9. előadás

2020. november 9.

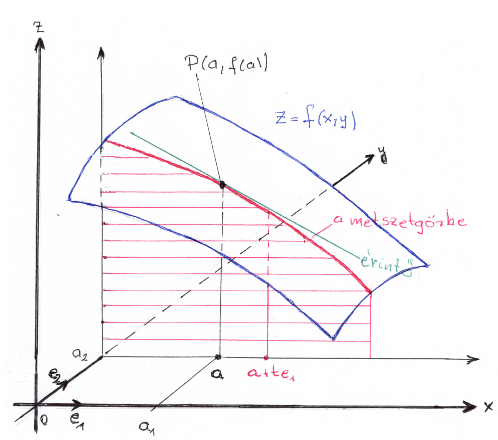
DIFFERENCIÁLSZÁMÍTÁS

A többváltozós függvények körében a deriválás több változatosságot kínál, mint a folytonosság vagy a határérték. Itt többféle derivált fogalommal kell megismerkednünk.

Parciális deriváltak $\mathbb{R}^n \to \mathbb{R}$ függvényekre

A parciális deriváltakat úgy kapjuk, hogy egy híján minden változót rögzítünk, és az így kapott egyváltozós függvényt deriváljuk.

A kétváltozós esetben a fogalomnak szemléletes jelentés adható. Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$ és $a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f$. A függvény grafikonja a térben a $z = f(x, y)$ egyenletű felület. Fektessünk az a ponton át az x tengellyel párhuzamos egyenest (ennek pontjai az xy síkban $(a_1 + t, a_2)$, $t \in \mathbb{R}$), majd vegyük a függvény értékeit ezekben a pontokban: $f(a_1 + t, a_2)$. Ekkor egy $F(t) := f(a_1 + t, a_2)$ valós-valós függvényt értelmezünk. Ennek képe egy, a felületen futó (metszet)görbe, vagyis a $z = f(x, y)$ egyenletű felület és az $y = a_2$ egyenletű sík metszésvonala. Az f függvény x változó szerinti parciális deriváltját a-ban $F'(0)$-val értelmezzük (feltéve, hogy a derivált létezik). $F'(0)$ a metszetgörbe $P\big(a, f(a)\big)$ pontbeli érintőjének a meredeksége. A fentieket szemlélteti a következő ábra:



A metszetgörbe a

$F : K(0) \ni t \mapsto f(a_1 + t, a_2) = f(a + te_1)$

függvény grafikonja.

$F'(0)$ az f függvény

x változó szerinti parciális deriváltja az a pontban.

Az y változó szerinti parciális deriváltat hasonló módon értelmezzük.

Egy $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \le n \in \mathbb{N})$ függvény mindegyik változója szerint képezhetjük a parciális deriváltat valamely $a = \big(a_1, a_2, \ldots, a_n\big) \in \operatorname{int} \mathcal{D}_f$ pontban. Rögzítsük az $a = \big(a_1, a_2, \ldots, a_n\big)$ pont koordinátáit az i-edik $(i = 1, 2, \ldots, n)$ kivételével. A

$$F : K(0) \ni t \mapsto f\big(a_1, a_2, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n\big) = f(a + te_i)$$

valós-valós függvény $(e_1, e_2, \ldots, e_n$ a kanonikus bázis $\mathbb{R}^n$-ben$)$ $t = 0$ pontban vett deriváltját (amennyiben létezik) nevezzük a f függvény a-ban vett i-edik parciális deriváltjának.

1. definíció. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$, $a \in \operatorname{int} \mathcal{D}_f$ és $e_1, \dots, e_n$ a kanonikus bázis $\mathbb{R}^n$-ben. A f függvénynek az a pontban létezik az i-edik $(i = 1, 2, \dots, n)$ változó szerinti **parciális deriváltja**, ha a*

$$F : K(0) \ni t \mapsto f(a + te_i)$$

valós-valós függvény deriválható a 0 pontban. A $F'(0)$ valós számot a f függvény a pontbeli, i-edik változó szerinti parciális deriváltjának nevezzük, és az alábbi szimbólumok valamelyikével jelöljük:

$$\partial_i f(a), \quad \partial_{x_i} f(a), \quad \frac{\partial f}{\partial x_i}(a), \quad f'_{x_i}(a), \quad D_i f(a).$$

