

# Belsőpontos algoritmusok önreguláris távolságfüggvénnyel

Pólik Imre\*, Terlaky Tamás†

2004. szeptember 15.

## Kivonat

A belsőpontos algoritmusok elméletének egyik érdekes kérdése a kis- és a nagylépéses módszerek komplexitásvizsgálata. Míg a gyakorlatban az utóbbiak működnek jobban, a kislépéses eljárások elméleti komplexitása jobb. Az utóbbi évek egyik jelentős eredménye Peng, Roos és Terlaky [5] nevéhez fűződik, akik önreguláris távolságfüggvények segítségével megmutatták, hogy a nagylépéses belsőpontos módszer legfeljebb  $\mathcal{O}(\sqrt{n} \log n \log(n/\varepsilon))$  lépésben  $\varepsilon$  pontossággal megoldja a lineáris optimalizálási feladatot, ezzel sokat javítottak az addigi  $\mathcal{O}(n \log(n/\varepsilon))$  eredményen, bár a komplexitás még mindig elmarad a kislépéses módszerek  $\mathcal{O}(\sqrt{n} \log(n/\varepsilon))$  lépésszámától.

Cikkünkben a monoton lineáris komplementaritási feladaton mutatjuk be az algoritmus egy újabb, egyszerűsített elemzését. A cikk alapjául Salahi és Terlaky [10] cikke szolgál.

## 1. Bevezetés

A belsőpontos algoritmusok 1984-es megjelenésével [2] lehetővé vált olyan nagy méretű lineáris optimalizálási feladatok megoldása is, amelyeket a szimplex módszerrel akkor nem, vagy csak kevésbé hatékonyan lehetett megoldani. Tíz évvel később Nesterov és Nemirovski [4] vezették be az önkorlátozó függvényeket, amelyek segítségével még általánosabb konvex optimalizálási feladatokat (mint pl. a szemidefinit vagy a másodrendűkúp-optimalizálási feladat) is polinomiális idő alatt lehetett megoldani. A szemidefinit optimalizálási feladatokat leghatékonyabban belsőpontos módszerekkel oldhatjuk meg. A belsőpontos algoritmusokról bővebben a [8, 9, 13, 14] munkákban, illetve magyarul a [3, 7] szakdolgozatokban olvashatunk.

---

\*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4L7, Kanada. Email: poliki@mcmaster.ca.

†Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4L7, Kanada. Email: terlaky@mcmaster.ca. Támogatva az NSERC Discovery Grant #5-48923 pályázat és a Canada Research Chair Program által. Mindkét szerző a „New Interior Point Methods and Software for Convex Conic-Linear Optimization and Their Application to Solve VLSI Circuit Layout Problems” MITACS kutatási projekt résztvevője.

Ebben a cikkben a monoton lineáris komplementaritási feladaton mutatjuk be egy adaptív lépéses, önreguláris függvényeket használó belsőpontos algoritmus teljes elemzését. A bizonyítás teljes, vagyis nem tartalmaz más belsőpontos cikkekre utaló külső hivatkozásokat.

## 2. A monoton lineáris komplementaritási feladat

A lineáris komplementaritási feladat általános alakja<sup>1</sup> a következő:

$$\begin{aligned} Mx + q &= s \\ x, s &\geq 0 \\ xs &= 0, \end{aligned} \tag{MLCP}$$

ahol  $M \in \mathbb{R}^{n \times n}$ ,  $x, s, q \in \mathbb{R}^n$ , az  $xs$  szorzatot pedig koordinátánként értjük (Hadamard-szorzat), vagyis  $xs = (x_1 s_1, \dots, x_n s_n)^T$ . A továbbiakban feltesszük, hogy az MLCP feladat *monoton*, vagyis  $M$  egy pozitív szemidefinit, de nem feltétlenül szimmetrikus mátrix. Ezzel az MLCP feladat tartalmazza a lineáris és a konvex kvadratikusan optimalizálási feladatot is.

Jelölje  $\mathbb{R}_+^n$  az  $n$ -dimenziós, minden koordinátájukban pozitív vektorokat,  $\mathbb{R}_\oplus^n$  pedig az  $n$ -dimenziós nemnegatív vektorokat. Definiáljuk a következő halmazokat:

**Megengedett megoldások:**  $\mathcal{F} = \{(x, s) \in \mathbb{R}_\oplus^{2n} : Mx + q = s\}$

**Szigorúan megengedett megoldások (belső pontok):**

$$\mathcal{F}_+ = \{(x, s) \in \mathbb{R}_+^{2n} : Mx + q = s\}$$

**Komplementáris (optimális) megoldások:**

$$\mathcal{F}_* = \{(x, s) \in \mathbb{R}_\oplus^{2n} : Mx + q = s, xs = 0\}$$

A továbbiakban feltesszük, hogy a szigorúan megengedett megoldások  $\mathcal{F}_+$  halmaza nem üres, vagyis létezik belső pont.

Az MLCP feladat centrális útján a szokásos módon az

$$\begin{aligned} Mx + q &= s \\ x, s &\geq 0 \\ xs &= \mu e \end{aligned} \tag{2.1}$$

---

<sup>1</sup>Szokás a lineáris komplementaritási feladatot a látszólag általánosabb

$$\begin{aligned} Qx + Rs &= -q \\ xs &= 0 \\ x, s &\geq 0 \end{aligned}$$

alakban is tárgyalni, ahol feltesszük, hogy a  $[Q|R] \in \mathbb{R}^{n \times 2n}$  mátrix teljes rangú. Ez a  $Q = M$ ,  $R = -I$  esetben éppen a fenti definíciót adja. Azonban mivel feltettük, hogy  $[Q|R]$  teljes rangú, ezért az  $x$  és  $s$  vektorok megfelelő koordinátáit kicserélve mindig elérhető, hogy az  $R$  mátrix invertálható legyen, vagyis a két felírás lényegében ekvivalens.

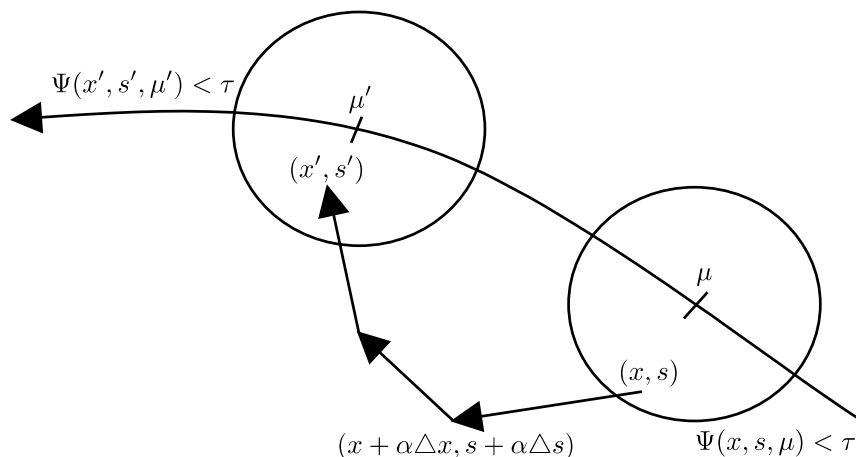
rendszer megoldáshalmazát értjük, ahol  $\mu > 0$  és  $e = (1, \dots, 1)$ . A belsőpontos algoritmusok egyik alapköve az a mára már klasszikusnak tekinthető eredmény, hogy ha létezik belső pont, akkor az  $M$  mátrixra tett feltevés miatt a fenti rendszer megoldása minden  $\mu > 0$  esetén létezik és egyértelmű, lásd például a [7, 9, 11, 12, 14] műveket. Jelölje ezt a megoldást  $(x(\mu), s(\mu))$ . A centrális út tehát egyparaméteres görbe a primál-duál változók terében. Az is könnyen látható, hogy a centrális út  $\mu \rightarrow 0$  esetén az MLCP feladat egy megoldásához tart.

