Martos Béla

Kvadratikus programozás kvázikonvex célfüggvénnyel

Gazdasági optimálási problémák gyakran nem modellezhetők valósághűen lineáris programozási feladatként, de megfogalmazhatók oly módon, hogy a korlátozó feltételek lineárisak legyenek, míg a célfüggvény nem lineáris. Ennek egyik legegyszerűbb esete az, amikor a célfüggvény kvadratikus, vagy legalábbis olyan típusú függvény, amely kvadratikussal jól közelíthető. Így érthető, hogy a kvadratikus programozás iránti érdeklődés viszonylag régi keletű és a témának kiterjedt irodalma van. [2], [4], [7], [11], [12] Ezek a vizsgálatok, amelyek számos hatékony algoritmust eredményeztek, megegyeztek abban, hogy feltételezték a minimálandó célfüggvény konvexitását. E cikkben megmutatjuk, hogy a korábban kidolgozott módszerek egyike, éspedig Frank és Wolfe [6] módszere, alkalmazható abban az esetben is, ha a célfüggvény nem konvex, de a változók nem-negatív értékére kvázikonvex.

Megjegyzendő azonban, hogy ez esetben az eljárás általában nem véges, míg a konvex kvadratikus programozási módszerek legtöbbje véges számú lépésben eljut az optimális megoldáshoz. Viszont egy példán meg fogjuk mutatni, hogy a véges algoritmusok közül az egyik legismertebb, Wolfe [12] módszere, az általunk vizsgált esetben már csődöt mond. E cikk 1. fejezetében a nem-negatív ortánsban kvázikonvex függvények néhány fontos tulajdonságával foglalkozunk. A 2. fejezetben leírjuk Frank és Wolfe algoritmusát, ahogyan az erre a függvénytípusra alkalmazható, és bizonyítjuk az eljárás konvergenciáját. A 3. fejezet egy példát tartalmaz és az imént említett ellenpéldát.

1. A nem-negatív ortánsban kvázikonvex kvadratikus függvények

Tekintsük a

$$\varphi(x) = \frac{1}{2} x'Cx + p'x = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_i x_j + \sum_{i=1}^{n} p_i x_i \tag{1}$$

kvadratikus függvényt, ahol $x = [x_1, \ldots, x_n]'$ és $p = [p_1, \ldots, p_n]'$ n elemű vektorok és $C = [c_{ij}]$ egy valós szimmetrikus $n \cdot n$-es mátrix. $E^n_+ = \{x \in E^n \mid x \geq 0\}$ az n dimenziós euklideszi tér nem-negatív ortánsát jelöli.

Ismeretes, hogy egy differenciálható $\psi(x)$ függvény az $X \subset E^n$ konvex halmazon akkor és csak akkor *konvex*, ha minden $x^1, x^2 \in X$-re

$$\psi(x^1) - \psi(x^2) \leq (x^1 - x^2)' \nabla \psi(x^1) \tag{2}$$

ahol $$\nabla\psi(x^1) = \left[\frac{\partial\psi(x)}{\partial x_1}, \frac{\partial\psi(x)}{\partial x_2}, \dots \frac{\partial\psi(x)}{\partial x_n}\right]_{x=x^1},$$

a $\psi(x)$ függvény x^1 pontbeli gradiense.

Hasonlóképpen $\psi(x)$ akkor és csak akkor *kvázikonvex* a X halmazon, [1] ha

$$\left.\begin{array}{r} x^1, x^2 \in X \\ \psi(x^1) \geq \psi(x^2) \end{array}\right\} \Rightarrow (x^1 - x^2)'\nabla\psi(x^1) \geq 0 \tag{3}$$

(2) és (3)-ból közvetlenül látható, hogy ha $\psi(x)$ konvex, akkor egyben kvázikonvex is X-en.

Ha (2)-t, ill. (3)-at az (1) alatti kvadratikus függvényre alkalmazzuk, némi számolással a következő kritériumokat kapjuk:
A $\varphi(x)$ kvadratikus függvény akkor és csak akkor *konvex* a konvex $X \subset E^n$ halmazon, ha minden x^1, $x^2 \in X$-re

$$(x^1 - x^2)'C(x^1 - x^2) \geq 0; \tag{4}$$

valamint akkor és csak akkor *kvázikonvex*, ha

$$\left.\begin{array}{r} x^1, x^2 \in X \\ \varphi(x^1) \geq \varphi(x^2) \end{array}\right\} \Rightarrow (x^1 - x^2)'(Cx^1 + p) \geq 0\,. \tag{5}$$

Már itt látható, hogy amíg a kvadratikus függvény konvexitása csak a C mátrixtól függ, addig kvázikonvexitása a lineáris rész p koefficiensvektorától is.

