3. előadás

PARCIÁLIS ÉS IRÁNYMENTI DERIVÁLTAK

Többváltozós függvények körében a derivált fogálmát nem kapjuk meg a valós-valós függvényeknél tanult fogalmából úgy, ahogy a folytonossággal és a határértékkel tettük. A problémát az okozza, hogy a differenciálhányados nem értelmezhető vektorok esetében, mivel osztani nem tudunk. Ez kétféle módon kerülhető el. Az egyik lehetőség az, hogy a függvényt csak bizonyos irányok mentén vizsgájuk, és az így kapott valós-valós függvényeket a már ismert módon deriváljuk. Ez nem lesz elegendő egy „teljes” deriváltfogalom megalkotásához, de számos gyakorlati előnnyel rendelkezik, ezért ezzel fogjuk kezdeni a téma tanulmányozását. Csak ezután kerül sor az „igazi” deriváltfogalom bevezetéséhez.

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

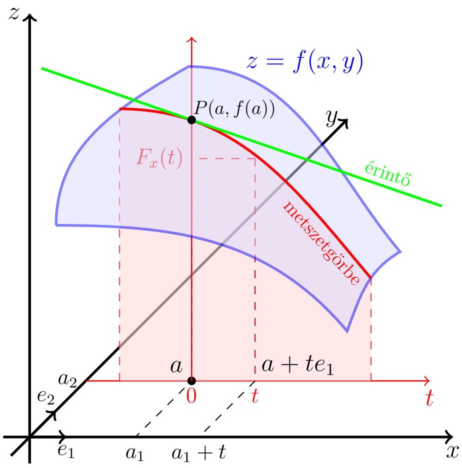
$$f \in \mathbb{R}^2 \to \mathbb{R} \quad \text{és} \quad a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f.$$

A függvény grafikonja a térben a $z = f(x, y)$ egyenletű felület. Fektessünk az a ponton át az x tengellyel párhuzamos egyenest (t tengelyt). Ennek pontjai az xy síkban

$$(a_1 + t, a_2) \qquad (t \in \mathbb{R}).$$

Vegyük a függvény értékeit ezekben a pontokban, és képezzük velük az

$$F_x(t) := f(a_1 + t, a_2) \qquad \big((a_1 + t, a_2) \in \mathcal{D}_f\big)$$



valós-valós függvényt. Ez a függvény értelmezhető a $t = 0$ pontnak egy $K(0)$ környezetében, mert $a \in \operatorname{int} \mathcal{D}_f$. Az F_x függvény képe egy, a felületen futó (metszet)görbe, vagyis a $z = f(x, y)$ egyenletű felület és az $y = a_2$ (x, z tetszőleges) egyenletű sík metszésvonala. ***Az f függvény x változó szerinti parciális deriváltját az a pontban*** (jele: $\boldsymbol{\partial_x f(a)}$) úgy értelmezzük, mint ***az F_x függvény deriváltja a 0 pontban***, feltéve, hogy a derivált létezik, azaz

$$\partial_x f(a) := F_x'(0) = \lim_{t \to 0} \frac{F_x(t) - F_x(0)}{t} = \lim_{t \to 0} \frac{f(a_1 + t, a_2) - f(a_1, a_2)}{t}.$$

$F_x'(0)$ a metszetgörbe $P(a, f(a))$ pontbeli érintőjének a meredeksége.

Az y változó szerinti parciális deriváltat hasonló módon értelmezzük. Ebben az esetben az a ponton átmenő, az y tengellyel párhuzamos egyenest vesszük. Ennek pontjai $(a_1, a_2 + t)$ $(t \in \mathbb{R})$. Vegyük a függvény értékeit ezekben a pontokban, és képezzük az

$$F_y(t) := f(a_1, a_2 + t) \qquad \big((a_1, a_2 + t) \in \mathcal{D}_f\big).$$

valós-valós függvényt. ***Az f függvény y változó szerinti parciális deriváltját az a pontban*** (jele: $\boldsymbol{\partial_y f(a)}$) így értelmezzük:

$$\partial_y f(a) := F_y'(0) = \lim_{t \to 0} \frac{F_y(t) - F_y(0)}{t} = \lim_{t \to 0} \frac{f(a_1, a_2 + t) - f(a_1, a_2)}{t}.$$

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

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

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

$$e_i := (0, 0, \ldots, 0, \underbrace{1}_{i\text{-edik}}, 0, \ldots, 0) \qquad (i = 1, 2, \ldots, n).$$

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

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

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

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

A parciális deriváltakat a következő módon számíthatjuk ki. Tekintsük a

$$G_i : K(a_i) \ni x_i \mapsto f(a_1, a_2, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n) \qquad (i = 1, 2, \ldots, n)$$