Ha például ki akarjuk számítani az

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

függvény első változója szerinti parciális deriváltját az $a = (a_1, a_2)$ pontban, akkor a második változót (y-t) konstansnak (a_2-nek) tekintjük, és az első változó (vagyis x) szerint derválunk:

$$\partial_1 f\big(a_1, a_2\big) = 3a_1^2 \cdot a_2^2 + 2 + 0.$$

Ugyanígy a második változó szerinti parciális deriválás során az első változót tekintjük konstansnak, és a második változó szerint deriválunk:

$$\partial_2 f\big(a_1, a_2\big) = a_1^3 \cdot 2a_2 + 0 + 1.$$

Legyen az f függvény értelmezve $\mathbb{R}^n$ egy részhalmzán. Az f függvény i-edik **parciális deriváltfüggvényén** azt a $\partial_i f$ függvényt értjük, amely azokban a pontokban van értelmezve, ahol az f függvény i-edik parciális deriváltja létezik és véges, és ott az értéke $\partial_i f(a)$.

• Magasabb rendű parciális deriváltak

Ha egy n-változós függvénynek az $a \in \mathbb{R}^n$ pont egy környezetében létezik a függvény valamely változó szerinti parciális deriváltja, akkor a parciális deriváltfüggvény szintén n-változós valós értékű függvény. Így ennek bármelyik másik változó szerinti parciális deriválhatóságát vizsgálhatjuk.

Legyen f értelmezve az $a \in \mathbb{R}^n$ pont egy környezetében. Ha rögzített $i = 1, 2, \dots, n$ esetén a $\partial_i f$ parciális derivált létezik az a pont egy környezetében és a $\partial_i f$ parciális deriváltfüggvénynek létezik a j-edik $(j = 1, 2, \dots, n)$ változó szerinti parciális deriváltja az a pontban, akkor a $\partial_j\big(\partial_i f\big)(a)$ számot (mint $\partial_i f \in \mathbb{R}^n \to \mathbb{R}$ függvény a-beli j-edik változó szerinti parciális deriváltját) a függvény a-beli ij-edik **másodrendű parciális deriváltjának** nevezzük és így jelöljük:

$$\partial_{ij} f(a) := \partial_i \partial_j f(a) := \partial_j\big(\partial_i f\big)(a)$$

$\partial_{ij} f(a)$ helyett használatosak még a következő jelölések is:

$$\partial_{x_i x_j} f(a), \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \quad f''_{x_i x_j}(a), \quad D_j D_i f(a), \quad D_{ji} f(a).$$

A $\partial_{ij} f(a)$ deriváltat $i = j$ esetén másodrendű *tiszta parciális deriváltnak* nevezzük, és a

$$\partial_i^2 f(a), \quad \frac{\partial^2 f}{\partial x_i^2}(a), \quad \dots .$$

szimbólumok valamelyikével jelöljük. Ha $i \neq j$, akkor $\partial_{ij} f(a)$-t másodrendű *vegyes parciális deriváltnak* is szokás nevezni.

Kettőnél magasabb rendre az s-edrendű ($2 < s \in \mathbb{N}$) parciális deriválhatóságot s szerinti indukcióval definiálhatjuk. Ha $1 \le i_1, i_2, \dots, i_s \le n$ tetszőleges indexek, akkor az f függvény a-beli s-edrendű i_1-edik, ..., i_s-edik változó szerinti parciális deriváltját az alábbi szimbólumok valamelyikével jelöljük:

$$\partial_{i_1} \partial_{i_2} \dots \partial_{i_s} f(a), \quad \frac{\partial^s f}{\partial x_{i_s} \dots \partial x_{i_1}}(a), \quad \dots .$$

Iránymenti deriváltak $\mathbb{R}^n \to \mathbb{R}$ függvényekre

Az iránymenti derivált a parciális derivált általánosítása. A $\partial_i f(a)$ számot tekinthetjük az f függvény a pontbeli e_i irányú *iránymenti deriváltjának.*

