10. előadás

2020. november 16.

Taylor-polinomok. Taylor-formula

Emlékeztető. Az egyváltozós analízisben a Taylor-polinomok motivációjaként azt a problémát vetettük fel, hogy egy adott „bonyolult” függvényt vajon meg lehet-e közelíteni egyszerű szerkezetű függvényekkel, például a jól kezelhető és könnyen számolható polinomokkal.

Megmutattuk, hogy egy függvény akkor és csak akkor differenciálható egy pontban, ha annak a pontnak egy környezetében a függvény lokálisan jól közelíthető elsőfokú polinommal. Ennek általánosításaként jutottunk el ahhoz az eredményhez, hogy ha egy függvény (mondjuk) m-szer deriválható a szóban forgó pont egy környezetében, akkor ebben a környezetben a függvény jól közelíthető egy alkalmasan megválasztott $(m-1)$-edfokú polinommal, nevezetesen a függvény adott ponthoz tartozó *Taylor-polinomjával*. Ezt állítja a *Taylor-formulára* vonatkozó alábbi tétel:

Ha $f \in \mathbb{R} \to \mathbb{R}$, $m \in \mathbb{N}$ és egy $K(a) \subset \mathcal{D}_f$ környezetben $f \in D^m\big(K(a)\big)$, akkor minden $h > 0$ $(a+h \in K(a))$ számhoz létezik olyan $\nu \in (0,1)$, amelyre az

$$f(a+h) = \sum_{k=0}^{m-1} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(m)}(a+\nu h)}{m!} h^m \tag{1}$$

egyenlőség teljesül.

A legfeljebb $(m-1)$-edfokú

$$T_{a,m-1}(f,h) := \sum_{k=0}^{m-1} \frac{f^{(k)}(a)}{k!} h^k \qquad (h \in \mathbb{R})$$

polinom az f függvény *a ponthoz tartozó $(m-1)$-edik Taylor-polinomja*, az (1) egyenlőség jobb oldalán álló összeg utolsó tagja pedig a *Taylor-formula Lagrange-féle maradéktagja*. Az elnevezést az a tény indokolja, hogy ez a tag $T_{a,m-1}(f,h)$-hoz képest kicsi, ha h közel van 0-hoz, azaz

$$f(a+h) \approx T_{a,m-1}(f,h) = \sum_{k=0}^{m-1} \frac{f^{(k)}(a)}{k!} h^k, \quad \text{ha } h \approx 0;$$

vagyis egy elég sima függvény az a pont környezetében lokálisan jól közelíthető egy elég magas fokszámú polinommal, nevezetesen az f függvény a ponthoz tartozó Taylor-polinomjával. $\square$

Most a Taylor-formulát általánosítjuk n-változós valós értékű függvényekre. Látni fogjuk, hogy a többváltozós esetben a Taylor-polinomok függvényértékek közelítése mellett (az egyváltozós esettel ellentétben) a szélsőérték-problémák vizsgálatánál is fontos szerepet játszanak.

Figyeljük meg, hogy a fenti egyváltozós eredmény $f \in \mathbb{R}^n \to \mathbb{R}$ függvényekre vonatkozó általánosításához *egyrészt* értelmezni kellene $h \in \mathbb{R}^n$ esetén a h^k hatványokat, *másrészt* az $f^{(k)}(a)$ deriváltakat, ha $k = 1, 2, \ldots, m$.

A továbbiakban a Taylor-formula kiterjesztést kétváltozós függvényekre és másodfokú Taylor-polinomokra *vázoljuk*. Legyen tehát $f \in \mathbb{R}^2 \to \mathbb{R}$ egy adott függvény, $a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f$ és $h = (h_1, h_2) \in \mathbb{R}^2$. Tegyük fel egyelőre még azt is, hogy egy $K(a) \subset \mathcal{D}_f$ környezetben az f függvény kétszer *folytonosan deriválható*, azaz $f \in C^2\big(K(a)\big)$. Vegyük a síkon az a és az $a+h \in K(a)$ pontokat összekötő egyenes $a+th$ $(t \in \mathbb{R})$ pontjait. Tekintsük a

$$\boxed{F(t) := f(a+th) \quad \big(t \in \mathbb{R}\big)}$$

valós-valós függvényt. Mivel $f \in C^2\big(K(a)\big)$, ezért az összetett függvény deriválására vonatkozó tételből következik, hogy a F és a F' függvény folytonos a $[0,1]$ intervallumon és F' differencálható $(0,1)$-en.

