Funkcije večih spremenljivk

Pri opisu fizikalnih pojavov pogosto uporabljamo funkcije večih realnih spremenljivk. Osnovni primeri količin, ki jih lahko tako opišemo, so temperatura, gostota, tlak in potencialna energija. Videli bomo, da lahko mnoge pojme, ki smo jih spoznali pri funkcijah ene realne spremenljivke brez problemov posplošimo na funkcije večih spremenljivk. Bolj podrobno si bomo pogledali zveznost, odvode in Taylorjeve vrste, integrale pa bomo pustili za kasneje.

Čeprav so najbolj uporabne funkcije dveh, treh in štirih spremenljivk, bomo zaradi enostavnosti študirali kar n-dimenzionalni evklidski prostor

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{n} = \{\, (v_1, v_2, \ldots, v_n) \,|\, v_1, v_2, \ldots, v_n \in \mathbb{R}\}.$$

Elementom prostora $\mathbb{R}^n$ rečemo vektorji, predstavimo pa jih z urejenimi n-tericami realnih števil. Ponavadi bomo uporabljali oznako

$$v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n.$$

Dodobra smo že spoznali prostor $\mathbb{R}^3$. Nekatere operacije med vektorji v $\mathbb{R}^3$, kot so seštevanje in skalarni produkt, lahko brez težav posplošimo na $\mathbb{R}^n$. Mešani produkt v višjih dimenzijah nadomestimo z determinanto, medtem ko ustreznega analoga vektorskega produkta v višjih dimenzijah ne poznamo. Razlog je v tem, da ima prostor $\mathbb{R}^3$ strukturo Liejeve algebre, ki pa je ne moremo naravno posplošiti.

Na prostoru $\mathbb{R}^n$ lahko kot v $\mathbb{R}^3$ definiramo seštevanje vektorjev po komponentah in pa množenje s skalarji. Če sta $v, w \in \mathbb{R}^n$ in $\alpha \in \mathbb{R}$, definiramo:

$$\begin{aligned} v + w &= (v_1 + w_1, v_2 + w_2, v_3 + w_3, \ldots, v_n + w_n), \\ \alpha v &= (\alpha v_1, \alpha v_2, \alpha v_3, \ldots, \alpha v_n). \end{aligned}$$

Tako postane $\mathbb{R}^n$ vektorski prostor. Skalarni produkt vektorjev $v, w \in \mathbb{R}^n$ je definiran s predpisom

$$v \cdot w = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n \in \mathbb{R},$$

dolžina vektorja $v \in \mathbb{R}^n$ pa je

$$|v| = \sqrt{v \cdot v}.$$

Podobno kot v $\mathbb{R}^3$ tudi v $\mathbb{R}^n$ veljata trikotniška in Cauchy-Schwartzeva neenakost:

$$\begin{aligned} |v + w| &\leq |v| + |w|, \\ |v \cdot w| &\leq |v||w|. \end{aligned}$$

Za vektorja $v, w \in \mathbb{R}^n$ rečemo, da sta pravokotna, če je $v \cdot w = 0$. Razdaljo med njima pa definiramo s predpisom

$$d(v, w) = |v - w|.$$

Razdalja med točkami nam omogoča, da tudi v $\mathbb{R}^n$ definiramo krogle, odprte množice, zaprte množice, okolice, limite...

Za poljubno točko $a \in \mathbb{R}^n$ in poljuben $\epsilon > 0$ je odprta krogla s središčem v a in s polmerom ϵ množica

$$K(a, \epsilon) = \{\, v \in \mathbb{R}^n \mid |a - v| < \epsilon\} \subset \mathbb{R}^n.$$

Naj bo sedaj $A \subset \mathbb{R}^n$ poljubna množica. Potem je:

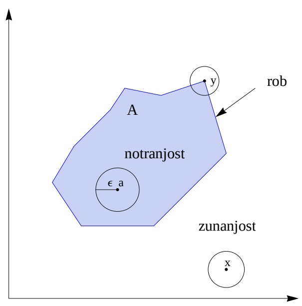
· Množica A je *okolica* točke $a \in \mathbb{R}^n$, če obstaja $\epsilon > 0$, da je $K(a, \epsilon) \subset A$.

· Množica A je *odprta*, če je okolica vsake svoje točke.

· Množica A je *zaprta*, če je množica $\mathbb{R}^n \setminus A$ odprta.

· Množica A je *omejena*, če obstaja tak $M \in \mathbb{R}^+$, da je $A \subset K(0, M)$.

Točka a je notranja točka množice A, če je A okolica točke a. Če je A^c okolica točke a, rečemo, da je a zunanja točka množice A. Če a ni niti notranja niti zunanja točka, je robna točka množice A.



Zaporedje v $\mathbb{R}^n$ je funkcija $a : \mathbb{N} \to \mathbb{R}^n$. Da ne bo zmede z indeksi, bomo zaporedja večinoma označevali z (a_m).

(i) Točka $a \in \mathbb{R}^n$ je *stekališče* zaporedja (a_m), če je v vsaki okolici točke a neskončno členov zaporedja (a_m).

(ii) Točka $a \in \mathbb{R}^n$ je *limita* zaporedja (a_m), če v vsaki okolici točke a ležijo vsi členi zaporedja (a_m), z izjemo morda končno mnogih. V tem primeru pišemo

$$a = \lim_{m \to \infty} (a_m).$$

Vse lastnosti limit in stekališč zaporedij števil veljajo tudi za zaporedja vektorjev v $\mathbb{R}^n$. Če ima zaporedje limito, ima natanko eno stekališče in je konvergentno. Tudi v $\mathbb{R}^n$ je zaporedje konvergentno natanko takrat, ko je Cauchyjevo. To pomeni, da je $\mathbb{R}^n$ poln prostor.

Če je (a_m) zaporedje vektorjev v $\mathbb{R}^n$, ga lahko zapišemo v obliki

$$a_m = ((a_m)_1, (a_m)_2, \ldots, (a_m)_n),$$

kjer so $((a_m)_i)$ zaporedja realnih števil za $i = 1, 2, \ldots, n$. Zaporedje vektorjev je potem konvergentno natanko takrat, ko so konvergentna vsa zaporedja njegovih komponent in velja

$$\lim_{m\to\infty}(a_m) = a \iff \lim_{m\to\infty}(a_m)_i = a_i, \quad i = 1, 2, \ldots, n.$$

Funkcija n spremenljivk je preslikava $f : D \subset \mathbb{R}^n \to \mathbb{R}$.

Zgled. V primeru, ko je $n = 2$, neodvisni spremenljivki ponavadi označimo z $(x, y) = (x_1, x_2)$. Poglejmo si funkcijo

$$f(x, y) = \frac{1}{x^2 + y^2}.$$

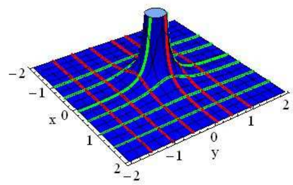
Njena domena je množica

$$D = \{\, (x, y) \in \mathbb{R}^2 \mid (x, y) \neq 0\} = \mathbb{R}^2 \setminus \{(0, 0)\},$$

njen graf pa množica

$$\operatorname{graf}(f) = \left\{ \left(x, y, \frac{1}{x^2 + y^2}\right) \mid (x, y) \in D \right\} \subset \mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3.$$

Geometrično si grafe funkcij dveh spremenljivk predstavljamo kot ploskve v prostoru.



Če graf prerežemo z ravnino $x = x_0$ ali $y = y_0$, dobimo krivuljo, ki ustreza grafom $y \mapsto f(x_0, y)$ oziroma $x \mapsto f(x, y_0)$. Če te krivulje zložimo skupaj, dobimo ploskev, ki predstavlja graf funkcije dveh spremenljivk. Risanje grafov funkcij dveh spremenljivk na roko je zahtevno, zato si raje pomagamo z računalniškimi orodji.

