# 10. előadás

# TÖBBVÁLTOZÓS FÜGGVÉNYEK DIFFERENCIÁLSZÁMÍTÁSA 2.

## Magasabb rendű deriváltak

Az $f \in \mathbb{R} \to \mathbb{R}$ függvények körében „nem okozott gondot" a 2-szer, 3-szor, ... való differenciálhatóság teljes indukcióval történő értelmezése. Először a kétszer deriválhatóság fogalmát definiáltuk. Azt mondtuk, hogy a szóban forgó függvény *kétszer differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *pontban* (röviden $f \in D^2\{a\}$), ha $f$ az $a$ pont egy $K(a) \subseteq \mathcal{D}_f$ környezetében deriválható ($f'$ értelmezve van a $K(a)$ halmazon) és $f'$ differenciálható az $a$ pontban, azaz $f' \in D\{a\}$. Ekkor az $f''(a) := \big(f'\big)'(a)$ számot az $f$ függvény *a pontbeli második deriváltjának neveztük.* Teljes indukcióval hasonlóan értelmeztük a 2-nél magasabb rendű deriválhatóság fogalmát.

A kétszeri differenciálhatóságnak ez az értelmezése átvihető az $f \in \mathbb{R}^n \to \mathbb{R}$ függvényekre akkor is, ha $n > 1$. Ti., ha $a \in \operatorname{int} \mathcal{D}_f$ és $\exists\, K(a) \subseteq \mathcal{D}_f \colon f \in D\big(K(a)\big)$, akkor az

$$f'(x) = \begin{pmatrix} \partial_1 f(x) & \partial_2 f(x) & \dots & \partial_n f(x) \end{pmatrix} \qquad \big(x \in K(a)\big),$$

deriváltmátrixok és a $\operatorname{grad} f(x) = \big(\partial_1 f(x), \partial_2 f(x), \dots, \partial_n f(x)\big)$ gradiens vektorok segítségével értelmezhető az

$$f' \colon \mathbb{R}^n \supseteq K(a) \ni x \mapsto \operatorname{grad} f(x) \in \mathbb{R}^n$$

deriváltfüggvény. Továbbá az $\mathbb{R}^n \to \mathbb{R}^n$ vektor-vektor függvények körében értelmeztük a deriválhatóságot. Tehát az $f \in \mathbb{R}^n \to \mathbb{R}$ függvényre minden további nélkül előírhatjuk, hogy $f' = \operatorname{grad} f \in D\{a\}$ teljesüljön. Ekkor

$$f''(a) := \big(f'\big)'(a) = \big(\operatorname{grad} f\big)'(a)$$

az $f$ függvény második deriváltja az $a$ pontban. Világos, hogy egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény esetén $f''(a)$ egy $(n \times n)$-es mátrix. Az is nyilvánvaló azonban, hogy ezen az úton nem tudunk eljutni a 2-nél magasabb rendű deriválhatósághoz, hiszen pl. $\mathbb{R}^n \to \mathbb{R}^{n \times n}$ típusú függvények differenciálhatóságát nem értelmeztük.

Emlékeztetünk ugyanakkor arra, hogy egy függvény akkor és csak akkor differenciálható egy pontban, ha a koordinátafüggvényei differenciálhatók ebben a pontban. Mivel a $\operatorname{grad} f$ függvény koordinátafüggvényei az $f$ függvény parciális deriváltjai, így

$$f' = \operatorname{grad} f = \big(\partial_1 f, \partial_2 f, \dots, \partial_n f\big) \in D\{a\} \qquad \Longleftrightarrow \qquad \partial_i f \in D\{a\} \quad (i = 1, 2, \dots, n).$$

Ez az észrevétel lehetőséget ad arra, hogy értelmezni tudjuk egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény magasabb rendű deriválhatóságát. Az értelmezés teljes indukcióval történik, első lépésként a kétszer deriválhatóságot definiálva.

**1. Definíció.** *Az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ függvény* **kétszer deriválható** (*vagy* **differenciálható**) *az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben: $f \in D^2\{a\}$), ha*

*a)* $\exists K(a) \subset \mathcal{D}_f$*, hogy $f \in D\{x\}$ minden $x \in K(a)$ pontban, és*

*b)* $\forall i = 1, 2, \dots, n$ *indexre $\partial_i f \in D\{a\}$.*

Az a) feltételt röviden úgy is írhatjuk, hogy a $\exists K(a) \subseteq \mathcal{D}_f$, hogy $f \in D\big(K(a)\big)$. Ebből következik, hogy $K(a)$ környezetben létezik az

$$f' = \Big(\partial_1 f, \partial_2 f, \ldots, \partial_n f\Big) = \operatorname{grad} f \in \mathbb{R}^n \to \mathbb{R}^n$$

deriváltfüggvény, ami tehát már egy vektor-vektor függvény.