ún. ***parciális függvényt***! A G_i függvényt úgy kapunk meg az f függvényből, hogy csak az i-edik változója marad változónak, a többi az a pont megfelelő koordinátájával egyenlő. Ekkor az összetett függvény deriválási szabálya szerint

$$F_i(t) = G_i(a_i + t) \quad \Longrightarrow \quad F_i'(t) = G_i'(a_i + t) \cdot (a_i + t)' = G_i'(a_i + t) \quad \Longrightarrow \quad F_i'(0) = G_i'(a_i),$$

és így $\partial_i f(a) = G_i'(a_i)$. Más szavakkal: $\partial_i f(a)$ értéket úgy is megkaphatjuk, hogy deriváljuk az f függvényt az i-edik változója szerint úgy, hogy a többi változóját konstansnak tekintjük, és az eredményben az a pont koordinátait behelyettesítjük. Így $n = 1$, azaz valós-valós függvények esetén, a parciális derivált megegyezik a „rendes” deriválttal.

Példa: Ha ki akarjuk számítani az

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

függvény x változója szerinti parciális deriváltját az $a = (a_1, a_2)$ pontban, akkor az y változót konstansnak tekintjük, és az x változó szerint deriválunk:

$$\begin{aligned}
\partial_x f(a_1, a_2) &= (x)'_x \cdot e^{x^2+y^2} + x \cdot (e^{x^2+y^2})'_x - (3y)'_x\Big|_{x=a_1,\, y=a_2} = \\
&= 1 \cdot e^{x^2+y^2} + x \cdot e^{x^2+y^2} \cdot (x^2 + y^2)'_x - 0\Big|_{x=a_1,\, y=a_2} = \\
&= e^{x^2+y^2} + x \cdot e^{x^2+y^2} \cdot 2x\Big|_{x=a_1,\, y=a_2} = \\
&= e^{a_1^2+a_2^2} + 2a_1^2\, e^{a_1^2+a_2^2}.
\end{aligned}$$

Ugyanígy az y változó szerinti parciális deriválás során az x változót tekintjük konstansnak, és az y változó szerint deriválunk:

$$\begin{aligned}\partial_y f(a_1,a_2) &= x\cdot \left(e^{x^2+y^2}\right)'_y - (3y)'_y\Big|_{x=a_1,\,y=a_2} = \\ &= x\cdot e^{x^2+y^2}\cdot (x^2+y^2)'_y - 3\Big|_{x=a_1,\,y=a_2} = \\ &= x\cdot e^{x^2+y^2}\cdot 2y - 3\Big|_{x=a_1,\,y=a_2} = \\ &= 2a_1a_2\, e^{a_1^2+a_2^2} - 3.\end{aligned}$$

A fenti számításaink szerint például

$$\partial_x f(0,0) = 1 \qquad \text{és} \qquad \partial_y f(0,0) = -3.$$

Legyen az f függvény értelmezve $\mathbb{R}^n$ egy részhalmazán. 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 $\partial_i f(a)$. A fenti értelmezés szerint világos, hogy $\partial_i f$ egy $\mathbb{R}^n \to \mathbb{R}$ típusú függvény.

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 $i = 1,2,\ldots,n$ esetén a $\partial_i f$ parciális derivált létezik az a pont egy környezetében és a $\partial_i f$ parciális deriváltfüggvénynek létezik a j-edik ($j = 1,2,\ldots,n$) 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:

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

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

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

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 $1 \le i_1, i_2, \ldots, i_s \le n$ 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:

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

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

A parciális deriváltaknál az e_i kanonikus vektorokkal párhuzamos „irányokban” deriváltuk az $f : \mathbb{R}^n \to \mathbb{R}$ függvény értékeiből keletkezett valós-valós függvényt az a pontban. Ezt úgy fogjuk általánosítani, hogy egy tetszőleges irányban csináljuk ugyanezt.

Egy $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \le n \in \mathbb{N})$ függvény minden $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ egységvektor szerint képezhetjük a v irányú iránymenti deriváltat valamely $a = (a_1, a_2, \dots, a_n) \in \operatorname{int} \mathcal{D}_f$ pontban. Mivel v egységvektor, így

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

Az

$$F_v : K(0) \ni t \mapsto f(a_1 + tv_1, a_2 + tv_2, \dots, a_n + tv_n) = f(a + tv)$$

valós-valós függvény $t = 0$ pontban vett deriváltját (amennyiben létezik) nevezzük az f függvény v irányú iránymenti deriváltjának az a pontban.

2. Definíció. *Legyen $n \in \mathbb{N}^+$, $f \in \mathbb{R}^n \to \mathbb{R}$ 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*

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

valós-valós függvény deriválható a 0 pontban. A $F_v'(0)$ valós számot az **f függvény a pontbeli v irányú iránymenti deriváltjának** *nevezzük, és a* $\boldsymbol{\partial_v f(a)}$ *vagy a* $\boldsymbol{f_v'(a)}$ *szimbólummal jelöljük.*

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 $F_v = F_{e_i} = F_i$, és így

$$\partial_v f(a) = \partial_{e_i} f(a) = F_{e_i}'(0) = F_i'(0) = \partial_i f(a).$$

Az iránymenti derivált tehát a parciális derivált általánosítása.

