $(x^0, y^0, s^0) \in \mathcal{N}_2(0.25), \epsilon > 0$

prirad' $k := 0$

repeat

ak k je párne

PREDIKTOR KROK

prirad' $\sigma_k := 0, \mu_k := (x^k)^T s^k / n$

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $\alpha_k := \max \{ \alpha \in [0, 1] \mid (x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_2(0.5) \}$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k(\alpha), y^k(\alpha), s^k(\alpha))$

ak k je nepárne

KOREKTOR KROK

prirad' $\sigma_k := 1, \mu_k := (x^k)^T s^k / n$

vyrieš

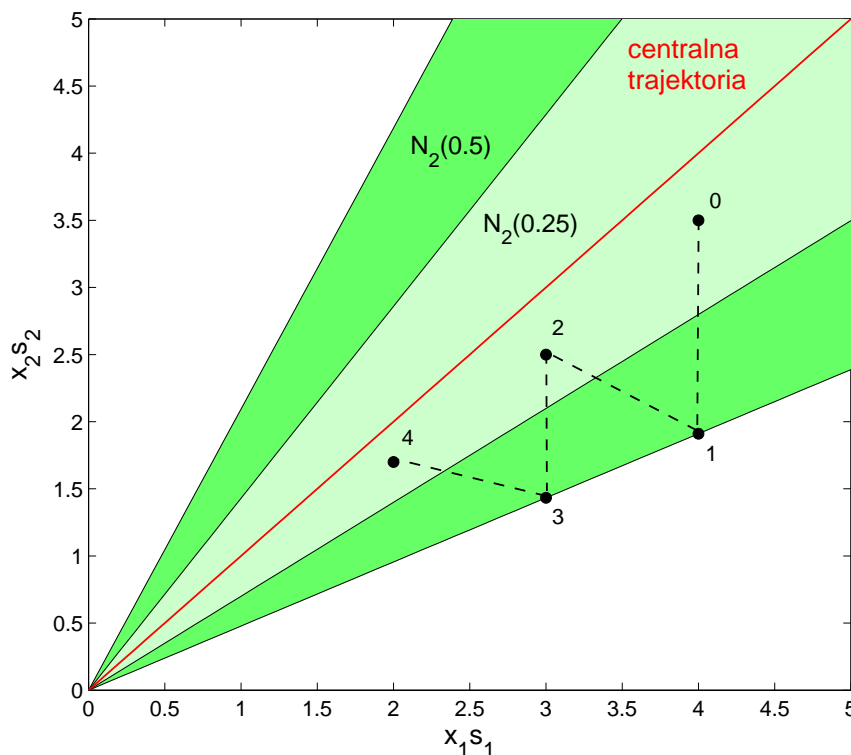
$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + (\Delta x, \Delta y, \Delta s)$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \epsilon$

Prediktor krok redukuje hodnotu parametra μ faktorom $(1 - \alpha)$, kde α je dĺžka kroku. Korektor krok síce neprispieva k redukcii duálnej medzery, ale pri návrate späť do vnútorného okolia dáva algoritmu väčší priestor na manévrovanie k ďalšej iterácii. V ďalšej časti opíšeme základné vlastnosti prediktor-korektor algoritmu a dokážeme jeho polynomiálnu zložitosť. Prvé tvrdenie definuje spodnú hranicu dĺžky prediktor kroku.



Obr. 3.6: Iterácie prediktor-korektor algoritmu alternujúce medzi špecifickými okoliami $\mathcal{N}_2(0.25)$ a $\mathcal{N}_2(0.5)$ ilustrované v priestore $x_1 s_1 \times x_2 s_2$.

Lema 3.6. Nech $(x, y, s) \in \mathcal{N}_2(0.25)$ a nech $(\Delta x, \Delta y, \Delta s)$ je riešením systému (1.11) s $\sigma = 0$. Potom $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_2(0.5)$, $\forall \alpha \in [0, \bar{\alpha}]$, kde

$$\bar{\alpha} = \min \left(\frac{1}{2}, \left(\frac{\mu}{8 \|\Delta X \Delta S e\|} \right)^{1/2} \right). \quad [29] \quad (3.13)$$

Dôkaz. Využitím Dôsledku 3.1 dostaneme

$$\begin{aligned} \|X(\alpha)S(\alpha) - \mu(\alpha)e\| &\leq (1 - \alpha)\|XSe - \mu e\| + \alpha^2 \|\Delta X \Delta S e\| \\ &\leq (1 - \alpha)\|XSe - \mu e\| + \frac{\mu}{8 \|\Delta X \Delta S e\|} \|\Delta X \Delta S e\| && \text{podľa (3.13)} \\ &\leq \frac{1}{4}(1 - \alpha)\mu + \frac{1}{8(1 - \alpha)}(1 - \alpha)\mu && \text{keďže } (x, y, s) \in \mathcal{N}_2(0.25) \\ &\leq \frac{1}{4}(1 - \alpha)\mu + \frac{1}{4}(1 - \alpha)\mu && \text{keďže } \alpha \leq \frac{1}{2} \\ &\leq \frac{1}{2}\mu(\alpha) && \text{podľa Lemy 3.2 s } \sigma = 0. \end{aligned}$$

To znamená, že pre $\alpha \in [0, \bar{\alpha}]$ bod $(x(\alpha), y(\alpha), s(\alpha))$ spĺňa podmienku presnosti pre okolie $\mathcal{N}_2(0.5)$. Dôkaz ostrej prípustnosti bodu $(x(\alpha), y(\alpha), s(\alpha))$ pre $\alpha \in [0, \bar{\alpha}]$ je analogický ako vo Vete 3.2. \square

Dôsledok 3.2. *Prediktor krok má dĺžku aspoň $\bar{\alpha}$ a nová hodnota parametra μ je najviac $(1 - \bar{\alpha})\mu$.* [29]

Dôkaz. Spodná hranica $\bar{\alpha}$ dĺžky prediktor kroku vyplýva priamo z Lemy 3.6 a zo spôsobu voľby dĺžky kroku

$$\alpha_k := \max \{ \alpha \in [0, 1] \mid (x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_2(0.5) \}.$$

Horné ohraničenie parametra μ v ďalšej iterácii vyplýva priamo z Lemy 3.2, ak vo vzťahu $\mu_{k+1} = [1 - \alpha_k(1 - \sigma)]\mu_k$ položíme $\sigma = 0$

$$\mu_{k+1} = [1 - \alpha_k(1 - \sigma)]\mu_k = (1 - \alpha_k)\mu_k \leq (1 - \bar{\alpha})\mu_k.$$

□

Poznamenajme, že pomocou Lemy 3.5 môžeme jednoducho odvodiť spodnú hranicu pre parameter $\bar{\alpha}$. Ak v Leme 3.5 položíme $\theta = 0.25$ a $\sigma = 0$, dostaneme

$$\frac{\mu}{8\|\Delta X \Delta S\|} \geq \frac{2^{3/2}(1 - 0.25)}{8[(0.25)^2 + n]} = \frac{3\sqrt{2}}{1 + 16n} \geq \frac{0.16}{n}.$$

Následne z nerovnice (3.13) odvodíme, že

$$\bar{\alpha} \geq \min \left(\frac{1}{2}, \left(\frac{0.16}{n} \right)^{1/2} \right) = \frac{0.4}{\sqrt{n}}. \quad (3.14)$$

Keďže každý krok typu prediktor zodpovedá párnej iterácii, tak Dôsledok 3.2 a ohraničenie (3.14) implikujú, že

$$\mu_{k+1} \leq (1 - \bar{\alpha}) \leq \left(1 - \frac{0.4}{\sqrt{n}} \right) \mu_k, \quad k = 0, 2, 4, \dots \quad (3.15)$$

Vlastnosti krokov typu korektor formuluje nasludujúca lema, ktorá hovorí o tom, že každý korektor krok vracia bod z vonkajšieho okolia $\mathcal{N}_2(0.5)$ do vnútorného okolia $\mathcal{N}_2(0.25)$ bez zmeny parametra μ .

Lema 3.7. *Nech $(x, y, s) \in \mathcal{N}_2(0.5)$ a nech $(\Delta x, \Delta y, \Delta s)$ je riešením systému (1.11) s $\sigma = 1$. Potom $(x(1), y(1), s(1)) \in \mathcal{N}_2(0.25)$ a $\mu(1) = \mu$.* [29]

Dôkaz. Dosadením $\sigma = 1$ do (3.2) dostaneme $\mu(1) = \mu$. Dosadením $\theta = 0.5$, $\alpha = 1$, $\sigma = 1$ do nerovnice (3.11) dostaneme

$$\|X(1)S(1)e - \mu(1)e\| \leq \frac{\mu}{4} = \frac{1}{4}\mu(1).$$

Tým sme dokázali, že bod $(x(1), y(1), s(1))$ spĺňa podmienku presnosti pre okolie $\mathcal{N}_2(0.25)$. Dôkaz ostrej prípustnosti bodu $(x(1), y(1), s(1))$ je analogický ako vo Vete (3.2). □

Nakoniec zostáva ukázať, že prediktor-korektor algoritmus má polynomiálny charakter. Všimnime si, že spojením nerovnice (3.15) a Lemy 3.7 môžeme odvodiť, že

$$\mu_{k+2} = \mu_{k+1} \leq \left(1 - \frac{0.4}{\sqrt{n}}\right) \mu_k, \quad k = 0, 2, 4, \dots \quad (3.16)$$

To znamená, že požadovaný spôsob znižovania bariérového parametra podľa Vety 3.1 je takmer splnený pre $\delta = 0.4$ a $\omega = 0.5$, až na to, že znižovanie parametra μ neprebíha v každej iterácii ale v dvojkrokovom cykle. Avšak kniha [29] uvádza, že dôkaz Vety 3.1 je možné upraviť tak, aby spôsob redukcie (3.16) vyhovoval predpokladom tvrdenia, čím je možné dokázať polynomiálnu zložitosť $O(\sqrt{n} \ln 1/\varepsilon)$ tohto algoritmu.