Definicija 1. *Funkcija $f : D \subset \mathbb{R}^n \to \mathbb{R}$ je zvezna v točki $a \in D$, če za vsak $\epsilon > 0$ obstaja tak $\delta > 0$, da za poljuben $y \in D \cap K(a, \delta)$ velja $f(y) \in K(f(a), \epsilon)$. Funkcija f je zvezna, če je zvezna v vsaki točki iz D.*

Če upoštevamo definicijo odprte krogle v $\mathbb{R}^n$, to pomeni, da je f zvezna v točki a, če za vsak $\epsilon > 0$ obstaja tak $\delta > 0$, da za vsak $y \in D$ velja

$$|y - a| < \delta \Longrightarrow |f(y) - f(a)| < \epsilon.$$

V tej obliki je definicija zveznosti funkcije večih spremenljivk enaka kot definicija zveznosti funkcije ene spremenljivke, le da moramo absolutne vrednosti realnih števil nadomestiti z dolžinami vektorjev.

Ista definicija zveznosti je smiselna tudi za funkcijo $g : D \subset \mathbb{R}^n \to \mathbb{R}^m$

$$g(x_1, x_2, \ldots, x_n) = (g_1(x_1, \ldots, x_n), g_2(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)) .$$

V tem primeru so $g_1, g_2, \ldots, g_m$ realne funkcije n spremenljivk, funkcija g pa je vektorska funkcija n spremenljivk. Pišemo tudi $g = (g_1, g_2, \ldots, g_m)$. Ni težko preveriti, da je funkcija g zvezna v točki a natanko takrat, ko so vse njene komponente $g_1, g_2, \ldots, g_m$ zvezne v točki a.

Če sta $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ in $h : E \subset \mathbb{R}^n \to \mathbb{R}^m$ zvezni funkciji, sta tudi njuna vsota in produkt

$$\begin{aligned} f + h &: D \cap E \to \mathbb{R}^m & \qquad (f + h)(a) &= f(a) + h(a), \\ f \cdot h &: D \cap E \to \mathbb{R} & \qquad (f \cdot h)(a) &= f(a) \cdot h(a), \end{aligned}$$

zvezni funkciji. Produkt funkcij je definiran samo v primeru $m = 1$. Če je še $g : D' \subset \mathbb{R}^m \to \mathbb{R}^k$ zvezna funkcija, je tudi kompozitum

$$g \circ f : f^{-1}(D') \to \mathbb{R}^k \qquad (g \circ f)(a) = g(f(a))$$

zvezna funkcija. Dokazi vseh teh lastnosti so isti kot pri realnih funkcijah ene spremenljivke.

Zgled. (1) Projekcije so funkcije $\mathrm{pr}_k : \mathbb{R}^n \to \mathbb{R}$, za $k = 1, 2, \ldots, n$, definirane s predpisi

$$\mathrm{pr}_k(x_1, x_2, \ldots, x_n) = x_k.$$

Iz neenakosti

$$|a - x| = \sqrt{(a_1 - x_1)^2 + \ldots + (a_n - x_n)^2} \geq \sqrt{(a_k - x_k)^2} = |a_k - x_k|$$

sledi implikacija

$$|a - x| < \delta \Longrightarrow |a_k - x_k| < \delta.$$

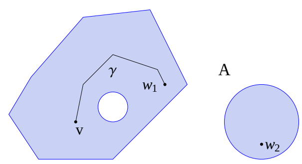
Če vzamemo $\delta = \epsilon$, torej vidimo, da so projekcije zvezne funkcije.

(2) Ker so zvezne projekcije in ker sta vsota in produkt zveznih funkcij spet zvezni funkciji, sta npr. zvezni tudi funkciji

$$\begin{aligned} f(x, y, z) &= xyz^3, \\ h(x, y, z) &= xy + 3yz^3 + 12yz^4 - xz^5. \end{aligned}$$

Bolj splošno so polinomi n spremenljivk zvezne funkcije. Projekcije lahko smatramo kot polinome stopnje ena oziroma linearne funkcije.

Naj bo sedaj $A \subset \mathbb{R}^n$. *Pot* v A od točke v do točke w je zvezna funkcija $\gamma : [0,1] \to A \subset \mathbb{R}^n$, za katero velja $\gamma(0) = v$ in $\gamma(1) = w$.



Pot si lahko predstavljamo kot zvezen sprehod med začetno in končno točko. V primeru množice A kot na zgornji sliki obstaja pot med točkama v in w_1, ni pa poti med točkama v in w_2.

Definicija 2. *Množica $A \subset \mathbb{R}^n$ je povezana s potmi, če za poljubni točki $v, w \in A$ obstaja pot v A od v do w.*

Trditev 3. *Naj bo $f : D \subset \mathbb{R}^n \to \mathbb{R}$ zvezna funkcija in naj bo D povezana s potmi. Potem velja:*

(i) Množica $f(D)$ je povezana s potmi.

(ii) Če sta $v, w \in D$ taki točki, da je $f(v) \cdot f(w) < 0$, potem obstaja točka $\xi \in D$, da je $f(\xi) = 0$.

Dokaz. (i) Vzemimo poljubni števili $x, y \in f(D)$. Potem obstajata taki točki $v, w \in D$, da velja $x = f(v)$ in $y = f(w)$. Ker je množica D povezana s potmi, obstaja pot γ v D od v do w. Ta pot $\gamma : [0,1] \to D$ je zvezna funkcija, za katero je

$$\gamma(0) = v,$$
$$\gamma(1) = w.$$

Kompozitum $f \circ \gamma$ je potem zvezna funkcija $f \circ \gamma : [0,1] \to f(D) \subset \mathbb{R}$. Torej lahko funkcijo $f \circ \gamma$ smatramo kot pot, za katero je

$$f(\gamma(0)) = f(v) = x,$$
$$f(\gamma(1)) = f(w) = y.$$

Torej je $f \circ \gamma$ pot od x do y, kar pa pomeni, da je množica $f(D)$ povezana s potmi.

(ii) Naj bosta sedaj $v, w \in D$ taki točki, da je $f(v) \cdot f(w) < 0$. Ker je D povezana s potmi, obstaja pot γ v D od v do w. Potem je $f \circ \gamma : [0,1] \to \mathbb{R}$ zvezna funkcija, za katero je

$$(f \circ \gamma)(0) \cdot (f \circ \gamma)(1) = f(v) \cdot f(w) < 0.$$

Zvezna funkcija $f \circ \gamma$ torej zavzame različno predznačeni vrednosti v krajiščih intervala $[0,1]$. Zato obstaja nek $c \in [0,1]$, da je $f(\gamma(c)) = 0$. Če definiramo $\xi = \gamma(c)$, potem velja $f(\xi) = 0$. $\square$

Primeri s potmi povezanih množic v $\mathbb{R}$ so intervali, medtem ko so s potmi povezane množice v $\mathbb{R}^2$ in $\mathbb{R}^3$ na primer liki in telesa.

Definicija 4. *Podmnožica $A \subset \mathbb{R}^n$ je kompaktna, če ima vsako zaporedje točk iz A vsaj eno stekališče, ki leži v A.*

Trditev 5. *Podmnožica $A \subset \mathbb{R}^n$ je kompaktna natanko takrat, ko je zaprta in omejena.*

Dokaz. ($\Longrightarrow$) Naj bo $A \subset \mathbb{R}^n$ kompaktna podmnožica. Najprej bomo pokazali, da je potem A omejena. Če namreč A ne bi bila omejena, bi lahko za vsak $m \in \mathbb{N}$ našli tak $a_m \in A$, da je $|a_m| > m$. Tako dobljeno zaporedje (a_m) potem ne bi imelo stekališča, kar pa je v protislovju s predpostavko.

Pokažimo še, da je A zaprta množica. Privzeli bomo, da A ni zaprta in pokazali, da nas to vodi v protislovje. Če A ne bi bila zaprta, bi lahko našli točko $t \in \mathbb{R}^n$, ki je v robu A, a ni v A. To bi pomenilo, da za vsak $m \in \mathbb{N}$ velja $K(t, 1/m) \cap A \neq \emptyset$, zato bi lahko za vsak $m \in \mathbb{N}$ našli tak $t_m \in \mathbb{R}^n$, da bi veljalo

$$t_m \in K(t, 1/m) \cap A.$$

Tako bi našli zaporedje (t_m) v A, ki bi konvergiralo k t. Ker t ni v A, bi to pomenilo, da (t_m) nima stekališča v A, kar pa je v nasprotju s predpostavko.

