FELTÉTELES OPTIMALIZÁLÁS

DR. NAGY TAMÁS
egyetemi docens

Miskolci Egyetem
Alkalmazott Matematikai Tanszék

„A bemutatott kutató munka a TÁMOP-4.2.1.B-10/2/KONV-2010-0001 jelű projekt részeként az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg"
„This research was carried out as part of the TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund."

Miskolc, 2012

Tartalomjegyzék

1. Bevezetés

Azt az optimalizálási feladatot, amelyben az $\mathbf{x} \in \mathbb{R}^n$ döntési változóra az $\mathbf{x} \in D_f \subset \mathbb{R}^n$, ill. az $\mathbf{x} \in X \subset \mathbb{R}^n$ kikötések (D_f az $f(\mathbf{x})$ függvény értelmezési tartománya, ill. az X egy nyílt halmaz) mellett egyéb előírásokat teszünk, feltételes optimalizálási feladatnak nevezzük. A feltételeket meghatározó halmazt S-el fogjuk jelölni és feltételi halmaznak vagy megengedett (lehetséges) megoldások halmazának nevezzük. A feladatunk az, hogy megkeressük az S halmaz azon pontját, amelyben az $f : \mathbb{R}^n \to \mathbb{R}$ függvény a legkisebb értéket veszi fel. Ezt a függvényt célfüggvénynek nevezzük. A feltételes optimalizálási feladat jelölésére az alábbiakat szokás használni:

$$\min\{f(\mathbf{x}) : \mathbf{x} \in S\} \qquad \text{vagy} \qquad \min_{\mathbf{x}\in S}\{f(\mathbf{x})\} \qquad \text{vagy} \qquad \underset{\mathbf{x}\in S}{f(\mathbf{x})} \;\to\; \min!$$

2. Javító irányok és lehetséges irányok kúpja

DEFINÍCIÓ (Javító irányok kúpjának definíciója) Legyen az $f : \mathbb{R}^n \to \mathbb{R}$ függvény. Az f függvény $\bar{\mathbf{x}} \in \mathbb{R}^n$ pontbeli javító irányok kúpjának (kónuszának) nevezzük és F-vel jelöljük az alábbiakban definiált halmazt:

$$F = \{\mathbf{d} : \mathbf{d} \in \mathbb{R}^n,\ \mathbf{d} \neq \mathbf{0},\ f(\bar{\mathbf{x}} + \lambda\mathbf{d}) < f(\bar{\mathbf{x}}) \text{ minden } \lambda \in (0,\delta)\text{-ra valamely } \delta > 0 \text{ esetén}\}.$$

Minden $\mathbf{d} \in F$ vektort javító (csökkentő) iránynak nevezünk. A definíció szerint tehát minden olyan $\mathbf{d}$ irányt javító (csökkentő) iránynak nevezünk, amely irányba az $\bar{\mathbf{x}}$ pontból legalább egy kicsit folytonosan elmozdulva a célfüggvény értéke mindig kisebb, mint az $\bar{\mathbf{x}}$ pontbeli célfüggvény érték.

DEFINÍCIÓ (Lehetséges irányok kúpjának definíciója)
Legyen $S \subset \mathbb{R}^n$ nemüres halmaz és legyen $\bar{\mathbf{x}} \in cl\ S$, azaz $\bar{\mathbf{x}}$ az S halmaz lezártjának egy pontja. Az S halmaz $\bar{\mathbf{x}}$ pontbeli lehetséges irányok kúpjának (kónuszának) nevezzük és D-vel jelöljük az alábbiakban definiált halmazt:

$$D = \{\mathbf{d} : \mathbf{d} \in \mathbb{R}^n,\ \mathbf{d} \neq \mathbf{0},\ \bar{\mathbf{x}} + \lambda\mathbf{d} \in S,\ \text{minden } \lambda \in (0,\delta)\text{-ra valamely } \delta > 0 \text{ esetén}\}.$$

Minden $\mathbf{d} \in D$ vektort lehetséges iránynak nevezünk. A definíció szerint tehát minden olyan $\mathbf{d}$ irányt lehetséges iránynak nevezünk, amely irányba az $\bar{\mathbf{x}}$ pontból legalább egy kicsit folytonosan elmozdulva mindig benne maradunk az S tartományban.

3. Geometriai szükséges feltétel

TÉTEL (geometriai szükséges feltétel)
Legyen $S \subset \mathbb{R}^n$ nemüres halmaz és $f : S \to \mathbb{R}$ függvény. Tekintsük az alábbi feltételes optimalizálási feladatot:

$$\min\{f(\mathbf{x}) : \mathbf{x} \in S\}.$$

Ha $\bar{\mathbf{x}} \in S$ lokális minimumpont, akkor

$$F \cap D = \emptyset,$$

ahol F az f függvény S halmaz $\bar{\mathbf{x}}$ pontbeli javító (csökkenő) irányok kúpja, D az S halmaz $\bar{\mathbf{x}}$ pontbeli lehetséges irányok kúpja.

A tétel azt fejezi ki, hogy lokális minimumpont esetén nincs olyan irány, amely irányban ha elmozdulnánk, akkor a tartományban is benne maradnánk és a célfüggvény is javulna.

Bizonyítás
Indirekte tegyük fel, hogy $F \cap D \neq \emptyset$. Ekkor létezik olyan $\mathbf{d} \neq \mathbf{0}$ vektor, amelyre $\mathbf{d} \in F$ és $\mathbf{d} \in D$. Ez azt jelenti, hogy a $\mathbf{d}$ vektor javító irány is és egyben lehetséges irány is. Ez pedig ellentmond annak, hogy az $\bar{\mathbf{x}}$ pont optimális (minimális) megoldás. **Q.e.d.**

4. Javító irányok és lehetséges irányok kúpjának karakterizációja

Az optimalitás geometriai szükséges feltételében különösebb megkötést nem tettünk sem az $f(\mathbf{x})$ célfüggvényre, sem az S feltételi halmazra, így a tétel túl általános. Az alábbi két tételben már bizonyos előírásokat teszünk a célfüggvényre és a feltételi halmazra, ezek felhasználásával adjuk meg az optimalizálásban fontos szerepet játszó javító és lehetséges irányok kúpját.

Az alábbi tételben feltesszük, hogy a célfüggvény differenciálható.

TÉTEL (javító irányok kúpjának karakterizációja)
Legyen $f : \mathbb{R}^n \to \mathbb{R}$ függvény differenciálható az $\bar{\mathbf{x}} \in \mathbb{R}^n$ pontban. Ekkor az

$$F_0 = \{\mathbf{d} : \boldsymbol{\nabla} f(\bar{\mathbf{x}})\mathbf{d} < 0\}$$

halmaz javító irányok kúpja és $F_0 \subseteq F$.

Bizonyítás
Legyen $\mathbf{d} \in F_0$ vektor és $\lambda > 0$ skalár. Az f függvény $\bar{\mathbf{x}}$ pontbeli differenciálhatósága miatt írhatjuk, hogy

$$f(\bar{\mathbf{x}} + \lambda\mathbf{d}) = f(\bar{\mathbf{x}}) + \lambda\boldsymbol{\nabla} f(\bar{\mathbf{x}})\mathbf{d} + \lambda \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}),$$

ahol $\lim_{\lambda\to 0} \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}) = 0$. Osszuk be az egyenletet $\lambda > 0$ számmal és rendezzük, ekkor

$$\frac{f(\bar{\mathbf{x}} + \lambda\mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda} = \boldsymbol{\nabla} f(\bar{\mathbf{x}})\mathbf{d} + \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}).$$

A $\boldsymbol{\nabla} f(\bar{\mathbf{x}})\mathbf{d} < 0$ és a $\lambda \to 0$ határátmenetből következik, hogy létezik $\delta > 0$ szám úgy, hogy

$$f(\bar{\mathbf{x}} + \lambda\mathbf{d}) - f(\bar{\mathbf{x}}) < 0 \text{ minden } \lambda \in (0, \delta) \text{ esetén.}$$

Ez pedig azt jelenti, hogy a $\mathbf{d} \in F_0$ vektor javító irány. Mivel a $\mathbf{d} \in F_0$ összefüggésből következik, hogy $\mathbf{d} \in F$, így $F_0 \subseteq F$. **Q.e.d.**

A továbbiakban az S halmazt függvények segítségével fogjuk megadni. A függvényekre folytonossági és differenciálhatósági előírást fogunk tenni.

Legyen az S feltételi halmaz az alábbi

$$S = \{\mathbf{x} : \mathbf{x} \in X, \quad g_i(\mathbf{x}) \leqq 0, \ i = 1, 2, \dots, m\}.$$

ahol $g_1, g_2, \dots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények és $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz.

DEFINÍCIÓ (Aktív és inaktív feltételek)
Tekintsük az $S = \{\mathbf{x} : \mathbf{x} \in X, \ g_i(\mathbf{x}) \leqq 0, i = 1, 2, \dots, m\}$ feltételi halmazt és egy $\bar{\mathbf{x}} \in S$ pontot. A $g_i(\mathbf{x})$ függvény aktív az $\bar{\mathbf{x}} \in S$ pontban, ha $g_i(\bar{\mathbf{x}}) = 0$, inaktív, ha $g_i(\bar{\mathbf{x}}) < 0$.

TÉTEL (**lehetséges irányok kúpjának karakterizációja**)
Legyen $S = \{\mathbf{x} : \mathbf{x} \in X, \quad g_i(\mathbf{x}) \leqq 0, \ i = 1, 2, \dots, m\}$, ahol $g_1, g_2, \dots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények és $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz. Legyen adott egy $\bar{\mathbf{x}} \in S$, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy a $g_i, i \in I$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók, a $g_i, i \notin I$ függvények az $\bar{\mathbf{x}}$ pontban folytonosak. Ekkor a

$$G_0 = \{\mathbf{d} : \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} < 0, \ i \in I\}$$

halmaz az S halmaz $\bar{\mathbf{x}}$ pontbeli lehetséges irányok kúpja és $G_0 \subseteq D$.

Bizonyítás
Legyen $\mathbf{d} \in G_0$ vektor és $\lambda > 0$ skalár. Ahhoz, hogy kimutassuk a $\mathbf{d}$ vektorról, hogy lehetséges irány, három oldalról kell vizsgálódnunk.
a) Mivel $\bar{\mathbf{x}} \in S$, így $\bar{\mathbf{x}} \in X$ is igaz. Az X halmaz nyílt, ebből következik, hogy létezik $\delta_1 > 0$ szám úgy, hogy

$$\bar{\mathbf{x}} + \lambda\mathbf{d} \in X \text{ minden } \lambda \in (0, \delta_1) \text{ esetén.}$$

b) Most vizsgáljuk meg a $g_i, i \notin I$ inaktív feltételi függvényeket. A $g_i(\bar{\mathbf{x}}) < 0$ és az $\bar{\mathbf{x}}$ pontbeli folytonosság miatt létezik $\delta_2 > 0$ szám úgy, hogy minden $i \notin I$ indexre

$$g_i(\bar{\mathbf{x}} + \lambda\mathbf{d}) \leqq 0 \text{ minden } \lambda \in (0, \delta_2) \text{ esetén.}$$

c) Végül pedig vizsgáljuk meg a $g_i, i \in I$ aktív feltételi függvényeket. Az aktív függvények $\bar{\mathbf{x}}$ pontbeli differenciálhatósága miatt írhatjuk, hogy minden $i \in I$ indexre

$$g_i(\bar{\mathbf{x}} + \lambda\mathbf{d}) = g_i(\bar{\mathbf{x}}) + \lambda\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} + \lambda \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}),$$

ahol $\lim_{\lambda\to 0} \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}) = 0$. Osszuk be az egyenletet $\lambda > 0$ számmal, rendezzük és vegyük figyelembe, hogy $g_i(\bar{\mathbf{x}}) = 0$, ekkor

$$\frac{g_i(\bar{\mathbf{x}} + \lambda\mathbf{d})}{\lambda} = \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} + \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}, \lambda\mathbf{d}).$$

A $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} < 0$ és a $\lambda \to 0$ határátmenetből következik, hogy létezik $\delta_3 > 0$ szám úgy, hogy

$$g_i(\bar{\mathbf{x}} + \lambda\mathbf{d}) < 0 \text{ minden } \lambda \in (0, \delta_3) \text{ esetén.}$$

Összefoglalva a vizsgálatainkat, legyen

$$\delta = \min\{\delta_1, \delta_2, \delta_3\} > 0,$$

ekkor pedig igaz minden $\lambda \in (0, \delta)$ esetén, hogy

$$\begin{array}{rcl} \bar{\mathbf{x}} + \lambda \mathbf{d} & \in & X \\ g_i(\bar{\mathbf{x}} + \lambda \mathbf{d}) & \leqq & 0, \quad i \notin I \\ g_i(\bar{\mathbf{x}} + \lambda \mathbf{d}) & < & 0, \quad i \in I \end{array}$$

Ezekből az összefüggésekből következik, hogy $\bar{\mathbf{x}} + \lambda \mathbf{d} \in S$, ez pedig azt jelenti, hogy a $\mathbf{d} \in G_0$ vektor lehetséges irány. Mivel a $\mathbf{d} \in G_0$ összefüggésből következik, hogy $\mathbf{d} \in D$, így $G_0 \subseteq D$. **Q.e.d.**

5. Szükséges feltételek egyenlőtlenségi előírások mellett

Először azt az optimalizálási feladatot vizsgáljuk, amelynél a feltételi halmaz leírásában csak egyenlőtlenségek szerepelnek, majd később egyenlőségeket is meg fogunk engedni.

A történeti hűség miatt meg kell jegyezni, hogy időben először az egyenlőséges feltételek vizsgálatával foglalkoztak. J. L. Lagrange (1736-1813) francia matematikus alkotta meg a róla elnevezett Lagrange multiplikátor (szorzó) módszert, majd 150 évvel később születtek meg az optimalizálás további elméleti eredményei.

Tekintsük a $g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvényeket és az $X \subseteq \mathbb{R}^n$ nemüres nyílt halmazt, ezek segítségével az alábbiak szerint adjuk meg az S halmazt:

$$S = \{\mathbf{x} : \mathbf{x} \in X, \quad g_i(\mathbf{x}) \leqq 0, \ i = 1, 2, \ldots, m\}.$$

Az optimalizálási feladat a következő formákban adható meg:

$$\min \{f(\mathbf{x}) : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\}$$

vagy más formában felírva

$$\begin{array}{rcl} f(\mathbf{x}) & \to & \min! \\ g_i(\mathbf{x}) & \leqq & 0, \ i = 1, 2, \ldots, m \\ \mathbf{x} & \in & X \subseteq \mathbb{R}^n, \ (X \text{ nyílt halmaz}) \end{array}$$

5.1. Geometriai-algebrai szükséges feltétel

TÉTEL (geometriai-algebrai szükséges feltétel)

Tekintsük a $\min \{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot, amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen adott egy $\bar{\mathbf{x}} \in S$, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i, i \in I$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók, a $g_i, i \notin I$ függvények az $\bar{\mathbf{x}}$ pontban folytonosak.

Ha $\bar{\mathbf{x}}$ lokális minimumpont, akkor

$$F_0 \cap G_0 = \emptyset,$$

ahol $F_0 = \{\mathbf{d} : \nabla f(\bar{\mathbf{x}})\mathbf{d} < 0\}$ és $G_0 = \{\mathbf{d} : \nabla g_i(\bar{\mathbf{x}})\mathbf{d} < 0, i \in I\}$.

Bizonyítás

Indirekte tegyük fel, hogy $F_0 \cap G_0 \neq \emptyset$. Ekkor létezik olyan $\mathbf{d} \neq \mathbf{0}$ vektor, amelyre $\mathbf{d} \in F_0$ és $\mathbf{d} \in G_0$. A karakterizációs tételek alapján a $\mathbf{d}$ vektor javító irány is és egyben lehetséges irány is. Ez pedig ellentmond annak, hogy az $\bar{\mathbf{x}}$ pont optimális (minimális) megoldás. **Q.e.d.**

Az alábbiakban a feltételes optimalizálás szükséges feltételeit fogalmazzuk meg tételek formájában. Ezek az összefüggések adnak lehetőséget az optimális megoldás megkeresésére.

5.2. Fritz John szükséges feltételek

TÉTEL (Fritz John szükséges feltételek I., 1948)

Tekintsük a $\min\{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot, amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen adott egy $\bar{\mathbf{x}} \in S$ pont, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i, i \in I$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók, a $g_i, i \notin I$ függvények az $\bar{\mathbf{x}}$ pontban folytonosak.

Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek $u_0, u_i, i \in I$ számok úgy, hogy

$$\begin{aligned} u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} u_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ (u_0, \mathbf{u}_I) &\geqq \mathbf{0} \\ (u_0, \mathbf{u}_I) &\neq \mathbf{0} \end{aligned}$$

ahol $\mathbf{u}_I$ egy olyan vektor, amelynek komponensei $u_i \geqq 0,\ i \in I$.

Bizonyítás

A geometriai-algebrai szükséges feltétel szerint, ha az $\bar{\mathbf{x}} \in S$ lokális optimális megoldás, akkor nincs olyan $\mathbf{d} \in \mathbb{R}^n$ vektor, hogy $\nabla f(\bar{\mathbf{x}})\mathbf{d} < 0$ és $\nabla g_i(\bar{\mathbf{x}})\mathbf{d} < 0, i \in I$. A gradiens vektorokból képezzünk egy olyan $\mathbf{A}$ mátrixot, amelynek első oszlopa $\nabla f(\bar{\mathbf{x}})$, a többi oszlopa pedig a $\nabla g_i(\bar{\mathbf{x}}), i \in I$ vektorok. A geometriai-algebrai szükséges feltétel szerint a $\mathbf{dA} < 0$ rendszernek nincs megoldása. A Gordan tétel szerint ekkor létezik olyan $\mathbf{p} \geqq \mathbf{0}$, $\mathbf{p} \neq \mathbf{0}$ vektor, hogy $\mathbf{Ap} = \mathbf{0}$. Jelölje a $\mathbf{p}$ vektor első komponensét u_0, a többit pedig u_i, $i \in I$. Innen pedig következik a tétel állítása. **Q.e.d.**

A tételben szereplő u_0, u_i számokat **Lagrange szorzóknak** nevezik.

1. Példa

Tekintsük az alábbi feltételes optimalizálási feladatot!

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 \to \min! \\ g_1(x_1, x_2) &= x_1 - x_2 - 2 \leqq 0 \\ g_2(x_1, x_2) &= -x_1 - x_2 + 4 \leqq 0 \\ g_3(x_1, x_2) &= x_1 + 2x_2 - 3 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

Tegyük fel, hogy ismerjük az optimális megoldást, amely $\bar{\mathbf{x}} = (2, 2)$. Az aktív feltételi függvények $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ indexhalmaza egyszerű ellenőrzéssel meghatározható. Mivel az első és a harmadik feltétel nem egyenlőséggel teljesül, a második igen, így $I = \{2\}$. Megállapítható, hogy az $\bar{\mathbf{x}} = (2, 2)$ pontban a célfüggvény és az aktív feltételi függvény (g_2) differenciálható, az inaktív feltételi függvények (g_1, g_3) pedig folytonosak, így a fenti tétel alkalmazható, azaz léteznek u_0, u_2 számok. A szükséges gradiens vektorok az alábbiak

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \qquad \nabla g_2(x_1, x_2) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

A u_0, u_2 Lagrange szorzó meghatározásához szükséges feltételek a tétel szerint

$$\begin{aligned} u_0 \begin{bmatrix} 4 \\ 4 \end{bmatrix} + u_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} &= \mathbf{0} \\ u_0, u_2 &\geqq 0 \\ (u_0, u_2) &\neq \mathbf{0} \end{aligned}$$

A tétel tehát azt állítja, hogy

$$\begin{aligned} 4u_0 - u_2 &= 0 \\ 4u_0 - u_2 &= 0 \end{aligned}$$

egyenletrendszernek van olyan megoldása, amelyben egyik szám sem negatív és legalább egyik pozitív. Ilyen megoldás mint látjuk valóban létezik, sőt végtelen sok létezik. Az összes megoldás $u_0 = \alpha, u_2 = 4\alpha$, ahol $\alpha > 0$.

A tétel azonban nem arra szolgál, hogy meghatározzuk az optimális megoldáshoz tartozó u_i számokat, hanem arra, hogy megkeressük azokat a megoldásokat, amelyek kielégítik az optimális megoldáshoz szükséges feltételeket. Mint láttuk a tételben szükség volt az aktív feltételek indexhalmazára, ez pedig nem ismert, hisz nem ismerjük az optimális megoldást. Ezt a problémát úgy oldjuk fel, hogy minden feltételi függvényhez hozzárendelünk u_i számokat, de az inaktív feltételekhez tartozóknak zérusnak kell lenni. Mivel nem tudjuk, melyek lesznek az aktív feltételek, így elő kell írni az összes feltételi függvény differenciálhatóságát. Most pedig következzen a tétel használható változata.

TÉTEL (Fritz John szükséges feltételek II., 1948)

Tekintsük a $\min \{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot, amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen adott egy $\bar{\mathbf{x}} \in S$ pont. Tegyük fel, hogy az $f, g_i, i = 1, 2, \ldots, m$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók.

Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek $u_0, u_i,\ i = 1, 2, \ldots, m$ számok úgy, hogy

$$\begin{aligned} u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\bar{\mathbf{x}}) &= 0 \quad i = 1, 2, \ldots, m \\ (u_0, \mathbf{u}) &\geqq \mathbf{0} \\ (u_0, \mathbf{u}) &\neq \mathbf{0} \end{aligned}$$

ahol $\mathbf{u}$ egy olyan vektor, amelynek komponensei $u_i \geqq 0,\ i = 1, 2, \ldots, m$.

Bizonyítás
Ha minden $i = 1, 2, \ldots, m$ indexre előírjuk, hogy $u_i g_i(\bar{\mathbf{x}}) = 0$, akkor ez a feltétel biztosítja, hogy az inaktív feltételek ($g_i(\bar{\mathbf{x}}) < 0$) esetén $u_i = 0$ legyen. **Q.e.d.**

Gyakran szokták az alábbi elnevezéseket használni.
Az $\mathbf{x} \in S$ lehetséges megoldásra vonatkozó

$$\begin{aligned} g_i(\mathbf{x}) &\leqq 0, \quad i = 1, 2, \ldots, m \\ \mathbf{x} &\in X \end{aligned}$$

feltételeket primál feltételeknek (PF) nevezik.
A tételben szereplő

$$\begin{aligned} u_0 \nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) &= \mathbf{0} \\ (u_0, \mathbf{u}) &\geqq \mathbf{0} \\ (u_0, \mathbf{u}) &\neq \mathbf{0} \end{aligned}$$

összefüggéseket duál feltételeknek (DF) nevezik.
A tételben szereplő

$$u_i g_i(\mathbf{x}) = 0, \quad i = 1, 2, \ldots, m$$

egyenlőségeket pedig komplementaritási feltételnek (KF) nevezik.
Együttesen a primál feltételeket, a duál feltételeket és a komplementaritási feltételeket Fritz John optimalitási feltételeknek hívjuk. Bármely olyan $\bar{\mathbf{x}}$ pontot, amelyre léteznek $(\bar{u}_0, \bar{\mathbf{u}})$ Lagrange szorzók úgy, hogy az $(\bar{\mathbf{x}}, \bar{u}_0, \bar{\mathbf{u}})$ vektor kielégíti a Fritz John optimalitási feltételeket Fritz John pontnak nevezünk.
Felhívjuk az olvasó figyelmét a duál feltételre, amely azt fejezi ki, hogy optimális esetben a célfüggvény és a feltételi függvények gradiens vektorainak lineáris kombinációja zérus vektor.

Az alábbi példában megmutatjuk, hogyan kell Fritz John pontot meghatározni.

2. Példa
Határozzuk meg az alábbi optimalizálási feladathoz tartozó Fritz John pontot!

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 \to \min! \\ g_1(x_1, x_2) &= x_1 - x_2 - 2 \leqq 0 \\ g_2(x_1, x_2) &= -x_1 - x_2 + 4 \leqq 0 \\ g_3(x_1, x_2) &= x_1 + 2x_2 - 3 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

Megoldás
Mivel nem ismerjük az $\bar{\mathbf{x}}$ FJ pontot, így az optimalizálási feladat függvényeinek minden pontban differenciálhatóknak kell lenniük. Ez teljesül a példánkban. A szükséges gradiens vektorok és a gradiens vektorokra vonatkozó lineáris kombináció az alábbiak:

$$\nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \nabla g_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla g_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

$$u_0 \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + u_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + u_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} + u_3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{0}.$$

A Fritz John pont meghatározásához szükséges Fritz John optimalitási feltételek (DF, KF, PF) az alábbiak szerint írhatók, vagyis az alábbi rendszert kell megoldani:

$$\left.\begin{aligned} u_0 2x_1 + u_1 - u_2 + u_3 = 0 \\ u_0 2x_2 - u_1 - u_2 + 2u_3 = 0 \\ u_0, u_1, u_2, u_3 \geqq 0 \\ (u_0, u_1, u_2, u_3) \neq \mathbf{0} \end{aligned}\right\} DF$$

$$\left.\begin{aligned} u_1(x_1 - x_2 - 2) = 0 \\ u_2(-x_1 - x_2 + 4) = 0 \\ u_3(x_1 + 2x_2 - 3) = 0 \end{aligned}\right\} KF$$

$$\left.\begin{aligned} x_1 - x_2 - 2 \leqq 0 \\ -x_1 - x_2 + 4 \leqq 0 \\ x_1 + 2x_2 - 3 \leqq 0 \end{aligned}\right\} PF$$

A fenti rendszer meglehetősen bonyolult, ezért javasoljuk a részekre bontását mégpedig úgy, hogy megoldását az u_0, u_1, u_2, u_3 Lagrange szorzók értékeitől függően 16 esetre bontjuk szét. A 16 eset közül csak kettőt vizsgálunk meg, a többit az olvasóra bízzuk. Csupán a megoldás lényegét mutatjuk be, mivel a későbbiekben sok feladatot fogunk megoldani, de azoknál már nem Fritz John pontot határozunk meg.

1. eset) $u_0 = u_1 = u_2 = 0, u_3 > 0$

A harmadik komplementaritási feltételből (KF) $x_1 + 2x_2 - 3 = 0$, mivel feltevésünk szerint $u_3 > 0$. Ehhez az egyenlethez hozzávéve a duál feltétel (DF) első két egyenletét, az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} u_3 &= 0 \\ 2u_3 &= 0 \\ x_1 + 2x_2 - 3 &= 0 \end{aligned}$$

Megoldásként az $u_3 = 0$ adódik, amely nem lehet, hisz legalább egy Lagrange szorzónak pozitívnak kell lennie. Ez az eset tehát nem ad Fritz John pontot.

2. eset) $u_0 > 0, u_1 = 0,\ u_2 > 0, u_3 = 0$

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} u_0 2x_1 - u_2 &= 0 \\ u_0 2x_2 - u_2 &= 0 \\ -x_1 - x_2 + 4 &= 0 \end{aligned}$$

Az első két egyenletből $x_1 = x_2$ adódik, a harmadikból pedig $x_1 = 2,\ x_2 = 2$. A Lagrange szorzók: $u_0 = \alpha, u_2 = 4\alpha$, ahol $\alpha > 0$. Ez a megoldás maradéktalanul teljesíti a többi Fritz John feltételt, így az $\bar{\mathbf{x}} = (2, 2)$ vektor Fritz John pont az $u_0 = \alpha, u_1 = 0,\ u_2 = 4\alpha, u_3 = 0$ $(\alpha > 0)$ Lagrange szorzókkal úgy, hogy a g_2 feltétel aktív, a g_1 és a g_3 feltétel inaktív.

Megjegyzés:

A többi esetet megvizsgálva azt fogjuk kapni, hogy nincs több Fritz John pont.

3. Példa
Határozzuk meg az alábbi optimalizálási feladathoz tartozó Fritz John pontot!

$$\begin{aligned} x_1^2 &\to \min! \\ (x_1-1)^3 &\geqq x_2 \\ x_2 &\geqq 0 \end{aligned}$$

Megoldás
A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1,x_2) &= x_1^2 \\ g_1(x_1,x_2) &= -(x_1-1)^3 + x_2 \leqq 0 \\ g_2(x_1,x_2) &= -x_2 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A példa függvényei differenciálhatók. A szükséges gradiens vektorok és a gradiens vektorokra vonatkozó lineáris kombináció:

$$\boldsymbol{\nabla} f = \begin{bmatrix} 2x_1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{\nabla} g_1 = \begin{bmatrix} -3(x_1-1)^2 \\ 1 \end{bmatrix}, \qquad \boldsymbol{\nabla} g_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

$$u_0 \begin{bmatrix} 2x_1 \\ 0 \end{bmatrix} + u_1 \begin{bmatrix} -3(x_1-1)^2 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \mathbf{0}.$$

A Fritz John pont meghatározásához szükséges Fritz John optimalitási feltételek az alábbiak szerint írhatók, vagyis az alábbi rendszert kell megoldani:

$$\begin{aligned} u_0 2x_1 - 3u_1(x_1-1)^2 &= 0 \\ u_1 - u_2 &= 0 \\ u_1\left[-(x_1-1)^3 + x_2\right] &= 0 \\ u_2(-x_2) &= 0 \\ -(x_1-1)^3 + x_2 &\leqq 0 \\ -x_2 &\leqq 0 \\ u_0, u_1, u_2 &\geqq 0 \\ (u_0, u_1, u_2) &\neq \mathbf{0} \end{aligned}$$

A 2. egyenlőség szerint $u_1 = u_2$. A megoldást két részre bontjuk, egyik eset a $u_1 = u_2 = 0$, másik eset az $u_1 = u_2 > 0$. Nézzük az egyes eseteket.

i) $u_1 = u_2 = 0$ eset
Mivel mindhárom Lagrange szorzó nem lehet zérus, így u_0-nak pozitívnak kell lenni. Ekkor az 1. egyenletből $x_1 = 0$ adódik. Az 5. sorbeli egyenlőtlenségből pedig $x_2 \leqq -1$, ami ellentmond a 6. sorbeli egyenlőtlenségnek ($x_2 \geqq 0$). Ez az eset tehát nem ad Fritz John pontot.

ii) $u_1 = u_2 > 0$ eset
A komplementaritási feltételekből azonnal adódik az $\mathbf{x} = (1,0)$ pont. Az 1. egyenletből pedig $u_0 = 0$ adódik. Az összes feltétel teljesül, így a Fritz John pont $\mathbf{x} = (1,0)$, a Lagrange szorzók pedig $u_0 = 0$, $u_1 = u_2 = \alpha$, ahol $\alpha > 0$.

5.3. Karush-Kuhn-Tucker szükséges feltételek

A fenti példában láttuk, hogy a Lagrange szorzók nem egyértelműen határozhatók meg, a következőkben már csak olyan Lagrange szorzókkal foglalkozunk, amelynél kikötjük, hogy a célfüggvény együtthatójának (u_0) értéke 1 legyen. Ehhez természetesen valamilyen előírásnak kell teljesülnie.

TÉTEL (Karush-Kuhn-Tucker szükséges feltételek I., Karush 1939; Kuhn, Tucker 1951)

Tekintsük a $\min\{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot, amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen adott egy $\bar{\mathbf{x}} \in S$ pont, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i, i \in I$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók, a $g_i, i \notin I$ függvények az $\bar{\mathbf{x}}$ pontban folytonosak. Tegyük fel továbbá, hogy a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})$, $i \in I$ vektorok lineárisan függetlenek.

Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek **egyértelműen** olyan u_i, $i \in I$ számok, hogy

$$\begin{aligned} \boldsymbol{\nabla} f(\bar{\mathbf{x}}) + \sum_{i \in I} u_i \boldsymbol{\nabla} g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i &\geqq 0, \quad i \in I \end{aligned}$$

Bizonyítás

A Fritz John tétel szerint léteznek olyan $u_0, \hat{u}_i$, $i \in I$ számok, amelyek közül legalább egy pozitív, úgy, hogy

$$\begin{aligned} u_0 \boldsymbol{\nabla} f(\bar{\mathbf{x}}) + \sum_{i \in I} \hat{u}_i \boldsymbol{\nabla} g_i(\bar{x}) &= \mathbf{0} \\ u_0, \hat{u}_i &\geqq 0, \quad i \in I \end{aligned}$$

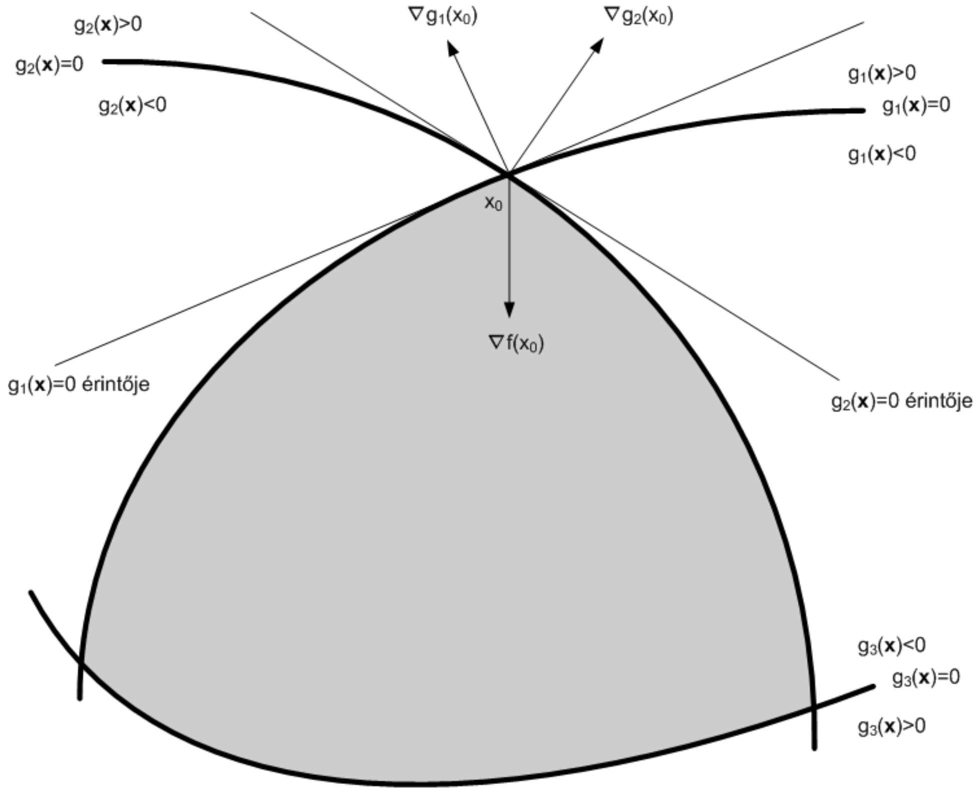
Amennyiben a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})$, $i \in I$ gradiens vektorok lineárisan függetlenek, akkor szükségképpen $u_0 > 0$. Ha ugyanis $u_0 = 0$ lenne, akkor a lineáris függetlenség miatt minden $\hat{u}_i$ számnak is zérusnak kellene lenni, pedig az összes között legalább egynek pozitívnak kell lennie. Az $u_0 > 0$ számmal való végigosztás és a $u_i = \frac{\hat{u}_i}{u_0}$ választással a tétel állítása azonnal adódik. A tétel tehát azt fejezi ki, hogy optimális megoldásnál a célfüggvény gradiens vektorának (-1)-szerese előállítható az aktív feltételi függvények nemnegatív lineáris kombinációjával. Az u_i, $i \in I$ számok egyértelmű létezése pedig abból az ismert tételből fakad, miszerint a lineárisan független vektorokkal történő előállítás egyértelmű. **Q.e.d.**

A tétel jobb megértése végett tekintsük azt a kétváltozós ($\mathbf{x} = (x_1, x_2)$ optimalizálási feladatot, amelyben három feltételi függvény van, a feltételek tehát: $g_1(\mathbf{x}) \leqq 0$, $g_2(\mathbf{x}) \leqq 0$, $g_3(\mathbf{x}) \leqq 0$. A feltételi függvények meghatározzák a feltételi halmazt, amelyet az (x_1, x_2) síkon az alábbi ábra mutat. Ezt úgy kell meghatározni, hogy ábrázoljuk a $g_i(\mathbf{x}) = 0$ $(i = 1, 2, 3)$ függvényeket, amelyek egy-egy görbét jelentenek. Megállapítjuk azokat a tartományokat, ahol $g_i(\mathbf{x}) \leqq 0$, ezek a tartományok a görbéknek és valamelyik oldaluknak a pontjaiból állnak. A három tartomány közös része alkotja az optimalizálási feladat feltételi halmazát.

Legyen adott az $\bar{\mathbf{x}}$ pont (az ábrában $\mathbf{x}_0$-al jelöltük), amely a $g_1(\mathbf{x}) = 0$ és a $g_2(\mathbf{x}) = 0$ görbék metszéspontjában van. Így az első és második feltétel aktív, a harmadik feltétel inaktív. Ez utóbbi az optimalizálás vizsgálatában semmiféle szerepet nem fog játszani. Most kiszámítjuk az $\bar{\mathbf{x}} = \mathbf{x}_0$ pontban a célfüggvény és az aktív feltételi függvények gradiensét, ezeket ábrázoljuk. Az aktív feltételek gradiens vektorát könnyű berajzolni, mert mint tudjuk, a gradiens vektor a szintvonalra merőleges, tehát a $g_1(\mathbf{x}) = 0$, ill. a $g_2(\mathbf{x}) = 0$ görbékre merőleges az aktív feltételek gradiens vektora. Az is tudjuk, hogy a tartomány belsejében $g_i(\mathbf{x}) < 0$, a határán $g_i(\mathbf{x}) = 0$, kívül $g_i(\mathbf{x}) > 0$. A gradiens pedig a függvény legnagyobb növekedésének irányába mutat. Ebből következik, hogy ezek a gradiens vektorok a tartományból kifelé mutatnak.

A példában láthatjuk, hogy az aktív feltételek gradiens vektorai lineárisan függetlenek, tehát a fenti tétel kikötése fennáll. A optimalizálás szükséges feltétele a következőt mondja ki. Ha az $\bar{\mathbf{x}} = \mathbf{x}_0$ pontról tudjuk, hogy az lokális minimumpont, akkor a célfüggvény gradiens vektorának (-1)-szerese előállítható az aktív feltételi függvények gradienseinek nemnegatív lineáris kombinációjával. Szemléletesen ez azt jelenti, hogy a $-\boldsymbol{\nabla} f(\bar{\mathbf{x}})$ vektor az aktív feltételi függvények gradiensei által meghatározott kúpban van.

Ne feledkezzünk meg arról, hogy a tétel azt nem állítja, hogy ilyen esetben egy $\bar{\mathbf{x}} = \mathbf{x}_0$ pont lokális minimumpont lenne. A kúpban való elhelyezkedés csak szükséges előírás, de nem biztos, hogy elegendő is ahhoz, hogy optimális pont legyen az $\bar{\mathbf{x}} = \mathbf{x}_0$ pont.



Mint láttuk az aktív feltételekhez tartozó gradiens vektorok lineáris függetlensége biztosította a célfüggvényhez tartozó u_0 szorzó pozitivitását. Más feltételek is biztosítják u_0 pozitivitását. Ezeket a feltételeket regularitási feltételeknek nevezzük. Az olyan lehetséges megoldást, amelynél ezek a regularitási feltételek is fennállnak reguláris megoldásnak (reguláris pontnak) nevezzük.

DEFINÍCIÓ (Reguláris pont egyik definíciója)

Az $\bar{\mathbf{x}} \in S$ lehetséges megoldást **regulárisnak** nevezzük, ha az aktív feltételekhez tartozó $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}}),\ i \in I$ vektorok lineárisan függetlenek.

A következőkben a fenti tétel (Karush-Kuhn-Tucker szükséges feltételek I.) használhatóbb változatát ismertetjük.

TÉTEL (Karush-Kuhn-Tucker szükséges feltételek II., Karush 1939; Kuhn, Tucker 1951)
Tekintsük a $\min\{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot, amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen adott egy $\bar{\mathbf{x}} \in S$ pont, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i, i = 1, 2, \ldots, m$ függvények az $\bar{\mathbf{x}}$ pontban differenciálhatók. Tegyük fel továbbá, hogy a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})$, $i \in I$ vektorok lineárisan függetlenek, azaz az $\bar{\mathbf{x}} \in S$ reguláris pont.
Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek egyértelműen olyan $u_i,\ i = 1, 2, \ldots, m$ számok, hogy

$$\begin{aligned}
\boldsymbol{\nabla} f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\
u_i g_i(\bar{\mathbf{x}}) &= 0, \qquad i = 1, 2, \ldots, m \\
u_i &\geqq 0, \qquad i = 1, 2, \ldots, m
\end{aligned}$$

Bizonyítás
Ha minden $i = 1, 2, \ldots, m$ indexre előírjuk, hogy $u_i g_i(\bar{\mathbf{x}}) = 0$, akkor ez a feltétel biztosítja, hogy az inaktív feltételek ($g_i(\bar{\mathbf{x}}) < 0$) esetén $u_i = 0$ legyen. **Q.e.d.**

Hasonlóan a Fritz John tételnél mondottakhoz a tételben szereplő u_i számokat Lagrange szorzóknak nevezik. Az $\mathbf{x} \in S$ lehetséges megoldásra vonatkozó feltételeket primál feltételeknek nevezzük. A tételben szereplő

$$\begin{aligned}
\boldsymbol{\nabla} f(\mathbf{x}) + \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\mathbf{x}) &= \mathbf{0} \\
u_i &\geqq 0 \quad i = 1, 2, \ldots, m
\end{aligned}$$

összefüggéseket duál feltételeknek nevezzük. Az $u_i g_i(\mathbf{x}) = 0$ egyenlőséget pedig komplementaritási feltételnek nevezzük. Együttesen a primál feltételeket (PF), a duál feltételeket (DF) és a komplementaritási feltételeket (KF) Karush-Kuhn-Tucker optimalitási feltételeknek (KKT) hívjuk. Bármely olyan $\bar{\mathbf{x}}$ pontot, amelyre léteznek $\bar{\mathbf{u}} = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m)$ Lagrange szorzók úgy, hogy az $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ vektor kielégíti a Karush-Kuhn-Tucker optimalitási feltételeket Karush-Kuhn-Tucker pontnak (röviden KKT pontnak) nevezünk. Amennyiben a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})$, $i \in I$ vektorok lineárisan függetlenek, akkor a Lagrange szorzók egyértelműen meghatározottak.

