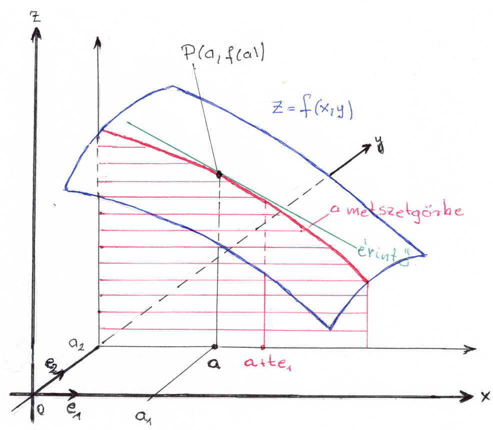
9. előadás

TÖBBVÁLTOZÓS FÜGGVÉNYEK DIFFERENCIÁLSZÁMÍTÁSA 1.

Többváltozós függvények esetén is megpróbálhatjuk átívelni a valós-valós függvényeknél alkalmazott differenciálhatóság fogalmát úgy, ahogy a folytonossággal és a határértékkel tettük. Azonban ez nem megy olyan egyszerűen. A problémát az okozza, hogy a differenciálhányados meghatározásához, mint a nevéből is kitűnik, elemek hányadosát kellene képezni, amit többdimenziós térben nem értelmezünk. A probléma elkerülésére először megpróbálunk olyan „metszeteket” kinyerni a többváltozós függvényből, amelyek már valós-valós függvényekkel leírhatók, és ezeket így a differenciálszámítás eszközeivel már tudjuk elemezni.

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. 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

$$F_x(t) := f(a_1 + t, a_2)$$

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 $P\big(a, f(a)\big)$ 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 kell figyelembe venni, amelynek pontjai $(a_1, a_2 + t)$ $(t \in \mathbb{R})$. Ez pedig az alábbi értelmezéshez vezet.

$$\partial_y f(a) := \lim_{t \to 0} \frac{f(a_1, a_2 + t) - f(a_1, a_2)}{t}.$$

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

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

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*

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

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 meg a

$$G_i \colon K(a_i) \ni x_i \mapsto f(a_1, a_2, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$

ún. ***parciális függvényt***! Mivel

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

így $\partial_i f(a) = G_i'(a_i)$. Más szavakkal, G_i az a függvény, amit úgy kapunk az f függvényből, hogy csak az i-edik változója marad változónak, a többi az a pont megfelelő koordinátájával egyenlő. Ezért $G_i'(a_i)$ az f függvény deriváltja az i-edik változója szerint úgy, hogy a többi változóját konstansnak kell tekinteni, és a végeredményben az a pont koordinátait kell behelyettesíteni. Ha $n = 1$, azaz f valós-valós függvény, akkor 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:

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

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 $\partial_i f(a)$.

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:

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

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

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

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 $1 \leq i_1, i_2, \ldots, i_s \leq n$ 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 $f\colon \mathbb{R}^n \to \mathbb{R}$ függvény értékeiből keletkezett valós-valós függvényt az a pontban. Ezt úgy fogjuk általánosítani, hogy egy tetszőleges irányban csináljuk ugyanezt.

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

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

Az

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

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 $i = 1, 2, \ldots, n$ egy rögzített index és $v = e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$ (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

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

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

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

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

Az esetek többségében az iránymenti derivált kiszámítása (vagyis F_v deriválhatóságának a vizsgálata) hosszadalmas feladat (ti. a parciális deriválttal ellentétben most f minden komponense változik). Bizonyos feltételek mellett gyorsabban tudunk eljutni a végeredményhez. Már láttuk, hogy a parciális deriváltak speciális iránymenti deriváltak, hiszen $\partial_{e_i} f(a) = \partial_i f(a)$ minden $i = 1, 2, \ldots, n$ esetén. A következő állítás egy fordított kapcsolatot mutat.

1. Tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$, illetve az f függvénynek léteznek a parciális deriváltjai egy $K(a) \subset \mathcal{D}_f$ környezetben, és ezek folytonosak az a pontban. Ekkor az f függvénynek az a pontból induló tetszőleges $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ 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.