Példa: Számítsuk ki az

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

függvény $v := \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ irányú iránymenti deriváltját az $a := (0, 0)$ pontban! Mivel

$$F_v(t) := f(a + tv) = f(a_1 + tv_1, a_2 + tv_2) = f\left(0 + t\,\tfrac{1}{2},\, 0 + t\,\tfrac{\sqrt{3}}{2}\right) = f\left(\tfrac{1}{2}\,t,\, \tfrac{\sqrt{3}}{2}\,t\right) =$$

$$= \frac{1}{2}\,t\,e^{\frac{1}{2}\,t + \frac{\sqrt{3}}{2}\,t} + \sin\left(\frac{1}{2}\,t \cdot \frac{\sqrt{3}}{2}\,t\right) = \frac{1}{2}\,t\,e^{\frac{1+\sqrt{3}}{2}\,t} + \sin\left(\frac{\sqrt{3}}{4}\,t^2\right),$$

így

$$F_v'(t) := \frac{1}{2}\,e^{\frac{1+\sqrt{3}}{2}\,t} + \frac{1}{2}\,t\,e^{\frac{1+\sqrt{3}}{2}\,t} \cdot \frac{1+\sqrt{3}}{2} + \cos\left(\frac{\sqrt{3}}{4}\,t^2\right) \cdot \frac{\sqrt{3}}{2}\,t.$$

Ekkor

$$\partial_v f(a) = F_v'(0) = \frac{1}{2}.$$

A fenti példából látható, hogy az iránymenti derivált kiszámítása a definíció alapján (vagyis az F_v függvény és 0-beli deriváltjának a meghatározása) hosszadalmas feladat lehet, ti. a parciális deriválttal ellentétben most f minden komponense változik. A következő állításban egy elégséges feltételt fogalmazunk meg az iránymenti deriválhatóságra, valamint egy olyan képletet adunk meg, amivel a parciális deriváltakból kiszámíthatjuk az iránymenti deriváltak értékeit.

1. Tétel. *Legyen* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ *és* $a \in \operatorname{int} \mathcal{D}_f$. *Tegyük fel, hogy az* a *pontnak van olyan* $K(a) \subset \mathcal{D}_f$ *környezete, amire minden* $i = 1, 2, \ldots, n$ *index esetén:*

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

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

Ekkor az f *függvénynek az* a *pontból induló tetszőleges* $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ *egységvektor irányban létezik az iránymenti deriváltja, és*

$$\partial_v f(a) = \sum_{i=1}^{n} \partial_i f(a) \cdot v_i = \partial_1 f(a) \cdot v_1 + \partial_2 f(a) \cdot v_2 + \cdots + \partial_n f(a) \cdot v_n.$$

Bizonyítás. Később.

Megjegyzés. A parciális és az iránymenti deriváltak kapcsolata a következő: Ha egy adott pontban egy függvénynek minden iránymenti deriváltja létezik, akkor minden parciális deriváltja is létezik, hiszen ezek speciális iránymenti deriváltak. Az állítás nem fordítható meg. Nem nehéz meggondolni, hogy az

$$f(x, y) = \begin{cases} 0 & (xy = 0) \\ 1 & (xy \neq 0) \end{cases}$$

függvénynek a $(0, 0)$ pontban csak a $\pm e_1$ és a $\pm e_2$ irányok mentén léteznek az iránymenti deriváltjai. Az 1. Tétel azt mondja ki, hogy ha a függvénynek léteznek a parciális deriváltjai a pont egy környezetében és folytonosak a pontban, akkor az állítás mégis megfordítható.

A tétel felhasználásával a számítások lényegesen egyszerűbbek lesznek. Az előző példánál:

$$f(x, y) := xe^{x+y} + \sin xy \qquad \big((x, y) \in \mathbb{R}^2\big), \qquad v := \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad a := (0, 0).$$

Mivel

$$\begin{aligned} \partial_x f(x, y) &= e^{x+y} + xe^{x+y} + y \cos xy, & \partial_x f(0, 0) &= 1, \\ \partial_y f(x, y) &= xe^{x+y} + x \cos xy, & \partial_y f(0, 0) &= 0, \end{aligned}$$

ezért

$$\partial_v f(a) = \partial_x f(0, 0) \cdot v_1 + \partial_y f(0, 0) \cdot v_2 = 1 \cdot \frac{1}{2} + 0 \cdot \frac{\sqrt{3}}{2} = \frac{1}{2}.$$