4. Példa

Határozzuk meg az alábbi optimalizálási feladathoz tartozó Karush-Kuhn-Tucker (KKT) pontot!

$$\begin{aligned} x_1x_2 &\rightarrow \text{min!} \\ 4x_1^2+x_2^2 &\leqq 8 \end{aligned}$$

Megoldás

A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1,x_2) &= x_1x_2 \\ g(x_1,x_2) &= 4x_1^2+x_2^2-8 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A példa függvényei differenciálhatók. A szükséges gradiens vektorok és a gradiens vektorokra vonatkozó lineáris kombináció:

$$\boldsymbol{\nabla} f = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}, \qquad \boldsymbol{\nabla} g = \begin{bmatrix} 8x_1 \\ 2x_2 \end{bmatrix},$$

$$\begin{bmatrix} x_2 \\ x_1 \end{bmatrix} + u \begin{bmatrix} 8x_1 \\ 2x_2 \end{bmatrix} = \mathbf{0}.$$

A KKT pont meghatározásához szükséges KKT optimalitási feltételek (DF, KF, PF) az alábbiak szerint írhatók, vagyis az alábbi rendszert kell megoldani:

$$\left.\begin{aligned} x_2+8ux_1=0 \\ x_1+2ux_2=0 \\ u\geqq 0 \end{aligned}\right\} DF$$
$$u(4x_1^2+x_2^2-8)=0\} \, KF$$
$$4x_1^2+x_2^2-8\leqq 0\} \, PF$$

Az u Lagrange szorzóra vonatkozó nemnegativitási feltételnek megfelelően a megoldást két részre bontjuk, egyik eset az $u>0$, másik eset az $u=0$. Nézzük az egyes eseteket.

i) $u>0$ eset

A komplementaritási feltételből következik, hogy az 5. sorbeli egyenlőtlenség egyenlőség lesz, azaz a feltétel aktív lesz. Ekkor az 1., 2. és az 5. sorban szereplő most már egyenletből az alábbi egyenletrendszert kell megoldani az x_1, x_2, u mennyiségek meghatározásához

$$\begin{aligned} x_2+8ux_1 &= 0 \\ x_1+2ux_2 &= 0 \\ 4x_1^2+x_2^2-8 &= 0 \end{aligned}$$

Az egyenletrendszerre négy megoldást kapunk, melyek az alábbiak

i	x_1	x_2	u
1	1	-2	$1/4$
2	-1	2	$1/4$
3	1	2	$-1/4$
4	-1	-2	$-1/4$

Ezekből a megoldásokból csak az első kettő KKT pont, mivel ezeknél pozitív a Lagrange szorzó.

ii) $u = 0$ eset
Az 1., 2. sorban szereplő egyenletekből egyszerűen adódik, hogy

$$x_1 = 0, \quad x_2 = 0.$$

Az 5. sorbeli egyenlőtlenség is fennáll, így az $\mathbf{x} = (0, 0)$ pont KKT pont az $u = 0$ Lagrange szorzóval.

Összefoglalva három KKT pont adódott, amelyek a Lagrange szorzóikkal együtt az alábbiak:
1. KKT pont: $x_1 = 1, \quad x_2 = -2, \quad u = \frac{1}{4}$, a feltétel aktív.
2. KKT pont: $x_1 = -1, \quad x_2 = 2, \quad u = \frac{1}{4}$, a feltétel aktív.
3. KKT pont: $x_1 = 0, \quad x_2 = 0, \quad u = 0$, a feltétel inaktív.

Mint tudjuk a KKT pontok csupán az optimum szükséges feltételeit elégítik ki, e három pont tehát nem biztos, hogy optimális megoldás. Az optimalitás elégséges feltételei fogják megmutatni, hogy ezek között melyek optimálisak.

5. Példa
Határozzuk meg az alábbi optimalizálási feladathoz tartozó Karush-Kuhn-Tucker (KKT) pontot!

$$\begin{aligned} x_1^2 &\rightarrow \min! \\ (x_1 - 1)^3 &\geqq x_2 \\ x_2 &\geqq 0 \end{aligned}$$

Megoldás
Ez a példa azonos a Fritz John pontnál tárgyalt 2. példával, de most KKT pontot akarunk meghatározni. A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1 - 3u_1(x_1 - 1)^2 &= 0 \\ u_1 - u_2 &= 0 \\ u_1\left[-(x_1 - 1)^3 + x_2\right] &= 0 \\ u_2(-x_2) &= 0 \\ -(x_1 - 1)^3 + x_2 &\leqq 0 \\ -x_2 &\leqq 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A 2. egyenlőség szerint $u_1 = u_2$. A megoldást két részre bontjuk, egyik eset a $u_1 = u_2 = 0$, másik eset az $u_1 = u_2 > 0$. Nézzük az egyes eseteket.

i) $u_1 = u_2 = 0$ eset
Az 1. egyenletből $x_1 = 0$ adódik. Az 5. sorbeli egyenlőtlenségből pedig $x_2 \leqq -1$, ami ellentmond a 6. sorbeli egyenlőtlenségnek.

ii) $u_1 = u_2 > 0$ eset

A komplementaritási feltételekből azonnal adódik, hogy $x_2 = 0, x_1 = 1$, azaz az $\mathbf{x} = (1, 0)$ pont. Azonban az 1. egyenlet nem teljesül.

A két esetet egybevetve tehát nincs KKT pont.

Egyébként ennek a feladatnak az optimális megoldása $\mathbf{x} = (1, 0)$. Valószínű az olvasó meglepődik azon, hogy nem találtunk KKT pontot, így az optimális megoldást sem sikerült a KKT ponton keresztül meghatározni. Mi volt az oka, hogy nem találtunk KKT pontot? Az $\mathbf{x} = (1, 0)$ pontban mindkét feltétel aktív, számítsuk ki ezekhez az aktív feltételekhez tartozó gradiens vektorokat, amelyek: $\boldsymbol{\nabla} g_1(\mathbf{x}) = (0, 1)$, ill. $\boldsymbol{\nabla} g_2(\mathbf{x}) = (0, -1)$. A KKT tétel csak akkor garantál KKT pontot, ha az aktív feltételekhez tartozó gradiens vektorok lineárisan függetlenek, esetünkben pedig ez nem áll fenn. Tehát nem teljesül a lineáris függetlenségi regularitási feltétel, ez okozta, hogy nem találtunk KKT pontot.

6. Szükséges feltételek egyenlőtlenségi és egyenlőségi előírások mellett

Ha az S feltételi halmaz megadásánál egyenlőségi feltételek is szerepelnek, akkor egy általánosabb matematikai programozási feladatot kapunk, amely az alábbiak szerint írható:

$$\min \{f(\mathbf{x}) : \mathbf{x} \in S\},$$

amelyben a S feltételi halmaz az alábbi:

$$S = \{\mathbf{x} : \mathbf{x} \in X, g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m;\ h_i(\mathbf{x}) = 0, i = 1, 2, \ldots, k\},$$

ahol $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények. A továbbiakban az ilyen feladatot az alábbi formában írjuk le:

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

6.1. Geometriai-algebrai szükséges feltétel

TÉTEL (geometriai-algebrai szükséges feltétel)

Legyen $X \subset \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Tegyük fel, hogy $\bar{\mathbf{x}}$ lokális optimális megoldás és legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív egyenlőtlenséges feltételi függvények indexhalmaza. Tegyük fel továbbá, hogy az $f, g_i,\ i \in I$ függvények differenciálhatók az $\bar{\mathbf{x}}$ pontban, a $g_i,\ i \notin I$ függvények folytonosak az $\bar{\mathbf{x}}$ pontban és minden $h_i,\ i = 1, 2, \ldots, k$ függvény folytonosan differenciálható az $\bar{\mathbf{x}}$ pontban.

Ha az $\bar{\mathbf{x}}$ lokális optimális megoldásban a $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ vektorok lineárisan függetlenek, akkor $F_0 \cap G_0 \cap H_0 = \emptyset$, ahol

$$
\begin{aligned}
F_0 &= \{\mathbf{d} : \nabla f(\bar{\mathbf{x}})\mathbf{d} < 0\} \\
G_0 &= \{\mathbf{d} : \nabla g_i(\bar{\mathbf{x}})\mathbf{d} < 0 \ \ i \in I\} \\
H_0 &= \{\mathbf{d} : \nabla h_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ \ i = 1, 2, \ldots, k\}
\end{aligned}
$$

Az alábbiakban két szükséges feltételt mondunk ki tétel (FJ és KKT) formájában, de az előzőektől eltérően az I. és II. sorszámú tételeket összevonjuk.

6.2. Fritz John szükséges feltételek

TÉTEL (Fritz John szükséges feltételek)

Legyen $X \subset \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények. Tekintsük az

$$
\begin{aligned}
&f(\mathbf{x}) \to \min! \\
&g_i(\mathbf{x}) \leqq 0 \qquad i = 1, 2, \ldots, m \\
&h_i(\mathbf{x}) = 0 \qquad i = 1, 2, \ldots, k \\
&\mathbf{x} \in X
\end{aligned}
$$

matematikai programozási feladatot. Legyen $\bar{\mathbf{x}}$ lehetséges megoldás, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív egyenlőtlenséges feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i,\ i \in I$ függvények differenciálhatók az $\bar{\mathbf{x}}$ pontban, a $g_i,\ i \notin I$ függvények folytonosak az $\bar{\mathbf{x}}$ pontban és minden $h_i,\ i = 1, 2, \ldots, k$ függvény folytonosan differenciálható az $\bar{\mathbf{x}}$ pontban.

Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek $u_0, u_i\ i \in I$ és $v_i\ i = 1, 2, \ldots, k$ számok úgy, hogy

$$
\begin{aligned}
u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\
(u_0, \mathbf{u}_I) &\geqq \mathbf{0} \\
(u_0, \mathbf{u}_I, \mathbf{v}) &\neq \mathbf{0}
\end{aligned}
$$

ahol $\mathbf{u}_I$ egy olyan vektor, amelynek komponensei $u_i \geqq 0,\ i \in I$, a $\mathbf{v}$ pedig egy olyan vektor, amelynek komponensei $v_i \in \mathbb{R},\ i = 1, 2, \ldots, k$.

Továbbá, ha $g_i,\ i \notin I$ függvények is differenciálhatók az $\bar{\mathbf{x}}$ pontban, akkor a fenti feltételeket az alábbi ekvivalens formában is írhatjuk:

$$
\begin{aligned}
u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\
u_i g_i(\bar{\mathbf{x}}) &= 0 \ \ i = 1, 2, \ldots, m \\
(u_0, \mathbf{u}) &\geqq \mathbf{0} \\
(u_0, \mathbf{u}, \mathbf{v}) &\neq \mathbf{0}
\end{aligned}
$$

ahol $\mathbf{u}$ egy olyan vektor, amelynek komponensei $u_i \geqq 0,\ i = 1, 2, \ldots, m$; a $\mathbf{v}$ pedig egy olyan vektor, amelynek komponensei $v_i \in \mathbb{R},\ i = 1, 2, \ldots, k$.

Bizonyítás

Két esetet vizsgálunk a $\nabla h_i(\bar{\mathbf{x}})$ vektorok szerint:

Ha a $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ vektorok lineárisan összefüggők, akkor a v_i számok között található nemzérus úgy, hogy $\sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) = \mathbf{0}$. Az $u_0, u_i, i = 1, 2, \ldots, m$ számokat, ha zérusra választjuk, akkor a tétel állítása igaz.

Ha a $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ vektorok lineárisan függetlenek. A gradiens vektorokból képezzük az $\mathbf{A}_1$ és az $\mathbf{A}_2$ mátrixokat az alábbiak szerint. Az $\mathbf{A}_1$ mátrix első oszlopvektora legyen $\nabla f(\bar{\mathbf{x}})$, a többi oszlopvektora pedig a $\nabla g_i(\bar{\mathbf{x}}), i \in I$ vektorok. Az $\mathbf{A}_2$ mátrix oszlopvektorai legyenek a $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ vektorok. Az $\bar{\mathbf{x}}$ optimalitásának geometriai szükséges feltétele szerint a

$$\begin{aligned} \mathbf{dA}_1 &< \mathbf{0} \\ \mathbf{dA}_2 &= \mathbf{0} \end{aligned}$$

rendszernek nincs megoldása. Mivel a feltevés szerint az $\mathbf{A}_2$ mátrix oszlopvektorai lineárisan függetlenek, így a szeparációs tétel egyik (másik tananyagban ismertetett) következménye szerint léteznek olyan $(\mathbf{x}_1, \mathbf{x}_2) \neq \mathbf{0}, \mathbf{x}_1 \geqq \mathbf{0}$ vektorok, hogy

$$\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{0}.$$

Jelölje az $\mathbf{x}_1$ vektor első komponensét u_0, a többit pedig $u_i, i \in I$, az $\mathbf{x}_2$ vektor komponenseit pedig $v_i, i = 1, 2, \ldots, k$, innen pedig következik a tétel első része.

Ha az összes indexre ki akarjuk terjeszteni a tételt, ez csak akkor lesz a tétel első részével ekvivalens, ha az u_i értékek az $i \notin I$ indexekre zérusak. Ezt úgy biztosíthatjuk, hogy minden $i = 1, 2, \ldots, m$ indexre előírjuk, hogy $u_i g_i(\bar{\mathbf{x}}) = 0$. Könnyen látható, hogy $i \notin I$ esetén, azaz, ha $g_i(\bar{\mathbf{x}}) < 0$, akkor $u_i = 0$ következik. **Q.e.d.**

6.3. Karush-Kuhn-Tucker szükséges feltételek

TÉTEL (Karush-Kuhn-Tucker szükséges feltételek)

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Legyen $\bar{\mathbf{x}}$ lehetséges megoldás, legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ az aktív egyenlőtlenséges feltételi függvények indexhalmaza. Tegyük fel, hogy az $f, g_i,\ i \in I$ függvények differenciálhatók az $\bar{\mathbf{x}}$ pontban, a $g_i,\ i \notin I$ függvények folytonosak az $\bar{\mathbf{x}}$ pontban és minden $h_i,\ i = 1, 2, \ldots, k$ függvény folytonosan differenciálható az $\bar{\mathbf{x}}$ pontban. Tegyük fel továbbá, hogy a $\nabla g_i(\bar{\mathbf{x}}),\ i \in I$ és a $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ vektorok lineárisan függetlenek, más szóval az $\bar{\mathbf{x}}$ lehetséges megoldás reguláris pont.

Ha $\bar{\mathbf{x}}$ lokális optimális megoldás, akkor léteznek egyértelműen olyan $u_i,\ i \in I$ és $v_i,\ i = 1, 2, \ldots, k$ számok, hogy

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i &\geqq 0, \quad i \in I \end{aligned}$$

Továbbá, ha $g_i,\ i \notin I$ függvények is differenciálhatók az $\bar{\mathbf{x}}$ pontban, akkor a fenti feltételeket az alábbi ekvivalens formában is írhatjuk:

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\bar{\mathbf{x}}) &= 0, \quad i = 1, 2, \ldots, m \\ u_i &\geqq 0, \quad i = 1, 2, \ldots, m \end{aligned}$$

Bizonyítás

A Fritz John tétel szerint léteznek olyan $u_0, \hat{u}_i,\ i \in I,\ \bar{v}_i\ i = 1, 2, \ldots, k$ számok, amelyek nem mindegyike zérus, úgy, hogy

$$\begin{aligned} u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} \hat{u}_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1 \bar{v} i}^{k} \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_0, \hat{u}_i &\geqq 0 \ \ i \in I \end{aligned}$$

Amennyiben a $\nabla g_i(\bar{\mathbf{x}}),\ i \in I$ és $\nabla h_i(\bar{\mathbf{x}}),\ i = 1, 2, \ldots, k$ gradiens vektorok lineárisan függetlenek, akkor $u_0 > 0$. Ha ugyanis $u_0 = 0$ lenne, akkor a lineáris függetlenség miatt minden $\hat{u}_i$ és $\bar{v}_i$ számnak is zérusnak kellene lenni, pedig az összes között legalább egynek zérustól különbözőnek kell lennie. A $u_0 > 0$ számmal való végigosztás, valamint az $u_i = \frac{\hat{u}_i}{u_0}$, $v_i = \frac{\bar{v}_i}{u_0}$ választással a tétel első része azonnal adódik. A lineáris függetlenségi feltételezés biztosítja a Lagrange szorzók egyértelműségét. Az ekvivalens második rész is nyilvánvaló. **Q.e.d.**

Az egyenlőségeket és egyenlőtlenségeket is tartalmazó optimalizálási probléma esetén is használjuk az eddig ismertetett fogalmakat, nevezetesen a primál feltételek (PF), duál feltételek (DF), komplementaritási feltételek (KF), Fritz-John (FJ) feltételek, ill. FJ pont, Karush-Kuhn-Tucker (KKT) feltételek, ill. KKT pont.

A későbbiekben legtöbbször a KKT feltételeket használjuk, ezért összefoglalásképpen közöljük ezeket

$$\left.\begin{array}{ll} g_i(\mathbf{x}) \leqq 0, & i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) = 0, & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & \end{array}\right\} PF$$

$$\left.\begin{array}{c} \nabla f(\mathbf{x}) + \sum\limits_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum\limits_{i=1}^{k} v_i \nabla h_i(\mathbf{x}) = \mathbf{0} \\ u_i \geqq 0, \quad i = 1, 2, \ldots, m \end{array}\right\} DF$$

$$\left. u_i g_i(\mathbf{x}) = 0, \quad i = 1, 2, \ldots, m \ \right\} KF$$

6. Példa

Határozzuk meg az alábbi optimalizálási feladathoz tartozó Karush-Kuhn-Tucker (KKT) pontot!

$$\begin{aligned} x_1^2 - 12x_1 + x_2^2 - 4x_2 + 5 &\rightarrow \text{min!} \\ x_1^2 + x_2^2 &\leqq 25 \\ x_1 + 4x_2 &= 16 \end{aligned}$$

Megoldás

A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1,x_2) &= x_1^2 - 12x_1 + x_2^2 - 4x_2 + 5 \\ g(x_1,x_2) &= x_1^2 + x_2^2 - 25 \leqq 0 \\ h(x_1,x_2) &= x_1 + 4x_2 - 16 = 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A példa függvényei differenciálhatók. A szükséges gradiens vektorok és a gradiens vektorokra vonatkozó lineáris kombináció:

$$\boldsymbol{\nabla} f = \begin{bmatrix} 2x_1 - 12 \\ 2x_2 - 4 \end{bmatrix}, \qquad \boldsymbol{\nabla} g = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \qquad \boldsymbol{\nabla} h = \begin{bmatrix} 1 \\ 4 \end{bmatrix},$$

$$\begin{bmatrix} 2x_1 - 12 \\ 2x_2 - 4 \end{bmatrix} + u \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + v \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \mathbf{0}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1 - 12 + u2x_1 + v &= 0 \\ 2x_2 - 4 + u2x_2 + 4v &= 0 \\ u(x_1^2 + x_2^2 - 25) &= 0 \\ x_1^2 + x_2^2 - 25 &\leqq 0 \\ x_1 + 4x_2 - 16 &= 0 \\ u &\geqq 0 \end{aligned}$$

Mint ahogy már láttuk, ilyen rendszer megoldására az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset a $u = 0$, másik eset a $u > 0$. Nézzük az egyes eseteket.

i) $u = 0$ eset
Az 1., 2. és az 5 sorban szereplő egyenletekből az x_1, x_2, v számokra kapjuk, hogy

$$x_1 = \frac{104}{17}, \quad x_2 = \frac{42}{17}, \quad v = -\frac{4}{17}.$$

A 3. egyenlet (komplementaritási feltétel) $u = 0$ esetén nem érdekes. Azonban a 4. egyenlőtlenségnek is fenn kell állnia, ezt ellenőrizzük

$$x_1^2 + x_2^2 - 25 = \frac{315}{17} \nleqq 0.$$

Mivel a 4. egyenlőtlenség nem áll fenn, így az $x_1 = \frac{104}{17}$, $x_2 = \frac{42}{17}$ megoldás nem KKT pont.

ii) $u > 0$ eset
A 4. egyenlőtlenségnek a 3. egyenlet miatt egyenlőséggel kell teljesednie. Ekkor az 1., 2., 5. sorbeli egyenletekből és a 4. sorban szereplő, most már egyenletből álló egyenletrendszert kell megoldani az x_1, x_2, u, v mennyiségek meghatározására

$$\begin{aligned} 2x_1 - 12 + u2x_1 + v &= 0 \\ 2x_2 - 4 + u2x_2 + 4v &= 0 \\ x_1^2 + x_2^2 - 25 &= 0 \\ x_1 + 4x_2 - 16 &= 0 \end{aligned}$$

A fenti egyenletrendszer megoldása:

$$x_1 = 4, \quad x_2 = 3, \quad u = \frac{9}{13}, \quad v = -\frac{20}{13}.$$

Mivel $u > 0$ és a $\nabla g = (8,6)$, ill. a $\nabla h = (1,4)$ vektorok lineárisan függetlenek, így az $x_1 = 4,\ x_2 = 3$ megoldás KKT pont az $u = \frac{9}{13}, v = -\frac{20}{13}$ Lagrange szorzókkal.

7. Elégséges feltételek

Az előzőekben részletesen bemutattuk az optimum szükséges feltételeit. Ezek azonban, mint ahogy a feltétel nélküli optimalizálásnál is láttuk, nem garantálják azt, hogy egy FJ pont, ill. egy KKT pont optimális megoldása az optimalizálási feladatnak. Bizonyos konvexitási feltételek garanciát adnak az optimumra. Az alábbiakban másodrendű elégséges feltételeket mutatunk be konvexitási feltevések nélkül. Mielőtt azonban ezeket bemutatnánk bevezetjük a Lagrange függvény fogalmát, amely nem más mint az optimalizálási feladatban szereplő függvények lineáris kombinációja.

7.1. Lagrange függvény

DEFINÍCIÓ (Lagrange függvény definíciója)

A Lagrange függvény az optimalizálási feladatban szereplő célfüggvény, az egyenlőtlenséges és az egyenlőséges feltételi függvények lineáris kombinációja, amelyben a célfüggvény együtthatója 1. A feltételi függvények együtthatóit Lagrange szorzóknak nevezzük. A Lagrange függvény képletben:

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

DEFINÍCIÓ (Aktualizált Lagrange függvény definíciója)

Legyen $\bar{\mathbf{x}}$ vektor az optimalizálási feladat KKT pontja és legyenek $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ vektorok a duál és a komplementaritási feltételt kielégítő Lagrange szorzók. Az adott $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ Lagrange szorzókhoz tartozó Lagrange függvényt aktualizált Lagrange függvénynek nevezzük, képletben az alábbi

$$L_a(\mathbf{x}) = L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\mathbf{x}) + \sum_{i=1}^{m} \bar{u}_i g_i(\mathbf{x}) + \sum_{i=1}^{k} \bar{v}_i h_i(\mathbf{x}).$$

Megjegyzés:

Itt használhatnánk csupán az aktív egyenlőtlenséges feltételeket is a szummában, mivel már ismert a KKT pont. Amíg azonban nem ismert a megoldás, csak maga a feladat, akkor természetes, hogy minden egyenlőtlenséges feltételt szerepeltetni kell.

LEMMA (az aktualizált Lagrange függvény és a célfüggvény kapcsolata)

Igazak az alábbi állítások:

a) $\nabla L_a(\bar{\mathbf{x}}) = \mathbf{0}$,
b) $L_a(\mathbf{x}) \leqq f(\mathbf{x})$ minden primál lehetséges $\mathbf{x}$ megoldásra,
c) $L_a(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$.

Bizonyítás

a) Tekintsük az ismert duál feltételt:

$$\nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} \bar{v}_i \nabla h_i(\bar{\mathbf{x}}) = \mathbf{0}.$$

Azonnal látható, hogy a baloldal ekvivalens az $L_a(\mathbf{x})$ függvény gradiensével az $\mathbf{x} = \bar{\mathbf{x}}$ helyen.

b) $L_a(\mathbf{x}) = f(\mathbf{x}) + \sum\limits_{i=1}^{m} \bar{u}_i g_i(\mathbf{x}) + \sum\limits_{i=1}^{k} \bar{v}_i h_i(\mathbf{x}) \leqq f(\mathbf{x})$, az $L_a(\mathbf{x})$ kifejezés első szummás tagja $\leqq 0$, ami a duál feltétel $\bar{u}_i \geqq 0$ részéből és a primál feltétel $g_i(\mathbf{x}) \leqq 0$ részéből következik, a második szummás tagja pedig zérus, ami a primál feltétel $h_i(\mathbf{x}) = 0$ részéből következik, így igaz az állítás.

c) $L_a(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) + \sum\limits_{i=1}^{m} \bar{u}_i g_i(\bar{\mathbf{x}}) + \sum\limits_{i=1}^{k} \bar{v}_i h_i(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$, az $\bar{u}_i g_i(\bar{\mathbf{x}}) = 0$ komplementaritási feltételből azonnal adódik az állítás. **Q.e.d.**

7.2. Elégséges feltételek

7.2.1. Tételek az optimalitás elégséges feltételeire

TÉTEL (gyenge elégséges feltétel)

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ kétszer differenciálható függvények. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Jelölje S a feltételi halmazt. Legyen $\bar{\mathbf{x}} \in S$ az optimalizálási feladat reguláris KKT pontja és legyenek $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ vektorok a duál és a komplementaritási feltételt kielégítő Lagrange szorzók. Jelölje $\nabla^2 L_a(\mathbf{x})$ az aktualizált Lagrange függvény Hesse mátrixát, amely a következő:

$$\nabla^2 L_a(\mathbf{x}) = \nabla^2 f(\mathbf{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla^2 g_i(\mathbf{x}) + \sum_{i=1}^{k} \bar{v}_i \nabla^2 h_i(\mathbf{x}).$$

Ha $\nabla^2 L_a(\bar{\mathbf{x}})$ pozitív definit, azaz $\mathbf{d}^T \nabla^2 L_a(\bar{\mathbf{x}}) \mathbf{d} > 0$ minden $\mathbf{d} \in \mathbb{R}^n$ ($\mathbf{d} \neq \mathbf{0}$) vektorra, akkor az $\bar{\mathbf{x}} \in S$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak.

Bizonyítás

A lemma a) állítása szerint $\nabla L_a(\bar{\mathbf{x}}) = \mathbf{0}$. Ez az $L_a(\mathbf{x})$ függvény minimumának szükséges feltétele. A feltétel nélküli optimalizálásnál megismert elégséges tétel értelmében, ha

$\boldsymbol{\nabla} L_a(\bar{\mathbf{x}}) = \mathbf{0}$ és $\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$ pozitív definit, akkor az $\bar{\mathbf{x}}$ pont az L_a aktualizált Lagrange függvénynek szigorú lokális minimumpontja, azaz

$$L_a(\bar{\mathbf{x}}) < L_a(\mathbf{x}) \quad \text{minden } \mathbf{x} \in S \cap K_\varepsilon(\bar{\mathbf{x}}) \text{ esetén.}$$

Most felhasználva az $L_a(\mathbf{x})$ függvény és az $f(\mathbf{x})$ függvény kapcsolatát kifejező lemma c) majd b) állításait, írhatjuk, hogy

$$f(\bar{\mathbf{x}}) = L_a(\bar{\mathbf{x}}) < L_a(\mathbf{x}) \leqq f(\mathbf{x}) \quad \text{minden } \mathbf{x} \in S \cap K_\varepsilon(\bar{\mathbf{x}}) \text{ esetén,}$$

ez pedig azt jelenti, hogy az $\bar{\mathbf{x}} \in S$ szigorú lokális minimumpontja az $f(\mathbf{x})$ függvénynek. **Q.e.d.**

Meg kell jegyezni, hogy ez az elégséges tétel csak ritka esetekben ad információt, mert a $\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$ pozitív definitsége azt követeli meg, hogy az $\bar{\mathbf{x}}$ KKT pontból kiinduló összes $\mathbf{d} \neq \mathbf{0}$ irány esetén legyen $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$. Holott elegendő lenne csak bizonyos $\mathbf{d}$ irányokra előírni az aktualizált Lagrange függvény pozitív definitségét, hiszen feltételes optimalizálási feladatról van szó. A következő elégséges tétel már sokkal jobban használható.

TÉTEL (erős elégséges feltétel)
Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ kétszer differenciálható függvények. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Jelölje S a feltételi halmazt. Legyen $\bar{\mathbf{x}}$ vektor az optimalizálási feladat reguláris KKT pontja és legyenek $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ vektorok a duál és a komplementaritási feltételt kielégítő Lagrange szorzók. Legyen $\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$ az aktualizált Lagrange függvény Hesse mátrixa az $\mathbf{x} = \bar{\mathbf{x}}$ pontban. A szokásos $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ indexhalmaz legyen felosztva az $I^+ = \{i \in I : \bar{u}_i > 0\}$ és az $I^0 = \{i \in I : \bar{u}_i = 0\}$ indexhalmazokra. Definiáljuk az alábbi K kúpot

$$K = \left\{\mathbf{d} \neq \mathbf{0} : \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} = 0,\ i \in I^+,\ \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} \leqq 0,\ i \in I^0,\ \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})\mathbf{d} = 0,\ i = 1, \ldots, k\right\}.$$

Ha $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$ minden $\mathbf{d} \in K$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak.

Megjegyzés:
Ha csak a $\mathbf{d} \in K$ irányokra írjuk elő, hogy $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$, akkor azt mondjuk, hogy az aktualizált Lagrange függvény Hesse mátrixa ($\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$) feltételesen pozitív definit. A definitséghez hasonló módon definiálható a feltételes negatív definitség, a feltételes szemidefinitség és a feltételes indefinitség is.

Ahhoz, hogy eldönthessük egy $\bar{\mathbf{x}}$ KKT pontról, hogy az szigorú lokális minimumpont vagy sem, kúpon való pozitív definitséget kell vizsgálnunk. Sajnos erre nem állnak rendelkezésünkre tesztek. A mátrixok definitségével foglalkozó tananyagban ismertetett tesztek ugyanis nem kúpra, hanem altérre vonatkoznak, ezért gyengítjük az optimalitás fenti erős elégséges feltételét úgy, hogy az aktualizált Lagrange függvény Hesse mátrixának pozitív definitségét kúp helyett egy altérre vonatkoztatjuk, így már a teszteket lehet alkalmazni, igaz az alkalmazási körüket ezzel egy kicsit leszűkítettük.

TÉTEL (használható elégséges feltétel)
Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ kétszer differenciálható függvények. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Jelölje S a feltételi halmazt. Legyen $\bar{\mathbf{x}}$ vektor az optimalizálási feladat reguláris KKT pontja és legyenek $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ vektorok a duál és a komplementaritási feltételt kielégítő Lagrange szorzók. Legyen $\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$ az aktualizált Lagrange függvény Hesse mátrixa az $\mathbf{x} = \bar{\mathbf{x}}$ pontban. A szokásos $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$ indexhalmaz legyen felosztva az $I^+ = \{i \in I : \bar{u}_i > 0\}$ és az $I^0 = \{i \in I : \bar{u}_i = 0\}$ indexhalmazokra. Definiáljuk az alábbi A alteret

$$A = \left\{ \mathbf{d} \neq \mathbf{0} : \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i \in I^+, \ \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i = 1, \ldots, k \right\}.$$

Ha $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak.

Bizonyítás
A K feltételi halmazból elhagytuk a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} \leqq 0, \ i \in I^0$ megkötéseket, ezáltal a K halmaznál bővebb A halmazon írtuk elő a pozitív definitséget. Ha ezen az A bővebb halmazon a feltételes pozitív definitség fennáll, akkor az eredeti K szűkebb halmazon is fennáll. Nyilvánvaló természetesen, hogy az utóbbi elégséges tétel állítása gyengébb, mint az előző, viszont ez utóbbi tétel állításának eldöntésére rendelkezésünkre állnak tesztek.

Megjegyzés:
Ha az aktív feltételekhez tartozó összes $\bar{u}_i > 0$, akkor az erős tétel érvényes, hiszen $K = A$ és ekkor biztosan eldönthető az optimalitás. Ha viszont van olyan aktív feltételünk, amelyhez tartozó Lagrange szorzó zérus és a használható elégséges feltétel szerint a $\boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})$ mátrix nem feltételesen pozitív definit, akkor ez még nem jogosít fel bennünket arra, hogy elvessük az adott pont optimalitását. Ebben az esetben egyéb vizsgálatok kellenek az optimalitás eldöntéséhez.

Most azt mutatjuk meg, hogyan használhatjuk a teszteket az optimalitás elégséges feltétele teljesülésének eldöntésére. Mint tudjuk egy $\mathbf{A}$ mátrix akkor feltételesen pozitív definit egy $\mathbf{B}$ mátrixra nézve, ha $\mathbf{d}^T \mathbf{A} \mathbf{d} > 0$ minden $\mathbf{d} \neq \mathbf{0}, \mathbf{B}\mathbf{d} = \mathbf{0}$ vektor esetén. Azt is tudjuk, hogy a feltételes pozitív definitséget egy $\mathbf{C}$ szegélyezett mátrix segítségével lehet eldönteni. A használható elégséges feltétel az $L_a(\bar{\mathbf{x}})$ mátrix pozitív definitségét írja elő minden olyan $\mathbf{d} \neq \mathbf{0}$ vektorra, amely merőleges a $\boldsymbol{\nabla} g_i(\bar{\mathbf{x}}), \ i \in I^+$ és a $\boldsymbol{\nabla} h_i(\bar{\mathbf{x}}), \ i = 1, \ldots, k$ lineárisan független vektorokra. Tehát az optimalizálási feladat esetében az $\mathbf{A}$ mátrix az $L_a(\bar{\mathbf{x}})$ mátrix, azaz az aktualizált Lagrange függvény $\bar{\mathbf{x}}$ pontbeli Hesse mátrixa. A $\mathbf{B}$ mátrix pedig olyan, amelynek sorvektorai az aktív egyenlőtlenséges feltételi függvények $\bar{\mathbf{x}}$ pontbeli gradiensei és az egyenlőséges feltételi függvények $\bar{\mathbf{x}}$ pontbeli gradiensei. Csak azoknak az aktív egyenlőtlenséges feltételi függvényeknek a gradienseit kell szerepeltetni, amelyeknél a Lagrange szorzó pozitív. A $\mathbf{B}$ mátrix teljes sorrangú, mivel a gradiens vektorok lineárisan függetlenek, hisz $\bar{\mathbf{x}}$ reguláris KKT pont. Az $\mathbf{A}$ mátrix $n \times n$-es, a $\mathbf{B}$ mátrix oszlopmérete n, sorméretét pedig az aktív egyenlőtlenséges feltételi függvények ($\bar{u}_i > 0$) és az egyenlőséges feltételi függvények darabszáma adja, jelöljük ezt a továbbiakban s-el. A $\mathbf{C}$ mátrix felépítésének sémája az alábbi:

$$\begin{aligned}
\mathbf{A} &= \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}}) \\
\mathbf{B} &= \begin{bmatrix} \boldsymbol{\nabla} g_i(\bar{\mathbf{x}}), & i \in I^+, (u_i > 0) \\ \boldsymbol{\nabla} h_i(\bar{\mathbf{x}}), & i = 1, \dots, k \end{bmatrix} \\
\mathbf{C} &= \begin{bmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix}
\end{aligned}$$

7. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned}
x_1^3 x_2^3 &\to \min! \\
x_1^2 + 4x_2^2 &\leqq 8
\end{aligned}$$

Megoldás
A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned}
f(x_1, x_2) &= x_1^3 x_2^3 \\
g(x_1, x_2) &= x_1^2 + 4x_2^2 - 8 \leqq 0 \\
X &= \mathbb{R}^2
\end{aligned}$$

1. lépés:
Legelőször a KKT pontokat határozzuk meg. A példa függvényei differenciálhatók. A szükséges gradiens vektorok és a gradiens vektorokra vonatkozó lineáris kombináció:

$$\boldsymbol{\nabla} f = \begin{bmatrix} 3x_1^2 x_2^3 \\ 3x_1^3 x_2^2 \end{bmatrix}, \qquad \boldsymbol{\nabla} g = \begin{bmatrix} 2x_1 \\ 8x_2 \end{bmatrix},$$

$$\begin{bmatrix} 3x_1^2 x_2^3 \\ 3x_1^3 x_2^2 \end{bmatrix} + u \begin{bmatrix} 2x_1 \\ 8x_2 \end{bmatrix} = \mathbf{0}.$$

A KKT pont meghatározásához szükséges KKT optimalitási feltételek az alábbiak szerint írhatók, vagyis az alábbi rendszert kell megoldani:

$$\begin{aligned}
3x_1^2 x_2^3 + 2ux_1 &= 0 \\
3x_1^3 x_2^2 + 8ux_2 &= 0 \\
u(x_1^2 + 4x_2^2 - 8) &= 0 \\
x_1^2 + 4x_2^2 - 8 &\leqq 0 \\
u &\geqq 0
\end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset az $u > 0$, másik eset az $u = 0$. Nézzük az egyes eseteket.

i) $u > 0$ eset

A 3. egyenletből, vagyis a komplementaritási feltételből következik, hogy a 4. egyenlőtlenség egyenlőség lesz, azaz a feltétel aktív lesz. Ekkor az 1., 2., és a 4. sorban szereplő most már egyenletből az alábbi egyenletrendszert kell megoldani az x_1, x_2, u mennyiségek meghatározásához

$$\begin{aligned} 3x_1^2x_2^3 + 2ux_1 &= 0 \\ 3x_1^3x_2^2 + 8ux_2 &= 0 \\ x_1^2 + 4x_2^2 - 8 &= 0 \end{aligned}$$

Az egyenletrendszernek 8 megoldása van ebből négy olyan megoldás, amelyekben $u \neq 0$, ezek a következők

x_1	x_2	u
2	-1	3
-2	1	3
2	1	-3
-2	-1	-3

Ezekből a megoldásokból csak az első kettő KKT pont, mivel ezeknél pozitív a Lagrange szorzó.

ii) $u = 0$ eset

Az 1., 2. sorban szereplő egyenletekből egyszerűen adódik az

$$x_1 = 0, \quad x_2 = 0$$

megoldás. A 4. sorbeli egyenlőtlenség is fennáll, így az $\mathbf{x} = (0, 0)$ pont KKT pont a $u = 0$ Lagrange szorzóval. Egyébként minden olyan megoldás is szóba jön, amelyekben valamelyik x változó zérus a másik pedig nem zérus. Ezek közül is természetesen csak azok, amelyekre a 4. sorbeli egyenlőtlenség is fennáll.

Összefoglalva, sok KKT pont adódott, amelyek a Lagrange szorzóikkal az alábbiak:
1. KKT pont: $\bar{x}_1 = 2, \quad \bar{x}_2 = -1, \quad \bar{u} = 3$, a feltétel aktív.
2. KKT pont: $\bar{x}_1 = -2, \quad \bar{x}_2 = 1, \quad \bar{u} = 3$, a feltétel aktív.
3. KKT pont: $\bar{x}_1 = 0, \quad \bar{x}_2 = 0, \quad \bar{u} = 0$, a feltétel inaktív.
a többi KKT pont:

$$\begin{aligned} &\bar{x}_1 = 0, \quad -\sqrt{2} \leqq \bar{x}_2 \leqq \sqrt{2}, \quad \bar{u} = 0. \\ &\bar{x}_2 = 0, \quad -2\sqrt{2} \leqq \bar{x}_1 \leqq 2\sqrt{2}, \quad \bar{u} = 0. \end{aligned}$$

2. lépés:

Most el kell dönteni, hogy a kapott KKT pontok optimális megoldásai-e a feltételes optimalizálási feladatnak. Ehhez képeznünk kell az aktualizált Lagrange függvény Hesse mátrixát. Ehhez a szükségesek az f és az aktív g függvény Hesse mátrixa:

$$\boldsymbol{\nabla}^2 f(x_1, x_2) = \begin{bmatrix} 6x_1x_2^3 & 9x_1^2x_2^2 \\ 9x_1^2x_2^2 & 6x_1^3x_2 \end{bmatrix}, \quad \boldsymbol{\nabla}^2 g(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}.$$

Az aktualizált Lagrange függvény Hesse mátrixa:

$$\boldsymbol{\nabla}^2 L_a(x_1, x_2) = \begin{bmatrix} 6x_1x_2^3 & 9x_1^2x_2^2 \\ 9x_1^2x_2^2 & 6x_1^3x_2 \end{bmatrix} + \bar{u} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}.$$

Az 1. KKT pont ($\bar{x}_1 = 2, \quad \bar{x}_2 = -1, \quad \bar{u} = 3$) vizsgálata:
Az aktualizált Lagrange függvény Hesse mátrixa az $\bar{\mathbf{x}} = (2, -1)$ pontban:

$$\nabla^2 L_a(\bar{x}_1, \bar{x}_2) = \begin{bmatrix} -12 & 36 \\ 36 & -48 \end{bmatrix} + 3 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} -6 & 36 \\ 36 & -24 \end{bmatrix}.$$

Erről a mátrixról kell eldönteni, hogy feltételesen pozitív definit vagy sem. Ehhez képezzük a $\mathbf{C}$ szegélyezett mátrixot. Mivel $\bar{u} > 0$, vagyis a feltétel aktív, így az aktualizált Lagrange függvény Hesse mátrixát a $\nabla g(\bar{\mathbf{x}}) = (4, -8)$ vektorral szegélyezzük, a $\mathbf{C}$ mátrix tehát

$$\mathbf{C} = \begin{bmatrix} -6 & 36 & 4 \\ 36 & -24 & -8 \\ 4 & -8 & 0 \end{bmatrix}.$$

Az aktualizált Lagrange függvény Hesse mátrixának feltételes pozitív definitségét az inerciateszt segítségével végezzük el. A számítás során keletkező Schur-komplemensek:

$$\begin{bmatrix} -6 & 36 & 4 \\ 36 & -24 & -8 \\ 4 & -8 & 0 \end{bmatrix}, \quad \begin{bmatrix} 192 & 16 \\ 16 & \frac{8}{3} \end{bmatrix}, \quad \begin{bmatrix} \frac{4}{3} \end{bmatrix}.$$

Az első pivotelem $-6 < 0$, az inercia $Iner = (1, 0, 0)$; következő pivotelem $192 > 0$, az inercia $Iner = (0, 0, 1)$; a végső blokk egy elemű, értéke $\frac{4}{3} > 0$, az inercia $Iner = (0, 0, 1)$. A kapott inerciákat összeadva, kapjuk, hogy $Iner(\mathbf{C}) = (1, 0, 2)$, tehát a szegélyezett aktualizált Lagrange függvény Hesse mátrixának a méreteknek megfelelő számú pozitív ill. negatív sajátértéke van, azaz $n = 2$ darab pozitív sajátértéke, $s = 1$ darab negatív sajátértéke és nincs zérus sajátértéke. Az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\bar{x}_1 = 2, \quad \bar{x}_2 = -1$ megoldás szigorú lokális minimumpont.

