Lineáris optimalizálás: elmélet és primál–duál belsőpontos algoritmusok

Diplomamunka

Írta: Nagy Marianna

Alkalmazott matematikus szak

Témavezető:

Illés Tibor, egyetemi docens
Operációkutatási Tanszék
Eötvös Loránd Tudományegyetem, Természettudományi Kar

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2004

Köszönetnyílvánítás

Köszönetet szeretnék mondani elsősorban témavezetőmnek, Illés Tibornak a belsőpontos algoritmusok világának megismertetéséért és a dolgozat elkészítésében nyújtott nélkülözhetetlen segítségéért, számos szakmai és technikai tanácsáért. Továbbá Stoyan Gisbertnek, aki szintén nagyfokú támogatást nyújtott tanulmányaim előremenetelében, lehetőséget teremtve a 17th ECMI Modelling Week rendezvényen való részvételre. Nem utolsó sorban köszönet illeti Brenyó Mihályt, középiskolai matematika tanáromat, aki rendkívül megszerettette velem a matematikát és megmutatta e tudományterület szépségeit.

Tartalomjegyzék

1. fejezet

Bevezetés

A szakdolgozat célja betekintést nyújtani a belsőpontos módszerek világába. A lineáris optimalizálás feladatainak megoldására legelterjedtebb eljárás a szimplex módszer (1947), ám egyre ismertebbé és elfogadottabbá válnak a belsőpontos algoritmusok (1984) is. A két különböző módszercsalád összevetését találhatjuk meg Illés és Terlaky cikkében [7], melyből néhány fontosabb gondolatot most itt is megemlítenénk.

Talán a legszembetűnőbb különbség az algoritmusok geometriájában jelentkezik. Míg a szimplex algoritmus a polinom csúcsait látogatja végig, addig a belsőpontos algoritmusok a megengedett tartomány (relatív) belsejében haladó centrális utat követve konvergálnak a megoldáshoz. Mindkét módszer a gyakorlatban jól alkalmazható, ezt igazolja az is, hogy a főbb programcsomagok részét képezi mindkét algoritmus-család egy-egy implementációja. Ennek ellenére a szimplex módszer esetében léteznek olyan feladatok, melyekre exponenciális a futásideje — a legelső exponenciális ellenpélda Klee és Minty nevéhez fűződik (1970). A belsőpontos algoritmusokkal polinom időben állítunk elő egy jól közelítő megoldást. Megfelelő pontosság esetén ebből a pontból egy erősen polinomiális eljárással nyerünk szigorúan komplementáris egzakt megoldást, sőt létezik olyan polinomiális algoritmus, mely ennek a pontnak az ismeretében optimális bázismegoldást nyújt. Ezzel szemben a szimplex algoritmussal kapott optimális bázismegoldásból szigorúan komplementáris megoldás előállítása degeneráció esetén közel olyan nehéz, mint az eredeti feladat. Minden algoritmus esetén fontos kérdés az induló pont megadása. Ez a szimplex algoritmus esetén az első fázis feladat megoldásával állítható elő. A belsőpontos módszerek esetén két lehetőség áll rendelkezésünkre, vagy mint egy infizíbilis feladatot oldjuk meg a problémát — ekkor belátható, hogy az algoritmus végére az eredeti feltételeinket kielégítő megoldást kapunk— vagy pedig egy beágyazást alkalmazunk, ám ekkor megnő a feladat dimenziója.

Mint láthatjuk mindkét módszernek vannak előnyei és hátrányai is a másikkal szemben. A gyakorlati problémák esetében ezeknek megfelelően kell eldönteni, melyik eljárással oldjuk meg a problémát, például megmutatható, hogy a vágósíkos módszerek esetén a szimplex algoritmus, míg az érzékenységvizsgálat esetén a belsőpontos algoritmusok a hatékonyabbak.

A dolgozat következő fejezetében a szükséges elméletet ismertetjük. Majd a harmadik fejezetben megmutatjuk, hogy létezik polinomiális belsőpontos módszer (Dikin algoritmus), és polinom időben pontos megoldás is előállítható. A negyedik fejezetben a primál–duál feladatpárt visszavezetjük a ferdén szimmetrikus önduális feladatosztályra — melyet a második fejezetben tárgyalunk — a Goldman–Tucker modellen keresztül. Továbbá a lineáris optimalizálás néhány alapvető tételére adunk bizonyítást az ismertetett eredményekre támaszkodva. A dolgozat második felében konkrét, a gyakorlatban is használt algoritmusokat mutatunk be egyfajta fejlődési utat követve. Ezen kívül minden fejezet bevezetésének végén kitekintést adunk, hogy az adott fejezet eredményeinek milyen általánosításai születtek már a belsőpontos algoritmusokra értelmezhető legbővebb, $\mathcal{P}_*$-mátrixokkal definiált lineáris komplementaritási feladatosztály esetén.

2. fejezet

A ferdén szimmetrikus önduális modell

A fejezetben a ferdén szimmetrikus önduális feladattal (SP) foglalkozunk, melyet először Tucker [29] vizsgált 1956-ban. Annak ellenére, hogy a $\mathbf{0}$ vektor megoldása, sőt optimális megoldása a feladatnak, nem érdektelen a további vizsgálódás, ugyanis általában a $\mathbf{0}$ nem az egyetlen optimális vektor, és célunk szigorúan komplementáris megoldás előállítása.

A feladatot iteratív módon oldjuk meg, melyhez szükségünk lesz a Newton–rendszerre illetve a Newton–lépésre, így a fejezetben ezekre is kitérünk. Az iteratív megoldás során a centrális utat követjük, melyet először Sonnevend György definiált [26]. Megmutatjuk, hogy ha létezik belső pont, akkor a centrális út létezik, sőt egyértelmű. Továbbá ekkor a centrális út a μ paraméter csökkentésével egy szigorúan komplementáris megoldáshoz tart (Goldman és Tucker [4] mutatták meg, hogy a feladatnak létezik szigorúan komplemntáris megoldása, a konvergencia kérdésével pedig Megiddo [15] foglalkozott legelőször), sőt a torlódási pontja megegyezik az (SP) feladat úgynevezett analitikus centrumával, amely fogalmat szintén Sonnevend [26] vezette be a szakirodalomba.

Végül az indexhalmaz optimális partíciójáról ejtünk pár szót. Ennek segítségével becslést adhatunk a centrális út mentén a változók méretére, aminek segítségével pontos megoldást lehet előállítani. Erre a következő fejezetben térünk ki.

Mint látni fogjuk a ferdén szimmetrikus feladatnak megfeleltethető egy lineáris komplementaritási feladat (LCP). Ennek segítségével definiálhatjuk a ferdén szimmetrikus feladat általánosításait, amiről akkor beszélhetünk, ha a feladat mátrixa nem ferdén szimmetrikus, hanem valamely tágabb mátrixosztály eleme. Kojimáék [12] vezették be a P_* mátrixosztályt, amely a legbővebb olyan feladatosztályhoz vezet, melyre a fejezet eredményei érvényben maradnak (lásd Pólik Imre szakdolgozata [21]).

2.1. A ferdén szimmetrikus önduális feladat alaptulajdonságai

Tekintsük a következő (speciális) lineáris programozási feladatot

$$\left.\begin{array}{lcr} \min & \mathbf{q}^T\mathbf{x} & \\ M\mathbf{x} & \geq & -\mathbf{q} \\ \mathbf{x} & \geq & \mathbf{0} \end{array}\right\} \quad (SP),$$

ahol $M \in \mathbb{R}^{n\times n}$ ferdén szimmetrikus mátrix ($M^T = -M$) és $\mathbf{q} \in \mathbb{R}^n_\oplus$ vektor. Az (SP) feladat megengedett megoldásainak halmazát

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{x} \geq \mathbf{0},\ M\mathbf{x} \geq -\mathbf{q}\ \}$$

jelöli. Felhasználva, hogy az M mátrix ferdén szimmetrikus és, hogy a jobb oldalon álló vektor $(-\mathbf{q})$ a célfüggvényvektor negatívja, az olvasó könnyen ellenőrizheti, hogy (SP) duálja ekvivalens az (SP) feladattal, azaz az (SP) *önduális* feladat. Az önduális tulajdonság miatt a következő eredmény triviális.

2.1.1. Lemma. *Az (SP) feladat optimális értéke nulla, továbbá az azonosan nulla vektor, $\mathbf{x} = \mathbf{0}$ megengedett és egyben optimális megoldása az (SP) feladatnak.* □

Ha $(\mathbf{x}, s(\mathbf{x}))$ adott megoldása az (SP) feladfatnak, ahol $s(\mathbf{x}) = M\mathbf{x} + \mathbf{q}$, akkor

$$\mathbf{q}^T\mathbf{x} = \mathbf{x}^T(s(\mathbf{x}) - M\mathbf{x}) = \mathbf{x}^T s(\mathbf{x}) = \mathbf{e}^T(\mathbf{x}\, s(\mathbf{x})),$$

azaz tetszőleges optimális megoldásra $\mathbf{e}^T(\mathbf{x}\, s(\mathbf{x})) = 0$, azaz $\mathbf{x}s(\mathbf{x}) = 0$, amiből az is következik, hogy az $\mathbf{x}$ és $s(\mathbf{x})$ vektorok komplementárisak. Jelölje

$$\begin{aligned} \mathcal{F}^* \ &:= \ \{\mathbf{x}^* \in \mathcal{F} \ : \ \mathbf{q}^T\mathbf{x}^* \leq \mathbf{q}^T\mathbf{x},\ \forall \mathbf{x} \in \mathcal{F}\} \\ &= \ \{\mathbf{x}^* \in \mathcal{F} \ : \ \mathbf{q}^T\mathbf{x}^* = 0\} = \{\mathbf{x}^* \in \mathcal{F} \ : \ \mathbf{x}^*\, s(\mathbf{x}^*) = \mathbf{0}\}, \end{aligned}$$

az (SP) *feladat optimális megoldásainak a halmazát*.

Az $\mathbf{x} \in \mathcal{F}$ esetén az $\mathbf{x}^T s(\mathbf{x}) = \mathbf{q}^T \mathbf{x}$ értéket *dualitás rés*nek nevezzük.[1] Az $\mathcal{F}^*$ halmaz definíciója alapján világos, hogy optimális megoldás esetén a dualitás rés nulla.

Az optimális megoldások egy, a továbbiakban gyakran használt tulajdonságát fogalmazzuk meg az alábbi lemmában.

2.1.2. Lemma. *Legyen* $\mathbf{x}$ *és* $\mathbf{y}$ *az* (SP) *feladat megengedett megoldása. Az* $\mathbf{x}$ *és* $\mathbf{y}$ *vektorok akkor és csak akkor optimálisak, ha*

$$\mathbf{x}\, s(\mathbf{y}) = \mathbf{y}\, s(\mathbf{x}) = \mathbf{x}\, s(\mathbf{x}) = \mathbf{y}\, s(\mathbf{y}) = \mathbf{0}.$$

Bizonyítás. Mivel M ferdén szimmetrikus, így $(\mathbf{x}-\mathbf{y})^T M(\mathbf{x}-\mathbf{y}) = 0$, amiből következik, hogy $(\mathbf{x}-\mathbf{y})^T(s(\mathbf{x}) - s(\mathbf{y})) = 0$. Ekkor $\mathbf{x}^T s(\mathbf{y}) + \mathbf{y}^T s(\mathbf{x}) = \mathbf{x}^T s(\mathbf{x}) + \mathbf{y}^T s(\mathbf{y})$ és ez akkor és csak akkor nulla, ha $\mathbf{x}$ és $\mathbf{y}$ is optimális, de ekkor

$$\mathbf{x}^T s(\mathbf{y}) + \mathbf{y}^T s(\mathbf{x}) = 0. \tag{2.1}$$

Figyelembe véve az $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ feltételt $\mathbf{x}^T s(\mathbf{y}) \geq 0$ és $\mathbf{y}^T s(\mathbf{x}) \geq 0$ teljesül, amelyből, az (2.1) összefüggés alapján

$$\mathbf{x}^T s(\mathbf{y}) = \mathbf{e}^T(\mathbf{x}\, s(\mathbf{y})) = 0 \qquad \text{és} \qquad \mathbf{y}^T s(\mathbf{x}) = \mathbf{e}^T(\mathbf{y}\, s(\mathbf{x})) = 0$$

következik. Tehát $\mathbf{x}\, s(\mathbf{y}) = \mathbf{y}\, s(\mathbf{x}) = \mathbf{0}$ adódik. □

Megállapíthatjuk, hogy az optimális megoldások általános értelemben is komplementárisak, azaz nem csak saját eltérés vektorukkal, hanem bármelyik más optimális megoldás eltérés vektorával is, komplementáris párt alkotnak.

Az (SP) feladat optimalitási kritériumát a következő alakban is megadhatjuk

$$\left.\begin{aligned} -M\mathbf{x} + \ \mathbf{s} &= \mathbf{q} \\ \mathbf{x} \geq \mathbf{0}, \quad \mathbf{s} &\geq \mathbf{0} \\ \mathbf{x}\,\mathbf{s} &= \mathbf{0} \end{aligned}\right\} \quad (LCP_{SP}),$$

amelyet az (SP) feladathoz tartozó *lineáris komplementaritási feladat*nak nevezünk.

Az összes eddigi eredmény, egy triviális optimális megoldás létezését is beleértve, szinte magától érthető volt az (SP) feladatra. Ebből talán arra következtethetnénk,

[1]Tesszük ezt annak ellenére, hogy az önduális feladat esetén az $\mathbf{x}^T s(\mathbf{x}) = \mathbf{q}^T \mathbf{x}$ nyilván csak a fele a klasszikus értelemben vett dualitás résnek.

hogy az (SP) feladat érdektelen, nagyon speciális esete a lineáris programozási feladatoknak.

Mielőtt megmutatnánk, hogy bármely primál és duál lineáris programozási feladatpár egy velük ekvivalens önduális feladattá transzformálható, építsük fel az önduális lineáris programozási feladatok elméletét. Az első felmerülő kérdés az, hogy létezik-e a triviálison kívül másmilyen optimális megoldása a feladatnak. Ennek a kérdésnek a megválaszolása elvezet az optimális megoldások bizonyos komplementaritási tulajdonságainak kérdéséhez, azaz a lineáris programozás elméletének egyik alapvető jelentőségű tételéhez, a Goldman–Tucker–tételhez. Szükségünk lesz a következő definícióra.

2.1.3. Definíció. *Legyen* $\mathbf{x} \in \mathcal{F}^*$*. Az* $\mathbf{x}$ *és* $s(\mathbf{x})$ *vektorokat* szigorúan komplementárisnak *nevezzük, ha* $\mathbf{x} + s(\mathbf{x}) > \mathbf{0}$ *feltétel teljesül.*

Az $\mathbf{x}$ és $s(\mathbf{x})$ vektorok szigorúan komplementaritásának az egyszerű következménye az, hogy bármely $i = 1, 2, \ldots, n$ index esetén az $x_i = 0$ és $s_i(x) = 0$ feltételek közül pontosan az egyik teljesül.

Vezessük be a *megengedett belsőpontos megoldások halmazát*

$$\mathcal{F}^0 := \{\mathbf{x} \in \mathcal{F} \;:\; (\mathbf{x},\, s(\mathbf{x})) > \mathbf{0}\}.$$

Ekkor az un. *belső pont (BP) feltétel*t az alábbi módon fogalmazhatjuk meg

$$\mathcal{F}^0 \neq \emptyset,$$

amelyet másként úgy is kimondhatunk, hogy létezik olyan $\bar{\mathbf{x}} \in \mathcal{F}$ vektor, amelyre

$$(\bar{\mathbf{x}},\ s(\bar{\mathbf{x}})) > \mathbf{0}$$

teljesül.

2.2. A Newton–lépés és tulajdonságai

Legyen adott $(\mathbf{x}, \mathbf{s}) > \mathbf{0}$, melyre $\mathbf{s} = M\mathbf{x} + \mathbf{q}$. Célunk, hogy adott $\mathbf{w} > \mathbf{0}$ vektorra meghatározzuk a $(\triangle\mathbf{x}, \triangle\mathbf{s})$ lépést úgy, hogy

$$\begin{aligned} M(\mathbf{x} + \triangle\mathbf{x}) + \mathbf{q} &= \mathbf{s} + \triangle\mathbf{s} \\ (\mathbf{x} + \triangle\mathbf{x})(\mathbf{s} + \triangle\mathbf{s}) &= \mathbf{w} \end{aligned}$$

teljesüljön. Sőt, azt is szeretnénk biztosítani, hogy

$$\mathbf{x} + \triangle\mathbf{x} > \mathbf{0} \quad \text{és} \quad \mathbf{s} + \triangle\mathbf{s} > \mathbf{0},$$

teljesüljön. Ekkor a ($\Delta\mathbf{x}$, $\Delta\mathbf{s}$) ismeretlenekre vonatkozóan a következő egyenletrendszert kapjuk

$$\begin{aligned} M\,\Delta\mathbf{x} - \Delta\mathbf{s} &= \mathbf{0} \\ \mathbf{x}\,\Delta\mathbf{s} + \mathbf{s}\,\Delta\mathbf{x} + \Delta\mathbf{x}\,\Delta\mathbf{s} &= \mathbf{w} - \mathbf{x}\,\mathbf{s}\ , \end{aligned}$$

amely még mindig nemlineáris. A másodrendű $\Delta\mathbf{x}\,\Delta\mathbf{s}$ tag elhagyásával a

$$\begin{aligned} M\,\Delta\mathbf{x} - \quad \Delta\mathbf{s} &= \mathbf{0} \\ S\,\Delta\mathbf{x} + X\,\Delta\mathbf{s} &= \mathbf{w} - \mathbf{x}\,\mathbf{s} \end{aligned}$$

Newton egyenletrendszer adódik, ahol $X = diag(\mathbf{x})$ és $S = diag(\mathbf{s})$ pozitív diagonális mátrixok. Ez már egy lineáris egyenletrendszer, amelynek a mátrixa

$$\bar{M} = \begin{pmatrix} M & -I \\ S & X \end{pmatrix} \in \mathbb{R}^{2n\times 2n} \tag{2.2}$$

alakú, míg az egyenlet, a $\Delta\mathbf{s} = M\,\Delta\mathbf{x}$ behelyettesítése után

$$(S + X\,M)\,\Delta\mathbf{x} = \mathbf{w} - \mathbf{x}\,\mathbf{s} \tag{2.3}$$

lesz.

2.2.1. Állítás. *Legyenek az X, S, I, $M \in \mathbb{R}^{n\times n}$ mátrixok. Az X és S pozitív diagonálisak, az I egység, míg az M ferdén szimmetrikus mátrix. Ekkor a (2.2) összefüggéssel definiált $\bar{M}$ mátrix reguláris, tehát a (2.3) lineáris egyenletrendszer megoldása létezik és egyértelmű.*

Bizonyítás. Indirekt, azaz tegyük fel, hogy az $\bar{M}$ mátrix szinguláris. Ez azt jelenti, hogy létezik $\mathbf{z} \in \mathbb{R}^{2n}\ :\ \bar{M}\,\mathbf{z} = \mathbf{0}$, de $\mathbf{z} \neq \mathbf{0}$.

Legyen $\mathbf{z} = (\mathbf{u},\, \mathbf{v})$ alakú, és ekkor az $\bar{M}\,\mathbf{z} = \mathbf{0}$ egyenletet

$$\begin{aligned} M\,\mathbf{u} - I\,\mathbf{v} &= \mathbf{0} \\ S\,\mathbf{u} + X\,\mathbf{v} &= \mathbf{0} \end{aligned}$$

alakban írhatjuk. Ekkor nyilván $\mathbf{v} = M\,\mathbf{u}$ lesz.

Mivel a $\mathbf{z} \neq \mathbf{0}$ vektor, ezért nyilván az $\mathbf{u} \neq \mathbf{0}$ összefüggés is teljesül, vagyis létezik olyan $1 \leq i \leq n$ index, amelyre $u_i \neq 0$. Figyelembe véve, az S és X mátrixok diagonalitását

$$s_i\, u_i + x_i\, v_i = 0$$

adódik, bármely i indexre. Mivel az $x_i > 0$ és az $s_i > 0$, ezért, ha $u_i = 0$ akkor $v_i = 0$ teljesül, viszont, ha $u_i \neq 0$ akkor

$$-\frac{v_i}{u_i} = \frac{s_i}{x_i} > 0$$

áll fenn, azaz az u_i és v_i előjele különböző. Az M mátrix ferdén szimmetrikussága miatt

$$0 = \mathbf{u}^T M\, \mathbf{u} = \mathbf{u}^T \mathbf{v} = \sum_{i=1}^{n} u_i\, v_i < 0,$$

ellentmondás adódik, tehát az $\bar{M}$ mátrix reguláris.

Ekkor a (2.3) egyenletnek is egyértelmű megoldása van, azaz az $S + X\, M$ mátrix is invertálható. □

A $\Delta\mathbf{x}$ és $\Delta\mathbf{s}$ *Newton–irány*ok az alábbi formában adhatók meg

$$\Delta\mathbf{x} \;=\; (S + X\, M)^{-1}\, (\mathbf{w} - \mathbf{x}\,\mathbf{s}) \quad \text{és} \quad \Delta\mathbf{s} \;=\; M\, (S + X\, M)^{-1}\, (\mathbf{w} - \mathbf{x}\,\mathbf{s}). \tag{2.4}$$

Miután a Newton irányban egy α lépéshosszú lépést teszünk, az új $(\mathbf{x}^+,\, \mathbf{s}^+)$ megoldásra a következő kifejezést kapjuk:

$$\begin{aligned} \mathbf{x}^+\mathbf{s}^+ \;&:=\; (\mathbf{x} + \alpha\Delta\mathbf{x})(\mathbf{s} + \alpha\,\Delta\mathbf{s}) = \mathbf{x}\,\mathbf{s} + \alpha\,(\mathbf{x}\Delta\mathbf{s} + \mathbf{s}\Delta\mathbf{x}) + \alpha^2\,\Delta\mathbf{x}\,\Delta\mathbf{s} \\ &=\; \mathbf{x}\,\mathbf{s} + \alpha\,(\mathbf{w} - \mathbf{x}\,\mathbf{s}) + \alpha^2\,\Delta\mathbf{x}\,\Delta\mathbf{s}. \end{aligned}$$

Ez az összefüggés világossá teszi, hogy az $\mathbf{x}\,\mathbf{s}$ vektor lokális megváltozását a $\mathbf{w} - \mathbf{x}\,\mathbf{s}$ vektor határozza meg. Szerencsére ezt a vektort explicit ismerjük, amikor a Newton–lépést alkalmazzuk. Így amikor α elegendően kicsi, akkor pontosan tudjuk, hogy $\mathbf{x}\,\mathbf{s}$ mely koordinátái csökkennek lokálisan (pontosan azok, amelyekre a $\mathbf{w} - \mathbf{x}\,\mathbf{s}$ vektor megfelelő koordinátái negatívak) és $\mathbf{x}\,\mathbf{s}$ mely koordinátái növekednek lokálisan (pontosan azok, amelyekre a $\mathbf{w} - \mathbf{x}\,\mathbf{s}$ vektor megfelelő koordinátái pozitívak).

Ezeket az észrevételeket fejezi ki pontosabban a következő állítás is.

2.2.2. Állítás. *Legyen az M ferdén szimmetrikus mátrix és $\mathbf{x} \in \mathcal{F}^0$, azaz $(\mathbf{x},\, s(\mathbf{x})) > \mathbf{0}$. Legyen továbbá, $\hat{\mathbf{w}} \in \mathbb{R}^n$, $\hat{\mathbf{w}} > \mathbf{0}$ és $\mathbf{w} = \mathbf{x}\, s(\mathbf{x})$. Definiáljuk a*

$$\mathcal{T}(\hat{\mathbf{w}},\, \mathbf{w}) = \{\, \mathbf{u} \in \mathbb{R}^n \;:\; \hat{w}_i \leq u_i \leq w_i \;\; \text{vagy} \;\; w_i \leq u_i \leq \hat{w}_i,\; \forall\, i \,\}$$

tégla halmazt. Ha $int\,\mathcal{T}(\hat{\mathbf{w}},\,\mathbf{w}) \neq \emptyset$ *akkor létezik* $\alpha \in (0,1)$ *és* $(\Delta\mathbf{x},\,\Delta\mathbf{s})$ *úgy, hogy az*

$$\mathbf{x}^+ := \mathbf{x} + \alpha\Delta\mathbf{x}, \quad \mathbf{s}^+ := \mathbf{s} + \alpha\,\Delta\mathbf{s}, \quad \mathbf{w}^+ := \mathbf{x}^+\mathbf{s}^+$$

és $\mathbf{w}^+ \in int\,\mathcal{T}(\hat{\mathbf{w}},\,\mathbf{w})$*, valamint* $\mathbf{x}^+ \in \mathcal{F}^0$ *teljesül.*

Bizonyítás. Az $int\,\mathcal{T}(\hat{\mathbf{w}},\,\mathbf{w}) \neq \emptyset$ feltétel miatt $\hat{\mathbf{w}} - \mathbf{w} \neq \mathbf{0}$ és így a (2.4) képletek miatt, a $\Delta\mathbf{x}$ és $\Delta\mathbf{s}$ vektoroknak nem lehet minden komponense nulla.

Az $\alpha > 0$ számot úgy kell meghatározni, hogy

$$\mathbf{x}^+ := \mathbf{x} + \alpha\Delta\mathbf{x} > \mathbf{0} \quad \text{és} \quad \mathbf{s}^+ := \mathbf{s} + \alpha\,\Delta\mathbf{s} > \mathbf{0} \tag{2.5}$$

teljesüljön, hiszen ekkor az $\mathbf{x}^+ \in \mathcal{F}^0$ fenáll, továbbá azt szeretnénk, hogy legyen $\mathbf{w}^+ \in int\,\mathcal{T}(\hat{\mathbf{w}},\,\mathbf{w})$, ehhez pedig az kell, hogy bármely $1 \leq i \leq n$ index esetén

$$\min\{\hat{w}_i,\,w_i\} \;<\; w_i^+ = x_i^+ s_i^+ \;<\; \max\{\hat{w}_i,\,w_i\} \tag{2.6}$$

egyenlőtlenségek teljesüljenek.

Világos, hogy az (2.5) egyenlőtlenségek teljesülése a $\Delta\mathbf{x}$ és $\Delta\mathbf{s}$ vektorok koordinátáinak az előjelétől függ.

Ha $\triangle x_i \geq 0$ akkor bármely $\alpha > 0$ esetén $x_i^+ = x_i + \alpha\,\triangle x_i > 0$ adódik. Ezzel szemben, ha $\triangle x_i < 0$ akkor

$$\alpha \leq \alpha_1 := \min\left\{-\frac{x_i}{\triangle x_i} \;:\; \triangle x_i < 0\right\}.$$

Hasonlóan

$$\alpha \leq \alpha_2 := \min\left\{-\frac{s_i}{\triangle s_i} \;:\; \triangle s_i < 0\right\}.$$

A (2.6) egyenlőtlenségek teljesülésének az elemzésekor szükségünk lesz a következő két index halmazra

$$\mathcal{I}_{\mathbf{w}} := \{i \,:\, w_i < \hat{w}_i\} \quad \text{és} \quad \mathcal{I}_{\hat{\mathbf{w}}} := \{i \,:\, \hat{w}_i < w_i\}.$$

Amikor az $i \in \mathcal{I}_{\mathbf{w}}$ akkor azt szeretnénk, hogy $w_i < w_i^+ < \hat{w}_i$ teljesüljön, azaz

$$w_i < w_i^+ = x_i^+\,s_i^+ = (x_i + \alpha\,\triangle x_i)\,(s_i + \alpha\,\triangle s_i) = w_i + \alpha\,(\hat{w}_i - w_i) + \alpha^2\,\triangle x_i\,\triangle s_i.$$

Figyelembe véve az $\alpha > 0$ összefüggést, a következő korlátot kapjuk

$$\alpha \leq \alpha_3 := \min_{i \in \mathcal{I}_{\mathbf{w}}}\left\{-\frac{\hat{w}_i - w_i}{\triangle x_i\,\triangle s_i} \;:\; \triangle x_i\,\triangle s_i < 0\right\}.$$

Amennyiben a $w_i^+ < \hat{w}_i$ egyenlőtlenséget vizsgáljuk akkor

$$\hat{w}_i > w_i^+ = x_i^+ \, s_i^+ = (x_i + \alpha \triangle x_i)\,(s_i + \alpha \triangle s_i) = w_i + \alpha\,(\hat{w}_i - w_i) + \alpha^2 \triangle x_i \triangle s_i,$$

adódik és ezt tovább alakítva a

$$0 < (\hat{w}_i - w_i) - \alpha\,(\hat{w}_i - w_i) - \alpha^2 \triangle x_i \triangle s_i$$

egyenlőtlenséget kapjuk, amely az α ismeretlenben másodfokú. Ha a $\triangle x_i \triangle s_i \leq 0$ akkor bármely $\alpha \in (0, 1)$ esetén, az előző egyenlőtlenség teljesül. Az ellenkező esetben, amikor $\triangle x_i \triangle s_i > 0$ akkor

$$\alpha \leq \alpha_4 := \min_{i \in \mathcal{I}_{\mathbf{w}}} \left\{ -\frac{(\hat{w}_i - w_i) - \sqrt{(\hat{w}_i - w_i)^2 + 4\,(\hat{w}_i - w_i) \triangle x_i \triangle s_i}}{2 \triangle x_i \triangle s_i} \; : \; \triangle x_i \triangle s_i > 0 \right\}.$$

Hasonlóan elemezhető ki az $i \in \mathcal{I}_{\hat{\mathbf{w}}}$ eset is. Ekkor az

$$\alpha \leq \alpha_5 := \min_{i \in \mathcal{I}_{\hat{\mathbf{w}}}} \left\{ \frac{(w_i - \hat{w}_i) - \sqrt{(w_i - \hat{w}_i)^2 - 4\,(w_i - \hat{w}_i) \triangle x_i \triangle s_i}}{2 \triangle x_i \triangle s_i} \; : \; \triangle x_i \triangle s_i < 0 \right\}$$

illetve

$$\alpha \leq \alpha_6 := \min_{i \in \mathcal{I}_{\hat{\mathbf{w}}}} \left\{ -\frac{\hat{w}_i - w_i}{\triangle x_i \triangle s_i} \; : \; \triangle x_i \triangle s_i > 0 \right\}.$$

Legyen az $\alpha^* := \min\{1, \alpha_1, \alpha_2, \ldots, \alpha_6\}$ és ekkor bármely $\alpha \in (0, \alpha^*)$ lépéshossz esetén

$$\mathbf{x}^+ \in \mathcal{F}^0 \quad \text{és} \quad \mathbf{w}^+ \in int\, \mathcal{T}(\hat{\mathbf{w}}, \, \mathbf{w})$$

teljesül. $\square$

2.3. Az (SP) feladat nívóhalmazai és tulajdonságaik

Vezessük be először a következő *szinthalmazt*

$$\mathcal{L}_K = \{\mathbf{x} \in \mathcal{F} \; : \; \mathbf{x}^T s(\mathbf{x}) \leq K\} = \{\mathbf{x} \in \mathcal{F} \; : \; \mathbf{q}^T \mathbf{x} \leq K\}$$

és bármely $\mathbf{w} \in \mathbb{R}^n_+$ vektor esetén az alábbi *általánosított szinthalmazt*

$$\mathcal{L}_{\mathbf{w}} = \{(\mathbf{x}, \, \mathbf{s}) \in \mathbb{R}^{2n}_{\oplus} \; : \; \mathbf{s} = M\mathbf{x} + \mathbf{q} \;\; \text{és} \;\; \mathbf{x}\,\mathbf{s} \in (\mathbf{w} - \mathbb{R}^n_{\oplus})\}$$

is. Nyilvánvaló, hogy ha az $(\mathbf{x}, \mathbf{s}) \in \mathcal{L}_{\mathbf{w}}$ teljesül, akkor $\mathbf{x}\mathbf{s} \leq \mathbf{w}$.

2.3.1. Lemma. *Legyen adott az* (SP) *feladat és tegyük fel, hogy* $\mathcal{F}^0 \neq \emptyset$. *Ekkor bármely* $K \in \mathbb{R}$, $K > 0$ *szám esetén az* $\mathcal{L}_K$ *szinthalmaz korlátos és zárt, azaz kompakt.*

Bizonyítás. Két eset lehetséges: $\mathcal{L}_K = \emptyset$[2] illetve $\mathcal{L}_K \neq \emptyset$. Az első eset triviális.
A $\mathcal{F}^0 \neq \emptyset$ azt jelenti, hogy létezik egy $\bar{\mathbf{x}} \in \mathcal{F} : (\bar{\mathbf{x}}, \bar{\mathbf{s}}(x)) > \mathbf{0}$ belső pont. Mivel az M mátrix ferdén szimmetrikus, így

$$0 = (\mathbf{x} - \bar{\mathbf{x}})^T M (\mathbf{x} - \bar{\mathbf{x}}) = (\mathbf{x} - \bar{\mathbf{x}})^T (\mathbf{s} - \bar{\mathbf{s}}) = \mathbf{x}^T\mathbf{s} + \bar{\mathbf{x}}^T\bar{\mathbf{s}} - \mathbf{x}^T\bar{\mathbf{s}} - \mathbf{s}^T\bar{\mathbf{x}},$$

ahol az $\mathbf{x} \in \mathcal{L}_K \subseteq \mathcal{F}$ tetszőleges vektor és az $\mathbf{s}$ a hozzátartozó eltérés vektor. Ebből, az $\mathcal{L}_K$ szinthalmaz definíciója alapján azt kapjuk, hogy

$$x_j \bar{s}_j \leq \mathbf{x}^T\bar{\mathbf{s}} + \mathbf{s}^T\bar{\mathbf{x}} = \mathbf{x}^T\mathbf{s} + \bar{\mathbf{x}}^T\bar{\mathbf{s}} \leq K + \bar{\mathbf{x}}^T\bar{\mathbf{s}}, \quad j = 1, \ldots, n$$

és így az $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) > \mathbf{0}$ belsőpont feltétel miatt, bármely $1 \leq j \leq n$ indexre

$$x_j \leq \frac{K + \bar{\mathbf{x}}^T\bar{\mathbf{s}}}{\bar{s}_j},$$

hasonlóan

$$s_j \leq \frac{K + \bar{\mathbf{x}}^T\bar{\mathbf{s}}}{\bar{x}_j},$$

azaz az $\mathcal{L}_K$ szinthalmaz korlátos. A zártsága egyszerű következménye a szinthalmaz második felírásának, azaz annak, hogy véges sok zárt féltér metszete. □

Az előzőnek megfelelő lemma igazolható az $\mathcal{L}_{\mathbf{w}}$ szinthalmazokra is. A szinthalmaz definíciójában szereplő nemlineáris kifejezés miatt kicsit bonyolultabb a zártság bizonyítása.

2.3.2. Lemma. *Legyen adott az* (SP) *feladat és tegyük fel, hogy* $\mathcal{F}^0 \neq \emptyset$. *Ekkor bármely* $\mathbf{w} \in \mathbb{R}^n_+$, *vektor esetén az* $\mathcal{L}_{\mathbf{w}}$ *szinthalmaz korlátos és zárt, azaz kompakt.*

Bizonyítás. Ha $\mathcal{L}_{\mathbf{w}} = \emptyset$ akkor igaz az állítás. (Később megmutatjuk, hogy a $\mathcal{F}^0 \neq \emptyset$ feltétel miatt ez egyetlen egy $\mathbf{w} > \mathbf{0}$ esetén sem fordulhat elő.)
Legyen $(\mathbf{x}, \mathbf{s}) \in \mathcal{L}_{\mathbf{w}}$, azaz $\mathbf{x} \in \mathcal{F}$ és $\mathbf{x}\,\mathbf{s} \leq \mathbf{w}$. Mivel M ferdén szimmetrikus, ezért

$$(\mathbf{x} - \bar{\mathbf{x}})^T (\mathbf{s} - \bar{\mathbf{s}}) = 0,$$

[2] Erről az esetről később megmutatjuk, hogy az $\mathcal{F}^0 \neq \emptyset$ feltétel mellett nem állhat fenn.

ahol $\bar{\mathbf{x}} \in \mathcal{F}^0$ és ekkor

$$\mathbf{x}^T\bar{\mathbf{s}} + \mathbf{s}^T\bar{\mathbf{x}} = \mathbf{x}^T\mathbf{s} + \bar{\mathbf{x}}^T\bar{\mathbf{s}},$$

azaz

$$x_j\,\bar{s}_j \le \mathbf{x}^T\bar{\mathbf{s}} \le \mathbf{x}^T\mathbf{s} + \bar{\mathbf{x}}^T\bar{\mathbf{s}} = \mathbf{e}^T(\mathbf{x}\,\mathbf{s}) + \mathbf{e}^T(\bar{\mathbf{x}}\,\bar{\mathbf{s}}) \le \mathbf{e}^T(\mathbf{w} + \bar{\mathbf{w}}),$$

ahol $\bar{\mathbf{w}} = \bar{\mathbf{x}}\,\bar{\mathbf{s}}$. Ekkor bármely $1 \le j \le n$ index esetén

$$x_j \le \frac{\mathbf{e}^T(\mathbf{w} + \bar{\mathbf{w}})}{\bar{s}_j} \qquad \text{illetve} \qquad s_j \le \frac{\mathbf{e}^T(\mathbf{w} + \bar{\mathbf{w}})}{\bar{x}_j}$$

teljesül, tehát az $\mathcal{L}_\mathbf{w}$ szinthalmaz korlátos.