Alkalmazzuk a F függvényre az (1) alatti Taylor-formulát a $[0,1]$ intervallumon: létezik tehát olyan $\nu \in (0,1)$, hogy

$$F(1) = F(0) + F'(0) + \frac{1}{2!}F''(\nu).$$

Most kiszámítjuk az $F(0)$, az $F'(0)$ és az $F''(\nu)$ értékeket. Mivel

$$F(t) = f(a+th) = f\big(a_1 + th_1, a_2 + th_2\big),$$

ezért

$$F(0) = f(a).$$

Az összetett függvény deriválására vonatkozó tétel (a láncszabály) alapján

$$F'(t) = \partial_1 f(a+th) \cdot h_1 + \partial_2 f(a+th) \cdot h_2 = \langle f'(a+th), h\rangle,$$

következésképpen

$$F'(0) = \langle f'(a), h\rangle.$$

A F függvény második deriváltja:

$$\begin{aligned} F''(t) &= (\partial_1 f(a+th) \cdot h_1 + \partial_2 f(a+th) \cdot h_2)' = \\ &= \Big(\partial_{11} f(a+th) \cdot h_1 + \partial_{12} f(a+th) \cdot h_2\Big) \cdot h_1 + \\ &\quad + \Big(\partial_{21} f(a+th) \cdot h_1 + \partial_{22} f(a+th) \cdot h_2\Big) \cdot h_2 = \end{aligned}$$

$$= \partial_{11} f(a+th) \cdot h_1^2 + \partial_{12} f(a+th) \cdot h_1 h_2 + \partial_{21} f(a+th) \cdot h_1 h_2 + \partial_{22} f(a+th) \cdot h_2^2 =$$
$$= \langle f''(a+th) \cdot h, h\rangle.$$

(Itt $f''(a+th)$ a (2×2)-es Hesse-féle mátrix, $h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$ oszlopvektor, ezért $f''(a+th)\cdot h$ egy kétdimenziós oszlopvektor.) Így

$$F''(\nu) = \langle f''(a+\nu h) \cdot h, h\rangle.$$

A feltételeinkből következik, hogy az $f'' \in \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ függvény folytonos az $a \in \mathcal{D}_{\mathbf{f}}$ pontban, ezért

$$f''(a+\nu h) = f''(a) + \eta(h),$$

ahol $\eta = \big[\eta_{ij}\big] \in \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ egy olyan függvény, amelyre $\lim\limits_{\mathbf{0}} \eta = \mathbf{0}$.

A fentiek alapján tehát

$$F(1) = f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2!}\langle f''(a)\cdot h, h\rangle + \tfrac{1}{2!}\langle \eta(h)\cdot h, h\rangle.$$

Az utolsó tagot még tovább alakítjuk:

$$\langle \eta(h)\cdot h, h\rangle = \sum_{i,j=1}^{2} \eta_{ij}(h)\cdot h_i h_j = \|h\|^2 \cdot \sum_{i,j=1}^{2} \eta_{ij}(h)\cdot \frac{h_i h_j}{\|h\|^2}.$$

Mivel $\dfrac{|h_i h_j|}{\|h\|^2} \leq M$ $(i,j=1,2)$ és $\lim\limits_{h\to\mathbf{0}} \eta_{ij} = 0$, ezért

$$\tfrac{1}{2!}\langle \eta(h)\cdot h, h\rangle = \varepsilon(h)\cdot \|h\|^2,$$

ahol $\varepsilon \in \mathbb{R}^2 \to \mathbb{R}$ olyan függvény, amelyre $\lim\limits_{\mathbf{0}} \varepsilon = 0$ teljesül.

Az eddigieket összefoglalva azt kaptuk, hogy van olyan $\varepsilon \in \mathbb{R}^2 \to \mathbb{R}$, a $\lim\limits_{\mathbf{0}} \varepsilon = 0$ feltételt kielégítő függvény, amelyre

$$\boxed{f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\|h\|^2}. \tag{2}$$

A fenti gondolatmenetet követve viszonylag egyszerűen belátható, hogy az (2) képlet tetszőleges $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 < n \in \mathbb{N})$ függvényre is teljesül. Nehezebb már annak a bizonyítása, hogy az állítás az $f \in C^2\big(K(a)\big)$ helyett az $f \in D^2\{a\}$ feltétel mellet is igaz.

