Komlósi Sándor

Matematikai programozási feladatok optimalitási kritériumairól

I. Bevezetés

A matematikai programozás viszonylag fiatal ága a gyakorlati (és ezen belül a gazdasági) problémák megoldása során sikerrel és hatékonyan alkalmazható matematikai elméleteknek. Alapfeladatát a következő feltételes szélsőérték-feladat formájában szokták megfogalmazni:

$$(1)\qquad \begin{aligned} &f(x) \to \max \\ &g_i(x) \leq 0,\ i = 1, \ldots, m \end{aligned}$$

ahol $x \in R^n$ (R^n az n-dimenziós euklideszi teret jelöli), az $f(x)$ célfüggvény és a $g_i(x)$ feltételi függvények tetszőleges számértékű függvények. Az (1) feladat legegyszerűbb közgazdasági interpretációja a következő: az $f(x)$ hozamfüggvény maximumát keressük az erőforrások korlátozottságát kifejező $g_i(x)$ függvények által meghatározott

$$L = \{x \in R^n \,|\, g_i(x) \leq 0,\ i = 1, \ldots, m\}$$

lehetséges programok halmazán. (lsd. pl. [3, 9. old.])

Az (1) feladat megoldása abban áll, hogy meghatározzuk az összes olyan $x_0 \in L$ lehetséges programot, amelyekre teljesül a következő feltétel:

$$(2)\qquad f(x_0) \geq f(x),\ \text{ha}\ x \in L.$$

A (2) feltételnek eleget tevő x_0 lehetséges programot az (1) feladat *optimális megoldás*ának, *optimumpont*jának nevezzük.

Ha a (2) feltétel az x_0 lehetséges program valamely G környezetében teljesül csupán, vagyis:

$$(3)\qquad f(x_0) \geq f(x),\ \text{ha}\ x \in L \cap G,$$

akkor az x_0 programot az (1) feladat *lokálisan optimális megoldás*ának, *lokális optimumpont*jának nevezzük.

Ha a (2), ill. (3) feltételekben $x \neq x_0$ esetén az egyenlőtlenségek szigorú formában teljesülnek, akkor az x_0 lehetséges programot az (1) feladat *szigorú optimumpont*jának nevezzük.

Az (1) feladattal kapcsolatban a legfontosabb elméleti és gyakorlati kérdés az optimális, ill. lokálisan optimális programok meghatározása, előállítása. Az e célhoz vezető úton az első jelentősebb lépés az *optimalitás szükséges feltételei*nek kidolgozása.

Differenciálható cél és feltételi függvények esetén jól ismertek az (1) feladatra vonatkozó *Kuhn-Tucker-Lagrange feltételek* (a továbbiakban, röviden, KTL-feltételek), amelyek meglehetősen tág feltételek mellett az (1) feladat lokálisan optimális x_0 megoldására adnak szükséges feltételeket [1, 3, 8, 12]:

Léteznek olyan $t_1, \ldots, t_m$ nem-negatív valós számok (ún. Lagrange-multiplikátorok), hogy

$$\nabla f(x_0) = \sum_{i=1}^{m} t_i \nabla g_i(x_0) \tag{4}$$

és

$$t_i g_i(x_0) = 0, \quad i = 1, \ldots, m. \tag{5}$$

A $\nabla f(x)$ szimbólum az $f(x)$ függvény gradiens-vektorát jelöli, melyet kényelmi okokból sorvektornak tekintünk. Az R^n tér elemeit viszont — szokás szerint — oszlopvektornak tekintjük.

A (4)—(5) KTL-feltételeknek eleget tevő lehetséges programot a szóban forgó feladat *KTL-stacionárius pont*jának nevezzük [12, Chapter 8., Definition 5]. Annak eldöntése, hogy egy KTL-stacionárius pont optimális vagy lokálisan optimális megoldása-e a tekintett feladatnak, további vizsgálódásokat igényel. Ismeretes, hogy bizonyos program-osztályokra (pl. lineáris, konvex, pszeudokonvex feladatokra) a KTL-feltételek az optimalitásnak egyúttal elegendő feltételei is [12, Chapter 8]. Ismeretesek továbbá olyan eredmények, amelyeknél a KTL-feltételeket még további feltételekkel egészítik ki oly módon, hogy a kiegészített feltételrendszer az adott program-osztályra nézve már elegendő legyen az optimalitáshoz vagy a lokális optimalitáshoz. Kétszer differenciálható cél-, illetve feltételi függvények esetén több ilyen ún. másodrendű optimalitási feltétel ismeretes [1, 2, 4, 10, 14].

Jelen dolgozat célja megadni olyan új feltételeket (első-, illetve másodrendűeket), amelyek a KTL-feltételekkel együtt az (1) feladat esetében a lokális optimalitás elegendő feltételéül szolgálnak.

