# Metoda unutrašnje točke u kvadratičnom programiranju

**Marodi, Petra**

**SVEUČILIŠTE U ZAGREBU**

**PRIRODOSLOVNO–MATEMATIČKI FAKULTET**

**MATEMATIČKI ODSJEK**

Petra Marodi

# METODA UNUTRAŠNJE TOČKE U KVADRATIČNOM PROGRAMIRANJU

Diplomski rad

Voditelj rada:
izv. prof. dr. sc. Marko Vrdoljak

Zagreb, rujan 2019.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom _______________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

*Zahvaljujem mentoru izv. prof. dr. sc. Marku Vrdoljaku na pomoći i savjetima.*
*Hvala mojim roditeljima, obitelji i Vedranu, bili ste mi velika podrška.*

# Sadržaj

# Uvod

U ovom radu obrađuje se tema iz područja optimizacije - metoda unutrašnje točke u linearnom i kvadratičnom programiranju. Metoda unutrašnje točke je metoda koja je, uz simpleks metodu, jedna od najuspješnijih za rješavanje zadaća linaernog programiranja, a uspješno je primjenjiva i na nelinearne zadaće. Naziv je dobila po tome što kada tražimo optimalno rješenje, svaka dopustiva točka na koju naiđemo se nalazi u unutrašnjosti dopustivog skupa. Najveći doprinos toj metodi dali su Yurii Nesterov i Arkadi Nemirovskii koji su ovu metodu primijenili na široku klasu problema, ali i Narendra K. Karmarkar koji je dokazao da je ovom metodom linearna zadaća rješiva u polinomijalnom vremenu.

U prvom poglavlju opisuje se navedena metoda za zadaće linearnog programiranja i uvodi se metoda centralnog puta kao specijalni slučaj metode unutrašnje točke. Za nalaženje novih točaka koristimo Newtonovu metodu i detaljnije analiziramo njezin određeni korak. Na kraju dajemo algoritam primarno-dualne metode unutrašnje točke, odnosno centralnog puta.

U drugom poglavlju primjenjujemo metodu unutrašnje točke na zadaće kvadratičnog programiranja. Kao i u prvom poglavlju, koristimo Newtonovu metodu za iteriranje novih točaka i fokusiramo se na metodu centralnog puta. Opisujemo okolinu centralnog puta, navodimo dva algoritma za nalaženje optimalnih točaka primarno-dualnom metodom centralnog puta i na kraju na primjeru demonstriramo jedan od njih.

# Poglavlje 1

# Linearno programiranje

## 1.1 Uvod

U ovom poglavlju promatrat ćemo zadaću linearnog programiranja u standardnoj formi:

$$\begin{aligned} c^T x &\to \min \\ Ax &= b \\ x &\geq 0\,, \end{aligned} \tag{1.1}$$

gdje je $A \in M_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ i $c \in \mathbb{R}^n$, pri čemu je $m < n$.

Dualni problem zadaće (1.1) je

$$\begin{aligned} b^T y &\to \max \\ A^T y &\leq c\,, \end{aligned}$$

koja se nakon uvođenja pomoćnih varijabli $s \in \mathbb{R}^n_+$ ekvivalentno može zapisati kao

$$\begin{aligned} b^T y &\to \max \\ A^T y + s &= c \\ s &\geq 0\,. \end{aligned} \tag{1.2}$$

Svaki vektor $x \in \mathbb{R}^n$ koji zadovoljava uvjete u (1.1) zove se dopustiva točka primarne zadaće. Slično, $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ koji zadovoljava uvjete u (1.2) naziva se dopustiva točka dualne zadaće. Ako je $x$ dopustiv za primarnu zadaću i $(y, s)$ dopustiv za dualnu zadaću, onda imamo

$$0 \leq x^T s = x^T(c - A^T y) = c^T x - (Ax)^T y = c^T x - b^T y. \tag{1.3}$$

Slijedi, $c^T x \geq b^T y$.

Sljedeći teorem govori da se "obično" optimalne vrijednosti zadaća (1.1) i (1.2) podudaraju.

**Teorem 1.** *Ako primarna zadaća* (1.1) *i dualna zadaća* (1.2) *imaju dopustive točke, onda obje zadaće imaju optimalna rješenja* $x^*$ *i* $(y^*, s^*)$ *respektivno. Nadalje, u tom slučaju vrijedi* $c^T x^* = b^T y^*$*, tj. optimalne vrijednosti za obje zadaće su jednake.*

Iz Teorema 1 (teorem dualnosti) lako se mogu izvesti uvjeti komplementarnosti za zadaće (1.1) i (1.2): ako je $x$ dopustiv za primarnu zadaću, a $(y, s)$ za dualnu, onda je $x$ optimalan za (1.1) i $(y, s)$ optimalan za (1.2) ako i samo ako

$$x_i s_i = 0, \; i = 1, \ldots, n. \tag{1.4}$$

Nadalje, označit ćemo s

$$\begin{aligned} P &= \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \\ D &= \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c, s \geq 0\} \end{aligned}$$

skupove dopustivih rješenja za zadaće (1.1) i (1.2) respektivno. Koristit ćemo i sljedeće oznake za skupove strogo dopustivih rješenja za (1.1) i (1.2):

$$\begin{aligned} P^+ &= \{x \in \mathbb{R}^n : Ax = b, x > 0\} \\ D^+ &= \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^n : A^T y + s = c, s > 0\}\,. \end{aligned}$$

U ovom poglavlju pretpostavit ćemo da su zadovoljene sljedeće tvrdnje: matrica $A$ je punog ranga, tj. $r(A) = m$ i obje zadaće (1.1) i (1.2) imaju strogo dopustiva rješenja, tj. pripadni skupovi $P^+$ i $D^+$ su neprazni. Teorem 1 tada implicira da obje zadaće (1.1) i (1.2) imaju optimalna rješenja. Kasnije ćemo pokazati da tom pretpostavkom ne gubimo na općenitosti.

Nadalje, koristit ćemo sljedeću notaciju: ako je $x \in \mathbb{R}^n$, tada je

$$X = diag(x) = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}.$$

Vektor jedinica označavat ćemo sa $e$, tj. $e = (1, \ldots, 1)^T$.

## 1.2 Centralni put

Iz teorema dualnosti u prethodnom poglavlju slijedi da su $x \in \mathbb{R}^n$ i $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ optimalna rješenja za (1.1) i (1.2) ako i samo ako zadovoljavaju

$$Ax = b \tag{1.5}$$
$$A^T y + s = c \tag{1.6}$$
$$Xs = 0 \tag{1.7}$$
$$x \geq 0\,, s \geq 0 \tag{1.8}$$

Jednadžba (1.5) zajedno s nenegativnošću od $x$ u (1.8) predstavlja dopustivost primarne, dok (1.6) sa nenegativnošću od $s$ predstavlja dopustivost dualne zadaće. Jednadžba (1.7) je zapravo uvjet komplementarnosti ekvivalentan s (1.4).

Za parametar $\mu \in \mathbb{R}$, $\mu \geq 0$ definiramo

$$P_\mu(x, y, s) = \begin{pmatrix} Ax - b \\ A^T y + s - c \\ Xs - \mu e \end{pmatrix}. \tag{1.9}$$

Tada je (1.5)-(1.8) ekvivalentno sa

$$P_0(x, y, s) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ x \geq 0, s \geq 0\,. \tag{1.10}$$

Za daljnu analizu problema koristit ćemo sljedeći rezultat koji daje uvjete regularnosti na Jacobijevu matricu.

**Teorem 2.** *Ako je matrica A ranga m, tada je za neki $\mu \in \mathbb{R}$, $\mu \geq 0$ Jacobijeva matrica*

$$DP_\mu(x, y, s) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix} \tag{1.11}$$

*regularna ako je $x > 0$ i $s > 0$.*

*Dokaz.* Neka su $(u, v, w) \neq (0, 0, 0)$ takvi da

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = DP_\mu(x, y, s) \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} Au \\ A^T v + w \\ Su + Xw \end{pmatrix},$$

pri čemu smo sa $S$ označili matricu $diag(s_1, \dots, s_n)$. Tada imamo

$$u^T w = u^T(-A^T v) = -(Au)^T v = 0\,, \tag{1.12}$$

jer je $Au = 0$.

Rješavanjem $Su + Xw = 0$ po $u$ slijedi $u = -S^{-1}Xw$, pa korištenjem (1.12) dobivamo

$$0 = u^T w = w^T u = -w^T S^{-1} X w. \tag{1.13}$$

Kako su $x > 0$ i $s > 0$, matrica $S^{-1}X$ je dijagonalna i pozitivno definitna, pa iz (1.13) zaključujemo da je $w = 0$. No tada iz $0 = Su + Xw = Su$ i $s > 0$ slijedi $u = 0$, a iz $0 = A^T v + w = A^T v$ slijedi $v = 0$ (pretpostavili smo na početku poglavlja da je matrica $A$ ranga $m$), što je u kontradikciji s pretpostavkom.

□

Promotrimo sada sljedeći relaksirani sustav za neki $\mu > 0$:

$$P_\mu(x, y, s) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; x \geq 0, s \geq 0. \tag{1.14}$$

Primijetimo da je (1.14) sustav sa $m + 2n$ jednadžbi i jednako toliko varijabli (ako zanemarimo uvjete nenegativnosti) i prirodno je pokušati ga riješiti Newtonovom metodom:

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ y^k \\ s^k \end{pmatrix} - DP_\mu(x^k, y^k, s^k)^{-1} \, P_\mu(x^k, y^k, s^k).$$

Iz $x^k > 0$ i $s^k > 0$ slijedi da su $x^{k+1}$, $y^{k+1}$ i $s^{k+1}$ dobro definirani. Nadalje, pokazat ćemo da sustav (1.14) ima jedinstveno rješenje.