Legyen

$$\hat{\mathcal{F}} := \{(\mathbf{x},\,\mathbf{s}) \in \mathbb{R}^{2\,n} \,:\, \mathbf{x} \in \mathcal{F},\, \mathbf{s} = M\,\mathbf{x} + \mathbf{q}\} \subset \mathbb{R}^{2\,n}$$

korlátos, zárt halmaz és az $f : \mathbb{R}^{2\,n} \to \mathbb{R}^n$, $f(\mathbf{x},\,\mathbf{s}) = \mathbf{x}\,\mathbf{s}$ folytonos leképezés. Ekkor az

$$f(\hat{\mathcal{F}}) \subseteq \mathbb{R}^n_\oplus \quad \text{zárt halmaz és így az} \quad f(\hat{\mathcal{F}}) \cap (\mathbf{w} - \mathbb{R}^n_\oplus) \quad \text{is zárt,}$$

tehát az őseképe $\mathcal{L}_\mathbf{w}$ halmaz is zárt. $\square$

2.3.3. Lemma. *Legyen adott az (SP) feladat és tegyük fel, hogy $\mathcal{F}^0 \neq \emptyset$. Definiáljuk a következő halmazt*

$$\mathcal{G} := \{\mathbf{w} \in \mathbb{R}^n_\oplus \,:\, \exists(\mathbf{x},\,\mathbf{s}) \in \hat{\mathcal{F}} \;\; \textit{amelyre} \;\; \mathbf{x}\,\mathbf{s} = \mathbf{w}\}.$$

Ekkor a $\mathcal{G}$ halmaz nem üres és zárt.

Bizonyítás. Mivel $\mathcal{F}^0 \neq \emptyset$, ezért létezik $(\bar{\mathbf{x}},\,\bar{\mathbf{s}}) > \mathbf{0}$ és $\bar{\mathbf{x}} \in \mathcal{F}^0$, tehát $\bar{\mathbf{x}}\,\bar{\mathbf{s}} = \bar{\mathbf{w}} > \mathbf{0}$, és így $\bar{\mathbf{w}} \in \mathcal{G}$. Azt kell még belátni, hogy a $\mathcal{G}$ halmaz tartalmazza az összes limeszpontját is. Legyen $\hat{\mathbf{w}} \ge \mathbf{0}$, amelyhez létezik $\mathbf{w}^k \in \mathcal{G}$ sorozat, úgy, hogy $\lim_{k\to\infty} \mathbf{w}^k = \hat{\mathbf{w}}$. Legyen továbbá $\tilde{\mathbf{w}}$ olyan, hogy $\tilde{\mathbf{w}} > \mathbf{w}^k$ teljesül, bármely k indexre és $\tilde{\mathbf{w}} > \hat{\mathbf{w}}$ is igaz. Az előző lemma és az $\mathcal{L}_{\tilde{\mathbf{w}}}$ szinthalmaz konstrukciója alapján, $\mathcal{L}_{\tilde{\mathbf{w}}}$ nem üres és kompkat halmaz. Továbbá létezik $(\mathbf{x}^k,\,\mathbf{s}^k) \in \mathcal{L}_{\tilde{\mathbf{w}}}$, amelyekre $\mathbf{x}^k\,\mathbf{s}^k = \mathbf{w}^k$. Az $\mathcal{L}_{\tilde{\mathbf{w}}}$ halmaz kompaktsága miatt, az általánosság korlátozása nélkül feltehető, hogy $(\mathbf{x}^k, \mathbf{s}^k)$ konvergens pontsorozat, és $\lim_{k\to\infty}(\mathbf{x}^k,\,\mathbf{s}^k) = (\hat{\mathbf{x}},\,\hat{\mathbf{s}})$.

Figyelembe véve a szorzás folytonosságát és a pontsorozatok definícióját kapjuk, hogy

$$\hat{\mathbf{x}}\,\hat{\mathbf{s}} = \lim_{k\to\infty} \mathbf{x}^k\,\mathbf{s}^k = \lim_{k\to\infty} \mathbf{w}^k = \hat{\mathbf{w}},$$

azaz a $\mathcal{G}$ halmaz zárt, mert tartalmazza a limeszpontjait. $\square$

Az előző lemma segítségével belátjuk, hogy belső pont létezése mellett tetszőleges $\mathcal{L}_\mathbf{w}$ szinthalmaz nem üres.

2.3.4. Tétel. *Legyen adott az (SP) feladat és tegyük fel, hogy $\mathcal{F}^0 \neq \emptyset$. Ekkor bármely $\mathbf{w} \in \mathbb{R}^n_+$, vektor esetén az $\mathcal{L}_\mathbf{w}$ szinthalmaz nem üres.*

Bizonyítás. Az 2.3.2 Lemmában már igazoltuk, hogy az $\mathcal{L}_\mathbf{w}$ halmaz kompkat. Indirekt módon tegyük fel, hogy létezik egy $\hat{\mathbf{w}} \in \mathbb{R}^n_+$ vektor, amelyre $\mathcal{L}_{\hat{\mathbf{w}}} = \emptyset$.

Az $\mathcal{F}^0 \neq \emptyset$ feltevés miatt létezik egy $\bar{\mathbf{x}} \in \mathcal{F}^0$ vektor és egy $\bar{\mathbf{w}} = \bar{\mathbf{x}}\,\bar{\mathbf{s}} > \mathbf{0}$, ahol $\bar{\mathbf{s}} = M\,\bar{\mathbf{x}} + \mathbf{q} > \mathbf{0}$ vektor, amelyek esetén az $\mathcal{L}_{\bar{\mathbf{w}}}$ szinthalmaz nem üres és kompakt.

Legyen $\mathbf{w}' > \bar{\mathbf{w}}$ és $\mathbf{w}' > \hat{\mathbf{w}}$. Ekkor az $\mathcal{A} := \{\mathbf{w} \in \mathbb{R}^n_\oplus \; : \; \mathbf{e}^T\mathbf{w} \leq \mathbf{e}^T\mathbf{w}'\}$ nem üres és kompakt halmaz. Továbbá az $\mathcal{A} \cap \mathcal{G}$ halmaz sem üres és kompakt. Definiáljuk az $f : \mathcal{A} \cap \mathcal{G} \to \mathbb{R}^n_\oplus$ függvényt a következő kifejezéssel

$$f_i(\mathbf{w}) := \begin{cases} 0, & \text{ha } w_i \leq \hat{w}_i \\ w_i - \hat{w}_i, & \text{különben} \end{cases}$$

Nyilván az f_i leképezés folytonos bármely i index esetén, tehát az f is folytonos. A $\|.\|_\infty$ folytonos függvény, ezért az $\|f(.)\|_\infty$ is az. Ekkor a Weirstrass–tétele miatt, létezik

$$\gamma = \|f(\tilde{\mathbf{w}})\|_\infty = \min_{\mathbf{w} \in \mathcal{A} \cap \mathcal{G}} \|f(\mathbf{w})\|_\infty \leq \|f(\bar{\mathbf{w}})\|_\infty.$$

Mivel az $\mathcal{L}_{\hat{\mathbf{w}}} = \emptyset$ (indirekt feltevés), ezért $\gamma = \|f(\tilde{\mathbf{w}})\|_\infty > 0$.

Ha $int\mathcal{T}(\tilde{\mathbf{w}},\, \hat{\mathbf{w}}) \neq \emptyset$ akkor létezik $\alpha \in (0,1)$, amelyre $\mathbf{x}^+ = \bar{\mathbf{x}} + \alpha\,\Delta\mathbf{x}$ és $\mathbf{s}^+ = \bar{\mathbf{s}} + \alpha\,\Delta\mathbf{s}$ olyan vektorok, amelyekre $\mathbf{w}^+ = \mathbf{x}^+\,\mathbf{s}^+ \in int\mathcal{T}(\tilde{\mathbf{w}},\, \hat{\mathbf{w}})$, azaz

$$\|f(\mathbf{w}^+)\|_\infty < \|f(\tilde{\mathbf{w}})\|_\infty = \gamma$$

ami ellentmond a Weirstrass–tételnek.

Ha $int\mathcal{T}(\tilde{\mathbf{w}},\, \hat{\mathbf{w}}) = \emptyset$ akkor legyen $\mathbf{w}^* \in \mathcal{B}_{\gamma/3}(\hat{\mathbf{w}}) \cap int\mathcal{T}(\mathbf{0},\, \hat{\mathbf{w}})$ vektor, amely esetén $int\mathcal{T}(\tilde{\mathbf{w}},\, \mathbf{w}^*) \neq \emptyset$ és megismételhetjük az előző gondolatmenetet. ($\mathcal{B}_\gamma(\mathbf{w}) := \{\mathbf{x} \in \mathbb{R}^n \; : \; \|\mathbf{x} - \mathbf{w}\| < \gamma\}$, azaz a $\mathbf{w}$ közepű γ sugarú gömböt jelöli.) $\square$

Ezek után egyszerűen bizonyítható, hogy belső pont feltevés mellett a ferdén szimmetrikus feladat optimális megoldásainak halmaza kompakt.

2.3.5. Következmény. *Legyen adott az (SP) feladat és tegyük fel, hogy $\mathcal{F}^0 \neq \emptyset$. Ekkor az $\mathcal{F}^*$ halmaz kompakt.*

Bizonyítás. Legyen $\mathbf{w}^i \in \mathbb{R}^n_\oplus$ és $\mathbf{w}^i > \mathbf{w}^{i+1}$, valamint $\lim_{i\to\infty} \mathbf{w}^i = \mathbf{0}$. Ekkor

$$\mathcal{L}_{\mathbf{w}^{i+1}} \subset \mathcal{L}_{\mathbf{w}^i}$$

teljesül. Ebből

$$\mathcal{F}^* = \mathcal{L}_{\mathbf{0}} = \cap_{i=1}^{\infty} \mathcal{L}_{\mathbf{w}^i}$$

adódik, azaz az $\mathcal{F}^*$ halmaz zárt és mivel $\mathcal{F}^* \subseteq \mathcal{L}_{\mathbf{w}^i}$ következik bármely i indexre, ezért korlátos is. Tehát az $\mathcal{F}^*$ kompakt halmaz is. $\square$

Belátható a következő állítások ekvivalenciája. Tekintettel arra, hogy ezeknek az állításoknak az ekvivalenciáját nem fogjuk használni a bizonyítás részleteit az olvasóra bízzuk.

2.3.6. Állítás. *Legyen adott az (SP) feladat és tegyük fel, hogy $\mathcal{F} \neq \emptyset$. Ekkor az alábbi állítások ekvivalensek:*

(i) *az* $\mathcal{F}^0 \neq \emptyset$

(ii) *az* $\mathcal{L}_{\mathbf{w}}$ *kompakt, nem üres halmaz;*

(iii) *az* $\mathcal{L}_K$ *kompakt, nem üres halmaz;*

(iv) *az* $\mathcal{F}^*$ *kompakt, nem üres halmaz.*

2.4. Centrális út

Az (SP) feladat optimális kritériumának a relaxáltja

$$\left.\begin{array}{r} -M\mathbf{x} + \ \mathbf{s} = \mathbf{q} \\ \mathbf{x} \geq \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \\ \mathbf{x}\,\mathbf{s} = \mu\mathbf{e} \end{array}\right\}$$

2.4.1. Definíció. *Az alábbi halmazt*

$$\mathcal{C} := \{(\mathbf{x}(\mu), s(\mathbf{x}(\mu)))\ :\ \mathbf{x}(\mu) \in \mathcal{F}^0,\ \mathbf{x}(\mu)\mathbf{s}(\mu) = \mu\mathbf{e},\ valamely \mu \in \mathbb{R}_+\}$$

az (SP) feladat centrális útján*ak nevezzük.*

A továbbiakban a centrális út tulajdonságaival foglalkozunk. Először egy segédtételt bizonyítunk be, amelyet a későbbiekben használunk fel.

2.4.1.1. Segédtétel. *Legyen $M \in \mathbb{R}^{n\times n}$ egy ferdén szimmetrikus mátrix. Ekkor minden $\mathbf{z} \in \mathbb{R}^n \setminus \{0\}$ vektor esetén létezik olyan j index, melyre $z_j \neq 0$ és $z_j(M\mathbf{z})_j \geq 0$.*

Bizonyítás. Indirekt tegyük fel, hogy létezik olyan $\overline{\mathbf{z}} \neq 0$, hogy minden j-re, amelyre $\overline{z}_j \neq 0$, $\overline{z}_j(M\overline{\mathbf{z}})_j < 0$. Ez esetben

$$0 = \overline{\mathbf{z}}^T M\overline{\mathbf{z}} = \sum_{j=1}^{n} \overline{z}_j(M\overline{\mathbf{z}})_j = \sum_{j:\overline{z}_j\neq 0} \overline{z}_j(M\overline{\mathbf{z}})_j < 0 \ ,$$

ami nyilvánvaló ellentmondás. □

Megmutatható, hogy belső pont létezése esetén a centrális út létezik és egyértelmű, sőt a centrális út egy környezete is egyértelmű. Ennél többet mutatunk meg a következő tételben, mégpedig a fenti tulajdonségok ekvivalenciáját.

2.4.2. Tétel. *Legyen adott az (SP) feladat és tegyük fel, hogy $\mathcal{F} \neq \emptyset$. Ekkor a következő állítások ekvivalensek:*

(i) $\mathcal{F}^0 \neq \emptyset$

(ii) $\forall \mathbf{w} \in \mathbb{R}^n_+$ *esetén* $\exists!\ (\mathbf{x}, \mathbf{s}) > 0\ :\ M\mathbf{x} + \mathbf{q} = \mathbf{s}$ *és* $\mathbf{xs} = \mathbf{w}$

(iii) $\forall \mu > 0$ *esetén* $\exists!\ (\mathbf{x}, \mathbf{s}) > 0\ :\ M\mathbf{x} + \mathbf{q} = \mathbf{s}$ *és* $\mathbf{xs} = \mu\mathbf{e}$

Bizonyítás. A (iii) az (ii) állítás speciális esete, így (ii) $\Rightarrow$ (iii).
A (iii) feltételből következik (i) ugyanis $\mathbf{x}(\mu)\mathbf{s}(\mu) = \mu\mathbf{e}$ esetén $\mathbf{x}(\mu) \in \mathcal{F}^0$
Az (i) $\Rightarrow$ (ii) implikációt indirekt bizonyítjuk. Tegyük fel, hogy $\exists\hat{\mathbf{w}} \in \mathbb{R}^n_+$ úgy, hogy $\nexists\hat{\mathbf{x}} \in \mathcal{F}^0$ és $\hat{\mathbf{s}} = M\hat{\mathbf{x}} + \mathbf{q}$, amire teljesülne $\hat{\mathbf{w}} = \hat{\mathbf{x}}\hat{\mathbf{s}}$. Mivel fennáll a belső pont feltétel, $\mathcal{L}_{\hat{\mathbf{w}}} \neq \emptyset$ és kompakt.

Legyen $f\ :\ \mathcal{L}_{\hat{\mathbf{w}}} \rightarrow \mathbb{R}_+$ függvény, $f(\mathbf{x}, \mathbf{s}) = \mathbf{x}^T\mathbf{s}$. Az indirekt feltevésünk miatt $f(\mathbf{x}, \mathbf{s}) = \mathbf{x}^T\mathbf{s} < \mathbf{e}^T\hat{\mathbf{w}}$ minden $(\mathbf{x}, \mathbf{s}) \in \mathcal{L}_{\hat{\mathbf{w}}}$. Az f folytonos függvény, így Weierstrass tétele miatt felveszi a maximumát az $\mathcal{L}_{\hat{\mathbf{w}}}$ kompakt halmazon, azaz

$$\mathbf{e}^T\overline{\mathbf{w}} = \overline{\mathbf{x}}^T\overline{\mathbf{s}} = f(\overline{\mathbf{x}}, \overline{\mathbf{s}}) = \max_{(\mathbf{x},\mathbf{s})\in\mathcal{L}_{\hat{\mathbf{w}}}} f(\mathbf{x}, \mathbf{s}) = \max_{(\mathbf{x},\mathbf{s})\in\mathcal{L}_{\hat{\mathbf{w}}}} \mathbf{x}^T\mathbf{s} < \mathbf{e}^T\hat{\mathbf{w}} \tag{2.7}$$

ahol $\overline{\mathbf{w}} := \overline{\mathbf{xs}}$ és $\overline{\mathbf{w}} \leq \hat{\mathbf{w}}$, de $\overline{\mathbf{w}} \neq \hat{\mathbf{w}}$ az indirekt feltevésünk miatt.
Két eset lehetséges:

1. Ha $int\mathcal{T}(\overline{\mathbf{w}}, \hat{\mathbf{w}}) \neq \emptyset$, akkor $\exists \mathbf{w}^+ \in \mathcal{T}(\overline{\mathbf{w}}, \hat{\mathbf{w}}) : \overline{\mathbf{w}} < \mathbf{w}^+ < \hat{\mathbf{w}}$ és $\mathbf{w}^+ = \mathbf{x}^+\mathbf{s}^+$, ahol $\mathbf{x}^+ \in \mathcal{F}^0$. Így $\mathbf{e}^T\overline{\mathbf{w}} < \mathbf{e}^T\mathbf{w}^+$, ez ellentmond a (2.7) egyenlőtlenségnek hiszen $(\mathbf{x}^+, \mathbf{s}^+) \in \mathcal{L}_{\hat{\mathbf{w}}}$

2. Ha $int\mathcal{T}(\overline{\mathbf{w}}, \hat{\mathbf{w}}) = \emptyset$, akkor legyen $\delta := \mathbf{e}^T(\hat{\mathbf{w}} - \overline{\mathbf{w}}) > 0$ és definiáljuk

$$\tilde{\mathbf{w}} \in \mathbb{R}^n_+ \ : \ \hat{\mathbf{w}} > \tilde{\mathbf{w}} > \hat{\mathbf{w}} - \frac{\delta}{n}\mathbf{e}.$$

Ekkor $\tilde{\mathbf{w}}$ választási szabálya miatt $\mathcal{T}(\overline{\mathbf{w}}, \tilde{\mathbf{w}}) \neq \emptyset$, azaz létezik $\mathbf{w}^+ \in \mathcal{T}(\overline{\mathbf{w}}, \tilde{\mathbf{w}})$ úgy, hogy $\mathbf{w}^+ = \mathbf{x}^+\mathbf{s}^+$, $\mathbf{x}^+ = \overline{\mathbf{x}} + \alpha\Delta\mathbf{x}$, $\mathbf{s}^+ = \overline{\mathbf{s}} + \alpha\Delta\mathbf{s}$, ahol $x^+ \in \mathcal{F}^0$ és $\alpha \in (0, 1)$. A továbbiakban megmutatjuk, hogy $\mathbf{e}^T\mathbf{w}^+ > \mathbf{e}^T\overline{\mathbf{w}}$, ami ellentmondás (2.7) miatt. Felhasználva a $\Delta\mathbf{x}$, $\Delta\mathbf{s}$ vektorok ortogonalitását, illetve az $\overline{\mathbf{s}}\Delta\mathbf{x} + \overline{\mathbf{x}}\Delta\mathbf{s} = \tilde{\mathbf{w}} - \overline{\mathbf{x}\mathbf{s}}$ egyenlőséget a következőt kapjuk:

$$\begin{aligned} \mathbf{e}^T(\mathbf{w}^+ - \overline{\mathbf{w}}) &= \overline{\mathbf{x}}^T\overline{\mathbf{s}} + \alpha(\overline{\mathbf{s}}^T\Delta\mathbf{x} + \overline{\mathbf{x}}^T\Delta\mathbf{s}) + \alpha^2(\Delta\mathbf{x})^T\Delta\mathbf{s} - \overline{\mathbf{x}}^T\overline{\mathbf{s}} \\ &= \alpha(\overline{\mathbf{s}}^T\Delta\mathbf{x} + \overline{\mathbf{x}}^T\Delta\mathbf{s}) = \alpha \ \mathbf{e}^T(\tilde{\mathbf{w}} - \overline{\mathbf{w}}) \end{aligned} \tag{2.8}$$

$\tilde{\mathbf{w}}$ definíciója miatt

$$\tilde{\mathbf{w}} - \overline{\mathbf{w}} > (\hat{\mathbf{w}} - \overline{\mathbf{w}}) - \frac{\delta}{n}\mathbf{e}, \quad \text{így} \quad \mathbf{e}^T(\tilde{\mathbf{w}} - \overline{\mathbf{w}}) > \mathbf{e}^T(\hat{\mathbf{w}} - \overline{\mathbf{w}}) - \frac{\delta}{n}\mathbf{e}^T\mathbf{e} = \delta - \delta = 0.$$

A (2.8) egyenlőséget figyelembe véve $\mathbf{e}^T\mathbf{w}^+ > \mathbf{e}^T\overline{\mathbf{w}}$, amit igazolni szerettünk volna, azaz ellentmondásra jutottunk.

A fentiekben beláttuk, $\forall \mathbf{w} \in \mathbb{R}^n_+$ esetén $\exists(\mathbf{x}, \mathbf{s}) > 0 : \ M\mathbf{x} + \mathbf{q} = \mathbf{s}, \ \mathbf{w} = \mathbf{x}\mathbf{s}$.

Az egyértelműség bizonyításához tegyük fel, hogy $\mathbf{w} \in \mathbb{R}^n_+$ vektornak létezik két különböző előállítása, azaz létezik $(\mathbf{x}^1, \mathbf{s}^1) > 0$ és $(\mathbf{x}^2, \mathbf{s}^2) > 0$, hogy

$$M\mathbf{x}^1 + \mathbf{q} = \mathbf{s}^1, \ \mathbf{w} = \mathbf{x}^1\mathbf{s}^1 \ \text{ és } \ M\mathbf{x}^2 + \mathbf{q} = \mathbf{s}^2, \ \mathbf{w} = \mathbf{x}^2\mathbf{s}^2, \ \text{továbbá } (\mathbf{x}^1, \mathbf{s}^1) \neq (\mathbf{x}^2, \mathbf{s}^2).$$

Legyen $\mathbf{z} = \mathbf{x}^1 - \mathbf{x}^2 \neq \mathbf{0}$. Ekkor a fenti segédtétel alapján létezik olyan j index, amelyre

$$x_j^1 \neq x_j^2 \quad \text{és} \quad 0 \leq (x_j^1 - x_j^2)(M(\mathbf{x}^1 - \mathbf{x}^2))_j = (x_j^1 - x_j^2)(s_j^1 - s_j^2).$$

Feltehető, hogy $x_j^1 > x_j^2$ (a másik eset hasonlóan bizonyítható), így $s_j^1 > s_j^2$. Mivel $x_j^1 s_j^1 = w_j = x_j^2 s_j^2$, ezért

$$1 < \frac{x_j^1}{x_j^2} = \frac{s_j^2}{s_j^1} \Rightarrow s_j^1 < s_j^2 \ , \text{ ami ellentmond az előbbi feltevésünknek.} \qquad \square$$

A következőkben megmutatjuk, hogy a centrális út torlódási pontja az (SP) feladat egy optimális, sőt szigorúan komplementáris megoldása.

2.4.3. Tétel. *Legyen adott az* (SP) *feladat, amelyre teljesül a belső pont feltétel. Ekkor létezik* $(\mathbf{x}^*, \mathbf{s}^*)$ *az alábbi tulajdonságokkal:*

(1) $(\mathbf{x}^*, \mathbf{s}^*) = \lim_{\mu\to 0}(\mathbf{x}(\mu), \mathbf{s}(\mu))$

(2) $\mathbf{x}^* \in \mathcal{F}^*$

(3) $(\mathbf{x}^*, \mathbf{s}^*)$ *szigorúan komlementáris megoldás.*

Bizonyítás. Legyen $\mu_k \to 0$ $(k = 1, 2, \dots)$ egy monoton csökkenő sorozat. Ekkor minden k indexre fennáll, hogy $(\mathbf{x}(\mu_k), \mathbf{s}(\mu_k)) \in \mathcal{L}_{\mathbf{w}^0}$, ahol $\mathbf{w}^0 = \mathbf{x}^0\mathbf{s}^0$. Mivel az $\mathcal{L}_{\mathbf{w}^0}$ halmaz kompakt, létezik a sorozatnak legalább egy $(\mathbf{x}^*, \mathbf{s}^*)$ torlódási pontja. Az általánosság korlátozása nélkül feltehető, hogy $(\mathbf{x}^*, \mathbf{s}^*) = \lim_{k\to\infty}(\mathbf{x}(\mu_k), \mathbf{s}(\mu_k))$.

A második tulajdonság nyilvánvaló, ugyanis

$$\mathbf{x}^*\mathbf{s}^* = \lim_{k\to\infty}(\mathbf{x}(\mu_k), \mathbf{s}(\mu_k)) = \lim_{k\to\infty}\mu_k\mathbf{e} = 0$$

Az utolsó állítás igazolásához vezessük be a következő jelölést. Legyen $\mathbf{x} \in \mathbb{R}^n_\oplus$ és jelölje $\sigma(\mathbf{x}) = \{i \ : \ x_i > 0\}$ az $\mathbf{x}$ vektor tartóját. Felhasználva az M mátrix ferdén szimetrikusságát

$$0 = (\mathbf{x}^* - \mathbf{x}(\mu))^T(\mathbf{s}^* - \mathbf{s}(\mu)) = (\mathbf{x}^*)^T\mathbf{s}^* + \mathbf{x}(\mu)^T\mathbf{s}(\mu) - (\mathbf{x}^*)^T\mathbf{s}(\mu) - (\mathbf{s}^*)^T\mathbf{x}(\mu)$$

egyenlet adódik. Figyelembe véve a második állítást és az $x(\mu)_i s(\mu)_i = \mu$ összefüggést

$$\begin{aligned}
\mu n &= (\mathbf{x}^*)^T\mathbf{s}(\mu) + (\mathbf{s}^*)^T\mathbf{x}(\mu) = \sum_{i\,:\,x_i^*>0} x_i^*(s(\mu))_i + \sum_{i\,:\,s_i^*>0} s_i^*(x(\mu))_i \\
n &= \sum_{i\,:\,x_i^*>0} x_i^*\frac{(s(\mu))_i}{\mu} + \sum_{i\,:\,s_i^*>0} s_i^*\frac{(x(\mu))_i}{\mu} = \sum_{i\,:\,x_i^*>0}\frac{x_i^*}{(x(\mu))_i} + \sum_{i\,:\,s_i^*>0}\frac{s_i^*}{(s(\mu))_i}
\end{aligned}$$

Határátmenettel, amikor $\mu \to 0$, akkor

$$|\sigma(\mathbf{x}^*)| + |\sigma(\mathbf{s}^*)| = n,$$

ami csak úgy teljesülhet, ha az $(\mathbf{x}^*, \mathbf{s}^*)$ egy szigorúan komplementáris megoldás. □

Az előző tétel harmadik állítása Goldman és Tucker eredménye, ami a következő formában ismert.

2.4.4. Következmény. (Goldman–Tucker tétel belső pont feltevés mellett) *Legyen adott az* (SP) *feladat és tegyük fel, hogy* $\mathcal{F}^0 \neq \emptyset$. *Ekkor létezik szigorúan komplementáris megoldás.*

2.5. Optimális partíció, analitikus centrum

Ebben a fejezetben a centrális út torlódási pontját vizsgaljuk meg. Ehhez definiáljuk a vektorok indexhalmazának egy partícióját, melyről a következő fejezetekben további tulajdonságokat látunk be.

2.5.1. Definíció. *Legyen*

$$\begin{aligned} B: &= \{i : x_i > 0, \quad \textit{valamely} \;\; x \in \mathcal{F}^* \;\; \textit{esetén}\} \\ N: &= \{i : s(x)_i > 0, \quad \textit{valamely} \;\; x \in \mathcal{F}^* \;\; \textit{esetén}\} \end{aligned}$$

Ekkor az indexeknek a (B, N) *felbontását* optimális partíciónak *nevezzük.*

A definícióban megelőlegeztük, hogy a (B, N) felbontás egy partíció, ezt az alábbiakban bebizonyítjuk.

2.5.2. Állítás. *Legyen adott az* (SP) *feladat, és tegyük fel, hogy teljesül a belső pont feltétel. Ekkor* $I = \{1, 2, \ldots, n\}$ *indexhalmaznak a* (B, N) *valóban partíciója.*

Bizonyítás. $B \cup N = I$, mert a 2.4.4.Következmény miatt létezik $(\mathbf{x}^*, \mathbf{s}^*)$ szigorúan komplementáris megoldás. A 2.1.2. Lemma miatt pedig $B \cap N = \emptyset$ is teljesül. □

A centrális út az optimális megoldások halmazának egy speciális pontjához tart amikor μ tart nullához, ez az úgynevezett analitikus centrum, amit a következőképpen definiálunk

2.5.3. Definíció. *Jelölje* $\overline{\mathbf{x}} \in \mathcal{F}^*$, $\overline{\mathbf{s}} = \overline{\mathbf{s}}(x)$ *azon vektorokat, melyek maximalizálják a*

$$\prod_{i \in B} x_i \prod_{i \in N} s_i$$

szorzatot az $\mathcal{F}^*$ *optimális halmazon. Ekkor az* $\overline{\mathbf{x}}$ *vektort az* $\mathcal{F}^*$ *optimális halmaz* analitikus centrum*ának nevezzük.*

Az 2.3.5. Következmény szerint belső pont feltevés mellett az $\mathcal{F}^*$ halmaz korlátos, így ekkor létezik az analitikus centrum. Továbbá létezik szigorúan komplementáris megoldás a 2.4.4. Következmény alapján, így a definícióban szereplő szorzat maximumértéke pozitív, tehát az analitikus centrum egyben szigorúan komplementáris megoldás.

2.5.4. Tétel. (Sonnevend) *Legyen a centrális út határpontja* $(\mathbf{x}^*, \mathbf{s}^*)$. *Ekkor* $\mathbf{x}^*$ *az* $\mathcal{F}^*$ *halmaz analitikus centruma.*

Bizonyítás. Legyen $\overline{\mathbf{x}} \in \mathcal{F}^*$ tetszőleges, és $\overline{\mathbf{s}} = s(\overline{\mathbf{x}})$. A 2.4.3 tétel bizonyításában leírtakhoz hasonlóan az alábbi összefüggésre jutunk:

$$\sum_{i\in B} \frac{\overline{x}_i}{x_i^*} + \sum_{i\in N} \frac{\overline{s}_i}{s_i^*} = n.$$

Mivel $x_i^* > 0 \ \forall i \in B$ és $s_i^* > 0 \ \forall i \in N$, alkalmazhatjuk a számtani-mértani közép közti egyenlőtlenséget:

$$\left(\prod_{i\in B} \frac{\overline{x}_i}{x_i^*} \prod_{i\in N} \frac{\overline{s}_i}{s_i^*}\right)^{\frac{1}{n}} \leq \frac{1}{n} \left(\sum_{i\in B} \frac{\overline{x}_i}{x_i^*} + \sum_{i\in N} \frac{\overline{s}_i}{s_i^*}\right) = 1.$$

Így

$$\prod_{i\in B} \overline{x}_i \prod_{i\in N} \overline{s}_i \leq \prod_{i\in B} x_i^* \prod_{i\in N} s_i^*,$$

□

2.6. Az optimális partíció meghatározása

A kondíció szám

Ahhoz, hogy a változók koordinátáira a centrális út mentén egy korlátot tudjunk adni, szükségünk lesz egy, az optimális megoldások halmazát jellemző mérőszámra. Az előzőek alapján az $\mathbf{x} \in \mathcal{F}$ vektorra pontosan akkor igaz, hogy $\mathbf{x} \in \mathcal{F}^*$, ha az $\mathbf{x}_N = \mathbf{0}$ és $\mathbf{s}_B = \mathbf{0}$ teljesül. Tetszőleges $x \in \mathcal{F}^*$ optimális megoldásra fennáll az

$$\mathbf{x}\, s(\mathbf{x}) = \mathbf{0} \quad \text{és} \quad \mathbf{x} + s(\mathbf{x}) \geq \mathbf{0}$$

összefüggés. Első lépésként az optimális megoldások nemnulla koordinátáinak nagyságára keresünk egy jellemző mennyiséget. Definiáljuk az (SP) feladat kondíciószámát, amely az $\mathcal{F}^*$ optimális halmazon a nemnulla koordináták nagyságrendjét jellemzi.

2.6.1. Definíció. *Legyen*

$$\sigma_{SP}^x = \begin{cases} \min\limits_{i\in B} \max\limits_{(\mathbf{x},\mathbf{s})\in\mathcal{F}^*} \{x_i\}, & \textit{ha } B \neq \emptyset \\ \infty, & \textit{különben} \end{cases} \quad \textit{és } \sigma_{SP}^s = \begin{cases} \min\limits_{i\in N} \max\limits_{(\mathbf{x},\mathbf{s})\in\mathcal{F}^*} \{s(x)_i\}, & \textit{ha } N \neq \emptyset \\ \infty, & \textit{különben} \end{cases}$$

Az (SP) *feladat* kondíció száma $\quad \sigma_{SP} = \min\{\sigma_{SP}^x, \sigma_{SP}^s\}$.

Mivel a B és N halmazok egyszerre nem lehetnek üresek, az $\mathcal{F}^*$ halmaz pedig korlátos, σ^x_{SP} és σ^s_{SP} számok közül legalább az egyik véges. Továbbá $\sigma^x_{SP}, \sigma^s_{SP} > 0$, azaz a kondíciószám egy véges pozitív szám. Valamint σ^x_{SP} és σ^s_{SP} esetén a maximumot $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^*$ komplementáris megoldásokra nézzük, következésképpen

$$\sigma_{SP} = \min_i \max_{(\mathbf{x},\mathbf{s})\in\mathcal{F}^*} \{x_i + s(x)_i\}$$

2.6.2. Lemma. *Tekintsük a következő megengedettségi feladatot*

$$\mathcal{M} = \{\mathbf{x} \in \mathbb{R}^n \; ; \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\},$$

ahol az $A \in \mathbb{Z}^{m\times n}$ *mátrix és* $\mathbf{b} \in \mathbb{Z}^m$, $\mathbf{b} \neq \mathbf{0}$ *vektor. Legyen az* $\mathcal{M}$ *korlátos halmaz és tegyük fel, hogy tartalmaz egy pozitív vektort. Ekkor bármely* $i \in I$ *index esetén*

$$\max_{\mathbf{x}\in\mathcal{M}} x_i \geq \frac{1}{\prod_{j=1}^n} \|\mathbf{a}_j\|.$$

Bizonyítás. $\mathcal{M}$ korlátos halmaz, így az A mátrixnak nem lehet csupa 0 oszlopa. Ez a tulajdonság szükséges ahhoz, hogy a Cramer szabályt alkalmazhassuk majd a bizonyítás során. Két eset lehetséges:

1.eset $rangA = m$ (így szükségképpen $m \leq n$)

Rögzítsük az $i \in I$ indexet és legyen $\mathbf{x} \in \mathcal{M}$ olyan vektor, amelyre x_i maximális (ilyen létezik, ui az $\mathcal{M}$ halmaz korlátos és zárt). Mivel az $\mathcal{M}$ halmaz tartalmaz pozitív vektort, nyilván $x_i > 0$. Legyen $J := supp(\mathbf{x}) = \{j \in I \; : \; x_j > 0\}$. Tegyük fel, hogy az $\mathbf{x}$ olyan vektor, melynek a tartója ($supp(\mathbf{x})$) minimális méretű. Ekkor az $\mathbf{a}_j$ vektorok lineárisan függetlenek (indirekt bizonyítható). Ha $|J| < m$, akkor J kiegészíthető egy K indexhalmazzá úgy, hogy $\{\mathbf{a}_j \; : \; j \in K\}$ lineárisan független, és $|K| = m$, ugyanis az A mátrix rangja m.

Ekkor $A_K\mathbf{x}_K = \mathbf{b}$ egyenletrendszer $\mathbf{x}_K$ megoldását kiegészítve nullákkal, azaz $x_j := 0, \; j \notin K$, a kapott x vektor az eredeti rendszernek megoldása. A Cramer szabály miatt

$$x_j = \frac{\det(A_K^{(j)})}{\det(A_K)} \qquad \forall j \in J \subset K,$$

ahol $A_K^{(j)}$ jelöli azt a mátrixot, amelyet úgy kapunk, hogy az A mátrix j. oszlopát $\mathbf{b}$-vel helyettesítjük. Mivel $x_j > 0$ $(j \in J)$, ezért $\det(A_K^{(j)}) \neq 0$, valamint az $A \in \mathbb{Z}^{m\times n}$ is figyelembe véve $|\det(A_K^{(j)})| \geq 1$. Tehát

$$x_j = |x_j| = \frac{|\det(A_K^{(j)})|}{|\det(A_K)|} \geq \frac{1}{|\det(A_K)|}$$

Alkalmazzuk a Hadamard egyenlőtlenséget: $\det(A_K) \leq |\det(A_K)| \leq \prod_{j=1}^{n} \|\mathbf{a}_j\|$, az $\mathbf{x}$ vektor koordinátáira az elvárt alsó becslést kapjuk:

$$x_j \geq \frac{1}{\prod_{j=1}^{n} \|\mathbf{a}_j\|}$$

2.eset $q = rang(A) < m$
A redundáns feltételek elhagyásával egy $A' \in \mathbb{Z}^{q \times n}$, $\mathbf{b} \in \mathbb{Z}^q$ feladatot kapunk, ahol a mátrixunk már teljes rangú, azaz az első esethez jutottunk.
Vegyük észre, hogy $\|\mathbf{a}'_j\| \leq \|\mathbf{a}_j\| \quad (j = 1, 2, \ldots, n)$, így a második esetben is igaz az állításunk. □

A fenti tétel akkor is igaz marad, ha a feladat adatai racionális számok, ugyanis ekkor a legkisebb közös többszörössel beszorozva az eredetivel ekvivalens, de már egészértékű feladatot kapunk.

