ENTRÓPIASZERŰ PROXIMÁLIS PONT MÓDSZER ALKALMAZÁSA A VALÓSZÍNŰSÉGGEL KORLÁTOZOTT LINEÁRIS PROGRAMOZÁSI FELADAT MEGOLDÁSÁBAN[1]

KOMÁROMI ÉVA
BKÁE Operációkutatás Tanszék

A következő feladatot vizsgáljuk: $ux \to \min, x \in X = \{x : \ F(x) \geq p\}$, ahol az F többdimenziós folytonos valószínűségi eloszlásfüggvény és adottak az $u > 0$, $u \in R^m$, és a $0 < p < 1$ megbízhatósági szint. Megmutatjuk, hogy e feladat a valószínűséggel korlátozott lineáris programozási feladat duálisának célfüggvényében jelenik meg. Elemezzük a feladat viselkedését az adott paraméterek függvényében. Megoldására proximális pont algoritmust mutatunk be, amelyben a kvadratikus eltéréstag helyett egy Csiszár által bevezetett φ-divergencia függvényt alkalmazunk. Bizonyítjuk az algoritmus konvergenciáját.

1 Bevezetés

A dolgozatban a következő feladatot vizsgáljuk:

$$\begin{array}{llcl} (P) & ux & \to & \min \\ & x & \in & X = \{x : \ F(x) \geq p\} \end{array}$$

ahol F m-dimenziós valószínűségi eloszlásfüggvény, $u \in R^m$ adott pozitív vektor, $0 < p < 1$, az x vektor tartalmazza a döntési változókat.

E feladat választását az motiválja, hogy (P) megjelenik a következő sztochasztikus programozási probléma célfüggvényében:

$$\begin{array}{lrcl} (DD) & \min\limits_{P(x \geq \beta) \geq p} ux + vb & \to & \max \\ & uA + vB & = & c, \quad u \geq 0, \ v \geq 0 \end{array}$$

ahol adottak a β m-dimenziós valószínűségi változó vektor folytonos együttes F eloszlásfüggvénnyel, az A és B $m \times n$ ill. $r \times n$-dimenziós determinisztikus mátrixok, a c n-dimenziós, b $k-$dimenziós vektorok és a $0 < p < 1$ megbízhatósági szint; $(u, v) \in R^{m+r}$ a döntési vektor. P valószínűséget jelöl, $P(x \geq \beta) = F(x)$ az eloszlásfüggvény definíciója szerint, ha F folytonos.

(DD) a lineáris programozási koncepció kiterjesztése. E feladatban a célfüggvény együtthatók egy részét a β valószínűségi változó vektor foglalja

[1]Beérkezett: 2001. december 2. E-mail: `konaromi@bkae.hu`. A dolgozat az FKFP 0231 program támogatásával készült.

magában, amelynek komponensei valamely előnyös tulajdonság (profit, árbevétel stb.) előre nem ismert fajlagos értékeit jelentik. Egy (u, v) lehetséges megoldásra a célfüggvény értéke a legrosszabb esetben $\min_{P(x \geq \beta) \geq p} ux + vb$, ha β legalább az előírt p valószínűséggel veszi fel értékét. A célfüggvény tehát a profit, árbevétel, stb. legkisebb feltételezett értékét képviseli. Így a feladatunk interpretálható úgy, hogy maximalizálni szeretnénk e legkisebb értéket (a profit, stb. kapcsolódó értékét a legrosszabb esetben) a lineáris feltételek által meghatározott lehetséges megoldások halmazán.

Látni fogjuk, hogy (DD) egyúttal duálisa az alábbi valószínűséggel korlátozott lineáris programozási feladatnak:

$$
\begin{aligned}
(PP) \qquad cz &\to \min \\
P(Az &\geq \beta) \geq p, \; Bz \geq b
\end{aligned}
$$

ahol adottak, mint a (DD) feladat leírásánál, az A és B determinisztikus mátrixok, c és b determinisztikus vektorok, a β m-dimenziós valószínűségi változó vektor ismert együttes eloszlásfüggvénnyel, a p megbízhatósági szint, $0 < p < 1$, a z n-dimenziós döntési vektor, P valószínűséget jelöl.

E feladat eredeti megfogalmazása Charnes és Cooper (1959) nevéhez fűződik, akik minden feltételhez egyenként írtak elő megbízhatósági szintet és megmutatták, hogy a feladat felírható ekvivalens lineáris program formájában. E koncepciót Miller és Wagner (1965) kiterjesztették együttes valószínűségi feltételre, ám a β komponenseiről feltették, hogy függetlenek. Itt bemutatott általános formájában a problémát Prékopa (1973) fogalmazta meg először. Prékopa logaritmikusan konkáv eloszlásfüggvényekre vonatkozó eredményei (1973) garantálják a feladat konvexitását az eloszlásfüggvények széles körére, beleértve a többdimenziós normális eloszlást is.

E feladat is kiterjesztése a lineáris programozási koncepciónak. Ha a jobb oldalon szereplő paraméterek egy része (pl. a termékek iránti jövőbeni kereslet, stb.) előre nem ismert, akkor a feladat megoldására a lineáris programozási modell nem alkalmas. Ha azonban e paraméterek valószínűségi változók ismert együttes eloszlással, akkor a lineáris programozási modellezésben alkalmazott megfontolásokhoz igazodóan azt írhatjuk elő, hogy az eredeti lineáris feltételeket a döntési változók legalább előre megadott p valószínűséggel elégítsék ki és ezen előírás teljesülése mellett legyen a cz célfüggvény maximális. Vegyük észre, hogy a (P) feladat a valószínűséggel korlátozott lineáris programozási feladat legegyszerűbb esetének tekinthető, annak az esetnek, amikor A az egységmátrix és B ill b elemei azonosan nullák.

A (PP) feladat megoldása konvexitás megléte esetén is nagy körültekintést igényel. Mayer (1998) könyvében az általa alkalmazott konvex programozási algoritmusokról és azoknak a (PP) feladatra történő számítógépes implementációiról számol be. A többdimenziós eloszlásfüggvény értékeinek kiszámításáról az érdeklődő olvasónak Deák (1990) és Szántai (1988) munkáit ajánljuk. Komáromi (1986) a (PP) feladat duális feladatát a (DD) feladatban fogalmazta meg és kifejlesztett egy prímál-duál tipusú algoritmust, amely a két feladatot egyidejűleg oldja meg. Ez az algoritmus minden iterációban igényli a (P) feladat megoldását.