**Lema 3.** *Pretpostavimo da je skup $D = \{(y, s) : A^T y + s = c, s \geq 0\}$ omeđen i $D^+ = \{(y, s) : A^T y + s = c, s > 0\}$ neprazan i omeđen. Tada za $\mu > 0$ sustav (1.14) ima jedinstveno rješenje $(x(\mu), y(\mu), s(\mu))$.*

*Dokaz.* Označimo stupce matrice $A$ sa $a_1, \ldots, a_n$ i promotrimo logaritamsku barijernu funkciju:

$$\phi(y) := -\sum_{i=1}^{n} ln \, (c_i - a_i^T \, y).$$

Tada je

$$D\phi(y) = \sum_{i=1}^{n} \frac{a_i}{c_i - a_i^T \, y} \tag{1.15}$$

$$D^2\phi(y) = \sum_{i=1}^{n} \frac{a_i a_i^T}{(c_i - a_i^T \, y)^2} \tag{1.16}$$

Iz (1.16) vidimo da je $D^2\phi(y)$ pozitivno definitna na $\mathbb{R}^m$. Naime, kako je $A$ ranga $m$, za bilo koji $z \neq 0$ postoji bar jedan $i$ takav da je $a_i^T z \neq 0$ (u suprotnom je $a_i^T z = 0$ za sve $i$, pa je $A^T z = 0$, što je u kontradikciji s činjenicom da je rang od $A$ jednak $m$). Za $z \neq 0$ tada imamo

$$z^T D^2\phi(y)z = \sum_{i=1}^{n} \frac{(a_i^T z)^2}{(c_i - a_i^T y)^2} > 0\,.$$

Kako je $D^2\phi(y)$ pozitivno definitna, slijedi da je $\phi$ strogo konveksna na

$$S_D := \{y : A^T y \leq c\}\,.$$

Nadalje, kako $\phi(y) \to +\infty$ za $y \to \partial S_D$ (bar jedan $c_i - a_i^T y$ teži u 0, dok su ostali ograničeni odozgo zbog ograničenosti skupa $D$), možemo zaključiti da $\phi$ ima jedinstveni minimum na $S_D$. Funkciji $\phi$ možemo dodati ili oduzeti linearni član bez narušavanja stroge konveksnosti i činjenice da teži u $+\infty$ kada promatramo rub skupa $S_D$. Zbog toga, funkcija

$$\Phi(y) := \phi(y) - \frac{1}{\mu} b^T y$$

ima jedinstveni minimum $\overline{y}$ u unutrašnjosti skupa $S_D$ koji zadovoljava

$$0 = D\Phi(\overline{y}) = \sum_{i=1}^{n} \frac{a_i}{c_i - a_i^T \overline{y}} - \frac{1}{\mu} b\,. \tag{1.17}$$

Definiramo $\overline{x}, \overline{s} \in \mathbb{R}^n$ kao

$$\overline{x_i} := \frac{\mu}{c_i - a_i^T \overline{y}} > 0 \tag{1.18}$$

$$\overline{s_i} := c_i - a_i^T \overline{y} > 0\,. \tag{1.19}$$

Iz definicije (1.19) imamo $A^T\overline{y} + \overline{s} - c = 0$. Štoviše, iz (1.17) i (1.18)

$$A\overline{x} = \sum_{i=1}^{n} a_i \overline{x_i} = \sum_{i=1}^{n} a_i \frac{\mu}{c_i - a_i^T \overline{y}} = \mu \sum_{i=1}^{n} \frac{a_i}{c_i - a_i^T \overline{y}} = b$$

Konačno, očito je $\overline{x_i}\,\overline{s_i} = \mu$, za $i = 1, \ldots, n$. Slijedi da $(\overline{x}, \overline{y}, \overline{s})$ rješava sustav (1.14) i taj sustav ima bar jedno rješenje.

Pretpostavimo sada da je $(x, y, s)$ rješenje sustava (1.14). Primijetimo da za zadani $y \in \mathbb{R}^n$ postoji najviše jedan $s > 0$ koji zajedno s $y$ zadovoljava $A^T y + s = c$ i $s > 0$. Jednom kad je određen $s > 0$, on jedinstveno određuje i $x > 0$. Dakle, ako pokažemo da se u $y$ postiže jedinstveni miminum od $\Phi$, pokazali smo da (1.14) ima jedinstveno rješenje. Imamo

$$\begin{aligned}
D\Phi(y) &= \sum_{i=1}^{n} \frac{a_i}{c_i - a_i^T y} - \frac{1}{\mu} b \\
&= \sum_{i=1}^{n} \frac{a_i}{s_i} - \frac{1}{\mu} b && \text{jer je } a_i^T y + s_i = c_i \\
&= \sum_{i=1}^{n} \frac{1}{\mu} a_i x_i - \frac{1}{\mu} b && \text{jer } s_i x_i = \mu \\
&= 0 && \text{jer } \sum_{i=1}^{n} a_i x_i = Ax = b\,.
\end{aligned}$$

Kako je $\Phi$ stogo konveksna, slijedi da se minimum postiže jedino u $y$, što dokazuje tvrdnju. □

**Definicija 4.** *Skup* $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$ *rješenja sustava (1.14) zove se (primarno-dualni) centralni put.*

Primijetimo da za $((x(\mu), y(\mu), s(\mu))$ na centralnom putu imamo

$$n\mu = \sum_{i=1}^{n} x_i(\mu)s_i(\mu) = x(\mu)^T s(\mu) = x(\mu)^T (c - A^T y(\mu)) = c^T x(\mu) - y^T Ax(\mu) = c^T x(\mu) - b^T y(\mu)\,.$$

## 1.3 Newtonova metoda za centralni put

Promotrimo sada rješenje sustava (1.14) za fiksirani $\mu > 0$ dobiveno Newtonovom metodom (kao što smo razmatrali iza jednadžbe (1.14), dobro je definirana $(k+1)$. iteracija jer je $DP_\mu$ regularna, a u Korolaru 11 ćemo pokazati da vrijedi $x^{k+1} > 0$ i $s^{k+1} > 0$).

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{pmatrix} = \begin{pmatrix} x^{k} \\ y^{k} \\ s^{k} \end{pmatrix} - DP_\mu(x^k, y^k, s^k)^{-1} P_\mu(x^k, y^k, s^k),$$

odnosno

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = -DP_\mu(x^k, y^k, s^k)^{-1} P_\mu(x^k, y^k, s^k).$$

Pretpostavimo da su u jednom koraku dobiveni $x := x^k, y := y^k$ i $s := s^k$ takvi da je

$$\begin{aligned} Ax &= b \\ A^T y + s &= c\,. \end{aligned}$$

Definiramo *reziduum* $r$ sa

$$r := r(x, y, s) := Xs - \mu e. \tag{1.20}$$

Newtonov korak $(\Delta x, \Delta y, \Delta s)$ iz $(x, y, s)$ je dan kao rješenje od

$$DP_\mu(x^k, y^k, s^k) \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix},$$

što se pomoću (1.11) može zapisati kao

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix}. \tag{1.21}$$

Sljedeću iteraciju Newtonove metode dobivamo kao

$$(x^+, y^+, s^+) := (x, y, s) + (\Delta x, \Delta y, \Delta s)\,.$$

Mi ćemo se kasnije usredotočiti na reziduum $r^+ := X^+ s^+ - \mu e$ koji odgovara sljedbeniku $(x^+, y^+, s^+)$ u odgovaraćoj iteraciji Newtonove metode.

**Napomena 5.** *Do jednadžbe (1.21) mogli smo ekvivalentno doći na drugi način:*

$$\begin{aligned} P_\mu(x + \Delta x, y + \Delta y, s + \Delta s) &= \begin{pmatrix} Ax + A\Delta x - b \\ A^T y + A^T \Delta y + s + \Delta s - c \\ Xs + X\Delta s + S\Delta x + \Delta X \Delta s - \mu e \end{pmatrix} \\ &= \begin{pmatrix} A\Delta x \\ A^T \Delta y + \Delta s \\ Xs + X\Delta s + S\Delta x + \Delta X \Delta s - \mu e \end{pmatrix} \\ &\approx \begin{pmatrix} A\Delta x \\ A^T \Delta y + \Delta s \\ Xs + X\Delta s + S\Delta x - \mu e \end{pmatrix} \\ &\overset{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \end{aligned}$$

*gdje druga jednakost vrijedi jer je $Ax = b$ i $A^T y + s = c$, dok u trećem redu zbog aproksimacije ispuštamo nelinearni član.*

## 1.4 Ortogonalne projekcije

Za analizu koraka Newtonove metode potrebne su ortogonalne projekcije na neke prostore, pa ćemo navesti rezultate koji će nam biti potrebni u daljnjim razmatranjima. Odsad pa nadalje pretpostavit ćemo da je $B$ $m \times n$ matrica ranga $m$.

**Lema 6.** *Neka je $R(B^T) \subseteq \mathbb{R}^n$ slika od $B^T$ i $N(B) \subseteq \mathbb{R}^n$ jezgra od $B$. Tada su $R(B^T)$ i $N(B)$ ortogonalni komplementi, tj.*

$$R(B^T) \perp N(B) \quad i \quad R(B^T) \oplus N(B) = \mathbb{R}^n \,.$$

*Dokaz.* Neka je $u \in R(B^T)$, tj. $u = B^T w$ i neka je $v \in N(B)$. Tada je $u^T v = w^T Bv = w^T 0 = 0$, dakle $R(B^T) \perp N(B)$. Uočimo da je $\dim R(B^T) = r(B^T) = r(B)$, dok je $\dim N(B) = d(B)$, pa po teoremu o rangu i defektu imamo $r(B) + d(B) = n$, iz čega slijedi tvrdnja. □

Za $x \in \mathbb{R}^n$ definiramo $r := B^T(BB^T)^{-1}Bx \in R(B^T)$. Matrica $BB^T$ je pozitivno definitna: $(BB^T x) \cdot x = B^T x \cdot B^T x = \|B^T x\|_2^2 > 0$ (jer je B punog ranga), pa je stoga invertibilna. Sada vidimo da je ortogonalna projekcija na potprostor $R := R(B^T)$ dana matricom

$$\Pi_R = B^T (BB^T)^{-1} B \,. \tag{1.22}$$

i dobro je definirana.

Ortogonalna projekcija na potprostor $N := N(B)$ dana je s

$$\Pi_N = I - \Pi_R. \tag{1.23}$$

**Lema 7.** *Neka su $\Pi_N$ i $\Pi_R$ ortogonalne projekcije na prostore $N := N(B) \subseteq \mathbb{R}^n$ i $R := R(B^T) \subseteq \mathbb{R}^n$ respektivno. Tada za svaki vektor $q \in \mathbb{R}^n$ vrijedi*

$$\|\Pi_N q\|_2 = \|q\|_2 \cos\theta \quad i \quad \|\Pi_R q\|_2 = \|q\|_2 \sin\theta$$

*gdje je $\theta = \angle(q, \Pi_N q)$ kut između $q$ i $\Pi_N q$.*

*Dokaz.* S jedne strane imamo

$$q^T \Pi_N q = \|q\|_2 \cdot \|\Pi_N q\|_2 \cdot \cos\angle(q, \Pi_N q) \,. \tag{1.24}$$

S druge strane

$$q^T \Pi_N q = (\Pi_R q + \Pi_N q)^T \Pi_N q = \|\Pi_N q\|_2^2 \,, \tag{1.25}$$

jer su $R$ i $N$ ortogonalni komplementi. Sada iz (1.24) i (1.25) slijedi
$\|\Pi_N q\|_2^2 = \|q\|_2 \|\Pi_N q\|_2 \cos\angle(q, \Pi_N q)$, što je tražena tvrdnja za $\Pi_N q$.