Annak megmutatására, hogy létezik olyan kvadratikus függvény, amely egy konvex halmazon kvázikonvex, de nem konvex, tekintsük a kétváltozós

$$\varphi(x, y) = -xy$$

függvényt E^2_+-on. Itt

$$C = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

$$p' = [0,\ 0]$$

és nyilván $[x^1, y^1] = [0, 0]$; $[x^2, y^2] = [1, 1]$-ra

$$[-1,\ -1]\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = -2 < 0$$

ellentétben (4)-gyel, tehát $\varphi(x, y)$ nem konvex E^2_+-on. Viszont ha x^1, x^2, y^1, $y^2 \geq 0$ és

$$-x^1\, y^1 \geq -x^2\, y^2$$

azaz $$x^2\, y^2 \geq x^1\, y^1$$

akkor

$$[x^1 - x^2,\ y^1 - y^2]\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x^1 \\ y^1 \end{bmatrix} = x^1y^2 + x^2y^1 - 2\,x^1y^1 \geq$$

$$\geq 2\sqrt{x^1x^2y^1y^2} - 2\,x^1y^1 = 2\sqrt{x^1y^1}\left(\sqrt{x^2y^2} - \sqrt{x^1y^1}\right) \geq 0,$$

tehát (5) áll, $\varphi(x, y)$ kvázikonvex E_+^2-on.

A következő két állítás viszonylag könnyen belátható:

a) Ha $\varphi(x)$ konvex E_+^n-ban, akkor konvex E^n-ben is

b) Ha $\varphi(x)$ kvázikonvex E^n-ben, akkor konvex is E^n-ben

E tételeket nem használjuk fel, ezért itt nem is bizonyítjuk, viszont motiválják azt, hogy a következőkben olyan kvadratikus függvényekkel foglalkozzunk, amelyek E_+^n-ban kvázikonvexek, de nem konvexek. Ezt készíti elő a következő tétel:

1\. *TÉTEL* A $\varphi(x) = \frac{1}{2}x'Cx + p'x$ függvény akkor és csak akkor kvázikonvex E_+^n-on, ha minden $v \in E^n$-re áll:

$$v'Cv < 0 \Rightarrow \begin{bmatrix} Cv \\ p'v \end{bmatrix} \gtreqless 0 \tag{6}$$

(ahol $\gtreqless$ azt jelenti, hogy vagy $\geq$ vagy $\leq$, tehát az $(n+1)$ elemű $\begin{bmatrix} Cv \\ p'v \end{bmatrix}$ vektornak nincs két ellentétes előjelű komponense).

Bizonyítás: Legyen x^1, $x^2 \in E_+^n$ és $\varphi(x^1) \geq \varphi(x^2)$, azaz

$$(x^1 - x^2)' \left[\frac{1}{2} C(x^1 + x^2) + p\right] \geq 0 . \tag{7}$$

$\varphi(x)$ pontosan akkor kvázikonvex E_+^n-on, ha

$$(x^1 - x^2)' (Cx^1 + p) \geq 0 . \tag{8}$$

Elégségesség Azt bizonyítjuk, hogy ha (6) áll, akkor (7) $\Rightarrow$ (8). Először tegyük fel, hogy x^1, $x^2 > 0$

a) Ha $\frac{1}{2}(x^1 - x^2)' C(x^1 - x^2) \geq 0$, akkor ezt (7)-hez adva (8)-at kapjuk.

b) Ha $\frac{1}{2}(x^1 - x^2)' C(x^1 - x^2) < 0$, akkor (6) alkalmazásával:

$$\begin{bmatrix} C(x^1 - x^2) \\ p'(x^1 - x^2) \end{bmatrix} \gtreqless 0 , \tag{9}$$

úgy hogy $C(x^1 - x^2) \neq 0$. Továbbá (7) némi átalakításával

$$\left[\frac{1}{2}(x^1 + x^2)', 1\right] \begin{bmatrix} C(x^1 - x^2) \\ p'(x^1 - x^2) \end{bmatrix} \geq 0 .$$

Itt az első tényező egy pozitív vektor, a másodiknak pedig van legalább egy 0-tól különböző komponense és nincs két különböző előjelű komponense. De 0-tól különböző komponensei nem lehetnek negatívak, mert akkor a szorzat is negatív lenne. Tehát (9) csak abban a változatban állhat, hogy

$$\begin{bmatrix} C(x^1 - x^2) \\ p'(x^1 - x^2) \end{bmatrix} \geq 0 .$$

Ezt balról az $[x^{1\prime}, 1]$ nem-negatív vektorral szorozva (8)-at kapjuk. Így $\varphi(x)$ kvázikonvex E^n_+ belsejében, tehát a folytonosság miatt kvázikonvex E^n_+-on.

Szükségesség Azt bizonyítjuk, hogy ha van olyan $v = v^0$, amelyre (6) nem áll, akkor van olyan x^1, $x^2 \in E^n_+$, amelyre (7) áll, de (8) nem. Tegyük fel tehát, hogy

$$v^{0\prime} C v^0 < 0, \tag{10}$$

de $\begin{bmatrix} C v^0 \\ p' v^0 \end{bmatrix}$-nak van két ellenkező előjelű komponense.