2. definíció. *Legyen $1 \le n \in \mathbb{N}$, $f \in \mathbb{R}^n \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy $v \in \mathbb{R}^n$ egy egységvektor: $\|v\| = 1$. A f függvénynek az a pontban létezik a v irányú* ***iránymenti deriváltja****, ha a*

$$F : K(0) \ni t \mapsto f(a + tv)$$

valós-valós függvény deriválható a 0 pontban. A $F'(0)$ valós számot a f függvény a pontbeli v irányú iránymenti deriváltjának nevezzük, és a $\partial_v f(a)$ szimbólummal jelöljük.

A F függvény f-nek az a-n átmenő v irányú egyenesre vonatkozó leszűkítése.

Ha $i = 1, 2, \dots, n$ egy rögzített index és $v = e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ (ahol tehát a v vektor i-edik koordinátája 1, a többi 0), akkor

$$\partial_v f(a) = \partial_{e_i} f(a) = \partial_i f(a).$$

Az iránymenti derivált tehát a parciális derivált általánosítása.

Legyen $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ egy tetszőleges $\mathbb{R}^n$-beli egységvektor, azaz

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} = 1.$$

Figyeljük meg, hogy a

$$F(t) = f(a + tv) = f\big(a_1 + tv_1, a_2 + tv_2, \dots, a_n + tv_n\big)$$

függvényértékek kiszámolása (vagyis F deriválhatóságának a vizsgálata) az „esetek többségében” hosszadalmas feladat (ti. a parciális deriválttal ellentétben most f minden komponense változik). A következő állításban egy elégséges feltételt fogalmazunk meg az iránymenti deriválhatóságra, valamint egy egyszerű képletet adunk $\partial_v f(a)$ kiszámolására.

1. tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$, $a \in \operatorname{int} \mathcal{D}_f$ és az f függvénynek léteznek és folytonosak a parciális deriváltjai egy $K(a) \subset \mathcal{D}_f$ környezetben. Ekkor az f függvénynek az a pontból induló tetszőleges $v \in \mathbb{R}^n$ irányban létezik az iránymenti deriváltja, és*

$$\partial_v f(a) = \langle f'(a), v \rangle = \sum_{k=1}^{n} \partial_k f(a) \cdot v_k,$$

ahol $f'(a) := \big(\partial_1 f(a), \partial_2 f(a), \ldots, \partial_n f(a)\big)$ az ún. ***gradiensvektor*** *és $v = (v_1, v_2, \ldots, v_n)$* ***egységvektor****, azaz*

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = 1.$$

A totális derivált $\mathbb{R}^n \to \mathbb{R}^m$ függvényekre

Emlékeztető. A totális deriválhatóság fogalma ugyan pontosan megfelel az egyváltozós deriválhatóság (egyik) definíciójának, de a totális derivált fogalma bonyolultabb, mint egy változóban.

Idézzünk fel a valós-valós függvények deriválhatóságával kapcsolatos néhány fontos ismeretet.

• Az $f \in \mathbb{R} \to \mathbb{R}$ függvény *differenciálható* vagy *deriválható* az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben: $f \in D\{a\}$), ha létezik és véges a

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \quad \text{határérték.}$$

Ezt a határértéket az $f'(a)$ szimbólummal jelöljük, és az f függvény a pontbeli *deriváltjának* vagy *differenciálhányadosának* nevezzük.

• A többváltozós függvények esetében a differenciahányadosnak nincs közvetlen megfelelője (hiszen vektorok körében nem tudunk osztani), ezért a deriválhatóságot nem tudjuk differenciahányadosok határértékeként értelmezni. Az egyváltozós analízisben azonban láttuk azt, hogy az elsőfokú plonimomokkal való lokális közelíthetőség ekvivalens a differenciálhatósággal. Nevezetesen: Ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

$$f \in D\{a\} \quad \Longleftrightarrow \quad \begin{cases} \exists\, A \in \mathbb{R} \ \text{ és } \ \exists\, \varepsilon : \mathcal{D}_f \to \mathbb{R},\ \lim\limits_{0} \varepsilon = 0 : \\ f(a+h) - f(a) = A \cdot h + \varepsilon(h) \cdot h \ \ \big(a + h \in \mathcal{D}_f\big). \end{cases}$$

Az A szám az f függvény $a \in \operatorname{int} \mathcal{D}_f$ pontbeli deriváltja, vagyis $A = f'(a)$.

