Uporaba odvodov funkcij več spremenljivk

Teden 10

Odvode funkcij več spremenljivk uporabimo za tri naloge: **lokalne ekstreme** (stacionarne točke in Hessejeva matrika), **vezane ekstreme** (Lagrangeova funkcija) in **Taylorjeve polinome** T_1 in T_2.

Stacionarne točke in lokalni ekstremi

Naj bo $f = f(x_1, \ldots, x_n)$ diferenciabilna funkcija.

Definicija (stacionarna točka)

Točka $a \in D_f$ je **stacionarna točka**, če je

$$\nabla f(a) = \vec{0} \in \mathbb{R}^n.$$

Ekvivalentno:

a) vsi parcialni odvodi v a so 0;

b) vsi smerni odvodi v a so 0;

c) tangentna ravnina v a je "vodoravna": $z = f(a)$.

Primer: 1 Poišči stacionarne točke funkcije $f(x,y) = x^2 + y^2 - 2x - 6y + 14$.

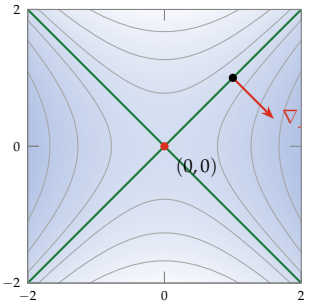
$$f_x = 2x - 2 = 0 \Rightarrow x = 1, \qquad f_y = 2y - 6 = 0 \Rightarrow y = 3.$$

Edina stacionarna točka je $\boxed{(1,3)}$.

Dopolnimo do popolnih kvadratov: $f = (x-1)^2 + (y-3)^2 + 4$. Torej je $(1,3)$ celo *globalni* minimum z vrednostjo 4.

Primer: 2 $f(x,y) = x^2 - y^2$: $f_x = 2x$, $f_y = -2y$, zato je edina stacionarna točka $(0,0)$. Nivojnica $f(x,y) = 0$ sta premici $y = \pm x$; npr. $\nabla f(1,1) = (2,-2)$.

$(0,0)$ zagotovo **ni** lokalni ekstrem!



Slika 1: Nivojnice funkcije $f = x^2 - y^2$ (temneje = višje). Nivojnica $f = 0$ sta zeleni premici $y = \pm x$. V $(0,0)$ je $\nabla f = \vec{0}$, a v vsaki okolici ima f tako pozitivne kot negativne vrednosti.

Definicija (lokalni ekstremi)

Točka $a \in D_f$ je

- **lokalni minimum**, če $\exists \varepsilon > 0\ \forall x : \|x - a\| < \varepsilon \Rightarrow f(x) \geq f(a)$;
- **lokalni maksimum**, če $\exists \varepsilon > 0\ \forall x : \|x - a\| < \varepsilon \Rightarrow f(x) \leq f(a)$.

Skupaj jih imenujemo **lokalni ekstremi**.

Trditev

Vsak lokalni ekstrem (diferenciabilne funkcije v notranji točki) je stacionarna točka.

Obrat ne velja: Primer 2 je stacionarna točka, ki ni ekstrem — sedlo.

Računalništvo: pri učenju nevronskih mrež je $\nabla L = \vec{0}$ tudi v sedlih. V visokih dimenzijah so sedla med stacionarnimi točkami pogosta: za minimum mora funkcija naraščati v *vseh* smereh hkrati, za sedlo zadošča ena smer navzdol.

Vprašanje: kako z drugim odvodom ugotoviti naravo stacionarne točke?

Drugi odvodi: $f_{xx},\ f_{xy},\ f_{yx},\ f_{yy}$ *(v splošnem jih je n^2).*

Trditev

Če sta f_{xy} in f_{yx} zvezni, potem $f_{xy} = f_{yx}$.

Velja tudi za višje odvode in za n spremenljivk: če so drugi odvodi zvezni, potem vrstni red odvajanja ni pomemben.

Druge odvode sestavimo v **Hessejevo matriko**:

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \quad \begin{array}{l} \leftarrow \text{odvisna od točke,} \\ \leftarrow \text{velikost } 2 \times 2. \end{array}$$

Po trditvi zgoraj je H_f **simetrična**. Pri n spremenljivkah je H_f matrika velikosti $n \times n$.

Test z drugim odvodom

Kakšen je pogoj za lokalni minimum? Če so v (x_0, y_0) *vsi smerni drugi odvodi pozitivni*, bo (x_0, y_0) lokalni minimum. Za smer $v = (v_1, v_2)$ naj bo

$$h(t) = f\big((x_0, y_0) + t(v_1, v_2)\big) = f(x_0 + tv_1,\ y_0 + tv_2).$$

Po verižnem pravilu:

$$\begin{aligned} h'(t) &= f_x(x_0 + tv_1, y_0 + tv_2) \cdot v_1 + f_y(\dots) \cdot v_2, \\ h''(t) &= f_{xx}\, v_1^2 + f_{xy}\, v_1 v_2 + f_{yx}\, v_1 v_2 + f_{yy}\, v_2^2 \ = \ v^{\mathsf{T}} H_f\, v. \end{aligned}$$

$$h''(0) > 0 \ \ \forall v \neq \vec{0} \quad \overset{\text{def. iz LA}}{\iff} \quad H_f(x_0, y_0) \text{ je } \textbf{pozitivno definitna}$$