Ekkor azonban van olyan $(n + 1)$ elemű pozitív

$$\begin{bmatrix} y^0 \\ \eta \end{bmatrix} > 0$$

vektor hogy

$$[y^{0\prime}, \eta] \begin{bmatrix} C v^0 \\ p' v^0 \end{bmatrix} = 0 .$$

Itt $\eta > 0$-val az első tényezőt osztva és $x^0 = \dfrac{y^0}{\eta}$-t téve, valamint a tényezők sorrendjét felcserélve azt kapjuk, hogy

$$v^{0\prime} (C x^0 + p) = 0 . \tag{11}$$

Emlékezve arra, hogy $x^0 = \dfrac{y^0}{\eta} > 0$, van egy olyan kicsiny $\alpha > 0$ szám, hogy

$$x^1 = x^0 + \frac{1}{2} \alpha v^0 \geq 0$$

és

$$x^2 = x^0 - \frac{1}{2} \alpha v^0 \geq 0 .$$

Így

$$\alpha v^0 = x^1 - x^2$$

$$x^0 = \frac{1}{2} (x^1 + x^2)$$

és (10)-ből, ill. (11)-ből

$$\alpha^2 v^{0\prime} C v^0 = (x^1 - x^2)' C (x^1 - x^2) < 0 \tag{12}$$

$$\alpha v^{0\prime}(C x^0 + p) = (x^1 - x^2)' \left[\frac{1}{2} C (x^1 + x^2) + p \right] = 0 . \tag{13}$$

(13) szerint (7) áll, de (12) $\dfrac{1}{2}$-szeresét (13)-hoz adva azt kapjuk, hogy

$$(x^1 - x^2)' (C x^1 + p) < 0,$$

azaz (8) nem áll. Ezzel a tételt bebizonyítottuk.

Megjegyzés: A tétel alapján, $p = 0$-t téve, könnyen belátható, hogy ha $\varphi(x) = \frac{1}{2}x'Cx + p'x$ kvázikonvex E^n_+-on, akkor a $Q(x) = \frac{1}{2}x'Cx$ kvadratikus forma is az. Továbbá azt is tudjuk, hogy ha $\varphi(x)$ nem konvex, akkor $Q(x)$ sem az. Ez a megjegyzés lehetővé teszi, hogy a további tételek bizonyításában felhasználjuk a szubdefinit mátrixoknak [9]-ben részletesebben kifejtett és bizonyított tulajdonságait.

2. *TÉTEL.* Ha a $\varphi(x) = \frac{1}{2}x'Cx + p'x$ függvény kvázikonvex E^n_+-on, de nem konvex, akkor a C együtthatómátrixról és a p vektorról a következőket mondhatjuk:

a) $C \leq 0$, $C \neq 0$,

b) C-nek pontosan egy (egyszeres) negatív sajátértéke van, ez a sajátérték domináns,

c) a domináns sajátértékhez tartozó sajátvektornak nincsenek különböző előjelű komponensei,

d) $p \leq 0$,

e) $Ce_i = 0 \Rightarrow p_i = 0$. ($e_i$ = az i-ik egységvektor)

Bizonyítás: A tétel *a)*, *b)* és *c)* alatti állításainak bizonyítását [9] Theorem 1-ben találhatjuk.

d) Az *a)* állítás és (6) következtében

$$v > 0 \Rightarrow \left\{\begin{matrix} v'Cv < 0 \\ Cv \leq 0,\ Cv \neq 0 \end{matrix}\right\} \Rightarrow p'v \leq 0$$

Mivel ez minden $v > 0$-ra áll, következik $p \leq 0$.

e) Legyen $e = [1, 1, \ldots, 1]'$ és $v = e + \beta e_i$, $\beta \neq 0$. Ekkor

$$Ce_i = 0 \Rightarrow \left\{\begin{matrix} v'Cv = e'Ce < 0 \\ Cv = Ce \leq 0, \neq 0 \end{matrix}\right\} \Rightarrow p'v = p'e + \beta p_i \leq 0$$

minden $\beta \neq 0$-ra. Tehát $p_i = 0$.

A 2. Tétel *e)* állítását a következőképpen is megfogalmazhatjuk: Ha valamelyik x_i változó a $\varphi(x)$ függvény kvadratikus részében nem lép fel, akkor a függvényben egyáltalán nem lép fel (azaz minden koefficiense 0). Az ilyen változókat *irreleváns* változóknak nevezzük, a többit *relevánsnak.* Most már bizonyítani tudjuk a következő tételt:

3. *TÉTEL.* Ha $\varphi(x) = \frac{1}{2}x'Cx + p'x$ kvázikonvex E^n_+-on, de nem konvex és x minden komponense releváns, akkor

$$\left.\begin{matrix} x^1,\ x^2 \geq 0,\ x^1 \neq 0 \\ (x^1 - x^2)'(Cx^1 + p) \leq 0 \end{matrix}\right\} \Rightarrow \varphi(x^1) \leq \varphi(x^2) \tag{14}$$

Bizonyítás: Másutt ([9], Theorem 2) már bizonyítottam, hogy ha $\varphi(x)$ eleget tesz e tétel feltételeinek, akkor minden $v \in E$-re

$$v'Cv < 0 \Rightarrow Cv \gtrless 0 \tag{15}$$

(ahol $\gtrless$ azt jelenti, hogy $>$ vagy $<$).

a) Ha $\frac{1}{2}(x^1-x^2)'\,C\,(x^1-x^2)\geqq 0$, akkor ennek (-1)-szeresét az $(x^1-x^2)'\,(Cx^1+p)\leqq 0$ egyenlőtlenséghez hozzáadva átrendezés után $\varphi(x^1)\leqq\varphi(x^2)$-t kapjuk. Tehát (14) áll.

b) Ha $1/2\,(x^1-x^2)'C\,(x^1-x^2)<0$, akkor (15) szerint

$$C(x^1-x^2)\gtreqless 0 \tag{16}$$

Ha most $C(x^1-x^2)>0$, akkor $x^1\geqq 0$, $x^1\neq 0$-ra tekintettel

$$(x^1-x^2)'C\,x^1>0.$$

Másrészt (6) szerint

$$(x^1-x^2)'p\geqq 0.$$

E kettőből $(x^1-x^2)'(Cx^1+p)>0$, feltevésünkkel ellentétben.

