PSZEUDOLINEÁRIS TÖRTFÜGGVÉNYEKRŐL[1]

KOMLÓSI SÁNDOR
PTE KTK

Martos Béla és Rapcsák Tamás emlékének ajánlom

Martos Béla a hiperbolikus programozásról írt 1960-as cikkében [8] új fejezetet nyitott a matematikai programozás terén, melyet Ő ugyan hiperbolikus programozásnak nevezett, de mivel ez az elnevezés csak a lineáris törtfüggvényekre találó, ezért a nemzetközi szakirodalom a törtprogramozás (fractional programming) elnevezést fogadta el. Martos Béla fedezte fel, hogy a lineáris törtfüggvények rendelkeznek számos olyan jó tulajdonsággal, mint a lineáris függvények, és ezért a hiperbolikus programozás joggal tekinthető a lineáris programozás egy általánosabb változatának. Martos Béla többek között azt is megmutatta, hogy a szimplex módszer hatóköre kiterjeszthető hiperbolikus programozási feladatokra is [8,9]. Ennek oka a lineáris törtfüggvénynek egy olyan tulajdonsága, melyet pszeudolinearitásnak nevezünk. A pszeudo-linearitás a pszeudokonvexitás és pszeudokonkavitás egyidejű fennállását jelenti. A pszeudolinearitással azóta is nagyon sokan foglalkoznak, főleg a törtprogramozásban betöltött szerepe miatt.

Rapcsák Tamás több cikkben foglalkozott a pszeudolinearitás vizsgálatával. Ezen a téren elért legjelentősebb eredménye 1991-ben jelent meg [10], amelyben differenciálgeometriai segédeszközökkel a pszeudolinearitás egy teljesen új jellemzését adta, melynek segítségével további szép eredményeket ért el kvadratikus törtfüggvények pszeudolinearitásának vizsgálatában [11, 12].

Ezzel a dolgozattal, mely kvadratikus törtfüggvények azon osztályának pszeudolinearitást vizsgálja, melyeket Rapcsák Tamás is vizsgált, szeretném Martos Béla és Rapcsák Tamás emléke előtt tisztelegni. Mindketten a közelmúltban hagytak itt bennünket, de gondolataik tovább munkálkodnak bennünk [17, 18].

Bevezetés

Ebben a dolgozatban $f(x)$ mindig n-változós függvényt jelöl, melynek $\nabla f(x)$ gradiensét oszlopvektornak tekintjük. A pszeudokonvexitás fogalmát Mangasarian vezette be differenciálható függvényekre. Ebből származik a pszeudolinearitás definíciója mely szerint $f(x)$ akkor pszeudolineáris, ha egyszerre pszeudokonvex és pszeudokonkáv.

1. Definíció. A differenciálható $f(x)$ függvényt pszeudolineárisnak nevezzük

[1] Beérkezett: 2009. február 15. E-mail: `komlosi@ktk.pte.hu`.

a $K \subseteq I\!\!R^n$ konvex halmazon, ha minden $x, y \in K$ esetén

$$f(x) < f(y) \Rightarrow \nabla f(y)^T(x-y) < 0 \quad \text{és} \quad f(x) > f(y) \Rightarrow \nabla f(y)^T(x-y) > 0 \ . \tag{PLIN}$$

A definíció közvetlen következménye, hogy ha $f(x)$ pszeudolineáris K-n, akkor $f(x)$ vagy konstans K-n (ekkor $\nabla f(x) = 0$ minden $x \in K$-ra), vagy $\nabla f(x) \neq 0$ minden $x \in K$ esetén. Emiatt a továbbiakban csak olyan $f(x)$ differenciálható függvényeket vizsgálunk, melyekre $\nabla f(x) \neq 0$ minden $x \in K$-ra teljesül.

A pszeudolinearitás lokális jellemzése

A pszeudolineáris függvények vizsgálatát egy olyan módszerrel fogjuk végezni, melyben kitüntetett szerepet játszik az implicitfüggvény tétel, illetve az implicit függvények. Mivel az implicit függvények lokális információt adnak a vizsgált függvényről, ezért célszerű bevezetni a lokális pszeudolinearitás fogalmát.

2. Definíció. Az $f(x)$ függvény lokálisan pszeudolineáris x_0-ban, ha van x_0-nak olyan G környezete, hogy

$$\begin{aligned} x \in G, f(x) < f(x_0) &\Rightarrow \nabla f(x_0)(x - x_0) < 0 \ , \\ x \in G, f(x) = f(x_0) &\Rightarrow \nabla f(x_0)(x - x_0) = 0 \ , \\ x \in G, f(x) > f(x_0) &\Rightarrow \nabla f(x_0)(x - x_0) > 0 \ . \end{aligned}$$

Ez a lokális fogalom azért is szerencsés, mert általa a globális tulajdonság is vizsgálható. Igaz ugyanis a következő tétel.

1. Tétel [5, 6]. Legyen $f(x)$ differenciálható a K nyílt konvex halmazon. $f(x)$ akkor és csak akkor pszeudolineáris K-n, ha lokálisan pszeudolineáris K valamennyi pontjában.

Megmutatható, hogy $\nabla f(x_0) \neq 0$ esetén a lokális pszeudolinearitás egyszerűbben is jellemezhető.

2. Tétel [5]. Legyen $f(x)$ differenciálható x_0 valamely környezetében és legyen $\nabla f(x_0) \neq 0$. Ekkor $f(x)$ akkor és csak akkor lokálisan pszeudolineáris x_0-ban, ha van x_0-nak olyan G környezete, hogy

$$\text{valahányszor } x \in G \text{ és } f(x) = f(x_0), \text{ mindannyiszor } \nabla f(x_0)(x - x_0) = 0 \ . \tag{LPLIN}$$

A lokális pszeudolinearitásnak ez az a formája, amely lehetővé teszi az implicit függvényekkel való jellemzést. Legyen $f(x)$ folytonosan differenciálható x_0 valamely környezetében, és legyen $\nabla f(x_0) \neq 0$. Tegyük fel, hogy $f'_{x_n}(x_0) \neq 0$. Vezessük be a következő jelöléseket:

$$(x_1, x_2, \dots, x_{n-1}) = u, \ x_n = v \ \text{ és } x = (u, v) \ .$$

Az implicitfüggvény tétel. A fenti feltételek teljesülése esetén az $\{\, x \in K : f(x) = f(x_0)\,\}$ szintvonal lokálisan parametrizálható x_0 közelében (x_0 egy G környezetében) egy egyértelműen meghatározott $p_{x_0}(u)$ függvény segítségével (mely u_0 egy N környezetén van értelmezve), a következő módon:

$x = (u, v) \in G$ és $f(x) = f(x_0)$ akkor és csak akkor teljesül, ha $v = p_{x_0}(u)$, $u \in N$.