A (PP) duálisa a következőképpen származtatható. A $P(Az \geq \beta) \geq p$ valószínűségi feltétel a valószínűség monotonitása miatt így írható fel:

$$Az \geq x, \quad F(x) \geq p, \quad x \in \operatorname{supp} F,$$

ahol F a β együttes eloszlásfüggvénye, supp F az F eloszlásfüggvény tartója (az R^m azon részhalmaza, amelynek az F függvény által meghatározott valószínűségi mértéke 1), x pedig m-komponensű változó vektor. Ekkor a (PP) feladat determinisztikus ekvivalens megfogalmazását kapjuk a

$$(PP1) \quad \begin{array}{l} \min_{Az \geq x, Bz \geq b} cz \to \min \\ x \in X = \{x : F(x) \geq p, x \in \operatorname{supp} F\} \end{array}$$

feladat formájában. A $\min_{Az \geq x, Bz \geq b} cz$ célfüggvény értéke egy x helyen maga egy lineáris programozási feladat optimális értéke, amely ha létezik, egyenlő duálisának optimális értékével, feltéve, hogy e duális lehetséges megoldásainak tartománya, a $V = \{u, v : uA + vB = c, u, v \geq 0\}$ poliedrikus halmaz nemüres. Tegyük fel tehát, hogy $V \neq \emptyset$. Ekkor $(PP1)$ ekvivalens, a lineáris programozás dualitási tétele értelmében, a

$$\begin{array}{l} \max_{uA+vB=c; u,v \geq 0} (ux + vb) \to \min \\ x \in X \end{array}$$

feladattal és fennáll, hogy a két feladat célfüggvényértéke egyenlő minden olyan x pontra, amelyre az $ux + vb$ felülről korlátos a V halmazon. A szokásoknak megfelelően $\min_{Az \geq x, Bz \geq b} cz = +\infty$, ha $\{z : Az \geq x, Bz \geq b\} = \emptyset$. Figyelembe véve, hogy a

$$\min_{x \in X} \{\max_{(u,v) \in V} (ux + vb)\} \geq \max_{(u,v) \in V} \{\min_{x \in X} (ux + vb)\}$$

egyenlőtlenség mindig fennáll, a (PP) feladat duálisát az alábbi feladatban találjuk meg:

$$\begin{array}{rcl} \min\limits_{F(x) \geq p; x \in \operatorname{supp} F} (ux) + vb & \to & \max \\ (u, v) & \in & V = \{u, v : uA + vB = c;\ u, v \geq 0\}, \end{array}$$

amely a fent leírt (DD) feladattal ekvivalens, ha supp $F = R^m$.

A két feladatot az itt következő dualitási tétel kapcsolja össze (Komáromi, 1986).

Tétel. *(a) Ha az $ux + vb$ függvénynek van $((\widehat{u}, \widehat{v}), \widehat{x})$ nyeregpontja a V halmazon történő maximalizálásra és az X halmazon történő minimalizálásra nézve, akkor $\widehat{x}$ optimális megoldása a $(PP1)$ feladatnak.*

(b) Legyen F kvázikonkáv és $\widehat{x}$ optimális megoldása a $(PP1)$ feladatnak. Ha az $\{x \in X : \exists y : Ay \geq x, By \geq b\}$ halmaznak van belső pontja, akkor létezik olyan $(u', v') \in V$, hogy $((u', v'), \widehat{x})$ nyeregpontja az $ux+vb$ függvénynek a V halmazon történő maximalizálásra és az X halmazon történő minimalizálásra nézve.

A (P) feladat vizsgálata és megoldása tehát hozzájárul a (PP) valószínűséggel korlátozott lineáris programozási feladat és duálisa, a (DD) sztochasztikus programozási feladat hatékony megoldásához.

A (P) feladat megoldására proximális pont módszert alkalmazunk.

E módszer (fő állomásairól ld. Martinet 1970, Rockafellar 1976 dolgozatokat) egy az n-dimenziós lineáris téren értelmezett konvex f függvény minimalizálására szolgál. A belsőpontos módszerek körébe tartozik, lényege az, hogy egy x_0 közelítő megoldásból kiindulva olyan $\{x_k\}$ sorozatot generál, amelynek $k-$dik eleme a szigorúan konvex $f(x) + \frac{1}{2c_k} |x - x_{k-1}|^2$ függvényt minimalizálja, ahol $\{c_k\}$ alkalmasan választott pozitív sorozat. A proximális pont módszert Teboulle (1992) oly módon alakította át, hogy a kvadratikus eltéréstag helyett egy „entrópiaszerű" eltéréstagot alkalmazott. Bemutatta, hogy az így származtatott entrópiaszerű proximális pont algoritmus egy egységesítő keretnek tekinthető abban az értelemben, hogy közéjük tartozik, az alkalmazott eltérésfüggvényeknek megfelelően, az ismert konvex programozási algoritmusok közül az exponenciális Lagrange szorzó módszer (Bertsekas, 1982), Polyak módosított barrier módszere és a Caroll féle módosított barrier módszer (Polyak, 1992) is.

Az entrópiaszerű eltérésfüggvények körébe két függvénycsaládot szoktak sorolni. Az egyik a Csiszár (1967) által bevezetett φ -divergencia függvénycsalád. Ez a következő: Ha adott a pozitív félegyenesen értelmezett φ függvény, az x pozitív n-dimenziós vektornak az y pozitív n-dimenziós vektortól való eltérését —a φ-divergenciát— az alábbi entrópia-függvénnyel definiáljuk:

$$d_\varphi(x,y) = \sum_{j=1}^{n} y_j \varphi\left(\frac{x_j}{y_j}\right)$$

A pozitív számegyenesen értelmezett olyan szigorúan konvex φ esetén, amelyre $\varphi(1) = 0$ és $\varphi'(1) = 0$, a d_φ függvény nemnegatív és 0 csak akkor, ha $x = y$; de d_φ általában nem teljesíti a távolságfüggvényre előírt háromszög-egyenlőtlenséget és nem szimmetrikus. Az *entrópia* terminológia abból származik, hogy ha speciálisan $\varphi(t) = t \ln t - t + 1$, akkor

$$d_\varphi(x,y) = \sum_{j=1}^{n} y_j \left[\left(\frac{x_j}{y_j}\right) \ln\left(\frac{x_j}{y_j}\right) - \left(\frac{x_j}{y_j}\right) + 1\right] = \sum_{j=1}^{n} \left[x_j \ln\left(\frac{x_j}{y_j}\right) - x_j + y_j\right],$$

amely éppen a Kullback-Leibler relatív entrópia.

A másik a Bregman-féle (1967) eltérésfüggvény család. Tekintsünk egy az R^n-n értelmezett szigorúan konvex differenciálható ψ függvényt. Az x n-dimenziós vektornak az y n-dimenziós vektortól való eltérését az alábbi függvénnyel definiáljuk:

$$D_\psi(x,y) = \psi(x) - \psi(y) - \nabla\psi(y)(x-y).$$

Minthogy ψ szigorúan konvex, ezért D_ψ nemnegatív és 0 csak akkor, ha $x = y$.

