

Eötvös Loránd Tudományegyetem

---

A lineáris programozási feladat optimális  
bázismegoldásának előállítása polinom időben

diplomamunka

*Majoros Csilla*

Témavezető:

*Illés Tibor*

egyetemi docens

ELTE-TTK Operációkutatási Tanszék



Budapest, 2014.

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>3</b>
<b>2. Alapvető eredmények áttekintése</b>	<b>6</b>
2.1. Lineáris programozási alapfeladat . . . . .	6
2.2. Lagrange-dualitás . . . . .	7
2.3. A Newton-módszer . . . . .	8
2.4. A barrier-módszer . . . . .	10
<b>3. Útkövető belsőpontos algoritmusok elmélete</b>	<b>12</b>
3.1. Az analitikus centrum . . . . .	12
3.2. Newton módszer az <i>LP</i> -feladatra . . . . .	13
3.3. Barrier módszer az <i>LP</i> -feladatra . . . . .	14
3.4. A centrális út . . . . .	16
3.5. A Newton-lépés . . . . .	19
3.6. A centrális út környezetei . . . . .	21
<b>4. Primál-duál útkövető algoritmusok</b>	<b>23</b>
4.1. A keretalgoritmus . . . . .	23
4.2. Egy rövid lépéses algoritmus . . . . .	24
4.3. Egy hosszú lépéses algoritmus . . . . .	28
4.4. Egy prediktor-korrektor algoritmus . . . . .	30
4.5. Pontos megoldás előállítása . . . . .	33
<b>5. Duális módszerek</b>	<b>37</b>
5.1. A szubgradiensek . . . . .	37
5.2. A szubgradiens módszer . . . . .	38
5.3. A vágósíkos módszer . . . . .	41
5.4. A volume algoritmus . . . . .	43
<b>6. Optimális bázismegoldás előállítása</b>	<b>49</b>
6.1. A perturbált feladat . . . . .	49
6.2. Kezdő bázis előállítása . . . . .	52
6.3. Primál fázis . . . . .	54
6.4. Duál fázis . . . . .	55

# 1. Bevezetés

Jelenleg két népszerű megközelítés van a lineáris programozási feladat megoldására: a pivot algoritmusok és a belsőpontos algoritmusok. Mindkét típusnak számos variációja lett kifejlesztve az utóbbi évtizedekben.

A szimplex módszernek megvan az a szép tulajdonsága, hogy ha a feladatnak van optimális megoldása, akkor az algoritmus egy optimális bázismegoldást talál. Azonban jelenleg nem ismert polinomiális idejű pivot algoritmus.

A szimplex módszerrel ellentétben a belsőpontos algoritmusok a megengedett tartomány belsejében lépkednek, a cél az optimális lap megközelítése. Azaz véges sok lépésben nem adnak egzakt optimális megoldást. Ye [10] javasolt egy eljárást, amit egy belsőpontos algoritmusba ágyazva pontos optimális megoldást állíthatunk elő polinom időben. Ez a megoldás azonban az optimális lap belső pontja lesz, azaz több optimális megoldás létezése esetén nem lesz bázismegoldás.

Számos esetben kívánatos, hogy az optimális megoldásunk bázismegoldás legyen:

- Bázismegoldást használ a klasszikus és olcsó érzékenység vizsgálat.
- Ha egymáshoz szorosan kapcsolódó feladat-sorozatunk van, akkor sok esetben az egyik optimális bázismegoldásából indítva néhány pivotlépéssel megkapjuk a következő optimális bázismegoldását.
- Egészértékű és vegyes programozási feladatok megoldási módszereinél is gyakran elengedhetetlen, hogy a relaxált feladat optimális bázismegoldása álljon rendelkezésre.

Felvetült tehát a kérdés, hogyan lehetne polinom időben optimális bázismegoldást előállítani. Erre a problémára az első releváns választ Megiddo [11] adta. Konstruktív módon bebizonyította, hogy ha adott egy optimális primál-duál megoldaspár, akkor erősen polinomiális időben elő tudunk állítani optimális bázismegoldást.

Természetes volt Megiddo módszerét kombinálni a belsőpontos algoritmusokkal, és Ye [10] pontos megoldást előállító kerekítési eljárásával. Ezt meg is tette Andersen és Ye [12], és egy elméleti és gyakorlati szempontból is jól működő algoritmust csináltak, ami polinomiális időben optimális bázismegoldást állít elő. Az eljárást röviden nevezzük *BI*-nek, azaz bázis identifikációnak.

A *BI* eljárás nem csak a belsőpontos algoritmusokkal kombinálva használható, hanem más LP-megoldó módszerekkel is, amik nem feltétlenül adnak olyan precíz közelítést. A lényeg, hogy az primál és duál közelítő megoldást is gyártson. Megiddo [11] ugyanis azt is bebizonyította, hogy ha csak egy primál, vagy csak egy duál optimális megoldás áll a rendelkezésünkre, akkor optimális bázist csak abban az esetben tudunk találni erősen polinomiális időben, ha az *LP*-feladat is megoldható erősen polinomiális időben.

A dolgozat célja, hogy ezt a folyamatot - közelítő megoldás előállítása, bázis identifikáció - kifejtse. Részletesen áttekintjük az útkövető belsőpontos algoritmusok elméletét, a nemlineáris programozás eredményeiből kiindulva. Majd megnézzük a Ye [10] által javasolt kerekítési eljárást, amivel egzakt, szigorúan komplementáris megoldaspárt állítunk elő.

A belsőpontos algoritmusok hátránya, hogy közeledve az optimális laphoz lelassul a folyamat, számítási nehézségek lépnek fel. Ezért mielőtt a *BI*-eljárást kifejtenénk, megnézzünk néhány közelítő algoritmust, amik számításigénye (sok esetben) jóval kedvezőbb a belsőpontos algoritmusokénál. Az első két vizsgált módszer csak duál-közelítő megoldást állít elő, ezért ezekkel a *BI*-eljárás nem kombinálható. Francisco Barahona és Ranga Anbil [8] dolgozták ki az előzőek továbbfejlesztésével a *volume* algoritmust, ami a duál közelítéssel párhuzamosan a primál közelítést is végzi. Ez az algoritmus egy szemléletes tételre alapszik, ami a primál változókat bizonyos térfogatok arányaként fejezi ki. Ez a módszer kis számításigényű, és kombinálva a *BI* eljárással sok esetben hatékonyan megkeres egy optimális bázist. Végül megnézzük a *BI*-eljárást, ami egy közelítő primál-duál megoldást adó algoritmussal kombinálva optimális bázismegoldást talál.

## **Köszönetnyilvánítás**

Ezúton is szeretnék köszönetet mondani Illés Tibor tanár úrnak a szakdolgozatomhoz nyújtott segítségéért, és azért, hogy a kurzusai során felkeltette az érdeklődésemet a téma iránt.

## 2. Alapvető eredmények áttekintése

### 2.1. Lineáris programozási alapfeladat

A primál-duál lineáris optimalizálási feladatpár standard alakja a következő:

$$(P) \quad \min\{c^T x : Ax = b, x \geq 0\},$$
$$(D) \quad \max\{b^T y : A^T y + s = c, s \geq 0\}.$$

ahol  $A \in \mathbb{R}^{m \times n}$  egy  $m$ -rangú mátrix;  $b, c, x, y, s$  pedig megfelelő méretű valós vektorok.

A vektorokat mindig oszlopvektornak képzeljük.

### Jelölések

$\mathcal{F}_P, \mathcal{F}_D$	a $(P), (D)$ feladatok megengedett megoldásai
$\mathcal{F}_P^0, \mathcal{F}_D^0$	a $(P), (D)$ feladatok szigorúan megengedett megoldásai
$\mathcal{F}_P^*, \mathcal{F}_D^*$	a $(P), (D)$ feladatok optimális megoldásai
$\mathcal{F}, \mathcal{F}^0, \mathcal{F}^*$	a megfelelő primál-duál megoldáspárok halmaza
$p^* = d^*$	a primál és duál optimum értéke
$S_\delta(x)$	az $x$ pont $\delta$ sugarú nyílt környezete
$Df$	az $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ függvény Jacobi-mátrixa
$\nabla f, \nabla^2 f$	az $f : \mathbb{R}^n \rightarrow \mathbb{R}$ függvény gradiense és Hesse-mátrixa
$X$	az a diagonális mátrix, amire $X_{ii} = x_i$ , ahol $x \in \mathbb{R}^n$
$e$	a megfelelő méretű csupa 1 vektor
$\mathcal{C}_P, \mathcal{C}_D, \mathcal{C}$	a primál, duál és primál-duál centrális út
$\mathbb{R}_\oplus^n, \mathbb{R}_+^n$	a nemnegatív és pozitív ortáns $\mathbb{R}^n$ -ben
$xs$	két vektor komponensenkénti szorzata ( $= XSe$ )

**2.1.1. Tétel (gyenge dualitás).** Legyen  $\bar{x} \in \mathcal{F}_P$  és  $(\bar{y}, \bar{s}) \in \mathcal{F}_D$ . Ekkor

$$c^T \bar{x} \geq b^T \bar{y}$$

és egyenlőség pontosan akkor van, ha

$$\bar{x}^T \bar{s} = 0.$$

**Biz.:** triviális □

**2.1.2. Tétel (erős dualitás).** Ha a primál és duál feladatnak is van megengedett megoldása, akkor mindkettőnek van optimális megoldása is. Tetszőleges  $x^*, y^*$  optimális megoldásra  $c^T x^* = b^T y^*$ .

Tehát az  $x^*$  primál, és  $(y^*, s^*)$  duál megengedett megoldások pontosan akkor optimálisak, ha  $(x^*)^T s^* = 0$ . A nemnegativitási feltételek miatt ez pontosan azt jelenti, hogy  $x_i^* s_i^* = 0$  minden  $i$ -re. Az optimális megoldások tehát pontosan a komplementárisak.

**2.1. Def.** Egy  $x^*$  és  $(y^*, s^*)$  optimális megoldaspárt erősen komplementárisnak nevezünk, ha  $x^* + y^* > 0$ .

**2.1.3. Tétel (Goldman-Tucker).** Ha a primál és duál feladatnak is van megengedett megoldása, akkor szigorúan komplementáris primál-duál optimális megoldaspár is létezik. Ha  $x^*, (y^*, s^*)$  egy szigorúan komplementáris megoldaspár, akkor az általa meghatározott  $\mathcal{P}^* := \{j : x_j^* > 0\}$  és  $\mathcal{D}^* := \{1, \dots, n\} \setminus \mathcal{P}^*$  partíció megegyezik minden egyes szigorúan komplementáris megoldás esetén.

## 2.2. Lagrange-dualitás

A következő ismeretekre a 5. fejezetben lesz majd szükség.

Tekintsük a következő optimalizálási feladatot

$$\left. \begin{array}{l} \min f(x) \\ h(x) = 0 \\ x \in M, \end{array} \right\}$$

ahol  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  adott függvények és  $M \subseteq \mathbb{R}^n$ .

A feladathoz rendelt-Lagrange függvény:

$$L(x, \pi) = f(x) + \pi^T h(x),$$

és duális függvény

$$q(\pi) = \inf_{x \in M} L(x, \pi).$$

A feladat Lagrange-duálisa a következő probléma:

$$\left. \begin{array}{l} \max q(\pi) \\ \pi \in \mathbb{R}^m \end{array} \right\}$$

A  $q(\pi)$  értéke  $-\infty$  is lehet néhány  $\mu$ -re. A duális feladat megoldása lényegében a következő feltételes maximalizálási feladat megoldását jelenti:

$$\left. \begin{array}{l} \max q(\pi) \\ \pi \in D \end{array} \right\}$$

ahol  $D = \{\pi | q(\pi) > -\infty\}$ . Nevezzük  $D$ -t a kiterjesztett értékű  $q$  függvény lényeges tartományának.

Függetlenül a primál feladat költségfüggvényétől és feltételeitől, a duális feladat szép konvexitási tulajdonságokkal rendelkezik, ezt fejezi ki a következő állítás.

**2.2.1. Állítás.** A duális függvény lényeges tartománya konvex halmaz, és  $q$  konkáv ezen a halmazon.

**Biz.:** Tetszőleges  $x, \pi, \bar{\pi}$  és  $\alpha \in [0, 1]$  esetén

$$L(x, \alpha\pi + (1 - \alpha)\bar{\pi}) = \alpha L(x, \pi) + (1 - \alpha)L(x, \bar{\pi}).$$

Mindkét oldal infimumát véve az  $M$  halmazon

$$\inf_{x \in M} L(x, \alpha\pi + (1 - \alpha)\bar{\pi}) \geq \alpha \inf_{x \in M} L(x, \pi) + (1 - \alpha) \inf_{x \in M} L(x, \bar{\pi}),$$

vagyis

$$\alpha q(\alpha\pi + (1 - \alpha)\bar{\pi}) \geq \alpha q(\pi) + (1 - \alpha)q(\bar{\pi}).$$

Ebből látszik, hogy  $\pi, \bar{\pi} \in D$  esetén  $\alpha\pi + (1 - \alpha)\bar{\pi} \in D$ , azaz  $D$  konvex halmaz. Továbbá az utolsó egyenlőtlenség épp  $q$  konkávitását mutatja.  $\square$

A duális függvény másik szép tulajdonságáról szól a következő állítás.

**2.2.2. Állítás.** *A duális függvény felülről féligfolytonos, azaz minden  $\pi^k \rightarrow \pi$  pontsorozatra*

$$\limsup_{k \rightarrow \infty} q(\pi^k) \leq q(\pi).$$

**Biz.:**

$$q(\pi^k) = \inf_{z \in M} L(z, \pi^k) \leq L(x, \pi^k), \quad \forall x \in M,$$

így

$$\limsup_{k \rightarrow \infty} q(\pi^k) \leq L(x, \pi), \quad \forall x \in M.$$

Infimumot véve az  $M$  halmazon

$$\limsup_{k \rightarrow \infty} q(\pi^k) \leq \inf_{x \in M} L(x, \pi) = q(\pi),$$

ami épp a bizonyítandó állítás.  $\square$

**2.2.1. Köv.** *A  $D$  tartomány zárt.*  $\square$

### 2.3. A Newton-módszer

A Newton-módszer az alapja a leghatékonyabb algoritmusoknak a lineáris és nemlineáris programozás területén. A következő fejezetben is erre fognak épülni az algoritmusaink. Névadója javasolta először, 1669-ben, polinomok gyökének meghatározására. Az  $f(x) = x^3 - 2x - 5 = 0$  példán mutatta be az eljárást, az  $x_0 = 2$  pontból indulva. Tetszőleges differenciálható függvényre *Raphson* alkalmazta először, 1690-ben, ezért gyakran *Newton-Raphson* módszerként hivatkoznak rá. Az eljárás kvadratikus konvergenciáját a gyökök környezetében *Fourier* bizonyította 1818-ban, majd *Cauchy* terjesztette ki több dimenzióra 1829-ben. A többdimenziós eset konvergenciáját *Fine* bizonyította 1916-ban. A huszadik század folyamán végtelen dimenzióra és általános függvényterekre is továbbfejlesztették. Bővebben pl. Polyak [3] cikkéből tájékozódhatunk.

Nézzük a következő feladatot:

$$\begin{aligned} g : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ g(x) &= 0 \\ x &=? \end{aligned}$$



A módszer egy olyan pontsorozatot fog generálni, ami a  $g$  függvény gyökéhez konvergál. Tegyük fel, hogy az  $x_0, x_1, \dots, x_k$  pontokat már kiszámoltuk.

A  $g(x)$ -et az  $x_k$  körül az  $x_k$ -beli elsőrendű Taylor-polinomjával közelítjük:

$$g(x) \approx g(x_k) + Dg(x_k)(x - x_k) =: h(x)$$

Az  $x_{k+1}$  pont a  $h(x)$  gyöke lesz:

$$\begin{aligned} h(x) = g(x_k) + Dg(x_k)(x - x_k) &= 0 \\ x_{k+1} &= x_k - Dg(x_k)^{-1}g(x_k) \end{aligned}$$

**2.3.1. Tétel.** (I.) Tegyük fel, hogy az  $x^*$  gyöke a  $g$ -nek, a  $g$  folytonosan differenciálható az  $x^*$  környezetében és a  $Dg(x^*)$  Jacobi-mátrix reguláris. Ekkor ha a Newton-módszert az  $x^*$ -hoz elég közel indítjuk, akkor a Newton-lépések jól definiáltak lesznek, és a generált pontsorozat az  $x^*$ -hoz konvergál. A konvergencia szuperlineáris.

(II.) Tegyük fel, hogy az  $L > 0, M > 0, \delta > 0$  számokra és minden  $x, y \in S_\delta(x^*)$ -ra:

$$\|Dg(x) - Dg(y)\| \leq L\|x - y\|, \quad \|(\nabla g(x))^{-1}\| \leq M.$$

Ekkor minden  $x^0 \in S_\delta(x^*)$  esetén:

$$\|x^{k+1} - x^*\| \leq \frac{LM}{2} \|x^k - x^*\|^2, \quad \forall k = 0, 1, \dots$$

Azaz ha  $\frac{LM\delta}{2} < 1$  és  $x^0 \in S_\delta(x^*)$ , akkor a konvergencia másodrendű.

**Biz.:** Lásd [2].

### Newton-módszer minimalizálásra

$$\begin{aligned} f : \mathbb{R}^n &\rightarrow \mathbb{R} \\ \min f(x) &=? \end{aligned}$$

Itt egy olyan pontsorozatot számolunk ki, ami egy lokális minimumhoz konvergál. Tegyük fel, hogy az  $x_0, x_1, \dots, x_k$  pontokat már kiszámoltuk.

Az  $f(x)$  függvényt az  $x_k$  közelében becsülhetjük a csonkított Taylor-sorával:

$$f(x) \approx f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k) =: h(x)$$

Az  $f(x)$  helyett a  $h(x)$ -et minimalizáljuk:

$$\nabla h(x) = \nabla f(x_k) + \nabla^2 f(x_k)(x - x_k) = 0.$$

A fenti egyenletből:

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k).$$

**Megjegyzés:** A Newton-módszer gyökkeresésre szélesebb körben alkalmazható. Az ott szereplő  $g(x)$  függvény nem feltétlenül egy másik függvény gradiense. Egy folytonosan differenciálható  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$  függvény pontosan akkor egy  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  függvény gradiense, ha a  $Dg(x)$  mátrix szimmetrikus minden  $x$ -re.

Ahhoz, hogy a lépés jól definiált legyen szükséges, hogy létezzen a Hesse-mátrix inverze. Ehhez tegyük fel, hogy az  $x^*$  lokális minimumban a  $\nabla^2 f(x^*)$  mátrix pozitív definit. Ekkor ha elég közel indítjuk az eljárást az  $x^*$ -hoz, akkor a  $\nabla^2 f(x_k)$  is pozitív definit lesz.

**2.3.2. Tétel.** Legyen  $f \in \mathcal{C}^3(\mathbb{R}^n)$ , és tegyük fel, hogy az  $x^*$  lokális minimumban a  $\nabla^2 f(x^*)$  Hesse-mátrix pozitív definit. Ekkor ha a Newton-módszert az  $x^*$ -hoz elég közel indítjuk, akkor a Newton-lépések jól definiáltak lesznek és a generált pontsorozat az  $x^*$ -hoz konvergál. A konvergencia rendje legalább kettő.