Ha pedig

$$C(x^1-x^2)<0 \tag{17}$$

akkor (6) miatt

$$p'(x^1-x^2)\leqq 0. \tag{18}$$

(17)-et $\frac{1}{2}(x^1+x^2)$-vel megszorozva és (18)-hoz adva, megfelelő átrendezés után $\varphi(x^1)\leqq\varphi(x^2)$-hez jutunk, azaz (14) ez esetben is áll. Ezzel a tételt bebizonyítottuk.

Megjegyzés: A (14) implikáció, mint (2) és (3) alapján könnyen látható, az x előjelének minden megkötése nélkül áll, ha $\varphi(x)$ konvex E^n-en. A tétel valójában azt fejezi ki, hogy az adott feltételek mellett $\varphi(x)$ pszeudokonvex [8] az $E^n_+-\{0\}$ halmazon (a szemipozitív ortánsban).

A gradiens módszerek alkalmazása szempontjából alapvető a 3. Tételből levonható alábbi korollárium:

4. *korollárium:* Legyen $X\subset E^n_+$ konvex, $\varphi(x)$ elégítse ki a 3. Tétel feltételeit Ha $x^*\in X$, $x^*\neq 0$ minden $x\in X$-re kielégíti az

$$(x^*-x)'(Cx^*+p)\leqq 0 \tag{19}$$

egyenlőtlenséget, akkor x^* optimális megoldása a $\min\{\varphi(x)\mid x\in X\}$ programozási feladatnak.

Mivel (19) azt fejezi ki, hogy $\varphi(x)$ függvénynek minden, az x^* pontból kiinduló megengedett irány menti deriváltja e pontban nem negatív, valójában optimalitási kritériumhoz jutottunk a 3. Tételben körülírt függvényekre.

2. A Frank-Wolfe-algoritmus

Tekintsük a következő kvadratikus programozási feladatot (QP): Keressük a

$$\varphi(x)=\frac{1}{2}x'\,Cx+p'x$$

kvadratikus függvény minimumát az

$$Ax + By = b$$
$$x \geq 0,\ y \geq 0$$

lineáris korlátozó feltételek mellett. Itt x a célfüggvény szempontjából releváns, y az irreleváns változók n ill. k elemű vektora, A egy $m \cdot n$, B egy $m \cdot k$ méretű mátrix, b egy m elemű vektor. A

$$z = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$L = \{z \mid [A \ \ B]\, z = b,\ z \geq 0\}$$

jelölés bevezetésével a feladat

$$QP: \qquad \min \{\varphi(x) \mid z \in L\}$$

alakban is írható. (L-et *megengedett halmaznak* nevezzük)

(Továbbiakban az azonos jelzéssel vagy felső indexszel ellátott z, x, y vektorok összetartoznak, pl. x^0, y^0 a z^0 részvektorai.)

Mint már említettük, számos módszer, éspedig az alábbinál hatékonyabb módszerek is rendelkezésre állnak a QP feladat megoldására, ha $\varphi(x)$ konvex. Ezért a továbbiakban a nem konvex esetre szorítkozhatunk. A következő feltevésekkel élünk:

A. $\varphi(x)$ kvázikonvex E_n^+-on, de nem konvex.

B. L korlátos.[1]

A Frank—Wolfe-algoritmus lényegében abból áll, hogy kiindulva egy megengedett bázismegoldásból, olyanból, amelynek x részvektora nem 0, felváltva egy lineáris programozási feladatot és egy egyváltozós kvadratikus szélsőértékfeladatot oldunk meg. Az LP feladat célfüggvény koefficienseit mindig $\varphi(x)$-nek az előző kvadratikus feladatot megoldó pontban vett gradiense szolgáltatja, azaz a $\varphi(x)$ célfüggvényt e pontban vett lineáris közelítésével helyettesítjük. Az egyváltozós kvadratikus feladat megoldásakor a $\varphi(x)$ célfüggvényt minimáljuk egy egyenes szakasz mentén, amely az előző LP feladat megoldásának megfelelő csúcspontot és az előző kvadratikus feladatot megoldó pontot köti össze.

Az algoritmus részletes leírása

1. fázis. Kiinduló megoldás keresése

Oldjuk meg a következő lineáris programozási feladatot

$$LP(0): \qquad \min \{-e'x \mid z \in L\}. \qquad e = [1, 1, \ldots, 1]'$$

a) Ha $\min(-e'x) = 0$, állj! Minden megengedett megoldás optimális, $\varphi(x) = 0$ célfüggvény-értékkel.

b) Ha $\min(-e'x) < 0$, legyen egy optimális bázis-megoldás $\hat{z}^{0\prime} = [\hat{x}^{0\prime}, \hat{y}^{0\prime}]$. Tegyük $z^1 = \hat{z}^0$ és térjünk át a 2. fázisra.