II. Egy elsőrendű optimalitási feltétel

Kuhn és *Tucker* immár klasszikussá vált cikkükben [9] bebizonyították, hogy amennyiben az (1) feladat célfüggvénye konkáv és a lehetséges programok L halmaza konvex, úgy a *KTL*-feltételek az optimalitásnak nem csak szükséges, hanem elégséges feltételei is. Ezt az eredményt O. L. *Mangasarian* általánosította, megmutatván a KTL-feltételek elegendőségét pszeudokonkáv célfüggvényű (1) feladat esetében is [11]. Vizsgálódásaink alapjául Mangasarian tételének egy általánosítása szolgál. Ennek megfogalmazásához szükségünk lesz a következő két definícióra.

1. *definíció.* Legyen $x_0 \in L$. Az L halmazt x_0-ban *lokálisan csillagszerű*nek nevezzük, ha x_0-nak van olyan G gömbkörnyezete, hogy minden $x \in L \cap G$ vektorra

$$[x, x_0] = \{tx + (1-t)x_0 \mid t \in [0, 1]\} \subset L. \tag{6}$$

2. *definíció.* A differenciálható $f(x)$ függvényt az x_0 *pontban lokálisan pszeudokonkávnak*, illetve *lokálisan szigorúan pszeudokonkávnak* nevezzük, ha van az

x_0-nak olyan G gömbkörnyezete, hogy a következő két feltétel közül a felsorolás sorrendjének megfelelő feltétel teljesül:

(7) ha $x \in G$ és $f(x) \geq f(x_0)$, akkor $\nabla f(x_0)(x - x_0) \geq 0$,

és $f(x) > f(x_0)$ esetén $\nabla f(x_0)(x - x_0) > 0$;

(8) ha $x \in G$, $x \neq x_0$ és $f(x) \geq f(x_0)$, akkor $\nabla f(x_0)(x - x_0) > 0$.

1. *tétel.* Tekintsük az (1) feladatot. Teljesüljenek az $x_0 \in L$ pontban a következő feltételek:

(a) x_0 KTL-stacionárius pontja az (1) feladatnak, továbbá a lehetséges programok L halmaza lokálisan csillagszerű x_0-ban,

(b) az $f(x)$ célfüggvény lokálisan pszeudokonkáv (lokálisan szigorúan pszeudokonkáv) x_0-ban.

Ekkor az x_0 lehetséges program lokális optimumpontja (szigorú lokális optimumpontja) az (1) feladatnak.

Bizonyítás. Legyen G egy olyan gömbkörnyezete x_0-nak, amelyre teljesül a (6) feltétel. A lokális pszeudokonkavitás esetét vizsgálva, az általánosság megszorítása nélkül feltehetjük, hogy a (7) feltétel teljesül a G környezeten. Okoskodjunk indirekt módon: tegyük fel, hogy van olyan $x_1 \in L \cap G$, hogy $f(x_1) > f(x_0)$. Ebből, a (7) feltétel folytán, a

(9) $$\nabla f(x_0)(x_1 - x_0) > 0$$

egyenlőtlenség adódik. Tekintettel arra, hogy az L halmaz lokálisan csillagszerű x_0-ban, ezért (6) szerint $[x_1, x_0] \subset L \cap G$. Most megmutatjuk, hogy

(10) $$t_i \nabla g_i(x_0)(x_1 - x_0) \leq 0$$

teljesül minden $i = 1, \ldots, m$ indexre. Mivel x_0-ra teljesül az (5) KTL-feltétel, ezért elegendő a (10) teljesülését csak az ún. x_0-ban aktív feltételekre igazolni. Legyen tehát i egy olyan index, amelyre $g_i(x_0) = 0$ teljesül. Mivel a $\nabla g_i(x_0)(x_1 - x_0) > 0$ feltételből az következik, hogy az $[x_1, x_0]$ szakasznak van olyan z belső pontja, amelyre $g_i(z) > 0$, ami viszont ellentmondásban van a tétel azon feltevésével miszerint $[x_1, x_0] \subset L$, ezért szükségképpen a $\nabla g_i(x_0)(x_1 - x_0) \leq 0$ egyenlőtlenségnek kell teljesülnie. Mivel $t_i \geq 0$, ezért (10) érvényessége bizonyítást nyert.

(10)-ből a (4) KTL-feltétel figyelembevételével

$$\nabla f(x_0)(x_1 - x_0) = \sum_{i=1}^{m} t_i \nabla g_i(x_0)(x_1 - x_0) \leq 0$$

következik, amely viszont ellentmond (9)-nek. Ez az ellentmondás bizonyítja, hogy $f(x) \leq f(x_0)$ teljesül minden $x \in L \cap G$ lehetséges programra.