2.6.3. Tétel. *Legyen adott az (SP) feladat, melyre teljesül a belső pont feltétel. Ha az M mátrix és a $\mathbf{q}$ vektor racionális értékű, akkor az (SP) feladat kondíciószámára a következő alsó becslés teljesül:*

$$\sigma_{SP} \geq \frac{1}{\prod_{j=1}^{n} \|\mathbf{m}_j\|}$$

Bizonyítás. Tegyük fel, hogy a (B, N) partíciót ismerjük, ekkor

$$\begin{pmatrix} \mathbf{s}_B \\ \mathbf{s}_N \end{pmatrix} = \begin{pmatrix} M_{BB} & M_{BN} \\ M_{NB} & M_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} + \begin{pmatrix} \mathbf{q}_B \\ \mathbf{q}_N \end{pmatrix}$$

Legyen $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}$ szigorúan komplementáris megoldás. A komplementaritás miatt a vektorunk optimális (4.1.3. Következmény), azaz $\mathbf{q}^T\mathbf{x} = 0$. Továbbá felhasználva a (B, N) partíció definícióját ($\mathbf{x}_N = \mathbf{0}$) kapjuk, hogy $\mathbf{q}_B^T\mathbf{x}_B = 0$. Figyelembe véve a szigorú komplementaritást $\mathbf{x}_B > \mathbf{0}$, így szükségszerűen $\mathbf{q}_B = \mathbf{0}$. Ezeket felhasználva a fenti rendszerünk a következő alakú:

$$\begin{pmatrix} 0 \\ \mathbf{s}_N \end{pmatrix} = \begin{pmatrix} M_{BB} & M_{BN} \\ M_{NB} & M_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{x}_B \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{q}_N \end{pmatrix} \Rightarrow \begin{pmatrix} 0 \\ \mathbf{s}_N \end{pmatrix} = \begin{pmatrix} M_{BB}\mathbf{x}_B \\ M_{NB}\mathbf{x}_B + \mathbf{q}_N \end{pmatrix}$$

Tehát

$$\begin{pmatrix} M_{BB} & 0_{BN} \\ M_{NB} & -I_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{x}_B \\ \mathbf{s}_N \end{pmatrix} = \begin{pmatrix} 0_B \\ -\mathbf{q}_N \end{pmatrix} \qquad \mathbf{x}_B \geq \mathbf{0},\ \mathbf{x}_N = \mathbf{0},\ \mathbf{s}_B = \mathbf{0},\ \mathbf{s}_N \geq \mathbf{0}$$

Az utóbbi rendszer éppen az $\mathcal{F}^*$ leírása, ami nem üres és tartalmaz szigorúan komplementáris megoldást, így az

$$A = \begin{pmatrix} M_{BB} & 0_{BN} \\ M_{NB} & -I_{NN} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{s}_N \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0_B \\ -\mathbf{q}_N \end{pmatrix}$$

választással alkalmazhatjuk az előző lemmát.

$$\max_{\mathbf{x}} x_i \geq \frac{1}{\prod_{j\in B}\|\mathbf{m}_j\|}, \quad \forall i \in B \qquad \max_{\mathbf{s}} s_i \geq \frac{1}{\prod_{j\in B}\|\mathbf{m}_j\|}, \quad \forall i \in N$$

$$\Rightarrow \quad \sigma_{SP} \geq \frac{1}{\prod_{j\in B}\|\mathbf{m}_j\|} \geq \frac{1}{\prod_{j=1}^{n}\|\mathbf{m}_j\|}$$

□

A változók mérete a centrális út mentén

A σ_{SP} kondíció szám segítségével alsó és felső korlátokat adhatunk a változókra a centrális út mentén. Legyen (B, N) az (SP) feladat optimális partíciója.

2.6.4. Lemma. *Minden* $(\mathbf{x}(\mu), \mathbf{s}(\mu)) \in \mathcal{C}$, $\mu > 0$ *esetén igazak az alábbi egyenlőtlenségek:*

$$\begin{aligned} x_i(\mu) &\geq \frac{\sigma}{n} \quad i \in B, & x_i(\mu) &\leq \frac{n\mu}{\sigma} \quad i \in N, \\ s_i(\mu) &\leq \frac{n\mu}{\sigma} \quad i \in B, & s_i(\mu) &\geq \frac{\sigma}{n} \quad i \in N. \end{aligned}$$

Bizonyítás. Legyen $(\mathbf{x}^*, \mathbf{s}^*)$ egy tetszőleges optimális megoldás. Ekkor az ortogonalitási tulajdonság miatt

$$\begin{aligned} (\mathbf{x}(\mu) - \mathbf{x}^*)^T(\mathbf{s}(\mu) - \mathbf{s}^*) &= 0, \\ \mathbf{x}(\mu)^T\mathbf{s}^* + \mathbf{s}(\mu)^T\mathbf{x}^* &= n\mu, \\ x_i(\mu)s_i^* \leq \mathbf{x}(\mu)^T\mathbf{s}^* &\leq n\mu, \ \ 1 \leq i \leq n. \end{aligned}$$

Mivel $s_i^* \geq \sigma$ és $x_i(\mu)s_i(\mu) = \mu$, minden $i \in N$ esetén, így

$$x_i(\mu) \leq \frac{n\mu}{s_i^*} \leq \frac{n\mu}{\sigma} \quad \text{és} \quad s_i(\mu) \geq \frac{\sigma}{n}, \ \ i \in N.$$

A többi korlát hasonlóan bizonyítható. □

A lemma alapján természetesen adódnak a következő elnevezések:

2.6.5. Definíció. *Legyen* $(\mathbf{x}(\mu), \mathbf{s}(\mu)) \in \mathcal{C}$, $\mu > 0$. *Ekkor az* $x_i(\mu)$ *koordinátát* nagy*nak mondjuk, ha* $i \in B$, *illetve az* $s_i(\mu)$ *koordinátát* nagy*nak mondjuk, ha* $i \in N$. *Ellenkező esetben* kicsi*nek nevezzük.*

Az elnevezés is mutatja, hogy az $\mathbf{x}$ vektor nagy koordinátái alkotják a B halmazt, míg az $\mathbf{s}$ nagy koordinátái az N halmazt. A lemma alapján ha a μ paraméter elég kicsi, akkor egyértelműen eldönthető, hogy a vektorok mely koordinátái nagyok és melyek kicsik, azaz a (B, N) partíció egyértelműen meghatározható.

2.6.6. Következmény. *Ha* $\mu < \frac{\sigma^2}{n^2}$ *centrális út paraméterhez ismert egy* $(\mathbf{x}(\mu), \mathbf{s}(\mu)) \in \mathcal{C}$ *megoldás, akkor meghatározhatjuk a* (B, N) *partíciót.*

Bizonyítás. A megadott μ érték mellett $\frac{n\mu}{\sigma} < \frac{\sigma}{n}$ teljesül, így az előző lemma alapján a következő szétosztás egyértelmű és megfelelő

$$B := \{i \ : \ x_i(\mu) \geq \frac{\sigma}{n}\} \qquad N := \{i \ : \ s_i(\mu) \geq \frac{\sigma}{n}\}$$

□

3. fejezet

Dikin–féle affin skálázású algoritmus

Az előző fejezetben megmutattuk, hogy ha létezik belső pont, akkor a centrális utat követve a ferdén szimmetrikus önduális feladat megoldásához tartunk. Ennek a fejezetnek a célja igazolni, hogy ezt a gondolatmenetet követve *polinom időben* a ferdén szimmetrikus önduális feladat megoldásához tetszőlegesen közeli pontot tudunk előállítani, azaz tetszőleges $\varepsilon > 0$ számhoz olyan $(\mathbf{x}, \mathbf{s})$ pontpárt, amely a megengedett tartományban van és $\mathbf{x}^T\mathbf{s} \leq \varepsilon$. E célból egy konkrét algoritmust, egy affin skálázású primál–duál módszert ismertetünk, melyet Dikin [2] 1967-ben fogalmazott meg. Ezt követően a (B, N) partíció meghatározásáról ejtünk pár szót, majd ennek segítségével ismertetjük a kerekítési eljárást Ye [30] 1992-es eredménye és Terlakyék könyve [24] alapján, azaz azt, hogy a Dikin módszerrel kapott jó közelítő megoldásból milyen módon lehet egy szigorúan komplementáris megoldást előállítani (a szakirodalomban például az ellipszoid módszer estén is találkozhatunk kerekítési eljárással [3, 25]). Ezzel eloszlatjuk azt a tévhitet ami a mai napig a köztudatban él, és amelyet Andersen és Ye [1] fogalmazott meg, miszerint lineáris optimalizálási feladat pontos megoldása belsőpontos módszerrel polinom időben nem állítható elő.

A Dikin algoritmust pozitív szemidefinit mátrixokkal adott lineáris komplementaritási feladatokra általánosította Jansen, Roos és Terlaky [9]. Ezt követően Illés, Roos, Terlaky [5] $P_*(\kappa)$–mátrixokkal meghatározott lineáris komplementaritási feladatokra is megmutatta az algoritmus hatékonyságát, valamint Illés, Peng, Roos és Terlaky [6] a kerekítési eljárást is elemezte ebben az esetben.

3.1. Dikin–féle affin skálázású algoritmus

A fejezetet két technikai jellegű lemmával kezdjük, melyeket majd a későbbi vizsgálatok során használunk fel.

3.1.1. Lemma. *Legyen* $D \in \mathbb{R}^{n\times n}$ *pozitív definit,* $M \in \mathbb{R}^{n\times n}$ *ferdén szimmetrikus mátrix. Ekkor* $D + M$ *és* $I + DMD$ *pozitív definit mátrixok.*

Bizonyítás. Bármely $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$ esetén

$$\mathbf{x}^T(D+M)\mathbf{x} = \mathbf{x}^T D\mathbf{x} + \mathbf{x}^T M\mathbf{x} = \mathbf{x}^T D\mathbf{x} > 0,$$

$$\mathbf{x}^T(I+DMD)\mathbf{x} = \mathbf{x}^T\mathbf{x} + \mathbf{x}^T DMD\mathbf{x} = \mathbf{x}^T\mathbf{x} + (\mathbf{x}^T D)M(D\mathbf{x}) = \mathbf{x}^T\mathbf{x} > 0.$$

□

Ezek után belátjuk, hogy két pozitív definit mátrixból az alábbi módon képezett mátrix is pozitív definit.

3.1.2. Lemma. *Legyenek* $A, B \in \mathbb{R}^{n\times n}$ *pozitív definit mátrixok. Ekkor* $A^{-1}B \in \mathbb{R}^{n\times n}$ *is pozitív definit mátrix.*

Bizonyítás. Legyen $A^{-1}B$ mátrix λ sajátértékéhez tartozó sajátvektora $\mathbf{q} \in \mathbb{R}^n$, azaz $A^{-1}B\,\mathbf{q} = \lambda\mathbf{q}$. Ekkor $B\,\mathbf{q} = \lambda\, A\,\mathbf{q}$. Szorozzuk meg balról $\mathbf{q}^T$ vektorral, adódik a

$$\mathbf{q}^T B\,\mathbf{q} = \lambda\,\mathbf{q}^T A\,\mathbf{q}$$

Mivel A és B pozitív definit, és $\mathbf{q} \neq \mathbf{0}$, ezért $\mathbf{q}^T A\,\mathbf{q} > 0$ illetve $\mathbf{q}^T B\,\mathbf{q} > 0$ és így

$$\lambda = \frac{\mathbf{q}^T B\,\mathbf{q}}{\mathbf{q}^T A\,\mathbf{q}} > 0$$ □

Csökkenési irány meghatározása

Tekintsünk az (SP) feladatot, és legyen $\mathbf{x} \in \mathcal{F}^0$, $\mathbf{s} = s(\mathbf{x}) > 0$.

Keressük $(\Delta\mathbf{x}, \Delta\mathbf{s}) \in \mathbb{R}^{2n}$:

$$\begin{aligned} -M\Delta\mathbf{x} + \Delta\mathbf{s} &= \mathbf{0} \\ \mathbf{x} + \Delta\mathbf{x},\ \mathbf{s} + \Delta\mathbf{s} &\geq \mathbf{0} \end{aligned}$$

Ennél mi többet szeretnénk elérni, mégpedig: $\mathbf{x}^+ = \mathbf{x} + \Delta\mathbf{x} > \mathbf{0}$ és $\mathbf{s}^+ = \mathbf{s} + \Delta\mathbf{s} > \mathbf{0}$, valamint, hogy közben csökkenjen a dualitásrés is. Az új dualitásrés értéke felhasználva, hogy M ferdén szimmetrikus mátrix:

$$\begin{aligned}\mathbf{q}^T\mathbf{x}^+ &= (\mathbf{x}^+)^T\mathbf{s}^+ = (\mathbf{x}+\Delta\mathbf{x})^T(\mathbf{s}+\Delta\mathbf{s}) = \mathbf{x}^T\mathbf{s} + \mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T\Delta\mathbf{s} + (\Delta\mathbf{x})^T\Delta\mathbf{s} = \\ &= \mathbf{x}^T\mathbf{s} + \mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T\Delta\mathbf{s} = \mathbf{q}^T\mathbf{x} + \mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T\Delta\mathbf{s}.\end{aligned}$$

A dualitás rés csökkenésének a feltétele az, hogy $\mathbf{q}^T\mathbf{x}^+ < \mathbf{q}^T\mathbf{x}$, azaz $\mathbf{s}^T\Delta\mathbf{x}+\mathbf{x}^T\Delta\mathbf{s} < 0$. Ebben az esetben a $(\Delta\mathbf{x}, \Delta\mathbf{s}) \in \mathbb{R}^{2n}$ vektort *csökkenési irány*nak nevezzük. Célunk csökkenési irány keresése, ami a következőképpen fogalmazható meg

$$\left.\begin{array}{r}\min\{\mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T\Delta\mathbf{s}\} \\[1em] -M\Delta\mathbf{x} + \Delta\mathbf{s} = \mathbf{0} \\[1em] \mathbf{x}+\Delta\mathbf{x} \geq \mathbf{0}\ ,\ \mathbf{s}+\Delta\mathbf{s} \geq \mathbf{0}\end{array}\right\}(IF)$$

iránymeghatározási segédfeladatot szeretnénk megoldani. De ez egy az eredetivel azonos „nehézségű” lineáris programozási feladat.

Dikin (1967) ötlete: Keressünk egy könnyebben megoldható feladatot, amely valamekkora dualitásréscsökkenést azért biztosít. Tegyük ezt úgy, hogy a pozitivitási megkötést „relaxáljuk”. A közelítést egy (konvex) zárt ellipszoiddal képzeli el.

Dikin ellipszoid: Legyenek

$$\begin{aligned}\mathcal{E}_D &:= \left\{\ (\mathbf{x}+\Delta\mathbf{x}, \mathbf{s}+\Delta\mathbf{s}) \in \mathbb{R}^{2n}\ :\ \left\|\frac{\Delta\mathbf{x}}{\mathbf{x}} + \frac{\Delta\mathbf{s}}{\mathbf{s}}\right\| \leq 1\right\} \\ \mathcal{F}_s^0 &:= \{(\mathbf{x},\mathbf{s}) \in \mathbb{R}^{2n}\ :\ \mathbf{x} \in \mathcal{F}^0\ ,\ \mathbf{s} > \mathbf{0}\}\end{aligned}$$

Ekkor $\mathcal{F}_s^0 \cap \mathcal{E}_D \neq \emptyset$, mert $(\mathbf{x},\mathbf{s}) \in \mathcal{F}_s^0 \cap int\ \mathcal{E}_D$. Belátható, hogy $\mathcal{F}_s^0 \cap \mathcal{E}_D$ egy nemüres kompakt halmaz.

Legyen $(\mathbf{x}+\Delta\mathbf{x}, \mathbf{s}+\Delta\mathbf{s}) \in \mathcal{F}_s^0 \cap \mathcal{E}_D$, ekkor

$$\begin{aligned}1 &\geq \left\|\frac{\Delta\mathbf{x}}{\mathbf{x}} + \frac{\Delta\mathbf{s}}{\mathbf{s}}\right\| = \left\|\frac{\mathbf{s}\,\Delta\mathbf{x} + \mathbf{x}\,\Delta\mathbf{s}}{\mathbf{x}\,\mathbf{s}}\right\| = \left\|\frac{\mathbf{s}\,\Delta\mathbf{x} + \mathbf{x}\,M\Delta\mathbf{x}}{\mathbf{x}\,\mathbf{s}}\right\| = \\ &= \left\|(XS)^{-1}(S+XM)\,\Delta\mathbf{x}\right\| = \left\|S^{-1}(X^{-1}S+M)\,\Delta\mathbf{x}\right\|\end{aligned}$$

Az előző lemmák alapján $S^{-1}(X^{-1}S + M)$ pozitív definit mátrix, ezért az

$$\|S^{-1}(X^{-1}S + M)\,\Delta\mathbf{x}\|$$

a $\Delta\mathbf{x}$ konvex kvadratikus függvénye, azaz

$$\{\Delta\mathbf{x} : \|S^{-1}(\mathbf{x}^{-1}S + M)\,\Delta\mathbf{x}\| \leq 1\}$$

korlátos konvex nemüres zárt halmaz. Vagyis az (IF) feladat relaxálható a következő módon:

$$\left.\begin{array}{r} \min\{\mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T M\Delta\mathbf{x}\} \\ \|S^{-1}(X^{-1}S + M)\,\Delta\mathbf{x}\| \leq 1 \end{array}\right\}(\overline{IF})$$

Az $(\overline{IF})$ lineáris célfüggvény minimalizálása konvex kompakt nemüres halmaz felett, így Weierstrass tétele miatt létezik megoldása.

Átskálázás

Az $(\overline{IF})$ feladat helyett tekinthetjük az

$$\left.\begin{array}{r} \min\{\mathbf{s}^T\Delta\mathbf{x} + \mathbf{x}^T\Delta\mathbf{s}\} \\ -M\Delta\mathbf{x} + \Delta\mathbf{s} = \mathbf{0} \\ \left\|\frac{\Delta\mathbf{x}}{\mathbf{x}} + \frac{\Delta\mathbf{s}}{\mathbf{s}}\right\| \leq 1 \end{array}\right\}(SF)$$

A célunk az, hogy olyan olyan skálázást (koordináta transzformációt) vezessünk be, amely után (SF) lineáris célfüggvényét egy gömbön kell minimalizálnunk.
Legyen

$$\mu := \frac{\mathbf{x}^T\mathbf{s}}{n}, \quad \mathbf{d} := \sqrt{\frac{\mathbf{x}}{\mathbf{s}}}, \quad \mathbf{u} := \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}}$$

Ekkor $\sqrt{\mathbf{x}\,\mathbf{s}} = \sqrt{\mu}\,\mathbf{u}$ és $\mathbf{d}^{-1}\mathbf{x} = \sqrt{\mu}\,\mathbf{u} = \mathbf{d}\,\mathbf{s}$. Vezessük be a $\mathbf{p}_x$ és $\mathbf{p}_s$ vektorokat úgy, hogy

$$\mathbf{d}^{-1}\Delta\mathbf{x} = \sqrt{\mu}\,\mathbf{p}_x \quad \text{és} \quad \mathbf{d}\,\Delta\mathbf{s} = \sqrt{\mu}\,\mathbf{p}_s$$

Ekkor

$$\frac{\Delta\mathbf{x}}{\mathbf{x}} = \frac{\mathbf{d}^{-1}\Delta\mathbf{x}}{\mathbf{d}^{-1}\mathbf{x}} = \frac{\sqrt{\mu}\,\mathbf{p}_x}{\sqrt{\mu}\,\mathbf{u}} = \frac{\mathbf{p}_x}{\mathbf{u}}$$

$$\text{és} \quad \frac{\Delta\mathbf{s}}{\mathbf{s}} = \frac{\mathbf{d}\,\Delta\mathbf{x}}{\mathbf{d}\,\mathbf{x}} = \frac{\sqrt{\mu}\,\mathbf{p}_s}{\sqrt{\mu}\,\mathbf{u}} = \frac{\mathbf{p}_s}{\mathbf{u}} \Rightarrow \frac{\Delta\mathbf{x}}{\mathbf{x}} + \frac{\Delta\mathbf{s}}{\mathbf{s}} = \frac{\mathbf{p}_x + \mathbf{p}_s}{\mathbf{u}}$$

Bevezetve a $\mathbf{p} := \mathbf{p}_x + \mathbf{p}_s$ jelölést

$$\frac{\mathbf{p}}{\mathbf{u}} = \frac{\Delta\mathbf{x}}{\mathbf{x}} + \frac{\Delta\mathbf{s}}{\mathbf{s}},$$

illetve $\mathbf{s}\,\Delta\mathbf{x}+\mathbf{x}\,\Delta\mathbf{s} = (\mathbf{s}\,\mathbf{d})(\mathbf{d}^{-1}\Delta\mathbf{x})+(\mathbf{x}\,\mathbf{d}^{-1})(\mathbf{d}\,\Delta\mathbf{s}) = \sqrt{\mu}\,\mathbf{u}\,\sqrt{\mu}\,\mathbf{p}_x+\sqrt{\mu}\,\mathbf{u}\,\sqrt{\mu}\,\mathbf{p}_s = \mu\,\mathbf{u}\,\mathbf{p}$

Tehát az (SF) átskálázott alakja rögzített μ paraméter mellett

$$\left.\begin{array}{r} \min\ \mu\,\mathbf{u}^T\mathbf{p} \\[2ex] \|\frac{\mathbf{p}}{\mathbf{u}}\| \leq 1 \end{array}\right\}$$

Legyen $\bar{\mathbf{p}} := \frac{\mathbf{p}}{\mathbf{u}}$. Ekkor az (SF) a következő alakot ölti:

$$\left.\begin{array}{r} \min\ \mu\,(\mathbf{u}^2)^T\bar{\mathbf{p}} \\[2ex] \|\bar{\mathbf{p}}\| \leq 1 \end{array}\right\}$$

ami gömb felett egy lineáris függvény minimalizázlása, tehát létezik egyértelmű minimuma:

$$\bar{\mathbf{p}} = -\frac{\mathbf{u}^2}{\|\mathbf{u}^2\|} \quad \text{és ekkor} \quad \mathbf{p} = -\frac{\mathbf{u}^3}{\|\mathbf{u}^2\|}$$

A $(\Delta\mathbf{x}, \Delta\mathbf{s})$ csökkenési irány meghatározása

Mivel $0 = -M\Delta\mathbf{x} + \Delta\mathbf{s} = -MD(D^{-1}\Delta\mathbf{x}) + D^{-1}(D\Delta\mathbf{s}) = -\sqrt{\mu}\,(MD\mathbf{p}_x - D^{-1}\mathbf{p}_s)$, ezért

$$\mathbf{p}_s = DMD\,\mathbf{p}_x \quad \text{és így a} \quad \mathbf{p} = \mathbf{p}_x + \mathbf{p}_s = (I + DMD)\,\mathbf{p}_x.$$

A 3.1.1. Lemma szerint $I + DMD$ pozitív definit mátrix, így

$$\mathbf{p}_x = (I + DMD)^{-1}\,\mathbf{p} \quad \text{és} \quad \mathbf{p}_s = DMD(I + DMD)^{-1}\,\mathbf{p}.$$

Figyelembe véve a $\mathbf{d}^{-1}\Delta\mathbf{x} = \sqrt{\mu}\,\mathbf{p}_x$ és $\mathbf{d}\Delta\mathbf{s} = \sqrt{\mu}\,\mathbf{p}_s$ összefüggéseket

$$\Delta\mathbf{x} = \sqrt{\mu}\,D(I + DMD)^{-1}\,\mathbf{p} \quad \text{és} \quad \Delta\mathbf{s} = \sqrt{\mu}\,MD(I + DMD)^{-1}\,\mathbf{p}$$

adódik. Az így kapott $(\Delta\mathbf{x}, \Delta\mathbf{s})$ vektort *Dikin irány*nak nevezzük.
Ezt meglépve a dualitásrés csökkenése:

$$\mu\,(\mathbf{u}^2)^T\bar{\mathbf{p}} = \mu\,(\mathbf{u}^2)^T\left(-\frac{\mathbf{u}^2}{\|\mathbf{u}^2\|}\right) = -\mu\,\frac{\|\mathbf{u}^2\|^2}{\|\mathbf{u}^2\|} = -\mu\,\|\mathbf{u}^2\| = -\mu\,\left\|\frac{\mathbf{x}\,\mathbf{s}}{\mu}\right\| = -\|\mathbf{x}\,\mathbf{s}\|.$$

A $\Delta\mathbf{x}$ vektorra kapott kifejezésbe behelyettesítve $\mathbf{d}$, $\mathbf{p}$, illetve $\mathbf{p}$ képletébe $\mathbf{u}$ definícióját, könnyen adódik a következő állítás.

3.1.3. Állítás. *A fenti jelölésekkel*

$$\Delta\mathbf{x} = \sqrt{\mu}\, D(I + DMD)^{-1}\mathbf{p} = -(S + XM)^{-1}\frac{\mathbf{x}^2\,\mathbf{s}^2}{\|\mathbf{x}\,\mathbf{s}\|} \qquad \textit{teljesül.}$$

□

Az eddigieket összefoglalva, az algoritmusunk a következő:

Dikin–féle affinskálázású algoritmus

Input:

megállási paraméter $\varepsilon > 0$

lépés paraméter α, $0 < \alpha \le 1$

$\mathbf{x}^0 \in \mathcal{F}^0 : s(\mathbf{x}^0) > \mathbf{0}$.

begin

Legyen $\mathbf{x} := \mathbf{x}^0$, $\quad \mathbf{s} := s(\mathbf{x})$.

while $\mathbf{x}^T\mathbf{s} \ge \varepsilon$ **do**

begin

$$\Delta\mathbf{x} := -(S + XM)^{-1}\frac{\mathbf{x}^2\mathbf{s}^2}{\|\mathbf{x}\,\mathbf{s}\|}$$

$\mathbf{x} := \mathbf{x} + \alpha\Delta\mathbf{x}$, $\mathbf{s} := s(\mathbf{x})$

end

end

end.

Az algoritmusunk elemzéséhez definiáljuk a következő centralitási mértéket:

$$\delta(\mathbf{x}) = \frac{\max(\mathbf{x}\cdot s(\mathbf{x}))}{\min(\mathbf{x}\cdot s(\mathbf{x}))} = \frac{\max \mathbf{u}^2}{\min \mathbf{u}^2}, \qquad \mathbf{x} \in \mathcal{F}^0$$

Ekkor $\delta(\mathbf{x}) \ge 1$ teljesül bármely $\mathbf{x} \in \mathcal{F}^0$, $\mathbf{s} = s(\mathbf{x}) > \mathbf{0}$ esetén, valamint $\delta(\mathbf{x}) = 1$ pontosan akkor, ha $\mathbf{x} \in \mathcal{F}^0 :\ \mathbf{x}\cdot s(\mathbf{x}) = \mu\mathbf{e}$.

Nevezzünk egy α *lépéshosszt megengedett*nek, ha az új pontunkra is teljesül a pozitivitási megkötés, azaz $\mathbf{x} + \alpha\Delta\mathbf{x} > 0$, $\mathbf{s} + \alpha\Delta\mathbf{s} > 0$.

3.1.4. Lemma. *Ha az α lépéshossz megengedett, akkor*

$$(\mathbf{x}^+)^T\mathbf{s}^+ \leq \left(1 - \frac{\alpha}{\sqrt{n}}\right)\mathbf{x}^T\mathbf{s}.$$

Bizonyítás. Mivel az

$$\mathbf{x}^+ = \mathbf{x} + \alpha\,\Delta\mathbf{x} = \mathbf{d}\,(\mathbf{d}^{-1}\,\mathbf{x}) + \alpha\,\mathbf{d}\,(\mathbf{d}^{-1}\,\Delta\mathbf{x}) = \mathbf{d}\,\sqrt{\mu}\,\mathbf{u} + \alpha\,\mathbf{d}\,\sqrt{\mu}\,\mathbf{p}_x = \sqrt{\mu}\,\mathbf{d}\,(\mathbf{u} + \alpha\,\mathbf{p}_x)$$

és

$$\mathbf{s}^+ = \mathbf{s} + \alpha\,\Delta\mathbf{s} = \sqrt{\mu}\,\mathbf{d}^{-1}(\mathbf{u} + \alpha\,\mathbf{p}_s),$$

ezért

$$\begin{aligned}(\mathbf{x}^+)\,s(\mathbf{x}^+) &= (\mathbf{x} + \alpha\,\Delta\mathbf{x})\,(\mathbf{s} + \alpha\,\Delta\mathbf{s}) = \mu\,(\mathbf{u} + \alpha\,\mathbf{p}_x)(\mathbf{u} + \alpha\,\mathbf{p}_s)\\ &= \mu\,(\mathbf{u}^2 + \alpha\,(\mathbf{p}_x + \mathbf{p}_s)\,\mathbf{u} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s).\end{aligned}$$

Figyelembe véve a

$$\mathbf{p} = \mathbf{p}_x + \mathbf{p}_s = -\frac{\mathbf{u}^3}{\|\mathbf{u}^2\|}$$

összefüggést

$$\mathbf{x}^+s(\mathbf{x}^+) = \mu\left(\mathbf{u}^2 - \alpha\,\frac{\mathbf{u}^4}{\|\mathbf{u}^2\|} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s\right) \tag{3.1}$$

adódik. Mivel az α megengedett lépéshossz, ezért

$$\mathbf{u}^2 - \alpha\frac{\mathbf{u}^4}{\|\mathbf{u}^2\|} + \alpha^2\mathbf{p}_x\mathbf{p}_s > \mathbf{0}.$$

Az M mátrix ferdén szimmetrikus, ezért $(\Delta\mathbf{x})^T\Delta\mathbf{s} = 0$ azaz

$$\mathbf{p}_x^T\mathbf{p}_s = \mathbf{e}^T(\mathbf{d}^{-1}\,\Delta\mathbf{x}\,\mathbf{d}\,\Delta\mathbf{s})\frac{1}{\mu} = \frac{1}{\mu}\,(\Delta\mathbf{x})^T\Delta\mathbf{s} = 0. \tag{3.2}$$

Számítsuk ki az új dualitás résnek és a μ paraméternek a hányadosát

$$\frac{(\mathbf{x}^+)^T\mathbf{s}^+}{\mu} = \mathbf{e}^T\mathbf{u}^2 - \alpha\,\frac{\mathbf{e}^T\mathbf{u}^4}{\|\mathbf{u}^2\|} + \alpha^2\,\mathbf{p}_x^T\mathbf{p}_s = \|\mathbf{u}\|^2 - \alpha\|\mathbf{u}^2\|,$$

ahol figyelembe vettük a $\mathbf{p}_x$ és $\mathbf{p}_s$ vektorok ortogonalitását és az $\mathbf{e}^T\mathbf{u}^4 = (\mathbf{u}^2)^T\mathbf{u}^2 = \|\mathbf{u}^2\|^2$ összefüggést. A Cauchy-Schwarz egyenlőtlenség alapján

$$\|\mathbf{u}^2\| = \frac{1}{\sqrt{n}}\|\mathbf{e}\|\,\|\mathbf{u}^2\| \geq \frac{\mathbf{e}^T\mathbf{u}^2}{\sqrt{n}} = \frac{\|\mathbf{u}\|^2}{\sqrt{n}},$$

behelyettesítve

$$(\mathbf{x}^+)^T\mathbf{s}^+ \le \mu\left(1-\frac{\alpha}{\sqrt{n}}\right)\|\mathbf{u}\|^2 = \left(1-\frac{\alpha}{\sqrt{n}}\right)\mathbf{x}^T\mathbf{s}$$

adódik. □

Legyen $\tau > 1, \tau \in \mathbb{R}$ és $\mathbf{x} \in \mathcal{F} : \delta(\mathbf{x}) \le \tau$. Továbbá legyen $\mathbf{u} := \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}}$ ekkor

$$\delta(\mathbf{x}) = \frac{\max \mathbf{u}^2}{\min \mathbf{u}^2} \le \tau \iff \max \mathbf{u}^2 \le \tau \min \mathbf{u}^2$$

összefüggésből következik, hogy létezik $\tau_1, \tau_2 > 0$ és $\tau_2 = \tau\tau_1$, amelyekre

$$\tau_1\mathbf{e} \le \mathbf{u}^2 \le \tau_2\mathbf{e}.$$

3.1.5. Állítás.

$$\|\mathbf{p}_x\,\mathbf{p}_s\|_\infty \le \frac{1}{4}\|\mathbf{p}_x + \mathbf{p}_s\|^2 = \frac{\|\mathbf{p}\|^2}{4}.$$

Bizonyítás. Írjuk fel a $\mathbf{p}_x\mathbf{p}_s$ szorzatot a következő alakban

$$\mathbf{p}_x\mathbf{p}_s = \frac{1}{4}\left((\mathbf{p}_x + \mathbf{p}_s)^2 - (\mathbf{p}_x - \mathbf{p}_s)^2\right),$$

ekkor triviális alsó és felső becslések nyerhetők

$$-\frac{1}{4}(\mathbf{p}_x - \mathbf{p}_s)^2 \le \mathbf{p}_x\mathbf{p}_s \le \frac{1}{4}(\mathbf{p}_x + \mathbf{p}_s)^2.$$

Az ortogonalitási tulajdonság (3.2) miatt $\|\mathbf{p}_x + \mathbf{p}_s\|^2 = \|\mathbf{p}_x - \mathbf{p}_s\|^2$. A

$$\mathbf{p}_x\mathbf{p}_s \le \frac{1}{4}\|\mathbf{p}_x + \mathbf{p}_s\|^2\mathbf{e},$$

hiszen bármely i indexre igaz a következő egyenlőtlenség

$$(\mathbf{p}_x)_i(\mathbf{p}_s)_i \le \frac{1}{4}((\mathbf{p}_x)_i + (\mathbf{p}_s)_i)^2,$$

tehát

$$\max_i\left((\mathbf{p}_x)_i(\mathbf{p}_s)_i\right) \le \frac{1}{4}\sum_{i=1}^{n}((\mathbf{p}_x)_i + (\mathbf{p}_s)_i)^2$$

teljesül, ami így is írható

$$\|\mathbf{p}_x\mathbf{p}_s\|_\infty \le \|\frac{1}{4}\|\mathbf{p}_x + \mathbf{p}_s\|^2\mathbf{e}\|_\infty = \frac{1}{4}\|\mathbf{p}_x + \mathbf{p}_s\|^2 = \frac{\|\mathbf{p}\|^2}{4},$$

ahol az utolsó előtti egyenlőség a végtelen norma definíciója miatt teljesül, míg az utolsó a $\mathbf{p}$ vektor értelmezése miatt igaz. □

Az iterációk során csak egy tompított Dikin lépést teszünk meg. A lépéshossz megfelelő megválasztásával biztosítjuk az új pontunk pozitivitását, így a megengedettségét, valamint a pontunkat nem engedjük ki a centrális út τ környezetéből. Ezen elvárásoknak megfelelő α lépéshosszra ad elégséges feltételt a következő lemma.

3.1.6. Lemma. *Legyen $\tau \in \mathbb{R}$, $\tau > 1$. Tegyük fel, hogy $\mathbf{x} > \mathbf{0}$, $\mathbf{s} := s(\mathbf{x}) > \mathbf{0}$, $\mathbf{x} \in \mathcal{F}$ és $\delta(\mathbf{x}) \le \tau$. Ekkor bármely α lépéshossz amely kielégíti az*

$$\alpha \le \frac{\|\mathbf{u}^2\|}{2\tau_2} \quad \text{\textit{és}} \quad \alpha < \frac{4\tau_1}{\|\mathbf{u}^2\|}$$

egyenlőtlenségeket, megengedett lépéshossz és az $\mathbf{x}^+ := \mathbf{x} + \alpha\Delta\mathbf{x}$ vektorra $\delta(\mathbf{x}^+) \le \tau$.

Bizonyítás. Az (3.1) összefüggésből, a baloldalon zárójelben lévő első két tagjában az $\mathbf{u}^2$ vektor helyett egy t változót írva az

$$f: \quad t \mapsto t - \alpha\frac{t^2}{\|\mathbf{u}^2\|} \tag{3.3}$$

függvényt nyerjük. Az α értékére kirótt első felső korlát alapján a (3.3) függvény növekvő a $t \in [0, \tau_2]$ intervallumon, mert az

$$f'(t) = 1 - 2\alpha\frac{t}{\|\mathbf{u}^2\|} \ge 0 \quad \text{egyenértékű az} \quad \alpha \le \frac{\|\mathbf{u}^2\|}{2t}$$

feltétellel, amely igaz a lemmában szereplő első felső korlát miatt.

Alkalmazzuk ezt a leképezést az $\mathbf{u}^2$ vektor minden komponensére, ekkor

$$\left(\tau_1 - \alpha\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} \le \mathbf{u}^2 - \alpha\frac{\mathbf{u}^4}{\|\mathbf{u}^2\|} \le \left(\tau_2 - \alpha\frac{\tau_2^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} \tag{3.4}$$

és így az előző egyenlőtlenség és a (3.1) alapján

$$\mu\left(\left(\tau_1 - \alpha\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s\right) \le \mathbf{x}^+ s(\mathbf{x})^+ \le \mu\left(\tau_2 - \alpha\frac{\tau_2^2}{\|\mathbf{u}^2\|} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s\right),$$

alsó és felső korlátot kapjuk.