A b) feltétel pedig azzal ekvivalens, hogy $f' \in D\{a\}$. Így minden $i = 1, 2, \ldots, n$ esetén a $\partial_i f \in \mathbb{R}^n \to \mathbb{R}$ parciális deriváltfüggvényeknek léteznek az $a$ pontban mindegyik változó szerinti parciális deriváltjai:

$$\partial_j\Big(\partial_i f\Big)(a) \quad (j = 1, 2, \ldots, n).$$

Ezeket a számokat az ***f* függvény *a* pontbeli, *i*-edik és *j*-edik változó szerinti másodrendű** (vagy ***második***) ***parciális deriváltjának*** nevezzük.

Az $f$ függvény $a$ ***pontbeli második deriváltját*** így értelmezzük:

$$f''(a) := \Big(\operatorname{grad} f\Big)'(a).$$

Világos, hogy egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény esetén $f''(a)$ egy $(n \times n)$-es mátrix.

**2. Definíció.** *Ha az $f \in \mathbb{R}^n \to \mathbb{R}$ $\ (n \in \mathbb{N}^+)$ függvény kétszer deriválható az $a \in \operatorname{int} \mathcal{D}_f$ pontban, akkor*

$$f''(a) = \begin{pmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \ldots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \ldots & \partial_{2n} f(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \ldots & \partial_{nn} f(a) \end{pmatrix} \in \mathbb{R}^{n \times n}$$

*az **f függvény a pontbeli Hesse-féle mátrixa**, ahol*

$$\partial_{ij} f(a) = \partial_j\Big(\partial_i f\Big)(a) \quad \Big(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n\Big).$$

A 2-nél magasabb rendű deriválhatóságot teljes indukcióval így értelmezzük: az $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \leq n \in \mathbb{N})$ függvény ***s-szer* $(2 \leq s \in \mathbb{N})$ *deriválható*** az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben $f \in D^s\{a\}$), ha

a) $\exists K(a) \subseteq \mathcal{D}_f$ környezet, hogy $f \in D^{s-1}\Big(K(a)\Big)$ és

b) minden $(s-1)$-edrendű $\partial_{i_1}\partial_{i_2}\ldots\partial_{i_{s-1}} f$ $\Big(1 \leq i_1, i_2, \ldots, i_{s-1} \leq n\Big)$ parciális deriváltfüggvény deriválható az $a$ pontban.

Az előbbihez kapcsolódik egy fontos fogalom: azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ függvény ***s-szer* $(s \in \mathbb{N}^+)$ *folytonosan deriválható*** az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben $f \in C^s\{a\}$), ha

a) $\exists K(a) \subseteq \mathcal{D}_f$ környezet, hogy $f \in D^s\Big(K(a)\Big)$ és

b) minden $s$-edrendű $\partial_{i_1}\partial_{i_2}\ldots\partial_{i_s} f$ $\Big(1 \leq i_1, i_2, \ldots, i_s \leq n\Big)$ parciális deriváltfüggvény folytonos az $a$ pontban.

***Megjegyzés.*** A totális deriválhatóságra vonatkozó elégséges feltétel azt jelenti, hogy $f$ (egyszer) folytonosan deriválható az $a$ pontban.

A magasabb rendű parciális deriváltakkal kapcsolatban joggal merül fel a *kérdés*, hogy azok kiszámításakor van-e szerepe a változók sorrendjének? A következő tétel azt állítja, hogy ha az $f$ függvény a szóban forgó $a$ helyen „elég sokszor” deriválható, akkor a sorrend elveszti a jelentőségét.