Csiszár a (1991) dolgozatában a lineáris inverz problémák (pl. képfelismerés) szempontjából fontos tulajdonságok szerint vizsgálja, csoportosítja és azonosítja az eltérésfüggvényeket és ily módon a két entrópia-függvény családot axiomatikusan vezeti be.

A φ illetve ψ függvény különböző választásaival eljutunk a statisztikában ismert jónéhány eltérésfüggvényhez. Ha például $\varphi(t) = (1-\sqrt{t})^2$, akkor d_φ a Hellinger távolság lesz. Ha $\psi(x) = -\ln x$, akkor D_ψ a Burg-féle entrópiához vezet. Számítógépes kísérleteinkben a $\varphi(t) = -\ln t + t - 1$ függvényt alkalmaztuk.

Az, hogy a proximális pont módszerben szereplő kvadratikus tag helyére egy φ-divergencia függvény kerül, biztosítja, hogy az algoritmus automatikusan pozitív x_k sorozatot generál. Teboulle bizonyította az algoritmus konvergenciáját a minimalizálandó konvex függvényre és az alkalmazott eltérésfüggvényre vonatkozó, az általánosság miatt kissé rigorózus feltevések mellett.

A következő részben a (P) feladat tulajdonságait vizsgáljuk abban az esetben, amikor F az m-dimenziós vektortéren értelmezett folytonosan differenciálható szigorúan logaritmikusan konkáv eloszlásfüggvény. Ezután, a harmadik részben, algoritmust ajánlunk a feladat megoldására oly módon, hogy a feladat Lagrange függvénye nyeregpontjának meghatározására az entrópiaszerű proximális pont módszert alkalmazzuk és az eltérésfüggvényt a $\varphi(t) = -\ln t + t - 1$ függvény generálja. Az utolsó részben bizonyítjuk az algoritmus konvergenciáját. Hangsúlyozzuk azonban, hogy mind az iterációs séma, mind a bizonyítások gondolatmenete tetszőleges Csiszár-féle φ-divergencia függvényre változtatás nélkül alkalmazható, ha fennáll, hogy φ a $(0, 1)$ félegyenesen értelmezett szigorúan konvex, folytonosan differenciálható függvény, $\varphi(1) = \varphi'(1) = 0, \lim_{t\to 0} \varphi'(t) = -\infty$. Annak az oka, hogy a tárgyalást korlátoztuk a $\varphi(t) = -\ln t + t - 1$ függvény esetére az, hogy az algoritmus koncepciójára szerettünk volna koncentrálni és elkerülni azt, hogy újabb konvex analízisbeli fogalmak részletezése elvegye az olvasó kedvét e szellemes, és konvex programozási feladatok megoldására általában is jól használható módszertől.

2 A (P) feladat tulajdonságai

Ha supp F felülről korlátos, akkor a lehetséges megoldások halmaza korlátos. A feladatot általánosabb formájában vizsgáljuk tehát azzal, hogy feltesszük, hogy supp $F = R^m$. Az alábbiakban azt is feltesszük, hogy az F eloszlásfüggvény szigorúan logaritmikusan konkáv és folytonosan differenciálható a supp F belsejében. Ez maga után vonja, hogy F minden komponensének szigorúan növekvő függvénye. Feltesszük továbbá, hogy $u > 0$, $0 < p < 1$.

A logaritmus szigorúan növekvő függvény volta miatt a (P) feladatot a következő ekvivalens és egyben klasszikus konvex minimalizációs formában vizsgálhatjuk:

$$(P) \qquad ux \quad \to \quad \min$$

$$x \quad \in \quad X = \left\{ x \in R^m : \ln \frac{p}{F(x)} \leq 0 \right\}.$$

A lehetséges megoldások X halmaza korlátos alulról abban az értelemben, hogy létezik olyan $w \in R^m$, hogy $x \in X$ maga után vonja, hogy $x \geq w$ - pl. $w_i = \arg\{F_i(x_i) = p\}, (i = 1, \ldots, m)$ választás mellett, ahol F_i az i-edik marginális eloszlás, ez az egyenlőtlenség fennáll. Így minden pozitív u mellett ux alulról korlátos az X halmazon. Mivel X zárt az F folytonossága miatt, ezért (P)-nek van optimális megoldása, és ez a megoldás egyetlen a $\ln F$ szigorú konkávitása miatt. A szereplő függvények konvexek és differenciálhatók.

A feladat tulajdonságait összefoglaljuk az alábbi lemmákban.

1. Lemma. *Az ux lineáris célfüggvény felveszi a minimális értékét a feltételi halmaz egyetlen pontjában.* □

A következő lemmákban $\widehat{x} = \arg\min_{x \in X} ux$.

2. Lemma. *A p választása miatt a $\ln \frac{p}{F(x)} \leq 0$ feltétel kielégíti a Slater feltételt: létezik olyan x' m-dimenziós vektor, amelyre* $\ln p - \ln F(x') < 0$. □

Tekintsük a (P) feladathoz tartozó Kuhn-Tucker feladatot:

$$(K-T) \qquad \begin{array}{lll} u_i = \delta \frac{1}{F(x)} \frac{\partial F(x)}{\partial x_i} & & (i = 1, \ldots, m) \\ \delta \geq 0, x \in R^m, & F(x) \geq p, & \delta(\ln p - \ln F(x)) = 0. \end{array}$$

3. Lemma. *Létezik olyan $\widehat{\delta} \in R$, hogy $\widehat{\delta}$ és $\widehat{x}$ megoldják a $(K-T)$ feladatot.*

Az állítás következik az 1. és 2. Lemmából a Kuhn-Tucker optimalitási tétel értelmében (ld. Mangasarian, 1969). □

4. Lemma. $\widehat{\delta} > 0$, *egyetlen és* $F(\widehat{x}) = p$.

Mivel $\ln F$ parciális deriváltjai pozitívok a függvény szigorúan növekvő volta miatt és mivel $u > 0$, ezért $\widehat{\delta} > 0$. Így az utolsó, az egyensúlyi feltétel miatt $F(\widehat{x}) = p$. □

5. Lemma. *A (P) feladat optimális célfüggvényértéke* $\ln p$ *szigorúan konvex, szigorúan növekvő függvénye.*

Bizonyítás. A konvexitást belátjuk. Legyen ugyanis $x' = \arg\min_{F(x) \geq p'} ux$, $x'' = \arg\min_{F(x) \geq p''} ux$, $x_\lambda = \arg\min_{F(x) \geq p(\lambda)} ux$, ahol $0 < p' < 1$, $0 < p'' < 1, 0 < \lambda < 1$, $\ln p(\lambda) = \lambda \ln p' + (1-\lambda) \ln p''$. Ekkor $\ln F$ szigorú konkávitása miatt

$$\ln F(\lambda x' + (1-\lambda)x'') > \lambda \ln F(x') + (1-\lambda) \ln F(x'') = \lambda \ln p' + (1-\lambda) \ln p'' = \ln p(\lambda).$$

De x_λ jelentése miatt $ux_\lambda < u(\lambda x' + (1-\lambda)x'') = \lambda ux' + (1-\lambda)ux''$, és ez az, amit bizonyítani akartunk.

