4. előadás

A TOTÁLIS DERIVÁLT

Most rátérünk a másik deriváltfogalom bevezetésére, de előtte 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.

Már említettünk, hogy 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 azt láttuk, hogy az elsőfokú polinomokkal való lokális közelíthetőség (lineáris közelítés) ekvivalens a differenciálhatósággal. Nevezetesen: ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

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

Ekkor az A szám az f függvény a pontbeli deriváltja, vagyis $A = f'(a)$.

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

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

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

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

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

A 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. Igazolható, hogy ebben az esetben egyértelműen létezik egy olyan $m \times n$-es A mátrix, amire

$$L(h) = A \cdot h \qquad (h \in \mathbb{R}^n)$$

teljesül. Itt $h \in \mathbb{R}^n$-t oszlopvektorként kell tekinteni, és a szorzás a mátrixszorzatot jelenti.

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

1. 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*

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

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

Megjegyzések.

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

2. Könnyű meggondolni, hogy

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

Egy $f \in \mathbb{R}^n \to \mathbb{R}^m$ vektor-vektor függvény $a \in \mathcal{D}_f$ pontbeli (totális) deriválhatósága az egyváltozós esethez hasonlóan tehát azt jelenti, hogy a függvény megváltozása az a pont környezetében „jól” közelíthető lineáris függvénnyel:

$$\mathbb{R}^m \ni f(a+h) - f(a) \approx A \cdot h \in \mathbb{R}^n, \qquad \text{ha } h \approx 0 \in \mathbb{R}^n.$$

1. Tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}^m$ és $f \in D\{a\}$. Ekkor az $f'(a) \in \mathbb{R}^{m\times n}$ deriváltmátrix egyértelműen meghatározott.*

Bizonyítás. Tegyük fel, hogy az A_1 és A_2 mátrixok kielégítik a totálisan deriválhatóság definíciójában szereplő feltételeket. Mivel $(A_1 - A_2) \cdot h = A_1 \cdot h - A_2 \cdot h$, így

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

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

A közrefogási elv miatt

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

$$= \left\|(A_1 - A_2)\cdot e_i\right\|.$$

Így $(A_1 - A_2)\cdot e_i = 0$, azaz $A_1 \cdot e_i = A_2 \cdot e_i$ 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

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

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

és így a közrefogási elv miatt a definícióban szereplő határérték tart nullához.

Az egyváltozós esethez hasonlóan a pontbeli differenciálhatóság a többváltozós függvények körében is „erősebb” tulajdonság a pontbeli folytonosságnál.

2. Tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}^m$ és $f \in D\{a\}$. Ekkor*

$$f \in D\{a\} \qquad \Longrightarrow \qquad f \in C\{a\}.$$

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

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

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

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

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

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

Tehát

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

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

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

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

így a közrefogási elvből $\lim\limits_{x \to a} \Big\|f(x) - f(a)\Big\|_\infty = 0$, amiből a tétel állítása következik.

Megjegyzés. Többváltozós esetben is igaz, hogy az előző tétel nem fordítható meg, azaz van olyan folytonos függvény, ami nem differenciálható egy adott pontban. Pl. az $f(x) = \|x\|$ $(x \in \mathbb{R}^n)$ függvény folytonos, de nem differenciálható az $a = 0$ pontban. A folytonosságot már igazoltuk. Továbbá, indirekt módon tegyük fel, hogy $\exists f'(0) = A = \begin{pmatrix} a_1 & a_1 & \dots & a_n \end{pmatrix}$ és legyen $h = te_1$ $(t \in \mathbb{R})$. Ekkor

$$0 = \lim_{h \to 0} \frac{\left| \|0 + h\| - \|0\| - A \cdot h \right|}{\|h\|} = \lim_{h \to 0} \left| \frac{\|h\| - A \cdot h}{\|h\|} \right| = \lim_{h \to 0} \left| 1 - \frac{A \cdot h}{\|h\|} \right| =$$

$$= \lim_{t \to 0} \left| 1 - \frac{t(A \cdot e_1)}{|t| \, \|e_1\|} \right| = \lim_{t \to 0} |1 - \operatorname{sgn}(t) a_1|,$$

Ebből azonban

$$0 = \lim_{t \to 0+0} \left| 1 - \operatorname{sgn}(t) a_1 \right| = |1 - a_1| \qquad \text{és} \qquad 0 = \lim_{t \to 0-0} \left| 1 - \operatorname{sgn}(t) a_1 \right| = |1 + a_1|,$$

amiből következik, hogy $a_1 = 1$ és $a_1 = -1$, de ez nem lehetséges.

A következő tétel azt állítja, hogy vektorértékű függvények totális deriválhatósága ekvivalens a koordinátafüggvények deriválhatóságával.

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

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

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 $\|y\|_\infty := \max\left\{ |y_j| \;\middle|\; j = 1, 2, \dots, m \right\}$. Vegyük észre, hogy

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

minden $j = 1, 2, \dots, m$, továbbá $e_j^\top \cdot (A \cdot h) = (e_j^\top \cdot A) \cdot h$. Így a normák ekvivalenciájából:

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

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

Így a közrefogási elv szerint $h \to 0$ határátmenettel, ha az egyik hányados tart nullához, akkor a másik hányados szintén tart nullához, amiből a tétel állítása már következik.

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

A totális és a parciális deriváltak kapcsolata

A totális derivált az erősebb.