A további vizsgálódások alapját képezi a következő tétel.

3. Lemma [5]. Legyen $f(x)$ folytonosan differenciálható x_0 valamely környezetében, és legyen $\nabla f(x_0) \neq 0$. $f(x)$ akkor és csak akkor pszeudolineáris x_0-ban, ha a $p_{x_0}(u)$ (röviden $p(u)$) implicitfüggvény lineáris N-en.

Bizonyítás. Mivel $f(u, p(u)) \equiv f(x_0)$ N-en, ezért mindkét oldal deriválásával adódik, hogy

$$f'_v(u, p(u))\nabla p(u) + \nabla_u f(u, p(u)) \equiv 0\ .$$

Ebből egyrészt minden $u \in N$-re

$$\nabla p(u) = -\frac{\nabla_u f(u, p(u))}{f'_v(u, p(u))}\ , \tag{1}$$

másrészt minden $x = (u, v) \in G$, $f(x) = f(x_0)$ esetén

$$\nabla f(x) = -f'_v(x) \begin{bmatrix} \nabla p(u) \\ -1 \end{bmatrix}\ , \tag{2}$$

illetve

$$\nabla f(x_0)^T (x - x_0) = f'_v(x_0)\big[p(u) - p(u_0) - \nabla p(u_0)^T (u - u_0)\big]$$

adódik. Ebből már nyilvánvaló, hogy (LPLIN) akkor és csak akkor teljesül, ha $p(u)$ lineáris N-en. □

Ebből a tételből következik a pszeudolinearitás Rapcsák-féle jellemzésének [10] lokális változata.

4. Tétel [5]. Legyen $f(x)$ folytonosan differenciálható x_0 valamely környezetében és legyen $\nabla f(x_0) \neq 0$. $f(x)$ akkor és csak akkor pszeudolineáris x_0-ban, ha van olyan $g \in \mathbb{R}^n$, $g \neq 0$ vektor és x_0 egy G környezetében értelmezett és ott előjeltartó folytonos $c(x)$ függvény, hogy

$$\text{valahányszor } x \in G \text{ és } f(x) = f(x_0), \text{ mindannyiszor } \nabla f(x) = c(x) g\ . \tag{R}$$

Bizonyítás. A 3. Lemma szerint $f(x)$ akkor és csak akkor pszeudolineáris x_0-ban, ha minden $u \in N$-re $\nabla p(u) = \nabla p(u_0) = r \in \mathbb{R}^{n-1}$. (2) szerint ez akkor és csak akkor teljesül, ha minden $x = (u, v) \in G$, $f(x) = f(x_0)$ esetén

$$\nabla f(x) = -f'_v(x) \begin{bmatrix} r \\ -1 \end{bmatrix} = c(x) g$$

teljesül, ahol $g^T = [r^T \ -1]$. □

A következő tétel, amely a 3. Lemma közvetlen következménye, másodrendű szükséges és elégséges feltételt ad lokális pszeudolinearitásra. Ez fogja képezni további vizsgálódásaink alapját.

5. Lemma. Legyen $f(x)$ kétszer folytonosan differenciálható x_0 valamely környezetében és legyen $\nabla f(x_0) \neq 0$. $f(x)$ akkor és csak akkor pszeudolineáris x_0-ban, ha a $p(u) = p_{x_0}(u)$ implicitfüggvény rendelkezik a következő tulajdonsággal: minden $u \in N$-re

$$\nabla^2_{uu} f(u,p(u)) + 2\nabla_u f'_v(u,p(u))\nabla p(u)^T + f''_{vv}(u,p(u))\nabla p(u)\nabla p(u)^T \equiv 0 \,. \quad (3)$$

Bizonyítás. A (3) azonosság, mint majd látni fogjuk, azzal ekvivalens, hogy a $p(u)$ implicitfüggvény $\nabla^2 p(u)$ Hesse mátrixa azonosan 0 N-en. Ez akkor és csak akkor teljesül, ha $p(u)$ lineáris N-en, következésképpen $f(x)$ lokálisan pszeudolineáris x_0-ban.

Deriváljuk kétszer az $f(u,p(u)) \equiv f(x_0)$, $u \in N$ azonosságot. Ebből az

$$f'_v(u,p(u))\nabla^2 p(u) =$$
$$= -\nabla^2_{uu} f(u,p(u)) - 2\nabla_u f'_v(u,p(u))\nabla p(u)^T - f''_{vv}(u,p(u))\nabla p(u)\nabla p(u)^T$$

azonosságot kapjuk, melyből nyilvánvaló, hogy (3) akkor és csak akkor teljesül, ha $\nabla^2 p(u) \equiv 0$ minden $u \in N$-re. □

Egy speciális függvényosztály vizsgálata

A továbbiakban egy többek által vizsgált függvényosztály pszeudolinearitását fogjuk vizsgálni az 5. Lemma (3) feltétele segítségével. Ezt a függvényosztályt vizsgálta Rapcsák Tamás is a [12, 13] cikkekben.

Tekintsük az

$$f(x) = \frac{\frac{1}{2}x^T A x + a^T x + \alpha}{b^T x + \beta}, \quad x \in K \subseteq \mathbb{R}^n \quad (4)$$

kvadratikus törtfüggvényt a $K \subseteq \{\, x \in \mathbb{R}^n : b^T x + \beta > 0 \,\}$ halmazon, ahol $A \neq 0$ n-edrendű szimmetrikus mátrix, $a, b \in \mathbb{R}^n$, $b \neq 0$ és $\alpha, \beta \in \mathbb{R}$.