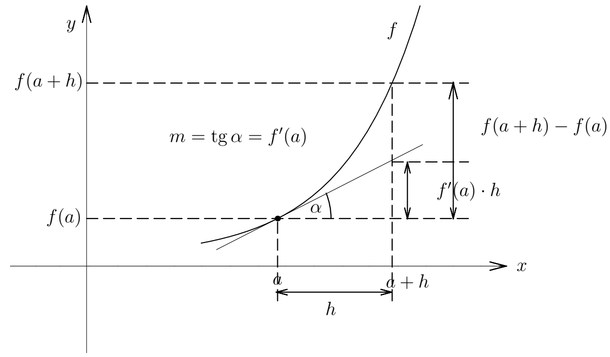
A függvényértékek megváltozása tehát

$$f(a+h) - f(a) = f'(a) \cdot h + \varepsilon(h) \cdot h \ \ \big(a + h \in \mathcal{D}_f\big).$$

A jobb oldalon álló első tag egy lineáris függvény, a második tag pedig a $\lim\limits_{0} \varepsilon = 0$ feltétel miatt az elsőhöz képest „kicsi”. Az f függvény *a pontbeli deriválhatósága tehát azt jelenti, hogy a függvény megváltozása az a pont környezetében „jól” közelíthető lineáris függvénnyel*:

$$f(a+h) - f(a) \approx f'(a) \cdot h, \quad \text{ha } h \approx 0.$$

Ezt szemlélteti a következő ábra.



• Vegyük észre, hogy a fentiekben az ε függvény szerepeltetése „kiküszöbölhető". Pontosabban: Ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

$$f \in D\{a\} \quad \Longleftrightarrow \quad \exists\, A \in \mathbb{R}: \quad \lim_{h\to 0} \frac{f(a+h)-f(a)-A\cdot h}{h} = 0.$$

Ha még azt is figyelembe vesszük, hogy egy $g \in \mathbb{R} \to \mathbb{R}$ függvény esetén $\lim\limits_{0} g = 0 \iff \lim\limits_{0} |g| = 0$, akkor végül azt kapjuk, hogy

(1) $$\boxed{f \in D\{a\} \quad \Longleftrightarrow \quad \exists\, A \in \mathbb{R}: \quad \lim_{h\to 0} \frac{|f(a+h)-f(a)-A\cdot h|}{|h|} = 0.}$$

Valós-valós függvény deriválhatóságára vonatkozó (1) ekvivalens átfogalmazás már kiterjeszthető vektor-vektor függvényre is, ha (1)-ben **az abszolút értéket a megfelelő *nomákkal*, az A valós számot pedig $m \times n$*-es mátrixszal* helyettesítjük**.

3. definíció. *Az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(1 \le n, m \in \mathbb{N})$ függvény **totálisan deriválható** az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben: $f \in D\{a\}$), ha*

$$\exists\, A \in \mathbb{R}^{m\times n}: \quad \lim_{h\to \mathbf{0}} \frac{\|f(a+h)-f(a)-A\cdot h\|^{(m)}}{\|h\|^{(n)}} = 0,$$

*ahol $\|\cdot\|^{(n)}$, illetve $\|\cdot\|^{(m)}$ az euklideszi norma az $\mathbb{R}^n$, illetve az $\mathbb{R}^m$ normált téren. Ekkor A egyértelmű, és $f'(a) := A$ az f függvény **deriváltmátrixa** az a pontban.*