**1. Tétel (Young-tétel).** *Ha* $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \leq n \in \mathbb{N})$ *és* $f \in D^2\{a\}$, *akkor*

$$\partial_{ij} f(a) = \partial_{ji} f(a) \quad \forall\, i, j = 1, \ldots, n \quad \textit{indexre.}$$

***Bizonyítás.*** A tételt nem bizonyítjuk.

***Példa.*** Az $f \in D^2\{a\}$ feltétel hiánya esetén a parciális deriváltak képzésének a sorrendje általában nem cserélhető fel. Ha például

$$f(x, y) := \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2}, & \text{ha } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} \\ 0, & \text{ha } (x, y) = (0, 0), \end{cases}$$

akkor

$$\partial_{12} f(0, 0) = -1, \qquad \partial_{21} f(0, 0) = 1.$$

Érdemes meggondolni azt, hogy $f \notin D^2\{a\}$ miért igaz.

***Megjegyzések.***

1. A Young-tételből következik, hogy ha $f \in D^2\{a\}$, akkor az $f$ függvény $a$ pontbeli Hesse-féle mátrixa *szimmetrikus* mátrix.

2. Teljes indukcióval igazolható a Young-tétel következő általánosítása: *Tegyük fel, hogy az* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ *függvény* $s$*-szer* $(s \in \mathbb{N})$ *differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *pontban. Legyen* $i_k \in \mathbb{N},\ 1 \leq i_k \leq n \quad (k = 1, 2, \ldots, s)$ *és* $j_1, j_2, \ldots, j_s$ *az* $i_1, i_2, \ldots, i_s$ *indexek egy tetszőleges permutációja. Ekkor*

$$\partial_{i_1} \partial_{i_2} \cdots \partial_{i_s} f(a) = \partial_{j_1} \partial_{j_2} \cdots \partial_{j_s} f(a).$$

## Taylor-polinomok

Induljunk ki az egyváltozós ismereteinkből. *Motivációként* akkor azt a fontos *problémát* vetettük fel, hogy egy adott „bonyolult” függvényt vajon meg lehet-e közelíteni egyszerű szerkezetű függvényekkel, például a jól kezelhető és könnyen számolható polinomokkal.

A valós-valós esetben beláttuk, 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 a fontos eredményhez, hogy ha egy függvény $m$-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.* Ezzel kapcsolatos alapvető eredményünk volt a *Taylor-formulára* vonatkozó alábbi állítás:

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

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

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 $T_{m-1,a}f(h)$-hoz képest kicsi, ha $h$ közel van 0-hoz.

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 ezt a kiterjesztést csak az $n = 2$ és az $m = 2$ speciális esetben igazoljuk. Legyen tehát $f \in \mathbb{R}^2 \to \mathbb{R}$, $a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f$ és $h = (h_1, h_2) \in \mathbb{R}^2$. Tegyük fel, hogy $f \in C^2\{a\}$, azaz egy $K(a) \subseteq \mathcal{D}_f$ környezetben az $f$ függvény kétszer deriválható, azaz $f \in D^2\big(K(a)\big)$, illetve a másodrendű parciális deriváltak folytonosak az $a$ pontban. Vegyük a síkon az $a$ és az $a + h \in K(a)$ pontokat összekötő egyenes $a + th \ \ (t \in \mathbb{R})$ pontjait. Tekintsük a

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

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

$$F(0) = f(a + 0 \cdot h) = f(a).$$

Másrészt az összetett függvény deriválására vonatkozó tétel alapján

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

következésképpen

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

A $F$ függvény második deriváltja:

$$\begin{aligned} F''(t) &= (\partial_1 f(a+th)h_1 + \partial_2 f(a+th)h_2)' = \Big[(\partial_1 f)'(a+th) \cdot h\Big]h_1 + \Big[(\partial_2 f)'(a+th) \cdot h\Big]h_2 = \\ &= \Big(\partial_{11} f(a+th)h_1 + \partial_{12} f(a+th)h_2\Big)h_1 + \Big(\partial_{21} f(a+th)h_1 + \partial_{22} f(a+th)h_2\Big)h_2 = \\ &= \partial_{11} f(a+th)h_1^2 + \partial_{12} f(a+th)h_1h_2 + \partial_{21} f(a+th)h_1h_2 + \partial_{22} f(a+th)h_2^2 = \\ &= \langle f''(a+th) \cdot h, h\rangle. \end{aligned}$$

(Az utolsó egyenlőséget gondoljuk végig a mátrixszorzás és a skaláris szorzat definícióinak birtokában. Itt $f''(a+th) = \begin{pmatrix} \partial_{11} f(a+th) & \partial_{12} f(a+th) \\ \partial_{21} f(a+th) & \partial_{22} f(a+th) \end{pmatrix}$ egy $(2 \times 2)$-es mátrix, $h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ oszlopvektor, tehát $f''(a+th) \cdot h \in \mathbb{R}^{2\times 1} \approx \mathbb{R}^2$ egy oszlopvektor.)