($\Longleftarrow$) Denimo sedaj, da je $A \subset \mathbb{R}^n$ zaprta in omejena in naj bo (a_m) zaporedje v A. Pokažemo lahko, da ima vsako omejeno zaporedje v $\mathbb{R}^n$ vsaj eno stekališče. V primeru $n = 1$ smo to že pokazali, v višjih dimenzijah pa lahko uporabimo podoben sklep za vsako komponento posebej in korak za korakom najdemo podzaporedje, ki konvergira na vseh komponentah. Ker je A omejena, ima torej zaporedje (a_m) vsaj eno stekališče $a \in \mathbb{R}^n$. Ker pa je A tudi zaprta, mora biti $a \in A$. $\square$

Interval v $\mathbb{R}$ je kompakten natanko takrat, ko je zaprt in omejen oziroma, ko je oblike $[a, b]$. Polodprti in odprti intervali pa niso kompaktni, ker ne vsebujejo robnih točk. Prav tako niso kompaktni neomejeni intervali, saj vsebujejo monotona, neomejena zaporedja.

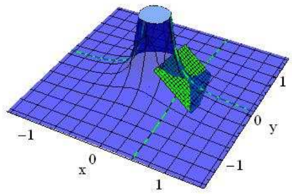
Trditev 6. *Naj bosta $D \subset \mathbb{R}^n$ kompaktna množica in $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ zvezna funkcija. Tedaj je tudi množica $f(D)$ kompaktna.*

Dokaz. Razložili bomo samo skico dokaza. Vzemimo zaporedje (b_m) v $f(D)$. Radi bi pokazali, da ima zaporedje (b_m) vsaj eno stekališče v $f(D)$.

Najprej lahko najdemo zaporedje (a_m) v D, tako da za vsak $m \in \mathbb{N}$ velja $b_m = f(a_m)$. Ker je množica D kompaktna, ima zaporedje (a_m) stekališče v D, zato lahko najdemo neko podzaporedje (a_{m_k}) zaporedja (a_m), ki konvergira k $a \in D$. Zaradi zveznosti funkcije f zaporedje $(b_{m_k}) = (f(a_{m_k}))$ konvergira k $f(a) \in f(D)$, kar pa pomeni, da je $f(a)$ stekališče zaporedja (b_m), kar smo želeli dokazati. $\square$

Odvodi funkcij večih spremenljivk

Poglejmo si še enkrat graf funkcije dveh spremenljivk, podane s predpisom $f(x, y) = \frac{1}{x^2+y^2}$. Če graf funkcije f presekamo z ravninama $x = x_0$ in $y = y_0$ dobimo dve krivulji, ki se sekata v točki $(x_0, y_0, f(x_0, y_0))$. Naklona teh krivulj sta načeloma različna, njuna tangentna vektorja pa napenjata tangentno ravnino na ploskev v dani točki.



Naklone ploskve v različnih smereh na dani ploskvi bomo računali s pomočjo parcialnih odvodov.

Naj bo $f : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}$ funkcija n spremenljivk in naj bo $v \in U$. Zanimale nas bodo funkcije oblike

$$x_j \mapsto f(v_1, v_2, \ldots, v_{j-1}, x_j, v_{j+1}, \ldots, v_n)$$

za nek $j \in \{1, 2, \ldots, n\}$. Vse koordinate razen j-te so fiksirane, kar pomeni, da graf takšne funkcije ustreza krivulji, ki poteka po grafu funkcije f v smeri j-te koordinate skozi točko $(v, f(v))$. Naklon te krivulje v dani točki je enak parcialnemu odvodu funkcije f na spremenljivko x_j v točki v, ki je definiran s predpisom (če seveda limita obstaja)

$$\frac{\partial f}{\partial x_j}(v) = \lim_{h \to 0} \frac{f(v_1, v_2, \ldots, v_j + h, \ldots, v_n) - f(v_1, v_2, \ldots, v_j, \ldots, v_n)}{h}.$$

Če označimo $e_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$, kjer je enica na j-tem mestu, lahko parcialni odvod definiramo tudi v vektorski obliki

$$\frac{\partial f}{\partial x_j}(v) = \lim_{h \to 0} \frac{f(v + h \cdot e_j) - f(v)}{h}.$$

Včasih uporabljamo tudi oznake

$$\frac{\partial f}{\partial x_j}(v) = f_{x_j}(v) = (D_j f)(v) = (D_{e_j} f)(v).$$

Parcialni odvod na j-to spremenljivko je spet funkcija večih spremenljivk

$$\frac{\partial f}{\partial x_j} : U \to \mathbb{R}.$$

Parcialni odvod interpretiramo kot odvod funkcije, ki jo dobimo z zožitvijo funkcije na premico v smeri ene izmed koordinatnih osi. Če bi namesto vektorja e_j izbrali nek drug smerni vektor $u \in \mathbb{R}$, bi dobili smerni odvod funkcije f v smeri u

$$(D_u f)(v) = \lim_{h \to 0} \frac{f(v + h \cdot u) - f(v)}{h}$$

Če je $g : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}^m$ vektorska funkcija, dobimo njen smerni odvod tako, da izračunamo smerne odvode vsake komponente posebej

$$D_u g = (D_u g_1, D_u g_2, \ldots, D_u g_m).$$

Gradient funkcije $f : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}$ je vektorska funkcija $\nabla f : U \to \mathbb{R}^n$, definirana s predpisom

$$\nabla f(v) = \left(\frac{\partial f}{\partial x_1}(v), \frac{\partial f}{\partial x_2}(v), \ldots, \frac{\partial f}{\partial x_n}(v) \right).$$

Zgled. Vzemimo polinom dveh spremenljivk, definiran s predpisom

$$f(x, y) = x^2 + 2xy^3 + 4x.$$

Njegova parcialna odvoda sta potem

$$\frac{\partial f}{\partial x} = 2x + 2y^3 + 4,$$
$$\frac{\partial f}{\partial y} = 6xy^2.$$

Naravno lahko definiramo tudi višje parcialne odvode. Če parcialni odvod na spremenljivko x_j obstaja, je spet funkcija $\frac{\partial f}{\partial x_j} : U \to \mathbb{R}$. Drugi parcialni odvod definiramo s predpisom

$$\frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_j}(f) \right) = \frac{\partial \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_k}.$$

Dobimo ga tako, da funkcijo f najprej odvajamo na spremenljivko x_j, nato pa dobljeno funkcijo še na spremenljivko x_k. Zaradi enostavnosti včasih uporabljamo krajšo oznako

$$\frac{\partial^2 f}{\partial x_k \partial x_j} = (f_{x_j})_{x_k} = f_{x_j x_k}.$$

Na podoben način lahko definiramo tudi parcialne odvode višjih redov. Z oznako

$$\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}} \left(\ldots \left(\frac{\partial}{\partial x_{i_2}} \left(\frac{\partial}{\partial x_{i_1}}(f) \right) \right) \right)$$

označimo k-ti parcialni odvod funkcije f na spremenljivke $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$. Če dvakrat ali večkrat odvajamo na isto spremenljivko, to krajše zapišemo z oznako

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial x \partial x}.$$

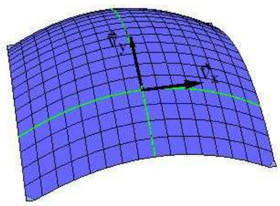
Zgled. Drugi parcialni odvodi funkcije $f(x, y) = x^2 + 2xy^3 + 4x$ so enaki:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= 2, \\
\frac{\partial^2 f}{\partial x \partial y} &= 6y^2, \\
\frac{\partial^2 f}{\partial y \partial x} &= 6y^2, \\
\frac{\partial^2 f}{\partial y^2} &= 12xy.
\end{aligned}$$

Vidimo, da sta mešana parcialna odvoda enaka. To ni slučaj, ni pa vedno res. Brez dokaza omenimo dejstvo, da ni važno, v kakšnem vrstnem redu računamo parcialne odvode, v kolikor vsi parcialni odvodi obstajajo in so zvezne funkcije.

Definicija 7. *Naj bo $f : U^{odp} \subset \mathbb{R}^n \to \mathbb{R}$ funkcija. Potem je $f \in C^1(U)$, če vsi parcialni odvodi $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}$ obstajajo in so zvezni na U. Funkcije v $C^1(U)$ imenujemo zvezno odvedljive funkcije.*

Če je $f : U^{\text{odp}} \subset \mathbb{R}^2 \to \mathbb{R}$ funkcija dveh spremenljivk, nam prva parcialna odvoda funkcije f omogočata izračunati enačbo tangentne ravnine v dani točki.