A második állítás nyilvánvalóan következik abból, hogy $F(\widehat{x}) = p$, F minden komponensének szigorúan növekvő függvénye és $u > 0$. □

A bevezetőben a (P) feladat fontosságát azzal indokoltuk, hogy a (DD) sztochasztikus programozási feladat (a (PP) duálisa) célfüggvényében megjelenik a nemnegatív u vektorokon értelmezett $\min_{F(x)\geq p} ux$ függvény. Hogy ez a motiváció meggyőző legyen, definiálnunk kell e függvényt minden nemnegatív, nemnulla u vektorra.

Jelöljük (P) optimális megoldását mint u függvényét $x(u)$-val és az optimális függvényértéket $ux(u)$-val. A valószínűségelméletből tudjuk, hogy az F parciális deriváltja és az i. változója szerinti feltételes eloszlásfüggvénye között fennáll az alábbi összefüggés:

$$\frac{\partial F(x_1,\ldots,x_m)}{\partial x_i} = F(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_m \mid x_i)\cdot f_i(x_i),$$

ahol f_i a β_i sűrűségfüggvénye. Ez az összefüggés maga után vonja, hogy

$$x_i(u) \to \infty, \quad \text{ha } u_i \to 0.$$

Definiáljuk $x(u)$-t tetszőleges $u \geq 0, u \neq 0$ esetre azzal, hogy vesszük ezt a határértéket.

Az $x(u)$ vektorfüggvény és az $ux(u)$ valós függvény figyelemre méltó tulajdonságokkal rendelkezik. A következő állításban ezeket foglaljuk össze. Az állítás kissé részletezőbb bizonyítása a (Komáromi 1986) dolgozatban megtalálható.

6. Lemma. *Tegyük fel, hogy* $\ln F$ *szigorúan konkáv és minden komponensének szigorúan növekvő függvénye. Ekkor*

(a) A (P) *optimális* $x(u)$ *megoldása folytonos az* $\{u : u > 0\}$ *halmazon. Ha* $\lim_{k\to\infty} u^k = \widehat{u}$ *egy* $u^1, u^2, \ldots$ *sorozatra (* $\widehat{u} \geq 0, \widehat{u} \neq 0, u^k > 0$ *minden k-ra), akkor* $\lim_{k\to\infty} x(u^k) = x(\widehat{u})$.

(b) $ux(u)$ *folytonos és szigorúan konkáv az* $\{u : u \geq 0, u \neq 0\}$ *halmazon abban az értelemben, hogy*

$$[\lambda u' + (1-\lambda)u'']x(\lambda u' + (1-\lambda)u'') > \lambda u'x(u') + (1-\lambda)u''x(u'')$$

minden $u' \geq 0, u' \neq 0, u'' \geq 0, u'' \neq 0,\ \frac{u'}{|u'|} \neq \frac{u''}{|u''|},\ 0 < \lambda < 1$ *esetén;*

(c) Az $\{u : u > 0\}$ *halmazon* $ux(u)$ *differenciálható,* $\nabla ux(u) = x(u)$.

Bizonyítás. (a) Minden $u > 0$ vektorra az $x(u)$ kielégíti a fenti $(K-T)$ feltételeket, szükségképpen $F(x(u)) = p, \delta > 0$, ténylegesen

$$\delta = p\frac{\sum_{i=1}^m u_i}{\sum_{i=1}^m \frac{\partial F(x(u))}{\partial x_i}}.$$

Ezért a $(K-T)$ feltételek folytonos egy-egy értelmű megfeleltetést képviselnek az $\{x \in R^m : F(x) = p\}$ és az $\{u > 0 : \sum_{i=1}^m u_i = 1\}$ halmazok között azon feltevés mellett, hogy az F függvény folytonosan differenciálható.

(b) Legyen $ux(u)$ értéke 0, ha $u = 0$ és $-\infty$, ha $u \not\geq 0$. Az $ux(u)$ folytonossága következik abból, hogy negatívja a $sup_{F(x)\geq p}(-u)x$ támaszfüggvénynek, mely folytonos (ld. Rockafellar, 1970). Konkávitását pedig az $x(u)$ definíciója vonja maga után: $ux' > ux(u)$ minden $x' \neq x(u)$ esetén.

(c) Definíció szerint egy $t \in R^m$ az $ux(u)$ konkáv függvény szubgradiense az $\widehat{u}$ helyen, ha

$$zx(z) \leq \widehat{u}x(\widehat{u}) + t(z - \widehat{u}) \quad \text{minden} \quad z \geq 0, z \neq 0 \quad \text{esetén} .$$

Rockafellar tétele szerint (1970, 25. fejezet) $ux(u)$ differenciálható $\widehat{u}$-ban és $\nabla\widehat{u}x(\widehat{u}) = x(\widehat{u})$, ha $x(\widehat{u})$ az $ux(u)$ egyetlen szubgradiense $\widehat{u}$-ban. A $t = x(\widehat{u})$ véges, ha $\widehat{u} > 0$, és kielégíti a fenti egyenlőtlenséget, mert bármely $z \geq 0, z \neq 0$ esetén

$$zx(z) - zt \leq \widehat{u}x(\widehat{u}) - \widehat{u}t = 0$$

az $x(z)$ definíciója szerint.

Hogy az állítást belássuk, megmutatjuk, hogy ha $t \neq x(\widehat{u})$, akkor van olyan $z > 0$, hogy

$$zx(z) - zt > \widehat{u}x(\widehat{u}) - \widehat{u}t.$$

Tegyük fel először, hogy $\widehat{u}x(\widehat{u}) - \widehat{u}t = \rho \neq 0$. Válasszuk z-t ily módon: $z = \lambda\widehat{u}$, ahol $\lambda > 1$, ha $\rho > 0$, és $0 < \lambda < 1$, ha $\rho < 0$. Figyelembe véve, hogy $x(\lambda\widehat{u}) = x(\widehat{u})$ minden $\lambda > 0$ mellett, azt kapjuk, hogy

$$zx(z) - zt = \lambda[\widehat{u}x(\widehat{u}) - \widehat{u}t] > \widehat{u}x(\widehat{u}) - \widehat{u}t.$$