1. tétel. (Taylor-formula a Peano-féle maradéktaggal.) *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \leq n \in \mathbb{N})$ 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 $\lim\limits_{\mathbf{0}} \varepsilon = 0$ feltételnek eleget tevő függvény, hogy*

$$f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\cdot\|h\|^2 \quad \big(h \in \mathbb{R}^n,\ a+h \in \mathcal{D}_f\big). \tag{3}$$

A jobb oldal első három tagjának az összegét (ez egy n-változós legfeljebb másodfokú polinom) az f függvény *a ponthoz tartozó második* ***Taylor-polinom****jának* nevezzük, és így jelöljük:

$$T_{a,m}(f,h) := f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle =$$

$$f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2}\left\langle \begin{bmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \dots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \dots & \partial_{2n} f(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \dots & \partial_{nn} f(a) \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \right\rangle =$$

$$= f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij} f(a) h_i h_j.$$

A (3) képletben az $\varepsilon(h)\cdot\|h\|^2$ tagot a *Taylor-formula* ***Peano-féle maradéktag****jának* nevezzük.

A tétel általánosítható arra az esetre is, ha $f \in D^s\{a\}$, aholl $s > 2$ tetszőleges természetes szám.

$\mathbb{R}^n \to \mathbb{R}$ függvények (feltétel nélküli) 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 $\mathbb{R}^n \to \mathbb{R}$ $(1 \leq n \in \mathbb{N})$ függvényekre.

• Fogalmak

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

Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \leq n \in \mathbb{N})$ 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 $f(x) \leq f(a)$ 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: 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

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

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

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

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.

2. tétel. (Elsőrendű szükséges feltétel a lokális szélsőértékre). *Legyen* $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$ *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 $\partial_i f(a) = 0$.

Az a 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** az

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

Peano-féle maradéktagos Taylor-formula alkalmazása.

Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$, $f \in D^2\{a\}$ és $f'(a) = \mathbf{0}$ (vagyis a az f függvény stacionárius pontja). 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

$$\langle f''(a)h, h\rangle > 0 \quad \text{minden } h \in \mathbb{R}^n \setminus \{\mathbf{0}\} \text{ vektorra,}$$

akkor $\lim\limits_{\mathbf{0}} \varepsilon = 0$ miatt a $(*)$ egyenlőség jobb oldala is pozitív egy alkalmas $K(a) \subset \mathcal{D}_f$ környezetben. Így

$$f(a+h) - f(a) > 0, \text{ azaz } f(a+h) > f(a), \quad \text{ha } a+h \in K(a).$$

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

Ha azt tesszük fel, hogy $\langle f''(a)h, h\rangle < 0$ minden $h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ vektorra, akkor az f függvénynek az a pontban lokális maximuma van.

A fentieket a következő állításban foglaljuk össze:

3. tétel. *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$ és $a \in \operatorname{int}\mathcal{D}_f$. Tegyük fel, hogy $f \in D^2\{a\}$, $f'(a) = \mathbf{0}$ és*

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

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

A $(**)$ feltételek nehezen ellenőrizhetők. A gyakorlatban már jól használható ekvivalens 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 szóban forgó fogalmakat $f''(a)$ helyett tetszőleges $A \in \mathbb{R}^{n\times n}$ **szimmetrikus** mátrixra definiáljuk.

1. definíció. *Legyen $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ $(1 \le n \in \mathbb{N})$ egy szimmetrikus mátrix. Ekkor a*

$$Q(h) := \langle A\cdot h, h\rangle = \left\langle \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}, \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \right\rangle = \sum_{i,j=1}^{n} a_{ij} h_i h_j, \quad h = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \in \mathbb{R}^n$$

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

Kvadratikus alakokat a felvett helyettesítési értékeik előjele alapján szokás osztályozni (az ún. „definitségi osztályokba” sorolni).

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

Világos, hogy minden Q kvadratikus alak az origóban nulla. A definíció szerint egy pozitív **definit** kvadratikus alak csak a $\mathbf{0}$ pontban veszi fel a 0 értéket. Pozitív **szemidefinit** kvadratikus alak azonban $\mathbf{0}$-tól különböző helyeken is lehet nulla. Például a

$$Q(h) = Q(h_1, h_2) = h_1^2 - 2h_1h_2 + h_2^2 = (h_1 - h_2)^2 \quad (h \in \mathbb{R}^2)$$

kvadratikus alak nem pozitív definit, de pozitív szemidefinit. Hasonló különbség van a negatív *definit* és negatív *szemidefinit* kvadratikus alakok között.