Naklona krivulj, ki jih dobimo, če graf funkcije f presekamo z ravninama oblike $y = $ konst in $x = $ konst, sta enaka $\frac{\partial f}{\partial x}$ in $\frac{\partial f}{\partial y}$. Smerna vektorja tangent na ti dve krivulji sta torej

$$\begin{aligned}
v_x &= \left(1, 0, \frac{\partial f}{\partial x}\right), \\
v_y &= \left(0, 1, \frac{\partial f}{\partial y}\right).
\end{aligned}$$

Normala tangentne ravnine mora biti pravokotna na oba ta dva vektorja, zato lahko vzamemo za normalo kar vektor

$$n = v_y \times v_x = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1 \right).$$

Od tod dobimo normalno obliko enačbe tangentne ravnine na graf funkcije f v točki (x_0, y_0)

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - (z - z_0) = 0,$$

kjer je $z_0 = f(x_0, y_0)$. Normalno premico na graf v točki (x_0, y_0) pa lahko opišemo z enačbo

$$\frac{x - x_0}{\frac{\partial f}{\partial x}(x_0, y_0)} = \frac{y - y_0}{\frac{\partial f}{\partial y}(x_0, y_0)} = \frac{z - z_0}{-1}.$$

Zgled. Izračunajmo tangentno ravnino in normalno premico na graf funkcije $f(x, y) = x^2 + 2xy^3 + 4x$ v točki $(x_0, y_0) = (1, 0)$. Izračunali smo že, da je $f_x = 2x + 2y^3 + 4$ in $f_y = 6xy^2$. Od tod dobimo $f_x(1, 0) = 6$ in $f_y(1, 0) = 0$. Ker je še $f(1, 0) = 5$, je enačba tangentne ravnine

$$\begin{aligned} 6(x - 1) + 0(y - 0) - (z - 5) &= 0, \\ 6x - z &= 1, \end{aligned}$$

enačba normalne premice pa

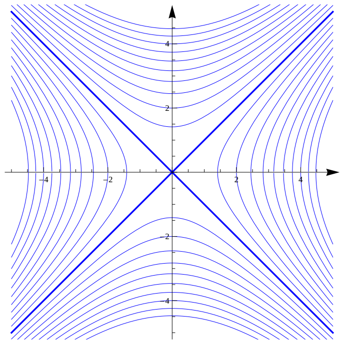
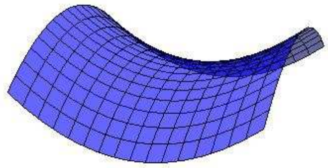
$$\frac{x - 1}{6} = \frac{z - 5}{-1},\ y = 0.$$

Namesto grafa včasih raje skiciramo izohipse oziroma nivojske krivulje funkcije dveh spremenljivk. To so krivulje, ki jih dobimo, če graf funkcije presekamo z ravninami oblike $z = \text{konst}$ in dobljeni presek nato projiciramo na xy ravnino. Če si predstavljamo graf funkcije kot relief neke pokrajine, sestavljajo nivojsko krivuljo točke, ki so na isti nadmorski višini.

Zgled. Poglejmo si nivojske krivulje funkcije $f(x, y) = x^2 - y^2$. To so krivulje, ki so implicitno določene z enačbami

$$x^2 - y^2 = c$$

za $c \in \mathbb{R}$. Če je $c \neq 0$, imajo nivojske krivulje obliko hiperbol, pri $c = 0$ pa dobimo izrojen primer $x = \pm y$, ki sestoji iz dveh premic. Graf funkcije f ima obliko sedla.

Kot smo že videli, je naklon grafa funkcije večih spremenljivk v različnih smereh različen. Naj bo f funkcija n spremenljivk in $u \in \mathbb{R}^n$ nek vektor, za katerega je $|u| = 1$. Kot bomo dokazali malce kasneje, je smerni odvod funkcije f v točki a in v smeri vektorja u enak

$$(D_u f)(a) = \nabla f(a) \cdot u = |u||\nabla f(a)| \cos \phi = |\nabla f(a)| \cos \phi.$$

Ta izraz bo maksimalen, ko bo $\cos \phi = 1$ oziroma, ko bo vektor u kazal v smeri gradienta funkcije f v točki a. Od tod sledi, da gradient $\nabla f(a)$ kaže v smeri najhitrejšega naraščanja funkcije f v točki a.

Za dokaz tega dejstva si bomo pomagali s pomožno trditvijo. Najprej bomo definirali limito funkcije večih spremenljivk. Naj bo $g : U \subset \mathbb{R}^n \to \mathbb{R}$ funkcija večih spremenljivk in b stekališče množice U. Potem rečemo, da je

$$\lim_{v \to b} g(v) = A,$$

če za vsak $\epsilon > 0$ obstaja tak $\delta > 0$, da za vsak $u \in K(b, \delta) \cap U$, $u \neq b$, velja $|g(u) - A| < \epsilon$.

Trditev 8. *Naj bo $f \in C^1(U)$, kjer je $U^{odp} \subset \mathbb{R}^n$, in naj bo $a \in U$. Potem velja*

$$\lim_{w \to 0} \frac{f(a + w) - f(a) - \nabla f(a) \cdot w}{|w|} = 0.$$

Dokaz. Izberimo poljuben $\epsilon > 0$. Ker so vsi parcialni odvodi funkcije f zvezni, lahko najdemo tak $r > 0$, da je

$$\left| \frac{\partial f}{\partial x_j}(v) - \frac{\partial f}{\partial x_j}(a) \right| < \frac{\epsilon}{n}$$

za vsak $v \in K(a, r)$ in za vsak $j = 1, 2, \ldots, n$. Izberimo sedaj poljuben $w \in \mathbb{R}^n$, za katerega je $|w| < r$, in naj bo $w = (h_1, h_2, \ldots, h_n)$. Definirajmo

$$\begin{aligned}
w_0 &= 0, \\
w_1 &= (h_1, 0, 0, \ldots, 0) = h_1 e_1, \\
w_2 &= (h_1, h_2, 0, \ldots, 0) = h_1 e_1 + h_2 e_2, \\
&\vdots \\
w_n &= (h_1, h_2, \ldots, h_n) = w.
\end{aligned}$$

Iz definicije sledi, da je $|w_j| < |w| < r$ za vsak $j = 1, 2, \ldots, n$, zato so vse točke oblike $a + w_j$ vsebovane v $K(a, r)$. Torej je v $K(a, r)$ vsebovana tudi pot, ki jo dobimo tako, da zaporedoma z daljicami povežemo točke $a, a + w_1, a + w_2, \ldots, a + w$. Zožitev funkcije f na daljico med točkama $a + w_{j-1}$ in $a + w_j$ je odvisna samo od koordinate x_j, ki teče od 0 do h_j, zato nam aplikacija Lagrangevega izreka da enakost

$$f(a + w_j) - f(a + w_{j-1}) = h_j \frac{\partial f}{\partial x_j}(a + w_{j-1} + \theta_j h_j e_j),$$

kjer je $\theta_j \in (0, 1)$. Če seštejemo izraze na levi strani zgornje enakosti, ko j teče od 1 do n, dobimo enakost

$$f(a + w) - f(a) = \sum_{j=1}^{n} \left(f(a + w_j) - f(a + w_{j-1})\right),$$

s pomočjo katere dobimo oceno

$$\begin{aligned}
|f(a + w) - f(a) - \nabla f(a) \cdot w| &= \left| f(a + w) - f(a) - \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) \right|, \\
&= \left| \sum_{j=1}^{n} \left(f(a + w_j) - f(a + w_{j-1})\right) - \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) \right|, \\
&= \left| \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a + w_{j-1} + \theta_j h_j e_j) - \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) \right|, \\
&= \left| \sum_{j=1}^{n} h_j \left(\frac{\partial f}{\partial x_j}(a + w_{j-1} + \theta_j h_j e_j) - \frac{\partial f}{\partial x_j}(a) \right) \right|.
\end{aligned}$$

Ko upoštevamo, da je $|h_j| \leq |w|$, da so parcialni odvodi funkcije f zvezni in uporabimo trikotniško neenakost, dobimo, da je

$$\begin{aligned}
|f(a + w) - f(a) - \nabla f(a) \cdot w| &= \left| \sum_{j=1}^{n} h_j \left(\frac{\partial f}{\partial x_j}(a + w_{j-1} + \theta_j h_j e_j) - \frac{\partial f}{\partial x_j}(a) \right) \right|, \\
&\leq \sum_{j=1}^{n} |h_j| \left| \frac{\partial f}{\partial x_j}(a + w_{j-1} + \theta_j h_j e_j) - \frac{\partial f}{\partial x_j}(a) \right|, \\
&\leq \sum_{j=1}^{n} |w| \frac{\epsilon}{n}, \\
&= |w|\epsilon.
\end{aligned}$$

Če je torej $|w| < r$, je

$$\frac{|f(a+w) - f(a) - \nabla f(a) \cdot w|}{|w|} < \epsilon,$$

to pa smo želeli pokazati. $\square$

Direktna posledica te trditve je naslednji rezultat.