Nyiván, ha az α megengedett lépéshossz, akkor $\delta(\mathbf{x}^+) \leq \tau$, ha

$$\tau\left(\left(\tau_1 - \alpha\,\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s\right) > \left(\tau_2 - \alpha\,\frac{\tau_2^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s \tag{3.5}$$

Másfelől, ha (3.5) teljesül, akkor az α nyilván megengedett lépéshossz. A (3.4) becslés miatt viszont

$$\tau\left(\left(\tau_1 - \alpha\,\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s\right) > \left(\tau_1 - \alpha\,\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s$$

adódik, és mivel $\tau > 1$, ezért

$$\left(\tau_1 - \alpha\,\frac{\tau_1^2}{\|\mathbf{u}^2\|}\right)\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s > \mathbf{0},$$

azaz $\mathbf{x}^+ s(\mathbf{x}^+)$ nem tünhet el egyetlen olyan α értékre sem, amely kielégíti a korlátokat, és ekkor α megengedett lépéshossz. Tehát azt kell megmutatni, hogy az α kielégíti a (3.5) összefüggést. Azaz

$$\tau\,\tau_1\mathbf{e} - \alpha\,\frac{\tau\tau_1^2}{\|\mathbf{u}^2\|}\mathbf{e} + \tau\alpha^2\mathbf{p}_x\mathbf{p}_s > \tau_2\mathbf{e} - \alpha\,\frac{\tau_2^2}{\|u^2\|}\mathbf{e} + \alpha^2\,\mathbf{p}_x\,\mathbf{p}_s,$$

teljesül, mivel $\tau_2 = \tau\tau_1$, ezért egyszerűsíteni lehet a kifejezést, és $\alpha > 0$ miatt α számmal elosztva az egyenlőtlenség mindkét oldalát a reláció iránya nem változik meg. Átrendezés után a következő egyenlőtlenséget nyerjük

$$\frac{\tau_2^2 - \tau\,\tau_1^2}{\|\mathbf{u}^2\|}\mathbf{e} + \alpha\,(\tau - 1)\mathbf{p}_x\,\mathbf{p}_s > \mathbf{0}. \tag{3.6}$$

A baloldali kifejezés első tagjának a számlálóját átalakíthajuk az alábbi módon

$$\tau_2^2 - \tau\,\tau_1^2 = \tau_2^2 - \tau_1\,\tau_2 = \left(\frac{\tau_2}{\tau_1} - 1\right)\tau_1\,\tau_2 = (\tau - 1)\,\tau_1\,\tau_2,$$

ahol az első és az utolsó azonosságnál a τ definícióját használtuk. Ekkor a (3.6) egyenlőtlenséget eloszthatjuk $(\tau - 1)$ kifejezéssel és az egyenlőtlenség iránya nem változik meg, mert $\tau > 1$

$$\frac{\tau_1\,\tau_2\mathbf{e}}{\|\mathbf{u}^2\|} + \alpha\,\mathbf{p}_x\,\mathbf{p}_s > \mathbf{0}. \tag{3.7}$$

Mivel a $\mathbf{p}_x^T\mathbf{p}_s = 0$, ezért a $\mathbf{p}_x\,\mathbf{p}_s$ vektor összes koordinátája nem lehet negatív. A 3.1.5. Állítás miatt

$$\|\mathbf{p}_x\,\mathbf{p}_s\|_\infty \leq \frac{1}{4}\|\mathbf{p}_x + \mathbf{p}_s\|^2 = \frac{\|\mathbf{p}\|^2}{4}, \tag{3.8}$$

illetve

$$\|\mathbf{p}\| = \left\| \frac{\mathbf{u}^3}{\|\mathbf{u}^2\|} \right\| \leq \|\mathbf{u}\|_\infty \left\| \frac{\mathbf{u}^2}{\|\mathbf{u}^2\|} \right\| = \|\mathbf{u}\|_\infty \leq \sqrt{\tau_2}. \tag{3.9}$$

A (3.7) egyenlőtlenségben szereplő vektorok végtelen normáját véve, továbbá alkalmazva a (3.8) és (3.9) becsléseket először az alábbi (szigorúbb) egyenlőtlenséget

$$\frac{\tau_1\,\tau_2}{\|\mathbf{u}^2\|} - \frac{\alpha\,\|\mathbf{p}\|^2}{4} > 0,$$

majd pedig a következő összefüggést kapjuk

$$\frac{\tau_1\,\tau_2}{\|\mathbf{u}^2\|} - \frac{\alpha\,\tau_2}{4} > 0,$$

amelyik ekvivalens az α értékére adott második felső korláttal. □

Az előző lemma szerint az $\alpha = 1/(\tau\sqrt{n})$ választás[1] megfelelő, azaz ebben az esetben az algoritmus jól definiált. Az alábbi tételben megmutatjuk, hogy ezen érték mellett az algoritmus polinomiális.

3.1.7. Tétel. *Legyen* $\tau := \max(2, \delta(\mathbf{x}^0))$ *és* $\alpha = 1/(\tau\sqrt{n})$*. Ha* $n \geq 2$ *akkor a Dikin–féle affin skálázású algoritmus a ferdén szimmetrikus lineáris programozási feladatot legfeljebb*

$$\left\lceil \tau n \log \frac{\mathbf{q}^T\mathbf{x}^0}{\varepsilon} \right\rceil$$

lépésben megoldja. A megoldás $\mathbf{x} \in \mathcal{F}$*, amelyre* $\delta(\mathbf{x}) \leq \tau$ *és* $\mathbf{q}^T\mathbf{x} \leq \varepsilon$*.*

Bizonyítás. Az $\mathbf{x}^0 \in \mathcal{F}^0$ induló megoldásra $\delta(\mathbf{x}^0) \leq \tau$ összefüggés teljesült. Az 3.1.6. Lemma alapján a lépéshossz megválasztható úgy, hogy minden egyes lépésben az új megoldás pozitív, megengedett megoldás legyen és a centralitás mértéke τ értékénél kisebb maradjon. Azt kell megmutatni, hogy a tétel feltételeiben megadott α érték kielégíti az 3.1.6. Lemmában adott korlátokat. Mivel $n \geq 2$ ezért

$$\alpha = \frac{1}{\tau\sqrt{n}} = \frac{\tau_1}{\tau_2\sqrt{n}} \leq \frac{\tau_1\sqrt{n}}{2\,\tau_2} = \frac{\|\tau_1\mathbf{e}\|}{2\,\tau_2} \leq \frac{\|\mathbf{u}^2\|}{2\,\tau_2},$$

ahol felhasználtuk azt is, hogy $\mathbf{0} \leq \tau_1\mathbf{e} \leq \mathbf{u}^2$. Az $\|\mathbf{u}^2\| \leq \tau_2\sqrt{n}$ miatt

$$\frac{4\,\tau_1}{\|\mathbf{u}^2\|} \geq \frac{4\,\tau_1}{\tau_2\sqrt{n}} = \frac{4}{\tau\sqrt{n}} > \alpha$$

[1] Mivel az α paraméter függ az n számtól, azaz a feladat dimenziójától, rövidlépéses algoritmusról beszélunk.

adódik, azaz a tétel feltételeiben megadott lépéshossz megengedett, mert teljesíti az 3.1.6. Lemmában adott korlátokat.

Az induló megoldásnál a célfüggvény értéke $\mathbf{q}^T\mathbf{x}^0$ volt. (Könnyen megmutatható, hogy $\mathbf{q}^T\mathbf{x}^0 = (\mathbf{x}^0)^T s(\mathbf{x}^0)$, ami pedig a dualitás rés.) A 3.1.4. Lemma alapján a dualitás rés minden lépésben $(1 - 1/(n\tau))$ nagyságrenddel csökken. Így k lépés után a célfüggvény értéke ε alá csökken, azaz

$$\left(1 - \frac{1}{n\tau}\right)^k \mathbf{q}^T\mathbf{x}^0 \leq \varepsilon.$$

Szeretnénk meghatározni a k értékét. Vegyük a logaritmusát az előző kifejezésnek, ekkor

$$k \log\left(1 - \frac{1}{n\tau}\right) + \log(\mathbf{q}^T\mathbf{x}^0) \leq \log \varepsilon.$$

Átrendezés után

$$\log \frac{\mathbf{q}^T\mathbf{x}^0}{\varepsilon} \leq -k \log\left(1 - \frac{1}{n\tau}\right)$$

adódik. Mivel $-\log(1-t) \geq t$, ezért a következő egyenlőtlenség erősebb feltétel a fentinél

$$\frac{k}{n\tau} \geq \log \frac{\mathbf{q}^T\mathbf{x}^0}{\varepsilon},$$

ebből pedig már következik a tétel állítása. $\square$

3.2. A (B,N) partíció meghatározása a centrális út egy környezetében

Legyen adott $\mathbf{x}^0 \in \mathcal{F}^0$, $\mathbf{s}^0 > 0$, melyek a $\mu^0 = 1$ paraméterhez tartoznak, azaz $(\mathbf{x}^0)^T\mathbf{s}^0 = \mathbf{q}^T\mathbf{x}^0 = n\mu^0 = n$. Legyen $\tau = 2$, ekkor a 3.1.7. Tétel alapján a Dikin lépéses algoritmus $\lceil 2n \log \frac{n}{\varepsilon} \rceil$ iterációban előállít egy $\mathbf{x} \in \mathcal{F}^0$ pontot, melyre $\delta(\mathbf{x}) \leq 2$ és $\mathbf{q}^T\mathbf{x} \leq \varepsilon$.

3.2.1. Lemma. *Legyen $\mathbf{x} \in \mathcal{F}^0$, $\mathbf{s} > 0$, melyekre $\delta(\mathbf{x}) \leq \tau$. Ekkor*

$$\begin{aligned} x_i &\geq \frac{\sigma}{\tau n} \quad i \in B, \qquad & x_i &\leq \frac{\mathbf{x}^T\mathbf{s}}{\sigma} \quad i \in N, \\ s_i &\leq \frac{\mathbf{x}^T\mathbf{s}}{\sigma} \quad i \in B, \qquad & s_i &\geq \frac{\sigma}{\tau n} \quad i \in N, \end{aligned}$$

ahol σ a feladat kondíciószáma.

Bizonyítás. $\tau > 1$, tehát létezik $\tau_1, \tau_2 > 0$, hogy $\tau_2 = \tau\tau_1$ és $\tau_1 \geq x_i s_i \geq \tau_2$ bármely i indexre. Legyen $\tilde{\mathbf{x}} \in \mathcal{F}^*$, $\tilde{\mathbf{s}} > 0$ olyan, amelyre $\tilde{x}_i$ maximális, azaz $\tilde{x}_i \geq \sigma > 0$, ezért $i \in B$. Felhasználva az $\tilde{\mathbf{x}}$ komplementaritását és az $(\mathbf{x} - \tilde{\mathbf{x}})^T(\mathbf{s} - \tilde{\mathbf{s}}) = 0$ egyenlőséget, adódik, hogy $\tilde{\mathbf{x}}^T\mathbf{s} + \mathbf{x}^T\tilde{\mathbf{s}} = \mathbf{x}^T\mathbf{s}$. Figyelembe véve a vektorok nemnegativitását kapjuk a lemma harmadik állítását.

$$s_i\tilde{x}_i \leq \mathbf{s}^T\tilde{\mathbf{x}} \leq \tilde{\mathbf{x}}^T\mathbf{s} + \mathbf{x}^T\tilde{\mathbf{s}} = \mathbf{x}^T\mathbf{s}, \quad \text{ezért} \quad s_i \leq \frac{\mathbf{x}^T\mathbf{s}}{\tilde{x}_i} \leq \frac{\mathbf{x}^T\mathbf{s}}{\sigma}.$$

Ezt felhasználva egyszerűen adódik az első egyenlőtlenség.

$$\tau_1 \leq x_i s_i \leq x_i\frac{\mathbf{x}^T\mathbf{s}}{\tilde{x}_i} \leq x_i\frac{n\tau_2}{\sigma}, \quad \text{azaz} \quad x_i \geq \frac{\tau_1\sigma}{n\tau_2} = \frac{\sigma}{n\tau}.$$

A második és a negyedik állítás hasonlóan bizonyítható. □

Megfelelő feltétel mellett a 3.2.1. Lemma segítségével az $\mathbf{x}$ illetve az $\mathbf{s}$ vektor koordinátáit nagyság szerint szétválogatva meghatározhatjuk a (B,N) partíciót.

3.2.2. Lemma. *Legyen* $\mathbf{x} \in \mathcal{F}^0$, $\mathbf{s} > 0$, *melyekre* $\delta(\mathbf{x}) \leq \tau$. *Ha* $\mathbf{x}^T\mathbf{s} \leq \frac{\sigma^2}{\tau n}$, *akkor a* (B,N) *partíciója az* (SP) *feladatnak a következő módon kapható meg:*

$$B := \{i \,:\, x_i > s_i\} \qquad \text{és} \qquad N := \{i \,:\, x_i < s_i\}.$$

Bizonyítás. A kicsi és nagy változókat szét tudjuk választani, ha $\frac{\mathbf{x}^T\mathbf{s}}{\sigma} < \frac{\sigma}{\tau n}$, azaz $\mathbf{x}^T\mathbf{s} < \frac{\sigma^2}{\tau n}$. Tehát az előző lemma szerint a fent megadott B és N halmazok jól definiáltak, és az indexhalmaz partícióját adják. □

A 3.2.2. Lemma feltételeit kielégítő $(\mathbf{x}, \mathbf{s})$ pontpár a 3.1.7. Tétel szerint $\lceil 2n \log \frac{2n^2}{\sigma^2} \rceil$ lépésben a Dikin algoritmussal előállítható, ennek tükrében az alábbi következmény könnyen adódik.

3.2.3. Következmény. *Legyen* $\mathbf{x}^0 \in \mathcal{F}^0$, $\mathbf{s}^0 > 0$, *melyekre* $\delta(\mathbf{x}^0) \leq \tau$ *és* $(\mathbf{x}^0)^T\mathbf{s}^0 = n$, $\tau = 2$. *Ekkor a Dikin lépéses algoritmus* $\lceil 2n \log \frac{2n^2}{\sigma^2} \rceil$ *lépésben megtalál egy olyan* $\mathbf{x} \in \mathcal{F}^0$ *megoldást, amely esetén* (B,N) *partíciót meg tudjuk határozni.* □

Mint már az első fejezetben említettük az $\mathbf{x} = \mathbf{0}$ mindig optimális megoldása az (SP) feladatnak. A következőkben megmutatjuk, hogy ez akkor és csak akkor az egyetlen optimális vektor, ha a B indexhalmaz üres. (Ezzel a speciális, triviális esettel a továbbiakban nem foglalkozunk.)

3.2.4. Tétel. *Legyen $\mathcal{F}^0 \neq \emptyset$. Ekkor $\mathbf{x} = \mathbf{0}$ az (SP) feladat egyetlen optimális megoldása akkor és csak akkor, ha $\mathbf{q} > 0$. Ez az eset pontosan akkor fordul elő, ha $B = \emptyset$.*

Bizonyítás. Ha $B = \emptyset$, akkor $\mathbf{x} = \mathbf{0}$ optimális, sőt szigorúan komplementáris is az $\mathcal{F}^0 \neq \emptyset$ mellett, azaz $s(\mathbf{x}) = \mathbf{q} > 0$ kell hogy teljesüljön.
Másfelől $s(\mathbf{0}) = \mathbf{q}$ és $s(\mathbf{0}) > 0$ miatt $\mathbf{q} > 0$, így $\mathbf{x} = \mathbf{0}$ egyértelmű optimum. □

Vizsgáljuk meg az $\mathbf{x} \in \mathcal{F}^0$, $s(\mathbf{x}) = \mathbf{s} > 0$ ε-optimális megoldást. Ha az ε kellően kicsi, akkor a (B, N) partíció ismert, így a rendszerünk a következő alakú:

$$\begin{pmatrix} \mathbf{s}_B \\ \mathbf{s}_N \end{pmatrix} = \begin{pmatrix} M_{BB} & M_{BN} \\ M_{NB} & M_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} + \begin{pmatrix} \mathbf{q}_B \\ \mathbf{q}_N \end{pmatrix}$$

Az előző tétel szerint a $B = \emptyset$ eset triviális, ezért tegyük fel, hogy $B \neq \emptyset$, ekkor az optimalitás miatt $\mathbf{q}_B = \mathbf{0}$.

$$\mathbf{s}_B = M_{BB}\mathbf{x}_B + M_{BN}\mathbf{x}_N + \mathbf{q}_B, \quad \text{tehát}$$

$$M_{BB}\mathbf{x}_B = \mathbf{s}_B - M_{BN}\mathbf{x}_N. \tag{3.10}$$

Ha $\tilde{\mathbf{x}} \in \mathcal{F}^*$, $\tilde{\mathbf{s}} = s(\tilde{\mathbf{x}})$, akkor $\tilde{\mathbf{x}}_N = \mathbf{0}$, $\tilde{\mathbf{s}}_B = \mathbf{0}$, és a szigorú komplementraitás miatt $\tilde{\mathbf{x}}_B > \mathbf{0}$, valamint (3.10) egyenlőség miatt $M_{BB}\tilde{\mathbf{x}}_B = \mathbf{0}$, azaz az M_{BB} mátrix szinguláris. Ennek következtében a

$$M_{BB}\xi = \mathbf{s}_B - M_{BN}\mathbf{x}_N \tag{3.11}$$

egyenlet megoldása nem egyértelmű és egy lehetséges megoldása a $\xi = \mathbf{x}_B$. Legyen ξ tetszőleges megoldása a (3.11) egyenletnek, ekkor definiálhatjuk, az $\overline{\mathbf{x}}_B = \mathbf{x}_B - \xi$, $\overline{\mathbf{x}}_N = \mathbf{0}$ megoldást és $\overline{\mathbf{s}} = s(\overline{\mathbf{x}})$ eltérésváltozót, ahol $\overline{\mathbf{s}}_B = M_{BB}\overline{\mathbf{x}}_B + M_{BN}\overline{\mathbf{x}}_N + \mathbf{q}_B = M_{BB}(\mathbf{x}_B - \xi) = \mathbf{0}$. Tehát az $(\overline{\mathbf{x}}, \overline{\mathbf{s}})$ komplementáris megoldás, de nem feltétlenül pozitív vektorok. Az $(\overline{\mathbf{x}}, \overline{\mathbf{s}})$ pozitivitásához szükséges:

$$\overline{\mathbf{x}}_B = \mathbf{x}_B - \xi > 0 \ , \quad \text{illetve}$$

$$\overline{\mathbf{s}}_N = M_{NB}\overline{\mathbf{x}}_B + M_{NN}\overline{\mathbf{x}}_N + \mathbf{q}_N = M_{NB}\mathbf{x}_B - M_{NB}\xi + \mathbf{q}_N = \mathbf{s}_N - M_{NN}\mathbf{x}_N - M_{NB}\xi > 0$$

Tehát ε megfelelő értéke mellett elő tudjuk állítani az $(\overline{\mathbf{x}}, \overline{\mathbf{s}})$ vektort, ami az előző feltételek teljesülése mellett szigorúan komplementáris megoldás lesz. A következő részben ezen megoldás előállításával foglalkozunk.

3.3. Szigorúan komplementáris megoldás előállítása

Vezessük be a következő jelöléseket:

$$w := \|M\|_\infty\,, \qquad B^* := \{i \in B \;:\; M_{Bi} \text{ oszlopvektor nem azonosan nulla }\}.$$

Definiáljuk a $\pi_B \in \mathbb{N}$ számot a következő módon

$$\pi_B := \begin{cases} 1, & \text{ha } B^* = \emptyset \\ \prod\limits_{j \in B^*} \|M_{Bj}\|, & \text{különben} \end{cases}$$

3.3.1. Lemma. *Legyen adott egy egész együtthatós (SP) feladat. Legyen $\mathbf{x} \in \mathcal{F}^0, \mathbf{s} > 0$, melyekre $\delta(\mathbf{x}) \le \tau = 2$. Ha $\mathbf{x}^T\mathbf{s} \le \varepsilon$, ahol $\varepsilon = \frac{\sigma^2}{6nw^2\sqrt{|B|\pi_B}}$, akkor a kerekítési eljárás $O(|B^*|^3)$ aritmetikai művelettel előállítja az (SP) feladat egy szigorúan komplementáris megoldását.*

Bizonyítás. A definíciókból közvetlenül következik, hogy w, π_B, $|B| \ge 1$. A pontunk a centrális út τ környezetében van, valamint

$$\varepsilon = \frac{\sigma^2}{6nw^2\sqrt{|B|\pi_B}} < \frac{\sigma^2}{2n} = \frac{\sigma^2}{\tau n},$$

azaz az ε elegendően kicsi ahhoz, hogy a (B, N) partíciót meg tudjuk határozni. Mivel a partíció ismert, az $M_{BB}\xi = \mathbf{s}_B - M_{BN}\mathbf{x}_N$ egyenletet fel tudjuk írni. Megmutatjuk, hogy létezik olyan ξ megoldása, amelyre $\mathbf{x}_B - \xi > 0$ és $\overline{\mathbf{s}}_N = \mathbf{s}_N - M_{NN}\mathbf{x}_N - M_{NB}\xi > 0$ teljesül. Két eset lehetséges:
(i) $M_{BB} = \mathbf{0}$, ekkor $\xi = 0$, azaz $\mathbf{s}_N - MNN\mathbf{x}_N > 0$ kell hogy teljesüljön. Ha $M_{NN} = 0$, akkor a feltétel mindig igaz. Ellenkező esetben ($M_{NN} \ne 0$) elegendő az $\mathbf{s}_N \ge \frac{\sigma}{2n}\mathbf{e} > M_{NN}\mathbf{x}_N$ egyenlőtlenséget belátni. Az első része az N halmaz definíciója miatt teljesül. A végtelen norma tulajdonsága miatt

$$\|M_{NN}\mathbf{x}_N\|_\infty \le \|M_{NN}\|_\infty \|\mathbf{x}_N\|_\infty \le w\frac{\varepsilon}{\sigma}$$

A lemmában megadottak szerint $\varepsilon < \frac{\sigma^2}{2nw}$, tehát $w\frac{\varepsilon}{\sigma} < \frac{\sigma}{2n}$, azaz az egyenlőtlenség második fele is igaz.
(ii) $M_{BB} \ne 0$. Oldjuk meg az $M_{BB}\xi = \mathbf{s}_B - M_{BN}\mathbf{x}_N$ egyenletet Gauss–Jordan eliminációval (ez a kerekítési eljárás). Ha M_{BB} szinguláris mátrix, akkor határozzuk

meg az M_{BB} maximális reguláris részmátrixát, jelölje $M_{B_1B_2}$. Majd ezek után az $M_{B_1B_2}\rho = \mathbf{s}_{B_1} - M_{B_1N}\mathbf{x}_N$ egyenlet megoldását keressük. Bármely $i \in B$ esetén a ρ_i értékét a Cramer–szabállyal is kiszámíthatjuk, azaz

$$\rho_i = \frac{\det(M^{(i)}_{B_1B_2})}{\det(M_{B_1B_2})},$$

ahol $M^{(i)}_{B_1B_2}$ az a mátrix, amelyet úgy kapunk, hogy az $M_{B_1B_2}$ mátrix i. oszlopát a jobb oldal vektorrával kicseréljük. Mivel az $M_{B_1B_2}$ mátrix egész számokból áll, ezért $|\det(M_{B_1B_2})| \geq 1$. A Hadamard egyenlőtlenséget használva

$$|\rho_i| \leq |\det(M^{(i)}_{B_1B_2})| \leq \|\mathbf{s}_{B_1} - M_{B_1N}\mathbf{x}_N\| \prod_{j\in B_2\setminus\{i\}} \|M_{Bj}\| \leq \|\mathbf{s}_{B_1} - M_{B_1N}\mathbf{x}_N\|\pi_B$$

Felhasználva a normák tulajdonságait és az eddigi eredményeket további becsléseket adhatunk

$$\begin{aligned}\|\mathbf{s}_B - M_{BN}\mathbf{x}_N\| &\leq \sqrt{|B|}\|\mathbf{s}_B - M_{BN}\mathbf{x}_N\|_\infty \leq \sqrt{|B|}\left(\|\mathbf{s}_B\|_\infty + \|M_{BN}\mathbf{x}_N\|_\infty\right)\\ &\leq \sqrt{|B|}\left(\frac{\varepsilon}{\sigma} + \frac{\varepsilon}{\sigma}M_{BN}\|_\infty\right) \leq \frac{\varepsilon}{\sigma}\sqrt{|B|}(1+w) \leq \frac{2w\varepsilon\sqrt{|B|}}{\sigma}\end{aligned}$$

Ezt felhasználva a ρ komponenseire a következő becslést kapjuk:

$$\rho_i \leq \frac{2w\varepsilon\sqrt{|B|}\pi_B}{\sigma}, \quad i \in B_2$$

Legyen

$$\xi := \begin{cases} \rho_i, & \text{ha } i \in B_2 \\ 0, & \text{különben} \end{cases}$$

Egyszerűen ellenőrizhető, hogy ξ a megadott egyenlet megoldása, így már csak a két pozitivitási feltételt kell ellenőrizni. Könnyen látható, hogy az $\mathbf{x}_B - \xi > 0$ feltétel az $i \notin B_2$ indexeknél mindig teljesül. A becsléseket felhasználva az $\mathbf{x}_B - \xi > 0$ elégséges feltétele az $i \in B_2$ indexekre

$$\frac{\sigma}{2n} - \frac{2w\sqrt{|B|}\pi_B}{\sigma}\varepsilon > 0,$$

ám ez a lemmában megadott ε értékre teljesül, ugyanis $\frac{\sigma^2}{4nw\sqrt{|B|}\pi_B} > \varepsilon$.
Az $\mathbf{s}_N - M_{NN}\mathbf{x}_N - M_{NB}\xi > 0$ feltétel helyett pedig az erősebb $\frac{\sigma}{2n} - \|M_{NN}\mathbf{x}_N\|_\infty - \|M_{NB}\xi\|_\infty > 0$ teljesülését vizsgáljuk.

$$\|M_{NN}\mathbf{x}_N\|_\infty \leq \|M_{NN}\|_\infty\|\mathbf{x}_N\|_\infty \leq w\frac{\varepsilon}{\sigma},$$

$$\|M_{NB}\xi\|_\infty \leq \|M_{NB}\|_\infty \|\xi\|_\infty \leq w\frac{2w\sqrt{|B|}\pi_B}{\sigma}\varepsilon.$$

A felső becsléseket beírva egy szigorúbb feltételt kapunk:

$$\frac{\sigma}{2n} - \frac{w}{\sigma}\varepsilon - \frac{2w^2\sqrt{|B|}\pi_B}{\sigma}\varepsilon > 0.$$

Mivel

$$\varepsilon = \frac{\sigma^2}{6nw^2\sqrt{|B|}\pi_B} < \frac{\sigma^2}{2nw(1+2w\sqrt{|B|}\pi_B)},$$

ezért

$$\frac{w}{\sigma}\left(1+2w\sqrt{|B|}\pi_B\right)\varepsilon < \frac{s}{2n},$$

azaz a megadott ε érték mellett teljesül a szigorúbb feltétel, tehát igaz az eredetileg kitűzött pozitivitási feltétel is.

Figyelembe véve a Gauss–Jordan elimináció lépésszámát a lemmát beláttuk. □

A most igazolt lemma állítása az előzőekben már említett okok miatt racionális együtthatók esetén is igaz marad. Az eddigi eredményeket a következő tételben foglaljuk össze.

3.3.2. Tétel. *A Dikin–féle affin skálázási algoritmussal egy* $\mathbf{x}^0 \in \mathcal{F}^0$ $(\mathbf{s}^0 > 0)$ *pontból, amelyre* $\delta(\mathbf{x}^0) \leq \tau = 2$ *és* $(\mathbf{x}^0)^T\mathbf{s}^0 = n$ *teljesül, előállítható egy olyan* $\mathbf{x} \in \mathcal{F}^0$ $(\mathbf{s} > 0)$ *pont, amelyből a kerekítési eljárással szigorúan komplementáris megoldás számolható ki. Az* $(\mathbf{x}, \mathbf{s})$ *pont előállításához* $\left\lceil 2n \log \frac{6n^2w^2\sqrt{|B|}\pi_B}{\sigma^2} \right\rceil$ *iteráció szükséges.* □

4. fejezet

Lineáris optimalizálás alapvető tételei

Az eddigiekben a ferdén szimmetrikus önduális feladatosztállyal foglalkoztunk. A fejezet célja megmutatni, hogy tetszőleges primál–duál feladatpárnak megfeleltethető egy speciális alakú ferdén szimmetrikus feladat, melyet a szakirodalomban a feladatpárhoz tartozó Goldman–Tucker modellnek neveznek (lásd például Roos, Terlaky és Vial [24] könyvét). Belátjuk, hogy elegendő a Goldman–Tucker modellt vizsgálni, ugyanis az erősen komplementáris megoldásaiból következtetni tudunk a primál és duál feladat megoldásaira, illetve a nem megoldhatóságukra. Ezt követően feloldjuk az igen erősnek tűnő belső pont feltételt, mely az eddig ismertetett eredmények szinte mindegyikéhez szükséges. Ezt a problémát a Goldman–Tucker modell beágyazásával oldjuk meg úgy, hogy a kapott feladatnak a csupa egyesekből álló vektor belső pontja lesz. Befejezésül pedig a jól ismert Farkas lemmára mutatunk egy újabb bizonyítást a fejezet eredményeinek segítsélgével.

4.1. Gyenge dualitás tétel

Az általános primál (P) és duál (D) lineáris programozási feladatot az alábbi kanonikus alakban tekintjük:

$$(P) \qquad \min\left\{\mathbf{c}^T\mathbf{x} \;:\; A\mathbf{x} \geq \mathbf{b},\;\; \mathbf{x} \geq \mathbf{0}\right\},$$

$$(D) \qquad \max\left\{\mathbf{b}^T\mathbf{y} \;:\; A^T\mathbf{y} \leq \mathbf{c},\;\; \mathbf{y} \geq \mathbf{0}\right\},$$

ahol $A \in \mathbb{R}^{m\times k}$ mátrix, $\mathbf{b}, \mathbf{y} \in \mathbb{R}^m$ és $\mathbf{c}, \mathbf{x} \in \mathbb{R}^k$.

A kanonikus lineáris programozási feladatra a gyenge dualitás tétel egyszerűen bizonyítható.

4.1.1. Tétel. *(Gyenge dualitás tétel) Tegyük fel, hogy* $\mathbf{x} \in \mathbb{R}^k$ *és* $\mathbf{y} \in \mathbb{R}^m$ *a primál* (P) *és duál* (D) *feladatok megengedett megoldásai. Ekkor*

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y},$$

ahol egyenlőség akkor és csak akkor áll fenn, ha
(i) $x_i(c - A^T y)_i = 0$ *minden* $i = 1, \cdots, k$ *és*
(ii) $y_j(Ax - b)_j = 0$ *minden* $j = 1, \cdots, m$.

Bizonyítás. Az $\mathbf{x}$ és $\mathbf{y}$ vektorok primál és duál megengedettségét használva kapjuk, hogy

$$(\mathbf{c} - A^T\mathbf{y})^T\mathbf{x} \geq 0 \qquad \text{és} \qquad \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) \geq 0,$$

ahol egyenlőség akkor és csak akkor áll fenn, ha (i) és (ii) teljesül. Ezen két egyenlőtlenséget összeadva a kivánt

$$0 \leq (\mathbf{c} - A^T\mathbf{y})^T\mathbf{x} + \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$$

egyenlőtlenség adódik. Tételünket bebizonyítottuk. □

4.1.2. Megjegyzés. *Az* (i) *és* (ii) *feltételeket komplementaritási feltételeknek nevezzük. A koordinátánkénti jelölést használva az* $\mathbf{x}(\mathbf{c} - A^T\mathbf{y}) = \mathbf{0}$ *és* $\mathbf{y}(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$ *egyenletekre jutunk. Tehát a Gyenge dualitás tétel szerint, komplementaritás és megengedettség a megoldások optimalitását garantálja.*

Az alábbi elégséges optimalitási feltétel könnyen levezethető.

4.1.3. Következmény. *(Gyenge equilibrium tétel) Legyenek* $\mathbf{x} \in \mathbb{R}^n$ *és* $\mathbf{y} \in \mathbb{R}^m$ *olyan primál és duál megengedett megoldások, melyekre* $\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$. *Ekkor* $\mathbf{x}$ *a primál* (P) *feladat egy optimális megoldása és* $\mathbf{y}$ *a duál* (D) *feladat egy optimális megoldása.* □

4.2. Goldman–Tucker modell

Célunk a lineáris programozási feladat olyan megoldásainak előállítása, melyekre a dualitási rés nulla, így azt az egyenlőtlenségrendszert kell megoldanunk, amely a primál és duál feltételekből áll, valamint előírjuk, hogy a duál célfüggvény értéke legalább akkora, mint a primál célfüggvény értéke ($\mathbf{b}^T\mathbf{y} \geq \mathbf{c}^T\mathbf{x}$). Ekkor ugyanis az 4.1.1. Gyenge dualitás tétel miatt ezen rendszer minden megengedett megoldása primál és duál megengedett megoldást ad, amelyre a célfüggvényértékek szükségképpen egyenlők, így az 4.1.3. Következmény szerint optimális megoldások.

A szükséges slack változók $(\mathbf{z}, \mathbf{s}, \zeta)$ bevezetésével az alábbi egyenlőtlenségrendszerre jutunk:

$$\begin{aligned} A\mathbf{x} - \mathbf{z} &= \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0} \\ A^T\mathbf{y} + \mathbf{s} &= \mathbf{c}, \quad \mathbf{y} \geq \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \\ \mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} - \zeta &= 0, \quad \zeta \geq 0. \end{aligned}$$

Homogenizálva az egyenleteket a *Goldman-Tucker feladat*ot kapjuk:

$$\left.\begin{array}{rrrrrrcl} & A\mathbf{x} & -\xi\mathbf{b} & -\mathbf{z} & & & = & \mathbf{0}, \quad \mathbf{x} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0} \\ -A^T\mathbf{y} & & +\xi\mathbf{c} & & -\mathbf{s} & & = & \mathbf{0}, \quad \mathbf{y} \geq \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \\ \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & & & -\zeta & = & 0, \quad \xi \geq 0, \quad \zeta \geq 0. \end{array}\right\} \quad (GT)$$

A triviális, azonosan nulla megoldás kielégíti ezt a homogén rendszert, de ez a céljaink szempontjából érdektelen. Ezen Goldman-Tucker rendszer bizonyos nemtriviális, speciális megoldását szeretnénk előállítani és ennek segítségével következtetünk a (P) és (D) feladatok megoldhatóságára, megoldásaira az alábbi tétel alapján.

4.2.1. Tétel. *Adott egy primál (P) és duál (D) lineáris programozási feladatpár. Az alábbi állítások igazak:*

1. *A (P) és (D) feladatok tetszőleges $(\mathbf{x}, \mathbf{y})$ optimális megoldás párja, melyre a dualitási rés nulla, a megfelelő Goldman-Tucker rendszer egy megoldását adja $\xi = 1,\ \zeta = 0$ választással.*

2. *Ha $(\mathbf{y}, \mathbf{x}, \xi, \mathbf{z}, \mathbf{s}, \zeta)$ a Goldman-Tucker rendszer egy megoldása, akkor vagy $\xi = 0$ vagy $\zeta = 0$, azaz $\xi\zeta > 0$ nem lehet igaz.*

3. *A Goldman-Tucker rendszer tetszőleges* $(\mathbf{y}, \mathbf{x}, \xi, \mathbf{z}, \mathbf{s}, \zeta)$ *megoldása, ahol* $\xi > 0$ *és* $\zeta = 0$, *a primál* (P) *és duál* (D) *feladatok egy* $(\frac{\mathbf{x}}{\xi}, \frac{\mathbf{y}}{\xi})$ *optimális megoldás párját adja, melyre a dualitási rés nulla.*

4. *Ha a Goldman-Tucker rendszernek van olyan megoldása* $(\bar{\mathbf{y}}, \bar{\mathbf{x}}, \bar{\xi}, \bar{\mathbf{z}}, \bar{\mathbf{s}}, \bar{\zeta})$, *ahol* $\bar{\xi} = 0$ *és* $\bar{\zeta} > 0$, *akkor megállapíthatjuk, hogy vagy a* (P), *vagy a* (D) *feladat, vagy mindkettő nem megengedett.*

Bizonyítás. Az első és a harmadik állítás behelyettesítéssel könnyen ellenőrizhető.

A második állítást indirekt bizonyítjuk. Ha $\xi\zeta$ pozitív lenne, akkor

$$0 < \xi\zeta = \xi\mathbf{b}^T\mathbf{y} - \xi\mathbf{c}^T\mathbf{x} = \mathbf{x}^T A\mathbf{y} - \mathbf{z}^T\mathbf{y} - \mathbf{x}^T A^T\mathbf{y} - \mathbf{s}^T\mathbf{x} = -\mathbf{z}^T\mathbf{y} - \mathbf{s}^T\mathbf{x} \le 0$$

egyenlőtlenséget kapnánk, ami nyilvánvaló ellentmondás.

Az utolsó állítás igazolásánál a $\bar{\xi} = 0$ feltételből következik, hogy $A\bar{\mathbf{x}} \ge \mathbf{0}$ és $A^T\bar{\mathbf{y}} \le \mathbf{0}$. Továbbá, ha $\bar{\zeta} > 0$, akkor vagy $\mathbf{b}^T\bar{\mathbf{y}} > 0$, vagy $\mathbf{c}^T\bar{\mathbf{x}} < 0$, vagy mindkettő fennáll. Ha $\mathbf{b}^T\bar{\mathbf{y}} > 0$, akkor, feltételezve, hogy a (P) feladatnak van egy $\mathbf{x} \ge \mathbf{0}$ megoldása a

$$0 < \mathbf{b}^T\bar{\mathbf{y}} \le \mathbf{x}^T A^T\bar{\mathbf{y}} \le 0$$

ellentmondáshoz jutunk. Tehát ha $\mathbf{b}^T\bar{\mathbf{y}} > 0$, akkor (P) nem megengedett.
Hasonlóan, ha $\mathbf{c}^T\bar{\mathbf{x}} < 0$, akkor a duál feladat nem megengedettségét kapjuk. □

Vegyük észre, hogy a Goldman-Tucker rendszer az alábbi kompakt alakba írható:

$$M\tilde{\mathbf{x}} \ge \mathbf{0}, \qquad \tilde{\mathbf{x}} \ge \mathbf{0}, \qquad s(\tilde{\mathbf{x}}) = M\tilde{\mathbf{x}}, \tag{4.1}$$

ahol

$$\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \\ \xi \end{pmatrix}, \qquad s(\tilde{\mathbf{x}}) = \begin{pmatrix} \mathbf{z} \\ \mathbf{s} \\ \zeta \end{pmatrix} \quad \text{és} \quad M = \begin{pmatrix} 0 & A & -\mathbf{b} \\ -A^T & 0 & \mathbf{c} \\ \mathbf{b}^T & -\mathbf{c}^T & 0 \end{pmatrix}$$

egy ferdén szimmetrikus mátrix, azaz tetszőleges (P) és (D) feladatpárhoz tartozó Goldmann–Tucker rendszer egy homogén (SP) feladat.

