Odvodi funkcij več spremenljivk

Teden 9

Izpeljave pokažemo za **dve spremenljivki**; teorija velja tudi za več spremenljivk. Povsod, kjer pišemo $f(x,y)$, si lahko mislimo tudi $g(x_1,\dots,x_n)$.

Parcialni odvod

Naj bo $D \subset \mathbb{R}^2$, $(a,b) \in D$ in

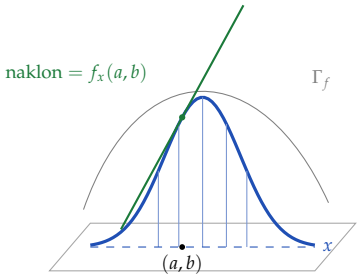
$$f\colon D \to \mathbb{R}, \qquad (x,y) \mapsto f(x,y)$$

funkcija dveh spremenljivk.

Naj bo $D_g \subset \mathbb{R}^n, (x_1,\dots,x_n) \in D_g$ in

$$g\colon D_g \to \mathbb{R}, \qquad (x_1,\dots,x_n) \mapsto g(x_1,\dots,x_n)$$

funkcija n spremenljivk.



Slika 1: Presek grafa Γ_f z ravnino $\{y = b\}$ in tangenta nanj.

Definicija (parcialni odvodi)

$$f_x(a,b) = \frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h},$$
$$f_y(a,b) = \frac{\partial f}{\partial y}(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h},$$
$$g_{x_i}(A) = \frac{\partial g}{\partial x_i}(A) = \lim_{h\to 0} \frac{g(a_1,\dots,a_i+h,\dots,a_n) - g(A)}{h}.$$

Pomen

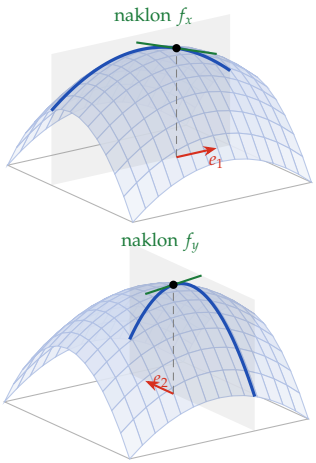
$\frac{\partial f}{\partial x}(a,b)$ je

a) odvod funkcije $t \mapsto f(a+t,b)$ pri $t = 0$;

b) naklonski koeficient preseka grafa Γ_f z ravnino $\{y = b\}$, tj. *naklon grafa v smeri x*;

c) naklon najboljše linearne aproksimacije v smeri x:

$$f(t,b) \doteq f(a,b) + \frac{\partial f}{\partial x}(a,b)\,(t-a).$$



Slika 2: Parcialna odvoda funkcije $f(x,y) = 3 - \frac{1}{4}x^2 - \frac{1}{2}y^2$ v točki $(a,b) = (0.8, 0.6)$: graf presekamo z ravnino $\{y = b\}$ (zgoraj) oziroma $\{x = a\}$ (spodaj) in odčitamo naklon tangente.

Če x spremenimo za δ, se f spremeni za približno $\delta \cdot f_x$.

Primer Izračunaj parcialna odvoda funkcije $f(x,y) = xy^2 + xe^y$ v točki $(1,0)$.

$f_x = y^2 + e^y$, $\quad f_y = 2xy + xe^y$, $\quad$ zato $f_x(1,0) = 1$ in $f_y(1,0) = 1$.

Smerni odvod

Definicija (smerni odvod)

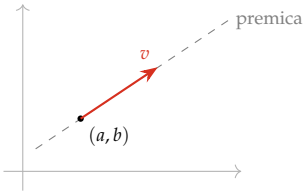
Za vektor $v \in \mathbb{R}^2$ oziroma $w \in \mathbb{R}^n$:

$$f_v(a,b) = \frac{\partial}{\partial t} f\big(\underbrace{(a,b) + t\,v}_{\text{premica}}\big)\Big|_{t=0},$$

$$g_w(A) = \frac{\partial}{\partial t} g(A + t\,w)\Big|_{t=0}.$$

Izraz pod odvodom je *funkcija ene spremenljivke* t.

Tako kot pri hoji po terenu: naklon je odvisen od smeri, v katero hodimo.



Slika 3: Funkcijo f opazujemo le vzdolž premice skozi (a,b) v smeri v.

Pomen

a) Kot pri parcialnih odvodih: odvod funkcije $t \mapsto f((a,b)+tv)$ pri $t = 0$.