**Biz.:** Lásd [1].

Ha a Hesse-mátrix nem jól-kondicionált (a legnagyobb és a legkisebb sajátérték hányadosa nagy), akkor a mátrix csak nehezen invertálható. A számításigény csökkentésére különféle technikákat dolgoztak ki. A [2]-ben található több ismertetőt.

## 2.4. A barrier-módszer

A barrier-módszer a következő típusú feladatokra alkalmazható:

$$\min_{x \in S} f(x)$$

ahol  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  függvény, a megengedett  $S$  halmaz belseje nem üres, és  $S$  bármely pontjához van tetszőlegesen közeli belső pont.

A barrier módszer lényege, hogy a megengedett tartomány határán olyan gátat emel, ami megakadályozza, hogy a keresőeljárás elhagyja a tartományt.

A  $B : \text{int}(S) \rightarrow \mathbb{R}$  függvényt barrier függvénynek nevezzük, ha:

- (1)  $B$  folytonos
- (2)  $B(x) \geq 0$
- (3)  $B(x) \rightarrow \infty$ , ha  $x$  az  $S$  határához tart

**Példák:**  $S := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, p\}$

$$B_1(x) := - \sum_{i=1}^p \frac{1}{g_i(x)}, \quad x \in \text{int}(S)$$

$$B_2(x) := - \sum_{i=1}^p \log(-g_i(x)), \quad x \in \text{int}(S)$$

Nézzük  $k = 1, 2, \dots$ -re a következő részfeladatokat:

$$\begin{aligned} \min(f(x) + \mu_k B(x)) \\ x \in \text{int}S, \end{aligned}$$

ahol  $\mu_k$  monoton fogyó nullsorozat. Legyen a megoldása  $x_k$ .

**Megjegyzés:** *Látszólag egy sereg, az eredetinél bonyolultabb, feltételes optimalizálási feladatot kaptunk. Viszont ezeknél a feladatoknál használhatóak lesznek a feltétel nélküli keresési eljárások, a megengedett tartományban maradáást a barrier tag fogja biztosítani.*

**2.4.1. Tétel.** *A barrier módszer által generált  $\{x_k\}$  sorozat tetszőleges határpontja megoldása lesz az eredeti feladatnak.*

**Biz.:** Lásd [2].

### 3. Útkövető belsőpontos algoritmusok elmélete

#### 3.1. Az analitikus centrum

Legyen az  $\mathcal{S}$  halmaz a következő módon megadva:

$$\mathcal{S} = \{x \in \mathcal{X} \subseteq \mathbb{R}^n : g_j(x) \geq 0, j = 1, \dots, m\},$$

ahol a  $g_j$  függvények folytonosak és  $\text{int}(\mathcal{S}) \neq \emptyset$ .

$\text{int}(\mathcal{S})$ -en definiálunk egy potenciálfüggvényt:

$$\psi(x) := - \sum_{j=1}^m \log g_j(x).$$

Az  $\mathcal{S}$  analitikus centruma az a pont(halmaz), ami minimalizálja a potenciálfüggvényt.

**3.1. Példa.** Legyen  $\mathcal{S} \subseteq \mathbb{R}^n$  a következő egyenlőtlenségekkel megadva:

$$\begin{aligned} x_i &\geq 0 \\ (1 - x_i) &\geq 0 \end{aligned} \quad i = 1, 2, \dots, n$$

Vagyis  $\mathcal{S} = [0, 1]^n$ , az egységkocka.

$$\psi(x) = - \sum_{j=1}^n \log x_j - \sum_{j=1}^n \log(1 - x_j),$$

Keressük meg a  $\psi(x)$  minimumát:

$$\begin{aligned} (\nabla\psi(x))_j &= -\frac{1}{x_j} + \frac{1}{1-x_j} = 0 \\ x_j &= \frac{1}{2} \end{aligned}$$

Az analitikus centrum tehát az  $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$  pont, ami valóban a kocka geometriai középpontja.

**3.2. Példa.** Az analitikus centrum függ attól, hogyan adjuk meg a halmazt. Az  $\mathcal{S} = [0, 1]^n$  egységkockát a következő egyenlőtlenségekkel is megadhatjuk:

$$\begin{aligned} x_i &\geq 0 \\ (1 - x_i)^d &\geq 0 \end{aligned} \quad i = 1, 2, \dots, n$$

ahol a  $d$  tetszőleges, egynél nagyobb szám. Ezekkel az egyenlőtlenségekkel kiszámolva az analitikus centrumot, az  $(\frac{1}{d+1}, \frac{1}{d+1}, \dots, \frac{1}{d+1})$  pontot kapjuk. Ez a pont nagy  $d$  esetén sokkal közelebb van a kocka origó-beli csúcsához, mint a többihez.

Az analitikus centrumot felesleges egyenlőtlenségek hozzávétele is módosítja. Például ha ugyanazt az egyenlőtlenséget többször hozzávesszük.

Az analitikus centrum értelmezését kiterjeszthetjük arra az esetre is, ha az  $\text{int}(\mathcal{S}) = \emptyset$ . Nézzük a következő speciális esetet:

$$\mathcal{S} = \{x \in \mathcal{X} \subseteq \mathbb{R}^n : x \geq 0\},$$

ahol az  $\mathcal{X}$  egy affin altér. Tegyük fel, hogy az  $\mathcal{S}$  korlátos (ekkor kompakt is).

Definiáljuk az  $\mathcal{S}$  tartóját a következőképpen:

$$\sigma(\mathcal{S}) := \{i : x_i > 0 \text{ valamely } x \in \mathcal{S} \text{ esetén}\}.$$

Az  $\mathcal{S}$  analitikus centruma  $\sigma(\mathcal{S}) = \emptyset$  esetén a 0 vektor, különben pedig az az  $x \in \mathcal{S}$  vektor, ami maximalizálja a

$$\prod_{i \in \sigma(\mathcal{S})} x_i$$

szorzatot.

Ha  $\sigma(\mathcal{S}) \neq \emptyset$ , akkor az  $\mathcal{S}$  kompaktsága miatt véges a maximum. Az  $\mathcal{S}$  konvexitása miatt pedig lesz olyan  $x \in \mathcal{S}$  vektor, amire  $x_{\sigma(\mathcal{S})} > 0$ , azaz a maximum pozitív.

A szorzat logaritmusá szigorúan konkáv, így a maximum értéke egyetlen  $\mathcal{S}$ -beli pontban vétetik fel, az analitikus centrum ebben az esetben egyértelmű.

**3.3. Példa.** Nézzük a primál-duál optimális lapot:

$$\mathcal{F}_{x,s}^* = \{(x, s) \in \mathbb{R}^n \times \mathbb{R}^n : Ax = b, A^T y + s = c, c^T x = p^*, x \geq 0, s \geq 0\}$$

Tegyük fel, hogy  $\mathcal{F}_{x,s}^* \neq \emptyset$  és korlátos. Vezessük be a következő indexhalmazokat:

$$\mathcal{P}^* = \{i : x_i > 0, \text{ valamely } x \in \mathcal{F}_{x,s}^* \text{ esetén}\}$$

$$\mathcal{D}^* = \{j : s_j > 0, \text{ valamely } (y, s) \in \mathcal{F}_{x,s}^* \text{ esetén}\}$$

Ekkor az  $\mathcal{F}_{x,s}^*$  tartója a  $\mathcal{P}^* \dot{\cup} \mathcal{D}^*$  indexhalmaz. Ekkor  $\mathcal{F}_{x,s}^*$  analitikus centruma az az egyértelmű  $(x^*, s^*)$  pont, ami maximalizálja a

$$\prod_{i \in \mathcal{P}^*} x_i \prod_{j \in \mathcal{D}^*} s_j$$

szorzatot. Ezt az eredményt és a  $\mathcal{P}^*$ ,  $\mathcal{D}^*$  halmazokat a későbbiekben még használni fogjuk.

## 3.2. Newton módszer az LP-feladatra

Ahhoz, hogy megkapjunk egy optimális primál-duál megoldáspárt, "elég" megoldanunk a következő rendszert:

$$\left. \begin{array}{l} Ax = b \quad x \geq 0, \\ A^T y + s = c \quad s \geq 0, \\ xs = 0. \end{array} \right\} OPT$$

Ugyanezt más formában felírva:

$$F_0(x, y, s) := \begin{pmatrix} A^T y + s - c \\ Ax - b \\ Xs \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{array}{l} x \geq 0 \\ s \geq 0 \end{array}$$

Próbáljuk megoldani ezt a nemlineáris rendszert a Newton-módszerrel:

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \\ s^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ y^k \\ s^k \end{pmatrix} - [DF_0(x^k, y^k, s^k)]^{-1} F_0(x^k, y^k, s^k)$$

ahol

$$DF_0 = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix},$$

az  $F_0$  Jacobi-mátrixa.

Legyen egy megoldás a  $p^* := (x^*, y^*, s^*)$ . Az  $Xs = 0$  feltétel miatt a  $p^*$  a megengedett tartomány határán lesz, és ott sajnos nem garantálható a Jacobi-mátrix invertálhatósága. Így itt a Newton-módszer nem vezet eredményre.

### 3.3. Barrier módszer az $LP$ -feladatra

Tekintsük az  $LP$ -feladat standard alakját:

$$\left. \begin{array}{l} \min c^T x \\ Ax = b \\ x \geq 0. \end{array} \right\} (P)$$

Tegyük fel, hogy  $\mathcal{F}_P^0 \neq \emptyset$  és  $\mathcal{F}_P^*$  korlátos.

Legyen a barrierfüggvénnyel módosított célfüggvény a következő:

$$f_P(x; \mu) := c^T x - \mu \sum_{i=1}^n \log x_i, \quad x \in \mathcal{F}_P^0$$

Minden  $\mu \geq 0$ -ra a következő barrier feladatot kapjuk:

$$\left. \begin{array}{l} \min f_P(x; \mu) \\ Ax = b \\ x > 0 \end{array} \right\} BP(\mu)$$

Világos, hogy  $\mu = 0$ -ra az eredeti  $LP$ -feladatot kapjuk. Ha a  $\mu \rightarrow \infty$ , akkor a  $BP(\mu)$  megoldásai az  $\mathcal{F}_P$  megengedett tartomány analitikus centrumához tartanak, hiszen ekkor a  $c^T x$  elhanyagolható, a barrier rész pedig éppen az  $\mathcal{F}_P^0$ -hoz tartozó potenciálfüggvény.

A későbbiekben belátjuk majd, hogy minden  $\mu > 0$ -ra egyértelmű  $x(\mu)$  megoldása lesz a  $BP(\mu)$  feladatnak. Az  $x(\mu)$  pontok adják a primál centrális utat, ami  $\mu \rightarrow 0$  esetén az optimális lap analitikus centrumához fog tartani. Tehát a primál centrális út az  $\mathcal{F}_P$  analitikus centrumától halad az  $\mathcal{F}_P^*$  analitikus centruma felé.

Most vizsgáljuk meg a duális feladatot:

$$\left. \begin{array}{l} \max b^T y \\ A^T y + s = c \\ s \geq 0. \end{array} \right\} (D)$$

Alkalmazva a barrier-módszert, készítsük el a hozzá tartozó barrier-feladatot:

$$f_D(y, s; \mu) := b^T y + \mu \sum_{i=1}^n \log s_i, \quad (y, s) \in \mathcal{F}_D^0$$

$$\left. \begin{array}{l} \max f_D(y, s; \mu) \\ A^T y + s = c \\ s > 0 \end{array} \right\} BD(\mu)$$

Tegyük fel, hogy  $\mathcal{F}_D^0 \neq \emptyset$  és  $\mathcal{F}_D^*$  korlátos. Ha a  $\mu$  folytonosan halad a 0 felé, akkor a  $BD(\mu)$  feladatok egyértelmű  $(y(\mu), s(\mu))$  megoldásai egy utat definiálnak, amit duál centrális útnak nevezünk.

Elmondhatjuk ugyanazt itt is, mint a primál esetben. A  $\mu = 0$ -ra az eredeti feladatot kapjuk. Ha a  $\mu \rightarrow \infty$ , akkor az  $(y(\mu), s(\mu))$   $\mu$ -centrumok az  $\mathcal{F}_D$  megengedett tartomány analitikus centrumához tartanak, míg  $\mu \rightarrow 0$  esetén az  $\mathcal{F}_D^0$  duál optimális lap analitikus centrumához.

**3.3.1. Tétel.** *Legyen  $\mu > 0$ . Ekkor a következő állítások ekvivalensek:*

- (1)  $\mathcal{F}_P^0 \neq \emptyset$  és  $\mathcal{F}_D^0 \neq \emptyset$
- (2) egyértelműen létezik megoldása a  $BP(\mu)$  feladatnak
- (3) egyértelműen létezik megoldása a  $BD(\mu)$  feladatnak
- (4) egyértelműen létezik megoldása a következő rendszernek:

$$F_\mu(x, y, s) := \left( \begin{array}{c} A^T y + s - c \\ Ax - b \\ Xs - \mu e \end{array} \right) = \left( \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right), \quad \left. \begin{array}{l} x \geq 0 \\ s \geq 0 \end{array} \right\} OPT(\mu)$$

**Biz.:** (2) $\Leftrightarrow$ (4) Az  $f_P$  függvény gradiense:

$$\nabla f_P(x; \mu) = c - \mu X^{-1}e,$$

és Hesse-mátrixa:

$$\nabla^2 f_P(x; \mu) = \mu X^{-2}.$$

A Hesse mátrix pozitív definit minden  $x \in \mathcal{F}_P^0$  pontban, azaz a függvény itt szigorúan konvex.

A  $BP(\mu)$  tehát egy konvex programozási feladat. A megengedett tartomány minden pontja reguláris a  $\text{rang}(A) = m$  feltétel miatt. Így  $BP(\mu)$  megoldása ekvivalens a hozzá tartozó  $KKT$ -rendszer megoldásával.

A  $BP(\mu)$  Lagrange-függvénye:

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

a  $KKT$ -rendszer:

$$\left. \begin{aligned} \nabla_x L(x, y) &= c - \mu X^{-1} e + y^T A = 0 \\ \nabla_y L(x, y) &= Ax - b = 0 \end{aligned} \right\} \quad (KKT)$$

Legyen  $s = \mu X^{-1} e$ , ekvivalensen  $Xs = \mu e$ . Azaz a  $KKT$ -rendszer ekvivalens az  $OPT(\mu)$  rendszerrel.

(3)  $\Leftrightarrow$  (4) Az  $f_D$  függvény gradiense:

$$\nabla f_D(y, s; \mu) = (b^T, e^T S^{-1}),$$

és Hesse-mátrixa:

$$\nabla^2 f_D(y, s; \mu) = \left[ \begin{array}{c|c} 0 & 0 \\ \hline 0 & -\mu S^{-2} \end{array} \right].$$

A Hesse-mátrix negatív szemidefinit minden  $(y, s) \in \mathcal{F}_D^0$  pontban, azaz a függvény itt konkáv. A  $BD(\mu)$  is egy konvex programozási feladat. Minden megengedett megoldás reguláris pont lesz, a slack változók jelenléte miatt. Ezért a  $BD(\mu)$  rendszer megoldása ekvivalens lesz a hozzá tartozó  $KKT$ -rendszer megoldásával.

A  $BD(\mu)$  Lagrange-függvénye:

$$L(y, s, x) = y^T b + \mu \sum_{i=1}^n \log s_i - (y^T A + s^T - c^T)x,$$

a  $KKT$ -rendszer:

$$\left. \begin{aligned} \nabla_{y,s} L(y, s, x) &= (b^T - (Ax)^T, \mu e S^{-1} - x^T) = 0 \\ \nabla_x L(y, s, x) &= -y^T - s^T + c^T = 0 \end{aligned} \right\} \quad (KKT).$$

Azaz a  $BD(\mu)$  feladathoz tartozó  $KKT$ -rendszer is éppen az  $OPT(\mu)$  rendszer.

(4)  $\Rightarrow$  (1): Az  $OPT(\mu)$  megoldásai nyilván megengedett megoldások, és az  $Xs = \mu e$  feltétel miatt belsőpontok is.

□

### 3.4. A centrális út

A következőkben a belsőpont-feltételt végig feltesszük. Az ekkor minden  $0 < \mu < \infty$  esetén egyértelműen létező  $(x^*(\mu), y^*(\mu), s^*(\mu))$  megoldások folytonos útját primál-duál centrális útnak nevezzük,  $\mathcal{C}$ -vel jelöljük. A  $\text{rang}(A) = m$  feltétel miatt kölcsönösen egyértelmű megfeleltetés van az  $y$  és  $s$  között. Ez lehetővé teszi, hogy az  $\{(x(\mu), s(\mu)) : 0 < \mu < \infty\}$  utat tekintsük primál-duál centrális útnak.



**3.4.1. Tétel.** *Tegyük fel, hogy  $\mathcal{F}_0 \neq \emptyset$ . Ekkor:*

- (1) *Az  $\{(x(\mu), s(\mu)) : 0 < \mu \leq \mu^0\}$  ponthalmaz korlátos minden  $0 < \mu^0 < \infty$  esetén.*
- (2)  *$\lim_{\mu \rightarrow 0} (x(\mu), s(\mu)) = (x^*, s^*)$ , ahol az  $(x^*, s^*)$  egy szigorúan komplementáris megoldáspár, ami megegyezik a  $\mathcal{F}_{x,s}^*$  optimális lap analitikus centrumával.*

**Biz.: (1)**

$$A(x(\mu^0) - x(\mu)) = b - b = 0 \text{ és } A^T(y(\mu^0) - y(\mu)) + (s(\mu^0) - s(\mu)) = 0,$$

azaz

$$(x(\mu^0) - x(\mu)) \in \mathcal{N}(A) \text{ és } (s(\mu^0) - s(\mu)) \in \mathcal{R}(A^T),$$

ezért:

$$(x(\mu^0) - x(\mu))^T (s(\mu^0) - s(\mu)) = 0$$

Kifejtve kapjuk:

$$\sum_{i=1}^n (s(\mu^0)_i x(\mu)_i + x(\mu^0)_i s(\mu)_i) = n(\mu^0 + \mu) \leq 2n\mu^0$$

Végigosztva  $\mu^0$ -lal:

$$\sum_{i=1}^n \left( \frac{x(\mu)_i}{x(\mu^0)_i} + \frac{s(\mu)_i}{s(\mu^0)_i} \right) \leq 2n,$$

hiszen  $s(\mu^0)_i x(\mu^0)_i = \mu^0$  minden  $i$ -re. Azaz  $x(\mu)$  és  $s(\mu)$  korlátosak.

(2) Az (1)-es miatt az  $\{(x(\mu), s(\mu)) : 0 < \mu \leq \mu^0\}$  halmaznak van legalább egy határpontja. Legyen az  $(x(0), s(0))$  egy tetszőleges határpont. Feltehető, hogy  $(x(0), s(0)) = \lim_{k \rightarrow \infty} (x(\mu_k), s(\mu_k))$ .