## 2.1. Klasszikus belsőpontos módszerek

A primál-duál belsőpontos algoritmusok a centrális utat követik, pontosabban a centrális út *környezetében* haladnak. A környezetet általában a következőképpen definiáljuk:

$$\mathcal{N}(n, \tau) := \{(x, s) : (x, s) > 0, Mx + q = s, \Phi(x, s, \mu) \leq \eta(n, \tau), \mu > 0\}, \quad (2.2)$$

ahol  $\Phi(x, s, \mu)$  függvény méri a távolságot,  $\eta(n, \tau)$  pedig a környezet nagyságát szabályozza.<sup>2</sup> Vegyük észre, hogy a távolság függ  $\mu$ -tól is.



1. ábra. Útkövető eljárás működése

A klasszikus belsőpontos módszerek (lásd az 2. ábra algoritmusát és a 1. ábrát) egy  $(x, s)$  belső pontból indulnak, sőt még azt is feltehetjük, hogy  $\mu = 1$  esetén  $(x, s) \in \mathcal{N}(n, \tau)$ . Az optimális halmazhoz akkor kerülünk közelebb, ha a centrális út  $\mu' := (1 - \theta)\mu$  paraméterhez tartozó pontja felé próbálunk ellépni. Ehhez a következő Newton-rendszert oldjuk meg:

$$\begin{aligned} M\Delta x - \Delta s &= 0 \\ s\Delta x + x\Delta s &= \mu'e - xs. \end{aligned} \quad (2.3)$$

<sup>2</sup>A különböző távolságfüggvényekkel kapcsolatban lásd még a [9, 13] könyveket.

---

**Input:**Pontosság:  $\varepsilon > 0$ Távolság a centrális úttól:  $\tau > \nu_1^{-1}$ Frissítési paraméter:  $0 < \theta < 1$  $(x^0, s^0)$  és  $\mu^0 = 1$ , amelyekkel  $\Phi(x^0, s^0, \mu^0) \leq \tau$ **Output:**  $x$  és  $s$ , amelyekre  $x^T s < \varepsilon$  $x := x^0, s := s^0, \mu := \mu^0$ **while**  $n\mu \geq \varepsilon$  **do** $\mu := (1 - \theta)\mu$ **while**  $\Phi(x, s, \mu) \geq \tau$  **do** $\Delta x$  és  $\Delta s$  kiszámítása (2.4) alapjánAz  $\alpha$  lépéshossz kiszámítása $x := x + \alpha\Delta x, s := s + \alpha\Delta s$ **end while****end while**

---

## 2. ábra. Klasszikus Newton-lépéses útkövető algoritmus

Könnyen ellenőrizhető, hogy a rendszer megoldása

$$\begin{aligned}\Delta x &= (S + XM)^{-1}(\mu'e - xs) \\ \Delta s &= M\Delta x,\end{aligned}\tag{2.4}$$

ahol  $X$  és  $S$  az  $x$  illetve az  $s$  vektorokat tartalmazó diagonális mátrixok. Ha létezik belső pont, akkor ez a rendszer minden  $\mu > 0$  esetén megoldható, továbbá az  $M$  mátrixra tett feltevés miatt a megoldás egyértelmű.<sup>3</sup> Mivel a Newton-rendszerben a kvadratikus tagot elhanyagoltuk, ezért nem biztos, hogy a teljes  $(\Delta x, \Delta s)$  lépést megtehetjük, a megoldás inkább csak egy irányt jelöl ki. A lépéshosszt az  $\alpha \in (0, 1]$  paraméter szabályozza, vagyis a lépés után kapott pont  $(x + \alpha\Delta x, s + \alpha\Delta s)$ . Célunk, hogy ez a pont minél közelebb legyen a centrális út  $\mu$ -höz tartozó pontjához, miközben a nemnegativitási feltételt is kielégíti. Ezt addig ismételjük, amíg elég közel nem kerültünk, ekkor  $\mu$  értékét  $(1 - \theta)\mu$ -re csökkentjük és megint Newton-lépésekkel folytatjuk. Az eljárás akkor ér véget, ha a dualitásrés elér egy megadott értéket, vagyis az aktuális pont bizonyos hibahatáron belül megoldja a feladatot.

A  $\mu$  paraméter értékét többféle módon csökkenthetjük:

- Legyen  $0 < \theta < 1$  a feladat méretétől, dimenziójától független konstans. A  $\mu$  paramétert minden lépésben  $(1 - \theta)\mu$ -re cseréljük, vagyis rögzített konstanssal szorzunk. Ezek a *nagylépéses* algoritmusok.
- A csökkentés mértéke függ a feladat dimenziójától, vagyis  $\mu$  helyett  $(1 - \theta_n)\mu$ -t veszünk, ahol például  $\theta_n = \frac{1}{2\sqrt{n}}$ . Így működnek a *kislépéses* eljárások.

---

<sup>3</sup>A létezés és az egyértelműség bizonyítása megtalálható pl. a [7, 9, 12] munkákban.

- Az előző lehetőség egyik változata az, amikor a csökkentés mértéke függ az algoritmus aktuális iterációjától, vagyis *adaptívan* változtatjuk a  $\mu$  paraméter értékét, erre mutatunk példát a következő szakaszban.

Furcsa ellentmondás, hogy a gyakorlatban a nagylépéses módszerek hatékonyabban, miközben a kislépéses eljárások elméleti komplexitása jobb:  $\mathcal{O}(\sqrt{n} \log \frac{n}{\varepsilon})$  a klasszikus Newton-lépéses algoritmusok  $\mathcal{O}(n \log \frac{n}{\varepsilon})$ , illetve az önreguláris függvényeket használó algoritmusok  $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\varepsilon})$  komplexitása helyett. Nyitott kérdés, hogy elérhető-e a jobb komplexitás a nagylépéses módszerekkel. A témáról bővebben lásd az [1, 9, 13, 14] műveket.

## 2.2. Az átskálázott rendszer

Legyen  $v = \sqrt{\frac{xs}{\mu}}$  és vezessük be a következő jelöléseket:

$$d_x := \frac{v}{x} \Delta x, \quad d_s := \frac{v}{s} \Delta s. \quad (2.5)$$

Tekintsük a következő Newton-rendszert:

$$\Lambda_M d_x = d_s, \quad d_x + d_s = -\nabla \Psi(v), \quad (2.6)$$

ahol  $\Lambda_M = \mu V S^{-1} M V S^{-1}$ . Ha  $\Psi(v) = \frac{\|v\|^2 - n}{2} - \sum_{i=1}^n \log v_i$  (a logaritmusos barrierfüggvény), akkor éppen a klasszikus (2.3) Newton-rendszert kapjuk. Innen az is látszik, hogy – megfelelő skálázás után – az eredeti (2.3) rendszer megoldása a  $\Psi(v)$  függvény vetített legmeredekebb csökkenési irányának felel meg.

Bizonyos feltételek teljesülése esetén természetesen választhatunk más  $\Psi$  távolságfüggvényt is. Először is célszerű a távolságfüggvényt az imént bevezetett  $v$  változó függvényének tekinteni, ekkor a centrális út pontjaira  $v = e$  teljesül. Másrészt, mivel minden koordinátát egyformán szeretnénk kezelni ezért kézenfekvő a  $\Psi(v) = \sum_{i=1}^n \psi(v_i)$  választás valamilyen alkalmas  $\psi$  függvénnyel. A  $\psi$  függvényről mindenképpen fel kell tennünk, hogy nemnegatív,  $\psi(1) = 0$ , valamint  $\psi(t) \rightarrow \infty$  ha  $t \rightarrow 0$  vagy  $t \rightarrow \infty$ . A komplexitás elemzéséhez ennél többre lesz szükségünk, erről szól a következő szakasz.

## 3. Önreguláris függvények és tulajdonságaik

Az önreguláris függvényeket Peng, Roos és Terlaky [5] definiálta először és használta belsőpontos algoritmusokban távolságfüggvényként.