A tétel segítségével lerövidülnek a számítások. A fenti 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

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

A totális derivált

A totális deriválhatóság fogalma ugyan pontosan megfelel az egyváltozós deriválhatóság definíciójának, de mégis bonyolultabb, mint egyváltozós esetben. 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 polinomokkal való lokális közelíthetőség ekvivalens a differenciálhatósággal. Nevezetesen: ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

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

Ekkor az A szám az f függvény a pontbeli deriváltja, vagyis $A = f'(a)$. Az f függvény *a pontbeli deriválhatósága tehát azt jelenti, hogy a függvény megváltozása az a pont környezetében „jól" közelíthető lineáris függvénnyel.*

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

$$f \in D\{a\} \qquad \Longleftrightarrow \qquad \exists A \in \mathbb{R}\colon \quad \lim_{h\to 0} \frac{f(a+h) - f(a) - A\cdot h}{h} = \lim_0 \varepsilon = 0.$$

Ha még azt is figyelembe vesszük, hogy $\lim\limits_0 \varepsilon = 0 \iff \lim\limits_0 |\varepsilon| = 0$, akkor végül azt kapjuk, hogy

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

Valós-valós függvény deriválhatóságára vonatkozó $(*)$ ekvivalens átfogalmazás már kiterjeszthető vektor-vektor függvényre is. Ehhez vegyük észre, hogy a $(*)$-ban szereplő $L(h) := A\cdot h$ tag felfogható, mint egy $L : \mathbb{R} \to \mathbb{R}$ *lineáris transzformáció*, ami azt jelenti, hogy

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

teljesül minden $x, y \in \mathbb{R}$ és $\alpha, \beta \in \mathbb{R}$ esetén. A lineáris transzformáció fogalma is hasonlóan megadható $L : \mathbb{R}^n \to \mathbb{R}^m$ típusú leképezések esetén, nevezetesen úgy, hogy $(\#)$ teljesüljön minden $x, y \in \mathbb{R}^n$ és $\alpha, \beta \in \mathbb{R}$ esetén. Tudjuk, hogy ekkor van egy $m \times n$-es A mátrix, amire $L(h) = A\cdot h$ teljesül. Itt $h \in \mathbb{R}^n$-t oszlopvektorként kell tekinteni, és a szorzás a mátrixszorzatot jelenti. ***Így $(*)$-ban az abszolút értéket a megfelelő normákkal, az A valós számot pedig egy $m \times n$-es mátrixszal helyettesítjük***.

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

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

Ekkor $f'(a) := A$ az f függvény **deriváltmátrixa** *az a pontban.*

Megjegyzések.

1. Az euklideszi normára mindig a $\|\cdot\|$ jelölést alkalmazzuk függetlenül attól, hogy hány dimenziós a benne szereplő vektor. $h \to 0$ azt jelenti, hogy a h vektor tart a 0 vektorhoz $\mathbb{R}^n$-ben, és így a határértékben lévő valós kifejezésnek tartania kell a 0 számhoz.

2. A fenti definíciót úgy lehet átfogalmazni, hogy

$$f \in D\{a\} \qquad \Longleftrightarrow \qquad \begin{cases} \exists A \in \mathbb{R}^{m\times n} \ \text{ és } \ \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}^m,\ \lim\limits_0 \varepsilon = 0 : \\ f(a+h) - f(a) = A\cdot h + \varepsilon(h)\cdot \|h\| \qquad (a+h \in \mathcal{D}_f). \end{cases}$$

2. Tétel. *Ha $f \in D\{a\}$, akkor az $f'(a)$ deriváltmátrix egyértelműen meghatározott.*

Bizonyítás. Tegyük fel, hogy az A_1 és A_2 mátrixok kielégítik a totálisan deriválhatóság definiciójában szereplő feltételeket. Ekkor

$$0 \le \frac{\|(A_1 - A_2)h\|}{\|h\|} = \frac{\|(f(a+h) - f(a) - A_2 h) - (f(a+h) - f(a) - A_1 h)\|}{\|h\|} \le$$

$$= \frac{\|f(a+h) - f(a) - A_2 h\|}{\|h\|} + \frac{\|f(a+h) - f(a) - A_1 h\|}{\|h\|} \to 0 + 0 = 0 \quad (h \to 0).$$

Ezért

$$\lim_{h\to 0}\frac{\|(A_1-A_2)h\|}{\|h\|}=0 \quad \overset{h=te_i}{\Longrightarrow} \quad 0=\lim_{t\to 0}\frac{\|(A_1-A_2)(te_i)\|}{\|te_i\|}=\lim_{t\to 0}\frac{|t|\cdot\|(A_1-A_2)e_i\|}{|t|\cdot\|e_i\|}=$$

$$=\|(A_1-A_2)e_i\|.$$

Ezért $(A_1-A_2)e_i=0$, azaz $A_1e_i=A_2e_i$ minden $i=1,2,\ldots,n$ esetén. Tehát a mátrixok mindegyik i-edik oszlopa megegyezik, és így $A_1=A_2$.