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

Mivel $f \in C^2\{a\}$, ezért a másodrendű parciális deriváltak folytonos az $a$ pontban, és így

$$f''(a+\nu h) = \begin{pmatrix} \partial_{11} f(a+\nu h) & \partial_{12} f(a+\nu h) \\ \partial_{21} f(a+\nu h) & \partial_{22} f(a+\nu h) \end{pmatrix} = \begin{pmatrix} \partial_{11} f(a)+\eta_{11}(h) & \partial_{12} f(a)+\eta_{12}(h) \\ \partial_{21} f(a)+\eta_{21}(h) & \partial_{22} f(a)+\eta_{22}(h) \end{pmatrix} =$$
$$= f''(a)+\eta(h),$$

ahol $\eta = \left[\eta_{ij}\right] \in \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ egy olyan függvény, amelyre $\lim\limits_0 \eta = \mathbf{0}$ .

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 $\dfrac{|h_i h_j|}{\|h\|^2} \le 1$ $(i,j=1,2)$ és $\lim\limits_0 \eta_{ij} = 0$, 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 $\lim\limits_0 \varepsilon = 0$ teljesül.

Az eddigieket összefoglalva azt kaptuk, hogy van olyan $\varepsilon \in \mathbb{R}^2 \to \mathbb{R}$, a $\lim\limits_0 \varepsilon = 0$ feltételt kielégítő függvény, amelyre

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

A fenti gondolatmenetet követve viszonylag egyszerűen belátható, hogy a (2) képlet tetszőleges $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 < n \in \mathbb{N})$ függvényre is teljesül.

**2. Tétel (A Taylor-formula a Peano-féle maradéktaggal).** *Legyen* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ *és tegyük fel, hogy* $a \in \operatorname{int} \mathcal{D}_f$ *és* $f \in C^2\{a\}$. *Ekkor van olyan* $\varepsilon \in \mathbb{R}^n \to \mathbb{R}$, *a* $\lim\limits_0 \varepsilon = 0$ *feltételnek eleget tevő függvény, hogy*

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

A formula jobb oldala első három tagjának az összege egy $n$-változós legfeljebb másodfokú polinom, amit az ***f* függvény *a* ponthoz tartozó második Taylor-polinomjának** nevezünk:

$$T_{2,a} f(h) := f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\left\langle f''(a)h, h\right\rangle = f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij} f(a) h_i h_j.$$

A formulában szereplő $\varepsilon(h)\cdot\|h\|^2$ tagot a ***Taylor-formula Peano-féle maradéktagjának*** nevezzük.

***Megjegyzések.***

1. A 2. Tétel jelentősége egyrészt abban van, hogy a felhasználásával „bonyolult” $n$-változós valós értékű függvények helyettesítési értékeire lehet „jó” közelítő értékeket adni. Másrészt, azt a következő órán látni fogjuk, hogy a 2. Tétel fontos szerepet játszik $\mathbb{R}^n \to \mathbb{R}$ típusú függvények szélsőérték-problémáinak a vizsgálatánál.
2. A 2. Tétel általánosítható magasabb rendű Taylor-polinomokra is.

***Példa.*** Számítsuk ki az
$$f(x,y) = e^x \cos y \qquad \big((x,y)\in\mathbb{R}^2\big)$$
függvény $a=(0,0)$ pont körüli második Taylor polinomját! A
$$T_{2,a}f(h) = f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij} f(a) h_i h_j$$
képletben legyen $a=(0,0)$ és $(x,y) = a+h = (h_1,h_2)$, azaz $h_1 = x$ és $h_2 = y$. Mivel
$$\begin{aligned}
&f(x,y) = e^x\cos y \qquad \to \qquad f(0,0)=1,\\
&\partial_1 f(x,y) = e^x\cos y,\quad \partial_2 f(x,y) = -e^x \sin y \qquad \to \qquad \partial_1 f(0,0)=1,\quad \partial_2 f(0,0)=0,\\
&\partial_{11} f(x,y) = e^x\cos y,\quad \partial_{12} f(x,y) = -e^x \sin y \qquad \to \qquad \partial_{11} f(0,0)=1,\quad \partial_{22} f(0,0)=0,\\
&\partial_{21} f(x,y) = -e^x\sin y,\quad \partial_{22} f(x,y) = -e^x \cos y \qquad \to \qquad \partial_{21} f(0,0)=0,\quad \partial_{22} f(0,0)=-1,
\end{aligned}$$
így a keresett Taylor polinom:
$$T_{2,a}f(x,y) = 1 + 1\cdot x + 0\cdot y + \frac{1}{2}\cdot\Big(1\cdot x^2 + 2\cdot 0\cdot xy - 1\cdot y^2\Big) = \frac{x^2}{2} - \frac{y^2}{2} + x + 1.$$