2×2-es szimmetrikus mátrixok definitségi osztályozása elemi úton könnyen „elintézhető". Legyen $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$ egy szimmetrikus mátrix. Ekkor

$$Q(h) := \langle Ah, h\rangle = \left\langle \begin{bmatrix} a & b \\ b & c \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right\rangle = ah_1^2 + 2bh_1h_2 + ch_2^2 \quad \left(h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \in \mathbb{R}^2\right)$$

Ha $0 \neq h_2 \in \mathbb{R}$, akkor a $t := h_1/h_2$ jelöléssel (minden $t \in \mathbb{R}$ szám felírható ilyen alakban) azt kapjuk, hogy

$$Q(h_1, h_2) = h_2^2 \cdot \left(a\left(\frac{h_1}{h_2}\right)^2 + 2b\left(\frac{h_1}{h_2}\right) + c \right) = h_2^2\big(at^2 + 2bt + c\big) = h_2^2 \cdot P(t) \quad (t \in \mathbb{R}).$$

Ha $0 \neq a \in \mathbb{R}$, akkor P másodfokú polinom, ezért

$$\begin{aligned} P(t) > 0 \ (\forall\, t \in \mathbb{R}) \quad &\Longleftrightarrow \quad a > 0 \ \text{ és } \ b^2 - ac = -\det A < 0, \\ P(t) < 0 \ (\forall\, t \in \mathbb{R}) \quad &\Longleftrightarrow \quad a < 0 \ \text{ és } \ b^2 - ac = -\det A < 0, \\ P \text{ pozitív és negatív értéket is felvesz} \quad &\Longleftrightarrow \quad b^2 - 4ac = -\det A > 0. \end{aligned}$$

A $h_2 = 0$ és az $a = 0$ esetek vizsgálatával adódik a következő állítás:

4. tétel. *Legyen* $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$ *egy szimmetrikus mátrix. Ekkor a*

$$Q(h) := \langle Ah, h\rangle = ah_1^2 + 2bh_1h_2 + ch_2^2 \quad \left(h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \in \mathbb{R}^2\right)$$

kvadratikus alak, illetve az $A \in \mathbb{R}^{2\times 2}$ *mátrix*

- *pozitív definit* $\Longleftrightarrow$ $a > 0$ *és* $\det A > 0$,
- *negatív definit* $\Longleftrightarrow$ $a < 0$ *és* $\det A > 0$,
- *indefinit* $\Longleftrightarrow$ $\det A < 0$.

Az általános esetben az A mátrix definitségét A bizonyos aldeterminánsainak az előjeleivel vizsgálhatjuk. Belátható a következő állítás:

5. tétel. **(Sylvester-kritérium.)** *Legyen* $A \in \mathbb{R}^{n\times n}$ $(1 \le n \in \mathbb{N})$ *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. Az A mátrix, illetve a Q kvadratikus alak pontosan akkor

pozitív definit, ha $D_1 > 0,\ D_2 > 0, D_3 > 0, \cdots,\ D_n > 0$,

negatív definit, ha $D_1 < 0,\ D_2 > 0, D_3 < 0, \cdots, (-1)^n D_n > 0$. $\square$

Térjünk most vissza a lokális szélsőérték-problémára. Egy $f \in D^2\{a\}$ függvény esetén értelmeztük az $f''(a)$ Hesse-mátrix definitségeit. Az előzőekből már egyszerűen következik az alábbi, a gyakorlatban már jól használható elégséges feltétel a lokális szélsőértékekre (vö. a 3. tétellel).

6. tétel. **(Másodrendű elégséges feltétel a lokális szélsőértékre $\mathbb{R}^n \to \mathbb{R}$ függvényekre.)** *Legyen* $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$ *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.*

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.

Meg lehet mutatni, hogy a 6. tételben megfogalmazott *elégséges* feltétel „nincs túl messze" egy *szükséges* feltételtől.

7. tétel. **(Másodrendű szükséges 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* f *függvénynek az* a *pontban lokális minimuma (maximuma) van.*

Ekkor $f'(a) = \mathbf{0}$, *és az* $f''(a)$ *Hesse-féle mátrix pozitív (negatív) szemidefinit.*

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.

8. 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*).

Az 1^o állítás a Sylvester-kritérium és a 6. 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.

9. 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

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

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

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

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:

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

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

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

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

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 $-\frac{1}{8}$. ■