Megjegyzés. A 3. definícióból következik, hogy egy $f \in \mathbb{R}^n \to \mathbb{R}^m$ vektor-vektor függvény $a \in \operatorname{int} \mathcal{D}_f$ pontbeli deriválhatósága az egyváltozós esethez hasonlóan azt jelenti, hogy a függvény megváltozása az a pont környezetében „jól" közelíthető lineáris függvénnyel:

$$f(a+h) - f(a) \approx A \cdot h, \quad \text{ha } h \approx \mathbf{0}.$$

Itt A egy $m \times n$-es mátrix, h egy n-dimenziós oszlopvektor és $\cdot$ a mátrixok közötti szorzás művelete. A jobb oldalon álló

$$L: \mathbb{R}^n \to \mathbb{R}^m, \qquad L(h) := A \cdot h$$

függvény *lineáris*, ami azt jelenti, hogy

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$$

teljesül minden $x, y \in \mathbb{R}^n$ és $\alpha, \beta \in \mathbb{R}$ esetén. $\square$

A definíció alapján a deriváltmátrix előállítása általában nem egyszerű feladat. A következő tétel azonban azt állítja, hogy a könnyen számolható parciális deriváltak segítségével egyszerűen felírható az $f \in \mathbb{R}^n \to \mathbb{R}^m$ függvény a pontbeli deriváltmátrixa.

2. tétel. (A deriváltmátrix előállítása). *Legyen* $1 \le n, m \in \mathbb{N}$,

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} \in \mathbb{R}^n \to \mathbb{R}^m,$$

ahol $f_i \in \mathbb{R}^n \to \mathbb{R}$ *az* f *függvény* i*-edik* $(i = 1, 2, \ldots, m)$ *koordinátafüggvénye. Ha* $f \in D\{a\}$, *akkor*

$$\exists\, \partial_j f_i(a) \quad (\forall\, i = 1, \ldots, m;\ j = 1, \ldots, n) \quad \textit{és}$$

$$f'(a) = \begin{bmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \cdots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \cdots & \partial_n f_2(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \cdots & \partial_n f_m(a) \end{bmatrix} \in \mathbb{R}^{m \times n}$$

a ***deriváltmátrix*** *vagy* ***Jacobi-mátrix***.

Speciális esetek:

• Ha $m = 1$, azaz f egy $\mathbb{R}^n \to \mathbb{R}$ függvény, akkor f deriváltmátrixa az a pontban

$$f'(a) = \begin{bmatrix} \partial_1 f(a) & \partial_2 f(a) & \ldots & \partial_n f(a) \end{bmatrix} \in \mathbb{R}^{1 \times n} \approx \mathbb{R}^n,$$

azaz ebben az esetben az $f'(a)$ Jacobi-mátrix tekinthető egy $\mathbb{R}^n$-beli sorvektornak, amit az f függvény a-beli **gradiensének** is nevezünk, és a $\operatorname{grad} f(a)$ szimbólummal is jelölünk.

• Ha $n = 1$, azaz $f = (f_1, f_2, \ldots, f_m)$ egy $\mathbb{R} \to \mathbb{R}^m$ függvény, akkor f deriváltmátrixa az a pontban

$$f'(a) = \begin{bmatrix} f_1'(a) \\ f_2'(a) \\ \vdots \\ f_m'(a) \end{bmatrix} \in \mathbb{R}^{m \times 1} \approx \mathbb{R}^m,$$

azaz ebben az esetben az $f'(a)$ Jacobi-mátrix tekinthető egy $\mathbb{R}^m$-beli oszlopvektornak.

A totális deriválhatóság definíciójából következik, hogy ha $f \in D\{a\}$, akkor f folytonos is az a pontban. Nem meglepő, hogy ennek az állításnak a megfordítása *nem igaz.* Másrészt az előző tétel alapján, ha $f \in D\{a\}$, akkor az összes parciális derivált létezik az a pontban. Az is sejthető azonban, hogy a parciális deriváltak létezéséből *nem következik* a totális deriválhatóság. Például az

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

függvény folytonos az origóban, itt léteznek a parciális deriváltak, de f nem totálisan deriválható az $a := (0,0)$ pontban.

Szerencsére az egyszerűen számolható parciális deriváltakból is következtethetünk a totális deriválhatóságra. Ehhez a fentiek alapján a pontbeli parciális deriváltak létezésénél többet kell feltennünk. A következő tétel egy ilyen gyakran alkalmazható *elégséges feltételt* ad a függvény totális deriválhatóságára.