Kako je $\|q\|_2^2 = \|\Pi_N q\|_2^2 + \|\Pi_R q\|_2^2$, upravo izvedena jednakost za $\Pi_N q$ daje

$$\|\Pi_R q\|_2^2 = \|q\|_2^2 - \|q\|_2^2 \cos^2\theta = \|q\|_2^2 (1 - \cos^2\theta) = \|q\|_2^2 \sin^2\theta \,,$$

što dokazuje lemu. □

## 1.5 Ortogonalne projekcije i korak Newtonove metode

U ovom poglavlju promatrat ćemo jednadžbu (1.21) i uvest ćemo neke formule i oznake koje će nam pomoći pri njenom rješavanju, odnosno pri dobivanju sljedeće iteracije Newtonove metode.

Definiramo dijagonalnu matricu $D$ sa

$$D^2 := XS^{-1}, \tag{1.26}$$

što je moguće jer su matrice $X$ i $S$ pozitivno definitne. Uvedimo oznaku

$$q := DX^{-1}r. \tag{1.27}$$

Primijetimo još da je $AD^2A^T$ pozitivno definitna matrica (jer je $A$ punog ranga).

**Lema 8.** *Rješenje sustava (1.21) dano je sljedećim jednakostima:*

$$q = DX^{-1}r \tag{1.28}$$

$$\Delta x = D^2A^T\Delta y - Dq \tag{1.29}$$

$$\Delta y = (AD^2A^T)^{-1}ADq \tag{1.30}$$

$$\Delta s = -D^{-1}q - D^{-2}\Delta x. \tag{1.31}$$

*Dokaz.* Ideja je reformulirati jednakosti (1.28–1.31) pomoću ortogonalnih projekcija tako da dobijemo jednakosti koje su njima ekvivalentne, ali iz kojih je lakše očitati da su rješenja jednadžbe (1.21).

Prema (1.22) i (1.23), ortogonalne projekcije na sliku od $B := DA^T$ i jezgru od $B^T = AD$ dane su sa

$$\begin{aligned}\Pi_R &= B(B^TB)^{-1}B^T = DA^T(AD^2A^T)^{-1}AD\\ \Pi_N &= I - \Pi_R = I - DA^T(AD^2A^T)^{-1}AD.\end{aligned}$$

Kako je $A$ punog ranga, možemo pomnožiti jednakost (1.30) tako da skup rješenja ostane isti, pa slijedi

$$DA^T\Delta y = DA^T(AD^2A^T)^{-1}ADq = \Pi_R q. \tag{1.32}$$

Kako je $\Pi_R + \Pi_N = I$, (1.29) možemo ekvivalentno pisati kao

$$\begin{aligned}\Delta x &= DDA^T\Delta y - Dq\\ &= D\Pi_R q - Dq\\ &= D(\Pi_R - I)q\\ &= -D\Pi_N q.\end{aligned} \tag{1.33}$$

Konačno, (1.31) ekvivalentno pišemo kao

$$\begin{aligned}\Delta s &= -D^{-1}(q + D^{-1}\Delta x)\\ &= -D^{-1}(q + D^{-1}(-D\Pi_N\, q))\\ &= -D^{-1}(q - \Pi_N\, q)\\ &= -D^{-1}\Pi_R\, q\,. \end{aligned} \tag{1.34}$$

Sada se pomoću dobivenih jednakosti (1.32–1.34) lako provjeri da vektori definirani u iskazu leme zadovoljavaju sustav (1.21): Iz (1.33) i činjenice da je $\Pi_N$ projekcija na jezgru od $AD$ slijedi $A\Delta x = -AD\Pi_N\, q = 0$. Iz (1.32) i (1.34) dobivamo

$$\begin{aligned}A^T\Delta y + \Delta s &= D^{-1}(DA^T\Delta y + D\Delta s)\\ &= D^{-1}(\Pi_R\, q + D(-D^{-1}\Pi_R\, q))\\ &= D^{-1}0\\ &= 0\,.\end{aligned}$$

Na kraju imamo

$$\begin{aligned}X\Delta s + S\,\Delta x &= X(-D^{-1}\Pi_R\, q) + S\,(-D\Pi_N\, q)\\ &= -XD^{-1}(\Pi_R\, q + X^{-1}S\; D^2\,\Pi_N\, q)\\ &= -XD^{-1}(\Pi_R\, q + \Pi_N\, q) && \text{(jer je } D^2 = XS^{-1}\text{)}\\ &= -XD^{-1}q\\ &= -r\,.\end{aligned}$$

Ovime smo dokazali tvrdnju.

□

Sljedeći korolar slijedi iz dokaza prethodne leme.

**Korolar 9.** *Rješenje jednadžbe (1.21) dano je rješenjem jednadžbi (1.28), (1.32), (1.33) i (1.34):*

$$\begin{aligned}q &= DX^{-1}r\\ DA^T\Delta y &= \Pi_R\, q\\ \Delta x &= -D\Pi_N\, q\\ \Delta s &= -D^{-1}\Pi_R\, q\,.\end{aligned}$$

## 1.6 Analiza koraka Newtonove metode

Promotrimo sada jedan korak Newtonove metode. Neka je

$$(x^+, y^+, s^+) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$$

sljedbenik od $(x, y, s)$. Novi reziduum $r^+ := X^+ s^+ - \mu e$ tada zadovoljava

$$\begin{aligned} r^+ &= (X + \Delta X)(s + \Delta s) - \mu e \\ &= Xs + \underbrace{X\Delta s + S\Delta x}_{= -r \text{ prema } (1.21)} + \Delta X \Delta s - \mu e \\ &= \underbrace{Xs - \mu e}_{= r} - r + \Delta X \Delta s \\ &= \Delta X \Delta s \,. \end{aligned} \tag{1.35}$$

Definiramo sada

$$\Delta\tilde{x} := -D^{-1}\Delta x \overset{(1.33)}{=} \Pi_N q \tag{1.36}$$

$$\Delta\tilde{s} := -D\Delta s \overset{(1.34)}{=} \Pi_R q \,, \tag{1.37}$$

iz čega slijedi $\Delta X \Delta s = \Delta\tilde{X}\Delta\tilde{s}$. Prema jednakosti (1.35) zaključujemo

$$r^+ = \Delta\tilde{X}\Delta\tilde{s} \,. \tag{1.38}$$

**Lema 10.** *Pretpostavimo da su $x > 0$, $s > 0$ i $y$ takvi da reziduum $r = Xs - \mu e$ zadovoljava $\|r\|_2 \leq \beta\mu$ za neki $\beta \in [0, \frac{1}{2}]$ . Ako napravimo jedan korak Newtonove metode krenuvši od $(x, y, s)$, novi reziduum $r^+$ tada zadovoljava $\|r^+\|_2 \leq \mu\beta^2$.*

*Dokaz.* Po definiciji je $q = DX^{-1}r$, pa je $\|q\|_2 \leq \|DX^{-1}\|_2 \|r\|_2$. Računamo

$$DX^{-1} = (\sqrt{XS^{-1}})X^{-1} = \sqrt{X^{-1}S^{-1}} = (\sqrt{R + \mu I})^{-1} = \begin{pmatrix} r_1 + \mu & & & \\ & r_2 + \mu & & \\ & & \ddots & \\ & & & r_n + \mu \end{pmatrix}^{-\frac{1}{2}} ,$$

što vrijedi jer je $r_i = x_i s_i - \mu$, za $i = 1, \ldots, n$. Prema tome,

$$\begin{aligned} \|DX^{-1}\|_2 &= \max \{ |\lambda| : \lambda \text{ svojstvena vrijednost od } DX^{-1} \} \\ &= \max_{i=1,\ldots,n} \frac{1}{\sqrt{|r_i + \mu|}} \end{aligned}$$

Slijedi da je $||q||_2 \leq \max \frac{1}{\sqrt{|r_i+\mu|}}||r||_2$.

Koristeći nejednakost trokuta i $|r_i| \leq ||r||_2 \leq \beta\mu$ imamo

$$|r_i + \mu| = |\mu - (-r_i)| \geq \mu - |r_i| \geq \mu - \beta\mu = (1-\beta)\mu\,,$$

pa slijedi

$$||DX^{-1}||_2 \leq \frac{1}{\sqrt{(1-\beta)\mu}}\,. \tag{1.39}$$

To nam daje sljedeću ocjenu za normu od $q$:

$$||q||_2 \leq ||DX^{-1}||_2||r||_2 \leq \frac{\beta\mu}{\sqrt{(1-\beta)\mu}} = \beta\sqrt{\frac{\mu}{1-\beta}}\,. \tag{1.40}$$

Iz (1.36), (1.37) i Leme 7 imamo

$$||\Delta\tilde{x}||_2 = ||\Pi_N q||_2 = ||q||_2 \cdot |\cos\theta| \tag{1.41}$$

$$||\Delta\tilde{s}||_2 = ||\Pi_R q||_2 = ||q||_2 \cdot |\sin\theta|\,. \tag{1.42}$$

Te jednakosti daju nam konačnu ocjenu za normu od $r^+$:

$$\begin{aligned}
||r^+||_2 &= ||\Delta\tilde{X}\Delta\tilde{s}||_2 \\
&\leq ||\Delta\tilde{x}||_2 \cdot ||\Delta\tilde{s}||_2 \\
&\leq ||q||_2^2 \cdot |\sin\theta\cos\theta| && \text{zbog (1.41) i (1.42)} \\
&= ||q||_2^2 \cdot \left|\frac{1}{2}\sin 2\theta\right| \\
&\leq \frac{1}{2}||q||_2^2 && \text{jer je } \sin 2\theta \in [0,1] \\
&\leq \frac{1}{2}\beta^2\mu\frac{1}{1-\beta} && \text{zbog (1.40)} \\
&\leq \frac{1}{2}\beta^2\mu && \text{jer je } \beta \in [0, \tfrac{1}{2}]\,,
\end{aligned}$$

što je bila tvrdnja leme. □

**Korolar 11.** *Uz pretpostavke kao u Lemi 10, vrijedi $x^+ > 0$ i $s^+ > 0$.*

*Dokaz.* Iz (1.41) zaključujemo $\|\Delta\tilde{x}\|_2 \leq \|q\|_2$. Računamo

$$\begin{aligned}
\|X^{-1}\Delta x\|_2 &= \|X^{-1}D\Delta\tilde{x}\|_2 && \text{iz (1.36)}\\
&\leq \|DX^{-1}\|_2\|\Delta\tilde{x}\|_2 \\
&\leq \frac{\|\Delta\tilde{x}\|_2}{\sqrt{(1-\beta)\mu}} && \text{iz (1.39)}\\
&\leq \frac{\beta\sqrt{\frac{\mu}{1-\beta}}}{\sqrt{(1-\beta)\mu}} && \text{iz (1.40)}\\
&= \frac{\beta}{1-\beta} \\
&\leq 1 && \text{jer je } \beta \in [0, \tfrac{1}{2}].
\end{aligned}$$

Slijedi da je $\left|\frac{\Delta x_i}{x_i}\right| \leq 1$, odnosno $|\Delta x_i| \leq |x_i|$ za $i = 1, \ldots, n$. Zato vrijedi $x_i + \Delta x_i \geq 0$ za $i = 1, \ldots, n$ i $x^+ = x + \Delta x \geq 0$.

