Sadržaj

Poglavlje

3 Diferencijabilnost funkcije više varijabli

U ovom poglavlju govorit ćemo o drugoj važnoj osobini proizvoljnog preslikavanja, o diferencijabilnosti. Ovdje ćemo pretpostavljati uvijek, ako drugačije nije naglašeno, da svaka tačka domena D_f posmatranog preslikavanja, pripada tom skupu zajedno sa nekom svojom okolinom, to jest pretpostavljat ćemo da je skup D_f otvoren. U nekim razmatranjima bit će neophodna i osobina povezanosti (koneksnosti) tog skupa. Za takav skup (otvoren i povezan) reći ćemo da je *oblast* u prostoru $\mathbb{R}^n$.

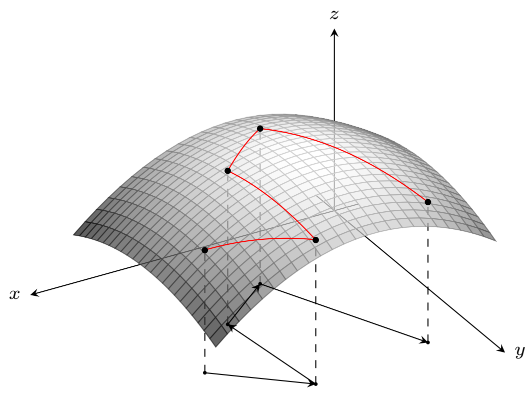
3.1 Izvod u pravcu

Za funkciju $\phi : \mathbb{R} \to \mathbb{R}$, izvod u tački $x_0 \in D_\phi$ definisali smo sa

$$\phi'(x_0) = \lim_{h\to 0} \frac{\phi(x_0+h) - \phi(x_0)}{h} , \tag{3.1}$$

i geometrijski, predstavljao je nagib tangente prema x osi (to jest najbolju linearnu aproksimaciju) na krivu ϕ u tački $(x_0, \phi(x_0))$ ili trenutnu mjeru promjene funkcije $\phi(x)$ u odnosu na varijablu x, kada je $x = x_0$. Kao uvod za nalaženje ovakve "najbolje linearne aproksimacije" za funkciju $f : \mathbb{R}^n \to \mathbb{R}$, pokušat ćemo iskoristiti, to jest generalizovati (3.1) da bi realizovali ideju "nagiba" i "mjere promjene" za ovakvo preslikavanje.

Kod funkcije jedne varijable, mogućnosti kretanja po grafu su bile samo "naprijed" i "nazad". Graf funkcije dvije varijeble je površ u $\mathbb{R}^3$ prostoru, a kretati iz neke tačke na tom grafu možemo uraditi na beskonačno mnogo načina, to jest krenuti iz tačke možemo izvesti u beskonačno mnogo pravaca. Proizvoljna tačka na grafu funkcije dvije varijable je oblika $(x, y, f(x, y))$, pri čemu je tačka (x, y) iz domena funkcije. Očigledno pomjeranje tačke na grafu funkcije je uzrokovano pomjeranjem tačke u domenu.

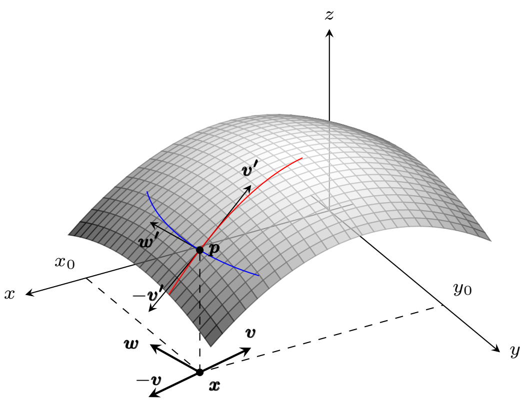


Posmatrajmo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, definisanu sa

$$f(x, y) = 4 - 2x^2 - y^2 ,$$

čiji je graf prikazan na gornjoj slici.

Ukoliko želimo da vizualiziramo kretanje po ovom grafu (površi), nagib puta po kome se krećemo prema xOy ravni, ovisi od polazne tačke ali i od pravca našeg kretanja. Naprimjer, neka je startna tačka $\boldsymbol{p} = (1,1,1)$ na površi i neka je pravac kretanja određen vektorom $\boldsymbol{v}' = (-1,-1,3)$. Ovo će uzrokovati kretanje ka "vrhu" grafa i jasno je da je "mjera promjene" rastuća. Međutim, ako se iz iste tačke krećemo u pravcu vektora $-\boldsymbol{v}'$, onda "silazimo niz graf", to jest "mjera promjene" je opadajuća. Obje ove mogućnosti naznačene su na slici crvenom bojom.



Ako iz iste tačke krenemo u pravcu vektora $\boldsymbol{w}' = \left(\frac{1}{2},-1,0\right)$, vidimo da je putanja kretanja po elipsi $2x^2 + y^2 = 3$ (pravac kretanja je paralelan xOy ravni), to jest "obilazimo" oko grafa, pa je "nagib" bez promjene, a time i "mjera promjene" je 0. Ova mogućnost kretanja je na slici prikazana plavom bojom. Dakle, govoriti o "nagibu" na graf funkcije f u tački, zahtijeva specificirati pravac kretanja.

Kretanju na grafu iz tačke $\boldsymbol{p}$, u pravcu vektora $\boldsymbol{v}'$, odgovara kretanje u domenu funkcije, iz tačke $\boldsymbol{x}$ u pravcu vektora $\boldsymbol{v} = (-1,-1)$ (projekcija tačke i vektora u 2D prostor). Analogno, kretanju u pravcu vektora $\boldsymbol{w}'$, odgovara kretanje iz $\boldsymbol{x}$ u pravcu $\boldsymbol{w} = \left(\frac{1}{2},-1\right)$. Dakle, ukoliko se krećemo iz tačke $\boldsymbol{x} = (1,1)$ u pravcu vektora

$$\boldsymbol{u} = \frac{\boldsymbol{v}}{||\boldsymbol{v}||} = -\frac{1}{\sqrt{2}}(1,1) \ ,$$

(normiranje vektora vršimo iz prostog razloga što se time pravac i smjer vektora ne mijenjaju, pa ćemo "veličinu" pomjeranja u pravcu takvog vektora diktirati sa veličinom t) tada izraz

$$\frac{f(\boldsymbol{x}+t\boldsymbol{u}) - f(\boldsymbol{x})}{t} \ ,$$

za proizvoljno t, će predstavljati aproksimaciju nagiba na graf funkcije f u tački $\boldsymbol{x}$, u pravcu $\boldsymbol{u}$. Uradimo malo računa.

$$\begin{aligned}
f(\boldsymbol{x}+t\boldsymbol{u}) - f(\boldsymbol{x}) &= f\left(1-\frac{t}{\sqrt{2}}, 1-\frac{t}{\sqrt{2}}\right) - f(1,1) \\
&= 4 - 2\left(1-\frac{t}{\sqrt{2}}\right)^2 - \left(1-\frac{t}{\sqrt{2}}\right)^2 - 1 \\
&= 3 - 3\left(1-\sqrt{2}t+\frac{t^2}{2}\right) \\
&= 3\sqrt{2}t - \frac{3t^2}{2} = t\left(3\sqrt{2}-\frac{3t}{2}\right) \ .
\end{aligned}$$

Kao što smo to radili sa funkcijama jedne varijable, dijeleći posljednji izraz sa t i puštajući da t teži ka 0, dobili bi egzaktan nagib na graf, u tački $\boldsymbol{x}$, u pravcu $\boldsymbol{u}$. Iz gornjeg onda imamo

$$\lim_{t\to 0}\frac{f(\boldsymbol{x}+t\boldsymbol{u}) - f(\boldsymbol{x})}{t} = \lim_{t\to 0}\left(3\sqrt{2}-\frac{3t}{2}\right) = 3\sqrt{2} \ .$$

Dakle, naš graf ima nagib od $3\sqrt{2}$ (naravno da ova veličina izražava tangens ugla pod kojim se krećemo) ukoliko startujemo iz tačke $\boldsymbol{x} = (1,1)$, u pravcu vektora $\boldsymbol{u}$. Sličnim računom bi dobili da je u pravcu $-\boldsymbol{u}$ nagib $-3\sqrt{2}$, odnosno u pravcu vektora

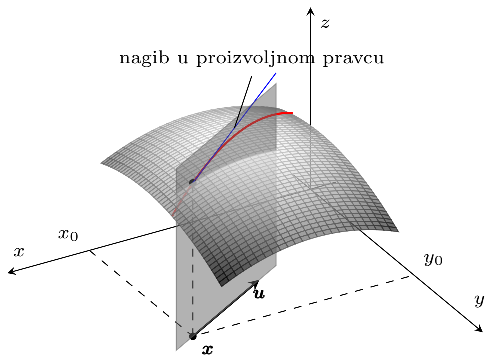
$$\frac{\boldsymbol{w}}{||\boldsymbol{w}||} = \frac{2}{\sqrt{5}}\left(\frac{1}{2},-1\right) \ ,$$

nagib je 0.

Definicija 3.1.1. *Neka je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ definisana u nekoj otvorenoj kugli oko tačke $\boldsymbol{x}$. Za dati vektor $\boldsymbol{u}$, izraz*

$$D_{\boldsymbol{u}} f(\boldsymbol{x}) = \lim_{t\to 0} \frac{f(\boldsymbol{x} + t\boldsymbol{u}) - f(\boldsymbol{x})}{t} , \tag{3.2}$$

ukoliko limes postoji, nazivamo izvod u pravcu funkcije f, u pravcu vektora $\boldsymbol{u}$, u tački $\boldsymbol{x}$.



Slika 3.1: Izvod u pravcu: $D_{\boldsymbol{u}} f(\boldsymbol{x})$.

Na primjer, za funkciju $f(\boldsymbol{x}) = f(x, y)$, izvod u pravcu vektora $\boldsymbol{u} = (u_1, u_2)$ je

$$D_{\boldsymbol{u}} f(\boldsymbol{x}) = \lim_{t\to 0} \frac{f((x, y) + t(u_1, u_2))}{t} = \lim_{t\to 0} \frac{f(x + tu_1, y + tu_2) - f(x, y)}{t} .$$

PRIMJER 1 : Prema gornjem razmatranju, za funkciju $f(x, y) = 4 - 2x^2 - y^2$, za tačku $\boldsymbol{x} = (1, 1)$ i vektor $\boldsymbol{u} = -\frac{1}{\sqrt{2}}(1, 1)$ je

$$D_{\boldsymbol{u}} f(1, 1) = \lim_{t\to 0} \frac{f\left(1 - \frac{t}{\sqrt{2}}, 1 - \frac{t}{\sqrt{2}}\right) - f(1, 1)}{t} = 3\sqrt{2},$$

$$D_{-\boldsymbol{u}} f(1, 1) = \lim_{t\to 0} \frac{f\left(1 + \frac{t}{\sqrt{2}}, 1 + \frac{t}{\sqrt{2}}\right) - f(1, 1)}{t} = -3\sqrt{2},$$

a za vektor $\boldsymbol{w} = \frac{2}{\sqrt{5}}\left(\frac{1}{2}, -1\right)$ je

$$D_{\boldsymbol{w}} f(1, 1) = \lim_{t\to 0} \frac{f\left(1 + \frac{t}{\sqrt{5}}, 1 - \frac{2t}{\sqrt{5}}\right) - f(1, 1)}{t} = 0.$$

3.2 Parcijalni izvod i parcijalni diferencijal

Kao što smo vidjeli u prethodnoj sekciji, za funkciju više varijabli ne možemo jednostavno govoriti o izvodu te funkcije, to jest možemo govoriti o izvodu ali pri tome moramo znati pravac kretanja, i tada ustvari govorimo o izvodu u pravcu. Pravac u kome nalazimo izvod funkcije više varijabli može biti proizvoljan, ali pravci određeni baznim vektorima prostora domena su od posebne važnosti. Neka su $\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n$ standardni vektori baze prostora $\mathbb{R}^n$,

$$\boldsymbol{e}_1 = (1, 0, 0, ..., 0) , \ \boldsymbol{e}_2 = (0, 1, 0, ..., 0) \ \cdots \ \boldsymbol{e}_n = (0, 0, 0, ..., 1) .$$

Posmatrajmo funkciju $f : \mathbb{R}^n \to \mathbb{R}$

$$f(\boldsymbol{x}) = f(x_1, x_2, ..., x_n) ,$$

koja je definisana u nekoj okolini $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a} = (a_1, a_2, ..., a_n) \in \mathbb{R}^n$. Razmotrimo za trenutak funkciju $g : \mathbb{R} \to \mathbb{R}$, uvedenu na sljedeći način

$$g(y) = f(y, x_2, x_3, ..., x_n) ,$$

to jest definišemo je preko funkcije f, tako što počev od druge, sve varijable držimo fiksnim (ne mjenjamo ih), a samo prvu shvatimo kao varijablu. Dakle, tada je g funkcija jedne varijable pa na nju možemo primjeniti jednakost (3.1),

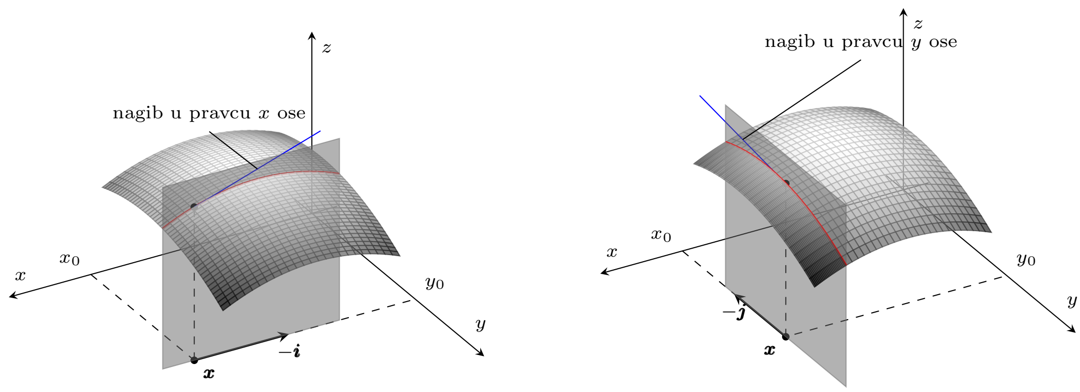
$$g'(y) = \lim_{t\to 0} \frac{g(y+t)-g(y)}{t} \ .$$

Ali tada imamo

$$\begin{aligned} g'(x_1) &= \lim_{t\to 0} \frac{g(x_1+t)-g(x_1)}{t} \\ &= \lim_{t\to 0} \frac{f(x_1+t, x_2, ..., x_n) - f(x_1, x_2, ..., x_n)}{t} \\ &= \lim_{t\to 0} \frac{f((x_1, x_2, ..., x_n) + (t, 0, ..., 0)) - f(x_1, x_2, ..., x_n)}{t} \\ &= \lim_{t\to 0} \frac{f(\boldsymbol{x} + t\boldsymbol{e}_1) - f(\boldsymbol{x})}{t} = D_{\boldsymbol{e}_1} f(\boldsymbol{x}) \ . \end{aligned}$$

Vidimo da je izvod funkcije g u tački x_1 u stvari izvod u pravcu funkcije f u tački $\boldsymbol{x}$, u pravcu vektora $\boldsymbol{e}_1$.
Na isti način smo mogli fiksirati proizvoljnu k-tu promjenljivu ($k \in \{1, 2, ..., n\}$) funkcije f, to jest staviti da je $g(y) = f(x_1, x_2, ..., x_{k-1}, y, x_{k+1}, ..., x_n)$ i zaključiti da bi vrijedilo

$$g'(x_k) = D_{\boldsymbol{e}_k} f(\boldsymbol{x}) \ .$$



(a) $\frac{\partial f}{\partial x}(\boldsymbol{x})$ (b) $\frac{\partial f}{\partial y}(\boldsymbol{x})$

Definicija 3.2.1. *Neka je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ definisana u nekoj okolini tačke $\boldsymbol{a}$ i neka je $\boldsymbol{e}_k$ ($k \in \{1, 2, ..., n\}$) k-ti vektor standardne baze u $\mathbb{R}^n$. Ukoliko postoji, izvod u pravcu $D_{\boldsymbol{e}_k} f(\boldsymbol{a})$ nazivamo parcijalni izvod funkcije f po promjenljivoj x_k, u tački $\boldsymbol{a}$.*

Naravno da smo pojam parcijalnog izvoda mogli uvesti i na mnogo formalniji način, uvodeći pojmove priraštaja.

Definicija 3.2.2. *Neka je $U_{\boldsymbol{a}} \subseteq \mathbb{R}^n$ okolina tačke $\boldsymbol{a} = (a_1, a_2, ..., a_n)$ i $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in U_{\boldsymbol{a}}$ proizvoljna. Razliku*

$$\triangle x_k = x_k - a_k \quad ; \quad k = 1, 2, ..., n$$

nazivamo priraštajem varijable x_k, a razliku

$$\triangle_{x_k} f(\boldsymbol{x}) = f(x_1, ..., x_k + \triangle x_k, ..., x_n) - f(x_1, ..., x_n)$$

nazivamo parcijalnim priraštajem funkcije f po promjenljivoj x_k, u tački $\boldsymbol{x}$.

Primjećujemo da parcijalni priraštaj funkcije n promjenljivih dobijamo tako što vršimo promjenu samo jedne varijable dok ostale varijable držimo fiksnim. Za funkciju dvije varijable imali bi parcijalne priraštaje,

$$\triangle_x f(x,y) = f(x+\triangle x, y) - f(x,y) \;\text{ i }\; \triangle_y f(x,y) = f(x, y+\triangle y) - f(x,y)\ ,$$

pri čemu posmatramo promjenu funkcije f samo u pravcu x-ose, odnosno samo u pravcu y-ose.

Definicija 3.2.3. *Ako postoji granična vrijednost*

$$\lim_{\triangle x_k \to 0} \frac{\triangle_{x_k} f(\boldsymbol{a})}{\triangle x_k} = \lim_{x_k \to a_k} \frac{f(a_1, ..., x_k, ..., a_n) - f(a_1, ...a_k, ..., a_n)}{x_k - a_k}\ ,$$

nazivamo je parcijalni izvod funkcije f po promjenljivoj x_k u tački $\boldsymbol{a}$.

U različitim knjigama matematičke analize nalazimo razne oznake za parcijalne izvode, kao npr.

$$f'_{x_k}\ ;\ f_{x_k}\ ;\ D_{x_k} f\ ;\ \frac{\partial f}{\partial x_k}.$$

Mi ćemo najčešće koristiti oznaku $\frac{\partial f}{\partial x_k}$, zato primjetimo da ovdje nismo koristili označavanje koje smo imali kod funkcije jedne promjenljive, to jest oznaku $\frac{df}{dx}$. Razlog za to je činjenica da izraz $\frac{\partial f}{\partial x_k}$ ni u kom slučaju ne možemo shvatiti kao dijeljenje (∂f sa ∂x), što jeste bio slučaj sa izrazom $\frac{df}{dx}$ ($df = f'(x)dx$).

PRIMJER 2 : Koristeći se definicijom parcijalnog izvoda odrediti iste za funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, zadatu sa $f(x,y) = xy$.

$$\frac{\partial f}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)y - xy}{\Delta x} = y\ .$$

$$\frac{\partial f}{\partial y}(x,y) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} = \lim_{\Delta y \to 0} \frac{x(y+\Delta y) - xy}{\Delta y} = x\ .$$

Osim izračunavanja parcijalnih izvoda po definiciji, tehnika određivanja parcijalnih izvoda se ni u čemu ne razlikuje od tehnike izračunavanja izvoda funkcije jedne promjenljive. Pri nalaženju parcijalnog izvoda po promjenljivoj x_k, sve ostale promjenljive shvatamo kao konstante, a nalazimo izvod po x_k, koristeći pravila i tablicu izvoda funkcija jedne promjenljive.