**I.  $(\mathbf{x}(0), \mathbf{s}(0)) \in \mathcal{F}_{\mathbf{x},\mathbf{s}}^*$**

Ez nyilvánvaló:

$$x(0)s(0) = \lim_{k \rightarrow \infty} (x(\mu_k)s(\mu_k)) = \lim_{k \rightarrow \infty} (\mu_k e) = 0$$

**II.  $(\mathbf{x}(0), \mathbf{s}(0))$  szigorúan komplementáris megoldás**

Az (1)-nél látottak alapján:

$$(x(0) - x(\mu))^T (s(0) - s(\mu)) = 0$$

Kifejtve:

$$\sum_{i=1}^n (s(0)_i x(\mu)_i + x(0)_i s(\mu)_i) = \mu n$$

Másképpen:

$$\sum_{i: x(0)_i > 0} x(0)_i \frac{s(\mu)_i}{\mu} + \sum_{i: s(0)_i > 0} s(0)_i \frac{x(\mu)_i}{\mu} = n$$

Felhasználva, hogy  $\mu = x(\mu)_i s(\mu)_i$  minden  $i$ -re:

$$\sum_{i:x(0)_i>0} \frac{x(0)_i}{x(\mu)_i} + \sum_{i:s(0)_i>0} \frac{s(0)_i}{s(\mu)_i} = n$$

$\mu \rightarrow 0$  határátmenettel kapjuk:

$$|\sigma(x(0))| + |\sigma(s(0))| = n,$$

amiből következik **II.**

**III. a korábban definiált  $\mathcal{P}^*$  és  $\mathcal{D}^*$ -re:**  $\mathcal{P}^* \cap \mathcal{D}^* = \emptyset$ ,  $\mathcal{P}^* \cup \mathcal{D}^* = \{1, 2, \dots, n\}$ .

Az állítás második része következik abból, hogy létezik szigorúan komplementáris megoldás. Az első rész bizonyításához indirekt tegyük fel, hogy  $i \in \mathcal{P}^* \cap \mathcal{D}^*$ . Azaz létezik  $x \in \mathcal{F}_P^*$ ,  $(y, s) \in \mathcal{F}_D^*$ , amikre  $x_i s_i > 0$ . Ez ellentmondás, hiszen az optimális megoldások pontosan a komplementárisak.

**IV. az  $(x(0), s(0))$  megegyezik az  $\mathcal{F}_{x,s}^*$  analitikus centrumával**

Legyen  $(x^*, s^*) \in \mathcal{F}_{x,s}^*$  tetszőleges. A **II.**-nél látott módon, az

$$(x^* - x(\mu))^T (s^* - s(\mu)) = 0$$

egyenletből kiindulva a következő összefüggésre jutunk:

$$\sum_{i \in \mathcal{P}^*} \frac{x_i^*}{x(0)_i} + \sum_{i \in \mathcal{D}^*} \frac{s_i^*}{s(0)_i} = n.$$

Mivel  $x_i > 0 \forall i \in \mathcal{P}^*$  és  $s_i > 0 \forall i \in \mathcal{D}^*$ , alkalmazhatjuk a számtani-mértani közép közti egyenlőtlenséget:

$$\left( \prod_{i \in \mathcal{P}^*} \frac{x_i^*}{x(0)_i} \prod_{i \in \mathcal{D}^*} \frac{s_i^*}{s(0)_i} \right)^{\frac{1}{n}} \leq \frac{1}{n} \left( \sum_{i \in \mathcal{P}^*} \frac{x_i^*}{x(0)_i} + \sum_{i \in \mathcal{D}^*} \frac{s_i^*}{s(0)_i} \right) = 1.$$

Így

$$\prod_{i \in \mathcal{P}^*} x_i^* \prod_{i \in \mathcal{D}^*} s_i^* \leq \prod_{i \in \mathcal{P}^*} x(0)_i \prod_{i \in \mathcal{D}^*} s(0)_i$$

ebből következik **IV.**

A korábbiakban láttuk, hogy az  $\mathcal{F}_{x,s}^*$  lapnak egyértelműen létezik az analitikus centruma, így  $\lim_{\mu \rightarrow 0} (x(\mu), s(\mu)) = (x(0), s(0))$ .  $\square$

Az előző tétel alapján a primál-duál centrális út  $\mu \rightarrow 0$  esetén az optimális lap analitikus centrumához konvergál, amennyiben ez a lap korlátos.

Könnyen látható, hogy ha a  $\mathcal{F}_{x,s}$  poliéder is korlátos, akkor a primál-duál centrális út  $\mu \rightarrow \infty$  esetén a  $\mathcal{F}_{x,s}$  analitikus centrumához konvergál.

Ezeket az állításokat könnyen ellenőrizhetjük külön-külön a primál és a duál centrális utakra, ugyanis az  $(x^*, s^*)$  pontosan akkor analitikus centruma  $\mathcal{F}_{x,s}^*$ -nak, ha  $x^*$  analitikus centruma  $\mathcal{F}_P^*$ -nak és  $(y^*, s^*)$  analitikus centruma  $\mathcal{F}_D^*$ -nak.

## Miért az analitikus jelző?

Vizsgáljuk az

$$(x(\mu), s(\mu)) : \mathbb{R}_+ \rightarrow \mathbb{R}^n \times \mathbb{R}^n$$

függvényt.

Az

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

függvény folytonosan differenciálható és hamarosan belátjuk, hogy a Jacobi-mátrixa nonszinguláris a belső pontokban. Tudjuk, hogy:

$$F_\mu(x(\mu), y(\mu), s(\mu)) = 0 \quad \forall \mu > 0.$$

Alkalmazva az implicitfüggvény-tételt kiderül, hogy a centrális út analitikus minden  $\mu > 0$  pontban.

Mivel a centrális út konvergál  $\mu \rightarrow 0$  esetén, kiterjeszthetjük az  $(x(\mu), s(\mu))$  függvényt az  $\mathbb{R}_\oplus$ -ra:

$$(x(0), s(0)) := \lim_{\mu \rightarrow 0} (x(\mu), s(\mu)).$$

A centrális út a  $\mu = 0$  pontban is analitikus lesz, ennek belátására viszont nem használhatjuk az implicitfüggvény-tételt a Jacobi-mátrix szingularitása miatt. A bizonyítás megtalálható [6]-ben.

## 3.5. A Newton-lépés

Amint korábban láttuk az  $OPT$  rendszer megoldása az  $\mathcal{F}$  megengedett megoldáshalmaz határán van. Ebben az esetben nem garantálható, hogy a  $DF_0$  Jacobi-mátrix invertálható, így itt nem működik a Newton-módszer. A relaxált  $OPT(\mu)$  rendszer megoldásai azonban mindig belső pontok.

### 3.5.1. Tétel. A

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

Jacobi-mátrix reguláris, ha  $x > 0$  és  $s > 0$  (azaz ha  $(x, y, s) \in \mathcal{F}^0$ ).

**Biz.:** Azt vizsgáljuk, hogy lehet-e nemtriviális megoldása a  $(DF_\mu)z = 0$  egyenletnek. Legyen  $z = (u, v, w) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ . Kifejtve a homogén egyenletet:

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

Ekkor

$$u^T w = u^T (-A^T v) = -(Au)^T v = 0.$$

Az  $Su + Xw = 0$  egyenletből

$$u = -S^{-1}Xw$$

A kettőt összerakva:

$$0 = w^T u = w^T (-S^{-1}X)w.$$

A belső pontokra a  $-S^{-1}X$  mátrix negatív definit, így a  $w$  csak a  $0$  lehet. A  $0 = Su + Xw = Su$  egyenletből kapjuk, hogy az  $u = 0$ , és a  $0 = A^T v + w = A^T v$  egyenletből a  $\text{rank}(A) = m$  miatt a  $v = 0$ .

Nincs nemtriviális megoldás, így a  $DF_\mu$  mátrix tényleg reguláris.  $\square$

Így az  $\text{OPT}(\mu)$  rendszer megoldására használhatjuk a Newton-módszert.

$$\begin{pmatrix} x^+ \\ y^+ \\ s^+ \end{pmatrix} = \begin{pmatrix} x \\ y \\ s \end{pmatrix} - \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix}$$

ahol

$$(\Delta x, \Delta y, \Delta s) = DF_\mu(x, y, s)^{-1} F_\mu(x, y, s),$$

azaz

$$DF_\mu(x, y, s)(\Delta x, \Delta y, \Delta s) = F_\mu(x, y, s).$$

Definiáljuk a következő maradékokat az  $(x, y, s)$  pontban:

$$\begin{aligned} r_d &:= A^T y + s + c \\ r_p &:= Ax - b \\ r_c &:= Xs - \mu e \end{aligned}$$

Ezekkel a jelölésekkel a  $(x, y, s)$ -beli Newton irány a következő rendszer megoldása:

$$DF_\mu(x, y, s) \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} r_d \\ r_p \\ r_c \end{pmatrix}$$

A Jacobi-mátrixot beírva a következőt kapjuk:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} r_d \\ r_p \\ r_c \end{pmatrix}$$

**3.5.2. Tétel.** *Az előző rendszer megoldása a következő:*

$$\begin{aligned} \Delta y &= (AXS^{-1}A^T)^{-1}(r_p - AS^{-1}(r_c - Xr_d)) \\ \Delta s &= r_d - A^T \Delta y \\ \Delta x &= S^{-1}(r_c - X\Delta s) \end{aligned}$$

**Biz.:** Ellenőrzés.  $\square$

### 3.6. A centrális út környezetei

#### Az $\mathcal{N}_2$ környezet

Szeretnénk mérni, hogy egy  $(x, y, s) \in \mathcal{F}^0$  pont milyen messze van a centrális úttól. Ehhez vizsgáljuk az  $\min_{\mu} \|F_{\mu}(x, y, s)\|$  kifejezést. Az értéke pontosan akkor lesz 0, ha az  $(x, y, s)$  pont a centrális úton van.

$$\min_{\mu} \|F_{\mu}(x, y, s)\| = \min_{\mu} \|XSe - \mu e\|$$

Az  $XSe$  vektor  $\mu e$  irányú merőleges vetülete  $\frac{x^T s}{n} e$ , ezért

$$\min_{\mu} \|F_{\mu}(x, y, s)\| = \|XSe - \frac{x^T s}{n} e\| = \|xs - \frac{x^T s}{n} e\|$$

Legyen  $\mu(x, s) = \frac{x^T s}{n}$ . Ekkor az  $(x, y, s)$ -hez legközelebbi centrum a  $\mu(x, s)$ -centrum.

Ezek alapján bevezetünk egy centralitási mértéket:

$$\delta_2(x, y, s) := \left\| \frac{xs}{\mu(x, s)} - e \right\|_2$$

A  $\delta_2$  távolság segítségével értelmezhetjük a centrális út környezetét:

$$\mathcal{N}_2(\gamma) := \{(x, y, s) \in \mathcal{F}^0 : \delta_2(x, y, s) \leq \gamma\}.$$

Könnyen látható, hogy ha  $0 \leq \gamma_1 \leq \gamma_2 \leq 1$ :

$$\mathcal{C} = \mathcal{N}_2(0) \subseteq \mathcal{N}_2(\theta_1) \subseteq \mathcal{N}_2(\theta_2) \subseteq \mathcal{N}_2(1) \subset \mathcal{F}^0.$$

A környezet szemléltetéséhez vezessük be a következő függvényt:

$$\omega : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \omega(x, y, s) := xs$$

Az  $\omega$  a  $\mathcal{C}$  centrális utat a  $\{\mu e : \mu > 0\}$  "átlós" félegyenesbe viszi. Használva az  $\omega = \omega(x, y, s)$  és  $\mu_{\omega} = \mu(x, s)$  jelölést:

$$(x, y, s) \in \mathcal{N}_2(\gamma) \Leftrightarrow \|\omega - \mu_{\omega} e\| \leq \gamma \mu_{\omega} \Leftrightarrow \|\omega - \mu_{\omega} e\|^2 \leq \gamma^2 \mu_{\omega}^2.$$

A jobboldali egyenlőtlenséget tovább alakítva kapjuk a

$$\langle \omega, \omega \rangle - \frac{n + \gamma^2}{n^2} \langle \omega, ee^T \omega \rangle \leq 0$$

feltételt.

Az  $A_{\gamma} := I - \left(\frac{n + \gamma^2}{n^2}\right) ee^T$  mátrix segítségével a következő egyszerű feltételt kapjuk:

$$\langle \omega, A_{\gamma} \omega \rangle \leq 0.$$

Az  $A_\gamma$  mátrixnak sajátvektora az  $e$ , a  $-\frac{\gamma^2}{n}$  sajátértékkel, és sajátvektora minden  $u \perp e$  vektor is, az 1 sajátértékkel. Legyen  $v_1, \dots, v_{n-1}, v_n$  ortonormált bázis, amire  $v_n = \frac{1}{\sqrt{n}}e$ . Ekkor minden  $\omega$  vektorra:

$$\omega = \sum_{i=1}^n \langle \omega, v_i \rangle v_i =: \sum_{i=1}^n \lambda_i v_i.$$

Ekkor:

$$\langle \omega, A_\gamma \omega \rangle = \lambda_1^2 + \dots + \lambda_{n-1}^2 - \frac{\gamma^2}{n} \lambda_n^2.$$

Azaz

$$(x, y, s) \in \mathcal{N}_2(\gamma) \Leftrightarrow \lambda_1^2 + \dots + \lambda_{n-1}^2 \leq \frac{\gamma^2}{n} \lambda_n^2.$$

Az  $\mathcal{N}_2(\gamma)$  környezet  $\omega$ -képe tehát egy körkép, aminek szimmetriatengelye éppen a centrális út  $\omega$ -képe.

### Az $\mathcal{N}_{-\infty}$ környezet

Más centralitási mértéket kapunk, ha nem az euklideszi normát használjuk.

Legyen:

$$\delta_\infty(x, y, s) = \left\| \frac{xs}{\mu(x, s)} - e \right\|_\infty,$$

és

$$\delta_{-\infty}(x, y, s) = \left\| \left[ \frac{xs}{\mu(x, s)} - e \right]^- \right\|_\infty,$$

ahol  $[x]_i^- = 0$ , ha  $x_i \geq 0$  és  $[x]_i^- = x_i$ , ha  $x_i \leq 0$ .

Az ebből származó környezetek:

$$\mathcal{N}_\infty(\gamma) := \{(x, y, s) \in \mathcal{F}^0 : \delta_\infty(x, y, s) \leq \gamma\},$$

$$\mathcal{N}_{-\infty}(\gamma) := \{(x, y, s) \in \mathcal{F}^0 : \delta_{-\infty}(x, y, s) \leq \gamma\},$$

valamely  $\gamma \in [0, 1]$ -re.

Mi a későbbiek folyamán az  $\mathcal{N}_{-\infty}$  környezetet fogjuk használni.

Világos, hogy ha  $0 \leq \gamma_1 \leq \gamma_2 \leq 1$ , akkor

$$\mathcal{C} = \mathcal{N}_{-\infty}(0) \subseteq \mathcal{N}_{-\infty}(\gamma_1) \subseteq \mathcal{N}_{-\infty}(\gamma_2) \subseteq \mathcal{N}_{-\infty}(1) = \mathcal{F}^0.$$

#### 3.6.1. Állítás.

$$\mathcal{N}_2(\gamma) \subseteq \mathcal{N}_\infty(\gamma) \subseteq \mathcal{N}_{-\infty}(\gamma).$$

**Biz.:** minden  $i$ -re teljesül, hogy

$$|x_i s_i - \mu(x, s)| \leq \|xs - \mu(x, s)e\|_2,$$

ebből látszik az első tartalmazás. A második tartalmazás triviális.  $\square$

## 4. Primál-duál útkövető algoritmusok

Egy keretalgoritmus három konkrét variációját nézzük meg: egy rövid lépéses, egy hosszú lépéses és egy prediktor-korrektor verziót. A hosszú lépéses algoritmustól eltekintve a lehető legjobb lépésszámot hozzák,  $O(\sqrt{n}L)$  iterációval. A hosszú lépéses verzió elméleti lépésszámbeceklése  $O(nL)$  iteráció, azonban a gyakorlatban hatékonyabb a rövid lépéses testvérénél. Ezt a jelenséget "az elmélet és a gyakorlat közti rés"-nek nevezik.

### 4.1. A keretalgoritmus

---

Belsőpontos keretalgoritmus

---

**Input**

$\epsilon > 0$  pontossági paraméter,  $(x^0, y^0, s^0)$  kezdő belső pont

**begin**

$(x, y, s) := (x^0, y^0, s^0)$ .

**while**  $n\mu > \epsilon$  **do**

$\mu := \mu(x, s)$

Oldjuk meg a következő egyenletet

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

Legyen  $(x, y, s) = (x, y, s) - \alpha(\Delta x, \Delta y, \Delta s)$ , ahol  $\alpha \in (0, 1]$  egy

olyan lépéshossz, amire  $(x, s) > 0$ ,

$\sigma \in [0, 1]$  a centralizáló paraméter.

**end**

**end**

---

A  $(\Delta x, \Delta y, \Delta s)$  egy Newton-lépést ad az  $(x(\sigma\mu), y(\sigma\mu), s(\sigma\mu)) \in \mathcal{C}$  centrum felé. A  $\sigma$  paraméter arra szolgál, hogy egyensúlyt tartsunk a centrális út felé történő, és a vele párhuzamos lépések között. Ha  $\sigma = 1$ , akkor a Newton-lépés épp azt a centrumot közelíti, amelyik a legközelebb van, azaz a centrális útra merőlegesen lépünk. Az ilyen lépést centralizáló lépésnek nevezik. A másik szélső eset, ha  $\sigma = 0$ . Ekkor egyenesen egy optimális pontot célozunk meg, azaz az optimális lap felé lépünk. Az ilyen lépést affin skálázási lépésnek nevezik.

A  $\sigma$  és  $\alpha$  választásától függően különböző algoritmusokat kapunk.

A következő lemma mutatja, hogy a  $\sigma$  és az  $\alpha$  függvényében mennyire csökken a  $\mu$ . Használjuk a következő jelöléseket:

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

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

**4.1.1. Lemma.** Az algoritmusban kiszámolt  $(\Delta x, \Delta y, \Delta s)$ -re fennállnak a következők:

$$i) \Delta x^T \Delta s = 0$$

$$ii) \mu^+ = (1 - \alpha(1 - \sigma))\mu$$

**Biz.:** i) Tudjuk, hogy

$$A\Delta x = 0 \tag{1}$$

$$A^T \Delta y + \Delta s = 0. \tag{2}$$

A (2)-est  $(\Delta x)^T$ -tal szorozva kapjuk:

$$0 = \langle \Delta x, A^T \Delta y \rangle + \langle \Delta x, \Delta s \rangle = \langle A\Delta x, \Delta y \rangle + \langle \Delta x, \Delta s \rangle,$$

így az (1) miatt  $\langle \Delta x, \Delta s \rangle = 0$ .

ii) Az  $S\Delta x + X\Delta s = XSe - \sigma\mu e$  egyenletet komponensenként összegezve:

$$\begin{aligned} \langle s, \Delta x \rangle + \langle x, \Delta s \rangle &= \langle x, s \rangle - \sigma \langle x, s \rangle \\ &= (1 - \sigma) \langle x, s \rangle \end{aligned}$$

.

Ezek alapján:

$$\begin{aligned} n\mu^+ &= \langle x^+, s^+ \rangle \\ &= \langle x, s \rangle - \alpha(\langle s, \Delta x \rangle + \langle x, \Delta s \rangle) \\ &= (1 - \alpha(1 - \sigma)) \langle x, s \rangle, \end{aligned}$$

amiből következik ii). □

## 4.2. Egy rövid lépéses algoritmus

Rövid lépéses algoritmusról beszélünk, ha a  $\sigma$  a dimenziótól függő konstans. A következő algoritmusnál  $\alpha = 1$  és  $\sigma = \left(1 - \frac{\nu}{\sqrt{n}}\right)$ .