(velja za poljuben n)

Za $n = 2$ se pogoj poenostavi:

Trditev

Naj bo (x_0, y_0) stacionarna točka funkcije $f(x, y)$ in

$$D = f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - \big(f_{xy}(x_0, y_0)\big)^2 = \det H_f(x_0, y_0).$$

a) Če je $D > 0$, je (x_0, y_0) lokalni ekstrem:

   i) $f_{xx}(x_0, y_0) > 0 \Rightarrow$ lokalni **minimum**;

   ii) $f_{xx}(x_0, y_0) < 0 \Rightarrow$ lokalni **maksimum**.

b) Če je $D < 0$, je (x_0, y_0) **sedlo**.

c) Če je $D = 0$, ne moremo nič sklepati — lahko je karkoli.

Pri $D > 0$ je $f_{xx} f_{yy} > f_{xy}^2 \geq 0$, zato $f_{xx} \neq 0$ in f_{yy} ima isti predznak kot f_{xx}.

Pri c): npr. $x^4 + y^4$ (min), $-x^4 - y^4$ (max), $x^3 + y^2$ (ni ekstrem) imajo v $(0,0)$ vse $D = 0$.

Primer: 1, nadaljevanje $f(x,y) = x^2+y^2-2x-6y+14$: $f_x = 2x-2$, $f_y = 2y-6$, stacionarna točka $(1,3)$.

$$H_f(1,3) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad D = 4 > 0, \quad f_{xx} = 2 > 0,$$

zato je $(1,3)$ lokalni **MIN**.

Primer: 2, nadaljevanje $f(x,y) = x^2 - y^2$:

$$H_f = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, \quad D = -4 < 0 \;\Rightarrow\; (0,0) \text{ je } \textbf{sedlo}.$$

Primer: 3: najmanjši kvadrati iz tedna 9 Pri premici $y = kx+n$ skozi točke $(1,2)$, $(2,5)$, $(3,1)$ smo minimizirali $F(k,n) = (k+n-2)^2 + (2k+n-5)^2 + (3k+n-1)^2$. Iz

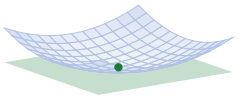
$$F_k = 28k + 12n - 30 = 0, \qquad F_n = 12k + 6n - 16 = 0$$

dobimo stacionarno točko $k = -\frac{1}{2}$, $n = \frac{11}{3}$. Hessejeva matrika

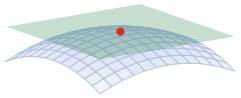
$$H_F = \begin{bmatrix} 28 & 12 \\ 12 & 6 \end{bmatrix}, \qquad D = 168 - 144 = 24 > 0, \quad F_{kk} = 28 > 0$$

pokaže, da gre za **minimum** — najboljšo premico $y = -\frac{1}{2}x + \frac{11}{3}$.

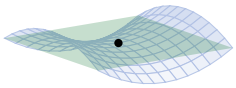
lokalni MIN: $D > 0$, $f_{xx} > 0$



lokalni MAX: $D > 0$, $f_{xx} < 0$



sedlo: $D < 0$



Slika 2: Tri možnosti v stacionarni točki. Zelena je vodoravna tangentna ravnina: graf je nad njo (minimum), pod njo (maksimum) ali jo seka (sedlo).

Vezani ekstremi

Definicija (vezani ekstremi)

Naj bosta $f, g\colon D \to \mathbb{R}$. **Vezani ekstremi** funkcije f pri pogoju $g = 0$ so ekstremi funkcije f, zožene na množico $\{g = 0\}$.

V naših primerih bo $D \subset \mathbb{R}^2$, množica $\{g = 0\}$ pa krivulja.

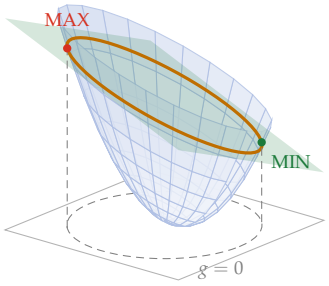
Potreben pogoj za lokalni ekstrem f na $\{g = 0\}$:

$$\boxed{\nabla f \parallel \nabla g}$$

Računski razlog. Želimo, da je smerni odvod f v smeri krivulje $\{g = 0\}$ enak 0:

$$\begin{aligned} \nabla f \cdot (\text{tangenta na } g = 0) = 0 &\implies \nabla f \perp \text{tangenta na } g = 0 \\ &\implies \nabla f \parallel \nabla g, \end{aligned}$$

saj je tudi ∇g pravokoten na svojo nivojnico $\{g = 0\}$ (teden 9). **Geometrijsko:** krivulja $\{g = 0\}$ se v ekstremu *dotakne* nivojnice funkcije f.