PRIMJER 3 : $f(x,y) = \frac{x}{y}$.

$$\frac{\partial f}{\partial x} = \frac{y\frac{\partial}{\partial x}x - x\frac{\partial}{\partial x}y}{y^2} = \frac{y-0}{y^2} = \frac{1}{y}\ ,$$

$$\frac{\partial f}{\partial y} = \frac{y\frac{\partial}{\partial y}x - x\frac{\partial}{\partial y}y}{y^2} = \frac{0-x}{y^2} = -\frac{x}{y^2}\ .$$

Kod funkcije jedne varijable $y = f(x)$, ako je $x = g(t)$, imali smo pravilo izvoda složene funkcije $y = f(g(t))$ (*pravilo kompozicije*ili *lančano pravilo*), koje glasi

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt}\ .$$

Pravilo kompozicije moramo takođe imati i kod funkcija više varijabli. Pokazaćemo to pravilo za funkciju dvije varijable, a ono se jednostavno prenosi na funkcije sa n varijabli. Kao prvo razmotrimo slučaj kada je f funkcija dviju varijabli i g funkcija jedne varijable, to jest posmatrajmo slučaj kompozicije $z = g(f(x,y))$. z je ovisna o dvije varijable pa njene parcijalne izvode računamo po pravilu:

$$\frac{\partial z}{\partial x} = \frac{dg}{df} \cdot \frac{\partial f}{\partial x} \quad , \quad \frac{\partial z}{\partial y} = \frac{dg}{df} \cdot \frac{\partial f}{\partial y} \ .$$

PRIMJER 4 : $f(x,y) = \sin(xy - y)$.

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{\partial}{\partial x}\sin(xy-y) \\ &= \cos(xy-y)\frac{\partial}{\partial x}(xy-y) \\ &= \cos(xy-y)\left(\frac{\partial}{\partial x}(xy) - \frac{\partial}{\partial x}y\right) \\ &= y\cos(xy-y) \ . \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial y}(x,y) &= \frac{\partial}{\partial y}\sin(xy-y) \\ &= \cos(xy-y)\frac{\partial}{\partial y}(xy-y) \\ &= \cos(xy-y)\left(\frac{\partial}{\partial y}(xy) - \frac{\partial}{\partial y}y\right) \\ &= (x-1)\cos(xy-y) \ . \end{aligned}$$

PRIMJER 5 : Neka je $f(x,y) = \sqrt{x^2+y^2}$. Ona je kompozicija polinomijalne funkcije (x^2+y^2) i korijene funkcije (funkcija jedne varijable).

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

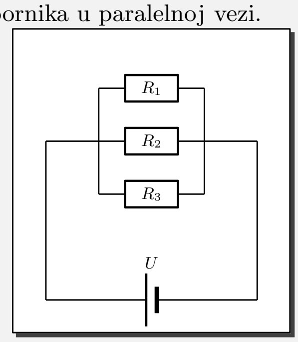
PRIMJER 6 : Posmatrajmo funkciju $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = \ln(x+yz)$.

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y,z) &= \frac{\partial}{\partial x}\ln(x+zy) = \frac{1}{x+zy}\frac{\partial}{\partial x}(x+zy) = \frac{1}{x+yz} \ , \\ \frac{\partial f}{\partial y}(x,y,z) &= \frac{\partial}{\partial y}\ln(x+yz) = \frac{1}{x+yz}\frac{\partial}{\partial y}(x+yz) = \frac{z}{x+yz} \ , \\ \frac{\partial f}{\partial z}(x,y,z) &= \frac{\partial}{\partial z}\ln(x+yz) = \frac{1}{x+yz}\frac{\partial}{\partial z}(x+yz) = \frac{y}{x+yz} \ . \end{aligned}$$

Parcijalni izvodi u konkretnoj tački, naprimjer $A(1,1,2)$ bili bi

$$\frac{\partial f}{\partial x}(1,1,2) = \frac{1}{3} \ , \ \frac{\partial f}{\partial y}(1,1,2) = \frac{2}{3} \ , \ \frac{\partial f}{\partial z}(1,1,2) = \frac{1}{3} \ .$$

PRIMJER 7 : Pravilo kompozicije možemo primjenjivati i u drugim situacijama. Naprimjer, posmatrajmo šemu otpornika u paralelnoj vezi.



Ukupan otpor kola dat je sa

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} \ . \tag{3.3}$$

Dakle, ukupan otpor je funkcija tri varijable, $R = R(r_1, r_2, r_3)$. Ako sada želimo naći parcijalne izvode po r_i $(i = 1, 2, 3)$, onda to možemo uraditi izračunavajući otpor R eksplicitno iz formule (3.3)

$$R = \frac{r_1 r_2 r_3}{r_1 r_2 + r_1 r_3 + r_2 r_3} \ .$$

Međutim, ako lijevu stranu u (3.3) shvatimo kao kompoziciju racionalne funkcije $(\frac{1}{R})$ i naše funkcije R, onda direktno imamo, primjenjujući gornje pravilo,

$$\frac{d\frac{1}{R}}{dR}\frac{\partial R}{\partial r_1} = \frac{\partial \frac{1}{r_1}}{\partial r_1} + \frac{\partial \frac{1}{r_2}}{\partial r_1} + \frac{\partial \frac{1}{r_3}}{\partial r_1} \ ,$$

odakle je

$$-\frac{1}{R^2}\frac{\partial R}{\partial r_1} = -\frac{1}{r_1^2}\ ,$$

to jest

$$\frac{\partial R}{\partial r_1} = \frac{R^2}{r_1^2}\ .$$

Iz ovoga imamo tumačenje: ako se otpor r_1 smanji ukupan otpor se poveća i obrnuto, povećanje otpora r_1 uzrokuje smanjenje ukupnog otpora.

Analogno nalazimo parcijalne izvode po ostalim promjenljivima.

Neka je $z = f(x,y)$ i neka su i x i y funkcije nekog parametra t, $x = x(t)$ i $y = y(t)$. Tada je funkcija $z = f(x(t), y(t))$, ustvari funkcija jedne varijable (t) i pri tome imamo:
Ako su funkcije $x(t)$ i $y(t)$ diferencijabilne u t i ako je funkcija f diferencijabilna u tački $(x(t), y(t))$, tada vrijedi

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\ .$$

$$\begin{tikzcd}
 & f(x,y) \arrow[dl, "\frac{\partial f}{\partial x}"'] \arrow[dr, "\frac{\partial f}{\partial y}"] & \\
x(t) \arrow[dr, "\frac{dx}{dt}"'] & & y(t) \arrow[dl, "\frac{dy}{dt}"] \\
 & t &
\end{tikzcd}$$

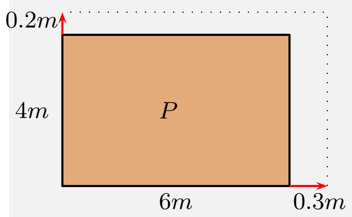
Primjer 8 : Neka je $f(x,y) = \sin x + \cos(xy)$ i neka su $x = t^2$ i $y = t^3$. Tada prema pravilu kompozicije imamo

$$\begin{aligned}
\frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\
&= (\cos x - \sin(xy)y)2t + (-\sin(xy)x)3t^2 \\
&= (\cos t^2 - t^3\sin t^5)2t - 3t^4\sin t^5\ .
\end{aligned}$$

U Primjeru 7 neka su svi pojedinačni otpori zavisni o vremenu t, to jest $r_i = r_i(t)$ $(i = 1,2,3)$. Tada je ukupni otpor takođe ovisan o vremenu t te imamo da je

$$\begin{aligned}
\frac{dR}{dt} &= \frac{\partial R}{\partial r_1}\frac{dr_1}{dt} + \frac{\partial R}{\partial r_2}\frac{dr_2}{dt} + \frac{\partial R}{\partial r_3}\frac{dr_3}{dt} \\
&= \frac{R^2}{r_1^2}\dot{r}_1 + \frac{R^2}{r_2^2}\dot{r}_2 + \frac{R^2}{r_3^2}\dot{r}_3.
\end{aligned}$$

Primjer 9 : Pravougaonik ima dužinu 6 m i širinu 4 m. U svakoj sekundi dužina se poveća za 0.3 m, a širina za 0.2 m. Odrediti promjenu površine pravougaonika u jednoj sekundi.



x - dužina pravougaonika
y - širina pravougaonika
P - površina pravougaonika
t - vrijeme

Dužina i širina pravougaonika su funkcije vremena, $x = x(t)$ i $y = y(t)$. Promjena dužine u jedinici vremena je $\frac{dx}{dt} = 0.3$, a promjena širine u jedinici vremena je $\frac{dy}{dt} = 0.2$.
Površina pravougaonika je $P(x,y) = x{\cdot}y$, a zbog zavisnosti dužine i širine od vremena imamo $P(x(t), y(t)) = x(t)\cdot y(t)$. Izračunajmo zavisnost površine o vremenu.

$$\frac{dP}{dt} = \frac{\partial P}{\partial x}\frac{dx}{dt} + \frac{\partial P}{\partial y}\frac{dy}{dt} = y\frac{dx}{dt} + x\frac{dy}{dt}\ .$$

Stepen promjene površine u prvoj sekundi je

$$\frac{dP}{dt}(6,4) = 4 \cdot 0.3 + 6 \cdot 0.2 = 2.4\ \frac{m^2}{s}\ .$$

Ukoliko su x i y zavisne od dvije varijable, to jest $x = x(t,s)$ i $y = y(t,s)$, tada pravilo kompozicije glasi:
Ako funkcije x i y imaju parcijalne izvode prvog reda u tački (t,s) i ako je funkcija $z = f(x,y)$ diferencijabilna u tački $(x(t,s), y(t,s))$, tada vrijedi

$$\frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

$$\begin{tikzcd}
& f(x,y) \arrow[dl, "\frac{\partial f}{\partial x}"'] \arrow[dr, "\frac{\partial f}{\partial y}"] & \\
x(t,s) \arrow[d, "\frac{\partial x}{\partial t}"'] \arrow[drr, "\frac{\partial x}{\partial s}"] & & y(t,s) \arrow[dll, "\frac{\partial y}{\partial t}"'] \arrow[d, "\frac{\partial y}{\partial s}"] \\
t & & s
\end{tikzcd}$$

Primjer 10 : Zadata je funkcija $z = f(x,y) = x^2 - y^2$ i pri tome je $x = \rho\cos\phi$, $y = \rho\sin\phi$. Odrediti parcijalne izvode funkcije f po promjenljivima ρ i ϕ.

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

$$\begin{aligned}\frac{\partial f}{\partial \phi} &= \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial \phi} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial \phi} = 2x(-\rho\sin\phi) - 2y\rho\cos\phi \\ &= -4\rho^2\cos\phi\sin\phi = -2\rho sin 2\phi\ .\end{aligned}$$

Primjer 11 : Neka je $f(x,y,z)$ funkcija tri varijable. Uvođenjem smjena: $x = \rho\cos\phi\sin\theta$, $y = \rho\sin\phi\sin\theta$ i $z = \rho\cos\theta$, dobijamo novu funkciju

$$F(\rho,\phi,\theta) = f(\rho\cos\phi\sin\theta, \rho\sin\phi\sin\theta, \rho\cos\theta)\ .$$

Sada imamo, koristeći se lančanim pravilom,

$$\begin{aligned}\frac{\partial F}{\partial \rho} &= \frac{\partial f}{\partial x}\cos\phi\sin\theta + \frac{\partial f}{\partial y}\sin\phi\sin\theta + \frac{\partial f}{\partial z}\cos\theta\ , \\ \frac{\partial F}{\partial \phi} &= -\frac{\partial f}{\partial x}\rho\sin\phi\sin\theta + \frac{\partial f}{\partial y}\rho\cos\phi\sin\theta\ , \\ \frac{\partial F}{\partial \theta} &= \frac{\partial f}{\partial x}\rho\cos\phi\cos\theta + \frac{\partial f}{\partial y}\rho\sin\phi\cos\theta - \frac{\partial f}{\partial z}\rho\sin\theta\ .\end{aligned}$$

Primjetimo da vrijedi jednakost (pokazati!)

$$(f'_x)^2 + (f'_y)^2 + (f'_z)^2 = (F'_\rho)^2 + \left(\frac{F'_\phi}{\rho\sin\theta}\right)^2 + \left(\frac{F'_\theta}{\rho}\right)^2\ .$$

U posljednjoj jednakosti korišteni su drugi zapisi za parcijalne izvode, praktičnosti zapisa radi.

3.3 Gradijent

Mnoge fizikalne veličine imaju različite vrijednosti u različitim tačkama prostora. Na primjer, temperatura u nekoj prostoriji nije jednaka u svim tačkama: zimi je visoka kraj izvora toplote, a niska pored otvorenog prozora. Električno polje oko tačkastog naboja veliko je pored naboja i smanjuje se kako se udaljavamo od naboja. Slično, Zemljina gravitacijska sila koja djeluje na neki satelit zavisi od udaljenosti satelita od Zemlje. Brzina toka vode u nekom potoku velika je u uskim

kanalima, a mala tamo gdje je potok širok.
U svim ovim primjerima postoji neko područje prostora koje nam je posebno zanimljivo za problem koji rješavamo; u svakoj tački prostora neka fizikalna veličina ima svoju vrijednost. Izraz polje znači često i područje i vrijednost fizikalne veličine u tom području (npr. elektično polje, gravitacijsko polje). Ako je fizikalna veličina koju promatramo skalar (npr. temperatura), tada govorimo o skalarnom polju. Ako je fizikalna veličina vektor (npr. električno polje, brzina, sila) tada govorimo o vektorskom polju.
Jedna od veličina koja karakteriše termin polja jeste pojam gradijenta.

Definicija 3.3.1. *Neka je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ definisana u okolini $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a}$ i neka postoje $\dfrac{\partial f}{\partial x_k}(\boldsymbol{a})$ za sve $k = 1, 2, ..., n$. Vektor*

$$\left(\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \frac{\partial f}{\partial x_2}(\boldsymbol{a}), ..., \frac{\partial f}{\partial x_n}(\boldsymbol{a})\right) ,$$

nazivamo gradijentni vektor funkcije f u tački $\boldsymbol{a}$.

U gornjoj definiciji posmatramo funkciju f čije su vrijednosti skalari, za koju u primjenama kažemo da je skalarno polje, a definisana veličina bi onda imala naziv *gradijent skalarnog polja*. Dakle, proizvoljnoj skalarnoj funkciji $f : \mathbb{R}^n \to \mathbb{R}$ pridružujemo n-dimenzionalni vektor,

$$f \;\mapsto\; \left(\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \frac{\partial f}{\partial x_2}(\boldsymbol{a}), ..., \frac{\partial f}{\partial x_n}(\boldsymbol{a})\right) .$$

To pridruživanje je očigledno vektorska funkcija, koju nazivamo *gradijent* i označavamo sa ∇ (čitamo: nabla). Ovakve funkcije uobičajeno nazivamo *vektorsko polje*, a sa čime ćemo se susresti u narednim matematičkim izučavanjima.

Vektorska funkcija ∇ (nabla) se u dekartovom pravouglom koordinatnom sistemu 2D i 3D definiše sa

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} \quad , \quad \nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} .$$

U ovakvoj formi uobičajeno se naziva *gradijent*, *diferencijalni operator*, *Hamiltonov operator* ili prosto ∇ (nabla) operator.

Kažemo da je to vektorska funkcija jer on skalarnoj funkciji f dodjeljuje vektorsku veličinu ∇f, po principu vektor puta skalar,

$$\nabla f = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \vec{i}\,\frac{\partial f}{\partial x} + \vec{j}\,\frac{\partial f}{\partial y} + \vec{k}\,\frac{\partial f}{\partial z} .$$

Iz ovoga zaključujemo da je gradijentni vektor funkcije f, to jest ∇f, ništa drugo do množenje vektora (∇) skalarom (f), a kao što ćemo vidjeti nešto kasnije, sa gradijentom (kao vektorom) možemo raditi i druge operacije.

Primjer 12 : Na osnovu Primjera 2, gradijent skalarnog polja $f(x, y) = xy$ je

$$\nabla f(x, y) = \left(\frac{\partial(xy)}{\partial x}, \frac{\partial(xy)}{\partial y}\right) = (y, x) ,$$

odnosno u konkretnoj tački $\boldsymbol{a} = (-2, 7)$ je $\nabla f(-2, 7) = (7, -2) = 7\vec{\boldsymbol{i}} - 2\vec{\boldsymbol{j}}$.

Primjer 13 : Iz Primjera 6 za funkciju $f(x, y, z) = \ln(x + yz)$ imamo

$$\nabla f(1, 1, 2) = \left(\frac{1}{3}, \frac{2}{3}, \frac{1}{3}\right) = \frac{1}{3}\vec{\boldsymbol{i}} + \frac{2}{3}\vec{\boldsymbol{j}} + \frac{1}{3}\vec{\boldsymbol{k}} .$$

PRIMJER 14 : Za funkciju $f(x,y) = 4 - 2x^2 - y^2$ imamo $\frac{\partial f}{\partial x}(x,y) = -4x$, $\frac{\partial f}{\partial y}(x,y) = -2y$, pa je gradijentni vektor u proizvoljnoj tački $\boldsymbol{x} = (x,y)$ dat sa

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f}{\partial x}(\boldsymbol{x}), \frac{\partial f}{\partial y}(\boldsymbol{x}) \right) = (-4x, 2y) \ .$$

Konkretno u tački $\boldsymbol{x} = (1,1)$ je $\nabla f(1,1) = (-4,2)$.

Nije teško pokazati da za gradijent skalarnog polja vrijede sljedeća pravila:

1. $\nabla(kf) = k\nabla f$, $(k = const.\ )$.
2. $\nabla(f \pm g) = \nabla f \pm \nabla g$.
3. $\nabla(fg) = g\nabla f + f\nabla g$.
4. $\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}$.

Gradijent skalarnog polja iznimno je važan u fizici gdje izražava vezu između polja i potencijala (gravitacijska polja), odnosno sile i potencijalne energije (električna polja). Ako se neko polje E može u cijelosti opisati konkretnom funkcijom $f(\boldsymbol{x})$, tako da je $E = -\nabla f(\boldsymbol{x})$, tada skalarnu funkciju f nazivamo njegovim potencijalom, što simbolički možemo zapisati sa

$$\text{polje} = -\nabla(\text{potencijal}) \ .$$

Specijalno, ako se neka sila F može napisati kao negativni gradijent neke funkcije V, tada skalarnu funkciju V nazivamo potencijalnom energijom.

PRIMJER 15 : Neka je $\nabla f = (3x^2y^2 + x^3, 2x^3y + \cos y)$. Odrediti skalarnu funkciju $f(x,y)$!

Rješenje: Iz definicije gradijenta imamo da je

$$\frac{\partial f}{\partial x} = 3x^2y^2 + x^3 \tag{3.4}$$

$$\frac{\partial f}{\partial y} = 2x^3y + \cos y \ . \tag{3.5}$$

Sada integralimo bilo koju od ovih jednakosti, recimo jednakost (3.4) integralimo po x. Vodeći računa da smo praveći parcijalni izvod po x ostale varijable posmatrali kao konstante (u našem slučaju samo varijabla y), onda integraleći $\frac{\partial f}{\partial x}$ po x, "konstanta integracije" će ovisiti o svim ostalim varijablama (u ovom slučaju o varijabli y). Dakle,

$$\int \frac{\partial f}{\partial x}(x,y)dx = x^3y^2 + \frac{1}{4}x^4 + \phi(y) = f(x,y) \ . \tag{3.6}$$

Nalaženjem parcijalnog izvoda po y iz 3.6 imamo

$$\frac{\partial f}{\partial y} = 2x^3y + \frac{d\phi}{dy} \ . \tag{3.7}$$

Izjednačavajući (3.5) sa (3.7), zaključujemo da mora biti $\frac{d\phi}{dy} = \cos y$, a ovo nam omogućava da odredimo nepoznatu funkciju ϕ, to jest $\phi(y) = \sin y + C$. Sa ovim imamo konačno da je

$$f(x,y) = x^3y^2 + \frac{1}{4}x^4 + \sin y + C \ .$$

3.4 Diferencijabilnost funkcija više promjenljivih

Neka je $f : \mathbb{R}^n \to \mathbb{R}$ definisana u nekoj okolini $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a} = (a_1, a_2, ..., a_n) \in \mathbb{R}^n$. Samo postojanje parcijalnih izvoda ne obezbjeđuje neke bitne osobine posmatrane funkcije, što vidimo iz sljedećeg primjera.