A (4) függvény tulajdonságai alapvetően függnek az A mátrixtól. A dolgozat további részeiben a következő jelöléseket fogjuk használni. $\mathrm{rang}(A)$ jelöli az A mátrix rangját, $\mathrm{image}(A)$ jelöli az A mátrix képterét, azaz $\mathrm{image}(A) = \{\, Ax : x \in \mathbb{R}^n \,\}$, $\mathrm{Iner}(A)$ pedig az A szimmetrikus mátrix inerciáját jelöli. Az A mátrix inerciája az a rendezett számhármas $(\nu(A), \zeta(A), \pi(A))$, ahol $\nu(A)$, $\zeta(A)$, $\pi(A)$ az A negatív, nulla illetve pozitív sajátértékeinek számát jelöli multiplicitásokkal számolva, vagyis $\mathrm{Iner}(A) = (\nu(A), \zeta(A), \pi(A))$.

R. Cambinitől és L. Carositól származik a következő jellemzés.

6. Tétel [1]. A (4) alatt definiált $f(x)$ függvény akkor és csak akkor pszeudolineáris a $K = \{\, x \in I\!R^n : b^T x + \beta > 0 \,\}$ nyílt konvex halmazon, ha $f(x)$ vagy lineáris, vagy pedig léteznek olyan $\hat{\alpha} \neq 0$ $\hat{\beta}$ és $\hat{\gamma}$ konstansok, hogy $\hat{\alpha}\hat{\gamma} < 0$ és

$$f(x) = \hat{\alpha} b^T x + \hat{\beta} + \frac{\hat{\gamma}}{b^T x + \beta} \,. \tag{5}$$

Később Rapcsák Tamás, egy teljesen más módszert alkalmazva, a következő eredményre jutott.

7. Tétel [12]. A (4) alatt definiált $f(x)$ függvény akkor és csak akkor pszeudolineáris a $K = \{\, x \in I\!R^n : b^T x + \beta > 0 \,\}$ nyílt konvex halmazon, ha $f(x)$ vagy lineáris, vagy pedig léteznek olyan $\hat{\alpha} \neq 0$ és $\hat{\beta}$ konstansok, hogy

$$A = \hat{\alpha} b b^T \quad \text{és} \quad a = \hat{\beta} b \,. \tag{6}$$

A továbbiakban megmutatjuk, hogy ezek az eredmények csupán lokális pszeudolinearitást feltételezve is kiadódnak, vagyis az (5), illetve (6) karakterizációk gyengébb feltételek esetén is fennállnak.

8. Tétel. Tegyük fel, hogy a (4) alatt definiált kvadratikus törtfüggvény lokálisan pszeudolineáris $x_0 \in K$-ban, ahol $\nabla f(x_0) \neq 0$. Az egyszerűség kedvéért tegyük fel, hogy $f'_{x_n}(x_0) \neq 0$. Ekkor létezik olyan $r \in I\!R^{n-1}$ vektor és léteznek olyan λ, $\mu \in I\!R$ skalárok, hogy

$$A = \begin{bmatrix} \lambda r r^T & \mu r \\ \mu r^T & -(\lambda + 2\mu) \end{bmatrix} . \tag{7}$$

Bizonyítás. Tekintsük a

$$\phi(x) = \frac{1}{2} x^T A x + a^T x + \alpha - f(x_0)(b^T x + \beta)$$

segédfüggvényt. Vegyük észre, hogy $f(x) = f(x_0)$ akkor és csak akkor teljesül, ha $\phi(x) = \phi(x_0)$, valamint $\nabla\phi(x_0) \neq 0$ és $\phi'_{x_n}(x_0) \neq 0$. Ennek az a fontos következménye, hogy $\phi(x)$ is lokálisan pszeudolineáris x_0-ban. Tekintsük a $p(u)$ implicitfüggvényt, amely x_0 egy környezetében parametrizálja a $\phi(x) = \phi(x_0)$ szintvonalat. A 3. és 5. Lemmák szerint ekkor minden $u \in N$-re

$$\nabla p(u) = -\frac{\nabla_u \phi(u, p(u))}{\phi'_v(u, p(u))} \equiv r = \text{const} \,, \tag{8}$$

és

$$\nabla^2_{uu}\phi(u, p(u)) \equiv -2\nabla_u \phi'_v(u, p(u))^T \nabla p(u) - \phi''_{vv}(u, p(u)) \nabla p(u)^T \nabla p(u) \,. \tag{9}$$

Használjuk az $x = (u, v)$ dekompozíciót és az A mátrix ennek megfelelő

$$A = \begin{bmatrix} A_{uu} & a_u \\ a_u^T & a_{vv} \end{bmatrix} \tag{10}$$

partícionált alakját. Ekkor igazak a következő összefüggések:

$$\nabla^2_{uu}\phi(u,p(u)) = A_{uu}\,, \quad \nabla_u\phi'_v(u,p(u))\nabla p(u)^T = a_u\nabla p(u)^T = a_u r^T\,,$$

$$\phi''_{vv}(u,p(u))\nabla p(u)\nabla p(u)^T = a_{vv}\nabla p(u)\nabla p(u)^T = a_{vv}rr^T\,.$$

Ezek, a (3) feltétellel együtt, azt adják, hogy

$$A_{uu} = -(2a_u + a_{vv}r)r^T\,. \tag{11}$$

Mivel A_{uu} szimmetrikus, ezért (11) csak úgy teljesülhet, ha $-2a_u - a_{vv}r = \lambda r$ valamely $\lambda \in I\!R$ mellett, következésképpen $A_{uu} = \lambda rr^T$ és $a_u = \mu r$, ahol $\mu = -(\lambda + a_{vv})/2$, és így $a_{vv} = -(\lambda + 2\mu)$. $\square$

Az A mátrix (10) alatti alakja különösen alkalmas arra, hogy az A mátrix inerciáját meghatározzuk. Az inerciaszámítás a Haynsworth féle inercia tétel segítségével könnyen elvégezhető [4]. Ez a tétel feltételezi a Schur-komplemens fogalmát, mely az A mátrix alábbi partícionált alakjával kapcsolatos, ahol P nemszinguláris kvadratikus részmátrix:

$$A = \begin{bmatrix} P & Q \\ Q^T & R \end{bmatrix}\,.$$

A determinánselméletből (Schur-lemma) jól ismert

$$S = R - Q^T P^{-1} Q \tag{12}$$

kvadratikus mátrixot (újabban) a P blokk A-beli Schur-komplemensének nevezzük. A Haynsworth-féle inercia tétel azt állítja, hogy amennyiben a nemszinguláris P mátrix principális (sorindexeinek halmaza megegyezik oszlopindexeinek halmazával), akkor

$$\mathrm{Iner}(A) = \mathrm{Iner}(P) + \mathrm{Iner}(S)\,, \tag{13}$$

ahol az összeadás komponensenként végzendő. A P blokk principalitása biztosítja azt, hogy mind a P, mind pedig az S mátrix szimmetrikus.