Tegyük fel most, hogy $\widehat{u}x(\widehat{u}) - \widehat{u}t = 0$. Legyen $I = \{i : x_i(\widehat{u}) - t_i > 0\} \subset \{1, \ldots, m\}$. $I \neq \oslash$, mert $t \neq x(\widehat{u})$ és $\widehat{u} > 0$. Válasszuk a $z > 0$ vektort és $0 < \lambda < 1$-t úgy, hogy

$$\begin{aligned} z_i &= (1-\lambda)\widehat{u}_i, && \textit{ha } x_i(\widehat{u}) - t_i \leq 0, \\ z_i &= \widehat{u}_i, && \textit{ha } x_i(\widehat{u}) - t_i > 0 \end{aligned}$$

fennálljon és $x_i(z) - t_i$ nemnegatív maradjon, ha $i \in I$. Ilyen λ létezik az (a) állításból következően. Akkor fennáll, hogy

$$zx(z) - zt = (1-\lambda)[\widehat{u}x(z) - \widehat{u}t] + \lambda \sum_{i \in I} \widehat{u}_i[x_i(z) - t_i].$$

De $\widehat{u}x(\widehat{u}) < \widehat{u}x(z)$ és $\sum_{i \in I} \widehat{u}_i[x_i(z) - t_i] \geq 0$ λ választása miatt. Így $zx(z) - zt > (1-\lambda)[\widehat{u}x(\widehat{u}) - \widehat{u}t] = 0$

$$zx(z) - zt > \widehat{u}x(\widehat{u}) - \widehat{u}t$$

Ezzel az állítást beláttuk. □

Tekintsük most a (P) feladat klasszikus duális feladatát (Mangasarian, 1969, 7. fejezet):

$$(D) \qquad \begin{aligned} & ux + \delta(\ln p - \ln F(x)) \to \max \\ & u_i = \delta \tfrac{1}{F(x)} \tfrac{\partial F(x)}{\partial x_i} \qquad i = 1, \ldots, m \\ & \delta \geq 0 \end{aligned}$$

A két feladat közötti szoros kapcsolatra világít rá az alábbi lemma (gyenge dualitási tétel):

7. Lemma. *Ha x' a (P) feladat, (δ'', x'') a (D) feladat lehetséges megoldása, akkor teljesül, hogy*

$$ux' \geq ux'' + \delta''(\ln p - \ln F(x''))$$

és egyenlőség pontosan akkor áll fenn, ha $x' = x''$ és $F(x') = p$.

Bizonyítás. Vegyük észre, hogy

$$ux' = ux'' + u(x' - x'') = ux'' + \delta'' \nabla \ln F(x'')(x' - x'')$$

mert (δ'', x'') a (D) lehetséges megoldása. De $\ln F(x)$ szigorú konkávitása miatt

$$\nabla \ln F(x'')(x' - x'') \geq \ln F(x') - \ln F(x'')$$

és egyenlő csak akkor, ha $x' = x''$. Így

$$\begin{aligned} ux' &\geq ux'' + \delta''(\ln F(x') - \ln F(x'')) \\ &= ux'' + \delta''(\ln \tfrac{F(x')}{p} - \ln \tfrac{F(x'')}{p}). \end{aligned}$$

Mivel x' lehetséges megoldása (P)-nek, ezért $F(x') \geq p$. Vagyis $\delta'' \ln \frac{F(x')}{p} \geq 0$ és mivel $\delta'' > 0$, ezért $\delta'' \ln \frac{F(x')}{p} = 0$ csak akkor, ha $F(x') = p$. Így $ux' \geq ux'' - \delta'' \ln \frac{F(x'')}{p}$ és egyenlőség teljesül csak akkor, ha $x' = x''$ és $F(x') = p$. □

Ebből, a 3. és 4. Lemmákat is figyelembe véve következik az alábbi

8. Lemma. *A (D) duális feladatnak létezik egyetlen $(\widehat{\delta}, \widehat{x})$ optimális megoldása, és ebben $\widehat{x}$ a (P) feladatnak is optimális megoldása.* □

Vegyük észre, hogy a (D) tetszőleges lehetséges (δ', x') megoldására fennáll, hogy egyben optimális megoldása annak a feladatnak, amelynek feltételei a (D) feltételei, célfüggvényében azonban p helyett $p' = F(x')$ szerepel. Ezt is tartalmazza a következő észrevétel.

9. Lemma. *Minden adott $\delta' > 0$ esetén van olyan x', hogy (δ', x') lehetséges megoldása a (D) feladatnak. Az x' egyetlen és (δ', x') optimális megoldása az*

$$\begin{aligned} &ux + \delta(\ln p' - \ln F(x)) \to \max \\ &u_i = \delta \frac{1}{F(x)} \frac{\partial F(x)}{\partial x_i} \qquad\qquad i = 1, \ldots, m \\ &\delta \geq 0 \end{aligned}$$

feladatnak, ahol $p' = F(x')$. □

A 8. Lemma a (P)

$$L(\delta, x) = ux + \delta(\ln p - \ln F(x))$$

Lagrange függvényére nézve a következőt jelenti:

10. Lemma. *Az $L(\delta, x) = ux + \delta(\ln p - \ln F(x))$ Lagrange függvénynek egyetlen (δ', x') nyeregpontja van az x-ben R^m-en történő minimalizálásra és δ-ban a nemnegatív valós számokon történő maximalizálásra nézve, $\delta' > 0$, (δ', x') a (D) feladat, x' a (P) feladat optimális megoldása. A* $\min_x\{ux + \delta(\ln p - \ln F(x))\}$ *mint* $\ln p$ *függvénye differenciálható és* δ' *az* $L(\delta, x)$ *nyeregértékének mint* $\ln p$ *függvényének* $\ln p$ *szerinti deriváltja.* □

Ez utóbbi állítás Rockafellar egy tételének (1969, 29.1 Tétel) alkalmazása feladatunkra.

A 9. és 10. Lemmákból következik, hogy minden $0 < p' < 1$ paraméterhez kölcsönösen egyértelműen tartozik egy δ' érték, és δ'-höz egyértelműen tartozik az az x', amellyel δ' és x' kielégítik a $(K-T)$ Kuhn-Tucker feltételeket. Konkávitása miatt a $\ln F(x)$ Hesse mátrixa negatív szemidefinit minden x esetén és ha az x' helyen negatív definit, akkor invertálható és így az implicit függvény tétel értelmében x' δ'-nek differenciálható függvénye. Ezt foglalja össze a következő

11. Lemma. *Tegyük fel, hogy az* $\ln F(x)$ *függvény* $H(x)$ *Hesse mátrixa létezik. Adott* $\delta' > 0$ *esetén egyetlen olyan* $x' \in R^m$ *létezik, amelyre fennáll, hogy* $\frac{u}{\delta'} = \nabla \ln F(x)$. *Ekkor* $x' = x(\delta') = (\nabla \ln F)^{-1}(\frac{u}{\delta'})$. *Ha ezen felül* $H(x)$ *invertálható az* $x' = x(\delta')$ *helyen, akkor* x' *deriválható a* δ' *helyen és* $\nabla x(\delta') = -H(x(\delta'))^{-1}\frac{u}{\delta'}$. □