### 3.1. Az önreguláris függvények definíciója

**3.1. Definíció (Önreguláris függvény).** A  $\psi(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$  függvényt önregulárnak (*self-regular*) nevezzük, ha kétszer folytonosan differenciálható és kielégíti az alábbi két feltételt:

**SR1** *A teljes értelmezési tartományán szigorúan konvex, a  $t = 1$  pontban globális minimuma van, amelyre  $\psi(1) = \psi'(1) = 0$ . Léteznek tovább  $\nu_2 \geq \nu_1 > 0$  és  $p, q \geq 1$  konstansok, hogy minden  $t > 0$ -ra*

$$\nu_1 (t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2 (t^{p-1} + t^{-1-q}). \quad (3.1)$$

**SR2** *Minden  $t_1, t_2 > 0$ -ra és  $r \in [0, 1]$ -re*

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1-r)\psi(t_2). \quad (3.2)$$

Az első feltétel egy szigorú konvexitási tulajdonságot követel meg a  $\psi$ -ről, valamint azt, hogy a 0 és a végtelen közelében megfelelően növekedjen, a második pedig szintén egy konvexitási tulajdonság, amely ekvivalens  $\psi(\exp(\xi))$  konvexitásával.

A szokásos módon legyen  $v = \sqrt{\frac{xs}{\mu}}$  és definiáljuk a centrális utat (2.1) alapján. Legyen továbbá  $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$  önreguláris függvény. Ekkor a centrális úttól mért távolságot a

$$\Phi(x, s, \mu) := \Psi(v) = \sum_{i=1}^n \psi(v_i) \quad (3.3)$$

képlettel definiálhatjuk.

Először vizsgáljuk meg, hogy értelmes-e ez a definíció. Egy  $(x, s)$  pontpár pontosan akkor van rajta a centrális úton, ha  $xs = \mu e$ , vagyis  $v = e$ . Ebben az esetben  $\Phi(x, s, \mu) = 0$ , sőt – a  $\psi$  függvény szigorú konvexitása miatt – a fordított állítás is igaz. Másrészt, ha az  $(x, s)$  vektor valamelyik koordinátája „közel van” 0-hoz, akkor  $\Phi(x, s, \mu)$  „nagy lesz”, ami azt jelenti, hogy a fenti  $\Phi$  függvény egyszerre bünteti a centrális úttól való eltávolodást és a megengedett tartomány határához való közeledést, megfelelően a fent kirótt követelményeknek. A kétféle büntetés arányát a  $\psi$  definíciójában szereplő  $p$  és  $q$  paraméterek szabályozzák.

Felmerülhet még a kérdés, hogy létezik-e egyáltalán önreguláris függvény. Erre könnyen válaszolhatunk, hiszen  $\psi(1) = \psi'(1) = 0$  miatt

$$\psi(t) = \int_1^t \int_1^\zeta \psi''(\xi) d\xi d\zeta. \quad (3.4)$$

Ha most a  $\nu_1 = \nu_2 = 1$  esetet tekintjük, akkor azt kapjuk, hogy  $\psi''(t) = t^{p-1} + t^{-1-q}$ , ami az SR1 tulajdonságnál megkövetelt  $\psi(1) = \psi'(1) = 0$  feltételekkel együtt egyértelműen meghatároz egy  $\psi(t)$  függvényt. Ezt a függvényt a továbbiakban  $\Upsilon_{p,q}(t)$ -vel jelöljük.

$$\begin{aligned} \Upsilon_{p,1}(t) &= \frac{t^{p+1} - 1}{p(p+1)} - \log t + \frac{p-1}{p}(t-1) \\ \Upsilon_{p,q}(t) &= \frac{t^{p+1} - 1}{p(p+1)} + \frac{t^{1-q} - 1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad q > 1. \end{aligned} \quad (3.5)$$

Vegyük észre, hogy  $\Upsilon_{p,q}(t) \rightarrow \Upsilon_{p,1}(t)$  ha  $q \rightarrow 1$ .

Az önreguláris függvények másik fontos osztálya a  $\Gamma_{p,q}(t)$  család:

$$\Gamma_{p,q}(t) = \frac{t^{p+1} - 1}{p+1} + \frac{t^{1-q} - 1}{q-1}, \quad p \geq 1, q > 1, \quad (3.6)$$

ahol<sup>4</sup>  $\nu_1 = 1$  és  $\nu_2 = q$ .

### 3.2. Az önreguláris függvények alapvető tulajdonságai

Először vizsgáljuk meg az önreguláris függvények definíciójában szereplő SR1 és SR2 feltétel kapcsolatát.

**3.2. Lemma ([5]).** *Legyen  $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$  kétszer folytonosan differenciálható függvény. Az alábbi állítások ekvivalensek:*

1. *A  $\psi(t)$  függvény kielégíti az SR2 feltételt.*
2. *A  $\psi(\exp(\xi))$ ,  $\xi \in \mathbb{R}$  függvény konvex.*
3. *Minden  $t > 0$  esetén  $\psi'(t) + t\psi''(t) \geq 0$ .*
4. *Minden  $t_1, t_2 > 0$  esetén  $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}\psi(t_1) + \frac{1}{2}\psi(t_2)$ .*

Szintén könnyen igazolhatók az alábbi tulajdonságok, részletes bizonyításaik megtalálhatók [5]-ben.

**3.3. Állítás ([5]).** *Ha a  $\psi_1(t), \psi_2(t)$  függvények önregulárisak, akkor önreguláris minden  $\beta_1\psi_1 + \beta_2\psi_2$ ,  $\beta_1, \beta_2 \geq 0$ ,  $\beta_1 + \beta_2 > 0$  pozitív lineáris kombinációjuk is, tehát az önreguláris függvények halmaza konvex kúp. Ha  $\psi(t)$  önreguláris, akkor  $-\psi(t)$  nem az, vagyis a kúp hegyes. Végül az önreguláris függvények kúpja nem zárt.*

A következő állítás az önreguláris függvények néhány alapvető és hasznos tulajdonságát tartalmazza.

**3.4. Állítás.** *Legyen  $\Omega_1$  és  $\Omega_2$  rendre az SR1 illetve az SR2 feltételt kielégítő függvények halmaza és legyen  $t > 0$ . A következők igazak:*

1. *Legyen  $\psi(t) \in \Omega_1$ , ekkor  $|\frac{1}{t}\psi'(t)| \leq \frac{\nu_2}{\nu_1}\psi''(t)$ .*
2. *Ha  $\psi(t) \in \Omega_1$  és  $\nu_1 = \nu_2$ , akkor  $\psi(t)$  önreguláris.*
3. *Ha  $\psi(t) \in \Omega_1$ , akkor  $2\nu_1\psi(t) \leq \psi'(t)^2$ .*
4. *Ha  $\psi(t) = \psi(t^{-1})$  és  $\psi(t) \in \Omega_1$ , akkor  $\psi(t)$  önreguláris.*
5. *Ha  $\psi(t) \in \Omega_2$ , akkor  $\psi(t^{-1}) \in \Omega_2$ .*

<sup>4</sup>Valójában csak annyit láttunk, hogy  $\Upsilon_{p,q}$  és  $\Gamma_{p,q}$  kielégítik az SR1 feltételt. A további bizonyításokhoz már ez is elég, azonban bizonyítható, hogy ezek a függvények tényleg önregulárisak.

6. Legyen  $N$  pozitív egész szám,  $\beta_0 \in \mathbb{R}$ ,  $\beta_i \geq 0$ ,  $\rho_i \in \mathbb{R}$ ,  $i = 1, \dots, N$ . Ekkor

$$\psi(t) = \beta_0 \log t + \sum_{i=1}^N \beta_i (t^{\rho_i} - 1) \quad (3.7)$$

függvény teljesíti az SR2 tulajdonságot.

Ha az irodalomban szokásos távolságfüggvényeket vizsgáljuk, akkor azt találjuk, hogy az  $\left\| \sqrt{\frac{xs}{\mu}} - e \right\|$  távolság nem önreguláris (nem bünteti közvetlenül a megengedett tartomány határához való közeledést, míg a  $\|v - v^{-1}\|^2$  távolság önreguláris ( $p = 1$ ,  $q = 3$ ). Ez talán szolgálhat némi magyarázattal az utóbbi távolsággal elérhető jobb eredményekre.

### 3.3. Önreguláris függvények és a lineáris komplementaritási feladat

A továbbiakban a lineáris komplementaritási feladatok szemszögéből tekintjük az önreguláris függvényeket. Először általában vizsgáljuk a  $\Gamma_{1,q}(t)$  függvényeket, majd részletesebben elemezzük a  $q = 3$  és a  $q = \log(n) + 1$  eseteket. Ezekre a függvényekre épül majd a következő szakasz belső pontos algoritmus, amelyet aztán részletesen elemezzünk.