3. tétel. (Elégséges feltétel a totális deriválhatóságra.)
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 az a pontnak létezik olyan $K(a) \subset \mathcal{D}_f$ környezete, amelyre minden $i = 1, \ldots, n$ index esetén a következők teljesülnek:

(a) *$\exists\, \partial_i f(x)$ minden $x \in K(a)$ pontban,*

(b) *a $\partial_i f : K(a) \to \mathbb{R}$ parciális deriváltfüggvény folytonos az a pontban.*

Ekkor az f függvény totálisan deriválható az a pontban.

Felület érintősíkja

Az egyváltozós analízisben láttuk, hogy ha $f \in D\{a\}$, akkor az f függvényt az a pont környezetében jól közelítő elsőfokú polinom nem más, mint a függvénygrafikon $\big(a, f(a)\big)$ pontbeli érintője. Most megvizsgáljuk, hogy mi felel meg ennek az állításnak az $\mathbb{R}^2 \to \mathbb{R}$ függvények körében.

Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$, $(x_0, y_0) \in \operatorname{int} \mathcal{D}_f$, és tegyük fel, hogy $f \in D\{(x_0, y_0)\}$. Ez azt jelenti, hogy

$$\begin{aligned} f(x,y) - f(x_0,y_0) &\approx \begin{bmatrix} \partial_1 f(x_0,y_0) & \partial_2 f(x_0,y_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} = \\ &= \partial_1 f(x_0,y_0)(x - x_0) + \partial_2 f(x_0,y_0)(y - y_0), \qquad \text{ha } (x,y) \approx (x_0,y_0). \end{aligned}$$

Legyen $z_0 := f(x_0, y_0)$. Ekkor

$$z - z_0 = \partial_1 f(x_0,y_0)(x - x_0) + \partial_2 f(x_0,y_0)(y - y_0) \tag{2}$$

esetén $f(x,y) \approx z$, ha $(x,y) \approx (x_0,y_0)$. A (2) egyenlet egy olyan sík egyenlete a térben, amelyik átmegy az $\big(x_0, y_0, f(x_0,y_0)\big)$ ponton és egy normálvektora

$$\mathbf{n} = \big(\partial_1 f(x_0,y_0), \partial_2 f(x_0,y_0), -1\big).$$

Megjegyzés. A háromdimenziós térben a sík általános egyenlete

$$Ax + By + Cz = D,$$

alakú, ahol az A, B, C együtthatók legalább egyike nem nulla. Ekkor a

$$\mathbf{n} = (A, B, C) \in \mathbb{R}^3$$

egy olyan vektor, amelyik a sík minden vektorára merőleges (ez a sík egy **normálvektora**).

Valóban: Tekintsünk a térben egy S síkot. Legyen $P_0 = (x_0, y_0, z_0)$ az S sík egy tetszőleges pontja, és $\mathbf{x_0}$ ebbe a pontba mutató helyvektor. Legyen $\mathbf{n} = (A, B, C)$ az S síkra merőleges nemnulla vektor (az S sík egy *normálvektora*). A tér geometriájából követlkezik, hogy egy $(x, y, z) \in \mathbb{R}^3$ koordinátájú pont (az ide mutató helyvektor $\mathbf{x}$) akkor és csak akkor eleme S-nek, ha az $\mathbf{x} - \mathbf{x_0}$ vektor merőleges az $\mathbf{n}$ vektorra, azaz

$$\langle \mathbf{x} - \mathbf{x_0}, \mathbf{n} \rangle = A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Ebből átrendezéssel adódik az állítás, ahol $D = \langle \mathbf{n}, \mathbf{x_0} \rangle$. $\square$

Most már könnyen definiálhatjuk az érintőnek megfelelő fogalmat kétváltozós függvényekre.