Példa: Az $f(x,y)=xy$ $\big((x,y)\in\mathbb{R}\big)$ függvény differenciálható az $a=(1,2)$ pontban és ebben a pontban vett deriváltmátrixa: $f'(1,2)=\begin{pmatrix}2 & 1\end{pmatrix}$. Valóban, ha $h=(h_1,h_2)\to(0,0)$ és $A=\begin{pmatrix}2 & 1\end{pmatrix}$, akkor

$$0\le\frac{\|f(a+h)-f(a)-A\cdot h\|}{\|h\|}=\frac{\left|(1+h_1)(2+h_2)-1\cdot 2-\begin{pmatrix}2 & 1\end{pmatrix}\cdot\begin{pmatrix}h_1\\h_2\end{pmatrix}\right|}{\|h\|}=$$

$$=\frac{|2+2h_1+h_2+h_1h_2-2-(2h_1+h_2)|}{\|h\|}=\frac{|h_1h_2|}{\|h\|}\le\left(|h_1h_2|\le\frac{h_1^2+h_2^2}{2}\right)\le\frac{\frac{\|h\|^2}{2}}{\|h\|}=\frac{\|h\|}{2}\to 0,$$

és így a definícióban szereplő határérték tart nullához.

3. Tétel. *Ha $f\in D\{a\}$, akkor $f\in C\{a\}$.*

Bizonyítás. A $\|.\|$ euklideszi és a $\|x\|_\infty:=\max\{|x_i|: i=1,2,\ldots,n\}$ normák ekvivalenciája miatt

$$f\in C\{a\}\quad\Longleftrightarrow\quad\lim_{x\to a}f(x)=f(a)\quad\Longleftrightarrow\quad\lim_{x\to a}\|f(x)-f(a)\|_\infty=0.$$

Ha $A\in\mathbb{R}^{m\times n}$ és $x\in\mathbb{R}^n$, akkor az $A\cdot x\in\mathbb{R}^m$ oszlopvektor i-edik koordinátára igaz, hogy

$$|(A\cdot x)_i|=\left|\sum_{j=1}^n a_{ij}x_j\right|\le\sum_{j=1}^n(|a_{ij}|\cdot|x_j|)\le\sum_{j=1}^n(|a_{ij}|\cdot\|x\|_\infty)=\|x\|_\infty\sum_{j=1}^n|a_{ij}|,$$

ahol a_{ij} az A mátrix i-edik sorában és j-edik oszlopában lévő eleme. Ezért

$$\|A\cdot x\|_\infty=\max\{|(A\cdot x)_i|: i=1,2,\ldots,m\}\le\|x\|_\infty\cdot\max\left\{\sum_{j=1}^n|a_{ij}|: i=1,2,\ldots,m\right\}.$$

Tehát

$$\alpha:=\max\left\{\sum_{j=1}^n|a_{ij}|: i=1,2,\ldots,m\right\}\quad\Longrightarrow\quad\|A\cdot x\|_\infty\le\alpha\|x\|_\infty.$$

Legyen $f\in D\{a\}$. Ekkor a $h=x-a\to 0$ helyettesítéssel

$$0\le\|f(x)-f(a)\|_\infty=\|f(a+h)-f(a)\|_\infty=\Big\|A\cdot h+\varepsilon(h)\cdot\|h\|\Big\|_\infty\le$$

$$\le\|A\cdot h\|_\infty+\Big\|\varepsilon(h)\cdot\|h\|\Big\|_\infty\le\alpha\|h\|_\infty+\|h\|\cdot\|\varepsilon(h)\|_\infty\to 0+0=0,$$

amiből a tétel állítása következik.

Megjegyzés. Az előző tétel nem fordítható meg. Nem nehéz igazolni, hogy az $f(x)=\|x\|$ $(x\in\mathbb{R}^n)$ függvény folytonos, de nem differenciálható az $a=0$ pontban.