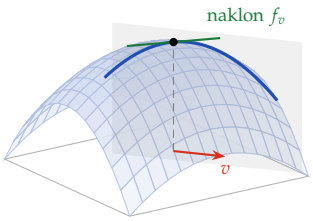
b) Geometrijski pomen (naklon, linearna aproksimacija) velja, če je $\|v\| = 1$ oziroma $\|w\| = 1$.

c) Parcialna odvoda sta posebna smerna odvoda: $f_x = f_{(1,0)}$, $f_y = f_{(0,1)}, \dots$

V splošnem velja

$$f_v(a,b) = f_{\frac{v}{\|v\|}}(a,b) \cdot \|v\|,$$

kar je približno sprememba f pri premiku $(a,b) \mapsto (a,b) + v$.



Slika 4: Smerni odvod: graf presekamo z navpično ravnino skozi (a,b) v smeri $v = (0.6, -0.8)$, $\|v\| = 1$. Tu je $f_v = f_x v_1 + f_y v_2 = 0.24 > 0$, zato f v smeri v narašča.

Dolžina v igra vlogo *hitrosti* premikanja. Primer: $f_{(2,0)} = f_{2e_1} = 2f_x$.

Primer Ali $f(x,y) = x^2 + 2xy - y^2$ v točki $(2,-1)$ v smeri $(1,-1)$ narašča?

$$\begin{aligned} f\big((2,-1)+(t,-t)\big) &= f(2+t,\,-1-t) \\ &= (2+t)^2 + 2(2+t)(-1-t) - (-1-t)^2. \end{aligned}$$

Odvajamo po t:

$$2(2+t) + 2(-1-t) - 2(2+t) + 2(-1-t)$$

in vstavimo $t = 0$:

$$f_{(1,-1)}(2,-1) = 4 - 2 - 4 - 2 = \boxed{-4}.$$

Funkcija pada.

Isti rezultat bomo kasneje dobili veliko hitreje z gradientom (razdelek *Primeri*).

Diferenciabilnost

Diferenciabilnost = odvedljivost v več spremenljivkah. Velja tudi za g.

Definicija (diferenciabilnost)

f je **diferenciabilna** v $(a,b) \in D$, če

$$f(a+h_1, b+h_2) = \underbrace{f(a,b) + f_x(a,b)\,h_1 + f_y(a,b)\,h_2}_{\text{parametrizacija ravnine } \Pi \text{ s } (h_1,h_2)} + o(h_1,h_2),$$

kjer je $\lim\limits_{(h_1,h_2)\to 0} \dfrac{o(h_1,h_2)}{\|(h_1,h_2)\|} = 0$ (*sublinearen ostanek*).

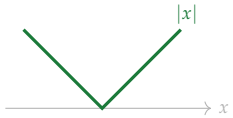
Diferenciabilnost pomeni:

a) Tangentna ravnina na Γ_f v (a,b) obstaja (v vsaki smeri) in je enaka Π. *Določena je le z f_x in f_y, torej so z njima določeni tudi vsi smerni odvodi.*

b) f_x in f_y določata linearno aproksimacijo funkcije.

Izrek

Če sta f_x in f_y **zvezni** v okolici (a,b), je f v (a,b) diferenciabilna.



Slika 5: Protiprimera. (1) $f(r,\varphi) = r = \sqrt{x^2+y^2}$ (stožec): presek z ravnino $y = 0$ je $|x|$, zato $f_x(0,0)$ ne obstaja. (2) $f(r,\varphi) = r\cos 2\varphi$: $f_x(0,0) = 1$, $f_y(0,0) = -1$ obstajata, smerni odvod v smeri φ pa je $\cos 2\varphi \neq \cos\varphi - \sin\varphi$ — ni linearen v smeri, zato f ni diferenciabilna.
+ velja tudi v $\mathbb{R}^n$.

Tangentni vektor na krivuljo

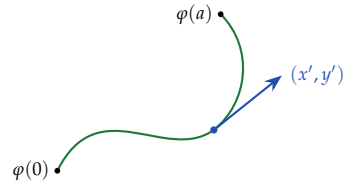
Naj bo $\varphi\colon [0,a] \to \mathbb{R}^2$ odvedljiva parametrizacija krivulje od $\varphi(0)$ do $\varphi(a)$:

$$t \mapsto \big(x(t), y(t)\big), \qquad x, y \text{ odvedljivi v } t.$$

Definicija (vektor hitrosti)

$$\frac{d}{dt}\big(x(t), y(t)\big) = \big(x'(t), y'(t)\big).$$

Velikost vektorja hitrosti je odvisna od parametrizacije, **smer** pa le od krivulje (če je $(x', y') \neq (0,0)$).