Az nyilvánvaló, hogy ha a G gömbkörnyezeten a (7) feltétel helyett a nála erősebb (8) feltétel teljesül, akkor egyetlen $x \in L \cap G$, $x \neq x_0$ vektorra sem lehet $f(x) = f(x_0)$. *Q. E. D.*

A továbbiakban feltételezzük, hogy az $f(x)$ célfüggvény folytonosan differenciálható az x_0 KTL-stacionárius pont valamely környezetében és $\nabla f(x_0) \neq 0$. Ezen feltételezések mellett megadunk egy, az 1. tétel (b) feltételével ekvivalens feltételt. Ezen új feltétel kimondásához és az ekvivalencia bizonyításához szükségünk van néhány fogalom és jelölés bevezetésére, valamint két segédtételre.

Legyen $d \in R^n$ olyan normált vektor, amelyre $\nabla f(x_0)d \neq 0$. Legyen $\mathcal{B} = \{b_1, \dots, b_{n-1}\} \subset R^n$ ortonormált vektoroknak egy olyan rendszere, amely d-vel kiegészítve R^n-ben bázist alkot. Jelölje B azt az $n \times (n-1)$-es mátrixot, melynek oszlopvektorai rendre a $b_1, \dots, b_{n-1}$ vektorok. Legyen $x \in R^n$ tetszőleges. Az $u = B^T x \in R^{n-1}$ vektor és a $v = d^T x \in R$ skalár egyértelműen meghatározzák x-et, ezért az x vektor és az (u, v) rendezett pár között nem teszünk különbséget. (^{T} a transzponálás jele.)

Tekintsük az $f(u, v) = f(x_0)$ egyenletet, ahol $x_0 = (u_0, v_0)$. Az implicit-függvény tétel szerint igazak a következő állítások:

(IFT): létezik x_0-nak olyan G környezete, u_0-nak olyan N környezete és egyetlen olyan N-en definiált differenciálható $h_{d,x_0}(u)$ függvény, hogy $(u, v) \in G$ és $f(u, v) = f(x_0)$ akkor és csak akkor teljesül, ha $u \in N$ és $h_{d,x_0}(u) = v$; továbbá minden $u \in N$-re érvényes a következő összefüggés:

$$\nabla h_{d,x_0}(u) = -\frac{\nabla f(u, h_{d,x_0}(u))B}{\nabla f(u, h_{d,x_0}(u))d}. \tag{11}$$

Vezessük be a következő függvényt:

$$H_{d,x_0}(u) = -(\nabla f(x_0)d)h_{d,x_0}(u), \quad u \in N.$$

A további vizsgálódásokban ez a *korrigált implicit-függvény* alapvető szerepet játszik. A (11) összefüggés felhasználásával egyszerűen igazolható a következő állítás.

1. *segédtétel.* Minden $u \in N$-re

$$\nabla f(x_0)(x - x_0) = \nabla H_{d,x_0}(u)(u - u_0) + H_{d,x_0}(u_0) - H_{d,x_0}(u),$$

ahol $x = (u, h_{d,x_0}(u))$.

A következő segédtétel az 1. tétel (b) feltételével ekvivalens.

2. *segédtétel.* Legyen $f(x)$ folytonosan differenciálható x_0 valamely környezetében. Legyen $\nabla f(x_0)d \neq 0$ és $x_0 = (u_0, v_0)$. Ahhoz, hogy $f(x)$ x_0-ban lokálisan pszeudokonkáv (lokálisan szigorúan pszeudokonkáv) legyen, szükséges és elegendő, hogy a $H_{d,x_0}(u)$ korrigált implicit-függvény u_0-ban lokálisan konkáv (lokálisan szigorúan konkáv) legyen.

Bizonyítás. Szükségesség: Legyen $f(x)$ lokálisan pszeudokonkáv az x_0 pontban. x_0-nak van olyan G és u_0-nak olyan N környezete, melyekre teljesül az (IFT) állítás. Az általánosság megszorítása nélkül feltehetjük, hogy G-n teljesül a (7) feltétel. Mivel minden $u \in N$ esetén $x = (u, h_{d,x_0}(u)) \in G$ és $f(x) = f(x_0)$, ezért az 1. segédtétel és a (7) összefüggés alapján minden $u \in N$-re igaz a következő egyenlőtlenség:

$$H_{d,x_0}(u) \leq H_{d,x_0}(u_0) + \nabla H_{d,x_0}(u_0)(u - u_0), \tag{12}$$

amely differenciálható függvényekre éppen a lokális konkavitás definíciója.