## $\mathbb{R}^n \to \mathbb{R}$ 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. Az egyváltozós esetben bevezetett fogalmak minden további nehézség nélkül átvihetők a többváltozós függvényekre.

**3. 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* **abszolút maximuma van** (*vagy másképp fogalmazva az $a$ pont az $f$ függvénynek* **abszolút maximumhelye**), *ha*
$$\forall x\in\mathcal{D}_f\colon\ f(x)\le f(a).$$
*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.

**4. Definíció.** *Azt mondjuk, hogy az $f\in\mathbb{R}^n\to\mathbb{R}$ $(n\in\mathbb{N}^+)$ függvénynek az $a\in\operatorname{int}\mathcal{D}_f$ pontban* **lokális maximuma van** (*vagy másképp fogalmazva az $a$ pont az $f$ függvénynek* **lokális maximumhelye**), *ha van olyan $K(a)\subseteq\mathcal{D}_f$ környezet, hogy*
$$\forall x\in K(a)\colon\ f(x)\le f(a).$$
*Ekkor az $f(a)$ függvényértéket az $f$ függvény* **lokális maximumának** *nevezzük.*

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

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

## Elsőrendű szükséges 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.

**3. Tétel (Elsőrendű szükséges feltétel a lokális szélsőértékre).** *Tegyük fel, hogy* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ *és* $a \in \operatorname{int} \mathcal{D}_f$. *Továbbá*

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

*Ekkor* $f'(a) = 0$, *azaz*

$$f'(a) = \big(\partial_1 f(a) \quad \partial_2 f(a) \quad \dots \quad \partial_n f(a)\big) = \big(0 \quad 0 \quad \dots \quad 0\big).$$

***Bizonyítás.*** Legyen $i = 1, 2, \dots, n$ rögzített és tekintsük meg a

$$G_i(t) = f\big(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n\big) \qquad \big(t \in K(a_i)\big)$$

parciális függvényt! Tudjuk, hogy $\exists \partial_i f(a) = G_i'(a_i)$, hiszen $f \in D\{a\}$. Minden $K(a) \in \mathcal{D}_f$ környezethez van olyan $r > 0$ szám, amire igaz, hogy

$$\big(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n\big) \in K(a) \qquad \big(t \in (a_i - r, a_i + r)\big).$$

Ezért, ha az $f$ függvénynek az $a$ pontban lokális szélsőértéke van, akkor a $G_i$ függvénynek az $a_i$ pontban is lokális szélsőértéke van, mivel $G_i(a_i) = f(a)$. Ekkor a valós-valós függvényeknél tanult, a lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel szerint $G_i'(a_i) = 0$, ami éppen azt jelenti, hogy $\partial_i f(a) = 0$.

**5. Definíció.** *Az* $a \in \operatorname{int} \mathcal{D}_f$ *pont az* $f \in \mathbb{R}^n \to \mathbb{R}$ *függvény* **stacionárius pontja**, *ha* $f \in D\{a\}$ *és* $f'(a) = 0$.

***Megjegyzések.***

1. A tétel tehát azt állítja, hogy a lokális szélsőértékhelyek szükségképpen a függvény stacionárius pontjai. Az $f'(a) = 0$ azonban csak *szükséges*, de *nem elégséges* feltétel a lokális szélsőértékre. Az $n = 1$ esetben például már ismerjük az $f(x) := x^3$ $(x \in \mathbb{R})$ függvényt, amelynek az $a = 0$ stacionárius pontjában nincs lokális szélsőértéke. Az $n = 2$ esetben az $f(x) := x^2 - y^2$ $\big((x, y) \in \mathbb{R}^2)\big)$ függvénynek az $a = (0, 0)$ pont stacionárius pontja, de itt sincs lokális szélsőérték.