PRIMJER 16 : Posmatrajmo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, zadatu sa

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & ; \quad (x,y) \neq (0,0) \\ 0 & ; \quad (x,y) = (0,0) \end{cases}$$

Nije teško pokazati da je f prekidna funkcija u tački $(0,0)$. S druge strane ona ima oba parcijalna izvoda u tački $(0,0)$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h} = 0 \ ,$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = \lim_{k\to 0} \frac{0}{k} = 0 \ .$$

Dakle, parcijalni izvodi postoje u tački $(0,0)$, a funkcija ima prekid u toj tački.

Jasno je dakle, da za razliku od funkcija jedne promjenljive gdje je postojanje izvoda značilo neprekidnost funkcije, postojanje parcijalnih izvoda kod funkcije više varijabli ne može garantovati određene "lijepe" osobine funkcije, nego moramo posmatrati neka svojstva koja uzimaju u obzir ponašanje funkcije u čitavoj okolini posmatrane tačke.

Definicija 3.4.1. *Razlika*

$$\triangle f(\boldsymbol{a}) = f(\boldsymbol{x}) - f(\boldsymbol{a}) \ ; \ (\boldsymbol{x} \in U_{\boldsymbol{a}}) \ ,$$

naziva se totalni priraštaj funkcije f u tački $\boldsymbol{a}$.

Totalni priraštaj funkcije u tački $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ izražavamo preko priraštaja nezavisnih promjenljivih, to jest

$$\triangle f(\boldsymbol{x}) = f(\boldsymbol{x} + \triangle \boldsymbol{x}) - f(\boldsymbol{x}) = f(x_1 + \triangle x_1, ..., x_n + \triangle x_n) - f(x_1, ..., x_n) \ .$$

Za razliku od parcijalnog priraštaja gdje jednu varijablu mijenjamo, a sve druge "držimo" fiksnim, kod totalnog priraštaja sve varijable istovremeno "doživljavaju" neku promjenu.

Definicija 3.4.2. *Za funkciju $f(\boldsymbol{x}) = f(x_1, ..., x_n)$ definisanu u okolini tačke $\boldsymbol{a} \in \mathbb{R}^n$, kažemo da je diferencijabilna u toj tački ako vrijedi*

$$\triangle f(\boldsymbol{a}) = L(\boldsymbol{x}) + \omega(\boldsymbol{x}) d(\boldsymbol{x}, \boldsymbol{a}) \ , \tag{3.8}$$

gdje je

$$L(\boldsymbol{x}) = \sum_{k=1}^{n} p_k (x_k - a_k) = \sum_{k=1}^{n} p_k \triangle x_k \ , \tag{3.9}$$

linearna funkcija priraštaja nezavisnih promjenljivih, p_k $(k = 1, 2, ..., n)$ su realni koeficijenti, $\omega(\boldsymbol{x})$ neprekidna funkcija u tački $\boldsymbol{a}$ takva da je

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \omega(\boldsymbol{x}) = \omega(\boldsymbol{a}) = 0 \ , \ a \ \ d(\boldsymbol{x}, \boldsymbol{a}) = \left(\sum_{k=1}^{n} (x_k - a_k)^2 \right)^{\frac{1}{2}} \ ,$$

rastojanje tačke $\boldsymbol{x}$ od tačke $\boldsymbol{a}$.

Definicija 3.4.3. *Linearnu funkciju $L(\boldsymbol{x})$ iz (3.9) nazivamo totalni diferencijal funkcije $f(\boldsymbol{x})$ u tački $\boldsymbol{a}$ i označavamo ga sa*

$$L(\boldsymbol{x}) = df(\boldsymbol{x}) = \sum_{k=1}^{n} p_k \triangle x_k = \sum_{k=1}^{n} p_k dx_k \ .$$

Teorem 3.4.1. ***(Potrebni uslovi diferencijabilnosti)***
Neka je funkcija $f(\boldsymbol{x})$ diferencijabilna u tački $\boldsymbol{a}$. Tada vrijedi:

1. *Postoje parcijalni izvodi po svakoj promjenljivoj funkcije f u tački $\boldsymbol{a}$.*
2. *Koeficijenti p_k $(k = 1, 2, ..., n)$ u izrazu za totalni diferencijal su upravo parcijalni izvodi funkcije, to jest*

$$p_k = \frac{\partial f}{\partial x_k}(\boldsymbol{a}) \ ; \ k = 1, 2, ..., n \ .$$

Dokaz : Ako je funkcija f diferencijabilna u tački $\boldsymbol{a}$, tada po Definiciji 3.4.2 vrijedi

$$\triangle f(\boldsymbol{a}) = f(x_1, x_2, ..., x_n) - f(a_1, a_2, ..., a_n) = \sum_{k=1}^{n} p_k \triangle x_k + \omega(\boldsymbol{x}) d(\boldsymbol{x}, \boldsymbol{a}) \ .$$

Ako fiksiramo $n-1$ promjenljivih

$$x_1 = a_1, ..., x_{k-1} = a_{k-1}, x_{k+1} = a_{k+1}, ..., x_n = a_n \ ,$$

imamo

$$\triangle f(\boldsymbol{a}) = f(a_1, ..., a_{k-1}, x_k, a_{k+1}, ..., a_n) - f(a_1, ..., a_k, ..., a_n) = p_k \triangle x_k + \omega(\boldsymbol{x}) |x_k - a_k| \ ,$$

odakle je

$$\lim_{x_k \to a_k} \frac{\triangle f(\boldsymbol{a})}{\triangle x_k} = p_k + sgn(x_k - a_k) \lim_{x_k \to a_k} \omega(a_1, ..., a_{k-1}, x_k, a_{k+1}, ..., a_n) \ .$$

Odavde vidimo da za proizvoljno $k \in \{1, 2, ..., n\}$ vrijedi

$$p_k = \frac{\partial f}{\partial x_k}(\boldsymbol{a}) \ ,$$

iz čega vidimo da parcijalni izvodi postoje i da su oni upravo koeficijenti p_k $(k = 1, 2, ..., n)$. ■

Na osnovu Teorema 3.4.1 vidimo da totalni diferencijal diferencijabilne funkcije $f(\boldsymbol{x})$ ima oblik

$$df(\boldsymbol{x}) = \frac{\partial f}{\partial x_1}(\boldsymbol{x}) dx_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{x}) dx_2 + ... + \frac{\partial f}{\partial x_n}(\boldsymbol{x}) dx_n \ . \tag{3.10}$$

Desnu stranu ove jednakosti možemo izraziti kao skalarni produkt dva vektora,

$$df(\boldsymbol{x}) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}), \frac{\partial f}{\partial x_2}(\boldsymbol{x}), ..., \frac{\partial f}{\partial x_n}(\boldsymbol{x}) \right) \cdot (dx_1, dx_2, ..., dx_n) = \nabla f(\boldsymbol{x}) \cdot d\boldsymbol{x} \ ,$$

gdje je $d\boldsymbol{x} = (dx_1, dx_2, ..., dx_n)$, vektor priraštaja nezavisnih varijabli. Ovo razmatranje iskazujemo u formi tvrđenja.

Teorem 3.4.2 (Veza diferencijala i gradijenta).
Neka je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u tački $\boldsymbol{x} \in \mathbb{R}^n$. Tada vrijedi

$$df(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) \cdot d\boldsymbol{x} \ . \tag{3.11}$$

Gornja veza nam već pokazuje važnost gradijenta funkcije. Naime, diferencijal funkcije se može prikazati kao skalarni produkt gradijenta funkcije i vektora priraštaja argumenata. U analogiji sa funkcijom jedne varijable $(df(x) = f'(x)dx)$ vidimo da ulogu izvoda funkcije preuzima gradijent.

PRIMJER 17 : Za funkciju $f(x,y) = 4 - 2x^2 - y^2$, totalni diferencijal u proizvoljnoj tački $\boldsymbol{x} = (x,y)$ računamo tako što prvo odredimo parcijalne izvode

$$\frac{\partial f}{\partial x} = -4x \quad , \quad \frac{\partial f}{\partial y} = -2y \ ,$$

a zatim iskoristimo (3.10)

$$df(\boldsymbol{x}) = \frac{\partial f}{\partial x}(\boldsymbol{x})dx + \frac{\partial f}{\partial y}(\boldsymbol{x})dy = -4xdx - 2ydy \ .$$

U konkretnoj tački $\boldsymbol{a} = (-1,2)$, totalni diferencijal glasi $df(\boldsymbol{a}) = 4dx - 4dy$.

S obzirom na uvedeno u gornjim definicijama, ako je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u tački $\boldsymbol{a} \in D_f$, priraštaj funkcije u tački $\boldsymbol{a}$ zbog (3.8) je oblika

$$\triangle f(\boldsymbol{a}) = f(\boldsymbol{x}) - f(\boldsymbol{a}) = df(\boldsymbol{x}) + \omega(\boldsymbol{x})d(\boldsymbol{x},\boldsymbol{a}) \ .$$

Ako se tačka $\boldsymbol{x}$ "približava" sve više tački $\boldsymbol{a}$, zbog $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \omega(\boldsymbol{x}) = 0$, vidimo da je priraštaj funkcije $\triangle f$ sve bolje aproksimiran diferencijalom funkcije df, to jest u blizini tačke $\boldsymbol{a}$ vrijedi $\triangle f(\boldsymbol{a}) \approx df(\boldsymbol{a})$, a to onda znači da se vrijednost funkcije u tački koja je blizu tačke $\boldsymbol{a}$ može aproksimativno izračunati sa

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + df(\boldsymbol{a}) \ ,$$

što nazivamo lokalnom aproksimacijom funkcije u tački.

PRIMJER 18 : Koristeći lokalnu aproksimaciju izračunati vrijednost funkcije $f(x,y) = \sqrt{x^2+y^2}$ u tački $\boldsymbol{x} = (3.04, 3.98)$.

Za zadatu funkciju njeni parcijalni izvodi su

$$\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} \quad , \quad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} \ .$$

Iz činjenice da je $\triangle f(\boldsymbol{x}_0) \approx df(\boldsymbol{x}_0)$ i $\triangle f(\boldsymbol{x}_0) = f(\boldsymbol{x}) - f(\boldsymbol{x}_0)$ za tačku $\boldsymbol{x}$ koja je "blizu" tački $\boldsymbol{x}_0$ imamo

$$\sqrt{x^2+y^2} \approx \sqrt{x_0^2+y_0^2} + \frac{x_0}{\sqrt{x_0^2+y_0^2}}(x-x_0) + \frac{y_0}{\sqrt{x_0^2+y_0^2}}(y-y_0) \ .$$

Uzmimo da je tačka $\boldsymbol{x}_0 = (3,4)$ i da je $\boldsymbol{x} = (3.04, 3.98)$. Tada imamo

$$\sqrt{(3.04)^2 + (3.98)^2} \approx 5 + \frac{3}{5}\cdot 0.04 + \frac{4}{5}\cdot(-0.02) = 5.008 \ .$$

Teorem 3.4.3. *Ako je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u tački $\boldsymbol{a} \in \mathbb{R}^n$, onda je i neprekidna u toj tački.*

Dokaz : Iz diferencijabilnosti funkcije imamo

$$\triangle f(\boldsymbol{a}) = f(\boldsymbol{x}) - f(\boldsymbol{a}) = L(\boldsymbol{x}) + \omega(\boldsymbol{x})d(\boldsymbol{x},\boldsymbol{a}) \ ,$$

a odavde onda imamo

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}}(f(\boldsymbol{x}) - f(\boldsymbol{a})) = \lim_{\boldsymbol{x}\to\boldsymbol{a}} L(\boldsymbol{x}) + \lim_{\boldsymbol{x}\to\boldsymbol{a}} \omega(\boldsymbol{x})d(\boldsymbol{x},\boldsymbol{a}) = 0$$

(jer je $L(\boldsymbol{a}) = 0$). Ovo ne znači ništa drugo do

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a}) \ ,$$

što u stvari znači neprekidnost funkcije f u tački $\boldsymbol{a}$. ■

Da iz neprekidnost ne slijedi diferencijabilnost, to jest da obrat u gornjem tvrđenju ne vrijedi, vidimo iz sljedećeg primjera.

Primjer 19 : Posmatrajmo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, zadatu sa

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & ; \quad (x,y) \neq (0,0) \\ 0 & ; \quad (x,y) = (0,0) \end{cases}$$

Data funkcija je neprekidna u tački $(0,0)$ (što je ostavljeno čitaocu za vježbu) i ima parcijalne izvode $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. Međutim, f nije diferencijabilna u tački $(0,0)$. Zaista, ako bi bila diferencijabilna imali bi smo

$$\Delta f(0,0) = f(\Delta x, \Delta y) - f(0,0) = \frac{\partial f}{\partial x}(0,0)\Delta x + \frac{\partial f}{\partial y}(0,0)\Delta y + \omega(\Delta x, \Delta y)d(X,O) \ ,$$

odnosno, odavde je zbog $d(X,O) = \sqrt{\Delta x^2 + \Delta y^2}$,

$$\omega(\Delta x, \Delta y) = \frac{\Delta x \Delta y^2}{(\Delta x^2 + \Delta y^2)^{\frac{3}{2}}} \ .$$

Zbog osobine funkcije ω, moralo bi biti $\lim\limits_{X \to O} \omega(X) = 0$, to jest

$$\lim_{\Delta x \to 0, \Delta y \to 0} \frac{\Delta x \Delta y^2}{(\Delta x^2 + \Delta y^2)^{\frac{3}{2}}} = 0 \ ,$$

što nije tačno jer za $\Delta x = \Delta y > 0$ je

$$\frac{\Delta x \Delta y^2}{(\Delta x^2 + \Delta y^2)^{\frac{3}{2}}} = \frac{1}{2\sqrt{2}} \not\to 0 \quad , \quad \Delta x, \Delta y \to 0 \ .$$

Uslov diferencijabilnosti u gornjem teoremu možemo zamijeniti nešto slabijim uslovima. Naime vrjedi,

Teorem 3.4.4. *Ako funkcija $f(\boldsymbol{x})$ u nekoj oblasti D ima ograničene parcijalne izvode po svakoj promjenljivoj, tada je ona neprekidna u toj oblasti.*

Šta više, sa još bližim informacijama o parcijalnim izvodima možemo imati još preciznije informacije o funkciji. Tako specijalno vrijedi

Teorem 3.4.5. *Ako funkcija $f(\boldsymbol{x})$ u oblasti D ima parcijalne izvode po svakoj promjenljivoj jednake nuli, onda je funkcija u toj oblasti konstanta.*

Ovo je analogon činjenici za funkciju jedne varijable, da ako je $f'(x) = 0$ za $x \in D$, da je tada f konstantna na D.

Sljedeći teorem je analogon Lagrangeovom teoremu za funkcije jedne promjenljive [1].

Teorem 3.4.6. ***(Lagrangeov teorem)***
Ako funkcija $f(\boldsymbol{x})$ u okolini $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a}$ ima konačne ili beskonačne parcijalne izvode po svakoj promjenljivoj, tada za proizvoljno $\boldsymbol{x} \in U_{\boldsymbol{a}}$ postoje tačke $\boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \ldots, \ \boldsymbol{x}_n \in U_{\boldsymbol{a}}$, takve da je

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\boldsymbol{x}_k) dx_k \ .$$

Iskažimo sada i dovoljne uslove diferencijabilnosti.

Teorem 3.4.7. ***(Dovoljni uslovi diferencijabilnosti)***

[1] Za diferencijabilnu funkciju na intervalu (a,b), za proizvoljan $[x,y] \subset (a,b)$, postoji $c \in (x,y)$, tako da je $f(y) - f(x) = f'(c)(y-x)$

Ako funkcija $f : \mathbb{R}^n \to \mathbb{R}$ ima u nekoj okolini tačke $\boldsymbol{a} \in \mathbb{R}^n$ parcijalne izvode po svakoj promjenljivoj i ako su ti parcijalni izvodi neprekidni u toj tački, tada je funkcija f diferencijabilna u tački $\boldsymbol{a}$.

Dokaz : Dokaz ćemo jednostavnosti zapisa radi, dati za funkciju dvije promjenljive i on se jednostavno može prenijeti na funkcije sa n promjenljivih.
Na osnovu Lagrangeovog teorema, priraštaj funkcije $f(x,y)$ ima oblik

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{\partial f}{\partial x}(\boldsymbol{x}_1)(x-a) + \frac{\partial f}{\partial y}(\boldsymbol{x}_2)(y-b) \ , \tag{3.12}$$

gdje su tačke $\boldsymbol{x} = (x,y), \boldsymbol{x}_1 = (\xi_1, b)$ i $\boldsymbol{x}_2 = (a, \xi_2)$ iz okoline $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a} = (a,b)$. Zbog pretpostavljene neprekidnosti parcijalnih izvoda, to jest neprekidnost funkcija $\frac{\partial f}{\partial x}(x,y)$ i $\frac{\partial f}{\partial y}(x,y)$ u tački $\boldsymbol{a}$, iz (3.12) imamo da vrijedi

$$\lim_{\substack{x \to a \\ y \to b}} f(x,y) = f(a,b) \ ,$$

pa važi

$$\frac{\partial f}{\partial x}(\boldsymbol{x}_1) = \frac{\partial f}{\partial x}(\boldsymbol{a}) + \varepsilon_1(\boldsymbol{x}) \ , \ \frac{\partial f}{\partial y}(\boldsymbol{x}_2) = \frac{\partial f}{\partial y}(\boldsymbol{a}) + \varepsilon_2(\boldsymbol{x}) \ ,$$

gdje $\varepsilon_1 \to 0$ i $\varepsilon_2 \to 0$, kada $\boldsymbol{x} \to \boldsymbol{a}$.
Ako posljednje dvije jednakosti pomnožimo sa $x-a$ i $y-b$ respektivno, i tako dobijene jednakosti saberemo, dobijamo

$$\begin{aligned} f(x,y) - f(a,b) &= \frac{\partial f}{\partial x}(\boldsymbol{x}_1)(x-a) + \frac{\partial f}{\partial y}(\boldsymbol{x}_2)(y-b) \\ &= \frac{\partial f}{\partial x}(\boldsymbol{a})(x-a) + \frac{\partial f}{\partial y}(\boldsymbol{a})(y-b) + \varepsilon_1(\boldsymbol{x})(x-a) + \varepsilon_2(\boldsymbol{x})(y-b) \ , \end{aligned}$$

odnosno

$$\triangle f(\boldsymbol{a}) = df(\boldsymbol{a}) + \varepsilon_1(\boldsymbol{x})(x-a) + \varepsilon_2(\boldsymbol{x})(y-b) \ ,$$

iz čega se, na osnovu Definicije 3.4.2, vidi da je funkcija f diferencijabilna u tački $\boldsymbol{a}$. ■

Za funkciju koja u nekoj tački ima neprekidne parcijalne izvode, reći ćemo da je *neprekidno diferencijabilna* u toj tački. Ako funkcija f zadovoljava taj uslov u svim tačkama nekog skupa D, onda kažemo da je f neprekidno diferencijabilna na D. Skup svih neprekidno diferencijabilnih funkcija na skupu D označavamo sa $C^1(D)$. Dakle, informacija $f \in C^1(D)$ znači da je f neprekidno diferencijabilna na D, odnosno da u svakoj tački skupa D funkcija f ima neprekidne parcijalne izvode. Ovo možemo tumačiti još i sa time da postoji gradijentni vektor funkcije f i da je on neprekidna vektorska funkcija u svakoj tački skupa D.