3.5 Algoritmus s dlhým krokom

Algoritmus s dlhým krokom generuje postupnosť iteračných bodov (x^k, y^k, s^k) v špeciickom okolí $\mathcal{N}_{-\infty}(\gamma)$, ktoré sa vyznačuje tým, že pre malé hodnoty parametra γ ($\approx 10^{-3}$) zaberá podstatnú časť množiny \mathcal{F}° . Tento algoritmus volí v každej iterácii centrujúci parameter σ_k z intervalu $[\sigma_{\min}, \sigma_{\max}]$, kde $0 < \sigma_{\min} < \sigma_{\max} < 1$. Algoritmus zároveň volí dĺžku kroku α_k maximálnu možnú tak, aby ďalší iteračný bod zotrval vo vnútri okolia $\mathcal{N}_{-\infty}(\gamma)$. Schéma algoritmu s dlhým krokom teda vyzerá nasledovne.

Schéma algoritmu s dlhým krokom

vstup $\gamma \in (0, 1)$, $\sigma_{\min}, \sigma_{\max} : 0 < \sigma_{\min} < \sigma_{\max} < 1$, $(x^0, y^0, s^0) \in \mathcal{N}_{\infty}(\gamma)$, $\varepsilon > 0$

prirad' $k := 0$

repeat

zvol' $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$

prirad' $\mu_k := (x^k)^T s^k / n$

vyrieš

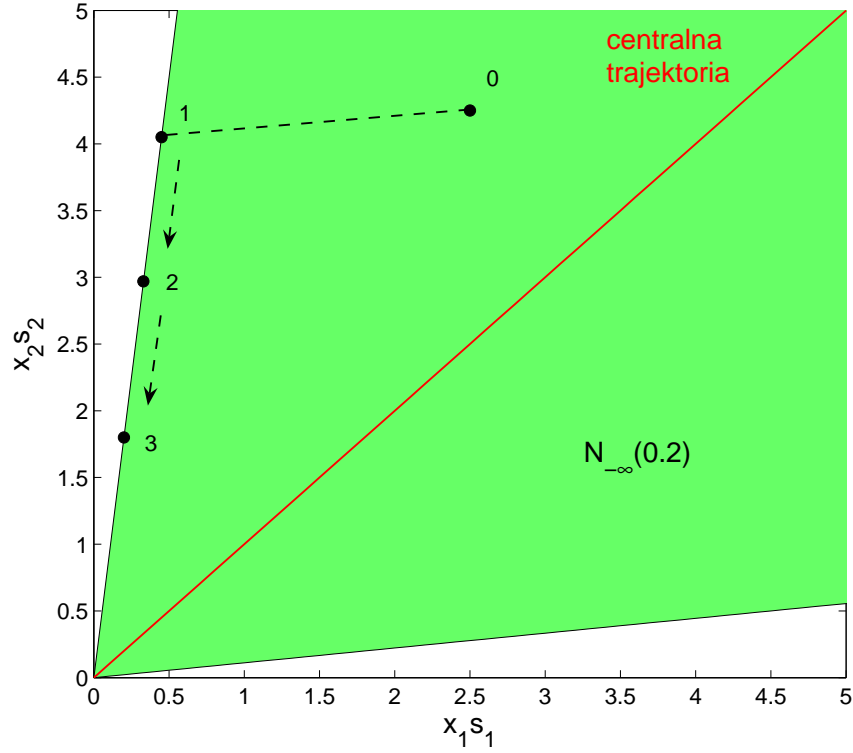
$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $\alpha_k := \max \{ \alpha \in [0, 1] \mid (x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_{-\infty}(\gamma) \}$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x(\alpha), y(\alpha), s(\alpha))$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \varepsilon$



Obr. 3.7: Iterácie algoritmu s dlhým krokom v špecifickom okolí $\mathcal{N}_{-\infty}(0.2)$ ilustrované v priestore $x_1s_1 \times x_2s_2$.

Lema 3.8. Ak $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$, potom $\|\Delta X \Delta S e\| \leq 2^{-3/2}(1 + 1/\gamma)n\mu$. [29]

Dôkaz. Z dôkazu Lemy 3.5 využijeme nerovnicu

$$\|\Delta X \Delta S e\| \leq 2^{-3/2} \|(XS)^{-1/2}(-XSe + \sigma\mu e)\|^2. \quad (3.17)$$

Rozpísaním výrazu na pravej strane nerovnice (3.17) dostaneme

$$\begin{aligned} \|\Delta X \Delta S e\| &\leq 2^{-3/2} \left\| -(XS)^{1/2}e + \sigma\mu(XS)^{-1/2}e \right\|^2 \\ &\leq 2^{-3/2} \left[x^T s - 2\sigma\mu e^T e + \sigma^2 \mu^2 \sum_{i=1}^n \frac{1}{x_i s_i} \right] \\ &\leq 2^{-3/2} \left[x^T s - 2\sigma\mu e^T e + \sigma^2 \mu^2 \frac{n}{\gamma\mu} \right] && \text{keďže } x_i s_i \geq \gamma\mu \\ &\leq 2^{-3/2} \left[1 - 2\sigma + \frac{\sigma^2}{\gamma} \right] n\mu \\ &\leq 2^{-3/2} \left(1 + \frac{1}{\gamma} \right) n\mu. && \text{keďže } \sigma \in (0, 1) \end{aligned}$$

□

Veta 3.3. Pre dané parametre $\gamma, \sigma_{\min}, \sigma_{\max}$ algoritmu s dlhým krokom existuje kladná konštanta $\delta > 0$ nezávislá od n taká, že

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right) \mu_k, \quad k = 1, 2, \dots \quad [29]$$

Dôkaz. Najprv dokážeme, že $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_{-\infty}(\gamma), \forall \alpha \in \left[0, 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}\right]$, čo implikuje dolné ohraničenie pre dĺžku kroku $\alpha_k \geq 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}$.

Z Lemy 3.8 vyplýva, že pre ľubovoľný index $i = 1, 2, \dots, n$ platí

$$|\Delta x_i \Delta s_i| \leq \|\Delta X^k \Delta S^k e\|_2 \leq 2^{-3/2} (1 + 1/\gamma) n \mu_k. \quad (3.18)$$

Rozpísaním výrazu $x_i^k(\alpha) s_i^k(\alpha)$ dostaneme

$$\begin{aligned} x_i^k(\alpha) s_i^k(\alpha) &= (x_i^k + \alpha \Delta x_i)(s_i^k + \alpha \Delta s_i) \\ &= x_i^k s_i^k + \alpha(x_i^k \Delta s_i + s_i^k \Delta x_i) + \alpha^2 \Delta x_i \Delta s_i \\ &\geq x_i^k s_i^k (1 - \alpha) + \alpha \sigma_k \mu_k - \alpha^2 |\Delta x_i \Delta s_i| && \text{podľa (1.11)} \\ &\geq \gamma(1 - \alpha) \mu_k + \alpha \sigma_k \mu_k - \alpha^2 2^{-3/2} (1 + 1/\gamma) n \mu_k. \quad x_i^k s_i^k \geq \gamma \mu_k, && \text{podľa (3.18)} \end{aligned}$$

Pripomeňme, že podľa Lemy 3.2 platí $\mu_k(\alpha) = [1 - \alpha(1 - \sigma_k)] \mu_k$. Z uvedených vzťahov vyplýva, že podmienka presnosti $x_i^k(\alpha) s_i^k(\alpha) \geq \gamma \mu_k(\alpha)$ pre okolie $\mathcal{N}_{-\infty}(\gamma)$ je splnená, ak

$$\gamma(1 - \alpha) \mu_k + \alpha \sigma_k \mu_k - \alpha^2 2^{-3/2} (1 + 1/\gamma) n \mu_k \geq \gamma [1 - \alpha(1 - \sigma_k)] \mu_k. \quad (3.19)$$

Elementárnymi úpravami môžeme ukázať, že nerovnica (3.19) platí pre $\alpha \leq 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}$. To znamená, že bod $(x^k(\alpha), y^k(\alpha), s^k(\alpha))$ spĺňa podmienku presnosti pre okolie $\mathcal{N}_{-\infty}(\gamma)$, ak $\alpha \in \left[0, 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}\right]$. Analogickým spôsobom, ako vo Vete 3.2, môžeme dokázať, že $(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{F}^\circ, \forall \alpha \in \left[0, 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}\right]$. Z toho jednoznačne vyplýva, že

$$(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma), \quad \forall \alpha \in \left[0, 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}\right],$$

a preto zrejme platí aj dolné ohraničenie pre parameter α_k

$$\alpha_k \geq 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \frac{\sigma_k}{n}. \quad (3.20)$$