Elegendőség: Az általánosság megszorítása nélkül feltehetjük, hogy minden $u \in N$-re teljesül a (12) egyenlőtlenség. Megmutatjuk, hogy ekkor a (7) összefüggés teljesül a G környezeten. *Rapcsák Tamás* egy eredménye alapján azonban ehelyett elegendő az

$$x \in G,\ f(x) \geq f(x_0) \Rightarrow \nabla f(x_0)(x - x_0) \geq 0$$

implikáció helyességét igazolni [13, 2.6. Lemma]. Indirekt módon okoskodva tegyük fel, hogy van olyan $x_1 \in G$, amelyre $f(x_1) \geq f(x_0)$ és mégis $\nabla f(x_0)(x_1 - x_0) < 0$. E két utóbbi egyenlőtlenségből következik, hogy az $f(x)$ függvény az $[x_0, x_1]$ szakasz mentén kezdetben csökken, x_0 közelében a függvényértékek az $f(x_0)$ értéknél kisebbek lesznek, az x_1 végpontban azonban az $f(x_1)$ függvényérték már ismét nem-kisebb, mint $f(x_0)$. Lévén az $f(x)$ folytonos az $[x_0, x_1]$ szakasz mentén, ezért az elmondottak alapján nyilvánvaló, hogy van olyan $x_2 \neq x_0$ pontja az $[x_0, x_1]$ szakasznak, amelyre $f(x_2) = f(x_0)$ teljesül. Mivel az $x_2 - x_0$ vektor csak egy pozitív szorzóban tér el az $x_1 - x_0$ vektortól, ezért x_2-re még a $\nabla f(x_0)(x_2 - x_0) < 0$ feltétel is teljesül. Legyen $x_2 = (u_2, v_2)$. Nyilvánvaló, hogy $u_2 \in N$ és $v_2 = h_{d,x_0}(u_2)$. Tekintettel az 1. segédtételre, azt kapjuk, hogy

$$\nabla H_{d,x_0}(u_0)(u_2 - u_0) + H_{d,x_0}(u_0) - H_{d,x_0}(u_2) < 0.$$

Ez viszont ellentmond kiindulási feltevésünknek.

A lokális szigorú pszeudokonkavitásra vonatkozó állítás teljesen hasonlóan bizonyítható. Q. E. D.

Az imént bizonyított segédtételt az 1. tétellel kombinálva nyerjük a következő optimalitási tételt.

2. *tétel.* Tekintsük az (1) feladatot. Teljesüljenek az $x_0 \in L$ pontban a következő feltételek:

(a) x_0 KTL-stacionárius pontja az (1) feladatnak, továbbá a lehetséges programok L halmaza lokálisan csillagszerű x_0-ban;

(c) $\nabla f(x_0)d \neq 0$ és a $H_{d,x_0}(u)$ korrigált implicit-függvény lokálisan konkáv (lokálisan szigorúan konkáv) az $u_0 = B^T x_0$ pontban.

Ekkor az x_0 lehetséges program lokális optimumpontja (szigorú lokális optimumpontja) az (1) feladatnak.

A következő részben kétszer differenciálható $f(x)$ célfüggvényt feltételezve, a 2. tétel (c) feltételéből kiindulva, az optimalitásnak egy másodrendű jellemzését nyerjük.

III. Egy másodrendű optimalitási feltétel és annak variánsai

Legyen az (1) feladat $f(x)$ célfüggvénye kétszer folytonosan differenciálható a feladat x_0 KTL-stacionárius pontjának valamely környezetében. Tegyük fel, hogy $\nabla f(x_0)d \neq 0$. Ismert tény, hogy ekkor a $h_{d,x_0}(u)$ implicit-függvény, és ezáltal a $H_{d,x_0}(u)$ korrigált implicit-függvény is kétszer folytonosan differenciálható az $u_0 = B^T x_0 \in R^{n-1}$ pont valamely környezetében. Jól ismert továbbá az is, hogy amennyiben a $H_{d,x_0}(u)$ függvény $\nabla^2 H_{d,x_0}(u)$ Hesse-mátrixa negatív definit az $u = u_0$ pontban, akkor a $H_{d,x_0}(u)$ függvény lokálisan szigorúan konkáv az $u = u_0$ pontban. Ezt összevetve a 2. tétellel a következő eredményre jutunk.

3. *tétel*. Tekintsük az (1) feladatot. Legyen $f(x)$ kétszer folytonosan differenciálható az $x_0 \in L$ pont valamely környezetében.
Teljesüljenek továbbá a következő feltételek:
(a) x_0 KTL-stacionárius pontja az (1) feladatnak és a lehetséges programok L halmaza lokálisan csillagszerű x_0-ban;
(d) a $\nabla^2 H_{d,x_0}(u_0)$ Hesse-mátrix negatív definit.
Ekkor az x_0 lehetséges program szigorú lokális optimumpontja az (1) feladatnak.