2. A tétel értelmében egy $f \in \mathbb{R}^n \to \mathbb{R}$ differenciálható függvény stacionárius pontjainak megkeresésére szükséges megoldani a

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

egyenletrendszert. Csak az így kapott $(x_1, \ldots, x_n)$ pontok *lehetnek* 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. Az elsőrendű elégséges feltételhez nem tudunk megfelelő állítást kimondani többváltozós függvények esetén. A másodrendű elégséges feltétel azt mondja ki, hogy 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. Ezt fogjuk általánosítani $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre.

A többváltozós esetben a kiindulópontunk *alapötlete* a Peano-féle maradéktagos Taylor-formula alkalmazása. Idézzük fel ezt az állítást: *Legyen* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ *és tegyük fel, hogy* $a \in \operatorname{int} \mathcal{D}_f$ *és* $f \in C^2\{a\}$. *Ekkor van olyan* $\varepsilon \in \mathbb{R}^n \to \mathbb{R}$, *a* $\lim_0 \varepsilon = 0$ *feltételnek eleget tevő függvény, hogy*

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

Ha az $a$ pont az $f$ függvény stacionárius pontja, azaz $f'(a) = 0$, akkor $\langle f'(a), h \rangle = 0$ minden $h \in \mathbb{R}^n$ esetén. Vizsgájuk meg a

$$\mathbb{R}^n \ni h \mapsto \langle f''(a)h, h \rangle \in \mathbb{R}$$

függvényt. Emlékezzük arra, hogy ha $f \in D^2\{a\}$, akkor a Young-tétel miatt az $f''(a) \in \mathbb{R}^{n \times n}$ másodrendű parciális deriváltakat tartalmazó Hesse-féle mátrix szimmetrikus.

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

$$Q(h) := \langle A \cdot h, h \rangle = \sum_{i,j=1}^{n} a_{ij} h_i h_j \in \mathbb{R} \qquad \big(h = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n\big)$$

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

A $Q\colon \mathbb{R}^n \to \mathbb{R}$ kvadratikus alakok többváltozós polinomfüggvények, ezért minden pontban folytonosak és differenciálhatóak. Másrészt minden $\lambda \in \mathbb{R}$ esetén

$$Q(\lambda h) = \sum_{i,j=1}^{n} a_{ij} \cdot (\lambda h_i) \cdot (\lambda h_j) = \lambda^2 \sum_{i,j=1}^{n} a_{ij} h_i h_j = \lambda^2 Q(h) \qquad (h \in \mathbb{R}^n).$$

Mivel a $Q\colon \mathbb{R}^n \to \mathbb{R}$ kvadratikus alakok folytonosak a $\{h \in \mathbb{R}^n\colon \|h\| = 1\}$ korlátos és zárt halmazon, így a Weierstrass-tétel szerint

$$\exists m_Q := \min\{Q(h)\colon h \in \mathbb{R}^n,\ \|h\| = 1\} \in \mathbb{R} \quad \text{és} \quad \exists M_Q := \max\{Q(h)\colon h \in \mathbb{R}^n,\ \|h\| = 1\} \in \mathbb{R}.$$

Ha $h \neq 0$, akkor

$$Q(h) = Q\left(\|h\| \cdot \frac{h}{\|h\|}\right) = \|h\|^2 \cdot Q\left(\frac{h}{\|h\|}\right),$$

és mivel $\frac{h}{\|h\|}$ egységvektor, így

$$(*) \qquad m_Q\|h\|^2 \le Q(h) \le M_Q\|h\|^2 \qquad (h \in \mathbb{R}^n).$$

**7. Definíció.** *Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix, illetve a hozzá tartozó $Q(h) = \langle A\cdot h, h\rangle \quad (h \in \mathbb{R}^n)$ kvadratikus alak*

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

A $(*)$ egyenlőtlenségből és az $m_Q$, $M_Q$ számok értelmezéséből következik, hogy $Q$ pontosan pozitív definit, ha $m_Q > 0$; negatív definit, ha $M_Q < 0$; és indefinit, ha $m_Q < 0 < M_Q$.