3.5 Veza izvoda u pravcu i gradijenta

Posmatrajmo sada $f : \mathbb{R}^2 \to \mathbb{R}$ i neka je f neprekidno diferencijabilna funkcija u nekoj okolini $U_{\boldsymbol{a}}$ tačke $\boldsymbol{a} = (a_1, a_2) \in \mathbb{R}^2$. Neka je $\boldsymbol{u} = (u_1, u_2)$ proizvoljan jedinični vektor i nađimo izvod u pravcu $D_{\boldsymbol{u}} f(\boldsymbol{a})$. Na osnovu definicije izvoda u pravcu imamo

$$\begin{aligned} D_{\boldsymbol{u}} f(\boldsymbol{a}) &= \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{u}) - f(\boldsymbol{a})}{h} \\ &= \lim_{h \to 0} \frac{f(a_1 + hu_1, a_2 + hu_2) - f(a_1, a_2)}{h} \\ &= \lim_{h \to 0} \frac{f(a_1 + hu_1, a_2 + hu_2) - f(a_1 + hu_1, a_2) + f(a_1 + hu_1, a_2) - f(a_1, a_2)}{h} \\ &= \lim_{h \to 0} \left(\frac{f(a_1 + hu_1, a_2 + hu_2) - f(a_1 + hu_1, a_2)}{h} + \frac{f(a_1 + hu_1, a_2) - f(a_1, a_2)}{h} \right) \ . \end{aligned}$$

Za fiksno $h \neq 0$, definišimo sada funkciju $\phi : \mathbb{R} \to \mathbb{R}$, sa

$$\phi(t) = f(a_1 + hu_1, a_2 + t) .$$

Pretpostavka o diferencijabilnosti funkcije f daje nam diferencijabilnost funkcije ϕ, te imamo

$$\begin{aligned} \phi'(t) &= \lim_{s\to 0} \frac{\phi(t+s)-\phi(t)}{s} \\ &= \lim_{s\to 0} \frac{f(a_1+hu_1, a_2+t+s) - f(a_1+hu_1, a_2+t)}{s} \\ &= \frac{\partial f}{\partial y}(a_1+hu_1, a_2+t) . \end{aligned}$$

Neka je sada $\alpha : \mathbb{R} \to \mathbb{R}$, definisana sa

$$\alpha(t) = \phi(u_2 t) = f(a_1 + hu_1, a_2 + tu_2) . \tag{3.13}$$

α je diferencijabilna i na osnovu izvoda složene funkcije imamo

$$\alpha'(t) = u_2\phi'(tu_2) = u_2\frac{\partial f}{\partial y}(a_1 + hu_1, a_2 + tu_2) . \tag{3.14}$$

Na osnovu teorema o srednjoj vrijednosti funkcije jedne varijable, postoji $\xi \in (0, h)$, takav da vrijedi

$$\frac{\alpha(h) - \alpha(0)}{h} = \alpha'(\xi) .$$

Stavljajući sada (3.13) i (3.14) u gornju jednakost dobijamo,

$$\frac{f(a_1 + hu_1, a_2 + hu_2) - f(a_1 + hu_1, a_2)}{h} = u_2\frac{\partial f}{\partial y}(a_1 + hu_1, a_2 + \xi u_2) . \tag{3.15}$$

Na isti način, posmatrajući funkciju $\beta : \mathbb{R} \to \mathbb{R}$, zadatu sa $\beta(t) = f(a_1 + tu_1, a_2)$, imamo da vrijedi

$$\beta'(t) = u_1\frac{\partial f}{\partial x}(a_1 + tu_1, a_2) ,$$

i opet koristeći teorem o srdnjoj vrijednosti, zaključili bi da postoji $\eta \in (0, h)$, tako da je

$$\frac{f(a_1 + hu_1, a_2) - f(a_1, a_2)}{h} = \frac{\beta(h) - \beta(0)}{h} = \beta'(\eta) = u_1\frac{\partial f}{\partial x}(a_1 + \eta u_1, a_2) . \tag{3.16}$$

Stavljajući sada (3.15) i (3.16) u izraz za $D_{\boldsymbol{u}}f(\boldsymbol{a})$, imamo

$$D_{\boldsymbol{u}}f(\boldsymbol{a}) = \lim_{h\to 0}\left(u_2\frac{\partial f}{\partial y}(a_1 + hu_1, a_2 + \xi u_2) + u_1\frac{\partial f}{\partial x}(a_1 + \eta u_1, a_2)\right) . \tag{3.17}$$

Kako su $\xi, \eta \in (0, h)$, kada $h \to 0$, to onda i $\xi, \eta \to 0$. Iskoristivši definitivno i pretpostavku o neprekidnosti parcijalnih izvoda $\frac{\partial f}{\partial x}$ i $\frac{\partial f}{\partial y}$, računajući limes u (3.17), dobijamo

$$D_{\boldsymbol{u}}f(\boldsymbol{a}) = u_1\frac{\partial f}{\partial x}(a_1, a_2) + u_2\frac{\partial f}{\partial y}(a_1, a_2) . \tag{3.18}$$

Generalizaciju tvrdnje iskazane u (3.18) iskazujemo za funkciju $f : \mathbb{R}^n \to \mathbb{R}$ sljedećim teoremom.

Teorem 3.5.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ neprekidno diferencijabilna u nekoj okolini tačke $\boldsymbol{a} \in \mathbb{R}^n$. Tada za proizvoljan jedinični vektor $\boldsymbol{u}$, postoji $D_{\boldsymbol{u}}f(\boldsymbol{a})$ i vrijedi*

$$D_{\boldsymbol{u}}f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{u} . \tag{3.19}$$

Primjetimo još jednom da biranjem vektora $\boldsymbol{u}$ kao baznog vektora, dobijamo parcijalne izvode. Naprimjer, neka je $f : \mathbb{R}^3 \to \mathbb{R}$. Tada je $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$. Ako izaberemo da je $\boldsymbol{u} = \boldsymbol{e}_1 = \boldsymbol{i} = (1,0,0)$, tada je

$$D_{\boldsymbol{i}} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{i} = \left(\frac{\partial f(\boldsymbol{a})}{\partial x}, \frac{\partial f(\boldsymbol{a})}{\partial y}, \frac{\partial f(\boldsymbol{a})}{\partial z}\right) \cdot (1,0,0) = \frac{\partial f(\boldsymbol{a})}{\partial x} \; .$$

Analogno će biti $D_{\boldsymbol{j}} f(\boldsymbol{a}) = \frac{\partial f(\boldsymbol{a})}{\partial y}$ i $D_{\boldsymbol{k}} f(\boldsymbol{a}) = \frac{\partial f(\boldsymbol{a})}{\partial z}$.

PRIMJER 20 : Neka je $f : \mathbb{R}^2 \to \mathbb{R}$, zadata sa $f(x,y) = 4 - 2x^2 - y^2$.

Sada je $\nabla f(x,y) = (-4x, -2y)$, pa za vektor $\boldsymbol{u} = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ imamo

$$D_{\boldsymbol{u}} f(1,1) = \nabla f(1,1) \cdot \boldsymbol{u} = (-4,-2) \cdot \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = 3\sqrt{2} \; ,$$

što možemo potvrditi sa ranije urađenim primjerom (Primjer 1). Sada jednostavno računamo i

$$D_{-\boldsymbol{u}} f(1,1) = \nabla f(1,1) \cdot (-\boldsymbol{u}) = (-4,-2) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -3\sqrt{2} \; .$$

Za vektor $\boldsymbol{w} = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ imamo

$$D_{\boldsymbol{w}} f(1,1) = (-4,-2) \cdot \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right) = 0 \; .$$

3.6 Veza gradijenta i promjene vrijednosti funkcije

Neka je sada $\boldsymbol{u}$ proizvoljan jedinični vektor ($||u|| = 1$), $f : \mathbb{R}^n \to \mathbb{R}$ i neka je $\boldsymbol{a} \in \mathbb{R}^n$. Koristeći Cauchy-Schwarzovu nejednakost[2] i vezu (3.19) možemo zaključiti sljedeće,

$$|D_{\boldsymbol{u}} f(\boldsymbol{a})| = |\nabla f(\boldsymbol{a}) \cdot \boldsymbol{u}| \leq ||\nabla f(\boldsymbol{a})|| \; ||\boldsymbol{u}|| = ||\nabla f(\boldsymbol{a})|| \; . \tag{3.20}$$

Ovo nam govori, bukvalno čitajući, da je apsolutna vrijednost izvoda funkcije u pravcu vektora $\boldsymbol{u}$, u tački $\boldsymbol{a}$, manja ili jednaka od intenziteta gradijentnog vektora funkcije u toj tački. Kako nam izvod u pravcu pokazuje promjenu vrijednosti funkcije, ovo znači da veličina promjene rasta funkcije u nekoj tački, u proizvoljnom pravcu nikad ne prelazi dužinu vektora gradijenta u toj tački. Šta više, znajući osobine Cauchy-Schwarzove nejednakosti, jednakost u (3.20) će se postići upravo u slučaju kada je vektor $\boldsymbol{u}$ kolinearan vektoru $\nabla f(\boldsymbol{a})$. Zaista, ako je $\nabla f(\boldsymbol{a}) \neq 0$, onda za vektor $\boldsymbol{u} = \frac{\nabla f(\boldsymbol{a})}{||\nabla f(\boldsymbol{a})||}$ imamo,

$$D_{\boldsymbol{u}} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{u} = \frac{\nabla f(\boldsymbol{a}) \cdot \nabla f(\boldsymbol{a})}{||\nabla f(\boldsymbol{a})||} = \frac{||\nabla f(\boldsymbol{a})||^2}{||\nabla f(\boldsymbol{a})||} = ||\nabla f(\boldsymbol{a})|| \; .$$

Šta više, vrijedi $D_{-\boldsymbol{u}} f(\boldsymbol{a}) = -||\nabla f(\boldsymbol{a})||$. Gornju tvrdnju iskazujemo narednim teoremom.

Teorem 3.6.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ neprekidno diferencijabilna funkcija u nekoj otvorenoj kugli koja sadrži tačku $\boldsymbol{a}$. Tada $D_{\boldsymbol{u}} f(\boldsymbol{a})$ ima maksimalnu vrijednost $||\nabla f(\boldsymbol{a})||$ kada je vektor $\boldsymbol{u}$ ort vektor vektora $\nabla f(\boldsymbol{a})$, a minimalnu vrijednost, $-||\nabla f(\boldsymbol{a})||$, kada je vektor $\boldsymbol{u}$ ort vektora $-\nabla f(\boldsymbol{a})$.*

Dakle, gradijentni vektor pokazuje pravac i smjer maksimalne promjene rasta funkcije, odnosno negativni gradijentni vektor pokazuje pravac i smjer maksimalne promjene opadanja funkcije. Šta više, intenzitet gradijentnog vektora nam govori o veličini rasta u smjeru maksimalnog rasta, odnosno njegova negativna vrijednos govori o veličini opadanja funkcije u smjeru maksimalnog opadanja.

[2] Za proizvoljne vektore $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ vrijedi

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq ||\boldsymbol{x}|| \; ||\boldsymbol{y}||$$

Primjer 21 : Posmatrajmo ponovo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$ zadatu sa

$$f(x, y) = 4 - 2x^2 - y^2 ,$$

za koju je

$$\nabla f(x, y) = (-4x, -2y) .$$

Ukoliko se nalazimo u tački domena $\boldsymbol{a} = (1, 1)$ (na grafu u tački $(1, 1, 1)$) i želimo krenuti u smjeru najvećeg rasta funkcije f, na osnovu gornjeg teoreme, trebamo krenuti u pravcu vektora

$$\boldsymbol{u} = \frac{\nabla f(1,1)}{||\nabla f(1,1)||} = \left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right) .$$

Ako tražimo pravac najbržeg opadanja funkcije, onda će to biti u pravcu vektora

$$-\boldsymbol{u} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) .$$

Šta više, veličina promjene rasta u pravcu vektora $\boldsymbol{u}$ je

$$D_{\boldsymbol{u}} f(1, 1) = ||\nabla f(1, 1)|| = \sqrt{20} ,$$

a veličina opadanja u pravcu $-\boldsymbol{u}$ je

$$D_{-\boldsymbol{u}} f(1, 1) = -||\nabla f(1, 1)|| = -\sqrt{20} .$$

Razmotrimo još jedan važan fakat vezan za gradijent funkcije. Pokazaćemo ga za funkciju dvije varijable, a isto rezonovanje imamo za proizvoljnu funkciju $f : \mathbb{R}^n \to \mathbb{R}$.

Dakle, neka je data funkcija $z = f(x, y)$ čiji je graf površ G u prostoru $\mathbb{R}^3$. Posmatrajmo poizvoljnu tačku $\boldsymbol{p} = (x_0, y_0, f(x_0, y_0))$ na grafu G i neka je l nivo linija na grafu G koja prolazi kroz tačku $\boldsymbol{p}$. Kako je za tu liniju zadovoljeno $f(x, y) = c$, za neko fiksno $c \in \mathbb{R}$, i kako je ona jednodimenzionalan objekat u prostoru, možemo je parametrizovati, to jest svaku tačku linije l možemo posmatrati kao vektorsku funkciju $(x(t), y(t), z(t))$ za $t \in [\alpha, \beta]$. Kretanje tačke po grafu funkcije $f(\boldsymbol{x})$ je uzrokovano kretanjem tačke $\boldsymbol{x}$ u domenu funkcije. Zato posmatrajmo projekciju nivo linije l u xOy ravan, to jest posmatrajmo njenu odgovarajuću konturnu liniju l'. Pri tome je svaka tačka na l' prikazana parametarski sa $\boldsymbol{r}(t) = (x(t), y(t))$, $t \in [\alpha, \beta]$ (kao projekcija tačke sa grafa to bi bio vektor $(x(t), y(t), 0)$).

Neka je t_0 ona vrijednost parametra koja odgovara tački $\boldsymbol{p}$, a samim tim je $\boldsymbol{x} = \boldsymbol{r}(t_0)$. Kako je nivo linija l na površi G, mora biti zadovoljena jednačina

$$f(\boldsymbol{r}(t)) = f(x(t), y(t)) = c \ , \ \text{za svako } t \in [\alpha, \beta] \ .$$

Diferenciranjem ove jednakosti po t, primjenom pravila kompozicije, imamo

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = 0 \ . \tag{3.21}$$

Nije teško vidjeti da se jednakost (3.21) može zapisati u vektorskoj notaciji,

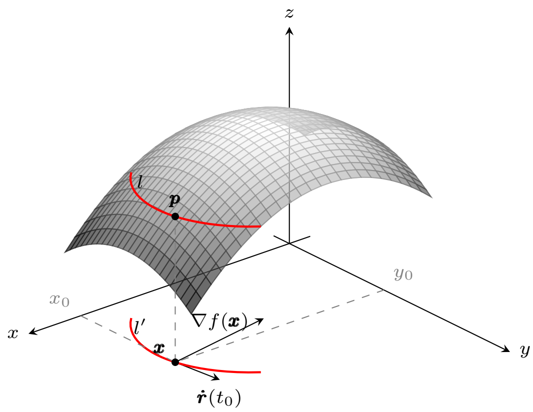
$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) = \nabla f \cdot \dot{\boldsymbol{r}} = 0 ,$$

gdje je $\dot{\boldsymbol{r}}(t) = \left(\frac{dx(t)}{dt}, \frac{dy(t)}{dt}\right)$ tangentni vektor na konturnu liniju l' u tački $\boldsymbol{x}$.

Gornje će vrijediti u proizvoljnoj tački konturne linije l', a time i u $\boldsymbol{x}$,

$$\nabla f(\boldsymbol{r}(t_0)) \cdot \dot{\boldsymbol{r}}(t_0) = \nabla f(x_0, y_0) \cdot \dot{\boldsymbol{r}}(t_0) = 0 \ .$$

Kako ovo vrijedi za bilo koju tačku konturne linije, zbog osobine da je skalarni produkt nenula vektora jednak 0 ako i samo ako su vektori ortogonalni, gornje razmatranje sumiramo generalnom tvrdnjom.

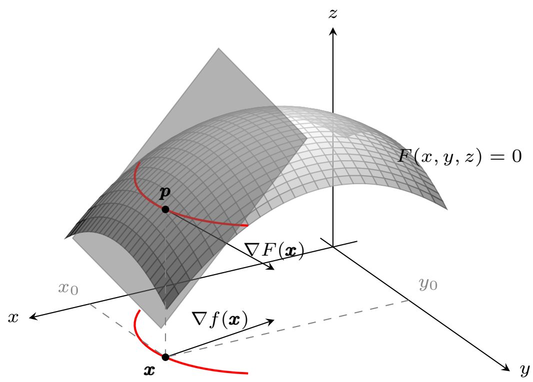


Slika 3.2: Položaj gradijentnog vektora.

Teorem 3.6.2. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna funkcija. Gradijentni vektor funkcije $z = f(\boldsymbol{x})$ u svakoj tački nivo površi $f(\boldsymbol{x}) = c$, ortogonalan je na odgovarajuću konturnu površ.*

Funkciju $z = f(x, y)$ možemo zapisati i kao implicitnu funkciju $F(x, y, z) = f(x, y) - z = 0$. Implicitno zadata funkcija tri varijable, $F(x, y, z) = 0$, u tom slučaju predstavlja jednu konkretnu konturnu površ funkcije $w = F(x, y, z)$ $(w = 0)$. Prema gornjem tvrđenju, gradijentni vektor funkcije w je ortogonalan na odgovarajuću konturnu površ.

$$\nabla w = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1\right).$$



Ovaj vektor kao vektor normale, definiše ravan u prostoru koja prolazi kroz neku tačku $\boldsymbol{p} = (x_0, y_0, z_0)$ na nivo površi $F(x, y, z) = 0$. Ta ravan data je sa

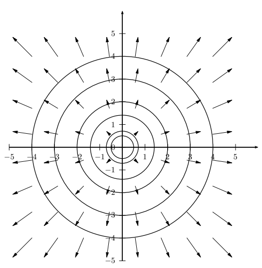
$$\nabla F(x_0, y_0, z_0) \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0,$$

odnosno ako izvršimo ovo skalarno množenje dobijamo

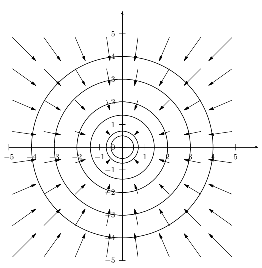
$$\frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0) - (z - z_0) = 0. \quad (3.22)$$

Dobijena ravan se naziva *tangencijalna ravan* na površ u prostoru, u tački (x_0, y_0, z_0).

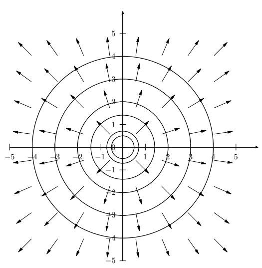
Na sljedećoj slici prikazano je nekoliko funkcija konturnim grafom (pomoću nivo linija, odnosno konturnih linija) i odgovarajućim "vektorskim poljem" ("strelice" na slici predstavljaju gradijentne vektore date funkcije u raznim tačkama). "Strelice" su usmjerene u pravcu najbržeg rasta funkcije, a veličina strelica odražava brzinu promjene funkcije u tom pravcu. Takođe uočavamo ortogonalnost gradijentnih vektora na odgovarajuće konturne linije.



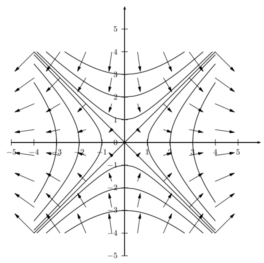
$f(x, y) = x^2 + y^2$



$f(x, y) = 1 - x^2 - y^2$



$f(x, y) = \sqrt{x^2 + y^2}$



$f(x, y) = x^2 - y^2$

Primjer 22 : Posmatrajmo površ u prostoru zadatu sa $f(x,y) = 4 - 2x^2 - y^2$. U tački $\boldsymbol{a} = \left(\frac{1}{2}, \frac{1}{2}, \frac{13}{4}\right)$ na površi prosuta je voda. Odrediti putanju kretanja vode!