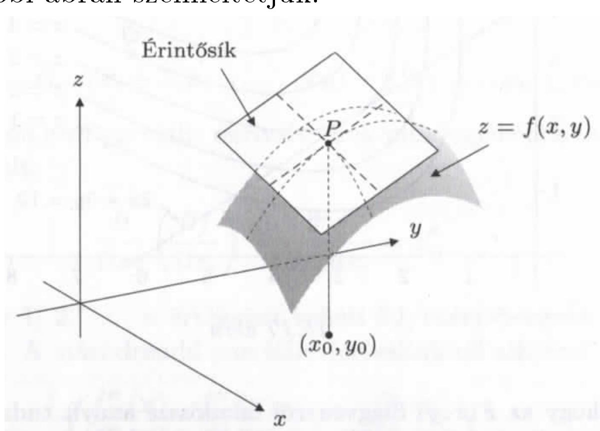
4. definíció. *Az $f \in \mathbb{R}^2 \to \mathbb{R}$ függvény grafikonjának az $\big(x_0, y_0, f(x_0, y_0)\big)$ pontban van érintősíkja, ha $f \in D\big\{(x_0, y_0)\big\}$. A érintősík egyenlete:*

$$z - f(x_0, y_0) = \partial_1 f(x_0, y_0) \cdot (x - x_0) + \partial_2 f(x_0, y_0) \cdot (y - y_0)$$

és egy normálvektora:

$$\mathbf{n} = \big(\partial_1 f(x_0, y_0), \partial_2 f(x_0, y_0), -1\big).$$

Az érintősíkot az alábbi ábrán szemléltetjük:



Deriválási szabályok

$\mathbb{R}^n \to \mathbb{R}^m$ típusú vektor-vektor függvények közötti algebrai műveletek és a totális derivált kapcsolatára az egyváltozós esethez hasonló állítások érvényesek:

• Ha $f, g \in \mathbb{R}^n \to \mathbb{R}^m$ $(1 \leq n, m \in \mathbb{N})$ és $f, g \in D\{a\}$, akkor $\forall\, \lambda, \mu \in \mathbb{R}$ esetén

$$\big(\lambda f + \mu g\big) \in D\{a\} \quad \text{és} \quad \big(\lambda f + \mu g\big)'(a) = \lambda f'(a) + \mu g'(a).$$

• Ha $m = 1$, akkor akkor az $f \cdot g$ és az f/g függvényekre az egyváltozós esethez hasonló deriválási szabályok teljesülnek.

A következő tétel függvények kompozíciójának differenciálhatóságára és deriváltjára vonatkozik.

4. tétel. (Az összetett függvény deriválhatósága, az ún. láncszabály.)
Legyen $1 \le n, m, s \in \mathbb{N}$. *Ha* $g \in \mathbb{R}^n \to \mathbb{R}^m$ *és* $g \in D\{a\}$, *továbbá* $f \in \mathbb{R}^m \to \mathbb{R}^s$ *és* $f \in D\{g(a)\}$, *akkor* $f \circ g \in D\{a\}$ *és*

$$(f \circ g)'(a) = f'\big(g(a)\big) \cdot g'(a), \tag{3}$$

ahol $\cdot$ *a mátrixok közötti szorzás műveletét jelöli.*

1. megjegyzés. Figyeljük meg a (3) képletben szereplő deriváltakat:

mivel $f \circ g \in \mathbb{R}^n \to \mathbb{R}^s$, ezért $\big(f \circ g\big)'(a) \in \mathbb{R}^{s \times n}$;

mivel $g \in \mathbb{R}^n \to \mathbb{R}^m$, ezért $g'(a) \in \mathbb{R}^{m \times n}$;

mivel $f \in \mathbb{R}^m \to \mathbb{R}^s$, ezért $f'\big(g(a)\big) \in \mathbb{R}^{s \times m}$.

Ez azt jelenti, hogy (3) bal oldalán egy $s \times n$-es mátrix áll. A jobb oldalon egy $s \times m$-es és egy $m \times n$-es mátrix ebben a sorrendben vett szorzata szerepel, ami egy $s \times n$-es mátrix.

