Budapesti Műszaki és Gazdaságtudományi Egyetem Gépészmérnöki Kar
Energetika és Mechatronika BSc szakok

Matematika A2H — előadás jegyzetek a XIII. hétre

Többváltozós differenciálás

Kiadva: 2016. május 21.

Múltkor felírtuk egy $f:\mathbb{R}^m\to\mathbb{R}^k$, $C^n(\mathbb{R}^m)$-beli függvény n-ed rendű l.j.n.r.k az $\underline{a}\in\mathbb{R}^m$ pont körül:

$$f(\underline{x})\approx f(\underline{a})+Df(\underline{a})(\underline{x}-\underline{a})+\frac{1}{2}D^{(2)}f(\underline{a})(\underline{x}-\underline{a})^2+\cdots+\frac{1}{n!}D^{(n)}f(\underline{a})(\underline{x}-\underline{a})^n.$$

A legnehezebb ennek megfelelő értelmezése. Itt $D^{(n)}f(a)$ egy homogén n-ed rendű polinom (n-ed rendű forma, n-ed rendű alak), amit a $\left[\frac{\partial^n f_j}{\partial x_{i_1},\dots\partial x_{i_n}}\right]_{j=1,i_k=1}^{k,m}$ $k\times m\times\cdots\times m$ rendű hipermátrixszal lehet felírni. Hatása a $\Delta\underline{x}=\underline{x}-\underline{a}=(h_1,\dots,h_m)$ megváltozásmennyiségre

$$D^{(n)}f(\underline{a})(\Delta\underline{x})^n=\left[(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m})^n f\right](\underline{a}).$$

Speciális esetek:

1.) Ha $\underline{x}-\underline{a}=t\underline{v}$ (azaz a rögzített $\underline{v}$ irányú egyenesre megszorítjuk f-et). Speciálisan, ha $k=1$ (vagy j-t is rögzítjük), akkor a $\varphi(t)=f_j(\underline{a}+t\underline{v})\in C^n(\mathbb{R})$ függvényt kapjuk.

   Felírható erre a Taylor-formula, Lagrange-maradéktag, általában a Lagrange- és Cauchy-k.é.t.-ek. Valójában ez egyváltozós elmélet – nem részletezzük. A többváltozós esetben-formuláknál, tételeknél ekkor iránymenti deriváltak vannak.

2.) Lineáris közelítés: $\exists$ l.j.l.k. $\underline{a}$ körül $\Leftrightarrow$ f differenciálható.

   Ekkor $f(\underline{x})-f(\underline{a})=Df(\underline{a})\Delta x$ ($\Delta\underline{x}:=\underline{x}-\underline{a}$) (csak $k=1$ estén: ez az érintősík).

   $Df(a)$ mindig egy lineáris leképezés. Általában $\mathbb{R}^{k\times m}$ mátrixszal szorzásként írható fel (adott – sztenderd – bázisban); $k=1$-re $\nabla f(a)$ sorvektor, skaláris szorzás. $\mathbb{R}^{k\times m}$-nél a j-edik sor $\nabla f_j(a)$– éppen emiatt lesz $Df(a)\cdot\Delta x=\begin{bmatrix}\cdots\\ \nabla f_j\cdot\Delta x\\ \cdots\end{bmatrix}_{j=1}^{n}$ mátrixszorzás.

3.) Kvadratikus közelítés: $k=1$ eseben a $D^{(2)}2f(a)=H$ Hesse-mátrix $H=\left[\frac{\partial^2 f(a)}{\partial x_i\partial x_j}\right]_{i=1,j=1}^{m,m}$; hatása kifejezhető mint $\Delta\underline{x}^T\cdot H\cdot\Delta\underline{x}=\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 f(a)}{\partial x_i\partial x_j}\Delta x_i\Delta x_j=\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 f(a)}{\partial x_i\partial x_j}h_ih_j.$

   (Young tétele szerint, ha $f\in C^2(\mathbb{R}^m)$, akkor H szimmetrikus mátrix.)