A (P) duálisa (P) Lagrange függvénye segítségével így fogalmazható meg:

$$(D1) \qquad g(\delta) = \min_x L(\delta, x) = \min_x\{ux + \delta(\ln p - \ln F(x))\} \to \max, \qquad \delta \geq 0.$$

A 8. Lemma értelmében $g(\delta)$ felveszi a maximumát a pozitív δ értékek halmazán. További fontos tulajdonság az alábbi:

12. Lemma. *$g(\delta)$ szigorúan konkáv a pozitív δ értékek halmazán.*

Ezt belátjuk. Legyen $x' = \arg\min_x \{ux + \delta'(\ln p - \ln F(x))\}$, $x'' = \arg\min_x \{ux + \delta''(\ln p - \ln F(x))\}$, $x(\lambda) = \arg\min_x \{ux + (\lambda\delta' + (1-\lambda)\delta'')(\ln p - \ln F(x))\}$, $0 < \lambda < 1$, $\delta' > 0$, $\delta'' > 0$. Ekkor

$$g(\delta') < ux(\lambda) + \delta'(\ln p - \ln F(x(\lambda))) \quad \textit{és} \quad g(\delta'') < ux(\lambda) + \delta''(\ln p - \ln F(x(\lambda))).$$

Így $\lambda g(\delta') + (1-\lambda)g(\delta'') < ux(\lambda) + (\lambda\delta' + (1-\lambda)\delta'')(\ln p - \ln F(x(\lambda))) = g(\lambda\delta' + (1-\lambda)\delta'')$. Ezzel az állítást igazoltuk. □

3 Algoritmus a (P) feladat megoldására

Az itt következő algoritmus a $(D1)$ feladat megoldására szolgál, olyan $\widehat{\delta} > 0$ értéket keresünk, amely $g(\delta)$-t maximalizálja. Ekkor $(\widehat{\delta}, \widehat{x})$, ahol $\widehat{x} = \arg\min_{x \in R^m}\{ux + \widehat{\delta}(\ln p - \ln F(x))\}$, a (D) feladat optimális megoldása lesz, és a 8. Lemma értelmében így $\widehat{x}$ a (P) feladat optimális megoldása.

Az algoritmus konstrukciójában elsősorban Teboulle (1992) és Eggermont (1990) eredményeire támaszkodtunk.

Tekintsük a (P) feladat Lagrange függvényének következő „entrópiaszerű kiterjesztését”:

$$\begin{aligned} E_{\delta_k}(\delta, x) &= ux + \delta(\ln p - \ln F(x)) - \omega_k^{-1} d_\varphi(\delta, \delta') \\ &= L(\delta, x) - \omega_k^{-1} d_\varphi(\delta, \delta'), \qquad \delta \geq 0,\ x \in R^m, \end{aligned}$$

ahol $\omega_k > 0, \delta_k > 0$ adottak. A továbbiakban δ és δ_k eltérését a $\varphi(t) = -\ln t + t - 1$ függvény generálja. Tehát

$$d_\varphi(\delta, \delta_k) = \delta_k(\ln \delta_k - \ln \delta) + \delta - \delta_k.$$

Az algoritmus lényege az, hogy a k. iterációban $\delta_k > 0$ birtokában meghatározzuk $E_{\delta_k}(\delta, x)$ nyeregpontját az x-ben az R^m téren történő minimalizálásra és δ-ban a nemnegatív δ értékeken történő maximalizálásra nézve.

Az eljárás kezdetén választunk egy kiinduló $\delta_0 > 0$ kezdeti értéket és megadjuk azt az $(\underline{\omega}, \overline{\omega})$, $0 < \underline{\omega} < \overline{\omega}$ intervallumot, amelyben az egyes iterációkban az ω_k $(k = 0, 1, \ldots)$ sorozat elemeit választjuk. Az algoritmus k-adik iterációjában meghatározzuk a

$$\delta_{k+1} = \arg\max_{\delta \geq 0} \{g(\delta) - \omega_k^{-1} d(\delta, \delta_k)\} = \arg\max_{\delta \geq 0} \{g(\delta) - \omega_k^{-1}[\delta_k(\ln \delta_k - \ln \delta) + \delta - \delta_k]\}$$

értéket a δ_k érték ismeretében, ahol $g(\delta) = \min_x L(\delta, x) = \min_x\{ux + \delta(\ln p - \ln F(x))\}$. Ez a következő iterációs sémához vezet:

x^{k+1} az alábbi egyenletrendszer megoldása:

$$(1) \qquad u_i = \frac{\delta_k}{1 + \omega_k \ln \frac{F(x)}{p}} \frac{\partial \ln F(x)}{\partial x_i} \qquad i = 1, \ldots, m,$$

δ_{k+1} az alábbi értékadás eredménye:

$$(2) \qquad \delta_{k+1} = \frac{\delta_k}{1 + \omega_k \ln \frac{F(x^{k+1})}{p}}$$

Ha $F(x^{k+1}) = p$ (vagy p-hez kellően közeli), akkor az eljárás befejeződik: x^{k+1} a (P) feladat optimális megoldása.[2]

[2]Egy megjegyzés erejéig itt hivatkozunk arra a bevezetés végi állításunkra, amely szerint az algoritmus alkalmazható és konvergens a φ–divergencia függvények széles körére. Amennyiben olyan Csiszár-féle φ-függvényt alkalmazunk, amely a bevezetőben leírt tulajdonságokkal rendelkezik, akkor mindkét formulában $\frac{\delta_k}{1+\omega_k \ln \frac{F(x^{k+1})}{p}}$ helyett $\delta_k \varphi^{*\prime}(\omega_k \ln \frac{p}{F(x^{k+1})})$ szerepel, ahol φ^* a φ függvény konjugáltja. Be lehet látni, a $\ln \frac{p}{F(x)} \leq 0$ feltétel ekkor ekvivalens az $\omega_k \varphi^*(\ln \frac{p}{F(x)}) \leq 0$ feltétellel. A $\varphi(t) = -\ln t + t - 1$ konjugáltja, mint könnyen kiszámítható, a $\varphi^*(s) = \ln \frac{1}{1-s})$ függvény, $s \leq 1$.

4 Az algoritmus konvergenciájáról

Először azt kell belátnunk, hogy az algoritmus jól definiált, azaz (x^{k+1}, δ_{k+1}) minden iterációban létezik.

13. Lemma. *Az (1)-(2) egyenletrendszert megoldó (x^{k+1}, δ_{k+1}) létezik és nyeregpontja az*

$$E_{\delta_k}(\delta, x) = ux + \delta(\ln p - \ln F(x)) - \omega^{-1} d_\varphi(\delta, \delta_k)$$

függvénynek $\omega = \omega_k > 0$ és $\delta_k > 0$ mellett a δ-ban a nemnegatív számokon történő maximalizálásra és x-ben az R^m-en történő minimalizálásra nézve.