2. fázis. r-ik iteráció

1. lépés: Legyen $r = 1$ és oldjuk meg a következő lineáris programozási feladatot

[1] Ez a feltétel regularizálással [5] mindig teljesíthető.

$$LP(r): \qquad \min \{x'(Cx^r + p) \mid z \in L\}$$

Egy optimális bázismegoldást jelöljön

$$\hat{z}^{r\prime} = [\hat{x}^{r\prime}, \hat{y}^{r\prime}]$$

a) Ha

$$\alpha^r = (x^r - \hat{x}^r)'(Cx^r + p) = 0$$

állj! z^r optimális megoldás.

b) Ha $\alpha^r > 0$, áttérünk a 2. lépésre

2. *lépés:*

a) Ha

$$\beta^r = (\hat{x}^r - x^r)'(C\hat{x}^r + p) \leq 0$$

legyen

$$z^{r+1} = \hat{z}^r$$

b) Ha $\beta^r > 0$, legyen

$$z^{r+1} = \frac{1}{\alpha^r + \beta^r} (\alpha^r \hat{z}^r + \beta^r z^r) .$$

Tegyük r helyébe $r + 1$-et és térjünk vissza az 1. lépésre.

Az algoritmus konvergenciájának bizonyítása

A bizonyítás több lépésből áll. Bebizonyítjuk, hogy az algoritmusban szereplő LP (r) feladatoknak van optimális megoldása (1^0) és hogy az adott megállási szabályok helyesek (2^0, 5^0). Megmutatjuk továbbá, hogy a z^r sorozat megengedett pontokból áll (3^0), úgy hogy $x^r \neq 0$ (4^0) és $\varphi(x^r)$ szigorúan monoton csökken (6^0). Végül bebizonyítjuk, hogy z^{r+1} minimumhely $[\hat{z}^r, z^r]$-en (7^0) és hogy az eljárás konvergens (8^0).

1^0 A QP feladattal kapcsolatban tett *B)* feltevés miatt az $LP(r)$, $r = 0, 1, 2, \ldots$ feladatoknak van optimális bázismegoldása.

2^0 Ha $\min (-e'x) = 0$, akkor minden megengedett z-nek az x részvektora 0. Tehát $\varphi(x) = 0$ minden megengedett pontban. Ezzel igazoltuk az 1. fázis *a)* esetének megállási szabályát. Ellenkező esetben (*b)* eset) olyan $z^1 = \hat{z}^0$ kiinduló megoldáshoz jutottunk, amelyben $x^1 = \hat{x}^0 \neq 0$.

3^0 Könnyen beláthatjuk, hogy a $z^1, z^2, \ldots$ sorozat megengedett pontokból áll. Abból ugyanis, hogy $\hat{z}^r$ optimális megoldása $LP(r)$-nek, következik, hogy $\alpha^r \geq 0$, minden r-re. $z^1 = \hat{z}^0 \in L$ és $\hat{z}^1, \hat{z}^2, \ldots \hat{z}^r$ is megengedettek. Tegyük fel, hogy $z^r \in L$, ekkor z^{r+1} is az, mert ha a 2. fázis 2. lépés *a)* esete szerint képeztük, akkor $\hat{z}^{r+1} = z^r \in L$, ha pedig a *b)* eset szerint képeztük, akkor $\alpha^r > 0$, $\beta^r > 0$, tehát $z^{r+1} \in [\hat{z}^r, z^r] \subset L$. Így teljes indukcióval beláttuk, hogy az algoritmus egy megengedett pontsorozatot generál.

4^0 Bebizonyítjuk továbbá, hogy $x^r \neq 0$, $r = 1, 2, \ldots$ Láttuk, hogy $x^1 \neq 0$. Tegyük fel, hogy $x^s \neq 0$. Ekkor $\hat{x}^s = 0$-ból következnek $\alpha^s = x^{s\prime} (C\, x^s + p) \geq$ ≥ 0. De mivel $C \leq 0$ és $p \leq 0$ (2. Tétel), $\alpha^s \leq 0$, tehát $\alpha^s = 0$. Ekkor z^{s+1} képzésére az 1. lépés *a)* szabálya szerint nem kerül sor. Tehát ha z^{s+1}-et egyáltalán képezzük, akkor $\hat{x}^s \neq 0$, és mivel $z^{s+1} \in [\hat{z}^s, z^s]$, tehát $x^{s+1} \neq 0$. Így teljes indukcióval beláttuk, hogy $x^r \neq 0$, $r = 1, 2, 3 \ldots$

5⁰ Tegyük most fel, hogy

$$\alpha^r = (x^r - \hat{x}^r)'(C\,x^r + p) = 0$$

Mivel $\hat{z}^r$ optimális megoldása $LP(r)$-nek

$$(\hat{x}^r - x)'(C\,x^r + p) \leqq 0 \quad \text{minden} \quad z \in L\text{-re.}$$

E kettő összegéből.

$$(x^r - x)'(C\,x^r + p) \leqq 0 \quad \text{minden} \quad z \in L\text{-re.}$$

Ebből — tekintve, hogy x minden komponense releváns és hogy $x^r \neq 0$, — a 4. korollárium alkalmazásával adódik, hogy z^r optimális megoldása QP-nek. Ezzel igazoltuk a 2. fázis 1. lépés *a)* megállási szabályát.

Eddigi megállapításainkat összefoglalva: az algoritmus vagy véges számú lépésben eljut az optimális megoldáshoz, vagy egy végtelen megengedett pontsorozatot állít elő, amelynek minden z^r tagjára áll: $x^r \neq 0$, és $\alpha^r > 0$.