4.3. Erős dualitás tétel

Az előzőekben már igazoltuk belső pont feltétel teljesülése mellett (SP) feladatokra a Goldman–Tucker tételt (2.4.4. Következmény). Célunk az általános esetben is belátni a tételt. Vegyük észre, hogy az (4.1) Goldman–Tucker modell nem teljesítheti a belső pont feltételt a 4.2.1. tétel miatt. Ezért szükségünk van a (4.1) homogén feladat beágyazására olyan ekvivalens (SP) feladatba, mely teljesíti a belső pont feltételt.

Legyen $\mathbf{x} = \mathbf{s} = \mathbf{e}$. Ezek a vektorok pozitívak, de nem elégítik ki a (4.1) feltételeket. Definiáljuk az $\mathbf{r}$ hibavektort

$$\mathbf{r} := \mathbf{e} - M\mathbf{e}\ , \quad \text{és legyen} \quad \lambda := n + 1.$$

Ekkor

$$\begin{pmatrix} M & \mathbf{r} \\ -\mathbf{r}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{e} \\ 1 \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \lambda \end{pmatrix} = \begin{pmatrix} M\mathbf{e} + \mathbf{r} \\ -\mathbf{r}^T\mathbf{e} + \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{e} \\ 1 \end{pmatrix}.$$

A fenti konstrukció alapján definiáljuk az $(\overline{SP})$ feladatot a következőképpen:

$$\min \left\{ \lambda\vartheta \ : \ \begin{pmatrix} M & \mathbf{r} \\ -\mathbf{r}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \vartheta \end{pmatrix} - \begin{pmatrix} \mathbf{s} \\ \nu \end{pmatrix} = - \begin{pmatrix} \mathbf{0} \\ \lambda \end{pmatrix} ; \ \begin{pmatrix} \mathbf{x} \\ \vartheta \end{pmatrix}, \begin{pmatrix} \mathbf{s} \\ \nu \end{pmatrix} \geq \mathbf{0} \right\}$$

feladat kielégíti a belső pont feltételt, mivel a csupa egyesből álló vektor megengedett megoldást ad. Ez a feladat (SP) alakú, ahol

$$\overline{M} = \begin{pmatrix} M & \mathbf{r} \\ -\mathbf{r}^T & 0 \end{pmatrix}, \quad \overline{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ \vartheta \end{pmatrix} \quad \text{és} \quad \overline{\mathbf{q}} = \begin{pmatrix} \mathbf{0} \\ \lambda \end{pmatrix}.$$

Mivel az $(\overline{SP})$ feladat célfüggvénye a ϑ változó egy pozitív többszöröse, így ennek a változónak minden optimális megoldásban, az 2.1.1. Lemma miatt, nulla az értéke. Tehát az $(\mathbf{x}, \vartheta, \mathbf{s}, \nu)$ szigorúan komplementáris megoldása az $(\overline{SP})$ feladatnak, akkor és csak akkor ha $(\mathbf{x}, \mathbf{s})$ szigorúan komplementáris megoldása a (4.1) Goldman–Tucker modellnek. Figyelembe véve, hogy az $(\overline{SP})$ feladat kielégíti a belső pont feltételt, az 2.4.3. tétel miatt létezik szigorúan komplementáris megoldása, ezzel beláttuk a Goldman–Tucker tételt általános esetben is.

Az eddigieket összefoglalva: Minden lineáris programozási feladat beágyazható egy az (SP) alakban adott önduális $(\overline{SP})$ feladatba olyan módon, hogy az $(\overline{\mathbf{x}}, \overline{\mathbf{s}}) = (e, e)$

vektor az $(\overline{SP})$ feladat megengedett megoldása. Továbbá a (4.1) ferdén szimmetrikus feladat bármely erősen komplementáris megoldása vagy az eredeti lineáris programozási feladat egy optimális megoldását adja, vagy bizonyítja, hogy az adott lineáris programozási feladatnak nincs optimális megoldása.

A fenti eredmények segítségével könnyen bebizonyítható az úgynevezett Erős dualitás tétel.

4.3.1. Tétel. (Erős dualitás tétel) *Legyenek a (P) és a (D) feladatok adottak. Ekkor a következő két alternatíva valamelyike teljesül:*

1. *$\mathcal{P} \neq \emptyset$ és $\mathcal{D} \neq \emptyset$, valamint létezik $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ szigorúan komplementáris megoldás.*
2. *$\mathcal{P}^* = \emptyset$ és $\mathcal{D}^* = \emptyset$, ez pontosan akkor fordulhat elő, ha létezik $(\overline{\mathbf{x}}, \overline{\mathbf{y}})$, melyre*

$$A\overline{\mathbf{x}} \geq \mathbf{0}, \quad A^T\overline{\mathbf{y}} \leq \mathbf{0}, \quad \mathbf{c}^T\overline{\mathbf{x}} \leq \mathbf{b}^T\overline{\mathbf{y}}.$$

Bizonyítás. Az általánosság korlátozása nélkül feltehető, hogy a lineáris programozási feladat Goldman-Tucker rendszerének van egy erősen komplementáris megoldása. Egy ilyen megoldásban vagy $\xi > 0$, és ebben az esetben a 4.2.1. Tétel 3. pontjából következik, hogy létezik optimális megoldás pár nulla dualitási réssel, vagy $\zeta > 0$ a Goldman-Tucker rendszer erősen komplementáris megoldásában. Az utóbbi esetben az 4.2.1. Tétel 4. pontja miatt (P) vagy (D) vagy mindkettő nemmegengedett. $\square$

A lineáris programozás fontos és alapvető eredménye a Farkas lemma, melyre a Goldman–Tucker tétel segítségével most egy újabb bizonyítást adunk.

4.3.2. Lemma. (Farkas lemma) *A következő két egyenlőtlenségrendszer közül pontosan az egyiknek van megoldása*

$$\left.\begin{aligned} A\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}\right\} \qquad \text{és} \qquad \left.\begin{aligned} A^T\mathbf{y} &\leq \mathbf{0} \\ \mathbf{b}^T\mathbf{y} &> 0 \\ \mathbf{y} &\geq \mathbf{0} \end{aligned}\right\},$$

ahol $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$.

Bizonyítás. Tekintsük a következő primál és duál lineáris programozási feladatot

$$(P) \qquad \min\left\{\mathbf{c}^T\mathbf{x} \ : \ A\mathbf{x} \geq \mathbf{b}, \ \ \mathbf{x} \geq \mathbf{0}\right\},$$

$$(D) \qquad \max\left\{\mathbf{b}^T\mathbf{y} \ : \ A^T\mathbf{y} \leq \mathbf{0}, \ \ \mathbf{y} \geq \mathbf{0}\right\}.$$

Elkészíthető a (P) és (D) feladathoz az (SP) ferdén szimmetrikus önduális feladat

$$\left.\begin{array}{lllr} \min\ \mathbf{0}^T\mathbf{y} & +\mathbf{0}^T\mathbf{x} & +0\cdot\xi & \\ & A\mathbf{x} & -\mathbf{b}\xi & \geq \mathbf{0} \\ -A^T\mathbf{y} & & +\mathbf{c}\xi & \geq \mathbf{0} \\ \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & \geq \mathbf{0} \\ & & \mathbf{y},\mathbf{x},\xi & \geq \mathbf{0} \end{array}\right\} (SP)$$

Az általánosság korlátozása nélkül feltehető, hogy az (SP) feladatnak van belső pontja, így létezik szigorúan komplementáris megoldása, $(\overline{\mathbf{y}},\overline{\mathbf{x}},\overline{\xi})$. Vezessük be a következő jelöléseket

$$s(\overline{\mathbf{y}}) := \mathbf{c}\overline{\xi} - A^T\overline{\mathbf{y}}, \qquad s(\overline{\mathbf{x}}) := A\overline{\mathbf{x}} - \mathbf{b}\overline{\xi}, \qquad s(\overline{\xi}) := \mathbf{b}^T\overline{\mathbf{y}} - \mathbf{c}^T\overline{\mathbf{x}}.$$

Ekkor két eset lehetséges: (i) $\overline{\xi} > 0$, illetve (ii) $\overline{\xi} = 0$. Vegyük figyelembe, hogy $\mathbf{c} = \mathbf{0}$.
(i) esetben $\overline{\mathbf{x}} \geq \mathbf{0}$, $s(\overline{\mathbf{x}}) \geq \mathbf{0}$, így $A\overline{\mathbf{x}} > \mathbf{b}\overline{\xi}$, tehát az $\overline{\mathbf{x}}/\overline{\xi}$ megoldása az első egyenletrendszernek.
(ii) esetben mivel $\overline{\xi} = 0$, ezért $s(\overline{\xi}) > 0$, tehát $\mathbf{b}^T\overline{\mathbf{y}} > 0$, $\overline{\mathbf{y}} \geq \mathbf{0}$. Valamint $s(\overline{\mathbf{y}}) \geq \mathbf{0}$ miatt $A^T\overline{\mathbf{y}} \leq 0$, azaz ebben az esetben az $\overline{\mathbf{y}}$ vektor megoldása a második rendszernek. A szigorú komplementaritás pedig biztosítja, hogy a két rendszernek nem lehet egyszerre megoldása. □

5. fejezet

Primál–duál belsőpontos algoritmusok

A primál–duál belsőpontos algoritmusok a gyakorlatban is hatékonyan használható módszerek. Az ilyen algoritmusoknak két sarkalatos pontjuk van. Az egyik a μ centralitási paraméter iterálási módja, a másik pedig, hogy milyen centralitási mértéket használunk a pontunk és a centrális út távolságának mérésére. Az első kérdés szempontjából három csoportba oszthatjuk a primál–duál belsőpontos algoritmusokat. Általánosan a centralitási paramétert $\mu^+ = (1-\theta)\mu$ szabály szerint frissítjük, és a θ értékének megfelelően osztályozunk. Ha θ egy konstans szám akkor hosszú lépéses, ha θ a feladat dimenziójától, azaz n értékétől függ, akkor pedig rövid lépéses módszerről beszélünk. A harmadik csoportot az adaptív algoritmusok alkotják, ahol a θ paraméter ellentétben az előző két esettel az iterációk során változik, függ az aktuális ponttól. A rövid lépéses módszerek $\mathcal{O}(\sqrt{n}\log\frac{n}{\varepsilon})$, míg a hosszú lépéses módszerek esetén $\mathcal{O}(n\log\frac{n}{\varepsilon})$ a bizonyított komplexitás [24], ez ellentmond a gyakorlati tapasztalatoknak, miszerint a hosszú lépéses módszerek hatékonyabbak.

A másik lényeges kérdés esetén is többfajta megoldással találkozunk a szakirodalomban. A centrális úton az $(\mathbf{x}, \mathbf{s})$ vektor minden koordinátája egyenlő, így kézenfekvő a

$$\delta_c(\mathbf{x}, \mathbf{s}) := \frac{\max(\mathbf{x}\,\mathbf{s})}{\min(\mathbf{x}\,\mathbf{s})}$$

módon definiálni a centralitási mértéket, ezt használtuk a Dikin–féle affin skálázású algoritmus elemzése során is [2, 24]. Ez a fajta centralitási mérték nem mindig felel meg céljainknak, mert a vektornak csak a maximális, illetve minimális elemét veszi figyelembe. Az alábbi mérték, mellyet Kojimáék [11] vezettek be, már a vektor összes koordinátájától függ, valamint a centrális úton lévő pontok esetén nulla az értéke, ami jobban tükrözi célunkat, miszerint a centrális út és az aktuális pont távolságát szeretnénk jellemezni vele.

$$\delta_0(\mathbf{x},\mathbf{s},\mu) := \frac{1}{2}\left\|\mathbf{e} - \frac{\mathbf{x}\,\mathbf{s}}{\mu}\right\|$$

A fejezetben használt centralitási mérték a szakirodalomban egyik leggyakrabban alkalmazott mérték, melyet Jansen [8] definiált, a következő

$$\delta_1 := \frac{1}{2}\left\|\left(\frac{\mathbf{x}\,\mathbf{s}}{\mu}\right)^{\frac{1}{2}} - \left(\frac{\mu}{\mathbf{x}\,\mathbf{s}}\right)^{\frac{1}{2}}\right\|.$$

Ez a fajta definíció még inkább megfelel az elvárásainknak. Nem csak hogy eltűnik a centrális úton, hanem ha egy bizonyos korlát alatt akarjuk tartani, akkor nem engedi a pontunkat a megengedett tartomány határhoz túl közel, ugyanis akár az $\mathbf{x}$, akár az $\mathbf{s}$ vektor valamely koordinátája tart nullához, vagy végtelenbe, akkor a pont centralitási mértéke is tart végtelenbe. (Ennek a tulajdonságnak csak a fele teljesül a δ_0 mérték esetén.) A δ_1 centralitási mérték általánosításának tekinthetők a legújabb, self–reguláris függvények segítségével definiált mértékek, melyekkel Terlakyék foglalkoznak [19].

A fejezetben először a Primál–duál logaritmikus büntetőfüggvényes algoritmust ismertetjük (Kojimáék [11] és Terlakyék [24] elemzése alapján), mely teljes Newton–lépések megtételével konvergál a megoldáshoz. Az eljárás polinomialitása egyszerűen bizonyítható, és Ling lemmái [13, 14] segítségével belátható az algoritmus $\mathcal{O}(\sqrt{n}\log\frac{n}{\varepsilon})$ komplexitása is. Az algoritmus továbbfejlesztett változatának tekinthetjük a prediktor–korrektor algoritmusokat, melyek egyik első változatát Sonnevend, Stoer és Zhao [26] vezette be lineáris programozási feladatokra. Sonnevendék algoritmusának a prediktor lépés után több korrektor lépésre volt szüksége ahhoz, hogy a centrális út megfelelő, előírt környezetébe visszajusson. Mizuno, Todd és Ye [17] olyan prediktor–korrektor belsőpontos algoritmust publikált lineáris programozási feladatra, amelynél egy prediktor lépést csupán egy korrektor lépés követett és az algoritmus komplexitása, a lineáris programozás szakirodalmából ismert legjobb. Ez utóbbi algoritmus egy változatát ismertetjük a fejezetben. A rövid lépéses prediktor–korrektor algoritmus elemzése után egy adaptív változatot is mutatunk, amely esetén már kvadratikus konvergencia is bizonyítható.

Mizunoék eredményét [17] Anstreicher és Ye [31] ültette át pozitív szemidefinit mátrixszal adott lineáris komplementaritási feladatokra megtartva az eljárás komplexitását. Az algoritmus további általánosítását adta Miao [16], illetve Potra és Sheng

[22] $P_*(\kappa)$ mátrixokkal adott lineáris komplementaritási feladatokra. Potra [23] újabb változatát adta az eredeti algoritmusnak pozitív szemidefinit mátrixok esetére új centralitási mértéket használva (a δ_0 helyett a δ_1 centralitási mértékkel a centrális út nagyobb környezetében mozog a pontunk). Ezt az új verziót $P_*(\kappa)$ mátrixokkal meghatározott lineáris komplementaritási feladatokra vizsgálja meg a TDK dolgozatom [18].

5.1. Primál–duál logaritmikus büntetőfüggvényes algoritmus

A lineáris programozási *primál* és *duál* feladatok legyenek a következő alakban adottak

$$\left.\begin{array}{rcl} \min & \mathbf{c}^T\mathbf{x} & \\ A\mathbf{x} & = & \mathbf{b} \\ \mathbf{x} & \geq & \mathbf{0} \end{array}\right\} \quad (P) \qquad\qquad \left.\begin{array}{rcl} \max & \mathbf{b}^T\mathbf{y} & \\ A^T\mathbf{y} & \leq & \mathbf{c} \end{array}\right\} \quad (D)$$

ahol $A \in \mathbb{R}^{m\times n}$, $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$ és $\mathbf{y}, \mathbf{b} \in \mathbb{R}^m$. Az általánosság korlátozása nélkül feltehető, hogy $rang(A) = m$. Legyen a *primál* illetve a *duál megengedett megoldás*ok halmaza rendre

$$\mathcal{P} := \{\mathbf{x} \in \mathbb{R}^n_\oplus \ : \ A\mathbf{x} = \mathbf{b}\} \ \text{ és } \ \mathcal{D} := \{(\mathbf{y}, \mathbf{s}) \in \mathbb{R}^{m+n} : \ A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \geq \mathbf{0}\}.$$

Jelölje a

$$\mathcal{P}_+ = \{\mathbf{x} \in \mathcal{P} : \ \mathbf{x} > \mathbf{0}\} \qquad \text{és a} \qquad \mathcal{D}_+ = \{(\mathbf{y}, \mathbf{s}) \in \mathcal{D} : \ \mathbf{s} > \mathbf{0}\}$$

primál- illetve *duál belsőpontok halmazát*.

Belső pont feltétel: létezik $\bar{\mathbf{x}} \in \mathcal{P}^+$ (létezik $(\bar{\mathbf{y}}, \bar{\mathbf{s}}) \in \mathcal{D}^+$)

5.1.1. Megjegyzés. *A belsőpontos feltételek teljesülése esetén, azaz ha $\mathcal{P}_+ \neq \emptyset$ és a $\mathcal{D}_+ \neq \emptyset$, akkor az erős dulitás tétel alapján létezik primál (és duál) optimális megoldás.*

A *primál* illetve a *duál optimális megoldások halmaza* a

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

$$\mathcal{D}^* := \{\mathbf{y}^* \in \mathcal{D} \,:\, \mathbf{b}^T\mathbf{y}^* \geq \mathbf{b}^T\mathbf{y},\ \ \forall \mathbf{y} \in \mathcal{D}\}.$$

Dualitás rés: $\mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} = \mathbf{y}^T(A\,\mathbf{x}) - (A^T\mathbf{y} + \mathbf{s})^T\mathbf{x} = \mathbf{s}^T\mathbf{x}.$

A következő állítás könnyen belátható.

5.1.2. Állítás. *Legyenek az* $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n$ *olyan vektorok, amelyekre* $\mathbf{x} \geq \mathbf{0}$ *és* $\mathbf{s} \geq \mathbf{0}$ *teljesül. Az* $\mathbf{x}$ *és* $\mathbf{s}$ *vektorok pontosan akkor ortogonálisak egymásra, ha* $x_i\, s_i = 0$ *teljesül* $i = 1, 2, ..., n$ *esetén.* □

Felhasználva az előző állítást az *optimalitási kritériumok* az alábbi formában írhatók

$$\begin{array}{rcrcll} & A\,\mathbf{x} & & = & \mathbf{b}, & \mathbf{x} \geq \mathbf{0}, \\ A^T\mathbf{y}+ & & \mathbf{s} & = & \mathbf{c}, & \mathbf{s} \geq \mathbf{0}, \\ & & \mathbf{x}\,\mathbf{s} & = & \mathbf{0}. & \end{array}$$

5.1.3. Állítás. *Ha az* $\bar{\mathbf{x}} \in \mathcal{P}_+$ *és* $(\bar{\mathbf{y}}, \bar{\mathbf{s}}) \in \mathcal{D}_+$ *akkor* $\bar{\mathbf{x}}\,\bar{\mathbf{s}} = \bar{\mathbf{w}} \in \mathbb{R}^n_+$*, azaz az optimum belső pontban nem vétetik fel.* □

Ezen tulajdonság miatt az előző feladatot relaxáljuk, olyan primál és duál belsőpontos megoldáspárt szeretnénk előállítani, amelyre:

$$\begin{array}{rcrcll} & A\,\mathbf{x} & & = & \mathbf{b}, & \mathbf{x} > \mathbf{0}, \\ A^T\mathbf{y}+ & & \mathbf{s} & = & \mathbf{c}, & \mathbf{s} > \mathbf{0}, \\ & & \mathbf{x}\,\mathbf{s} & = & \mu\,\mathbf{e}. & \end{array}$$

teljesül, ahol $\mu > 0$ adott. Az $(\bar{\mathbf{x}}, \bar{\mathbf{s}}) > 0$ primál-duál megoldásból szeretnénk az előző *relaxált optimalitási kritériumok*nak egy megoldását előállítani oly módon, hogy valamilyen ismeretlen $\Delta\mathbf{x}$, $\Delta\mathbf{y}$, $\Delta\mathbf{s}$ értékkel módosítjuk a megengedett megoldást:

$$\begin{array}{rcrcll} & A\,(\bar{\mathbf{x}} + \Delta\mathbf{x}) & & = & \mathbf{b}, & \bar{\mathbf{x}} + \Delta\mathbf{x} > \mathbf{0}, \\ A^T(\bar{\mathbf{y}} + \Delta\mathbf{y})+ & & (\bar{\mathbf{s}} + \Delta\mathbf{s}) & = & \mathbf{c}, & \bar{\mathbf{s}} + \Delta\mathbf{s} > \mathbf{0}, \\ & & (\bar{\mathbf{x}} + \Delta\mathbf{x})\,(\bar{\mathbf{s}} + \Delta\mathbf{s}) & = & \mu\,\mathbf{e}. & \end{array}$$

Egyszerű számítások után kapjuk, hogy

$$\begin{aligned} A\,\Delta\mathbf{x} \quad\quad\quad &= \mathbf{0}, & \bar{\mathbf{x}} + \Delta\mathbf{x} > \mathbf{0}, \\ A^T\Delta\mathbf{y} + \quad\quad\quad \Delta\mathbf{s} &= \mathbf{0}, & \bar{\mathbf{s}} + \Delta\mathbf{s} > \mathbf{0}, \\ \bar{\mathbf{s}}\,\Delta\mathbf{x} + \; \bar{\mathbf{x}}\,\Delta\mathbf{s} + \; \Delta\mathbf{x}\,\Delta\mathbf{s} &= \mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}, \end{aligned}$$

ahol az utolsó egyenlet nemlineáris (kvadratikus). Ezt a nemlineáris egyenlőtlenség-rendszert egyszerűsítjük. Két dolgot hagyunk el: a pozitivitási megkötéseket és a másodrendű tagokat. Ekkor a harmadik egyenlet a következő alakúvá válik

$$\bar{\mathbf{s}}\,\Delta\mathbf{x} + \bar{\mathbf{x}}\,\Delta\mathbf{s} = \mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}.$$

Newton-rendszer: keresendő a $(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s}) \in \mathbb{R}^{n+m+n}$ vektor úgy, hogy

$$\begin{aligned} A\,\Delta\mathbf{x} \quad\quad\quad &= \mathbf{0}, \\ A^T\Delta\mathbf{y} + \quad\quad\quad \Delta\mathbf{s} &= \mathbf{0}, \\ \bar{\mathbf{s}}\,\Delta\mathbf{x} + \; \bar{\mathbf{x}}\,\Delta\mathbf{s} &= \mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}. \end{aligned}$$

Legyenek $\bar{S} = diag(\bar{\mathbf{s}})$ és $\bar{X} = diag(\bar{\mathbf{x}})$ pozitív mátrixok. Ekkor az utolsó egyenletet a következő formára hozzuk

$$\bar{S}\,\Delta\mathbf{x} + \bar{X}\,\Delta\mathbf{s} = \mu\,\mathbf{e} - \bar{S}\,\bar{\mathbf{x}}, \quad \text{azaz} \quad \Delta\mathbf{x} + \bar{S}^{-1}\bar{X}\,\Delta\mathbf{s} = \mu\,\bar{\mathbf{s}}^{-1} - \bar{\mathbf{x}}.$$

Alkalmazzuk az A mátrixot a kapott n-dimenziós vektorokra és használjuk fel az $\bar{\mathbf{x}} \in \mathcal{P}$ összefüggést illetve a Newton-rendszer első egyenletét. Ekkor

$$A\bar{X}\bar{S}^{-1}\Delta\mathbf{s} = A\,\Delta\mathbf{x} + A\bar{X}\bar{S}^{-1}\Delta\mathbf{s} = \mu\,A\,\bar{\mathbf{s}}^{-1} - A\,\bar{\mathbf{x}} = \mu\,A\,\bar{\mathbf{s}}^{-1} - \mathbf{b}.$$

A Newton-rendszer második egyenletére alkalmazzuk az $A\bar{X}\bar{S}^{-1}$ mátrixot, majd pedig rendezzük át az egyenletet. Ekkor

$$-A\,\bar{X}\,\bar{S}^{-1}\,A^T\Delta\mathbf{y} = A\bar{X}\bar{S}^{-1}\Delta\mathbf{s} = \mu\,A\bar{\mathbf{s}}^{-1} - \mathbf{b}.$$

Mivel $rang(A) = m$, ezért $A\bar{X}\bar{S}^{-1}A^T \in \mathbb{R}^{m\times m}$ és nem szinguláris mátrix, hiszen az A teljes rangú, ezért $\Delta\mathbf{y}$ egyértelműen kiszámítható:

$$\Delta\mathbf{y} = (A\bar{X}\bar{S}^{-1}A^T)^{-1}(\mathbf{b} - \mu\,A\bar{\mathbf{s}}^{-1}).$$

Ezek után egyszerűen és egyértelműen kiszámíthatók a $\Delta\mathbf{s}$ és a $\Delta\mathbf{x}$ vektorok, azaz

$$\Delta\mathbf{s} = -A^T(A\bar{X}\bar{S}^{-1}A^T)^{-1}(\mathbf{b} - \mu\, A\,\bar{\mathbf{s}}^{-1}) \quad \text{és}$$

$$\Delta\mathbf{x} = \mu\,\bar{\mathbf{s}}^{-1} - \bar{\mathbf{x}} + \bar{X}\bar{S}^{-1}A^T(A\bar{X}\bar{S}^{-1}A^T)^{-1}(\mathbf{b} - \mu\, A\,\bar{\mathbf{s}}^{-1}).$$

5.1.4. Lemma. *$\Delta\mathbf{x}$ és $\Delta\mathbf{s}$ ortogonális vektorok.*

Bizonyítás. $\Delta\mathbf{x} \in Null(A)$ és $\Delta\mathbf{s} \in rowsp(A^T) = columnsp(A) = R(A)$ és ismert, hogy $Null(A) \perp R(A)$. □

A $\Delta\mathbf{x}$ és a $\Delta\mathbf{s}$ Newton lépések merőlegességét felhasználva szükséges és elégséges feltételt adhatunk a teljes Newton lépés megengedettségére.

5.1.5. Lemma. *A (P) és (D) feladatok esetén, $\bar{\mathbf{x}} + \Delta\mathbf{x} \geq \mathbf{0}$ és $\bar{\mathbf{s}} + \Delta\mathbf{s} \geq \mathbf{0}$ pontosan akkor teljesül, ha $\mu\,\mathbf{e} + \Delta\mathbf{x}\,\Delta\mathbf{s} \geq \mathbf{0}$, ahol $\bar{\mathbf{x}} \in \mathcal{P}_+$ és $(\bar{\mathbf{y}}, \bar{\mathbf{s}}) \in \mathcal{D}_+$.*

Bizonyítás. 1. Ha $\bar{\mathbf{x}} + \Delta\mathbf{x} \geq \mathbf{0}$ és $\bar{\mathbf{s}} + \Delta\mathbf{s} \geq \mathbf{0}$, akkor

$$\begin{aligned} \mathbf{0} \leq (\bar{\mathbf{x}} + \Delta\mathbf{x})(\bar{\mathbf{s}} + \Delta\mathbf{s}) &= \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \bar{\mathbf{x}}\,\Delta\mathbf{s} + \bar{\mathbf{s}}\,\Delta\mathbf{x} + \Delta\mathbf{x}\,\Delta\mathbf{s} = \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \mu\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \Delta\mathbf{x}\,\Delta\mathbf{s} \\ &= \mu\,\mathbf{e} + \Delta\mathbf{x}\,\Delta\mathbf{s}. \end{aligned}$$

2. Legyen $\mu\,\mathbf{e} + \Delta\mathbf{x}\,\Delta\mathbf{s} \geq \mathbf{0}$, azaz $\Delta\mathbf{x}\,\Delta\mathbf{s} \geq -\mu\mathbf{e}$ és legyen

$$\mathbf{x}^\alpha := \bar{\mathbf{x}} + \alpha\,\Delta\mathbf{x}, \quad \mathbf{s}^\alpha := \bar{\mathbf{s}} + \alpha\,\Delta\mathbf{s}, \qquad \text{ahol } \alpha \in [0,1].$$

Ha $\alpha = 0$, akkor az $\mathbf{x}^0 = \bar{\mathbf{x}}$ és az $\mathbf{s}^0 = \bar{\mathbf{s}}$, tehát $\mathbf{x}^0\,\mathbf{s}^0 = \bar{\mathbf{x}}\,\bar{\mathbf{s}} > \mathbf{0}$. Másfelől, ha $\alpha = 1$ akkor az $\mathbf{x}^1 = \bar{\mathbf{x}} + \Delta\mathbf{x}$ és az $\mathbf{s}^1 = \bar{\mathbf{s}} + \Delta\mathbf{s}$. Számítsuk ki az $\mathbf{x}^\alpha\,\mathbf{s}^\alpha$ vektort:

$$\begin{aligned} \mathbf{x}^\alpha\,\mathbf{s}^\alpha &= \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \alpha\,(\bar{\mathbf{x}}\,\Delta\mathbf{s} + \bar{\mathbf{s}}\,\Delta\mathbf{x}) + \alpha^2\,\Delta\mathbf{x}\,\Delta\mathbf{s} = \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \alpha\,(\mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}) + \alpha^2\,\Delta\mathbf{x}\,\Delta\mathbf{s} \\ &\geq \bar{\mathbf{x}}\,\bar{\mathbf{s}} + \alpha\,(\mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}) - \alpha^2\,\mu\,\mathbf{e} = (1-\alpha)\,\bar{\mathbf{x}}\,\bar{\mathbf{s}} + \alpha\,(1-\alpha)\,\mu\,\mathbf{e} = \\ &= (1-\alpha)\,(\bar{\mathbf{x}}\,\bar{\mathbf{s}} + \alpha\,\mu\,\mathbf{e}) > \mathbf{0} \quad \forall \alpha \in [0,1) \end{aligned}$$

Határátmenetet véve $(\alpha \to 1)$: $\mathbf{x}^1\mathbf{s}^1 \geq \mathbf{0}$, azaz $\mathbf{x}^1 \geq \mathbf{0}$, $\mathbf{s}^1 \geq \mathbf{0}$. □

A továbbiakban jelölje

$$\mathbf{x}^+ := \mathbf{x}^1 = \overline{\mathbf{x}} + \Delta\mathbf{x}, \quad \mathbf{y}^+ := \mathbf{y}^1 = \overline{\mathbf{y}} + \Delta\mathbf{y} \quad \text{és} \quad \mathbf{s}^+ := \mathbf{s}^1 = \overline{\mathbf{s}} + \Delta\mathbf{x}$$

a Newton–lépéssel kapott új pontokat.

5.1.6. Lemma. *A (P) és (D) feladatok esetén, ha a* $(\Delta\mathbf{x},\, \Delta\mathbf{s})$ *olyan, hogy* $\mathbf{x}^+ \in P_+$, $(\mathbf{y}^+, \mathbf{s}^+) \in D_+$, *akkor* $(\mathbf{x}^+)^T\mathbf{s}^+ = \mu\, n$.

Bizonyítás. Figyelembe véve, hogy $(\Delta\mathbf{x})^T\Delta\mathbf{s} = 0$.

$$\begin{aligned}(\mathbf{x}^+)^T\mathbf{s}^+ &= (\bar{\mathbf{x}} + \Delta\mathbf{x})^T(\bar{\mathbf{s}} + \Delta\mathbf{s}) = \bar{\mathbf{x}}^T\bar{\mathbf{s}} + \bar{\mathbf{s}}^T\Delta\mathbf{x} + \bar{\mathbf{x}}^T\Delta\mathbf{s} + (\Delta\mathbf{x})^T\Delta\mathbf{s} = \\ &= \bar{\mathbf{x}}^T\bar{\mathbf{s}} + \mathbf{e}^T(\mu\,\mathbf{e} - \bar{\mathbf{x}}\,\bar{\mathbf{s}}) = \mu\,\mathbf{e}^T\mathbf{e} = \mu\, n.\end{aligned}$$ □

Átskálázás

Célunk olyan jelölések bevezetése, melyekkel a Newton rendszer egyszerűbb alakra hozható, ezzel kiemelve a feladat struktúráját. Az alábbi vektorok segítségével átírt Newton–rendszer harmadik feltételének jobb oldalán megjelenik az $\mathbf{u}^{-1} - \mathbf{u}$ vektor, melynek normája fogja mérni a pontunk távolságát a centrélis úttól, ugyanis ha a pontunk rajta van a centrális úton, akkor ez a kifejezés nullával egyenlő. Legyen $\bar{\mathbf{x}} \in \mathcal{P}_+$ és $(\bar{\mathbf{y}},\, \bar{\mathbf{s}}) \in \mathcal{D}_+$. Definiáljuk a következő vektorokat

$$\mathbf{d} := \sqrt{\frac{\bar{\mathbf{x}}}{\bar{\mathbf{s}}}}, \quad \mathbf{u} := \sqrt{\frac{\bar{\mathbf{x}}\,\bar{\mathbf{s}}}{\mu}}, \qquad \text{ekkor} \qquad \frac{\mathbf{d}^{-1}\,\bar{\mathbf{x}}}{\sqrt{\mu}} = \frac{\mathbf{d}\,\bar{\mathbf{s}}}{\sqrt{\mu}} = \mathbf{u}.$$

Továbbá

$$\mathbf{d}_x := \frac{\mathbf{d}^{-1}\,\Delta\mathbf{x}}{\sqrt{\mu}} \qquad \text{és} \qquad \mathbf{d}_s := \frac{\mathbf{d}\,\Delta\mathbf{s}}{\sqrt{\mu}}.$$

Ekkor

$$\mathbf{x}^+ = \bar{\mathbf{x}} + \Delta\mathbf{x} = \sqrt{\mu}\,\mathbf{d}\,(\mathbf{u} + \mathbf{d}_x) \ \text{ és } \ \mathbf{s}^+ = \bar{\mathbf{s}} + \Delta\mathbf{s} = \sqrt{\mu}\,\mathbf{d}^{-1}\,(\mathbf{u} + \mathbf{d}_s), \ \text{ valamint}$$

$$\mathbf{x}^+\,\mathbf{s}^+ = \mu\,\mathbf{e} + \Delta\mathbf{x}\,\Delta\mathbf{s} = \mu\,\mathbf{e} + \sqrt{\mu}\,\mathbf{d}\,\mathbf{d}_x\,\sqrt{\mu}\,\mathbf{d}^{-1}\,\mathbf{d}_s = \mu\,(\mathbf{e} + \mathbf{d}_x\,\mathbf{d}_s).$$

Az új jelölésekkel a 5.1.5. Lemma eredménye a következő alakban írható.