Fenti tétel alkalmazhatósága felveti a $\nabla^2 H_{d,x_0}(u_0)$ Hesse-mátrix kiszámíthatóságának kérdését, valamint azt, hogy a $h_{d,x_0}(u)$ implicit-függvény által meghatározott $\nabla^2 H_{d,x_0}(u_0)$ Hesse-mátrix milyen kapcsolatban van magával az $f(x)$ függvénnyel? Erről a kapcsolatról szól a következő segédtétel, melynek igazságáról az $f(u, h_{d,x_0}(u)) = \text{const.}$ egyenletből kétszeri differenciálással lehet meggyőződni.

3. *segédtétel*. Legyen $f(x)$ kétszer differenciálható az x_0 pont valamely környezetében és $\nabla f(x_0)d \neq 0$. Ekkor érvényes a következő összefüggés:

$$\nabla^2 H_{d,x_0}(u_0) = B^T(Q_1(x_0) - Q_2(x_0, d) + Q_3(x_0, d))B, \tag{13}$$

ahol

$$Q_1(x_0) = \nabla^2 f(x_0),$$

$$Q_2(x_0, d) = \frac{1}{\nabla f(x_0)d}(\nabla f(x_0)^T d^T \nabla^2 f(x_0) + \nabla^2 f(x_0) d \nabla f(x_0)),$$

$$Q_3(x_0, d) = \frac{d^T \nabla^2 f(x_0) d}{(\nabla f(x_0) d)^2} \nabla f(x_0)^T \nabla f(x_0).$$

A (13) összefüggés jobb oldalára vezessük be a $Q_{\mathfrak{B}}(f;\, x_0, d)$ jelölést, és ezt az $n-1$-ed rendű szimmetrikus mátrixot nevezzük az $f(x)$ függvény x_0 pontbeli, a $\{d, \mathfrak{B}\}$ bázishoz asszociált *kvázi Hesse-mátrix*ának [5, 7].
Valamely $\{d, \mathfrak{B}\}$ bázist x_0-ban megengedettnek nevezünk, ha a d vektorra teljesül a $\nabla f(x_0)d \neq 0$ feltétel. x_0-ban megengedett bázis akkor és csak akkor létezik, ha $\nabla f(x_0) \neq 0$. E feltétel teljesülése esetén viszont végtelen sok x_0-ban megengedett bázis van. A következő segédtétel különböző megengedett bázisokhoz asszociált kvázi Hesse-mátrixok kapcsolatáról szól.

4. *segédtétel*. [6, 3.9. Lemma] Legyen $f(x)$ kétszer differenciálható az x_0 pont valamely környezetében és $\nabla f(x_0) \neq 0$. Legyenek $\{d_1, \mathfrak{B}_1\}$ és $\{d_2, \mathfrak{B}_2\}$ x_0-ban megengedett bázisok. Ekkor a $Q_{\mathfrak{B}_1}(f;\, x_0, d_1)$ és a $Q_{\mathfrak{B}_2}(f;\, x_0, d_2)$ kvázi Hesse-mátrixok hasonlóak.

Ebből a segédtételből közvetlenül következik, hogy az $f(x)$ függvény x_0 pontbeli valamennyi $Q_{\mathfrak{B}}(f;\, x_0, d)$ kvázi Hesse-mátrixa vagy egyidejűleg negatív definit, vagy egyik sem az.
Ha a 3. tételben a $\nabla^2 H_{d,x_0}(u_0)$ Hesse-mátrixot a $Q_{\mathfrak{B}}(f;\, x_0, d)$ kvázi Hesse-mátrixszal helyettesítjük, akkor *Rapcsák Tamás* eredményével lényegében megegyező következő tételhez jutunk.

4. *tétel.* [14, 4.1. Tétel] Tekintsük az (1) feladatot. Legyen $f(x)$ kétszer folytonosan differenciálható az $x_0 \in L$ pont valamely környezetében. Teljesüljenek továbbá a következő feltételek:

(a) x_0 KTL-stacionárius pontja az (1) feladatnak és a lehetséges programok L halmaza lokálisan csillagszerű x_0-ban,

(e) a $Q_{\mathcal{B}}(f;\, x_0,\, d)$ kvázi Hesse-mátrix negatív definit valamely x_0-ban megengedett $\{d,\, \mathcal{B}\}$ bázisra.

Ekkor az x_0 lehetséges program szigorú lokális optimumpontja az (1) feladatnak.

A következő segédtételből kiderül, hogy a meglehetősen bonyolult képzési szabályú kvázi Hesse-mátrixok között vannak egyszerűbb szerkezetűek is.