Most térjünk vissza a Taylor-formulához. Ha $f'(a) = 0$ és $Q(h) := \langle f''(a)h, h\rangle$ akkor

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

Ha $Q$ pozitív definit, akkor $(*)$ miatt

$$f(a+h) - f(a) \ge \frac{m_Q}{2}\cdot\|h\|^2 + \varepsilon(h)\cdot\|h\|^2 = \left(\frac{m_Q}{2} + \varepsilon(h)\right)\cdot\|h\|^2.$$

Mivel $\lim\limits_{0} \varepsilon = 0$, így van olyan $K_\delta(0)$ környezet, ahol $\frac{m_Q}{2} + \varepsilon(h) > 0$ teljesül. Ezért, ha $h \in K_\delta(0)$, akkor $f(a+h) - f(a) \ge 0$. Az $x = a + h$ helyettesítéssel azt kapjuk, hogy

$$\forall x \in K_\delta(a)\colon f(x) \ge f(a),$$

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

Hasonlóan igazolható, hogy ha $Q$ negatív definit, akkor az $f$ függvénynek lokális maximuma van az $a$ pontban. Az eddigi eredményeket a következő állításban tudjuk összefoglalni.

**4. Tétel (Másodrendű elégséges feltétel a lokális szélsőértékre).** *Legyen $f \in \mathbb{R}^n \to \mathbb{R} \quad (n \in \mathbb{N}^+)$, $a \in \operatorname{int}\mathcal{D}_f$ és $f \in C^2\{a\}$. Tegyük fel, hogy*

- *$f'(a) = 0$,*
- *az $f''(a)$ Hesse-féle mátrix pozitív (negatív) definit.*

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

## Másodrendű szükséges feltétel a lokális szélsőértékre

**8. Definíció.** *Azt mondjuk, hogy az* $A \in \mathbb{R}^{n\times n}$ *szimmetrikus mátrix, illetve a hozzá tartozó* $Q(h) = \langle A \cdot h, h\rangle \quad (h \in \mathbb{R}^n)$ *kvadratikus alak*

- **pozitív szemidefinit**, *ha* $\forall\, h \in \mathbb{R}^n$ *esetén* $Q(h) \geq 0$,
- **negatív szemidefinit**, *ha* $\forall\, h \in \mathbb{R}^n$ *esetén* $Q(h) \leq 0$.

A $(*)$ egyenlőtlenségből és az $m_Q$, $M_Q$ számok értelmezéséből következik, hogy $Q$ pontosan pozitív szemidefinit, ha $m_Q \geq 0$, és negatív szemidefinit, ha $M_Q \leq 0$. Ha $Q$ nem (pozitív vagy negatív) szemidefinit, akkor indefinit.

**5. Tétel (Másodrendű szükséges feltétel a lokális szélsőértékre).** *Legyen* $f \in \mathbb{R}^n \to \mathbb{R} \quad (n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ *és* $f \in C^2\{a\}$. *Ha az* $f$ *függvénynek az* $a$ *pontban lokális minimuma (maximuma) van, akkor*

- $f'(a) = 0$,
- *az* $f''(a)$ *Hesse-féle mátrix pozitív (negatív) szemidefinit.*

***Bizonyítás.*** Az elsőrendű szükséges feltétel miatt $f'(a) = 0$. A Hesse-féle mátrixszal kapcsolatos állítás igazolásához tegyük fel indirekt módon, hogy

$$Q(h) := \langle f''(a)h, h\rangle$$

nem pozitív szemidefinit, azaz $m_Q < 0$. Ekkor $\exists u \in \mathbb{R}^n$ úgy, hogy $\|u\| = 1$ és $Q(u) < 0$. Legyen $h = tu$. Ekkor $(\#)$ miatt

$$f(a+tu) - f(a) = \tfrac{1}{2}Q(tu) + \varepsilon(tu) \cdot \|tu\|^2 = \frac{t^2}{2}Q(u) + \varepsilon(tu) \cdot t^2 = \Big(\frac{1}{2}Q(u) + \varepsilon(tu)\Big) \cdot t^2.$$

Mivel $Q(u) < 0$ és $\lim_0 \varepsilon = 0$, így van olyan $\delta > 0$, hogy $\frac{1}{2}Q(u) + \varepsilon(tu) < 0$ minden $|t| < \delta$ esetén, és így $f(a+tu) - f(a) < 0$. Azonban minden $K(a)$ környezet tartalmaz egy olyan $x = a + tu$ pontot, amire $|t| < \delta$ teljesül. Ez azt jelenti, hogy $a$ nem lokális minimumhelye az $f$ függvénynek.

Hasonlóan igazolható, hogy a Hesse-féle mátrix negatív szemidefinit, ha az $f$ függvénynek az $a$ pontban lokális maximuma van.