4. Tétel. *Legyen az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvény koordinátafüggvényei $f_j \colon \mathbb{R}^n \supseteq \mathcal{D}_f \to \mathbb{R}$, ahol $j = 1, 2, \dots, m$, illetve $a \in \operatorname{int} \mathcal{D}_f$. Ekkor*

$$f \in D\{a\} \qquad \Longleftrightarrow \qquad f_j \in D\{a\} \quad \text{minden } j = 1, 2, \dots, m \text{ esetén.}$$

Továbbá $f_j'(a) = e_j^\top \cdot f'(a)$, azaz $f_j'(a)$ az $f'(a)$ mátrix j-edik sora.

Bizonyítás. Alkalmazzuk az

$$\|y\|_\infty \le \|y\| \le \sqrt{m}\, \|y\|_\infty \qquad (y \in \mathbb{R}^m).$$

normák ekvivalenciáját, ahol $\|x\|_\infty := \max\{|y_j| \colon j = 1, 2, \dots, m\}$. Vegyük észre, hogy

$$(f(a+h) - f(a) - A \cdot h)_j = f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h$$

minden $j = 1, 2, \dots, m$. Így a normák ekvivalenciájából:

$$0 \le \frac{\max\limits_{1 \le j \le m} |f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h|}{\|h\|} \le \frac{\|f(a+h) - f(a) - A \cdot h\|}{\|h\|}$$
$$= \sqrt{m}\, \frac{\max\limits_{1 \le j \le m} |f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h|}{\|h\|}.$$

Ebből következik a tétel állítása, hiszen ha az egyik hányados nullához tart $h \to 0$ esetén, akkor a másik is nullához tart.

Megjegyzés. Az előző tétel szerint elég lenne valós értékű függvényekkel foglalkozni.

A totális és a parciális derivált kapcsolata

A totális derivált az erősebb.

5. Tétel. *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ és $a \in \operatorname{int} \mathcal{D}_f$. Ha $f \in D\{a\}$, akkor $\forall v \in \mathbb{R}^n$ egységvektor esetén az f függvénynek van v irányú iránymenti deriváltja az a pontban, és*

$$\partial_v f(a) = f'(a) \cdot v$$

Bizonyítás. Ha $f \in D\{a\}$, akkor $\exists \varepsilon : \mathcal{D}_f \to \mathbb{R}$, $\lim\limits_{0} \varepsilon = 0$ úgy, hogy

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

Legyen $h = tv$ $(t \in \mathbb{R})$. Ekkor $\|h\| = |t| \cdot \|v\| = |t|$, hiszen $\|v\| = 1$, és így

$$f(a+tv) - f(a) = f'(a) \cdot tv + \varepsilon(tv) \cdot |t|.$$

Ezért

$$\partial_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = f'(a) \cdot v + \lim_{t \to 0} \big(\varepsilon(tv) \operatorname{sgn}(t)\big) = f'(a) \cdot v + 0 = f'(a) \cdot v,$$

hiszen $\operatorname{sgn}(t)$ korlátos függvény és $\lim\limits_{t \to 0} \varepsilon(tv) = \lim\limits_{h \to 0} \varepsilon(h) = 0$.

A totális derivált definíciója nem nyújt információt a deriváltmátrix előállításáról, de az előző tétel alapján igazolni tudjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}^m$ függvény a pontbeli deriváltmátrixa egyszerűen felírható a koordinátafüggvények a pontban vett parciális deriváltjai segítségével.

6. Tétel (A deriváltmátrix előállítása). *Legyen az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvény koordinátafüggvényei $f_j \colon \mathbb{R}^n \supseteq \mathcal{D}_f \to \mathbb{R}$, ahol $j = 1, 2, \ldots, m$, illetve $a \in \operatorname{int} \mathcal{D}_f$. Ha $f \in D\{a\}$, akkor*

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

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

az ún. **Jacobi-mátrix***.*

Bizonyítás. A 4. Tétel szerint, ha $f \in D\{a\}$, akkor $f_j \in D\{a\}$ minden $j = 1, 2, \ldots, m$ esetén, és az $f'(a)$ mátrix j-edik sora az $f_j'(a)$ sormátrix. Alkalmazzuk az 5. tételt az f_j koordinátafüggvényekre a $v = e_i$ irányok esetén! Ekkor

$$\exists\, \partial_i f_j(a) = \partial_{e_i} f_j(a) = f'(a) \cdot e_i = a_i \qquad (i = 1, 2, \ldots, n).$$

Ez azt jelenti, hogy

$$f_j'(a) = \begin{pmatrix} \partial_1 f_j(a) & \partial_2 f_j(a) & \ldots & \partial_n f_j(a) \end{pmatrix},$$

amiből a tétel állítása következik.