Rješenje: Neka je $\vec{\boldsymbol{x}}(t) = (x(t), y(t))$ vektorska funkcija realne promjenljive koja opisuje putanju kretanja vode u xOy ravni (projekciju u xOy ravan stvarne putanje vode na datoj površi). Svakome je poznato da će se voda kretati "linijom najmanjeg otpora", to jest kretat će se niz brdo u najstrmijem pravcu (ako je to moguće), u svakoj tački putanje. Ovo znači da će se voda kretati u pravcu $-\nabla f$ (ovdje ćemo zanemariti fizikalne efekte kao što je naprimjer moment kretanja). Kako je $\dot{\vec{\boldsymbol{x}}}(t)$ vektor brzine (izvod puta je brzina), jasno je da je njegov pravac onda paralelan sa ∇f. Ovo opet znači da postoji skalar c takav da je

$$c\nabla f = \dot{\vec{\boldsymbol{x}}}(t) = (x'(t), y'(t)) \ . \tag{3.23}$$

Kako je $\nabla f = (-4x, -2y)$ i uzimajući da je $x'(t) = \frac{dx}{dt}$ i $y'(t) = \frac{dy}{dt}$, iz (3.23) imamo

$$(-4cx, -2cy) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \ ,$$

iz čega dobijamo sistem

$$-4cx = \frac{dx}{dt} \ ; \ -2cy = \frac{dy}{dt} \quad \Longleftrightarrow \quad c = -\frac{1}{4x}\frac{dx}{dt} \ ; \ c = -\frac{1}{2y}\frac{dy}{dt} \ .$$

Posljednje nam daje vezu

$$-\frac{1}{4x}\frac{dx}{dt} = -\frac{1}{2y}\frac{dy}{dt} \ .$$

Želimo li pronaći eksplicitnu vezu između x i y, gornju jednakost ćemo integraliti po promjenljivoj t.

$$\begin{aligned}
\int \frac{1}{4x}\frac{dx}{dt}dt &= \int \frac{1}{2y}\frac{dy}{dt}dt \\
\int \frac{1}{4x}dx &= \int \frac{1}{2y}dy \\
\frac{1}{4}\ln|x| + C_1 &= \frac{1}{2}\ln|y| + C_2 \\
\ln|x| &= 2\ln|y| + C_3 \quad (C_3 = C_2 - C_1) \\
\ln|x| &= \ln|y^2| + C_3 \ .
\end{aligned}$$

Eksponenciranjem obje strane posljednje jednakosti sa bazom e imamo,

$$x = e^{\ln y^2 + C_3} = e^{\ln y^2} \cdot \underbrace{e^{C_3}}_{C} = y^2 \cdot C \ .$$

Kako je polazna tačka kretanja vode u $\left(\frac{1}{2}, \frac{1}{2}\right)$, možemo izračunati vrijednost konstante C.

$$x = y^2 \cdot C \implies \frac{1}{2} = \frac{1}{4} \cdot C \implies C = 2 \ .$$

Dakle, putanja vode, posmatrano u xOy ravni, se odvija po krivoj $x = 2y^2$.

3.7 Pravila diferenciranja

Kao što smo već mogli primjetiti, pravila nalaženja diferencijala funkcija više varijabli neće se razlikovati od tih pravila kod funkcije jedne varijabe.

Teorem 3.7.1. *Neka su funkcije $f, g : D \to \mathbb{R}$ $(D \subseteq \mathbb{R}^n)$ diferencijabilne u tački $\boldsymbol{a} \in D$ i neka su $\alpha, \beta \in \mathbb{R}$ proizvoljni. Tada je i funkcija $\alpha f + \beta g$ diferencijabilna u tački $\boldsymbol{a}$ i vrijedi*

$$d(\alpha f + \beta g)(\boldsymbol{a}) = \alpha df(\boldsymbol{a}) + \beta dg(\boldsymbol{a}) \ .$$

Teorem 3.7.2. *Neka su funkcije $f, g : D \to \mathbb{R}$ ($D \subseteq \mathbb{R}^n$) diferencijabilne u tački $\boldsymbol{a} \in D$. Tada su i funkcije $f \cdot g$ i $\frac{f}{g}$ (posljednja uz uslov $g(\boldsymbol{a}) \neq 0$) diferencijabilne u tački $\boldsymbol{a}$ i vrijedi*

$$d(fg)(\boldsymbol{a}) = g(\boldsymbol{a})df(\boldsymbol{a}) + f(\boldsymbol{a})dg(\boldsymbol{a}) \ ,$$

$$d\left(\frac{f}{g}\right)(\boldsymbol{a}) = \frac{g(\boldsymbol{a})df(\boldsymbol{a}) - f(\boldsymbol{a})dg(\boldsymbol{a})}{(g(\boldsymbol{a}))^2} \ .$$

PRIMJER 23 : Odrediti prvi diferencijal funkcije $f(x,y) = \frac{1}{2}xe^{x^2+y^2}$.

Rješenje: Naravno da određivanju prvog diferencijala ove funkcije možemo pristupiti na klasičan način, određivanjem parcijalnih izvoda i primjenom relacija (3.10). Međutim, poslužimo se u tom određivanju gornjim rezultatima. Ako zamislimo funkciju f kao proizvod dviju funkcija, to jest $f(x,y) = g(x,y)h(x,y)$, gdje su $g(x,y) = x$ i $h(x,y) = \frac{1}{2}e^{x^2+y^2}$, mi treba da odredimo diferencijal proizvoda. Kako je

$$dg = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = dx \ ,$$

$$dh = \frac{\partial h}{\partial x}dx + \frac{\partial h}{\partial y}dy = xe^{x^2+y^2}dx + ye^{x^2+y^2}dy \ ,$$

prema prvoj formuli gornjeg teorema je

$$\begin{aligned} df = d(gh) = gdh + hdg &= x\left(xe^{x^2+y^2}dx + ye^{x^2+y^2}dy\right) + \frac{1}{2}e^{x^2+y^2}dx \\ &= \left(x^2 + \frac{1}{2}\right)e^{x^2+y^2}dx + xye^{x^2+y^2}dy \ . \end{aligned}$$

3.8 Izvodi višeg reda, Hesseova matrica

Ukoliko funkcija $f : \mathbb{R}^n \to \mathbb{R}$ ima parcijalne izvode koji postoje na nekom otvorenom skupu U, tada za svako $i \in \{1, 2, ..., n\}$, $\frac{\partial f}{\partial x_i}$ je takođe funkcija sama za sebe, to jest $\frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$. Tada naravno ima smisla tražiti parcijalne izvode i ovakve funkcije. Parcijalni izvodi funkcije $\frac{\partial f}{\partial x_i}$, ukoliko postoje, nazivaju se *parcijalni izvodi drugog reda* funkcije f ili jednostavno drugi parcijalni izvodi.

Kao i za prve parcijalne izvode i za druge parcijalne izvode postoje razne oznake kao naprimjer: $\frac{\partial^2 f}{\partial x_i \partial x_j}$, $f''_{x_i x_j}$, $D_{x_i x_j}f$ ili jednostavno $f_{x_i x_j}$. Mi ćemo se služiti uglavnom prvom navedenom notacijom, ali po potrebi skraćivanja zapisa, često ćemo upotrebljavati i posljednju navedenu notaciju. Tako za funkciju $z = f(x,y)$ imamo sljedeće parcijalne izvode drugog reda, zapisane i sa prvom i sa posljednjom notacijom:

$$f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} \ ; \ f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$$

$$f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x \partial y} \ ; \ f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

Tehnika nalaženja parcijalnih izvoda drugog reda sadržana je u simboličkom zapisivanju tih izvoda. Naprimjer, f_{xy} znači da od izvoda f'_x (prvi s lijeva indeks nam govori od koga pravimo parcijalni izvod) nalazimo parcijalni izvod po y (drugi indeks s lijeva nam govori po čemu radimo drugi parcijalni izvod).

PRIMJER 24 : Odrediti parcijalne izvode drugog reda funkcije $z = x^2y$.

Rješenje: Prvo odredimo parcijalne izvode prvog reda:

$$\frac{\partial z}{\partial x} = 2xy \ ; \ \frac{\partial z}{\partial y} = x^2 \ .$$

Odredimo sada parcijalne izvode drugog reda, koristeći gornje objašnjenje. Za nalaženje f_{xx}, uzimamo $\frac{\partial z}{\partial x}$ i od njega tražimo parcijalni izvod po x. Tako dobijamo

$$f_{xx} = (2xy)'_x = 2y \ .$$

Analogno, za f_{xy} uzimamo prvi parcijalni izvod po x, pa od njega tražimo izvod po y

$$f_{xy} = (2xy)'_y = 2x \ .$$

Istu logiku koristimo kod nalaženja ostala dva parcijalna izvoda drugog reda,

$$f_{yx} = 2x \quad ; \quad f_{yy} = 0 \ .$$

PRIMJER 25 : $z = f(x,y) = e^{x^2+y^2}$

$$f_x = 2xe^{x^2+y^2} \quad ; \quad f_y = 2ye^{x^2+y^2} \ .$$

$$f_{xx} = 2e^{x^2+y^2} + 2x2xe^{x^2+y^2} = 2e^{x^2+y^2}(1+2x^2) \ ; \ f_{xy} = 2x2ye^{x^2+y^2} = 4xye^{x^2+y^2} \ ;$$

$$f_{yx} = 2y2xe^{x^2+y^2} = 4xye^{x^2+y^2} \quad ; \ f_{yy} = 2e^{x^2+y^2} + 2y2ye^{x^2+y^2} = 2e^{x^2+y^2}(1+2y^2) \ .$$

Za funkciju $f : \mathbb{R}^n \to \mathbb{R}$, druge parcijalne izvode $f_{x_i x_j}$ i $f_{x_j x_i}$ $(i \neq j)$ nazivamo *mješoviti parcijalni izvodi* i na osnovu opisanog postupka, jasna nam je razlika između njih istaknuta poretkom indeksa.
U pokazana dva primjera primijećujemo da su mješoviti parcijalni izvodi jednaki, $f_{xy} = f_{yx}$. Postavlja se pitanje da li je to tako u opštem slučaju? Kao što ćemo kasnije vidjeti taj uslov je veoma bitan, a ovdje ćemo dati uslove pod kojima su ti parcijalni izvodi jednaki za funkciju dvije promjenljive. Prije toga, odgovor na postavljeno pitanje nam daje sljedeći primjer.

PRIMJER 26 : Posmatrajmo funkciju $f : \mathbb{R}^2 \to \mathbb{R}$ zadatu sa

$$f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & ; \quad (x,y) \neq (0,0) \\ 0 & ; \quad (x,y) = (0,0) \end{cases}$$

Tada je

$$\frac{\partial f}{\partial x}(x,y) = y\frac{x^4 - y^4 + 4x^2y^2}{(x^2+y^2)^2} \ , \ (x,y) \neq (0,0) \ ; \ \frac{\partial f}{\partial x}(0,0) = 0 \ .$$

Onda je

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{h\to 0} \frac{1}{h}\left(\frac{\partial f}{\partial x}(0,h) - \frac{\partial f}{\partial x}(0,0)\right) = \lim_{h \to 0} \frac{-h}{h} = -1 \ .$$

Na sličan način određujući, imamo da je

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = 1 \ ,$$

pa očigledno u opštem slučaju mješoviti izvodi nisu jednaki.

Definicija 3.8.1. *Za funkciju $f : \mathbb{R}^n \to \mathbb{R}$ kažemo da je dva puta neprekidno diferencijabilna na otvorenom skupu $U \subseteq \mathbb{R}^n$, i pišemo $f \in C^2(U)$, ako su funkcije $f_{x_i x_j}$ neprekidne na U, za sve $i, j \in \{1, 2, ..., n\}$.*

Pod određenim uslovima koji su dati u narednom teoremu, mješoviti izvodi će biti jednaki.

Teorem 3.8.1. *Neka je $U \subseteq \mathbb{R}^n$ otvoren skup koji sadrži tačku $\boldsymbol{a}$ i neka je funkcija $f \in C^2(U)$.*

Tada vrijedi

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{a}) \ ,$$

za sve $i, j \in \{1, 2, ..., n\}$ *takve da je* $i \neq j$.

Gornje rezonovanje o parcijalnim izvodima drugog reda sada možemo proširiti na parcijalne izvode trećeg, četvrtog i višeg reda, kao i na funkcije tri, četiri i više promjenljivih. Za funkciju dvije varijable, vidjeli smo, postoje četiri parcijalna izvoda drugog reda. Praveći od njih ponovo parcijalne izvode, dobijamo parcijalne izvode trećeg reda, kojih će tada biti osam. Za funkciju tri varijable, parcijalnih izvoda drugog reda ima devet, a trećeg reda 27.

U opštem slučaju, funkcija $f : \mathbb{R}^n \to \mathbb{R}$ ima n^2 parcijalnih izvoda drugog reda, od kojih onda možemo formirati kvadratnu matricu reda $n \times n$.

Definicija 3.8.2. *Neka svi parcijalni izvodi drugog reda funkcije* $f : \mathbb{R}^n \to \mathbb{R}$ *postoje u tački* $\boldsymbol{a} \in \mathbb{R}^n$. *Matricu reda* $n \times n$

$$\boldsymbol{H}f(\boldsymbol{a}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_3 \partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{a}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\boldsymbol{a}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_3}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_3^2}(\boldsymbol{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_3}(\boldsymbol{a}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\boldsymbol{a}) & \frac{\partial^2 f}{\partial x_3 \partial x_n}(\boldsymbol{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\boldsymbol{a}) \end{bmatrix} \tag{3.24}$$

nazivamo Hesseova matrica ili Hessijan funkcije f *u tački* $\boldsymbol{a}$.

Primjetimo da je i-ta kolona Hesseove matrice, ništa drugo do gradijent funkcije $\frac{\partial f}{\partial x_i}$. Formalno, hesijan možemo zapisati sa

$$\boldsymbol{H}f(\boldsymbol{a}) = \left[\nabla \frac{\partial f}{\partial x_1}(\boldsymbol{a}) \ \nabla \frac{\partial f}{\partial x_2}(\boldsymbol{a}) \ \ldots \ \nabla \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \right],$$

zbog čega ćemo češće za hesijan $\boldsymbol{H}f(\boldsymbol{a})$ koristiti oznaku $\nabla^2 f(\boldsymbol{a})$ (kao gradijent gradijenta).

Primjer 27 : Neka je $f(x, y) = x^2 y - xy^2$. Tada je

$$\nabla^2 f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{yx}(x, y) \\ f_{xy}(x, y) & f_{yy}(x, y) \end{bmatrix} = \begin{bmatrix} 2y & 2x - 2y \\ 2x - 2y & -2x \end{bmatrix} .$$

Sada naprimjer, u tački $\boldsymbol{a} = (2, 1)$, Hessijan glasi

$$\nabla^2 f(2, 1) = \begin{bmatrix} 2 & 2 \\ 2 & -4 \end{bmatrix} .$$

3.9 Taylorova formula

Neka je $f : \mathbb{R}^2 \to \mathbb{R}$ dva puta neprekidno diferencijabilna na nekoj otvorenoj kugli $B(\boldsymbol{a}, r) \subseteq \mathbb{R}^2$ i neka je $\boldsymbol{h} = (h_1, h_2)$ vektor, takav da je $||\boldsymbol{h}|| < r$. Definišimo novu funkciju $\varphi : \mathbb{R} \to \mathbb{R}$, na sljedeći način

$$\varphi(t) = f(\boldsymbol{a} + t\boldsymbol{h}) \ .$$

(Veličinu $\boldsymbol{a} + t\boldsymbol{h}$ shvatamo tako da se iz tačke $\boldsymbol{a}$ pomjerimo u pravcu vektora $\boldsymbol{h}$, za dužinu $t||\boldsymbol{h}||$). Funkcija φ je funkcija jedne varijable i pri tome je specijalno $\varphi(0) = f(\boldsymbol{a})$ i $\varphi(1) = f(\boldsymbol{a} + \boldsymbol{h})$. Na osnovu Taylorovog teorema[3] za funkciju jedne varijable sada imamo

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(\xi)\ , \tag{3.25}$$

gdje je $\xi \in (0,1)$. Kako je $d\varphi = df$, koristeći pravilo izvoda kompozicije, imamo

$$\varphi'(t) = \nabla f(\boldsymbol{a}+t\boldsymbol{h}) \cdot \frac{d}{dt}(\boldsymbol{a}+t\boldsymbol{h}) = \nabla f(\boldsymbol{a}+t\boldsymbol{h}) \cdot \boldsymbol{h} = f_x(\boldsymbol{a}+t\boldsymbol{h})h_1 + f_y(\boldsymbol{a}+t\boldsymbol{h})h_2\ . \tag{3.26}$$

(jasno, u izrazu $\nabla f(\boldsymbol{a}+t\boldsymbol{h}) \cdot \boldsymbol{h}$ imamo skalarno množenje). Analogno nalazimo i drugi izvod

$$\begin{aligned}\varphi''(t) &= \nabla\left(h_1 f_x(\boldsymbol{a}+t\boldsymbol{h}) + h_2 f_y(\boldsymbol{a}+t\boldsymbol{h})\right) \cdot \boldsymbol{h} \\ &= \left(h_1 \nabla f_x(\boldsymbol{a}+t\boldsymbol{h}) + h_2 \nabla f_y(\boldsymbol{a}+t\boldsymbol{h})\right) \cdot (h_1, h_2) \\ &= [h_1\ h_2] \begin{bmatrix} f_{xx}(\boldsymbol{a}+t\boldsymbol{h}) & f_{xy}(\boldsymbol{a}+t\boldsymbol{h}) \\ f_{yx}(\boldsymbol{a}+t\boldsymbol{h}) & f_{yy}(\boldsymbol{a}+t\boldsymbol{h}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}\ .\end{aligned}$$

(Zadnji zapis dobijamo nakon jednostavnog matričnog računa). Koristeći sada oznake

$$\boldsymbol{h} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

i

$$\boldsymbol{h}^T = [h_1\ h_2]\ ,$$

posljednje možemo zapisati sa

$$\varphi''(t) = \boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a}+t\boldsymbol{h}) \cdot \boldsymbol{h}\ . \tag{3.27}$$

Stavljajući (3.26) i (3.27) u izraz (3.25), dobijamo sljedeću vezu

$$f(\boldsymbol{a}+\boldsymbol{h}) = \varphi(1) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a})^T \cdot \boldsymbol{h} + \frac{1}{2}\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a}+\xi\boldsymbol{h}) \cdot \boldsymbol{h}\ .$$

Ovaj rezultat predstavlja verziju Taylorove teoreme za funkcije više varijabli, koga generalizujemo sljedećim teoremom

Teorem 3.9.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ i neka je $f \in C^2(B(\boldsymbol{a}, r))$ $(r > 0)$. Neka je $\boldsymbol{h}$ vektor, takav da je $||\boldsymbol{h}|| < r$. Tada postoji realan broj $\xi \in (0,1)$, takav da vrijedi*

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a})^T \cdot \boldsymbol{h} + \frac{1}{2}\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a}+\xi\boldsymbol{h}) \cdot \boldsymbol{h}\ . \tag{3.28}$$

Uvedemo li oznake $\boldsymbol{x} = \boldsymbol{a}+\boldsymbol{h}$ i izračunamo li Hessijan u tački $\boldsymbol{a}$, izraz (3.28) predstavlja polinomijalnu aproksimaciju funkcije f.

Definicija 3.9.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ dva puta neprekidno diferencijabilna na nekoj otvorenoj kugli oko tačke $\boldsymbol{a}$. Funkciju*

$$P_2(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a})^T \cdot (\boldsymbol{x}-\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T \cdot \nabla^2 f(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a})\ ,$$

nazivamo Taylorov polinom drugog reda, funkcije f u tački $\boldsymbol{a}$.

[3] Neka je $f : \mathbb{R} \to \mathbb{R}$ $n+1$ puta neprekidno diferencijabilna na otvorenom intervalu koji sadrži tačku a. Tada za svako x iz tog intervala vrijedi,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n+1}(x),$$

gdje izraz $R_{n+1}(x)$ predstavlja grešku i zadovoljava $R_{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$, za neko c između a i x.

PRIMJER 28 : Odredimo Taylorov polinom drugog reda za funkciju $f(x,y) = e^{-2x+y}$, u tački $(0,0)$.