Trditev 9. *Naj bo $f \in C^1(U)$, kjer je $U^{odp} \subset \mathbb{R}^n$, in naj bo $u \in \mathbb{R}^n$. Potem smerni odvod $D_u f$ obstaja v vsaki točki $a \in U$ in velja*

$$(D_u f)(a) = \nabla f(a) \cdot u.$$

Dokaz. Iz prejšnje trditve sledi, da velja

$$f(a+w) - f(a) = \nabla f(a) \cdot w + r(w),$$

kjer je $w : U \to \mathbb{R}$ neka funkcija, za katero je $\lim_{w \to 0} \frac{r(w)}{|w|} = 0$. Ta limita velja za poljuben $w \in \mathbb{R}^n$, ki se približuje 0. V posebnem torej ta limita velja, če vzamemo $w = hu$, kjer je $h \in \mathbb{R}$, $u \in \mathbb{R}^n$ pa nek fiksni neničelni vektor. Če v zgornjo enakost vstavimo $w = hu$, tako dobimo

$$\begin{aligned} f(a+hu) - f(a) &= h\,(\nabla f(a) \cdot u) + r(hu), \\ \frac{f(a+hu) - f(a)}{h} &= \nabla f(a) \cdot u + \frac{r(hu)}{h}. \end{aligned}$$

Ko pošljemo $h \to 0$, dobimo na levi strani v limiti smerni odvod $(D_u f)(a)$. Na desni strani pa z upoštevanjem limite

$$0 = \lim_{w \to 0} \frac{r(w)}{|w|} = \lim_{h \to 0} \frac{r(hu)}{|h||u|}$$

dobimo v limiti izraz $\nabla f(a) \cdot u$. Sledi $(D_u f)(a) = \nabla f(a) \cdot u$. $\square$

Pravila za odvajanje funkcij večih spremenljivk so enaka kot pri funkcijah ene spremenljivke. Naj bosta $g : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}^m$ vektorska funkcija in $f : E^{\text{odp}} \subset \mathbb{R}^m \to \mathbb{R}$ skalarna funkcija. Denimo, da je $g = (g_1, g_2, \ldots, g_m)$, kjer so $g_1, g_2, \ldots, g_m \in C^1(U)$, da je $g(U) \subset E$ in da velja $f \in C^1(E)$. Potem je $f \circ g \in C^1(U)$ in velja

$$\frac{\partial (f \circ g)}{\partial x_j}(a) = (\nabla f(g(a))) \cdot ((D_j g)(a)),$$

kjer je $(D_j g)(a) = \left(\frac{\partial g_1}{\partial x_j}(a), \ldots, \frac{\partial g_m}{\partial x_j}(a)\right)$. Če označimo z $x_1, x_2, \ldots, x_n$ spremenljivke v $\mathbb{R}^n$, z $u_1, u_2, \ldots, u_m$ pa spremenljivke v $\mathbb{R}^m$, lahko to na dolgo zapišemo v obliki

$$
\begin{aligned}
\frac{\partial(f\circ g)}{\partial x_j}(a) &= \sum_{i=1}^{m} \frac{\partial f}{\partial u_i}(g(a))\frac{\partial g_i}{\partial x_j}(a),\\
&= \frac{\partial f}{\partial u_1}(g(a))\frac{\partial g_1}{\partial x_j}(a) + \frac{\partial f}{\partial u_2}(g(a))\frac{\partial g_2}{\partial x_j}(a) + \ldots + \frac{\partial f}{\partial u_m}(g(a))\frac{\partial g_m}{\partial x_j}(a).
\end{aligned}
$$

To je verižno pravilo za odvajanje funkcij večih spremenljivk. Dokazali ga bomo pri Matematiki 2.

Zgled. Vzemimo funkcijo $g : \mathbb{R}^2 \to \mathbb{R}^2$, definirano s predpisom

$$g(r,\phi) = (r\cos\phi, r\sin\phi).$$

Ta funkcija nam predstavlja prehod iz polarnih v kartezične koordinate, saj bi lahko pisali tudi $g = (g_1, g_2) = (x, y)$, kjer je $x(r,\phi) = r\cos\phi$ in $y(r,\phi) = r\sin\phi$.

Naj bo sedaj $f : \mathbb{R}^2 \to \mathbb{R}$ funkcija dveh spremenljivk. Mislimo si, da je $f = f(x,y)$ odvisna od kartezičnih koordinat, zanima pa nas, kakšni so parcialni odvodi te funkcije, če jo podamo v polarnih koordinatah. Najprej je $f\circ g = f(r\cos\phi, r\sin\phi)$, od tod pa z uporabo verižnega pravila dobimo

$$
\begin{aligned}
\frac{\partial(f\circ g)}{\partial r}(r,\phi) &= \frac{\partial f}{\partial x}(r\cos\phi, r\sin\phi)\cdot\frac{\partial x}{\partial r}(r,\phi) + \frac{\partial f}{\partial y}(r\cos\phi, r\sin\phi)\cdot\frac{\partial y}{\partial r}(r,\phi),\\
&= \frac{\partial f}{\partial x}(r\cos\phi, r\sin\phi)\cdot\cos\phi + \frac{\partial f}{\partial y}(r\cos\phi, r\sin\phi)\cdot\sin\phi,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial(f\circ g)}{\partial \phi}(r,\phi) &= \frac{\partial f}{\partial x}(r\cos\phi, r\sin\phi)\cdot\frac{\partial x}{\partial \phi}(r,\phi) + \frac{\partial f}{\partial y}(r\cos\phi, r\sin\phi)\cdot\frac{\partial y}{\partial \phi}(r,\phi),\\
&= \frac{\partial f}{\partial x}(r\cos\phi, r\sin\phi)\cdot(-r\sin\phi) + \frac{\partial f}{\partial y}(r\cos\phi, r\sin\phi)\cdot r\cos\phi.
\end{aligned}
$$

Poglejmo si še čisto konkreten zgled. Vzemimo funkcijo

$$f(x,y) = x^2 + y^2.$$

Potem je

$$(f\circ g)(r,\phi) = f(r\cos\phi, r\sin\phi) = r^2\cos^2\phi + r^2\sin^2\phi = r^2.$$

Vidimo, da je v polarnih koordinatah funkcija f odvisna od spremenljivke r, ne pa od spremenljivke ϕ. Zato pričakujemo, da bo njen parcialni odvod na spremenljivko ϕ enak nič. Ker je $f_x = 2x$ in $f_y = 2y$, je

$$
\begin{aligned}
\frac{\partial(f\circ g)}{\partial r}(r,\phi) &= 2x\cdot\cos\phi + 2y\cdot\sin\phi = 2r\cos^2\phi + 2r\sin^2\phi = 2r,\\
\frac{\partial(f\circ g)}{\partial \phi}(r,\phi) &= 2x\cdot(-r\sin\phi) + 2y\cdot r\cos\phi,\\
&= -2r^2\cos\phi\sin\phi + 2r^2\sin\phi\cos\phi = 0.
\end{aligned}
$$

Pri računanju s parcialnimi odvodi nam pogosto pride prav naslednji rezultat, ki ga bomo omenili brez dokaza. Za funkcijo $f : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}$ rečemo, da je dvakrat zvezno odvedljiva, če obstajajo vsi njeni prvi in drugi parcialni odvodi in so zvezne funkcije. Označimo $f \in C^2(U)$.