Slika 6: Vektor hitrosti je tangenten na krivuljo.
+ velja tudi v $\mathbb{R}^n$.

Verižno pravilo in gradient

Naj bo f diferenciabilna, x, y odvedljivi in $h(t) = f\big(x(t), y(t)\big)$.

Verižno pravilo

$$h'(t) = f_x\big(x(t), y(t)\big) \cdot x'(t) + f_y\big(x(t), y(t)\big) \cdot y'(t)$$

Alternativna zapisa.

a) Z diferenciali (kot $dF = F'\,dx$ oz. $F' = \frac{dF}{dx}$):

$$dh = f_x\,dx + f_y\,dy. \qquad \leftarrow \text{ linearna aproksimacija}$$

b) S skalarnim produktom:

$$h' = \underbrace{(f_x, f_y)}_{\text{grad } f} \cdot (x', y').$$

Definicija (gradient)

$$\operatorname{grad} f = \nabla f = (f_x, f_y) \in \mathbb{R}^2,$$
$$\operatorname{grad} g = \nabla g = (g_{x_1}, g_{x_2}, \ldots, g_{x_n}) \in \mathbb{R}^n.$$

Pozor, nekaj drugega: $\Delta f = f_{xx} + f_{yy}$ je *Laplaceov operator*, ne gradient.

Verižno pravilo v $\mathbb{R}^n$

$$\frac{d}{dt} g(x_1(t), \ldots, x_n(t)) = \operatorname{grad} g \cdot (x_1', \ldots, x_n') = \sum_{i=1}^{n} g_{x_i} x_i'$$

Lastnosti gradienta (veljajo tako za f kot g)

a) V vsaki točki je $\nabla f(a,b) \in \mathbb{R}^2$ vektor.

b) **Formula za smerni odvod.** Za $x(t) = a + tv_1$, $y(t) = b + tv_2$ je $h(t) = f\big((a,b) + t(v_1, v_2)\big)$ in po verižnem pravilu

$$\boxed{h'(0) = \nabla f(a,b) \cdot (v_1, v_2) = f_{(v_1,v_2)}(a,b).}$$

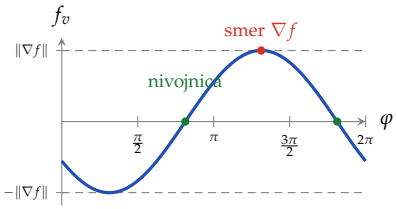
$$\boxed{g_{(v_1,v_2,\ldots,v_n)}(x_1, \ldots, x_n) = \nabla g(x_1, \ldots, x_n) \cdot (v_1, \ldots, v_n).}$$

c) **Smer največje rasti.**

$$f_v = \nabla f \cdot v = \|\nabla f\| \, \|v\| \cos\varphi.$$

Za $\|v\| = 1$ je f_v maksimalen pri $\varphi = 0$, tj. ko je $\nabla f \parallel v$.

$\nabla f(a,b)$ kaže v **smer največje rasti** f v točki (a,b); njegova **velikost** je smerni odvod v tej smeri, Glej Sliko 7.



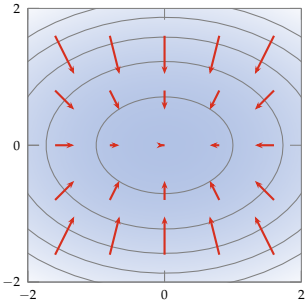
Slika 7: Smerni odvod v $(0.8, 0.6)$ za enotske $v = (\cos\varphi, \sin\varphi)$ v odvistnosti od kota φ: $f_v = \|\nabla f\| \cos(\varphi - \varphi_{\nabla f})$. Največji je v smeri gradienta, enak 0 v smereh tangente na nivojnico.

Primer s terenom: gradient nadmorske višine kaže naravnost v breg, kjer je pot najbolj strma, njegova velikost pa je naklon v dotični smeri.

d) **Gradient je pravokoten na nivojnice.** Naj $(x(t), y(t))$ parametrizira nivojnico funkcije f:

$$f(x(t), y(t)) = \text{konst.} \xrightarrow{d/dt} \nabla f \cdot \underbrace{(x', y')}_{\text{tangenta na nivojnico}} = 0.$$

Torej grad $f \perp$ nivojnice v vsaki točki.