Využitím ohraničenia (3.20) vo vzťahu $\mu_k(\alpha) = [1 - \alpha(1 - \sigma_k)] \mu_k$ dostaneme

$$\begin{aligned} \mu_{k+1} &= [1 - \alpha_k(1 - \sigma_k)] \mu_k \\ &\leq \left[1 - \frac{2^{3/2}}{n} \gamma \frac{1-\gamma}{1+\gamma} \sigma_k (1 - \sigma_k)\right] \mu_k. \end{aligned} \quad (3.21)$$

Je zrejmé, že výraz $\sigma(1 - \sigma)$ je konkávnou funkciou v premennej σ , a preto na ľubovoľnom uzavretom intervale dosahuje svoje minimum v niektorom z dvojice krajných bodov daného intervalu. Z toho dôvodu platí

$$\sigma_k(1 - \sigma_k) \geq \min \{ \sigma_{\min}(1 - \sigma_{\min}), \sigma_{\max}(1 - \sigma_{\max}) \}, \forall \sigma_k \in [\sigma_{\min}, \sigma_{\max}].$$

To znamená, že nerovnicu (3.21) môžeme ďalej rozpísať do tvaru

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right) \mu_k,$$

kde $\delta = 2^{3/2} \gamma \frac{1-\gamma}{1+\gamma} \min \{ \sigma_{\min}(1 - \sigma_{\min}), \sigma_{\max}(1 - \sigma_{\max}) \}$. □

Polynomiálna zložitosť $O(n \ln 1/\varepsilon)$ je opäť priamym dôsledkom Vety 3.1 a Vety 3.3.

3.6 Konvergencia

V tejto časti uvedieme konvergenčné vlastnosti algoritmov sledovania centrálnej trajektórie. V diplomovej práci [7] sme dokázali, že centrálna trajektória \mathcal{C} konverguje v smere zmenšovania parametra μ k optimálnemu riešeniu $(x^*, y^*, s^*) \in \mathcal{F}^*$. Dôkaz konvergenzie iteračných bodov v špecifickom okolí centrálnej trajektórie je podobný. Postupnosť iteračných bodov (x^k, y^k, s^k) je pozdĺž špecifického okolia ohraničená, a preto má aspoň jeden limitný bod (x^*, y^*, s^*) . Je zrejmé, že pre ľubovoľný iteračný bod (x^k, y^k, s^k) z okolia $\mathcal{N}_2(\theta)$ alebo $\mathcal{N}_{-\infty}(\gamma)$ platí

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

Iteračné body v špecifickom okolí $\mathcal{N}_2(\theta)$ navyše spĺňajú podmienku

$$0 \leq \|X^k S^k e - \mu_k e\| \leq \theta \mu_k.$$

Keďže množiny $\{x \in \mathbf{R}^n \mid Ax = b\}$, $\{(y, s) \in \mathbf{R}^m \times \mathbf{R}^n \mid A^T y + s = c\}$ sú uzavreté a $\mu_k \rightarrow 0$, tak limitným prechodom pre $k \rightarrow \infty$ dostaneme

$$Ax^* = b, \quad A^T y^* + s^* = c, \quad (x^*, s^*) \geq 0, \quad X^* S^* e = 0.$$

V prípade špecifického okolia $\mathcal{N}_{-\infty}(\gamma)$ je tiež možné ukázať, že limitný bod (x^*, y^*, s^*) spĺňa podmienku komplementarity $X^* S^* e = 0$, čo však vyžaduje pojem optimálneho rozkladu. To znamená, že limitný bod (x^*, y^*, s^*) v oboch prípadoch vyhovuje nutným a postačujúcim podmienkam optimality z Vety 1.2, a preto bod (x^*, y^*, s^*) je optimálnym riešením dvojice úloh (P) , (D) . Navyše, dvojica (x^*, s^*) zodpovedá ostro komplementárnemu primárno-duálnemu optimálnemu riešeniu. Tento výsledok úvádza aj nesledujúce tvrdenie, ktorého dôkaz je možné nájsť v knihe [29].

Veta 3.4. *Nech $\{(x^k, y^k, s^k)\}$ je postupnosť iteračných bodov generovaná algoritmom s krátkym krokom, prediktor-korektor algoritmom alebo algoritmom s dlhým krokom. Nech $\mu_k \rightarrow 0$ pre $k \rightarrow \infty$. Potom postupnosť $\{(x^k, y^k, s^k)\}$ je ohraničená, a preto má aspoň jeden limitný bod (x^*, y^*, s^*) , pričom každá dvojica (x^*, s^*) je ostro komplementárnym primárno-duálnym riešením úloh (P) , (D) . [29]*

3.7 Porovnanie algoritmov

Algoritmus s krátkym krokom je najjednoduchšou verziou primárno-duálnych algoritmov. Tento algoritmus volí v každej iterácii konštantnú hodnotu centrujúceho parametra σ a jednotkovú dĺžku kroku α . Určitou nadstavbou tohto algoritmu je prediktor-korektor algoritmus, ktorého iterácie alternujú medzi dvomi krokmi. Zatiaľ čo prediktor krok ($\sigma = 0$) pozdĺž afinno-škálovacieho smeru znižuje duálnu medzeru, korektor krok ($\sigma = 1$) pozdĺž centrujúceho smeru zlepšuje centrovanie smerom k centrálnej trajektórii. Oba algoritmy využívajú špecifické okolie $\mathcal{N}_2(\theta)$, ktorého reštriktívny charakter poskytuje algoritmom menej priestoru na manévrovanie v množine \mathcal{F}° . Dôsledkom toho je, že oba algoritmy generujú kratšie kroky, a preto sa iteračné body nachádzajú bližšie k centrálnej trajektórii. Z toho dôvodu sú algoritmy s krátkym krokom a prediktor-korektor algoritmy presnejšie, rýchlejšie konvergujú, pričom sa vyznačujú zložitou $O(\sqrt{n} \ln 1/\varepsilon)$.

Opačný charakter má algoritmus s dlhým krokom. Tento algoritmus zvyčajne volí centrujúci parameter σ blízko nuly, čo spôsobuje výraznú redukciu duálnej medzery v každej iterácii. Zároveň volí najväčšiu možnú dĺžku kroku α tak, aby ďalší iteračný bod zotrval v špecifickom okolí $\mathcal{N}_{-\infty}(\gamma)$. Okolie $\mathcal{N}_{-\infty}(\gamma)$ poskytuje algoritmu viac priestoru na manévrovanie, čo mechanizmus algoritmu prirodzene využíva, a preto generuje iteračné body ďalej od centrálnej trajektórie a bližšie k hranici množiny \mathcal{F}° . Z toho dôvodu sa môže stať, že iteračné body zostanú zaseknuté pri vrcholoch množiny \mathcal{F}° , čím sa konvergencia môže spomaliť, a preto algoritmus nemusí vždy spĺňať teoretické výsledky. Napriek tomu sa algoritmus s dlhým krokom v praxi často využíva, aj keď vykazuje o niečo horšiu zložitú $O(n \ln 1/\varepsilon)$.

NEPRÍPUSTNÉ ALGORITMY

Doteraz sme sa zaoberali algoritmami, pri ktorých sme vždy vychádzali zo štartovacieho bodu, ktorý bol ostro prípustným riešením z množiny \mathcal{F}° . Táto skutočnosť zaručila, že každý ďalší iteračný bod algoritmu bol tiež ostro prípustným riešením z množiny \mathcal{F}° . Avšak vo všeobecnosti ostro prípustné riešenie nemusí vždy existovať, a aj keď existuje, tak pri inicializácii algoritmu nemusí byť jednoduché nájsť ho. V tejto kapitole opíšeme triedu primárno-duálnych neprípustných algoritmov, ktoré síce využívajú podobný prístup ako algoritmy sledovania centrálnej trajektórie, ale nevyžadujú, aby štartovací bod bol ostro prípustným riešením. Uvedieme všeobecnú schému takéhoto algoritmu a sformulujeme tvrdenia, ktoré dokazujú jeho konvergenciu a polynomiálnu zložitosť. Zároveň opíšeme konkrétny variant neprípustného algoritmu – Mehrotrov prediktor-korektor algoritmus.

4.1 Neprípustnosť

Uvažujme konkrétnu primárnu úlohu

$$\min \{2x_1 - 3x_3 \mid 3x_1 + x_2 + x_3 = 4, x_2 + x_3 = 4, x \geq 0\},$$

kde množinou prípustných riešení je $\{(0, \eta, 4 - \eta) \mid \eta \in [0, 4]\}$. Keďže $x_1 = 0$, tak je zrejmé, že množina ostro prípustných riešení \mathcal{F}° je prázdna, a preto na riešenie takejto úlohy nie je možné aplikovať žiadny z doteraz prezentovaných algoritmov. Jedným zo spôsobov, ako predísť tomuto problému, je previesť danú úlohu do ekvivalentného tvaru vo väčšom rozmere, kde ostro prípustné riešenie existuje [14], [17], [20]. Iným prístupom je riešiť úlohu neprípustným algoritmom, ktorý nevyžaduje, aby štartovací bod (x^0, y^0, s^0) bol ostro prípustným riešením, požaduje iba splnenie podmienky $(x^0, s^0) > 0$. Zdôraznime, že v prípade ľubovoľného neprípustného bodu (x, y, s) podmienky prípustnosti nie sú splnené, a preto v prvých dvoch zložkách na pravej strane systému (1.11) vznikajú rezíduá r_P, r_D , kde

$$r_P = Ax - b, \quad r_D = A^T y + s - c.$$

Rezíduá vyčíslené v neprípustnom iteračnom bode (x^k, y^k, s^k) budeme označovať r_P^k, r_D^k .