5.1.7. Következmény. *A (P) és (D) feladatok esetén,* $\bar{\mathbf{x}} + \Delta\mathbf{x} \geq \mathbf{0}$ *és* $\bar{\mathbf{s}} + \Delta\mathbf{s} \geq \mathbf{0}$ *pontosan akkor teljesül, ha* $\mathbf{e} + \mathbf{d}_x\,\mathbf{d}_s \geq \mathbf{0}$, *ahol* $\bar{\mathbf{x}} \in \mathcal{P}_+$ *és* $(\bar{\mathbf{y}}, \bar{\mathbf{s}}) \in \mathcal{D}_+$. □

Mivel $\bar{\mathbf{s}}\,\mathbf{d} = \bar{\mathbf{x}}\,\mathbf{d}^{-1} = \sqrt{\mu}\,\mathbf{u}$, ezért

$$\bar{\mathbf{x}}\,\Delta\mathbf{s} + \bar{\mathbf{s}}\,\Delta\mathbf{x} = \bar{\mathbf{x}}\,\sqrt{\mu}\,\mathbf{d}^{-1}\,\mathbf{d}_s + \bar{\mathbf{s}}\,\sqrt{\mu}\,\mathbf{d}\,\mathbf{d}_x = \sqrt{\mu}\,(\bar{\mathbf{s}}\,\mathbf{d}\,\mathbf{d}_x + \bar{\mathbf{x}}\,\mathbf{d}^{-1}\,\mathbf{d}_s) = \mu\,\mathbf{u}\,(\mathbf{d}_x + \mathbf{d}_s).$$

Másfelől $\mu\,\mathbf{e}-\bar{\mathbf{x}}\,\bar{\mathbf{s}} = \bar{\mathbf{x}}\,\Delta\mathbf{s}+\bar{\mathbf{s}}\Delta\mathbf{x} = \mu\,\mathbf{u}\,(\mathbf{d}_x+\mathbf{d}_s)$ és $\mu\,\mathbf{e}-\bar{\mathbf{x}}\,\bar{\mathbf{s}} = \mu\,\mathbf{e}-\mu\,\mathbf{u}^2 = \mu\,\mathbf{u}\,(\mathbf{u}^{-1}-\mathbf{u})$, tehát $\mathbf{d}_x+\mathbf{d}_s = \mathbf{u}^{-1}-\mathbf{u}$. Bevezetve a $\mathbf{d}_y := \Delta\mathbf{y}/\sqrt{\mu}$ vektort, az átskálázott Newton-rendszer a következő lesz

$$\begin{array}{rrcl} AD\,\mathbf{d}_x & & = & \mathbf{0}, \\ (AD)^T\mathbf{d}_y + & \mathbf{d}_s & = & \mathbf{0}, \\ \mathbf{d}_x + & \mathbf{d}_s & = & \mathbf{u}^{-1}-\mathbf{u}. \end{array}$$

Tehát a $\mathbf{d}_x \in Null(AD)$ és a $\mathbf{d}_s \in rowsp(AD) = Null(HD^{-1})$, ahol H tetszőleges olyan mátrix, melynek nulltere megegyezik az A mátrix sorterével. Ekkor

$$\mathbf{d}_x = P_{AD}(\mathbf{u}^{-1}-\mathbf{u}) \quad \text{és} \quad \mathbf{d}_s = P_{HD^{-1}}(\mathbf{u}^{-1}-\mathbf{u}), \tag{5.1}$$

és mivel a $(\Delta\mathbf{x})^T\Delta\mathbf{s} = \mathbf{d}_x^T\mathbf{d}_s = 0$ ezért

$$\|\mathbf{d}_x\|^2 + \|\mathbf{d}_s\|^2 = \|\mathbf{u}^{-1}-\mathbf{u}\|^2.$$

A $\mathbf{d}_x+\mathbf{d}_s = \mathbf{0}$ pontosan akkor teljesül, ha $\mathbf{u}^{-1}-\mathbf{u} = \mathbf{0}$, ami azzal ekvivalens, hogy $\mathbf{u}^{-1} = \mathbf{u}$, azaz $\mathbf{u} = \mathbf{e}$. Vezessük be a *centralitás mérték*et az alábbi módon

$$\delta(\bar{\mathbf{x}},\bar{\mathbf{s}},\mu) = \frac{1}{2}\,\|\mathbf{u}^{-1}-\mathbf{u}\| = \frac{1}{2}\,\left\|\sqrt{\frac{\bar{\mathbf{x}}\,\bar{\mathbf{s}}}{\mu}} - \sqrt{\frac{\mu}{\bar{\mathbf{x}}\,\bar{\mathbf{s}}}}\right\|.$$

5.1.8. Lemma. *Legyen* $\mathbf{a}$ *és* $\mathbf{b}$ *tetszőleges egymásra merőleges vektor. Ekkor*

$$\|\mathbf{a}\,\mathbf{b}\|_\infty \le \frac{1}{4}\,\|\mathbf{a}+\mathbf{b}\|^2 \quad \textit{és} \quad \|\mathbf{a}\,\mathbf{b}\| \le \frac{\sqrt{2}}{4}\,\|\mathbf{a}+\mathbf{b}\|^2.$$

Bizonyítás. Az első egyenlőtlenség a 3.1.5. Állításhoz hasonlóan igazolható. A második állítás esetén figyelembe véve, hogy $\|\mathbf{a}+\mathbf{b}\| = \|\mathbf{a}-\mathbf{b}\|$, mert $\mathbf{a}^T\mathbf{b} = 0$, adódik a

$$\begin{aligned} \|\mathbf{a}\,\mathbf{b}\|^2 &= \mathbf{e}^T(\mathbf{a}\,\mathbf{b})^2 = \frac{1}{16}\,\mathbf{e}^T((\mathbf{a}+\mathbf{b})^2-(\mathbf{a}-\mathbf{b})^2)^2 \le \frac{1}{16}\,\mathbf{e}^T((\mathbf{a}+\mathbf{b})^4+(\mathbf{a}-\mathbf{b})^4) \\ &\le \frac{1}{16}\,(\|\mathbf{a}+\mathbf{b}\|^4+\|\mathbf{a}-\mathbf{b}\|^4) = \frac{1}{8}\,\|\mathbf{a}+\mathbf{b}\|^4. \end{aligned}$$

Ebből pedig kapjuk, hogy $\|\mathbf{a}\,\mathbf{b}\| \le \frac{\sqrt{2}}{4}\,\|\mathbf{a}+\mathbf{b}\|^2$ □

Az előző lemmát $\mathbf{a} = \mathbf{d}_x$ és $\mathbf{b} = \mathbf{d}_s$ vektorokra felírva, valamint figyelembe véve a $\mathbf{d}_x+\mathbf{d}_s = \mathbf{u}^{-1}+\mathbf{u}$ összefüggést, a következőket kapjuk.

5.1.9. Következmény. *Legyen* $\bar{\mathbf{x}} \in \mathcal{P}_+$, $(\bar{\mathbf{y}}, \bar{\mathbf{s}}) \in \mathcal{D}_+$ *és* $\mu > 0$. *Ha* $\delta := \delta(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \mu)$ *akkor*

$$\|\mathbf{d}_x\, \mathbf{d}_s\|_\infty \leq \delta^2 \quad \textit{és} \quad \|\mathbf{d}_x\, \mathbf{d}_s\| \leq \delta^2 \sqrt{2}. \qquad \square$$

A következőkben megmutatjuk, hogy ha a centrális út $\delta(\mathbf{x}, \mathbf{s}, \mu) < 1$ környezetében vagyunk, akkor a teljes Newton–lépés megengedett illetve az így kapott pont is ebben a környezetben lesz.

5.1.10. Tétel. *Ha* $\delta := \delta(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \mu) \leq 1$ *akkor* $\mathbf{x}^+ \in \mathcal{P}$, $(\mathbf{y}^+, \mathbf{s}^+) \in \mathcal{D}$. *Továbbá, ha* $\delta < 1$ *akkor* $\mathbf{x}^+ > \mathbf{0}$, $\mathbf{s}^+ > \mathbf{0}$ *és*

$$\delta(\mathbf{x}^+, \mathbf{s}^+, \mu) \leq \frac{\delta^2}{\sqrt{2\,(1-\delta^2)}}.$$

Bizonyítás. Az $\mathbf{x}^+$ és s^+ pontosan akkor megengedett, ha $\mathbf{e} + \mathbf{d}_x\, \mathbf{d}_s \geq \mathbf{0}$, aminek elégséges feltétele $\|\mathbf{d}_x\, \mathbf{d}_s\|_\infty \leq 1$. Legyen $\mathbf{u}^+ := \sqrt{\frac{\mathbf{x}^+ \mathbf{s}^+}{\mu}}$, ekkor $(\mathbf{u}^+)^2 = \frac{\mathbf{x}^+ \mathbf{s}^+}{\mu} = (\mathbf{e} + \mathbf{d}_x\, \mathbf{d}_s)$. Másfelől a

$$\begin{aligned}
2\,\delta^+ &= \left\|\left(\mathbf{u}^+\right)^{-1} - \mathbf{u}^+\right\| = \left\|\left(\mathbf{u}^+\right)^{-1}\left(\mathbf{e} - \left(\mathbf{u}^+\right)^2\right)\right\| = \left\|\left(\mathbf{u}^+\right)^{-1}\left(\mathbf{e} - \mathbf{e} - \mathbf{d}_x\, \mathbf{d}_s\right)\right\| \\
&= \left\|\frac{\mathbf{d}_x\, \mathbf{d}_s}{\sqrt{\mathbf{e} + \mathbf{d}_x\, \mathbf{d}_s}}\right\| \leq \frac{\|\mathbf{d}_x\, \mathbf{d}_s\|}{\sqrt{1 - \|\mathbf{d}_x\, \mathbf{d}_s\|_\infty}} \leq \frac{\sqrt{2}\,\delta^2}{\sqrt{1-\delta^2}}, \quad \text{és így} \\
\delta^+ &\leq \frac{\delta^2}{\sqrt{2\,(1-\delta^2)}}. \qquad \square
\end{aligned}$$

Világos, hogy ha $\delta \leq 1/\sqrt{2}$, akkor a tétel szerint $\delta^+ \leq \delta^2$, azaz ez esetben az algoritmus lokálisan kvadratikusan konvergens, ám ha $\delta \geq \sqrt{2/3}$, akkor a tétel semmitmondó, mert még a $\delta^+ \leq \delta$ egyenlőtlenség se teljesül.

A következő két Lingtől származó lemma teszi lehetővé, hogy az előző tételben az új pont centralitási mértékére adott felső korlátot élesítsük.

5.1.11. Lemma. (Ling 1. lemmája (1993)) *Legyen* $\mathbf{u} \in \mathbb{R}^p$ *olyan vektor, amelyre* $\mathbf{u} > -\mathbf{e}$ *és legyen* $\sigma := \mathbf{e}^T\mathbf{u}$. *Ekkor, ha* $\mathbf{u} \geq \mathbf{0}$ *vagy* $\mathbf{u} \leq \mathbf{0}$ *akkor*

$$\sum_{i=1}^{p} \frac{-u_i}{1+u_i} \leq \frac{-\sigma}{1+\sigma}$$

és egyenlőség pontosan akkor teljesül, ha legfeljebb egy nem nulla koordinátája van az $\mathbf{u}$ *vektornak.*

Bizonyítás. Ha az $\mathbf{u} = \mathbf{0}$ akkor $\sigma = 0$ és az egyenlőség triviálisan teljesül. Legyen

$$f : (-1, \infty)^p \to \mathbb{R} \quad \text{és} \quad f(\mathbf{u}) := \sum_{i=1}^{p} \frac{-u_i}{1+u_i}.$$

Ekkor

$$\frac{\partial f}{\partial u_i}(\mathbf{u}) = \frac{-(1+u_i) - (-u_i)}{(1+u_i)^2} = \frac{-1 - u_i + u_i}{(1+u_i)^2} = -\frac{1}{(1+u_i)^2},$$

továbbá

$$\frac{\partial^2 f}{\partial u_j \partial u_i}(\mathbf{u}) = 0, \quad \text{ha} \quad i \neq j, \quad \text{és} \quad \frac{\partial^2 f}{\partial u_i^2}(\mathbf{u}) = \frac{2}{(1+u_i)^3}.$$

Tehát bármely $\mathbf{u} > -\mathbf{e}$ vektor esetén $\frac{\partial^2 f}{\partial u_i^2}(\mathbf{u}) > 0$ teljesül, ami azt jelenti, hogy az f szigorúan konvex függvény, mert $H_f(\mathbf{u})$, az f Hesse mátrixa, pozitív definit mátrix. Figyelembe véve azt, hogy $\mathbf{u} \neq \mathbf{0}$ vektor, azt kapjuk, hogy $\mathbf{e}^T\mathbf{u} = \sigma$ ekvivalens az $\mathbf{e}^T\left(\frac{1}{\sigma}\mathbf{u}\right) = 1$ feltétellel, ami pontosan azt jelenti, hogy $\sum_{i=1}^{p} \frac{u_i}{\sigma} = 1$, és mivel $\mathbf{u} \geq \mathbf{0}$ vagy $\mathbf{u} \leq \mathbf{0}$, ezért $\frac{u_i}{\sigma} \geq 0$. Kihasználva az f függvény konvexitását

$$f(\mathbf{u}) = f\left(\sum_{i=1}^{p} \frac{u_i}{\sigma}\,\sigma\,\mathbf{e}_i\right) \leq \sum_{i=1}^{p} \frac{u_i}{\sigma} f(\sigma\,\mathbf{e}_i) = \sum_{i=1}^{p} \frac{u_i}{\sigma}\left(\frac{-\sigma}{1+\sigma}\right) = \frac{-\sigma}{1+\sigma},$$

ahol $\mathbf{e}_i \in \mathbb{R}^p$ az i. egységvektor. Egyenlőség pontosan akkor áll fenn, ha $\mathbf{u} = \sigma\,\mathbf{e}_i$, valamely i indexre. □

Az első Ling lemmát alkalmazva két egymásra merőleges vektor koordinátánkénti szorzatára a következőt kapjuk.

5.1.12. Lemma. (Ling 2. lemmája (1993)) *Legyenek az* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *ortogonális vektorokok, és tegyük fel, hogy* $\|\mathbf{u} + \mathbf{v}\| = 2\,\rho$, *ahol* $\rho < 1$. *Ekkor*

$$\mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{e} + \mathbf{u}\,\mathbf{v}} - \mathbf{e}\right) \leq \frac{2\,\rho^4}{1 - \rho^4}.$$

Bizonyítás. 5.1.8. Lemma alapján $\|\mathbf{u}\,\mathbf{v}\|_\infty \leq \frac{1}{4}\,\|\mathbf{u} + \mathbf{v}\|^2 = \frac{1}{4}\,4\,\rho^2 = \rho^2 < 1$. Jelölje $\mathbf{w} = \mathbf{u}\,\mathbf{v}$, ekkor $\mathbf{e}^T\mathbf{w} = 0$ és $-\mathbf{e} < \mathbf{w} < \mathbf{e}$ teljesül az $\|\mathbf{u}\,\mathbf{v}\|_\infty \leq \rho^2 < 1$ miatt. Jelölje

$$I_+ := \{i : \ w_i > 0\} \quad \text{és} \quad I_- := \{i : \ w_i < 0\},$$

ekkor $\sigma := \sum_{i \in I_+} w_i = -\sum_{i \in I_-} w_i$. Egyszerű számolással és az előző lemma alkalmazásával

$$\mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{e} + \mathbf{u}\,\mathbf{v}} - \mathbf{e}\right) \;=\; \mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{e} + \mathbf{w}} - \mathbf{e}\right) = \sum_{i=1}^{n}\left(\frac{1}{1+w_i} - 1\right) = \sum_{i=1}^{n} \frac{-w_i}{1+w_i}$$

$$= \sum_{i\in I_+} \frac{-w_i}{1+w_i} + \sum_{i\in I_-} \frac{-w_i}{1+w_i} \le \frac{-\sigma}{1+\sigma} + \frac{\sigma}{1-\sigma} = \frac{2\,\sigma^2}{1-\sigma^2} \quad \text{adódik.}$$

Az utolsó kifejezés σ monoton növő függvénye, ezért további felső becslést kapunk, ha σ számot a felső korlátjával helyettesítjük

$$\sigma = \frac{1}{2}\sum_{i=1}^{n} |w_i| = \frac{1}{2}\sum_{i=1}^{n} |u_i\, v_i| \le \frac{1}{4}\sum_{i=1}^{n} \left(u_i^2 + v_i^2\right) = \frac{1}{4}\,\|\mathbf{u}+\mathbf{v}\|^2 = \rho^2 \quad \text{és így}$$

$$\mathbf{e}^T \left(\frac{\mathbf{e}}{\mathbf{e}+\mathbf{u}\,\mathbf{v}} - \mathbf{e} \right) \le \frac{2\,\sigma^2}{1-\sigma^2} \le \frac{2\,\rho^4}{1-\rho^4}.$$

□

Most már rendelkezésünkre áll az a becslés, melynek segítségével pontosabb felső korlátot tudunk adni az új pontunk centralitási mértékére.

5.1.13. Tétel. *Ha* $\delta := \delta(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \mu) < 1$ *akkor*

$$\delta(\mathbf{x}^+, \mathbf{s}^+, \mu) \le \frac{\delta^2}{\sqrt{2\,(1-\delta^4)}}.$$

Bizonyítás. Legyen $\delta^+ := \delta\,(\mathbf{x}^+, \mathbf{s}^+, \mu)$, $\mathbf{x}^+ = \bar{\mathbf{x}} + \Delta\mathbf{x}$, $\mathbf{s}^+ = \bar{\mathbf{s}} + \Delta\mathbf{s}$ és $\mathbf{u}^+ = \sqrt{\frac{\mathbf{x}^+\,\mathbf{s}^+}{\mu}} = \sqrt{\mathbf{e} + \mathbf{d}_x\,\mathbf{d}_s}$. Ekkor a $\mathbf{d}_x^T \mathbf{d}_s = 0$ összefüggést is figyelembe véve

$$\left\|\mathbf{u}^+\right\|^2 = \mathbf{e}^T(\mathbf{u}^+\,\mathbf{u}^+) = \mathbf{e}^T(\mathbf{u}^+)^2 = \mathbf{e}^T\left(\sqrt{\mathbf{e}+\mathbf{d}_x\,\mathbf{d}_s}\right)^2 = \mathbf{e}^T(\mathbf{e}+\mathbf{d}_x\,\mathbf{d}_s) = \mathbf{e}^T\mathbf{e}+\mathbf{d}_x^T\mathbf{d}_s = n$$

adódik. Felhasználva az előző számítás eredményét, azt kapjuk, hogy

$$\begin{aligned} 4\,\left(\delta^+\right)^2 &= \left\|\left(\mathbf{u}^+\right)^{-1} - \mathbf{u}^+\right\|^2 = \left(\left(\mathbf{u}^+\right)^{-1} - \mathbf{u}^+\right)^T \left(\left(\mathbf{u}^+\right)^{-1} - \mathbf{u}^+\right) = \\ &= \left\|\left(\mathbf{u}^+\right)^{-1}\right\|^2 + \left\|\mathbf{u}^+\right\|^2 - 2\,n = \left\|\left(\mathbf{u}^+\right)^{-1}\right\|^2 - n = \left\|\frac{\mathbf{e}}{\sqrt{\mathbf{e}+\mathbf{d}_x\,\mathbf{d}_s}}\right\|^2 - n = \\ &= \mathbf{e}^T\left(\frac{\mathbf{e}}{\mathbf{e}+\mathbf{d}_x\,\mathbf{d}_s} - \mathbf{e}\right). \end{aligned}$$

Alkalmazva a 5.1.12. Lemmát az $\mathbf{u} = \mathbf{d}_x$ és $\mathbf{v} = \mathbf{d}_s$ vektorokkal

$$4\,\left(\delta^+\right)^2 \le \frac{2\delta^4}{1-\delta^4}$$

adódik, ugyanis $\|\mathbf{d}_x + \mathbf{d}_s\| = 2\delta$, a centralitás mértékének a definíciója szerint, ahol $\delta < 1$. Átrendezés után pedig megkapjuk a tétel állítását. □

A 5.1.10. valamint a 5.1.13. Tételekben beláttuk, hogy a centrális út megfelelő környezetéből indulva a teljes Newton–lépés sem vezet ki a környezetből, tehát az alábbi algoritmus megfelelő τ (például $\tau < 1$) esetén jól definiált lesz.

Teljes Newton-lépéses primál-duál logaritmikus büntetőfüggvényes algoritmus

Input:

τ az eltérés paraméter, $0 \leq \tau < 1$
$\varepsilon > 0$ számítási pontosság
kezdeti megoldás: $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{s}^0) \in \mathcal{P}_+ \times \mathcal{D}_+$, $(\mathbf{x}^0)^T \mathbf{s}^0 = \mu^0 n$ és $\delta(\mathbf{x}^0, \mathbf{s}^0, \mu^0) \leq \tau$
büntető paraméter θ, $0 < \theta < 1$

begin
  $\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$
  while $\mu n \geq (1-\theta)\,\varepsilon$
  begin
    $\mathbf{x} := \mathbf{x} + \Delta\mathbf{x}$
    $\mathbf{s} := \mathbf{s} + \Delta\mathbf{s}$
    $\mu := (1-\theta)\,\mu$
  end
end.

Az algoritmus komplexitásának meghatározása érdekében először vizsgáljuk meg, hogy hogyan változik a centralitási mérték a μ paraméter változtatása következtében.

5.1.14. Lemma. *Legyen* $\mathbf{x} \in \mathcal{P}_+$, $(\mathbf{y}, \mathbf{s}) \in \mathcal{D}_+$ *és* $\mu > 0$, *amelyre* $\mathbf{x}^T\mathbf{s} = \mu n$. *Továbbá, legyen* $\delta := \delta(\mathbf{x}, \mathbf{s}, \mu)$ *és* $\mu^+ = (1-\theta)\,\mu$. *Ekkor*

$$\delta(\mathbf{x}, \mathbf{s}, \mu^+)^2 = (1-\theta)\,\delta^2 + \frac{\theta^2 n}{4\,(1-\theta)}.$$

Bizonyítás. Legyen $\delta^+ := \delta(\mathbf{x}, \mathbf{s}, \mu^+)$ és $u = \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}}$. Először belátjuk, hogy az $\mathbf{u}^{-1} - \mathbf{u}$ és az $\mathbf{u}$ vektorok ortogonálisak. Mivel $\mathbf{x}^T\mathbf{s} = \mu\, n$, ezért $\|\mathbf{u}\|^2 = n$, továbbá

$$\left(\mathbf{u}^{-1} - \mathbf{u}\right)^T \mathbf{u} = \mathbf{e}^T \left(\left(\mathbf{u}^{-1} - \mathbf{u}\right)\mathbf{u}\right) = \mathbf{e}^T \left(\mathbf{e} - \mathbf{u}^2\right) = \mathbf{e}^T\mathbf{e} - \mathbf{u}^T\mathbf{u} = n - n = 0.$$

Ekkor

$$\begin{aligned}
4\left(\delta^+\right)^2 &= \left\|\sqrt{\frac{\mu^+}{\mu}}\,\mathbf{u}^{-1} - \sqrt{\frac{\mu}{\mu^+}}\,\mathbf{u}\right\|^2 = \left\|\sqrt{1-\theta}\,\mathbf{u}^{-1} - \frac{\mathbf{u}}{\sqrt{1-\theta}}\right\|^2 \\
&= \left\|\sqrt{1-\theta}\,\left(\mathbf{u}^{-1} - \mathbf{u}\right) + \frac{\theta\,\mathbf{u}}{\sqrt{1-\theta}}\right\|^2 = (1-\theta)\,\left\|\mathbf{u}^{-1} - \mathbf{u}\right\|^2 + \frac{\theta^2\,\|\mathbf{u}\|^2}{1-\theta} \\
&= (1-\theta)\,4\,\delta^2 + \frac{n\,\theta^2}{1-\theta},
\end{aligned}$$

ahol az utolsó egyenlőségnél felhasználtuk a $\delta := \delta(\mathbf{x}, \mathbf{s}, \mu) = \frac{1}{2}\,\|\mathbf{u}^{-1} - \mathbf{u}\|$ és $\|\mathbf{u}\|^2 = n$ összefüggéseket. □

Ezen eredmény ismeretében már könnyen igazolható, hogy az előzőekben ismertetett Teljes Newton–lépéses primál–duál büntetőfüggvényes algoritmus $\mathcal{O}(\sqrt{n}\log\frac{n\mu}{\varepsilon})$ lépésben előállít egy legfeljebb ε dualitásrésű pontpárt.

5.1.15. Tétel. *Legyen* $\tau = \frac{1}{\sqrt{2}}$ *és* $\theta = \frac{1}{2\sqrt{n}}$ *akkor a primál-duál logaritmikus büntetőfüggvényes algoritmus a teljes Newton–lépéssel legfeljebb*

$$\left\lceil 2\sqrt{n}\,\log\frac{n\,\mu^0}{\varepsilon}\right\rceil$$

lépést igényel ahhoz, hogy előállítson egy $\mathbf{x} \in \mathcal{P}_+$, $(\mathbf{y}, \mathbf{s}) \in \mathcal{D}_+$, *amelyre* $\mathbf{x}^T\mathbf{s} \le \varepsilon$.

Bizonyítás. Az algoritmus kezdetén $\delta(\mathbf{x}, \mathbf{s}, \mu) \le \tau = \frac{1}{\sqrt{2}}$ teljesül. Egy iteráció után

$$\delta\left(\mathbf{x}^+, \mathbf{s}^+, \mu\right) \le \frac{\delta^2}{\sqrt{2\,(1-\delta^4)}} \le \frac{\tau^2}{\sqrt{2\,(1-\tau^4)}} = \frac{\frac{1}{2}}{\sqrt{\frac{3}{2}}} < \frac{1}{2}$$

adódik. Továbbá $\left(\mathbf{x}^+\right)^T\mathbf{s}^+ = \mu\, n$ a 5.1.6. Lemma alapján. Miután $\mu^+ = (1-\theta)\,\mu$, ahol $\theta = \frac{1}{2\sqrt{n}}$, 5.1.14. Lemma alapján

$$\begin{aligned}
\delta\left(\mathbf{x}^+, \mathbf{s}^+, \mu^+\right)^2 &= (1-\theta)\,\delta\left(\mathbf{x}^+, \mathbf{s}^+, \mu\right)^2 + \frac{n\,\theta^2}{4(1-\theta)} < \frac{1-\theta}{4} + \frac{n\,\theta^2}{4(1-\theta)} \\
&< \frac{1-\theta}{4} + \frac{n\,\left(\frac{1}{4\,n}\right)}{4\,(1-\theta)} = \frac{1-\theta}{4} + \frac{1}{16\,(1-\theta)} \le \\
&\le \frac{1}{4} + \frac{1}{8} = \frac{3}{8} < \frac{1}{2} = \tau^2,
\end{aligned}$$

mivel $0 < \theta = \frac{1}{2\sqrt{n}} \leq \frac{1}{2}$ teljesül, bármely n esetén. Tehát az iteráció után az új pont ismét a centrális út megfelelő környezetében van. Az algoritmus generál egy

$$(\mathbf{x}_0, \mathbf{s}_0),\ (\mathbf{x}_1, \mathbf{s}_1),\ \ldots,\ (\mathbf{x}_k, \mathbf{s}_k),\ \ldots$$

pontsorozatot, amelynek az elemei rendre a $\mu_0,\ \mu_1,\ \ldots,\ \mu_k,\ \ldots$ paraméterű centrumokhoz tartoznak, és az azoktól mért távolsága nem nagyobb, mint τ. Sőt a pontpárokhoz tartozó dualitás rés egyre kisebb, hiszen $\mathbf{x}_k^T\,\mathbf{s}_k = \mu_k\, n = (1-\theta)^k\,\mu_0\, n$, ahol $1 - \theta < 1$, tehát ez a sorozat tart a nullához. Azaz bármely $\varepsilon > 0$ adott pontossághoz létezik k index úgy, hogy

$$\mathbf{x}_k^T\,\mathbf{s}_k = \mu_k\, n = (1-\theta)^k\,\mu_0\, n \leq \varepsilon, \qquad \text{ekkor} \qquad k\,\log(1-\theta) + \log(\mu_0\, n) \leq \log\varepsilon,$$

a logaritmus monoton növekedése miatt. Figyelembe véve a $-\log(1-\theta) \geq \theta$ összefüggést, az előző egyenlőtlenségnél egy erősebbet követelünk meg, amely valamely, az előzőleg megadott k indexnél nagyobbra teljesül

$$k\,\log(1-\theta) + \log(\mu_0\, n) \leq -k\,\theta + \log(\mu_0\, n) \leq \log\varepsilon,$$

tehát

$$\log(\mu_0\, n) - \log\varepsilon \leq k\,\theta,$$

ezért

$$\frac{1}{\theta}\,\log\frac{\mu_0\, n}{\varepsilon} \leq k, \qquad \text{valamely } k \in \mathbb{N} \text{ esetén.}$$

Behelyettesítve a $\theta = \frac{1}{2\sqrt{n}}$ értéket azt kapjuk, hogy

$$2\sqrt{n}\,\log\frac{\mu_0\, n}{\varepsilon} \leq k, \quad \text{azaz} \quad \left\lceil 2\sqrt{n}\,\log\frac{\mu_0\, n}{\varepsilon}\right\rceil = k.$$

□

5.2. Prediktor–korrektor algoritmus

Az előző fejezetben ismertetett primál–duál algoritmust továbbfejlesztve jutunk el a prediktor–korrektor algoritmushoz. A bevezetett jelöléseket használjuk a továbbiakban is. Mint már láttuk, a $\mathbf{d}_x$ és $\mathbf{d}_s$ vektorok merőlegesek egymásra, valamint $\mathbf{d}_x$ az AD, míg $\mathbf{d}_s$ a HD^{-1} mátrixok nullterére vett vetülete az $\mathbf{u}^{-1} - \mathbf{u}$ vektornak (lásd (5.1)).

Szedjük szét a $\mathbf{d}_x$ és $\mathbf{d}_s$ vektorokat egy affin skálázású és egy centralizáló résznek nevezett komponensre a következőképpen:

centralizáló lépés: $\quad \mathbf{d}_x^c := P_{AD}(\mathbf{u}^{-1}) \quad$ és $\quad \mathbf{d}_s^c := P_{HD^{-1}}(\mathbf{u}^{-1}),$

affin skálázású lépés: $\quad \mathbf{d}_x^a := P_{AD}(-\mathbf{u}) \quad$ és $\quad \mathbf{d}_s^a := P_{HD^{-1}}(-\mathbf{u}).$

Azaz

$$\mathbf{d}_x = \mathbf{d}_x^c + \mathbf{d}_x^a \qquad \text{és} \qquad \mathbf{d}_s = \mathbf{d}_s^c + \mathbf{d}_s^a,$$

valamint az $\mathbf{u}^{-1}$ és $\mathbf{u}$ vektorok felbontását kapjuk a következő értelemben

$$\mathbf{d}_x^c + \mathbf{d}_s^c = \mathbf{u}^{-1}, \tag{5.2}$$

$$\mathbf{d}_x^a + \mathbf{d}_s^a = -\mathbf{u}. \tag{5.3}$$

Az eredeti skálázással analóg módon definiáljuk az affin skálázású illetve a centralizáló lépésben a Newton irányokat:

$$\Delta^a\mathbf{x} := \sqrt{\mu}\,\mathbf{d}\,\mathbf{d}_x^a, \qquad \Delta^a\mathbf{s} := \sqrt{\mu}\,\mathbf{d}^{-1}\mathbf{d}_s^a,$$

$$\Delta^c\mathbf{x} := \sqrt{\mu}\,\mathbf{d}\,\mathbf{d}_x^c, \qquad \Delta^c\mathbf{s} := \sqrt{\mu}\,\mathbf{d}^{-1}\mathbf{d}_s^c.$$

Az affin skálázású lépésben θ hosszúságú Newton–lépést teszünk meg, míg a centralizáló lépésben meglépjük a teljes Newton–lépést, így az $(\mathbf{x}, \mathbf{s})$ pontból a következő pontokba lépünk:

$$\mathbf{x}^a(\theta) = \mathbf{x} + \theta\Delta^a\mathbf{x}, \qquad \mathbf{s}^a(\theta) = \mathbf{s} + \theta\Delta^a\mathbf{s},$$

$$\mathbf{x}^c = \mathbf{x} + \Delta^c\mathbf{x}, \qquad \mathbf{s}^c = \mathbf{s} + \Delta^c\mathbf{s}.$$

A fenti azonosságok alapján könnyen beláthatók az alábbi összefüggések az affin skálázású illetve a centralizáló lépésekre:

$$\mathbf{x}\,\Delta^a\mathbf{s} + \mathbf{s}\,\Delta^a\mathbf{x} = -\mathbf{x}\,\mathbf{s} \tag{5.4}$$

$$\mathbf{x}\,\Delta^c\mathbf{s} + \mathbf{s}\,\Delta^c\mathbf{x} = \mu\,\mathbf{e} \tag{5.5}$$

ugyanis $\mathbf{x}\Delta^a\mathbf{s} = \sqrt{\mu}\mathbf{x}\mathbf{d}^{-1}\mathbf{d}_s^a = \mu\mathbf{u}\mathbf{d}_s^a$, valamint $\mathbf{s}\Delta^a\mathbf{x} = \sqrt{\mu}\mathbf{s}\mathbf{d}\mathbf{d}_x^a = \mu\mathbf{u}\mathbf{d}_x^a$. Az utóbbi két egyenlőséget összeadva, és figyelembe véve (5.3) összefüggést, a (5.4) egyenlőséget kapjuk

$$\mathbf{x}\,\Delta^a\mathbf{s} + \mathbf{s}\,\Delta^a\mathbf{x} = \mu\,\mathbf{u}\,\mathbf{d}_s^a + \mu\,\mathbf{u}\,\mathbf{d}_x^a = \mu\,\mathbf{u}(-\mathbf{u}) = -\mathbf{x}\,\mathbf{s}.$$

A centralizáló lépésre vonatkozó (5.5) összefüggést hasonlóan igazolhatjuk felhasználva a definíciókat és a (5.2) egyenlőséget.

5.2.1. Lemma. *Legyen $\mathbf{x}^T\mathbf{s} = n\mu$. Tegyük fel, hogy megengedett lépéseket teszünk, ekkor a θ hosszúságú affin skálázású lépésben $(1-\theta)$-szorosára csökken a dualitás rés, míg a centralizáló lépés során megduplázódik.*

Bizonyítás. Először vizsgáljuk meg az affin skálázású lépés esetén a dualitás rést.

$$\mathbf{x}^a(\theta)\,\mathbf{s}^a(\theta) = (\mathbf{x} + \theta\Delta^a\mathbf{x})(\mathbf{s} + \theta\Delta^a\mathbf{s}) = \mathbf{x}\,\mathbf{s} + \theta(\mathbf{x}\,\Delta^a\mathbf{s} + \mathbf{s}\,\Delta^a\mathbf{x}) + \theta^2\Delta^a\mathbf{x}\,\Delta^a\mathbf{s}.$$

Felhasználva (5.4) egyenlőséget,

$$\mathbf{x}^a(\theta)\,\mathbf{s}^a(\theta) = (1-\theta)\,\mathbf{x}\,\mathbf{s} + \theta^2\Delta^a\mathbf{x}\,\Delta^a\mathbf{s} \tag{5.6}$$

adódik. Mivel $\Delta^a\mathbf{x}$ és $\Delta^a\mathbf{s}$ merőlegesek egymásra (ugyanis $\mathbf{d}_x^a$ és $\mathbf{d}_s^a$ vektorok merőlegesek a 5.1.4. Lemma bizonyításához hasonló megfontolások alapján), a következőt kapjuk

$$(\Delta^a\mathbf{x})^T\Delta^a\mathbf{s} = \mathbf{e}^T((1-\theta)\mathbf{x}\,\mathbf{s} + \theta^2\Delta^a\mathbf{x}\,\Delta^a\mathbf{s}) = (1-\theta)\mathbf{x}^T\mathbf{s},$$

amivel beláttuk az első állításunkat. A centralizáló lépésnél hasonlóan járunk el.

$$\mathbf{x}^c\,\mathbf{s}^c = (\mathbf{x} + \Delta^c\mathbf{x})(\mathbf{s} + \Delta^c\mathbf{s}) = \mathbf{x}\,\mathbf{s} + (\mathbf{x}\,\Delta^c\mathbf{s} + \mathbf{s}\,\Delta^c\mathbf{x}) + \Delta^c\mathbf{x}\,\Delta^c\mathbf{s}.$$

Figyelembe véve a (5.5) összefüggést

$$\mathbf{x}^c\,\mathbf{s}^c = \mathbf{x}\,\mathbf{s} + \mu\,\mathbf{e} + \Delta^c\mathbf{x}\,\Delta^c\mathbf{s}$$

kapjuk. Hasonlóan az előzőekhez a $\Delta^c\mathbf{x}$ és $\Delta^c\mathbf{s}$ vektorok is merőlegesek egymásra, így

$$(\mathbf{x}^c)^T\mathbf{s}^c = \mathbf{e}^T(\mathbf{x}\,\mathbf{s} + \mu\,\mathbf{e} + \Delta^c\mathbf{x}\,\Delta^c\mathbf{s}) = \mathbf{x}^T\mathbf{s} + \mu\,\mathbf{e}^T\mathbf{e} = 2n\mu,$$

ezzel igazoltuk a második állítást is. □

Az előző lemma világossá teszi, hogy a dualitás rés csökkenését az affin skálázású lépéssel biztosítjuk, míg a fenti módon definiált centralizáló lépés következtében a dualitás rés megduplázódik, ezzel elrontva az algoritmus konvergenciáját, ezért ehelyett egy teljes Newton lépést teszünk meg az aktuális μ érték felé. Ennek következtében, mint már az előző alfejezetben is láttuk (5.1.6. Lemma) a pontunk dualitásrése változatlan marad, de közelebb kerülünk a centrális úthoz. Ezen tulajdonsága miatt nevezzük ezt a lépést centralizáló lépésnek. Ennek megfelelően az affin lépést prediktor, míg a centralizáló lépést korrektor lépésnek is nevezik.

Célunk olyan megoldáspár előállítása, melynek a dualitás rése közel nulla, azaz a dualitás rést szeretnénk minél gyorsabban csökkenteni, vagyis az affin skálázású lépés hatékonysága fontos kérdése az algoritmusnak. Ezen a gondolaton alapszik a prediktor–korrektor algoritmus. Megjegyezzük, hogy ha az affin skálázású lépés esetén a teljes Newton–lépés megengedett, akkor e lépéssel egy nulla dualitásrésű ponthoz jutunk, azaz egy optimális megoldáshoz, ám ez az esetek többségében nem következik be.

Prediktor–korrektor algoritmus

Input:

τ az eltérés paraméter, $0 \leq \tau < 1$
$\varepsilon > 0$ számítási pontosság
kezdeti megoldás: $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{s}^0) \in \mathcal{P}_+ \times \mathcal{D}_+$, $(\mathbf{x}^0)^T \mathbf{s}^0 = \mu^0 \, n$ és $\delta(\mathbf{x}^0, \mathbf{s}^0, \mu^0) \leq \tau$
büntető paraméter θ, $0 < \theta < 1$

begin
  $\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$
  while $\mu \, n \geq (1-\theta)\,\varepsilon$
  begin
    Korrektor lépés (Newton lépés)
      $\mathbf{x} := \mathbf{x} + \Delta\mathbf{x}$
      $\mathbf{s} := \mathbf{s} + \Delta\mathbf{s}$
    Prediktor lépés (affin skálázási lépés)
      $\mathbf{x} := \mathbf{x} + \theta \Delta^a \mathbf{x}$
      $\mathbf{s} := \mathbf{s} + \theta \Delta^a \mathbf{s}$
      $\mu := (1-\theta)\,\mu$
  end
end.