A Schur-komplemens fontos szerepet játszik minden pivot algoritmusban, ezért a (10) képlet alapján bármely szimmetrikus mátrix inerciája kiszámítható speciális pivot elem választási szabályon alapuló pivot transzformációk véges sorozatával. Ezt a numerikus módszert R. W. Cottle publikálta [2]-ben. Erről részletes leírás található [7]-ben.

9. Tétel. Tegyük fel, hogy a (4) alatt definiált kvadratikus törtfüggvény lokálisan pszeudolineáris $x_0 \in K$ -ban, ahol $\nabla f(x_0) \neq 0$. Ha $r \neq 0$, akkor igazak a következő állítások:

(i) Ha $\lambda + \mu \neq 0$, akkor $\mathrm{Iner}(A) = (1, n-2, 1)$.

(iia) Ha $\lambda + \mu = 0$ és $\lambda + 2\mu > 0$, akkor $\mathrm{Iner}(A) = (1, n-1, 0)$.

(iib) Ha $\lambda + \mu = 0$ és $\lambda + 2\mu < 0$, akkor $\mathrm{Iner}(A) = (0, n-1, 1)$.

Ha $r = 0$, akkor $\mathrm{Iner}(A) = (1, n-1, 0)$, ha $\lambda + 2\mu > 0$ és $\mathrm{Iner}(A) = (0, n-1, 1)$, ha $\lambda + 2\mu < 0$.

Bizonyítás. (ia) Vizsgáljuk először azt az esetet, amikor $\lambda + 2\mu \neq 0$, és legyen $P = [-(\lambda + 2\mu)]$ a (13) inercia-formulában. A (12) Schur-komplemens formula szerint

$$S = \left[\frac{(\lambda + \mu)^2}{\lambda + 2\mu} rr^T\right] .$$

$r \neq 0$ miatt $\mathrm{Iner}(rr^T) = (0, n-2, 1)$, és így

$$\mathrm{Iner}(S) = \mathrm{Iner}\left[\frac{(\lambda + \mu)^2}{\lambda + 2\mu}\right] + (0, n-2, 0) .$$

Mivel $\mathrm{Iner}(A) = \mathrm{Iner}(P) + \mathrm{Iner}(S) = \mathrm{Iner}\,[-(\lambda + 2\mu)] + \mathrm{Iner}\left[\frac{(\lambda+\mu)^2}{\lambda+2\mu}\right] + (0, n-2, 0)$, ezért az (ia) állítás bizonyítást nyert.

(ib) Vizsgáljuk most a $\lambda + 2\mu = 0$ esetet. Ekkor szükségképpen $\lambda \neq 0$. Legyen $P_1 = \left[\lambda r_i^2\right]$, ahol $r_i \neq 0$ valamely i-re. Az egyszerűség kedvéért tegyük fel, hogy $i = 1$. Ekkor (12) a következőt adja:

$$S_1 = \begin{bmatrix} 0 & 0 \\ 0^T & -\lambda/4 \end{bmatrix} .$$

Válasszuk S_1-ben a $P_2 = [-\lambda/4]$ blokkot. Mivel P_2-nek S_1-ben az $(n-2)$-ed rendű nullmátrix a Schur-komplemense, ezért

$$\mathrm{Iner}(A) = \mathrm{Iner}\left[\lambda r_i^2\right] + \mathrm{Iner}\,[-\lambda/4] + (0, n-2, 0) = (1, n-2, 1) .$$

(ii) Legyen $P = [-(\lambda + 2\mu)]$. (12) szerint, $\lambda + \mu = 0$ miatt

$$S = \left[\frac{(\lambda + \mu)^2}{\lambda + 2\mu} rr^T\right] = 0 .$$

Mivel $\mathrm{Iner}(A) = \mathrm{Iner}(P) + \mathrm{Iner}(S) = \mathrm{Iner}\,[-(\lambda + 2\mu)] + (0, n-1, 0)$, ezért (ii) bizonyítást nyert.

Az A mátrix (7) alakja miatt az $r = 0$ eset bizonyítása a (ii) állítás bizonyításával megegyező. □

A következő tétel a lokális pszeudolinearitás további fontos következményeit fogalmazza meg.

10. Tétel. Tegyük fel, hogy a (4) alatt definiált kvadratikus törtfüggvény lokálisan pszeudolineáris $x_0 \in K$-ban, ahol $\nabla f(x_0) \neq 0$. Legyen $g^T = \left[r^T \ -1\right]$. Ekkor igazak a következő állítások:

(i) $\mathrm{rang}(A) = 1$ vagy 2.

(ii) Ha $\mathrm{rang}(A) = 1$, akkor $\nabla f(x_0), a - f(x_0)b \in \mathrm{Lin}\{g\} = \mathrm{image}(A)$.

(iii) Ha $\mathrm{rang}(A) = 2$, akkor $\nabla f(x_0), a - f(x_0)b \in \mathrm{Lin}\{g\} \subseteq \mathrm{Lin}\{g, Ag\} = \mathrm{image}(A)$.

(iv) Ha $f(x)$ lokálisan pszeudolineáris egy x_0-tól különböző x_1 pontban is, ahol $\nabla f(x_1) \neq 0$ és $f(x_1) \neq f(x_0)$, akkor $a, b \in \mathrm{image}(A)$.