2. megjegyzés. Tekintsük a láncszabályt a $2 \le n, m \in \mathbb{N}$ és $s = 1$ speciális esetben. Ha $g = (g_1, \ldots, g_m) \in \mathbb{R}^n \to \mathbb{R}^m$ és $g \in D\{a\}$, továbbá $f \in \mathbb{R}^m \to \mathbb{R}^1$ és $f \in D\{g(a)\}$, akkor a

$$F(x) := \big(f \circ g\big)(x) = f(g(x)) = f\big(g_1(x), \ldots, g_m(x)\big) \quad \big(x \in \mathcal{D}_g\big)$$

$\mathbb{R}^n \to \mathbb{R}$ típusú függvény differenciálható az a pontban, és a (3) képlet alapján

$$\partial_j F(a) = \partial_1 f\big(g(a)\big) \cdot \partial_j g_1(a) + \partial_2 f\big(g(a)\big) \cdot \partial_j g_2(a) + \cdots + \partial_m f\big(g(a)\big) \cdot \partial_j g_m(a) \tag{4}$$

minden $j = 1, 2, \ldots, n$ esetén.

A (4) összefüggés könnyebben megjegyezhető a következő alakban. Jelöljük f, illetve g változóit $y_1, \ldots, y_m$-mel, illetve $x_1, \ldots, x_n$-nel. Ekkor azt kapjuk, hogy ha $j = 1, 2, \ldots, n$, akkor

$$\frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial y_1} \cdot \frac{\partial g_1}{\partial x_j} + \frac{\partial f}{\partial y_2} \cdot \frac{\partial g_2}{\partial x_j} + \cdots + \frac{\partial f}{\partial y_m} \cdot \frac{\partial g_m}{\partial x_j}.$$

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* (ezt röviden az $f \in D^2\{a\}$ szimbólummal jelöltük), ha f az a pont egy $K(a) \subset \mathcal{D}_f$ környezetében deriválható (értelmezve van tehát az $f' : K(a) \to \mathbb{R}$ deriváltfüggvény) és az f' deriváltfüggvény 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 $2 \le n \in \mathbb{N}$. Ti., ha $a \in \operatorname{int} \mathcal{D}_f$ és $\exists\, K(a) \subset \mathcal{D}_f$, hogy $f \in D\big(K(a)\big)$, akkor értelmezhető 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, és $\mathbb{R}^n \to \mathbb{R}^n$ 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

$$f' = \operatorname{grad} f = \big(\partial_1 f, \partial_2 f, \ldots, \partial_n f\big) \in D\{a\}$$

azzal ekvivalens, hogy a $\partial_i f$ $(i = 1, 2, \ldots, n)$ parciális deriváltfüggvények differenciálhatók az a pontban, azaz

$$\partial_i f \in D\{a\} \qquad (i = 1, 2, \ldots, n).$$

Éppen ezért az alábbiak szerint definiáljuk egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény magasabb rendű deriválhatóságát. Az értelmezés teljes indukcióval történik, első lépésként a kétszer deriválhatóságot definiálva.

5. definíció. *Az $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \le n \in \mathbb{N})$ 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) *$\forall\, i = 1, 2, \ldots, n$ indexre $\partial_i f \in D\{a\}$.*

Az (a) feltételt röviden úgy is írhatjuk, hogy $\exists\, K(a) \subset \mathcal{D}_f$, 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.

6. definíció. *Ha az $f \in \mathbb{R}^n \to \mathbb{R}$ $(1 \leq n \in \mathbb{N})$ függvény kétszer deriválható az $a \in \operatorname{int} \mathcal{D}_f$ pontban, akkor*

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

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

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

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

(b) minden $(s-1)$-edrendű $\partial_{i_1} \partial_{i_2} \dots \partial_{i_{s-1}} f$ $\big(1 \leq i_1, i_2, \dots, i_{s-1} \leq n\big)$ parciális deriváltfüggvény 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.

5. 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*

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

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.

Megjegyzés. Teljes indukcióval igazolható a Young-tétel következő általánosítása: *Legyen 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. Tegyük fel, hogy $i_k \in \mathbb{N}$, $1 \leq i_k \leq n$ $(k = 1, 2, \dots, s)$ és $j_1, j_2, \dots, j_s$ az $i_1, i_2, \dots, i_s$ indexek egy tetszőleges permutációja. Ekkor*

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