7. előadás

2020. március 30.

$\boxed{\mathbb{R}^n \to \mathbb{R} \text{ típusú függvények szélsőértékei}}$

Amint azt már az „egyváltozós analízisben” is hangsúlyoztuk, a matematikai alkalmazások egyik legfontosabb fejezete a függvények szélsőértékeinek a vizsgálata. Valós-valós függvényeknél megismerkedtünk az abszolút- és a lokális szélsőértékek fogalmával, a lokális szélsőértékekre vonatkozó szükséges feltétellel, valamint több elégséges feltétellel. Most ezeket az ismereteket terjesztjük ki $\mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ típusú függvényekre.

- **Fogalmak**

Az egyváltozós esetben bevezetett fogalmak minden további nehézség nélkül átvihetők a többváltozós függvényekre.

Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ egy adott függvény.

Azt mondjuk, hogy az $a \in \mathcal{D}_f$ pontban az f függvénynek *abszolút maximuma van* (vagy másképp fogalmazva az a pont az f függvénynek *abszolút maximumhelye*), ha az $f(x) \le f(a)$ egyenlőtlenség igaz $\forall\, x \in \mathcal{D}_f$ pontban. Ekkor az $f(a)$ függvényértéket az f függvény *abszolút maximumának* nevezzük.

Analóg módon értelmezzük az *abszolút minimumhely* és az *abszolút minimum* fogalmát.

Az abszolút maximumhelyet, illetve az abszolút minimumhelyet közösen *abszolút szélsőértékhelynek*, az abszolút maximumot, illetve az abszolút minimumot közösen *abszolút szélsőértéknek* nevezzük.

Minden további nehézség nélkül definiálhatjuk ezeknek a fogalmaknak a *lokális* változatait. Például:

1. definíció. Azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ függvénynek az $a \in \mathcal{D}_f$ pontban *lokális maximuma* van (más szóval az a pont *lokális maximumhely*), ha

$$\exists\, K(a): \quad \forall\, x \in K(a) \cap \mathcal{D}_f \ \text{ pontban } \ f(x) \le f(a).$$

Ekkor az $f(a)$ függvényértéket az f függvény *lokális maximumának* nevezzük.

A *lokális minimumhely* és a *lokális minimum* definíciója hasonló. Ha ui. $a \in \mathcal{D}_f$ és egy $K(a)$ környezet esetén igaz az

$$f(x) \ge f(a) \quad \big(x \in K(a) \cap \mathcal{D}_f\big)$$

becslés, akkor az $f(a)$ függvényértéket az f függvény *lokális minimumának*, az a pontot pedig az f *lokális minimumhelyének* nevezzük. (Más szóval ekkor az f függvénynek az a pontban *lokális minimuma van.*)

A lokális maximumhelyet, illetve a lokális minimumhelyet közösen *lokális szélsőértékhelynek*, a lokális maximumot, illetve a lokális minimumot közösen *lokális szélsőértéknek* nevezzük.

• Szükséges feltétel a lokális szélsőértékre

A valós-valós függvények lokális szélsőértékeire vonatkozó szükséges feltétel *lényeges* nehézség nélkül átvihető az $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre.

1. tétel. (Elsőrendű szükséges feltétel a lokális szélsőértékre). Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy

- $f \in D\{a\}$ és
- az f függvénynek az a pontban lokális szélsőértéke van.

Ekkor $f'(a) = \mathbf{0}$, azaz

$$f'(a) = \big(\partial_1 f(a), \partial_2 f(a), \ldots, \partial_n f(a)\big) = (0, 0, \ldots, 0).$$

Az állítás bizonyításaként elég arra gondolni, hogy ha az f függvénynek az a pontban például lokális minimuma van és $i = 1, 2, \ldots, n$ egy rögzített index, akkor a

$$g_i(t) = f\big(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n\big) \quad \big(t \in K(a_i)\big)$$

valós-valós függvénynek a $t = a_i$ pontban lokális minimuma van. Mivel $f \in D\{a\}$, ezért $g_i \in D\{a_i\}$ és $g_i'(a_i) = 0$, ami éppen azt jelenti, hogy $\partial_i f(a) = 0$.