(v) Ha $Ax = \nabla f(x_0)$, akkor $\nabla f(x_0)^T x = 0$ és, ha $\nabla f(x_0)^T x = 0$, akkor $x^T Ax = 0$.

Bizonyítás. (i) Mivel $\mathrm{rang}(A) = \nu(A)+\pi(A)$, ezért a 9. Tételből következik, hogy $\mathrm{rang}(A) = 1$ vagy 2.

(ii) A 9. Tétel szerint, $\mathrm{rang}(A) = 1$ ekvivalens a következő állítással: vagy $r \neq 0$ és $\lambda + \mu = 0$, vagy pedig $r = 0$. A 8. Tétel szerint ekkor

$$A = \kappa \begin{bmatrix} rr^T & -r \\ -\,r^T & 1 \end{bmatrix} = \kappa \begin{bmatrix} r \\ -\,1 \end{bmatrix} [\, r^T \quad -1\,] \ ,$$

ahol az első esetben $\kappa = \lambda$, a másodikban pedig $\kappa = -(\lambda + 2\mu)$. Ebből közvetlenül adódik, hogy $A = \kappa gg^T$, és $\mathrm{image}(A) = \mathrm{Lin}\{g\} = \{tg : t \in I\!R\}$. A (8) összefüggés szerint

$$\nabla\phi(x_0)^T = -\phi_v'(x_0)\,[\, r^T \quad -1\,] = -\phi_v'(x_0) g^T \ ,$$

ahol $\phi_v'(x_0) \neq 0$. A $\phi(x)$ függvény definícióját is figyelembe véve, igaz a következő:

$$\nabla f(x_0) = \frac{1}{b^T x_0 + \beta} \nabla\phi(x_0) = -\frac{\phi_v'(x_0)}{b^T x_0 + \beta} g \ ,$$

és ezért $\nabla\phi(x_0), \nabla f(x_0) \in \mathrm{Lin}\{g\}$. Másfelől viszont $\nabla\phi(x_0) - Ax_0 = a - f(x_0)b$, és ezért $a - f(x_0)b \in \mathrm{Lin}\{g\} = \mathrm{image}(A)$.

(iii) A 9. Tétel szerint $\mathrm{rang}(A) = 2$ azzal ekvivalens, hogy $r \neq 0$ és $\lambda + \mu \neq 0$. Először azt mutatjuk meg, hogy $g \in \mathrm{image}(A)$, vagyis az $Ax = g$ egyenlet megoldható x-ben. Tekintsük az x és g vektorokat, valamint az A mátrixot az (u, v) felbontásukban, ahol $u \in I\!R^{n-1}$ és $v \in I\!R$. Ekkor

$$Ax = \begin{bmatrix} \lambda rr^T & \mu r \\ \mu r^T & -(\lambda + 2\mu) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \lambda(r^T u)r + \mu vr \\ \mu(r^T u) - (\lambda + 2\mu)v \end{bmatrix} .$$

Ebből a felbontásból nyilvánvaló, hogy az $Ax = g$ egyenlet akkor és csak akkor teljesül, ha az x vektor u és v komponensei kielégítik a következő, $r^T u$-ban és v-ben lineáris egyenletrendszert:

$$\begin{aligned} \lambda(r^T u) + \mu v &= 1 \ , \\ \mu(r^T u) - (\lambda + 2\mu)v &= -1 \ . \end{aligned} \tag{14}$$

Mivel az egyenletrendszer determinánsa $\lambda + \mu \neq 0$, ezért (14)-nek egyetlen megoldása van:

$$r^T u = v = \frac{1}{\lambda + \mu} \ , \tag{15}$$

amiből $g \in \mathrm{image}(A)$ adódik.

Most megmutatjuk, hogy g és Ag lineárisan függetlenek. Első lépésben azt mutatjuk meg, hogy $Ag \neq 0$. Indirekt módon okoskodva, tegyük fel, hogy $Ag = 0$. Ez azt jelenti, hogy

$$Ag = \begin{bmatrix} \lambda rr^T & \mu r \\ \mu r^T & -(\lambda + 2\mu) \end{bmatrix} \begin{bmatrix} r \\ -\,1 \end{bmatrix} = \begin{bmatrix} \lambda(r^T r)r - \mu r \\ \mu(r^T r) + (\lambda + 2\mu) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

Ez azonban akkor és csak akkor teljesül, ha

$$\lambda r^T r - \mu = 0 \,,$$
$$\mu r^T r + \lambda + 2\mu = 0 \,,$$

és ennek következtében $\lambda\mu r^T r = \mu^2 = -\lambda^2 - 2\lambda\mu$. Ez azonban ekvivalens a $(\lambda+\mu)^2 = 0$ feltétellel, ami $\lambda + \mu \neq 0$ miatt lehetetlen. Ez az ellentmondás bizonyítja állításunkat.

A következő lépésben megmutatjuk, hogy g és Ag lineárisan függetlenek. Indirekt módon okoskodva tegyük fel, hogy $Ag = \gamma g$ valamely $\gamma \neq 0$ valós számmal. Legyen $\delta = 1/\gamma \neq 0$. Feltevésünk szerint ekkor $A(\delta g) = g$ kell, hogy teljesüljön. Ez azonban azt jelenti, hogy az $x = \delta g = (u, v)$ vektornak teljesítenie kell a (15) egyenletet, ami ekvivalens azzal, hogy $g^T x = g^T(\delta g) = \delta g^T g = 0$, ami azonban $\delta \neq 0$ és $g \neq 0$ miatt lehetetlen. Ez bizonyítja azt, hogy g és Ag lineárisan függetlenek. Mivel $g, Ag \in \text{image}(A)$ és $\text{rang}(A) = 2$, ezért $\text{image}(A) = \text{Lin}\{g, Ag\}$, ahogy azt állítottuk.

A (ii) állítás bizonyítása során alkalmazott gondolatmenet ebben az esetben is azt eredményezi, hogy $\nabla f(x_0), a - f(x_0)b \in \text{Lin}\{g\} \subseteq \text{Lin}\{g, Ag\} = \text{image}(A)$.