A tétel fontos következménye: *Ha* $f \in \mathbb{R}^n \to \mathbb{R} \quad (n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ *és* $f \in C^2\{a\}$, *továbbá* $f'(a) = 0$ *és az* $f''(a)$ **Hesse-féle mátrix indefinit**, *akkor az* $f$ *függvénynek az* $a$ *pontban* **nincs lokális szélsőértéke**.

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

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

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

*az $A$ mátrix „bal felső sarokmátrixainak” a determinánsát. Ekkor az $A$ mátrix, illetve a $Q$ kvadratikus alak*

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

***Bizonyítás.*** A tételt nem bizonyítjuk.

***Megjegyzés.*** A Sylvester-kritériumból nem lehet megtudni mikor indefinit egy szimmetrikus mátrix. Azonban $n = 2$ esetén van egy egyszerű elégséges feltétel erre az esetre is, nevezetesen ha $\det A < 0$. Ez elemi úton, a másodfokú függvény tulajdonságai alapján könnyen igazolható.

Foglaljuk össze a kapott eredményeket kétváltozós függvények esetében:

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

$$\partial_1 f(a) = 0 \quad \textit{és} \quad \partial_2 f(a) = 0.$$

*Jelölje*

$$D(a) := \det \begin{pmatrix} \partial_{11} f(a) & \partial_{12} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) \end{pmatrix}.$$

*Ekkor*

1. *ha $D(a) > 0$ és $\partial_{11} f(a) > 0$ [illetve $\partial_{11} f(a) < 0$], akkor az $f$ függvénynek $a$-ban lokális minimuma [illetve maximuma] van.*
2. *ha $D(a) < 0$, akkor $f$-nek $a$-ban nincs lokális szélsőértéke (ezt nevezzük nyeregpontnak)*
3. *ha $D(a) = 0$, akkor így nem tudjuk megállapítani, hogy az $a$ pont vajon lokális szélsőértékhely-e vagy sem.*

***Példa.*** Legyen

$$f(x, y) = x^2 + xy + 2y^2 \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Ekkor $f \in C^2(\mathbb{R}^2)$. Továbbá minden $(x, y) \in \mathbb{R}^2$ esetén

$$\partial_x f(x, y) = 2x + y = 0, \quad \partial_y f(x, y) = x + 4y = 0 \qquad \Longrightarrow \qquad x = y = 0.$$

Ez azt jelenti, hogy $f$-nek csak a $P(0, 0)$ pontban lehet lokális szélsőértéke. Másrészt

$$\partial_{xx} f(x, y) = 2, \quad \partial_{xy} f(x, y) = \partial_{yx} f(x, y) = 1, \quad \partial_{yy} f(x, y) = 4 \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Mivel

$$D(0, 0) = \det \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} = 7 > 0 \qquad \text{és} \qquad \partial_{xx} f(x, y) = 2 > 0,$$

így $f$-nek a $P(0, 0)$ pontban lokális minimuma van.

***Példa.*** Legyen
$$f(x,y) = x^2 - y^2 \qquad \big((x,y) \in \mathbb{R}^2\big).$$
Ekkor $f \in C^2(\mathbb{R}^2)$. Továbbá minden $(x,y) \in \mathbb{R}^2$ esetén
$$\partial_x f(x,y) = 2x = 0, \quad \partial_y f(x,y) = -2y = 0 \qquad \Longrightarrow \qquad x = y = 0.$$
Ez azt jelenti, hogy $f$-nek csak a $P(0,0)$ pontban lehet lokális szélsőértéke. Másrészt
$$\partial_{xx} f(x,y) = 2, \quad \partial_{xy} f(x,y) = \partial_{yx} f(x,y) = 0, \quad \partial_{yy} f(x,y) = -2 \qquad \big((x,y) \in \mathbb{R}^2\big).$$
Mivel
$$D(0,0) = \det \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = -4 < 0$$
így $f$-nek nincs a $P(0,0)$ pontban lokális szélsőértéke.

## Abszolút szélsőértékek

Az a megfigyelés, hogy a lokális szélsőértékhelyeken a függvény deriváltja nulla (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 stacionárius pont, azaz $f'(c) = 0$.

Ha tehát megkeressük $f$ összes $c \in (a,b)$ stacionárius pontját, 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.

A fenti gondolatmenetet könnyen általánosíthatjuk többváltozós függvényekre.

**7. 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.*