Bizonyítás. Az $\frac{1}{1+\omega_k \ln \frac{F(x)}{p}} \nabla \ln F(x)$ függvény az $\left\{x : F(x) > pe^{-\omega_k^{-1}}\right\}$ halmazon értelmezett szigorúan konkáv $\ln(1 + \omega_k \ln \frac{F(x)}{p})$ függvény gradiense, amely 0-hoz tart, ha $x \to \infty$, vagyis ha $F(x) \to 1$. A $\ln(1 + \omega_k \ln \frac{F(x)}{p})$ függvény esszenciálisan sima: értelmezési tartománya belsejében folytonosan differenciálható és $\left|\nabla \ln(1 + \omega_k \ln \frac{F(x)}{p})\right| \to +\infty$, ha x az értelmezési tartomány egy határpontjához tart. Ezért $\nabla \ln(1 + \omega_k \ln \frac{F(x)}{p})$ invertálható, ld. Rockafellar (1970). Az F gradiense pozitív. Ezért minden pozitív $\frac{u}{\delta_k}$ vektor esetén van egyetlen olyan x^{k+1}, amelyre $\frac{u}{\delta_k} = \nabla \ln(1 + \omega_k \ln \frac{F(x)}{p})$, vagyis olyan x^{k+1}, amely kielégíti az (1) feltételeket. Vegyük észre, hogy egyúttal $x^{k+1} = \arg\min_x E_{\delta_k}(\delta_{k+1}, x)$. Végül vegyük észre, hogy δ_{k+1} maximalizálja az $E_{\delta_k}(\delta, x^{k+1})$ szigorúan konkáv függvényt a $\delta \geq 0$ feltétel mellett. Így (x^{k+1}, δ_{k+1}) az $E_{\delta_k}(\delta, x)$ függvény nyeregpontja. □

Mivel $F(x^{k+1}) > pe^{-\omega_k^{-1}}$, ezért $\delta_{k+1} > 0$. Fontossága miatt ezt külön állításban fogalmazzuk meg.

14. Lemma. $\delta_{k+1} > 0$.

15. Lemma. *A $g(\delta_k)$ szigorúan növekvő konvergens sorozat.*

Bizonyítás. Vegyük figyelembe, hogy

$$\begin{aligned} x^{k+1} &= \arg\min_x \{ux + \delta_{k+1}(\ln p - \ln F(x))\}, \\ \delta_{k+1} &= \arg\max \left\{\delta(\ln p - \ln F(x_{k+1})) - \omega_k^{-1} d_\varphi(\delta, \delta_k)\right\} . \end{aligned}$$

Vagyis $g(\delta_{k+1}) = ux^{k+1} + \delta_{k+1}(\ln p - \ln F(x^{k+1}))$. Így

$$\begin{aligned} g(\delta_{k+1}) \ &> ux^{k+1} + \delta_{k+1}(\ln p - \ln F(x^{k+1})) - \omega_k^{-1} d_\varphi(\delta_{k+1}, \delta_k) \\ &= \max\nolimits_{\delta \geq 0} \left\{ux^{k+1} + \delta(\ln p - \ln F(x^{k+1})) - \omega_k^{-1} d_\varphi(\delta, \delta_k)\right\} \\ &> ux^{k+1} + \delta_k(\ln p - \ln F(x^{k+1})) - \omega_k^{-1} d_\varphi(\delta_k, \delta_k) \\ &= ux^{k+1} + \delta_k(\ln p - \ln F(x^{k+1})) \\ &> \min\nolimits_x \{ux + \delta_k(\ln p - \ln F(x))\} \\ &\ \ ux^k + \delta_k(\ln p - \ln F(x^k)) = g(\delta_k). \end{aligned}$$

A $\{g(\delta_k)\}$ sorozat konvergenciája ezután abból következik, hogy a sorozat felülről korlátos, hiszen $g(\delta)$ felveszi a maximumát a pozitív δ értékek halmazán. $\square$

16. Lemma. *A $\{\delta_k\}$ sorozat konvergens.*

Bizonyítás. A 10. Lemma szerint $g(\delta)$ felveszi a maximumát. Mivel $g(\delta)$ konkáv, ez azt jelenti, hogy a $\{\delta \geq 0 : g(\delta) \geq \gamma\}$ halmaz korlátos minden $\gamma \in R$ esetén, beleértve a $\gamma = g(\delta_0)$ esetet is. Így a $\{\delta_k\}$ sorozat is korlátos, van tehát torlódási pontja. Tegyük fel, hogy az állítással ellentétben a $\{\delta_k\}$ sorozatnak két torlódási pontja van, jelölje ezeket $\overline{\delta}$ és $\overline{\overline{\delta}}$. Minthogy az $\{\omega_k\}$ sorozat korlátos, szintén van torlódási pontja, jelölje ezt ω, $\omega > 0$. Legyen $\{\delta_{j_k}\}$ a $\{\delta_k\}$ olyan végtelen részsorozata, amelyre fennáll, hogy

$$\lim_{k\to\infty} \omega_{j_k} = \omega, \ \lim_{k\to\infty} \delta_{j_k} = \overline{\overline{\delta}}, \ \lim_{k\to\infty} \delta_{j_k+1} = \overline{\delta}.$$

A 15. Lemma bizonyításában beláttuk, hogy

$$g(\delta_{j_k+1}) \quad > g(\delta_{j_k+1}) - \omega_k^{-1} d_\varphi(\delta_{j_k+1}, \delta_{j_k}) > g(\delta_{j_k})$$

minden k esetén, vagyis a

$$g(\overline{\delta}) \quad \geq g(\overline{\delta}) - \omega^{-1} d_\varphi(\overline{\delta}, \overline{\overline{\delta}}) \geq g(\overline{\overline{\delta}})$$

egyenlőtlenségnek teljesülnie kell a szereplő függvények folytonossága miatt. Mivel $g(\overline{\delta})$ =g$(\overline{\overline{\delta}})$ a 15. Lemma szerint, ezért $d_\varphi(\overline{\delta}, \overline{\overline{\delta}}) = 0$, ami implikálja, hogy $\overline{\delta} = \overline{\overline{\delta}}$. Ellentmondásra jutottunk tehát azzal a feltevéssel, hogy $\overline{\delta} \neq \overline{\overline{\delta}}$. $\square$