A legfontosabb kérdés a kvadratikus közelítés kapcsán, hogy az érintősík alatt, felett lesz-e.
a.) támaszsíkok, konvexitás-konkávitás   b.) vízsz. érintősík esetén: maximum-minimum?

Ha $f \in C^2(B)$ $\underline{a}$ egy $B \subset \mathbb{R}^m$ környezetében, akkor a másodrendű közelítésből következtethetünk magának f-nek a viselkedésére is.

1. Tétel. *Ha $Q = [\underline{u}_1, \dots, \underline{u}_m]$ a H Hesse-mátrix lineárisan független sajátvektoraiból képzett ortogonális (ortonormált) mátrix, akkor $H = QDQ^T$, ahol* $D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & \\ & \dots & & \\ 0 & \dots & & \lambda_n \end{bmatrix}$, *$\lambda_j$ sajátértékek, $H\underline{u}_j = \lambda_j \underline{u}_j$, és így $\Delta\underline{x}^T H \Delta\underline{x} = (\Delta\underline{x}^T Q) D (Q^T \Delta\underline{x}) = (Q^T\Delta\underline{x})^T D (Q^T \Delta\underline{x}) = \underline{y} D \underline{y}$ diagonális alakban írható a kvadratikus alak (főtengely-transzformáció).*

Ezért $\underline{x}^T H \underline{x} > \underline{0}$ $(\forall \underline{x} \neq \underline{0}) \Leftrightarrow \underline{y}^T D \underline{y} > 0$ $(\forall \underline{y} \neq 0) \Leftrightarrow \sum_{j=0}^{m} \lambda_j \underline{y}_j^2 > 0 (\forall \underline{y} \neq 0)$.

Ezt nevezzük (szigorúan) pozitív definit mátrixnak; (Jelölés: $H \gg 0$).

A H szimmetrikus mátrixra $H \gg 0 \Leftrightarrow \forall j = 1, \dots, m, \lambda_j > 0$.

Hasonlóan pozitív szemidefinit (≥ 0), negatív definit (< 0), negatív szemidefinit (≤ 0) vagy indefinit $(\exists < 0,\ \exists > 0)$ esetekre.

A sajátértékek előjeléből tehát leolvashatjuk a kvadratikus alak jellegét, és így T_2-nek és f-nek is az érintősíkhoz képesti menetét.

2. Tétel. *Ha egy D konvex tartományon $f \in C^2(D)$, akkor f konvex (konkáv) a D-n $\Leftrightarrow$ Hesse mátrixa pozitív (negatív) szemidefinit $\Leftrightarrow$ Hesse mátrixa (minden pontban) $\lambda_j \geq 0$ (≤ 0) sajátértékű.*

1. Megjegyzés. *Ha $m = 1$; f konvex I-n $\Leftrightarrow f'' \geq 0$ (konkáv: $f'' \leq 0$).*

2. Megjegyzés. *Egyetlen pontban $f''(a) \geq 0$ már $m = 1$ változóban sem volt elég; ekkor nem tudjuk, hogy ha $f''(a) = 0$, akkor f alatta vagy felette lesz az érintőnek. Több változóban még inkább: lehet nyeregpont stb.*

1. Következmény. *(tétel következménye) Ha $H = D^{(2)} f(\underline{a}) \underset{=}{\gg} 0$ (pozitív szemidefinit) egy $\underline{a} \in \mathbb{R}^m$ egy B környezetében, akkor ott $T_2(f)$, és vele együtt f, az érintő felett halad – az érintősík egyúttal támaszsík lesz.*

2. Következmény. *Szélsőértékek meghatározhatóak.*

Szükséges feltétel: Vízszintes érintősík – kritikus pont, $\nabla f(\underline{a}) = \underline{0}$.
Ha ez megvan, elégséges

a) $D^2 f(a) \gg 0$ (min.) $\ll 0$ (max.)

b) $D^2 f \underset{=}{\gg} 0$ B-ben (min.) $\underset{=}{\ll}$ (max.)

Ha $f \in C^2(B)$, akkor ugye H folytonos: a) $\Rightarrow$ b).

Az a) feltétel elégséges, de nem szükséges – van példa, ahol csak b) van.