**Input**

$\epsilon > 0$  pontossági paraméter,

$\gamma$  és  $\nu$  konstansok, amelyek kielégítik a következőket:

$$\begin{aligned} 0 \leq \gamma < \frac{1}{2}, \quad 0 < \nu < \sqrt{n}, \\ \frac{\gamma^2 + \nu^2}{2(1-\gamma)} \leq \gamma \left(1 - \frac{\nu}{\sqrt{n}}\right) \end{aligned} \quad (3)$$

$(x^0, y^0, s^0) \in \mathcal{N}_2(\gamma)$ .

**begin**

$(x, y, s) := (x^0, y^0, s^0)$ ,  $\mu := \mu(x, s)$

**while**  $n\mu > \epsilon$  **do**

Határozzuk meg azt az  $(x, y, s)$ -beli  $(\Delta x, \Delta y, \Delta s)$  Newton-irányt,

ami a  $\sigma\mu$ -centrumot közelíti, ahol  $\sigma = \left(1 - \frac{\nu}{\sqrt{n}}\right)$ .

Legyen  $(x, y, s) = (x, y, s) - (\Delta x, \Delta y, \Delta s)$  és  $\mu := \mu(x, s)$ .

**end****end**

**4.2.1. Tétel.** *Legyenek  $\gamma$  és  $\nu$  a (3) feltételeket kielégítő konstansok, és tegyük fel, hogy  $(x, y, s) \in \mathcal{N}_2(\gamma)$ . Ekkor:*

i)  $(x^+, y^+, s^+) \in \mathcal{N}_2(\gamma)$

ii)  $\mu^+ = \left(1 - \frac{\nu}{\sqrt{n}}\right)\mu$

**Biz.:** i) A szakasz végén bizonyítjuk, a következő lemmák felhasználásával.

ii) A 4.1.1 lemma ii) részéből egyenesen következik.

**4.2.1. Köv.** *Az algoritmus során nem lépünk ki az  $\mathcal{N}_2(\gamma)$  környezetből és a  $k$ . iteráció után:*

$$\mu^k = \left(1 - \frac{\nu}{\sqrt{n}}\right)^k \mu^0$$

□

**4.2.2. Tétel.** *Az algoritmus legfeljebb  $K = \lceil \log(n\epsilon^{-1}\mu^0) \frac{\sqrt{n}}{\nu} \rceil$  iteráció után  $\epsilon$ -optimális megoldást ad.*

**Biz.:** Az algoritmus leáll, ha  $n\mu^k \leq \epsilon$ , belátjuk, hogy ez teljesül a  $k = K$ -ra:

$$\log(n\mu^K) = \log \left[ n\mu^0 \left(1 - \frac{\nu}{\sqrt{n}}\right)^K \right] = \log(n\mu^0) + K \log \left(1 - \frac{\nu}{\sqrt{n}}\right) \leq$$

$$\leq \log(n\mu^0) - K \frac{\nu}{\sqrt{n}} \leq \log(n\mu^0) - \left[ (\log(n\mu^0) - \log \varepsilon) \frac{\sqrt{n}}{\nu} \right] \frac{\nu}{\sqrt{n}} = \log \varepsilon.$$

Az első egyenlőség az előző következmény miatt áll fenn, míg az első egyenlőtlenségnél azt a tényt használjuk, hogy  $\log(1-x) \leq -x$  ha  $x < 1$ .  $\square$

A 4.2.1 tétel i) részének bizonyításához szükség van néhány segédállításra.

**4.2.1. Lemma.** *Legyen  $u, v \in \mathbb{R}^n$  két vektor, melyekre  $u^T v \geq 0$ . Ekkor*

$$\|uv\| \geq \sqrt{2^{-3}} \|u+v\|^2.$$

**Biz.:** Bontsuk szét az  $I = \{1, 2, \dots, n\}$  indexhalmazt két részre:

$$P := \{i \in I \mid u_i v_i \geq 0\}, \quad M := \{i \in I \mid u_i v_i < 0\},$$

és egy  $T \subseteq I$  indexhalmazra legyen

$$z_T = \sum_{i \in T} u_i v_i e_i.$$

Ekkor

$$\begin{aligned} \|uv\|^2 &= \|z_P\|^2 + \|z_M\|^2 \leq \|z_P\|_1^2 + \|z_M\|_1^2 \\ &\leq 2\|z_P\|_1^2 = 2\left(\sum_{i \in P} u_i v_i\right)^2 \\ &\leq 2\left(\sum_{i \in P} \frac{1}{4}(u_i + v_i)^2\right)^2 \\ &\leq 2^{-3}\left(\sum_{i=1}^n (u_i + v_i)^2\right)^2 = 2^{-3}\|u+v\|^4, \end{aligned}$$

Ami bizonyítja az állítást.  $\square$

**4.2.1. Megjegyzés.** *A 4.2.1 lemma egyenlőtlensége éles, például a következő két vektor esetén egyenlőség van:  $u := (r, -r, 0, \dots, 0) \in \mathbb{R}^n$  és  $v := (r, r, 0, \dots, 0) \in \mathbb{R}^n$ .*  $\square$

**4.2.2. Lemma.** *Ha  $(x, y, s) \in \mathcal{N}_2(\gamma)$ , akkor*

$$\min_i x_i s_i \geq (1 - \gamma)\mu.$$

**Biz.:** Minden  $i$  indexre

$$\mu - x_i s_i \leq |x_i s_i - \mu| \leq \|xs - \mu e\| \leq \gamma\mu,$$

azaz  $(1 - \gamma)\mu \leq x_i s_i$ .  $\square$

**4.2.3. Lemma.** *Ha  $(x, y, s) \in \mathcal{N}_2(\gamma)$ , akkor*

$$\|\Delta x \Delta s\| \leq \frac{\gamma^2 + \nu^2}{\sqrt{2^3}(1 - \gamma)} \mu.$$

**Biz.:** Legyen  $D := X^{\frac{1}{2}}S^{-\frac{1}{2}}$ . A Newton-rendszer harmadik egyenletét  $(XS)^{-\frac{1}{2}}$ -del szorozva a következő egyenletet kapjuk:

$$D^{-1}\Delta x + D\Delta s = (XS)^{-\frac{1}{2}}(xs - \sigma\mu e).$$

Ezt az egyenlőséget, és az 4.2.1 lemmát felhasználva

$$\begin{aligned} \|\Delta x \Delta s\| &= \|(D^{-1}\Delta x)(D\Delta s)\| \\ &\leq \sqrt{2^{-3}}\|D^{-1}\Delta x + D\Delta s\|^2 \\ &= \sqrt{2^{-3}}\|(XS)^{-\frac{1}{2}}(xs - \sigma\mu e)\|^2 \\ &= \sqrt{2^{-3}} \sum_i i = 1^n \frac{(x_i s_i - \sigma\mu)^2}{x_i s_i} \leq \sqrt{2^{-3}} \frac{\sum_i i = 1^n (x_i s_i - \sigma\mu)^2}{\min_i x_i s_i} \\ &= \sqrt{2^{-3}} \frac{\|xs - \sigma\mu e\|^2}{\min_i x_i s_i} \end{aligned} \quad (4)$$

A számlálót kifejtve és felhasználva, hogy  $(x, y, s) \in \mathcal{N}_2(\gamma)$  kapjuk:

$$\begin{aligned} \|xs - \sigma\mu e\|^2 &= \|(xs - \mu e) + (1 - \sigma)\mu e\|^2 \\ &= \|xs - \mu e\|^2 + 2(1 - \sigma)\mu e^T(xs - \mu e) + (1 - \sigma)^2\mu^2 e^T e \\ &= \|xs - \mu e\|^2 + 2(1 - \sigma)\mu(e^T(xs) - \mu e^T e) + (1 - \sigma)^2\mu^2 n \\ &\leq \gamma^2\mu^2 + (1 - \sigma)^2\mu^2 n \end{aligned} \quad (5)$$

A (4) tört számlálóját az (5) alapján, a nevezőjét pedig a 4.2.2 lemma alapján becsülve a következőt kapjuk

$$\|\Delta x \Delta s\| \leq \frac{\gamma^2 + n(1 - \sigma)^2}{\sqrt{2^3}(1 - \gamma)}\mu,$$

amibe behelyettesítve a  $\sigma = \left(1 - \frac{\nu}{\sqrt{n}}\right)$  képletet, következik az állítás.  $\square$

**4.2.4. Lemma.** Ha  $(x, y, s) \in \mathcal{N}_2(\gamma)$ , ekkor minden  $\alpha \in [0, 1]$  lépéshosszra

$$\|x^+ s^+ - \mu^+ e\| \leq (1 - \alpha)\|xs - \mu e\| + \alpha^2\|\Delta x \Delta s\|.$$

**Biz.:** A 4.1.1 lemma ii) állítása és egyszerű számítás alapján

$$\begin{aligned} x^+ s^+ - \mu e &= (x - \alpha\Delta x)(s - \alpha\Delta s) - (1 - \alpha(1 - \sigma))\mu e \\ &= xs - \alpha(s\Delta x + x\Delta s) + \alpha^2\Delta x \Delta s - (1 - \alpha(1 - \sigma))\mu e \\ &= xs - \alpha(xs - \sigma\mu e) + \alpha^2\Delta x \Delta s - (1 - \alpha + \alpha\sigma)\mu e \\ &= (1 - \alpha)(xs - \mu e) + \alpha^2\Delta x \Delta s, \end{aligned}$$

amiből következik az állítás.  $\square$

### 4.3. Egy hosszú lépéses algoritmus

Hosszú-lépéses algoritmusokról beszélünk, ha a  $\sigma$  a dimenziótól független konstans. A következő algoritmusban a  $\sigma$ -t minden iterációnál egy  $[\sigma_{\min}, \sigma_{\max}]$  intervallumból választjuk. Az előző algoritmussal ellentétben itt a bővebb  $\mathcal{N}_{-\infty}(\gamma)$  környezetet használjuk, ami a csaknem a teljes  $\mathcal{F}^0$ -at jelenti, ha a  $\gamma$  közel van az 1-hez. Az  $\alpha$  lépéshosszt a lehető legnagyobbak választjuk, ügyelve arra, hogy ne lépjünk ki a  $\mathcal{N}_{-\infty}(\gamma)$  környezetből.

---

Hosszú-lépéses algoritmus

---

**Input**

$\gamma \in (0, 1)$ ,  $\sigma_{\min}, \sigma_{\max}$ :  $0 < \sigma_{\min} \leq \sigma_{\max} < 1$ ,  
 $\epsilon > 0$  pontossági paraméter,  $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$ .

**begin**

$(x, y, s) := (x^0, y^0, s^0)$ .

**while**  $n\mu > \epsilon$  **do**

$\mu := \mu(x, s)$

Válasszunk  $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ .

Határozzuk meg azt a  $(x, y, s)$ -beli  $(\Delta x, \Delta y, \Delta s)$  Newton-irányt, ami a  $\sigma\mu$ -centrumot közelíti.

Legyen  $\alpha \in (0, 1]$  az a legnagyobb érték, amire

$$(x^+, y^+, s^+) \in \mathcal{N}_{-\infty}(\gamma).$$

Legyen  $(x, y, s) = (x^+, y^+, s^+)$ .

**end**

**end**

---

Ahhoz, hogy a lépésszámról mondani tudjunk valamit, találni kellene egy a 4.2.1 következményhez hasonló összefüggést a  $\mu^k$  és  $\mu^0$  között. Ehhez két lemmára van szükségünk.

**4.3.1. Lemma.** *Ha  $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ , akkor:*

$$\|\Delta x \Delta s\| \leq 2^{-\frac{2}{3}} \left(1 + \frac{1}{\gamma}\right) n\mu(x, s).$$

**Biz.:** A 4.2.3 lemmánál látottak alapján, és felhasználva, hogy  $x_i s_i \geq \gamma\mu$  és  $x^T s = n\mu$

$$\begin{aligned} \|\Delta x \Delta s\| &\leq \sqrt{2^{-3}} \|(XS)^{-\frac{1}{2}}(xs - \sigma\mu e)\|^2 \\ &= \sqrt{2^{-3}} \|(XS)^{-\frac{1}{2}}e - \sigma\mu(XS)^{-\frac{1}{2}}e\| \\ &= \sqrt{2^{-3}}(x^T s - 2\sigma\mu n + \sigma^2\mu^2 \frac{n}{\gamma\mu}) \\ &= \sqrt{2^{-3}}n\mu(1 - 2\sigma + \frac{\sigma^2}{\gamma}) \\ &\leq \sqrt{2^{-3}}n\mu(1 + \frac{1}{\gamma},) \end{aligned}$$

ami bizonyítja az állítást. □

A következő lemma egy  $\bar{\alpha}$  alsó korlátot ad a lépéshosszakra.

**4.3.2. Lemma.** *Ha  $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ , akkor  $(x^+, y^+, s^+) \in \mathcal{N}_{-\infty}(\gamma)$  minden  $\alpha \in (0, \bar{\alpha}]$  esetén, ahol*

$$\bar{\alpha} = \sqrt{8}\gamma\sigma \frac{1-\gamma}{n(1+\gamma)}.$$

**Biz.:** Az előző lemma alapján minden  $i$ -re

$$|\Delta x_i \Delta s_i| \leq \|\Delta x \Delta s\| \leq \sqrt{2^{-3}} n \mu \left(1 + \frac{1}{\gamma}\right).$$

Felhasználva az előző egyenlőtlenséget, a Newton rendszer harmadik egyenletét, és azt, hogy  $x_i s_i \geq \gamma \mu$

$$\begin{aligned} x^+ s^+ &= xs - \alpha(s\Delta x + x\Delta s) + \alpha^2 \Delta x \Delta s \\ &\geq xs(1-\alpha) + \alpha\sigma\mu e - \alpha^2 |\Delta x \Delta s| \\ &\geq \gamma\mu e(1-\alpha) + \alpha\sigma\mu e - \alpha^2 \sqrt{2^{-3}} n \mu \left(1 + \frac{1}{\gamma}\right) e \end{aligned} \quad (6)$$

Az állítás belátásához azt kell igazolnunk, hogy  $x^+ s^+ \geq \gamma \mu^+ e$ . Felhasználva, hogy  $\mu^+ = (1 - \alpha(1 - \sigma))\mu$  és a (6) becslést könnyen ellenőrizhető az állítás.  $\square$

**4.3.1. Tétel.** *Létezik  $\nu > 0$  az  $n$ -től független konstans, amire*

$$\mu^{k+1} \leq \left(1 - \frac{\nu}{n}\right) \mu^k \quad k = 1, 2, 3, \dots$$

**Biz.:** A 4.1.1 lemma alapján  $\mu^{k+1} = (1 - \alpha(1 - \sigma))\mu^k$ . Kihasználva, hogy  $\alpha \geq \bar{\alpha}$  minden iterációban:

$$\mu^{k+1} \leq \left(1 - \frac{\sqrt{8}}{n} \gamma \frac{1-\gamma}{1+\gamma} \sigma(1-\sigma)\right) \mu^k. \quad (7)$$

A  $\sigma \rightarrow \sigma(1 - \sigma)$  függvény szigorúan konkáv, így

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

Ha a

$$\nu = 2^{\frac{2}{3}} \gamma \frac{1-\gamma}{1+\gamma} \min\{\sigma_{\min}(1 - \sigma_{\min}), \sigma_{\max}(1 - \sigma_{\max})\} > 0$$

értéket behelyettesítjük az (7) egyenlőtlenségbe, akkor megkapjuk a bizonyítandó állítást.  $\square$

**4.3.2. Tétel.** *A hosszú lépéses algoritmus legfeljebb  $K = \lceil \log(n\varepsilon^{-1}\mu^0) \frac{n}{\nu} \rceil$  iteráció után  $\varepsilon$ -optimális megoldást ad.*

**Biz.:** A 4.2.2 tételhez hasonlóan bizonyítható, az előző lemma felhasználásával.  $\square$

## 4.4. Egy prediktor-korrektor algoritmus

---

Prediktor-korrektor algoritmus

---

**Input**

$\epsilon > 0$  pontossági paraméter,  
 $(x^0, y^0, s^0) \in \mathcal{N}_2(\gamma)$ .

**begin**

$(x, y, s) := (x^0, y^0, s^0)$ ,  $\mu := \mu(x, s)$ .

**while**  $n\mu > \epsilon$  **do**

PREDIKTOR lépés

Határozzuk meg azt az  $(x, y, s)$ -beli  $(\Delta x, \Delta y, \Delta s)$  Newton-irányt,  
ami a 0-centrumot közelíti  
Legyen  $\alpha \in (0, 1]$  az a legnagyobb érték, amire

$$(x^+, y^+, s^+) \in \mathcal{N}_2(\bar{\gamma})$$

Legyen  $(x, y, s) = (x^+, y^+, s^+)$  és  $\mu := (1 - \alpha)\mu$ .

KORREKTOR lépés

Határozzuk meg azt az  $(x, y, s)$ -beli  $(\Delta x, \Delta y, \Delta s)$  Newton-irányt,  
ami a  $\mu$ -centrumot közelíti  
Legyen  $(x, y, s) = (x, y, s) - (\Delta x, \Delta y, \Delta s)$

**end**

**end**

---

Az algoritmus működését igazolják a következő lemmák. A prediktor lépést a  $\gamma = \frac{1}{4}$  és  $\bar{\gamma} = \frac{1}{2}$  esetben vizsgáljuk.

**4.4.1. Lemma.** *Tegyük fel, hogy  $(x, y, s) \in \mathcal{N}_2(\frac{1}{4})$ , és legyen  $(\Delta x, \Delta y, \Delta s)$  a prediktor lépésben kiszámított Newton-irány. Ekkor  $(x^+, y^+, s^+) \in \mathcal{N}_2(\frac{1}{2})$  minden  $\alpha \in [0, \bar{\alpha}]$ -ra, ahol*

$$\bar{\alpha} := \min \left\{ \frac{1}{2}, \left( \frac{\mu}{8 \|\Delta x \Delta s\|} \right)^{\frac{1}{2}} \right\}.$$

**Biz.:** A rövid lépéses algoritmusnál látottak alapján

$$\begin{aligned} \|x^+ s^+ - \mu^+ e\| &\leq (1 - \alpha) \|xs - \mu e\| + \alpha^2 \|\Delta x \Delta s\| \\ &\leq (1 - \alpha) \|xs - \mu e\| + \frac{\mu}{8 \|\Delta x \Delta s\|} \|\Delta x \Delta s\| \\ &\leq \frac{1}{4} (1 - \alpha) \mu + \frac{\mu}{8} \\ &\leq \frac{1}{4} (1 - \alpha) \mu + \frac{1}{4} (1 - \alpha) \mu \\ &= \frac{1}{2} \mu^+, \end{aligned}$$

ami épp azt jelenti, hogy  $(x^+, y^+, s^+) \in \mathcal{N}_2(\frac{1}{2})$ . □

**4.4.1. Köv.** A prediktor lépés legalább  $\bar{\alpha}$  nagyságú és  $\mu^+ \leq (1 - \bar{\alpha})\mu$ .  $\square$

A korrektor lépést általánosabban, a  $\gamma := \bar{\gamma}^2$  esetén vizsgáljuk.

**4.4.2. Lemma.** Tegyük fel, hogy  $(x, y, s) \in \mathcal{N}_2(\bar{\gamma})$  valamilyen  $\bar{\gamma} \in (0, \frac{\sqrt{8}-1}{\sqrt{8}}]$ , és legyen  $(\Delta x, \Delta y, \Delta s)$  a korrektor lépésben kiszámított Newton-irány. Ekkor  $(x^+, y^+, s^+) \in \mathcal{N}_2(\bar{\gamma}^2)$  és  $\mu^+ = \mu$ .