2. definíció. Az $a \in \operatorname{int} \mathcal{D}_f$ pont a deriválható $f \in \mathbb{R}^n \to \mathbb{R}$ függvény *stacionárius pontja*, ha $f'(a) = \mathbf{0} \in \mathbb{R}^n$.

Megjegyzések

1. A tétel tehát azt állítja, hogy a lokális szélsőértékhelyek szükségképpen a függvény stacionárius pontjai. Az $f'(a) = \mathbf{0}$ azonban csak *szükséges*, de *nem elégséges* feltétel a lokális szélsőértékre. Az $n = 1$ esetben például az $f(x) := x^3$ $(x \in \mathbb{R})$ függvénynek az $a = 0$ pont stacionárius pontja, mivel $f'(0) = 0$, de ez a pont nyilván nem lokális szélsőértékhely.

2. Mivel $f \in \mathbb{R}^n \to \mathbb{R}$ esetén $f'(x) = \big(\partial_1 f(x), \ldots, \partial_n f(x)\big)$, ahol $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, ezért f stacionárius pontjai az $x_1, x_2, \ldots, x_n$ ismeretlenekre vonatkozó

$$\begin{aligned} \partial_1 f(x_1, \ldots, x_n) &= 0, \\ \partial_2 f(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ \partial_n f(x_1, \ldots, x_n) &= 0 \end{aligned}$$

egyenletrendszer megoldásai. Az így kapott $(x_1, \ldots, x_n)$ pont(ok)ban *lehet(nek)* tehát az f függvény lokális szélsőértékhelyei.

• Másodrendű elégséges feltétel a lokális szélsőértékre

A fentiek alapján a stacionárius pontok között lehetnek olyanok, amelyekben a függvénynek nincs lokális szélsőértéke. Fontos kérdés tehát annak eldöntése, hogy egy stacionárius hely vajon lokális szélsőértékhely-e. Ennek eldöntéséhez a valós-valós esetben az elsőrendű- vagy a másodrendű elégséges feltételt használtuk. Ez utóbbi állítás szerint, ha $f \in D^2\{a\}$, $f'(a) = 0$ és $f''(a) > 0$ (illetve $f''(a) < 0$), akkor az f függvénynek az a pontban lokális minimuma (illetve lokális maximuma) van. Világos, hogy ennek az állításnak a bizonyításánál alkalmazott utat $\mathbb{R}^n \to \mathbb{R}$ típusú függvényre nem használhatjuk.

A többváltozós esetben a kiindulópontunk *alapötlete* a Peano-féle maradéktagos Taylor-formula alkalmazása. Idézzük fel ezt az állítást: *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ és az $a \in \operatorname{int} \mathcal{D}_f$ pontban $f \in D^2\{a\}$. Ekkor van olyan $\varepsilon \in \mathbb{R}^n \to \mathbb{R}$, a $\lim_{\mathbf{0}} \varepsilon = 0$ feltételnek eleget tevő függvény, hogy*

$$f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\cdot \|h\|^2 \quad \big(h \in \mathbb{R}^n,\ a+h \in \mathcal{D}_f\big).$$

Tegyük fel, hogy az a pont az f függvény stacionárius pontja, azaz $f'(a) = \mathbf{0}$. Ekkor $\langle f'(a), h\rangle = 0$ minden $h \in \mathbb{R}^n$ vektorra, tehát

$$(*) \qquad f(a+h) - f(a) = \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\cdot \|h\|^2 \quad \big(h \in \mathbb{R}^n,\ a+h \in \mathcal{D}_f\big).$$

A következő fontos *észrevétel* az, hogy ha $\langle f''(a)h, h\rangle > 0$ minden $h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ vektorra, ekkor a $(*)$ egyenlőség jobb oldala is pozitív. (Fogadjuk el bizonyítás nélkül ezt az állítás.) Ugyanez igaz tehát $(*)$ bal oldalára is, tehát

$$f(a+h) - f(a) > 0, \text{ vagyis } f(a+h) > f(a), \quad \text{ha } h \in \mathbb{R}^n \text{ és } a+h \in \mathcal{D}_f.$$

Ez azt jelenti, hogy az f függvénynek az a pontban lokális minimuma van.