A  $\Gamma_{1,q}(t) = \frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}$  függvény alapján definiáljuk a következő távolságfüggvényt:

$$\Phi_q(x, s, \mu) := \Psi_q(v) = \sum_{i=1}^n \Gamma_{1,q}(v_i) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{1-q} - n}{q-1}. \quad (3.8)$$

A következő állításban meghatározzuk a függvény  $\mu$ -szerinti globális minimumát:

**3.5. Állítás.** [10] Legyen  $(x, s) > 0$  rögzített, ekkor a  $\Phi_q(x, s, \mu)$  függvény  $\mu$  szerinti globális minimumhelye:

$$\mu_q^* = \left( \frac{x^T s}{\left( x^{\frac{1-q}{2}} \right)^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q+1}}.$$

Innen egyszerűen adódik az alábbi következmény:

**3.6. Következmény.** [10] Minden rögzített  $(x, s) > 0$  esetén a  $\Phi_q(x, s, \mu)$  függvény  $\mu$ -ben csökken, ha  $\mu \leq \mu_q^*$ , és növekszik, ha  $\mu > \mu_q^*$ .

A centrális út környezetének definiálásakor lesz fontos szerepe a következő lemmának:



**3.7. Lemma.** [10] Legyen  $\tau \in \mathbb{R}, \tau \geq 2$ . Legyen

$$\mu_q^h = \left( \frac{n}{\left( x^{\frac{1-q}{2}} \right)^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q-1}}$$

az  $xs$  vektor komponenseinek általánosított harmonikus közepe és

$$\mu_g := \frac{x^T s}{n} \quad (3.9)$$

az aktuális dualitásréshez tartozó  $\mu$  érték. A következő állítások ekvivalensek:

1.  $\frac{\mu_g}{\mu_q^h} \leq \tau$ ,
2.  $\Phi_q(x, s, \frac{\mu_g}{\tau}) \leq \frac{(\tau-1)n}{2}$ ,
3.  $\Phi_q(x, s, \mu_g) \leq \frac{\tau^{\frac{q-1}{2}} - 1}{q-1} n$ .

A távolságfüggvény speciális  $\Psi(v) = \sum_{i=1}^n \Gamma_{1,q}(v_i)$  választása miatt a (2.6) Newton-rendszer az eredeti  $\Delta x$  és  $\Delta s$  változókat használva a következő alakban írható fel:

$$\begin{aligned} M\Delta x - \Delta s &= 0 \\ s\Delta x + x\Delta s &= \mu^{\frac{q+1}{2}} \left( x^{\frac{1-q}{2}} \right) s^{\frac{1-q}{2}} - xs. \end{aligned} \quad (3.10)$$

Jelölje ennek a rendszernek az egyértelmű megoldását, másnéven a Newton-lépést,  $(\Delta x(\mu), \Delta s(\mu))$ . A következő két lemma azt adja meg, hogy különböző  $\mu$  értékek esetén mennyivel csökken a dualitásrés a Newton-lépés után.

**3.8. Lemma.** [10] Legyen  $(\Delta x(\mu_q^*), \Delta s(\mu_q^*))$  a (3.10) rendszer megoldása  $\mu = \mu_q^*$  mellett. Ekkor

$$x^T \Delta s(\mu_q^*) + s^T \Delta x(\mu_q^*) = 0.$$

**Bizonyítás:** Egyszerűen adódik  $\mu_q^*$  definíciójából. □

Ebből a lemmából a következő becslést kapjuk a dualitásrésre:

**3.9. Következmény.** [10] Ha  $\mu = \mu_q^*$ , akkor a dualitásrés semmilyen  $\alpha$  esetén nem csökken, vagyis

$$(x + \alpha \Delta x(\mu_q^*))^T (s + \alpha \Delta s(\mu_q^*)) \geq x^T s.$$

Hasonló eredményeket kapunk  $\mu = \mu_q^h$  esetén.

**3.10. Lemma.** [10] Legyen  $(\Delta x(\mu_q^h), \Delta s(\mu_q^h))$  a (3.10) rendszer megoldása  $\mu = \mu_q^h$  mellett. Ekkor

$$x^T \Delta s(\mu_q^h) + s^T \Delta x(\mu_q^h) = n\mu_q^h - x^T s.$$

**Bizonyítás:** Egyszerűen adódik  $\mu_q^h$  definíciójából. □

**3.11. Következmény.** Ha  $\mu = \mu_q^h$ , akkor a dualitásrés a következőképpen változik:

$$(x + \alpha \Delta x (\mu_q^h))^T (s + \alpha \Delta s (\mu_q^h)) \geq x^T s \left( 1 - \alpha + \frac{\mu_q^h \alpha}{\mu_g} \right).$$

A (3.8) definiáló egyenlőségből egyszerű behelyettesítéssel adódik, hogy a  $q = 3$  esetben

$$\Phi_3(x, s, \mu) := \Psi_3(v) = \frac{1}{2} \|v - v^{-1}\|^2. \quad (3.11)$$

**3.12. Állítás.** [6] Rögzített  $(x, s) > 0$  esetén a  $\Phi_3(x, s, \mu)$  távolságfüggvény  $\mu$  szerinti  $\mu_3^*$  minimumhelyére teljesül, hogy

$$\mu_3^* = \sqrt{\frac{x^T s}{x^{-T} s^{-1}}} = \sqrt{\mu_g \mu_3^h}.$$

**Bizonyítás:** Egyszerűen következik a 3.5. állításból és a 3.7 lemmából. □

**3.13. Állítás.** [6] Rögzített  $(x, s) > 0$  mellett

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^h),$$

és

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^*) + \frac{\Phi(x, s, \mu_3^*)^2}{2n}.$$

A  $q = \log n + 1$  esetben a távolságfüggvény:

$$\Phi_\ell(x, s, \mu) := \Psi_\ell(v) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{-\log n} - n}{\log n}. \quad (3.12)$$

**3.14. Állítás.** Rögzített  $(x, s) > 0$  mellett a  $\Phi_\ell(x, s, \mu)$  függvény  $\mu$  szerinti minimumhelye  $\mu_\ell^* := \mu_{\log n + 1}^*$ .

**Bizonyítás:** Egyszerűen adódik a 3.5. állításból. □

## 4. Adaptív lépéses útkövető algoritmus önreguláris távolsággal

Az itt bemutatott algoritmus újdonsága, hogy a  $\mu$  paraméter értékét adaptívan csökkentti. Hasonló rugalmasságot már régóta alkalmaznak a gyakorlatban, ám viszonylag kevés ismert a módszer elméleti tulajdonságairól.

Legyen  $\tau \geq 2$  és legyen  $\eta(n, \tau) = \frac{(\tau-1)n}{2}$  a távolságfüggvény maximális megengedett értéke. A  $\mu = \mu_q^t$  paraméter értékét minden iterációban úgy módosítjuk, hogy a  $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$  feltétel teljesüljön. Könnyen látható, hogy ez pontosan akkor teljesül, ha  $\mu_q^t$  az alábbi egyenlet megoldása:

$$2 \left( x^{\frac{1-q}{2}} \right)^T s^{\frac{1-q}{2}} \mu^{\frac{q+1}{2}} - (2n + \tau(q-1)n) \mu + (q-1)x^T s = 0. \quad (4.1)$$

Ennek az egyenletnek két megoldása van, az egyik kisebb, mint  $\mu_q^*$ , a másik nagyobb. Mivel mi  $\mu$  értékét minél kisebbnek szeretnénk választani, ezért a két gyök közül a kisebbet fogjuk használni. Könnyen belátható, hogy ha  $\mu_g \leq \tau \mu_q^h$ , akkor  $\mu_q^t \leq \mu_q^h$  és egyenlőség pontosan akkor áll fenn, ha  $\mu_g = \tau \mu_q^h$ . A másik irányú becslésről szól a következő lemma:

**4.1. Lemma.** *Legyen  $\mu_q^t$  az (4.1) egyenlet kisebbik gyöke. Ekkor*

$$\mu_q^h \leq 2\tau \mu_q^t.$$

**Bizonyítás:** Az (4.1) egyenlet bal oldalán szereplő függvény a  $\mu$  változó konvex függvénye. A  $\mu = \frac{\mu_q^h}{2\tau}$  értéket behelyettesítve és felhasználva, hogy  $\mu_g \geq \mu_q^h$  azt kapjuk, hogy a függvényérték alulról becsülhető a

$$\frac{2}{(2\tau)^{\frac{q+1}{2}}} - \frac{1}{\tau} + \frac{q-1}{2} \quad (4.2)$$

mennyiséggel. A konvexitás miatt elég megmutatni, hogy ez a kifejezés nemnegatív. Mivel  $\tau \geq 2$ , így a (4.2) kifejezés  $q \geq 1$  esetén a  $q$  paraméter szigorúan konvex növekvő függvénye, továbbá  $q = 1$  esetén a függvényérték 0, tehát  $q > 1$  esetén pozitív. Ezzel a lemmát beláttuk.  $\square$

A következő lemma a  $v - v^{-q}$  vektor normájára ad alsó becslést:

**4.2. Lemma.** *Ha  $\tau \geq 2$  akkor  $\sigma = \|v - v^{-q}\| \geq 1$ .*

**Bizonyítás:** A 3.4 állítás 3. pontjának felhasználásával kapjuk, hogy

$$\sigma^2 \geq 2\Phi(x, s, \mu_q^t) = (\tau-1)n \geq 1, \quad \forall n \geq 1.$$

$\square$

A következő lemmában azt vizsgáljuk, hogyan változik a dualitásrés, miközben a centrális út  $\mu_q^t$ -hez tartozó pontja felé lépünk.

**4.3. Lemma.** *Legyen  $(\Delta x(\mu_q^t), \Delta s(\mu_q^t))$  a (3.10) rendszer megoldása  $\mu = \mu_q^t$  esetén, ahol  $\mu_q^t$  a (4.1) egyenlet kisebbik gyöke. Ekkor*

$$x^T \Delta s(\mu_q^t) + s^T \Delta x(\mu_q^t) = \mu_q^t \frac{q+1}{2} \left( x^{\frac{1-q}{2}} \right)^T s^{\frac{1-q}{2}} - x^T s,$$

és

$$(x + \alpha \Delta x (\mu_q^t))^T (s + \alpha \Delta s (\mu_q^t)) \geq x^T s \left( 1 - \alpha + \frac{\mu_q^{t \frac{q+1}{2}} \alpha}{\mu_g \mu_q^{h \frac{q-1}{2}}} \right). \quad (4.3)$$

**Bizonyítás:** Egyszerűen kiszámolható a definíciók és a rendszer monotonitásának felhasználásával.  $\square$

**4.4. Megjegyzés.** Ha  $\mu_q^t \approx \mu_q^h$ , akkor az önreguláris függvényből kapott keresési irány hasonló dualitásrés-csökkenést ér el mint a Newton-lépés. Ha viszont  $\mu_q^t \ll \mu_q^h$ , akkor az iménti eredmény szerint az önreguláris keresési irány sokkal hatékonyabb. További motivációk találhatók a [10] cikkben.

Az új, adaptív algoritmus (lásd a 3. ábrát) az 2 algoritmus speciális esete. Ebben az eljárásban – függetlenül az aktuális pont centrális úttól való távolságától – mindig nagy lépést teszünk, vagyis  $\mu$  értékét  $\mu = \mu_q^t \leq \mu_g/\tau$ -ra csökkentjük, majd egy Newton-lépést végzünk. Végül definiáljuk a centrális út környezetét a

---

**Input:**

Pontosság:  $\varepsilon > 0$

Távolság a centrális úttól:  $\tau > 2$

$(x^0, s^0)$ , amelyekkel  $\mu_g/\mu_q^h \leq \tau$

$x := x^0, s := s^0$

**while**  $x^T s \geq \varepsilon$  **do**

$\mu := \mu_q^t$  a (4.1) egyenlet kisebbik gyöke

$\Delta x$  és  $\Delta s$  kiszámítása (3.10) alapján

Olyan  $\alpha$  lépéshossz kiszámítása, amelyre

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{24q}$$

és  $\mu_g(\alpha) \leq \tau \mu_q^h(\alpha)$

$x := x + \alpha \Delta x, s := s + \alpha \Delta s$

**end while**

**Output:**  $x$  és  $s$ , amelyekre  $x^T s < \varepsilon$

---

3. ábra. Adaptív útkövető algoritmus

következő módon:

$$\mathcal{N}_q(n, \tau) := \{(x, s) : (x, s) > 0, Mx + q = s, \Phi_q(x, s, \mu) \leq \eta_q(n, \tau), \mu > 0\}, \quad (4.4)$$

ahol  $\Phi_q(x, s, \mu)$ -t a (3.8) egyenlet definiálja,  $\tau \geq 2$  és legyen

$$\eta_q(n, \tau) := \frac{\left( \tau^{\frac{q-1}{2q}} - 1 \right) n}{q-1}. \quad (4.5)$$

Az algoritmus komplexitásának elemzése során az  $\alpha$  lépéshosszra fogunk alsó becslést adni. Célunk annak bizonyítása, hogy ezáltal a  $\mu_g$  paramétert kellő mértékben tudjuk csökkenteni. A továbbiakban tehát a  $\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha))$  és a  $\Phi_q(x(\alpha), s(\alpha), \mu_q^t)$  függvények növekedési tulajdonságait vizsgáljuk, ahol  $\mu_g(\alpha) = (x + \alpha\Delta x)^T (s + \alpha\Delta s)/n$ .

Kezdjük egy technikai lemmával.

**4.5. Lemma.** *Legyen  $\sigma = \|\nabla\Psi_q(v)\|$ . Ekkor*

$$v_{\min} \geq (1 + \sigma)^{-\frac{1}{q}}.$$

**Bizonyítás:** Ha  $v_{\min} \geq 1$ , akkor a lemma állítása triviális. Másrészt ha  $v_{\min} < 1$ , akkor  $\sigma = \|v - v^{-q}\| \geq v_{\min}^{-q} - v_{\min} \geq v_{\min}^{-q} - 1$ . Ezzel a lemmát beláttuk.  $\square$

Ebből az eredményből egyszerű alsó becslést kapunk az  $\alpha$  lépéshosszra.

**4.6. Lemma.** *Legyen  $(\Delta x, \Delta s)$  a (3.10) rendszer megoldása, ahol  $\mu = \mu_q^t$  a (4.1) egyenlet kisebbik gyöke és  $\sigma = \|\nabla\Psi_q(v)\|$ . Legyen  $\alpha_{\max}$  az a maximális lépéshossz, amelyre  $x + \alpha_{\max}\Delta x \geq 0$  és  $s + \alpha_{\max}\Delta s \geq 0$ . Ekkor*

$$\alpha_{\max} \geq \bar{\alpha} := \frac{1}{\sigma(1 + \sigma)^{\frac{1}{q}}}.$$

**Bizonyítás:** A  $v(\alpha)$  változó definíciója alapján írhatjuk, hogy

$$\begin{aligned} v(\alpha_{\max}) &= \sqrt{\frac{x(\alpha_{\max})s(\alpha_{\max})}{\mu}} = \sqrt{\frac{(x + \alpha_{\max}\Delta x)(s + \alpha_{\max}\Delta s)}{\mu}} = \\ &= \sqrt{\frac{xs}{\mu} \left(e + \alpha_{\max}\frac{x}{\Delta x}\right) \left(e + \alpha_{\max}\frac{s}{\Delta s}\right)} = v \left(e + \alpha_{\max}v^{-1}d_x\right)^{\frac{1}{2}} \left(e + \alpha_{\max}v^{-1}d_s\right)^{\frac{1}{2}}. \end{aligned} \quad (4.6)$$

A  $v(\alpha_{\max}) \geq 0$  feltételhez elég, ha

$$e + \alpha_{\max}v^{-1}d_x \geq 0 \quad \text{és} \quad e + \alpha_{\max}v^{-1}d_s \geq 0.$$

Ezekből az egyenletekből azt kapjuk, hogy

$$\alpha_{\max} \geq \frac{1}{\|(v^{-1}d_x, v^{-1}d_s)\|}.$$

Az előző lemmát felhasználva kapjuk, hogy

$$\|(v^{-1}d_x, v^{-1}d_s)\| \leq \frac{\|(d_x, d_s)\|}{v_{\min}} \leq \sigma(1 + \sigma)^{\frac{1}{q}}.$$