A 2. KKT pont ($\bar{x}_1 = -2, \quad \bar{x}_2 = 1, \quad \bar{u} = 3$) vizsgálata:
Az aktualizált Lagrange függvény Hesse mátrixa az $\bar{\mathbf{x}} = (-2, 1)$ pontban:

$$\nabla^2 L_a(\bar{x}_1, \bar{x}_2) = \begin{bmatrix} -12 & 36 \\ 36 & -48 \end{bmatrix} + 3 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} -6 & 36 \\ 36 & -24 \end{bmatrix}.$$

A $\mathbf{C}$ szegélyezett mátrix előállításához az aktualizált Lagrange függvény Hesse mátrixát a $\nabla g(\bar{\mathbf{x}}) = (-4, 8)$ vektorral szegélyezzük, a $\mathbf{C}$ mátrix tehát

$$\mathbf{C} = \begin{bmatrix} -6 & 36 & -4 \\ 36 & -24 & 8 \\ -4 & 8 & 0 \end{bmatrix}.$$

Az előzőekhez hasonlóan megállapíthatjuk, hogy $Iner(\mathbf{C}) = (1, 0, 2)$, tehát az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\bar{x}_1 = -2, \quad \bar{x}_2 = 1$ megoldás szigorú lokális minimumpont.

A 3. KKT pont ($\bar{x}_1 = 0, \quad \bar{x}_2 = 0, \quad \bar{u} = 0$) vizsgálata:

Mivel az $\bar{\mathbf{x}} = (0,0)$ pontban a feltétel inaktív, így az aktualizált Lagrange függvény Hesse mátrixa nem más mint a célfüggvény Hesse mátrixa, amely az $\bar{\mathbf{x}} = (0,0)$ pontban az alábbi

$$\nabla^2 L_a(\bar{x}_1, \bar{x}_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Mivel ez a mátrix nem pozitív definit, így az $\bar{\mathbf{x}} = (0,0)$ pontról a rendelkezésünkre álló tételek alapján nem tudjuk eldönteni az optimalitást. Ehhez tekintsük az $\bar{\mathbf{x}} = (0,0)$ KKT pont környezetében az

$$\mathbf{x} = (\lambda d_1, \lambda d_2)$$

vektort, ahol $\lambda > 0$ skalár, $\mathbf{d} = (d_1, d_2) \in \mathbb{R}^2, \mathbf{d} \neq \mathbf{0}$. Továbbá a $\mathbf{d}$ olyan vektor, amelynél a primál feltétel tejesül, azaz $(\lambda d_1)^2 + (\lambda d_2)^2 \leqq 8$. Az $\mathbf{x}$ vektort behelyettesítve a célfüggvénybe megkapjuk a KKT pont környezetében lévő pontokban a célfüggvény értékét, amely rendezés után az alábbi

$$f(x_1, x_2) = \lambda^6 d_1^3 d_2^3.$$

A KKT pontban a függvény értéke $f(0,0) = 0$, a környezetében pedig a fenti képletből és a feltételből látható, hogy felvehet akár pozitív, akár negatív értékeket is, így az $\bar{\mathbf{x}} = (0,0)$ KKT pont nem lehet minimumpont. Maximumpont sem lehet a fentiek miatt.

Javasoljuk az olvasónak, hogy a többi KKT pontról is döntse el, hogy se nem minimumpont se nem maximumpont.

7.2.2. Amikor a szükséges feltétel elégséges is

Az alábbiakban azt a feltételes optimalizálási feladatot mutatjuk be, amelynél nem kell semmiféle elégséges tétel ahhoz, hogy a KKT pont minimumpont legyen, azaz a szükséges feltétel egyben elégséges is. A feltétel nélküli optimalizálásból tudjuk, ha a minimalizálandó függvény konvex, akkor a $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ szükséges feltétel elégséges is. Ismert tény, hogy a konvex függvényekre igaz az alábbi egyenlőtlenség

$$f(\mathbf{x}) - f(\bar{\mathbf{x}}) \geqq \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{minden } \mathbf{x}\text{-re},$$

ebből a $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ szükséges feltétel miatt

$$f(\mathbf{x}) \geqq f(\bar{\mathbf{x}}) \quad \text{minden } \mathbf{x}\text{-re},$$

ez pedig azt jelenti, hogy $\bar{\mathbf{x}}$ minimumpont, sőt az összefüggésből az is látható, hogy $\bar{\mathbf{x}}$ globális minimumpont.

Most térjünk vissza a feltételes optimalizáláshoz, amelynél ismert tény, hogy $\nabla L_a(\bar{\mathbf{x}}) = \mathbf{0}$. Ha az $L_a(\mathbf{x})$ függvény konvex az S konvex halmazon, akkor az $\bar{\mathbf{x}}$ pont az L_a aktualizált Lagrange függvénynek globális minimumpontja, azaz

$$L_a(\bar{\mathbf{x}}) \leqq L_a(\mathbf{x}) \quad \text{minden } \mathbf{x} \in S.$$

Most felhasználva az $L_a(\mathbf{x}) \leqq f(\mathbf{x})$ minden $\mathbf{x} \in S$ és az $L_a(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$ ismert összefüggéseket, kapjuk, hogy

$$f(\bar{\mathbf{x}}) = L_a(\bar{\mathbf{x}}) \leqq L_a(\mathbf{x}) \leqq f(\mathbf{x}) \quad \text{minden } \mathbf{x} \in S,$$

ez pedig azt jelenti, hogy az $\bar{\mathbf{x}}$ globális minimumpontja az $f(\mathbf{x})$ függvénynek az S feltételi halmazon.

Az előzőek alapján, ha az S feltételi halmaz konvex és az aktualizált Lagrange függvény is konvex, akkor ennek a feladatnak a KKT pontja garantáltan minimumpont. A következő tétel adja meg a választ arra, hogy milyen optimalizálási feladat esetén lehet ezt a garanciát biztosítani.

TÉTEL

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz, az $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények konvexek, a $h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények lineárisak. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Legyen $\bar{\mathbf{x}}$ az optimalizálási feladat KKT pontja. Ekkor az $\bar{\mathbf{x}}$ pont globális minimumpontja a fenti feladatnak.

Bizonyítás

a) Először belátjuk, hogy az S feltételi halmaz konvex. Mivel $g_1, g_2, \ldots, g_m$ függvények konvexek, ezért az alsó nívóhalmazuk konvex halmaz, azaz az egyenlőtlenséges feltételek ($g_i(\mathbf{x}) \leqq 0$) mindegyike konvex halmazt definiál. A $h_1, h_2, \ldots, h_k$ függvények lineárisak, így az egyenlőséges feltételek ($h_i(\mathbf{x}) = 0$) mindegyike egy hipersíkot definiál, amiről tudjuk, hogy konvex halmaz. Az X halmaz is konvex. Tehát az S feltételi halmaz konvex halmazok metszete. Ismert tétel alapján a konvex halmazok metszete konvex, így S valóban konvex halmaz.

b) Most pedig azt látjuk be, hogy az aktualizált Lagrange függvény konvex függvény, ehhez idézzük az aktualizált Lagrange függvényt

$$L_a(\mathbf{x}) = L(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\mathbf{x}) + \sum_{i=1}^{m} \bar{u}_i g_i(\mathbf{x}) + \sum_{i=1}^{k} \bar{v}_i h_i(\mathbf{x}).$$

Az első két tagban az $f, g_1, g_2, \ldots, g_m$ függvények nemnegatív lineáris kombinációja szerepel, azt pedig tudjuk, hogy konvex függvények nemnegatív kombinációja is konvex. A harmadik tagban lineáris függvények tetszőleges kombinációja szerepel, ami lineáris függvényt eredményez. Egy konvex és egy lineáris függvény összege konvex függvény, így az aktualizált Lagrange függvény konvex függvény. Ekkor pedig a szükséges feltétel egyben elégséges is. **Q.e.d.**

8. Lagrange függvény használata a példamegoldásban

A fenti példákat megoldva jártasságot szereztünk a feltételes optimalizálási feladat megoldásában. Láttuk azt is, hogy a Lagrange függvény milyen fontos szerepet játszik a megoldás során. Az alábbiakban bemutatjuk a megoldásnak egy sablonos formáját. Tekintsük az alábbi feltételes optimalizálási feladatot:

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

Írjuk fel a feladat Lagrange függvényét, ahol a Lagrange szorzók az $\mathbf{u}$, $\mathbf{v}$ vektorok elemei:

$$L(\mathbf{x},\mathbf{u},\mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

Tekintsük a Lagrange függvénynek az összes változója szerinti gradiensét:

$$\boldsymbol{\nabla} L(\mathbf{x},\mathbf{u},\mathbf{v}) = \begin{bmatrix} \boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x},\mathbf{u},\mathbf{v}) \\ \boldsymbol{\nabla}_{\mathbf{u}} L(\mathbf{x},\mathbf{u},\mathbf{v}) \\ \boldsymbol{\nabla}_{\mathbf{v}} L(\mathbf{x},\mathbf{u},\mathbf{v}) \end{bmatrix},$$

ahol $\boldsymbol{\nabla}_{\mathbf{x}}, \boldsymbol{\nabla}_{\mathbf{u}}, \boldsymbol{\nabla}_{\mathbf{v}}$ a Lagrange függvény megfelelő változók szerinti elsőrendű parciális deriváltjait (gradienst) jelentik. Könnyen ellenőrizhetjük, hogy $\boldsymbol{\nabla}_{\mathbf{u}} L(\mathbf{x},\mathbf{u},\mathbf{v})$ vektor i-edik komponense a $g_i(\mathbf{x})$ feltételi függvény, a $\boldsymbol{\nabla}_{\mathbf{v}} L(\mathbf{x},\mathbf{u},\mathbf{v})$ vektor i-edik komponense pedig a $h_i(\mathbf{x})$ feltételi függvény. A $\boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x},\mathbf{u},\mathbf{v})$, azaz a Lagrange függvény $\mathbf{x}$ szerinti gradiense a duál feltétel baloldalát adja. Tehát a KKT pont meghatározásához szükséges rendszer egy részét a Lagrange függvény gradiensének a segítségével is felírhatjuk, ehhez még a komplementaritási és a Lagrange szorzókra előírt nemnegativitást kell hozzávenni, azaz a KKT pont meghatározásához szükséges rendszer:

$$\begin{aligned} \boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x},\mathbf{u},\mathbf{v}) &= \mathbf{0} \\ \boldsymbol{\nabla}_{\mathbf{u}} L(\mathbf{x},\mathbf{u},\mathbf{v}) &\leqq \mathbf{0} \\ \boldsymbol{\nabla}_{\mathbf{v}} L(\mathbf{x},\mathbf{u},\mathbf{v}) &= \mathbf{0} \\ \mathbf{u}\left[\boldsymbol{\nabla}_{u} L(\mathbf{x},\mathbf{u},\mathbf{v})\right] &= 0 \\ \mathbf{u} &\geqq \mathbf{0} \end{aligned}$$

A $\mathbf{C}$ mátrix előállítását is egyszerűen végezhetjük a Lagrange függvény segítségével. Képezzük a Lagrange függvénynek az összes változója szerinti Hesse mátrixát:

$$\boldsymbol{\nabla}^2 L(\mathbf{x},\mathbf{u},\mathbf{v}) = \begin{bmatrix} \boldsymbol{\nabla}^2_{\mathbf{xx}} L & \boldsymbol{\nabla}^2_{\mathbf{xu}} L & \boldsymbol{\nabla}^2_{\mathbf{xv}} L \\ \boldsymbol{\nabla}^2_{\mathbf{ux}} L & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\nabla}^2_{\mathbf{vx}} L & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

ahol $\boldsymbol{\nabla}^2_{\mathbf{pq}} L$ a Lagrange függvény megfelelő változók szerinti másodrendű parciális deriváltjait (Hesse mátrix) jelentik. Könnyen ellenőrizhetjük, hogy a $\boldsymbol{\nabla}^2_{\mathbf{xu}} L$, ill. a $\boldsymbol{\nabla}^2_{\mathbf{xv}} L$ blokkokban rendre a $\boldsymbol{\nabla} g_i$, ill. $\boldsymbol{\nabla} h_i$ vektorok vannak. A $\boldsymbol{\nabla}^2_{\mathbf{xu}} L$, ill. a $\boldsymbol{\nabla}^2_{\mathbf{xv}} L$ blokkok oszlopvektorai, míg a $\boldsymbol{\nabla}^2_{\mathbf{ux}} L$, ill. a $\boldsymbol{\nabla}^2_{\mathbf{vx}} L$ blokkok sorvektorai a gradiens vektorok. Arról is könnyen meggyőződhetünk, hogy az aktualizált Lagrange függvényt kapjuk, ha a $\boldsymbol{\nabla}^2_{\mathbf{xx}} L$ blokkba behelyettesítjük a KKT ponthoz tartozó Lagrange szorzókat. Tehát, ha a fenti $\boldsymbol{\nabla}^2 L$ mátrixba behelyettesítjük a KKT pontot és a hozzá tartozó Lagrange szorzókat, akkor a feltételes definitség eldöntéséhez szükséges szegélyezett $\mathbf{C}$ mátrix bővített változatát kapjuk meg. Ebből a mátrixból esetleg el kell hagyni bizonyos oszlopokat és sorokat, ennek indoklását az alábbi megjegyzésben tesszük meg.

Megjegyzés:

Nagyon fontos megjegyezni, hogy ezzel az egyszerüsített, sablonizált eljárással óvatosan kell eljárni, ha $\bar{u}_i = 0$. A $\mathbf{C}$ mátrixban szereplő aktualizált Lagrange függvény Hesse mátrixának részével ($\boldsymbol{\nabla}^2_{\mathbf{xx}} L$) nincs probléma, a szegélyekkel viszont van. A szegélyek ugyanis az összes feltételi függvény gradiensét tartalmazzák. Az elégséges feltétel tárgyalásánál megismertük, hogy csak az olyan aktív feltételek gradiensét kell tekinteni, amelyeknél $\bar{u}_i > 0$. Tehát a $\mathbf{C}$ mátrixot nem egyszerű behelyettesítéssel nyerjük a $\nabla^2 L$ mátrixból, hanem azokat az oszlopokat és sorokat törölni kell, amelyek inaktív feltételekhez tartoznak, ill. aktív feltételhez tartoznak ugyan, de $\bar{u}_i = 0$ Lagrange szorzóval.

8. Példa
A fentiek bemutatására oldjuk meg a 7. példa feladatát, amely az alábbi:

$$\begin{aligned} f(x_1, x_2) &= x_1^3 x_2^3 \to \min! \\ g(x_1, x_2) &= x_1^2 + 4x_2^2 - 8 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

1. lépés:
A feladat Lagrange függvénye:

$$L(x_1, x_2, u) = x_1^3 x_2^3 + u(x_1^2 + 4x_2^2 - 8).$$

Meghatározzuk a Lagrange függvény gradiensét és Hesse mátrixát (az összes változó szerint deriválva), amelyek a következők:

$$\nabla L(x_1, x_2, u) = \begin{bmatrix} 2ux_1 + 3x_1^2 x_2^3 \\ 8ux_2 + 3x_1^3 x_2^2 \\ x_1^2 + 4x_2^2 - 8 \end{bmatrix}, \qquad \nabla^2 L(x_1, x_2, u) = \begin{bmatrix} 2u + 6x_1 x_2^3 & 9x_1^2 x_2^2 & 2x_1 \\ 9x_1^2 x_2^2 & 8u + 6x_1^3 x_2 & 8x_2 \\ 2x_1 & 8x_2 & 0 \end{bmatrix}.$$

2. lépés:
KKT feltéételek felírása és a KKT pont meghatározása:
A ∇L vektorban az x_i szerinti deriváltak zérusok, ezek a duál feltételek, az u szerinti derivált a primál feltétel, amely $\leqq 0$, így a ∇L vektorból könnyen felírható a KKT pont meghatározásához szükséges rendszer. Ezeket még a komplementaritási feltétellel és a Lagrange szorzó nemnegativitási feltételével kell kiegészíteni. Tehát a megoldandó rendszer:

$$\begin{aligned} 3x_1^2 x_2^3 + 2ux_1 &= 0 \\ 3x_1^3 x_2^2 + 8ux_2 &= 0 \\ x_1^2 + 4x_2^2 - 8 &\leqq 0 \\ u(x_1^2 + 4x_2^2 - 8) &= 0 \\ u &\geqq 0 \end{aligned}$$

Láthatjuk, hogy az előzőekben szereplő KKT feltételeket kaptuk.

3. lépés:
A KKT pont optimalitásának vizsgálata:
Miután meghatároztuk a KKT pontot a hozzá tartozó Lagrange szorzóval együtt, jöhet az elégségességi vizsgálat. Ehhez képezni kell a Lagrange függvény Hesse mátrixát (az összes változó szerint deriválva), amelyet az 1. lépésben meghatároztunk. A $\mathbf{C}$ mátrix előállításához be kell helyettesíteni a KKT pontot és a Lagrange szorzót. Mivel a feltétel aktív, nem kell sorokat, oszlopokat törölni. Ezután eldöntjük a $\mathbf{C}$ mátrix segítségével valamilyen teszt alkalmazásával, hogy a KKT pont optimális vagy sem.

Az $\bar{\mathbf{x}} = (2, -1)$ KKT pont és az $\bar{u} = 3$ Lagrange szorzó esetén a $\mathbf{C}$ mátrix:

$$\mathbf{C} = \nabla^2 L(\bar{x}_1 \bar{x}_2, \bar{u}) = \begin{bmatrix} -6 & 36 & 4 \\ 36 & -24 & -8 \\ 4 & -8 & 0 \end{bmatrix}.$$

Láthatjuk, hogy az előző (7.) példában szereplő $\mathbf{C}$ mátrixot kaptuk.

Feladat:
A többi KKT pont esetében is végezze el az elégségességi vizsgálatot ezzel a sablonizált módszerrel!

9. Regularitási feltételek

A Fritz-John szükséges feltételnél a célfüggvény gradiensének is volt együtthatója, u_0 volt a jele. A Karush-Kuhn-Tucker szükséges feltételnél a célfüggvény gradiensének együtthatója 1 volt. A u_0 együttható pozitivitását mint láttuk az biztosította, hogy az aktív egyenlőtlenséges és az egyenlőséges feltételi függvények gradiensei az $\bar{\mathbf{x}}$ pontban lineárisan függetlenek voltak. Felmerül a kérdés, hogy milyen más előírás biztosítja az u_0 együttható pozitivitását. Ez a kérdés azért is fontos, mert a lineáris függetlenségi feltételt csak a KKT pont ismeretében tudjuk ellenőrizni. Van-e olyan feltétel, amely alapján a KKT pont ismerete nélkül is el tudjuk dönteni, hogy használhatjuk-e a KKT feltételeket az optimalitás szükséges feltételeként. A szakirodalomban számos regularitási feltétel ismeretes, mi három regularitási feltételt ismertetünk, egyik a linearitási regularitási feltétel, a másik pedig a Slater-féle regularitási feltétel, míg a harmadik a lineáris függetlenséget és a Slater-féle regularitást egybekapcsoló feltétel. Leginkább az első kettőt használják a gyakorlatban. Ezeket az előírásokat regularitási feltételeknek nevezzük.

9.1. Linearitási regularitási feltétel

TÉTEL (linearitási regularitási feltétel)
Legyen az általános alakú optimalizálási feladatban az összes feltételi függvény lineáris függvény. Tekintsük tehát a

$$\begin{gathered} f(\mathbf{x}) \to \min! \\ \mathbf{A}\mathbf{x} \leqq \mathbf{b} \\ \mathbf{B}\mathbf{x} = \mathbf{c} \\ \mathbf{x} \in X \end{gathered}$$

matematikai programozási feladatot, ahol $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{k \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^k$.

Ekkor a szükséges feltételnél a célfüggvény gradiensének u_0 együtthatója pozitív, azaz az ilyen feladatoknál a szükséges feltétel a KKT feltétel.

Bizonyítás
Általánosan azt mondtuk, hogy egyenlőtlenséges feltételnél a $\mathbf{d}$ akkor lehetséges irány az $\bar{\mathbf{x}}$ pontban, ha $\nabla g_i(\bar{\mathbf{x}})\mathbf{d} < 0$. Lineáris feltételeknél $\nabla g_i(\bar{\mathbf{x}})\mathbf{d} \leqq 0$ is lehetséges irányt ad. Könnyen meggyőződhetünk, hogy lineáris feltételnél a $\nabla g_i(\bar{\mathbf{x}})$ vektorok az $\mathbf{A}$ mátrix i-edik sorvektorával, a $\nabla h_i(\bar{\mathbf{x}})$ vektorok pedig a $\mathbf{B}$ mátrix i-edik sorvektorával azonosak. Legyen az

aktív egyenlőtlenséges feltételekhez tartozó együtthatómátrix $\mathbf{A}_1$, az inaktív egyenlőtlenséges feltételekhez tartozó együtthatómátrix pedig $\mathbf{A}_2$. Hasonlóan partícionáljuk a $\mathbf{b}$ vektort $\mathbf{b}_1, \mathbf{b}_2$ vektorokra. Így az $\mathbf{A}\mathbf{x} \leqq \mathbf{b}$ feltétel két részre bontható, nevezetesen $\mathbf{A}_1\bar{\mathbf{x}} = \mathbf{b}_1, \mathbf{A}_2\bar{\mathbf{x}} < \mathbf{b}_2$. Ekkor $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$ esetén a $\mathbf{d}$ lehetséges irány, ugyanis $\mathbf{A}_1(\bar{\mathbf{x}}+\lambda\mathbf{d}) = \mathbf{A}_1\bar{\mathbf{x}}+\lambda\mathbf{A}_1\mathbf{d} = \mathbf{b}_1+\lambda\mathbf{A}_1\mathbf{d} \leqq \mathbf{b}_1$, akkor igaz minden $\lambda > 0$ esetén, ha $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$. A $\mathbf{B}\mathbf{d} = \mathbf{0}$ esetén a $\mathbf{d}$ szintén lehetséges irány, ugyanis $\mathbf{B}(\bar{\mathbf{x}} + \lambda\mathbf{d}) = \mathbf{B}\bar{\mathbf{x}} + \lambda\mathbf{B}\mathbf{d} = \mathbf{c} + \lambda\mathbf{B}\mathbf{d} = \mathbf{c}$, akkor igaz minden $\lambda > 0$ esetén, ha $\mathbf{B}\mathbf{d} = \mathbf{0}$. Mivel $\mathbf{A}_2\bar{\mathbf{x}} < \mathbf{b}_2$, így ez nem ad szigorítást a $\mathbf{d}$ irányvektorra. Tehát az $\bar{\mathbf{x}}$ optimalitásának geometriai szükséges feltétele lineáris feltételek esetén az, hogy a

$$\begin{aligned} \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0} \\ \mathbf{B}\mathbf{d} &= \mathbf{0} \\ \boldsymbol{\nabla} f(\bar{\mathbf{x}})\mathbf{d} &< 0 \end{aligned}$$

rendszernek ne legyen megoldása. Ekkor pedig a Farkas tétel 4. verziója szerint, vannak olyan $\mathbf{u} \geqq \mathbf{0}$ és $\mathbf{v} \in \mathbb{R}^k$ vektorok, hogy

$$\mathbf{A}_1^T\mathbf{u} + \mathbf{B}^T\mathbf{v} = -\boldsymbol{\nabla} f(\bar{\mathbf{x}}).$$

Ezt zérusra rendezve olyan összefüggés adódik, amelyben a célfüggvény gradiense 1, tehát lineáris feltételek esetén szükséges feltételként mindig használható a KKT feltétel.

Most a Slater regularitási feltételt tárgyaljuk. Először az általánosított Slater regularitási feltételt ismertetjük, ez a regularitási feltétel egyenlőtlenséges és egyenlőséges feltételekkel rendelkező optimalizálási feladatokra vonatkozik. Ezután a gyakorlatban sűrűbben alkalmazott egyenlőtlenséges változatot is bemutatjuk.

9.2. Slater-féle regularitási feltétel

TÉTEL (általánosított Slater-féle regularitási feltétel)

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz, a $g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények konvexek, a $h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények affin függvények, azaz a $\mathbf{h}(\mathbf{x}) = (h_1, h_2, \ldots, h_k) = \mathbf{A}\mathbf{x}-\mathbf{b}$. Tegyük fel, hogy a következő általánosított Slater-féle regularitási feltétel fennáll, azaz létezik olyan $\mathbf{z} \in X$ pont, hogy $g_i(\mathbf{z}) < 0$ minden $i = 1, 2, \ldots, m$ indexre és $h_i(\mathbf{z}) = 0$ minden $i = 1, 2, \ldots, k$ indexre, továbbá $\mathbf{0} \in \mathbf{h}(X)$, ahol $\mathbf{h}(X) = \{\mathbf{h}(\mathbf{x}) : \mathbf{x} \in X\}$. Tekintsük az

$$\begin{aligned} & f(\mathbf{x}) \to \min! \\ & g_i(\mathbf{x}) \leqq 0 \qquad i = 1, 2, \ldots, m \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, 2, \ldots, k \\ & \mathbf{x} \in X \end{aligned}$$

matematikai programozási feladatot. Ekkor a szükséges feltételnél a célfüggvény gradiensének u_0 együtthatója pozitív, azaz az ilyen feladatoknál a szükséges feltétel a KKT feltétel.

Bizonyítás

A Fritz John tétel szerint az $\bar{\mathbf{x}} \in S$ optimalitása esetén léteznek olyan $u_0, u_i,\ i = 1, \ldots, m$; $v_i,\ i = 1, \ldots, k$ számok, amelyek nem mindegyike zérus, úgy, hogy

$$\begin{aligned} u_0\boldsymbol{\nabla} f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i\boldsymbol{\nabla} g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i\boldsymbol{\nabla} h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\bar{\mathbf{x}}) &= 0, \quad i = 1, \ldots, m \\ u_0, u_i &\geqq 0, \quad i = 1, \ldots, m \end{aligned}$$

Tegyük fel, hogy $u_0 = 0$, ekkor

$$\sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \boldsymbol{\nabla} h_i(\bar{\mathbf{x}}) = \mathbf{0}.$$

Szorozzuk a fenti egyenletet az $(\mathbf{x} - \bar{\mathbf{x}})$ vektorral, ahol $\mathbf{x} \in X$ ekkor

$$\sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) = 0, \quad \text{minden } \mathbf{x} \in X \text{ esetén.}$$

Most használjuk fel a g_i függvények konvexitását, miszerint

$$g_i(\mathbf{x}) - g_i(\bar{\mathbf{x}}) \geqq \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \quad \text{minden } \mathbf{x} \in X \text{ esetén,}$$

és a h_i függvények linearitását, miszerint

$$h_i(\mathbf{x}) - h_i(\bar{\mathbf{x}}) = \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \quad \text{minden } \mathbf{x} \in X \text{ esetén.}$$

Felhasználva ezeket és hogy $g_i(\bar{\mathbf{x}}) = 0$ az aktív feltételeknél, az inaktívaknál pedig $u_i = 0$, valamint $h_i(\bar{\mathbf{x}}) = 0$ kapjuk, hogy minden $\mathbf{x} \in X$ esetén

$$\begin{aligned} 0 &= \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \leqq \\ &\leqq \sum_{i=1}^{m} u_i \left[g_i(\mathbf{x}) - g_i(\bar{\mathbf{x}})\right] + \sum_{i=1}^{k} v_i \left[h_i(\mathbf{x}) - h_i(\bar{\mathbf{x}})\right] = \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}) \end{aligned}$$

azaz

$$\sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}) \geqq 0, \text{ minden } \mathbf{x} \in X.$$

Mivel ez minden $\mathbf{x} \in X$ vektorra fennáll, így arra a $\mathbf{z} \in X$ vektorra is, amelyre $g_i(\mathbf{z}) < 0$ és $h_i(\mathbf{z}) = 0$, azaz

$$\sum_{i=1}^{m} u_i g_i(\mathbf{z}) + \sum_{i=1}^{k} v_i h_i(\mathbf{z}) \geqq 0.$$

Mivel minden $g_i(\mathbf{z}) < 0$, $h_i(\mathbf{z}) = 0$ és $u_i \geqq 0$, így a minden $\mathbf{x} \in X$ vektorra vonatkozó fenti egyenlőtlenség baloldala $\leqq 0$, ezért a fenti egyenlőtlenség csak akkor állhat fenn, ha a baloldal zérus, ez pedig csak akkor lehet, ha minden $u_i = 0$. Így a fenti egyenlőtlenség egyenlőségre redukálódott és csak a h_i függvények szerepelnek benne, azaz

$$\sum_{i=1}^{k} v_i h_i(\mathbf{x}) = 0, \text{ minden } \mathbf{x} \in X.$$

Használjuk fel a Slater regularitás azon részét, hogy $\mathbf{0} \in \mathbf{h}(X)$, ahol $\mathbf{h}(X) = \{\mathbf{h}(\mathbf{x}) : \mathbf{x} \in X\}$. A h_i lineáris függvények az X halmazt a $\mathbf{h}(X)$ vektorok halmazába képezik le, mivel X nyílt konvex halmaz, így $\mathbf{h}(X)$ szintén nyílt konvex halmaz, a regularitási feltétel szerint a $\mathbf{h}(X)$ halmaz tartalmazza az origót, így az origó tetszőleges környezete is eleme a $\mathbf{h}(X)$-nek, ebből pedig következik, hogy a $\mathbf{h}$ vektor biztosan megválasztható úgy, hogy a v_i Lagrange szorzók által alkotott $\mathbf{v} = (v_1, v_2, \ldots, v_k)$ vektor irányába mutasson, azaz $\mathbf{h} = \lambda\mathbf{v}$, ahol $\lambda > 0$. Ekkor pedig

$$0 = \sum_{i=1}^{k} v_i h_i(\mathbf{x}) = \lambda \sum_{i=1}^{k} v_i^2.$$

Ez pedig csak akkor állhat fenn, ha a $\mathbf{v}$ vektor zérus. Eljutottunk tehát oda, hogy $u_0 = 0$ esetén az összes Lagrange szorzó zérus, ami ellentmondás. Tehát u_0 nem lehet zérus, csak pozitív, ekkor pedig választható 1-re is, így a KKT feltételt kapjuk. **Q.e.d.**

TÉTEL (Slater-féle regularitási feltétel)

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz, a $g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ függvények konvexek. Tegyük fel, hogy a következő Slater-féle regularitási feltétel fennáll, azaz létezik olyan $\mathbf{z} \in X$ pont, hogy $g_i(\mathbf{z}) < 0$ minden $i = 1, 2, \ldots, m$ indexre. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ \mathbf{x} &\in X \end{aligned}$$

egyenlőtlenséges matematikai programozási feladatot. Ekkor a szükséges feltételnél a célfüggvény gradiensének u_0 együtthatója pozitív, azaz az ilyen feladatoknál a szükséges feltétel a KKT feltétel.

Bizonyítás

A bizonyítás hasonló, de sokkal egyszerűbb az előző tétel bizonyításánál, mivel itt nincsenek egyenlőséges feltételek. A bizonyítást az olvasóra bízzuk.

Q.e.d.

9.3. Lineáris függetlenségi és Slater-féle regularitási feltétel

TÉTEL (lineáris függetlenségi és Slater-féle regularitási feltétel)

Legyen $X \subseteq \mathbb{R}^n$ nemüres nyílt halmaz, $g_1, g_2, \ldots, g_m; h_1, h_2, \ldots, h_k : \mathbb{R}^n \to \mathbb{R}$ függvények. Legyen $\bar{\mathbf{x}}$ minimumpont és legyen $I = \{i : g_i(\bar{\mathbf{x}}) = 0\}$. Tegyük fel, hogy az f, g_i $i \in I$ függvények differenciálhatók az $\bar{\mathbf{x}}$ pontban, a g_i $i \notin I$ függvények folytonosak az $\bar{\mathbf{x}}$ pontban és minden h_i $i = 1, 2, \ldots, k$ függvény folytonosan differenciálhatók az $\bar{\mathbf{x}}$ pontban. Tegyük fel továbbá, hogy érvényes az alábbi regularitási feltétel: a $\nabla h_i(\bar{\mathbf{x}})$, $i = 1, 2, \ldots, k$ vektorok lineárisan függetlenek, valamint létezik olyan $\mathbf{z} \in X$ vektor, hogy $\nabla g_i(\bar{\mathbf{x}})\mathbf{z} < 0$, $i \in I$ és $\nabla h_i(\bar{\mathbf{x}})\mathbf{z} = 0$, $i = 1, 2, \ldots, k$. Tekintsük az

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

matematikai programozási feladatot. Ekkor a szükséges feltételnél a célfüggvény gradiensének u_0 együtthatója pozitív, azaz az ilyen feladatoknál a szükséges feltétel a KKT feltétel.

Bizonyítás

A Fritz John tétel szerint az $\bar{\mathbf{x}} \in S$ optimalitása esetén léteznek olyan $u_0, u_i,\ i = 1, \ldots, m$; $v_i,\ i = 1, \ldots, k$ számok, amelyek nem mindegyike zérus úgy, hogy

$$\begin{aligned} u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\bar{\mathbf{x}}) &= 0 \quad i = 1, \ldots, m \\ u_0, u_i &\geqq 0 \quad i = 1, \ldots, m \end{aligned}$$

Tegyük fel, hogy $u_0 = 0$, ekkor

$$\sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}}) = \mathbf{0}.$$

Ha minden $u_i = 0$, akkor a $\nabla h_i(\bar{\mathbf{x}})$ vektorok lineáris függetlensége miatt minden v_i is zérus, mivel legalább egy Lagrange szorzónak nem zérusnak kell lennie, ezért legalább egy $u_i > 0$. Most szorozzuk a fenti egyenletet $\mathbf{z}$ vektorral, ahol $\mathbf{z} \in X$ ekkor

$$\sum_{i=1}^{m} u_i \nabla g_i(\bar{\mathbf{x}})\mathbf{z} + \sum_{i=1}^{k} v_i \nabla h_i(\bar{\mathbf{x}})\mathbf{z} = 0.$$

Most használjuk fel a regularitási feltételt, a második tag $\nabla h_i(\bar{\mathbf{x}})\mathbf{z} = 0$ miatt zérus, az első tag pedig negatív, mivel minden $\nabla g_i(\bar{\mathbf{x}})\mathbf{z} < 0$ és legalább egy $u_i > 0$, így a fenti egyenlet nem állhat fenn, tehát u_0 nem lehet zérus, csak pozitív, ekkor pedig választható 1-re is, és ekkor a KKT feltételt kapjuk. **Q.e.d.**

Összefoglalva tehát mindegyik regularitási feltétel a feltételi függvényekre ró ki valamilyen előírást. Elsőnél a gradiensek lineáris függetlenségét, másodiknál a linearitást, míg a harmadiknál a konvexséget és belső pont létezését kötöttük ki.

10. Nem standard optimalizálási feladatok

Az előző fejezetekben csak olyan feltételes optimalizálási feladatokat vizsgáltunk, amelyekben a célfüggvény minimumát keressük és az egyenlőtlenséges feltételek $\leqq$ formában voltak megfogalmazva. Ezt a feladatot nevezhetjük **standard feltételes optimalizálási feladatnak**. A gyakorlatban sokszor előfordul, amikor a célfüggvény maximumát keressük vagy vannak $\geqq$ típusú feltételek is. Azok a feladatok sem ritkák, amelyekben a döntési változók nemnegativitását is előírjuk. Az alábbiakban a különböző feladatokra vonatkozó állításokat három részben mondjuk ki: a) Lagrange függvény, b) KKT feltételek, c) elégségességi feltétel.

10.1. Maximum feladat

A maximum feladatot legegyszerűbben úgy kezelhetjük, hogy a célfüggvény (-1)-szeresét vesszük és így visszavezetjük minimum feladatra, amelynek már minden elméleti vonatkozását ismerjük. A teljesség kedvéért azonban felírjuk a maximum feladat KKT feltételeit, majd az elégségességi feltételt is. A maximum feladatban az egyenlőséges feltételeket $\geqq$ formában szokás megadni, mert akkor a Lagrange szorzók nemnegatívak lesznek.

TÉTEL (maximum feladat)
Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{array}{ll} f(\mathbf{x}) \to \max! & \\ g_i(\mathbf{x}) \geqq 0 & i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) = 0 & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & \end{array}$$

a) A feladathoz tartozó Lagrange függvény:

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

b) A feladathoz tartozó KKT feltételek az alábbiak:

$$\begin{array}{rcll} \nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x}) & = & \mathbf{0} & \\ u_i g_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, m \\ u_i & \geqq & 0, & i = 1, 2, \ldots, m \\ g_i(\mathbf{x}) & \geqq & 0, & i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, k \\ \mathbf{x} & \in & X & \end{array}$$

c) Legyen $\bar{\mathbf{x}}$ KKT pont az $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ Lagrange szorzókkal. Legyen I^+ azon aktív feltételek indexhalmaza, amelyekhez tartozó $\bar{u}_i > 0$ és definiáljuk az alábbi alteret:

$$A = \left\{ \mathbf{d} \neq \mathbf{0} : \nabla g_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i \in I^+, \ \nabla h_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i = 1, \ldots, k \right\}.$$

Ha $\mathbf{d}^T \nabla^2 L_a(\bar{\mathbf{x}})\mathbf{d} < 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális maximumpontja az optimalizálási feladatnak. Más szavakkal: Ha az aktualizált Lagrange függvény Hesse mátrixa feltételesen negatív definit, akkor a KKT pont szigorú lokális maximumpont.

Bizonyítás
Írjuk át a feladatot standard feladatra

$$\begin{array}{ll} -f(\mathbf{x}) \to \min! & \\ -g_i(\mathbf{x}) \leqq 0 & i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) = 0 & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & \end{array}$$

A standard feladathoz tartozó Lagrange függvény

$$L^S(\mathbf{x}, \mathbf{u}, \mathbf{v}) = -f(\mathbf{x}) + \sum_{i=1}^{m} u_i^S(-g_i(\mathbf{x})) + \sum_{i=1}^{k} v_i^S h_i(\mathbf{x}).$$

Az eredeti feladat Lagrange függvénye (-1)-szerese a standard feladathoz tartozó Lagrange függvénynek az $u_i^S = u_i \geqq 0$ és $v_i^S = -v_i \in \mathbb{R}$ választással. Ebből pedig a KKT feltételek egyszerűen következnek. Az elégséges feltétel is könnyen belátható, hiszen tudjuk, hogy egy feltételesen pozitív definit mátrix (-1)-szerese feltételesen negatív definit. **Q.e.d.**

9. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} 20\ln\left(\sqrt{x_1}x_2\right) - 3x_1 - x_1x_2 - x_2 &\rightarrow \text{max!} \\ x_1 + x_2 &\geqq 10 \end{aligned}$$

Megoldás
A logaritmus függvény értelmezési tartománya miatt ki kell kötni, hogy $x_1, x_2 > 0$, ami egy nyílt halmaz, így a korábbi jelöléseknek megfelelően a maximum feladat megszokott formája:

$$\begin{aligned} f(x_1,x_2) &= 20\ln\left(\sqrt{x_1}x_2\right) - x_1 - x_1x_2 - 3x_2 \\ g(x_1,x_2) &= x_1 + x_2 - 10 \geqq 0 \\ (x_1,x_2) &\in X = \{(x_1,x_2) : x_1 > 0, x_2 > 0\} \subset \mathbb{R}^2 \end{aligned}$$

A feladat Lagrange függvénye:

$$L(x_1,x_2,u) = 20\ln\left(\sqrt{x_1}x_2\right) - x_1 - x_1x_2 - 3x_2 + u(x_1 + x_2 - 10)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} \frac{10}{x_1} - 1 - x_2 + u \\ \frac{20}{x_2} - x_1 - 3 + u \\ x_1 + x_2 - 10 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} -\frac{10}{x_1^2} & -1 & 1 \\ -1 & -\frac{20}{x_2^2} & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} \tfrac{10}{x_1} - 1 - x_2 + u &= 0 \\ \tfrac{20}{x_2} - x_1 - 3 + u &= 0 \\ x_1 + x_2 - 10 &\geqq 0 \\ u(x_1 + x_2 - 10) &= 0 \\ u &\geqq 0 \\ x_1, x_2 &> 0 \end{aligned}$$

1. eset: $u = 0$
Az első két egyenletből a következő egyenletrendszert kell megoldani:

$$\begin{aligned} \tfrac{10}{x_1} - 1 - x_2 &= 0 \\ \tfrac{20}{x_2} - x_1 - 3 &= 0 \end{aligned}$$

Két megoldás adódik, a megoldások: $x_1 = 2, x_2 = 4$ és $x_1 = -15, x_2 = -\frac{5}{3}$. Az első megoldás a primál feltételt, a második megoldás pedig a döntési változókra előírt pozitivitási feltételt nem teljesíti.

2. eset: $u > 0$
A megoldandó egyenletrendszer:

$$\begin{aligned} \tfrac{10}{x_1} - 1 - x_2 + u &= 0 \\ \tfrac{20}{x_2} - x_1 - 3 + u &= 0 \\ x_1 + x_2 - 10 &= 0 \end{aligned}$$

Három megoldás adódik, a megoldások az alábbiak:

i	x_1	x_2	u
1	7.7	2.3	2
2	1.3	8.7	2
3	5	5	4

Mindhárom megoldás teljesíti a KKT feltételeket, így a feladatnak három KKT pontja van és mindegyikben aktív a feltétel.

Az elégségességi vizsgálatban szereplő $\mathbf{C}$ mátrixot a maximum feladat esetén is a Lagrange függvény Hesse mátrixából képezzük, ahogy ezt a standard feladatnál láttuk. A főminor teszttel fogjuk eldönteni a feltételes definitséget. Jelen példában $n = 2, s = 1$. Csak egyetlen determinánst kell kiszámítani ($k = 2$), mégpedig a teljes $\mathbf{C}$ mátrixét. A maximum feladatról van szó, ezért az aktualizált Lagrange függvény Hesse mátrixának feltételes negatív definitségét kell eldönteni. Mint tudjuk, ha $(-1)^k \det(\mathbf{C}) > 0$, azaz ha $\det(\mathbf{C}) > 0$, akkor a szóbanforgó mátrix feltételesen negatív definit, így ekkor a KKT pont maximumpont.