Viszont ha $f \notin C^2(B)$, csak $\exists D^{(2)} f(a)$, az a) akkor is elégséges.

1. Definíció. *Legyen $H = [h_{ij}]_{i,j=1}^n$ szimmetrikus mátrix. Ekkor H k-adrendű főminorai az $M_k := \det[h_{ij}]_{i,j=1}^k$ determináns-értékek.*

3. Tétel (Kritérium pozitív / negatív definitségre). *Legyen H szimmetrikus mátrix. Ekkor a) $H \gg 0 \Leftrightarrow M_k > 0$ $(\forall\ 1 \leq k \leq n)$;   és   b) $H \ll 0 \Leftrightarrow (-1)^k M^k > 0$ $(\forall\ 1 \leq k \leq n)$.*

Függvények kompozíciója – az általános láncszabály

Tegyük fel, hogy $f : \mathbb{R}^m \to \mathbb{R}^k$ és $g : \mathbb{R}^n \to \mathbb{R}^m$; $f \in C^1(\mathcal{E}), g \in C^1(\mathcal{D})$, $\underline{a} \in \mathcal{D}$, $R_g \subset \mathcal{E}$, $g(\underline{a}) = \underline{b} \in \mathcal{E}$. Legyen $h := f \circ g : \mathbb{R}^n \to \mathbb{R}^k$. Írjuk fel az $\underline{a} \in \mathcal{D}$-ben a $h := f \circ g$ függvény lineáris közelítését!

1. Ha h differenciálható: $h(\underline{x}) \approx h(\underline{a}) + Dh(\underline{a})\Delta\underline{x} \qquad (Dh(\underline{a}) \in \mathbb{R}^{k\times n})$
2. $h(\underline{x}) = f(g(\underline{x})) \approx f(\underline{b}) + Df(\underline{b})\Delta\underline{y} \approx f(\underline{b}) + Df(\underline{b}) \cdot (Dg(\underline{a}) \cdot \Delta\underline{x})$

A kettőt egybevetve: fennáll ($\exists$ l.j.l.k.) és $Dh(\underline{a}) = Df(\underline{b})Dg(\underline{a})$, azaz $D(f \circ g)(\underline{a}) = Df(g(\underline{a})) \cdot Dg(\underline{a})$ (mátrixszorzás asszociatív!).

Inverz függvény deriváltja: Tegyük fel, hogy $f \in C^1(\mathcal{D})$, $\mathcal{D} \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}^n$ és $\underline{a} \in \mathcal{D}$. Ekkor $\exists\underline{a}$-nak olyan $\mathcal{B}$ környezete, amelyen f injektív (invertálható), és az $\mathcal{E} := f(\mathcal{B})$ halmazon $f^{-1} \in C^1(\mathcal{E}) \Leftrightarrow Df(\underline{a})$ nem szinguláris.

Ekkor $Df^{-1}(\underline{b}) = Df(\underline{a})^{-1}$, $Df^{-1}(\underline{y}) = Df(\underline{x})^{-1}$ $(f(\underline{a}) = \underline{b},\ f(\underline{x}) = \underline{y})$.

Bizonyítás: C^1 miatt $Df(\underline{a})$ nem szinguláris $\Leftrightarrow \exists\mathcal{B},\mathcal{E}, Df(x)$ nem szinguláris $\Leftrightarrow \lambda = 0$ nem sajátérték, nen lehet szemidefinit.

Ha $\exists f^{-1} \Rightarrow h := f^{-1} \circ f =$identitás $\approx Dh(a) \cdot (\underline{x} - \underline{a}) + \underline{a}$, $Dh(\underline{a}) = I_n$.

Láncszabályból: ha $f^{-1} \in C^1$, $f \in C^1 \Rightarrow I_n = Dh(\underline{a}) = Df^{-1}(\underline{b}) \cdot Df(\underline{a})$.