6⁰ Most bebizonyítjuk, hogy e pontsorozaton haladva a célfüggvény értéke szigorúan monoton csökken. Ugyanis, ha z^{r+1}-et a 2. lépés *a)* szabálya szerint képeztük, akkor (némi számolással)

$$\varphi(x^r) - \varphi(x^{r+1}) = \varphi(x^r) - \varphi(\hat{x}^r) = \frac{1}{2}(\alpha^r - \beta^r) > 0$$

mivel $\alpha^r > 0$, $\beta^r \leqq 0$. Ha pedig *b)* szabályt alkalmaztuk, akkor

$$\varphi(x^r) - \varphi(x^{r+1}) = \varphi(x^r) - \varphi\left(\frac{\alpha^r\,\hat{x}^r + \beta^r\,x^r}{\alpha^r + \beta^r}\right) = \frac{(\alpha^r)^2}{2(\alpha^r + \beta^r)} > 0$$

mivel $\alpha^r > 0$, $\beta^r > 0$. Tehát mindkét esetben

$$\varphi(x^r) > \varphi(x^{r+1}) \tag{20}$$

azaz a z^r sorozaton $\varphi(x^r)$ szigorúan monoton csökken.

7⁰ Továbbá bebizonyítjuk, hogy z^{r+1} minimumhelye a $\varphi(x)$ függvénynek, a $[\hat{z}^r, z^r]$ szakaszon. Tekintsük ugyanis az egyváltozós

$$\psi(\lambda) = \varphi[\lambda\hat{x}^r + (1 - \lambda)\,x^r] \qquad 0 \leqq \lambda \leqq 1$$

függvényt. Ekkor, némi számolással:

$$\frac{d\psi}{d\lambda} = \lambda\beta^r - (1 - \lambda)\,\alpha^r$$

$$\frac{d^2\psi}{d\lambda^2} = \alpha^r + \beta^r\,.$$

Mivel $\alpha^r > 0$, ha $\beta^r \leqq 0$, akkor $\frac{d\psi}{d\lambda} < 0$ a $[0, 1]$ szakaszon. Tehát minimumát a $\lambda = 1$, $z = \hat{z}^r = z^{r+1}$ helyen veszi fel. (2. lépés *a)* eset.) Ha viszont $\beta^r > 0$, akkor $\frac{d^2\psi}{d\lambda^2} > 0$, így $\psi(\lambda)$ a minimumát ott veszi fel, ahol $\frac{d\psi}{d\lambda}$ eltűnik, azaz a $\lambda = \frac{\alpha^r}{\alpha^r + \beta^r}$, $z = z^{r+1}$ helyen. Mindkét esetben

(21) $$\varphi(x^{r+1}) \leq \varphi(x), \text{ ha } z \in [\hat{z}^r, z^r]$$

8° Következik annak bizonyítása,[2] hogy az algoritmus előállít egy, az optimális megoldáshoz konvergáló pontsorozatot. Tekintve, hogy L korlátos, $\varphi(x)$ alulról korlátos L-en, tehát a monoton csökkenő $\varphi(x^r)$ sorozatnak van véges határértéke. Legyen

$$\omega = \lim_{r\to\infty} \varphi(x^r).$$

Ekkor azonban a $\{z^r\}$ sorozatnak van konvergens részsorozata (ennek indexhalmazát jelölje R), úgy, hogy

$$z^r \to \tilde{z} \in L, \qquad \text{ha} \qquad r \in R,\ r \to \infty$$

és
$$\varphi(\tilde{z}) = \omega$$

Bebizonyítjuk, hogy ω a célfüggvény minimuma és $\tilde{z}$ egy optimális megoldás. E célból válasszunk ki R-ből egy olyan $\hat{R}$ részhalmazt, hogy

$$z^r_{\,} \to \tilde{z}, \text{ ha } r \in \hat{R} \text{ és } \hat{z}^r = \bar{z} \text{ minden } r \in \hat{R}\text{-re}.$$

Ilyen $\hat{R}$ részhalmaz létezik, mivel L csúcspontjainak száma véges, tehát legalább egyikük a $\hat{z}^r$ sorozatban végtelen sokszor ismétlődik. Egy ilyet jelöltünk $\bar{z}$-vel.

Következésképp $\bar{z}$ optimális megoldása minden $LP(r)$-nek, ha $r \in \hat{R}$, azaz

$$(x-\bar{x})'(Cx^r + p) \geq 0, \text{ minden } r \in \hat{R},\ z \in L\text{-re},$$

és ha
$$z^r \to \tilde{z}:$$

(22) $$(x-\bar{x})'(C\tilde{x} + p) \geq 0, \qquad \text{minden } z \in L\text{-re}.$$

Legyen most $s, r \in \hat{R}$ és $s > r$. Ekkor (20) miatt

(23) $$\varphi(x^s) \leq \varphi(x^{r+1}) < \varphi(x^r)$$

és (21) miatt

(24) $$\varphi(x^{r+1}) \leq \varphi[\lambda\bar{x} + (1-\lambda)\, x^r], \quad \text{ha} \quad 0 \leq \lambda \leq 1.$$