Az 1. KKT pont esetén a $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} -0.168 & -1 & 1 \\ -1 & -3.786 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Mivel $\det(\mathbf{C}) = 1.954 > 0$, így az $x_1 = 7.7, x_2 = 2.3$ szigorú lokális maximumpont. A célfüggvény értéke: 4.76 .

A 2. KKT pont esetén a $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} -5.931 & -1 & 1 \\ -1 & -0.264 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Mivel $\det(\mathbf{C}) = 4.195 > 0$, így az $x_1 = 1.3, x_2 = 8.7$ szigorú lokális maximumpont. A célfüggvény értéke: 7.18 .

A 3. KKT pont esetén a $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} -\frac{2}{5} & -1 & 1 \\ -1 & -\frac{4}{5} & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Mivel $\det(\mathbf{C}) = -\frac{4}{5} \not> 0$, így az $x_1 = 5, x_2 = 5$ nem maximumpont.

Összefoglalva megállapíthatjuk, hogy két szigorú lokális maximumpont adódott, a szigorú globális maximumpont: $x_1 = 1.3, x_2 = 8.7$, a célfüggvény maximális értéke: 7.18 .

Feladat:
Vizsgálja meg, hogy az $x_1 = 5, x_2 = 5$ pont minimumpont vagy nyeregpont!

10.2. Vegyes egyenlőtlenséges feltételű feladat

Olyan optimalizálási feladathoz tartozó KKT feltételeket írunk fel, amelyben különbséget teszünk az egyenlőtlenségek között.

TÉTEL (vegyes egyenlőtlenséges feltételű feladat)
Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{array}{rl} f(\mathbf{x}) \to \min! & \\ g_i(\mathbf{x}) \leqq 0 & i = 1, 2, \ldots, m_1 \\ g_i(\mathbf{x}) \geqq 0 & i = m_1 + 1, m_1 + 2, \ldots, m \\ h_i(\mathbf{x}) = 0 & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & \end{array}$$

a) A feladat Lagrange függvénye:

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i g_i(\mathbf{x}) + \sum_{i=m_1+1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

b) A feladathoz tartozó KKT feltételek az alábbiak:

$$\begin{array}{rcll} \boldsymbol{\nabla} f(\mathbf{x}) + \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \boldsymbol{\nabla} h_i(\mathbf{x}) & = & \mathbf{0} & \\ u_i g_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, m \\ u_i & \geqq & 0, & i = 1, 2, \ldots, m_1 \\ u_i & \leqq & 0, & i = m_1 + 1, m_1 + 2, \ldots, m \\ g_i(\mathbf{x}) & \leqq & 0, & i = 1, 2, \ldots, m_1 \\ g_i(\mathbf{x}) & \geqq & 0, & i = m_1 + 1, m_1 + 2, \ldots, m \\ h_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, k \\ \mathbf{x} & \in & X & \end{array}$$

c) Legyen $\bar{\mathbf{x}}$ KKT pont az $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ Lagrange szorzókkal. Legyen $I^{\pm}$ azon aktív feltételek indexhalmaza, amelyekhez tartozó $\bar{u}_i \neq 0$ és definiáljuk az alábbi alteret:

$$A = \left\{ \mathbf{d} \neq \mathbf{0} : \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i \in I^{\pm}, \ \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})\mathbf{d} = 0 \ i = 1, \ldots, k \right\}.$$

Ha $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak. Más szavakkal: Ha az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, akkor a KKT pont szigorú lokális minimumpont.

Bizonyítás
Írjuk át a feladatot standard feladatra és képezzük ennek Lagrange függvényét:

$$L^S(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i^S g_i(\mathbf{x}) + \sum_{i=m_1+1}^{m} u_i^S(-g_i(\mathbf{x})) + \sum_{i=1}^{k} v_i^S h_i(\mathbf{x}).$$

Az eredeti feladat Lagrange függvénye azonos a standard feladathoz tartozó Lagrange függvénnyel az $u_i^S = u_i \geqq 0$ $(i = 1, 2, \ldots, m_1)$, $u_i^S = -u_i \geqq 0$ $(i = m_1 + 1, m_1 + 2, \ldots, m)$ és $v_i^S = v_i \in \mathbb{R}$ választással. Ebből pedig a KKT feltételek egyszerűen következnek. Az elégséges feltétel is könnyen belátható. **Q. e. d.**

10.3. Extrémum feladat

Olyan optimalizálási feladathoz tartozó KKT feltételeket írunk fel, amelyben extrémum pontokat keresünk, azaz a minimum és a maximum pontokra is kíváncsiak vagyunk.

TÉTEL (extremum feladat)
Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{array}{rl} f(\mathbf{x}) \to \text{extremum!} & (\,\text{min! vagy max!}) \\ g_i(\mathbf{x}) \leqq 0 & i = 1, 2, \ldots, m_1 \\ g_i(\mathbf{x}) \geqq 0 & i = m_1 + 1, m_1 + 2, \ldots, m \\ h_i(\mathbf{x}) = 0 & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & \end{array}$$

a) A feladat Lagrange függvénye:

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i g_i(\mathbf{x}) + \sum_{i=m_1+1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

b) A minimum és a maximum feladatra együttesen vonatkozó KKT feltételek (a Lagrange szorzókra szokásos előjelkötést nem szerepeltetjük):

$$\begin{array}{rcll} \boldsymbol{\nabla} f(\mathbf{x}) + \sum_{i=1}^{m} u_i \boldsymbol{\nabla} g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \boldsymbol{\nabla} h_i(\mathbf{x}) & = & \mathbf{0} & \\ u_i g_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, m \\ g_i(\mathbf{x}) & \leqq & 0, & i = 1, 2, \ldots, m_1 \\ g_i(\mathbf{x}) & \geqq & 0, & i = m_1 + 1, m_1 + 2, \ldots, m \\ h_i(\mathbf{x}) & = & 0, & i = 1, 2, \ldots, k \\ \mathbf{x} & \in & X & \end{array}$$

Legyen a fenti rendszert kielégítő megoldás $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$.
Minimum feladat esete:
Ha $\bar{u}_i \geqq 0$ minden $i = 1, 2, \ldots, m_1$ esetén és $\bar{u}_i \leqq 0$ minden $i = m_1 + 1, m_1 + 2, \ldots, m$ esetén, akkor az $\bar{\mathbf{x}}$ a minimum feladat KKT pontja.

Maximum feladat esete:
Ha $\bar{u}_i \leqq 0$ minden $i = 1, 2, \ldots, m_1$ esetén és $\bar{u}_i \geqq 0$ minden $i = m_1 + 1, m_1 + 2, \ldots, m$ esetén, akkor az $\bar{\mathbf{x}}$ a maximum feladat KKT pontja.

c) Legyen $\bar{\mathbf{x}}$ KKT pont az $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ Lagrange szorzókkal. Legyen $I^{\pm}$ azon aktív feltételek indexhalmaza, amelyekhez tartozó $\bar{u}_i \neq 0$ és definiáljuk az alábbi alteret:

$$A = \left\{ \mathbf{d} \neq \mathbf{0} : \boldsymbol{\nabla} g_i(\bar{\mathbf{x}})\mathbf{d} = 0 \; i \in I^{\pm}, \; \boldsymbol{\nabla} h_i(\bar{\mathbf{x}})\mathbf{d} = 0 \; i = 1, \ldots, k \right\}.$$

Ha $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak.
Ha $\mathbf{d}^T \boldsymbol{\nabla}^2 L_a(\bar{\mathbf{x}})\mathbf{d} < 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális maximumpontja az optimalizálási feladatnak.

Más szavakkal: Ha az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív (negatív) definit, akkor a KKT pont szigorú lokális minimumpont (maximumpont).

Bizonyítás
Írjuk át a minimum feladatot standard feladatra és képezzük ennek Lagrange függvényét:

$$L^S(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i^S g_i(\mathbf{x}) + \sum_{i=m_1+1}^{m} u_i^S(-g_i(\mathbf{x})) + \sum_{i=1}^{k} v_i^S h_i(\mathbf{x}).$$

Az eredeti minimum feladat Lagrange függvénye azonos a standard feladathoz tartozó Lagrange függvénnyel az $u_i^S = u_i \geqq 0$ $(i = 1, 2, \dots, m_1)$, $u_i^S = -u_i \geqq 0$ $(i = m_1{+}1, m_1{+}2, \dots, m)$ és $v_i^S = v_i \in \mathbb{R}$ választással.

Most írjuk át a maximum feladatot standard feladatra és képezzük ennek Lagrange függvényét:

$$L^S(\mathbf{x}, \mathbf{u}, \mathbf{v}) = -f(\mathbf{x}) + \sum_{i=1}^{m_1} u_i^S g_i(\mathbf{x}) + \sum_{i=m_1+1}^{m} u_i^S(-g_i(\mathbf{x})) + \sum_{i=1}^{k} v_i^S h_i(\mathbf{x}).$$

Az eredeti maximum feladat Lagrange függvénye (-1)-szerese a standard feladathoz tartozó Lagrange függvénynek az $u_i^S = -u_i \geqq 0$ $(i = 1, 2, \dots, m_1)$, $u_i^S = u_i \geqq 0$ $(i = m_1 + 1, m_1 + 2, \dots, m)$ és $v_i^S = -v_i \in \mathbb{R}$ választással. Ezekből pedig a KKT feltételek következnek. Az elégséges feltétel is könnyen belátható. **Q. e. d.**

10.4. Nemnegativitási feltételes feladat

Nagyon gyakran előfordul, hogy a döntési változókra előírjuk az $\mathbf{x} \geqq \mathbf{0}$ nemnegativitási feltételt is. Ez a nemnegatív ortánst jelenti, amely nem nyílt halmaz, így nem tekinthetjük az X halmaznak, vagyis egyenlőtlenségként kell kezelnünk. Mivel ezek speciális feltételek, így a KKT feltételek felírásában némi egyszerűsítést tesznek lehetővé.

TÉTEL (Nemnegativitási feltételes feladat)
Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{array}{ll} f(\mathbf{x}) \to \min! & \\ g_i(\mathbf{x}) \leqq 0 & i = 1, 2, \dots, m \\ h_i(\mathbf{x}) = 0 & i = 1, 2, \dots, k \\ x_i \geqq 0 & i = 1, 2, \dots, n \\ \mathbf{x} \in X & \end{array}$$

a) A feladat Lagrange függvénye:

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i h_i(\mathbf{x}).$$

b) A feladathoz tartozó KKT feltételek az alábbiak:

$$
\begin{aligned}
\nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x}) &\geqq \mathbf{0} \\
\left[\nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x})\right] \mathbf{x} &= 0 \\
u_i g_i(\mathbf{x}) &= 0, \quad i = 1, 2, \ldots, m \\
u_i &\geqq 0, \quad i = 1, 2, \ldots, m \\
g_i(\mathbf{x}) &\leqq 0, \quad i = 1, 2, \ldots, m \\
h_i(\mathbf{x}) &= 0, \quad i = 1, 2, \ldots, k \\
x_i &\geqq 0, \quad i = 1, 2, \ldots, n \\
\mathbf{x} &\in X
\end{aligned}
$$

c) Legyen $\bar{\mathbf{x}}$ KKT pont az $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ Lagrange szorzókkal. Legyen I_1^+ azon aktív feltételek indexhalmaza, amelyekhez tartozó $\bar{u}_i > 0$ és legyen I_2^0 azon x_i változók indexhalmaza, amelyre $\bar{x}_i = 0$, azaz az $\bar{x}_i \geqq 0$ speciális feltétel aktív. Definiáljuk az alábbi alteret:

$$
A = \left\{\mathbf{d} \neq \mathbf{0} : \nabla g_i(\bar{\mathbf{x}})\mathbf{d} = 0,\ i \in I_1^+,\ \nabla h_i(\bar{\mathbf{x}})\mathbf{d} = 0,\ i = 1, \ldots, k,\ -\mathbf{e}_i\mathbf{d} = 0, i \in I_2^0\right\}.
$$

Ha $\mathbf{d}^T \nabla^2 L_a(\bar{\mathbf{x}})\mathbf{d} > 0$ minden $\mathbf{d} \in A$ irányra, akkor az $\bar{\mathbf{x}}$ KKT pont szigorú lokális minimumpontja az optimalizálási feladatnak. Más szavakkal: Ha az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, akkor a KKT pont szigorú lokális minimumpont.

Bizonyítás

A tétel KKT feltételi része azt mutatja, hogy a speciális $x_i \geqq 0,\ i = 1, 2, \ldots, n$ egyenlőtlenségekhez tartozó Lagrange szorzók elhagyhatók. A standard alakban szereplő $-x_i \leqq 0$ feltételhez tartozó Lagrange szorzó legyen w_i. A $-x_i \leqq 0$ feltételi függvény gradiense: $\nabla(-x_i) = -\mathbf{e}_i$, ahol $\mathbf{e}_i$ az i-edik egységvektor. Ekkor a standard feladat KKT feltételei az alábbiak:

$$
\begin{aligned}
\nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x}) + \sum_{i=1}^{n} w_i(-\mathbf{e}_i) &= \mathbf{0} \\
u_i g_i(\mathbf{x}) &= 0, \quad i = 1, 2, \ldots, m \\
u_i &\geqq 0, \quad i = 1, 2, \ldots, m \\
w_i(-x_i) &= 0, \quad i = 1, 2, \ldots, n \\
w_i &\geqq 0, \quad i = 1, 2, \ldots, n \\
g_i(\mathbf{x}) &\leqq 0, \quad i = 1, 2, \ldots, m \\
h_i(\mathbf{x}) &= 0, \quad i = 1, 2, \ldots, k \\
x_i &\geqq 0, \quad i = 1, 2, \ldots, n
\end{aligned}
$$

A duál feltételben az $\mathbf{e}_i$ vektorok lineáris kombinációja szerepel, ez a lineáris kombináció írható $-\mathbf{E}\mathbf{w} = -\mathbf{w}$ alakban is, ahol $\mathbf{E}$ egységmátrix. A $\mathbf{w}$ Lagrange szorzó nemnegativitásából következik az eredeti feladathoz tartozó duál feltétel:

$$
\mathbf{w} = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x}) \geqq \mathbf{0}.
$$

A $w_i(-x_i) = 0$ komplementaritási feltételek pedig a

$$\left[\nabla f(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\mathbf{x})\right] \mathbf{x} = 0$$

alakban írhatók, amely valójában n darab egyenletnek felel meg.

Az elégségességi feltételt pedig az alábbiakban láthatjuk be. Ha a feladat standardját tekintjük, akkor a $-x_i \leqq 0$ feltételek közül azok indexei szerepelnének az A altér definíciójában, amelyekhez tartozó $\bar{w}_i > 0$. A $w_i(-x_i) = 0$ komplementaritási feltétel miatt ezek pedig azok az indexek, amelyekre $x_i = 0$. A $-x_i$ feltétel gradiense pedig $\nabla(-x_i) = -\mathbf{e}_i$. **Q.e.d.**

10. Példa

Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} x_1^2 - 6x_1 + x_2^2 - 2x_2 &\rightarrow \text{min!} \\ x_1 + 2x_2 &\leqq 2 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

Megoldás

A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1, x_2) &= x_1^2 - 6x_1 + x_2^2 - 2x_2 \\ g(x_1, x_2) &= x_1 + 2x_2 - 2 \leqq 0 \\ x_1 &\geqq 0 \\ x_2 &\geqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A feladat Lagrange függvénye:

$$L(x_1, x_2, u_1, u_2, u_3) = x_1^2 - 6x_1 + x_2^2 - 2x_2 + u(x_1 + 2x_2 - 2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} 2x_1 - 6 + u \\ 2x_2 - 2 + 2u \\ x_1 + 2x_2 - 2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1 - 6 + u &\geqq 0 \\ 2x_2 - 2 + 2u &\geqq 0 \\ x_1 + 2x_2 - 2 &\leqq 0 \\ (2x_1 - 6 + u)x_1 &= 0 \\ (2x_2 - 2 + 2u)x_2 &= 0 \\ u(x_1 + 2x_2 - 2) &= 0 \\ x_1 &\geqq 0 \\ x_2 &\geqq 0 \\ u &\geqq 0 \end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk.

i) $u = 0$ eset
Az 1., 2. sorban lévő egyenlőtlenségekből

$$x_1 \geqq 3, \quad x_2 \geqq 1,$$

a 4., 5. sorban szereplő egyenletekből pedig

$$x_1 = 3, \quad x_2 = 1$$

adódik, azonban ez nem teljesíti a 3. sorban lévő primál feltételt, így az $\mathbf{x} = (3, 1)$ pont nem KKT pont.

ii) $u > 0$ eset
A 3. sorban lévő primál feltétel ekkor egyenlőséggel teljesül. Hozzávéve a 4. és az 5. sorbeli egyenleteket, az x_1, x_2, u változókra az alábbi egyenletrendszert kapjuk:

$$\begin{aligned} x_1 + 2x_2 - 2 &= 0 \\ (2x_1 - 6 + u)x_1 &= 0 \\ (2x_2 - 2 + 2u)x_2 &= 0 \end{aligned}$$

Három megoldás adódik, a megoldások az alábbiak:

i	x_1	x_2	u
1	$\frac{12}{5}$	$-\frac{1}{5}$	$\frac{6}{5}$
2	2	0	2
3	0	1	0

Az első megoldásban x_2 negatív. A 3. megoldás nem teljesíti az első egyenlőtlenséget. A második megoldás a fennmaradó KKT feltételeket is teljesíti, így az $\bar{\mathbf{x}} = (2, 0)$ KKT pont az $\bar{u} = 2$ Lagrange szorzóval és a feltétel aktív.

A következőkben megállapítjuk, hogy az $\bar{\mathbf{x}} = (2, 0)$ KKT pont minimumpont. A $\mathbf{C}$ szegélyezett mátrix felírásánál most is segítségünkre lesz a Lagrange függvény Hesse mátrixa. Mivel a Lagrange függvényben nem szerepeltettük a nemnegativitási feltételeket, így a Hesse mátrixban nem szerepelnek a $-x_i \leqq 0$ feltételek gradiensei, a $-\mathbf{e}_i$ vektorok. Ezekkel tehát ki kell egészíteni, de csak azokkal, amelynél $x_i = 0$. Esetünkben ez a $-\mathbf{e}_2$ vektor, mert $\bar{x}_2 = 0$. A szegélyezett $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & -1 \\ 1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}.$$

Az inerciatesztet használva az alábbi Schur-komplemensek adódtak.

$$\begin{bmatrix} 2 & 2 & -1 \\ 2 & -\frac{1}{2} & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} -\frac{5}{2} & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}, \quad [2].$$

Az egyes lépésekben kapott inerciákat összeadva, kapjuk, hogy $Iner(\mathbf{C}) = (2,0,2)$, tehát a szegélyezett aktualizált Lagrange függvény Hesse mátrixának $n = 2$ darab pozitív sajátértéke, $s = 2$ darab negatív sajátértéke és nincs zérus sajátértéke. Az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\bar{\mathbf{x}} = (2,0)$ pont szigorú minimumpont.

Megjegyzés:
Mivel az aktualizált Lagrange függvény Hesse mátrixa ($\mathbf{C}$ mátrix 2×2-es főminor mátrixa) pozitív definit, így a KKT pont szigorú minimumpont. Vegyük észre azt is, hogy a feladat összes feltétele lineáris és a célfüggvény konvex, így a KKT pont egyben globális minimumpont is. A célfüggvény Hesse mátrixa ugyanis pozitív definit, ez pedig a vonatkozó tétel szerint szigorúan konvex függvényt jelent. Az elégségességi vizsgálatot tehát jelen feladatnál el is hagyhattuk volna, ha figyelembe vettük volna a fentieket.

Feladat:
Oldja meg a feladatot standard alakra hozva is!

11. Példamegoldások

Az alábbiakban számos példát mutatunk be a feltételes optimalizálási feladatok megoldására. Összefoglalva tehát az első lépésben KKT pontot határozunk meg, majd a második lépésben eldöntjük, hogy ez a pont optimális megoldás-e.

11. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} x_1^2 - 12x_1 + x_2^2 - 4x_2 + 5 &\rightarrow \text{min!} \\ x_1^2 + x_2^2 &\leqq 25 \\ x_1 + 4x_2 &= 16 \end{aligned}$$

Megoldás
Ennek az optimalizálási feladatnak a 6. példában meghatároztuk a KKT pontját, de ott nem használtuk a Lagrange függvényt. Most újra meghatározzuk a KKT pontot, de használjuk a Lagrange függvényt. A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1,x_2) &= x_1^2 - 12x_1 + x_2^2 - 4x_2 + 5 \\ g(x_1,x_2) &= x_1^2 + x_2^2 - 25 \leqq 0 \\ h(x_1,x_2) &= x_1 + 4x_2 - 16 = 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1,x_2,u,v) = x_1^2 - 12x_1 + x_2^2 - 4x_2 + 5 + u(x_1^2 + x_2^2 - 5) + v(x_1 + 4x_2 - 16)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2x_1 - 12 + 2ux_1 + v \\ 2x_2 - 4 + 2ux_2 + 4v \\ x_1^2 + x_2^2 - 25 \\ x_1 + 4x_2 - 16 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2+2u & 0 & 2x_1 & 1 \\ 0 & 2+2u & 2x_2 & 4 \\ 2x_1 & 2x_2 & 0 & 0 \\ 1 & 4 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned}
2x_1 - 12 + 2ux_1 + v &= 0\\
2x_2 - 4 + 2ux_2 + 4v &= 0\\
x_1^2 + x_2^2 - 25 &\leqq 0\\
x_1 + 4x_2 - 16 &= 0\\
u(x_1^2 + x_2^2 - 25) &= 0\\
u &\geqq 0
\end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset az $u = 0$, másik eset az $u > 0$. Nézzük az egyes eseteket.

i) $u = 0$ eset
Az 1., 2. és a 4. sorban szereplő egyenletekből az x_1, x_2, v számokra kapjuk, hogy

$$x_1 = \frac{104}{17}, \quad x_2 = \frac{42}{17}, \quad v = -\frac{4}{17}.$$

Az 5. megkötés (komplementaritási feltétel) $u = 0$ esetén nem érdekes. Azonban a 3. sorbeli egyenlőtlenségnek is fenn kell állnia, ezt ellenőrizzük

$$x_1^2 + x_2^2 - 25 = \frac{315}{17} > 0.$$

Mivel a 3. egyenlőtlenség nem áll fenn, így az $x_1 = \frac{104}{17}$, $x_2 = \frac{42}{17}$ megoldás nem KKT pont.

ii) $u > 0$ eset
A 3. sorbeli egyenlőtlenségnek az 5. sor miatt egyenlőséggel kell teljesednie. Ekkor az 1., 2., 4. sorbeli egyenletekből és az 5. sorban szereplő most már egyenletből az x_1, x_2, u, v számokra kapjuk, hogy

$$x_1 = 4, \quad x_2 = 3, \quad u = \frac{9}{13}, \quad v = -\frac{20}{13}.$$

Tehát az $\bar{x}_1 = 4$, $\bar{x}_2 = 3$ megoldás KKT pont, mert teljesíti a KKT feltételek mindegyikét.

Most meg kell győződni arról, hogy ez a megoldás optimális megoldás. A $\mathbf{C}$ mátrixot a $\nabla^2 L$ mátrixból építjük fel a megoldások behelyettesítésével. Nem kell sorokat és oszlopokat elhagyni a $\nabla^2 L$ mátrixból, mert $u > 0$.

$$\mathbf{C} = \begin{bmatrix} \frac{44}{13} & 0 & 8 & 1\\ 0 & \frac{44}{13} & 6 & 4\\ 8 & 6 & 0 & 0\\ 1 & 4 & 0 & 0 \end{bmatrix}.$$

Megállapíthatjuk, hogy $Iner(\mathbf{C}) = (2, 0, 2)$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, amiből következik, hogy az $\bar{x}_1 = 4$, $\bar{x}_2 = 3$ KKT pont szigorú lokális minimumpont.

Megjegyzés:
Az elégséges feltétel gyenge változatát is használhattuk volna ebben a példában. Mivel a $\nabla^2 L_a(\bar{\mathbf{x}})$ mátrix (a $\mathbf{C}$ mátrix 2×2-es főminormátrixa) pozitív definit, így a gyenge elégségességi feltétel is választ ad az optimalitásra.

12. Példa

Adott az $x_1 - x_2 = 2$ és az $x_1 + x_2 = 4$ egyenes. A két egyenes az (x_1, x_2) síkot négy részre osztja. Tekintsük azt a tartományt, amely tartalmazza a $(3, 3)$ pontot és ez a tartomány a félegyeneseket is tartalmazza. Határozzuk meg a tartomány azon pontját, amely legközelebb van az origóhoz!

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. Megjegyezzük, hogy a távolságot definiáló négyzetgyök függvény szigorúan monoton növekvő, így elegendő célfüggvényként a távolság négyzetét használni, amely a jelen példában $x_1^2 + x_2^2$. A halmazt célszerű felrajzolni és ekkor könnyen felírhatók a feltételek. Természetesen rajzolás nélkül is egyszerűen megállapíthatók a feltételek a következő módon. Helyettesítsük be a $(3, 3)$ pontot az egyenesek egyenletébe: az $x_1 - x_2 = 2$ egyenes esetén $3 - 3 < 2$, így a feltétel $x_1 - x_2 \leqq 2$, az $x_1 + x_2 = 4$ egyenes esetén $3 + 3 > 4$, így a feltétel $x_1 + x_2 \geqq 4$. Az optimalizálási feladat a szokásos alakjában az alábbi:

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 \to \min! \\ g_1(x_1, x_2) &= x_1 - x_2 - 2 \leqq 0 \\ g_2(x_1, x_2) &= -x_1 - x_2 + 4 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u_1, u_2) = x_1^2 + x_2^2 + u_1(x_1 - x_2 - 2) + u_2(-x_1 - x_2 + 4)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2x_1 + u_1 - u_2 \\ 2x_2 - u_1 - u_2 \\ x_1 - x_2 - 2 \\ -x_1 - x_2 + 4 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2 & 0 & 1 & -1 \\ 0 & 2 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1 + u_1 - u_2 &= 0 \\ 2x_2 - u_1 - u_2 &= 0 \\ x_1 - x_2 - 2 &\leqq 0 \\ -x_1 - x_2 + 4 &\leqq 0 \\ u_1(x_1 - x_2 - 2) &= 0 \\ u_2(-x_1 - x_2 + 4) &= 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását az u_1, u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1 = u_2 = 0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x} = (0, 0)$ pont. Ez azonban nem teljesíti a második primál feltételt, vagyis nem lehetséges megoldás.

ii) $u_1 = 0, u_2 > 0$ eset

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - u_2 &= 0 \\ 2x_2 - u_2 &= 0 \\ -x_1 - x_2 + 4 &= 0 \end{aligned}$$

A megoldás: $x_1 = 2$, $x_2 = 2$, $u_2 = 4$. Ez a megoldás maradéktalanul teljesíti a többi KKT feltételt, így az $\mathbf{x} = (2,2)$ vektor KKT pont az $u_1 = 0, u_2 = 4$ Lagrange szorzókkal úgy, hogy a g_1 feltétel inaktív, a g_2 feltétel aktív.

iii) $u_1 > 0, u_2 = 0$ eset

Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 + u_1 &= 0 \\ 2x_2 - u_1 &= 0 \\ x_1 - x_2 - 2 &= 0 \end{aligned}$$

A megoldás: $x_1 = 1$, $x_2 = -1$, $u_1 = -2$. Ez a megoldás nem teljesíti a Lagrange szorzókra vonatkozó nemnegativitási feltételt, így az $\mathbf{x} = (1,-1)$ vektor nem KKT pont.

iv) $u_1 > 0, u_2 > 0$ eset

Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 + u_1 - u_2 &= 0 \\ 2x_2 - u_1 - u_2 &= 0 \\ x_1 - x_2 - 2 &= 0 \\ -x_1 - x_2 + 4 &= 0 \end{aligned}$$

A megoldás: $x_1 = 3$, $x_2 = 1$, $u_1 = -2$, $u_2 = 4$. Az $\mathbf{x} = (3,1)$ vektor ugyan teljesíti a primál feltételeket, de az első Lagrange szorzó $u_1 = -2$ negatív, így az $\mathbf{x} = (3,1)$ vektor nem lehet KKT pont.

Összefoglalva az eseteket, az $\mathbf{x} = (2,2)$ vektor KKT pont az $u_1 = 0, u_2 = 4$ Lagrange szorzókkal úgy, hogy a g_1 feltétel inaktív, a g_2 feltétel aktív.

Arról, hogy ez valóban minimumpont, az optimum elégséges feltétele segítségével győződünk meg. A szükséges $\mathbf{C}$ szegélyezett mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixból a 3. sor és a 3. oszlop elhagyásával nyerjük, mivel az első feltétel inaktív:

$$\mathbf{C} = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 0 \end{bmatrix}.$$

Az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, így az $x_1 = 2$, $x_2 = 2$ megoldás szigorú lokális minimumpont. Itt is használható a gyenge elégséges feltétel, mert $\boldsymbol{\nabla}^2 L_a$ pozitív definit.

13. Példa

Legyen az (x_1, x_2) síkon egy tartomány, amely az $x_2 = x_1^2 - 3$ parabola és az $x_2 = 2 - \frac{x_1}{2}$ egyenes által határolt kompakt halmaz. Határozzuk meg a halmaz azon pontját, amely a $(3, 2)$ ponthoz a legközelebb van.

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. A halmazt célszerű felrajzolni és ekkor könnyen felírhatók a feltételek. Az optimalizálási feladat a szokásos alakjában az alábbi:

$$\begin{aligned} f(x_1, x_2) &= (x_1 - 3)^2 + (x_2 - 2)^2 \to \min! \\ g_1(x_1, x_2) &= x_1^2 - x_2 - 3 \leqq 0 \\ g_2(x_1, x_2) &= x_1 + 2x_2 - 4 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u_1, u_2) = (x_1 - 3)^2 + (x_2 - 2)^2 + u_1(x_1^2 - x_2 - 3) + u_2(x_1 + 2x_2 - 4)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2x_1 - 6 + 2u_1x_1 + u_2 \\ 2x_2 - 4 - u_1 + 2u_2 \\ -x_2 + x_1^2 - 3 \\ x_1 + 2x_2 - 4 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2u_1 + 2 & 0 & 2x_1 & 1 \\ 0 & 2 & -1 & 2 \\ 2x_1 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1 - 6 + 2u_1x_1 + u_2 &= 0 \\ 2x_2 - 4 - u_1 + 2u_2 &= 0 \\ x_1^2 - x_2 - 3 &\leqq 0 \\ x_1 + 2x_2 - 4 &\leqq 0 \\ u_1(x_1^2 - x_2 - 3) &= 0 \\ u_2(x_1 + 2x_2 - 4) &= 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását az u_1, u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1 = u_2 = 0$ eset

Az első két egyenletből azonnal adódik az $x_1 = 3$ és $x_2 = 2$ megoldás, azaz az $\mathbf{x} = (3, 2)$ pont. Ez azonban nem teljesíti a primál feltételeket, vagyis nem lehet KKT pont.

ii) $u_1 = 0, u_2 > 0$ eset

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 6 + u_2 &= 0 \\ 2x_2 - 4 + 2u_2 &= 0 \\ x_1 + 2x_2 - 4 &= 0 \end{aligned}$$

A megoldás: $x_1 = 12/5$, $x_2 = 4/5$, $u_2 = 6/5$. Azonban az $\mathbf{x} = (12/5, 4/5)$ pont nem esik a tartományba.

iii) $u_1 > 0, u_2 = 0$ eset

Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 6 + 2u_1x_1 &= 0 \\ 2x_2 - 4 - u_1 &= 0 \\ x_1^2 - x_2 - 3 &= 0 \end{aligned}$$

Kézi számolással nehéz megoldani a fenti egyenletrendszert, hisz rendezés után az x_1 változóra a $2x_1^3 - 9x_1 - 3 = 0$ harmadfokú egyenlet adódik. Gépi megoldással az alábbi három megoldást kapjuk:

i	x_1	x_2	u_1
1	2.2716	2.1603	0.3206
2	-0.3422	-2.8829	-9.7657
3	-1.9294	0.7225	-2.5549

Az utolsó két megoldás u_1 negatívitása miatt nem jöhet szóba, az első pedig azért nem, mert az $x_1 = 2.271\,6, x_2 = 2.1603$ nem esik a tartományba. Tehát ez az eset sem ad KKT pontot.

iv) $u_1 > 0, u_2 > 0$ eset

Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 6 + 2u_1x_1 + u_2 &= 0 \\ 2x_2 - 4 - u_1 + 2u_2 &= 0 \\ x_1^2 - x_2 - 3 &= 0 \\ x_1 + 2x_2 - 4 &= 0 \end{aligned}$$

Célszerű a megoldást két részre bontani, először az utolsó két egyenletből meghatározzuk x_1, x_2 értékét, majd az első két egyenletből az u_1, u_2 értékeket. A fenti egyenletrendszerre két megoldás is adódik, az első megoldás

$$x_1 = 2, \quad x_2 = 1, \quad u_1 = \frac{2}{9}, \quad u_2 = \frac{10}{9},$$

a második megoldás pedig

$$x_1 = -\frac{5}{2}, \quad x_2 = \frac{13}{4}, \quad u_1 = -\frac{49}{18}, \quad u_2 = -\frac{47}{18}.$$

Azonnal láthatjuk, hogy a második megoldásban a Lagrange szorzókra előírt nemnegatívitás nem teljesül, így nem lehet KKT pont. Az első megoldás maradéktalanul teljesíti a KKT feltételeket, így az $\bar{\mathbf{x}} = (2, 1)$ vektor KKT pont az $\bar{\mathbf{u}} = (\frac{2}{9}, \frac{10}{9})$ Lagrange szorzókkal és mindkét feltétel aktív.

Az eseteket összevetve egyetlen KKT pont adódott, amely az $\bar{\mathbf{x}} = (2, 1)$ vektor az $\bar{\mathbf{u}} = (\frac{2}{9}, \frac{10}{9})$ Lagrange szorzókkal.

Most azt kell eldönteni, hogy a KKT pont optimális megoldás-e. Ehhez képeznünk kell a $\mathbf{C}$ mátrixot. Mivel mindkét feltétel aktív, ezért a $\boldsymbol{\nabla}^2 L$ mátrixból nem kell elhagyni oszlopokat és sorokat, hanem egyszerű behelyettesítést kell végrehajtani, a $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} \frac{22}{9} & 0 & 4 & 1 \\ 0 & 2 & -1 & 2 \\ 4 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}.$$

Az inerciatesztet használjuk, a számítás során keletkező Schur-komponensek a következők:

$$\begin{bmatrix} \frac{22}{9} & 0 & 4 & 1 \\ 0 & 2 & -1 & 2 \\ 4 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{22}{9} & 4 & 1 \\ 4 & -\frac{1}{2} & 1 \\ 1 & 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} \frac{53}{18} & \frac{9}{2} \\ \frac{9}{2} & 0 \end{bmatrix}, \quad \begin{bmatrix} -\frac{27}{106} \end{bmatrix}.$$

Az első pivotelem $c_{22} = 2 > 0$, az inercia $Iner = (0,0,1)$; a második pivotelem $-2 < 0$, az inercia $Iner = (1,0,0)$; a harmadik pivotelem $\frac{53}{18} > 0$, az inercia $Iner = (1,0,0)$; a végső blokk egy elemű, értéke $-\frac{27}{106} < 0$, az inercia $Iner = (1,0,0)$. Az inerciákat összeadva, kapjuk, hogy $Iner(\mathbf{C}) = (2,0,2)$, tehát a szegélyezett aktualizált Lagrange függvény Hesse mátrixának az x ismeretlennek megfelelő számú ($n = 2$) pozitív sajátértéke van , az aktív feltételeknek megfelelő számú ($s = 2$) negatív sajátértéke van és nincs zérus sajátértéke. Az inerciateszt alapján tehát az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\bar{\mathbf{x}} = (2,1)$ megoldás szigorú lokális minimum.

Arról, hogy ez valóban minimumpont úgy is meggyőződhetünk, hogy a tartományt és a kívülálló pontot felrajzoljuk az (x_1, x_2) síkon.

A példa kapcsán felhívjuk a figyelmet arra, hogy az optimalitást a gyenge elégségességi tétel alapján is eldönthettünk volna.

Feladat:
Határozza meg a tartomány azon pontját, amely legtávolabb van a $(3,2)$ ponttól!

14. Példa

Tervezzünk egy egyenes körhenger alakú alul-felül zárt konzervdobozt, amelynek a térfogata legalább $V = 16\pi$ térfogategység és az elkészítéséhez szükséges lemez területe minél kisebb legyen!

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. Legyen a két keresett mennyiség a doboz alapkörének sugara, ill. a magassága, jelölje ezeket rendre x_1 és x_2. Megjegyezzük, hogy a keresett sugár és magasság pozitív értékek, ezeket a feltételeket az X nyílt halmaz fogja tartalmazni. Az optimalizálási feladat a szokásos alakjában az alábbi:

$$\begin{aligned} f(x_1, x_2) &= 2x_1^2\pi + 2x_1\pi x_2 \to \min! \\ g(x_1, x_2) &= 16\pi - x_1^2\pi x_2 \leqq 0 \\ (x_1, x_2) &\in X = \{(x_1, x_2) : x_1 > 0, x_2 > 0\} \subset \mathbb{R}^2 \end{aligned}$$

A feltételi függvényt is és a célfüggvényt is oszthatjuk π-vel, így könnyebben kezelhető feladatot kapunk, amely az alábbi:

$$\begin{aligned} f(x_1,x_2) &= 2x_1^2 + 2x_1x_2 \to \min! \\ g(x_1,x_2) &= 16 - x_1^2x_2 \leqq 0 \\ (x_1,x_2) &\in X = \{(x_1,x_2) : x_1 > 0, x_2 > 0\} \subset \mathbb{R}^2 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1,x_2,u) = 2x_1^2 + 2x_1x_2 + u(16 - x_1^2x_2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 4x_1 + 2x_2 - 2ux_1x_2 \\ 2x_1 - ux_1^2 \\ 16 - x_1^2x_2 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} -2ux_2 + 4 & -2ux_1 + 2 & -2x_1x_2 \\ -2ux_1 + 2 & 0 & -x_1^2 \\ -2x_1x_2 & -x_1^2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 4x_1 + 2x_2 - 2x_1x_2u &= 0 \\ 2x_1 - x_1^2u &= 0 \\ 16 - x_1^2x_2 &\leqq 0 \\ u(16 - x_1^2x_2) &= 0 \\ u &\geqq 0 \\ x_1, x_2 &> 0 \end{aligned}$$

A fenti rendszer megoldását az u Lagrange szorzó értékeitől függően 2 esetre bontjuk szét, amelyek a következők.

i) $u = 0$ eset
Az első két egyenletből azonnal adódik az $x_1 = 0$ és $x_2 = 0$ megoldás, ami nem primál lehetséges megoldás, mert nem esik bele az X nyílt halmazba és a primál feltétel sem teljesül.

ii) $u > 0$ eset
Az első két egyenletből és a komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 4x_1 + 2x_2 - 2x_1x_2u &= 0 \\ 2x_1 - x_1^2u &= 0 \\ 16 - x_1^2x_2 &= 0 \end{aligned}$$

Egyszerű számolással adódik, hogy

$$x_1 = 2, \quad x_2 = 4, \quad u = 1.$$

Láthatjuk, hogy a megoldás maradéktalanul teljesíti az összes feltételt, így megkaptuk a KKT pontot és a hozzátartozó Lagrange szorzót.

Az optimum elégséges feltételének vizsgálatához a $\mathbf{C}$ mátrixot egyszerű behelyettesítéssel kapjuk, mert a feltétel aktív és $u > 0$:

$$\mathbf{C} = \begin{bmatrix} -4 & -2 & -16 \\ -2 & 0 & -4 \\ -16 & -4 & 0 \end{bmatrix}.$$

Gyakorlásképpen a főminor tesztet végezzük el. Jelen példában $n = 2, s = 1$. Csak egyetlen determinánst kell kiszámítani, hisz a vonatkozó tétel szerint a determinánsok indexe: $k = s + 1, \ldots, n$, azaz jelen példában $k = 2$. A keresett determináns:

$$\det(\mathbf{C}) = \det\left(\begin{bmatrix} -4 & -2 & -16 \\ -2 & 0 & -4 \\ -16 & -4 & 0 \end{bmatrix}\right) = -192.$$

Mivel $(-1)^s \det(\mathbf{C}) = 192 > 0$, így a teszt szerint az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $x_1 = 2, x_2 = 4$ megoldás szigorú lokális minimum.

Most gyakorlásképpen végezzük el az inerciatesztet is. A számítás során keletkező Schur-komplemensek:

$$\begin{bmatrix} -4 & -2 & -16 \\ -2 & 0 & -4 \\ -16 & -4 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 \\ 4 & 64 \end{bmatrix}, \quad [48].$$

Az első pivotelem $-4 < 0$, az inercia $Iner = (1, 0, 0)$; következő pivotelem $1 > 0$, az inercia $Iner = (0, 0, 1)$; a végső blokk egy elemű, értéke $48 > 0$, az inercia $Iner = (0, 0, 1)$. A számítás során kapott inerciákat összeadva, kapjuk, hogy $Iner(\mathbf{C}) = (1, 0, 2)$, tehát a szegélyezett aktualizált Lagrange függvény Hesse mátrixnak a méreteknek megfelelő számú sajátértéke van, azaz $n = 2$ darab pozitív sajátértéke, $s = 1$ darab negatív sajátértéke és nincs zérus sajátértéke. Az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $r = 2$, $m = 4$ megoldás szigorú lokális minimum. Tehát ugyanarra az eredményre jutottunk mindkét teszt segítségével.

Megjegyezzük, hogy e példában a gyenge elégségességi feltétel nem alkalmazható, mivel az aktualizált Lagrange függvény Hesse mátrix ($\mathbf{C}$ mátrix 2×2-es főminormátrixa) nem pozitív definit (indefinit).

Feladat:
Tervezze meg azt a konzervdobozt, amelynek a felszíne legfeljebb egy adott $F > 0$ érték és a térfogata minél nagyobb legyen!