Analogno se provodi račun i za $s^+$. Primijetimo da je $S^{-1}D^{-1} = S^{-1}\sqrt{X^{-1}S} = \sqrt{X^{-1}S^{-1}} = DX^{-1}$, pa vrijedi

$$\begin{aligned}
\|S^{-1}\Delta s\|_2 &= \|S^{-1}D^{-1}\Delta\tilde{s}\|_2 && \text{iz (1.37)}\\
&\leq \|DX^{-1}\|_2\|\Delta\tilde{s}\|_2 \\
&\leq \frac{\|\Delta\tilde{s}\|_2}{\sqrt{(1-\beta)\mu}} && \text{iz (1.39)}\\
&\leq 1 && \text{jer je } \beta \in [0, \tfrac{1}{2}].
\end{aligned}$$

Zaključujemo kao i u prijašnjem slučaju da je $s^+ = s + \Delta s \geq 0$. Konačno, kako je

$$\left|x_i^+ s_i^+ - \mu\right| \leq \|X^+ s^+ - \mu e\|_2 = \|r^+\|_2 \overset{Lema10}{\leq} \beta^2\mu < \mu$$

za $i = 1, \ldots, n$, nužno slijedi $x_i^+ > 0$ i $s_i^+ > 0$ za $i = 1, \ldots, n$. □

## 1.7 Primarno-dualna metoda unutrašnje točke

Koristeći rezultate prethodne točke, možemo naslutiti ideju algoritma primarno - dualne metode. Krenemo od $(x^0, y^0, s^0)$ i $\mu_0$ tako da početni reziduum $r^0$ zadovoljava $\|r^0\|_2 \leq \frac{1}{2}\mu_0$. Napravimo jedan korak Newtonove metode, reduciramo $\mu$ i nastavimo analogno dalje.

Preciznije, pretpostavimo da imamo dva problema linearnog programiranja s početka poglavlja (1.1) i (1.2), početne vektore $x^0 > 0$ i $s^0 > 0$ i vrijednost $\mu_0 > 0$ takve da vrijede sljedeće jednakosti:

$$\begin{aligned} Ax^0 &= b \\ A^T y^0 + s^0 &= c \\ r^0 &= X^0 s^0 - \mu_0 e \\ \|r^0\|_2 &\leq \frac{1}{2}\mu_0\,, \end{aligned}$$

te neka je $\epsilon > 0$. Algoritam glasi:

Neka je $k := 0$.
Sve dok vrijedi $\mu_k > \frac{\epsilon}{n}$:

1. Odredimo korak Newtonove metode $(\Delta x, \Delta y, \Delta s)$ iz $(x^k, y^k, s^k)$ koristeći (1.21)
2. Postavimo $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k + \Delta x, y^k + \Delta y, s^k + \Delta s)$
3. Reduciramo $\mu_k$ na $\mu_{k+1} := (1 - \frac{1}{6\sqrt{n}})\mu_k$
4. $k = k + 1$.

Primijetimo kako su nam poteškoće u Newtonovoj metodi zadavali uvjeti nenegativnosti, odnosno nelinearni uvjeti. Zbog toga kod metode unutrašnje točke u svakoj iteraciji tražimo rješenje relaksiranog sustava, te uz reduciranje parametra $\mu$ dolazimo do tražene točnosti. Strogi uvjeti nejednakosti na početku $(x^0 > 0, s^0 > 0)$ i u svakoj sljedećoj iteraciji omogućavaju da se u svakom koraku algoritma nalazimo u unutrašnjosti dopustivog skupa, a ne na njegovom rubu. Zbog toga je i metoda unutrašnje točke i dobila taj naziv.

Sljedeći rezultat govori o broju iteracija koje su potrebne da se dođe do optimalnih točaka i o njihovoj točnosti.

**Teorem 12.** *Navedeni algoritam daje niz strogo dopustivih točaka primarne, odnosno dualne zadaće. Završava nakon najviše* $6\sqrt{n}\ln\frac{n\mu_0}{\epsilon}$ *iteracija i za dobivene optimalne točke x i (y,s) vrijedi* $c^T x - b^T y \leq 2\epsilon$.

*Dokaz.* Za dokaz prvog dijela teorema, prema Korolaru 11 dovoljno je pokazati da je $\|r^k\|_2 \leq \frac{\mu_k}{2}$ za $k = 0, 1, \ldots$, što ćemo pokazati indukcijom po broju iteracija $k$.

Za $k = 0$ tvrdnja očito vrijedi. U koraku indukcije potrebno je pokazati

$$\|r^{k+1}\|_2 = \|X^{k+1}s^{k+1} - \mu_{k+1}e\|_2 \leq \frac{1}{2}\mu_{k+1}\,.$$

Vrijedi

$$\begin{aligned} r^{k+1} &= X^{k+1}s^{k+1} - \mu_{k+1}e \\ &= \underbrace{X^{k+1}s^{k+1} - \mu_k e}_{:=\,r^+} + (\mu_k - \mu_{k+1})e \\ &= r^+ + \frac{\mu_k}{6\sqrt{n}}e\,, \end{aligned}$$

pri čemu smo u zadnjoj jednakosti koristili definiciju od $\mu_{k+1}$.

Nadalje, zbog Leme 10 vrijedi $\|r^+\|_2 \leq \frac{\mu_k}{4}$, iz čega slijedi

$$\begin{aligned} \|r^{k+1}\|_2 &\leq \|r^+\|_2 + \frac{\mu_k}{6\sqrt{n}}\|e\|_2 \\ &\leq \frac{\mu_k}{4} + \frac{\mu_k}{6} \\ &= \frac{5}{12}\mu_k\,. \end{aligned}$$

Kako je $\mu_{k+1} \geq \frac{5}{6}\mu_k$, zaključujemo $\|r^{k+1}\|_2 \leq \frac{5}{12} \cdot \frac{6}{5}\mu_{k+1} = \frac{\mu_{k+1}}{2}$.

Za dokaz drugog dijela teorema, primijetimo da algoritam staje kad je $\mu_k \leq \frac{\epsilon}{n}$. Računamo

$$\begin{aligned} &\frac{\epsilon}{n} \geq \mu_k = (1 - \frac{1}{6\sqrt{n}})^k \mu_0 \\ &\Leftrightarrow ln\frac{\epsilon}{n\mu_0} \geq k \cdot ln(1 - \frac{1}{6\sqrt{n}}) \\ &\Leftrightarrow -ln\frac{n\mu_0}{\epsilon} \geq k \cdot ln(1 - \frac{1}{6\sqrt{n}}) \\ &\Leftrightarrow \frac{ln\frac{n\mu_0}{\epsilon}}{-ln(1 - \frac{1}{6\sqrt{n}})} \leq k\,. \end{aligned} \tag{1.43}$$

Kako općenito za $a < 1$ vrijedi $ln(1-a) \leq -a$, slijedi $-ln(1-\frac{1}{6\sqrt{n}}) \geq \frac{1}{6\sqrt{n}}$, a da bi relacija (1.43) bila zadovoljena, dovoljno je da je $k \geq 6\sqrt{n}\, ln\, \frac{n\mu_0}{\epsilon}$.

Konačno, pretpostavimo da su $x, y$ i $s$ točke koje daje zadnji korak algoritma i označimo sa $\mu$ posljednju vrijednost koju poprima $\mu_k$. Tada je

$$\begin{aligned} c^T x - b^T y &= x^T s && \text{iz (1.3)} \\ &= e^T(Xs) \\ &= e^T(\mu e + r) && \text{po definiciji reziduuma r} \\ &\leq \underbrace{\mu n}_{\leq \epsilon} + \underbrace{||r||_1}_{\leq \sqrt{n}\,||r||_2 \leq \sqrt{n}\mu/2 \leq \epsilon} \\ &\leq 2\epsilon\,, \end{aligned}$$

što u potpunosti dokazuje teorem.

□

Primijetimo da smo u ovom poglavlju zanemarili problem nalaženja početnih vrijednosti $(x^0, y^0, s^0)$, $\mu_0$ i $r^0$. Detaljnije o tome i o složenosti algoritma opisano je u skripti Sven O. Krumke, (2004), *Interior Point Methods*.

# Poglavlje 2

# Kvadratično programiranje

## 2.1 Egzistencija rješenja zadaće kvadratičnog programiranja

U ovom ćemo poglavlju govoriti o kvadratičnom programiranju, odnosno o optimizaciji kvadratne funkcije uz linearne uvjete. Takav problem zapisujemo u sljedećem obliku:

$$\begin{aligned}\frac{1}{2}x^T Qx + c^T x &\rightarrow \min\\ Ax &= b\\ x &\geq 0\,,\end{aligned} \tag{2.1}$$

gdje su $A \in M_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ i $x \in \mathbb{R}^n$, te $Q \in \mathbb{R}^{n\times n}$, pri čemu je $m < n$.

Dualni problem zadaće (2.1) glasi

$$\begin{aligned}b^T y - \frac{1}{2}x^T Qx &\rightarrow \max\\ A^T y - Qx &\leq c\\ x &\geq 0\,,\end{aligned}$$

što nakon uvođenja pomoćne varijable $s \in \mathbb{R}^n_+$ postaje

$$\begin{aligned}&b^T y - \frac{1}{2}x^T Qx \rightarrow \max\\ &A^T y - Qx + s = c\\ &s \geq 0\,,\ x \geq 0\,.\end{aligned} \tag{2.2}$$

Primijetimo da kada je $Q$ pozitivno semidefinitna matrica, odnosno kada je $y^T Qy \geq 0$ za svaki $y$, tada je funkcija cilja primarne zadaće konveksna te je svaki njezin lokalni minimum ujedno i globalni minimum. Tu tvrdnju ćemo dokazati u sljedećoj lemi.