(iv) Tekintsük most az $x_1 \in K$ pontot, melyre $f(x_1) \neq f(x_0)$, $\nabla f(x_1) \neq 0$ teljesül, és amelyben $f(x)$ ugyancsak lokálisan pszeudolineáris. (ii) és (iii) szerint ekkor

$$\nabla\phi(x_i) - Ax_i = a - f(x_i)b \in \text{image}(A)$$

$i = 0, 1$ esetén. Mivel $f(x_0) \neq f(x_1)$, ezért nyilvánvaló, hogy $a, b \in \text{image}(A)$.

(v) Megmutatjuk, hogy, ha $Ax = \nabla f(x_0)$, akkor $\nabla f(x_0)^T x = 0$ és, ha $\nabla f(x_0)^T x = 0$, akkor $x^T Ax = 0$.

Nézzük először azt az esetet, amikor $\text{rang}(A) = 1$. Mivel ekkor $A = \kappa g g^T$ és $\nabla f(x_0) = \gamma g$, ahol $\kappa, \gamma \neq 0$, ezért mindkét állítás nyilvánvaló.

Nézzük most azt az esetet, amikor rang(A) $= 2$. Legyen $x = (u, v)$ megoldása az $Ax = g$ egyenletnek. Ekkor (15) szükségképpen teljesül, ami pontosan a $g^T x = r^T u - v = 0$ feltételt szolgáltatja. Mivel $\nabla f(x_0) = \gamma g$, ezért, ha $Ax = \nabla f(x_0)$, akkor $\nabla f(x_0)^T x = 0$. Tegyük most fel, hogy $\nabla f(x_0)^T x = 0$, amely ekvivalens azzal, hogy $g^T x = r^T u - v = 0$. (15) szerint ekkor $Ax = (\lambda + \mu)g$ és így $x^T Ax = (\lambda+\mu) g^T x = 0$. □

A következő tétel talán túlságosan technikainak tűnik, de fontos szerepe lesz a dolgozat fő eredményeinek bizonyításában.

11. Segédtétel. Tegyük fel, hogy a (4) alatt definiált kvadratikus törtfüggvény lokálisan pszeudolineáris $x_0 \in K$-ban, ahol $\nabla f(x_0) \neq 0$. Tegyük fel, hogy $\text{rang}(A) = 2$. Ekkor léteznek olyan $\hat{a}, \hat{b} \in \mathbb{R}^n$ vektorok, hogy

$$a = A\hat{a} \quad \text{és} \quad b = A\hat{b} \,, \tag{16}$$

és $\lambda = f(x_0)$ kielégíti a következő egyenletet:

$$b^T \hat{b} \lambda^2 + 2(\beta - a^T \hat{b})\lambda + a^T \hat{a} - 2\alpha = 0 \,. \tag{17}$$

Ha $f(x)$ lokálisan pszeudolineáris legalább három olyan pontban, ahol $f(x)$ gradiense nem tűnik el, és a függvényértékek különbözőek, akkor az $\hat{a}$, $\hat{b}$ vektorokra teljesülniük kell a következő feltételeknek:

$$a = A\hat{a}\,, \quad b = A\hat{b}\,, \quad b^T\hat{b} = 0\,, \quad a^T\hat{b} = \beta \quad \text{és} \quad a^T\hat{a} = 2\alpha\,. \tag{18}$$

Bizonyítás. Mivel $\operatorname{rang}(A) = 2$ akkor és csak akkor teljesül, ha $\operatorname{Iner}(A) = (1, n-2, 1)$, ezért a 10. Tétel (ii) és (iii) állításaiból következik (16)-nak eleget tevő $\hat{a}, \hat{b} \in I\!R^n$ vektorok létezése. (16) felhasználásával kapjuk a következőt:

$$\nabla f(x_0) = \frac{Ax_0 + a - f(x_0)b}{b^T x_0 + \beta} = \frac{Ax_0 + A\hat{a} - f(x_0)A\hat{b}}{b^T x_0 + \beta} = \frac{A(x_0 + \hat{a} - f(x_0)\hat{b})}{b^T x_0 + \beta}\,.$$

A 10. Tétel szerint, ha $Ax = \nabla f(x_0)$, akkor $\nabla f(x_0)^T x = 0$, ezért

$$\big(Ax_0 + a - f(x_0)b\big)^T \big(x_0 + \hat{a} - f(x_0)\hat{b}\big) = 0\,.$$

Elvégezve a beszorzást és felhasználva az $a = A\hat{a}$, $b = A\hat{b}$ összefüggéseket, azt kapjuk, hogy

$$b^T\hat{b}f^2(x_0) - (2b^T x_0 + 2b^T\hat{a})f(x_0) + x_0^T Ax_0 + 2a^T x_0 + a^T\hat{a} = 0\,.$$

Végrehajtva az

$$x_0^T Ax_0 + 2a^T x_0 = 2f(x_0)(b^T x_0 + \beta) - 2\alpha$$

helyettesítést, a

$$b^T\hat{b}f(x_0)^2 + 2(\beta - a^T\hat{b})f(x_0) + a^T\hat{a} - 2\alpha = 0$$

egyenletre jutunk, amely pontosan azt adja, hogy $\lambda = f(x_0)$ kielégíti a (17) egyenletet. Ha $f(x)$ lokálisan pszeudolineáris legalább három olyan pontban, ahol $f(x)$ gradiense nem tűnik el, és a függvényértékek különbözőek, akkor a (17) egyenletnek legalább három különböző megoldása van, ami csak úgy lehetséges, hogy valamennyi együttható 0-val egyenlő. □

A következő tétel tisztán mátrixalgebrai, de a további vizsgálatokban szükségünk lesz rá.

12. Segédtétel. Legyen az A szimmetrikus mátrix inerciája $(1, n-2, 1)$. Ekkor minden olyan $d \in I\!R^n$, $d \neq 0$ vektorhoz, mely rendelkezik azzal a tulajdonsággal, hogy $d = A\hat{d}$ és $d^T\hat{d} = 0$ teljesül valamely $\hat{d} \in I\!R^n$ vektorral, található egyetlen olyan $p \in I\!R^n$ vektor, hogy

$$A = pd^T + dp^T\,,$$

továbbá p rendelkezik a következő tulajdonsággal: $p = A\hat{p}$ és $p^T\hat{p} = 0$ teljesül valamely p vektorral, és $p^T\hat{d} = d^T\hat{p} = 1$.