4.2 Algoritmus

Neprípustný algoritmus sledovania centrálnej trajektórie je určitou modifikáciou algoritmu s dlhým krokom. Princíp sledovania centrálnej trajektórie naznačuje, že algoritmus vychádza z predpokladu $\mathcal{F}^\circ \neq \emptyset$, no napriek tomu štartovací bod (x^0, y^0, s^0) volí mimo

množiny \mathcal{F}° , pričom zachováva podmienku $(x^0, s^0) > 0$. Dôležitou súčasťou neprípustného algoritmu je špecifické okolie $\mathcal{N}_{-\infty}(\gamma, \varrho)$, ktoré zodpovedá rozšíreniu špecifického okolia $\mathcal{N}_{-\infty}(\gamma)$ obsahujúce aj neprípustné primárno-duálne riešenia. Okolie $\mathcal{N}_{-\infty}(\gamma, \varrho)$ je definované ako

$$\mathcal{N}_{-\infty}(\gamma, \varrho) = \left\{ (x, y, s) \mid \|(r_P, r_D)\| \leq \frac{\|(r_P^0, r_D^0)\|}{\mu_0} \varrho \mu, (x, s) > 0, x_i s_i \geq \gamma \mu, i = 1, 2, \dots, n \right\},$$

kde $\gamma \in (0, 1)$, $\varrho \geq 1$ a (r_P^0, r_D^0) a $\mu_0 > 0$ sú vyčíslené v štartovacom bode (x^0, y^0, s^0) . Všimnime si, že pre všetky body v okolí $\mathcal{N}_{-\infty}(\gamma, \varrho)$ je neprípustnosť rovnomerne ohraničená nejakým násobkom parametra μ . Z toho dôvodu môžeme predpokladať, že znižovaním parametra μ_k a udržiavaním iteračných bodov (x^k, y^k, s^k) v okolí $\mathcal{N}_{-\infty}(\gamma, \varrho)$ rezíduá konvergujú do nuly $(r_P^k, r_D^k) \rightarrow 0$ pre $k \rightarrow \infty$. Pri voľbe dĺžky kroku α_k implementujeme dodatočnú podmienku

$$\mu_k(\alpha_k) \leq (1 - 0.01\alpha_k)\mu_k, \quad (4.1)$$

ktorá zabezpečí, že hodnota parametra μ_k bude klesať aspoň nejakou minimálnou mierou. Pritom v každej iterácii algoritmu kontrolujeme okrem duálnej medzery aj primárnu a duálnu prípustnosť, a to pomocou vopred určených tolerančných konštánt $\epsilon_P > 0$, $\epsilon_D > 0$. Schéma neprípustného algoritmu teda vyzerá nasledovne.

Schéma neprípustného algoritmu

vstup $\gamma \in (0, 1)$, $\varrho \geq 1$, $\sigma_{\min}, \sigma_{\max} : 0 < \sigma_{\min} < \sigma_{\max} < 1$,
 $(x^0, y^0, s^0) : (x^0, s^0) > 0$, $(\epsilon, \epsilon_P, \epsilon_D) > 0$

prirad' $k := 0$

repeat

zvoľ' $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$

prirad' $\mu_k := (x^k)^T s^k / n$

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} r_P^k \\ r_D^k \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}$$

prirad' $\alpha_k := \max \{ \alpha \in [0, 1] \mid (x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \varrho), \mu_k(\alpha) \leq (1 - 0.01\alpha)\mu_k \}$

prirad' $(x^{k+1}, y^{k+1}, s^{k+1}) := (x(\alpha), y(\alpha), s(\alpha))$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \epsilon$ a $\|r_P^k\| \leq \epsilon_P$ a $\|r_D^k\| \leq \epsilon_D$

4.3 Konvergenca a polynomiálna zložitosť

Na základe uvedenej schémy neprípustného algoritmu môžeme pre rezíduá r_P^k, r_D^k odvodiť, že

$$\begin{aligned}
 r_P^k &= Ax^k - b \\
 &= A(x^{k-1} + \alpha_{k-1}\Delta x^{k-1}) - b \\
 &= (Ax^{k-1} - b) + \alpha_{k-1}(A\Delta x^{k-1}) \\
 &= r_P^{k-1} + \alpha_{k-1}r_P^{k-1} \\
 &= (1 - \alpha_{k-1})r_P^{k-1} \\
 \\
 r_D^k &= A^T y^k + s^k - c \\
 &= A^T(y^{k-1} + \alpha_{k-1}\Delta y^{k-1}) + s^{k-1} + \alpha_{k-1}\Delta s^{k-1} - c \\
 &= (A^T y^{k-1} + s^{k-1} - c) + \alpha_{k-1}(A^T \Delta y^{k-1} + \Delta s^{k-1}) \\
 &= r_D^{k-1} - \alpha_{k-1}r_D^{k-1} \\
 &= (1 - \alpha_{k-1})r_D^{k-1}.
 \end{aligned}$$

Z toho vyplýva, že pre dvojicu rezíduí (r_P^k, r_D^k) platí

$$\begin{aligned}
 (r_P^k, r_D^k) &= (1 - \alpha_{k-1})(r_P^{k-1}, r_D^{k-1}) \\
 &= (1 - \alpha_{k-1}) \dots (1 - \alpha_0)(r_P^0, r_D^0) \\
 &= \psi_k(r_P^0, r_D^0),
 \end{aligned} \tag{4.2}$$

kde $\psi_k = \prod_{j=0}^{k-1} (1 - \alpha_j)$ a $\psi_0 = 1$. Rovnica (4.2) je len technickým výsledkom, ktorý vedie k dôkazu konvergenencie [11] a polynomiálnej zložitosti [31] neprípustného algoritmu. Avšak formálny dôkaz oboch vlastností vyžaduje náročnejšie a rozsiahle technické postupy, a preto uvedieme len kľúčové výsledky s náznakom dôkazu, pričom celý dôkaz je možné nájsť v knihe [29].

Veta 4.1. *Postupnosťi $\{\mu_k\}$ a $\{\|(r_P^k, r_D^k)\|\}$ generované neprípustným algoritmom konvergujú do nuly.*

Kľúčovým medzikrokom je dokázať, že existuje kladná konštanta $\bar{\alpha} > 0$ taká, že $\alpha_k \geq \bar{\alpha}$ pre $k = 1, 2, \dots$. Potom z podmienky (4.1) a z rovnice (4.2) môžeme odvodiť, že

$$\begin{aligned}
 \mu_{k+1} &\leq (1 - 0.01\alpha_k)\mu_k \leq (1 - 0.01\bar{\alpha})\mu_k, \quad k = 1, 2, \dots, \\
 \|(r_P^{k+1}, r_D^{k+1})\| &\leq (1 - \bar{\alpha})\|(r_P^k, r_D^k)\|, \quad k = 1, 2, \dots,
 \end{aligned}$$

z čoho môžeme usúdiť, že ak $k \rightarrow \infty$, potom $\mu_k \rightarrow 0$ a $\|(r_P^k, r_D^k)\| \rightarrow 0$. Nepriamym dôsledkom tohto pozorovania je nasledujúce tvrdenie, ktoré formuluje konvergenčné vlastnosti neprípustného algoritmu.

Veta 4.2. *Nech $\{(x^k, y^k, s^k)\}$ je postupnosť iteračných bodov generovaná neprípustným algoritmom. Predpokladajme, že množina \mathcal{F}° je neprázdna. Potom postupnosť $\{(x^k, y^k, s^k)\}$ je ohraňovaná, a preto má aspoň jeden limitný bod (x^*, y^*, s^*) , pričom každá dvojica (x^*, s^*) je ostro komplementárnym primárno-duálnym riešením úloh (P), (D). [29]*

O polynomiálnej zložitosti hovorí nasledujúca veta.