**Biz.:** A 4.1.1 lemmát felhasználva azonnal adódik, hogy  $\mu^+ = \mu$ . A 4.2.4 lemmát felhasználva pedig

$$\|x^+ s^+ - \mu^+ e\| \leq \bar{\gamma}^2 \mu = \bar{\gamma}^2 \mu^+,$$

ami épp azt jelenti, hogy  $(x^+, y^+, s^+) \in \mathcal{N}_2(\bar{\gamma}^2)$ .  $\square$

A 4.4.1 következmény, és a az  $\bar{\alpha}$  definíciója miatt könnyen felírhatunk egy a 4.3.1 tétel állításához hasonló formulát:

**4.4.1. Állítás.** Ha  $\nu \leq 1 - \frac{1}{\sqrt{2}}$ , akkor  $\mu^{k+2} \leq (1 - \frac{\nu}{\sqrt{n}})^2 \mu^k$ .  $\square$

Az előző állítás és a 4.2.2 tétel alapján könnyen bizonyíthatjuk a következőt:

**4.4.1. Tétel.** A prediktor-korrektor algoritmus legfeljebb  $K = \lceil \log(n\varepsilon^{-1}\mu^0) \frac{\sqrt{n}}{\nu} \rceil$  iteráció után  $\varepsilon$ -optimális megoldást ad.  $\square$

## Az algoritmusok összehasonlítása

A hosszú-lépéses algoritmusok elméletei lépésszáma elmarad a másik két verziótól, a gyakorlatban azonban a hosszú-lépéses algoritmusok gyakran jobban teljesítenek. Nézzünk meg egy példát az algoritmusok szemléltetésére!

**4.1. Példa.** Legyen  $A = (1, 1, 1)$ ,  $b = 1$  és  $c = (-2, 1, -3)^T$ . A következő feladatpárt kapjuk:

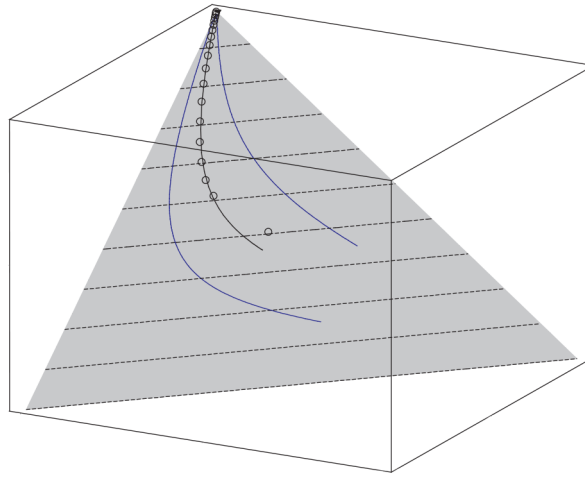
$$(P) \quad \begin{cases} -2x_1 + x_2 - 3x_3 \rightarrow \min \\ x_1 + x_2 + x_3 = 1 \\ x \geq 0 \end{cases}$$

$$(D) \quad \begin{cases} y \rightarrow \max \\ y + s_1 = -2 \\ y + s_2 = 1 \\ y + s_3 = -3 \\ s \geq 0 \end{cases}$$

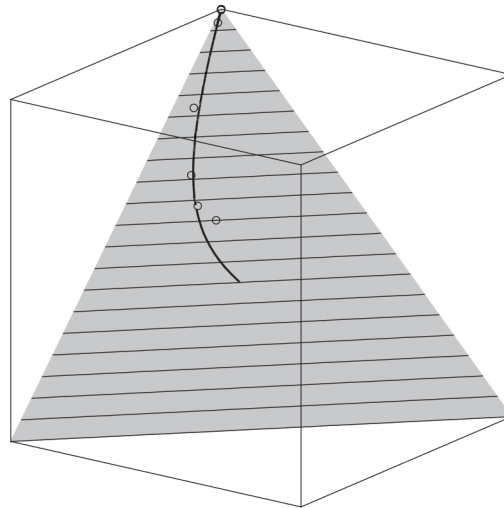
Az  $OPT(\mu)$  rendszer a következő lesz:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ y + s_1 &= -2 \\ y + s_2 &= 1 \\ y + s_3 &= -3 \\ x_i s_i &= \mu \end{aligned}$$

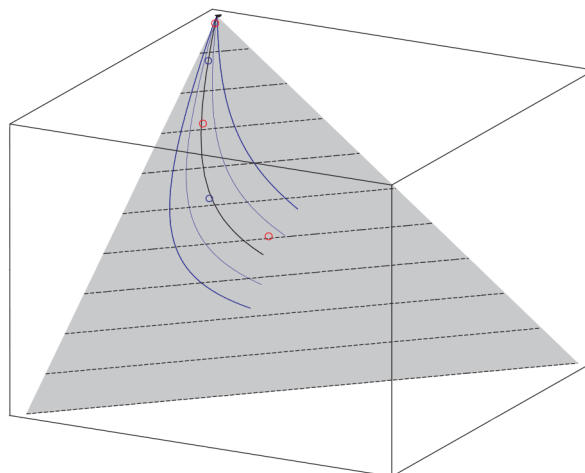
### A rövid lépéses algoritmus működése



### A hosszú lépéses algoritmus működése



### A prediktor-korrektor algoritmus működése





## 4.5. Pontos megoldás előállítása

Az előző fejezetekben tárgyalt algoritmusok segítségével olyan primál-duál megengedett megoldaspárt kaphatunk, ami tetszőlegesen közel van az optimális megoldáspárhoz. Ebben a fejezetben azt tárgyaljuk, hogy hogyan juthatunk el egy pontos megoldásig.

Elevenítsük fel a következő indexhalmazokat:

$$\begin{aligned}\mathcal{P}^* &= \{i : x_i > 0, \text{ valamely } u \in \mathcal{F}_P^* \text{ esetén}\} \\ \mathcal{D}^* &= \{i : s_i > 0, \text{ valamely } (y, s) \in \mathcal{F}_D^* \text{ esetén}\}\end{aligned}$$

Az eddig tárgyaltak alapján  $\mathcal{P}^* \cap \mathcal{D}^* = \emptyset$ , és  $\mathcal{F}_P \neq \emptyset, \mathcal{F}_D \neq \emptyset$  esetén  $\mathcal{P}^* \cup \mathcal{D}^* = \{1, \dots, n\}$ .

Legyen  $(x^k, y^k, s^k)$  valamelyik fenti algoritmus  $k$ . iterációjában kiszámolt pont. Ekkor

$$\begin{aligned}\mathcal{P}^k &= \mathcal{P}(x^k, s^k) := \{j \in \{1, \dots, n\} : x_j^k \geq s_j^k\} \\ \mathcal{D}^k &= \mathcal{D}(x^k, s^k) := \{1, \dots, n\} \setminus \mathcal{P}^k.\end{aligned}$$

Legyen

$$\begin{aligned}\xi_P &= \min_{j \in \mathcal{P}^*} \{\max\{x_j : x \in \mathcal{F}_P^*\}\}, \\ \xi_D &= \min_{j \in \mathcal{D}^*} \{\max\{s_j : (y, s) \in \mathcal{F}_D^*\}\},\end{aligned}$$

és legyen

$$\xi = \min(1, \xi_P, \xi_D).$$

A  $\xi$ -t az  $LP$  feladat kondíciós számának nevezzük.

Az  $LP$  feladat tárigényének felső korlátja az

$$L(A, b, c) = \sum_{i=1}^m \sum_{j=1}^n ((\log_2(|a_{ij}|+1)+1)) + \sum_{i=1}^m ((\log_2(|b_i|+1)+1)) + \sum_{j=1}^n ((\log_2(|c_j|+1)+1))$$

érték.

**4.5.1. Állítás.** *Ha a  $LP$  feladat adatai racionálisak, akkor  $\xi \geq 2^{-L}$ .*

**Biz.:**

**4.5.1. Tétel.** *Származzon az  $\{x^k, s^k\}$  pontsorozat egy olyan belsőpontos algoritmusból, ami az  $x^0 = e, s^0 = e$  belső pontból indul, az iterációs száma  $O(\sqrt{n} \log(\frac{n}{\varepsilon} \mu^0))$  és kielégíti a következő feltételeket*

$$(x^{k+1})^T s^{k+1} \leq (x^k)^T s^k \quad \text{és} \quad \frac{\min_i x_i^k s_i^k}{(x^k)^T s^k} \geq \Omega\left(\frac{1}{n}\right). \quad (8)$$

*Ekkor:*

$$i) 0 < e^T x^k + e^T s^k \leq 2n, \quad \forall k.$$

ii)

$$x_j^k \geq \Omega\left(\frac{\xi}{n}\right), \quad s_j^k \leq \frac{(x^k)^T s^k}{\xi} \quad \forall k \text{ és } \forall j \in \mathcal{P}^*$$

és

$$s_j^k \geq \Omega\left(\frac{\xi}{n}\right), \quad x_j^k \leq \frac{(x^k)^T s^k}{\xi} \quad \forall k \text{ és } \forall j \in \mathcal{D}^*.$$

iii) ha  $(x^k)^T s^k \leq O\left(\frac{\xi^2}{n}\right)$ , akkor

$$s_j^k < \Omega\left(\frac{\xi}{n}\right) \forall j \in \mathcal{P}^* \quad \text{és} \quad x_j^k < \Omega\left(\frac{\xi}{n}\right) \forall j \in \mathcal{D}^*,$$

így

$$\mathcal{P}^k = \mathcal{P}^* \quad \text{és} \quad \mathcal{D}^k = \mathcal{D}^*.$$

**Biz.: i)**

$$(x^k - x^0) \in \mathcal{N}(A), \quad (s^k - s^0) \in \mathcal{R}(A^T),$$

így

$$0 = (x^k - x^0)^T (s^k - s^0) = (x^k - e)^T (s^k - e),$$

amiből

$$e^T x^k + e^T s^k = n + (x^k)^T s^k \leq 2n.$$

ii) Tetszőleges  $(x^*, s^*)$  optimális megoldásra  $(x^k - x^*)^T (s^k - s^*) = 0$ , amiből

$$\sum_{j \in \mathcal{D}^*} x_j^k s_j^* + \sum_{j \in \mathcal{P}^*} s_j^k x_j^* = (x^k)^T s^k,$$

tovább alakítva

$$\sum_{j \in \mathcal{D}^*} \frac{s_j^*}{s_j^k} \frac{x_j^k s_j^k}{(x^k)^T s^k} + \sum_{j \in \mathcal{P}^*} \frac{x_j^*}{x_j^k} \frac{s_j^k x_j^k}{(x^k)^T s^k} = 1.$$

Így minden  $j \in \mathcal{P}^*$  és  $(x^*, s^*)$  optimális megoldás esetén

$$s_j^k x_j^* \leq (x^k)^T s^k \quad \text{és} \quad \frac{x_j^*}{x_j^k} \frac{s_j^k x_j^k}{(x^k)^T s^k} \leq 1.$$

Adott  $j \in \mathcal{P}^*$  esetén legyen  $x^* = \arg \max\{x_j : x \in \mathcal{F}_P^*\}$ . Használva a fenti eredményt, a  $\xi$  definícióját és a (8) feltételeket, a következőt kapjuk

$$s_j^k \xi \leq s_j^k x_j^* < (x^k)^T s^k \quad \text{és} \quad \frac{\xi}{x_j^k} \leq \frac{x_j^*}{x_j^k} < \frac{(x^k)^T s^k}{x_j^k s_j^k} \leq O(n),$$

amiből következik ii) minden  $j \in \mathcal{P}^*$ . Az állítás hasonlóan bizonyítható minden  $j \in \mathcal{D}^*$

iii) Egyenesen következik a második állításból. □

**4.5.1. Köv.** A  $(\mathcal{P}^k, \mathcal{D}^k)$  partíció  $O(\sqrt{n} \log(\frac{n}{\xi}))$  iteráció után az optimális partíciót adja. Ha az LP feladat adatai racionálisak, akkor ez  $O(\sqrt{n}L)$  iterációt jelent. □

**4.5.2. Köv.** Ha  $(x^k)^T s^k = \mu^k \rightarrow 0$ , akkor az  $x$  és  $s$  "kicsi" változói egyenletesen tartanak a 0-hoz, a "nagy" változói pedig egy egyenletes alsó korláttal el vannak szeparálva a 0-tól. Az  $(x^k, s^k)$  pontsorozat tehát egy szigorúan komplementáris megoldáshoz tart.  $\square$

**4.5.1. Lemma.** A bemutatott belsőpontos algoritmusok által generált pontsorozatra teljesülnek a (8) feltételek.

**Biz.:** Az első feltétel teljesülése nyilvánvaló a 4.2.1 és 4.3.1 tételekből. A rövid lépéses algoritmusnál minden előállított  $(x^k, y^k, s^k)$  pont benne van az  $\mathcal{N}_2(\gamma)$  környezetben valamilyen  $\gamma$  konstansra, ezért érvényes lesz a következő minden  $i$ -re:

$$\frac{x^T s}{n} - x_i s_i \leq \left| x_i s_i - \frac{x^T s}{n} \right| \leq \left\| x s - \frac{x^T s}{n} e \right\| \leq \gamma \frac{x^T s}{n}.$$

A hosszú lépéses algoritmusnál minden előállított pont az  $\mathcal{N}_{-\infty}(\gamma)$  környezetben lesz, így ebben az esetben minden  $i$ -re:

$$\frac{x^T s}{n} - x_i s_i \leq \left| x_i s_i - \frac{x^T s}{n} \right| \leq \gamma \frac{x^T s}{n}.$$

Mindkét esetben következik az

$$\frac{\min_i x_i^k s_i^k}{(x^k)^T s^k} \geq \frac{1 - \gamma}{n}.$$

$\square$

## Optimális megoldás előállítása merőleges vetítéssel

Az optimális partíció segítségével a következő módon tudjuk felírni az optimális lapokat:

$$\begin{aligned} \mathcal{F}_P^* &= \{x : Ax = b, x \geq 0, x_j = 0 \quad \forall j \in \mathcal{D}^*\} \\ \mathcal{F}_D^* &= \{(y, s) : A^T y + s = c, s \geq 0, s_j = 0 \quad \forall j \in \mathcal{P}^*\} \end{aligned}$$

Ha a feltételek közül elhagyjuk a nemnegativitást, akkor affin altereket kapunk. A pontos megoldást úgy akarjuk előállítani, hogy a közelítő megoldást merőlegesen vetítjük ezekben az affin alterekbe. Ha elég közel voltunk az optimális halmazhoz, akkor a vetület optimális lesz.

A algoritmusunk a következő módon fog működni:

1. futtatjuk valamelyik útkövető algoritmusunkat néhány iteráció erejéig, így kapjuk az  $(x^k, y^k, s^k) \in \mathcal{F}^0$  pontot
2. kiszámítjuk a  $\mathcal{P}^k$  és  $\mathcal{D}^k$  indexhalmazokat, mint becslést az optimális partícióra

3. megoldjuk az

$$\begin{aligned} \|x^* - x^k\| &\rightarrow \min \\ Ax^* &= b \\ x_i^* &= 0 \quad i \in \mathcal{D}^k \end{aligned} \tag{9}$$

és

$$\begin{aligned} \|y^* - y^k\| &\rightarrow \min \\ A^T y^* + s^* &= c \\ s_i^* &= 0 \quad i \in \mathcal{P}^k \end{aligned} \tag{10}$$

rendszeret

4. ha a megoldásokra  $x_i^* > 0 \quad i \in \mathcal{P}^k$  és  $s_i^* > 0 \quad i \in \mathcal{D}^k$ , akkor  $(x^*, y^*, s^*)$  optimális megoldás
5. ha a pozitivitási feltételek nem teljesülnek, akkor visszatérünk az útkövető algoritmushoz, végzünk még egy iterációt

---

Optimális megoldás vetítéssel

---

**Input**  $(x^k, y^k, s^k) \in \mathcal{F}^0$

**begin**

meghatározzuk a  $\mathcal{P}^k$  és  $\mathcal{D}^k$  halmazokat

megoldjuk a (9) és (10) rendszereket

**if**  $(x_i^* > 0 \text{ ha } i \in \mathcal{P}^k \text{ és } s_i^* > 0 \text{ ha } i \in \mathcal{D}^k)$  **than**

az  $(x^*, y^*, s^*)$  optimális megoldás

**else**

az útkövető algoritmus folytatása

**end**

---

## 5. Duális módszerek

A következő fejezetek célja, hogy megismerjük a volume algoritmust, ami egy önmagában is érdekes, primál-duál közelítő megoldáspárt adó módszer. Kedvező számításgényű, és sok esetben hatékonyan alkalmazható az ismertetésre kerülő bázis identifikációs eljárással.

Tekintsük a következő lineáris programot

$$\left. \begin{array}{l} \min c^T x \\ Ax = b \\ Dx = e \\ x \geq 0, \end{array} \right\}$$

ahol  $c \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $D \in \mathbb{R}^{d \times n}$ ,  $e \in \mathbb{R}^d$  és  $\text{rang}(D) = d$ . Úgy tekintjük, hogy az  $Ax = b$  feltételek nehezen kezelhetőek, a  $Dx = e$  feltételek könnyen. Ekkor lehetséges megközelítés, hogy a Lagrange-relaxációból származó duális feladatot oldjuk meg. Két kérdés megválaszolása áll előttünk:

1. hogyan oldjuk meg a nemdifferenciálható duális feladatot
2. hogyan állítunk elő egy primál megoldást

A feladathoz tartozó Lagrange-függvény

$$L(x, \pi) = c^T x + \pi^T (Ax - b).$$

A duális függvény ekkor

$$q(\pi) = \min_{x \in M} L(x, \pi),$$

ahol  $M = \{x \in \mathbb{R}^n \mid Dx = e, x \geq 0\}$ .

A duális feladat a következő

$$\max_{\pi \in \mathbb{R}^m} q(\pi).$$

Mielőtt folytatnánk a feladat megoldását, megnézzük a duális módszereket általánosabb szituációban.

### 5.1. A szubgradiensek

**5.1. Def.** Legyen  $q : \mathbb{R}^m \rightarrow \mathbb{R}$  konkáv függvény. Azt monjuk, hogy a  $v \in \mathbb{R}^m$  szubgradiense  $q$ -nak a  $\pi$ -ben, ha

$$q(\pi') \leq q(\pi) + (\pi' - \pi)^T v, \quad \forall \pi' \in \mathbb{R}^m.$$

A  $\pi$ -beli szubgradiensek halmazát  $\partial q(\pi)$ -vel jelöljük.

**5.2. Def.** Legyen  $q : \mathbb{R}^m \rightarrow \mathbb{R}$  konkáv függvény. Azt monjuk, hogy a  $w \in \mathbb{R}^m$   $\varepsilon$ -szubgradiense  $q$ -nak a  $p$ -ben, ha

$$q(\pi') \leq q(p) + (\pi' - p)^T w, \quad \forall \pi' \in \mathbb{R}^m.$$

A  $p$ -beli  $\varepsilon$ -szubgradiensek halmazát  $\partial_\varepsilon q(p)$ -vel jelöljük.

**5.1.1. Megjegyzés.** A  $\partial_\varepsilon q(p)$  hamaz szép folytonossági tulajdonságokkal rendelkezik. Ha  $(\varepsilon_t, p_t, w_t \in \partial_{\varepsilon_t} q(p_t)) \rightarrow (\varepsilon^*, p^*, w^*)$ , akkor  $w^* \in \partial_{\varepsilon^*} q(p^*)$ .