A parciális deriváltak létezéséből *nem következik* a totális deriválhatóság. Például gyakorlaton fogjuk igazolni, hogy az

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

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

Azonban, ha a parciális deriváltak létezésénél valamivel többet feltételezünk, akkor már tudjuk garantálni a totális deriválhatóságot. A következő tétel egy ilyen gyakran alkalmazható *elégséges feltételt* ad a függvény totális deriválhatóságára.

7. Tétel (Elégséges feltétel a totális deriválhatóságra). *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 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.

Bizonyítás. A tételt nem bizonyítjuk.

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.

Tudjuk, hogy az a pontbeli totális deriválhatóságot át lehet fogalmazni a következő módon:

$$\exists A \in \mathbb{R}^{m\times n} \;\text{ és }\; \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}^m,\; \lim_{0} \varepsilon = 0 : f(a+h) - f(a) = A \cdot h + \varepsilon(h) \cdot \|h\| \quad (a+h \in \mathcal{D}_f),$$

ahol $A = f'(a)$. Ezért

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

Az $\approx$ jelölés azt jelenti, hogy két vektor távolsága (különbségük normája) kicsi.

Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$, $a = (x_0, y_0) \in \operatorname{int} \mathcal{D}_f$ és $f \in D\{a\}$. Ha egy a-hoz közeli $(x, y) \in \mathcal{D}_f$ pontot felírunk $(x, y) = a + h$ alakban, akkor a fentiek szerint

$$(*) \qquad \begin{aligned} f(x,y) - f(x_0,y_0) &\approx \big(\partial_x f(x_0,y_0) \quad \partial_y f(x_0,y_0)\big) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = \\ &= \partial_x f(x_0,y_0)(x-x_0) + \partial_y f(x_0,y_0)(y-y_0), \qquad \text{ha } (x,y) \approx (x_0,y_0). \end{aligned}$$

Legyen $z_0 := f(x_0, y_0)$ és tekintsük a

$$z - z_0 = \partial_x f(x_0,y_0)(x-x_0) + \partial_y f(x_0,y_0)(y-y_0)$$

egyenletű síkot. Ez egy olyan sík, ami átmegy az $\big(x_0, y_0, f(x_0, y_0)\big)$ ponton és egyik normálvektora

$$\vec{n}\big(\partial_x f(x_0,y_0), \partial_y f(x_0,y_0), -1\big).$$

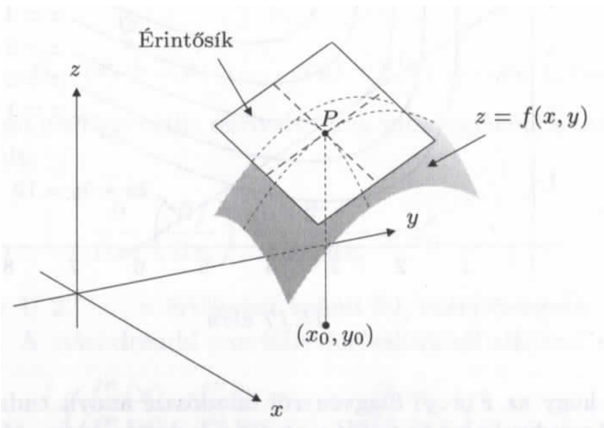
Mivel $(*)$ miatt $f(x, y) \approx z$, ha $(x, y) \approx (x_0, y_0)$, ezért érdemes a felület érintősíkját az előbbi síkkal értelmezni.

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 $f \in D\big\{(x_0, y_0)\big\}$.* **Az érintősík egyenlete**:

$$z - f(x_0,y_0) = \partial_x f(x_0,y_0) \cdot (x-x_0) + \partial_y f(x_0,y_0) \cdot (y-y_0),$$

amelynek egyik **normálvektora**: $\vec{n}\big(\partial_x f(x_0,y_0), \partial_y f(x_0,y_0), -1\big)$.