**Lema 13.** *Neka je $Q$ pozitivno semidefinitna. Ako je $x$ lokalni minimum zadaće (2.1), onda je i globalni minimum.*

*Dokaz.* Kao što smo već zaključili, funkcija cilja $f$ zadaće (2.1) je konveksna. Dopustiv skup je također konveksan, označimo ga s $D$. Pretpostavimo da je $x$ lokalno rješenje, ali ne i globalno, tj. da postoji $\bar{x} \in D$ takav da vrijedi $f(\bar{x}) < f(x)$. Neka je $y := \lambda x + (1 - \lambda)\bar{x}$, $\lambda \in [0, 1]$. Vrijedi $y \in D$. Zbog konveksnosti i pretpostavke imamo

$$f(y) = f(\lambda x + (1 - \lambda)\bar{x}) \leq \lambda f(x) + (1 - \lambda)f(\bar{x}) < \lambda f(x) + (1 - \lambda)f(x) = f(x)\,.$$

Dobili smo da je $f(y) < f(x)$ za $\lambda \in \langle 0, 1\rangle$, tj. $x$ nije lokalni minimum, što je kontradikcija s pretpostavkom. □

Sada ćemo reći nešto više o egzistenciji rješenja zadaće (2.1). Sljedeći teorem daje dovoljne uvjete:

**Teorem 14.** *(Weierstrass)*
*Ako je funkcija cilja f neprekidna i dopustivi skup $\Omega$ kompaktan, tada postoji rješenje zadaće (2.1).*

*Dokaz.* Neka je $v^* = \inf_{x\in\Omega} f(x)$ (neprekidna funkcija poprima minimum na kompaktnom skupu). Definiramo niz $(x_n) \subseteq \Omega$ takav da $f(x_n) \to v^*$. Kako je $(x_n)$ ograničen, on ima konvergentan podniz $(x_{n_k})$. Označimo njegov limes s $x^*$. Vrijedi $x^* \in \Omega$ jer je $\Omega$ zatvoren, a zbog neprekidnosti od $f$ imamo $f(x_{n_k}) \to f(x^*)$. Niz i podniz imaju isti limes, pa vrijedi $f(x^*) = v^*$, odnosno $x^*$ je optimalno rješenje. □

Vrijedi i sljedeća tvrdnja:

**Korolar 15.** *Ako je funkcija cilja $f$ neprekidna i koercitivna, tj. ako vrijedi*

$$\lim_{\|x\|\to+\infty} f(x) = +\infty\,,$$

*i dopustiv skup $\Omega$ je zatvoren i neprazan, tada postoji rješenje zadaće (2.1).*

Kako bismo mogli iskazati teorem o egzistenciji rješenja polazne zadaće, uvest ćemo neke pojmove. Prisjetimo se, skup $S \subseteq \mathbb{R}^n$ je *konus* ako

$$(\forall x \in S)(\forall t \geq 0)\ tx \in S.$$

Definirajmo sada *recesivni konus* dopustivog skupa. Označit ćemo ga s $rec(\Omega)$ i definirati kao

$$rec(\Omega) = \{d \in \mathbb{R}^n : Ad = 0 \text{ i } d \geq 0\}.$$

Vektor $d \in rec(\Omega)$ zovemo *recesivni smjer* dopustivog skupa. Primijetimo da za recesivni smjer $d$ i svaki dopustivi $x$ te svaki $t \geq 0$ točka $x + td$ također pripada dopustivom skupu.

Teorem o egzistenciji rješenja zadaće kvadratičnog programiranja (2.1) sada glasi:

**Teorem 16.** *Zadaća (2.1) ima rješenje ako i samo ako je dopustivi skup neprazan te vrijede sljedeći uvjeti:*

$$\begin{aligned}
&d^T Qd \geq 0, \quad d \in rec(\Omega) \\
&d^T(Qx + c) \geq 0, \quad x \in \Omega,\ d \in rec(\Omega) \ t.d.\ d^T Qd = 0.
\end{aligned}$$

**Napomena 17.** *Spomenimo neke specijalne slučajeve:*

- *Ako je $Q = 0$, tj. ako promatranu zadaću svedemo na zadaću linearnog programiranja, tada ona ima rješenje ako i samo ako je*

$$d^T c \geq 0, \quad d \in rec(\Omega).$$

- *Ako je $Q$ pozitivno definitna, tada su uvjeti iz Teorema 16 zadovoljeni i zadaća (2.1) ima rješenje. Isto možemo zaključiti i iz Korolara 15.*

- *Ako je $rec(\Omega)$ trivijalan, onda je $\Omega$ kompaktan pa prema Teoremu 14 postoji rješenje.*

## 2.2 Nužni i dovoljni uvjeti za optimalnost

Prije nego što dokažemo teorem o nužnim i dovoljnim uvjetima za optimalnost, dokazat ćemo sljedeći teorem:

**Teorem 18.** *(Karush-Kuhn-Tucker)*
*Neka je $K \subseteq \mathbb{R}^n$ konveksan i otvoren, $f, g_1, \ldots, g_n : K \to \mathbb{R}$ konveksne klase $C^1$ i $h_1, \ldots, h_m : K \to \mathbb{R}$ afine. Neka je dana zadaća*

$$\begin{aligned}
&f(x) \to \min \\
&g_i(x) \leq 0, \quad i = 1, \ldots, n \\
&h_j(x) = 0, \quad j = 1, \ldots, m.
\end{aligned}$$

*Ako postoje* $(\lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m$ *takvi da za dopustivu točku* $x^*$ *vrijedi*

$$\begin{aligned}
&\lambda_i \geq 0, \quad i = 1, \dots, n \\
&\lambda_i g_i(x^*) = 0, \quad i = 1, \dots, n \\
&\nabla f(x^*) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*) = 0\,,
\end{aligned}$$

*onda je* $x^*$ *rješenje dane zadaće.*

*Dokaz.* Zapišimo Lagrangeovu funkciju dane zadaće:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x).$$

Funkcije $f$ i $g_i$ su konveksne, a $h_j$ su afine pa su konveksne. Zbroj konveksnih funkcija je opet konveksna funkcija, kao i produkt konveksne funkcije s nenegativnim skalarom, pa slijedi da je $L(x, \lambda, \mu)$ konveksna. Po karakterizaciji konveksnih funkcija za dopustivi $x^*$ imamo

$$L(x, \lambda, \mu) \geq L(x^*, \lambda, \mu) + \nabla L(x^*, \lambda, \mu)^T (x - x^*),$$

odnosno

$$L(x, \lambda, \mu) \geq L(x^*, \lambda, \mu) + \underbrace{\left( \nabla f(x^*) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*) \right)^T}_{=0 \text{ po Teoremu 18}} (x - x^*).$$

Za dopustivi $x$ sada slijedi

$$f(x) + \sum_{i=1}^{n} \underbrace{\lambda_i g_i(x)}_{=0} + \sum_{j=1}^{m} \mu_j \underbrace{h_j(x)}_{=0} \geq f(x^*) + \sum_{i=1}^{n} \underbrace{\lambda_i g_i(x^*)}_{=0} + \sum_{j=1}^{m} \mu_j \underbrace{h_j(x^*)}_{=0},$$

tj.

$$f(x) \geq f(x^*).$$

Dakle, $x^*$ je točka minimuma. □

**Napomena 19.** *Vrijedi i obrat Teorema 18, ali uz dodatni zahtjev regularnosti. Primjer takvog zahtjeva je Slaterov uvjet koji glasi: postoji dopustiva točka* $\bar{x}$ *takva da za svaki nelinearni uvjet vrijedi* $g_i(\bar{x}) < 0$. *Međutim, u zadaći (2.1) su uvjeti linearni, pa obrat Teorema 18 vrijedi bez dodatnih zahtjeva regularnosti.*

**Teorem 20.** *(Nužni i dovoljni uvjeti za optimalnost)*
*Neka je $x$ rješenje zadaće (2.1) i pretpostavimo da je $Q$ pozitivno semidefinitna matrica. Tada postoje vektori $y$ i $s$ takvi da vrijedi*

$$A^T y - Qx + s = c \tag{2.3}$$

$$s \geq 0 \tag{2.4}$$

$$x_i s_i = 0\,,\ i = 1, \ldots, n\,. \tag{2.5}$$

*Obratno, ako vektori $x$, $y$ i $s$ zadovoljavaju uvjete dopustivosti primarne zadaće*

$$Ax = b \tag{2.6}$$

$$x \geq 0 \tag{2.7}$$

*i uvjete (2.3)-(2.5), tada je $x$ rješenje zadaće (2.1).*

*Dokaz.* Ako je $x$ rješenje zadaće (2.1), po Napomeni 19 postoje $(\lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^m$ takvi da je

$$\begin{aligned}
&\lambda_i \geq 0, \quad i = 1, \ldots, n \\
&\lambda_i(-x_i) = 0, \quad i = 1, \ldots, n \\
&Qx + c - \lambda + A^T \mu = 0\,.
\end{aligned}$$

Ako označimo $y := -\mu$ i $s := \lambda$, gornje relacije su upravo one iz tvrdnje teorema.

Obratno, prema Teoremu 18 dovoljno je pokazati da se uvjeti (2.3)-(2.5) svode na uvjete iz spomenutog teorema. Mora vrijediti:

$$\begin{aligned}
&\lambda_i \geq 0, \quad i = 1, \ldots, n \\
&\lambda_i g_i(x) = -\lambda_i x_i = 0, \quad i = 1, \ldots, n \\
&\nabla f(x) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x) + \sum_{j=1}^{m} \mu_j \nabla h_j(x) = Qx + c - \lambda + A^T \mu = 0\,.
\end{aligned}$$

Za $\lambda := s$ i $\mu := -y$ to su upravo uvjeti (2.3)-(2.5), pa po Teoremu 18 slijedi tvrdnja. □

Sada imamo nužne i dovoljne uvjete za optimalno rješenje zadaće (2.1): $x \in \mathbb{R}^n$ i $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ su optimalna rješenja za (2.1) i (2.2) ako i samo ako zadovoljavaju uvjete (2.3)-(2.7).

## 2.3 Metoda unutrašnje točke

Sada ćemo opisati metodu unutrašnje točke za kvadratično programiranje. Analogno kao u prethodnom poglavlju, govorit ćemo o centralnom putu i koristeći Newtonovu metodu pokušat ćemo doći do optimalnog rješenja.