Ha azt tesszük fel, hogy $\langle f''(a)h, h\rangle < 0$, akkor az f függvénynek az a pontban lokális maximuma van.

A fentieket a következő állításban foglaljuk össze:

2. tétel. Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy $f \in D^2\{a\}$, $f'(a) = \mathbf{0}$ és

$$(**) \qquad \langle f''(a)h, h\rangle > 0 \ \ (\text{illetve } \ \langle f''(a)h, h\rangle < 0), \quad \text{ha } \ h \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Ekkor az f függvénynek az a pontban lokális minimuma (illetve lokális maximuma) van.

Világos, hogy a $(**)$ feltételek nehezen ellenőrizhetők. A gyakorlatban már jól használható ekvivales átfogalmazásukhoz azonban további fogalmak bevezetésére lesz szükségünk. Az $f \in D^2\{a\}$ feltétel miatt a Young-tétel szerint az $f''(a) \in \mathbb{R}^{n\times n}$ Hesse-féle mátrix *szimmetrikus*, ezért a továbbiakban $f''(a)$ helyett egy tetszőleges $A \in \mathbb{R}^{n\times n}$ *szimmetrikus* mátrixot fogunk tekinteni.

2. definíció. Legyen $n \in \mathbb{N}$ és $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ egy szimmetrikus mátrix. Ekkor a

$$Q(h) := \langle A\cdot h, h\rangle = \sum_{i,j=1}^{n} a_{ij}h_ih_j, \qquad h = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \in \mathbb{R}^n$$

függvényt az A mátrix által meghatározott *kvadratikus alaknak* nevezzük.

3. definíció. Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix, illetve a $Q(h) = \langle A \cdot h, h\rangle$ $\big(h \in \mathbb{R}^n\big)$ kvadratikus alak

- *pozitív definit*, ha $\forall\, h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ esetén $Q(h) > 0$;
- *negatív definit*, ha $\forall\, h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ esetén $Q(h) < 0$;
- *indefinit*, ha Q pozitív és negatív értéket is felvesz.

Megjegyzés. Ha $f \in D^2\{a\}$, akkor a fentiek alapján már értelmeztük azt is, hogy az $f''(a) \in \mathbb{R}^{n\times n}$ Hesse-féle mátrix pozitív, illetve negatív definit.

Egy mátrix, illetve kvadratikus alak definitségének az eldöntése nem egyszerű feladat. A következő állításban a gyakorlatban jól használható eredményt fogalmazunk meg.

3. tétel. (Sylvester-kritérium). Legyen $n \in \mathbb{N}$, $A \in \mathbb{R}^{n\times n}$ egy szimmetrikus mátrix és $Q(h) = \langle A \cdot h, h\rangle$ $\big(h \in \mathbb{R}^n\big)$ az A által meghatározott kvadratikus alak. Jelölje

$$D_k := \det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \dots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix} \qquad (k = 1, 2, \dots, n)$$

az A mátrix „bal felső sarokmátrixának” a determinánsát.

Ekkor az A mátrix, illetve a Q kvadratikus alak

- *pozitív definit* $\iff$ ha $D_k > 0$ $\;(\forall\, k = 1, 2, \dots, n)$,
- *negatív definit* $\iff$ ha $(-1)^k D_k > 0$ $\;(\forall\, k = 1, 2, \dots, n)$.

Megjegyzés. A gyakorlaton megmutatjuk, hogy ez az algebrai jellegű tétel az $n = 2$ esetben elemi úton könnyen belátható.

4. tétel. (Másodrendű elégséges feltétel a lokális szélsőértékre $\mathbb{R}^n \to \mathbb{R}$ függvényekre.) Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy

- $f \in D^2\{a\}$,
- az a pont az f függvény stacionárius pontja, azaz $f'(a) = \mathbf{0}$,
- az $f''(a)$ Hesse-féle mátrix pozitív (negatív) definit.

Ekkor az f függvénynek az a pontban lokális minimuma (maximuma) van.