Veta 4.3. *Nech $\varepsilon > 0$. Predpokladajme, že štartovací bod je tvaru $(x^0, y^0, s^0) = (\zeta e, 0, \zeta e)$, kde ζ spĺňa podmienku $\zeta^2 \leq \vartheta/\varepsilon^\kappa$ pre nejaké kladné konštanty $\vartheta, \kappa > 0$. Potom existuje index $\bar{k} = O(n^2 |\ln \varepsilon|)$ taký, že iteračné body (x^k, y^k, s^k) generované neprípustným algoritmom spĺňajú $\mu_k \leq \varepsilon, \forall k \geq \bar{k}$. [29]*

4.4 Mehrotrov prediktor-korektor algoritmus

Mehrotrov prediktor-korektor algoritmus [18] patrí k najúspešnejším primárno-duálnym algoritmom metód vnútorného bodu. Tento algoritmus sa odlišuje od generického algoritmu tým, že zlepšuje štandardný Newtonov smer o korekciu, ktorá je výpočtovo nenáročná a vedie k lepším numerickým výsledkom. Algoritmus v sebe zahŕňa niekoľko heuristických prvkov, ktoré boli do algoritmu implementované počas niekoľkých rokov výpočtových skúseností z praxe. Konkrétne ide o spojenie neprípustného algoritmu sledovania centrálnej trajektórie [15] a použitia aproximácie centrálnej trajektórie pomocou vyšších rádov [17], [21]. Navyše, Mehrotra doplnil túto kombináciu o adaptívny spôsob voľby centrujúceho parametra σ v každej iterácii. Výsledkom bol vysoko efektívny algoritmus, ktorý tvorí základný koncept väčšiny súčasných primárno-duálnych algoritmov metód vnútorného bodu.

Algoritmus

Mehrotrov prediktor-korektor algoritmus generuje postupnosť neprípustných iteračných bodov (x^k, y^k, s^k) , kde $(x^k, s^k) > 0$. Výsledný smer v každej iterácii je kombináciou niekoľkých prvkov – afinno-škálovací prediktor smer, centrujúca zložka zodpovedajúca adaptívne zvolenému parametru σ , korektor smer.

Pre iteračný bod (x^k, y^k, s^k) , $(x^k, s^k) > 0$ vypočítame afinno-škálovací prediktor smer $(\Delta x^{\text{af}}, \Delta y^{\text{af}}, \Delta s^{\text{af}})$, ktorý zodpovedá riešeniu systému

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^{\text{af}} \\ \Delta y^{\text{af}} \\ \Delta s^{\text{af}} \end{bmatrix} = - \begin{bmatrix} r_P^k \\ r_D^k \\ X^k S^k e \end{bmatrix}. \quad (4.3)$$

Afinno-škálovací smer $(\Delta x^{\text{af}}, \Delta y^{\text{af}}, \Delta s^{\text{af}})$ je teda štandardným Newtonovým smerom neprípustného algoritmu, ktorý získame voľbou $\sigma = 0$. Pre primárnu a duálnu zložku vypočítame samostatne dĺžku kroku pozdĺž smeru $(\Delta x^{\text{af}}, \Delta y^{\text{af}}, \Delta s^{\text{af}})$

$$\begin{aligned}\alpha_{af}^P &= \arg \max \{ \alpha \in [0, 1] \mid x^k + \alpha \Delta x^{af} \geq 0 \}, \\ \alpha_{af}^D &= \arg \max \{ \alpha \in [0, 1] \mid s^k + \alpha \Delta s^{af} \geq 0 \}.\end{aligned}$$

Aby sme mohli posúdiť kvalitu získaného afinno-škálovacieho smeru $(\Delta x^{af}, \Delta y^{af}, \Delta s^{af})$, definujme μ_{af} ako hypotetickú hodnotu parametra μ , ktorú by sme získali plným krokom pozdĺž smeru $(\Delta x^{af}, \Delta y^{af}, \Delta s^{af})$ až k hranici, teda

$$\mu_{af} = (x^k + \alpha_{af}^P \Delta x^{af})^T (s^k + \alpha_{af}^D \Delta s^{af}) / n.$$

Ak μ_{af} je výrazne menšie ako μ , potom afinno-škálovací smer $(\Delta x^{af}, \Delta y^{af}, \Delta s^{af})$ je vhodným smerom, pozdĺž ktorého sa duálna medzera výrazne zmenší, centrovanie nie je potrebné, a preto volíme parameter σ blízko 0. Na druhej strane, ak μ_{af} je len o niečo menšie ako μ , potom potrebujeme výraznejšie centrovanie, a preto volíme parameter σ blízko 1. Táto voľba vedie k centrovaniu blízko centrálnej trajektórie, čo algoritmu poskytuje lepšiu pozíciu na redukciu duálnej medzery v ďalšej iterácii. Mehrotra vo svojom pôvodnom algoritme [18] navrhol nasledovný heuristický spôsob voľby centrujúceho parametra

$$\sigma = \left(\frac{\mu_{af}}{\mu} \right)^3,$$

od ktorého sa ďalej odvíja korekčná zložka algoritmu. Aby sme vysvetlili jej význam, analyzujeme súčin $x_i^k s_i^k$, ktorý by sme získali plným krokom pozdĺž afinno-škálovacieho smeru $(\Delta x^{af}, \Delta y^{af}, \Delta s^{af})$. Dostaneme

$$(x_i^k + \Delta x_i^{af})(s_i^k + \Delta s_i^{af}) = x_i^k s_i^k + \underbrace{x_i \Delta s_i^{af} + s_i \Delta x_i^{af}}_{=-x_i^k s_i^k, \text{ podľa (4.3)} + \Delta x_i^{af} \Delta s_i^{af} = \Delta x_i^{af} \Delta s_i^{af}. \quad (4.4)$$

To znamená, že plným krokom pozdĺž afinno-škálovacieho smeru sa súčin $x_i^k s_i^k$ transformuje na $\Delta x_i^{af} \Delta s_i^{af}$ namiesto nulovej hodnoty. Korektor smer $(\Delta x^{kor}, \Delta y^{kor}, \Delta s^{kor})$ sa snaží korigovať túto odchýlku, čím by sa súčiny $x_i^k s_i^k$ priblížili k nule. Korektor smer $(\Delta x^{kor}, \Delta y^{kor}, \Delta s^{kor})$ zodpovedá riešeniu systému

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^{kor} \\ \Delta y^{kor} \\ \Delta s^{kor} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ \Delta X^{af} \Delta S^{af} e \end{bmatrix} \quad (4.5)$$

kde $\Delta X^{af} = \text{diag}(\Delta x^{af})$ a $\Delta S^{af} = \text{diag}(\Delta s^{af})$. Aby sme mohli posúdiť efektívnosť korektor smeru, uvažujme opäť súčin $x_i^k s_i^k$, ktorý by sme získali plným krokom pozdĺž kombinovaného afinno-škálovacieho-korektor smeru. Dostaneme

$$\begin{aligned}
 & (x_i^k + \Delta x_i^{\text{af}} + \Delta x_i^{\text{kor}})(s_i^k + \Delta s_i^{\text{af}} + \Delta s_i^{\text{kor}}) \\
 &= \underbrace{(x_i^k + \Delta x_i^{\text{af}})(s_i^k + \Delta s_i^{\text{af}})}_{=\Delta x_i^{\text{af}} \Delta s_i^{\text{af}}, \text{ podľa (4.4)}} + \underbrace{x_i \Delta s_i^{\text{kor}} + s_i \Delta x_i^{\text{kor}}}_{=-\Delta x_i^{\text{af}} \Delta s_i^{\text{af}}, \text{ podľa (4.5)}} + \Delta x_i^{\text{af}} \Delta s_i^{\text{kor}} + \Delta x_i^{\text{kor}} \Delta s_i^{\text{af}} + \Delta x_i^{\text{kor}} \Delta s_i^{\text{kor}} \\
 &= \Delta x_i^{\text{af}} \Delta s_i^{\text{kor}} + \Delta x_i^{\text{kor}} \Delta s_i^{\text{af}} + \Delta x_i^{\text{kor}} \Delta s_i^{\text{kor}}. \tag{4.6}
 \end{aligned}$$

Ak je matica koeficientov na ľavej strane systému (4.3) a (4.5) v limite pre $k \rightarrow \infty$ regulárna, potom výraz (4.6) je bližšie k nule ako výraz (4.6), z čoho pramo vyplýva efektivita korektor smeru. Avšak vo všeobecnosti matica koeficientov v limite nemusí byť vždy regulárna, no napriek tomu použitie korekcie prostredníctvom korektor smeru aj v takejto situácii zvyčajne zlepšuje celkovú efektivitu algoritmu v praxi. Nakoniec vypočítame kombinovaný centrujúci-korektor smer $(\Delta x^{\text{ck}}, \Delta y^{\text{ck}}, \Delta s^{\text{ck}})$, ktorý zodpovedá riešeniu systému

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^{\text{ck}} \\ \Delta y^{\text{ck}} \\ \Delta s^{\text{ck}} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ \Delta X^{\text{af}} \Delta S^{\text{af}} e - \sigma \mu e \end{bmatrix}.$$

Uvedená metóda tvorí podstatu pôvodného Mehrotrovho prediktor-korektor algoritmu [18] z roku 1992. Vysoká efektivita a presvedčivé numerické výsledky tohto algoritmu viedli prirodzene k snahám o jeho zlepšenie. Výsledkom boli rôzne modifikácie [16], [33], ktoré vznikali predovšetkým za účelom praktických implementácií. Na záver kapitoly uvedieme špecifický variant Mehrotrovho algoritmu, ktorý je implementovaný napríklad aj v programovom balíku LIPSOL [32].