Trditev 10. *Naj bo $f : U^{odp} \subset \mathbb{R}^n \to \mathbb{R}$ dvakrat zvezno odvedljiva funkcija. Tedaj velja*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

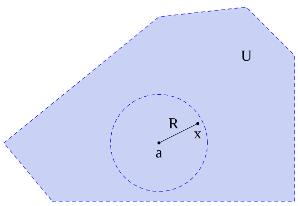
za vse $i, j = 1, 2, \ldots, n$.

Podobne enakosti veljajo tudi za višje mešane odvode, če je funkcija dovolj lepa. Funkcija $f : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}$ je m-krat zvezno odvedljiva, če obstajajo vsi parcialni odvodi do reda m in so zvezne funkcije. Če je f m-krat zvezno odvedljiva, so mešani odvodi do reda m odvisni le od spremenljivk na katere odvajamo, nič pa od vrstnega reda po katerem odvajamo.

Taylorjeva formula za funkcije večih spremenljivk

V tem poglavju bomo Taylorjev razvoj funkcije ene spremnljivke posplošili na funkcije večih spremenljivk. V tem primeru bomo funkcije aproksimirali s polinomi večih spremenljivk.

Naj bo $f : U^{\text{odp}} \subset \mathbb{R}^n \to \mathbb{R}$ m-krat zvezno odvedljiva funkcija in naj bo $a \in U$. Ker je U odprta, lahko najdemo tak $R > 0$, da je $K(a, R) \subset U$. Za poljuben $x \in K(a, R)$ potem cela daljica od a do x leži v $K(a, R)$.



Označimo s $h = x - a$ vektor, ki kaže od točke a do točke x. Daljico od a do x lahko potem parametriziramo z vektorsko funkcijo $\gamma : [0, 1] \to \mathbb{R}^n$, definirano s predpisom

$$\gamma(t) = a + th.$$

Funkcija γ je poljubnokrat odvedljiva, zanjo pa velja $\gamma(0) = a$, $\gamma(1) = x$ in $\gamma'(t) = h$. Vrednost $f(a)$ poznamo, radi pa bi izračunali (približno) vrednost $f(x)$. Če zožimo funkcijo f na daljico med a in x, bomo dobili funkcijo ene spremenljivke, za katero pa že poznamo Taylorjevo formulo.

Definirajmo sedaj funkcijo (ene realne spremenljivke) $g : [0, 1] \to \mathbb{R}$ s predpisom

$$g(t) = f(\gamma(t)).$$

Ker je funkcija f m-krat zvezno odvedljiva, je tudi funkcija g m-krat zvezno odvedljiva in velja $g(0) = f(a)$ ter $g(1) = f(x)$. Taylorjeva formula nam sedaj pove, da obstaja tak $\theta \in (0,1)$, da velja

$$g(1) = \sum_{k=0}^{m-1} \frac{g^{(k)}(0)(1-0)^k}{k!} + \frac{g^{(m)}(\theta)}{m!}1^m = \sum_{k=0}^{m-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m)}(\theta)}{m!}.$$

Da bi lahko zapisali Taylorjevo formulo za funkcijo f, moramo najprej podrobno analizirati odvode funkcije g in jih izraziti s pomočjo parcialnih odvodov funkcije f.

Omenili smo že, da je $\gamma'(t) = h$. Če pišemo $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ in $h = (h_1, h_2, \ldots, h_n)$, to pomeni, da je $\gamma_j'(t) = h_j$ za vsak $j = 1, 2, \ldots, n$ in za vsak $t \in (0,1)$. Z uporabo verižnega pravila sedaj dobimo

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma(t)) \cdot \gamma_i'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+th) \cdot h_i = \sum_{i=1}^{n} (D_i f)(a+th)h_i.$$

V zgornji vsoti so h_i konstante, medtem ko so členi tipa $(D_i f)(a+th)$ ponovno funkcije podobnega tipa kot funkcija g, le da namesto funkcije f nastopa odvod $D_i f$. S ponovno uporabo verižnega pravila dobimo

$$g''(t) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} (D_j(D_i f))(a+th) \cdot h_j \right) \cdot h_i = \sum_{i,j=1}^{n} (D_{ji} f)(a+th) \cdot h_j \cdot h_i,$$

v splošnem pa

$$g^{(k)}(t) = \sum_{i_1,\ldots,i_k=1}^{n} (D_{i_1 \ldots i_k} f)(a+th) \cdot h_{i_1} \cdot \ldots \cdot h_{i_k}.$$

Če vstavimo $t = 0$, dobimo Taylorjevo formulo za funkcijo večih spremenljivk

$$f(x) = \sum_{k=0}^{m-1} \frac{1}{k!} \cdot \left(\sum_{i_1,\ldots,i_k=1}^{n} (D_{i_1 \ldots i_k} f)(a) \cdot h_{i_1} \cdot \ldots \cdot h_{i_k} \right) + \frac{1}{m!} \sum_{i_1,\ldots,i_m=1}^{n} (D_{i_1 \ldots i_m} f)(a+\theta h) \cdot h_{i_1} \cdot \ldots \cdot h_{i_m}$$

za nek $\theta \in (0,1)$. V zgornjem izrazu tečejo indeksi $i_1, i_2, \ldots, i_k$ od 1 do n, s tem da se lahko vsak indeks ponovi večkrat. Ker je funkcija f m-krat zvezno odvedljiva, pa vrstni red odvajanja ni pomemben, zato lahko privzamemo, da pri vsakem odvodu najprej nekajkrat odvajamo na prvo spremenljivko, nato nekajkrat na drugo spremenljivko itd. Vzemimo na primer parcialni odvod

$$D_{i_1 \ldots i_k} f = \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}.$$

Recimo, da jih je med temi indeksi r_1 enakih 1, r_2 enakih 2 in v splošnem r_j enakih j za $3 \leq j \leq n$. Potem je

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_1^{r_1} \partial x_2^{r_2} \dots \partial x_n^{r_n}}.$$

Vidimo torej, da različna zaporedja i-jev lahko dajo isto zaporedje r-ov. Sedaj nas bo zanimalo, koliko različnih zaporedij nam da iste r-e. Z nekaj kombinatorike lahko preverimo, da je takšnih možnosti natanko

$$\frac{k!}{r_1! r_2! \dots r_n!}.$$

Če so vsi indeksi med sabo različni, je imenovalec enak 1, različnih možnosti pa je toliko, kot je permutacij k elementov. Če pa se kakšen indeks ponovi, pa moramo upoštevati, da istih indeksov ne ločimo med sabo, zato je izraz treba deliti s številom ustreznih permutacij. Sedaj lahko pišemo

$$\sum_{i_1,\dots,i_k=1}^{n} (D_{i_1 \dots i_k} f)(a) \cdot h_{i_1} \cdot \dots \cdot h_{i_k} = \sum_{\substack{0 \leq r_1,\dots,r_n \leq k \\ r_1+\dots+r_n=k}} \frac{k!}{r_1! \dots r_n!} \frac{\partial^k f}{\partial x_1^{r_1} \dots \partial x_n^{r_n}}(a) \cdot h_1^{r_1} \dots h_n^{r_n},$$

od koder sledi, da je

$$\begin{aligned} f(x) = \sum_{k=0}^{m-1} & \left(\sum_{\substack{0 \leq r_1,\dots,r_n \leq k \\ r_1+\dots+r_n=k}} \frac{\partial^k f}{\partial x_1^{r_1} \dots \partial x_n^{r_n}}(a) \cdot \frac{h_1^{r_1}}{r_1!} \cdot \dots \cdot \frac{h_n^{r_n}}{r_n!} \right) \\ & + \sum_{\substack{0 \leq r_1,\dots,r_n \leq m \\ r_1+\dots+r_n=m}} \frac{\partial^k f}{\partial x_1^{r_1} \dots \partial x_n^{r_n}}(a+\theta h) \cdot \frac{h_1^{r_1}}{r_1!} \cdot \dots \cdot \frac{h_n^{r_n}}{r_n!} \end{aligned}$$

za nek $\theta \in (0,1)$.