Rješenje: Kao prvo nalazimo gradijent i hesijan zadate funkcije,

$$\nabla f(x,y) = \left(-2e^{-2x+y}, e^{-2x+y}\right) \ \text{ i } \ \nabla^2 f(x,y) = \begin{bmatrix} 4e^{-2x+y} & -2e^{-2x+y} \\ -2e^{-2x+y} & e^{-2x+y} \end{bmatrix},$$

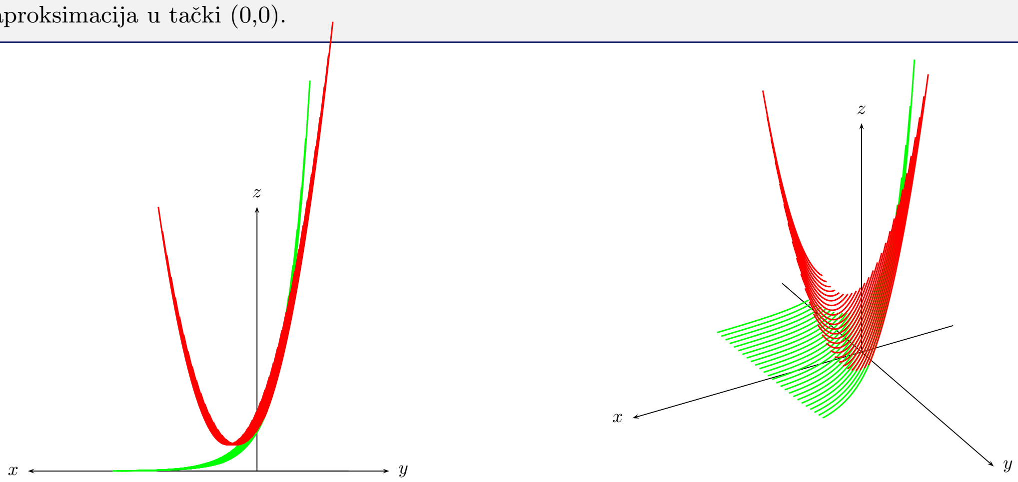
odnosno

$$\nabla f(0,0) = (-2,1)\ ,\ \nabla^2 f(0,0) = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}.$$

Sada imamo

$$\begin{aligned} P_2(x,y) &= f(0,0) + \nabla f(0,0)\cdot(x,y) + \frac{1}{2}[x\ y]\mathbf{H}f(0,0)\begin{bmatrix} x \\ y \end{bmatrix} \\ &= 1 + (-2,1)\cdot(x,y) + \frac{1}{2}[x\ y]\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= 1 - 2x + y + \frac{1}{2}[x\ y]\begin{bmatrix} 4x-2y \\ -2x+y \end{bmatrix} \\ &= 1 - 2x + y + \frac{1}{2}(4x^2 - 2xy - 2xy + y^2) \\ &= 2x^2 + \frac{1}{2}y^2 - 2xy - 2x + y + 1\ . \end{aligned}$$

Svrha Taylorovog polinoma je da se funkcija njime dovoljno dobro aproksimira u okolini neke tačke. Na slici (3.3) dat je prikaz te aproksimacije iz dva ugla posmatranja, da bi se bolje uočila istaknuta aproksimacija u tački (0,0).



Slika 3.3: Aproksimacija funkcije $f(x,y) = e^{-2x+y}$ (zelena) u tački (0,0), Taylorovim polinomom $P_2(x,y) = 2x^2 + \frac{1}{2}y^2 - 2xy - 2y + y + 1$ (crvena)

3.10 Kvadratne forme i definitnost matrice

U dijelu linearne algebre, koga smo izučavali ranije, upoznali smo pojam *simetrične matrice*, to jest kvadratne matrice $M = [a_{ij}]_{n\times n}$ za koju vrijedi

$$M = M^T\ ,$$

ili za čije elemente vrijedi $a_{ij} = a_{ji}$.

PRIMJER 29 : Matrica

$$M = \begin{bmatrix} -1 & 2 & -2 \\ 2 & 5 & 0 \\ -2 & 0 & 3 \end{bmatrix},$$

primjer je simetrične matrice, a matrica

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} ,$$

je primjer nesimetrične matrice.

Ako je $f \in C^2$, tada na osnovu Teorema (3.8.1), imamo da su mješoviti izvodi jednaki, to jest

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

a to će onda na osnovu definicije Hessijana značiti da je za svaku dva puta neprekidno diferencijabilnu funkciju, njen Hessijan simetrična matrica.

Neka je sada M proizvoljna simetrična matrica reda $n \times n$. Za proizvoljnu matricu kolonu $\boldsymbol{x}$ (možemo reći i vektor $\boldsymbol{x} = (x_1, x_2, ..., x_n)$), definišimo funkciju $q : \mathbb{R}^n \to \mathbb{R}$ na sljedeći način

$$q(\boldsymbol{x}) = \boldsymbol{x}^T M \boldsymbol{x} . \tag{3.29}$$

Funkcija q je polinom drugog reda po promjenljivima $x_1, x_2, ..., x_n$ i nazivamo je *kvadratna forma* po promjenljivima $x_1, x_2, ..., x_n$, a matricu M nazivamo *matrica kvadratne forme q*.

PRIMJER 30 : Neka je

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} .$$

Kvadratnu formu dobijamo iz (3.29),

$$q_2(\boldsymbol{x}) = \boldsymbol{x}^T M \boldsymbol{x} = [x_1 \ x_2] \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 + 2x_2 \ \ 2x_1 + x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + x_2^2 + 4x_1x_2 .$$

Za matricu

$$M = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 2 \end{bmatrix} ,$$

kvadratna forma u tački $\boldsymbol{x} = (x_1, x_2, x_3)$ glasi

$$q_3(\boldsymbol{x}) = 2x_1^2 + 5x_2^2 + 2x_3^2 - 2x_1x_2 + 2x_1x_3 + 2x_2x_3 .$$

Definicija 3.10.1. *Za kvadratnu formu $q(\boldsymbol{x}) = \boldsymbol{x}^T M \boldsymbol{x}$ kažemo da je*

- *pozitivno poludefinitna, ako je za svako $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, zadovoljeno $q(\boldsymbol{x}) \geq 0$.*
- *pozitivno definitna, ako je za svako $\boldsymbol{x} \neq 0$, zadovoljeno $q(\boldsymbol{x}) > 0$.*
- *negativno poludefinitna, ako je za svako $\boldsymbol{x} \in \mathbb{R}^n$, zadovoljeno $q(\boldsymbol{x}) \leq 0$.*
- *negativno definitna, ako je za svako $\boldsymbol{x} \neq 0$, zadovoljeno $q(\boldsymbol{x}) < 0$.*
- *indefinitna ili promjenljivog znaka, ako postoje $\boldsymbol{x}', \boldsymbol{x}'' \in \mathbb{R}^n$, tako da je $q(\boldsymbol{x}') > 0$ i $q(\boldsymbol{x}'') < 0$.*
- *kažemo da je kvadratna forma nedefinitna u tački, ako je $q(\boldsymbol{x}) = 0$ za $\boldsymbol{x} \neq \mathbf{0}$.*

PRIMJER 31 : Kvadratnu formu q_2 iz gornjeg primjera gdje je

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} ,$$

možemo zapisati

$$q_2(\boldsymbol{x}) = x_1^2 + x_2^2 + 4x_1x_2 = (x_1 + x_2)^2 + 2x_1x_2 ,$$

pa za $\boldsymbol{x} = (1,0)$ imamo $q_2(\boldsymbol{x}) = 1 > 0$, a za $\boldsymbol{x} = (1,-1)$ imamo $q_2(\boldsymbol{x}) = -2 < 0$. Na osnovu definicije, kvadratna forma q_2 je indefinitna.

Kvadratnu formu q_3 možemo nakon malo računa zapisati sa

$$q_3(\boldsymbol{x}) = (x_1 + x_2 + x_3)^2 + (x_1 - 2x_2)^2 + x_3^2 \ ,$$

pa je očigledno ova kvadratna forma pozitivno definitna (kao suma kvadrata) odnosno, za svako $\boldsymbol{x} = (x_1, x_2, x_3) \neq 0$ je $q_3(\boldsymbol{x}) > 0$.

Kao što ćemo uskoro vidjeti, od velikog je interesa imati način određivanja definitnosti neke kvadratne forme. Najjednostavniji način bio bi obrazovati tu kvadratnu formu, a onda je svesti na neki "pogodan" oblik iz koga "lagano" možemo ocijeniti njenu definitnost (ovo smo primjenili u posljednjem primjeru). Nađimo taj način u za nas važnom slučaju 2×2 matrice. Neka je zadata simetrična matrica

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \ .$$

Kvadratna forma određena ovom matricom je

$$q(x,y) = [x\ y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 \ .$$

Ako je $a \neq 0$, poznatim postupkom svođenja trinoma na kanonski oblik dobijamo

$$\begin{aligned} q(x,y) &= a\left(x^2 + \frac{2b}{a}xy + \frac{c}{a}y^2\right) \\ &= a\left(\left(x + \frac{b}{a}y\right)^2 - \frac{b^2}{a^2}y^2 + \frac{c}{a}y^2\right) \\ &= a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2 \\ &= a\left(x + \frac{b}{a}y\right)^2 + \frac{det(M)}{a}y^2 \ . \end{aligned}$$

Sada imamo diskusiju:

1. Ako je $a > 0$ i $det(M) > 0$, tada je za svako $(x,y) \neq (0,0)$, $q(x,y) > 0$, to jest kvadratna forma je pozitivno definitna.

2. Ako je $a < 0$ i $det(M) > 0$, tada je za svako $(x,y) \neq (0,0)$, $q(x,y) < 0$, to jest kvadratna forma je negativno definitna.

3. Ako je $det(M) < 0$, tada u tačkama $(x,y) = (1,0)$ i $(x,y) = (-\frac{b}{a}, 1)$ imamo različite znakove kvadratne forme, pa je ona indefinitna.

4. Ako je $det(M) = 0$, tada imamo

$$q(x,y) = a\left(x + \frac{b}{a}y\right)^2 \ .$$

Ako je $x \neq -\frac{b}{a}y$, kvadratna forma ima znak koga ima parametar a. Dakle, $q(x,y)$ je ili pozitivno ili negativno poludefinitan. Ako je $x = -\frac{b}{a}y$, onda je $q(x,y) = 0$, te je u ovoj tački kvadratna forma nedefinitna.

Jasno nam je da bi ovakav postupak određivanja definitosti kvadratnih formi, određenih matrica viših dimenzija, bio poprilično težak posao. Zato sljedećim teoremom dajemo veoma jednostavan kriterij za utvrđivanje definitnosti kvadratne forme.

Teorem 3.10.1. ***(Sylvesterov kriterijum)***

Neka je

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} ,$$

proizvoljna kvadratna matrica koja određuje kvadratnu formu $q : \mathbb{R}^n \to \mathbb{R}$. *Označimo sa*

$$A_i = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix} , \quad (i = 1, 2, ..., n)$$

glavne minore matrice M, *to jest*

$$A_1 = a_{11} \ , \ A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} , \ \cdots \ , \ A_n = det(M) \ .$$

Kvadratna forma q *je pozitivno definitna ako i samo ako su svi glavni minori matrice* M *pozitivni, to jest ako vrijedi* $A_i > 0$ *za* $i = 1, 2, \ldots n$.
Kvadratna forma je negativno definitna ako i samo ako su glavni minori alternativnih znakova, tako da je $A_1 < 0$, *to jest ako je* $A_i < 0$ *za neparne* i, *a* $A_i > 0$ *za parne* i.
Kvadratna forma je indefinitna ako je $A_n \neq 0$, *a nije ni pozitivno definitna ni negativno definitna.*
Ako je $A_n = 0$ *kvadratna forma je nedefinitna.*

Uobičajeno je i za matricu M reći da je pozitivno definitna, negativno definitna ili indefinitna kad god je takva kvadratna forma koja je njome određena.

Primjer 32 : Za matricu

$$M = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 2 \end{bmatrix} ,$$

glavni minori su $A_1 = 2 > 0$, $A_2 = 9 > 0$ i $A_3 = det(M) = 9 > 0$, pa je kvadratna forma određena ovom matricom pozitivno definitna.

Za matricu

$$M = \begin{bmatrix} -2 & 1 \\ 1 & -4 \end{bmatrix} ,$$

glavni minori su $A_1 = -2 < 0$ i $A_2 = det(M) = 7 > 0$, pa je kvadratna forma negativno definitna.

Za matricu

$$M = \begin{bmatrix} -3 & 1 \\ 1 & 2 \end{bmatrix} ,$$

glavni minori su $A_1 = -3 < 0$ i $A_2 = det(M) = -7 < 0$, pa je kvadratna forma indefinitna. Za matricu

$$M = \begin{bmatrix} -3 & 1 & 2 \\ 1 & 2 & -2 \\ 1 & 2 & -2 \end{bmatrix} ,$$

glavni minori su $A_1 = -3 < 0$ i $A_2 = det(M) = -7 < 0$ i $A_3 = det(M) = 0$, pa je kvadratna forma nedefinitna.

3.11 Diferencijali višeg reda

Neka je u oblasti $D \subset \mathbb{R}^n$ definisana funkcija $f(\boldsymbol{x})$ koja ima neprekidne parcijalne izvode do k-tog reda. Ranije smo vidjeli da totalni diferencijal ima oblik

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + ... + \frac{\partial f}{\partial x_n} dx_n \ ,$$

gdje su dx_i $(i = 1, 2, ..., n)$ priraštaji, odnosno diferencijali nezavisnih promjenljivih. Totalni diferencijal drugog reda ili kraće diferencijal drugog reda, definiše se kao diferencijal prvog diferencijala, to jest $d^2 f = d(df)$.

$$d^2 f = d\left(\frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + ... + \frac{\partial f}{\partial x_n}dx_n\right) .$$

Kako je $d(dx_i) = 0$ za svako $i = 1, 2, ..., n$, to sada imamo:

$$d^2 f = d\left(\frac{\partial f}{\partial x_1}\right)dx_1 + d\left(\frac{\partial f}{\partial x_2}\right)dx_2 + ... + d\left(\frac{\partial f}{\partial x_n}\right)dx_n ,$$

odakle sada primjenjujući formulu za diferencijal funkcije imamo

$$d^2 f = \frac{\partial^2 f}{\partial x_1^2}dx_1^2 + ... + \frac{\partial^2 f}{\partial x_n^2}dx_n^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}dx_1 dx_2 + ... + 2\frac{\partial^2 f}{\partial x_{n-1}\partial x_n}dx_{n-1}dx_n .$$

Ovaj postupak možemo generalizovati na diferencijale proizvoljnog reda, to jest vrijedi relacija,

$$d^k f = d(d^{k-1} f) , \ k \in \mathbb{N} .$$

Specijalno, za funkciju dvije varijable $f(x, y)$, drugi diferencijal je dat sa

$$d^2 f = \frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial y^2}dy^2 + \frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y \partial x}dxdy .$$

Primjer 33 : Odrediti drugi diferencijal funkcije $f(x, y) = x^3y^2 - x^2y + 2xy - 3x + 4$. Određujemo prve parcijalne izvode:

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

Drugi parcijalni izvodi su:

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

Konačno, izraz za drugi diferencijal je,

$$d^2 f(x, y) = (6xy^2 - 2y)dx^2 + 2(6x^2y - 2x + 2)dxdy + 2x^3dy^2 .$$

Primjer 34 : Odrediti drugi diferencijal funkcije $g(x, y, z) = \sin x + \sin y + \sin z$.
Prvi parcijalni izvodi su:

$$\frac{\partial g}{\partial x} = \cos x \ , \ \frac{\partial g}{\partial y} = \cos y \ , \ \frac{\partial g}{\partial z} = \cos z .$$

Drugi parcijalni izvodi su:

$$\frac{\partial^2 g}{\partial x^2} = -\sin x \ , \ \frac{\partial^2 g}{\partial y^2} = -\sin y \ , \ \frac{\partial^2 g}{\partial z^2} = -\sin z$$

$$\frac{\partial^2 g}{\partial x \partial y} = \frac{\partial^2 g}{\partial x \partial z} = \frac{\partial^2 g}{\partial y \partial x} = \frac{\partial^2 g}{\partial y \partial z} = \frac{\partial^2 g}{\partial z \partial x} = \frac{\partial^2 g}{\partial z \partial y} = 0 .$$

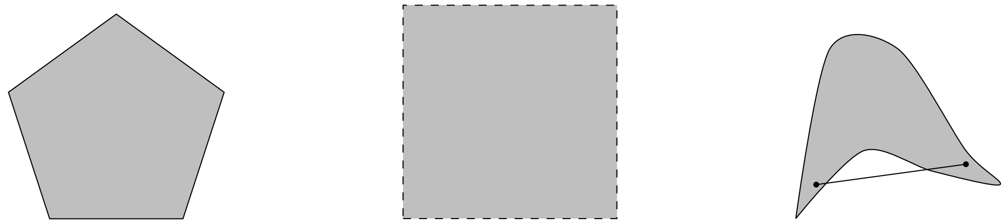
Dakle, drugi diferencijal je

$$d^2 g = -\sin x dx^2 - \sin y dy^2 - \sin z dz^2 .$$

3.12 Konveksne funkcije

Definicija 3.12.1. *Za skup $C \subseteq \mathbb{R}^n$ kažemo da je konveksan skup ako i samo ako duž koja spaja bilo koje dvije tačke skupa C leži u C, to jest*

$$(\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in C)(\forall \lambda \in [0,1])\ \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in C\ .$$



(a) Konveksan skup (b) Konveksan skup (c) Nekonveksan skup

Pojednostavljeno govoreći, skup je konveksan ako su sve tačke tog skupa "vidljive" iz bilo koje tačke tog skupa. Pri tome se pod "vidljivo" podrazumijeva po nesmetanom pravcu između tačaka (koji leži unutar skupa).

Definicija 3.12.2. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ i $C \subseteq \mathbb{R}^n$. Kažemo da je f konveksna na C ako i samo ako je skup C konveksan i za proizvoljne $\boldsymbol{x}, \boldsymbol{y} \in C$ i $\lambda \in [0,1]$ vrijedi*

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})\ . \tag{3.30}$$

Ukoliko u (3.30) vrijedi stroga nejednakost kažemo da je funkcija strogo konveksna.
Funkcija f je (strogo) konkavna ako i samo ako je $-f$ (strogo) konveksna.

Geometrijski gledano konveksna funkcija ima osobinu da tetiva između bilo koje dvije tačke na grafu funkcije se nalazi iznad odgovarajućeg dijela grafa te funkcije, Slika 3.4.

(d) (e)

Slika 3.4: (a) konveksna funkcija ; (b) nekonveksna funkcije.

Teorem 3.12.1. *Neka je $C \subseteq \mathbb{R}^n$ konveksan skup i neka su $f, g : C \to \mathbb{R}$ konveksne funkcije. Tada je i linearna kombinacija $af + bg$ takođe konveksna funkcija za proizvoljne $a, b \geq 0$.*

Teorem 3.12.2. *Neka su $C_1 \subseteq \mathbb{R}$ i $C_2 \subseteq \mathbb{R}^n$ konveksni skupovi. Neka je $g : C_1 \to \mathbb{R}$ neopadajuća konveksna funkcija na C_1 i neka je $h : C_2 \to \mathbb{R}$ konveksna funkcija takva da je $h(C_2) \subseteq C_1$, tada je i funkcija $f : C_2 \to \mathbb{R}$, definisana sa $f = g \circ h$, konveksna na C_2.*

Konveksnost funkcije se može okarakterisati na razne načine, a najčešći je u terminima izvoda. Za to su nam potrebne pretpostavke o diferencijabilnosti funkcije.

Teorem 3.12.3 (Uslovi konveksnosti prvog reda I).
Neka je $C \subseteq \mathbb{R}^n$ neprazan, otvoren i konveksan skup i neka je $f : \mathbb{R}^n \to \mathbb{R}$ neprekidno diferencijabilna na C. Funkcija f je konveksna na C ako i samo ako za proizvoljne $\boldsymbol{x}, \boldsymbol{y} \in C$ vrijedi

$$f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}) \ . \tag{3.31}$$

Dokaz : ($\Longrightarrow$) Neka su $\boldsymbol{x}, \boldsymbol{y} \in C$ proizvoljni i $\lambda \in (0,1)$. Zbog konveksnosti funkcije imamo

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \ .$$

Ovaj izraz, zbog $\lambda > 0$ možemo transformisati u ekvivalentan izraz

$$\frac{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - f(\boldsymbol{y})}{\lambda} \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) \ ,$$

ili

$$\frac{f(\boldsymbol{y} + \lambda(\boldsymbol{x} - \boldsymbol{y})) - f(\boldsymbol{y})}{\lambda} \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) \ .$$

Puštajući da $\lambda \to 0$ lijeva strana postaje izvod funkcije f u tački $\boldsymbol{y}$ u pravcu vektora $\boldsymbol{x} - \boldsymbol{y}$,

$$\lim_{\lambda \to 0} \frac{f(\boldsymbol{y} + \lambda(\boldsymbol{x} - \boldsymbol{y})) - f(\boldsymbol{y})}{\lambda} = D_{\boldsymbol{x}-\boldsymbol{y}} f(\boldsymbol{y}) \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) \ .$$

Kako je $D_{\boldsymbol{u}} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a})^T \cdot \boldsymbol{u}$, zaključujemo da vrijedi

$$\nabla f(\boldsymbol{y})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) \ ,$$

iz čega dobijamo traženu nejednakost (3.31).