Az érintősíkot az alábbi ábrán szemléltetjük:



Megjegyzés. A háromdimenziós térben a sík általános egyenlete

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

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

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

egy olyan vektor, amelyik a sík minden vektorára merőleges (ez a sík egyik normálvektora). Ez azért igaz, mert ha $P_0 = (x_0, y_0, z_0)$ a sík egyik pontja, akkor $P = (x, y, z)$ akkor és csak akkor van rajta a síkon, ha merőleges az $\vec{n}$ vektorra, azaz $\langle \overrightarrow{P_0P}, \vec{n} \rangle = 0$. Ebből

$$\langle \overrightarrow{P_0P}, n \rangle = A(x - x_0) + B(y - y_0) + C(z - z_0) = 0,$$

amiből átrendezéssel adódik az állítás, ha $D = Ax_0 + By_0 + Cz_0$.

Példa. Írjuk fel a $z = xy$ egyenletű felület $P_0(1, 2)$ pontjához tartozó érintősíkjának az egyenletét! Legyen $f(x, y) = xy \quad \big((x, y) \in \mathbb{R}^2\big)$ és $a = (x_0, y_0)$, ahol $x_0 = 1$ és $y_0 = 2$. Ekkor

$$\partial_x f(x, y) = y \quad \text{és} \quad \partial_y f(x, y) = x \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Mivel ezek folytonos függvények, ezért f differenciálható minden pontban, azaz $f \in D\{a\}$. Másrészt

$$\partial_x f(1, 2) = 2, \quad \partial_y f(1, 2) = 1, \quad f(1, 2) = 2.$$

Ezért az érintősík egyenlete:

$$z - f(1, 2) = \partial_x f(1, 2) \cdot (x - 1) + \partial_y f(1, 2) \cdot (y - 2) \qquad \Longrightarrow \qquad z - 2 = 2(x - 1) + (y - 2),$$

azaz $2x + y - z = 2$.

Deriválási szabályok

$\mathbb{R}^n \to \mathbb{R}^m$ típusú vektor-vektor függvényekre hasonló deriválási szabályok érvényesek, mint valós-valós függvények esetén.

8. Tétel (Algebrai műveletekre vonatkozó deriválási szabályok).

- *Ha* $f, g \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ *és* $f, g \in D\{a\}$, *akkor* $\forall \alpha, \beta \in \mathbb{R}$ *esetén*

$$\big(\alpha f + \beta g\big) \in D\{a\} \qquad \text{és} \qquad \big(\alpha f + \beta g\big)'(a) = \alpha f'(a) + \beta g'(a).$$

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

Bizonyítás. A tétel állításai igazolhatók követlenül a totális derivált definíciójából a valós-valós esethez hasonló átalakításokkal.

Az összetett függvény deriválhatóságára vonatkozó állítás meglepő hasonlóságot mutat a valós-valós függvényekre jól ismert állítással. De vigyázzunk, mert a többváltozós képletben mátrix-szorzás szerepel.

9. Tétel (Az összetett függvény deriválhatósága). *Legyen* $n, m, s \in \mathbb{N}^+$*. 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*

$$(\#) \qquad \big(f \circ g\big)'(a) = f'\big(g(a)\big) \cdot g'(a),$$

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

Bizonyítás. Valós-valós esetben a bizonyítás a derivált lineáris közelítéssel történő átfogalmazásán alapszik. $\mathbb{R}^n \to \mathbb{R}^m$ típusú függvények esetén is megvan ennek a megfelelője, így a bizonyítás lényegében hasonló, de az egyes lépéseket adaptálni kell többváltozós függvények esetére.

Megjegyzések.