3. Megjegyzés. *Ha $Df(\underline{a})$ szinguláris, akkor esetleg létezhet még f^{-1}, de biztosan nem lesz differenciálható; $\nexists M$ mátrix, hogy $MDf(\underline{a}) = I_n$ legyen.*

Példák ($n = 1$-re): $f(x) = x$ – $\exists$ inverz, differenciálható; $f(x) = x^2$ – $\nexists$ inverz 0 körül; $f(x) = x^3$ – $f'(0) = 0$, de mégis $\exists$ inverz, viszont $\sqrt[3]{x}$ nem differenciálható.

Implicit deriválás

Tegyük fel, hogy az $F(x_1,\dots,x_n)$ függvény csak implicit módon, egy egyenlettel adott, mégis ki akajuk számítani $\nabla F(\underline{a})$-t. Jelölje $z := F(x_1,\dots,x_n)$, és legyen az adott egyenlet $f(x_1,\dots,x_n,z) = 0$. Ezt kompozícióval felírhatjuk: $g : x \to (x, F(x))$ $(\mathbb{R}^n \to \mathbb{R}^{n+1})$ és $f : (x,z) \to f(x_1,\dots,x_n,z) \in \mathbb{R}$ $(\mathbb{R}^{n+1} \to \mathbb{R})$ kompozíciója $h := f \circ g : \mathbb{R}^n \to \mathbb{R}$, és a feltétel szerint az adott $D \subset \mathbb{R}^n$-en ez azonosan konstans 0. Jelölje $g(\underline{a}) = (\underline{a}, F(\underline{a})) = (\underline{a}, b) = \underline{b}$.

Ekkor $\nabla h(\underline{a}) = \mathbf{0} = \nabla f(\underline{b}) \cdot Dg(\underline{a})$. $\nabla f(\underline{b}) = \nabla f(\underline{a}, F(\underline{a})) = \nabla f(\underline{a}, b) \in \mathbb{R}^{1\times(n+1)}$.

$$Dg(a) \in \mathbb{R}^{(n+1)\times n} = \begin{bmatrix} I_n \\ \nabla F \end{bmatrix} = \begin{bmatrix} \dots & e_{ij} & \dots \\ \dots & \frac{\partial F}{\partial x_j}(\underline{a}) & \dots \end{bmatrix}$$

Tehát $\mathbf{0} = \nabla f(\underline{b})Dg(\underline{a}) \Leftrightarrow j = 1,\dots,n$-re $\nabla f(\underline{b}) \cdot Dg(\underline{a})^{(j)} = 0$, azaz

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{b})e_{ij} + \frac{\partial f}{\partial z}(\underline{b})\frac{\partial F}{\partial x_j}(\underline{a}) = \frac{\partial f}{\partial x_j}(\underline{b}) + \frac{\partial f}{\partial z}(\underline{b})\frac{\partial F}{\partial x_j}(\underline{a}) = 0.$$

Ebből: $\nabla F(a) = \left(\dots, \dfrac{\partial F}{\partial x_j}(\underline{a}), \dots\right) = \left(\dots, \dfrac{-\frac{\partial f}{\partial x_j}(\underline{b})}{\frac{\partial f}{\partial z}(\underline{b})}, \dots\right).$

4. Megjegyzés. *Ha $\underline{z} = F(\underline{x})$ is vektor, és impliciten az $f(\underline{x}, \underline{z}) = 0$ egyenlettel van megadva, akkor $-\dfrac{\partial f}{\partial x_j}(\underline{b}) = \sum_{k=1}^{m} \dfrac{\partial f}{\partial z_k}(\underline{b})\dfrac{\partial F_k}{\partial x_j}(\underline{a})$ $(j = 1,\dots,n)$ adódik. Ez azonban nem határozza meg egyértelműen F-et, illetve az F_k-kat, ∇F_k-t.*

Ha több egyenletünk is van F-re, azaz $f = (\dots, f_\ell, \dots)$ is vektor értékű függvény — tipikusan a megfelelő meghatározottsághoz ugyanannyi egyenlet van, ahány koordinátafüggvény, tehát $F : \mathbb{R}^b \to \mathbb{R}^m$, és $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$, – akkor ezek az egyenletek f_ℓ koordinátafüggvényenként fennállnak, azaz

$$\left[-\frac{\partial f_\ell}{\partial x_j}(\underline{b}) \right]_{\ell=1,j=1}^{m,n} = \left[\sum_{k=1}^{m} \frac{\partial f_\ell}{\partial z_k}(\underline{b}) \cdot \frac{\partial F_k}{\partial x_j}(\underline{a}) \right]_{\ell=1,j=1}^{m,n}.$$