Zapišimo najprije uvjete optimalnosti iz Teorema 20 u matričnom obliku:

$$F(x,y,s) = \begin{pmatrix} A^T y - Qx + s - c \\ Ax - b \\ XSe \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (x,s) \geq 0\,. \tag{2.8}$$

Kao i ranije, $X$ i $S$ su dijagonalne matrice koje na dijagonalama imaju vrijednosti $x_1, \ldots, x_n$ i $s_1, \ldots, s_n$ respektivno, dok je vektor $e$ $n$-dimenzionalni vektor jedinica.

Uvedimo oznake za skupove dopustivih i strogo dopustivih rješenja primarno-dualne zadaće (uzimamo u obzir uvjete dopustivosti i primarne i dualne zadaće):

$$\mathcal{F} := \{(x,y,s) : Ax = b, A^T y - Qx + s = c, x \geq 0, s \geq 0\} \tag{2.9}$$

$$\mathcal{F}^+ := \{(x,y,s) : Ax = b, A^T y - Qx + s = c, x > 0, s > 0\} \tag{2.10}$$

Sustav (2.8) ima $m + 2n$ jednadžbi te jednako toliko nepoznanica i za njegovo rješavanje koristit ćemo Newtonovu metodu. Jacobijeva matrica dana je sa

$$DF(x^k, y^k, s^k) = \begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix}, \tag{2.11}$$

pri čemu su $S^k$ i $X^k$ dijagonalne matrice sa odgovarajućim elementima koji su dobiveni $k$-tom iteracijom.

Da bi Newtonova metoda bila dobro definirana, Jacobijeva matrica mora biti regularna. Sljedeća lema pokazuje pod kojim je uvjetima to zadovoljeno.

**Lema 21.** *Jacobijeva matrica*

$$DF(x,y,s) = \begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}$$

*je regularna ako su $x > 0$, $s > 0$, $A$ punog ranga $m$ i $Q$ pozitivno semidefinitna.*

*Dokaz.* Neka je $A$ punog ranga i $Q$ pozitivno semidefinitna, te neka su $(u, v, w) \neq (0, 0, 0)$ takvi da je

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

odnosno

$$\begin{pmatrix} -Qu + A^T v + w \\ Au \\ Su + Xw \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Računamo

$$u^T w = u^T(Qu - A^T v) = u^T Qu - (Au)^T v = u^T Qu \geq 0\,,$$

jer je $Au = 0$ i $Q$ pozitivno semidefinitna . Nadalje, iz gornje relacije te iz $Su + Xw = 0$ slijedi

$$0 \leq u^T w = w^T u = -w^T S^{-1} X w,$$

iz čega zaključujemo da je $w = 0$ jer je $S^{-1}X$ pozitivno definitna (dijagonalna je i svi elementi su pozitivni). Dakle, $u = -S^{-1}Xw = 0$ . Konačno,

$$A^T v = Qu - w = 0\,,$$

iz čega slijedi da je $v = 0$ jer je $A$ punog ranga. Dakle, jezgra polazne matrice je trivijalna pa zaključujemo da je matrica regularna.

□

Uz pretpostavke iz Leme 21, Newtonova metoda za sustav (2.8) glasi:

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ y^k \\ s^k \end{pmatrix} - DF(x^k, y^k, s^k)^{-1} F(x^k, y^k, s^k),$$

odnosno

$$\begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = -DF(x^k, y^k, s^k)^{-1} F(x^k, y^k, s^k). \tag{2.12}$$

Ako u $k$-toj iteraciji dobijemo $(x^k, y^k, s^k) \in \mathcal{F}^+$ , tj. strogo dopustivo rješenje, tada iz (2.8) slijedi

$$F(x^k, y^k, s^k) = \begin{pmatrix} 0 \\ 0 \\ X^k S^k e \end{pmatrix}, \tag{2.13}$$

pa iz (2.11)-(2.13) dobivamo

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^k S^k e \end{pmatrix} . \tag{2.14}$$

Sljedeću iteraciju bi onda dobili kao $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + (\Delta x^k, \Delta y^k, \Delta s^k)$, no tu se javlja problem jer takvom iteracijom ne moramo nužno dobiti točku koja zadovoljava uvjete $x^{k+1} > 0$ i $s^{k+1} > 0$. Zbog toga modificiramo iteraciju i nalazimo parametar $\alpha_k \in \langle 0, 1]$ tako da dobijemo

$$\begin{aligned} x^k + \alpha_k \Delta x^k &> 0 \quad \text{i} \\ s^k + \alpha_k \Delta s^k &> 0 \,. \end{aligned}$$

Tada je sljedeća iteracija Newtonove metode dana sa

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

U praksi se pokazuje da takav pristup nije dobar jer $\alpha_k$ postaju jako mali i to otežava postupak. Zbog toga ćemo promatrati metodu centralnog puta.

## 2.4 Centralni put

Za parametar $\tau > 0$ uvodimo sljedeći relaksirani sustav:

$$F(x, y, s) = \begin{pmatrix} A^T y - Qx + s - c \\ Ax - b \\ XSe \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tau e \end{pmatrix}, \quad (x, s) > 0 \,. \tag{2.15}$$

Kada je skup $\mathcal{F}^+$ neprazan, sustav (2.15) ima jedinstveno rješenje za svaki $\tau > 0$ i označavamo ga s $(x_\tau, y_\tau, s_\tau)$. Centralni put je tada skup

$$C = \{(x_\tau, y_\tau, s_\tau) : \tau > 0\} \,. \tag{2.16}$$

Trajektorija $(x_\tau, y_\tau, s_\tau)$ konvergira prema optimalnom rješenju zadaće (2.1). Kako $\tau \to 0$, uvjeti u (2.15) koji određuju točke na centralnom putu će sve bolje aproksimirati skup uvjeta optimalnosti danih u (2.8).

Primijetimo da treći redak u (2.15) možemo ekvivalentno zapisati kao

$$(x_\tau)_i (s_\tau)_i = \tau \,, \; i = 1, \ldots, n.$$

Uz gore opisnu relaksaciju, sustav (2.8) sada postaje

$$F(x, y, s) = \begin{pmatrix} A^T y - Qx + s - c \\ Ax - b \\ XSe - \tau e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{2.17}$$

dok (2.14) modificiramo kao

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tau e - X^k S^k e \end{pmatrix}. \tag{2.18}$$

Kako bismo odredili $\tau$, najprije definiramo mjeru koju nazivamo *prosječna komplementarnost* $\mu$ kao

$$\mu(x, s) := \frac{\sum_{i=1}^{n} x_i s_i}{n} = \frac{x^T s}{n}. \tag{2.19}$$

Primijetimo da kada $(x, y, s)$ zadovoljava uvjete $Ax = b$, $x \geq 0$ i $A^T y - Qx + s = c$, tada je $(x, y, s)$ optimalno rješenje ako i samo ako vrijedi $\mu(x, s) = 0$. Ako je $\mu$ velik, znači da smo daleko od optimalnog rješenja. Zbog toga je $\mu$ mjera koja opisuje optimalnost dopustivih točaka - što je $\mu$ manji, bliže smo optimalnom rješenju.

Za točku centralnog puta $(x_\tau, y_\tau, s_\tau)$ slijedi

$$\mu(x_\tau, s_\tau) = \frac{\sum_{i=1}^{n} (x_\tau)_i (s_\tau)_i}{n} = \frac{\sum_{i=1}^{n} \tau}{n} = \tau.$$

Zbog te jednakosti, točke centralnog puta $(x_\tau, y_\tau, s_\tau)$ su one dopustive točke $(x, y, s)$ za koje vrijedi $\mu(x, s) = \tau$. Ako se u jednoj iteraciji naše metode nalazimo u točki $(x, y, s)$, tada možemo izabrati $\tau$, odnosno sljedeću točku centralnog puta na tri načina: da bude na nižoj razini od trenutne (tj. da je $\tau < \mu(x, s)$), da bude na jednakoj razini kao trenutna točka ($\tau = \mu(x, s)$) ili da bude na višoj razini ($\tau > \mu(x, s)$). Treći slučaj najčešće nije dobar izbor jer će dovesti do točke koja je dalja optimalnoj točki od trenutne. Dakle, uvijek ćemo izabrati takav $\tau$ da vrijedi $\tau \leq \mu(x, s)$.

Sada ćemo dodatno modificirati sustav (2.18) i uvesti još jedan parametar $\sigma$, pa imamo

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sigma^k \mu^k e - X^k S^k e \end{pmatrix}, \tag{2.20}$$

pri čemu je $\mu^k := \mu(x^k, s^k) = \frac{(x^k)^T s^k}{n}$, dok $\sigma^k \in [0, 1]$ određuje utjecaj parametra $\mu^k$. Naime, ako je $\sigma^k = 0$, tada govorimo o ranije spomenutoj čistoj Newtonovoj metodi koja nije dobra u praksi. Ako je $\sigma^k = 1$, odnosno ako je $\tau = \mu^k$, u sljedećoj iteraciji ćemo dobiti točku s istom prosječnom komplementarnosti, no ponekad je to poželjno jer kasnije može doprinjeti značajnom pomaku prema optimalnom rješenju. U praksi se najčešće biraju vrijednosti $\sigma^k$ koje nisu na rubu dozvoljenog intervala.

Opći algoritam za nalaženje optimalnog rješenja metodom unutrašnje točke sad glasi:

Neka je $k := 0$ i $\epsilon > 0$ . Odaberemo početne vrijednosti $(x^0, y^0, s^0) \in \mathcal{F}^+$ .
Sve dok vrijedi $\mu^k > \frac{\epsilon}{n}$:

1. Odaberemo $\sigma^k \in [0, 1]$ te neka je $\mu^k = \frac{(x^k)^T s^k}{n}$. Riješimo sustav (2.20).
2. Odaberemo $\alpha_k$ tako da vrijedi $x^k + \alpha_k \Delta x^k > 0$ i $s^k + \alpha_k \Delta s^k > 0$.
3. Postavimo $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k(\Delta x^k, \Delta y^k, \Delta s^k)$.
4. $k = k + 1$.

Više o samom odabiru parametara u algoritmu reći ćemo nešto kasnije.