Špecifická schéma Mehrotrovho prediktor-korektor algoritmu

vstup $(x^0, y^0, s^0) : (x^0, s^0) > 0, (\epsilon, \epsilon_P, \epsilon_D) > 0$

prirad' $k := 0$

repeat

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^{\text{af}} \\ \Delta y^{\text{af}} \\ \Delta s^{\text{af}} \end{bmatrix} = - \begin{bmatrix} r_P^k \\ r_D^k \\ X^k S^k e \end{bmatrix}$$

prirad' $\alpha_{\text{af}}^P := \arg \max \{ \alpha \in [0, 1] \mid x^k + \alpha \Delta x^{\text{af}} \geq 0 \}$

$\alpha_{\text{af}}^D := \arg \max \{ \alpha \in [0, 1] \mid s^k + \alpha \Delta s^{\text{af}} \geq 0 \}$

prirad' $\mu_{\text{af}} := (x^k + \alpha_{\text{af}}^P \Delta x^{\text{af}})^T (s^k + \alpha_{\text{af}}^D \Delta s^{\text{af}}) / n$

prirad' $\sigma := (\mu_{\text{af}} / \mu)^3$

vyrieš

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^{\text{ck}} \\ \Delta y^{\text{ck}} \\ \Delta s^{\text{ck}} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ \Delta X^{\text{af}} \Delta S^{\text{af}} e - \sigma \mu e \end{bmatrix}$$

prirad' $(\Delta x^k, \Delta y^k, \Delta s^k) := (\Delta x^{\text{af}}, \Delta y^{\text{af}}, \Delta s^{\text{af}}) + (\Delta x^{\text{ck}}, \Delta y^{\text{ck}}, \Delta s^{\text{ck}})$

prirad' $\alpha_{\max}^P := \arg \max \{ \alpha \geq 0 \mid x^k + \alpha \Delta x^k \geq 0 \}$

$\alpha_{\max}^D := \arg \max \{ \alpha \geq 0 \mid s^k + \alpha \Delta s^k \geq 0 \}$

prirad' $\alpha_k^P := \min(0.99 \alpha_{\max}^P, 1)$

$\alpha_k^D := \min(0.99 \alpha_{\max}^D, 1)$

prirad' $x^{k+1} = x^k + \alpha_k^P \Delta x^k$

$(y^{k+1}, s^{k+1}) = (y^k, s^k) + \alpha_k^D (\Delta y^k, \Delta s^k)$

prirad' $k := k + 1$

until $(x^k)^T s^k \leq \epsilon$ a $\|r_P^k\| \leq \epsilon_P$ a $\|r_D^k\| \leq \epsilon_D$

Paradoxom Mehrotrovho algoritmu je, že napriek jeho preukázanej výnimočnej efektívitě sa stále nepodarilo získať teoretické výsledky bežné pri iných, v praxi menej efektívnych algoritmoch. Teoretická analýza sa podarila zatiaľ len pri niektorých jeho miernych modifikáciách [1], [25].

DODATOK

5.1 Dôkaz Lemy 2.1

(i) Pre všetky $\beta > -1$ platí $\ln(1 + \beta) \leq \beta$, pričom rovnosť nastáva práve vtedy, keď $\beta = 0$.

Definujme pomocné funkcie $f(\beta) = \ln(1 + \beta)$, $g(\beta) = \beta$ pre $\beta > -1$. Graf funkcie g je dotyčnicou ku grafu funkcie f v bode $\beta = 0$. Je zrejmé, že funkcia f je konkávna, a preto dotyčnica ku grafu funkcie f v ľubovoľnom bode $\bar{\beta} > -1$ leží nad grafom funkcie f , pričom jediným spoločným bodom grafov je bod $(\bar{\beta}, f(\bar{\beta})) = (\bar{\beta}, g(\bar{\beta}))$. Z toho priamo vyplýva, že

$$\ln(1 + \beta) = f(\beta) \leq g(\beta) = \beta, \quad \forall \beta > -1,$$

pričom $\ln(1 + \beta) = \beta$ práve vtedy, keď $\beta = 0$.

(ii) Pre ľubovoľný vektor $z \in \mathbf{R}^n$, $\|z\|_\infty \leq \tau < 1$ platí

$$-\sum_{i=1}^n \ln(1 + z_i) \leq -e^T z + \frac{\|z\|^2}{2(1 - \tau)}.$$

Definujme pomocnú funkciu $f(z) = -\sum_{i=1}^n \ln(1 + z_i)$, ktorá je hladkou funkciou pre $\|z\|_\infty \leq \tau < 1$. Pre gradient ∇f a Hessovu maticu $\nabla^2 f$ zrejme platí

$$(\nabla f(z))_i = \frac{-1}{1 + z_i}, \quad (\nabla^2 f(z))_{ij} = \begin{cases} \frac{1}{(1+z_i)^2} & \text{ak } i = j \\ 0 & \text{ak } i \neq j. \end{cases}$$

Pomocou Taylorovho rozvoja funkcie f dostaneme

$$f(z) = f(0) + z^T \nabla f(0) + \frac{1}{2} \int_0^1 z^T \nabla^2 f(\eta z) z d\eta = -e^T z + \frac{1}{2} \int_0^1 z^T \nabla^2 f(\eta z) z d\eta.$$

Pre integrand $z^T \nabla^2 f(\eta z) z$ platí

$$z^T \nabla^2 f(\eta z) z = \sum_{i=1}^n \frac{z_i^2}{(1 + \eta z_i)^2} \leq \sum_{i=1}^n \frac{z_i^2}{(1 - \eta \tau)^2} = \frac{\|z\|^2}{(1 - \eta \tau)^2}.$$

Elementárnym integrovaním dostaneme

$$\int_0^1 z^T \nabla^2 f(\eta z) z d\eta \leq \|z\|^2 \int_0^1 \frac{d\eta}{(1 - \eta \tau)^2} = \|z\|^2 \frac{1}{\tau} \int_{1-\tau}^1 \frac{d\nu}{\nu^2} = \|z\|^2 \frac{1}{\tau} \left[-\frac{1}{\nu} \right]_{1-\tau}^1 = \frac{\|z\|^2}{1 - \tau},$$

z čoho nakoniec vyplýva, že

$$f(z) = -\sum_{i=1}^n \ln(1+z_i) = -e^T z + \frac{1}{2} \int_0^1 z^T \nabla^2 f(\eta z) z d\eta \leq -e^T z + \frac{\|z\|^2}{2(1-\tau)}.$$

5.2 Dôkaz Lemy 2.4

V dôkaze Lemy 2.4 sme využili skutočnosť, že vektory v a $V^{-1}e - v/\mu$ sú navzájom ortogonálne, teda že $v^T(V^{-1}e - v/\mu) = 0$. Tento výsledok dokážeme. Pripomeňme, že $v = Ve$, $V^T = V$, $V^T V = XS$ a $x^T s = n\mu$.

$$\begin{aligned} v^T \left(V^{-1}e - \frac{v}{\mu} \right) &= v^T V^{-1}e - \frac{v^T v}{\mu} \\ &= (Ve)^T V^{-1}e - \frac{(Ve)^T Ve}{\mu} \\ &= e^T V V^{-1}e - \frac{e^T V^T V e}{\mu} \\ &= n - \frac{e^T X S e}{\mu} \\ &= n - \frac{x^T s}{\mu} = 0 \end{aligned}$$

5.3 Odvodenie špecifického okolia pre $n = 2$

V ilustráciách Obr. 3.3 a Obr. 3.4 sme špecifické okolie centrálnej trajektórie zobrazili dvomi polpriamkami. Ukážeme, že v dvojrozmernom prípade ($n = 2$) špecifické okolia $\mathcal{N}_2(\theta)$, $\mathcal{N}_{-\infty}(\gamma)$ naozaj zodpovedajú prieniku dvoch polrovín v priestore $x_1 s_1 \times x_2 s_2$.

Špecifické okolie $\mathcal{N}_2(\theta)$

Podmienku presnosti $\|XSe - \mu e\|_2 \leq \theta\mu$ môžeme vyjadriť v tvare

$$(x_1 s_1 - \mu)^2 + (x_2 s_2 - \mu)^2 \leq (\theta\mu)^2. \quad (5.1)$$

Pre jednoduchosť označme $\phi_1 = x_1 s_1$, $\phi_2 = x_2 s_2$. Dosadením $\mu = \frac{x^T s}{n} = \frac{\phi_1 + \phi_2}{2}$ do nerovnice (5.1) dostaneme

$$\begin{aligned} \left(\frac{\phi_1 - \phi_2}{2} \right)^2 + \left(\frac{\phi_1 - \phi_2}{2} \right)^2 &\leq \theta^2 \left(\frac{\phi_1 + \phi_2}{2} \right)^2 \\ 2(\phi_1 - \phi_2)^2 &\leq \theta^2 (\phi_1 + \phi_2)^2 \\ \sqrt{2}|\phi_1 - \phi_2| &\leq \theta(\phi_1 + \phi_2). \end{aligned}$$

Z poslednej nerovnice môžeme elementárnymi úpravami odvodiť, že

$$\phi_2 \leq \left(\frac{\sqrt{2} + \theta}{\sqrt{2} - \theta} \right) \phi_1, \quad \phi_2 \geq \left(\frac{\sqrt{2} - \theta}{\sqrt{2} + \theta} \right) \phi_1.$$