A következő feladatból indulunk ki

$$\left. \begin{array}{l} \min f(x) \\ h(x) = 0 \\ x \in M. \end{array} \right\}$$

A hozzá tartozó Lagrange-függvény

$$L(x, \pi) = f(x) + \pi^T h(x),$$

és duális függvény

$$q(\pi) = \inf_{x \in M} L(x, \pi).$$

Legyen

$$x_\pi = \arg \min_{x \in M} L(x, \pi).$$

**5.1.1. Állítás.** A fenti jelölésekkel a  $h(x_\pi)$  a szubgradiense a duális függvénynek a  $\pi$  helyen.

**Biz.:** Minden  $\pi' \in \mathbb{R}^m$ -re egyszerű számolással adódik:

$$\begin{aligned} q(\pi') &= \inf_{x \in M} L(x, \pi') \leq L(x_\pi, \pi') = \\ &= f(x_\pi) + (\pi')^T h(x_\pi) = \\ &= f(x_\pi) + \pi^T h(x_\pi) + (\pi' - \pi)^T h(x_\pi) = \\ &= q(\pi) + (\pi' - \pi)^T h(x_\pi). \end{aligned}$$

□

Két algoritmust fogunk tárgyalni, ami a szubgradienseket használja. Az első a *szubgradiens módszer*, ami minden iterációban az éppen aktuális szubgradienst használja. A második pedig a *vágósíkos módszer*, ami minden iterációban a korábban kiszámolt összes szubgradienst használja.

## 5.2. A szubgradiens módszer

A szubgradiens módszer nagyon hasonlít a differenciálható függvények esetében használható gradiens-módszerre. Lényeges eltérés, hogy a lépéshossz előre rögzített, nem alkalmaz egyenes menti keresést minden iterációban. A gradiens módszerrel ellentétben a célfüggvény értéke növekedhet egy-egy lépés után. Azonban a módszer előnye, hogy jóval egyszerűbb és szélesebb körben alkalmazható.

A következő típusú feladatra alkalmazzuk a módszert

$$\begin{array}{l} \max q(\pi) \\ \pi \in X \end{array}$$

ahol  $q : \mathbb{R}^m \rightarrow [-\infty, \infty)$  egy felülről féligfolytonos, konkáv függvény,  $X \subseteq \mathbb{R}^m$  konvex és zárt halmaz. Ezek a feltételek a korábbiak alapján fennállnak a Lagrange-duális feladat esetén. és  $q$  lényeges tartományának minden pontjában könnyen ki tudunk számolni egy szubgradienst.

A feladatot a  $q$  függvény  $D$  lényeges tartománya alapján a következő formában is felírhatjuk:

$$\max_{\pi \in X \cap D} q(\pi)$$

### Megjegyzések:

1. A feltételek miatt valójában a vetített szubgradiens algoritmusra lesz szükség.
2. Ha  $q$  valós értékű függvény, és  $X = \emptyset$ , akkor nem kell vetíteni.
3. Valós értékű konkáv függvény automatikusan folytonos, így ha nem engedünk meg kiterjesztett értékű függvényeket, akkor elég a konkávitást feltenni.
4. Euklideszi vetítés alapvetően konvex és zárt halmazokra működik. A  $q$  konkávitása és felülről féligfolytonossága biztosítja, hogy a  $D$  konvex és zárt halmaz legyen. Az  $X$ -re vonatkozó feltételek miatt pedig az  $X \cap D$  is konvex és zárt lesz, azaz lehet bele vetíteni.

A módszer a következő iteráció alapján működik

$$\mu^{k+1} := [\mu^k + \alpha^k h^k]^+,$$

ahol  $[\cdot]^+$  jelöli a merőleges vetítést a konvex és zárt  $M := X \cap D$  halmazba és  $h^k$  a  $q$  függvény  $\pi^k$ -beli szubgradiense.

Mivel a célfüggvény az iterációk során nem monoton növekvő, ezért minden lépésben számontartjuk az eddigi legjobb értéket

$$q_*^k = \max\{q_*^{k-1}, q(\pi^k)\}.$$

Az esetleges célfüggvény csökkenés ellenére az teszi a módszert használhatóvá, hogy elég kicsi lépéshossz esetén közelebb kerülünk az optimális megoldás halmazhoz. Erről szól a következő állítás.

**5.2.1. Állítás.** *Ha  $\pi^k$  nemoptimális, akkor minden  $\pi^*$  optimális megoldásra*

$$\|\pi^{k+1} - \pi^*\| < \|\pi^k - \pi^*\|,$$

*minden olyan  $\alpha^k$  lépéshosszra, amire*

$$0 < \alpha^k < \frac{2(q(\pi^*) - q(\pi^k))}{\|h^k\|^2}. \quad (11)$$

**Biz.:**

$$\|\pi^{k+1} - \pi^*\| = \|[\pi^k + \alpha^k h^k]^+ - \pi^*\| \leq \|\pi^k + \alpha^k h^k - \pi^*\|,$$

ahol az utolsó egyenlőtlenség azért áll fenn, mert  $\pi^* \in M$  és a vetítés az nem növeli a távolságot.

Tovább becüljük a jobb oldalt:

$$\|\pi^k + \alpha^k h^k - \pi^*\|^2 = \|\pi^k - \pi^*\|^2 - 2\alpha^k (\pi^* - \pi^k)^T h^k + (\alpha^k)^2 \|h^k\|^2.$$

Használva a

$$(\pi^* - \pi^k)^T h^k \geq q(\pi^*) - q(\pi^k)$$

szubgradiens egyenlőtlenséget:

$$\begin{aligned} \|\pi^k + \alpha^k h^k - \pi^*\|^2 &\leq \\ &\leq \|\pi^k - \pi^*\|^2 - 2\alpha^k (q(\pi^*) - q(\pi^k)) + (\alpha^k)^2 \|h^k\|^2. \end{aligned} \quad (12)$$

Legyen

$$\gamma^k = \frac{\alpha^k \|h^k\|^2}{q(\pi^*) - q(\pi^k)}.$$

A  $\gamma^k$  segítségével a következő formára hozhatjuk a (12) egyenlőtlenséget:

$$\|\pi^k + \alpha^k h^k - \pi^*\|^2 \leq \|\pi^k - \pi^*\|^2 - \frac{\gamma^k (2 - \gamma^k) (q(\pi^*) - q(\pi^k))^2}{\|h^k\|^2} \quad (13)$$

Ha az  $\alpha^k$  lépéshossz kielégíti a tételben megadott feltételt, akkor  $0 < \gamma^k < 2$  teljesül. Így a (13) egyenlőtlenség alapján

$$\|\pi^k + \alpha^k h^k - \pi^*\| \leq \|\pi^k - \pi^*\|,$$

amivel kész a bizonyítás.  $\square$

### Használatos lépéshosszak

- Konstans lépés.  $\alpha_k = h$ , függetlenül a  $k$ -tól.
- Konstans lépéshossz.  $\alpha_k = h/\|g^k\|$ . Ekkor  $\|\mu^{k+1} - \mu^k\| = h$ .
- Négyzetesen összegezhető, de nem összegezhető. Vagyis a következőket elégítik ki

$$\sum_{k=1}^{\infty} \alpha_k^2 < \infty, \quad \sum_{k=1}^{\infty} \alpha_k = \infty.$$

Tipikus példa az  $\alpha_k = \frac{a}{b+k}$ , ahol  $a > 0$  és  $b \geq 0$ .

- Nem összegezhető, de nullsorozat. Vagyis kielégítik a következőket

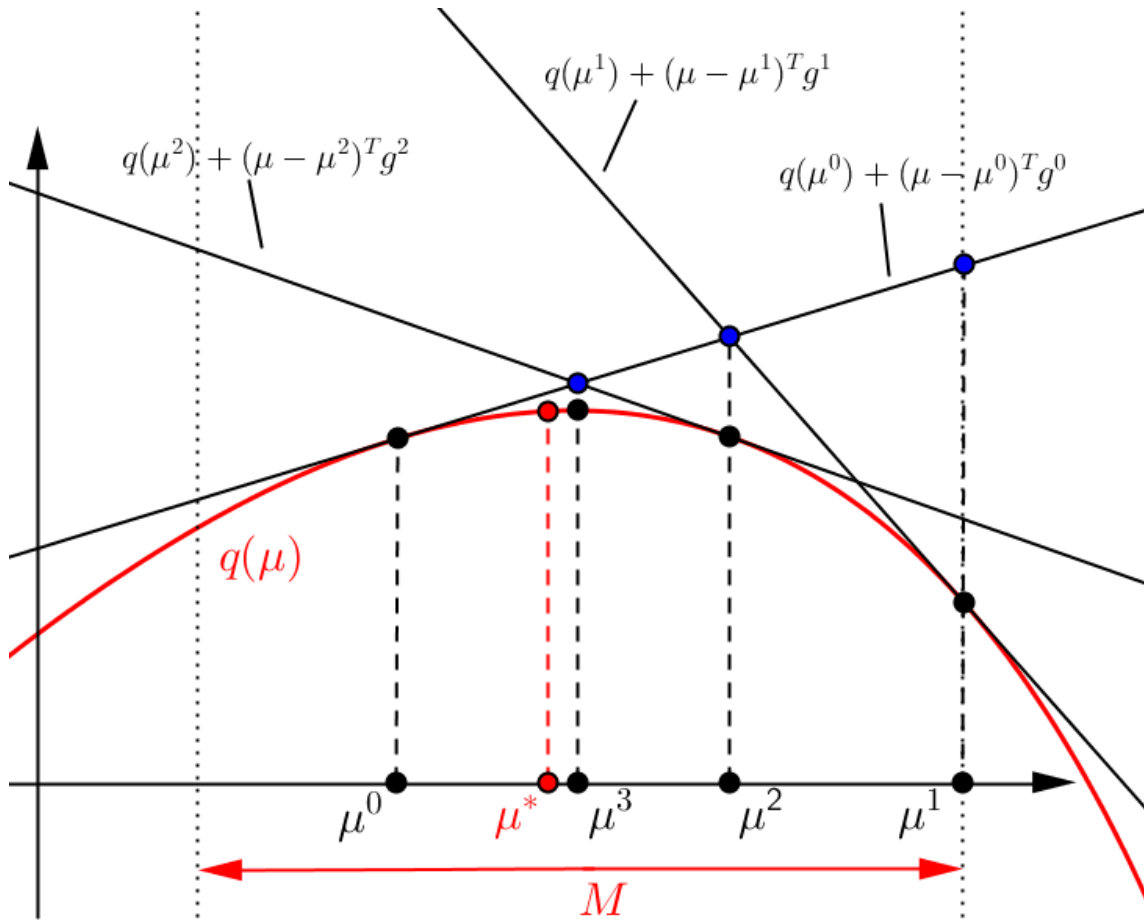
$$\lim_{k \rightarrow \infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty.$$

Tipikus példa az  $\alpha_k = \frac{a}{\sqrt{k}}$ , ahol  $a > 0$ .

A szubgradiens algoritmusok konvergencia viselkedése egyelőre nem jól megértett, annak ellenére, hogy számos konvergencia tétel kapcsolódik hozzájuk. Sok esetben nagyon hatékonyan működik, annak ellenére, hogy nincs elméleti konvergencia; más-kor éppen az ellenkezője történik.



### 5.3. A vágósíkos módszer



Tekintsük újra az előző feladatot

$$\left. \begin{array}{l} \max q(\pi) \\ \pi \in M. \end{array} \right\}$$

A vágósíkos eljárás a  $k$ . iterációban a következő részfeladatot oldja meg

$$\left. \begin{array}{l} \max Q^k(\pi) \\ \pi \in M, \end{array} \right\}$$

ahol  $Q^k$  a  $q$  lineáris approximációja, amit az addig generált  $\pi^k$  pontok és  $h^k$  szubgradiensek segítségével készítünk el a következő módon

$$Q^k(\pi) = \min\{q(\pi^0) + (\pi - \pi^0)^T g^0, \dots, q(\pi^{k-1}) + (\pi - \pi^{k-1})^T g^{k-1}\},$$

és

$$\pi^k = \arg \max_{\pi \in M} Q^k(\pi).$$

A következő állítás az algoritmus konvergenciájáról szól.

**5.3.1. Állítás.** *Tegyük fel, hogy a  $(h^k)$  szubgradiens sorozat korlátos. Ekkor a  $(\pi^k)$  sorozat tetszőleges határpontja optimális megoldás lesz.*

**Biz.:** A  $h^k$  szubgradiense a  $q$ -nak a  $\pi^k$ -ban, azaz

$$q(\pi) \leq q(\pi^i) + (\pi - \pi^i)^T h^i, \quad \forall \pi \in M,$$

így a  $Q^k$  és  $\pi^k$  definíciója alapján

$$Q^k(\pi^k) \geq Q^k(\pi) \geq q(\pi) \quad \forall \pi \in M. \quad (14)$$

Tegyük fel, hogy a  $\{\pi^k\}_K$  részsorozat a  $\bar{\pi}$ -hoz konvergál. Mivel az  $M$  zárt,  $\bar{\pi} \in M$ . Felhasználva (14) és a  $Q^k$  és  $\pi^k$  definícióját minden  $k$  és  $i < k$  esetén

$$q(\pi^i) + (\pi^k - \pi^i)^T h^i \geq Q^k(\pi^k) \geq Q^k(\bar{\pi}) \geq q(\bar{\pi}). \quad (15)$$

Határértéket fogunk venni  $i \rightarrow \infty$ ,  $k \rightarrow \infty$ ,  $i \in K$ ,  $k \in K$  esetén. Kihasználjuk a  $q$  felülről féligfolytonosságát:

$$\limsup_{i \rightarrow \infty, i \in K} q(\pi^i) \leq q(\bar{\pi}), \quad (16)$$

és a  $\{h^k\}$  sorozat korlátosságát:

$$\lim_{\substack{i \rightarrow \infty, \\ i \in K, \\ k \rightarrow \infty, \\ k \in K}} (\pi^k - \pi^i)^T h^i = 0. \quad (17)$$

Felhasználva az (15), (16) és (17) sorokat

$$q(\bar{\pi}) \geq \limsup_{k \rightarrow \infty, k \in K} Q^k(\pi^k) \geq \liminf_{k \rightarrow \infty, k \in K} Q^k(\pi^k) \geq q(\bar{\pi}),$$

azaz

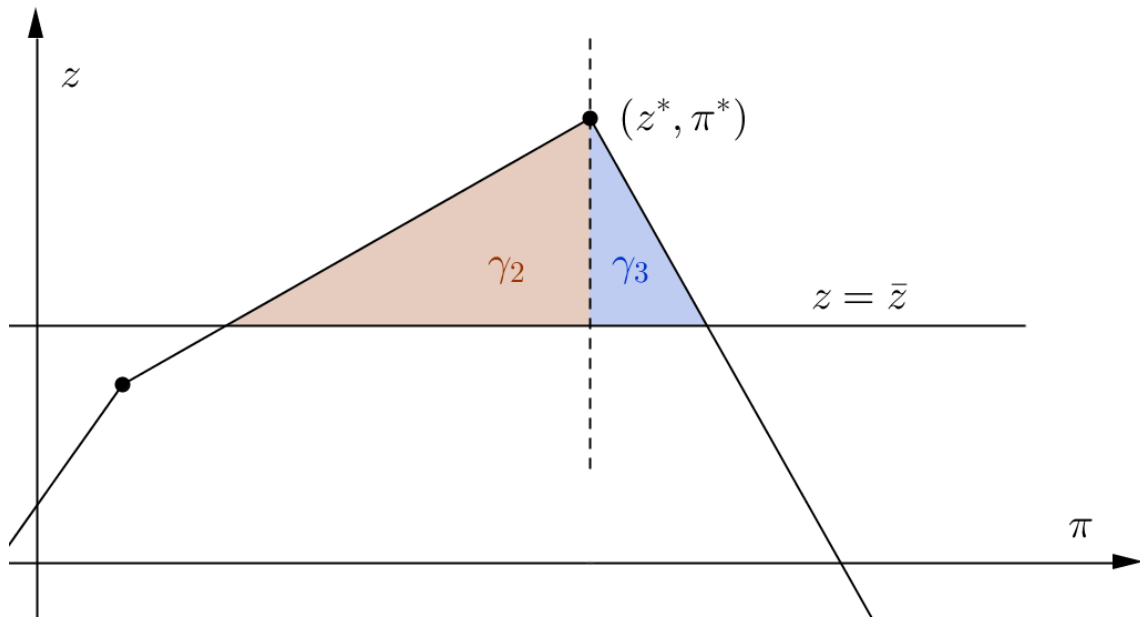
$$\lim_{k \rightarrow \infty, k \in K} Q^k(\pi^k) = q(\bar{\pi}).$$

Az utóbbi eredményt kombinálva az (14) sorral kapjuk

$$q(\bar{\pi}) \geq q(\pi), \quad \forall \pi \in M,$$

ami mutatja, hogy a  $\bar{\pi}$  optimális megoldás. □

## 5.4. A volume algoritmus



A szubgradiens algoritmus sok esetben jó közelítést ad a duál feladat megoldására, de primál megoldást nem állít elő. A volume algoritmus kiterjeszti a szubgradiens módszert, hogy orvosolja ezt a hiányosságot.

Az algoritmus működése a következő tételen alapszik:

**5.4.1. Tétel.** *Tekintsük a következő lineáris programot*

$$\left. \begin{array}{l} \max z \\ z + a_i^T \pi \leq b_i \quad i = 1, \dots, m \end{array} \right\}$$

Ahol  $z \in \mathbb{R}$ ,  $\pi \in \mathbb{R}^{n-1}$ .

A feladat duálisa a következő

$$\left. \begin{array}{l} \min \lambda^T y \\ \sum_{i=1}^n \lambda_i = 1 \\ \sum_{i=1}^n \lambda_i a_i = 0 \\ \lambda \geq 0. \end{array} \right\}$$

Legyen  $(z^*, \pi^*)$  a primál feladat optimális megoldása, és tegyük fel, hogy a  $(z^*, \pi^*)$  pontban az  $1, 2, \dots, m'$  egyenlőtlenségek aktívak. Legyen  $\bar{z} < z^*$ , és tegyük fel, hogy a következő poliéder korlátos

$$\left. \begin{array}{l} z + a_i^T \pi \leq b_i \quad i = 1, \dots, m' \\ z \geq \bar{z}. \end{array} \right\} P$$

Legyen  $\gamma_i$  az a térfogat, amit a  $z + a_i^T \pi \leq b_i$  lap és a  $z = \bar{z}$  hipersík határoz meg (ilyen térfogatok láthatóak az ábrán). Ekkor az optimális duális megoldás a következő módon kapható meg:

$$\lambda_i = \frac{\gamma_i}{\sum_{j=1}^{m'} \gamma_j}.$$

**Biz.:** A tétel kimondása során definiált  $P$  poliéder egy teljes dimenziós politóp. Legyen  $F_0$   $P$ -nek a  $z \geq \bar{z}$  egyenlőtlenség által definiált lapja,  $F_i$  pedig a  $z + a_i^T \pi \leq b_i$  által definiált. Gauss-Osztrogradszkij tétele szerint, ha  $P$  egy korlátos és zárt tartomány, aminek a határa egy szakaszonként folytonosan differenciálható irányítható  $S$  felület, akkor tetszőleges  $v$  vektorra

$$\int_S v^T n \, dS = 0, \quad (18)$$

ahol az  $n$  az  $S$  felület egységnormális mezője. Ha alkalmazzuk a tételt a  $v = e_j$   $j = 1, 2, \dots, n$  esetben, akkor a következőt kapjuk

$$\int_S n \, dS = \mathbf{0},$$

ahol a jobb oldalon a csupa 0 vektor áll. A  $P$  politóp egységnormális mezőjét ismerjük, ez alapján átírva az utóbbi egyenletet

$$\sum_{i=1}^{m'} \delta_i \frac{(1, a_i)}{\|(1, a_i)\|} - \delta_0 (1, 0, \dots, 0) = \mathbf{0},$$

ahol a  $\delta_i$  jelöli az  $F_i$  lap területét. Átrendezve

$$(1, 0, \dots, 0) = \sum_{i=1}^{m'} \frac{\delta_i}{\delta_0 \|(1, a_i)\|} (1, a_i).$$

Ebből kiolvasható a duális feladat megoldása:

$$\lambda_i = \begin{cases} \frac{\delta_i}{\delta_0 \|(1, a_i)\|}, & i = 1, \dots, m' \\ 0, & \text{különben.} \end{cases}$$

A következőkben belátjuk, hogy  $\gamma_i = C \frac{\delta_i}{\|(1, a_i)\|}$ , ahol  $C$  egy konstans.