1. Figyeljük meg, hogy a (#) képletben szereplő mátrixszorzat elvégezhető és az eredmény olyan típusú mátrix, mint az $\big(f \circ g\big)'(a)$ deriváltmátrix, hiszen

$$\begin{array}{ccc} \big(f \circ g\big)'(a) \in \mathbb{R}^{s \times n}, & f'\big(g(a)\big) \in \mathbb{R}^{s \times m}, & g'(a) \in \mathbb{R}^{m \times n}. \\ (f \circ g \in \mathbb{R}^n \to \mathbb{R}^s) & (f \in \mathbb{R}^m \to \mathbb{R}^s) & (g \in \mathbb{R}^n \to \mathbb{R}^m) \end{array}$$

2. Koordinátafüggvényekkel felírva az összetett függvény általános alakja:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \xrightarrow{g} y = \begin{pmatrix} y_1 := g_1(x_1, x_2, \dots, x_n) \\ y_2 := g_2(x_1, x_2, \dots, x_n) \\ \vdots \\ y_m := g_m(x_1, x_2, \dots, x_n) \end{pmatrix} \xrightarrow{f} z = \begin{pmatrix} z_1 := f_1(y_1, y_2, \dots, y_m) \\ z_2 := f_2(y_1, y_2, \dots, y_m) \\ \vdots \\ z_s := f_s(y_1, y_2, \dots, y_m) \end{pmatrix}.$$

Az összetett függvény deriválhatóságáról szóló tétel szerint, ha $g \in D\{x\}$ és $f \in D\{g(x)\}$, akkor $f \circ g \in D\{x\}$, és így léteznek az összetett függvény parciális deriváltjai az x pontban. Ekkor (#) alapján minden $i = 1, 2, \dots, n$ és $j = 1, 2, \dots, s$ rögzített indexpár esetén

$$\frac{\partial z_j}{\partial x_i}(x) = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k}\big(y(x)\big) \cdot \frac{\partial y_k}{\partial x_i}(x).$$

Példa. Legyen $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x + yz$, és

$$g\colon \mathbb{R}^3 \to \mathbb{R}^3, \qquad g(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 \\ xyz \\ x \end{pmatrix}.$$

Az f és g függvények parciális deriváltjai folytonosak minden $\mathbb{R}^3$-beli pontban, ezért $f \in D(\mathbb{R}^3)$ és $g \in D(\mathbb{R}^3)$, és így $f \circ g \in D(\mathbb{R}^3)$. Vezessük be a következő jelöléseket:

$$\begin{aligned} y_1 &:= g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 \\ y_2 &:= g_2(x_1, x_2, x_3) = x_1 x_2 x_3 \qquad \text{és} \qquad z = f(y_1, y_2, y_3) = y_1 + y_2 y_3. \\ y_3 &:= g_3(x_1, x_2, x_3) = x_1 \end{aligned}$$

Legyen $F := f \circ g$. Ekkor az összetett függvény deriválhatóságáról szóló tétel szerint

$$\frac{\partial F}{\partial x_1} = \frac{\partial z}{\partial x_1} = \frac{\partial z}{\partial y_1}\cdot\frac{\partial y_1}{\partial x_1} + \frac{\partial z}{\partial y_2}\cdot\frac{\partial y_2}{\partial x_1} + \frac{\partial z}{\partial y_3}\cdot\frac{\partial y_3}{\partial x_1} =$$

$$= 1\cdot 2x_1 + y_3\cdot x_2x_3 + y_2\cdot 1 = 2x_1 + x_1\cdot x_2x_3 + x_1x_2x_3 = 2x_1 + 2x_1x_2x_3,$$

$$\frac{\partial F}{\partial x_2} = \frac{\partial z}{\partial x_2} = \frac{\partial z}{\partial y_1}\cdot\frac{\partial y_1}{\partial x_2} + \frac{\partial z}{\partial y_2}\cdot\frac{\partial y_2}{\partial x_2} + \frac{\partial z}{\partial y_3}\cdot\frac{\partial y_3}{\partial x_2} =$$

$$= 1\cdot 2x_2 + y_3\cdot x_1x_3 + y_2\cdot 0 = 2x_2 + x_1\cdot x_1x_3 = 2x_2 + x_1^2x_3,$$

$$\frac{\partial F}{\partial x_3} = \frac{\partial z}{\partial x_3} = \frac{\partial z}{\partial y_1}\cdot\frac{\partial y_1}{\partial x_3} + \frac{\partial z}{\partial y_2}\cdot\frac{\partial y_2}{\partial x_3} + \frac{\partial z}{\partial y_3}\cdot\frac{\partial y_3}{\partial x_3} =$$

$$= 1\cdot 2x_3 + y_3\cdot x_1x_2 + y_2\cdot 0 = 2x_3 + +x_1\cdot x_1x_2 = 2x_3 + x_1^2x_2.$$