15. Példa

Adott az $x_2 = 4 + (x_1 + 4)^2$ parabola. A parabola az (x_1, x_2) síkot két tartományra osztja. Tekintsük azt a tartományt, amely az origót nem tartalmazza. Határozza meg a tartomány azon pontjait, amelyeknél az $f(x_1, x_2) = x_1 x_2$ függvény értéke a legnagyobb! A megoldást az alábbi lépésekben végezze el:

a) Írja fel az optimalizálási feladatot matematikai formában!

b) Határozza meg az összes KKT pontot!

c) Döntse el, hogy az egyes KKT pontok közül melyik lokális maximumpont!
d) Határozza meg a globális maximumpontot!

a) Az optimalizálási feladat matematikai megfogalmazása:

$$\begin{aligned} f(x_1, x_2) &= x_1 x_2 \to \max! \\ g(x_1, x_2) &= x_2 - (x_1+4)^2 - 4 \geqq 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u) = x_1 x_2 + u(x_2 - (x_1+4)^2 - 4)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2 - 2u\,(x_1+4) \\ x_1 + u \\ x_2 - (x_1+4)^2 - 4 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} -2u & 1 & -2(x_1+4) \\ 1 & 0 & 1 \\ -2(x_1+4) & 1 & 0 \end{bmatrix}.$$

b) A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 - 2u\,(x_1+4) &= 0 \\ x_1 + u &= 0 \\ x_2 - (x_1+4)^2 - 4 &\geqq 0 \\ u(x_2 - (x_1+4)^2 - 4) &= 0 \\ u &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását az u Lagrange szorzó értékeitől függően 2 esetre bontjuk szét, amelyek a következők.

i) $u = 0$ eset
Az első két egyenletből azonnal adódik az $x_1 = 0$ és $x_2 = 0$ megoldás, de ez nem teljesíti a primál feltételt, így nem KKT pont.

ii) $u > 0$ eset
A megoldandó rendszer:

$$\begin{aligned} x_2 - 2u\,(x_1+4) &= 0 \\ x_1 + u &= 0 \\ x_2 - (x_1+4)^2 - 4 &= 0 \end{aligned}$$

A második egyenletből $x_1 = -u$, az elsőből $x_2 = 8u - 2u^2$ és ezeket a harmadik egyenletbe behelyettesítjük a $3u^2 - 16u + 20 = 0$ egyenletet kapjuk, amelynek megoldásai: $u_1 = 2$ és a $u_2 = \frac{10}{3}$, mindkettő pozitív.
Az $u_1 = 2$ esetben $\mathbf{x} = (-2, 8)$ és ez KKTpont.
Az $u_2 = \frac{10}{3}$ esetben $\mathbf{x} = (-\frac{10}{3}, \frac{40}{9})$ és ez KKTpont.

c) Lokális maximumpontok meghatározása:
A $\mathbf{C}$ mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixba történő behelyettesítéssel kapjuk, mivel a feltétel aktív, pozitív Lagrange szorzóval.

Az $\mathbf{x} = (-2, 8), u = 2$ eset vizsgálata:

$$\mathbf{C} = \begin{bmatrix} -4 & 1 & -4 \\ 1 & 0 & 1 \\ -4 & 1 & 0 \end{bmatrix}.$$

Inerciateszt segítségével számolva:

$$\begin{bmatrix} -4 & 1 & -4 \\ 1 & 0 & 1 \\ -4 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 4 \end{bmatrix}, \quad [4].$$

Mivel $Iner(\mathbf{C}) = (1, 0, 2)$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen nem negatív definit, tehát az $\mathbf{x} = (-2, 8)$ KKT pont nem maximumpont.

Az $\mathbf{x} = (-\frac{10}{3}, \frac{40}{9}), u = \frac{10}{3}$ eset vizsgálata:

$$\mathbf{C} = \begin{bmatrix} -\frac{20}{3} & 1 & -\frac{4}{3} \\ 1 & 0 & 1 \\ -\frac{4}{3} & 1 & 0 \end{bmatrix}.$$

A főminor teszt segítségével számolva: Mivel $n = 2, s = 1$, csak egyetlen determinánst kell kiszámítani, amely $\det(\mathbf{C}) = 4$ és $k = 2$. Mivel $(-1)^k \det(\mathbf{C}) = 4 > 0$, így a teszt szerint az aktualizált Lagrange függvény Hesse mátrixa feltételesen negatív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\mathbf{x} = (-\frac{10}{3}, \frac{40}{9})$ megoldás szigorú lokális maximumpont.

d) Az $\mathbf{x} = (-\frac{10}{3}, \frac{40}{9})$ lokális maximumpontban a célfüggvény értéke $-\frac{400}{27}$. Ha a tartományt és a célfüggvényt megvizsgáljuk, akkor könnyen találunk olyan pontokat (természetesen nem a pont környezetében), ahol a célfüggvény ennél az értéknél nagyobb, például az $\mathbf{x} = (1, 30)$ tartománypontban, ahol a célfüggvényérték 30. Az is könnyen látható, hogy a tartományban (ha $x_1 > 0$) a célfüggvényérték minden határon túl növelhető. Ez azt jelenti, hogy a feladatnak globális maximuma nincs. Tehát általában nem lehet egyszerűen azt mondani, hogy ha egyetlen lokális optimum van, akkor az globális is.

16. Példa

Adott a $(2, 1)$ középpontú, 5 sugarú körtartomány a határokkal. Határozza meg a tartomány azon pontjait, amelyek a $(10, 7)$ ponthoz legközelebb, ill. ugyanettől a ponttól legtávolabb vannak.

a) Írja fel az optimalizálási feladatot matematikai formában!
b) Határozza meg az összes KKT pontot!
c) Döntse el, hogy az egyes KKT pontok közül melyik lokális maximumpont!
d) Határozza meg a globális maximumpontot!

a) Az optimalizálási feladat matematikai megfogalmazása:

$$\begin{aligned} f(x_1, x_2) &= (x_1 - 10)^2 + (x_2 - 7)^2 \to \text{extrémum} \quad (\text{min!} \quad \text{vagy} \quad \text{max!}) \\ g(x_1, x_2) &= (x_1 - 2)^2 + (x_2 - 1)^2 - 25 \leqq 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u) = (x_1 - 10)^2 + (x_2 - 7)^2 + u((x_1 - 2)^2 + (x_2 - 1)^2 - 25)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2(x_1-10)+2u\,(x_1-2) \\ 2(x_2-7)+2u\,(x_2-1) \\ (x_1-2)^2+(x_2-1)^2-25 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2+2u & 0 & 2(x_1-2) \\ 0 & 2+2u & 2(x_2-1) \\ 2(x_1-2) & 2(x_2-1) & 0 \end{bmatrix}.$$

b) A minimum és a maximum feladat KKT pontjának meghatározásához az alábbi rendszert kell megoldani. Az u Lagrange szorzó előjele lehet pozitív és negatív is. Mivel a feltétel $\leqq 0$ típusú volt, így ha $u \geqq 0$, akkor a minimum feladat, ha $u \leqq 0$, akkor a maximum feladat KKT pontját kapjuk.

$$\begin{aligned} 2(x_1-10)+2u\,(x_1-2) &= 0 \\ 2(x_2-7)+2u\,(x_2-1) &= 0 \\ (x_1-2)^2+(x_2-1)^2-25 &\leqq 0 \\ u((x_1-2)^2+(x_2-1)^2-25) &= 0 \end{aligned}$$

A fenti rendszer megoldását az u Lagrange szorzó értékeitől függően 2 esetre bontjuk szét, amelyek a következők.

i) $u=0$ eset
Az első két egyenletből azonnal adódik az $x_1=10$ és $x_2=7$ megoldás, de ez nem teljesíti a primál feltételt, így nem KKT pont.

ii) $u \neq 0$ eset
A megoldandó rendszer:

$$\begin{aligned} 2(x_1-10)+2u\,(x_1-2) &= 0 \\ 2(x_2-7)+2u\,(x_2-1) &= 0 \\ (x_1-2)^2+(x_2-1)^2-25 &= 0 \end{aligned}$$

Érdemes bevezetni az $y_1=x_1-2$ és az $y_2=x_2-1$ változókat, ekkor az első két egyenletből $y_1=\frac{8}{1+u}, y_2=\frac{6}{1+u}$. A harmadik egyenletből pedig $(1+u)^2=4$ adódik. A Lagrange szorzó tehát $u_1=1$ és a $u_2=-3$.

Az $u_1=1>0$ Lagrange szorzó a minimum feladathoz tartozik és ennek KKT pontja $\mathbf{x}=(6,4)$.

Az $u_2=-3<0$ Lagrange szorzó a maximum feladathoz tartozik és ennek KKT pontja $\mathbf{x}=(-2,-2)$.

c) Lokális extrémumpontok meghatározása:
A $\mathbf{C}$ mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixba történő behelyettesítéssel kapjuk, mivel a feltétel aktív, egyik Lagrange szorzó sem zérus.

Az $\mathbf{x}=(6,4), u=1$ eset vizsgálata:

$$\mathbf{C} = \begin{bmatrix} 4 & 0 & 8 \\ 0 & 4 & 6 \\ 8 & 6 & 0 \end{bmatrix}.$$

Az $\mathbf{x} = (-2, -2), u = -3$ eset vizsgálata:

$$\mathbf{C} = \begin{bmatrix} -4 & 0 & -8 \\ 0 & -4 & -6 \\ -8 & -6 & 0 \end{bmatrix}.$$

Egyszerűen látható, hogy az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminormátrixa) az első esetben pozitív definit, a második esetben pedig negatív definit, így feltételesen is az. Ebből következik, hogy az $\mathbf{x} = (6, 4)$ KKT pont szigorú minimumpont, az $\mathbf{x} = (-2, -2)$ KKT pont szigorú maximumpont.

Feladat:
Végezze el az elégségességi vizsgálatot az inerciateszt és főminorteszt segítségével is!

d) A fentebb megállapított pontok egyben globális optimumpontok is, a célfüggvény és a feltétel vizsgálata alapján.

17. Példa
Adott az origó középpontú, $\sqrt{2}$ sugarú körtartomány a határokkal. Határozza meg a tartomány azon pontjait, amelyekben az $f(x_1, x_2) = x_1 x_2$ mennyiség a legkisebb, ill. legnagyobb értékű.
a) Írja fel az optimalizálási feladatot matematikai formában!
b) Határozza meg az összes KKT pontot!
c) Döntse el, hogy az egyes KKT pontok közül melyik lokális maximumpont!
d) Határozza meg a globális maximumpontot!

a) Az optimalizálási feladat matematikai megfogalmazása:

$$\begin{aligned} f(x_1, x_2) &= x_1 x_2 \to \text{extrémum} \quad (\text{min!} \quad \text{vagy} \quad \text{max!}) \\ g(x_1, x_2) &= x_1^2 + x_2^2 - 2 \leqq 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u) = x_1 x_2 + u(x_1^2 + x_2^2 - 2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2 + 2ux_1 \\ x_1 + 2ux_2 \\ x_1^2 + x_2^2 - 2 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2u & 1 & 2x_1 \\ 1 & 2u & 2x_2 \\ 2x_1 & 2x_2 & 0 \end{bmatrix}.$$

b) A minimum és a maximum feladat KKT pontjának meghatározásához az alábbi rendszert kell megoldani. Az u Lagrange szorzó előjele lehet pozitív és negatív is. Mivel a feltétel $\leqq 0$ típusú volt, így ha $u \geqq 0$, akkor a minimum feladat, ha $u \leqq 0$, akkor a maximum feladat KKT pontját kapjuk.

$$\begin{aligned} x_2 + 2ux_1 &= 0 \\ x_1 + 2ux_2 &= 0 \\ x_1^2 + x_2^2 - 2 &\leqq 0 \\ u(x_1^2 + x_2^2 - 2) &= 0 \end{aligned}$$

A fenti rendszer megoldását az u Lagrange szorzó értékeitől függően 2 esetre bontjuk szét, amelyek a következők.

i) $u = 0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x} = (0,0)$ megoldás és ez teljesíti a primál feltételt, így KKT pont, elviekben lehet a minimum és a maximum feladat KKT pontja is.

ii) $u \neq 0$ eset

A megoldandó rendszer:

$$\begin{aligned} x_2 + 2ux_1 &= 0 \\ x_1 + 2ux_2 &= 0 \\ x_1^2 + x_2^2 - 2 &= 0 \end{aligned}$$

Az első két egyenletből $x_1^2 = x_2^2$. A harmadik egyenletből pedig 4 megoldás adódik, amelyek a következők:

1. $\mathbf{x}_1 = (1,1), \quad u = -\frac{1}{2}$
2. $\mathbf{x}_2 = (-1,-1), \; u = -\frac{1}{2}$
3. $\mathbf{x}_3 = (1,-1), \quad u = \frac{1}{2}$
4. $\mathbf{x}_4 = (-1,1), \quad u = \frac{1}{2}$

Az $u = -\frac{1}{2} < 0$ Lagrange szorzó esetén a maximum feladat két KKT pontját kapjuk.
Az $u = \frac{1}{2} > 0$ Lagrange szorzó esetén a minimum feladat két KKT pontját kapjuk.

c) Lokális extrémumpontok meghatározása:

Először az $\mathbf{x} = (0,0)$ KKT pontot vizsgáljuk. A feltétel inaktív, így az aktualizált Lagrange függvény Hesse mátrixa a $\nabla^2 L$ mátrix 2×2-es főminormátrixa, amely behelyettesítés $(u = 0)$ után az alábbi

$$\nabla^2 L_a = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Mivel az aktualizált Lagrange függvény Hesse mátrixa indefinit, így az $\mathbf{x} = (0,0)$ KKT pont sem a minimum feladatnak, sem a maximum feladatnak nem lehet megoldása.

Másodszor pedig a négy pontot vizsgáljuk.

A $\mathbf{C}$ mátrixot a $\nabla^2 L$ mátrixba történő behelyettesítéssel kapjuk, mivel a feltétel aktív, egyik Lagrange szorzó sem zérus. A méretek miatt $(n = 2, s = 1, k = 2)$ alkalmazhatjuk a főminor tesztet is, csak a $\mathbf{C}$ mátrix determinánsára van szükség:

A maximum feladat KKT pontjainak vizsgálata:

$$\mathbf{C}_1 = \begin{bmatrix} -1 & 1 & 2 \\ 1 & -1 & 2 \\ 2 & 2 & 0 \end{bmatrix}, \qquad \mathbf{C}_2 = \begin{bmatrix} -1 & 1 & -2 \\ 1 & -1 & -2 \\ -2 & -2 & 0 \end{bmatrix}.$$

A determinánsok: $\det(\mathbf{C}_1) = 16, \quad \det(\mathbf{C}_2) = 16.$

Mivel $(-1)^k \det(\mathbf{C}_1) = (-1)^k \det(\mathbf{C}_2) > 0$, így a teszt szerint az aktualizált Lagrange függvény Hesse mátrixa feltételesen negatív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\mathbf{x}_1 = (1,1), \mathbf{x}_2 = (-1,-1)$ megoldások szigorú lokális maximumpontok.

A minimum feladat KKT pontjainak vizsgálata:

$$\mathbf{C}_3 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ 2 & -2 & 0 \end{bmatrix}, \qquad \mathbf{C}_4 = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 1 & 2 \\ -2 & 2 & 0 \end{bmatrix}.$$

A determinánsok: $\det(\mathbf{C}_3) = -16$, $\det(\mathbf{C}_4) = -16$.

Mivel $(-1)^s \det(\mathbf{C}_3) = (-1)^s \det(\mathbf{C}_4) > 0$, így a teszt szerint az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az erős elégségességi tétel alapján azt jelenti, hogy a kapott $\mathbf{x}_3 = (1, -1)$, $\mathbf{x}_3 = (-1, 1)$ megoldások szigorú lokális minimumpontok.

d) A fentebb megállapított lokális pontok egyben globális optimumpontok is.

18. Példa

Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} x_1x_2 + x_2x_3 + x_1x_3 &\rightarrow \text{min!} \\ -x_1 + 2x_2 - 3x_3 &\leqq 28 \\ x_1 + x_2 - x_3 &= 5 \end{aligned}$$

Megoldás

A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1, x_2, x_3) &= x_1x_2 + x_2x_3 + x_1x_3 \\ g(x_1, x_2, x_3) &= -x_1 + 2x_2 - 3x_3 - 28 \leqq 0 \\ h(x_1, x_2, x_3) &= x_1 + x_2 - x_3 - 5 = 0 \\ X &= \mathbb{R}^3 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, x_3, u, v) = x_1x_2 + x_2x_3 + x_1x_3 + u(-x_1 + 2x_2 - 3x_3 - 28) + v(x_1 + x_2 - x_3 - 5)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2 + x_3 - u + v \\ x_1 + x_3 + 2u + v \\ x_1 + x_2 - 3u - v \\ -x_1 + 2x_2 - 3x_3 - 28 \\ x_1 + x_2 - x_3 - 5 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & -3 & -1 \\ -1 & 2 & -3 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 + x_3 - u + v &= 0 \\ x_1 + x_3 + 2u + v &= 0 \\ x_1 + x_2 - 3u - v &= 0 \\ -x_1 + 2x_2 - 3x_3 - 28 &\leqq 0 \\ x_1 + x_2 - x_3 - 5 &= 0 \\ u(-x_1 + 2x_2 - 3x_3 - 28) &= 0 \\ u &\geqq 0 \end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset az $u > 0$, másik eset az $u = 0$. Nézzük az egyes eseteket.

i) $u > 0$ eset
Ekkor az 1., 2., 3., 5. és a 4. sorban szereplő most már egyenletből az x_1, x_2, x_3, u, v számokra kapjuk, hogy

$$x_1 = -8, \quad x_2 = 19, \quad x_3 = 6, \quad u = 9, \quad v = -16.$$

Az $\bar{\mathbf{x}} = (-8, 19, 6)$ pont KKT pont a $\bar{u} = 9, \bar{v} = -16$ Lagrange szorzókkal.

ii) $u = 0$ eset
Az 1., 2., 3. és az 5. sorban szereplő egyenletekből az x_1, x_2, x_3, v számokra kapjuk, hogy

$$x_1 = 1, \quad x_2 = 1, \quad x_3 = -3, \quad v = 2.$$

A 4. sorbeli egyenlőtlenség is fennáll, így az $\bar{\mathbf{x}} = (1, 1, -3)$ pont KKT pont a $\bar{u} = 0, \bar{v} = 2$ Lagrange szorzókkal.

Most meg kell győződni arról, hogy ezek optimális megoldások-e. Ezt az elégséges feltételek fennállásának ellenőrzésével fogjuk végezni.

Először az $u > 0$ esettel foglalkozunk. Mivel mindkét feltétel aktív és az $\bar{\mathbf{x}} = (-8, 19, 6)$ KKT pont reguláris, mivel a $\boldsymbol{\nabla} g = (-1, 2, -3)$ és a $\boldsymbol{\nabla} h = (1, 1, -1)$ vektorok, azaz a $\boldsymbol{\nabla}^2 L$ mátrix szegélyein található vektorok lineárisan függetlenek, így a $\mathbf{C}$ mátrix a $\boldsymbol{\nabla}^2 L$ mátrix lesz, azaz

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & -3 & -1 \\ -1 & 2 & -3 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{bmatrix}.$$

Gyakorlásképpen a főminor tesztet és az inerciatesztet is alkalmazzuk. A főminor tesztnél, mivel $n = 3$, $s = 2$, és $k = s + 1, \dots n$, így $k = 3$, azaz csak egy mátrix, mégpedig a teljes $\mathbf{C}$ mátrix determinánsát kell kiszámítani, amely nem kellemes számolás után $\det(\mathbf{C}) = 10$. Mivel $(-1)^s \det(\mathbf{C}) = 10 > 0$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, az erős elégségességi tételből pedig az következik, hogy ekkor az $\bar{\mathbf{x}} = (-8, 19, 6)$ megoldás szigorú lokális minimumpont.

Most gyakorlásképpen az inerciateszttel is ellenőrizzük, hogy a KKT pont szigorú lokális minimumpont. Mivel minden főelem zérus, ezért egy 2×2-es blokkot választunk, legyen ez a bal felső blokk, kettős pivotálást kell végezni, az inercia $Iner = (1, 0, 1)$; következő pivotelem $-2 < 0$, az inercia $Iner = (1, 0, 0)$; következő pivotelem $12 > 0$, az inercia $Iner = (0, 0, 1)$; a végső blokk egy elemű, értéke pozitív, az inercia $Iner = (0, 0, 1)$. A számítás során keletkező Schur-komplemensek:

$$\begin{bmatrix} 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & -3 & -1 \\ -1 & 2 & -3 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} -2 & -4 & -3 \\ -4 & 2 & -1 \\ -3 & -1 & -2 \end{bmatrix}, \quad \begin{bmatrix} 12 & 5 \\ 5 & \frac{5}{2} \end{bmatrix}, \quad \begin{bmatrix} \frac{5}{12} \end{bmatrix}.$$

Az egyes lépésekben kapott inerciákat összeadva, kapjuk, hogy $Iner(\mathbf{C}) = (2, 0, 3)$, tehát a szegélyezett aktualizált Lagrange függvény Hesse mátrixának $n = 3$ darab pozitív sajátértéke, $s = 2$ darab negatív sajátértéke és nincs zérus sajátértéke. Az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami az elégségességi tétel alapján azt jelenti, hogy a kapott $\bar{\mathbf{x}} = (-8, 19, 6)$ megoldás szigorú lokális minimum. Tehát ugyanarra az eredményre jutottunk.

Most a $u = 0$ esettel foglalkozunk. Az $\bar{\mathbf{x}} = (1, 1, -3)$ KKT pontban $g(\bar{\mathbf{x}}) = -10 < 0$, azaz inaktív, tehát az erős elégséges feltételt használhatjuk, hogy eldöntsük, hogy az adott KKT pont minimumpont-e. Itt most a $\mathbf{C}$ mátrix 4×4-es lesz, mert az inaktív egyenlőtlenségi feltétel gradiensét a szegélyen nem kell szerepelteni. Az inerciateszt végrehajtásánál keletkező Schur-komplemensek:

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} -2 & -3 \\ -3 & -2 \end{bmatrix}, \quad \begin{bmatrix} \frac{5}{2} \end{bmatrix}.$$

Az egyes lépésekben kapott inerciák rendre $Iner = (1, 0, 1)$; $Iner = (1, 0, 0)$; $Iner = (0, 0, 1)$; összegük $Iner(\mathbf{C}) = (2, 0, 2)$. Mivel nem igaz az, hogy $n = 3$ darab pozitív sajátérték és $s = 1$ darab negatív sajátérték van, így az $\bar{\mathbf{x}} = (1, 1, -3)$ KKT pont nem szigorú minimumpont.

Feladat:
Döntse el, hogy az $\bar{\mathbf{x}} = (1, 1, -3)$ pont maximumpont-e!

19. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} x_1 x_2 &\to \min! \\ 4x_1^2 + x_2^2 &\leqq 8 \end{aligned}$$

Megoldás
A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1, x_2) &= x_1 x_2 \\ g(x_1, x_2) &= 4x_1^2 + x_2^2 - 8 \leqq 0 \\ X &= \mathbb{R}^n \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, x_3, u, v) = x_1 x_2 + u(4x_1^2 + x_2^2 - 8)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2 + 8ux_1 \\ x_1 + 2ux_2 \\ 4x_1^2 + x_2^2 - 8 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 8u & 1 & 8x_1 \\ 1 & 2u & 2x_2 \\ 8x_1 & 2x_2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 + 8ux_1 &= 0 \\ x_1 + 2ux_2 &= 0 \\ 4x_1^2 + x_2^2 - 8 &\leqq 0 \\ u(4x_1^2 + x_2^2 - 8) &= 0 \\ u &\geqq 0 \end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset az $u > 0$, másik eset az $u = 0$. Nézzük az egyes eseteket.

i) $u > 0$ eset
Ekkor az 1., 2., és a 3. sorban szereplő most már egyenletből az x_1, x_2, u számokra négy megoldást kapunk, melyek az alábbiak

i	x_1	x_2	u
1	1	-2	$1/4$
2	-1	2	$1/4$
3	1	2	$-1/4$
4	-1	-2	$-1/4$

Ezekből a megoldásokból az első kettő KKT pont, mivel ezeknél pozitív a Lagrange szorzó, a feltétel mindkettőben aktív.

ii) $u = 0$ eset
Az 1., 2. sorban szereplő egyenletekből egyszerűen adódik, hogy

$$x_1 = 0, \quad x_2 = 0.$$

A 3. sorbeli egyenlőtlenség is fennáll, így az $\bar{\mathbf{x}} = (0, 0)$ pont KKT pont az $\bar{u} = 0$ Lagrange szorzóval.

Most meg kell győződni arról, hogy ezek optimális megoldások-e. Ezt az elégséges feltételek fennállásának ellenőrzésével fogjuk végezni.

Először a $u = 0$ esettel foglalkozunk. Mivel a feltétel inaktív, így a $\mathbf{C}$ mátrixot úgy kapjuk, hogy a 3. sort és 3. oszlopot elhagyjuk a $\boldsymbol{\nabla}^2 L$ mátrixból, vagyis az aktualizált Lagrange függvény Hesse mátrixát kapjuk, mivel $u = 0$, így ez a célfüggvény Hesse mátrixával azonos:

$$\mathbf{C} = \boldsymbol{\nabla}^2 L_a(\bar{x}) = \boldsymbol{\nabla}^2 f(\bar{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Ez a mátrix indefinit, így az $\bar{\mathbf{x}} = (0, 0)$ pont KKT nem minimumpont.

Most az $u > 0$ esettel foglalkozunk, megállapíthatjuk, hogy az $\bar{\mathbf{x}} = (1, -2)$ és az $\bar{\mathbf{x}} = (-1, 2)$ KKT pontokban a feltétel aktív. A szegélyezett aktualizált Lagrange függvény Hesse mátrixa:

$$\mathbf{C} = \begin{bmatrix} 8u & 1 & 8x_1 \\ 1 & 2u & 2x_2 \\ 8x_1 & 2x_2 & 0 \end{bmatrix}.$$

Két KKT pont is van, ezért együttesen végezzük el az elégségességi tesztet. A főminor tesztet fogjuk alkalmazni. Mivel $n = 2$, $s = 1$, és $k = s+1, \dots n$, így $k = 2$, azaz csak egy mátrix, mégpedig a teljes $\mathbf{C}$ mátrix determinánsát kell kiszámítani, rövid számolás után kapjuk, hogy

$$\det(\mathbf{C}) = 32x_1x_2 - 32u(4x_1^2 + x_2^2).$$

Az $\bar{\mathbf{x}} = (1, -2)$ és az $\bar{\mathbf{x}} = (-1, 2)$ KKT pontokban (mindkettőnél $\bar{u} = \frac{1}{4}$) $\det(\mathbf{C}) = -128$. Mivel $(-1)^s \det(\mathbf{C}) = 128 > 0$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, az elégségesség tételéből pedig az következik, hogy ekkor az $\bar{\mathbf{x}} = (1, -2)$ és az $\bar{\mathbf{x}} = (-1, 2)$ megoldások mindegyike szigorú lokális minimumpont.

Javasoljuk az olvasónak az inerciateszttel történő ellenőrzést.

Feladat:
Igazolja, hogy a fennmaradó $\bar{\mathbf{x}} = (1, 2)$ és az $\bar{\mathbf{x}} = (-1, -2)$ pontok maximumpontok.

20. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} 2x_1^3 + x_2^3 &\rightarrow \text{min!} \\ x_1 + 2x_2 &\geqq 15 \end{aligned}$$

Megoldás
A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(x_1, x_2) &= 2x_1^3 + x_2^3 \\ g(x_1, x_2) &= -x_1 - 2x_2 + 15 \leqq 0 \\ X &= \mathbb{R}^2 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u) = 2x_1^3 + x_2^3 + u(-x_1 - 2x_2 + 15)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 6x_1^2 - u \\ 3x_2^2 - 2u \\ -x_1 - 2x_2 + 15 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 12x_1 & 0 & -1 \\ 0 & 6x_2 & -2 \\ -1 & -2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 6x_1^2 - u &= 0 \\ 3x_2^2 - 2u &= 0 \\ -x_1 - 2x_2 + 15 &\leqq 0 \\ u(-x_1 - 2x_2 + 15) &= 0 \\ u &\geqq 0 \end{aligned}$$

Az u-ra vonatkozó egyenlőtlenségnek megfelelően a megoldást két részre bontjuk, egyik eset az $u > 0$, másik eset az $u = 0$. Nézzük az egyes eseteket.

i) $u = 0$ eset
Az 1., 2. sorban szereplő egyenletekből

$$x_1 = 0, \quad x_2 = 0.$$

A 3. sorbeli egyenlőtlenség nem áll fenn, így az $\bar{\mathbf{x}} = (0, 0)$ pont nem KKT pont.

ii) $u > 0$ eset
Ekkor az 1., 2. és a 3. sorban szereplő most már egyenletből az x_1, x_2, u számokra két megoldást kapunk, melyek az alábbiak

i	x_1	x_2	u
1	3	6	54
2	-5	10	125

Mindkettő KKT pont, mivel a Lagrange szorzó pozitív.

Most meggyőződünk arról, hogy ezek optimális megoldások-e. Mindkét pontban a feltétel aktív, így a szegélyezett aktualizált Lagrange függvény Hesse mátrixát kell vizsgálnunk, ami a következő:

$$\mathbf{C} = \begin{bmatrix} 12x_1 & 0 & -1 \\ 0 & 6x_2 & -2 \\ -1 & -2 & 0 \end{bmatrix}.$$

A főminor tesztet fogjuk alkalmazni, mert csak egy determinánst, a teljes $\mathbf{C}$ mátrix determinánsát kell kiszámítani ($n = 2$, $s = 1$, és $k = s + 1, \ldots n$, így $k = 2$). A determináns

$$\det(\mathbf{C}) = -48x_1 - 6x_2.$$

Az $\bar{\mathbf{x}} = (3, 6)$ KKT pontban $\det(\mathbf{C}) = -180$. Mivel $(-1)^s \det(\mathbf{C}) = 180 > 0$, így a szegélyezett aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, az erős elégségességi tételéből pedig az következik, hogy ekkor az $\bar{\mathbf{x}} = (3, 6)$ megoldás szigorú lokális minimumpont.

Az $\bar{\mathbf{x}} = (-5, 10)$ KKT pontban $\det(\mathbf{C}) = 180$. Mivel $(-1)^s \det(\mathbf{C}) = -180 < 0$, így a szegélyezett aktualizált Lagrange függvény Hesse mátrixa csak azt tudja eldönteni, hogy az $\bar{\mathbf{x}} = (-5, 10)$ megoldás nem lehet szigorú lokális minimumpont.

Javasoljuk az olvasónak az inerciateszttel történő ellenőrzést.

Feladat:
Vizsgálja meg, hogy az $\bar{\mathbf{x}} = (-5, 10)$ megoldás maximumpont vagy nem!

21. Példa
Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} \sum_{j=1}^{n} \frac{c_j}{x_j} &\rightarrow \text{min!} \\ \sum_{j=1}^{n} a_j x_j &= b \\ x_1, x_2, \ldots, x_n &> 0 \end{aligned}$$

ahol a_j, c_j $(j = 1, \ldots, n)$ és b pozitív számok.

Megoldás
A korábbi jelöléseknek megfelelően a feladat megszokott formája:

$$\begin{aligned} f(\mathbf{x}) &= \sum_{j=1}^{n} \frac{c_j}{x_j} \\ h(\mathbf{x}) &= \sum_{j=1}^{n} a_j x_j - b = 0 \\ X &= \{\mathbf{x} \in \mathbb{R}^n : x_j > 0,\ j = 1, \ldots, n\} \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, \ldots, x_n, v) = \frac{c_1}{x_1} + \frac{c_2}{x_2} + \ldots + \frac{c_n}{x_n} + v\,(a_1 x_1 + a_2 x_2 + a_n x_n - b)$$

A feladat függvényei differenciálhatók az $x_j > 0$ tartományban, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} -\frac{c_1}{x_1^2} + v a_1 \\ -\frac{c_2}{x_2^2} + v a_2 \\ \vdots \\ -\frac{c_n}{x_n^2} + v a_n \\ \sum_{j=1}^{n} a_j x_j - b \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} \frac{2c_1}{x_1^3} & 0 & \ldots & 0 & a_1 \\ 0 & \frac{2c_2}{x_2^3} & \ldots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \frac{2c_n}{x_n^3} & a_n \\ a_1 & a_2 & \ldots & a_n & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} -\frac{c_j}{x_j^2} + v a_j &= 0, \qquad j = 1, \ldots, n \\ \sum_{j=1}^{n} a_j x_j &= b \end{aligned}$$

Az első sor egyenleteiből x_j-re kapjuk, hogy

$$x_j = \frac{1}{\sqrt{v}} \sqrt{\frac{c_j}{a_j}}, \qquad j = 1, \ldots, n$$

amelyet a második sorbeli, azaz a feltételi egyenletbe behelyettesítve a v Lagrange szorzóra az alábbi adódik

$$v = \left(\frac{\sum_{k=1}^{n} \sqrt{c_k a_k}}{b} \right)^2,$$

ezt visszahelyettesítve az első egyenletbe, az alábbi KKT pontot kapjuk:

$$\bar{x}_j = \sqrt{\frac{c_j}{a_j}} \frac{b}{\sum_{k=1}^{n} \sqrt{c_k a_k}}, \qquad j = 1, \ldots, n$$

Most meggyőződünk arról, hogy a fenti KKT pont optimális megoldás-e. A szegélyezett mátrix:

$$\mathbf{C} = \begin{bmatrix} \frac{2c_1}{\bar{x}_1^3} & 0 & \dots & 0 & a_1 \\ 0 & \frac{2c_2}{\bar{x}_2^3} & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{2c_n}{\bar{x}_n^3} & a_n \\ a_1 & a_2 & \dots & a_n & 0 \end{bmatrix}.$$

Az inerciateszttel ellenőrizzük, hogy a KKT pont lokális minimumpont. Az aktualizált Lagrange függvény Hesse mátrixának blokkja diagonális mátrix és a főátlóbeli elemek pozitívak. Ha a pivotelemet mindig a főátlóban fentről lefelé választjuk, akkor azt tapasztaljuk, hogy így a Schur-komplemensek egy elem kivételével azonosak lesznek a $\mathbf{C}$ mátrix megfelelő elemeivel, a nem egyező elem a jobb alsó elem, amely mindig csökken. Az alábbi 1×1-es Schur-komplemenst kapjuk n darab principális pivotálás után

$$\left[-\frac{a_1^2\bar{x}_1^3}{2c_1} - \frac{a_2^2\bar{x}_2^3}{2c_2} - \dots - \frac{a_n^2\bar{x}_n^3}{2c_n}\right].$$

Az n darab principális pivotálás során az inercia $Iner = (0, 0, n)$, mert mindegyik pivotelem pozitív; az 1×1-es blokk pedig negatív, így $Iner = (1, 0, 0)$. Mivel $Iner(\mathbf{C}) = (1, 0, n)$, így az inerciateszt alapján az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, ami azt jelenti, hogy a kapott KKT pont szigorú lokális minimum.

Megjegyezzük, hogy az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix $n \times n$-es főminormátrixa) pozitív definit, mivel diagonális mátrix és a főátlóban lévő összes elem pozitív, így a gyenge elégségességi feltételt is alkalmazhattuk volna.

22. Példa

Legyen az (x_1, x_2) síkon adva a $-x_1 + x_2 = 3$ egyenes és az egyenes egyik oldalán két pont, amelyek legyenek $A(2, 1)$ és $B(8, 7)$. Az A pontból szeretnénk eljutni a B pontba egyenes vonalú, egyenletes sebességű mozgással úgy, hogy érintsük az egyenest és a megtett út hossza a lehető legrövidebb legyen.

Megoldás

Legyen a keresett pont az egyenesen a $P(x_1, x_2)$ pont. Feladatunk úgy meghatározni az egyenesen lévő $P(x_1, x_2)$ pontot, hogy az AP és a PB szakaszok hosszának összege a legkisebb legyen. Az alábbi feltételes optimalizálási feladatot kell megoldani:

$$\begin{aligned} \sqrt{(x_1-2)^2 + (x_2-1)^2} + \sqrt{(x_1-8)^2 + (x_2-7)^2} &\rightarrow \min! \\ -x_1 + x_2 - 3 &= 0 \end{aligned}$$

A feladat Lagrange függvénye:

$$L(x_1, x_2, v) = \sqrt{(x_1-2)^2 + (x_2-1)^2} + \sqrt{(x_1-8)^2 + (x_2-7)^2} + v(-x_1 + x_2 - 3)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} \frac{x_1-2}{\sqrt{(x_1-2)^2+(x_2-1)^2}} + \frac{x_1-8}{\sqrt{(x_1-8)^2+(x_2-7)^2}} - v \\ \frac{x_2-1}{\sqrt{(x_1-2)^2+(x_2-1)^2}} + \frac{x_2-7}{\sqrt{(x_1-8)^2+(x_2-7)^2}} + v \\ -x_1+x_2-3 \end{bmatrix},$$

$$\nabla^2 L = \begin{bmatrix} \frac{(x_2-1)^2}{\left((x_1-2)^2+(x_2-1)^2\right)^{\frac{3}{2}}} + \frac{(x_2-7)^2}{\left((x_1-8)^2+(x_2-7)^2\right)^{\frac{3}{2}}} & -\frac{(x_1-2)(x_2-1)}{\left((x_1-2)^2+(x_2-1)^2\right)^{\frac{3}{2}}} - \frac{(x_1-8)(x_2-7))}{\left((x_1-8)^2+(x_2-7)^2\right)^{\frac{3}{2}}} & -1 \\ -\frac{(x_1-2)(x_2-1)}{\left((x_1-2)^2+(x_2-1)^2\right)^{\frac{3}{2}}} - \frac{(x_1-8)(x_2-7)}{\left((x_1-8)^2+(x_2-7)^2\right)^{\frac{3}{2}}} & \frac{(x_1-2)^2}{\left((x_1-2)^2+(x_2-1)^2\right)^{\frac{3}{2}}} + \frac{(x_1-8)^2}{\left((x_1-8)^2+(x_2-7)^2\right)^{\frac{3}{2}}} & 1 \\ -1 & 1 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} \frac{x_1-2}{\sqrt{(x_1-2)^2+(x_2-1)^2}} + \frac{x_1-8}{\sqrt{(x_1-8)^2+(x_2-7)^2}} - v = 0 \\ \frac{x_2-1}{\sqrt{(x_1-2)^2+(x_2-1)^2}} + \frac{x_2-7}{\sqrt{(x_1-8)^2+(x_2-7)^2}} + v = 0 \\ -x_1+x_2-3 = 0 \end{aligned}$$

A KKT pont és a Lagrange szorzó: $x_1 = 3,\ x_2 = 6,\ v = -0.784$. A szegélyezett mátrix

$$\mathbf{C} = \begin{bmatrix} 0.19612 & -0.07543 & -1 \\ -0.07543 & 0.29341 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

Mivel $n = 2$ és $s = 1$, így a $\mathbf{C}$ mátrix determinánsát kell csak kiszámítani, amelynek értéke $\det(\mathbf{C}) = -0.24138$. Mivel $(-1)^s \det(C) > 0$, így a főminorteszt alapján megállapítható, hogy a Lagrange függvény Hesse mátrixa feltételesen pozitív definit. Tehát a kapott $P(3,6)$ ponton való áthaladás esetén tesszük meg a legrövidebb utat az A és B pontok között. A minimális úthossz pedig: $2\sqrt{26} \approx 10.2$.

A példánál maradva, tegyük fel, hogy az AP szakaszt egyenletes $v = 2$ sebességgel, a PB szakaszt pedig szintén egyenletes, de $v = 3$ sebességgel tesszük meg. Az A pontból most úgy szeretnénk eljutni a B pontba, hogy érintsük az egyenest és az utat a lehető legrövidebb idő alatt tegyük meg. Ekkor az alábbi feltételes optimalizálási feladatot kell megoldani:

$$\begin{aligned} \frac{1}{2}\sqrt{(x_1-2)^2+(x_2-1)^2} + \frac{1}{3}\sqrt{(x_1-8)^2+(x_2-7)^2} &\to \min! \\ -x_1+x_2-3 &= 0 \end{aligned}$$

A feladat megoldása: $P(1.53, 4.53)$, a Lagrange szorzó: -0.38, a szükséges minimális idő: 4.1, a megtett út hossza: 10.5 .

Most pedig újabb sebességviszonyokkal is oldjuk meg a feladatot, az AP szakaszt $v = 4$, a PB szakaszt $v = 3$ sebességgel tegyük meg.

A feladat megoldása: $P(4.16, 7.16)$, a Lagrange szorzó: -0.25, a szükséges minimális idő: 2.9, a megtett út hossza: 10.4 .

A feladat grafikusan is megoldható az alábbi módon: valamelyik pontot tükrözzük az egyenesre merőlegesen, a tükörpontot és a másik pontot összekötő egyenes szakasznak az egyenessel való metszéspontja lesz a megoldás.

Megfigyelhetjük, ha nem azonos sebességgel haladunk, akkor a megoldás eltér ettől a ponttól, mégpedig a kisebb sebességű szakasz irányába. Természetesen a megtett távolság növekszik.

23. Példa

Adott $R > 0$ sugarú gömbbe írható egyenes körhengererek közül melyiknek a térfogata maximális?

Megoldás

Jelölje x_1 a körhenger sugarát, x_2 a körhenger magasságának a felét. Ha a gömbbe be akarjuk írni a hengert, akkor annak alapköre és fedőköre is a gömbön van, ekkor pedig a körhenger sugara és a magasságának fele között az $x_1^2 + x_2^2 = R^2$ összefüggésnek fenn kell állnia. A feltételes optimalizálási feladat tehát az alábbi:

$$\begin{aligned} x_1^2 \pi x_2 &\to \max! \\ x_1^2 + x_2^2 &= R^2 \\ x_1, x_2 &> 0 \end{aligned}$$

A feladat Lagrange függvénye (oszthatjuk a célfüggvényt π-vel):

$$L(x_1, x_2, v) = x_1^2 x_2 + v(R^2 - x_1^2 - x_2^2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} 2x_1x_2 - 2x_1v \\ x_1^2 - 2x_2v \\ R^2 - x_1^2 - x_2^2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} 2x_2 - 2v & 2x_1 & -2x_1 \\ 2x_1 & -2v & -2x_2 \\ -2x_1 & -2x_2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2x_1x_2 - 2x_1v &= 0 \\ x_1^2 - 2x_2v &= 0 \\ R^2 - x_1^2 - x_2^2 &= 0 \end{aligned}$$

A KKT pont és a Lagrange szorzó:

$$x_1 = \sqrt{2}\frac{R}{\sqrt{3}}, \qquad x_2 = \frac{R}{\sqrt{3}}, \qquad v = \frac{R}{\sqrt{3}}.$$

A szegélyezett mátrix:

$$\mathbf{C} = \begin{bmatrix} 0 & 2\sqrt{2}\frac{R}{\sqrt{3}} & -2\sqrt{2}\frac{R}{\sqrt{3}} \\ 2\sqrt{2}\frac{R}{\sqrt{3}} & -2\frac{R}{\sqrt{3}} & -2\frac{R}{\sqrt{3}} \\ -2\sqrt{2}\frac{R}{\sqrt{3}} & -2\frac{R}{\sqrt{3}} & 0 \end{bmatrix}$$

A példában $n = 2$ és $s = 1$, tehát $k = 2$, így a $\mathbf{C}$ mátrix determinánsát kell csak kiszámítani, amelynek értéke $\det(\mathbf{C}) = \frac{16}{\sqrt{3}}R^3$. Mivel $(-1)^k \det(\mathbf{C}) > 0$ minden $R > 0$ esetén, így a főminorteszt alapján megállapítható, hogy a Lagrange függvény Hesse mátrixa feltételesen negatív definit, ami azt jelenti, hogy a KKT pont maximumpont.