Ezzel a lemmát beláttuk.  $\square$

Végül szükségünk lesz még a következő technikai lemmára:

**4.7. Lemma.** *Tegyük fel, hogy  $h(t)$  kétszer differenciálható konvex függvény, amelyre*

$$h(0) = 0, \quad h'(0) < 0. \quad (4.7)$$

*Tegyük fel továbbá, hogy  $h(t)$ -nek globális minimuma van a  $t_* > 0$  pontban és hogy  $h''(t)$  monoton növekvő függvény. Ekkor  $\forall t \in [0, t_*]$  esetén*

$$h(t) \leq \frac{h'(0)t}{2}. \quad (4.8)$$

**Bizonyítás:** Mivel  $h(0) = 0$ , így

$$\begin{aligned} h(t) &= \int_0^t h'(s) ds = h'(0)t + \int_0^t \int_0^s h''(z) dz ds \leq h'(0)t + \int_0^t s h''(s) ds = \\ &= h'(0)t + [s h'(s)]_0^t - \int_0^t h'(s) ds \leq h'(0)t - h(t), \end{aligned} \quad (4.9)$$

ahonnan átrendezés után kapjuk a kívánt állítást.  $\square$

Most már meg tudjuk becsülni a távolságfüggvény csökkenését.

**4.8. Tétel.** *Legyen  $(\Delta x, \Delta s)$  a (3.10) rendszer megoldása, ahol  $\mu = \mu_q^t$  a (4.1) egyenlet kisebbik gyöke. Ekkor az  $\alpha^* = \bar{\alpha}/4q$  lépéshosszra teljesül, hogy*

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}.$$

**Bizonyítás:** Legyen

$$\begin{aligned} h(\alpha) &:= \Phi_q(x(\alpha), s(\alpha), \mu_q^t) - \Phi_q(x, s, \mu_q^t) = \\ &= \frac{\|v(\alpha)\|^2 - n}{2} + \frac{\|v(\alpha)^{\frac{1-q}{2}}\|^2 - n}{q-1} - \frac{\|v\|^2 - n}{2} - \frac{\|v^{\frac{1-q}{2}}\|^2 - n}{q-1}. \end{aligned} \quad (4.10)$$

A tagokat külön becsüljük. Könnyen láthatjuk, hogy

$$\begin{aligned} v(\alpha) &= \sqrt{\frac{x(\alpha)s(\alpha)}{\mu}} = \sqrt{\frac{(x + \alpha\Delta x)(s + \alpha\Delta s)}{\mu}} = \\ &= \sqrt{\frac{xs}{\mu} \left(e + \alpha \frac{\Delta x}{x}\right) \left(e + \alpha \frac{\Delta s}{s}\right)} = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}. \end{aligned} \quad (4.11)$$

Hasonló módon győződhetünk meg arról is, hogy  $x + \alpha\Delta x \geq 0$  ekvivalens  $v + \alpha d_x \geq 0$ -val, tehát valóban felbonthatjuk a gyökjel alatti szorzatot. A számtani-

mértani közepek közötti egyenlőtlenséget használva kapjuk, hogy

$$\begin{aligned} \left\| v(\alpha)^{\frac{1-q}{2}} \right\|^2 &= \sum_{i=1}^n v_i(\alpha)^{1-q} = \sum_{i=1}^n \sqrt{(v_i + \alpha(d_x)_i)^{1-q} (v_i + \alpha(d_s)_i)^{1-q}} \leq \\ &\leq \frac{1}{2} \sum_{i=1}^n \left( (v_i + \alpha(d_x)_i)^{1-q} + (v_i + \alpha(d_s)_i)^{1-q} \right). \end{aligned} \quad (4.12)$$

Szintén  $v(\alpha)$  (4.11) alatti alakját használva kapjuk, hogy

$$\|v(\alpha)\|^2 - \|v\|^2 = (v + \alpha d_x)^T (v + \alpha d_s) - v^T v = \alpha v^T (d_x + d_s) + \alpha^2 d_x^T d_s. \quad (4.13)$$

Az  $\alpha^2 d_x^T d_s$  mennyiség becsléséhez használjuk fel, hogy bármilyen  $a, b \in \mathbb{R}$  esetén  $ab \leq \frac{(a+b)^2}{4}$ , így:

$$d_x^T d_s = \sum_{i=1}^n (d_x)_i (d_s)_i \leq \sum_{i=1}^n \frac{((d_x)_i + (d_s)_i)^2}{4} = \frac{\sigma^2}{4}, \quad (4.14)$$

miel a Newton-rendszer miatt  $d_x + d_s = -\nabla \Psi(v)$ .

A  $h(\alpha)$  függvényt tehát a következőképpen becsülhetjük:

$$\begin{aligned} h(\alpha) &\leq \frac{1}{2} \alpha v^T (d_x + d_s) + \frac{\alpha^2 \sigma^2}{4} + \frac{1}{2(q-1)} \sum_{i=1}^n (v_i + \alpha(d_x)_i)^{1-q} + \\ &+ \frac{1}{2(q-1)} \sum_{i=1}^n (v_i + \alpha(d_s)_i)^{1-q} - \frac{1}{q-1} \left\| v^{\frac{1-q}{2}} \right\|^2 =: h_1(\alpha). \end{aligned} \quad (4.15)$$

A  $h_1(\alpha)$  függvény kétszer folytonosan differenciálható, a deriváltak:

$$\begin{aligned} h_1'(0) &= -\frac{\sigma^2}{2} \\ h_1''(\alpha) &\leq \frac{q\sigma^2}{2} (v_{\min} - \alpha\sigma)^{-1-q} + \frac{\sigma^2}{2}, \end{aligned} \quad (4.16)$$

így a következő becslést kapjuk:

$$h(\alpha) \leq -\frac{\alpha\sigma^2}{2} + \frac{\alpha^2\sigma^2}{4} + \frac{q\sigma^2}{2} \int_0^\alpha \int_0^\zeta (v_{\min} - \eta\sigma)^{-1-q} d\eta d\zeta =: h_2(\alpha).$$

A  $h_2(\alpha)$  függvény kétszer folytonosan differenciálható, szigorúan konvex függvény a  $[0, \bar{\alpha}]$  intervallumon, ahol  $\bar{\alpha} := \frac{1}{\sigma(1+\sigma)^{\frac{1}{q}}}$ , tehát létezik globális minimumhelye, jelölje ezt  $\alpha_1^*$ . Nyilvánvaló, hogy  $\alpha_1^*$  az alábbi egyenlet egyértelmű megoldása:

$$h_3(\alpha) = \frac{2h_2'(\alpha)}{\sigma^2} = -1 + \alpha + q \int_0^\alpha (v_{\min} - \eta\sigma)^{-q-1} d\eta = 0.$$

Azt szeretnénk megmutatni, hogy  $\alpha^* < \alpha_1^*$ . Mivel a  $h_3(\alpha)$  függvény monoton növekvő és  $h_3(0) < 0$  ezért ez azzal ekvivalens, hogy  $h_3(\alpha^*) < 0$ . A következőt írhatjuk tehát:

$$h_3(\alpha^*) = -1 + \alpha^* + q \int_0^{\alpha^*} (v_{\min} - \eta\sigma)^{-q-1} d\eta \leq -1 + \frac{1}{4} + q\alpha^*(v_{\min} - \alpha^*q)^{-q-1},$$

ahol az integrálandó függvényt az intervallumon felvett maximumával helyettesítettük. A további felső becsléshez felhasználjuk, hogy  $v_{\min} \geq (1 + \sigma)^{-1/q}$ , és

$$(1 + \sigma)^{-1/q} - \alpha^*\sigma = (1 + \sigma)^{-1/q} \left(1 - \frac{1}{4q}\right) > 0,$$

így

$$h_3(\alpha^*) \leq -\frac{3}{4} + q\alpha^* \left((1 + \sigma)^{-1/q} - \alpha^*\sigma\right)^{-q-1} = -\frac{3}{4} + \frac{\sigma + 1}{4\sigma} \left(1 - \frac{1}{4q}\right)^{-q-1}.$$