Nem nehéz ellenőrizni a kapott eredmények helyességét, hiszen

$$f(g(x_1,x_2,x_3)) = z = y_1 + y_2y_3 = (x_1^2 + x_2^2 + x_3^2) + (x_1x_2x_3)\cdot x_1 = x_1^2 + x_2^2 + x_3^2 + x_1^2x_2x_3.$$

Az eredeti jelölés alkalmazásával:

$$\frac{\partial(f\circ g)}{\partial x} = 2x + 2xyz, \qquad \frac{\partial(f\circ g)}{\partial y} = 2y + x^2z, \qquad \frac{\partial(f\circ g)}{\partial z} = 2z + x^2y,$$

Gradiens vektor

A 3. előadáson megismertük az $\mathbb{R}^n \to \mathbb{R}$ típusú függvények iránymenti deriváltját. Egy olyan tételt mondtunk ki, amely szerint bizonyos feltételek mellett a $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ egységvektor irányú iránymenti derivált az a pontban kiszámolható a parciális deriváltak segítségével a következő módon:

$$(\star) \qquad \partial_v f(a) = \sum_{i=1}^{n} \partial_i f(a)\cdot v_i.$$

Ehhez az kellet, hogy létezzenek az f függvény parciális deriváltjai egy $K(a) \subset \mathcal{D}_f$ környezetben, és folytonosak legyenek az a pontban. Ez éppen az elégséges feltétel a totális deriválhatóságra, tehát $f \in D\{a\}$. De ekkor $(\star)$ következik a 4. előadás 4. tételéből, ahol azt igazoltuk, hogy ha f totális deriváltható az a pontban, akkor ott bármilyen v irányban van iránymenti deriváltja, és

$$\partial_v f(a) = f'(a)v.$$

Ez azért igaz, mert valós értékű függvények esetében

$$f'(a) = \begin{pmatrix}\partial_1 f(a) & \partial_1 f(a) & \ldots & \partial_n f(a)\end{pmatrix},$$

és így $(\star)$ adódik a fenti sormátrix és v koordinátaiból álló oszlopmátrix szorzatából. Érdekes, hogy $(\star)$ is igazolható az összetett függvény deriválhatóságáról szóló tétel alapján. Valóban, a definíció szerint $\partial_v f(a)$ nem más, mint az $F_v(t) = f(a+tv)$ függvény deriváltja a $t = 0$ pontban. De $F_v = f \circ g$, ahol

$$g\colon \mathbb{R} \supseteq K(0) \ni t \mapsto a + tv \in \mathbb{R}^n.$$

Látható, hogy $g(0) = a$, illetve nem nehéz igazolni, hogy $g'(0) = v$ (g konstans + lineáris függvény). Ezért

$$\partial_v f(a) = F_v'(0) = \big(f\circ g\big)'(0) = f'\big(g(0)\big)\cdot g'(0) = f'(a)v.$$

A $(\star)$ képletből következik, hogy az összes v egységvektor közül az, amelynek iránya olyan, mint a

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

(ún. ***gradiens vektor***) vektor iránya, azzal a tulajdonsággal rendelkezik, hogy az ebbe az irányba mutató iránymenti derivált a legnagyobb. A szóban forgó egységvektor:

$$w := \frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|} = \left(\frac{\partial_1 f(a)}{\|\operatorname{grad} f(a)\|}, \frac{\partial_2 f(a)}{\|\operatorname{grad} f(a)\|}, \ldots, \frac{\partial_n f(a)}{\|\operatorname{grad} f(a)\|}\right).$$

Ezért $(\star)$ szerint

$$\max\{\partial_v f(a) \colon v \in \mathbb{R}^n, \|v\| = 1\} = \partial_w f(a) = \sum_{i=1}^{n} \left(\partial_i f(a) \cdot \frac{\partial_i f(a)}{\|\operatorname{grad} f(a)\|}\right) =$$

$$= \frac{1}{\|\operatorname{grad} f(a)\|} \sum_{i=1}^{n} \big(\partial_i f(a)\big)^2 = \frac{1}{\|\operatorname{grad} f(a)\|} \cdot \|\operatorname{grad} f(a)\|^2 = \|\operatorname{grad} f(a)\|.$$

Megjegyzés. A gradiens irányában a leggyorsabb a függvény növekedése, ellentétes irányban a leggyorsabb a csökkenése. A gradiensre merőleges irányban az iránymenti derivált nulla.