Slika 3: Presek paraboloida $z = x^2 + y^2$ (modro) in ravnine $x + y + z = 12$ (zeleno) je oranžna krivulja. Njena projekcija na tla je krožnica $g = 0$. Iščemo najvišjo (rdeče) in najnižjo (zeleno) točko krivulje.

Alternativna formulacija.

Definicija (Lagrangeova funkcija)

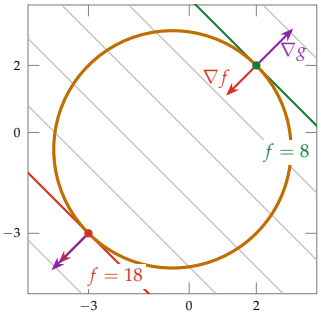
$$L\colon D\times\mathbb{R}\to\mathbb{R},\qquad L(x,y,\lambda)=f(x,y)-\lambda\, g(x,y).$$

Definirana načeloma na odprtem območju. Številu λ rečemo *Lagrangev multiplikator*.

Stacionarne točke (x_0,y_0,λ_0) funkcije L so **kandidati** za vezane ekstreme (x_0,y_0) funkcije f pri pogoju $g=0$:

$$\left.\begin{aligned}L_x&=f_x-\lambda g_x=0\\ L_y&=f_y-\lambda g_y=0\end{aligned}\right\}\ \text{pogoj } \nabla f=\lambda\nabla g,$$

$$L_\lambda=-g=0\qquad\text{pogoj } g=0.$$



Slika 4: Pogled od zgoraj: nivojnice višine $f=12-x-y$ so premice $x+y=$ konst., oranžna je krožnica $g=0$. V ekstremih se krožnica nivojnice *dotakne*, zato sta ∇f (rdeče) in ∇g (vijolično) vzporedna.

Primer: najvišja in najnižja točka preseka Poiščimo najvišjo in najnižjo točko preseka ravnine $\Pi\colon x+y+z=12$ in paraboloida $P\colon x^2+y^2=z$, glej sliko 3.

$P\cap\Pi$: iz $z=12-x-y$ in $z=x^2+y^2$ sledi

$$x^2+y^2+x+y=12\qquad\leftarrow \textit{krožnica}$$

in pogoj je

$$g(x,y)=x^2+y^2+x+y-12=0.$$

Iščemo ekstreme *višine* $f(x,y)=12-x-y$ vzdolž $g=0$:

$$\nabla f=(-1,-1),\qquad \nabla g=(2x+1,\,2y+1).$$

Pogoj $t\,\nabla f=\nabla g$:

$$\left.\begin{aligned}-t&=2x+1\\ -t&=2y+1\end{aligned}\right\}\Rightarrow 2x-2y=0\Rightarrow \boxed{x=y}.$$

Vstavimo v pogoj: $g(x,x)=2x^2+2x-12=0$, $x^2+x-6=0$, $(x+3)(x-2)=0$.

stacionarna točka	f	
$(2,2)$	8	lokalni **MIN**
$(-3,-3)$	18	lokalni **MAX**

Iz problema vidimo, da sta to tudi **globalna** ekstrema, a to v splošnem ni nujno res (npr. če bi bila $\{g=0\}$ neomejena ali z robom ...).

Taylorjevi polinomi

Taylorjevi polinomi obstajajo v splošnem; omenili bomo le T_1 in T_2 za $n=2$. Polinoma okrog (x_0,y_0):

Polinom stopnje 1 — tangentna ravnina

$$T_1(x,y)=f(x_0,y_0)+f_x(x_0,y_0)\,(x-x_0)+f_y(x_0,y_0)\,(y-y_0)$$

Dobra aproksimacija za diferenciabilne funkcije — to je ravno definicija diferenciabilnosti iz tedna 9.

Polinom stopnje 2

$$T_2(x,y) = T_1(x,y) + A(x-x_0)^2 + B(x-x_0)(y-y_0) + C(y-y_0)^2$$

Želimo, da se drugi odvodi f in T_2 ujemajo v (x_0, y_0):

$$(T_2)_{xx} = 2A = f_{xx} \quad \Rightarrow \quad \boxed{A = \tfrac{1}{2} f_{xx}(x_0, y_0)},$$
$$(T_2)_{yy} = 2C = f_{yy} \quad \Rightarrow \quad \boxed{C = \tfrac{1}{2} f_{yy}(x_0, y_0)},$$
$$(T_2)_{xy} = B = f_{xy} \quad \Rightarrow \quad \boxed{B = f_{xy}(x_0, y_0)}.$$

Za T_n bi uporabili še odvode do reda n.

Kviz

Koliko je odvodov reda 3, 5, n (funkcije dveh spremenljivk)?

Rešitev: Če štejemo vse vrstne rede odvajanja, jih je $2^3 = 8$, $2^5 = 32$, 2^n. Če sta mešana odvoda zvezna, je *različnih* le 4, 6, $n+1$: določa jih le, kolikokrat odvajamo po x.