5. *segédtétel.* Ha a $\{d,\, \mathcal{B}\}$ x_0-ban megengedett bázis olyan, hogy

$$b_i^T \nabla^2 f(x_0) d = 0$$

teljesül minden $b_i \in \mathcal{B}$-re, akkor

$$Q_{\mathcal{B}}(f;\, x_0,\, d) = B^T \left[\nabla^2 f(x_0) + \frac{d^T \nabla^2 f(x_0) d}{(\nabla f(x_0) d)^2} \nabla f(x_0)^T \nabla f(x_0)\right] B.$$

Bizonyítás. Az állítás nyilvánvaló a kvázi Hesse-mátrix definíciója, a (13) formula alapján. A segédtétel feltétele ugyanis nem más, mint a $B^T \nabla^2 f(x_0) d = 0$ összefüggés, amiből közvetlenül adódik a $B^T Q_2(x_0,\, d) B = 0$ egyenlőség, illetve a bizonyítandó állítás.

Ha a d vektor történetesen sajátvektora a $\nabla^2 f(x_0)$ Hesse-mátrixnak, akkor az 5. segédtétel feltételei nyilvánvaló módon teljesülnek. Ez az oka annak, hogy a továbbiakban olyan megengedett bázisokat vizsgálunk, melyek elemei a $\nabla^2 f(x_0)$ Hesse-mátrix sajátvektorai.

Legyenek $u_1, \dots, u_n \in R^n$ a $\nabla^2 f(x_0)$ Hesse-mátrix ortonormált sajátvektorai, $h_1, \dots, h_n \in R$ pedig a megfelelő sajátértékek. Legyen U az az ortogonális mátrix, melynek oszlopvektorai $u_1, \dots, u_n$. Mivel $\nabla f(x_0) \neq 0$, ezért van olyan i index, hogy $\nabla f(x_0) u_i \neq 0$. Legyen $\mathcal{B}_i = \{u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n\}$. A továbbiakban a $Q_{\mathcal{B}_i}(f;\, x_0,\, u_i)$ kvázi Hesse-mátrixot fogjuk vizsgálni.

Legyen E_i az az $n \times (n-1)$-es mátrix, melynek oszlopvektorai rendre az $e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n \in R^n$ egységvektorok. Az 5. segédtételből közvetlenül adódik a következő összefüggés:

$$(14) \qquad Q_{\mathcal{B}_i}(f;\, x_0,\, u_i) = E_i^T U^T \left[\nabla^2 f(x_0) + \frac{h_i}{(\nabla f(x_0) u_i)^2} \nabla f(x_0)^T \nabla f(x_0)\right] U E_i.$$

Ennek az összefüggésnek a segítségével a $Q_{\mathcal{B}_i}(f;\, x_0,\, u_i)$ kvázi Hesse-mátrix negatív definitségének érdekes jellemzését adhatjuk, amely egyúttal az optimalitás egy új másodrendű elégséges feltételét is szolgáltatja.

5. *tétel.* Tekintsük az (1) feladatot. Legyen $f(x)$ kétszer folytonosan differenciálható az $x_0 \in L$ pont valamely környezetében és $\nabla f(x_0) \neq 0$. Legyenek $u_1, \dots, u_n \in R^n$ a $\nabla^2 f(x_0)$ Hesse-mátrix ortonormált sajátvektorai, $h_1, \dots, h_n \in R$ pedig a megfelelő sajátértékek. Teljesüljenek továbbá a következő feltételek:

(a) x_0 KTL-stacionárius pontja az (1) feladatnak és a lehetséges programok L halmaza lokálisan csillagszerű x_0-ban,

(f) a $\nabla^2 f(x_0)$ Hesse-mátrix vagy negatív definit, vagy egyetlen egyszeres nem-negatív sajátértéke van (legyen ez h_k) és teljesül a $\nabla f(x_0)u_k \neq 0$ összefüggés, amely a $h_k > 0$ esetben még kiegészül a következő egyenlőtlenséggel:

$$\sum_{j=1}^{n} \frac{(\nabla f(x_0)u_j)^2}{h_j} > 0. \tag{15}$$

Ekkor az x_0 lehetséges program szigorú lokális optimumpontja az (1) feladatnak.

Bizonyítás. Ha a $\nabla^2 f(x_0)$ Hesse-mátrix negatív definit, akkor a (14) összefüggés szerint a $Q_{\mathfrak{B}_i}(f;\, x_0,\, u_i)$ kvázi Hesse-mátrix ugyancsak negatív definit. (Itt i egy olyan indexet jelöl, amelyre $\nabla f(x_0)u_i \neq 0$.)