Tehát a maximális térfogatú körhenger sugara $\sqrt{2}\frac{R}{\sqrt{3}}$, magassága $2\frac{R}{\sqrt{3}}$, térfogata pedig $\frac{1}{\sqrt{3}}\frac{4R^3\pi}{3}$. Mint tapasztalhatjuk a a magasság a sugár $\sqrt{2}$-szöröse, a maximális térfogat pedig a gömb térfogatának $\frac{1}{\sqrt{3}}$-szorosa.

Feladat:
Adott $R > 0$ sugarú gömbbe írható egyenes körhengererek közül melyiknek a teljes felszíne maximális?

24. Példa
Mekkora annak a téglatestnek a három kiterjedése, amelynek az egy csúcsban találkozó élek hosszúságának összege legfeljebb $s > 0$, a térfogata pedig maximális?

Megoldás
Jelölje x_1, x_2, x_3 az oldalak hosszát. Ekkor a feltételes optimalizálási feladat:

$$\begin{aligned} x_1x_2x_3 &\rightarrow \max! \\ x_1 + x_2 + x_3 &\leqq s \\ x_1, x_2, x_3 &> 0 \end{aligned}$$

A feladatot maximum feladatként kezeljük, így a feltétel: $g(\mathbf{x}) = s - x_1 - x_2 - x_3 \geqq 0$.
A feladat Lagrange függvénye:

$$L(x_1, x_2, x_3, u) = x_1x_2x_3 + u(s - x_1 - x_2 - x_3)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2x_3 - u \\ x_1x_3 - u \\ x_1x_2 - u \\ s - x_1 - x_2 - x_3 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 0 & x_3 & x_2 & -1 \\ x_3 & 0 & x_1 & -1 \\ x_2 & x_1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2x_3 - u &= 0 \\ x_1x_3 - u &= 0 \\ x_1x_2 - u &= 0 \\ s - x_1 - x_2 - x_3 &= 0 \\ u(s - x_1 - x_2 - x_3) &= 0 \\ u &\geqq 0 \end{aligned}$$

A KKT pont és a Lagrange szorzó, amelyet az olvasó is könnyen ellenőrizhet:

$$x_1 = \frac{s}{3}, \qquad x_2 = \frac{s}{3}, \qquad x_3 = \frac{s}{3}, \qquad u = \frac{s^2}{9} > 0.$$

A szegélyezett mátrix:

$$\mathbf{C} = \begin{bmatrix} 0 & \frac{1}{3}s & \frac{1}{3}s & -1 \\ \frac{1}{3}s & 0 & \frac{1}{3}s & -1 \\ \frac{1}{3}s & \frac{1}{3}s & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

Az inerciatesztet érdemes alkalmazni, először kettős pivotálást kell végezni, célszerűen valamelyik -1-es elemet választva pivotelemnek. A $\mathbf{C}$ mátrix inerciája $Iner(\mathbf{C}) = (3,0,1)$ minden $s > 0$ esetén, így a Lagrange függvény Hesse mátrixa feltételesen negatív definit, ami azt jelenti, hogy a KKT pont maximumpont.

Tehát azt kaptuk, hogy a három oldal azonos, azaz kockát kaptunk, az oldalak hossza $\frac{s}{3}$, a maximális térfogat $\frac{s^3}{27}$.

Feladat:

Mekkora annak a téglatestnek a három kiterjedése, amelynek a térfogata legalább $V > 0$, a felszíne pedig minimális?

25. Példa

Adottak az $x_1 - x_2 + 1 = 0$ és a $2x_1 - x_2 - 2 = 0$ egyenesek. A két egyenes az (x_1, x_2) síkot négy tartományra osztja. Tekintsük azt a tartományt, amely az origót tartalmazza. Határozza meg a tartomány azon pontjait, amelyeknél az $f(x_1, x_2) = x_1x_2$ függvény értéke a legkisebb! A megoldást az alábbi lépésekben végezze el:

a) Írja fel az optimalizálási feladatot matematikai formában!

b) Határozza meg az összes KKT pontot!

c) Döntse el, hogy az egyes KKT pontok közül melyik lokális minimumpont!

d) Határozza meg a globális minimumpontot!

Megoldás

a) A matematikai modell felírása:

Ahhoz, hogy felírhassuk a feladatot matematikai formában meg kell határozni a feltételi halmazt. Egyszerű behelyettesítéssel megállapítható, hogy az origó, az $\mathbf{x} = (0,0)$ pont egyik egyenesen sincs rajta, tehát akkor az origó az egyenesek által kijelölt valamelyik féltérben van. Mivel az origóban $x_1 - x_2 + 1 > 0$ és $2x_1 - x_2 - 2 < 0$, így a keresett félterek az $x_1 - x_2 + 1 \geqq 0$ és a $2x_1 - x_2 - 2 \leqq 0$ feltételekkel írhatók le. Tehát a korábbi jelöléseknek megfelelően a feladat matematikai formája:

$$\begin{aligned}
f(x_1, x_2) &= x_1x_2 \to \min!\\
g_1(x_1, x_2) &= -x_1 + x_2 - 1 \leqq 0\\
g_2(x_1, x_2) &= 2x_1 - x_2 - 2 \leqq 0\\
X &= \mathbb{R}^2
\end{aligned}$$

b) KKT pont meghatározása

A feladat Lagrange függvénye:

$$L(x_1, x_2, u_1, u_2) = x_1x_2 + u_1(-x_1 + x_2 - 1) + u_2(2x_1 - x_2 - 2)$$

A függvények differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} x_2 - u_1 + 2u_2 \\ x_1 + u_1 - u_2 \\ -x_1 + x_2 - 1 \\ 2x_1 - x_2 - 2 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 0 & 1 & -1 & 2 \\ 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{bmatrix}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 - u_1 + 2u_2 &= 0 \\ x_1 + u_1 - u_2 &= 0 \\ -x_1 + x_2 - 1 &\leqq 0 \\ 2x_1 - x_2 - 2 &\leqq 0 \\ u_1(-x_1 + x_2 - 1) &= 0 \\ u_2(2x_1 - x_2 - 2) &= 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását az u_1, u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1 = u_2 = 0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x} = (0,0)$ pont. Mivel ez a megoldás teljesíti a két primál feltételt is, így az $\mathbf{x} = (0,0)$ megoldás KKT pont.

ii) $u_1 > 0, u_2 = 0$ eset

Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 - u_1 &= 0 \\ x_1 + u_1 &= 0 \\ -x_1 + x_2 - 1 &= 0 \end{aligned}$$

A megoldás: $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{2}$, $u_1 = \frac{1}{2}$. Ez a megoldás maradéktalanul teljesíti a többi KKT feltételt, így az $\mathbf{x} = (-\frac{1}{2}, \frac{1}{2})$ vektor KKT pont az $u_1 = \frac{1}{2}, u_2 = 0$ Lagrange szorzókkal úgy, hogy a g_1 feltétel aktív, a g_2 feltétel inaktív.

iii) $u_1 = 0, u_2 > 0$ eset

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 + 2u_2 &= 0 \\ x_1 - u_2 &= 0 \\ 2x_1 - x_2 - 2 &= 0 \end{aligned}$$

A megoldás: $x_1 = \frac{1}{2}$, $x_2 = -1$, $u_2 = \frac{1}{2}$. Ez a megoldás maradéktalanul teljesíti a többi KKT feltételt, így az $\mathbf{x} = (\frac{1}{2}, -1)$ vektor KKT pont az $u_1 = 0, u_2 = \frac{1}{2}$ Lagrange szorzókkal úgy, hogy a g_1 feltétel inaktív, a g_2 feltétel aktív.

iv) $u_1 > 0, u_2 > 0$ eset

Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 - u_1 + 2u_2 &= 0 \\ x_1 + u_1 - u_2 &= 0 \\ -x_1 + x_2 - 1 &= 0 \\ 2x_1 - x_2 - 2 &= 0 \end{aligned}$$

A megoldás: $x_1 = 3$, $x_2 = 4$, $u_1 = -10$, $u_2 = -7$. Az $\mathbf{x} = (3,4)$ nem lehet KKT pont, mert a Lagrange szorzókra nem teljesül a nemnegativitás.

Összefoglalva az eseteket, három KKT pont adódott, amelyek az alábbiak

1. KKT pont: $x_1 = 0$, $x_2 = 0$, $u_1 = 0$, $u_2 = 0$, és egyik feltétel sem aktív,

2. KKT pont: $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{2}$, $u_1 = \frac{1}{2}$, $u_2 = 0$, és az első feltétel aktív, a második feltétel inaktív.

3. KKT pont: $x_1 = \frac{1}{2}$, $x_2 = -1$, $u_1 = 0$, $u_2 = \frac{1}{2}$, és az első feltétel inaktív, a második feltétel aktív.

c) Lokális minimum meghatározása

Most megvizsgáljuk az elégséges feltétel segítségével, hogy a KKT pontok optimális megoldások-e. Ehhez felépítjük a szükséges $\mathbf{C}$ mátrixot a Lagrange függvény Hesse mátrixa segítségével.

Az 1. KKT pont esetén a Hesse mátrixból törölnünk kell a 3., 4. sort és a 3., 4. oszlopot, a maradék mátrixba pedig be kell helyettesíteni a KKT pont meghatározásánál kapott x_1, x_2, u_1, u_2 értékeket.

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

A $\mathbf{C}$ mátrix jelen esetben a célfüggvény Hesse mátrixa, mert mindkét feltétel inaktív és $u_1 = 0$, $u_2 = 0$. Könnyen meggyőződhetünk arról, hogy a $\mathbf{C}$ mátrix indefinit, így az $\mathbf{x} = (0, 0)$ nem minimumpont, hanem nyeregpont.

A 2. KKT pont esetén a Hesse mátrixból törölnünk kell a 4. sort és a 4. oszlopot, a maradék mátrixba pedig be kell helyettesíteni a KKT pont meghatározásánál kapott x_1, x_2, u_1, u_2 értékeket.

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

Az inercia teszt alapján $Iner(\mathbf{C}) = (1, 0, 2)$ adódik, amely szerint az $\mathbf{x} = (-\frac{1}{2}, \frac{1}{2})$ szigorú lokális minimumpont.

A 3. KKT pont esetén a Hesse mátrixból törölnünk kell a 3. sort és a 3. oszlopot, a maradék mátrixba pedig be kell helyettesíteni a KKT pont meghatározásánál kapott x_1, x_2, u_1, u_2 értékeket.

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}$$

Az inercia teszt alapján $Iner(\mathbf{C}) = (1, 0, 2)$ adódik, amely szerint az $\mathbf{x} = (\frac{1}{2}, -1)$ szigorú lokális minimumpont.

d) Globális minimum meghatározása

Példánkban két szigorú lokális minimumpont is adódott, ezekből egyszerűen kiválaszthatjuk a globális minimumpontot, mégpedig a célfüggvénybe történő behelyettesítéssel. Az $f(-\frac{1}{2}, \frac{1}{2}) = -\frac{1}{4}$ és az $f(\frac{1}{2}, -1) = -\frac{1}{2}$ értékek közül az utóbbi a kisebb, így a feladat szigorú globális minimumpontja az $\mathbf{x} = (\frac{1}{2}, -1)$ pont, a célfüggvény minimuma:$-\frac{1}{2}$.

26. Példa

Tekintsünk egy gazdasági problémát. Legyen két termelési tényező, mondjuk tőke és munka. A termelés során a termelési tényezők felhasznált mennyiségét jelölje rendre x_1, x_2. A termelési tényezők egységára legyen $p_1 = 3, p_2 = 4$. Legyen továbbá adott az $f(x_1, x_2) = x_1^3 x_2^2$

ún. Cobb-Douglas-féle termelési függvény, amely a termelési tényezőkkel előállítható termékmennyiséget fejezi ki a termelési tényezők mennyiségének függvényében. Határozzuk meg a termelési tényezők azon optimális mennyiségét, amellyel a legkevesebb költséggel legalább $q = 8$ termékmennyiség állítható elő!

Megoldás

A probléma matematikai modelljét egyszerűen felírhatjuk, hisz a termelés költsége $3x_1 + 4x_2$, a termelési tényezők mennyisége pedig nem lehet negatív, így a matematikai modell a következő:

$$
\begin{aligned}
3x_1 + 4x_2 &\rightarrow \min! \\
x_1^3 x_2^2 &\geqq 8 \\
x_1, x_2 &\geqq 0
\end{aligned}
$$

A termékmennyiségre előírt feltételből azonnal látható, hogy egyik döntési változó sem lehet zérus, így a második feltétel $x_1, x_2 > 0$ alakban írható. A szokásos jelöléssel az optimalizálási feladat alakja:

$$
\begin{aligned}
f(x_1, x_2) &= 3x_1 + 4x_2 \rightarrow \min! \\
g(x_1, x_2) &= 8 - x_1^3 x_2^2 \leqq 0 \\
(x_1, x_2) &\in X = \{(x_1, x_2) : x_1, x_2 > 0\}
\end{aligned}
$$

A Lagrange függvény:

$$L(x_1, x_2, u) = 3x_1 + 4x_2 + u(8 - x_1^3 x_2^2)$$

A Lagrange függvény gadiense és Hesse mátrixa:

$$
\nabla L = \begin{bmatrix} 3 - 3ux_1^2 x_2^2 \\ 4 - 2ux_1^3 x_2 \\ 8 - x_1^3 x_2^2 \end{bmatrix}, \qquad
\nabla^2 L = \begin{bmatrix} -6ux_1 x_2^2 & -6ux_1^2 x_2 & -3x_1^2 x_2^2 \\ -6ux_1^2 x_2 & -2ux_1^3 & -2x_1^3 x_2 \\ -3x_1^2 x_2^2 & -2x_1^3 x_2 & 0 \end{bmatrix}.
$$

A KKT feladat:

$$
\begin{aligned}
3 - 3ux_1^2 x_2^2 &= 0 \\
4 - 2ux_1^3 x_2 &= 0 \\
8 - x_1^2 x_2^3 &\leqq 0 \\
u(8 - x_1^2 x_2^3) &= 0 \\
u &\geqq 0 \\
(x_1, x_2) &\in X
\end{aligned}
$$

Az u Lagrange szorzó csak pozitív lehet, amely az első két egyenletből azonnal látható.Az $u > 0$ esetben az alábbi egyenletrendszert kell megoldani:

$$
\begin{aligned}
3ux_1^2 x_2^2 &= 3 \\
2ux_1^3 x_2 &= 4 \\
x_1^2 x_2^3 &= 8
\end{aligned}
$$

Az első két egyenletet elosztjuk egymással és a kapott $x_1 = 2x_2$ összefüggést a harmadik egyenletbe behelyettesítve a megoldás, amely kielégíti az $(x_1, x_2) \in X$ feltételt, a következő:

$$x_1 = 2, \qquad x_2 = 1, \qquad u = \frac{1}{4}.$$

Most meggyőződünk arról, hogy az $\mathbf{x} = (2, 1)$ KKT pont valóban minimumpont. Ehhez a $\mathbf{C}$ mátrixot kell előállítani, amely a Lagrange függvény Hesse mátrixából egyszerű behelyettesítéssel adódik.

$$\mathbf{C} = \begin{bmatrix} -3 & -6 & -12 \\ -6 & -4 & -16 \\ -12 & -16 & 0 \end{bmatrix}.$$

Az inerciateszthez szükséges számítások:

$$\begin{bmatrix} -3 & -6 & -12 \\ -6 & -4 & -16 \\ -12 & -16 & 0 \end{bmatrix} \qquad \begin{bmatrix} 8 & 8 \\ 8 & 48 \end{bmatrix} \qquad [40]\,.$$

Mivel az $Iner(\mathbf{C}) = (1, 0, 2)$, így a KKT pont szigorú lokális minimumpont, tehát az első termelési tényezőből 2, a második termelési tényezőből pedig 1 mennyiséget kell felhasználni, hogy az előírt termékmennyiséget minimális költséggel tudjuk előállítani.

27. Példa

Tekintsünk egy másik gazdasági problémát. Legyen két termék és ezeket egy fogyasztó meg akarja vásárolni. A termékek egységára rendre $3, 4$ pénzegység. A fogyasztó által vásárolt termékmennyiségeket jelölje x_1, x_2. Legyen továbbá adott az $f(x_1, x_2) = x_1^3 x_2^2$ ún. Cobb-Douglas-féle függvény, amely a fogyasztó által megvásárolt termékek hasznosságát fejezi ki. Mennyit vásároljon a fogyasztó az egyes termékekből (mi legyen a fogyasztói kosár), ha maximális hasznosságot akar biztosítani, de a vásárlásra 10 pénzegységnél többet nem akar fordítani?

Megoldás

A probléma matematikai modelljét egyszerűen felírhatjuk, hisz a vásárlásra fordított pénzmennyiség $3x_1 + 4x_2$, a termékmennyiség nem lehet negatív, így a matematikai modell a következő:

$$\begin{aligned} x_1^3 x_2^2 &\rightarrow \text{max!} \\ 3x_1 + 4x_2 &\leqq 10 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

Egyik döntési változó sem lehet zérus, hisz ekkor a célfüggvény zérus lenne, a feltétel pedig megenged ennél nagyobb hasznosságot is (pl. az $\mathbf{x} = (1, 1)$ fogyasztói kosár esetén 1), így a második feltétel $x_1, x_2 > 0$ alakban írható. A szokásos jelöléssel az optimalizálási feladat alakja:

$$\begin{aligned} f(x_1, x_2) &= x_1^3 x_2^2 \rightarrow \text{max!} \\ g(x_1, x_2) &= 10 - 3x_1 - 4x_2 \geqq 0 \\ (x_1, x_2) &\in X = \{(x_1, x_2) : x_1, x_2 > 0\} \end{aligned}$$

A Lagrange függvény:

$$L(x_1, x_2, u) = x_1^3 x_2^2 + u(10 - 3x_1 - 4x_2)$$

A Lagrange függvény gadiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} 3x_1^2 x_2^2 - 3u \\ 2x_1^3 x_2 - 4u \\ 10 - 4x_2 - 3x_1 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} 6x_1 x_2^2 & 6x_1^2 x_2 & -3 \\ 6x_1^2 x_2 & 2x_1^3 & -4 \\ -3 & -4 & 0 \end{bmatrix}.$$

A KKT feladat:

$$\begin{aligned} 3x_1^2 x_2^2 - 3u &= 0 \\ 2x_1^3 x_2 - 4u &= 0 \\ 10 - 4x_2 - 3x_1 &\geqq 0 \\ u(10 - 4x_2 - 3x_1) &= 0 \\ u &\geqq 0 \\ (x_1, x_2) &\in X \end{aligned}$$

Mivel előírtuk, hogy $x_1, x_2 > 0$, így az u Lagrange szorzó csak pozitív lehet, amely az első két egyenletből azonnal látható. Az $u > 0$ esetben az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 3x_1^2 x_2^2 - 3u &= 0 \\ 2x_1^3 x_2 - 4u &= 0 \\ 10 - 4x_2 - 3x_1 &= 0 \end{aligned}$$

Egyetlen megoldás adódik, amely a következő:

$$x_1 = 2, \qquad x_2 = 1, \qquad u = 4.$$

Most meggyőződünk arról, hogy az $\mathbf{x} = (2, 1)$ KKT pont valóban maximumpont. Ehhez a $\mathbf{C}$ mátrixot kell előállítani, amely a Lagrange függvény Hesse mátrixából egyszerű behelyettesítéssel adódik.

$$\mathbf{C} = \begin{bmatrix} 12 & 24 & -3 \\ 24 & 16 & -4 \\ -3 & -4 & 0 \end{bmatrix}.$$

Az $n = 2$ és $s = 1$ miatt a főminorteszt is könnyen eredményt ad, hiszen csak a $\mathbf{C}$ mátrix determinánsát kell kiszámítani ($k = 2$), amely $\det(\mathbf{C}) = 240$. Mivel $(-1)^k \det(\mathbf{C}) > 0$, így a KKT pont szigorú lokális maximumpont, tehát a fogyasztói kosár tartalma az első termékből 2, a második termékből pedig 1 mennyiség. A maximális hasznosság 8 egység.

28. példa

Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} x_1 x_2^2 x_3 &\to \text{max!} \\ x_1 + x_2 + x_3^2 &\leqq 7 \\ x_1, x_2, x_3 &> 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, , x_3, u) = x_1 x_2^2 x_3 + u(7 - x_1 - x_2 - x_3^2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} x_2^2 x_3 - u \\ 2x_1 x_2 x_3 - u \\ x_1 x_2^2 - 2u x_3 \\ 7 - x_1 - x_2 - x_3^2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} 0 & 2x_2x_3 & x_2^2 & -1 \\ 2x_2x_3 & 2x_1x_3 & 2x_1x_2 & -1 \\ x_2^2 & 2x_1x_2 & -2u & -2x_3 \\ -1 & -1 & -2x_3 & 0 \end{bmatrix}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2^2 x_3 - u &= 0 \\ 2x_1 x_2 x_3 - u &= 0 \\ x_1 x_2^2 - 2u x_3 &= 0 \\ 7 - x_1 - x_2 - x_3^2 &\geqq 0 \\ u(7 - x_1 - x_2 - x_3^2) &= 0 \\ u &\geqq 0 \\ x_1, x_2, x_3 &> 0 \end{aligned}$$

Könnyen belátható, hogy $u = 0$ nem lehet, ha pozitív megoldásokat keresünk. Az $u > 0$ esetben az alábbi egyenletrendszer pozitív megoldásait kell megkeresni

$$\begin{aligned} x_2^2 x_3 - u &= 0 \\ 2x_1 x_2 x_3 - u &= 0 \\ x_1 x_2^2 - 2u x_3 &= 0 \\ u(7 - x_1 - x_2 - x_3^2) &= 0 \end{aligned}$$

Az egyenletrendszer megoldása $x_1 = 2, x_2 = 4, x_3 = 1, u = 16$. A KKT pont $\mathbf{x} = (2, 4, 1)$, $u = 16$ Lagrange szorzóval.

A $\mathbf{C}$ mátrix a Lagrange függvény Hesse mátrixából

$$\mathbf{C} = \begin{bmatrix} 0 & 8 & 16 & -1 \\ 8 & 4 & 16 & -1 \\ 16 & 16 & -32 & -2 \\ -1 & -1 & -2 & 0 \end{bmatrix}.$$

Az inerciához szükséges Schur komplemensek

$$\begin{bmatrix} 0 & 8 & 16 & -1 \\ 8 & 4 & 16 & -1 \\ 16 & 16 & -32 & -2 \\ -1 & -1 & -2 & 0 \end{bmatrix}, \quad \begin{bmatrix} -16 & -16 & 1 \\ -16 & -96 & 2 \\ 1 & 2 & \frac{1}{4} \end{bmatrix}, \quad \begin{bmatrix} -80 & 1 \\ 1 & -\frac{3}{16} \end{bmatrix}, \quad \begin{bmatrix} -\frac{14}{80} \end{bmatrix}.$$

A pivotelemválasztás a főátlóban rendre: $4, -16, -80, -\frac{14}{80}$, így $Iner(\mathbf{C}) = (3, 0, 1)$, tehát az aktualizált Lagrange függvény feltételesen negatív definit, így az $\mathbf{x} = (2, 4, 1)$ KKT pont maximumpont.

29. Példa

Oldjuk meg az alábbi optimalizálási feladatot!

$$\begin{aligned} 2x_1 - x_2 &\rightarrow \min! \\ x_1^2 + x_2^2 &\leqq 1 \\ (x_1-3)^2 + x_2^2 &\leqq 4 \end{aligned}$$

Megoldás
A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u_1, u_2) = 2x_1 - x_2 + u_1(x_1^2 + x_2^2 - 1) + u_2((x_1-3)^2 + x_2^2 - 4)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2 + 2u_1x_1 + 2u_2(x_1-3) \\ -1 + 2u_1x_2 + 2u_2x_2 \\ x_1^2 + x_2^2 - 1 \\ x_2^2 + (x_1-3)^2 - 4 \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2u_1 + 2u_2 & 0 & 2x_1 & 2x_1 - 6 \\ 0 & 2u_1 + 2u_2 & 2x_2 & 2x_2 \\ 2x_1 & 2x_2 & 0 & 0 \\ 2x_1 - 6 & 2x_2 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 2 + 2u_1x_1 + 2u_2(x_1-3) &= 0 \\ -1 + 2u_1x_2 + 2u_2x_2 &= 0 \\ x_1^2 + x_2^2 - 1 &\leqq 0 \\ x_2^2 + (x_1-3)^2 - 4 &\leqq 0 \\ u_1(x_1^2 + x_2^2 - 1) &= 0 \\ u_2((x_1-3)^2 + x_2^2 - 4) &= 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását a u_1, u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1 = u_2 = 0$ eset
Az első két egyenletből azonnal ellentmondás adódik.

ii) $u_1 > 0, u_2 = 0$ eset
Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2u_1x_1 + 2 &= 0 \\ 2u_1x_2 - 1 &= 0 \\ x_1^2 + x_2^2 - 1 &= 0 \end{aligned}$$

Két megoldás is adódik, de az egyikben $u_1 < 0$, a másik megoldás: $x_1 = -\frac{2}{\sqrt{5}}, x_2 = \frac{1}{\sqrt{5}}$. Ez a megoldás azonban nem teljesíti a második primál feltételt.

iii) $u_1 = 0, u_2 > 0$ eset
Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2u_2(x_1-3) + 2 &= 0 \\ 2u_2x_2 - 1 &= 0 \\ x_2^2 + (x_1-3)^2 - 4 &= 0 \end{aligned}$$

Szintén két megoldás adódik, a megoldás, amelynél $u_2 > 0$: $x_1 = 3 - \frac{4}{\sqrt{5}}, x_2 = \frac{2}{\sqrt{5}}$. Ez a megoldás pedig az első primál feltételt nem teljesíti.

iv) $u_1 > 0, u_2 > 0$ eset
Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2 + 2u_1x_1 + 2u_2(x_1 - 3) &= 0 \\ -1 + 2u_1x_2 + 2u_2x_2 &= 0 \\ x_1^2 + x_2^2 - 1 &= 0 \\ x_2^2 + (x_1 - 3)^2 - 4 &= 0 \end{aligned}$$

Az egyenletrendszernek nincs megoldása.

Összefoglalva tehát nem kaptunk KKT pontot. Az (x_1, x_2) síkon ábrázolva a tartományt azt látjuk, hogy az első feltételt az origó középpontú 1 sugarú kör, a második feltételt pedig a $(3, 0)$ középpontú 2 sugarú kör pontjai és belseje elégíti ki. A két körlap metszete egyetlen pont, az $(1, 0)$ pont. Ez pedig biztosan optimális megoldás. Tehát a KKT ponton keresztül nem tudtuk meghatározni az optimális megoldást. Ez nem véletlen, hiszen a tartomány nem Slater reguláris, a feltételi függvények ugyan konvexek, de nincs ugyanis olyan pont, amelyben a $\leqq$ feltételek szigorú egyenlőtlenséggel ($<$) teljesülnének, azaz nincs belső pont.

Feladat:
A Fritz-John ponton keresztül határozza meg az optimális megoldást!

30. Példa

Tekintsünk n befektetési eszközt, amelyek várható hozama $r_1, r_2, \ldots, r_n$. Az egyes befektetési eszközök hozamai között ismert a $\mathbf{C} = [c_{ij}]$ kovarianciamátrix. A befektetési eszközöknek egy olyan portfólióját kívánjuk összeállítani, amely a befektetőnek maximális hozamot biztosít a portfólió egy megadott szintű varianciája (σ_p^2) mellett. Ezt a modellt Markowitz-féle portfóliókiválasztási modellnek nevezik. Jelölje $x_1, x_2, \ldots, x_n$ a befektetési eszközök arányát a portfólióban. A modell matematikai formája az alábbi:

$$\begin{aligned} \sum_{i=1}^{n} r_i x_i &\rightarrow \text{max!} \\ \sum_{i=1}^{n} \sum_{j=1}^{n} x_i c_{ij} x_j &= \sigma_p^2 \\ x_1 + x_2 + \ldots + x_n &= 1 \\ x_1, x_2, \ldots, x_n &\geqq 0 \end{aligned}$$

Oldjuk meg a feladatot az alábbi adatokkal. Legyen három befektetési eszközünk. Az egyes hozamok (r_i) legyenek rendre: $30, 40, 50$, a portfólió előírt varianciája legyen $\sigma_p^2 = 0.26$, a kovarianciamátrix pedig legyen

$$\mathbf{C} = \begin{bmatrix} 0.8 & -0.2 & 0.1 \\ -0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}.$$

Az optimalizálási feladat tehát az alábbi:

$$\begin{aligned} 30x_1 + 40x_2 + 50x_3 &\to \max! \\ 0.8x_1^2 - 0.4x_1x_2 + 0.2x_1x_3 + 0.5x_2^2 + 0.6x_2x_3 + 0.4x_3^2 &= 0.26 \\ x_1 + x_2 + x_3 &= 1 \\ x_1, x_2, x_3 &\geqq 0 \end{aligned}$$

Megoldás

A feladathoz tartozó Lagrange függvény:

$$\begin{aligned} L(x_1, x_2, x_3, v_1, v_2) = {} & 30x_1 + 40x_2 + 50x_3 + \\ & +v_1(0.8x_1^2 - 0.4x_1x_2 + 0.2x_1x_3 + 0.5x_2^2 + 0.6x_2x_3 + 0.4x_3^2 - 0.26) + \\ & +v_2(x_1 + x_2 + x_3 - 1) \end{aligned}$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa rendre a következő:

$$\begin{bmatrix} 30 + v_1(1.6x_1 - 0.4x_2 + 0.2x_3) + v_2 \\ 40 + v_1(-0.4x_1 + x_2 + 0.6x_3) + v_2 \\ 50 + v_1(0.2x_1 + 0.6x_2 + 0.8x_3) + v_2 \\ 0.8x_1^2 - 0.4x_1x_2 + 0.2x_1x_3 + 0.5x_2^2 + 0.6x_2x_3 + 0.4x_3^2 - 0.26 \\ x_1 + x_2 + x_3 - 1 \end{bmatrix}$$

$$\begin{bmatrix} 1.6v_1 & -0.4v_1 & 0.2v_1 & 1.6x_1 - 0.4x_2 + 0.2x_3 & 1 \\ -0.4v_1 & v_1 & 0.6v_1 & -0.4x_1 + x_2 + 0.6x_3 & 1 \\ 0.2v_1 & 0.6v_1 & 0.8v_1 & 0.2x_1 + 0.6x_2 + 0.8x_3 & 1 \\ 1.6x_1 - 0.4x_2 + 0.2x_3 & -0.4x_1 + x_2 + 0.6x_3 & 0.2x_1 + 0.6x_2 + 0.8x_3 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 30 + v_1(1.6x_1 - 0.4x_2 + 0.2x_3) + v_2 &\leqq 0 \\ 40 + v_1(-0.4x_1 + x_2 + 0.6x_3) + v_2 &\leqq 0 \\ 50 + v_1(0.2x_1 + 0.6x_2 + 0.8x_3) + v_2 &\leqq 0 \\ 0.8x_1^2 - 0.4x_1x_2 + 0.2x_1x_3 + 0.5x_2^2 + 0.6x_2x_3 + 0.4x_3^2 - 0.26 &= 0 \\ x_1 + x_2 + x_3 - 1 &= 0 \\ [30 + v_1(1.6x_1 - 0.4x_2 + 0.2x_3) + v_2]\, x_1 &= 0 \\ [40 + v_1(-0.4x_1 + x_2 + 0.6x_3) + v_2]\, x_2 &= 0 \\ [50 + v_1(0.2x_1 + 0.6x_2 + 0.8x_3) + v_2]\, x_3 &= 0 \\ x_1, x_2, x_3 &\geqq 0 \end{aligned}$$

Meglehetősen bonyolult rendszer adódott. Az egyenleteket tekintve az alábbi öt egyenletből álló egyenletrendszert kell megoldani az öt ismeretlenre, a megoldandó egyenletrendszer:

$$\begin{aligned} 0.8x_1^2 - 0.4x_1x_2 + 0.2x_1x_3 + 0.5x_2^2 + 0.6x_2x_3 + 0.4x_3^2 - 0.26 &= 0 \\ x_1 + x_2 + x_3 - 1 &= 0 \\ [30 + v_1(1.6x_1 - 0.4x_2 + 0.2x_3) + v_2]\, x_1 &= 0 \\ [40 + v_1(-0.4x_1 + x_2 + 0.6x_3) + v_2]\, x_2 &= 0 \\ [50 + v_1(0.2x_1 + 0.6x_2 + 0.8x_3) + v_2]\, x_3 &= 0 \end{aligned}$$

Négy megoldás adódik, amelyek az alábbiak:

i	x_1	x_2	x_3	v_1	v_2
1	0.58	0.42	0	17.46	−43.27
2	0.24	0.76	0	−17.46	−28.48
3	0.62	0.96	−0.58	77.15	−68.02
4	0.22	0.27	0.51	−77.15	−2.74

Az első két megoldás nem teljesíti a KKT feltételek közül az első három egyenlőtlenséget, a harmadik megoldásban pedig x_3 negatív. A negyedik megoldás minden további KKT feltételt teljesít, így az $x_1 = 0.22, x_2 = 0.27, x_3 = 0.51$ KKT pont, a Lagrange szorzók pedig: $v_1 = -77.15, v_2 = -2.74$. Tehát a portfólióban az egyes befektetési eszközök aránya rendre 22%, 27%, 51%. Még meg kell győződni arról, hogy ez valóban optimális megoldás. Mivel minden döntési változó pozitív, így a Lagrange függvény Hesse mátrixa lesz a szegélyezett $\mathbf{C}$ mátrix. Nem kell tehát kiegészíteni egységvektorokkal a $\boldsymbol{\nabla}^2 L$ mátrixot.

$$\mathbf{C} = \begin{bmatrix} -123.44 & 30.86 & -15.43 & 0.35 & 1 \\ 30.86 & -77.15 & -46.29 & 0.48 & 1 \\ -15.43 & -46.29 & -61.72 & 0.61 & 1 \\ 0.35 & 0.48 & 0.61 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Mivel $Iner(\mathbf{C}) = (3, 0, 2)$, tehát az aktualizált Lagrange függvény feltételesen negatív definit, így a fentebb adódott KKT pont maximumpont.

31. Példa

Egy vállalat két terméket állít elő két erőforrás igénybevételével. Az első termék egységnyi mennyiségű gyártásához az első erőforrásból 1, a második erőforrásból 0.2 mennyiséget használ fel. A második termék egységnyi mennyiségű gyártásához az első erőforrásból 0.5, a második erőforrásból is 0.5 mennyiséget használ fel. Az első erőforrás egységára $p_1^e = 375$, a másodiké $p_2^e = 750$ pénzegység. Az első termék eladási egységára $p_1^t = 2000$, a másodiké $p_2^t = 3000$ pénzegység. Egy adott időszakban az első erőforrásból legfeljebb 980, a másodikból pedig legfeljebb 220 mennyiség használható fel. Tegyük fel, hogy a megtermelt termékeket el is tudjuk adni. Határozzuk meg az adott időszakban egyes termékekből termelhető mennyiséget, ha maximális nyereségre törekszik a vállalat!

Megoldás

Jelölje x_1, x_2 az egyes termékekből gyártott mennyiségeket. A nyereség a termékek eladásából származó árbevétel és a felhasznált erőforrásokra fordított kiadás különbsége.

A termékek eladásából származó árbevétel: $p_1^t x_1 + p_2^t x_2$.

Az x_1, x_2 mennyiségű termék gyártásához az első erőforrásból $x_1 + 0.5x_2$, a második erőforrásból $0.2x_1 + 0.5x_2$ mennyiséget használnak fel a technológiai adatok szerint.

A felhasznált erőforrásokra fordított kiadás: $p_1^e(x_1 + 0.5x_2) + p_2^e(0.2x_1 + 0.5x_2)$.

A nyereség a termékek eladásából származó árbevétel és a felhasznált erőforrásokra fordított kiadás különbsége: $[p_1^t x_1 + p_2^t x_2] - [(p_1^e(x_1 + 0.5x_2) + p_2^e(0.2x_1 + 0.5x_2)]$, amely az áradatokkal: $1475x_1 + 2437.5x_2$.

A modell matematikai formája az alábbi:

$$\begin{aligned} 1475x_1 + 2437.5x_2 &\rightarrow \text{max!} \\ x_1 + 0.5x_2 &\leqq 980 \\ 0.2x_1 + 0.5x_2 &\leqq 220 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

Az optimális megoldás: $x_1 = 950, x_2 = 60$, a maximális nyereség 1547500, az árbevétel: 2080000, a kiadás: 532500.

Feladat:
Oldja meg a fenti lineáris programozási (LP) feladatot szimplex módszerrel és a KKT pont meghatározásán keresztül! Hasonlítsa össze a megoldások időigényét!

Módosítsuk a feladatot úgy, hogy a termékek eladási egységára függ a gyártott (eladott) mennyiségtől, mégpedig az alábbi módon: az első terméké $p_1^t = 2000 - 0.5x_1 - 0.15x_2$, a második terméké $p_2^t = 3000 - 0.2x_1 - 1.5x_2$. Az erőforrások egységára pedig az adott erőforrásból felhasznált mennyiségtől függ az alábbiak szerint: $p_1^e = 375 - 0.05(x_1 + 0.5x_2)$, $p_2^e = 750 - 0.1(0.2x_1 + 0.5x_2)$. Ekkor a célfüggvény a következőképpen módosul:

$$\begin{gathered} \left[p_1^t x_1 + p_2^t x_2\right] - \left[(p_1^e(x_1 + 0.5x_2) + p_2^e(0.2x_1 + 0.5x_2)\right] = \\ = \left[(2000 - 0.5x_1 - 0.15x_2)x_1 + (3000 - 0.2x_1 - 1.5x_2)x_2\right] - \\ - \{[(375 - 0.05(x_1 + 0.5x_2))(x_1 + 0.5x_2)] + [(750 - 0.1(0.2x_1 + 0.5x_2))(0.2x_1 + 0.5x_2)]\} \end{gathered}$$

A szorzásokat elvégezve a célfüggvény egyszerűbb formája:

$$1475x_1 + 2437.5x_2 - 0.446x_1^2 - 0.28x_1x_2 - 1.4625x_2^2$$

Összefoglalva a megoldandó optimalizálási feladat a következő:

$$\begin{aligned} 1475x_1 + 2437.5x_2 - 0.446x_1^2 - 0.28x_1x_2 - 1.4625x_2^2 &\rightarrow \text{max!} \\ x_1 + 0.5x_2 &\leqq 980 \\ 0.2x_1 + 0.5x_2 &\leqq 220 \\ x_1, x_2 &\geqq 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L = 1475x_1 + 2437.5x_2 - 0.446x_1^2 - 0.28x_1x_2 - 1.4625x_2^2 + u_1(980 - x_1 - 0.5x_2) + u_2(220 - 0.2x_1 - 0.5x_2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\begin{aligned} \boldsymbol{\nabla} L &= \begin{bmatrix} 1475 - 0.892x_1 - 0.28x_2 - u_1 - 0.2u_2 \\ 2437.5 - 0.28x_1 - 2.925x_2 - 0.5u_1 - 0.5u_2 \\ 980 - x_1 - 0.5x_2 \\ 220 - 0.2x_1 - 0.5x_2 \end{bmatrix}, \\ \boldsymbol{\nabla}^2 L &= \begin{bmatrix} -0.892 & -0.28 & -1 & -0.2 \\ -0.28 & -2.925 & -0.5 & -0.5 \\ -1 & -0.5 & 0 & 0 \\ -0.2 & -0.5 & 0 & 0 \end{bmatrix}. \end{aligned}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned}
1475 - 0.892x_1 - 0.28x_2 - u_1 - 0.2u_2 &\leqq 0\\
2437.5 - 0.28x_1 - 2.925x_2 - 0.5u_1 - 0.5u_2 &\leqq 0\\
980 - x_1 - 0.5x_2 &\geqq 0\\
220 - 0.2x_1 - 0.5x_2 &\geqq 0\\
u_1(980 - x_1 - 0.5x_2) &= 0\\
u_2(220 - 0.2x_1 - 0.5x_2) &= 0\\
(1475 - 0.892\,x_1 - 0.28x_2 - u_1 - 0.2u_2)\,x_1 &= 0\\
(2437.5 - 0.28x_1 - 2.925x_2 - 0.5u_1 - 0.5u_2)\,x_2 &= 0\\
x_1, x_2 &\geqq 0\\
u_1, u_2 &\geqq 0
\end{aligned}$$

Az alábbi négy esetre bontjuk a megoldást:

1. eset: $u_1 = u_2 = 0$

A megoldandó egyenletrendszer:

$$\begin{aligned}
(1475 - 0.892x_1 - 0.28x_2)\,x_1 &= 0\\
(2437.5 - 0.28x_1 - 2.925x_2)\,x_2 &= 0
\end{aligned}$$

Az alábbi négy megoldás adódott:

i	x_1	x_2
1	1435.1	695.95
2	0	0
3	1653.6	0
4	0	833.33

A megoldások közül egyik sem lehet KKT pont, mert nem teljesedik az összes többi KKT feltétel.

2. eset: $u_1 > 0, u_2 = 0$

A megoldandó egyenletrendszer:

$$\begin{aligned}
980 - x_1 - 0.5x_2 &= 0\\
(1475 - 0.892x_1 - 0.28x_2 - u_1)\,x_1 &= 0\\
(2437.5 - 0.28x_1 - 2.925x_2 - 0.5u_1)x_2 &= 0
\end{aligned}$$

Az alábbi három megoldás adódott:

i	x_1	x_2	u_1
1	655.26	649.47	708.65
2	980	0	600.84
3	0	1960	-6591

A megoldások közül egyik sem lehet KKT pont, mert nem teljesedik az összes többi KKT feltétel.