(23) és (24)-ből:

$$\varphi[\lambda\bar{x} + (1-\lambda)\, x^r] \geq \varphi(x^s), \qquad \text{ha} \quad 0 \leq \lambda \leq 1,$$

és így

$$\frac{\varphi[\lambda\bar{x} + (1-\lambda)\, x^r] - \varphi(x^r)}{\lambda} \geq \frac{\varphi(x^s) - \varphi(x^r)}{\lambda}, \text{ ha } 0 < \lambda \leq 1.$$

Ha most r és s tart a $+\infty$-hez, akkor x^r és x^s tart $\tilde{x}$-hez, tehát

$$\frac{\varphi[\lambda\bar{x} + (1-\lambda)\,\tilde{x}] - \varphi(\tilde{x})}{\lambda} \geq 0, \text{ ha } 0 < \lambda \leq 1.$$

[2] Vö. [3]

$\lambda \to 0$ határátmenettel

$$(\bar{x} - \tilde{x})' \nabla \varphi(\tilde{x}) = (\bar{x} - \tilde{x})' (C\tilde{x} + p) \geq 0 \tag{25}$$

(22) és (25)-ből

$$(x - \tilde{x})'(C\tilde{x} + p) \geq 0 \qquad \text{minden } z \in L\text{-re.}$$

Tehát a 4. korollárium értelmében $\tilde{x}$ optimális megoldása QP-nek, hacsak $\tilde{x} \neq 0$. Ez a feltétel azonban fennáll, hiszen

$$\varphi(\tilde{x}) < \varphi(x^1) \leq 0.$$

3. Két példa

Példa Frank és Wolfe *módszerére*

Minimáljuk a

$$\varphi(x_1, x_2, x_3) = -\frac{1}{2}[(x_1)^2 + 4x_1 x_2 + 14x_1 x_3] - 50x_1$$

függvényt a következő feltételek mellett

$$2x_1 + x_2 + x_3 \leq 16$$
$$x_2 + 2x_3 \leq 12$$
$$x_1, x_2, x_3 \leq 0$$

A célfüggvény nyilván nem konvex, de az 1. tételt alkalmazva belátható, hogy $x_1, x_2, x_3 \geq 0$-ra kvázikonvex. Irreleváns változót nem tartalmaz. A megengedett halmaz korlátos és így a Frank–Wolfe algoritmus alkalmazható. Az eljárás egyes lépéseiben kapott megoldásokat az alábbi táblázat tartalmazza:

r	x_1	x_2	x_3	$\hat{x}_1$	$\hat{x}_2$	$\hat{x}_3$	α	β	$\varphi(x)$
0	—	—	—	2	12	0	—	—	—
1	2	12	0	8	0	0	408	−156	−150
2	8	0	0	5	0	6	162	81	−432
3	6	0	4	8	0	0	0	—	−486

Ez esetben tehát az eljárás véges számú lépésben eljutott az optimális [6, 0, 4] megoldáshoz.

Ellenpélda Wolfe *módszeréhez*

Wolfe [12] véges algoritmust adott a konvex kvadratikus programozási feladat megoldására. Ez azon alapszik, hogy az

$$Ax = b$$
$$Cx + Au + p \geq 0$$
$$x \geq 0, \; u \text{ előjelben nem korlátozott}$$

feltételekkel meghatározott poliédrikus halmaz olyan csúcspontjai, amelyek kielégítik az

$$x'C\,x + p'x + b'u = 0$$

komplementaritási feltételt, egyben megoldásai a

$\min \left\{ \varphi(x) = \frac{1}{2} x'Cx + p'x \mid Ax = b,\; x \geq 0 \right\}$ programozási feladatnak is. Ezért mesterséges változók bevezetésével és a bázisba való belépés szabályainak olyan korlátozásával, amely a komplementaritási feltétel állandó kielégítését biztosítja, szimplex módszerrel megolja a feladatot. Ugyanezek a KUHN–TUCKER-feltételek szükségesek és elégségesek az E^n_+-on kvázikonvex kvadratikus függvény minimálására is (lásd [1] 1. tétel), ha az $x = 0$ pontot kizárjuk. Ennek ellenére Wolfe módszere a kvázikonvex esetben nem biztosítja e feladat megoldását. A továbbiakban a módszert ismertnek tételezzük fel, az eredeti publikációból vesszük a jelölésmódot és a „rövid formát" alkalmazzuk.[3] Az ellenpélda az előbbi példától csak abban különbözik, hogy a célfüggvény lineáris tagját elhagyjuk és a korlátozó feltételekben egyenlőséget írunk elő. Az így keletkező

$$\min \left\{ -\frac{1}{2} [(x_1)^2 + 4x_1x_2 + 14x_1x_3] \mid 2x_1 + x_2 + x_3 + 16;\; x_2 + 2x_3 = 12; x_1, x_2, x_3 \geq 0 \right\}$$

feladat optimális megoldása [5, 0, 6] és a célfüggvény minimuma (−222,5). Az induló táblázat:

bázis	bv. értéke	x_1	x_2	x_3	v_1	v_2	v_3	z_1	z_2	z_3	u_1	u_2	w_1	w_2
w_1	16	2	1	1									1	
w_2	12	0	1	2										1
z_1	0	−1	−2	−7	−1			1			2	0		
z_2	0	−2	0	0		−1			1		1	1		
z_3	0	−7	0	0			−1			1	1	2		

bázis	bv. értéke	x_1	x_2	x_3	v_1	v_2	v_3	z_1	z_2	z_3	u_1	u_2
x_1	2	1	0	−1/2								
x_2	12	0	1	2								
z_1	26	0	0	−7/2	−1			1			2	0
z_2	4	0	0	−1		−1			1		1	1
z_3	14	0	0	−7/2			−1			1	1	2

[3] Ezt lehetővé teszi, hogy példánkban $p = 0$.