Most tegyük fel, hogy a $\nabla^2 f(x_0)$ mátrixnak h_k az egyetlen egyszeres nem-negatív sajátértéke és $\nabla f(x_0)u_k \neq 0$. Ez utóbbi feltétel biztosítja a $Q_{\mathfrak{B}_k}(f;\, x_0,\, u_k)$ kvázi Hesse-mátrix létezését. Most megmutatjuk, hogy a feladat feltételei mellett ez a mátrix negatív definit. Legyen $z \in R^{n-1}$ tetszőleges és legyen $E_k^T z = (z_1, \ldots, z_n)^T \in R^n$. A (14) összefüggésből következik, hogy

$$z^T Q_{\mathfrak{B}_k}(f;\, x_0,\, u_k) z = \sum_{j\neq k} h_j z_j^2 + \frac{h_k}{(\nabla f(x_0)u_k)^2} \left(\sum_{j \neq k} (\nabla f(x_0)u_j) z_j \right)^2 . \tag{16}$$

Ha $h_k = 0$, akkor (16)-ból közvetlenül kiadódik, hogy a $Q_{\mathfrak{B}_k}(f;\, x_0,\, u_k)$ mátrix negatív definit. Ha $h_k > 0$, akkor a

$$\left(\sum_{j\neq k} (\nabla f(x_0)u_j) z_j \right)^2 \leq \left(\sum_{j\neq k} \frac{(\nabla f(x_0)u_j)^2}{h_j} \right) \left(\sum_{j\neq k} h_j z_j^2 \right)$$

egyenlőtlenség (melynek érvényessége a *Cauchy-Schwarz-Bunyakovszkij*-féle egyenlőtlenség segítségével belátható) és a (16) formula figyelembevételével a következő egyenlőtlenséghez jutunk:

$$z^T Q_{\mathfrak{B}_k}(f;\, x_0,\, u_k) z \leq \left(\sum_{j\neq k} h_j z_j^2 \right) \left(1 + \frac{h_k}{(\nabla f(x_0)u_k)^2} \sum_{j\neq k} \frac{(\nabla f(x_0)u_j)^2}{h_j} \right).$$

Itt a jobb oldali szorzat első tényezője $z \neq 0$ esetén nyilván negatív, a második tényező viszont (15) miatt pozitív, következésképpen a bal oldal értéke negatív, ami azt jelenti, hogy a $Q_{\mathfrak{B}_k}(f;\, x_0,\, u_k)$ mátrix negatív definit. Kimutattuk tehát, hogy bizonyítandó tételünk (*f*) feltétele maga után vonja a 4. tétel (e) feltételét, melynek alapján nyilvánvaló az 5. tétel helyessége. Q. E. D.

Megemlítjük, hogy az (e) és (f) feltételek valójában ekvivalensek [7].

Míg az imént bizonyított tétel másodrendű optimalitási feltételében — a nem-konkáv esetben — a $\nabla^2 f(x_0)$ mátrix minden sajátvektorának és sajátértékének ismeretére szükség van, addig a következő tétel másodrendű feltételében — a nem-konkáv esetben — csak egy sajátérték és a hozzátartozó sajátvektor ismeretére van szükség.

6. *tétel.* Az 5. tétel állítása érvényben marad, ha az (f) feltételt a következő feltétellel helyettesítjük:

(g) a $\nabla^2 f(x_0)$ Hesse-mátrix vagy negatív definit, vagy egyetlen egyszeres nem-negatív sajátértéke van (legyen ez h_k) és teljesül a $\nabla f(x_0)u_k \neq 0$ összefüggés, amely a $h_k > 0$ esetben kiegészül még a következővel: az

$$r = (\nabla^2 f(x_0))^{-1} \nabla f(x_0)^T$$

vektorra $\nabla f(x_0)r > 0$ teljesül.

Bizonyítás. Megmutatjuk, hogy a (g) feltétel teljesülése maga után vonja az (f) feltétel teljesülését. Ehhez csupán a $h_k > 0$ esetet kell vizsgálnunk. Legyen U a $\nabla^2 f(x_0)$ mátrix sajátvektoraiból a már leírt módon képezett mátrix és $p \in R^n$ az a vektor, amelyre $r = Up$. Ekkor az r vektorra vonatkozó feltétel miatt fennáll az

$$U^T \nabla^2 f(x_0) U p = U^T \nabla f(x_0)^T$$

összefüggés. Ennek az egyenletnek a koordinátás alakja a következő:

$$h_j p_j = \nabla f(x_0) u_j, \quad j = 1, \ldots, n. \tag{17}$$

Tekintettel arra, hogy

$$\nabla f(x_0) r = \nabla f(x_0) U p = \sum_{j=1}^{n} (\nabla f(x_0) u_j) p_j,$$

ezért ebből, a (17) és a $\nabla f(x_0)r > 0$ feltételekből közvetlenül kiadódik az (f) feltételben szereplő

$$\sum_{j=1}^{n} \frac{(\nabla f(x_0) u_j)^2}{h_j} > 0$$

egyenlőtlenség. Q. E. D.