3. eset: $u_1 = 0, u_2 > 0$

A megoldandó egyenletrendszer:

$$\begin{aligned} 220 - 0.2x_1 - 0.5x_2 &= 0 \\ (-0.892x_1 - 0.28x_2 + 1475 - 0.2u_2)\, x_1 &= 0 \\ (-0.28x_1 - 2.925x_2 + 2437.5 - 0.5u_2)\, x_2 &= 0 \end{aligned}$$

Az alábbi három megoldás adódott:

i	x_1	x_2	u_2
1	784.86	126.06	3698
2	1100	0	2469
3	0	440	2301

A megoldások közül csak az első teljesíti az összes többi KKT feltételt, így a KKT pont: $x_1 = 784.86, x_2 = 126.06$, az $u_1 = 0, u_2 = 3698$ Lagrange szorzókkal. Az első feltétel inaktív, a második pedig aktív.

4. eset: $u_1 > 0, u_2 > 0$
A megoldandó egyenletrendszer:

$$\begin{aligned} 980 - x_1 - 0.5x_2 &= 0 \\ 220 - 0.2x_1 - 0.5x_2 &= 0 \\ (1475 - 0.892x_1 - 0.28x_2 - u_1 - 0.2u_2)\, x_1 &= 0 \\ (2437.5 - 0.28x_1 - 2.925x_2 - 0.5u_1 - 0.5u_2)\, x_2 &= 0 \end{aligned}$$

Egyetlen megoldás adódott: $x_1 = 950.0, x_2 = 60.0, u_1 = -234.5, u_2 = 4226.5$, de ebből az u_1 negatívitása miatt nem kapunk KKT pontot.

Összefoglalva a négy esetet, az optimalizálási feladat KKT pontja: $x_1 = 784.86, x_2 = 126.06$, az $u_1 = 0, u_2 = 3698$ Lagrange szorzókkal és az első feltétel inaktív, a második pedig aktív.

Az elégségesség vizsgálatához szükséges $\mathbf{C}$ mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixból úgy kapjuk, hogy a 3. sort és a 3. oszlopot töröljük, mivel az első feltétel inaktív. A

$$\mathbf{C} = \begin{bmatrix} -0.892 & -0.28 & -0.2 \\ -0.28 & -2.925 & -0.5 \\ -0.2 & -0.5 & 0 \end{bmatrix}$$

mátrix inerciája: $Iner(\mathbf{C}) = (2, 0, 1)$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen negatív definit, így a KKT pont szigorú maximumpont. Megjegyezzük, hogy az aktualizált Lagrange függvény Hesse mátrixa negatív definit, így feltételesen is az.

Ha az árak a mennyiségtől is függnek, akkor az optimális termelés (kerekítve) az első termékből 785, a második termékből pedig 126 mennyiség. A maximális nyereség 1139250, az árbevétel 1581455, a kiadás 442205 pénzegység.

Megjegyezzük, hogy az olyan optimalizálási feladatokat, amelyek feltételi függvényei lineárisak, célfüggvénye pedig kvadratikus függvény kvadratikus programozási (QP) feladatnak nevezzük. A fenti feladat is ilyen típusú. Ennek a feladattípusnak a megoldására hasonlóan a lineáris programozáshoz számos megoldási módszert dolgoztak ki.

32. Példa

Adott az $x_1 + x_2 = 2$ egyenes és az origó középű 2 sugarú kör. Az egyenes a körlapot két tartományra bontja, tekintsük azt a tartományt, amely az origót tartalmazza. Határozzuk meg a tartománynak azt a pontját, amelynél a $\ln \frac{1}{x_1} - x_2$ függvény értéke a legkisebb.

Megoldás

Könnyen eldönthetjük, hogy a tartomány $x_1 + x_2 \leqq 2$ és $x_1^2 + x_2^2 \leqq 4$ feltételekkel írható fel. A célfüggvény értelmezési tartománya megköveteli az $x_1 > 0$ kikötést is, de ez egy nyílt halmazt jelent. Az optimalizálási feladat:

$$\begin{aligned} \ln \frac{1}{x_1} - x_2 &\rightarrow \text{min!} \\ x_1 + x_2 - 2 &\leqq 0 \\ x_1^2 + x_2^2 - 4 &\leqq 0 \\ x_1 &> 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u_1, u_2) = -\ln x_1 - x_2 + u_1(x_1 + x_2 - 2) + u_2(x_1^2 + x_2^2 - 4)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} -\frac{1}{x_1} + u_1 + 2u_2x_1 \\ -1 + u_1 + 2u_2x_2 \\ x_1 + x_2 - 2 \\ x_1^2 + x_2^2 - 4 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} \frac{1}{x_1^2} + 2u_2 & 0 & 1 & 2x_1 \\ 0 & 2u_2 & 1 & 2x_2 \\ 1 & 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} -\tfrac{1}{x_1} + u_1 + 2u_2x_1 &= 0 \\ -1 + u_1 + 2u_2x_2 &= 0 \\ x_1 + x_2 - 2 &\leqq 0 \\ x_1^2 + x_2^2 - 4 &\leqq 0 \\ u_1(x_1 + x_2 - 2) &= 0 \\ u_3(x_1^2 + x_2^2 - 4) &= 0 \\ x_1 &> 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

Négy esetet különböztetünk meg a Lagrange szorzók előjelétől függően.

1. eset: $u_1 = u_2 = 0$

Az első egyenlet miatt ez a megoldás nem jöhet szóba.

2. eset: $u_1 > 0, u_2 = 0$

A megoldandó egyenletrendszer:

$$\begin{aligned} -\tfrac{1}{x_1} + u_1 &= 0 \\ -1 + u_1 &= 0 \\ x_1 + x_2 - 2 &= 0 \end{aligned}$$

Az alábbi megoldás adódik: $x_1 = 1, x_2 = 1, u_1 = 1$. Ez KKT pont, mert a többi feltételt is teljesíti.

3. eset: $u_1 = 0, u_2 > 0$
A megoldandó egyenletrendszer:

$$\begin{aligned}
-\tfrac{1}{x_1} + 2u_2x_1 &= 0\\
-1 + 2u_2x_2 &= 0\\
x_1^2 + x_2^2 - 4 &= 0
\end{aligned}$$

Az első egyenletből $x_1^2 = \frac{1}{2u_2}$, a másodikból $x_2 = \frac{1}{2u_2}$, azaz $x_1^2 = x_2$. A harmadikból x_2 re egy másodfokú egyenlet adódik, amelynek megoldása $x_2 = \frac{\sqrt{17}-1}{2} \approx 1.56$. Ebből egyszerűen adódik, hogy $u_2 \approx 0.32, \quad x_1 \approx 1.25$. A megoldás nem lehet KKT pont, mert nem teljesedik az $x_1 + x_2 \leqq 2$ feltétel.

4. eset: $u_1 > 0, u_2 > 0$
A megoldandó egyenletrendszer:

$$\begin{aligned}
-\tfrac{1}{x_1} + u_1 + 2u_2x_1 &= 0\\
-1 + u_1 + 2u_2x_2 &= 0\\
x_1 + x_2 - 2 &= 0\\
x_1^2 + x_2^2 - 4 &= 0
\end{aligned}$$

Egyetlen megoldás adódott: $x_1 = 2, x_2 = 0, u_1 = 1, u_2 = -\frac{1}{8}$. Mivel $u_2 < 0$, így nem lehet KKT pont.

Összefoglalva a négy esetet, az optimalizálási feladatnak egyetlen KKT pontja van: $x_1 = 1, x_2 = 1$, az $u_1 = 1, u_2 = 0$ Lagrange szorzókkal és az első feltétel aktív, a második pedig inaktív.

Az elégségesség vizsgálatához szükséges $\mathbf{C}$ mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixból úgy kapjuk, hogy a 4. sort és a 4. oszlopot töröljük, mivel a második feltétel inaktív. A $\mathbf{C}$ mátrix

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

A mátrix inerciája: $Iner(\mathbf{C}) = (2, 0, 1)$, így az aktualizált Lagrange függvény Hesse mátrixa feltételesen pozitív definit, tehát a KKT pont szigorú minimumpont. A főminor teszt is könnyen eredményt ad, mert csak a $\mathbf{C}$ mátrix determinánsát kell meghatározni, amely $\det(\mathbf{C}) = -1$, és igaz, hogy $(-1)^1 \det(\mathbf{C}) > 0$, így az $\mathbf{x} = (1, 1)$ pont valóban szigorú minimumpont.

Megjegyezzük, hogy a példában az elégségességi vizsgálat helyett hivatkozhattunk volna arra a tételre, amely szerint, ha a minimalizálási feladatban minden függvény konvex, akkor a szükséges feltétel egyben elégséges is, tehát a KKT pont optimális megoldást szolgáltat. A célfüggvény konvex, mert a $-\ln x_1$ függvény konvex és ahhoz hozzáadva vagy levonva lineáris függvényt konvex függvényt kapunk. Az első feltételi függvény lineáris, tehát konvex, a második pedig szintén konvex.

33. Példa

Legyen adott az $\mathbb{R}^3$ térben egy gömb, amelynek középpontja az origóban van, sugara pedig $\sqrt{2}$ és adott továbbá az $x_3 = x_1x_2$ nyeregfelület. Tekintsünk két térbeli tartományt: az egyik a gömbön lévő és a belsejében lévő pontok összessége, a másik pedig a nyeregfelületen és a nyeregfelület fölött elhelyezkedő pontok összessége. Határozzuk meg a két térbeli tartomány metszettartományának azon pontját, amely legmesszebb van a $P(2,2,0)$ ponttól!

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. Az egyik tartomány pontjait a $x_1^2+x_2^2+x_3^2 \leqq 2$, a másik tartomány pontjait pedig az $x_3 \geqq x_1x_2$ egyenlőtlenségek írják le. Célfüggvényként a távolság helyett elegendő a távolság négyzetét használni. Az optimalizálási feladat az alábbi:

$$\begin{aligned}(x_1-2)^2+(x_2-2)^2+x_3^2 &\rightarrow \max!\\ 2-x_1^2-x_2^2-x_3^2 &\geqq 0\\ x_3-x_1x_2 &\geqq 0\end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1,x_2,x_3,u_1,u_2)=(x_1-2)^2+(x_2-2)^2+x_3^2+u_1(2-x_1^2-x_2^2-x_3^2)+u_2(x_3-x_1x_2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L=\begin{bmatrix} 2x_1-4-2u_1x_1-u_2x_2\\ 2x_2-4-2u_1x_2-u_2x_1\\ 2x_3-2u_1x_3+u_2\\ 2-x_1^2-x_2^2-x_3^2\\ x_3-x_1x_2\end{bmatrix},\quad \nabla^2 L=\begin{bmatrix} 2-2u_1 & -u_2 & 0 & -2x_1 & -x_2\\ -u_2 & 2-2u_1 & 0 & -2x_2 & -x_1\\ 0 & 0 & 2-2u_1 & -2x_3 & 1\\ -2x_1 & -2x_2 & -2x_3 & 0 & 0\\ -x_2 & -x_1 & 1 & 0 & 0\end{bmatrix}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned}2x_1-4-2u_1x_1-u_2x_2&=0\\ 2x_2-4-2u_1x_2-u_2x_1&=0\\ 2x_3-2u_1x_3+u_2&=0\\ 2-x_1^2-x_2^2-x_3^2&\geqq 0\\ x_3-x_1x_2&\geqq 0\\ u_1(2-x_1^2-x_2^2-x_3^2)&=0\\ u_2(x_3-x_1x_2)&=0\\ u_1,u_2&\geqq 0\end{aligned}$$

A fenti rendszer megoldását a u_1,u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1=u_2=0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x}=(2,2,0)$ pont. Ez azonban nem teljesíti a primál feltételt. KKT pont nem adódik.

ii) $u_1=0, u_2>0$ eset

Az első három egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned}2x_1-4-u_2x_2&=0\\ 2x_2-4-u_2x_1&=0\\ 2x_3+u_2&=0\\ x_3-x_1x_2&=0\end{aligned}$$

A harmadik egyenletből $u_2 > 0$ miatt következik, hogy $x_3 < 0$. A negyedik egyenletből következik, hogy x_1, x_2 ellenkező előjelű, nemzérus számok. Ezeket figyelembe véve az első kettőből az következik, hogy $u_2 < 0$, ami nem lehet. KKT pont nem adódik.

iii) $u_1 > 0, u_2 = 0$ eset
Az első három egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 4 - 2u_1x_1 &= 0 \\ 2x_2 - 4 - 2u_1x_2 &= 0 \\ 2x_3 - 2u_1x_3 &= 0 \\ 2 - x_1^2 - x_2^2 - x_3^2 &= 0 \end{aligned}$$

A harmadik egyenletből $x_3 = 0$ vagy $u_1 = 1$ lehet. Az $u_1 = 1$ esetén az első két egyenlet nem teljesülhet. Az $x_3 = 0$ esetén az első két egyenletből kifejezve x_1, x_2-t és behelyettesítve a negyedikbe, kapjuk, hogy $u_1 = 3$, ebből pedig $x_1 = -1, x_2 = -1$. Ez a megoldás azonban nem teljesíti a második feltételt. KKT pont nem adódik.

iv) $u_1 > 0, u_2 > 0$ eset
Az első három egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 4 - 2u_1x_1 - u_2x_2 &= 0 \\ 2x_2 - 4 - 2u_1x_2 - u_2x_1 &= 0 \\ 2x_3 - 2u_1x_3 + u_2 &= 0 \\ 2 - x_1^2 - x_2^2 - x_3^2 &= 0 \\ x_3 - x_1x_2 &= 0 \end{aligned}$$

A megoldás: $x_1 = -0.855,\ x_2 = -0.855,\ x_3 = 0.732,\ u_1 = 2.349,\ u_2 = 1.976$.
Összefoglalva egy KKT pont adódott, az $\mathbf{x} = (-0.855,\ -0.855,\ 0.732)$ vektor az $\mathbf{u} = (2.349,\ 1.976)$ Lagrange szorzókkal úgy, hogy mindkét feltétel aktív.

Arról, hogy ez valóban maximumpont, az optimum elégséges feltétele segítségével győződünk meg. A szükséges $\mathbf{C}$ szegélyezett mátrixot a $\boldsymbol{\nabla}^2 L$ mátrixból behelyettesítéssel kapjuk:

$$\mathbf{C} = \begin{bmatrix} -2.7 & -1.976 & 0 & 1.71 & 0.855 \\ -1.976 & -2.7 & 0 & 1.71 & 0.855 \\ 0 & 0 & -2.7 & 1.46 & 1 \\ 1.71 & 1.71 & 1.46 & 0 & 0 \\ 0.855 & 0.855 & 1 & 0 & 0 \end{bmatrix}.$$

Az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 3×3-as főminormátrixa) negatív definit, így a feltételes definitségre nincs szökség. A KKT pont szigorú maximumpont.

Megjegyzés:
Látható, hogy ennél a feladatnál bonyolult egyenletrendszereket kellett megoldani. A feltételes optimalizálási feladatok negoldásának megkeresésére különböző numerikus módszereket dolgoztak ki. Ezeket a módszereket A feltételes optimalizálás algoritmusai c. tananyagban ismertetjük.

Feladat:

Ha az olvasó szereti a kihívásokat, határozzuk meg a két térbeli tartomány metszettartományának azon pontját, amely legközelebb van a $P(2,2,0)$ ponthoz!

34. Példa

Adott az $x_2 = x_1^2$ parabola és az $x_2 = 2x_1 + 3$ egyenes. Tekintsük a parabola és az egyenes által határolt kompakt tartományt. Határozzuk meg a tartomány azon pontját, amelyben az $x_1 x_2$ mennyiség a legnagyobb!

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. A tartomány pontjait az $x_2 \geqq x_1^2$ és az $x_2 \leqq 2x_1 + 3$ egyenlőtlenségek írják le. Az optimalizálási feladat az alábbi:

$$\begin{aligned} x_1 x_2 &\rightarrow \max! \\ x_2 - x_1^2 &\geqq 0 \\ 2x_1 - x_2 + 3 &\geqq 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, u_1, u_2) = x_1 x_2 + u_1(x_2 - x_1^2) + u_2(2x_1 - x_2 + 3)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} x_2 - 2u_1 x_1 + 2u_2 \\ x_1 + u_1 - u_2 \\ x_2 - x_1^2 \\ 2x_1 - x_2 + 3 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} -2u_1 & 1 & -2x_1 & 2 \\ 1 & 0 & 1 & -1 \\ -2x_1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 - 2u_1 x_1 + 2u_2 &= 0 \\ x_1 + u_1 - u_2 &= 0 \\ x_2 - x_1^2 &\geqq 0 \\ 2x_1 - x_2 + 3 &\geqq 0 \\ u_1(x_2 - x_1^2) &= 0 \\ u_2(2x_1 - x_2 + 3) &= 0 \\ u_1, u_2 &\geqq 0 \end{aligned}$$

A fenti rendszer megoldását a u_1, u_2 Lagrange szorzók értékeitől függően 4 esetre bontjuk szét, amelyek a következők.

i) $u_1 = u_2 = 0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x} = (0,0)$ pont. Ez teljesíti az összes KKT feltételt is, így KKT pont.

ii) $u_1 = 0, u_2 > 0$ eset

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 + 2u_2 &= 0 \\ x_1 - u_2 &= 0 \\ 2x_1 - x_2 + 3 &= 0 \end{aligned}$$

A megoldás: $x_1 = -\frac{3}{4}, x_2 = \frac{3}{2}, u_2 = -\frac{3}{4}$. Az u_2 negatívitása miatt nem kaptunk KKT pontot.

iii) $u_1 > 0, u_2 = 0$ eset

Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 - 2u_1x_1 &= 0 \\ x_1 + u_1 &= 0 \\ x_2 - x_1^2 &= 0 \end{aligned}$$

A megoldás: $x_1 = 0, x_2 = 0, u_1 = 0$. Ez azonos az i) esetben kapott $\mathbf{x} = (0,0)$ KKT ponttal.

iv) $u_1 > 0, u_2 > 0$ eset

Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} x_2 - 2u_1x_1 + 2u_2 &= 0 \\ x_1 + u_1 - u_2 &= 0 \\ x_2 - x_1^2 &= 0 \\ 2x_1 - x_2 + 3 &= 0 \end{aligned}$$

Az utolsó két egyenletből két megoldás adódik: $x_1 = 3, x_2 = 9$, ill. $x_1 = -1, x_2 = 1$. A megoldásokhoz tartozó Lagrange szorzókat behelyettesítés után az első két egyenletből határozhatjuk meg, amelyek az alábbiak: $u_1 = \frac{15}{4}, u_2 = \frac{27}{4}$, ill. $u_1 = \frac{1}{4}, u_2 = -\frac{3}{4}$. A második nem lehet KKT pont $u_2 < 0$ miatt.

A kapott KKT pont: $x_1 = 3, x_2 = 9$ az $u_1 = \frac{15}{4}, u_2 = \frac{27}{4}$ Lagrange szorzókkal.

Most a két KKT pontról kell eldönteni, hogy maximumpontok-e.

Az $\mathbf{x} = (0,0)$ KKT pont ($\mathbf{u} = (0,0)$ Lagrange szorzókkal) vizsgálata:

Az első feltétel aktív, de $u_1 = 0$, a második feltétel nem aktív. Az erős elégségességi tétel nem alkalmazható, a gyenge tétel pedig nem dönt az optimalitásról. Viszont könnyen belátható, hogy az $\mathbf{x} = (0,0)$ pont környezetében van a zérusnál nagyobb és kisebb függvényérték, így nem lehet az origó lokális maximumpont. Az $\mathbf{x} = (0,0)$ környezetbeli pont $\mathbf{x} = (\lambda d_1, \lambda d_2)$, a célfüggvényérték $\lambda^2 d_1 d_2$, amely tetszőleges előjelű lehet a környezet és a tartomány közös részén.

Az $\mathbf{x} = (3,9)$ KKT pont ($\mathbf{u} = (\frac{15}{4}, \frac{15}{4})$ Lagrange szorzókkal) vizsgálata:

Mindkét feltétel aktív és a Lagrange szorzók pozitívak, így az erős elégségességi tétel használható. A $\mathbf{C}$ mátrix a $\boldsymbol{\nabla}^2 L$ mátrix lesz a behelettesítés után:

$$\mathbf{C} = \begin{bmatrix} -\frac{15}{2} & 1 & -6 & 2 \\ 1 & 0 & 1 & -1 \\ -6 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{bmatrix}.$$

Mivel $Iner(\mathbf{C}) = (2,0,2)$, így az $\mathbf{x} = (3,9)$ KKT pont szigorú maximumpont.

Feladat:
Határozzuk meg a tartomány azon pontját, amelyben az x_1x_2 mennyiség a legkisebb!

35. Példa
Legyen adott egy ellipszis, amelynek centruma az origó, x_1 tengely irányú féltengelyének hossza 2, az x_2 tengely irányú féltengelyének hossza pedig 1. Legyen adott továbbá egy kör, amelynek középpontja $(0,1)$, sugara 2. Tekintsük azt a tartomány, amelynek pontja az ellipszis és a kör belsejének közös pontjai beleértve a határpontokat is. Határozzuk meg a tartomány azon pontját, amelyek a $(2,1)$ ponthoz legközelebb, ill. legmesszebb van!

Megoldás
Először matematikailag megfogalmazzuk az optimalizálási feladatot. Az ellipszis egyenlete:

$$\frac{x_1^2}{2^2}+\frac{x_2^2}{1^2}=1,$$

a kör egyenlete:

$$x_1^2+(x_2-1)^2=2^2.$$

Az optimalizálási feladat az alábbi:

$$\begin{array}{rcl}(x_1-2)^2+(x_2-1)^2 & \to & \text{min!} \quad \text{vagy} \quad \text{max!}\\ x_1^2+4x_2^2-4 & \leqq & 0\\ x_1^2+x_2^2-2x_2-3 & \leqq & 0\end{array}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1,x_2,u_1,u_2)=(x_1-2)^2+(x_2-1)^2+u_1(x_1^2+4x_2^2-4)+u_2(x_1^2+x_2^2-2x_2-3)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\begin{aligned}\nabla L &= \begin{bmatrix} 2x_1-4+2u_1x_1+2u_2x_1\\ 2x_2-2+8u_1x_2+u_2\,(2x_2-2)\\ x_1^2+4x_2^2-4\\ x_1^2+x_2^2-2x_2-3\end{bmatrix},\\ \nabla^2 L &= \begin{bmatrix} 2+2u_1+2u_2 & 0 & 2x_1 & 2x_1\\ 0 & 2+8u_1+2u_2 & 8x_2 & 2x_2-2\\ 2x_1 & 8x_2 & 0 & 0\\ 2x_1 & 2x_2-2 & 0 & 0\end{bmatrix}.\end{aligned}$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned}2x_1-4+2u_1x_1+2u_2x_1&=0\\ 2x_2-2+8u_1x_2+u_2\,(2x_2-2)&=0\\ x_1^2+4x_2^2-4&\leqq 0\\ x_1^2+x_2^2-2x_2-3&\leqq 0\\ u_1(x_1^2+4x_2^2-4)&=0\\ u_2(x_1^2+x_2^2-2x_2-3)&=0\end{aligned}$$

Az $u_1,u_2=0$ Lagrange szorzókra nem tettünk előjelkötést. Szerencsére a minimum és a maximum feladatnak is ugyanaz a KKT rendszere, csupán a Lagrange szorzók előjelkötésében különböznek. Mint ismeretes, a pozitív u_i értékek a minimum feladat Lagrange szorzói, a negatív u_i értékek pedig a maximum feladat Lagrange szorzói. Az u_1, u_2 értékeitől függően 4 esetre bontjuk szét a KKT feladatot, amelyek a következők.

i) $u_1 = u_2 = 0$ eset

Az első két egyenletből azonnal adódik az $\mathbf{x} = (2, 1)$ pont. Ez azonban nem teljesíti az első primál feltételt, mert az ellipszisen kívülre esik a pont.

ii) $u_1 = 0, u_2 \neq 0$ eset

Az első két egyenletből és a második komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 4 + 2u_2x_1 &= 0 \\ 2x_2 - 2 + u_2(2x_2 - 2) &= 0 \\ x_1^2 + x_2^2 - 2x_2 - 3 &= 0 \end{aligned}$$

Két megoldás adódik: $x_1 = 2, x_2 = 1, u_2 = 0$ és az $x_1 = -2, x_2 = 1, u_2 = -2$. Egyik sem teljesíti az első feltételt.

iii) $u_1 \neq 0, u_2 = 0$ eset

Az első két egyenletből és az első komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 4 + 2u_1x_1 &= 0 \\ 2x_2 - 2 + 8u_1x_2 &= 0 \\ x_1^2 + 4x_2^2 - 4 &= 0 \end{aligned}$$

Két megoldás adódik: $x_1 = 1.66, x_2 = 0.55, u_1 = 0.20$ és az $x_1 = -1.98, x_2 = -0.14, u_1 = -2.01$. Az első megoldás teljesíti a második feltételt, mert a kör belsejében van az $\mathbf{x} = (1.66, 0.55)$ pont. Mivel $u_1 > 0$ ezért a minimum feladat KKT pontja. A második megoldás a körön kívül van.

iv) $u_1 \neq 0, u_2 \neq 0$ eset

Az első két egyenletből és a két komplementaritási feltételből adódóan az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2x_1 - 4 + 2u_1x_1 + 2u_2x_1 &= 0 \\ 2x_2 - 2 + 8u_1x_2 + u_2(2x_2 - 2) &= 0 \\ x_1^2 + 4x_2^2 - 4 &= 0 \\ x_1^2 + x_2^2 - 2x_2 - 3 &= 0 \end{aligned}$$

Két megoldás adódik:

$x_1 = 1.886, \quad x_2 = 0.333, u_1 = 0.354, \quad u_2 = -0.293$ és az

$x_1 = -1.886, x_2 = 0.333, u_1 = -0.354, u_2 = -1.707.$

Az első megoldás nem KKT pont, mert a Lagrange szorzók nem azonos előjelüek. A második megoldás KKT pont, mert a Lagrange szorzók azonos előjelüek, mégpedig negatívok. Ezért a $(-1.886, 0.333)$ pont a maximum feladat KKT pontja.

Összefoglalva eredményeinket, az alábbiakat kaptuk:

a) A minimum feladatra egy KKT pont adódott: $\mathbf{x} = (1.66, 0.55)$, a Lagrange szorzók: $u_1 = 0, u_2 = 0.20$, az első feltétel aktív, a második feltétel inaktív.

b) A maximum feladatra egy KKT pont adódott: $\mathbf{x} = (-1.886, 0.333)$, a Lagrange szorzók: $u_1 = -0.354, u_2 = -1.707$, az első és a második feltétel is aktív.

Elégségesség vizsgálat:
a) Minimum feladat esetén: A szegélyezett $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} 2.4 & 0 & 3.32 \\ 0 & 2.4 & 4.4 \\ 3.32 & 4.4 & 0 \end{bmatrix}$$

Mivel az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminor mátrixa) pozitív definit, így az $\mathbf{x} = (1.66, 0.55)$ pont szigorú lokális minimumpont.

b) Maximum feladat esetén: A szegélyezett $\mathbf{C}$ mátrix:

$$\mathbf{C} = \begin{bmatrix} -2.122 & 0 & -3.772 & -3.772 \\ 0 & -4.246 & 2.664 & -1.334 \\ -3.772 & 2.664 & 0 & 0 \\ -3.772 & -1.334 & 0 & 0 \end{bmatrix}$$

Mivel az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminor mátrixa) negatív definit, így az $\mathbf{x} = (-1.886, 0.333)$ pont szigorú lokális maximumpont.

36. Példa

Legyen adott egy ellipszis, amelynek centruma az origó, x_1 tengely irányú féltengelyének hossza 1, az x_2 tengely irányú féltengelyének hossza pedig 2.

a) Határozzuk meg az ellipszisbe írható legnagyobb területű téglalapot!

b) Határozzuk meg az ellipszisbe írható legnagyobb kerületű téglalapot!

Megoldás

a) Először matematikailag megfogalmazzuk az optimalizálási feladatot. Legyen x_1 a téglalap vizszintes oldalhosszának a fele, x_2 pedig a téglalap függőleges oldalhosszának a fele. Az (x_1, x_2) pont az ellipszisen van, így ki kell elégíteni az ellipszis egyenletét:

$$\frac{x_1^2}{1^2} + \frac{x_2^2}{2^2} = 1.$$

A téglalap területete: $4x_1x_2$.

Az optimalizálási feladat az alábbi:

$$\begin{aligned} x_1x_2 &\rightarrow \text{max!} \\ 4 - 4x_1^2 - x_2^2 &= 0 \\ (x_1, x_2) &\in X = \{(x_1, x_2) : x_1, x_2 > 0\} \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, v) = x_1x_2 + v(4 - 4x_1^2 - x_2^2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} x_2 - 8vx_1 \\ 4x_1 - 2vx_2 \\ 4 - 4x_1^2 - x_2^2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} -8v & 1 & -8x_1 \\ 1 & -2v & -2x_2 \\ -8x_1 & -2x_2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} x_2 - 8vx_1 &= 0 \\ x_1 - 2vx_2 &= 0 \\ 4 - 4x_1^2 - x_2^2 &= 0 \end{aligned}$$

Csak a $v \neq 0$ megoldás jöhet szóba, mert ellenkező esetben ellentmondásra jutunk. Négy megoldás is adódik, de a döntési változókra előírt pozitivitás miatt csak az alábbi jöhet szóba: $x_1 = \frac{1}{2}\sqrt{2}, x_2 = \sqrt{2}$, a Lagrange szorzó értéke $v = \frac{1}{4}$.

Elégségesség vizsgálathoz szükséges szegélyezett $\mathbf{C}$ mátrix, amely a $\nabla^2 L$ mátrixból származtatható:

$$\mathbf{C} = \begin{bmatrix} -2 & 1 & -4\sqrt{2} \\ 1 & -\frac{1}{2} & -2\sqrt{2} \\ -4\sqrt{2} & -2\sqrt{2} & 0 \end{bmatrix}.$$

Mivel az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminor mátrixa) negatív definit, így az $x_1 = \frac{1}{2}\sqrt{2}, x_2 = \sqrt{2}$ megoldás szigorú lokális maximumpont. Javasoljuk az olvasónak, hogy gyakorlásképpen az aktualizált Lagrange függvény Hesse mátrixának feltételes negatív definitségét is ellenőrizze.

Tehát a maximális területű téglalap oldalhosszúságai: vizszintes tengely irányú: $\sqrt{2} \approx 1.41$, függőleges tengely irányú: $2\sqrt{2} \approx 2.83$, a maximális terület: 4

b) A maximális kerületű téglalap meghatározásához tartozó optimalizálási feladat a következő:

$$\begin{aligned} x_1 + x_2 &\rightarrow \text{max!} \\ 4 - 4x_1^2 - x_2^2 &= 0 \\ (x_1, x_2) &\in X = \{(x_1, x_2) : x_1, x_2 > 0\} \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, v) = x_1 + x_2 + v(4 - 4x_1^2 - x_2^2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} 1 - 8vx_1 \\ 1 - 2vx_2 \\ 4 - 4x_1^2 - x_2^2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} -8v & 0 & -8x_1 \\ 0 & -2v & -2x_2 \\ -8x_1 & -2x_2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} 1 - 8vx_1 &= 0 \\ 1 - 2vx_2 &= 0 \\ 4 - 4x_1^2 - x_2^2 &= 0 \end{aligned}$$

A döntési változókra előírt pozitivitás miatt csak az alábbi megoldás jöhet szóba: $x_1 = \frac{1}{5}\sqrt{5}, x_2 = \frac{4}{5}\sqrt{5}$, a Lagrange szorzó értéke $v = \frac{1}{8}\sqrt{5}$.

Elégségesség vizsgálathoz szükséges szegélyezett $\mathbf{C}$ mátrix, amely a $\nabla^2 L$ mátrixból származtatható:

$$\mathbf{C} = \begin{bmatrix} -\sqrt{5} & 0 & -\frac{8}{5}\sqrt{5} \\ 0 & -\frac{1}{4}\sqrt{5} & -\frac{8}{5}\sqrt{5} \\ -\frac{8}{5}\sqrt{5} & -\frac{8}{5}\sqrt{5} & 0 \end{bmatrix}$$

Mivel az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminor mátrixa) negatív definit, így az $x_1 = \frac{1}{5}\sqrt{5}, x_2 = \frac{4}{5}\sqrt{5}$ megoldás szigorú lokális maximumpont. Javasoljuk az olvasónak, hogy gyakorlásképpen az aktualizált Lagrange függvény Hesse mátrixának feltételes negatív definitségét is ellenőrizze.

Tehát a maximális kerületű téglalap oldalhosszúságai: vizszintes tengely irányú: $\frac{2}{5}\sqrt{5} \approx 0.89$, függőleges tengely irányú: $\frac{8}{5}\sqrt{5} \approx 3.58$, a maximális kerület: $4\sqrt{5} \approx 8.9$.

Az a) és b) feladat eredményei összefoglalva:

A maximális területű téglalap oldalhosszúságai: $\sqrt{2} \approx 1.41$, $2\sqrt{2} \approx 2.83$; területe: 4; kerülete: 2.83.

A maximális kerületű téglalap oldalhossúzságai: $\frac{2}{5}\sqrt{5} \approx 0.89$, $\frac{8}{5}\sqrt{5} \approx 3.58$; területe: $\frac{16}{5} = 3.2$; kerülete: $4\sqrt{5} \approx 8.9$.

37. Példa

Legyen adott két parabolaív: $y = \sqrt{x}$ és az $y = \sqrt{9-2x}$. A két parabolaív és az x tengely meghatároz egy kompakt tartományt. Határozzuk meg a tartományba írható legnagyobb területű téglalapot!

Megoldás

Először matematikailag megfogalmazzuk az optimalizálási feladatot. Legyen x_1 a keresett téglalap vízszintes tengelyen lévő baloldali csúcspontjának x koordinátája, x_2 pedig a jobboldali csúcspontjának x koordinátája. A téglalap bal felső csúcspontja az első parabolaíven, a jobb felső csúcspontja pedig a második parabolaíven van. Ahhoz, hogy téglalap alakuljon ki, a két felső csúcspontnak azonos távolságra kell lenni az x tengelytől, azaz fenn kell állni a következő összefüggésnek:

$$\sqrt{x_1} = \sqrt{9-2x_2},$$

amely négyzetre emelés után a $9 - x_1 - 2x_2 = 0$ feltételt adja. A téglalap területetete: $(x_2 - x_1)\sqrt{x_1}$.

Az optimalizálási feladat az alábbi:

$$\begin{aligned} (x_2 - x_1)\sqrt{x_1} &\to \max! \\ 9 - x_1 - 2x_2 &= 0 \\ (x_1, x_2) &\in X = \{(x_1, x_2) : 0 < x_1 < 3, \quad 3 < x_2 < 4.5\} \end{aligned}$$

A feladathoz tartozó Lagrange függvény:

$$L(x_1, x_2, v) = x_2\sqrt{x_1} - x_1\sqrt{x_1} + v(9 - x_1 - 2x_2)$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense és Hesse mátrixa:

$$\nabla L = \begin{bmatrix} \frac{x_2}{2\sqrt{x_1}} - \frac{3}{2}\sqrt{x_1} - v \\ \sqrt{x_1} - 2v \\ 9 - x_1 - 2x_2 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} -\frac{1}{4x_1\sqrt{x_1}}x_2 - \frac{3}{4\sqrt{x_1}} & \frac{1}{2\sqrt{x_1}} & -1 \\ \frac{1}{2\sqrt{x_1}} & 0 & -2 \\ -1 & -2 & 0 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi rendszert kell megoldani:

$$\begin{aligned} \frac{x_2}{2\sqrt{x_1}} - \frac{3}{2}\sqrt{x_1} - v &= 0 \\ \sqrt{x_1} - 2v &= 0 \\ 9 - x_1 - 2x_2 &= 0 \end{aligned}$$

Figyelembe véve az X halmazt, az alábbi megoldás adódott: $x_1 = 1, x_2 = 4$, a Lagrange szorzó értéke $v = 0.5$.

Elégségesség vizsgálathoz szükséges szegélyezett $\mathbf{C}$ mátrix, amely a $\nabla^2 L$ mátrixból származtatható az alábbi:

$$\mathbf{C} = \begin{bmatrix} -\frac{7}{4} & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & -2 \\ -1 & -2 & 0 \end{bmatrix}.$$

Főminor teszt segítségével: Mivel $n = 2$, $s = 1$, így $k = 2$, azaz csak a $\mathbf{C}$ mátrix determinánsát kell meghatározni. amely $\det(\mathbf{C}) = 9$. Mivel $(-1)^k \det(\mathbf{C}) > 0$, így az aktualizált Lagrange függvény Hesse mátrixa (a $\mathbf{C}$ mátrix 2×2-es főminor mátrixa) feltételesen negatív definit, amelyből következik, hogy a kapott megoldás szigorú lokális maximum.

Inercia teszt segítségével: Egyszerű számolással kapjuk, hogy $Iner(\mathbf{C}) = (2, 0, 1)$, ugyanarra a végeredményre jutottunk.

Tehát a maximális területű téglalap oldalhossúságai: vízszintes tengely irányú: 3, függőleges tengely irányú: 1, a maximális terület: 3.

Feladat:
Határozza meg az ellipszisbe írható legnagyobb kerületű téglalapot!

38. Példa

Adott az $\mathbf{A} \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix segítségével az $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ kvadratikus függvény és legyen $\mathbf{x}^T\mathbf{x} = 1$, azaz $\sum x_i^2 = 1$. Tekintsük az alábbi optimalizálási feladatot:

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T\mathbf{A}\mathbf{x} \to \text{min! vagy} \quad \text{max!} \\ \mathbf{x}^T\mathbf{x} &= 1 \end{aligned}$$

Igazoljuk, hogy az optimalizálási feladathoz tartozó Lagrange szorzó értéke a minimum feladatnál az $\mathbf{A}$ mátrix legkisebb sajátértéke, maximum feladatnál pedig a legnagyobb sajátértéke! Az $\mathbf{x}$ vektor optimális értéke pedig a sajátértékhez tartozó sajátvektor.

Megoldás:
A feladathoz tartozó Lagrange függvény:

$$L(\mathbf{x}, v) = \mathbf{x}^T\mathbf{A}\mathbf{x} + v(1 - \mathbf{x}^T\mathbf{x})$$

A feladat függvényei differenciálhatók, a Lagrange függvény gradiense:

$$\nabla L = \begin{bmatrix} 2\mathbf{A}\mathbf{x} - 2v\mathbf{x} \\ 1 - \mathbf{x}^T\mathbf{x} \end{bmatrix}.$$

A KKT pont meghatározásához akár minimum, akár maximum feladatról van szó az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 2\mathbf{A}\mathbf{x} - 2v\mathbf{x} &= \mathbf{0} \\ 1 - \mathbf{x}^T\mathbf{x} &= 0 \end{aligned}$$

Az első egyenletből $\mathbf{A}\mathbf{x} = v\mathbf{x}$ adódik, ami szerint a v Lagrange szorzó az $\mathbf{A}$ mátrix sajátértéke, az $\mathbf{x}$ vektor pedig a v sajátértékhez tartozó sajátvektor. Szorozzuk be az egyenletet balról $\mathbf{x}^T$ vektorral, kapjuk, hogy $\mathbf{x}^T\mathbf{A}\mathbf{x} = v\mathbf{x}^T\mathbf{x}$, figyelembe véve KKT rendszer utolsó egyenletét, adódik, hogy a célfüggvény egyenlő a Lagrange szorzóval. Ebből következik, hogy az optimalizálási feladatban a célfüggvény maximuma a legnagyobb sajátérték, minimuma pedig a legkisebb sajátérték.

Tekintsünk az elmondottakra egy szánmpéldát, legyen adott az alábbi $\mathbf{A}$ mátrix és oldjuk meg a feladatot!

$$\mathbf{A} = \begin{bmatrix} 6 & 3 \\ 3 & -2 \end{bmatrix}.$$

Az optimalizálási feladat:

$$\begin{aligned} f(x_1, x_2) &= \ = \mathbf{x}^T\mathbf{A}\mathbf{x} = 6x_1^2 + 6x_1x_2 - 2x_2^2 \to \text{min! vagy} \quad \text{max!} \\ h(x_1, x_2) &= 1 - x_1^2 - x_2^2 = 0 \end{aligned}$$

A feladathoz tartozó Lagrange függvény és gradiense:

$$L(\mathbf{x}, v) = 6x_1^2 + 6x_1x_2 - 2x_2^2 + v(1 - x_1^2 - x_2^2), \qquad \nabla L = \begin{bmatrix} 12x_1 + 6x_2 - 2vx_1 \\ 6x_1 - 4x_2 - 2vx_2 \\ 1 - x_1^2 - x_2^2 \end{bmatrix}.$$

A KKT pont meghatározásához az alábbi egyenletrendszert kell megoldani:

$$\begin{aligned} 12x_1 + 6x_2 - 2vx_1 &= 0 \\ 6x_1 - 4x_2 - 2vx_2 &= 0 \\ 1 - x_1^2 - x_2^2 &= 0 \end{aligned}$$

Négy megoldás adódott, ezek a következők:
$v_1 = 7, \quad \mathbf{x}_{11} = \left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)$
$v_1 = 7, \quad \mathbf{x}_{12} = \left(-\frac{3}{\sqrt{10}}, -\frac{1}{\sqrt{10}}\right)$
$v_2 = -3, \quad \mathbf{x}_{21} = \left(\frac{1}{\sqrt{10}}, -\frac{3}{\sqrt{10}}\right)$
$v_2 = -3, \quad \mathbf{x}_{22} = \left(-\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right)$
A maximális célfüggvényérték 7, a minimális pedig -3.

12. Érzékenységvizsgálat

Hasonlóan a feltétel nélküli optimalizáláshoz, a feltételes optimalizálási feladat esetében is sok esetben nemcsak az optimális megoldásra vagyunk kíváncsiak, hanem arra is, hogy a feladatban szereplő adatok változása milyen hatást gyakorol az optimális megoldásra. A vizsgált adat lehet csak a célfüggvényben, csak a feltételi függvényekben, de az is lehet, hogy mindenütt szerepel. A képletek bonyolultsága miatt itt is először csak egy paramétert tartalmazó optimalizálási feladattal foglalkozzunk.