Teda špecifickým okolím $\mathcal{N}_2(\theta)$ je množina ostro prípustných riešení z množiny \mathcal{F}° , ktoré spĺňajú podmienku

$$x_2 s_2 \leq \left(\frac{\sqrt{2} + \theta}{\sqrt{2} - \theta} \right) x_1 s_1, \quad x_2 s_2 \geq \left(\frac{\sqrt{2} - \theta}{\sqrt{2} + \theta} \right) x_1 s_1.$$

Špecifické okolie $\mathcal{N}_{-\infty}(\gamma)$

Keďže $\mu = \frac{x^T s}{2}$, tak podmienku presnosti $x_i s_i \geq \gamma \mu$, $i = 1, 2$ môžeme vyjadriť v tvare

$$x_1 s_1 \geq \gamma \frac{x_1 s_1 + x_2 s_2}{2}, \quad x_2 s_2 \geq \gamma \frac{x_1 s_1 + x_2 s_2}{2}.$$

Po elementárnych uprávách zistíme, že špecifickým okolím $\mathcal{N}_{-\infty}(\gamma)$ je množina ostro prípustných riešení z množiny \mathcal{F}° , ktoré vyhovujú podmienkam

$$x_2 s_2 \leq \left(\frac{2 - \gamma}{\gamma} \right) x_1 s_1, \quad x_2 s_2 \geq \left(\frac{\gamma}{2 - \gamma} \right) x_1 s_1.$$

5.4 Vzťah parametrov θ a σ v algoritme s krátkym krokom

V rámci algoritmov sledovania centrálnej trajektórie sme v schéme algoritmu s krátkym krokom použili konkrétne parametre $\theta = 0.4$ a $\sigma = 1 - 0.4/\sqrt{n}$. Dokážeme, že takto zvolené parametre θ a σ spĺňajú podmienku (3.12) pre $n \geq 1$.

$$\begin{aligned} \frac{\theta^2 + n(1 - \sigma)^2}{2^{3/2}(1 - \theta)} &\leq \sigma\theta \\ \frac{0.16 + n(0.4/\sqrt{n})^2}{0.6\sqrt{8}} &\leq 0.4 - \frac{0.16}{\sqrt{n}} \\ \frac{0.8}{0.6\sqrt{8}} &\leq 1 - \frac{0.4}{\sqrt{n}} \\ \frac{0.2}{0.3\sqrt{2}} &\leq 1 - \frac{0.4}{\sqrt{n}} \\ \frac{0.4}{\sqrt{n}} &\leq \frac{0.3\sqrt{2} - 0.2}{0.3\sqrt{2}} \\ n &\geq \left(\frac{0.12\sqrt{2}}{0.3\sqrt{2} - 0.2} \right)^2 \approx 0.76 \end{aligned}$$

Dokázali sme, že ak $\theta = 0.4$ a $\sigma = 1 - 0.4/\sqrt{n}$, potom podmienka (3.12) platí pre $n \geq 0.76$, a preto zrejme platí aj pre $n \geq 1$.

Záver

V rigoróznejšej práci sme sa zaoberali primárno-duálnymi algoritmi metód vnútorného bodu v lineárnom programovaní. V prvej kapitole sme zhrnuli teoretické výsledky z lineárneho programovania a metód vnútorného bodu. Bez ujmy na všeobecnosti sme vychádzali z primárno-duálnej dvojice lineárnych úloh (P) , (D) v štandardnom tvare. Sformulovali sme základné predpoklady $(P1)$, $(P2)$ a na základe systému (1.4)-(1.6) sme definovali centrálnu trajektóriu. V závere prvej kapitoly sme opísali všeobecný princíp primárno-duálnych metód, uviedli sme schému generického algoritmu a vysvetlili sme spôsob označenia zložitosti algoritmov. V druhej kapitole sme sa zaoberali algoritmi redukcie potenciálu, ktorých charakteristickým znakom je nepriamy prístup zmenšovania duálnej medzery pomocou potenciálovej funkcie. Definovali sme primárno-duálnu logaritmickejšiu potenciálovú funkciu Φ_ρ a uvideli sme jej dôležité špecifické vlastnosti. Dokázali sme, že mechanizmus algoritmu spôsobuje redukciu funkcie Φ_ρ , čo vedie k redukcii duálnej medzery. Odvodili sme horný odhad funkcie Φ_ρ pozdĺž prípustného smeru, čo nám umožnilo bližšie určiť odhad redukcie funkcie Φ_ρ v každej iterácii. Výsledkom bolo zistenie, že potenciálová funkcia Φ_ρ je v každej iterácii redukovaná aspoň konštantnou mierou. Na záver kapitoly sme dokázali, že algoritmus redukcie potenciálu má polynomiálnu zložitosť. Vo štvrtej kapitole sme sa venovali primárno-duálnym algoritmom sledovania centrálnej trajektórie. Charakteristickým znakom týchto algoritmov je princíp sledovania centrálnej trajektórie v jej špecifickom okolí, v ktorom postupnosť iteračných bodov konverguje k optimálnemu riešeniu. Zdefinovali sme dve štandardné špecifické okolia $\mathcal{N}_2(\theta)$, $\mathcal{N}_{-\infty}(\gamma)$ a uviedli sme ich vlastnosti. V jednorozmernom prípade sme ilustrovali kvalitatívny rozdiel špecifických okolí a dvojrozmernom prípade sme ilustrovali ich vzájomný vzťah. Pre algoritmy sledovania centrálnej trajektórie sme sformulovali a dokázali všeobecné tvrdenie o polynomiálnej zložitosti. V závislosti od voľby jednotlivých parametrov sme uviedli tri základné varianty algoritmu – algoritmus s krátkym krokom, prediktor-korektor algoritmus a algoritmus s dlhým krokom. Podrobne sme vysvetlili ich princíp, uviedli ich schému a dokázali ich vlastnosti. Zároveň sme ilustrovali ich mechanizmus a na základe Vety 3.1 sme dokázali ich polynomiálny charakter. V závere kapitoly sme sformulovali konvergenčné vlastnosti týchto algoritmov a uviedli sme ich vzájomné porovnanie. Vo štvrtej kapitole sme opísali triedu neprípustných primárno-duálnych algoritmov. Opísali sme všeobecnú schému neprípustného algoritmu, uviedli sme konvergenčné výsledky a sformulovali tvrdenie o polynomiálnej

zložitosti. V rámci tejto kapitoly sme tiež opísali pôvodný Mehrotrov prediktor-korektor algoritmus. Uviedli sme jeho princíp, ktorý spočíva v kombinácii korekcie smeru a adaptívneho spôsobu voľby centrujúceho parametra. V závere kapitoly sme uviedli aj schému špecifického variantu Mehrotrovho algoritmu.

V rigoróznej práci sme splnili ciele sformulované v úvode. Uviedli sme základnú kategorizáciu primárno-duálnych algoritmov metód vnútorného bodu v lineárnom programovaní, vysvetlili ich princíp a uviedli ich charakteristické znaky a vlastnosti. V práci sme sa zamerali na podrobnú analýzu algoritmov a dôkladné vysvetlenie, pričom sme dbali na jednotný spôsob prezentácie. V práci sme kládli dôraz na grafickú interpretáciu, o čom svedčí niekoľko zaujímavých ilustrácií. Poznamenajme, že rigorózna práca poskytuje len základný prehľad zaužívaných primárno-duálnych algoritmov. Z toho dôvodu je možné prácu ďalej doplniť o ďalšie podtriedy algoritmov, alebo ju rozšíriť o algoritmy, ktoré využívajú iný prístup.

Zoznam použitej literatúry

- [1] CARTIS, C. *Some disadvantages of a Mehrotra-type primal-dual corrector interior point algorithm for linear programming*. In *Applied Numerical Mathematics* [online]. 2009, Vol. 59, Iss. 5 [cit. 15.1.2014], 1110-1119 . Dostupné na internete: [〈http://www.optimization-online.org/DB_FILE/2005/02/1062.pdf〉](http://www.optimization-online.org/DB_FILE/2005/02/1062.pdf).
- [2] DANTZIG, G. *Lineárne programovanie a jeho rozvoj*. Bratislava: SVTL, 1966.
- [3] FREUND, R. *Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function*. In *Mathematical Programming* [online]. 1991, Vol. 51, Iss. 1-3 [cit. 8.1.2014], 203-222. Dostupné na internete: [〈http://link.springer.com/article/10.1007%2F01586933〉](http://link.springer.com/article/10.1007%2F01586933). ISSN 1436-4646.
- [4] GILL, P. E., MURRAY, W., SAUNDERS, M. A., TOMLIN, J. A., WRIGHT, M. H. *On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method*. In *Mathematical Programming* [online]. 1986, Vol. 36, Iss. 2 [cit. 8.1.2014], 183-209. Dostupné na internete: [〈http://link.springer.com/article/10.1007%2F02592025〉](http://link.springer.com/article/10.1007%2F02592025). ISSN 1436-4646.
- [5] GONZAGA, C. *Polynomial affine algorithms for linear programming*. In *Mathematical Programming* [online]. 1990, Vol. 49, Iss. 1-3 [cit. 8.1.2014], 7-21. Dostupné na internete: [〈http://link.springer.com/article/10.1007/BF01588776〉](http://link.springer.com/article/10.1007/BF01588776). ISSN 1436-4646.
- [6] GONZAGA, C., TODD, M. J. *An $O(\sqrt{n}L)$ -iteration large-step primal-dual affine algorithm for linear programming*. In *SIAM Journal on Optimization* [online]. 1992, Vol. 2, Iss. 3 [cit. 8.1.2014], 349-359. Dostupné na internete: [〈http://epubs.siam.org/doi/abs/10.1137/0802017〉](http://epubs.siam.org/doi/abs/10.1137/0802017). ISSN 1095-7189.
- [7] KABÁT, M. *Metódy vnútorného bodu v lineárnom programovaní a ich aplikácie vo financiách* [Diplomová práca]. Bratislava: Univerzita Komenského, Fakulta matematiky, fyziky a informatiky, 2013. Dostupné na internete: [〈http://www.marekabat.sk〉](http://www.marekabat.sk).