Itt u_1 és u_2 előjelben nem korlátozott Lagrange-szorzók, w_1 és w_2 pedig mesterséges változók. Ezeket x_1 és x_2 változókkal felcserélve a következő, megengedett táblát kapjuk:
A két utolsó oszlopot, mint feleslegest elhagytuk és így fogunk tenni a továbbiakban az u_1, u_2 előjelben nem korlátozott (bázis) változóknak megfelelő sorokkal és oszlopokkal is. Most ugyanis $(z_1 + z_2 + z_3)$-at minimáljuk és e célból z_1 ill. z_2 helyett u_1 és u_2 lesznek a bázisváltozók. Így a következő táblázathoz jutunk:

bázis	bv. értéke	x_1	x_2	x_3	v_1	v_2	v_3	z_1	z_2	z_3
x_1	2	1	0	$-1/2$						
x_2	12	0	1	2		O			O	
z_3	19	0	0	$-\frac{13}{4}$	$-\frac{1}{2}$	2	-1	$\frac{1}{2}$	-2	1
	19	0	0	$-\frac{13}{4}$	$-\frac{1}{2}$	2	-1	$-\frac{1}{2}$	-3	0

Itt v_2 a WOLFE-algoritmus szabálya szerint nem vonható be a bázisba, mivel x_2 bázisváltozó, a többi változó bevonása pedig vagy a $\sum z_i$ célfüggvény növekedéséhez vagy meg nem engedett megoldáshoz vezetne. Az algoritmus tehát megáll, mielőtt $\sum z_i = 0$-t, illetőleg az optimális megoldást elértük volna. Az utolsó táblázatban szereplő [2, 12, 0] megoldáshoz tartozó függvényérték: (-50).

(Beérkezett: 1969. *augusztus* 5.*)*

IRODALOM

1. ARROW, K. J.—ENTHOVEN, H. C.: Quasi-concave programming. Econometrica, 1961. 29., 779—800 pp.
2. BEALE, E. M. L.: On minimizing a convex function subject to linear inequalities. Journ. Royal Stat. Soc., 1955. 17., 173—184 pp.
3. BERGE, C.—GHOUILA-HOURI, A.: Programmes, jeux et réseaux de transport. Paris, 1962. Dunod.
4. BOOT, J. C. G.: Quadratic programming. Amsterdam, 1964. North-Holland Publ. Co.
5. CHARNES, A.—COOPER, W. W.: Management models and industrial applications of linear programming. Vol. 1.—2. New York, 1961. Wiley.
6. FRANK, M.—WOLFE, PH.: An algorithm for quadratic programming. Naval Research Logistic Quarterly, 1956. 3., 95—110 pp.
7. KÜNZI, H. P.—KRELLE, W.: Nichtlineare Programmierung. Berlin, 1962. Springer Verlag.
8. MANGASARIAN, O. L.: Pseudo-convex functions. J. SIAM. Control Ser. A. 1965. 3., 281—290 pp.
9. MARTOS, B.: Subdefinite matrices and quadratic forms. SIAM. J. Appl. Math. 1969. 17., (Megjelenés alatt.)
10. PONSTEIN, J.: Seven kinds of convexity. SIAM Review, 1967. 9., 115—119 pp.
11. VAN DE PANNE, C.—WHINSTON, A.: The simplex and the dual method for quadratic programming. Operational Res. Quart. 1964. 15., 355—388. pp.
12. WOLFE, PH.: The simplex method for quadratic programming. Econometrica, 1959. 27., 382—398 pp.

QUADRATIC PROGRAMMING WITH QUASICONVEX OBJECTIVE FUNCTION

The linearly constrained quadratic programming problem can be solved by the method of Frank and Wolfe not only if the objective function is convex but also if it is only quasiconvex in the nonnegative orthant.

In the proof of the convergence of the algorithm (Chap. 2) some properties of the mentioned class of functions find application, these properties are also discussed here at first (Chap. 1). An example illustrates the application of the method and an other shows that Wolfe's quadratic simplex method is inappropriate for this kind of problems. (Chap. 3.)

КВАДРАТИЧНОЕ ПРОГРАММИРОВАНИЕ ПРИ ПОМОЩИ КВАЗИВЫПУКЛОЙ ЦЕЛЕВОЙ ФУНКЦИИ

Задача квадратичного прграммирования с линейными условиями как ограничениями может быть решена методом Франка и Вольфе не только в случае выпуклой целевой функции, но и если целевая функция является квазивыпуклой в неотрицательном орте.

Чтобы доказать конвергенцию алгоритма (глава 2), автор использует некоторые специфические черты рассматриваемого класса функций, излагаемые также впервые в данной статье (глава 1). Применение метода представляется на примере, на другом же примере показывается, что метод квадратичного симплексного программирования Вольфе непригоден для решения такого рода проблем (глава 3).