A prediktor–korrektor algoritmus esetén ahelyett, hogy az affin skálázású és a centralizáló lépésből egy Newton–lépést készítenénk, az előző primál–duál algoritmusunk Newton–lépését két részre szedjük szét. Először egy centralizáló lépést, majd egy affin skálázású lépést teszünk meg. Mivel a teljes affin lépés nem megengedett, egy θ paraméterrel tompítjuk. A θ megfelelő megválasztásával egyrészről biztosítjuk a lépésünk megengedettségét, másrészről korlátozzuk az új pontunk túlzott eltávolodását a centrális úttól. A centralizáló lépésünk célja visszajutni a centrális út megfelelő környezetébe, ezzel biztosítva az algoritmus konvergenciáját.

Affin skálázású lépés

A továbbiakban megvizsgáljuk, hogy a pontunk centralitási mértéke hogyan változik θ függvényében az affin skálázású lépés során. Legyen $(\mathbf{x}, \mathbf{s})$ pozitív megengedett primál–duál pár. Legyen továbbá $\mu > 0$, melyre $\mathbf{x}^T\mathbf{s} = n\mu$. Az elemzésünk során használjuk az előző fejezetben definiált centralitás mértéket

$$\delta = \delta(\mathbf{x}, \mathbf{s}, \mu) = \frac{1}{2}\|\mathbf{u}^{-1} - \mathbf{u}\|, \text{ ahol} \quad \mathbf{u} = \sqrt{\frac{\mathbf{x}\,\mathbf{s}}{\mu}}.$$

A következő lemmában az $\mathbf{u}$ vektor koordinátáira adunk alsó és felső korlátot.

5.2.2. Lemma. *Legyen* $\rho(\delta) := \delta + \sqrt{1+\delta^2}$. *Ekkor*

$$\frac{1}{\rho(\delta)} \le u_i \le \rho(\delta), \qquad 1 \le i \le n.$$

Bizonyítás. A centralitás mérték definíciója és az $\mathbf{u} > 0$ miatt a következő egyenlőtlenség áll fenn

$$-2\delta u_i \le 1 - u_i^2 \le 2\delta u_i.$$

Átrendezés után

$$u_i^2 - 2\delta u_i - 1 \le 0 \le u_i^2 + 2\delta u_i - 1$$

adódik, azaz

$$(u_i - \delta)^2 - 1 - \delta^2 \le 0 \le (u_i + \delta)^2 - 1 - \delta^2.$$

Átalakítások után

$$u_i - \delta \le |u_i - \delta| \le \sqrt{1+\delta^2} \le u_i + \delta$$

kapjuk, ebből az u_i komponensre az alábbi egyenlőtlenség adódik

$$-\delta + \sqrt{1+\delta^2} \le u_i \le \delta + \sqrt{1+\delta^2} = \rho(\delta).$$

Felhasználva az alábbi azonosságot rögtön adódik a lemma állítása

$$-\delta + \sqrt{1+\delta^2} = \frac{1}{\delta + \sqrt{1+\delta^2}} = \frac{1}{\rho(\delta)}. \qquad \square$$

Az $\mathbf{u}$ vektorra adott korlátok ismeretében a θ lépéshosszra a következő elégséges feltételt kapjuk.

5.2.3. Lemma. *Legyen* $(\mathbf{x}, \mathbf{s})$ *megengedett pont, melyre* $\mathbf{x}^T\mathbf{s} = n\mu$ *és* $\delta := \delta(\mathbf{x}, \mathbf{s}, \mu) < \tau$*. Jelölje továbbá* $\mathbf{x}^+ := \mathbf{x} + \theta\Delta^a\mathbf{x}$ *és* $\mathbf{s}^+ := \mathbf{s} + \theta\Delta^a\mathbf{s}$ *az affin lépés után kapott pontot. Ekkor* $\delta^+ := \delta(\mathbf{x}^+, \mathbf{s}^+, (1-\theta)\mu) \le \tau$*, ha* θ *paraméterre teljesül az alábbi egyenlőtlenség*

$$\frac{\theta^2 n}{1-\theta} \le 2\sqrt{2}\left(\tau\sqrt{\frac{4}{\rho(\delta)^2} + 4\delta\rho(\delta)\sqrt{2} + 2\tau^2} - 2\delta\rho(\delta) - \tau^2\sqrt{2}\right).$$

Továbbá rögzített τ *mellett a fenti kifejezés jobb oldala* δ *monoton csökkenő függvénye.*

Bizonyítás. Felhasználva a (5.6) egyenlőséget a következőt kapjuk

$$\mathbf{x}^+\mathbf{s}^+ = (1-\theta)\mathbf{x}\,\mathbf{s} + \theta^2\Delta^a\mathbf{x}\,\Delta^a\mathbf{s} = \mu\left((1-\theta)\mathbf{u}^2 + \theta^2\mathbf{d}_x^a\,\mathbf{d}_s^a\right).$$

Vezessük be az alábbi jelölést

$$\mathbf{u}^+ := \sqrt{\frac{\mathbf{x}^+\mathbf{s}^+}{(1-\theta)\mu}}, \text{ ekkor } \quad (\mathbf{u}^+)^2 = \mathbf{u}^2 + \frac{\theta^2\mathbf{d}_x^a\,\mathbf{d}_s^a}{1-\theta}.$$

Az új pontunk centralitási mértékére

$$\delta^+ = \frac{1}{2}\left\|(\mathbf{u}^+)^{-1}\left(\mathbf{e} - \mathbf{u}^+\right)\right\| \le \frac{1}{2}\left\|(\mathbf{u}^+)^{-1}\right\|_\infty \left\|\mathbf{e} - (\mathbf{u}^+)^2\right\| \text{ áll fenn.}$$

További becsléshez felső korlátot adunk a fenti szorzat két tényezőjére külön-külön, először a második kifejezést vizsgáljuk.

$$\begin{aligned}\left\|\mathbf{e} - (\mathbf{u}^+)^2\right\| &= \left\|\mathbf{e} - \mathbf{u}^2 - \frac{\theta^2\mathbf{d}_x^a\,\mathbf{d}_s^a}{1-\theta}\right\| \le \left\|\mathbf{e} - \mathbf{u}^2\right\| + \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\| \\ &\le \left\|\mathbf{e} - \mathbf{u}^2\right\| + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}.\end{aligned}$$

Az utolsó egyenlőtlenség a 5.1.9. Következmény miatt igaz, figyelembe véve az $\|\mathbf{d}_x^a + \mathbf{d}_s^a\|^2 = \|\mathbf{u}\|^2 = n$ egyenlőséget. A 5.2.2. Lemmabeli eredmények segítségével kapjuk az alábbiakat

$$\left\|\mathbf{e} - \mathbf{u}^2\right\| = \left\|\mathbf{u}\,\frac{\mathbf{e} - \mathbf{u}^2}{\mathbf{u}}\right\| \le \|\mathbf{u}\|_\infty \left\|\frac{\mathbf{e} - \mathbf{u}^2}{\mathbf{u}}\right\| \le 2\delta\rho(\delta).$$

Tehát a második tényezőre a következő felső becslés teljesül:

$$\left\|\mathbf{e} - (\mathbf{u}^+)^2\right\| \le 2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}.$$

Az első tényező becsléséhez ismét használjuk fel a 5.2.2. Lemmát és a 5.1.9. Következményt.

$$(u_i^+)^2 \geq u_i^2 - \frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\|_\infty \geq \frac{1}{\rho(\delta)^2} - \frac{\theta^2 n}{4\,(1-\theta)},$$

azaz

$$\left\|(\mathbf{u}^+)^{-1}\right\|_\infty \leq \frac{1}{\sqrt{\frac{1}{\rho(\delta)^2} - \frac{\theta^2 n}{4\,(1-\theta)}}}.$$

A kapott felső korlátokat behelyettesítve a következőre jutunk

$$\delta^+ \leq \frac{2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}}{2\sqrt{\frac{1}{\rho(\delta)^2} - \frac{\theta^2 n}{4\,(1-\theta)}}}.$$

A fentiek alapján a $\delta^+ \leq \tau$ fennállásának elégséges feltétele

$$\left(2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}\right)^2 \leq \frac{4\,\tau^2}{\rho(\delta)^2} - \frac{\theta^2 n \tau^2}{1-\theta}. \tag{5.7}$$

Átrendezések után

$$\left(2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}\right)^2 + 2\,\tau^2\sqrt{2}\left(2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)}\right) \leq \frac{4\,\tau^2}{\rho(\delta)^2} + 4\,\tau^2\delta\rho(\delta)\sqrt{2},$$

$$\left(2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)} + \tau^2\sqrt{2}\right)^2 \leq \frac{4\,\tau^2}{\rho(\delta)^2} + 4\,\tau^2\delta\rho(\delta)\sqrt{2} + 2\,\tau^4.$$

Mindkét oldalból gyököt vonva

$$2\delta\rho(\delta) + \frac{\theta^2 n}{2\sqrt{2}(1-\theta)} + \tau^2\sqrt{2} \leq \tau\sqrt{\frac{4}{\rho(\delta)^2} + 4\delta\rho(\delta)\sqrt{2} + 2\,\tau^2},$$

és θ paraméterre nézve rendezve

$$\frac{\theta^2 n}{2\sqrt{2}\,(1-\theta)} \leq \tau\sqrt{\frac{4}{\rho(\delta)^2} + 4\delta\rho(\delta)\sqrt{2} + 2\,\tau^2} - 2\delta\rho(\delta) - \tau^2\sqrt{2}$$

megkaptuk a lemma első állítását. A második állítás igazolásához vegyük észre, hogy a lemmabeli egyenlőtlenség ekvivalens a (5.7) egyenlőtlenséggel, így elegendő az utóbbit vizsgálni. Könnyen belátható, hogy a (5.7) egyenlőtlenség bal oldali kifejezése mind θ, mind δ változóban monoton nő, míg a jobb oldali kifejezés mindkettő szerint monoton fogy. Azaz ha θ kielégíti az egyenlőtlenséget valamely δ érték mellett, akkor ugyanezen θ kielégíti kisebb δ értékre is. Mivel az előbb említett két egyenlőtlenség ekvivalens, ugyanez igaz a lemmabeli egyenlőtlenségre is, tehát a lemma második állítását is beláttuk. $\square$

A továbbiakban legyen $\tau = 1/2$ és $\theta = 1/(2\sqrt{n})$. Megmutatjuk, hogy minden iteráció után olyan $(\mathbf{x}, \mathbf{s})$ pontba érkezünk, melyre $\delta(\mathbf{x}, \mathbf{s}, \mu) \leq \tau$, azaz az algoritmus jól definiált.
A korrektor lépés egy teljes Newton–lépés, így a 5.1.13. Tétel alapján a kapott $(\mathbf{x}, \mathbf{s})$ pont centralitási mértéke

$$\delta := \delta(\mathbf{x}, \mathbf{s}, \mu) \leq \frac{\frac{1}{4}}{\sqrt{2\left(1 - \frac{1}{4}\right)}} = \frac{1}{\sqrt{24}}.$$

Tehát a korrektor lépés biztosítja, hogy a pontunk visszajusson a centrális út $1/\sqrt{24}$ környezetébe. Ezután vizsgáljuk meg a prediktor lépés során kapott pontot. A megadott θ érték kielégíti a 5.2.3. Lemma feltételét, ugyanis

$$\rho(\delta) = \frac{1}{\sqrt{24}} + \sqrt{1 + \frac{1}{24}} = \sqrt{\frac{3}{2}}.$$

Ezt behelyettesítve θ paraméterre a következő egyenlőtlenségnek kell teljesülnie

$$\frac{\theta^2 n}{1 - \theta} \leq 0.612626,$$

ami a megadott θ érték mellett minden $n \geq 1$ mellett fennáll. Ennek köszönhetően a 5.2.3. Lemma miatt a prediktor lépésben kapott pontunk centralitási mértéke is kisebb marad τ értékénél.

Ezek után megfogalmazhatjuk fő tételünket, mely a Prediktor–korrektor algoritmus komplexitását határozza meg. A bizonyítás a teljes Newton–lépéses primál–duál logaritmikus büntetőfüggvényes algoritmus lépésszámának bizonyításához (5.1.15. Tétel) teljesen hasonlóan megy.

5.2.4. Tétel. *Legyen $\tau = 1/2$ és $\theta = 1/(2\sqrt{n})$, ekkor a Prediktor–korrektor algoritmus legfeljebb*

$$\left\lceil 2\sqrt{n} \log \frac{n\mu^0}{\varepsilon} \right\rceil$$

lépésben előállít egy olyan $\mathbf{x} \in \mathcal{P}_+$, $(\mathbf{y}, \mathbf{s}) \in \mathcal{D}_+$ primál-duál pontpárt, amelyre $\mathbf{x}^T\mathbf{s} \leq \varepsilon$. $\square$

A prediktor–korrektor algoritmus adaptív változata

A prediktor–korrektor algoritmus hatékonyságát tovább növelhetjük, ha a μ paraméter iterálási stratégiáját megváltoztatjuk. A fenti algoritmusban a θ értéke egy fix szám. Ezzel szemben, ha az algoritmus során a pontunktól függően változtatjuk a θ értékét, azaz a μ paramétert, akkor egy hatékonyabb, gyorsabb algoritmust kapunk.

Először a θ megengedett értékére adunk egy jobb korlátot a 5.2.3. Lemma eredményének a javításával.

5.2.5. Lemma. *Legyen $(\mathbf{x},\mathbf{s})$ megengedett pont, melyre $\mathbf{x}^T\mathbf{s} = n\mu$ és $\delta := \delta(\mathbf{x},\mathbf{s},\mu) < \tau$. Jelölje továbbá $\mathbf{x}^+ := \mathbf{x} + \theta\Delta^a\mathbf{x}$ és $\mathbf{s}^+ := \mathbf{s} + \theta\Delta^a\mathbf{s}$ az affin lépés után kapott pontot. Ekkor $\delta^+ := \delta(\mathbf{x}^+,\mathbf{s}^+,(1-\theta)\mu) \leq \tau$, ha θ paraméterre teljesül az alábbi egyenlőtlenség*

$$\frac{\theta^2}{1-\theta}\,\|\,\mathbf{d}_x^a\,\mathbf{d}_s^a\,\| \leq 2\,\tau\left(\sqrt{\frac{1}{\rho(\delta)^2} + 2\delta\rho(\delta) + \tau^2} - \tau\right) - 2\delta\rho(\delta).$$

Bizonyítás. A bizonyítás szinte megegyezik a 5.2.3. Lemma bizonyításával, felhasználjuk az ottani eredményeket és jelöléseket. Mint már láttuk az affin lépés után a pontunk centralitási mértéke

$$\delta^+ = \frac{1}{2}\left\|(\mathbf{u}^+)^{-1}\left(\mathbf{e} - \mathbf{u}^+)\right)\right\| \leq \frac{1}{2}\left\|(\mathbf{u}^+)^{-1}\right\|_\infty \left\|\mathbf{e} - (\mathbf{u}^+)^2\right\|.$$

Továbbá

$$\left\|\mathbf{e} - (\mathbf{u}^+)^2\right\| \leq \left\|\mathbf{e} - \mathbf{u}^2\right\| + \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|,$$

és

$$\left\|\mathbf{e} - \mathbf{u}^2\right\| \leq 2\delta\rho(\delta).$$

A centralitási mérték felső korlátjában szereplő első tényező becslését is élesítjük.

$$(u_i^+)^2 \geq u_i^2 - \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|_\infty \geq u_i^2 - \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|.$$

Behelyettesítve a felső korlátokat

$$\delta^+ \leq \frac{2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|}{2\sqrt{\frac{1}{\rho(\delta)^2} - \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|}}.$$

A $\delta^+ \leq \tau$ teljesül, ha

$$\left(2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|\right)^2 \leq \frac{4\,\tau^2}{\rho(\delta)^2} - \frac{4\,\theta^2\tau^2}{1-\theta}\left\|\mathbf{d}_x^a\,\mathbf{d}_s^a\right\|.$$

Átrendezés után kapjuk a következőt

$$\left(2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\|\right)^2 + 4\,\tau^2\left(2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\|\right) \le \frac{4\,\tau^2}{\rho(\delta)^2} + 8\,\tau^2\delta\rho(\delta),$$

azaz

$$\left(2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\| + 2\,\tau^2\right)^2 \le \frac{4\,\tau^2}{\rho(\delta)^2} + 8\,\tau^2\delta\rho(\delta) + 4\,\tau^4.$$

Négyzetgyököt vonva

$$2\delta\rho(\delta) + \frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\| + 2\,\tau^2 \le 2\,\tau\sqrt{\frac{1}{\rho(\delta)^2} + 2\delta\rho(\delta) + \tau^2}$$

egyenlőtlenséghez jutunk, melyet a θ változóra rendezve megkapjuk a lemma állítását.

□

A következő tételben adunk egy, a $\tau = \frac{1}{3}$ értékhez tartozó megfelelő θ paramétert, azaz egy olyan választást, mely mellett az algoritmus jól definiált, vagyis minden egyes iterációban olyan pontba jutunk, amely benne van a centrális út τ környezetében.

5.2.6. Tétel. *Legyen $\tau = \frac{1}{3}$. Ekkor a*

$$\theta = \frac{2}{1 + \sqrt{1 + 13\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\|}}$$

választás esetén minden iteráció után fennáll a $\delta(\mathbf{x}, \mathbf{s}, \mu) \le \tau$ egyenlőtlenség.

Bizonyítás. Induljunk el egy $(\mathbf{x}, \mathbf{s})$ pontból, melyre $\delta(\mathbf{x}, \mathbf{s}, \mu) \le \frac{1}{3}$. Ekkor a centralizáló lépés után a 5.1.13. Lemma miatt egy olyan $(\overline{\mathbf{x}}, \overline{\mathbf{s}}, \mu)$ pontba jutunk, melyre a centralitási mérték

$$\delta := \delta(\overline{\mathbf{x}}, \overline{\mathbf{s}}, \mu) \le \frac{\frac{1}{9}}{\sqrt{2\left(1 - \frac{1}{9}\right)}} = \frac{1}{12}.$$

A 5.2.5. Lemma alapján ha a θ kielégíti a lemma feltételét, akkor az affin lépéssel kapott pont centralitási mértéke se fogja meghaladni a τ értékét. Az említett feltétel a θ paraméterre az ismert értékek behelyettesítése után a következő

$$\frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\| \le 0,308103.$$

Egyszerű behelyettessel igazolható, hogy

$$\frac{\theta^2}{1-\theta}\,\|\mathbf{d}_x^a\,\mathbf{d}_s^a\| \le \frac{4}{13},$$

azaz a feltétel teljesül. Ezzel a tételt beláttuk. □

Kvadratikus konvergencia

Könnyen látható, hogy a konvergencia rendje θ paraméter választásától függ. A kvadratikus konvergencia feltétele

$$(1-\theta)\mu = \mathcal{O}(\mu^2), \qquad \text{azaz} \qquad 1-\theta = \mathcal{O}(\mu).$$

Megmutatjuk, hogy az előző adaptív algoritmus estén megadott θ érték teljesíti ezt a feltételt, azaz a fent tárgyalt adaptív prediktor–korrektor algoritmus konvergenciája kvadratikus.

5.2.7. Lemma. *A büntető paraméter* $\theta = \frac{2}{1+\sqrt{1+13\|\mathbf{d}_x^a \mathbf{d}_s^a\|}}$ *választása esetén*

$$1-\theta \leq \frac{13}{4}\|\mathbf{d}_x^a\, \mathbf{d}_s^a\| \quad \textit{áll fenn.}$$

Azaz az adaptív prediktor–korrektor algoritmus konvergenciája kvadratikus, ha $\|\mathbf{d}_x^a \mathbf{d}_s^a\| = \mathcal{O}(\mu)$ *teljesül.*

Bizonyítás. Vizsgáljuk meg az

$$f : [0,\infty) \to \mathbb{R}_+ \qquad\qquad f(x) = 1 - \frac{2}{1+\sqrt{1+13x}} \quad \text{függvényt.}$$

Ehhez számoljuk ki a függvény első és második deriváltját

$$f'(x) = \frac{13}{\sqrt{1+13x}(1+\sqrt{1+13x})^2} \quad \text{illetve} \quad f''(x) = \frac{-169\,(1+3\sqrt{1+13x})}{(1+13x)^{\frac{3}{2}}(1+\sqrt{1+13x})^3}.$$

Ezek alapján f monoton növő konkáv függvény, és mivel $f'(0) = 13/4$, ezért $f(x) \leq 13x/4$ fennáll minden $x \geq 0$ számra. Alkalmazva a kapott eredményt $x = \|\mathbf{d}_x^a \mathbf{d}_s^a\|$ értékre kapjuk a lemma állítását. □

A továbbiakban az egyes bevezetett vektorok nagyságrendjét fogjuk megbecsülni, pontosabban a koordinátáikat aszerint, hogy az indexük a B vagy az N halmazba esik. Az állításokat a következő táblázatban foglalhatjuk össze

Először az $\mathbf{u}$ vektort vizsgáljuk meg. Figyelembe véve, hogy $\delta \leq 1/2$ a 5.2.2. Lemma alapján

$$0,61833 \leq \frac{1}{\rho(\delta)} \leq u_i \leq \rho(\delta) \leq 1,618033\ , \quad \text{azaz} \quad \mathbf{u} = \Theta(1).$$

	vektor	B	N
1	$\mathbf{x}$	$\Theta(1)$	$\Theta(\mu)$
2	$\mathbf{s}$	$\Theta(\mu)$	$\Theta(1)$
3	$\mathbf{u}$	$\Theta(1)$	$\Theta(1)$
4	$\mathbf{d}$	$\Theta(\frac{1}{\sqrt{\mu}})$	$\Theta(\sqrt{\mu})$
5	$\mathbf{d}_x^a$	$\mathcal{O}(\mu)$	$\mathcal{O}(1)$
6	$\mathbf{d}_s^a$	$\mathcal{O}(1)$	$\mathcal{O}(\mu)$
7	$\Delta^a\mathbf{x}$	$\mathcal{O}(\mu)$	$\mathcal{O}(\mu)$
8	$\Delta^a\mathbf{s}$	$\mathcal{O}(\mu)$	$\mathcal{O}(\mu)$

Ezután foglakozzunk az $\mathbf{x}$ és $\mathbf{s}$ vektorok becslésével. Legyen σ a (P) és D feladatpár kondíciószáma[1], ekkor a 2.6.4. Lemmához teljesen hasonlóan belátható

$$\sigma_i \leq \frac{n\mu}{\sigma}, \qquad \text{tehát} \quad s_B = \mathcal{O}(\mu).$$

Ennek segítségével, felhasználva az $\mathbf{x}_B\mathbf{s}_B = \mu\mathbf{u}_B^2 = \Theta(\mu)$ összefüggést

$$(\mathbf{x}_B)^{-1} = \frac{\mathbf{s}_B}{\Theta(\mu)} = \frac{\mathcal{O}(\mu)}{\Theta(\mu)} = \mathcal{O}(1),$$

tehát az $\mathbf{x}_B$ elválasztható a nullától, azaz létezik egy pozitív szám, ami alulról korlátozza. Másrészről az olyan $(\mathbf{x}, \mathbf{s})$ pontok halmaza melyekre a dualitás rés $n\mu$ korlátos, így $\mathbf{x}_B$ felülről is korlátos, tehát $\mathbf{x}_B = \Theta(1)$. Továbbá mivel $\mathbf{x}_B\mathbf{s}_B = \Theta(\mu)$ kapjuk, hogy $\mathbf{s}_B = \Theta(\mu)$. Hasonlóan bizonyíthatók az $\mathbf{x}_N = \Theta(\mu)$ és $\mathbf{s}_N = \Theta(1)$ összefüggések.

A $\mathbf{d}$ vektor koordinátáinak nagysága közvetlenül következik a $\mathbf{d}$ vektor definíciójából és az $\mathbf{x}$ és $\mathbf{s}$ vektorokra adott becslésekből.

Definíció szerint $\mathbf{d}_x^a$ a $-\mathbf{u}$ vektor vetülete az AD mátrix nullterére, így a projekció tulajdonsága miatt

$$\|\mathbf{d}_x^a\| = \|P_{AD}(-\mathbf{u})\| \leq \|\mathbf{u}\| = \sqrt{n}, \quad \text{azaz} \quad \mathbf{d}_x^a = \mathcal{O}(1).$$

[1]Megjegyezzük, hogy a kondíciószám csak a feladat adataitól függ, azaz független a μ paramétertől, minden iteráció esetén ugyanaz.

Mivel $(\mathbf{d}_x^a)_N$ a $\mathbf{d}_x^a$ vektor részvektora, ezért $(\mathbf{d}_x^a)_N = \mathcal{O}(1)$ teljesül. Hasonlóan igazolható a $(\mathbf{d}_s^a)_B = \mathcal{O}(1)$. A $(\mathbf{d}_x)_B$ illetve a $(\mathbf{d}_s)_N$ esetén ennél többet szeretnénk belátni, ezért további elemzéseket teszünk. A $\mathbf{d}_x^a$ definíciója és a projekció linearitása miatt

$$\mathbf{d}_x^a = -P_{AD}(\mathbf{u}) = -\frac{1}{\sqrt{\mu}} P_{AD}(\sqrt{\mathbf{x}\,\mathbf{s}}) = -\frac{1}{\sqrt{\mu}} P_{AD}(\mathbf{d}\,\mathbf{s}). \tag{5.8}$$

Legyen $(\tilde{\mathbf{y}}, \tilde{\mathbf{s}})$ optimális duál megoldás. Ekkor

$$\mathbf{s} = \mathbf{c} - A^T\mathbf{y} = A^T\tilde{\mathbf{y}} + \tilde{\mathbf{s}} - A^T\mathbf{y} = \tilde{\mathbf{s}} + A^T(\tilde{\mathbf{y}} - \mathbf{y}),$$

így $\mathbf{d}\,\mathbf{s} = \mathbf{d}\,\tilde{\mathbf{s}} + (AD)^T(\tilde{\mathbf{y}} - \mathbf{y})$, azaz a $\mathbf{d}\,\mathbf{s} - \mathbf{d}\,\tilde{\mathbf{s}}$ vektor az AD mátrix sorterében van. Mivel a sortér merőleges a nulltérre, $P_{AD}(\mathbf{d}\,\mathbf{s}) = P_{AD}(\mathbf{d}\,\mathbf{s} + (\mathbf{d}\,\tilde{\mathbf{s}} - \mathbf{d}\,\mathbf{s})) = P_{AD}(\mathbf{d}\,\tilde{\mathbf{s}})$. Ezt behelyettesítve a (5.8) összefüggésbe

$$\mathbf{d}_x^a = -\frac{1}{\sqrt{\mu}} P_{AD}(\mathbf{d}\,\tilde{\mathbf{s}}) \tag{5.9}$$

kapjuk. Mivel $\tilde{\mathbf{s}}$ optimális, ezért $\tilde{\mathbf{s}}_B = \mathbf{0}$, így a (5.9) összefüggéssel ekvivalens a következő

$$-\sqrt{\mu}\,\mathbf{d}_x^a = \arg\min_{\mathbf{h}} \left(\|\mathbf{d}\,\tilde{\mathbf{s}} - \mathbf{h}\|^2 \ : \ AD\mathbf{h} = \mathbf{0} \right), \quad \text{illetve}$$

$$-\sqrt{\mu}\,\mathbf{d}_x^a = \arg\min_{\mathbf{h}} \left(\|\mathbf{d}_B\,\tilde{\mathbf{s}}_B - \mathbf{h}_B\|^2 + \|\mathbf{d}_N\,\tilde{\mathbf{s}}_N - \mathbf{h}_N\|^2 \ : \ A_B D_B \mathbf{h}_B + A_N D_N \mathbf{h}_N = \mathbf{0} \right),$$

és $\mathbf{h}_B = -\sqrt{\mu}\,(\mathbf{d}_x^a)_B$, valamint $\mathbf{h}_N = -\sqrt{\mu}\,(\mathbf{d}_x^a)_N$ a minimalizálás megoldása. Behelyettesítés után a következőt kapjuk

$$-\sqrt{\mu}\,(\mathbf{d}_x^a)_B = \arg\min_{\mathbf{h}_B} \left(\|\mathbf{h}_B\|^2 \ : \ A_B D_B \mathbf{h}_B = \sqrt{\mu}\,A_N D_N (\mathbf{d}_x^a)_N \right),$$

ami azt jelenti, hogy a $-\sqrt{\mu}\,(\mathbf{d}_x^a)_B$ a legkisebb normájú vektor az

$$\mathcal{S} := \{\xi : A_B D_B \xi = \sqrt{\mu}\,A_N D_N (\mathbf{d}_x^a)_N\}$$

affin altérben. Az $A_B\mathbf{z} = \sqrt{\mu}\,A_N D_N (\mathbf{d}_x^a)_N$ egyenlet legkisebb normájú megoldása az $\mathbf{z}^* = \sqrt{\mu} A_B^+ A_N D_N (\mathbf{d}_x^a)_N$ vektor, ahol A_B^+ jelöli az A_B mátrix pseudo inverzét (azaz jelen esetben $A_B^+ = A_B^T (A_B A_B^T)^{-1}$, bővebben lásd [28]). Nyilvánvaló, hogy $D_B^{-1}\mathbf{z}^* \in \mathcal{S}$, így

$$\|\sqrt{\mu}\,(\mathbf{d}_x^a)_B\| \leq \left\|D_B^{-1} z^*\right\| = \sqrt{\mu}\left\|D_B^{-1} A_B^+ A_N D_N (\mathbf{d}_x^a)_N\right\|,$$

azaz

$$\|(\mathbf{d}_x^a)_B\| \leq \left\|D_B^{-1}\right\| \left\|A_B^+\right\| \|A_N\| \, \|D_N\| \, \|(\mathbf{d}_x^a)_N\|,$$

ahol $\left\|D_B^{-1}\right\|$, $\|D_N\|$ $\Theta(\sqrt{\mu})$ nagyságrendű, $\left\|A_B^{+}\right\|$, $\|A_N\|$ és $\|(\mathbf{d}_x^a)_N\|$ pedig $\mathcal{O}(1)$. Ezzel beláttuk, hogy $(\mathbf{d}_x^a)_B = \mathcal{O}(\sqrt{\mu})$. Hasonlóan adódik $(\mathbf{d}_s^a)_N$ becslése is.

A $\Delta^a\mathbf{x}$ és $\Delta^a\mathbf{s}$ vektorok esetén a bizonyítás egyszerűen következik a

$$\Delta^a\mathbf{x} = \sqrt{\mu}\,\mathbf{d}\,\mathbf{d}_x \qquad \text{és} \qquad \Delta^a\mathbf{s} = \sqrt{\mu}\,\mathbf{d}^{-1}\mathbf{d}_s$$

összefüggésekből és az előző eredményekből.

Ezek után megfogalmazhatjuk az alfejezetünk fő tételét

5.2.8. Tétel. *Az adaptív prediktor–korrektor algoritmus kvadratikusan konvergens.*

Bizonyítás. Az előzőekben belátottak alapján $\mathbf{d}_x^a\,\mathbf{d}_s^a = \mathcal{O}(\mu)$, ez pedig a 5.2.7. Lemma szerint pont azt jelenti, hogy az algoritmus kvadratikus. $\square$

6. fejezet

Self-reguláris belsőpontos algoritmusok

A következőkben ismertetett elmélet kiindulópontja az előző fejezet bevezetőjében említett ellentmondás, miszerint a rövid lépéses algoritmusok $\mathcal{O}(\sqrt{n}\log\frac{n}{\varepsilon})$ bizonyított komplexitása jobb, mint a hosszú lépéses algoritmusok esetén ismert $\mathcal{O}(n\log\frac{n}{\varepsilon})$, ezzel szemben a gyakorlatban a hosszú lépéses módszerek hatékonyabbnak bizonyultak. Terlaky, Roos és Peng [19] ennek a feloldása céljából a hosszú lépéses eljárásokat új centralitási mérték bevezetése mellett vizsgálta meg. Az új mértékeket a self-reguláris függvények segítségével definiálták. A megváltoztatott centralitási mértéknek megfelelően a Newton–rendszert is módosították. Az így kapott új algoritmus tekinthető az előző fejezetben bevezetett δ_1 centralitási mérték illetve a hozzá tartozó Newton–rendszer általánosításának. A fejezetben a self-reguláris algoritmusok egy speciális esete kerül ismertetésre Peng, Roos és Terlaky [20] alapján, melyben a módosított eljárás komplexitása $\mathcal{O}(\sqrt{n}\log n\log\frac{n}{\varepsilon})$, amely a rövid lépéses algoritmusok komplexitását igen jól megközelíti.

6.1. Self-reguláris algorimtus

6.1.1. Definíció. *Egy $\psi : (0,\infty) \to \mathbb{R}$ kétszer folytonosan deriválható függvényt* self-reguláris függvény*nek nevezünk, ha teljesíti a következő feltételeket:*

1. *$\psi(t)$ szigorúan konvex függvény és a globális minimumpontjában, $t = 1$ pontban eltűnik, azaz $\psi(1) = \psi'(1) = 0$. Továbbá léteznek $\nu_1 \geq \nu_2 > 0$ és $p \geq 1$, $q \geq 1$*

konstansok úgy, hogy

$$\nu_1(t^{p-1} + t^{-1-q}) \le \psi''(t) \le \nu_2(t^{p-1} + t^{-1-q}), \quad \forall t \in (0, \infty)$$

2. *Bármely* $t_1, t_2 > 0$ *esetén*

$$\psi(t_1^r t_2^{1-r}) \le r\psi(t_1) + (1-r)\psi(t_2), \quad \forall r \in [0,1].$$

Ha a ψ self-reguláris függvény, akkor a q paramétert büntetés mértékének (a szám nullához tartását büntetjük), míg a p paramétert növekedés mértékének nevezzük.

Két speciális családját emelnénk ki a self-reguláris függvényeknek. Az első családot definiálja

$$\Upsilon_{p,q}(t) = \frac{t^{p+1}-1}{p(p+1)} + \frac{t^{1-q}-1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad \text{ahol } p \ge 1,\ q > 1,$$

ekkor a definícióban szereplő konstansok $\nu_1 = \nu_2 = 1$. A második család

$$\Gamma_{p,q}(t) = \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, \quad \text{ahol } p \ge 1,\ q > 1,$$

és a konstansok $\nu_1 = 1$ illetve $\nu_2 = q$.
Ezen függvények segítségével definiáljuk az eljárás során használt büntetőfüggvényt

$$\Psi(\mathbf{u}) = \sum_{i=1}^{n} \psi(u_i)$$

egyenlőséggel.

Ebben a fejezetben a második családba tartozó

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

függvénnyel foglalkozunk. Ekkor a büntetőfüggvény

$$\Psi_q(\mathbf{x}, \mathbf{s}, \mu) = \Psi_q(\mathbf{u}) = \frac{\mathbf{e}^T\mathbf{u}^2 - n}{2} + \frac{\mathbf{e}^T\mathbf{u}^{1-q} - n}{q-1}.$$

A Ψ függvény első tagja az $\mathbf{u}$ vektor koordinátáinak növekedését, míg a második tag a nullához tartásukat bünteti.

Az előző fejezethez hasonlóan definiáljuk a Newton–rendszert, és annak átskálázását, csak a Newton irányt változtatjuk meg a következőképpen:

$$
\begin{array}{rcrcl}
 & & AD\,\mathbf{d}_x & = & \mathbf{0}, \\
(AD)^T\mathbf{d}_y & + & \mathbf{d}_s & = & \mathbf{0}, \\
 & & \mathbf{d}_x + \mathbf{d}_s & = & -\nabla\Psi(\mathbf{u}) = \mathbf{u}^{-q} - \mathbf{u}.
\end{array}
$$

Self–reguláris primál–duál belsőpontos algoritmus

Input:

τ az eltérés paraméter, $0 \le \tau < 1$
$\varepsilon > 0$ számítási pontosság
kezdeti megoldás: $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{s}^0) \in \mathcal{P}_+ \times \mathcal{D}_+$, $(\mathbf{x}^0)^T\mathbf{s}^0 = \mu^0\, n$
büntető paraméter θ, $0 < \theta < 1$

begin
  $\mathbf{x} := \mathbf{x}^0$, $\mathbf{s} := \mathbf{s}^0$, $\mu := \mu^0$
  while $\mu\, n \ge \varepsilon$
  begin
    $\mu := (1-\theta)\,\mu$
    while
      $\mathbf{x} := \mathbf{x} + \alpha\Delta\mathbf{x}$
      $\mathbf{s} := \mathbf{s} + \alpha\Delta\mathbf{s}$
    end
  end
end.

Az új büntetőfüggvényünknek és az új Newton–rendszernek megfelelően definiálja az elemzésnél használt centralitási mértéket

$$\sigma(\mathbf{u}) := \|\nabla\Psi(\mathbf{u})\| = \left\|\mathbf{u}^{-q} - \mathbf{u}\right\| = \|\mathbf{d}_x + \mathbf{d}_s\| = \|(\mathbf{d}_x, \mathbf{d}_s)\| \tag{6.1}$$

kifejezés, ahol az utolsó egyenlőség $\mathbf{d}_x$ és $\mathbf{d}_s$ vektorok merőlegessége miatt igaz (ezt az előző fejezetben már beláttuk).

Korlátok u vektorra és a lépéshosszra

Első lépésként alsó és felső korlátot adunk az $\mathbf{u}$ vektor koordinátáira.