## 2.5 Okolina centralnog puta

U metodi unutrašnje točke parametar $\tau$ iz sustava (2.15) zapravo parametrizira trajektoriju centralnog puta $C$. Međutim, umjesto da generiramo točke koje se nalaze točno na centralnom putu, ideja je generirati točke koji ih aproksimiraju. Te aproksimacije postižemo pomnim odabirom parametara $\sigma$ i $\alpha$. Kako centralni put vodi optimalnom rješenju, tako će i naše aproksimacije dovesti do željenog rješenja. Na slici 2.1 prikazan je centralni put, njegova okolina te pripadne iteracije. Glavni razlog zbog kojeg ne generiramo točke na centralnom putu jest taj da to može biti zahtjevno i teško izvedivo u praksi. U ovoj cjelini reći ćemo nešto više o okolini centralnog puta.

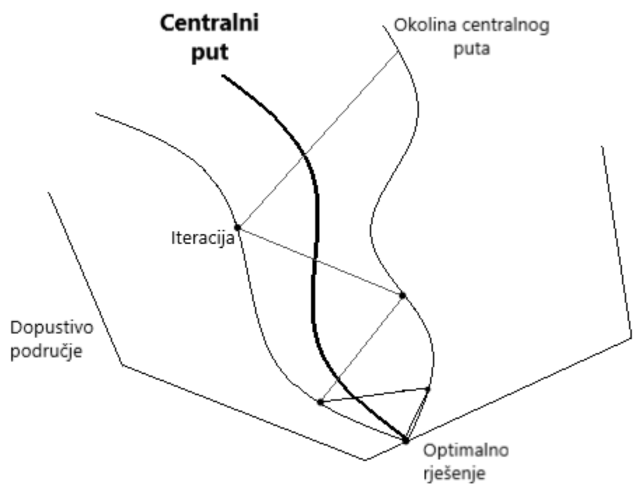
Prisjetimo se, točke na centralnom putu su one koje se nalaze u skupu $\mathcal{F}^+$ i koje zadovoljavaju dodatni uvjet

$$x_i s_i = \tau\,, \quad i = 1, \ldots, n\,,$$

za neki $\tau > 0$. Promotrimo sada točku $(x_\tau, y_\tau, s_\tau)$ na centralnom putu. Ako točka $(x, y, s)$ aproksimira navedenu točku, očekujemo da je njihova udaljenost $\|(x, y, s) - (x_\tau, y_\tau, s_\tau)\|_2$ mala. Zato skup točaka koje aproksimiraju točku $(x_\tau, y_\tau, s_\tau)$ definiramo kao

$$\{(x, y, s) \in \mathcal{F}^+ : \|(x, y, s) - (x_\tau, y_\tau, s_\tau)\|_2 \leq \epsilon\}\,, \tag{2.21}$$

za neki $\epsilon \geq 0$.




Slika 2.1: Okolina centralnog puta

Međutim, u praksi je teško egzaktno odrediti točku na centralnom putu $(x_\tau, y_\tau, s_\tau)$, pa definicija (2.21) nije od velikog značaja. Zbog toga koristimo sustav (2.17) kako bi opisali te točke i promatramo njihove okoline, odnosno skupove koji opisuju njihovu neposrednu blizinu. Primjer najčešće korištenih takvih skupova su

$$\mathcal{N}_2(\theta) := \left\{(x, y, s) \in \mathcal{F}^+ : \|XSe - \mu e\|_2 \le \theta\mu\,, \quad \mu = \frac{x^T s}{n}\right\} \tag{2.22}$$

za neki $\theta \in \langle 0, 1\rangle$ i

$$\mathcal{N}_{-\infty}(\gamma) := \left\{(x, y, s) \in \mathcal{F}^+ : x_i s_i \ge \gamma\mu\,, \quad i = 1, \dots, n\,, \quad \mu = \frac{x^T s}{n}\right\}, \tag{2.23}$$

za neki $\gamma \in \langle 0, 1\rangle$.

Točka koju dobijemo iteracijom je "blizu" centralnom putu ako se nalazi u jednom od skupova (2.22) ili (2.23). Ako je $\theta = 0$ u (2.22) ili $\gamma = 1$ u (2.23), okolina postaje upravo centralni put $C$.

Promotrimo sada nejednakost iz definicije (2.22).

$$\|XSe - \mu e\|_2 \leq \theta\mu \quad \Leftrightarrow \quad \left\| \begin{matrix} \frac{x_1 s_1}{\mu} - 1 \\ \frac{x_2 s_2}{\mu} - 1 \\ \vdots \\ \frac{x_n s_n}{\mu} - 1 \end{matrix} \right\|_2 \leq \theta\,, \tag{2.24}$$

što je ekvivalentno s

$$\sum_{i=1}^{n} \left( \frac{x_i s_i}{\mu} - 1 \right)^2 \leq \theta^2\,.$$

Posljednja jednakost zapravo predstavlja odstupanje pojedinog $x_i s_i$ od prosječne vrijednosti $\mu$. Zbog toga će točka zadovoljavati uvjete dane u (2.22) samo ako je suma kvadratnih odstupanja od $\mu$ mala, iz čega zaključujemo da skup $\mathcal{N}_2(\theta)$ ne sadrži puno točaka. S druge strane, da bi se točka nalazila u skupu $\mathcal{N}_{-\infty}(\gamma)$, dovoljno je samo da za svaki $i$ vrijednost od $x_i s_i$ nije puno manja od prosječne vrijednosti $\mu$.

Prema tome, za uobičajene vrijednosti $\theta$ i $\gamma$, okolina definirana skupom $\mathcal{N}_2(\theta)$ je puno uža od okoline definirane skupom $\mathcal{N}_{-\infty}(\gamma)$. Zato razlikujemo metode, odnosno algoritme koji koriste navedene okoline. Ukoliko metoda koristi okolinu $\mathcal{N}_2(\theta)$, zovemo ju metoda kratkog koraka, dok za okolinu $\mathcal{N}_{-\infty}(\gamma)$ imamo metodu dugog koraka.

Razlike između navedenih metoda očituju se u brzini konvergencije prema optimalnom rješenju. Ukoliko metoda koristi užu okolinu, garantira nam da ćemo svakom iteracijom dobiti točku koja je relativno blizu centralnog puta. S druge strane, ako koristimo širu okolinu, u najgorem slučaju ćemo se dosta udaljiti od centralnog puta i bit će potrebno puno više iteracija da dođemo do optimalnog rješenja. Međutim, u praksi se pokazuje da u prosjeku brže konvergiraju metode koje koriste širu okolinu.

Sada ćemo navesti algoritam koji koristi širu okolinu oko centralnog puta, odnosno algoritam metode dugog koraka.

Neka su dani $\epsilon > 0$, $\gamma \in \langle 0, 1\rangle$ i $0 < \theta_{min} < \theta_{max} < 1$.
Neka je $k := 0$. Odaberemo početne vrijednosti $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$.
Sve dok vrijedi $\mu^k > \frac{\epsilon}{n}$:

1. Odaberemo $\sigma^k \in [\theta_{min}, \theta_{max}]$ te neka je $\mu^k = \frac{(x^k)^T s^k}{n}$. Riješimo sustav (2.20).
2. Odaberemo $\alpha_k$ tako da vrijedi $(x^k, y^k, s^k) + \alpha_k(\Delta x^k, \Delta y^k, \Delta s^k) \in \mathcal{N}_{-\infty}(\gamma)$
3. Postavimo $(x^{k+1}, y^{k+1}, s^{k+1}) := (x^k, y^k, s^k) + \alpha_k(\Delta x^k, \Delta y^k, \Delta s^k)$.
4. $k = k + 1$.

Što se tiče odabira parametara i u ovom i u prethodnom algoritmu, uvijek biramo najveći mogući $\alpha_k$ za koji vrijedi uvjet u algoritmu, dok za $\sigma$ postoji nekoliko mogućnosti: uz rubne vrijednosti (0 ili 1), jedan on najčešćih odabira je $\sigma^k = 1 - \frac{1}{\sqrt{n}}, \forall k$. Također, za $\sigma_{min}$ najčešće uzimamo 0.01, dok za $\sigma_{max}$ uzimamo 0.75.

Sada ćemo reći nešto o nalaženju početne točke $(x^0, y^0, s^0)$. Primijetimo da smo u oba algoritma koja smo naveli (i opći i long-step) počeli sa strogo dopustivom točkom iz skupa $\mathcal{F}^+$. Pronalaženje takve točke nije trivijalno, no možemo prilagoditi početnu točku koja nije dopustiva pomoću nekih modifikacija sustva kojeg rješavamo u svakom koraku. Zato ćemo za početnu točku zahtijevati da zadovoljava samo $x^0 > 0$ i $s^0 > 0$, što nije teško generirati. Promotrimo ponovo sustav (2.17):

$$F(x, y, s) = \begin{pmatrix} A^T y - Qx + s - c \\ Ax - b \\ XSe - \tau e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Newtonov korak sada opet određujemo iz sustava (2.12), a nakon uvrštavanja Jacobijeve matrice i matrice $F$ dobivamo

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} Qx^k - A^T y^k - s^k + c \\ b - Ax^k \\ \tau e - X^k S^k e \end{pmatrix}. \tag{2.25}$$

Usporedimo li dobiveni sustav sa sustavom (2.18), vidimo da s desne strane u (2.25) u prva dva retka matrice nemamo nule jer ne pretpostavljamo da točka $(x^k, y^k, s^k)$ zadovoljava uvjete

$$Ax^k = b \quad i$$
$$A^T y^k - Qx^k + s^k = c.$$

Dakle, ako u našim algoritmima umjesto sustava (2.20) rješavamo gornji sustav, odnosno sustav

$$\begin{pmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} Qx^k - A^T y^k - s^k + c \\ b - Ax^k \\ \sigma^k \mu^k e - X^k S^k e \end{pmatrix}, \tag{2.26}$$

osigurat ćemo da oni rade i sa točkama koje nisu dopustive. U tom slučaju istovremeno tražimo i dopustivu i optimalnu točku.