Ha  $\delta_i = 0$ , akkor  $\gamma_i = 0$ , így feltehető, hogy  $\delta_i > 0$ . A Gauss-Osztrogradszkij tételt, vagyis a (18) formulát szeretnénk újra alkalmazni. Legyen  $Q_i$  a konvex burka az  $F_i$  lapnak és a  $(\bar{z}, \pi^*)$  pontnak. Ekkor  $\text{vol}(Q_i) = \gamma_i$ .

Legyen  $\bar{F}_0 = F_0 \cap Q_i$ . A  $Q_i$  politóp  $\bar{F}_0$ -tól és  $F_i$ -től különböző lapjait olyan egyenlőtlenségek határozzák meg, ahol hiányzik a  $z$ , vagyis ezeknek a lapoknak a normálisa merőleges a  $z$  tengelyre.

Alkalmazzuk a 18 formulát a  $Q_i$  politópra. Ha  $v = e_i$ , akkor a  $v^T n$  szorzat csak az  $F_i$  és az  $\bar{F}_0$  lapok esetén ad 0-tól különböző értéket. Ezt kihasználva

$$\delta_i \frac{(1, a_i)}{\|(1, a_i)\|} + A_i (1, 0, \dots, 0) = \mathbf{0},$$

ahol  $A_i$  az  $\bar{F}_i$  lap területe. Ebből kapjuk:

$$A_i = \frac{\delta_i}{\|(1, a_i)\|}.$$

Legyen  $h = z^* - \bar{z}$ , ekkor a

$$\gamma_i = \text{vol}(Q_i) = \frac{A_i h}{2} = \frac{1}{2} h \frac{\delta_i}{\|(1, a_i)\|}.$$

Mivel  $\delta_0 = \sum A_j$ , felírhatjuk a következőt:

$$\lambda_i = \frac{\delta_i}{\delta_0 \|(1, a_i)\|} = \frac{A_i}{\sum A_j} = \frac{\gamma_i}{\sum \gamma_j},$$

amivel készen vagyunk. □

Térjünk vissza a megoldandó feladathoz:

$$\left. \begin{array}{l} \min c^T x \\ Ax = b \\ Dx = e \\ x \geq 0, \end{array} \right\} \iff \left. \begin{array}{l} \min c^T x \\ Ax = b \\ x \in M \end{array} \right\}$$

A Lagrange-relaxáltját szeretnénk megoldani a volume-módszerrel. Minden olyan módszer, ami nemdifferenciálható függvényt optimalizál, felhasznál egy orákulumot. Az orákulum egy adott  $\pi$ -re megadja a  $q(\pi)$  értéket, és ad egy  $\pi$ -beli szubgradienst.

A mi esetünkben az orákulum feladata kiszámolni a

$$q(\pi) = \min_{x \in M} L(x, \pi) = \min_{x \in M} (c^T x + \pi^T (Ax - b)) \quad (19)$$

értéket, és az

$$x_\pi = \arg \min_{x \in M} L(x, \pi)$$

pontot.

Korábban láttuk, hogy ekkor a  $v = h(x_\pi) = Ax_\pi - b$  vektor szubgradiense lesz a  $q$ -nak a  $\pi$  pontban, azaz a szubgradiens kiszámítása már nem időigényes. Az 19 részfeladat megoldásának nehézsége az  $M$  halmazt definiáló feltételeken múlik. Minél könnyebb megoldani ezt a részfeladatot, annál jobban működik a duális megközelítés.

A volume-algoritmus a következő dolgokat generálja:

- $\pi_t$  alappontok
- $\hat{\pi}_k$  "stabil centrumok", ami a  $\pi_t$  alappontok egy olyan részsorozata, ami elég jó előrehaladást biztosít az optimalizálás során
- a következő  $\pi_t$  alappontot az aktuális stabil centrumból, és egy bizonyos  $w_t$  irányú és  $s_t$  hosszúságú lépés alapján kapjuk

- $z_t$  a primál megoldás approximációja, a korábbi primál pontok konvex kombinációjaként kapjuk, ugyanazokkal az együtthatókkal, amelyekkel a  $w_t$ -t számítjuk a korábbi szubgradiensekből

---

### Volume-algoritmus

---

**Input**

$\pi_0 \in \mathbb{R}^n$  kezdő duális alappont.

**Inicializálás**

Megoldjuk a 19 részfeladatot,  $x_0 \in \arg \min L(x, \pi_0)$ .

$v_0 := Ax_0 - b$ ,  $z_1 := x_0$ ,  $w_1 := v_0$  és  $\hat{\pi}_1 := \pi_0$ .

$k = t = 1$  és  $T := \emptyset$ .

**begin**

**loop**

Teszünk egy duális lépést:  $\pi_t = \hat{\pi}_k + s_t w_t$ .

Megoldjuk a 19 részfeladatot,  $x_t \in \arg \min L(x, \pi_t)$ .

$v_t := Ax_t - b$ .

**if** ( $q(\pi_t) > q(\hat{\pi}_k)$  és  $\langle w_t, v_t \rangle$ )

$\hat{\pi}_{k+1} := \pi_t$ ,  $t_k := t$ ,  $T := T \cup \{t_k\}$ ,  $k := k + 1$ .

Kiszámítjuk az  $s_{t+1}$  lépéshosszat.

Adott  $0 \leq \alpha_t \leq 1$  paraméterre legyen

$$\begin{aligned} z_{t+1} &= \alpha_t x_t + (1 - \alpha_t) z_t, \\ w_{t+1} &= \alpha_t v_t + (1 - \alpha_t) w_t, \end{aligned}$$

$t := t + 1$ .

**end**

**end**

---

**5.4.1. Megjegyzés.** A [8] cikkben az  $s_t = \mu \frac{UB - q(\pi_t)}{\|w_t\|^2}$  lépéshosszt javasolják, ahol  $UB$  egy felső korlát az optimumra és  $\mu \in (0, 2)$  egy paraméter, de más lehetőségek is vannak.

### Néhány összefüggés

**5.4.1. Lemma.** Legyen  $\{z_t\}$ ,  $\{v_t\}$  és  $\{w_t\}$  a volume-algoritmus által generált sorozat. Ekkor

$$v_t \in \partial q(\pi_t) \text{ és } w_t = Az_t - b.$$

□

**5.4.2. Lemma.** Minden  $t \geq 1$  esetén vezettük be a következő együtthatókat

$$\mu_{t,j} := \alpha_{t-j} \prod_{i=t-j+1}^t (1 - \alpha_i) \quad j = 0, 1, \dots, t,$$

ahol  $\alpha_0 := 1$  és  $i_f < i_0$  esetén  $\prod_{i=i_0}^{i_f} (1 - \alpha_i) := 1$ . Ekkor:

i) Minden  $j \leq t$  esetén  $\mu_{t,j} \geq 0$  és  $\sum_{j=0}^t \mu_{t,j} = 1$ .

ii)  $z_{t+1} = \sum_{j=0}^t \mu_{t,j} x_{t-j}$  és  $w_{t+1} = \sum_{j=0}^t \mu_{t,j} v_{t-j}$ .

**Biz.: i)** Az  $\alpha_t \in [0, 1]$  miatt a  $\mu_{t,j}$  nemnegativitása nyilvánvaló. Az összegük vizsgálatához

$$\begin{aligned} \sum_{j=0}^t \mu_{t,j} &= \alpha_t 1 + \alpha_{t-1}(1 - \alpha_t) + \alpha_{t-2}(q - \alpha_t)(1 - \alpha_{t-1} + \dots) \\ &= \alpha_t + (1 - \alpha_t)[\alpha_{t-1} + (1 - \alpha_{t-1})(\alpha_{t-2} + (1 - \alpha_{t-2})(\dots(1 - \alpha_2)(\alpha_1 + 1 - \alpha_1)\dots))] \\ &= \alpha_t + (1 - \alpha_t)[\alpha_{t-1} + (1 - \alpha_{t-1})(\alpha_{t-2} + (1 - \alpha_{t-2})1)] \\ &= \alpha_t + (1 - \alpha_t)[\alpha_{t-1} + (1 - \alpha_{t-1})1] \\ &= 1 \end{aligned}$$

ii) Felhasználva az i) állítást és a  $z_{t+1}$ ,  $w_{t+1}$  definícióját indukcióval könnyen következik.  $\square$

A következő tétel arról szól, hogy a  $w_t$  egy  $\varepsilon$ -szubgradiense a  $q$  függvénynek a  $p_t$  pontban, ahol a  $p_t$ -t a következő rekurzív formulával kapjuk

$$\begin{aligned} p_1 &:= \pi_0 \\ p_{t+1} &:= \alpha_t \pi_t + (1 - \alpha_t) p_t, \end{aligned}$$

vagy ekvivalensen

$$p_{t+1} := \sum_{j=0}^t \mu_{t,j} \pi_{t-j}.$$

**5.4.2. Tétel.**  $\varepsilon_1 := 0$  és minden  $t \geq 1$  esetén legyen

$$\varepsilon_t = \alpha_t(1 - \alpha_t)\langle v_t - w_t, p_t - \pi_t \rangle + (1 - \alpha_t)\varepsilon_t.$$

Ekkor

$$\varepsilon_t \geq 0 \text{ és } w_t \in \partial_{\varepsilon_t} q(p_t).$$

**Biz.:**  $t = 1$  esetén

$$w_1 = v_0 = Ax_0 - b \in \partial q(\pi_0) = \partial_{\varepsilon_1} q(\pi_0) = \partial_{\varepsilon_1} q(p_1)$$

Tegyük fel, hogy az állítás igaz a  $t \geq 1$  esetén. Ekkor kihasználva a  $w_t \in \partial_{\varepsilon_t} q(p_t)$  és  $v_t \in \partial q(\pi_t)$  összefüggéseket minden  $\pi' \in \mathbb{R}^m$  esetén

$$q(\pi') \leq q(\pi_t) + (\pi' - \pi_t)^T v_t \tag{20}$$

$$q(\pi') \leq q(p_t) + (\pi' - p_t)^T w_t + \varepsilon_t \tag{21}$$

Véve az  $\alpha_t \cdot (20) + (1 - \alpha_t) \cdot (21)$  konvex kombinációt, és kihasználva a  $q$  konkávitását és a  $p_t$  definícióját, a következőt állíthatjuk minden  $\pi' \in \mathbb{R}^m$  esetén

$$q(\pi') \leq q(p_{t+1}) + \langle w_{t+1}, \pi' \rangle - \langle \alpha_t v_t, \pi_t \rangle - \langle (1 - \alpha_t) w_t, p_t \rangle + (1 - \alpha_t) \varepsilon_t.$$

Az  $\varepsilon_{t+1}$  definícióját kihasználva a következő alakra hozhatjuk

$$q(\pi') \leq q(p_{t+1}) + (\pi' - p_{t+1})^T w_{t+1} + \varepsilon_{t+1},$$

ami éppen azt jelenti, hogy  $w_{t+1} \in \partial_{\varepsilon_{t+1}} q(p_{t+1})$ , azaz az állítás öröklődik.

Kihasználva a  $v_t \in \partial q(\pi_t)$  és  $w_t \in \partial_{\varepsilon_t} q(p_t)$  tényeket, a következő becslést kapjuk

$$\varepsilon_{t+1} \geq -\alpha_t(1 - \alpha_t)\varepsilon_t + (1 - \alpha_t)\varepsilon_t = (1 - \alpha_t)^2\varepsilon_t,$$

azaz az  $\varepsilon_t \geq 0$  is teljesül. □



## 6. Optimális bázismegoldás előállítása

Ebben a fejezetben egy bázis identifikációs eljárást (röviden *BIP*) tekintünk át, Megiddo [11] algoritmusára támaszkodva, ami erősen polinomiális időben állít elő egy optimális bázismegoldást, egy adott primál-duál komplementáris megoldáspárból indulva. A *BIP* olyan *LP* megoldó módszerekkel együtt alkalmazható, ami egyszerre állít elő primál és duál közelítő megoldást. Egy  $(x^k, s^k)$  jó közelítő megoldáspár "kevésbé" sérti csak a komplementaritási feltételeket. A *BIP* a következő módon működik:

1. Perturbáljuk az *LP*-feladatot, megkapva az  $LP^k$  feladatot, majd kerekítjük az  $(x^k, s^k)$  közelítő megoldást, hogy az  $LP^k$ -nak komplementáris megoldáspárja legyen.
2. A perturbált feladatra alkalmazzuk Megiddo algoritmusát:
  - (a) Keresünk egy kezdő bázist.
  - (b) Pivotalással eljutunk egy primál optimális bázishoz.
  - (c) További pivotalással eljutunk egy optimális bázishoz,  $\mathcal{B}_k$ -hoz.
  - (d) Ellenőrizzük, hogy a  $\mathcal{B}_k$  bázis optimális-e az eredeti *LP*-feladatra.

Azt reméljük, hogy ha a  $k$ . iterációban kapott  $(x^k, s^k)$  pont elég közel van az optimális laphoz, akkor az  $LP^k$  probléma  $\mathcal{B}^k$  optimális bázisa egyben az eredeti *LP*-feladat optimális bázisa is.

### 6.1. A perturbált feladat

Jelöljük  $(\mathcal{B}, \mathcal{N})$ -nel a változók felbontását bázis és nem bázis változókra, legyenek  $A_{\mathcal{B}}$  és  $A_{\mathcal{N}}$  az  $A$  mátrix megfelelő részmátrixai. A  $\mathcal{B}$  pontosan akkor optimális bázis, ha  $A_{\mathcal{B}}$  nonsinguláris,

$$x_B = A_{\mathcal{B}}^{-1}b \geq 0,$$

és

$$y = A_{\mathcal{B}}^{-T}c_B, \quad s_N = c_N - A_{\mathcal{N}}^T y \geq 0.$$

Ha az *LP*-feladat megengedett és korlátos, akkor létezik optimális bázismegoldása. Egy bázismegoldás primál degenerált, ha az  $x_B$ -nek van 0 komponense, duál degenerált, ha az  $s_N$ -nek van 0 komponense.

Tegyük fel, hogy egy primál-duál közelítő megoldást adó algoritmus generálja az  $(x^k, y^k, s^k)$  pontsorozatot és a  $(\mathcal{P}^k, \mathcal{D}^k)$  partícióssorozatot.

$$\left. \begin{array}{l} \min(c^k)^T x \\ Ax = b^k, \\ x \geq 0, \end{array} \right\} (LP^k)$$

ahol

$$b^k = P^k x_{\mathcal{P}^k}^k, \quad c_{\mathcal{P}^k}^k = (P^k)^T y^k, \quad c_{\mathcal{D}^k}^k = (D^k)^T y^k + s_{\mathcal{D}^k}^k.$$

**6.1.1. Állítás.** Az  $x = (x_{\mathcal{P}^k}^k, 0)$  és  $(y, s) = (y^k, (0, s_{\mathcal{D}^k}^k))$  egy szigorúan komplementáris megoldása az  $LP^k$  feladatnak.

**Biz.:** Ellenőrzés. □

### Bázis identifikáció belsőpontos algoritmus segítségével

Az  $LP^k$  feladatra a  $\mathcal{P}^k, \mathcal{D}^k$  adja az optimális partíciót, így a 4.5.1 tétel alapján bizonyos számú iteráció után az  $LP^k$  és  $LP$  feladatokra megegyezik az optimális partíció.

Tegyük fel, hogy a  $k$ -ra már  $(\mathcal{P}^k, \mathcal{D}^k) = (\mathcal{P}^*, \mathcal{D}^*)$ . Felhasználva a 4.5.1 tételt kapjuk:

$$\|b^k - b\| = \|D^k x_{\mathcal{D}^k}^k\| = \left\| \sum_{j \in \mathcal{D}^k} s_j^k a_j \right\| \leq \frac{(x^k)^T s^k}{\xi} \sum_{j \in \mathcal{D}^k} \|x_j\| \leq \frac{(x^k)^T s^k}{n} \eta, \quad (22)$$

és

$$\|c^k - c\| = \|s_{\mathcal{P}^k}^k\| \leq \frac{(x^k)^T s^k}{\xi} \|e\| = \frac{(x^k)^T s^k}{\xi} \sqrt{n}, \quad (23)$$

amiből látszik, hogy  $(x^k)^T s^k \rightarrow 0$  esetén  $b^k \rightarrow b$  és  $c^k \rightarrow c$ . Továbbá ha

$$(x^k)^T s^k \leq \min \left\{ O\left(\frac{\xi^2}{n}\right), 2^{-t \frac{\xi}{n\eta}} \right\},$$

akkor

$$(\mathcal{P}^k, \mathcal{D}^k) = (\mathcal{P}^*, \mathcal{D}^*), \quad \|b^k - b\| \leq 2^{-t}, \quad \|c^k - c\| \leq 2^{-t}.$$

**6.1.1. Megjegyzés.** Ha általában tekintünk egy  $LP(A, b, c)$  feladatot, és egy  $LP(A, b^k, c^k)$  feladatrendszerrel, amire  $b^k \rightarrow b$  és  $c^k \rightarrow c$ , akkor nem mondhatjuk el, hogy elég nagy  $k$ -ra meg fog egyezni az  $LP$  és  $LP^k$  optimális partíciója.

**6.1.1. Tétel.** Tegyük fel, hogy az  $(x^k, y^k, s^k)$  pontsorozat egy  $O(\sqrt{n}L)$ -es belsőpontos algoritmusból származik. Legyen  $\mathcal{B}$  az  $LP^k$  feladat optimális bázisa. Ekkor létezik  $0 < \bar{t} < \infty$ , amire  $(x^k)^T s^k < 2^{-\bar{t}}$  esetén  $\mathcal{B}$  az eredeti  $LP$  feladat optimális bázisa is. Ha az  $LP$ -feladat racionális, akkor  $\bar{t} \leq O(L)$ .