Megjegyzés. A Sylvester-kritérium segítségével vizsgálhatjuk az $f''(a)$ Hesse-mátrix pozitív, illetve negatív definitségét. Ha a Sylvester-kritérium feltételei nem teljesülnek, akkor a fenti elégséges feltétel *nem használható*. Ilyenkor *egyedi vizsgálatokkal* lehet eldönteni, hogy egy stacionárius pont vajon lokális szélsőértékhely-e vagy sem.

Kétváltozós függvények esetén azonban van egy egyszű elégséges feltétel arra az esetre is, amikor a függvénynek *nincs lokális szélsőértékhelye* egy stacionárius pontan. Ezért célszerű megmegjegyezni ezt a speciális esetet.

5. tétel. Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy

(a) $f \in D^2\{a\}$,

(b) $f'(a) = \big(\partial_1 f(a), \partial_2 f(a)\big) = (0, 0)$.

Ekkor:

1^0 Ha

$$\det f''(a) = \det \begin{bmatrix} \partial_{11} f(a) & \partial_{12} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) \end{bmatrix} > 0$$

és $\partial_{11} f(a) > 0$ [illetve $\partial_{11} f(a) < 0$], akkor az f függvénynek a-ban lokális minimuma [illetve maximuma] van.

2^0 Ha $\det f''(a) < 0$, akkor f-nek a-ban nincs lokális szélsőértéke (az a pont nyeregpont).

Megjegyzés. Világos, hogy az 1^o állítás a Sylvester-kritériumból következő 4. tétel az $n = 2$ speciális esetben.

Ha $n > 2$, akkor nincs a 2^o állításnak megfelelő „szép" elégséges feltétel.

- **Abszolút szélsőértékek**

Az a megfigyelés, hogy a lokális szélsőértékhelyeken a függvény deriváltja eltűnik (feltéve, hogy létezik), lehetővé tette olyan f egyváltozós függvény abszolút szélsőértékeinek meghatározását, amelyik *folytonos egy korlátos és zárt $[a, b]$ intervallumban*, és differenciálható annak (a, b) belsejében. Ekkor ui. f-nek van legnagyobb és legkisebb értéke a *Weierstrass-tétel* szerint. Ha f ezek valamelyikét egy c pontban veszi fel, akkor vagy $c = a$, vagy $c = b$, vagy pedig $c \in (a, b)$. Ez utóbbi esetben lokális szélsőértékről van szó, és így $f'(c) = 0$. Ha tehát megkeressük az összes olyan $c \in (a, b)$ pontot, amelyben f' eltűnik, akkor biztos, hogy az abszolút szélsőértékhelyek ezek közül, valamint az a és a b végpontok közül kerülnek ki. Például az abszolút maximumhelyet úgy határozzuk meg, hogy kiszámítjuk f értékeit ezekben a pontokban (nem feledkezve meg az a és b végpontokról sem), és kiválasztjuk azokat, amelyekben f értéke a legnagyobb.

Ezt a gondolatmenetet könnyen általánosíthatjuk többváltozós függvényekre.

6. tétel. Legyen $H \subset \mathbb{R}^n$ korlátos és zárt halmaz. Tegyük fel, hogy az $f : H \to \mathbb{R}$ függvény folytonos és léteznek a parciális deriváltjai H belsejének minden pontjában. Ekkor f a legnagyobb (legkisebb) értékét vagy a H halmaz határán veszi fel, vagy pedig egy olyan $a \in \operatorname{int} \mathcal{D}_f$ belső pontban, ahol $\partial_i f(a) = 0$ teljesül minden $i = 1, 2, \ldots, n$ indexre.

A fentieket a következő példával illusztráljuk.

Példa. Határozzuk meg az

$$f(x, y) := xy(x^2 + y^2 - 1) \quad \big((x, y) \in \mathbb{R}^2\big)$$

függvény abszolút szélsőértékhelyeit és abszolút szélsőértékeit a

$$H := \big\{(x, y) \mid x^2 + y^2 \le 1\big\} \subset \mathbb{R}^2$$

zárt körlapon.