Megemlítjük, hogy a (g) és (f) feltételek valójában ekvivalensek [7].

(Beérkezett: 1984. szeptember 15-én.)

IRODALOM

1. Bazaraa, M. S.—Shetty, C. M.: *Nonlinear programming: theory and algorithms* Wiley, New York, 1979.
2. Fiacco, A. V.—McCormick, G. P.: *Nonlinear programming: Sequential unconstrained minimization techniques,* Wiley, New York, 1968.
3. Forgó F.: *Nemkonvex és diszkrét programozás,* KJK, Bp., 1978.
4. Hestenes, M. R.: *Optimization theory—The finite dimensional case,* Wiley, New York, 1975.
5. Komlósi, S.: Second order characterisation of pseudoconvex and strictly pseudoconvex functions in terms of quasi-Hessians, in: *Contributions to the theory of optimization,* Ed.: Forgó F., Karl Marx University of Economics Budapest, DM 83-2. (1983) 19—45.
6. Komlósi S.: Néhány adalék a kvázikonvex függvények elméletéhez, *Alkalmazott Matematikai Lapok,* 10 (1984), 103—113.
7. Komlósi S.: A kvázikonvex programozás elmélete és módszerei egyes kérdéseihez, Baku, 1984. Kandidátusi disszertáció. (Orosz nyelven.)
8. Krekó B.: *Optimumszámítás,* KJK. BP. 1973.

9. KUHN, H. W.—TUCKER, A. W.: Non-linear programming, in: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Ed.: J. Neyman). Univ. California Press, Berkeley, Calif. 1951.
10. LUENBERGER, D. G.: *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading, MA, 1973.
11. MANGASARIAN, O. L.: Pseudo-convex functions, *SIAM Journal on Control* 3 (1965) 281—290.
12. MARTOS, B.: *Nonlinear programming: theory and methods*, Akadémiai Kiadó, Budapest, 1975.
13. RAPCSÁK T.: A SUMT-módszer alkalmazása nem konvex programozási feladatok esetén, *Alk. Mat. Lapok*, 2 (1976) 427—437.
14. RAPCSÁK T.: Az optimalitás másodrendű feltételeiről, *Alk. Mat. Lapok*, 4 (1978) 109—116.

ON OPTIMALITY CRITERIA OF MATHEMATICAL PROGRAMMING PROBLEMS

The paper deals with sufficient optimality conditions of inequality constrained nonlinear programming problems. Assuming differentiability of the objective and the constraint functions first and second order conditions, respectively, are elaborated which, together with the well-known Kuhn-Tucker conditions, are sufficient for local optimality and strict local optimality, respectively. Investigations of the paper are closely connected with function properties of local pseudo-concavity and local strict pseudo-concavity, respectively. The local variant of a well-known theorem of Mangasarian serves as a starting point (Theorem 1). From this results the first order characterization of local optimality in terms of the implicit function (Theorem 2). In the following, assuming continuous bidifferentiability of the objective function second order conditions are submitted which ensure the fulfilment of the first order optimality conditions (Theorems 3—6). Second order optimality criteria are fundamentally based on the notion of the quasi-Hessian matrix (Theorem 4) and its various properties. Theorem 5 gives such a sufficient optimality condition that may be described by means of the eigenvalues and eigenvectors of the Hessian as well as of the gradient-vector of the objective function. This theorem is a generalization of the classical optimality theorem where the definiteness of the Hessian matrix is required.

О КРИТЕРИЯХ ОПТИМАЛЬНОСТИ В ЗАДАЧАХ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

В статье рассматриваются вопросы достаточности условий оптимальности задач нелинейного программирования с нелинейными ограничениями. Предполагая дифференцируемость целевой функции и ограничений, в статье разработаны условия первого и второго порядка, которые совместно с условиями Куна-Таккера служат условиями достаточности локальной и строго локальной оптимальности. Исследования тесно связаны со свойствами локальной псевдовогнутости и строго локальной псевдовогнутости функций. Исходной точкой является локальный вариант одной из известных теорем Мангасаряна (1. теорема). Исходя из неё и используя теорему о неявной функции, получаем анализ первого порядка локальной оптимальности с помощью неявных функций (2. теорема).

В дальнейшем, предполагая дважды непрерывную дифференцируемость целевой функций, задаются такие условия второго порядка, которые обеспечивают выполнение условий оптимальности первого порядка (3.—6. теорема). Условия оптимальности второго порядка опираются в основном на понятия квази-матрицы Хессе (4. теорема) и на её различные свойства. 5. теорема даёт такое условие достаточности, которое может быть описано с помощью собственных значений и собственных векторов матрицы Хессе целевой функции. Эта теорема является обобщением той классической теоремы оптимальности, в которой требуется определённость матрицы Хессе.