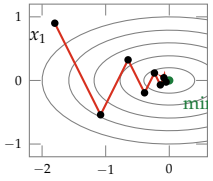


Slika 8: Gradient $\nabla f = (-\frac{1}{2}x, -y)$ funkcije $f(x,y) = -x^2/4 - y^2/2$ (temneje = višje). Puščice so pravokotne na nivojnice in kažejo proti vrhu.

e) **Gradientni spust.**[1] Iščemo $\min f$. Začnemo z začetnim približkom x_1 in naredimo induktivni korak

$$x_{n+1} = x_n - \gamma \cdot \nabla f(x_n),$$

kjer je $\gamma > 0$ dolžina koraka.



Slika 9: Gradientni spust za $\phi(x,y) = \frac{1}{2}x^2 + 2y^2$ s korakom $\gamma = 0.4$: vsak korak gre v smeri $-\nabla\phi$, pravokotno na nivojnico, zato pot "cikcaka" proti minimumu.

[1] *"Motor" nevronskih mrež — več v zapiskih MAT2.*

Primeri

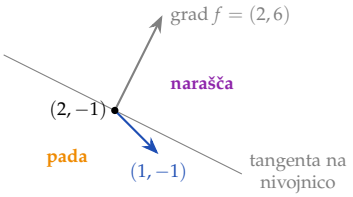
Primer: prejšnji primer z verižnim pravilom Ali f narašča ali pada, če se iz $(2,-1)$ premaknemo v smeri $v=(1,-1)$?

$$f(x,y)=x^2+2xy-y^2, \qquad \operatorname{grad} f=(2x+2y,\ 2x-2y),$$

$$\operatorname{grad} f(2,-1)=(2,6).$$

Po formuli za smerni odvod (glej Sliko 10):

$$f_v=\nabla f\cdot v=(2,6)\cdot(1,-1)=\boxed{-4}.$$



Slika 10: Vektor $(1,-1)$ s gradientom oklepa top kot, zato f v tej smeri pada.

Primer $f(x,y,z,w)=x^4-2xyz-3z\cos w+yz^5$. V katero smer se moramo premakniti iz $(1,1,1,0)$, da se bo f najbolj zmanjšala?

Rešitev: $\operatorname{grad} f=(4x^3-2yz,\ -2xz+z^5,\ -2xy-3\cos w+5yz^4,\ 3z\sin w)$, $\operatorname{grad} f(1,1,1,0)=(2,-1,0,0)$. Najhitreje pada v smeri $-\operatorname{grad} f=(-2,1,0,0)$.

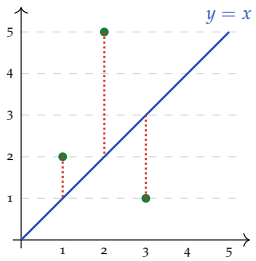
Primer: metoda najmanjših kvadratov Kako naj malo spremenimo premico $y=kx+n$ (začnemo z $y=x$, tj. $k=1$, $n=0$, glej Sliko 11), da bo kvadratna napaka do točk $(1,2)$, $(2,5)$, $(3,1)$ manjša?

$$\begin{aligned}F(k,n)&=(k+n-2)^2+(2k+n-5)^2+(3k+n-1)^2,\\ F_k&=2(k+n-2)+4(2k+n-5)+6(3k+n-1),\\ F_n&=2\big((k+n-2)+(2k+n-5)+(3k+n-1)\big),\end{aligned}$$

$$\begin{aligned}F_k(1,0)&=2\cdot(-1)+4\cdot(-3)+6\cdot 2=-2,\\ F_n(1,0)&=2(-1-3+2)=-4.\end{aligned}$$

$\operatorname{grad} F(1,0)=(-2,-4)$ je smer največjega naraščanja napake. Če bi radi zmanjšali napako, za nek majhen $\varepsilon>0$ naredimo

$$k\mapsto k+2\varepsilon, \qquad n\mapsto n+4\varepsilon.$$



Slika 11: Točke in začetna premica; rdeče so odstopanja, katerih kvadrate seštejemo.

Primeri odvodov iz prakse

Oglejmo si GPS sled kolesarjenja:

funkcija	↔ nadmorska višina
trajektorija poti	↔ izris poti na zemljevidu
parametrizacija poti	↔ potovanje po sledi v odvisnosti od časa
gradientni vektor	↔ kaže v smer najbolj strmega klanca, dolžina je naklon v tej smeri
naklon vzdolž poti	↔ naklon grafa nadmorske višine v odvisnosti od poti
smerni odvod po tangentnem vektorju	↔ naklon grafa nadmorske višine v odvisnosti od časa
vezani ekstremi	↔ lokalni vrhovi in doline po poti