Megoldás. A $H \subset \mathbb{R}^2$ halmaz korlátos és zárt, továbbá az f polinomfüggvény folytonos a H halmazon. Ezért Weierstrass tétele szerint f-nek a H halmazon van legnagyobb és legkisebb értéke. Az abszolút szélsőértékhelyek vagy a körlap határán (ez az $x^2 + y^2 = 1$ egyenletű körvonal), vagy pedig a H halmaz belsejében helyezkednek el.

Világos, hogy az f függvény értéke nulla a H halmaz határának minden pontjában. Mivel $(x, y) \in \operatorname{int} H$ (azaz $x^2 + y^2 < 1$), $x > 0$, $y < 0$ esetén f pozitív, továbbá $x > 0$ és $y > 0$ esetén f negatív, ezért f abszolút szélsőértékhelyei szükségképpen H belsejében helyezkednek el, és az abszolút szélsőértékek nullától különbözőek.

Legyen $(x, y) \in \operatorname{int} H$ egy olyan pont, ahol f-nek abszolút szélsőértéke van. Ez a pont egyúttal lokális szélsőértékhely is. Az

$$f(x, y) := xy(x^2 + y^2 - 1) = x^3y + xy^3 - xy \quad \big((x, y) \in \mathbb{R}^2\big)$$

egy polinomfüggvény, ezért $f \in D(\operatorname{int} H)$. A lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel szerint a szóban forgó helyen a parciális deriváltak 0-val egyenlőek:

$$\begin{aligned} \partial_x f(x, y) &= 3x^2y + y^3 - y = y(3x^2 + y^2 - 1) = 0, \\ \partial_y f(x, y) &= x^3 + 3xy^2 - x = x(x^2 + 3y^2 - 1) = 0. \end{aligned}$$

Ha $y = 0$, akkor a második egyenletből $x = 0$ adódik (hiszen $|x| < 1$ miatt $x^2 - 1 \neq 0$). Az origóban a függvény értéke nulla, ezért a fentiek alapján a $(0, 0)$ pont *nem* abszolút szélsőértékhely. Így $x \neq 0$, $y \neq 0$, tehát

$$\left.\begin{aligned} 3x^2 + y^2 - 1 = 0 \\ x^2 + 3y^2 - 1 = 0 \end{aligned}\right\} \quad \Longleftrightarrow \quad 3x^2 + y^2 - 1 = x^2 + 3y^2 - 1 \quad \Longleftrightarrow \quad x^2 = y^2 \quad \Longleftrightarrow$$

$$x^2 = \frac{1}{4} \text{ és } y^2 = \frac{1}{4} \quad \Longleftrightarrow \quad x = \pm\frac{1}{2} \text{ és } y = \pm\frac{1}{2}.$$

A lehetséges szélsőértékhelyek tehát az

$$\Big(\frac{1}{2}, \frac{1}{2}\Big), \quad \Big(\frac{1}{2}, -\frac{1}{2}\Big), \quad \Big(-\frac{1}{2}, \frac{1}{2}\Big), \quad \Big(-\frac{1}{2}, -\frac{1}{2}\Big)$$

pontok. Az itt felvett helyettesítési értékek:

$$f\Big(\frac{1}{2}, \frac{1}{2}\Big) = -\frac{1}{8}, \quad f\Big(\frac{1}{2}, -\frac{1}{2}\Big) = \frac{1}{8}, \quad f\Big(-\frac{1}{2}, \frac{1}{2}\Big) = \frac{1}{8}, \quad f\Big(-\frac{1}{2}, -\frac{1}{2}\Big) = -\frac{1}{8}.$$

A függvényértékeket összehasonlítva azt kaptuk, hogy a H halmazon az f függvény legnagyobb értéke $\frac{1}{8}$, és ezt az értéket az $(\frac{1}{2}, -\frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$ pontokban veszi fel. Az abszolút minimumhelyek pedig az $(\frac{1}{2}, \frac{1}{2})$, $(-\frac{1}{2}, -\frac{1}{2})$ pontok, és az abszolút minimum $-\frac{1}{8}$. ■