17. Lemma. *Az $\{F(x^k)\}$ sorozat konvergens,* $\lim_{k\to\infty} F(x^k) = p$.

Bizonyítás. Legyen $\delta_0 = \lim_{k\to\infty} \delta^k$. Ha $\delta_0 > 0$, akkor a

$$\delta_{k+1} = \frac{\delta_k}{1 + \omega_k \frac{F(x^{k+1})}{p}}$$

összefüggés miatt és azért, mert az ω_k sorozatot úgy választjuk, hogy minden eleme egy adott pozitív intervallumban legyen következik, hogy az $F(x^{k+1})$ sorozat p–hez tart. De $\delta_0 > 0$, különben a $\nabla \ln F(x)$ minden komponense $+\infty$-hez tartana, hiszen (x^k, δ_k) kielégíti az $u = \delta_k \nabla \ln F(x)$ egyenletrendszert. Ez F eloszlásfüggvény volta miatt csak akkor lehet, ha az $F(x^k)$ sorozat 0-hoz tart. De ekkor, figyelembe véve azt is, hogy ω_k-t egy korlátos pozitív intervallumban választjuk, az $1 + \omega_k(\ln F(x^k) - \ln p)$ szükségképpen $-\infty$-hez tart, ellentmondásban azzal, hogy $\delta_k > 0$. $\square$

5 A módszer implementálhatóságáról

Végül néhány aggodalmaskodó megjegyzés következik a módszer implementálhatóságáról. Minden iterációban, mint látható, meg kell oldanunk az (1) egyenletrendszert (ez maga egy minimalizálási feladat). A szereplő függvények az F gradiensei, amelyek függvényértékei többváltozós integrálás eredményei (többváltozós normális eloszlásfüggvény esetében e gradiensek is többváltozós normális eloszlásfüggvények), pontosan kiszámítani őket nem lehet, csak közelíteni. Az egyenletrendszer megoldására, majd az új δ értékének pontos meghatározására már emiatt sincs mód. A (P) feladat optimális megoldása tehát szükségképpen pontatlan, amely tény további pontatlanságokat eredményez, amikor e megoldásokat a (PP) ill. (DD) feladatok megoldására szolgáló már hivatkozott (Komáromi, 1986) prímál-duál algoritmusban használjuk fel vagy ha más konvex programozási algoritmust alkalmazunk a (DD) feladat megoldására. (E szempontok természetesen nem csak az entrópiaszerű proximális pont módszerrel kapcsolatban merülnek fel.) Ilyen körülmények között felvethető, vajon érdemes-e az eljárás konvergenciájának vizsgálatával bajlódni. A válasz természetesen igenlő, hiszen óvintézkedéseket is csak akkor tudunk beépíteni az eljárásba, ha tudjuk, hogy „illik" viselkednie.

A szerző tapasztalatai a pontatlansággal kapcsolatos aggodalmakat nem igazolták. A (P) feladat megoldására végzett számítógépes kisérleteinkben legfeljebb 10-változós nemdegenerált normális eloszlásfüggvényeket és a $\varphi(t) = t\ln t - t + 1$ függvényt választottuk. Az (1) egyenletrendszer megoldására a ciklikus csökkentés módszerét alkalmaztuk, amely abban áll e feladat esetében, hogy a $(K-T)$ feladat i-edik egyenletében az x_i értékét növeljük vagy csökkentjük oly mértékben, hogy egyenlőséget kapjunk – ezt folytatjuk egészen addig, amíg a pontossággal megelégszünk. Ez az eljárás is konvergens, a bizonyítás megtalálható Zangwill (1969) könyvében. Ilyen választás mellett az entrópiaszerű proximális pont módszer biztonságosnak és gyorsnak bizonyult.

Irodalom

1. Bertsekas, D., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, NY, 1982.

2. Bregman, L. M., The Relaxation Method of Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming, *USSR Comput. Math. and Math Phys.* 7, 1967, pp. 200–217.

3. Csiszár, I., Information-Type Measures of Difference of Probability Distributions and Indirect Observations, *Studia Sci. Math. Hungar.* 2, 1967, pp. 299–318.

4. Csiszár, I., Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems, *Annals of Statistics* 19, 1991, pp. 2032–68.

5. Deák, I., *Random Numbers Generators and Simulation*, Akadémiai Kiadó, 1990.

6. Eggermont, P. P. B., Multiplicative Iterative Algorithms for Convex Programming *Linear Algebra and Appl.* 130, 1990, pp. 25–42.

7. Iusem, A. N., Svaiter, B. F., Teboulle, M., Entropy-Like Proximal Methods in Convex Programming, *Mathematics of Operations Research* 19, 1994, pp. 790–814.

8. Komáromi, É., A Dual Method for the Probabilistic Constrained Problem, *Mathematical Programming Study* 28, 1986, pp. 94–112.
9. Mangasarian, O. L., *Nonlinear Programming,* McGraw-Hill, New York, 1969.
10. Martinet, B., Regularisation d'inéquations variationnelles par approximations successive, *Rev. Francaise d'Automatique et Inform. Rech. Opér.* 4, 1970, pp. 154–159.
11. Mayer, J., *Stochastic Linear Programming Algorithms*, Gordon and Breach Science Publishers, 1998.
12. Polyak, R. A., Modified Barrier Functions (Theory and Methods), *Mathematical Programming* 54, 1992, pp. 177–222
13. Prékopa, A., Contributions to the Theory of Stochastic Programming, *Mathematical Programming* 4, 1973, pp. 202–221.
14. Prékopa, A., On Logarithmic Concave Measures and Functions, *Acta Scientiarum Mathematicarum* 34, 1973, pp. 335–343.
15. Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
16. Rockafellar, R. T., Monotone Operators and the Proximal Point Algorithm, *SIAM J. Control Optimization* 14, 1976, pp. 877–898.
17. Szántai, T., A Computer Code for Solution of Probabilistic Constrained Stochastic Programming Problems, In: *Numerical Techniques for Stochastic Optimization* (Y. Ermoliev and R. J.-B. Wets eds), Springer-Verklag, 1988, pp. 229–235.
18. Teboulle, M., Entropic Proximal Mappings with Applications to Nonlinear Programming, *Mathematics of Operations Research* 17, 1992, pp. 670–690.
19. Zangwill, I. W., *Nonlinear Programming*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

AN APPLICATION OF THE ENTROPY-LIKE PROXIMAL POINT METHOD IN THE SOLUTION OF THE PROBABILISTIC CONSTRAINED LINEAR PROGRAMMING PROBLEM

We are concerned with the following problem: $ux \to \min, x \in X = \{x : \ F(x) \geq p\}$, where F is a given joint continuous probability distribution function, $u > 0$, $u \in R^m$, and p is a reliability level, $0 < p < 1$. It will be shown that this problem appears in the objective function of the dual of the probabilistic constrained linear programming problem. We will investigate the behavior of the problem as the function of its parameters. For solving it we apply a proximal point algorithm in which instead of a quadratic term we employ a φ-divergence function introduced by Csiszár. We prove the convergence of the presented algorithm.