12.1. Célfüggvény érzékenységvizsgálata (burkolótétel)

Legyen a döntési változók vektora $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, az egyetlen paraméter pedig $p \in \mathbb{R}$. Legyen az optimalizálási feladadat a következő:

$$\begin{array}{rl} f(\mathbf{x}, p) \to \min! & \\ g_i(\mathbf{x}, p) \leqq 0 & i = 1, 2, \ldots, m \\ h_i(\mathbf{x}, p) = 0 & i = 1, 2, \ldots, k \\ \mathbf{x} \in X & X \subset \mathbb{R}^n, \text{ nyílt} \end{array}$$

Mielőtt belekezdenénk az érzékenységvizsgálat tárgyalásába, felírjuk a feladat Lagrange függvényét, amely mint látni fogjuk, fontos szerepet fog játszani a vizsgálatainkban. A feltétel nélküli optimalizálásban megismert összefüggésekhez hasonló összefüggések fognak itt is szerepelni, de az ottani célfüggvény szerepét itt a Lagrange függvény veszi át. A Lagrange függvény felírásához az egyszerűbb vektoros formát használjuk.

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}, p) = f(\mathbf{x}, p) + \mathbf{u}^T \mathbf{g}(\mathbf{x}, p) + \mathbf{v}^T \mathbf{h}(\mathbf{x}, p),$$

ahol a $\mathbf{g}(\mathbf{x}, p)$ és a $\mathbf{h}(\mathbf{x}, p)$ oszlopvektorok, elemei pedig a feltételi függvények. A képletek bonyolutsága miatt itt használjuk a traszponálás jelét.

A KKT pont meghatározásához megoldjuk a KKT feltételeket leíró rendszert, amelynek megoldása általában mind a döntési változók, mind a Lagrange szorzók tekintetében függ a paramétertől. Tegyük fel, hogy fennállnak az implicit függvény tétel feltételei és így a döntési változók és a Lagrange szorzók optimális értékei felírhatók a paraméter függvényében, azaz

$$\begin{aligned} \mathbf{x} &= \mathbf{x}(p) \\ \mathbf{u} &= \mathbf{u}(p) \\ \mathbf{v} &= \mathbf{v}(p) \end{aligned}$$

Ha a döntési változók optimális értékeit behelyettesítjük a célfüggvénybe, akkor megkapjuk a célfüggvény optimális értékét a paraméter függvényében, mint tudjuk ezt a függvényt optimum érték függvénynek is szokás nevezni és M-el jelöljük. Az optimum érték függvény tehát

$$M(p) = f(\mathbf{x}(p), p).$$

Hasonló módon a döntési változók és a Lagrange szorzók optimális értékeit behelyettesíthetjük a Lagrange függvénybe, ekkor megkapjuk a Lagrange függvényt a paraméter függvényében. Jelölje ezt L_M, amelyet a következő módon írhatunk:

$$L_M(p) = L(\mathbf{x}(p), \mathbf{u}(p), \mathbf{v}(p), p) = f(\mathbf{x}(p), p) + \mathbf{u}(p)^T \mathbf{g}(\mathbf{x}(p), p) + \mathbf{v}(p)^T \mathbf{h}(\mathbf{x}(p), p).$$

ÁLLÍTÁS:

$$L_M(p) \equiv M(p)$$

Ezt az állítást nem kell bebizonyítani, mivel korábbi ismereteink szerint tudjuk, hogy optimális esetben a célfüggvény és a Lagrange függvény értéke azonos. Most is optimális esetről van szó, de most minden paraméter értékre lesz ugyanaz a két oldal értéke, így a két függvény megegyezik.

TÉTEL (Burkolótétel):
Legyen adott a $\min\{f(\mathbf{x},p) : \mathbf{g}(\mathbf{x},p) \leqq \mathbf{0}, \mathbf{h}(\mathbf{x},p) = \mathbf{0}, \mathbf{x} \in X\}$ optimalizálási feladat $M(p)$ optimum érték függvénye. Ekkor az optimum érték függvény p paraméter szerinti deriváltja

$$\frac{dM(p)}{dp} = \left.\frac{\partial L(\mathbf{x},\mathbf{u},\mathbf{v},p)}{\partial p}\right|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)}$$

Szavakban: Az $M(p)$ optimum érték függvény p paraméter szerinti deriváltja megegyezik az $L(\mathbf{x},\mathbf{u},\mathbf{v},p)$ Lagrange függvény p paraméter szerinti parciális deriváltjának az optimumhelyen vett értékével.

Bizonyítás:
A fentebb leírt $M(p) = L_M(p)$ állítás értelmében írhatjuk, hogy

$$\begin{aligned}\frac{dM(p)}{dp} &= \frac{dL_M(p)}{dp} = \frac{dL(\mathbf{x}(p),\mathbf{u}(p),\mathbf{v}(p),p)}{dp} = \\ &= \frac{df(\mathbf{x}(p),p)}{dp} + \frac{d\left[\mathbf{u}(p)^T\mathbf{g}(\mathbf{x}(p),p)\right]}{dp} + \frac{d\left[\mathbf{v}(p)^T\mathbf{h}(\mathbf{x}(p),p)\right]}{dp}\end{aligned}$$

A fenti képletnél alkalmaztuk az összegfüggvényre vonatkozó differenciálási szabályt, most alkalmazzuk a szorzatfüggvényre vonatkozó szabályt és a láncszabályt, majd pedig egy kis rendezés után a következőket kapjuk:

$$\begin{aligned}\frac{dM(p)}{dp} &= \frac{df(\mathbf{x}(p),p)}{dp} + \sum_{j=1}^{n}\frac{\partial f(\mathbf{x}(p),p)}{\partial x_j}\frac{dx_j(p)}{dp} \\ &\quad + \left[\frac{d\mathbf{u}(p)^T}{dp}\mathbf{g}(\mathbf{x}(p),p) + \mathbf{u}(p)^T\left(\frac{d\mathbf{g}(\mathbf{x}(p),p)}{dp} + \sum_{j=1}^{n}\frac{\partial \mathbf{g}(\mathbf{x}(p),p)}{\partial x_j}\frac{dx_j(p)}{dp}\right)\right] \\ &\quad + \left[\frac{d\mathbf{v}(p)^T}{dp}\mathbf{h}(\mathbf{x}(p),p) + \mathbf{v}(p)^T\left(\frac{d\mathbf{h}(\mathbf{x}(p),p)}{dp} + \sum_{j=1}^{n}\frac{\partial \mathbf{h}(\mathbf{x}(p),p)}{\partial x_j}\frac{dx_j(p)}{dp}\right)\right] \\ &= \frac{df(\mathbf{x}(p),p)}{dp} + \mathbf{u}(p)^T\frac{d\mathbf{g}(\mathbf{x}(p),p)}{dp} + \mathbf{v}(p)^T\frac{d\mathbf{h}(\mathbf{x}(p),p)}{dp} \\ &\quad + \sum_{j=1}^{n}\left(\frac{\partial f(\mathbf{x}(p),p)}{\partial x_j} + \mathbf{u}(p)^T\frac{\partial \mathbf{g}(\mathbf{x}(p),p)}{\partial x_j} + \mathbf{v}(p)^T\frac{\partial \mathbf{h}(\mathbf{x}(p),p)}{\partial x_j}\right)\frac{dx_j(p)}{dp} \\ &\quad + \frac{d\mathbf{u}(p)^T}{dp}\mathbf{g}(\mathbf{x}(p),p) + \frac{d\mathbf{v}(p)^T}{dp}\mathbf{h}(\mathbf{x}(p),p)\end{aligned}$$

Az utolsó előtti sorban szereplő

$$\frac{\partial f(\mathbf{x}(p),p)}{\partial x_j} + \mathbf{u}(p)^T\frac{\partial \mathbf{g}(\mathbf{x}(p),p)}{\partial x_j} + \mathbf{v}(p)^T\frac{\partial \mathbf{h}(\mathbf{x}(p),p)}{\partial x_j}$$

mennyiség valójában az $\boldsymbol{\nabla}_{\mathbf{x}}L$ gradiens vektor egyes elemei. Az optimalitás egyik szükséges feltétele (duál feltétel) szerint $\boldsymbol{\nabla}_{\mathbf{x}}L = \mathbf{0}$, így ezek mindegyike zérus az optimális esetben. A fenti hosszú képlet utolsó sorában lévő tagok mindegyike zérus, hisz $\mathbf{h}(\mathbf{x}(p),p) = \mathbf{0}$, mivel az $\mathbf{x}(p)$ optimális megoldás lehetséges megoldás is. A $\mathbf{g}(\mathbf{x}(p),p) = \mathbf{0}$ is igaz, a Lagrange függvényben ugyanis csak az aktív feltételi függvények szerepelnek, hiszen az inaktív feltételeknél $u_i = 0$. Tehát a keresett mennyiség a fenti hosszú képlet alulról harmadik sorában lévő összefüggésre redukálódik, amely nem más mint a Lagrange függvény p paraméter szerinti deriváltja az optimumhelyen. **Q.e.d.**

A burkolótétel itt azt mondja, hogy a p paraméter megváltozásának az $M(p)$ optimum érték függvényre gyakorolt hatása azonos a Lagrange függvényre gyakorolt közvetlen hatással.

12.2. Döntési változók érzékenységvizsgálata

TÉTEL

Legyen adott a $\min\{f(\mathbf{x},p) : \mathbf{g}(\mathbf{x},p) \leqq \mathbf{0}, \mathbf{h}(\mathbf{x},p) = \mathbf{0}, \mathbf{x} \in X\}$ optimalizálási feladat. Legyenek ismertek a döntési változó $\mathbf{x} = \mathbf{x}(p)$ és a Lagrange szorzók $\mathbf{u} = \mathbf{u}(p),\ \mathbf{v} = \mathbf{v}(p)$ optimális értékei a p paraméter függvényében. Ekkor a döntési változók és a Lagrange szorzók vektorának p paraméter szerinti deriváltja

$$\frac{d}{dp}\begin{bmatrix} \mathbf{x}(p) \\ \mathbf{u}(p) \\ \mathbf{v}(p) \end{bmatrix} = -\left.\begin{bmatrix} \boldsymbol{\nabla}^2_{\mathbf{xx}}L & \boldsymbol{\nabla}\mathbf{g}^T & \boldsymbol{\nabla}\mathbf{h}^T \\ \boldsymbol{\nabla}\mathbf{g} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\nabla}\mathbf{h} & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1}\right|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} \cdot \frac{\partial}{\partial p}\left.\begin{bmatrix} \boldsymbol{\nabla}_{\mathbf{x}}L \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}\right|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)}$$

ahol a $\boldsymbol{\nabla}^2_{\mathbf{xx}}L$ a Lagrange függvény döntési változók szerinti másodrendű parciális deriváltjait (Hesse mátrix) jelentik. A $\boldsymbol{\nabla}\mathbf{g}$, $\boldsymbol{\nabla}\mathbf{h}$ pedig olyan mátrixok, amelyek sorvektorai az aktív egyenlőtlenséges és az egyenlőséges feltételi függvények gradiensei.

Bizonyítás:

Induljunk ki az optimalitás szükséges feltételéből, amely a duál feltétel, valamint az aktív egyenlőtlenséges és az egyenlőséges feltételi függvények segítségével az alábbi egyenletrendszert szolgáltatja:

$$\begin{aligned} \boldsymbol{\nabla}_{\mathbf{x}}L(\mathbf{x},\mathbf{u},\mathbf{v},p) &= \mathbf{0} \\ \mathbf{g}(\mathbf{x},p) &= \mathbf{0} \\ \mathbf{h}(\mathbf{x},p) &= \mathbf{0} \end{aligned}$$

A döntési változók vektora $\mathbf{x}(p)$, a Lagrange szorzók vektora pedig $\mathbf{u}(p),\ \mathbf{v}(p)$. Az egyenletrendszer Jacobi-mátrixa az alábbi:

$$\mathbf{J} = \begin{bmatrix} \boldsymbol{\nabla}_{\mathbf{x}}(\boldsymbol{\nabla}_{\mathbf{x}}L) & \boldsymbol{\nabla}_{\mathbf{u}}(\boldsymbol{\nabla}_{\mathbf{x}}L) & \boldsymbol{\nabla}_{\mathbf{v}}(\boldsymbol{\nabla}_{\mathbf{x}}L) \\ \boldsymbol{\nabla}_{\mathbf{x}}\mathbf{g} & \boldsymbol{\nabla}_{\mathbf{u}}\mathbf{g} & \boldsymbol{\nabla}_{\mathbf{u}}\mathbf{g} \\ \boldsymbol{\nabla}_{\mathbf{x}}\mathbf{h} & \boldsymbol{\nabla}_{\mathbf{u}}\mathbf{h} & \boldsymbol{\nabla}_{\mathbf{v}}\mathbf{h} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\nabla}^2_{\mathbf{xx}}L & \boldsymbol{\nabla}\mathbf{g}^T & \boldsymbol{\nabla}\mathbf{h}^T \\ \boldsymbol{\nabla}\mathbf{g} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\nabla}\mathbf{h} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Az optimalitás miatt a Jacobi-mátrix determinánsa nem zérus, így alkalmazható az implicit függvény tétel, amelyből kapjuk, hogy

$$\frac{d}{dp}\begin{bmatrix} \mathbf{x}(p) \\ \mathbf{u}(p) \\ \mathbf{v}(p) \end{bmatrix} = -\left.\begin{bmatrix} \boldsymbol{\nabla}^2_{\mathbf{xx}}L & \boldsymbol{\nabla}\mathbf{g}^T & \boldsymbol{\nabla}\mathbf{h}^T \\ \boldsymbol{\nabla}\mathbf{g} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\nabla}\mathbf{h} & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1}\right|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} \cdot \frac{\partial}{\partial p}\left.\begin{bmatrix} \boldsymbol{\nabla}_{\mathbf{x}}L \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}\right|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)}$$

Q.e.d.

Megjegyzés:

A fenti tételbeli képlet sok esetben az alábbi ekvivalens formában is használatos. Könnyen meggyőződhetünk arról, hogy $\boldsymbol{\nabla}_{\mathbf{u}} L = \mathbf{g}$, $\boldsymbol{\nabla}_{\mathbf{v}} L = \mathbf{h}$. Arról is könnyen meggyőződhetünk, hogy $\boldsymbol{\nabla}^2_{\mathbf{ux}} L = (\boldsymbol{\nabla}^2_{\mathbf{xu}} L)^T = \boldsymbol{\nabla}\mathbf{g}$, $\boldsymbol{\nabla}^2_{\mathbf{vx}} L = (\boldsymbol{\nabla}^2_{\mathbf{xv}} L)^T = \boldsymbol{\nabla}\mathbf{h}$. E jelölésekkel az ekvivalens formula a következő:

$$\frac{d}{dp}\begin{bmatrix} \mathbf{x}(p) \\ \mathbf{u}(p) \\ \mathbf{v}(p) \end{bmatrix} = -\begin{bmatrix} \boldsymbol{\nabla}^2_{\mathbf{xx}} L & \boldsymbol{\nabla}^2_{\mathbf{xu}} L & \boldsymbol{\nabla}^2_{\mathbf{xv}} L \\ \boldsymbol{\nabla}^2_{\mathbf{ux}} L & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\nabla}^2_{\mathbf{vx}} L & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1}\Bigg|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} \cdot \frac{\partial}{\partial p}\begin{bmatrix} \boldsymbol{\nabla}_{\mathbf{x}} L \\ \boldsymbol{\nabla}_{\mathbf{u}} L \\ \boldsymbol{\nabla}_{\mathbf{v}} L \end{bmatrix}_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)}$$

vagy még egyszerűbben, ha a Lagrange függvény összes változója szerint képezzük a deriváltakat ($\mathbf{H}_L = \boldsymbol{\nabla}^2 L$):

$$\frac{d}{dp}\begin{bmatrix} \mathbf{x}(p) \\ \mathbf{u}(p) \\ \mathbf{v}(p) \end{bmatrix} = -\left(\boldsymbol{\nabla}^2 L\right)^{-1}\Big|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} \cdot \frac{\partial}{\partial p}\boldsymbol{\nabla} L\Big|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)}.$$

A tételt ebben az utóbbi formában jegyezhetjük meg legkönnyebben, azaz a döntési változók és a Lagrange szorzók p paraméter szerinti deriváltvektora egyenlő a Lagrange függvény Hesse mátrixa inverzének (-1)-szerese szorozva a Lagrange függvény gradiensvektorának p paraméter szerinti parciális deriváltjával. Mindkét mennyiséget az optimumhelyen vett értékekkel kell figyelembe venni.

12.3. Lagrange szorzók értelmezése

Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_1(\mathbf{x}) &\leqq p \\ g_i(\mathbf{x}) &\leqq 0 \qquad i = 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \qquad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

ahol most a paraméter az első egyenlőtlenséges feltétel jobboldala. Ekkor a feladat szokásos formában való felírásánál az első feltétel $g_1(\mathbf{x}) - p \leqq 0$ alakra módosul. A megoldás után ismertek az $\mathbf{x} = \mathbf{x}(p), \mathbf{u} = \mathbf{u}(p), \mathbf{v} = \mathbf{v}(p)$, $M(p)$ optimális értékek. Ha most alkalmazzuk a burkolótételt, akkor

$$\begin{aligned} \frac{dM(p)}{dp} &= \frac{\partial\left[f(\mathbf{x}) + u_1(g_1(\mathbf{x}) - p) + \sum_{i=2}^m u_i g_i(\mathbf{x}) + \mathbf{v}^T\mathbf{h}\right]}{\partial p}\Bigg|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} \\ &= -u_1(p)\big|_{\mathbf{x}=\mathbf{x}(p),\mathbf{u}=\mathbf{u}(p),\mathbf{v}=\mathbf{v}(p)} = -u_1 \end{aligned}$$

A többi Lagrange szorzó is hasonlóan értelmezhető. Az i-edik Lagrange szorzó a célfüggvénynek az i-edik feltétel jobboldala szerinti deriváltjának (-1)-szerese, tehát a jobboldal változásának a célfüggvényre kifejtett hatását fejezi ki.

39. Példa (paraméteres feladat)

Adott az alábbi paraméteres feltételes optimalizálási feladat. Végezzük el az érzékenységvizsgálatot a $p = 5$ paraméterérték esetén!

$$\begin{aligned} f(x_1, x_2, p) &= (x_1 - p)^2 + (x_2 - 2p)^2 \to \min! \\ g(x_1, x_2, p) &= 2x_1 + x_2 - p \leqq 0 \end{aligned}$$

ahol $p \in \mathbb{R}$ paraméter.

Megoldás

A $p = 5$ paraméterérték esetén az optimális megoldás: $x_1(p) = -1, x_2(p) = 7$, a Lagrange szorzó: $u(p) = 6$, a célfüggvény optimális értéke: $M(p) = 45$.

A Lagrange függvény

$$L(x_1, x_2, u, p) = (x_1 - p)^2 + (x_2 - 2p)^2 + u(2x_1 + x_2 - p)$$

A Lagrange függvény gradiense és Hesse mátrixa

$$\boldsymbol{\nabla} L = \begin{bmatrix} 2(x_1 - p) + 2u \\ 2(x_2 - 2p) + u \\ 2x_1 + x_2 - p \end{bmatrix}, \qquad \boldsymbol{\nabla}^2 L = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix}.$$

a) A célfüggvény érzékenységvizsgála:

A burkolótétel alkalmazásával végezzük. A számítás úgy történik, hogy a Lagrange függvényt deriváljuk p szerint és a döntési változók, ill. a Lagrange szorzó optimális értékét behelyettesítjük a deriváltba:

$$\begin{aligned} \frac{dM(p)}{dp} &= \left.\frac{\partial L(x_1, x_2, u, p)}{\partial p}\right|_{\mathbf{x}=\mathbf{x}(p), u=u(p)} = \\ &= \left.\frac{\partial}{\partial p}\left[(x_1 - p)^2 + (x_2 - 2p)^2 + u(2x_2 + x_2 - p)\right]\right|_{\mathbf{x}=\mathbf{x}(p), u=u(p)} = \\ &= -2(x_1 - p) - 4(x_2 - 2p) - u\big|_{x_1=-1, x_2=7, u=6} = 18. \end{aligned}$$

Ha a $p = 5$ számadat, mint paraméter megváltozik mondjuk ε nagyon kicsi értékkel, akkor a célfüggvény értéke 45-ről megközelítőleg $45 + 18\varepsilon$ értékre változik meg.

b) A döntési változók érzékenységvizsgálata:

Az ismert tétel alapján végezzük. A számítás úgy történik, hogy a Lagrange függvény Hesse mátrixa inverzének (-1)-szeresét megszorozzuk a Lagrange függvény gradiens vektorának p szerinti derivált vektorával:

$$\begin{aligned} \frac{d}{dp}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} &= -\left(\boldsymbol{\nabla}^2 L\right)^{-1}\Big|_{\mathbf{x}=\mathbf{x}(p), \mathbf{u}=\mathbf{u}(p), \mathbf{v}=\mathbf{v}(p)} \cdot \frac{\partial}{\partial p}\boldsymbol{\nabla} L\Big|_{\mathbf{x}=\mathbf{x}(p), \mathbf{u}=\mathbf{u}(p), \mathbf{v}=\mathbf{v}(p)} \\ &= -\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix}^{-1} \cdot \frac{\partial}{\partial p}\left.\begin{bmatrix} 2(x_1 - p) + 2u \\ 2(x_2 - 2p) + u \\ 2x_1 + x_2 - p \end{bmatrix}\right|_{x_1=-1, x_2=7, u=6} \\ &= \frac{1}{10}\begin{bmatrix} 1 & -2 & 4 \\ -2 & 4 & 2 \\ 4 & 2 & -4 \end{bmatrix}\begin{bmatrix} -2 \\ -4 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} \\ \frac{7}{5} \\ \frac{6}{5} \end{bmatrix}. \end{aligned}$$

Az eredmények értelmezése:

A $\frac{dx_1}{dp} = -\frac{1}{5} = -0.2$ értelmezése:

Ha a $p = 5$ számadat, mint paraméter megváltozik, mondjuk ε nagyon kicsi értékkel, akkor az első döntési változó értéke -1-ről megközelítőleg $-1 - 0.2\varepsilon$ értékre változik meg.

A $\frac{dx_2}{dp} = \frac{7}{5} = 1.4$ értelmezése:

Ha a $p = 5$ számadat, mint paraméter megváltozik, mondjuk ε nagyon kicsi értékkel, akkor a második döntési változó értéke 7-ről megközelítőleg $7 + 1.4\varepsilon$ értékre változik meg.

A $\frac{du}{dp} = \frac{6}{5} = 1.2$ értelmezése:

Ha a $p = 5$ számadat, mint paraméter megváltozik, mondjuk ε nagyon kicsi értékkel, akkor a Lagrange szorzó értéke 6-ról megközelítőleg $6 + 1.2\varepsilon$ értékre változik meg.

Meghatározhatjuk természetesen a p paraméter minden elemére is az optimális megoldást, majd az érzékenységvizsgálatot is. Az alábbiakban ezt végezzük el.

A KKT pont meghatározásához szükséges rendszer:

$$\begin{aligned}
2(x_1 - p) + 2u &= 0 \\
2(x_2 - 2p) + u &= 0 \\
2x_1 + x_2 - p &\leqq 0 \\
u(2x_1 + x_2 - p) &= 0 \\
u &\geqq 0
\end{aligned}$$

A rendszer megoldása a p paraméter függvényében:

$$\begin{aligned}
x_1(p) &= -\frac{1}{5}p \\
x_2(p) &= \frac{7}{5}p \\
u(p) &= \frac{6}{5}p \\
M(p) &= \frac{9}{5}p^2
\end{aligned}$$

A fenti megoldás optimális megoldás, mivel a szegélyezett Hesse mátrix feltételesen pozitív definit, sőt az aktualizált Lagrange függvény Hesse mátrixa pozitív definit.

Az érzékenységvizsgálat közvetlen deriválással az alábbiak szerint alakul.

a) A célfüggvény érzékenységvizsgála deriválással:

$$\frac{dM(p)}{dp} = \frac{18}{5}p$$

b) A döntési változók érzékenységvizsgálata deriválással:

$$\frac{d}{dp}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} \\ \frac{7}{5} \\ \frac{6}{5} \end{bmatrix}.$$

Feladat:
Az érzékenységvizsgálatnál megismert képletek begyakorlása miatt azokkal a képletekkel is végezze el az érzékenységvizsgálatot! Természetesen a fenti eredményeket fogja megkapni.

A példa kapcsán vizsgáljuk meg a Lagrange szorzó jelentését. Tekintsük a $p = 5$ paraméterhez tartozó optimalizálási feladatot:

$$\begin{aligned} f(x_1, x_2) &= (x_1 - 5)^2 + (x_2 - 10)^2 \to \min! \\ g(x_1, x_2) &= 2x_1 + x_2 - 5 \leqq 0 \end{aligned}$$

A feladat megoldása: a célfüggvény minimuma $f_{\min} = 45$, a Lagrange szorzó értéke $u = 6$. Mi a jelentése a Lagrange szorzónak, mit mutat az $u = 6$ érték? A feladat feltétele ekkor $2x_2 + x_2 - 5 \leqq 0$. A jobboldal értékének (a 0 értéknek) kis megváltozása a célfüggvény optimális értékében (45-ben) közel $u = 6$-szoros változást okoz, pontosabban, ha megváltozik a feltétel jobboldala, a 0 érték, mondjuk ε kis értékkel, akkor a célfüggvény értéke 45-ről megközelítőleg $45+(-6)\varepsilon$ értékre változik. Ha nő a jobboldal (a 0 érték), akkor a célfüggvény csökken és fordítva. Ha a feltételt $2x_1 + x_2 \leqq 5$ formában értelmezzük (a gyakorlatban ez a megszokottabb), akkor a fentebb említett jobboldal változása tulajdonképpen az 5 számérték megváltozását jelenti. Ha tehát az 5 jobboldal megváltozik ε kis értékkel $5+\varepsilon$ értékre, akkor a célfüggvény 45-ről megközelítőleg $45 + (-6)\varepsilon$ értékre változik.

Többparaméteres eset
A többparaméteres optimalizálási feladat:

$$\min \{ f(\mathbf{x}, \mathbf{p}) : \mathbf{g}(\mathbf{x}, \mathbf{p}) \leqq \mathbf{0}, \mathbf{h}(\mathbf{x}, \mathbf{p}) = \mathbf{0}, \mathbf{x} \in X \},$$

ahol $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ a paraméterek vektora. Az egyparaméteres esethez hasonlóan számíthatjuk a paraméterek megváltozásának hatását az optimum érték függvényre és a döntési változók, ill. a Lagrange szorzók optimális értékeire.

A célfüggvény érzékenységvizsgálata (burkolótétel) több paraméter esetén:

$$\frac{\partial M(\mathbf{p})}{\partial p_i} = \left. \frac{\partial L(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{p})}{\partial p_i} \right|_{\mathbf{x}=\mathbf{x}(\mathbf{p}), \mathbf{u}=\mathbf{u}(\mathbf{p}), \mathbf{v}=\mathbf{v}(\mathbf{p})} \qquad i = 1, 2, \ldots, k$$

A döntési változók és a Lagrange szorzók érzékenységvizsgálata több paraméter esetén:

$$\frac{\partial}{\partial p_i} \begin{bmatrix} \mathbf{x}(\mathbf{p}) \\ \mathbf{u}(\mathbf{p}) \\ \mathbf{v}(\mathbf{p}) \end{bmatrix} = - \left. \left(\boldsymbol{\nabla}^2 L \right)^{-1} \right|_{\mathbf{x}=\mathbf{x}(p), \mathbf{u}=\mathbf{u}(p), \mathbf{v}=\mathbf{v}(p)} \cdot \left. \frac{\partial}{\partial p_i} \boldsymbol{\nabla} L \right|_{\mathbf{x}=\mathbf{x}(p), \mathbf{u}=\mathbf{u}(p), \mathbf{v}=\mathbf{v}(p)} \qquad i = 1, 2, \ldots, k$$

40. Példa (Lagrange szorzó értelmezésére)
Oldjuk meg az alábbi paraméteres feltételes optimalizálási feladatot és értelmezzük a Lagrange szorzót!

$$\begin{aligned} x_1^2 + x_2^2 &\to \min! \\ -2x_1 - x_2 &\leqq -5 \end{aligned}$$

Megoldás

A feladat Lagrange függvénye, a Lagrange függvény döntési változók és a Lagrange szorzó szerint vett gradiense és Hesse mátrixa

$$L(x_1, x_2, u) = x_1^2 + x_2^2 + u(-2x_1 - x_2 + 5)$$

$$\nabla L = \begin{bmatrix} 2x_1 - 2u \\ 2x_2 - u \\ -2x_1 - x_2 + 5 \end{bmatrix}, \qquad \nabla^2 L = \begin{bmatrix} 2 & 0 & -2 \\ 0 & 2 & -1 \\ -2 & -1 & 0 \end{bmatrix}.$$

A KKT pont: $x_1 = 2, x_2 = 1$, a hozzá tartozó Lagrange szorzó $u = 2$ és ez egyben optimális megoldás is. Az optimális célfüggvényérték: $f_{\min} = 5$.

Az $u = 2$ jelentése:

Ha a feladat feltételének jobboldalát (-5) tekintjük paraméternek, akkor a célfüggvény optimális értékének változása a burkolótétel szerint $\frac{dM}{dp} = -u$. Ha tehát a jobboldal megváltozása nagyon kicsi ε szám, akkor a célfüggvény megváltozása közel $(-2)\varepsilon$. Ha $\leqq$ feltételünk van, akkor ismeretes, hogy a Lagrange szorzó általában pozitív (lehet zérus is). Tehát általában elmondható, hogy minimum feladat és $\leqq$ feltétel esetén, ha a jobboldal nő, akkor a célfüggvény értéke csökken és fordítva.

Gyakorlásképpen a fenti példa kapcsán levezetjük a fent közölt eredményt. Legyen a jobboldal a paraméter. A feladat ekkor

$$\begin{aligned} x_1^2 + x_2^2 &\to \min! \\ -2x_1 - x_2 - p &\leqq 0 \end{aligned}$$

Az optimális megoldás a paraméter függvényében

$$x_1 = -\frac{2}{5}p, \quad x_2 = -\frac{1}{5}p, \quad u = -\frac{2}{5}p, \quad M(p) = \frac{1}{5}p^2$$

A célfüggvény változása

$$\frac{dM(p)}{dp} = \frac{2}{5}p,$$

ami valóban az u Lagrange szorzó optimális értékének (-1)-szeresét adja, amely $p = -5$ esetén a korábban kiszámított $u = 2$ érték.

Feladat:

Végezze el az érzékenységvizsgálatot és értelmezze a Lagrange szorzót a feltételes optimalizálás 14. és 15. példájában!

41. példa:

Tekintsük újra a 26. példában bemutatott gazdasági problémát, amely a következő volt: "Legyen két termelési tényező, mondjuk tőke és munka. A termelés során a termelési tényezők felhasznált mennyiségét jelölje rendre x_1, x_2. A termelési tényezők egységára legyen $p_1 = 3, p_2 = 4$. Legyen továbbá adott az $f(x_1, x_2) = x_1^3 x_2^2$ ún. Cobb-Douglas-féle termelési függvény, amely a termelési tényezőkkel előállítható termékmennyiséget fejezi ki a termelési tényezők mennyiségének függvényében. Határozzuk meg a termelési tényezők azon optimális mennyiségét, amellyel a legkevesebb költséggel legalább $q = 8$ termékmennyiség állítható elő!"

Most tekintsük a példában paraméternek a termelési tényezők egységárait (p_1, p_2), a Cobb-Douglas-féle termelési függvényben szereplő kitevőket (a_1, a_2), valamint a termékmennyiségre vonatkozó korlátot (q).

Ekkor az alábbi paraméteres matematikai programozási feladat adódik:

$$\begin{aligned} p_1x_1 + p_2x_2 &\rightarrow \min! \\ x_1^{a_1}x_2^{a_2} &\geqq q \\ x_1, x_2 &\geqq 0 \end{aligned}$$

A $p_1 = 3, p_2 = 4, a_1 = 3, a_2 = 2, q = 8$ esetén ismert a megoldás, amely a következő: $x_1 = 2, x_2 = 1, u = \frac{1}{4}, \min = 10.$

Mielőtt az érzékenységvizsgálatot elvégezzük, értelmezzük a Lagrange szorzót. A feltétel minimum feladatra átírt formája $q - x_1^{a_1}x_2^{a_2} \leqq 0$. Ha a jobboldal (0) ε-nal megváltozik, akkor a 10 értékű célfüggvény megközelítőleg $10 - \frac{1}{4}\varepsilon$-ra változik. Más szavakkal, ha a $-q$ jobboldal ε-nal megváltozik, akkor a 10 értékű célfüggvény megközelítőleg $10 - \frac{1}{4}\varepsilon$-ra változik. Másképpen, ha az eredeti $\geqq$ feltétel jobboldala $(q = 8)$ ε-nal megváltozik, akkor a 10 értékű célfüggvény megközelítőleg $10 + \frac{1}{4}\varepsilon$-ra változik. Kiolvasható ebből az is, hogy ha nő az előírt termékmennyiség, akkor nő a felhasznált termelőeszközök költsége is és fordítva, az arányossági tényező megközelítőleg $\frac{1}{4}$.

Most végezzük el a célfüggvény, a döntési változók és a Lagrange szorzó érzékenységvizsgálatát a paraméterek függvényében!

Megoldás

A feladat Lagrange függvénye:

$$L(x_1, x_2, u, p_1, p_2, a_1, a_2, q) = p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})$$

A Lagrange függvény x_1, x_2 döntési változók és az u Lagrange szorzó szerint vett gradiense és Hesse mátrixa a következő:

$$\begin{aligned} \nabla L &= \begin{bmatrix} p_1 - ua_1x_2^{a_2}x_1^{a_1-1} \\ p_2 - ua_2x_1^{a_1}x_2^{a_2-1} \\ q - x_1^{a_1}x_2^{a_2} \end{bmatrix} \\ \nabla^2 L &= \begin{bmatrix} -ua_1x_2^{a_2}(a_1-1)x_1^{a_1-2} & -ua_1a_2x_1^{a_1-1}x_2^{a_2-1} & -a_1x_2^{a_2}x_1^{a_1-1} \\ -ua_1a_2x_1^{a_1-1}x_2^{a_2-1} & -ua_2x_1^{a_1}(a_2-1)x_2^{a_2-2} & -a_2x_1^{a_1}x_2^{a_2-1} \\ -a_1x_2^{a_2}x_1^{a_1-1} & -a_2x_1^{a_1}x_2^{a_2-1} & 0 \end{bmatrix} \end{aligned}$$

A célfüggvény érzékenységvizsgálatát a burkolótétel alkalmazásával végezzük, amely szerint a paraméterektől függő $M(p_1, p_2, a_1, a_2, q)$ optimális célfüggvény deriváltja a paraméterek szerint rendre az alábbi:

$$\begin{aligned} \frac{\partial M(p_1,p_2,a_1,a_2,q)}{\partial p_1} &= \frac{\partial L}{\partial p_1} = \frac{\partial}{\partial p_1}\left[p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})\right] = x_1 = 2, \\ \frac{\partial M(p_1,p_2,a_1,a_2,q)}{\partial p_2} &= \frac{\partial L}{\partial p_2} = \frac{\partial}{\partial p_2}\left[p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})\right] = x_2 = 1, \\ \frac{\partial M(p_1,p_2,a_1,a_2,q)}{\partial a_1} &= \frac{\partial L}{\partial a_1} = \frac{\partial}{\partial a_1}\left[p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})\right] = -u\,(\ln x_1)\,x_1^{a_1}x_2^{a_2} = -1.3863, \\ \frac{\partial M(p_1,p_2,a_1,a_2,q)}{\partial a_2} &= \frac{\partial L}{\partial a_2} = \frac{\partial}{\partial a_2}\left[p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})\right] = -u\,(\ln x_2)\,x_1^{a_1}x_2^{a_2} = 0, \\ \frac{\partial M(p_1,p_2,a_1,a_2,q)}{\partial q} &= \frac{\partial L}{\partial q} = \frac{\partial}{\partial q}\left[p_1x_1 + p_2x_2 + u(q - x_1^{a_1}x_2^{a_2})\right] = u = \frac{1}{4}. \end{aligned}$$

Most a döntési változók és a Lagrange szorzó érzékenységvizsgálatát végezzük el. Először szükségünk van a Lagrange függvény Hesse mátrixának inverzére. A $\mathbf{H}$ Hesse mátrixot egyszerű behelyettesítéssel kapjuk, a $\boldsymbol{\nabla}^2 L$ mátrixba az $x_1 = 2, x_2 = 1, u = \frac{1}{4}, p_1 = 3, p_2 = 4, a_1 = 3, a_2 = 2, q = 8$ értékeket kell behelyettesíteni. A Hesse mátrix és annak inverze a következő:

$$\mathbf{H} = \begin{bmatrix} -3 & -6 & -12 \\ -6 & -4 & -16 \\ -12 & -16 & 0 \end{bmatrix}, \qquad \mathbf{H}^{-1} = \begin{bmatrix} \frac{4}{15} & -\frac{1}{5} & -\frac{1}{20} \\ -\frac{1}{5} & \frac{3}{20} & -\frac{1}{40} \\ -\frac{1}{20} & -\frac{1}{40} & \frac{1}{40} \end{bmatrix}.$$

Ezután ki kell számolni a $\boldsymbol{\nabla} L$ vektornak az egyes paraméterek szerinti deriváltját. Az öt paraméter közül részletesen csak az a_1 szerint végezzük el az érzékenységvizsgálatot, a többi szerint csak közöljük a végeredményt.

$$\frac{\partial}{\partial a_1}\boldsymbol{\nabla} L = \frac{\partial}{\partial a_1}\begin{bmatrix} p_1 - u a_1 x_2^{a_2} x_1^{a_1-1} \\ p_2 - u a_2 x_1^{a_1} x_2^{a_2-1} \\ q - x_1^{a_1} x_2^{a_2} \end{bmatrix} = \begin{bmatrix} -u x_2^{a_2} x_1^{a_1-1} - u a_1 (\ln x_1) x_2^{a_2} x_1^{a_1-1} \\ -u a_2 (\ln x_1) x_1^{a_1} x_2^{a_2-1} \\ -(\ln x_1) x_1^{a_1} x_2^{a_2} \end{bmatrix} = \begin{bmatrix} -3.0794 \\ -2.7726 \\ -5.5452 \end{bmatrix}.$$

$$\frac{\partial}{\partial a_1}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = -\mathbf{H}^{-1} \cdot \frac{\partial}{\partial a_1}\boldsymbol{\nabla} L = -\begin{bmatrix} \frac{4}{15} & -\frac{1}{5} & -\frac{1}{20} \\ -\frac{1}{5} & \frac{3}{20} & -\frac{1}{40} \\ -\frac{1}{20} & -\frac{1}{40} & \frac{1}{40} \end{bmatrix} \begin{bmatrix} 3.0794 \\ 2.7726 \\ -5.5452 \end{bmatrix} = \begin{bmatrix} -0.54391 \\ 0.06136 \\ 0.36192 \end{bmatrix}.$$

A további paraméterek szerinti deriváltak a következők:

$$\frac{\partial}{\partial a_2}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ \frac{3}{10} \\ -\frac{1}{20} \end{bmatrix}, \qquad \frac{\partial}{\partial q}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} \frac{1}{20} \\ \frac{1}{40} \\ -\frac{1}{40} \end{bmatrix},$$

$$\frac{\partial}{\partial p_1}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} -\frac{4}{15} \\ \frac{1}{5} \\ \frac{1}{20} \end{bmatrix}, \qquad \frac{\partial}{\partial p_2}\begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} \frac{1}{5} \\ -\frac{3}{20} \\ \frac{1}{40} \end{bmatrix}.$$

Végezetül megadjuk néhány érzékenységvizsgálati eredmény gazdasági tartalmát.

A $\frac{\partial M}{\partial p_1} = 2$ jelentése:
Ha csak a p_1 paraméter, azaz az első termelési tényező ára változik meg, mondjuk ε nagyon kicsi értékkel, akkor a minimális termelési költség 10-ről megközelítőleg $10 + 2\varepsilon$ értékre változik.

A $\frac{\partial x_2}{\partial p_1} = \frac{1}{5}$ jelentése:
Ha csak a p_1 paraméter, azaz az első termelési tényező ára változik meg, mondjuk ε nagyon kicsi értékkel, akkor a második termelési tényezőből felhasznált mennyiség optimális értéke 1-ről megközelítőleg $1 + \frac{1}{5}\varepsilon$ értékre változik.

A $\frac{\partial x_1}{\partial a_2} = -\frac{2}{5}$ jelentése:
Ha csak az a_2 paraméter, azaz a termelési függvényben a második termelési tényező kitevője változik meg, mondjuk ε nagyon kicsi értékkel, akkor az első termelési tényezőből felhasznált mennyiség optimális értéke 2-ről megközelítőleg $2 - \frac{2}{5}\varepsilon$ értékre változik.

A $\frac{\partial x_1}{\partial q} = \frac{1}{20}$ jelentése:

Ha csak a q paraméter, azaz a termékmennyiségre előírt alsó korlát változik meg, mondjuk ε nagyon kicsi értékkel, akkor az első termelési tényezőből felhasznált mennyiség optimális értéke 2-ről megközelítőleg $2 + \frac{1}{20}\varepsilon$ értékre változik.

A $\frac{\partial u}{\partial p_2} = \frac{1}{40}$ jelentése:

Ha csak a p_2 paraméter, azaz a második termelési tényező ára változik meg, mondjuk ε nagyon kicsi értékkel, akkor a Lagrange szorzó optimális értéke $\frac{1}{4}$-ről megközelítőleg $\frac{1}{4} + \frac{1}{40}\varepsilon$ értékre változik. Mit jelent a Lagrange szorzó változása? Ha tehát a p_2 paraméter megváltozik, akkor a Lagrange szorzó megközelítőleg $\frac{1}{4} + \frac{1}{40}\varepsilon$. A Lagrange szorzó értelmezéséből pedig tudjuk, hogy ha a q megváltozik λ-val, akkor a célfüggvény optimális értéke megközelítőleg $10 + \frac{1}{4}\lambda$. Ezeket összefoglalva a következőket mondhatjuk. A p_2 paraméter megváltozott értékéhez (a változás ε) tartozó optimalizálási feladatban, ha a q megváltozik λ-val, akkor a célfüggvény optimális értéke megközelítőleg $10 + \left(\frac{1}{4} + \frac{1}{40}\varepsilon\right)\lambda$.