**Biz.:** Legyen  $\mathcal{B}$  tetszőleges bázisa az  $A$  mátrixnak, amire  $A_{\mathcal{B}} \bar{x}_B = b$  és  $A_{\mathcal{B}}^T \bar{y} = c_B$ . Legyen  $\tau_B$  az  $\bar{x}_B$  és  $\bar{s} = c - A^T \bar{y}$  legkisebb pozitív abszolút értékű komponense. Legyen  $\tau = \min\{1, \min_B \tau_B\}$ . Ekkor a  $\tau > 0$ , mert az  $A$  mátrixnak véges sok bázisa van. Legyen  $\zeta = \max\{1, \max_B \|A_{\mathcal{B}}^{-1}\|\}$ . Ha az  $LP(A, b, c)$  feladat adatai racionálisak, akkor  $\tau \leq 2^{-L}$  és  $\zeta \leq 2^L$ . Először a következő lemmát bizonyítjuk.

**6.1.1. Lemma.** Legyen  $\mathcal{B}$  optimális bázisa az  $LP(A, b^k, c^k)$  feladatnak, ahol  $\|b^k - b\| < \frac{\tau}{\zeta}$  és  $\|c^k - c\| < \frac{\tau}{\eta\zeta}$ . Ekkor  $\mathcal{B}$  az  $LP(A, b, c)$  feladat optimális bázisa is, azaz a  $\mathcal{B}$ -hez tartozó  $\bar{x}_B$  és  $\bar{s}$  nemnegatív.

**Biz.:** Legyen az  $LP^k$  feladat  $\mathcal{B}$ -hez tartozó primál és duál optimális bázis megoldása  $\bar{x}_B^k$  és  $(\bar{y}^k, \bar{s}^k)$ . Tegyük fel, hogy az  $\bar{x}_B$ -nek van negatív komponense, az  $i$ . komponens. Ekkor a  $\tau$  definíciója miatt  $\bar{x}_i \leq -\tau$ . Tudjuk, hogy

$$\bar{x}_B^k - \bar{x}_B = A_B^{-1}(b^k - b),$$

azaz

$$\|\bar{x}_B^k - \bar{x}_B\| = \|A_B^{-1}(b^k - b)\| \leq \|A_B^{-1}\| \cdot \|b^k - b\|.$$

A legutóbbi egyenlőtlenségből

$$|\bar{x}_i^k - \bar{x}_i| \leq \|A_B^{-1}\| \cdot \|b^k - b\| < \tau,$$

azaz

$$\bar{x}_i^k < \tau + \bar{x}_i \leq \tau - \tau = 0,$$

ami ellentmond annak, hogy  $\bar{x}_B^k \geq 0$ .

Az  $\bar{s}_B$  nemnegativitása hasonlóan jön ki. Tegyük fel, hogy  $\bar{s}_j$  negatív. Ekkor  $\bar{s}_j \leq -\tau$  is teljesül.

$$|\bar{s}_j^k - \bar{s}_j| = |a_j^T(\bar{y}^k - \bar{y})| \leq |a_j^T A_B^{-T}(c_B^k - c_B)| \leq \|a_j^T\| \cdot \|A_B^{-1}\| \cdot \|c^k - c\| < \tau,$$

azaz

$$\bar{s}_j^k < \tau + \bar{s}_j \leq \tau - \tau = 0,$$

ami szintén ellentmondás. □

A tétel belátásához úgy kell megválasztanunk a  $\bar{t}$ -t, hogy

$$(P^k, D^k) = (P^*, D^*)$$

és

$$\|b^k - b\| < \frac{\tau}{\zeta}, \quad \|c^k - c\| < \frac{\tau}{\eta\zeta}.$$

A (22) és (23) sorok alapján ez teljesül, ha

$$2^{-\bar{t}} \leq \min \left\{ O\left(\frac{\xi^2}{n}\right), \frac{\xi\tau}{n\eta\zeta} \right\}.$$

Ha az  $LP(A, b, c)$  feladat adatai racionálisak, akkor  $\bar{t} \leq O(L)$ . □

**6.1.1. Köv.** Származzon az  $\{x^k, s^k\}$  pontsorozat egy olyan belsőpontos algoritmusból, ami az  $x^0 = e$ ,  $s^0 = e$  belső pontból indul, az iterációs száma  $O(\sqrt{n}L)$  és kielégíti a (8) feltételeket. Ekkor legfeljebb  $O(\sqrt{n}L)$  iteráció után az  $LP^k$  feladat optimális bázisa megegyezik az eredeti LP feladat optimális bázisával. □

## Bázis identifikáció

A következő részben megnézzük Megiddo [11] algoritmusát, ami az  $LP(A, b^k, c^k)$  feladat optimális bázisát állítja elő, az ismert komplementáris megoldáspárból.

A korábbi eredményeket felhasználva, a következőket tesszük fel az  $LP^k$  feladatról:

- (1) A  $\mathcal{P}^k, \mathcal{D}^k$  optimális partíció ismert.
- (2) Ismert az  $\tilde{x}$  primál optimális megoldás, amire  $A\tilde{x} = b^k$ ,  $\tilde{x}_{\mathcal{D}^k} = 0$  és  $\tilde{x}_{\mathcal{P}^k} > 0$ .
- (3) Ismert az  $(\tilde{y}, \tilde{s})$  duális optimális megoldás, amire  $A^T\tilde{y} + \tilde{s} = c^k$ ,  $\tilde{s}_{\mathcal{D}^k} > 0$  és  $\tilde{s}_{\mathcal{P}^k} = 0$ .

Az egyszerűség kedvéért a következőkben elhagyjuk a  $k$  indexet.

Legyen a  $(\mathcal{B}, \mathcal{N})$  a változók tetszőleges bázis-nembázis partíciója. Ekkor

$$\tilde{x}_B = A_B^{-1}(b - A_N\tilde{x}_N).$$

Az  $(\tilde{x}_B, \tilde{x}_N)$  megoldást primál szuper-bázismegoldásnak nevezhetjük, hiszen vannak olyan nembázis változói amik nem nullák. Ha  $x_i > 0$  egy  $i \in \mathcal{N}$  indexre, akkor az  $i$  egy primál szuper-bázisváltozó. Hasonlóan ha  $s_i > 0$  egy  $i \in \mathcal{B}$  indexre, akkor az  $i$  egy duális szuper-bázisváltozó. A cél a szuper-bázisváltozók kiküszöbölése. Először azonban szükségünk van egy kezdő bázisra.

## 6.2. Kezdő bázis előállítása

Legyen az  $[A_1, A_2, A_3]$  az  $A$  mátrix következő indexhalmazokhoz tartozó partíciója:

$$I_1 = \{i | \tilde{x}_i > 0, \tilde{s}_i = 0\}$$

$$I_2 = \{i | \tilde{x}_i = 0, \tilde{s}_i = 0\}$$

$$I_3 = \{i | \tilde{x}_i = 0, \tilde{s}_i > 0\}$$

---

Kezdő bázis keresés

---

**Input**

$A = [A_1, A_2, A_3]$  mátrix a fenti partícióval,  $A \in \mathbb{R}^{m \times n}$ ,  $\text{rang}(A) = m$ .

**Output**

$A_B$  kezdő bázis.

**begin**

$A' := [ ]$ ,  $A_B := [ ]$ ,  $i := 1$ .

**while**  $\text{rang}(A_B) < m$  **do**

$A' := A' \cup A_i$ .

**while**  $\text{rang}(A_B) < \text{rang}(A')$  **do**

Keresünk egy  $a_j$   $A'$ -beli oszlopot, ami független a  $A_B$ -től.

$A_B := A_B \cup a_j$ .

**end**

$i := i + 1$ .

**end**

**end**

---

Ha a  $A_{\mathcal{B}}$ -hez az algoritmus sorrendjében vesszük hozzá az oszlopokat, akkor nyilvánvalóan egy olyan kezdő bázist kapunk, amiben minimális a szuper-bázisváltozók száma. Ha a primál és duál szuper-bázisváltozókat  $\tilde{x}_{\mathcal{N}}^+$  és  $\tilde{s}_{\mathcal{B}}^+$  jelöli, akkor a következő formájú kezdő pivottáblát kapjuk:

Kezdő bázis struktúra

$\tilde{x}_{\mathcal{B}}^+$	$\tilde{x}_{\mathcal{N}}^+$	$\tilde{x}_{\mathcal{B}}^0$	$\tilde{x}_{\mathcal{N}}^0$	$\tilde{x}_{\mathcal{B}}^0$	$\tilde{x}_{\mathcal{N}}^0$
$\tilde{s}_{\mathcal{B}}^0$	$\tilde{s}_{\mathcal{N}}^0$	$\tilde{s}_{\mathcal{B}}^0$	$\tilde{s}_{\mathcal{N}}^0$	$\tilde{s}_{\mathcal{B}}^+$	$\tilde{s}_{\mathcal{N}}^+$
$I$	$*$	$0$	$*$	$0$	$*$
$0$	$0$	$I$	$*$	$0$	$*$
$0$	$0$	$0$	$0$	$I$	$*$

Mielőtt folytatnánk a bázis identifikáció folyamatát, fel kell elevenítenünk a bázistáblák egy fontos tulajdonságát.

Az aktuális  $\mathcal{B}$  bázishoz tartozó bázistábla a következő módon néz ki:

	$\mathcal{B}$	$\mathcal{N}$
$\mathcal{B}$	$I$	$A_{\mathcal{B}}^{-1}A_{\mathcal{N}}$

Legyen a tábla  $i \in \mathcal{B}$ ,  $j \in \mathcal{N}$  indexekhez tartozó eleme a  $t_{ij}$ . Definiálunk két típusú vektort: minden  $i \in \mathcal{B}$ -re a  $t^{(i)}$  vektort és minden  $j \in \mathcal{N}$ -re a  $t_j$  vektort.

$$t_k^{(i)} = \begin{cases} t_{ik} & \text{ha } k \in \mathcal{N} \\ 1 & \text{ha } k = i \\ 0 & \text{különben} \end{cases}$$

és

$$(t_j)_k = \begin{cases} t_{kj} & \text{ha } k \in \mathcal{B} \\ -1 & \text{ha } k = j \\ 0 & \text{különben} \end{cases}$$

**6.2.1. Tétel.**  $t_j \perp t^{(i)}$  minden  $i \in \mathcal{B}$ ,  $j \in \mathcal{N}$  esetén.

**Biz.:** Könnyen adódik a vektorok definíciójából:

$$t^{(i)} = \begin{array}{c} \mathcal{B} \\ i \end{array} \begin{array}{|c|c|c|c|c|c|} \hline 0 \dots 0 & 1 & 0 \dots 0 & * \dots * & t_{ij} & * \dots * \\ \hline \end{array}$$

$$t_j = \begin{array}{|c|c|c|c|c|c|} \hline * \dots * & t_{ij} & * \dots * & 0 \dots 0 & -1 & 0 \dots 0 \\ \hline \end{array}$$

□

### 6.3. Primál fázis

Az algoritmus a korábban előállított kezdő bázisból indul. A primál szuper-bázisváltozókat úgy lehet eltüntetni, ha nullára csökkentjük az értéküket, vagy bevisszük őket a bázisba. Ezt végzi el a következő algoritmus.

---

Megiddo - primál fázis

---

**Input**

$(\tilde{x}, \tilde{y}, \tilde{s})$  komplementáris megoldáspár,  $A_{\mathcal{B}}$  kezdő bázis.

**Output**

$A_{\mathcal{B}^*}$  primál optimális bázis, és a hozzá tartozó  $x^P$  bázismegoldás.

**begin**

$x^P := \tilde{x}$ ,  $\mathcal{B}^* := \mathcal{B}$ ,  $\mathcal{N}^* := \mathcal{N}$ .

$\mathcal{J}^+ := \{j \in \mathcal{N}^* \mid 0 < x_j^P\}$ .

**while**  $|\mathcal{J}^+| > 0$  **do**

Legyen  $j \in \mathcal{J}^+$ .

Készítsük el a  $t_j$  vektort.

$$\alpha := \min \left\{ \frac{x_i^P}{|t_{ji}|} : t_{ji} < 0 \right\},$$

$$x^P := x^P + \alpha t_j.$$

**for each**  $k \in \arg \min_i \left\{ \frac{x_i^P}{|t_{ji}|} : t_{ji} < 0 \right\}$  **do**

**if**  $(k \in \mathcal{B}^*)$

Pivotálás a  $(k, q)$  pozícióban, amire a  $t_{kq} \neq 0$ .

**end**

$\mathcal{J}^+ := \mathcal{J}^+ - j$ .

**end**

**end**

---

**6.3.1. Tétel.** *A primál fázis erősen polinomiális időben előállít egy  $x_P$  optimális primál bázismegoldás.*

**Biz.:** Az algoritmus minden iterációjában  $\mathcal{J}^+$  számossága legalább eggyel csökken. Így a pivotlépések száma legfeljebb a kiinduló megoldás szuper-bázisváltozóinak száma, ami azt jelenti, hogy az algoritmus erősen polinomiális.

Az  $(x^P, s)$  megengedett, és komplementáris megoldáspárból indul az algoritmus. Belátjuk, hogy ez végig fennmarad, amivel az optimalitás bizonyítást nyer.

**Megengedettség:** Az 6.2.1 tétel alapján a  $t_j$  vektor merőleges az  $A$  mátrix sortérére. Ezt felhasználva:

$$A(x^P + \alpha t_j) = Ax^P + \alpha A t_j = b + 0 = b$$

A hányadoseszt miatt pedig  $0 \leq x^P + \alpha t_j$  is teljesül.

**Komplementaritás:** Azt kell belátni, hogy a  $t_j \perp s$ . Ez a  $t_j$  definíciója, és a kezdő bázishoz tartozó bázistábla alapján könnyen látható. Mivel a  $j$  primál szuper-bázisváltozó, ezért a  $t_j$  vektor kizárólag az  $\tilde{x}_{\mathcal{N}^+}$ -nek megfelelő oszlopokból készülhet.

A  $t_j$  nemnulla elemei a kezdő bázis struktúrából leolvashatóan az  $\tilde{s}_B^0$  résznek felelnek meg. A kezdő bázisnak megfelelő struktúra az algoritmus során végig fennáll.

A végén nem lesz primál szuper-bázisváltó, azaz primál bázismegoldást kapunk.  $\square$

**6.3.1. Megjegyzés.** *A primál fázis során, ha a kezdő bázis duál megengedett, akkor az algoritmus során végig az marad.*  $\square$

## 6.4. Duál fázis

---

Megiddo - duál fázis

---

**Input**

$(\tilde{x}, \tilde{y}, \tilde{s})$  komplementáris megoldáspár,  $A_B$  kezdő bázis.

**Output**

$A_{B^*}$  duál optimális bázis, és a hozzá tartozó  $(y^D, s^D)$  bázismegoldás.

**begin**

$s^D := \tilde{s}, y^D := \tilde{y}, B^* := B, N^* := N.$

$\mathcal{J}^+ := \{j \in B^* \mid 0 < s_j^D\}.$

**while**  $|\mathcal{J}^+| > 0$  **do**

Legyen  $j \in \mathcal{J}^+.$

Keressük meg azt az  $i$ -t, amire  $(A_{B^*}^{-1}A)_{ij} = 1.$

Készítsük el a  $t^{(i)}$  vektort.

$$\alpha := \min \left\{ \frac{s_j^D}{t_j^{(i)}} : s_j^D > 0, t_j^{(i)} > 0 \right\},$$

$$s^D := s^D - \alpha t^{(i)}$$

$$y^D := y^D + \alpha A_{B^*}^{-T} e_i. \quad (24)$$

**for each**  $k \in \arg \min_j \left\{ \frac{s_j^D}{t_j^{(i)}} : s_j^D > 0, t_j^{(i)} > 0 \right\}$  **do**

**if**  $(k \notin B^*)$

Pivotálás a  $(k, q)$  pozícióban, amire a  $t_{kq} \neq 0.$

**end**

$\mathcal{J}^+ := \mathcal{J}^+ - j.$

**end**

**end**

---

**6.4.1. Tétel.** *A duál fázis erősen polinomiális időben előállít egy  $(y^D, s^D)$  optimális duális bázismegoldást.*

**Biz.:** Csakúgy, mint a primál fázisnál, a  $\mathcal{J}^+$  elemszáma minden iteráció után legalább eggyel csökken, azaz az algoritmus erősen polinomiális, legfeljebb  $|\mathcal{J}^+|$  pivotlépéssel.

Belátjuk, hogy az  $(y^D, s^D)$  végig megengedett megoldás és az  $(x, s^D)$  végig komplementáris, amivel az optimalitás bizonyítást nyer.

**Megengedettség:** Az  $s^D$  nemnegativitását a hányadosteszt biztosítja minden lépésben.

$$A^T(y^D + \alpha(A_{\mathcal{B}^*}^{-T})e_i) + (s^D - \alpha t^{(i)}) = c + \alpha(A^T A_{\mathcal{B}^*}^{-T})e_i - \alpha t^{(i)} = c + \alpha((A_{\mathcal{B}^*}^{-1}A)^T e_i - t^{(i)}) = c$$

Az utolsó egyenlőségnél kihasználtuk, hogy  $t^{(i)} = (A_{\mathcal{B}^*}^{-1}A)^T e_i$ .

**Komplementaritás:** A komplementaritáshoz azt kell látni, hogy  $t^{(i)} \perp x$ , ami a kezdő bázis struktúrából kiolvasható.

Mivel a végén nem marad duális szuper-bázisváltó, ezért a kapott megoldás bázismegoldás.  $\square$

**6.4.1. Megjegyzés.** *A duál fázis során, ha a kezdő bázis primál megengedett, akkor az algoritmus során végig az marad.*  $\square$

A primál és duál fázisok egymástól függetlenek, bármelyikkel lehet kezdeni. Összességében egy optimális bázis eléréséig legfeljebb annyi pivotiterációra van szükség, mint amennyi a primál és duál szuper-bázisváltók összege a kezdő bázisra vonatkozóan.



## Hivatkozások

- [1] David G. Luenberg, Yinyu Ye: *Linear and Nonlinear Programming*, Springer, 2008.
- [2] Dimitri P. Bertsekas: *Nonlinear programming*, Athena Scientific, 1999.
- [3] B. T. Polyak: *Newton's method and its use in optimization*, European Journal of Operational Research, 181: 1086-1096, 2007.
- [4] R. D. C. Monteiro, I. Adler: *Interior path following primal-dual algorithms. Part I: Linear Programming*, Mathematical Programming, 44: 27-41, 1989.
- [5] Julia Vogt: *Primal-Dual Path-Following Methods for Linear Optimization*, Diploma Thesis, University of Konstanz, 2008.
- [6] Margaréta Halická: *Two simple proofs for analyticity of the central path in linear programming*, Operation Research Letters 28: 9-19, 2001.
- [7] 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, 964-981, 1993.
- [8] Francisco Barahona, Ranga Anbil: *The Volume Algorithm: producing primal solutions with a subgradient method*, Mathematical Programming, 87: 385-399, 2000.
- [9] L. Bahiense, N. Maculan, C. Sagastizábal: *The volume algorithm revisited: relation with bundle methods*, Mathematical Programming, 94: 41-69, 2002.
- [10] Y. Ye: *On the finite convergence of interior-point algorithms for linear programming*, Mathematical Programming, 57: 325-335, 1992.
- [11] N. Megiddo: *On finding primal- and dual-optimal bases*, ORSA Journal on Computing, 3(1): 63-65, 1991.
- [12] Erling D. Andersen, Yinyu Ye: *Combining interior-point and pivoting algorithms for linear programming*, Management Science, 42:1719-1731, 1996.
- [13] Erling D. Andersen: *On exploiting problem structure in a basis identification procedure for linear programming*, Technical report, Odense University, Denmark, 1996.