Zgled. Zgornja formula izgleda precej zapletena, zato si bomo pogledali nekaj najpogostejših posebnih primerov. Taylorjev razvoj funkcije do členov drugega reda se glasi

$$f(x) = f(a) + \left(\frac{\partial f}{\partial x_1}(a) h_1 + \dots + \frac{\partial f}{\partial x_n}(a) h_n \right) + \left(\frac{1}{2!} \sum_{1 \leq i,j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j \right) + \dots$$

V primeru, ko je $n = 2$, spremenljivki ponavadi označimo z x in y. Če razvijamo okoli točke (a,b), je $h_x = x-a$, $h_y = y-b$ in

$$\begin{aligned} f(x,y) = f(a,b) &+ \left(\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)\right) + \\ &+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2 + \frac{\partial^2 f}{\partial x \partial y}(a,b)(x-a)(y-b) + \frac{\partial^2 f}{\partial y^2}(a,b)(y-b)^2 + \right. \\ &\left. + \frac{\partial^2 f}{\partial y \partial x}(a,b)(y-b)(x-a)\right) + \ldots \end{aligned}$$

Zgled. Poglejmo si še konkreten zgled. Razvijmo funkcijo dveh spremenljivk $f(x,y) = x^3 + y^2 + xy$ do členov reda 3 okoli točke $(1,2)$.

Parcialni odvodi funkcije f so: $f_x = 3x^2 + y$, $f_y = 2y + x$, $f_{xx} = 6x$, $f_{xy} = f_{yx} = 1$, $f_{yy} = 2$ in $f_{xxx} = 6$. Ostali odvodi reda tri so ničelni. Ko izračunamo vrednosti teh odvodov v točki $(1,2)$, dobimo Taylorjev razvoj

$$\begin{aligned} f(x,y) \approx 7 &+ \Big(5(x-1) + 5(y-2)\Big) + \\ &+ \frac{1}{2}\Big(6(x-1)^2 + 2(y-2)^2 + 2 \cdot 1(x-1)(y-2)\Big) + \\ &+ \frac{1}{3!}6(x-1)^3 \end{aligned}$$

oziroma

$$f(x,y) \approx 7 + 5(x-1) + 5(y-2) + 3(x-1)^2 + (y-2)^2 + (x-1)(y-2) + (x-1)^3.$$

Ekstremi funkcij večih spremenljivk

V tem poglavju bomo obrazložili potrebne in zadostne pogoje za nastop ekstrema funkcije večih spremenljivk. Videli bomo, da so ti pogoji posplošitev pogojev pri funkcijah ene spremenljivke. Poleg lokalnih maksimumov in minimumov v večih spremenljivkah obstajajo tudi stacionarne točke sedlastega tipa, ki niso ekstremi, a imajo vseeno netrivialne druge odvode.

Definicija 11. *Naj bo $f : U^{odp} \subset \mathbb{R}^n \to \mathbb{R}$ in $a \in U$.*

(i) Funkcija f ima v točki a (strogi) lokalni maksimum, če obstaja $\epsilon > 0$, da za vsak $v \in (U \cap K(a,\epsilon)) \setminus \{a\}$ velja:

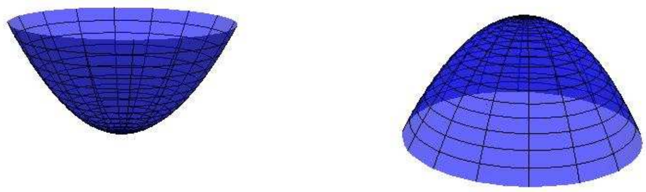
$$(f(v) < f(a)) \quad f(v) \leq f(a).$$

(ii) Funkcija f ima v točki a (strogi) lokalni minimum, če obstaja $\epsilon > 0$, da za vsak $v \in (U \cap K(a,\epsilon)) \setminus \{a\}$ velja:

$$(f(v) > f(a)) \quad f(v) \geq f(a).$$

Če velja bodisi (i) ali (ii), pravimo, da ima funkcija f v točki a (strogi) lokalni ekstrem.

V primeru funkcij dveh spremenljivk izgleda lokalni maksimum kot vrh gore, lokalni minimum pa kot dno kotanje.



Definicija 12. *Naj bo $f : U^{odp} \subset \mathbb{R}^n \to \mathbb{R}$ zvezno odvedljiva funkcija. Točka $a \in U$ je stacionarna točka funkcije f, če je $\nabla f(a) = 0$.*

Trditev 13 (Potreben pogoj za ekstrem)**.** *Naj bo $f : U^{odp} \subset \mathbb{R}^n \to \mathbb{R}$ zvezno odvedljiva funkcija. Če ima f lokalni ekstrem v točki $a \in U$, je a stacionarna točka funkcije f.*

Dokaz. Vzemimo poljuben neničeln vektor $u \in \mathbb{R}^n$ in definirajmo vektorsko funkcijo $\gamma \in C^1(\gamma^{-1}(U))$ s predpisom

$$\gamma(t) = a + tu.$$

Potem je funkcija $f \circ \gamma : \gamma^{-1}(U) \subset \mathbb{R} \to \mathbb{R}$ zvezno odvedljiva. Ker ima f v točki a lokalni ekstrem, ima $f \circ \gamma$ v točki 0 lokalni ekstrem, kar pa pomeni, da je $(f \circ \gamma)'(0) = 0$. Verižno pravilo nam pove, da mora torej veljati

$$(f \circ \gamma)'(0) = \nabla f(\gamma(0)) \cdot \gamma'(0) = 0.$$

Funkcijo γ smo definirali tako, da je $\gamma(0) = a$ in $\gamma'(0) = u$, zato mora veljati

$$\nabla f(a) \cdot u = 0.$$

Ker je bil vektor u poljuben, je vektor $\nabla f(a)$ pravokoten na cel $\mathbb{R}^n$, kar pa pomeni, da je $\nabla f(a) = 0$. $\square$

Za zvezno odvedljive funkcije je torej ničeln gradient potreben pogoj za nastop ekstrema. Pri funkcijah ene spremenljivke smo videli, da je zadosten pogoj za nastop ekstrema neničeln drugi odvod, medtem ko so pri funkcijah večih spremenljivk stvari bolj komplicirane. Nastop ekstrema funkcije je odvisen od Hessejeve matrike drugih odvodov, ki jo bomo spoznali pri Matematiki 2. Pri funkcijah dveh spremenljivk lahko ekstreme klasificiramo s pomočjo naslednje pomožne funkcije.

Definicija 14. *Naj bo $f : U^{odp} \subset \mathbb{R}^2 \to \mathbb{R}$ dvakrat zvezno odvedljiva funkcija dveh spremenljivk. Potem definiramo funkcijo*

$$K(f)(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y) \cdot \frac{\partial^2 f}{\partial y^2}(x,y) - \left(\frac{\partial^2 f}{\partial x \partial y}(x,y)\right)^2.$$

Trditev 15 (Zadostni pogoj za ekstrem)**.** *Naj bo $f \in C^2(U)$, kjer je $U^{odp} \subset \mathbb{R}^2$, in naj bo (x_0, y_0) stacionarna točka funkcije f. Potem velja:*

(i) Če je $K(f)(x_0,y_0) > 0$, ima f v (x_0,y_0) strogi lokalni ekstrem:

(a) strogi lokalni maksimum, če je $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) < 0$, oziroma

(b) strogi lokalni minimum, če je $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) > 0$.

(ii) Če je $K(f)(x_0,y_0) < 0$, funkcija f v (x_0,y_0) nima lokalnega ekstrema (ima sedlo).

Dokaz. Če je (x_0,y_0) stacionarna točka funkcije f, nam Taylorjev razvoj funkcije f okoli te točke pove, da obstaja tak $\theta \in (0,1)$, da velja

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \frac{1}{2}\Bigg(\frac{\partial^2 f}{\partial x^2}(x_0+\theta h, y_0+\theta k)h^2 + \\ + 2\frac{\partial^2 f}{\partial x \partial y}(x_0+\theta h, y_0+\theta k)hk + \frac{\partial^2 f}{\partial y^2}(x_0+\theta h, y_0+\theta k)k^2\Bigg).$$

Označimo zaradi enostavnosti

$$\begin{aligned}
\alpha(h,k) &= \frac{\partial^2 f}{\partial x^2}(x_0+\theta h, y_0+\theta k), \\
\beta(h,k) &= \frac{\partial^2 f}{\partial x \partial y}(x_0+\theta h, y_0+\theta k), \\
\gamma(h,k) &= \frac{\partial^2 f}{\partial y^2}(x_0+\theta h, y_0+\theta k).
\end{aligned}$$

Potem lahko pišemo

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \frac{1}{2}\left(\alpha(h,k)h^2 + 2\beta(h,k)hk + \gamma(h,k)k^2\right).$$

Zanimal nas bo predznak oklepaja na desni strani. Če je ves čas pozitiven ali pa ves čas negativen na neki majhni okolici točke (x_0,y_0), imamo v točki (x_0,y_0) ekstrem, sicer pa ekstrema ni.