Az  $(1 + \frac{1}{n})^n$  sorozat konvergenciatulajdonságait felhasználva egyszerűen megmutatható, hogy

$$\left(1 - \frac{1}{4q}\right)^{-q} \leq \sqrt[4]{\frac{4e}{3}},$$

így végül azt kapjuk, hogy

$$h_3(\alpha^*) \leq -\frac{3}{4} + \frac{\sigma + 1}{4\sigma} \left(1 - \frac{1}{4q}\right)^{-q-1} \leq -\frac{3}{4} + \frac{1}{3} \sqrt[4]{\frac{4e}{3}} < 0,$$

azaz  $\alpha^* < \alpha_1^*$ . A 4.7 lemmát a  $h_2(\alpha)$  függvényre alkalmazva kapjuk, hogy

$$h(\alpha^*) \leq h_2(\alpha^*) \leq -\frac{\sigma^2\alpha^*}{2} = \frac{-\sigma}{8q(1 + \sigma)^{1/q}}.$$

Végül, mivel  $\sigma \geq 1$  és  $\sigma^2 \geq \Psi^2$  azt kapjuk, hogy

$$h(\alpha^*) \leq \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}.$$

Ezzel a tételt beláttuk. □

Most azt vizsgáljuk hogyan változik a  $\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha))$  távolságfüggvény (vagy ami ezzel ekvivalens, a  $\Phi_q(x(\alpha), s(\alpha), \mu_q^*(\alpha))$  függvény) a Newton-lépés során. A  $\Phi_q(x(\alpha), s(\alpha), \mu_q^*(\alpha)) \leq \Phi_q(x(\alpha), s(\alpha), \mu_q^*)$  egyenlőtlenség miatt elég a  $\Phi_q(x(\alpha), s(\alpha), \mu_q^*)$  függvényt vizsgálni.

**4.9. Tétel.** *Legyen  $(\Delta x, \Delta s)$  a (3.10) rendszer megoldása, ahol  $\mu = \mu_q^t$  a (4.1) egyenlet kisebbik gyöke. Ekkor minden  $\alpha \leq \alpha^*$  esetén*

$$\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha)) \leq \frac{\left(\tau^{\frac{q-1}{2}} - 1\right) n}{q - 1}.$$



**Bizonyítás:** A 4.6 lemma miatt minden  $\alpha \leq \alpha^* < \bar{\alpha}$  lépéshossz megengedett, továbbá  $\mu_q^t$  választása miatt

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^t(\alpha)) = \frac{(\tau - 1)n}{2}.$$

A  $\Phi_q(x, s, \mu)$  függvény a  $\mu$  változó konvex függvénye, a minimuma  $\mu_q^*$ -ban van. Az általánosított számtani és mértani közepek közötti egyenlőtlenség felhasználásával kapjuk, hogy  $\mu_q^h(\alpha) \leq \mu_q^*(\alpha)$ , így minden  $\alpha \leq \alpha^*$  esetén

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^h(\alpha)) \leq \frac{(\tau - 1)n}{2},$$

vagyis a 3.7 lemmát felhasználva

$$\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha)) \leq \frac{\left(\tau^{\frac{q-1}{2}} - 1\right)n}{q-1}.$$

Ezzel a tételt beláttuk.  $\square$

Az algoritmus lépésszámának meghatározásához becslést kell adnunk az  $\alpha^*$  lépéshosszra, vagyis meg kell becsülnünk mennyivel csökken  $\mu_q^t$  egy lépés során. A következő technikai lemma a távolságfüggvény növekedését vizsgálja, miközben  $\mu$  csökken.

**4.10. Lemma.** *Legyen  $v^+ = \frac{v}{\sqrt{1-\theta}}$ , ahol  $\theta \in (0, 1)$ . Ekkor*

$$\Psi_q(v^+) \leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{1-\theta} \left( \frac{1 - \left(\frac{1}{2\tau}\right)^{\frac{q-1}{2}}}{q-1} \right).$$

**Bizonyítás:** A távolságfüggvény definíciójából kapjuk, hogy

$$\begin{aligned} \Psi_q(v^+) &= \frac{\|v^+\|^2 - n}{2} + \frac{\|v^{+\frac{1-q}{2}}\|^2 - n}{q-1} = \frac{\frac{1}{1-\theta}\|v\|^2 - n}{2} + \\ &+ \frac{(1-\theta)^{\frac{q-1}{2}}\|v^{\frac{1-q}{2}}\|^2 - n}{q-1} = \frac{1}{1-\theta} \left( \frac{\|v\|^2 - n}{2} + \frac{\|v^{\frac{1-q}{2}}\|^2 - n}{q-1} \right) + \\ &+ \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{(q-1)(1-\theta)} + \left( (1-\theta)^{\frac{q-1}{2}} - \frac{1}{1-\theta} \right) \frac{\|v^{\frac{1-q}{2}}\|^2}{q-1} \leq \\ &\leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{\theta}{1-\theta} \left( \frac{n - \|v^{\frac{1-q}{2}}\|^2}{q-1} \right) \leq \\ &\leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{\theta n}{1-\theta} \left( \frac{1 - \left(\frac{1}{2\tau}\right)^{\frac{q-1}{2}}}{q-1} \right), \end{aligned} \tag{4.17}$$

ahol az utolsó egyenlőtlenséget a 4.1 lemmából kapjuk.  $\square$

A 4.10 lemmát és a 4.8 tételt felhasználva a következő tételt láthatjuk be:

**4.11. Tétel.** *Legyen  $\tau \geq 2$  és legyen  $(\Delta x, \Delta s)$  a (3.10) rendszer megoldása. Legyen  $\alpha^*$  a 4.8 tételben definiált lépéshossz, ekkor*

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1-\theta)\mu_q^t) \leq \Phi_q(x, s, \mu_q^t),$$

ahol

$$\theta = \frac{(\tau-1)^{\frac{q-1}{2q}}}{16q \left( \frac{\tau}{2} + 1 + \frac{\log \tau}{2} \right) n^{\frac{q+1}{2q}}}.$$

**Bizonyítás:** A 4.10 lemmát felhasználva elég egy olyan  $\theta$  értéket választanunk, amelyre teljesül, hogy

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) + \frac{n\theta}{2} + \frac{n\theta \left( 1 - \left( \frac{1}{2\tau} \right)^{\frac{q-1}{2}} \right)}{q-1} \leq (1-\theta)\Phi_q(x, s, \mu_q^t).$$

A 4.8 tétel miatt ehhez elég, ha

$$\theta\Phi_q(x, s, \mu_q^t) + \frac{n\theta}{2} + \frac{n\theta \left( 1 - \left( \frac{1}{2\tau} \right)^{\frac{q-1}{2}} \right)}{q-1} \leq \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}. \quad (4.18)$$

Felhasználva, hogy  $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$ , a következő ekvivalens alakot kapjuk:

$$\theta \left( \frac{(\tau-1)n}{2} + \frac{n}{2} + \frac{n \left( 1 - \left( \frac{1}{2\tau} \right)^{\frac{q-1}{2}} \right)}{q-1} \right) \leq \frac{(\tau-1)^{\frac{q-1}{2q}} n^{\frac{q-1}{2q}}}{16q}.$$

Ebből már következik, hogy a

$$\theta = \frac{(\tau-1)^{\frac{q-1}{2q}}}{16q \left( \frac{\tau}{2} + 1 + \frac{\log \tau}{2} \right) n^{\frac{q+1}{2q}}}$$

választás esetén

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1-\theta)\mu_q^t) \leq \Phi_q(x, s, \mu_q^t).$$

Ezzel a bizonyítást befejeztük.  $\square$

Most már minden a kezünkben van ahhoz, hogy az algoritmus komplexitását megadjuk. A  $\mu_q^t$  választása miatt a  $\Phi_q(x, s, \mu_q^t)$  távolságfüggvény nem változik

az iterációk során. Legyen  $\mu_q^{t+}$  a  $\mu_q^t$  paraméter értéke a következő iterációban, ekkor

$$\Phi_q(x, s, \mu_q^t) = \Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^{t+}).$$

Másrészt a 4.11 tétel miatt

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1 - \theta)\mu_q^t) \leq \Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^{t+}).$$

Mivel a távolságfüggvény a  $\mu$  paraméter konvex függvénye, így

$$\mu_q^{t+} \leq \left(1 - \frac{(\tau - 1)^{\frac{q-1}{2q}} n^{-\frac{q-1}{2q}}}{16q \left(\frac{\tau}{2} + 1 + \frac{\log \tau}{2}\right)}\right) \mu_q^t. \quad (4.19)$$

A következő tételben megadjuk az algoritmus komplexitását.