($\Longleftarrow$) Neka je zadovoljena nejednakost (3.31) za proizvoljne i različite $\boldsymbol{x}, \boldsymbol{y} \in C$. Za proizvoljno $\lambda \in [0,1]$ posmatrajmo $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C$. Nejednakost (3.31) primjenimo na $\boldsymbol{x}$ i $\boldsymbol{z}$,

$$\begin{aligned}
f(\boldsymbol{x}) &\geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \boldsymbol{z}) \\
&= f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \lambda \boldsymbol{x} - (1-\lambda)\boldsymbol{y}) \\
&= f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot ((1-\lambda)\boldsymbol{x} - (1-\lambda)\boldsymbol{y}) \\
&= f(\boldsymbol{z}) + (1-\lambda)\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) \ ,
\end{aligned}$$

i na $\boldsymbol{y}$ i $\boldsymbol{z}$

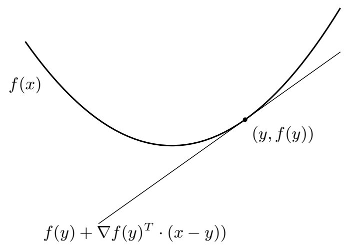
$$\begin{aligned}
f(\boldsymbol{y}) &\geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \boldsymbol{z}) \\
&= f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \lambda \boldsymbol{x} - (1-\lambda)\boldsymbol{y}) \\
&= f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T \cdot (\lambda \boldsymbol{y} - \lambda \boldsymbol{x}) \\
&= f(\boldsymbol{z}) + \lambda \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \ .
\end{aligned}$$

Pomnožimo prvu nejednakost sa λ, drugu sa $1-\lambda$ i sve to saberimo,

$$\begin{aligned}
&\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq \\
&\geq \lambda f(\boldsymbol{z}) + \lambda(1-\lambda)\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) + (1-\lambda) f(\boldsymbol{z}) + \lambda(1-\lambda)\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \\
&= \lambda f(\boldsymbol{z}) + \lambda(1-\lambda)\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) + (1-\lambda) f(\boldsymbol{z}) - \lambda(1-\lambda)\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) \\
&= f(\boldsymbol{z}) = f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \ .
\end{aligned}$$

Dakle, f je konveksna funkcija na C. ■

Kako izraz $f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y})$ predstavlja Taylorov razvoj prvog reda funkcije f u tački $\boldsymbol{y}$ (linearna aproksimacija), rezultat gornjeg teorema možemo tumačiti na način da je kod konveksne funkcije f njena linearna aproksimacija uvijek (u svakoj tački) "ispod" funkcije f. To možemo iskazati i sa time da je linearna aproksimacija (Taylorov razvoj prvog reda) globalno donje ograničenje konveksne funkcije. Oba ova tumačenja vrijede i obrnuto.



Slika 3.5: Interpretacija nejednakosti (3.31).

Teorem 3.12.4 (Uslovi konveksnosti prvog reda II).
Neka je $C \subseteq \mathbb{R}^n$ neprazan, otvoren i konveksan skup i neka je $f : \mathbb{R}^n \to \mathbb{R}$ neprekidno diferencijabilna na C. Funkcija f je konveksna na C ako i samo ako za proizvoljne $\boldsymbol{x}, \boldsymbol{y} \in C$ vrijedi

$$(\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}))^T \cdot (\boldsymbol{x} - \boldsymbol{y}) \geq 0 \ . \tag{3.32}$$

Dokaz : ($\Longrightarrow$) Za proizvoljne $\boldsymbol{x}, \boldsymbol{y} \in C$, na osnovu Teorema 3.12.3 imamo

$$f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T \cdot (\boldsymbol{x} - \boldsymbol{y})$$
$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \cdot (\boldsymbol{y} - \boldsymbol{x})$$

Sabiranjem ovih dviju nejednakosti dobijamo

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T \cdot (\boldsymbol{x} - \boldsymbol{y}) + \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \ ,$$

što nakon sređivanja daje nejednakost (3.32).

($\Longleftarrow$) Prema teoremu o srednjoj vrijednosti za $\boldsymbol{x}, \boldsymbol{y} \in C$ imamo

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) = \nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \ , \tag{3.33}$$

gdje je $\boldsymbol{z} = \lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}$, za neko $\lambda \in (0, 1)$. Po pretpostavci , za $\boldsymbol{z}$ i $\boldsymbol{x}$ vrijedi

$$(\nabla f(\boldsymbol{z}) - \nabla f(\boldsymbol{x}))^T \cdot (\boldsymbol{z} - \boldsymbol{x}) \geq 0 \ ,$$

a odavde imamo

$$(1 - \lambda) (\nabla f(\boldsymbol{z}) - \nabla f(\boldsymbol{x}))^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \geq 0 \ .$$

Odavde zaključujemo da vrijedi

$$\nabla f(\boldsymbol{z})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \geq \nabla f(\boldsymbol{x})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \ .$$

Koristeći (3.33) imamo

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \geq \nabla f(\boldsymbol{x})^T \cdot (\boldsymbol{y} - \boldsymbol{x}) \ ,$$

što daje traženi uslov teorema. ∎

Rezultat Teorema 3.12.4 nam govori da je gradijent, kao preslikavanje, monotono rastuće preslikavanje. Ako bi posmatrali funkciju jedne varijable, onda nam ovo tvrđenje kaže da je f konveksna funkcija ako i samo ako je f' neopadajuća funkcija.

Napomena: Ako u Teoremu 3.12.3 i Teoremu 3.12.4 u relacijama (3.31) i (3.32) respektivno vrijede stroge nejednakosti, tada zaključujemo u oba slučaja strogu konveksnost funkcije

U oba uslova prvog reda (Teorem 3.12.3 i Teorem 3.12.4) primjećujemo da je konveksnost obezbjeđena na otvorenom skupu. Naravno da se postavlja pitanje, šta ako je posmatrani skup C zatvoren? Odgovor nam daje naredno tvrđenje koje navodimo bez dokaza.

Teorem 3.12.5. *Neka je C konveksan skup u $\mathbb{R}^n$ sa nepraznom unutrašnjošću i neka je $f : C \to \mathbb{R}$. Ako je f neprekidna funkcija na C i konveksna na $int(C)$, tada je f konveksna i na C.*

Teorem 3.12.6 (Uslovi konveksnosti drugog reda).
Neka je $C \subseteq \mathbb{R}^n$ otvoren i konveksan skup i neka je $f : C \to \mathbb{R}$ dva puta neprekidno diferencijabilna na C. Tada vrijedi:

1. *Funkcija f je konveksna na C ako i samo ako je $\nabla^2 f(\boldsymbol{x})$ pozitivno poludefinitna za svako $\boldsymbol{x} \in C$.*
2. *Ako je $\nabla^2 f(\boldsymbol{x})$ pozitivno definitna za svako $\boldsymbol{x} \in C$ tada je f strogo konveksna na C.*

Komentar: Ako u Teoremu 3.12.6 uzmemo specijalno da je $n = 1$ i da je C otvoreni interval, tvrđenje tada glasi
(a) f je konveksna na C ako i samo ako je $f''(x) \geq 0$, za svako $x \in C$.
(b) Ako je $f''(x) > 0$ za svako $x \in C$ tada je funkcija strogo konveksna.

Primjetimo da u tvrdnji (b) Teorema 3.12.6 vrijedi samo implikacija s lijeva na desno. Da obrat ne vrijedi poslužimo se jednostavnim primjerom funkcije $f : \mathbb{R} \to \mathbb{R}$, zadate sa $f(x) = x^4$. Ona je strogo konveksna na čitavom $\mathbb{R}$, ali $f''(x) = 12x^2 \geq 0$, to jest nemamo pozitivnu definitnost drugog izvoda jer je $f''(0) = 0$.

Poglavlje

4 Ekstremumi funkcija više promjenljivih

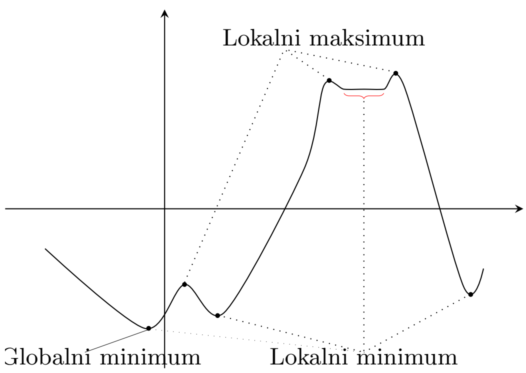
Sada ćemo naš rad iz prethodnih sekcija primjeniti na problem nalaženja minimalne i maksimalne vrijednosti funkcija više varijabli. Primjetit ćemo da je tehnika određivanja ekstremnih vrijednosti funkcije više varijabli veoma slična tehnici koju smo izučavali kod funkcija jedne varijable.

Definicija 4.0.1. *Neka je funkcija* $f : \mathbb{R}^n \to \mathbb{R}$, *definisana na skupu* D_f. *Kažemo da funkcija* f *ima maksimalnu vrijednost* M *u tački* $\boldsymbol{x}_0$, *ako je* $f(\boldsymbol{x}_0) = M$ *i za sve* $\boldsymbol{x} \in D_f$ *vrijedi* $f(\boldsymbol{x}) \leq M$. *Kažemo da funkcija* f *ima minimalnu vrijednost* m *u tački* $\boldsymbol{x}_0$, *ako je* $f(\boldsymbol{x}_0) = m$ *i za sve* $\boldsymbol{x} \in D_f$, *vrijedi* $f(\boldsymbol{x}) \geq m$.

Često maksimalnu i minimalnu vrijednost funkcije, uvedene gornjom definicijom, nazivamo *globalni maksimum* i *globalni minimum*, za razliku od pojmova lokalni maksimum i minimum, koje uvodimo sljedećom definicijom.

Definicija 4.0.2. *Neka je* $f : \mathbb{R}^n \to \mathbb{R}$ *definisana na otvorenom skupu* U. *Kažemo da funkcija* f *ima lokalnu maksimalnu vrijednost* M *u tački* $\boldsymbol{x}_0$, *ako je* $f(\boldsymbol{x}_0) = M$ *i za sve* $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$, *za neko* $r > 0$, *vrijedi* $f(\boldsymbol{x}) \leq M$. *Ukoliko za sve* $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$ *vrijedi* $f(\boldsymbol{x}) < m$ *govorimo o strogom lokalnom maksimumu.*
Kažemo da funkcija f *ima lokalnu minimalnu vrijednost* m *u tački* $\boldsymbol{x}_0$, *ako je* $f(\boldsymbol{x}_0) = m$ *i za sve* $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$, *za neko* $r > 0$, *vrijedi* $f(\boldsymbol{x}) \geq m$. *Ukoliko za sve* $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$ *vrijedi* $f(\boldsymbol{x}) > m$ *govorimo o strogom lokalnom minimumu.*

Često ćemo upotrebljavati i termin *globalni ekstrem* ili *globalna ekstremna vrijednost*, bilo da govorimo o globalnom maksimumu ili globalnom minimumu, a takođe i lokalni ekstrem ili *lokalna ekstremna vrijednost*, kada govorimo o lokalnom maksimumu ili minimumu.



Teorem 4.0.1. ***(Teorem o ekstremnoj vrijednosti)***
Neka je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ neprekidna na nekom otvorenom skupu U. Ako je D ograničen i zatvoren podskup skupa U, tada funkcija f dostiže maksimalnu i minimalnu vrijednost na skupu D.

Sa gornjim teoremom imamo odličan rezultat koji nam govori o egzistenciji ekstremne vrijednosti funkcije, ali ne i kako locirati tu vrijednost. Naš sljedeći posao je pronaći kriterije za lociranje tačaka koje su kandidati u kojima će se postizati ekstremne vrijednosti, a onda i kriterije za njihovu klasifikaciju, to jest da li se u njima postiže ili ne postiže ekstrem i ako se postiže, koja je vrsta ekstrema, maksimum ili minimum.

PRIMJER 35 : Ispitati postojanje ekstremne vrijednosti funkcije $f(x, y) = x^2 + y^2$ na skupu $D = \{(x, y) \in \mathbb{R}^2 |\ x^2 + 4y^2 \leq 4\}$.

Rješenje: Kako je skup D ograničen i zatvoren (unutrašnjost elipse zajedno sa svojim rubom), a funkcija f neprekidna na $\mathbb{R}^2$ (kao polinomijalna funkcija), na osnovu Teorema 4.0.1 zaključujemo da funkcija ima i maksimum i minimum na skupu D.

4.1 Bezuslovna ekstremizacija

Kada radimo ekstremizaciju funkcije $f : \mathbb{R}^n \to \mathbb{R}$, bez ikakvih ograničenja na nezavisnu vaerijablu, to jest kada posmatramo problem

$$\begin{aligned} f(\boldsymbol{x}) &\ \to\ ext \\ \boldsymbol{x} &\in \mathbb{R}^n, \end{aligned} \tag{4.1}$$

tada radimo takozvanu *bezuslovnu* ekstremizaciju. Razmatranje ćemo podjeliti na egzistenciju i nalaženje lokalnog i globalnog ekstrema funkcija više varijabli.

4.1.1 Nalaženje lokalnog ekstrema

Za početak, posmatrajmo funkciju $f : \mathbb{R}^n \to \mathbb{R}$ koja je diferencijabilna na otvorenom skupu U i koja ima ekstrem u tački $\boldsymbol{x}_0$. Neka je $\boldsymbol{u}$ proizvoljan jedinični vektor, tada će očigledno, funkcija $\varphi : \mathbb{R} \to \mathbb{R}$, definisana sa

$$\varphi(t) = f(\boldsymbol{x}_0 + t\boldsymbol{u})\ ,$$

takođe imati ekstremnu vrijednost i to upravo za $t = 0$. Kako je φ funkcija jedne varijable, to onda mora biti

$$\varphi'(0) = 0\ . \tag{4.2}$$

Ali u sekciji 3.1 smo vidjeli da ovaj izvod nije ništa drugo do izvod funkcije f u pravcu vektora $\boldsymbol{u}$, to jest

$$\varphi'(0) = D_{\boldsymbol{u}} f(\boldsymbol{x}_0)\ . \tag{4.3}$$

Na osnovu veze izvoda u pravcu i gradijenta zaključujemo da vrijedi,

$$\varphi'(0) = D_{\boldsymbol{u}} f(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{u} = 0\ .$$

Skalarni produkt jednak je nuli ako je bar jedan od vektora tog produkta nula-vektor ili ako su vektori ortogonalni. Ortogonalnost otpada jer gornje vrijedi za proizvoljan jedinični vektor $\boldsymbol{u}$. Koristeći proizvoljnost vektora $\boldsymbol{u}$, uzmimo specijalno vektore baze. Tada imamo

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{e}_i = \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = 0\ ,\ i = 1, 2, ..., n.$$

Ovo nam govori da mora biti $\nabla f(\boldsymbol{x}_0) = \boldsymbol{0}$. Primjetimo da ovo znači i to da je nagib grafa funkcije f jednak 0 u tački $\boldsymbol{x}_0$, u pravcima svih baznih vektora. Međutim, to znači mnogo više. Naime, nagib grafa je 0 u svim pravcima $\boldsymbol{u}$ jer je $D_{\boldsymbol{u}} f(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{u}$. Ovo razmatranje sumiramo teoremom.

Teorem 4.1.1 (Neophodan uslov za ekstrem 1).
Neka je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna na otvorenom skupu U i neka ima lokalnu ekstremnu vrijednost u tački $\boldsymbol{x}^* \in U$, tada je $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$.

Kako je totalni diferencijal funkcije jednak umnošku gradijenta i diferencijala argumenta, to jest

$$df(\boldsymbol{x}) = \frac{\partial f}{\partial x_1}(\boldsymbol{x})dx_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{x})dx_2 + \cdot \frac{\partial f}{\partial x_n}(\boldsymbol{x})dx_n = \nabla f(\boldsymbol{x}) \cdot d\boldsymbol{x} \ ,$$

to onda za diferencijabilnu funkciju koja ima ekstremnu vrijednost u tački $\boldsymbol{x}_0$, vrijedi

$$df(\boldsymbol{x}_0) = 0 \ ,$$

a takvu situaciju smo imali i kod funkcije jedne varijable jer je neophodan uslov bio $f'(x) = 0$, a vrijedilo je $df(x) = f'(x)dx$. Teorem 4.1.1 nam daje kandidate za ekstreme diferencijabilne funkcije. U opštem slučaju nediferencijabilne funkcije kandidata može biti i više.

Teorem 4.1.2 (Neophodan uslov za ekstrem 2).
Ako funkcija $f : \mathbb{R}^n \to \mathbb{R}$ ima ekstrem u tački $\boldsymbol{x}^*$, tada vrijedi, ili je $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ ili prvi parcijalni izvodi funkcije u tački $\boldsymbol{x}^*$ ne postoje.

Dakle, kandidati za ekstremnu vrijednost proizvoljne funkcije su sve one tačke u kojima je gradijent jednak $\mathbf{0}$ i sve one u kojima funkcija nije diferencijabilna. Ovo nas navodi da ove tačke definišemo precizno.

Definicija 4.1.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u tački $\boldsymbol{x}_0$ i neka je $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$. Tada tačku $\boldsymbol{x}_0$ nazivamo stacionarnom tačkom funkcije f.*
Tačke u kojima funkcija f nije diferencijabilna, nazivamo singularnim tačkama funkcije f.

Često se za obje gore pomenute vrste tačaka kaže da su *kritične tačke* funkcije.

PRIMJER 36 : Funkcija $f(x, y) = x^2 + y^2$ je diferencijabilna funkcija na $\mathbb{R}^2$ i $\nabla f(x, y) = (2x, 2y)$. Jedine kandidate za ekstremne vrijednosti dobijamo rješavanjem sistema

$$\begin{aligned} 2x &= 0 \\ 2y &= 0 \end{aligned}$$

Dakle, jedina kritična tačka je stacionarna tačka $\boldsymbol{x}_0 = (0, 0)$.

PRIMJER 37 : Funkcija $f(x, y) = 1 - x^2 - y^2$ je diferencijabilna i $\nabla f(x, y) = (-2x, -2y)$. Rješavanjem sistema

$$\begin{aligned} -2x &= 0 \\ -2y &= 0 \end{aligned}$$

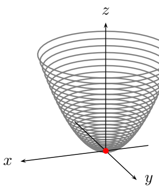
dobijamo stacionarnu tačku $\boldsymbol{x}_0 = (0, 0)$.

PRIMJER 38 : Za funkciju $f(x, y) = \sqrt{x^2 + y^2}$, gradijent je

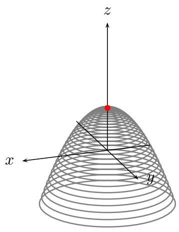
$$\nabla f(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right) \ .$$

Parcijalni izvodi ne postoje u tački $\boldsymbol{x}_0 = (0, 0)$ i to je jedina kritična tačka funkcije f, i ona je singularna tačka.

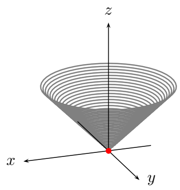
Primjer 39 : Funkcija $f(x, y) = x^2 - y^2$ ima gradijent $\nabla f(x, y) = (2x, -2y)$, pa je jedina kritična tačka, stacionarna tačka $\boldsymbol{x}_0 = (0, 0)$.