- [8] KARMARKAR, N. *A new polynomial-time algorithm for linear programming*. In *Combinatorica* [online]. 1984, Vol. 4, Iss. 4 [cit. 31.3.2013], 373-395. Dostupné na internete: [⟨http://link.springer.com/article/10.1007/BF02579150⟩](http://link.springer.com/article/10.1007/BF02579150). ISSN 1439-6912.
- [9] KHACHIYAN, L. *A polynomial algorithm in linear programming*. In *Doklady Akademii Nauk SSSR*. 1979, Vol. 244, No. 5, 1093-1096. ISSN 0002-3264.
- [10] KLEE, V., MINTY, G. *How good is the simplex algorithm?* In SHISHA, O. *Inequalities III*. New York: Academic Press, 1972. 159-175.
- [11] KOJIMA, M., MEGIDDO, N., MIZUNO. S. *A primal-dual infeasible-interior-point algorithm for linear programming*. In *Mathematical Programming* [online]. 1993, Vol. 61, Iss. 1-3 [cit. 12.1.2014], 263-280. Dostupné na internete: [⟨http://link.springer.com/article/10.1007%2FBF01582151⟩](http://link.springer.com/article/10.1007%2FBF01582151). ISSN 1436-4646.
- [12] KOJIMA, M., MIZUNO. S., YOSHISE, A. *An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems*. In *Mathematical Programming* [online]. 1991, Vol. 50, Iss. 1-3 [cit. 29.12.2013], 331-342. Dostupné na internete: [⟨http://link.springer.com/article/10.1007/BF01594942⟩](http://link.springer.com/article/10.1007/BF01594942). ISSN 1436-4646.
- [13] KOJIMA, M., MIZUNO. S., YOSHISE, A. *A polynomial-time algorithm for a class of linear complementarity problems*. In *Mathematical Programming* [online]. 1989, Vol. 44, Iss. 1-3 [cit. 11.1.2014], 1-26. Dostupné na internete: [⟨http://link.springer.com/article/10.1007/BF01587074⟩](http://link.springer.com/article/10.1007/BF01587074). ISSN 1436-4646.
- [14] KOJIMA, M., MIZUNO. S., YOSHISE, A. *A primal-dual interior point algorithm for linear programming*. In *Progress in Mathematical Programming: Interior-Point and Related Methods*. New York: Springer, 1989, ISBN 978-1-4613-9617-8, 29-47.
- [15] LUSTIG, I. J., MARSTEN, R. E., SHANNO, D. F. *Computational experience with a primal-dual interior point method for linear programming*. In *Linear Algebra and its Applications* [online]. 1991, Vol. 152 [cit. 13.1.2014], 191-222. Dostupné na internete: [⟨http://www.sciencedirect.com/science/article/pii/0024379591902752⟩](http://www.sciencedirect.com/science/article/pii/0024379591902752). ISSN 0024-3795.
- [16] LUSTIG, I. J., MARSTEN, R. E., SHANNO, D. F. *On implementing Mehrotra's predictor-corrector interior point method for linear programming*. In *SIAM Journal on Optimization* [online]. 1992, Vol. 2, No. 3 [cit. 31.3.2013], 435-449. Dostupné na internete: [⟨http://epubs.siam.org/toc/sjope8/2/3⟩](http://epubs.siam.org/toc/sjope8/2/3). ISSN 1095-7189.
- [17] MEGIDDO, N. *Pathways to the optimal set in linear programming*. In *Progress in Mathematical Programming: Interior-Point and Related Methods* [online]. New York: Springer, 1989 [cit. 11.1.2014], 131-158. Dostupné na internete: [⟨http://theory.stanford.edu/~megiddo/bio.html⟩](http://theory.stanford.edu/~megiddo/bio.html). ISBN 978-1-4613-9619-2.

- [18] MEHROTRA, S. *On the implementation of a primal-dual interior point method*. In *SIAM Journal on Optimization* [online]. 1992, Vol. 2, No. 4 [cit. 31.3.2013], 576-601. Dostupné na internete: <http://epubs.siam.org/toc/sjope8/2/4>. ISSN 1095-7189.
- [19] MIZUNO, S., TODD, M., YE, Y. *On adaptive step primal-dual interior-point algorithms for linear programming*. In *Mathematics of Operations Research* [online]. 1993, Vol. 18, Iss. 4 [cit. 11.1.2014], 964-981. Dostupné na internete: <http://www.jstor.org/stable/3690133>.
- [20] MONTEIRO, R. D. C., ADLER, I. *Interior path-following primal-dual algorithms. Part I: Linear programming*. In *Mathematical Programming* [online]. 1989, Vol. 44, Iss. 1-3 [cit. 11.1.2014], 27-41. Dostupné na internete: <http://www2.isye.gatech.edu/~monteiro/publications>. ISSN 1436-4646.
- [21] MONTEIRO, R. D. C., ADLER, I., RESENDE, M. G. C. *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*. In *Mathematics of Operations Research* [online]. 1990, Vol. 15, Iss. 2 [cit. 13.1.2014], 191-214. Dostupné na internete: <http://dx.doi.org/10.1287/moor.15.2.191>.
- [22] RENEGAR, J. *A polynomial-time algorithm, based on Newton's method, for linear programming*. In *Mathematical Programming* [online]. 1988, Vol. 40, Iss. 1-3 [cit. 8.1.2014], 59-93. Dostupné na internete: <http://link.springer.com/article/10.1007%2F01580724>. ISSN 1436-4646.
- [23] ROOS, C., TERLAKY, T., VIAL, J-Ph. *Interior point methods for linear optimization*. New York: Springer, 2005. ISBN 978-0387-26379-3.
- [24] ROOS, C., TERLAKY, T., VIAL, J-Ph. *Theory and algorithms for linear optimization: an interior point approach*. New York: John Wiley & Sons, 1997. ISBN 978-0471-95676-1.
- [25] SALAHI, M., TERLAKY, T. *Mehrotra-type predictor-corrector algorithm revisited*. In *Optimization Methods and Software* [online]. 2008, Vol. 23, Iss. 2 [cit. 15.1.2014], 259-273. Dostupné na internete: http://www.optimization-online.org/DB_FILE/2007/02/1577.pdf.
- [26] TANABE, K. *Centered Newton method for mathematical programming*. In *System Modeling and Optimization: Proceedings of the 13th IFIP Conference Tokyo, Lecture Notes in Control and Information Sciences*, New York: Springer-Verlag, 1988. ISBN 978-3-540-39164-7, Vol. 113, 197-206.

- [27] TODD, M. J., BURRELL, B. P. *An extension of Karmarkar's algorithm for linear programming using dual variables*. In *Algorithmica I*. [online]. 1986, Vol. 1, Iss. 1-4 [cit. 8.1.2014], 409-424. Dostupné na internete: <http://link.springer.com/article/10.1007%2F01840455>. ISSN 1432-0541.
- [28] TODD, M. J., YE, Y. *A centered projective algorithm for linear programming*. In *Mathematics of Operations Research* [online]. 1990, Vol. 15, Iss. 3 [cit. 29.12.2013], 508-529. Dostupné na internete: <http://www.jstor.org/discover/10.2307/3689994>.
- [29] WRIGHT, J. S. *Primal-Dual Interior-Point Methods*. Philadelphia, USA: Society for Industrial and Applied Mathematics, 1997. ISBN 0-89871-328-X.
- [30] YE, E. *An $O(n^3L)$ potential reduction algorithm for linear programming*. In *Mathematical Programming* [online]. 1991, Vol. 50, Iss. 1-3 [cit. 8.1.2014], 239-258. Dostupné na internete: <http://link.springer.com/article/10.1007%2F01594937>. ISSN 1436-4646.
- [31] ZHANG, Y. *On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problems*. In *SIAM Journal on Optimization* [online]. 1994, Vol. 4, Iss. 1 [cit. 12.1.2014], 208-227. Dostupné na internete: <http://epubs.siam.org/doi/abs/10.1137/0804012>. ISSN 1095-7189.
- [32] ZHANG, Y. *User's Guide to LIPSOL: Linear-programming Interior Point Solvers v0.4* [online]. 1998. [cit. 31.3.2013]. Dostupné na internete: <http://www.caam.rice.edu/~zhang/lipsol>.
- [33] ZHANG, Y., ZHANG, D. *On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms*. In *Mathematical Programming* [online]. 1995, Vol. 68, Iss. 1-3 [cit. 31.3.2013], 303-318. Dostupné na internete: <http://link.springer.com/article/10.1007%2F01585769>. ISSN 1436-4646.