Mátrix alakban formálisan tehát $-\frac{\partial f}{\partial \underline{x}}(\underline{b}) = \frac{\partial f}{\partial \underline{z}}(\underline{b}) \cdot \frac{\partial \underline{F}}{\partial \underline{x}}(\underline{a})$, amiből jó esetben megoldható az egyenlet $\frac{\partial \underline{F}}{\partial \underline{x}}(\underline{a})$-ra. Itt $\frac{\partial f}{\partial \underline{z}}(\underline{b}) \in \mathbb{R}^{m \times m}$ négyzetes mátrix, és az egyenlet egyértelműen megoldható $\Leftrightarrow$ $\frac{\partial f}{\partial \underline{z}}(\underline{a})$ nem szinguláris.

Teljes differenciál

2. Definíció. *(formális definíció, jelölés)*
Egy változóban: $df = f'dx$
Két változóban $df = f_x dx + f_y dy$.
n változóban: $df = f'_{x_1} dx_1 + \dots + f'_{x_n} dx_n$.
Pontosan ugyanazt jelenti, minthogy $\nabla f = (f'_{x_1}, \dots, f'_{x_n})$.

Megfordítás – antiderivált, primitív függvény/ potenciál függvény keresése.
$n = 1$ – ha $g \in C(I)$, akkor $\exists f$, hogy $df = gdx$, $f' = g$.
$n \geq 2$ – ha $(g_1, \dots, g_n) \in C(D)$, akkor $\exists ? f \in C^1(D)$, $\nabla f = g$?

Probléma: ha $f_j \in C^1(D)$, akkor $f \in C^2(D) \Rightarrow$ a Young-tételnek teljesülnie kell!

4. Tétel. *Ha* $g_1, \dots, g_n \in C^1(D)$, *és* $\frac{\partial g_j}{\partial x_i} = \frac{\partial g_i}{\partial x_j}$ $1(\forall i, j)$ *(Young-feltétel stimmel), akkor kis környezetben (minden egyszeresen összefüggő* $B \subset D$*-ben)* $\exists f \in C^2(B)$, *amelyre* $\nabla f = \underline{g}$, *azaz* $df = g_1 dx_1 + \dots + g_n dx_n$, *(tehát* $\exists f$, *amelynek* B*-n* $(g_1, \dots, g_n)$ *a teljes differenciálja).*

Erről többet A3-ban a vonalintegráloknál tanulunk majd.

Fontos példa: Legyen $(P, Q) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right) \in C^1(\mathbb{R}^2 \setminus \{0\})$. Ez teljesíti a Young-feltételt: $P'_y = \frac{-1}{x^2+y^2} + \frac{y \cdot 2y}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$, $Q'_x = \frac{1}{x^2+y^2} - \frac{x \cdot 2x}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$.

Például ha $B_1 := \{(x, y) : x > 0\}$, akkor $F(x, y) := \operatorname{arctg}(\frac{y}{x})$ jó potenciálfüggvény:

$$F'_x = \frac{1}{1 + (\frac{y}{x})^2} \cdot \frac{-y}{x^2} = P(x, y), \ F'_y = \frac{1}{1 + (\frac{y}{x})^2} \cdot \frac{1}{x} = \frac{x}{x^2+y^2} = Q(x, y).$$

Ha $B_2 := \{x < 0\}$, akkor ez ugyanígy megy. Ha $B_3 := \{y > 0\}$, $B_4 := \{y < 0\}$, akkor is van potenciálfüggvény. $G(x, y) := -\operatorname{arctg}(\frac{x}{y})$-ra $G'_x = P$, $G'_y = Q$. De az egész $D = \mathbb{R}^2 \setminus \{0\}$-ra nem lehet egyetlen jó potenciált találni!