**4.12. Tétel.** *Legyen  $\tau \geq 2$ . Ekkor a 3 algoritmus legfeljebb*

$$\left\lceil \frac{8q(\tau + 2 + \log \tau) n^{\frac{q+1}{2q}}}{(\tau - 1)^{\frac{q-1}{2q}}} \log \frac{2n\tau^2}{\epsilon} \right\rceil$$

*iteráció után olyan megengedett megoldást talál, amelyre  $x^T s \leq \epsilon$ .*

**Bizonyítás:** A (4.19) egyenlőtlenség alapján legfeljebb

$$\left\lceil \frac{8q(\tau + 2 + \log \tau) n^{\frac{q+1}{2q}}}{(\tau - 1)^{\frac{q-1}{2q}}} \log \frac{2n\tau^2}{\epsilon} \right\rceil$$

iteráció után  $\mu_q^t \leq \frac{\epsilon}{2n\tau^2}$ . A 3.7 lemma és a 4.1 lemma miatt  $\mu_g \leq 2\tau^2 \mu_q^t \leq \frac{\epsilon}{n}$ , vagyis  $x^T s \leq \epsilon$ .  $\square$

A  $q$  paramétert optimálisan választva a következő komplexitást kapjuk.

**4.13. Következmény.** *Ha  $q = \log n$  akkor a 4.12 tétel a következő lépésszámot adja:*

$$\mathcal{O}\left(\sqrt{n} \log n \log \frac{n}{\epsilon}\right).$$

A nagylépéses eljárások között ez a jelenleg ismert legjobb lépésszámkorlát, nemcsak a monoton lineáris komplementaritási, hanem a klasszikus lineáris optimalizálási feladat esetén is.

## 5. Összefoglalás

Ebben a cikkben egy önreguláris függvényekre épülő belsőpontos algoritmus teljes elemzését adtuk meg a monoton lineáris komplementaritási feladatra. Az

algoritmus adaptív lépést használ, komplexitása  $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ , ami megegyezik a [5, 6] munkákban elért komplexitási eredményekkel. Az eredményhez nagyban hozzájárult a

$$\psi(t) = \frac{t^2 - 1}{2} + \frac{t^{1-q} - 1}{q - 1}$$

függvényosztály tulajdonságainak vizsgálata. Ezen tulajdonságok arra engednek következtetni, hogy ha az algoritmus során akár távol, akár közel vagyunk a centrális úthoz, a nagylépéses eljárás jó keresési irányt szolgáltat és a távolságfüggvényt is megfelelően csökkenti.

Érdeemes megjegyezni, hogy az itt leírt adaptív belsőpontos módszer nem használ belső iterációkat. Ez egyáltalán nem szokatlan a gyakorlatban implementált algoritmusoknál, ugyanakkor eltér a [5]-ben közölt sémától. Az algoritmus gyakorlati vizsgálata még hátravan, de kezdeti eredmények arra mutatnak, hogy az önreguláris függvényekből származó keresési irányt használva kevesebb iterációra van szükség, mint a szokásos Newton-lépés esetén.

Az önreguláris függvények további tulajdonságainak felderítése izgalmas kutatási terület. Páratlan  $q \geq 3$  esetén például a következő összefüggés igaz:

$$\Phi_q(x, s, \mu_g) = \Phi_q(x, s, \mu^*) + \frac{\Phi_q(x, s, \mu^*)^2}{\frac{(q+1)n}{2(q-1)}} + \dots + \frac{\Phi_q(x, s, \mu^*)^{\frac{q+1}{2}}}{\left(\frac{(q+1)}{2(q-1)}\right)^{\frac{q+1}{2}} (q-1)n^{\frac{q+1}{2}}}.$$

Nyitott kérdés, hogy páros  $q$  esetén igaz-e hasonló formula. Ezek az összefüggések nagyban leegyszerűsítik a komplexitás elemzését.

További lehetőség hasonló eljárások elemzése általánosabb lineáris komplementaritási feladatok, illetve másodrendű és szemidefinit optimalizálási feladatok esetén. Nem megengedett iterációkat használó (infeasible) és prediktor-korrektor algoritmusok is természetes módon konstruálhatók. Végül az eredmény esetleg kiterjeszhető az önreguláris függvények nagyobb osztályára is, bár ez nehezebb feladatnak tűnik, hiszen a komplexitás elemzése során többször is felhasználtuk, hogy a  $\Phi_q(x, s, \mu)$  függvény a  $\mu$  változó konvex (vagy legalább kvázikonvex) függvénye. A  $\Gamma_{1,q}(t)$  függvényekből kapott  $\psi(1/\sqrt{t})$  függvényekre ez könnyen ellenőrizhetően teljesül, azonban általában nem következik az önregularitásból.

## Hivatkozások

- [1] E. D. ANDERSEN, J. GONDZIO, MÉSZÁROS Cs. és X. XU, Implementation of interior point methods for large scale linear programming. In *Interior Point Methods of Mathematical Programming*, T. Terlaky, Ed. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996, pp. 189–252.
- [2] N. K. KARMAKAR, A new polynomial-time algorithm for linear programming. *Combinatorica* 4 (1984), 373–395.

- [3] NAGY M., *Lineáris optimalizálás: elmélet és primál-duál belsőpontos algoritmusok.* Diplomamunka, ELTE TTK Alkalmazott Matematikus szak, Budapest, 2004. Elérhető a <http://www.cs.elte.hu/alkmatdiploma/nmariann.pdf> címről.
- [4] Y. E. NESTEROV ÉS A. S. NEMIROVSKI, *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Studies in Applied Mathematics, Volume 13. SIAM, Philadelphia, 1994.
- [5] J. PENG, C. ROOS ÉS TERLAKY T., *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Methods.* Princeton University Press, Princeton, NY, 2002.
- [6] J. PENG ÉS TERLAKY T., A dynamic large-update primal-dual interior-point method for linear optimization. *Optimization Methods and Software* 17 (2002), 1077–1104.
- [7] PÓLIK I., *Lineáris optimalizálás belsőpontos módszereinek újszerű vizsgálata.* Diplomamunka, ELTE TTK Matematikus szak, Budapest, 2002. Elérhető a <http://www.cs.elte.hu/matdiploma/polik.pdf> címről.
- [8] J. RENEGAR, *A Mathematical View of Interior-Point Methods in Convex Optimization.* SIAM, Philadelphia, PA, 2001.
- [9] C. ROOS, TERLAKY T. ÉS J.-PH. VIAL, *Theory and Algorithms for Linear Optimization. An Interior Point Approach.* John Wiley & Sons, Chichester, UK, 1997.
- [10] M. SALAHİ ÉS TERLAKY T., An adaptive self-regular proximity based large-update interior point methods for linear optimization. *Optimization Methods and Software* (2003). Megjelenés előtt.
- [11] SONNEVEND GY., An ‘analytic center’ for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In *System modelling and optimization: Proceedings of the 12th IFIP conference held in Budapest, Hungary, September, 1985* (Berlin, West Germany, 1986), Prékopa A., Szelezsán J. and Strazicky B., Eds., vol. 84 of *Lecture notes in Control and Information Sciences*, Springer-Verlag, pp. 866–876.
- [12] TERLAKY T., A belsőpontos módszerek elméletének egy elemi tárgyalása. In *Új utak a magyar operációkutatásban: In memoriam Farkas Gyula*, Komlósi S., Szántai T., Eds. Dialóg Campus Kiadó, Pécs, 1999, pp. 97–132.
- [13] S. J. WRIGHT, *Primal-dual Interior-Point Methods.* SIAM, Philadelphia, PA, 1997.
- [14] Y. YE, *Interior Point Algorithms, Theory and Analysis.* John Wiley & Sons, Chichester, UK, 1997.