Bizonyítás. Legyen $s_1, s_2, s_3, \ldots, s_n$ az A sajátvektoraiból álló ortonormált bázis $I\!R^n$-ben. Tegyük fel, hogy $As_1 = \sigma_1$, $\sigma_1 < 0$ és $As_2 = \sigma_2$, $\sigma_2 > 0$. A feltevések alapján nyilvánvaló, hogy $\mathrm{image}(A) = \mathrm{Lin}\{\, s_1, s_2 \,\}$. Rendelkezzen a $d \in I\!R^n$ vektor a megkívánt tulajdonságokkal. Ekkor $d \in \mathrm{image}(A) = \mathrm{Lin}\{\, s_1, s_2 \,\}$, és ezért d következő előállítása egyértelmű:

$$d = \delta_1 s_1 + \delta_2 s_2 \,.$$

Mivel $d = A\hat{d}$ és $d^T\hat{d} = 0$, ezért az általánosság megsértése nélkül feltehetjük, hogy $\hat{d} \in \mathrm{Lin}\{\, s_1, s_2 \,\}$, következésképpen

$$\hat{d} = \frac{\delta_1}{\sigma_1} s_1 + \frac{\delta_2}{\sigma_2} s_2 \,,$$

és ennélfogva

$$d^T\hat{d} = \frac{\delta_1^2}{\sigma_1} + \frac{\delta_2^2}{\sigma_2} = 0 \,.$$

Mivel $d \neq 0$, ezért ez csak úgy lehet, ha $\delta_1 \neq 0$ és $\delta_2 \neq 0$. Legyen

$$p = \frac{\sigma_1}{2\delta_1} s_1 + \frac{\sigma_2}{2\delta_2} s_2 \quad \text{és} \quad \hat{p} = \frac{1}{2\delta_1} s_1 + \frac{1}{2\delta_2} s_2 \,.$$

Könnyen meggyőződhetünk arról, hogy $p = A\hat{p}$,

$$p^T\hat{p} = \frac{\sigma_1}{4\delta_1^2} + \frac{\sigma_2}{4\delta_2^2} = \frac{4\sigma_1\sigma_2}{\delta_1^2\delta_2^2}\left(\frac{\delta_1^2}{\sigma_1} + \frac{\delta_2^2}{\sigma_2}\right) = \frac{4\sigma_1\sigma_2}{\delta_1^2\delta_2^2} d^T\hat{d} = 0$$

és

$$d^T\hat{p} = \delta_1 \frac{1}{2\delta_1} + \delta_2 \frac{1}{2\delta_2} = 1 \quad \text{és} \quad p^T\hat{d} = \frac{\sigma_1}{2\delta_1}\frac{\delta_1}{\sigma_1} + \frac{\sigma_2}{2\delta_2}\frac{\delta_2}{\sigma_2} = 1 \,.$$

Most megmutatjuk, hogy $A = pd^T + dp^T$. Tekintsük az $S = [s_1, s_2, \ldots, s_n]$ bázis mátrixot. S konstrukciója folytán $S^{-1} = S^T$, és így $A = pd^T + dp^T$ akkor és csak akkor áll fenn, ha

$$S^T A S = (S^T p)(d^T S) + (S^T d)(p^T S) \,.$$

Írjuk mindkét mátrixot azonosan particionálva, a következő módon:

$$S^T A S = \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}, \qquad \text{ahol} \quad P = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

és

$$(S^T p)(d^T S) + (S^T d)(p^T S) = \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix},$$

ahol

$$Q = \begin{pmatrix} 2\dfrac{\sigma_1}{2\delta_1}\delta_1 & \dfrac{\sigma_1}{2\delta_1}\delta_2 + \dfrac{\sigma_2}{2\delta_2}\delta_1 \\ \dfrac{\sigma_1}{2\delta_1}\delta_2 + \dfrac{\sigma_2}{2\delta_2}\delta_1 & 2\dfrac{\sigma_2}{2\delta_2}\delta_2 \end{pmatrix} .$$

Nyilvánvaló, hogy $P = Q$ akkor és csak akkor teljesül, ha

$$\frac{\sigma_1}{2\delta_1}\delta_2 + \frac{\sigma_2}{2\delta_2}\delta_1 = 0 \,.$$

Vegyük észre, hogy

$$\frac{\sigma_1}{2\delta_1}\delta_2 + \frac{\sigma_2}{2\delta_2}\delta_1 = \frac{\sigma_1\sigma_2}{2\delta_1\delta_2}\left(\frac{\delta_1^2}{\sigma_1} + \frac{\delta_2^2}{\sigma_2}\right) = \frac{\sigma_1\sigma_2}{2\delta_1\delta_2} d^T\hat{d} = 0 \ ,$$

ami azt bizonyítja, hogy $P = Q$ és $A = pd^T + dp^T$. □

Most bebizonyítjuk a (4) típusú kvadratikus törtfüggvények Cambini-Carosi, illetve Rapcsák féle karakterizációját (lásd 6. és 7. Tételek) a globális pszeudolinearitás feltételezését lokálisra gyengítve. A bizonyítást két lépésben végezzük el, annak megfelelően, hogy a (4) függvényben szereplő A mátrixnak 1 vagy 2 a rangja.