## 2.6 Primjer

**Primjer 22.** *U ovom ćemo primjeru demonstrirati metodu dugog koraka. Pokazat ćemo da su zadovoljeni određeni uvjeti i provesti jednu iteraciju algoritma.*

*Neka je dan problem oblika*

$$\begin{aligned} \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix} \begin{pmatrix} 0.01 & 0.005 & 0 & 0 \\ 0.005 & 0.01 & 0 & 0 \\ 0 & 0 & 0.04 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} &\to \min \\ x_1 + x_2 + x_3 &= 1 \\ -x_2 + x_3 + x_4 &= 0.1 \\ x_i &\geq 0\,, i = 1, 2, 3, 4 \end{aligned} \tag{2.27}$$

*te neka je dana sljedeća iteracija:*

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0.1 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.001 \\ -0.001 \end{pmatrix}, \quad s = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} 0.004 \\ 0.003 \\ 0.0133 \\ 0.001 \end{pmatrix}.$$

*Vidimo da je* $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{pmatrix}$ $m \times n$ *matrica ranga* $m = 2$*, dok je* $b = \begin{pmatrix} 1 \\ 0.1 \end{pmatrix}$ *i* $c = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

*Matrica Q je pozitivno semidefinitna jer za svaki* $x \geq 0$ *vrijedi*

$$x^T Qx = 0.01x_1^2 + 0.01x_1x_2 + 0.01x_2^2 + 0.04x_3^2 \geq 0\,.$$

*Pokažimo sada je zadana točka* $(x, y, s) \in \mathcal{F}^+$ . *Računamo*

$$Ax = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0.1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.1 \end{pmatrix} = b\,,$$

$$\begin{aligned} A^T y - Qx + s &= \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0.0001 \\ -0.001 \end{pmatrix} - \begin{pmatrix} 0.01 \cdot \frac{1}{3} + 0.005 \cdot \frac{1}{3} \\ 0.005 \cdot \frac{1}{3} + 0.01 \cdot \frac{1}{3} \\ 0.04 \cdot \frac{1}{3} \\ 0 \end{pmatrix} + \begin{pmatrix} 0.004 \\ 0.003 \\ 0.0133 \\ 0.001 \end{pmatrix} \\ &= \begin{pmatrix} 0.001 \\ 0.002 \\ 0 \\ -0.001 \end{pmatrix} - \begin{pmatrix} 0.005 \\ 0.005 \\ 0.0133 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.004 \\ 0.003 \\ 0.0133 \\ 0.001 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = c\,. \end{aligned}$$

*Nadalje, očito vrijedi još* $x > 0$ *i* $s > 0$*, pa iz navedenog slijedi da je* $(x, y, s) \in \mathcal{F}^+$ . *Spomenuta točka nije na centralnom putu jer ne vrijedi* $x_i s_i = \tau$ *za svaki* $i = 1, 2, 3, 4$*, npr.*

$$x_1 s_1 = \frac{1}{3} \cdot 0.004 \neq \frac{1}{3} \cdot 0.003 = x_2 s_2\,.$$

*Promotrimo sada širu okolinu oko centralnog puta zadanu parametrom* $\gamma = 0.05$*, tj.* $\mathcal{N}_{-\infty}(0.05)$ . *Imamo*

$$\mu = \frac{x^T s}{n} = \frac{0.006867}{4} = 0.0017167\,,$$

*pa je* $\gamma\mu = 0.00008533$ . *Sada računamo*

$$\begin{aligned} x_1 s_1 &= 0.00133\,, \\ x_2 s_2 &= 0.001\,, \\ x_3 s_3 &= 0.00443\,, \\ x_4 s_4 &= 0.0001\,, \end{aligned}$$

*iz čega se vidi da je* $x_i s_i \geq \gamma\mu$ *za svaki* $i = 1, 2, 3, 4$ *, pa zaključujemo da je zadana točka* $(x, y, s) \in \mathcal{N}_{-\infty}(0.05)$ .

*Sada ćemo izračunati sljedeću iteraciju. Za parametre ćemo uzeti* $\gamma = 0.05$*,* $\sigma_{min} = 0.2$ *i* $\sigma_{max} = 0.8$*. Polazna točka bit će ona koju smo ranije zadali, dakle* $(x^k, y^k, s^k) = (x, y, s)$ *i* *kao što smo već vidjeli, ona zadovoljava pretpostvaku da se nalazi u* $\in \mathcal{N}_{-\infty}(0.05)$. *Uzet ćemo* $\sigma^k = 0.5$, *dok smo* $\mu^k$ *već izračunali,* $\mu^k = \mu = 0.0017167$. *Prema algoritmu, sada rješavamo sustav (2.20), koji nakon uvrštavanja vrijednosti glasi*

$$\begin{pmatrix} -0.01 & -0.005 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ -0.005 & -0.01 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -0.04 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.004 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0.003 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0.0133 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0.001 & 0 & 0 & 0 & 0 & 0 & 0.1 \end{pmatrix} \begin{pmatrix} \Delta x_1^k \\ \Delta x_2^k \\ \Delta x_3^k \\ \Delta x_4^k \\ \Delta y_1^k \\ \Delta y_2^k \\ \Delta s_1^k \\ \Delta s_2^k \\ \Delta s_3^k \\ \Delta s_4^k \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -0.000475 \\ -0.000142 \\ -0.003575 \\ 0.000758 \end{pmatrix} .$$

*Gornji sustav ima rješenje jer je rang matrice sustava jednak rangu proširene matrice sustava (kada joj dodamo stupac rješenja) i taj rang iznosi* $10$. *Kako je matrica sustava kvadratna, tj. imamo isti broj jednadžbi i nepoznanica i regularna je (jer je punog ranga), slijedi da za navedeni sustav imamo jedinstveno rješenje i ono je dano s*

$$\Delta x^k = \begin{pmatrix} -0.071903 \\ 0.237648 \\ -0.165745 \\ 0.403393 \end{pmatrix},$$

$$\Delta s^k = \begin{pmatrix} -0.000562 \\ -0.002564 \\ -0.004112 \\ 0.003549 \end{pmatrix} \quad i$$

$$\Delta y^k = \begin{pmatrix} 0.001031 \\ -0.003549 \end{pmatrix}.$$

*Sada moramo naći* $\alpha_k$ *takav da je* $(x^k, y^k, s^k) + \alpha_k(\Delta x^k, \Delta y^k, \Delta s^k) \in \mathcal{N}_{-\infty}(0.05)$. *U praksi se često bira najveći mogući* $\alpha_k$ *koji zadovoljava dani uvjet, pa primjetimo da za* $\alpha_k = 1$

*imamo*

$$x^{k+1} = x^k + \Delta x^k = \begin{pmatrix} 0.261431 \\ 0.570981 \\ 0.167589 \\ 0.503393 \end{pmatrix},$$

$$y^{k+1} = y^k + \Delta y^k = \begin{pmatrix} 0.002031 \\ -0.004549 \end{pmatrix},$$

$$s^{k+1} = s^k + \Delta s^k = \begin{pmatrix} 0.003438 \\ 0.000436 \\ 0.009188 \\ 0.004549 \end{pmatrix}$$

*i u tom slučaju mora vrijediti* $x_i^{k+1} s_i^{k+1} \geq \gamma \cdot \mu^{k+1} = 0.05 \cdot \frac{(x^{k+1})^T s^{k+1}}{4} = 0.05 \cdot 0.001244 = 0.0000622$. *Računamo*

$$\begin{aligned} x_1^{k+1} s_1^{k+1} &= 0.0008987547 \\ x_2^{k+1} s_2^{k+1} &= 0.0002490447 \\ x_3^{k+1} s_3^{k+1} &= 0.0015398402 \\ x_4^{k+1} s_4^{k+1} &= 0.0022901390, \end{aligned}$$

*iz čega vidimo da je za novu točku zadovoljen gornji uvjet.*

*Primijetimo da je* $\mu^{k+1} = 0.001244 < 0.0017167 = \mu^k$, *pa smo ovom iteracijom dobili točku koja je bliža optimalnom rješenju od polazne. Time je dovršena jedna iteracija algoritma.*

# Bibliografija


[1] S. O. Krumke, *Interior Point Methods*, Technische Universität Kaiserslautern, 2004.

[2] G. Cornuejols, R. Tütüncu, *Optimization Methods in Finance*, Cambridge, 2007.

[3] L. Čaklović, *Geometrija linearnog programiranja*, Element, 2010.

[4] A. Geletu, Quadratic programming problems - a review on algorithms and applications, dostupno na `https://www.tu-ilmenau.de/fileadmin/media/simulation/Lehre/Vorlesungsskripte/Lecture_materials_Abebe/QPs_with_IPM_and_ASM.pdf` (kolovoz 2019.)

[5] R. M. Freund, Introduction to Optimization, and Optimality Conditions for Unconstrained Problems, dostupno na `http://ocw.nur.ac.rw/NR/rdonlyres/Sloan-School-of-Management/15-084JSpring2004/8E57ABB0-7419-42AF-BB86-A6ACD754E493/0/lec1_unconstr_opt.pdf` (kolovoz 2019.)

[6] M. Passacantando, Existence of optimal solutions and optimality conditions, dostupno na `http://pages.di.unipi.it/passacantando/om/2-existence_and_opt_conditions.pdf` (rujan 2019.)

# Sažetak

U ovom radu opisuje se metoda unutrašnje točke, odnosno primarno-dualna metoda centralnog puta i primjenjuje se na zadaće linearnog i kvadratičnog programiranja.

Za linearno programiranje navode se uvjeti optimalnosti i uvodi se pojam centralnog puta. Detaljno se analizira Newtonova metoda na kojoj se temelji metoda unutrašnje točke i u tu svrhu koriste se ortogonalne projekcije. Navodi se algoritam nalaženja optimalne točke i dokazuje se rezultat o njegovoj točnosti i broju iteracija.

Za kvadratično programiranje najprije se dokazuju nužni i dovoljni uvjeti za optimalnost i opisuje se metoda centralnog puta uz modifikaciju Newtonove metode. Uvodi se pojam okoline centralnog puta i na temelju toga navode se dva algoritma primarno-dualne metode centralnog puta: metoda kratkog koraka i metoda dugog koraka, koja se demonstrira na konkretnom primjeru.

# Summary

In this paper we explain the interior point method, i.e. primal-dual path-following method and apply it to linear and quadratic optimization problems.

For linear programming we formulate optimality conditions and introduce the term of central path. We analyze Newton's method which is usually used in many interior point methods and for this purpose we use orthogonal projections. We describe algorithm of primal-dual interior point method and prove the result of it's accuracy and number of iterations.

For quadratic programming we first prove necessary and sufficient optimality conditions and describe the central path method with some modifications of Newton's method. We introduce the term of neighborhood of the central path and based on that we describe two algorithms: short-step path-following method and long-step path-following method, which we demonstrate on example.

# Životopis

Rođena sam 6.7.1994. u Čakovcu. Osnovnu školu pohađala sam u Domašincu, a nakon toga 2009. godine upisujem Gimnaziju Josipa Slavenskog Čakovec, opći smjer. Na Prirodoslovno-matematičkom fakultetu u Zagrebu 2013. godine upisujem preddiplomski sveučilišni studij, smjer Matematika. Godine 2017. stječem titulu sveučilišnog prvostupnika matematike i upisujem diplomski sveučilišni studij Financijska i poslovna matematika.