Denimo, da je $K(f)(x_0,y_0) \neq 0$. Ker je $f \in C^2(U)$, so vsi drugi parcialni odvodi funkcije f zvezni, zato je zvezna tudi funkcija $K(f)$. Torej bo $K(f)$ istega predznaka kot $K(f)(x_0,y_0)$ na neki majhni okolici točke (x_0,y_0).

i) Denimo najprej, da je $K(f)(x_0, y_0) > 0$. Če je $k \neq 0$, lahko pišemo

$$\alpha h^2 + 2\beta hk + \gamma k^2 = k^2(\alpha u^2 + 2\beta u + \gamma),$$

kjer je $u = h/k$. Izraz v oklepaju je kvadratna funkcija spremenljivke u z diskriminanto

$$D = (2\beta)^2 - 4\alpha\gamma = 4(\beta^2 - \alpha\gamma) = -4K(f) < 0.$$

Diskriminanta je negativna za dovolj majhne (h, k), kar pa pomeni, da je izraz v oklepaju istega predznaka kot α. Tudi v primeru, ko je $k = 0$, je izraz

$$\alpha h^2 + 2\beta hk + \gamma k^2 = \alpha h^2$$

istega predznaka kot α. Za dovolj majhne (h, k) je torej izraz

$$\alpha h^2 + 2\beta hk + \gamma k^2$$

ves čas negativen ali pa ves čas pozitiven. Od tod sledi, da ima funkcija f v točki (x_0, y_0) ekstrem. Tip ekstrema je odvisen od predznaka α. Ker je α zvezna funkcija, je dovolj pogledati vrednosti pri $h = k = 0$. Tam dobimo

$$\alpha(0, 0) = \frac{\partial^2 f}{\partial x^2}(x_0, y_0),$$

od koder vidimo, da je v primeru $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ v točki (x_0, y_0) strogi lokalni minimum, v primeru $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ pa strogi lokalni maksimum.

ii) Naj bo sedaj $K(f)(x_0, y_0) < 0$. Radi bi pokazali, da potem obstajajo poljubno majhni pari (h_1, k_1) in (h_2, k_2), da velja

$$f(x_0 + h_1, y_0 + k_1) > f(x_0, y_0),$$
$$f(x_0 + h_2, y_0 + k_2) < f(x_0, y_0).$$

Iz pogoja $K(f)(x_0, y_0) < 0$ sledi, da ima kvadratna funkcija

$$\alpha(0, 0)u^2 + 2\beta(0, 0)u + \gamma(0, 0)$$

pozitivno diskriminanto, zato obstajata taka $u_1, u_2 \in \mathbb{R}$, da velja

$$\alpha(0, 0)u_1^2 + 2\beta(0, 0)u_1 + \gamma(0, 0) > 0,$$
$$\alpha(0, 0)u_2^2 + 2\beta(0, 0)u_2 + \gamma(0, 0) < 0.$$

Zaradi zveznosti drugih parcialnih odvodov veljata tudi za dovolj majhne pare (h, k) neenakosti

$$\alpha(h, k)u_1^2 + 2\beta(h, k)u_1 + \gamma(h, k) > 0,$$
$$\alpha(h, k)u_2^2 + 2\beta(h, k)u_2 + \gamma(h, k) < 0.$$

Izberimo sedaj h_1 in k_1 tako, da velja $u_1 = h_1/k_1$. Množica takšnih parov tvori premico v ravnini (h,k), če pa sta h_1 in k_1 dovolj majhna in $k_1 \neq 0$, pa bo veljalo

$$\begin{aligned} \alpha(h_1,k_1)u_1^2 + 2\beta(h_1,k_1)u_1 + \gamma(h_1,k_1) &> 0, \\ \alpha(h_1,k_1)h_1^2 + 2\beta(h_1,k_1)h_1k_1 + \gamma(h_1,k_1)k_1^2 &> 0. \end{aligned}$$

Našli smo torej poljubno majhne pare (h_1,k_1), za katere je

$$\begin{aligned} f(x_0+h_1,y_0+k_1) &= f(x_0,y_0) + \frac{1}{2}\left(\alpha(h_1,k_1)h_1^2 + 2\beta(h_1,k_1)h_1k_1 + \gamma(h_1,k_1)k_1^2\right), \\ &> f(x_0,y_0). \end{aligned}$$

Na podoben način lahko najdemo tudi poljubno majhne pare (h_2,k_2), ki zadoščajo pogoju $u_1 = h_2/k_2$, in za katere velja

$$f(x_0+h_2,y_0+k_2) < f(x_0,y_0).$$

Oboje skupaj nam pove, da v točki (x_0,y_0) ni ekstrema. $\square$

Zgled. (1) Poiščimo najprej ekstreme funkcije $f(x,y) = x^3 + y^3 - 3xy$. Parcialna odvoda sta $f_x = 3x^2 - 3y$ in $f_y = 3y^2 - 3x$. Stacionarne točke morajo torej zadoščati sistemu enačb sistem enačb

$$\begin{aligned} f_x &= 3x^2 - 3y = 0, \\ f_y &= 3y^2 - 3x = 0. \end{aligned}$$

Iz prve enačbe dobimo $x^2 = y$, iz druge pa $y^2 = x$. Če drugo enačbo kvadriramo in nato upoštevamo prvo enačbo, dobimo, da je $y^4 = y$. Ta enačba ima realni rešitvi $y = 0$ in $y = 1$, zato ima funkcija f dve stacionarni točki $T_1(0,0)$ in $T_2(1,1)$.

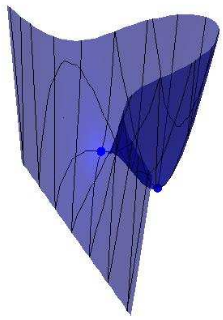
Drugi parcialni odvodi so $f_{xx} = 6x$, $f_{xy} = -3$ in $f_{yy} = 6y$. Sledi

$$K(f)(x,y) = 6x \cdot 6y - (-3)^2 = 36xy - 9.$$

V stacionarnih točkah je

$$\begin{aligned} K(f)(0,0) &= -9 < 0, \\ K(f)(1,1) &= 27 > 0, \end{aligned}$$

zato v točki T_1 ni ekstrema, v točki T_2 pa imamo strogi lokalni minimum, ker je $f_{xx}(1,1) = 6 > 0$.



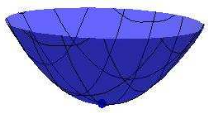
V okolici stacionarne točke T_1 ima graf obliko sedla, v točki T_2 pa je dno kotline.

(2) Poglejmo si sedaj funkcijo $f(x,y) = 2x^2+3y^2-4x+2 = 2(x-1)^2+3y^2$. Nivojske krivulje funkcije f so elipse, preseki grafa funkcije f z ravninami oblike $x =$ konst in $y =$ konst pa so parabole. Parcialna odvoda funkcije f sta $f_x = 4x-4$ in $f_y = 6y$. Stacionarna točka je ena sama, in sicer $T(1,0)$.

Drugi parcialni odvodi so $f_{xx} = 4$, $f_{xy} = 0$ in $f_{yy} = 6$, zato je

$$K(f)(x,y) = 4 \cdot 6 - 0 = 24 > 0.$$

V točki T je torej strogi lokalni minimum, ker je $f_{xx} = 4 > 0$. Graf funkcije f ima obliko paraboloida.



(3) Za konec si poglejmo še funkcijo $f(x,y) = y^2 - x^2$. Parcialna odvoda sta $f_x = -2x$ in $f_y = 2y$, kar pomeni, da imamo stacionarno točko $T(0,0)$.

Drugi parcialni odvodi so $f_{xx} = -2$, $f_{xy} = 0$ in $f_{yy} = 2$, od koder dobimo

$$K(f)(x,y) = (-2) \cdot 2 - 0 = -4 < 0.$$

V točki T torej ni ekstrema, pač pa sedlo.