13. Tétel. Legyen $f(x)$ (4) típusú kvadratikus törtfüggvény, és az A mátrix rangja legyen 1. Ekkor $f(x)$ akkor és csak akkor lokálisan pszeudolineáris $x_0 \in K$-ban, ahol $\nabla f(x_0) \neq 0$, ha léteznek olyan $\tilde{\alpha} \neq 0$ és $\tilde{\beta}$, valamint $\hat{\alpha} \neq 0$, $\hat{\beta}$ és $\hat{\gamma}$ konstansok, melyekre fennállnak a következő összefüggések:

$$A = \tilde{\alpha} bb^T \quad \text{és} \quad a = \tilde{\beta} b \ , \tag{19}$$

valamint

$$f(x) = \hat{\alpha} b^T x + \hat{\beta} + \frac{\hat{\gamma}}{b^T x + \beta} \ , \tag{20}$$

illetve

$$\hat{\alpha} \neq \hat{\gamma}(b^T x_0 + \beta)^2 \ . \tag{21}$$

Bizonyítás. *Szükségesség.* A 10. Tétel (ii) állításának közvetlen következménye olyan $\tilde{\alpha} \neq 0$, $\tilde{\beta}$ és $\tilde{\gamma}$ konstansok létezése, melyekkel (19) fennáll. Ezeket az összefüggéseket (4)-be helyettesítve megkapjuk (20)-at, ahol

$$\hat{\alpha} = \frac{\tilde{\alpha}}{2} \ , \quad \hat{\beta} = \tilde{\beta} - \beta\frac{\tilde{\alpha}}{2} \quad \text{és} \quad \hat{\gamma} = \alpha - \beta\tilde{\beta} + \beta^2\frac{\tilde{\alpha}}{2} \ .$$

(20)-ból egyszerű számolással adódik, hogy

$$\nabla f(x) = \left(\hat{\alpha} - \frac{\hat{\gamma}}{(b^T x + \beta)^2}\right) b \ . \tag{22}$$

A $\nabla f(x_0) \neq 0$ feltétel csak úgy teljesülhet, ha (21) fennáll.

Elegendőség. Ha (21) teljesül, akkor folytonossági okokból a (20) függvény (22) gradiense x_0 egy egész környezetén különbözik 0-tól, ennélfogva teljesül a 4. Tétel Rapcsák-féle (R) feltétele, amely elegendő az x_0-beli lokális pszeudolinearitáshoz. □

14. Tétel. Legyen $f(x)$ (4) típusú kvadratikus törtfüggvény, és legyen az A mátrix rangja 2. Tegyük fel, hogy $f(x)$ lokálisan pszeudolineáris legalább három olyan különböző pontban, ahol $f(x)$ gradiense nem tűnik el, és a függvényértékek különbözőek. Ekkor létezik olyan $p \in \mathbb{R}^n$ vektor és $\pi \in \mathbb{R}$ skalár, hogy

$$f(x) = p^T x + \pi \ . \tag{23}$$

Bizonyítás. A $\mathrm{rang}(A) = 2$ feltétel esetünkben azzal ekvivalens, hogy $\mathrm{Iner}(A) = (1, n-2, 1)$. A 11. Segédtétel szerint ekkor léteznek olyan $\hat{a}$, $\hat{b} \in \mathbb{R}^n$ vektorok, melyekkel teljesül (18). A 12. Segédtétel szerint ekkor létezik olyan $p \in \mathbb{R}^n$ vektor, hogy

$$A = bp^T + pb^T \ .$$

A-nak ez a felbontása a (16) és (18) feltételekkel együtt azt adja, hogy

$$a = A\hat{a} = (p^T\hat{a})b + (b^T\hat{a})p = \pi b + \beta p$$

és

$$\alpha = \frac{1}{2}a^T\hat{a} = \frac{1}{2}\left(\pi b^T\hat{a} + \beta p^T\hat{a}\right) = \frac{1}{2}(\pi\beta + \beta\pi) = \pi\beta \ ,$$

ahol $\pi = p^T\hat{a}$. Ezeket az összefüggéseket (4)-be helyettesítve megkapjuk (23)-at. □

Irodalom

1. R. Cambini and L. Carosi, On generalized linearity of quadratic functions, *Journal of Global Optimization*, 30 (2004) 235–251.
2. R. W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.*, 8 (1974) 189–211.
3. Frenk, J. B. G. and S. Schaible, Fractional programming, *Handbook of Generalized Convexity and Generalized Monotonicity*, Chapter 8, eds. Hadjisavvas, N., Komlosi, S. and S. Schaible, Springer-Verlag, Berlin-Heidelberg-New York, 2005, 335–386.
4. E. V. Haynsworth, Determination of the inertia of a partitioned hermitian matrix, *Linear Algebra Appl.*, 1 (1968) 73-81.
5. S. Komlósi, First and second order characterization of pseudolinear functions, *European Journal of Operational Research*, 67 (1993) 278–286.
6. S. Komlósi, On pseudoconvex functions, *Acta. Sci. Math.* (Szeged), 57 (1993) 569–586.
7. Komlósi, S. *Az optimalizáláselmélet alapja*, Dialóg Campus Kiadó, Budapest-Pécs, 2001.
8. Martos, B. Hiperbolikus programozás, *MTA Matematikai Kutató Intézet Közleményei*, 5 (1960) 383-406.
9. B. Martos, Hyperbolic Programming, *Naval Research Logistic Quarterly*, 11 (1964) 135–155.
10. B. Martos, *Nonlinear Programming: Theory and Methods*, North Holland, Amsterdam, 1975.
11. T. Rapcsák, On pseudolinear functions, *European Journal of Operational Research*, 50 (1991) 353–360.
12. T. Rapcsák, On pseudolinearity of quadratic fractional programming, *Optimization Letters*, 1 (2007) 193-200.
13. T. Rapcsák and M. Újvári, Some results on pseudolinear quadratic functions, *CEJOR*, 16 (2008) 415–424.

14. S. Schaible and T. Ibaraki, Fractional programming, *European Journal of Operational Research*, 12 (1983) 325–338.
15. S. Schaible, Fractional Programming, in: *Handbook of Global Optimization*, R. Horst and P. Pardalos (eds.), Kluwer Academic Publishers, Dordrecht, 495–608.
16. M. Újvári, Simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope, *Pu.M.A.*, 18 (2007) 189–202.
17. Simonovits, A., Martos Béla (1920-2007), *Szigma*, 38 (2007) 75–77.
18. Rapcsák Tamás (1947-2008), (szerkesztőségi megemlékezés) *Alkalmazott Matematikai Lapok*, 26 (2009) 1–14.

ON PSEUDOLINEAR FRACTIONAL FUNCTIONS

The aim of the paper is to present a local approach called "the implicit function approach" in characterizing pseudolinearity in case of quadratic fractional functions. The class of functions investigating in the paper has already been intensively studied by several authors, including Cambini-Carosi and Rapcsák. It is shown in the paper that results known from the literature can be obtained under weaker hypothesis (local pseudolinearity) on the fractional function under consideration.