4. 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 $\exists \varepsilon \in \mathbb{R}^n \to \mathbb{R}$, $\lim\limits_{0} \varepsilon = 0$ úgy, hogy

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

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

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

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

Megjegyzés. Az előző tétel állítása nem fordítható meg. A gyakorlaton meg fogjuk mutatni, hogy például az

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

függvény folytonos a $(0,0)$ pontban, itt minden irányban deriválható, de f nem totálisan differenciálható ebben a pontban.

Az eddig bemutatott deriváltak között a totális derivált az, ami betölti az egyváltozós derivált szerepét a többváltozós függvényekkel kapcsolatos állításokban. Azonban a definíció alapján a deriváltmátrix előállítása általában nem egyszerű feladat. A következő tétel azt állítja, hogy a deriváltmátrixban parciális deriváltak állnak, ami jelentősen leegyszerűsíti annak meghatározását.

5. Tétel (A deriváltmátrix előállítása). *Legyen* $f = (f_1, f_2, \ldots, f_m) \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$, *ahol* $f_j \in \mathbb{R}^n \to \mathbb{R}$ $(j = 1, 2, \ldots, m)$ *az* f *függvény* j*-edik, koordinátafüggvénye. Ha* $f \in D\{a\}$, *akkor*

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

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

az ún. **Jacobi-mátrix**.

Bizonyítás. Alkalmazzuk a 4. Tétel állítását az f_j koordinátafüggvényekre a $v = e_i$ irányok mentén! Ekkor

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

azaz $\partial_i f_j(a)$ az $f_j'(a)$ sormátrix i-edik eleme, és így

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

Másrészt a 3. 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. Ebből már következik a tétel állítása.

A parciális deriváltak létezéséből *nem következik* a totális deriválhatóság. Például látni fogjuk gyakorlaton, 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. A tételt nem bizonyítjuk.

6. 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 van olyan $K(a) \subset \mathcal{D}_f$ környezete, amire minden $i = 1, 2, \ldots, n$ index esetén a következők teljesülnek:*

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 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 $(a, f(a))$ 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 deriválthatóságot át lehet fogalmazni a következő módon:

$$\exists A \in \mathbb{R}^{m \times n} \ \text{ és } \ \exists \varepsilon \in \mathbb{R}^n \to \mathbb{R}^m, \ \lim_0 \varepsilon = 0 : f(a + h) - f(a) = A \cdot h + \varepsilon(h)\, \|h\| \quad (a + h \in \mathcal{D}_f),$$

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

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

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.

2. 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***:*

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

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

Megjegyzés. Egy $P_0 = (x_0, y_0, z_0)$ ponton átmenő és $\vec{n}(A, B, C)$ normálvektorral rendelkező sík egyenlete a háromdimenziós térben:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0,$$

hiszen egy $P = (x, y, z)$ pont akkor és csak akkor van rajta ezen a síkon, ha $\overrightarrow{P_0P}$ merőleges az $\vec{n}$ vektorra, azaz $\langle \overrightarrow{P_0P}, \vec{n} \rangle = 0$. Ezért a sík általános egyenlete

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

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

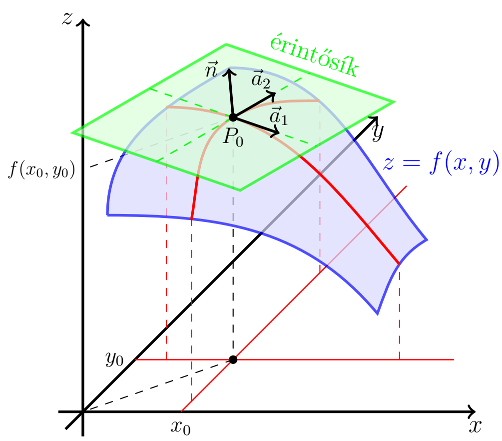
A parciális deriváltak geometriai jelentéséből is megállapítható, hogy az érintősíkot a megadott egyenlettel érdemes értelmezni. Ti. elvárás, hogy az x és y metszetgörbék $P_0 = (x_0, y_0, f(x_0, y_0))$ pontbeli érintői rajta legyenek az érintősíkon. Ezért az

$$\vec{a}_1(1, 0, \partial_x f(x_0, y_0)) \quad \text{és az} \quad \vec{a}_2(0, 1, \partial_y f(x_0, y_0))$$

nem egymással párhuzamos vektorok az érintősíkra illeszkednek. Ekkor a vektorok vektoriális szorzata

$$\vec{a}_1 \times \vec{a}_2 = \vec{n}\Big(\partial_x f(x_0, y_0), \partial_y f(x_0, y_0), -1\Big)$$

az érintősík normálvektora.



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

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

Ezek folytonos függvények, ezért f differenciálható minden pontban, azaz $f \in D\{a\}$. Ezért az érintősík egyenlete:

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

Másrészt $\partial_x f(1, 2) = 2$, $\partial_y f(1, 2) = 1$, $f(1, 2) = 2$. Behelyettesítés után:

$$z - 2 = 2(x - 1) + (y - 2) \quad \Longrightarrow \quad 2x + y - z = 2.$$

Deriválási szabályok

$\mathbb{R}^n \to \mathbb{R}^m$ típusú vektor-vektor függvényekre hasonló deriválási szabályok érvényesek, mint valós-valós függvények esetén. A tételeket nem igazoljuk, bizonyításuk hasonló technikákkal történik, mint valós-valós esetben.

7. 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.*

8. 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.

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

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

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

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