6.1.2. Lemma. *Legyen $\sigma = \sigma(\mathbf{u})$. Ekkor*

$$u_{\min} \geq (1+\sigma)^{-\frac{1}{q}}, \qquad u_{\max} \leq 1+\sigma.$$

Bizonyítás: A lemma két állítása triviális, ha $u_{\min} \geq 1$, illetve $u_{\max} \leq 1$. Először vizsgáljuk meg a $u_{\min} < 1$ esetet, ekkor

$$\sigma = \left\|\mathbf{u}^{-q} - \mathbf{u}\right\| \geq \|u_i^{-q} - u_i\|, \quad 1 \leq i \leq n,$$

azaz

$$\sigma \geq \frac{1}{u_{\min}^q} - u_{\min} \geq \frac{1}{u_{\min}^q} - 1.$$

Másrészről, ha $u_{\max} > 1$,

$$\sigma \geq u_{\max} - \frac{1}{u_{\max}^q} \geq u_{\max} - 1.$$

A lemma következik a két kapott egyenlőség átrendezésével. □

Vezessük be a következő jelöléseket:

$$\Delta_x = \frac{\Delta \mathbf{x}}{\mathbf{x}} = \frac{\mathbf{d}_x}{\mathbf{u}}, \qquad \Delta_s = \frac{\Delta \mathbf{s}}{\mathbf{s}} = \frac{\mathbf{d}_s}{\mathbf{u}}.$$

Továbbá jelölje az új pontunkat az α paraméterrel tompított Newton lépés után $(\mathbf{x}^+, \mathbf{s}^+)$, azaz $\mathbf{x}^+ = \mathbf{x} + \alpha\, \Delta\mathbf{x}$ és $\mathbf{s}^+ = \mathbf{s} + \alpha\, \Delta\mathbf{s}$. Ekkor a következőt írhatjuk

$$\mathbf{x}^+ = \mathbf{x}\,(\mathbf{e} + \alpha\,\Delta_x), \qquad \mathbf{s}^+ = \mathbf{s}\,(\mathbf{e} + \alpha\,\Delta_s).$$

Könnyen látható, hogy az α pontosan akkor megengedett lépéshossz, ha $\mathbf{e} + \alpha\,\Delta_x \geq 0$ és $\mathbf{e} + \alpha\,\Delta_s \geq 0$ teljesül, amiknek elégséges feltétele

$$1 - \alpha\,\|(\Delta_x, \Delta_s)\| \geq 0. \tag{6.2}$$

Ezután már adhatunk alsó becslést a legnagyobb megengedett lépéshosszra.

6.1.3. Lemma.

$$\|(\Delta_x, \Delta_s)\| \leq \sigma(1+\sigma)^{\frac{1}{q}}.$$

Valamint az $\alpha_{\max}$ legnagyobb megengedett lépéshosszra fennáll az

$$\alpha_{\max} \geq \frac{1}{\sigma(1+\sigma)^{\frac{1}{q}}}$$

egyenlőtleneg.

Bizonyítás: A 6.1.2. Lemma és σ (6.1) definíciója alapján

$$\|(\Delta_x, \Delta_s)\| = \left\|\left(\frac{\mathbf{d}_x}{\mathbf{u}}, \frac{\mathbf{d}_s}{\mathbf{u}}\right)\right\| \leq \frac{\|(\mathbf{d}_x, \mathbf{d}_s)\|}{u_{\min}} = \frac{\sigma}{u_{\min}} \leq \sigma(1+\sigma)^{\frac{1}{q}}.$$

A (6.2) egyenlőtlenséget átrendezve

$$\alpha_{\max} \geq \frac{1}{\|(\Delta_x, \Delta_s)\|}$$

alsó becslést kapjuk, amiből az elkőző egyenlőtlenség figyelembevételével következik a lemma második állítása. □

Ezután vizsgáljuk meg a büntetőfüggvény csökkenési mértékét. Jelölje a Ψ büntetőfüggvény második tagját, azaz a nullához tartást büntető részt

$$\Psi_0(\mathbf{u}) = \sum_{i=1}^{n} \psi_0(u_i), \text{ ahol} \qquad \psi_0(t) = \frac{t^{1-q}-1}{q-1}.$$

Ezekkel a jelölésekkel a büntetőfüggvényünk magfüggvénye a következő alakban írható

$$\psi(t) = \frac{t^2-1}{2} + \psi_0(t) \tag{6.3}$$

Egy Newton lépés megtétele után jelölje $\mathbf{u}$ vektor új értékét $\mathbf{u}^+ = \sqrt{\mathbf{x}^+\mathbf{s}^+/\mu}$. Ekkor

$$\begin{aligned}(\mathbf{u}^+)^2 &= \frac{(\mathbf{x}+\alpha\Delta\mathbf{x})(\mathbf{s}+\alpha\Delta\mathbf{s})}{\mu} = \frac{1}{\mu}\frac{\mathbf{x}}{\mathbf{u}}\left(\mathbf{u}+\alpha\frac{\mathbf{u}\,\Delta\mathbf{x}}{\mathbf{x}}\right)\frac{\mathbf{s}}{\mathbf{u}}\left(\mathbf{u}+\alpha\frac{\mathbf{u}\,\Delta\mathbf{s}}{\mathbf{s}}\right)\\ &= (\mathbf{u}+\alpha\,\mathbf{d}_x)(\mathbf{u}+\alpha\,\mathbf{d}_s).\end{aligned}$$

Figyelembe véve a $\mathbf{d}_x$ és $\mathbf{d}_s$ vektorok ortogonalitását

$$\mathbf{e}^T(\mathbf{u}^+)^2 = \mathbf{e}^T\mathbf{u}^2 + \alpha\,\mathbf{u}^T(\mathbf{d}_x+\mathbf{d}_s) = \mathbf{e}^T\mathbf{u}^2 + \alpha\,\mathbf{u}^T(\mathbf{u}^{-q}-\mathbf{u}).$$

A büntetőfüggvény értéke az új pontban

$$\begin{aligned}\Psi(\mathbf{x}^+, \mathbf{s}^+, \mu) &= \Psi(\mathbf{u}^+) = \frac{\mathbf{e}^T(\mathbf{u}^+)^2 - n}{2} + \Psi_0(\mathbf{u}^+) \\ &= \frac{\mathbf{e}^T\mathbf{u}^2 + \alpha\, \mathbf{u}^T(\mathbf{u}^{-q} - \mathbf{u}) - n}{2} + \Psi_0\left(\sqrt{(\mathbf{u} + \alpha\, \mathbf{d}_x)(\mathbf{u} + \alpha\, \mathbf{d}_s)}\right) \qquad (6.4)\end{aligned}$$

A következő lemmában egy fontos tulajdonságát adjuk meg a Ψ_0 függvénynek.

6.1.4. Lemma. *Legyen $t_1 > 0$ és $t_2 > 0$. Ekkor*

$$\psi_0\left(\sqrt{t_1 t_2}\right) \leq \frac{1}{2}\left(\psi(t_1) + \psi_0(t_2)\right).$$

Bizonyítás: Könnyen belátható, hogy az állítás pontosan akkor igaz, ha a $\psi_0(e^z)$ mint z függvénye konvex, ami akkor és csak akkor áll fenn, ha $\psi_0'(t) + t\,\psi_0''(t) \geq 0$ teljesül minden $t \geq 0$ esetén. Mivel

$$\psi_0'(t) = -t^{-q}, \qquad \psi_0''(t) = q\, t^{-1-q},$$

így $\psi_0'(t) + t\,\psi_0''(t) = (q-1)t^{-q} > 0$, tehát a lemmát beláttuk. □

A lemma eredményének segítségével felső becslést adhatunk a ψ_0 függvényre

$$\Psi_0(\mathbf{u}^+) = \Psi\left(\sqrt{(\mathbf{u} + \alpha\, \mathbf{d}_x)(\mathbf{u} + \alpha\, \mathbf{d}_s)}\right) \leq \frac{1}{2}\sum_{i=1}^{n}\Big(\psi_0(u_i + \alpha\, d_{xi}) + \psi_0(u_i + \alpha\, d_{si})\Big).$$

Behelyettesítve a fenti becslést a (6.4) egyenlőségbe

$$\Psi(\mathbf{x}^+, \mathbf{s}^+, \mu) \leq \frac{\mathbf{e}^T\mathbf{u}^2 + \alpha\, \mathbf{u}^T(\mathbf{u}^{-q} - \mathbf{u}) - n}{2} + \frac{1}{2}\sum_{i=1}^{n}\Big(\psi_0(u_i + \alpha\, d_{xi}) + \psi_0(u_i + \alpha\, d_{si})\Big)$$

egyenlőtlenség adódik. Legyen $f(\alpha)$ a Ψ függvény megváltozása egy lépés során a lépéshossz függvényében, azaz

$$f(\alpha) := \Psi(\mathbf{x}^+, \mathbf{s}^+, \mu) - \Psi(\mathbf{x}, \mathbf{s}, \mu) \leq f_1(\alpha),$$

ahol

$$f_1(\alpha) := -\Psi(\mathbf{x}, \mathbf{s}, \mu) + \frac{\mathbf{e}^T\mathbf{u}^2 + \alpha\, \mathbf{u}^T(\mathbf{u}^{-q} - \mathbf{u}) - n}{2} + \frac{1}{2}\sum_{i=1}^{n}\Big(\psi_0(u_i + \alpha d_{xi}) + \psi_0(u_i + \alpha d_{si})\Big).$$

Vegyük észre, hogy $f(0) = f_1(0) = 0$. Számoljuk ki a az f_1 függvény első illetve második deriváltját α lépéshosszra nézve.

$$f_1'(\alpha) = \frac{1}{2}\mathbf{u}^T(\mathbf{u}^{-q} - \mathbf{u}) + \frac{1}{2}\sum_{i=1}^{n}\Big(\psi_0'(u_i + \alpha d_{xi})d_{xi} + \psi_0'(u_i + \alpha d_{si})d_{si}\Big).$$

Felhasználva a Newton–rendszer harmadik egyenletét valamint a σ (6.1) definícióját

$$f_1'(0) = \frac{1}{2}\mathbf{u}^T(\mathbf{u}^{-q} - \mathbf{u}) - \frac{1}{2}\sum_{i=1}^{n} u_i^{-q}(d_{xi} + d_{si}) = \frac{1}{2}(\mathbf{d}_x + \mathbf{d}_s)^T(\mathbf{u} - \mathbf{u}^{-q}) = -\frac{1}{2}\sigma^2. \quad (6.5)$$

Továbbá felhasználva, hogy a $\psi''(t) = qt^{-1-q}$ függvény monoton csökken

$$\begin{aligned} f_1''(\alpha) &= \frac{1}{2}\sum_{i=1}^{n}\Big(\psi_0''(u_i + \alpha d_{xi})d_{xi}^2 + \psi_0''(u_i + \alpha d_{si})d_{si}^2\Big) \\ &= \frac{1}{2}\sum_{i=1}^{n}\Big(\left(q(u_i + \alpha d_{xi})^{-q-1}\right)d_{xi}^2 + \left(q(u_i + \alpha d_{si})^{-1-q}\right)d_{si}^2\Big) \\ &\leq \frac{q}{2}\sum_{i=1}^{n}(u_{\min} - \alpha\sigma)^{-q-1}(d_{xi}^2 + d_{si}^2). \end{aligned}$$

Az utolsó egyenlőtlenségnél a következőket használtuk fel

$$u_i + \alpha d_{xi} \geq u_{\min} - \alpha\left\|\mathbf{d}_x\right\| \geq u_{\min} - \alpha\sigma,$$

$$u_i + \alpha d_{si} \geq u_{\min} - \alpha\left\|\mathbf{d}_s\right\| \geq u_{\min} - \alpha\sigma.$$

Figyelembe véve (6.1) azonosságot

$$f_1''(\alpha) \leq h(\alpha) := \frac{1}{2}q\sigma^2(u_{\min} - \alpha\sigma)^{-1-q} \quad (6.6)$$

adódik. Integrálással kapjuk minden olyan α lépéshosszra, melyre $u_{\min} - \alpha\sigma \geq 0$ teljesül, hogy

$$f_1'(\alpha) = f_1'(0) + \int_0^\alpha f_1''(\xi)d\xi \leq f_1'(0) + \int_0^\alpha h(\xi)d\xi.$$

Még egyszer integrálva az egyenlőtlenséget, valamint felhasználva az $f(\alpha) \leq f_1(\alpha)$ egyenlőtlenséget és hogy $f_1(0) = 0$

$$f(\alpha) \leq f_1(\alpha) = \int_0^\alpha f_1'(\zeta)d\zeta \leq f_2(\alpha) := f_1'(0)\alpha + \int_0^\alpha\int_0^\zeta h(\xi)d\xi d\zeta$$

kapjuk. Nyilván $f_2''(\alpha) = h(\alpha) > 0$, így az f_2 konvex függvény. Mivel $f_2(0) = 0$, $f_2'(0) = f_1'(0) < 0$ és $f_2''(\alpha) = h(\alpha)$ végtelenhez tart, ha α tart az $u_{\min}/\sigma$ értékhez, ezért az f_2 függvény a minimumát valamely $\tilde{\alpha} > 0$ helyen veszi fel, mely az

$$f_2'(\alpha) = f_1'(0) + \int_0^\alpha h(\xi)d\xi = 0$$

egyenlet megoldása. Felhasználva a (6.6) összefüggést az előző egyenlet a következővel ekvivalens

$$f_1'(0) + \frac{1}{2}q\sigma^2 \int_0^\alpha (u_{\min} - \xi\sigma)^{-1-q} d\xi = 0.$$

Megoldva α paraméterre nézve az egyenletet

$$\tilde{\alpha} = \frac{u_{\min}}{\sigma}\left(1 - (1 + \sigma u_{\min}^q)^{\frac{1}{-q}}\right) \tag{6.7}$$

kifejezésre jutunk. A 6.1.2. Lemma alapján

$$u_{\min}^q \geq \frac{1}{\sigma + 1}.$$

Behelyettesítve a (6.7) összefüggésbe a következőre jutunk

$$\tilde{\alpha} \geq \frac{u_{\min}}{\sigma}\left(1 - \left(1 + \frac{\sigma}{\sigma + 1}\right)^{\frac{1}{-q}}\right).$$

A következő lemmára a további becslésekhez van szükségünk.

6.1.5. Lemma. *Legyen $p \in [0, 1]$, ekkor*

$$(1 - t)^p \leq 1 - p\,t, \qquad \textit{minden } t \in [0, 1].$$

Bizonyítás: Tetszőleges rögzétett $p \leq 1$ esetén a $(t - 1)^p - 1 + p\,t$ monoton fogyó függvény a $t \in [0, 1]$ intervallumon, és nulla $t = 0$ esetén. Ebből már következik a lemma állítása. □

Az előző lemma becslését alkalmazva

$$\left(1 + \frac{\sigma}{\sigma + 1}\right)^{-\frac{1}{q}} = \left(1 - \frac{\sigma}{2\sigma + 1}\right)^{\frac{1}{q}} \leq 1 - \frac{1}{\sigma} q(2\sigma + 1)$$

adódik, így még egyszer használva a 6.1.2. Lemmát az $\tilde{\alpha}$ lépéshosszra a következő alsó korlátot nyerjük

$$\tilde{\alpha} \geq \frac{u_{\min}}{\sigma}\frac{\sigma}{q(2\sigma + 1)} = \frac{u_{\min}}{q(2q + 1)} \geq \frac{1}{q(2\sigma + 1)(\sigma + 1)^{\frac{1}{\sigma}}}.$$

A továbbiakban feltesszük, hogy $\sigma \geq 1$ és az

$$\alpha^* = \frac{1}{3q\sigma(\sigma+1)^{\frac{1}{q}}} \tag{6.8}$$

lépéshosszal számolunk. Vegyük észre, hogy $\alpha^* \leq \tilde{\alpha}$. Az $f_2(\alpha^*)$ kifejezés becsléséhez szükségünk lesz a következő lemmára.

6.1.6. Lemma. *Legyen $h(t)$ kétszer differenciálható konvex függvény, továbbá $h(0) = 0$, $h'(0) < 0$ és a h függvény a $t^* > 0$ pontban veszi fel a (globális) minimumát. Ekkor ha h'' függvény monoton nő a $[0, t^*]$ intervallumon, akkor*

$$h(t) \leq \frac{th'(0)}{2}, \qquad 0 \leq t \leq t^*.$$

Bizonyítás: A lemma feltételei alapján

$$\begin{aligned} h(t) &= \int_0^t h'(\xi)d\xi = h'(0)t + \int_0^t \int_0^\xi h''(\zeta)d\zeta d\xi \leq h'(0)t + \int_0^t \xi h''(\xi)d\xi \\ &= h'(0)t + [\xi h'(\xi)]_0^t - \int_0^t h'(\xi)d\xi \leq h'(0)t - h(t). \end{aligned}$$

Ezt átrendezve megkapjuk a lemma állítását. □

Alkalmazzuk az előző lemmát $h = f_2$ függvényre. Először ellenőrizzük, hogy az előző lemma feltételei teljesülnek-e: $f_2(0) = 0$ és $f_2'(0) < 0$. Továbbá $f_2''(\alpha) = h(\alpha) > 0$ és

$$f_2'''(\alpha) = h'(\alpha) = \frac{q(q+1)\sigma^3}{2}(u_{\min} - \alpha\sigma)^{-2-q} > 0.$$

Tehát alkalmazhatjuk a lemmát, továbbá a (6.5) azonosságot figyelembe véve $f(\alpha^*)$ értékére a következő felső becslést kapjuk

$$f(\alpha^*) \leq f_2(\alpha^*) \leq \frac{\alpha^* f_2'(0)}{2} = \frac{\alpha^* f_1'(0)}{2} = -\frac{\alpha^*\sigma^2}{4}.$$

Végül behelyettesítve α^* (6.8) definícióját és figyelembe véve $\sigma \geq 1$ feltevésünket

$$f(\alpha^*) \leq \frac{-\sigma}{12q(\sigma+1)^{\frac{1}{q}}} \leq \frac{-\sigma}{12q(2\sigma)^{\frac{1}{q}}} \leq \frac{-\sigma^{\frac{q-1}{q}}}{24q}. \tag{6.9}$$

A μ paraméter iterálásának hatása

A következőkben néhány technikai eredményt ismertetünk melyekre a konkrét elemzések során lesz szükségünk. Mivel

$$\psi'(t) = t - t^{-q}, \qquad \psi''(t) = 1 + qt^{-q-1},$$

ezért $\psi''(t) \geq 1$ minden $t > 0$ esetén.

6.1.7. Lemma.

$$\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}\psi'(t)^2, \qquad t > 0.$$

Bizonyítás: Mivel $\psi(1) = \psi'(1) = 0$,

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \geq \int_1^t \int_1^\xi d\zeta d\xi = \frac{1}{2}(t-1)^2$$

áll fenn, amiből következik az egyenlőtlenség első fele. A második egyenlőtlenséghez felülről becsüljük a ψ függvényt

$$\begin{aligned} \psi(t) &= \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \leq \int_1^t \int_1^\xi \psi''(\xi)\psi''(\zeta) d\zeta d\xi \\ &= \int_1^t \psi''(\xi)\psi'(\xi) d\xi = \frac{1}{2}\psi'(t)^2, \end{aligned}$$

amiből következik a lemma állítása. □

A fenti lemmában a büntetőfüggvény magfüggvényére kapott alsó becslés segítségével az $\mathbf{u}$ vektor normájára felső becslést adhatunk.

6.1.8. Következmény. *A következő egyenlőtlenség teljesül*

$$\|\mathbf{u}\| \leq \sqrt{n} + \sqrt{2\Psi(\mathbf{u})}.$$

Bizonyítás: Felhasználva a 6.1.7. Lemmát és a Cauchy–Schwarz egyenlőtlenséget kapjuk

$$2\Psi(\mathbf{u}) = 2\sum_{i=1}^n (u_i - 1)^2 = \|\mathbf{u}\|^2 - 2\mathbf{e}^T\mathbf{u} + n \geq \|\mathbf{u}\|^2 - 2\|\mathbf{e}\|\,\|\mathbf{u}\| + \|\mathbf{e}\|^2 = (\|\mathbf{u}\| - \|\mathbf{e}\|)^2.$$

Ezt átrendezve

$$\|\mathbf{u}\| \leq \|\mathbf{e}\| + \sqrt{2\Psi(\mathbf{u})} = \sqrt{n} + \sqrt{\Psi(\mathbf{u})}$$

adódik, amiből már következik az állítás. □

Az előző lemmában a ψ függvényre adott felső korlát összevetve a σ definíciójával a büntetőfüggvényre felső becslést eredményez.

6.1.9. Következmény. $\Psi(\mathbf{u}) \leq \frac{1}{2}\sigma(\mathbf{u})^2$.

Bizonyítás: Ismét a 6.1.7. Lemmát használva, illetve a (6.1) alapján

$$\Psi(\mathbf{u}) = \sum_{i=1}^{n} \psi(u_i) \leq \frac{1}{2} \sum_{i=1}^{n} \psi'(u_i)^2 = \frac{1}{2} \|\nabla\Psi(\mathbf{u})\|^2 = \frac{1}{2}\, \sigma(\mathbf{u})^2.$$

Ezzel a bizonyítást befejeztük. □

A következő lemmában megvizsgáljuk, hogyan változik $\psi(t)$ értéke ha a t változót β paraméterszeresére növeljük, ahol $\beta \geq 1$.

6.1.10. Lemma. *Legyen* $\beta \geq 1$. *Ekkor*

$$\psi(\beta t) \leq \psi(t) + \frac{1}{2}(\beta - 1)t^2.$$

Bizonyítás: Használva a bevezetett (6.3) jelölést

$$\psi(\beta t) = \frac{\beta^2 t^2 - 1}{2} + \psi_0(\beta t) = \psi(t) + \frac{1}{2}(\beta^2 t^2 - t^2) + \psi_0(\beta t) - \psi_0(t).$$

Mivel a ψ_0 függvény monoton fogyó, $\psi_0(\beta t) - \psi_0(t) \leq 0$. Ebből pedig már következik a bizonyítandó egyenlőtlenség. □

A fenti lemmában kapott egyenlőtlenség módot nyújt a μ paraméter változásának a büntetőfüggvény értékére gyakorolt hatásának vizsgálatára.

6.1.11. Lemma. *Legyen* $\theta \in (0, 1)$ *és* $\mu^+ = (1 - \theta)\mu$. *Ekkor*

$$\Psi(\mathbf{x}, \mathbf{s}, \mu^+) \leq \Psi(\mathbf{x}, \mathbf{s}, \mu) + \frac{\theta}{2(1-\theta)} \left(2\Psi(\mathbf{u}) + \sqrt{2n\Psi(\mathbf{u})} + n\right).$$

Bizonyítás: Vegyük észre, mikor a μ paraméter értéke μ^+ értékre változik, az $\mathbf{u}$ vektor $\mathbf{u}^+ = \sqrt{\mathbf{xs}/\mu^+} = 1/\sqrt{1-\theta}\,\mathbf{u}$ vektorra módosul. Alkalmazzuk a 6.1.10. Lemmát $\beta = 1/\sqrt{1-\theta}$ értékre

$$\Psi(\mathbf{u}^+) = \sum_{i=1}^{n} \psi(\beta u_i) \leq \sum_{i=1}^{n} \left(\psi(u_i) + \frac{1}{2}(\beta^2 - 1)u_i^2\right) = \Psi(\mathbf{u}) + \frac{\theta \|\mathbf{u}\|^2}{2(1-\theta)}.$$

Az 6.1.8. Következmény segítségével folytathatjuk a becslést

$$\Psi(\mathbf{u}^+) \leq \Psi(\mathbf{u}) + \frac{\theta\left(\sqrt{n} + \sqrt{2\Psi(\mathbf{u})}\right)^2}{2(1-\theta)} = \Psi(\mathbf{u}) + \frac{\theta}{2(1-\theta)} \left(2\Psi(\mathbf{u}) + \sqrt{2n\Psi(\mathbf{u})} + n\right).$$

Ezzel az állítást beláttuk. □

A szükséges becslések felállítása után most rátérünk az algoritmus lépésszámának meghatározására. Az algorimus külső ciklusa elején a μ paraméter iterálása előtt $\Psi(\mathbf{x}, \mathbf{s}, \mu) \leq \tau$ teljesül. Az 6.1.11. Lemma alapján a μ frissítése után

$$\Psi(\mathbf{x}, \mathbf{s}, \mu^+) \leq \tau + \frac{\theta}{2(1-\theta)} \left(2\tau + \sqrt{2n\tau} + n\right) \tag{6.10}$$

áll fenn. Az (6.8) alatt meghatározott α^* értéket használva (6.9) alapján a büntetőfüggvény értékének csökkenése legalább

$$\frac{\sigma^{\frac{q-1}{q}}}{24q}. \tag{6.11}$$

Megjegyeznénk, hogy ez az eredmény csak a $\sigma \geq 1$ feltétel teljesülése mellett igaz, ám 6.1.9. Következménybeli egyenlőtlenség alapján ennek elégséges feltétele $\Psi(\mathbf{x}, \mathbf{s}, \mu) \geq 1$. Ami pedig $\tau \geq 1$ mellett biztosan fennáll minden belső ciklus megkezdésekor. A 6.1.9. Következmény és (6.11) alapján minden belső iteráció során a büntetőfüggvény csökkenése legalább

$$\frac{\Psi^{\frac{q-1}{2q}}}{24q}, \tag{6.12}$$

ahol Ψ a büntetőfüggvény értéke a belső iteráció megkezdése előtt.

6.1.12. Lemma. *Minden belső ciklusban legfeljebb*

$$\left\lceil \frac{48q^2\left(\tau + \frac{\theta}{2(1-\theta)}(2\tau + \sqrt{2n\tau} + n)\right)^{\frac{q+1}{2q}}}{q+1} \right\rceil$$

iteráció történik.

Bizonyítás: Jelölje Ψ a büntetőfüggvény értékét a belső ciklus elején. Egy iteráció után az új értéket jelölje Ψ^+. Ekkor (6.12) alapján

$$\Psi^+ \leq \Psi - \frac{\Psi^{\frac{q-1}{2q}}}{24q}.$$

Emeljük az egyenlőtlenséget $(q+1)/(2q)$ hatványra, majd alkalmazzuk az $(1-t)^\alpha \leq 1 - \alpha t$, $(\alpha,\, t \in [0,1])$ összefüggést

$$\begin{aligned} (\Psi^+)^{\frac{q+1}{2q}} &\leq \left(\Psi - \frac{\Psi^{\frac{q-1}{2q}}}{24q}\right)^{\frac{q+1}{2q}} = \Psi^{\frac{q+1}{2q}} \left(1 - \frac{\Psi^{-\frac{q+1}{2q}}}{24q}\right)^{\frac{q+1}{2q}} \\ &\leq \Psi^{\frac{q+1}{2q}} \left(1 - \frac{q+1}{2q}\,\frac{\Psi^{-\frac{q+1}{2q}}}{24q}\right) = \Psi^{\frac{q+1}{2q}} - \frac{q+1}{48q^2}. \end{aligned}$$

Ha $\Psi^{(k)}$ jelöli a k. belső iteráció után a büntetőfüggvény értékét, akkor a következő becslés áll fenn

$$\Psi^{(k)} \leq \Psi^{\frac{q+1}{2q}} - k\,\frac{q+1}{48q^2}.$$

Ha a $\Psi^{(k)}$ értékre adott felső korlát kisebb, mint τ, akkor a belső iterációk száma legfeljebb k, tehát

$$\Psi^{\frac{q+1}{2q}} - k\,\frac{q+1}{48q^2} \leq \tau,$$

átrendezés után

$$\frac{48q^2}{q+1}\Psi^{\frac{q+1}{2q}} - \tau \leq k$$

adódik. Az alsó korlátot felülről becsülve $\tau > 0$ illetve (6.10) figyelembevételével megkapjuk a lemma állítását. □

A belső iterációk maximális számának és a külső iterációk maximális számának szorzata felső becslést ad az algoritmus lépésszámára.

6.1.13. Tétel. *Az algoritmus legfeljebb*

$$\left\lceil \frac{48q^2(\tau + \frac{\theta}{2(1-\theta)}(2\tau + \sqrt{2n\tau} + n))^{\frac{q+1}{2q}}}{q+1} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil$$

lépésben megad egy $(\mathbf{x}, \mathbf{s})$ *megengedett pontpárt, melyre* $\mathbf{x}^T\mathbf{s} < \varepsilon$.

Bizonyítás: Mivel hosszú lépéses az algoritmus, a külső iterációk száma legfeljebb

$$\left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil$$

(lásd Roos, Terlaky, Vial könyv [24]). Ezt megszorozva a 6.1.12. Lemma eredményével, adódik a tétel állítása. □

A tétel eredménye egészértékjelek nélkül a $\theta = \frac{2}{3}$, $\tau = n$ választás esetén

$$\frac{72q^2(4n + n\sqrt{2})^{\frac{q+1}{2q}}}{q+1} \log \frac{n}{\varepsilon} \leq 72(4 + \sqrt{2}\frac{q^2 n^{\frac{q+1}{2q}}}{q+1} \log \frac{n}{\varepsilon} \leq 390qn^{\frac{q+1}{2q}} \log \frac{n}{\varepsilon}.$$

Rögzített $q > 1$ esetén ez a hosszú lépéses algoritmusokra ismert $\mathcal{O}(n \log \frac{n}{\varepsilon})$ komplexitás javítását jelenti. A $q = \frac{1}{2} \log n$ választás minimalizálja az előző felső korlátot. Ekkor az iterációszám

$$\mathcal{O}\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right),$$

ami már igen közel van a belsőpontos algoritmusok elméletében ismert legjobb lépésszámhoz, azaz a $\mathcal{O}(\sqrt{n} \log \frac{n}{\varepsilon})$ komplexitáshoz.

Jelölések

$\mathbb{R}^n_+$	az n-dimenziós pozitív vektorok halmaza
$\mathbb{R}^n_\oplus$	az n-dimenziós nemnegatív vektorok halmaza
$\mathbf{x}, x_i$	vektorokat vastag betű, skalárokat normál betű jelzi
$\mathbf{x}\,\mathbf{s}$	az $\mathbf{x}$ és az $\mathbf{s}$ vektorok koordinátánkénti (Hadamard) szorzata
$\mathbf{x}^\alpha$	a hatványozás koordinátánkénti elvégzésével kapott n-dimenziós vektor
$\mathbf{x}^T\mathbf{s}$	a két vektor skalárszorzata
X	az $\mathbf{x}$ vektorból előállított diagonális mátrix, azaz $X = diag(\mathbf{x})$
$\|\cdot\|$	az Euklideszi norma
$\|\cdot\|_\infty$	az l_∞ norma
$\mathbf{e}$	a csupa egyesből álló, n-dimenziós vektor
I	az egységmátrix
$\mathbf{a}_j$	az A mátrix j. oszlopa
$\lvert J\rvert$	a $J \subset \{1, 2, \dots, n\}$ indexhalmaz számossága
A_J	$A \in \mathbb{R}^{m\times n}$ azon részmátrixa, melynek oszlopindexei $J \subset \{1, 2, \dots, n\}$-hez tartoznak
A_{KJ}	az A_J azon részmátrixa, melynek sorindexei $K \subset \{1, 2, \dots, m\}$-beliek

Definíció. Legyen f, $g : \mathbb{R}_+ \to \mathbb{R}_+$ függvények. Ekkor
$f(x) = \mathcal{O}(g(x))$, ha $\exists\, c > 0$ konstans, melyre

$$f(x) \le c\, g(x) \quad \forall x > 0.$$

$f(x) = \Theta(g(x))$, ha $\exists\, c_1, c_2 > 0$ konstansok, melyekre

$$c_1\, g(x) \le f(x) \le c_2\, g(x)\ \forall x > 0.$$

Tárgymutató

Irodalomjegyzék

[1] E.D. Andersen, Y. Ye. *Combining interior-point and pivot algorithms for linear programming.* Technical report, Department of Management Sciences, University of Iowa, 1995.

[2] I.I. Dikin. *Iterative solution of problems of linear and quadratic programming.* Doklady Akademii Nauk SSSR, 174:747-748, 1967. Translated in Soviet Mathematics Doklady, 8:674-675, 1967.

[3] P. Gács, L. Lovász. *Khachiyan's algorithm for linear programming.* Mathematical Programming Study 14, 61-68, 1981.

[4] A.J. Goldman and A.W. Tucker. *Theory of linear programming.* In H.W. Kuhn and A.W.Tucker, editors, Linear Inequalities and Related Systems, Annals of Mathematical Studies, No. 38, pp. 53- 97. Princeton University Press, Princeton, New Jersey, 1956.

[5] T. Illés, C. Roos, and T. Terlaky. *Polynomial Affine–Scaling Algorithms for $P_*(\kappa)$ Linear Complementarity Problems.* In P. Gritzmann, R. Horst, E. Sachs, R. Tichatschke, editors, *Recent Advances in Optimization*, Proceedings of the 8^{th} French-German Conference on Optimization, Trier, July 21-26, 1996, Lecture Notes in Economics and Mathematical Systems 452, pp. 119-137, Springer Verlag, 1997.

[6] T. Illés, J. Peng, C. Roos, T. Terlaky. *A Strongly Polynomial Rounding Procedure Yielding A Maximally Complementary Solution for $P_*(\kappa)$ Linear Complementarity Problems.* SIAM Journal on Optimization, 11:320-340, 2000.

[7] T. Illés and T. Terlaky. *Pivot Versus Interior Point Methods: Pros and Cons*, European Journal of Operations Research, 140:6-26, 2002.

[8] B. Jansen, C. Roos, T. Terlaky. *Primal–dual algorithms for linear programming based on the logarithmic barrier method.* Journal of Optimization Theory and Applications, 83:1-26, 1994.

[9] B. Jansen, C. Roos, T. Terlaky. *A Family of Polynomial Affine Scaling Algorithms for Positive Semidefinite Linear Complementarity Problems.* SIAM Journal on Optimization Vol. 7. No. 1. 126–140, 1997.

[10] M. Kojima, S. Mizuno and A. Yoshise. *A Polynomial-time Algorithm for a Class of Linear Complementarity Problems.* Mathematical Programming, 44 (1989) 1-26.

[11] M. Kojima, S. Mizuno and A. Yoshise. *A primal–dual interior point algorithm for linear programming.* In N. Megiddo, editor, Prograss in Mathematical Programming: Interior Point and Related Methods, pp. 29-47. Springer Verlag, New York, 1989.

[12] M. Kojima, N. Megiddo, T. Noma, A. Yoshise. *A unified approach to interior point algorithms for linear complementarity problems.* volume 538 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Germany, 1991.

[13] P.D. Ling. *A new proof of convergence for the new primal–dual affine scaling interior–point algorithm of Jansen, Roos and Terlaky.* Working paper, University of East–Anglia, Norwich, England, 1993.

[14] P.D. Ling. *A predictor–corrector algorithm.* Working paper, University of East–Anglia, Norwich, England, 1993.

[15] N. Megiddo. *Pathways to the optimal set in linear programming.* In N. Megiddo, editor, Progress in Mathematical Programming: Interior Point and Related Methods, pp. 131-158. Springer Verlag, New York, 1989.

[16] J. Miao. *A quadratically convergent $O((\kappa + 1)\sqrt{n}L)$-iteration algorithm for the $P_*(\kappa)$-matrix linear complementarity problem.* Mathematical Programming 69 (1995) 355-368.

[17] S. Mizuno, M. J. Todd and Y. Ye. *On Adaptive-step Primal-dual Interior-point Algorithms for Linear Programming.* Mathematics of Operations Research, Vol. 18, No. 4 (1993) 964-981.

[18] Nagy M. *Mizuno–Todd–Ye prediktor–korrektor algoritmus $P_*(\kappa)$ mátrixokra.* Témavezető: Illés Tibor. TDK dolgozat, Eötvös Loránd Tudományegyetem Természettudományi Kar, Budapest, 2003.

[19] J. Peng, C. Roos, T. Terlaky. *Self–Regularity: A New Paradigm for Primal–Dual Interior–Point Algorithm.* Princeton University Press, Princeton, New Jersey, 2002.

[20] J. Peng, C. Roos, T. Terlaky. *A New and Efficient Large/Update Interior/Point Method for Linear Optimization.* Journal of Computational Technologies, 6(4):61-80, 2001. ISSN 1560-7534.

[21] Pólik I. *Lineáris optimalizálás belsőpontos módszereinek újszerű vizsgálata.* Szakdolgozat, Témavezető: Illés Tibor, Eötvös Loránd Tudományegyetem Természettudományi Kar, Budapest, 2002.

[22] F.A. Potra, R. Sheng. *Predictor–corrector algorithm for solving $P_*(\kappa)$-matrix LCP from arbitrary positive starting points.* Mathematical Programming 76(1) (1996) 223-244.

[23] F.A. Potra. *The Mizuno-Todd-Ye algorithm in a Larger Neighborhood of the Central Path.* Europian Journal of Operational Research 143 (2002) 257-267.

[24] C. Roos, T. Terlaky, J.-Ph. Vial. *Theory and Algorithms for Linear Optimization, An Interior Point Approach.* Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, USA, 1997.

[25] A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley & Sons, New York, 1986.

[26] Gy. Sonnevend. *An „analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming.* In A. Prékopa, J. Szelezsán and B. Strazicky, editors, System Modelling and Optimization: Proceedings of the 12th IFIP-Conference in Budapest, Hungary, September 1985, volume 84 of Lecture Notes in Control and Information Sciences, pp. 866-876. Springer Verlag, Berlin, West–Germany, 1986.

[27] Gy. Sonnevend, J. Stoer and G. Zhao. *On the Complexity of Following the Central Path of Linear Programs by Linear Extrapolation.* Methods of Operations Research 63 (1989) 19-31.

[28] Stoyan G., Takó G. *Numerikus módszerek I.*. ELTE-TYPOTEX, Budapest, 1993.

[29] A.W. Tucker. *Dual system of homogeneous linear relations.* In H.W. Kuhn and A.W. Tucker, editors, Linear Inequalities and Related System, Annals of Mathematical Studies, No. 38, pp. 3-18. Princeton University Press, Princeton, New Jersey, 1956.

[30] Y. Ye. *On the finite convergence of interior point algorithms for linear programming.* Mathematical Programming, 57:325-335, 1992.

[31] Y. Ye, K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor–corrector algorithm for LCP.* Mathematical Programming 62 (1993) 537-551.