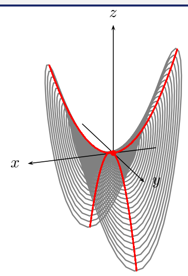
$f(x, y) = x^2 + y^2$
Minimum u (0,0)
$\nabla f(0, 0) = (0, 0)$



$f(x, y) = 1 - x^2 - y^2$
Maksimum u (0,0)
$\nabla f(0, 0) = (0, 0)$



$f(x, y) = \sqrt{x^2 + y^2}$
Minimum u (0,0)
f_x i f_y ne postoje



$f(x, y) = x^2 - y^2$
Nema ekstrema,
$\nabla f(0, 0) = (0, 0)$

Sada kada smo u mogućnosti utvrditi postojanje ekstremne vrijednosti funkcije (Teorem 4.0.1) i identifikovati kandidate za te vrijednosti (Teorem 4.1.2) ostaje nam pronaći kriterije za utvrđivanje da li ti kandidati jesu ekstremi i klasificirati ih. Prisjetimo se opet funkcija jedne varijable, da je jedan od kriterija za identifikaciju lokalnih ekstrema bio test drugog izvoda. Naime, ako je c bila stacionarna tačka funkcije $\varphi : \mathbb{R} \to \mathbb{R}$, tada ako je $\varphi''(c) > 0$, funkcija je imala minimum u c, a ako je $\varphi''(c) < 0$, funkcija je imala maksimum u tački c. Taylorov polinom nam na najvidljiviji način pokazuje zašto je to tako. Naprimjer, neka je c stacionarna tačka funkcije φ i neka je $\varphi''(c)$ neprekidna na otvorenom intervalu koji sadrži c, i neka je $\varphi''(c) > 0$. Tada za neko $\varepsilon > 0$, postoji interval $I = (c - \varepsilon, c + \varepsilon)$ na kome je $\varphi''(c)$ neprekidna i $\varphi''(t) > 0$, za sve $t \in I$. Na osnovu Taylorove teoreme, za proizvoljno h, takav da je $|h| < \varepsilon$, postoji $s \in (c, c + h)$, takav da je

$$\varphi(c + h) = \varphi(c) + \varphi'(c)h + \frac{1}{2}\varphi''(s)h^2 \ . \tag{4.4}$$

Zbog stacionarnosti je $\varphi'(c) = 0$. Takođe smo imali $\varphi''(s) > 0$, pa koristeći to u (4.4) dobijamo da je za proizvoljno h, takav da je $|h| < \varepsilon$, zadovoljeno

$$\varphi(c + h) > \varphi(c) \ ,$$

a ovo znači da je u tački c lokalni minimum.

Veoma slično razmatranje sada možemo sprovesti i za funkciju $f : \mathbb{R}^n \to \mathbb{R}$. Na osnovu Teorema 3.9.1 znamo da vrijedi formula

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a})^T \cdot \boldsymbol{h} + \frac{1}{2}\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a} + \xi\boldsymbol{h}) \cdot \boldsymbol{h} \ ,$$

gdje je f dva puta neprekidno diferencijabilna funkcija u nekoj kugli $B(\boldsymbol{a}, r)$ i $\xi \in (0, 1)$. Neka je sada $\boldsymbol{a}$ stacionarna tačka funkcije f i neka je Hesijan $\nabla^2 f(\boldsymbol{x})$ pozitivno definitna matrica u kugli $B(\boldsymbol{a}, r)$. Tada je $\nabla f(\boldsymbol{a}) = \boldsymbol{0}$, pa vrijedi

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + \frac{1}{2}\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a} + \xi\boldsymbol{h}) \cdot \boldsymbol{h} \ .$$

Kako je još $\boldsymbol{a} + \xi\boldsymbol{h} \in B(\boldsymbol{a}, r)$ (za $||\boldsymbol{h}|| \leq r$), zbog pozitivne definitnosti hesijana je $\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a} + \xi\boldsymbol{h}) \cdot \boldsymbol{h} > 0$, te imamo

$$f(\boldsymbol{a} + \boldsymbol{h}) > f(\boldsymbol{a}) \ ,$$

u pravcu proizvoljnog vektora $\boldsymbol{h}$, za koga je $||\boldsymbol{h}|| < r$. Ali ovo onda upravo znači da funkcija f ima lokalni minimum u tački $\boldsymbol{a}$.

Istim argumentima bi rezonovali da smo pretpostavili negativnu definitnost Hesijana i naravno, zaključili bi da funkcija ima lokalni maksimum u tački $\boldsymbol{a}$.

Ako je Hesijan indefinitan, to bi značilo da postoji vektor $\boldsymbol{h}$, takav da je kvadratna forma

$$\boldsymbol{h}^T \cdot \nabla^2 f(\boldsymbol{a} + \xi\boldsymbol{h}) \cdot \boldsymbol{h} > 0 \ ,$$

i postoji vektor $\boldsymbol{k}$ takav da je

$$\boldsymbol{k}^T \cdot \nabla^2 f(\boldsymbol{a} + \xi\boldsymbol{k}) \cdot \boldsymbol{k} < 0 \ .$$

Ovo bi onda uzrokovalo da za neki vektor $\boldsymbol{h}$ vrijedi $f(\boldsymbol{a}+\boldsymbol{h}) > f(\boldsymbol{a})$, a istovremeno za neki drugi vektor $\boldsymbol{h}$ je $f(\boldsymbol{a}+\boldsymbol{h}) < f(\boldsymbol{a})$. U ovom slučaju jasno je da u tački $\boldsymbol{a}$ ne može biti niti lokalni minimum niti lokalni maksimum. Tada tačka $\boldsymbol{a}$ predstavljala takozvanu *sedlastu tačku* funkcije f. Na osnovu gornjeg rečenog sada možemo iskazati dovoljne uslove za ekstremnu vrijednost funkcije $f : \mathbb{R}^n \to \mathbb{R}$.

Teorem 4.1.3 (Test druge derivacije).
Neka je $f : \mathbb{R}^n \to \mathbb{R}$ i $f \in C^2(U)$, gdje je U otvoren skup. Ako je $\boldsymbol{x}^* \in U$ stacionarna tačka funkcije f, tada je

1. *$f(\boldsymbol{x}^*)$ lokalni minimum funkcije f, ako je $\nabla^2 f(\boldsymbol{x}^*)$ pozitivno definitna matrica.*
2. *$f(\boldsymbol{x}^*)$ lokalni maksimum funkcije f, ako je $\nabla^2 f(\boldsymbol{x}^*)$ negativno definitna matrica.*
3. *tačka $\boldsymbol{x}^*$ sedlasta tačka funkcije f, ako je $\nabla^2 f(\boldsymbol{x}^*)$ indefinitna matrica.*

Ukoliko je $\nabla^2 f(\boldsymbol{x}^*)$ nedefinitna matrica, potrebna su dodatna ispitivanja za klasifikaciju tačke $\boldsymbol{x}^*$.

PRIMJER 40 : Odrediti lokalne ekstremne vrijednosti funkcije $f(x,y) = xye^{-x^2-y^2}$.

Rješenje: Nalazimo prvo gradijent

$$\nabla f(x,y) = e^{-x^2-y^2}\left(y - 2x^2y, x - 2xy^2\right) .$$

Kako je $e^{-x^2-y^2} > 0$ za sve $(x,y) \in \mathbb{R}^2$, činjenica da je $\nabla f(x,y) = (0,0)$, svodi se na sistem

$$\begin{aligned} y(1-2x^2) &= 0 , \\ x(1-2y^2) &= 0 . \end{aligned}$$

Prva jednačina će biti tačna ako je $y = 0$ ili $x = -\frac{1}{\sqrt{2}}$ ili $x = \frac{1}{\sqrt{2}}$. Ako je $y = 0$, onda iz druge jednačine vidimo da mora biti i $x = 0$, a time smo dobili prvu stacionarnu tačku $M_1(0,0)$. Ako je $x = \pm\frac{1}{\sqrt{2}}$, onda je druga jednačina zadovoljena ako je $1 - 2y^2 = 0$, odnosno ako je $y = \frac{1}{\sqrt{2}}$ ili $y = -\frac{1}{\sqrt{2}}$, pa na taj način dobijamo još četiri stacionarne tačke: $M_2\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $M_3\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $M_4\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ i $M_5\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$.

Drugi korak u rješavanju problema ovog tipa je određivanje hesijana funkcije

$$\nabla^2 f(x,y) = e^{-x^2-y^2}\begin{bmatrix} 4x^3y - 6xy & 4x^2y^2 - 2x^2 - 2y^2 + 1 \\ 4x^2y^2 - 2x^2 - 2y^2 + 1 & 4y^3yx - 6xy \end{bmatrix} .$$

Sada nakon kraćeg računa dobijamo

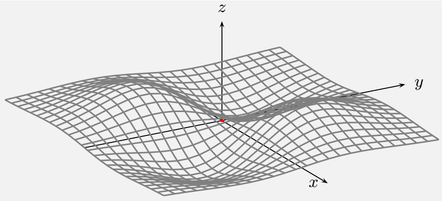
$$\nabla^2 f(\boldsymbol{M}_2) = \nabla^2 f(\boldsymbol{M}_5) = e^{-1}\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} .$$

Kako je $A_1 = -2e^{-1} < 0$ i

$$A_2 = det\begin{bmatrix} -2e^{-1} & 0 \\ 0 & -2e^{-1} \end{bmatrix} = 4e^{-1} > 0 ,$$

na osnovu testa druge derivacije zaključujemo da funkcija u tačkama $\boldsymbol{M}_2$ i $\boldsymbol{M}_5$ ima lokalni maksimum, i pri tome je $f(\boldsymbol{M}_2) = f(\boldsymbol{M}_5) = f_{\max} = \frac{1}{2}e^{-1}$.

Dalje imamo



$$\nabla^2 f(\boldsymbol{M}_3) = \nabla^2 f(\boldsymbol{M}_4) = e^{-1}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} .$$

Sada je $A_1 = 2e^{-1} > 0$ i

$$A_2 = det\begin{bmatrix} 2e^{-1} & 0 \\ 0 & 2e^{-1} \end{bmatrix} = 4e^{-1} > 0 ,$$

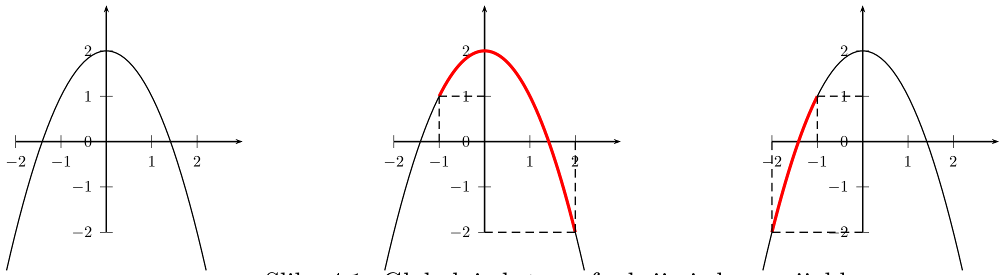
pa opet na osnovu testa druge derivacije zaključujemo da funkcija u tačkama $\boldsymbol{M}_3$ i $\boldsymbol{M}_4$ ima lokalni minimum i pri tome je $f(\boldsymbol{M}_3) = f(\boldsymbol{M}_4) = f_{\min} = -\frac{1}{2}e^{-1}$. Ostala nam je još taka $\boldsymbol{M}_1(0,0)$ u kojoj je

$$\nabla^2 f(\boldsymbol{M}_1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

Sada je $A_1 = 0$ i $A_2 = -1$, pa je Hesijan indefinitan, a to znači da je tačka $M_1(0,0)$ sedlasta tačka.

4.1.2 Nalaženje globalnog ekstrema

Posmatrajmo funkciju $f(x) = 2 - x^2$. Posmatramo li je na čitavom $\mathbb{R}$, ona ima lokalni maksimum u tački $x = 0$, koji je i globalni maksimum, a globalnog minimuma nema (slika lijevo). Ako je posmatramo na skupu $[-1, 2]$ i dalje je globalni maksimum u $x = 0$, ali sada je globalni minimum u tački $x = 2$ (slika u sredini). Ako je posmatramo za vrijednosti iz $[-2, -1]$, njen globalni minimum je u $x = -2$, a globalni maksimum je u $x = -1$ (slika desno). Dakle, globalni ekstrem funkcije direktno zavisi od područja na kom tu funkciju posmatramo.



Slika 4.1: Globalni ekstrem funkcije jedne varijable

Nešto slično imamo i kod funkcija više promjenljivih.

Primjer 41 : Odrediti globalne ekstreme funkcije $f(x, y) = x^2 + y^2$ na skupu

$$D = \{(x, y)|\ x^2 + 4y^2 \leq 4\} .$$

Rješenje: Skup D je zatvoren i ograničen, pa na osnovu Teoreme 4.0.1, funkcija f dostiže svoju najmanju i najveću vrijednost.
Kako je funkcija diferencijabilna (kao polinomijalna funkcija), njene jedine kritične tačke su stacionarne tačke, koje dobijamo iz uslova

$$\nabla f(x, y) = (2x, 2y) = 0 .$$

U tački $(0, 0)$ je moguć lokalni ekstrem funkcije ali moramo sada posmatrati šta se događa sa našom funkcijom na rubu oblasti D, to jest na skupu

$$\partial D = \{(x, y)|\ x^2 + 4y^2 = 4\} .$$

S obzirom da su na ∂D nezavisne varijable vezane relacijom $x^2+4y^2 = 4$, uvodeći polarne koordinate, to jest smjene $x(t) = 2\cos t$ i $y(t) = \sin t$, gje je $t \in [0, 2\pi]$, naša funkcija f postaje funkcija jedne varijable

$$\begin{aligned} g(t) = f(x(t), y(t)) &= f(2\cos t, \sin t) \\ &= 4\cos^2 t + \sin^2 t \\ &= 3\cos^2 t + 1 , \end{aligned}$$

gdje je $t \in [0, 2\pi]$. Ekstremne vrijednosti funkcije g će biti i ekstremne vrijednosti funkcije f. Zato posmatrajmo jednačinu

$$g'(t) = -6\cos t \sin t = 0 \ .$$

Stacionarne tačke će biti $t = 0$, $t = \frac{\pi}{2}$, $t = \pi$, $t = \frac{3\pi}{2}$ i $t = 2\pi$. Dakle, pored tačke $(0, 0)$ imamo još četiri kandidata za globalni ekstrem, a to su tačke $(2, 0)$, $(0, 1)$, $(-2, 0)$ i $(0, -1)$, određene gornjim vrijednostima za t.
Izračunavajući sada vrijednost funkcije u svakoj od ovih pet tačaka, određujemo globalne ekstreme.

$$f(0, 0) = 0 \ , \ f(2, 0) = 4 \ , \ f(0, 1) = 1 \ , \ f(-2, 0) = 4 \ , \ f(0, -1) = 1 \ .$$

Upoređujući gornje vrijednosti, zaključujemo da funkcija f ima globalnu maksimalnu vrijednost 4 u tačkama $(2, 0)$ i $(-2, 0)$ i globalnu minimalnu vrijednost 0 u tački $(0, 0)$.

Kao što nam pokazuje upravo urađeni primjer, za funkciju zadatu na zatvorenoj i ograničenoj oblasti određivanje globalnih ekstrema se svodi na to da pronađemo lokalne ekstreme i ekstreme funkcije na rubu te oblasti, a onda određujemo šta će od njih biti globalne ekstremne vrijednosti. Ako funkciju ne posmatramo na zatvorenoj i ograničenoj oblasti, onda se problem određivanja globalnih ekstrema svodi na to da pronađemo lokalne ekstreme, a onda nekom metodom ispitamo da li su oni ujedno i globalni ekstremi. Posmatrajmo sljedeći primjer.

Primjer 42 : U nekoj firmi žele da naprave pravougaonu posudu bez krova, zapremine 500 m^3 i da pritome utroše što je manje moguće materijala.

Rješenje: Označimo sa x i y dužine stranica te posude u osnovi i sa z visinu te posude (sve veličine su izražene u metrima), tada u stvari treba pronaći minimalnu vrijednost funkcije

$$M'(x, y, z) = xy + 2xz + 2yz \ ,$$

pri čemu zapremina mora biti $xyz = 500$. Izražavajući z iz ove jednakosti i uvrštavanjem u funkciju M', dobijamo funkciju

$$M(x, y) = xy + \frac{1000}{y} + \frac{1000}{x} \ ,$$

kojoj treba odrediti minimalnu vrijednost na beskonačnom pravougaoniku

$$R = \{(x, y)| \ x > 0 \ , \ y > 0\} \ .$$

Rješavanjem sistema

$$\begin{aligned} \frac{\partial M}{\partial x} &= y - \frac{1000}{x^2} = 0 \\ \frac{\partial M}{\partial y} &= x - \frac{1000}{y^2} = 0 \end{aligned}$$

dobijamo jedinu stacionarnu tačku $A(10, 10)$.
Hesijan funkcije M glasi

$$\nabla^2 M(x, y) = \begin{bmatrix} \frac{2000}{x^3} & 1 \\ 1 & \frac{2000}{y^3} \end{bmatrix} \ ,$$

odnosno

$$\nabla^2 M(10, 10) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \ .$$

Kako je $det(\nabla^2 M(10, 10)) = 3$, zaključujemo da je Hesijan pozitivno definitan, a to znači da funkcija M ima lokalni minimum u tački $A(10, 10)$ i pri tome je

$$M_{\min} = M(10, 10) = 10 \cdot 10 + \frac{1000}{10} + \frac{1000}{10} = 300 \ .$$

Ostaje nam ispitati da li je ovo i globalni minimum funkcije M?
Ako je bilo koja varijabla manja od jedan, to jest $0 < x < 1$ ili $0 < y < 1$, tada je $\frac{1000}{x} > 1000$, odnosno $\frac{1000}{y} > 1000$, pa je očigledno vrijednost funkcije M veća od 300. Ako je sada $x \geq 400$ i $y \geq 1$, onda

je $xy \geq 400$, pa bi opet vrijednosti naše funkcije bile veće od 300 (analogno i slučaj $y \geq 400$ i $x \geq 1$). Dakle, ako posmatramo skup

$$D = \{(x,y) |\ 1 \leq x \leq 400\ ,\ 1 \leq y \leq 400\}\ ,$$

izvan skupa D vrijednosti funkcije M su veće od 300. Na skupu D naša funkcija ima globalni minimum u tački $(10,10)$, pa je to onda očigledno globalni minimum funkcije na čitavom skupu R.

Ostaje samo još zaključiti da je tada

$$z = \frac{500}{10 \cdot 10} = 5\ ,$$

odnosno da posuda treba biti dimenzija $10 \times 10 \times 5$, da bi imala odgovarajuću zapreminu i da bi imali minimalne troškove.

4.1.3 Ekstremizacija konveksnih funkcija

Istaknimo neke specifične rezultate vezane za ekstremizaciju konveksnih funkcija.

Teorem 4.1.4. *Neka je $\boldsymbol{x}^*$ lokalni ekstrem konveksne funkcije $f : D_f \to \mathbb{R}$, definisane na konveksnom skupu D_f. Tada je $\boldsymbol{x}^*$ globalni ekstrem funkcije f.*

Dokaz : Neka je $\boldsymbol{x}^*$ lokalni minimum funkcije f. Pretpostavimo da to nije globalni minimum. To znači da postoji $\boldsymbol{y} \in D_f$ takav da je $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$. Neka je $\lambda \in (0,1)$. Kako je D_f konveksan skup, onda $\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{x}^*$ pripada skupu D_f, a kako je f konveksna to je onda

$$f(\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{x}^*) \leq \lambda f(\boldsymbol{y}) + (1-\lambda)\boldsymbol{x}^* < \lambda f(\boldsymbol{x}^*) + (1-\lambda) f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)\ .$$

Ovo znači da je za proizvoljno $\lambda \in (0,1)$, $f(\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{x}^*) = f(\boldsymbol{x}^* + \lambda(\boldsymbol{y} - \boldsymbol{x}^*)) < f(\boldsymbol{x}^*)$. Jasno je da će za dovoljno malo λ ovo biti u suprotnosti sa činjenicom da je $\boldsymbol{x}^*$ lokalni minimum. Dakle, polazna pretpostavka je neodrživa. ■

Teorem 4.1.5. *Neka je f strogo konveksna funkcija. Ako je $\boldsymbol{x}^*$ lokalni ekstrem funkcije, tada je on strogi globalni ekstrem funkcije.*

Teorem 4.1.6. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna i diferencijabilna na $\mathbb{R}^n$. Tačka $\boldsymbol{x}^* \in \mathbb{R}^n$ je globalni ekstrem funkcije f na $\mathbb{R}^n$ ako i samo ako je $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$.*

Dokaz : ($\Longrightarrow$) Ovo je dokazano u Teoremu 4.1.1,

($\Longleftarrow$) Neka je $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$. Zbog konveksnosti funkcije imamo da za proizvoljan $\boldsymbol{y} \in \mathbb{R}^n$ vrijedi

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^T \cdot (\boldsymbol{y} - \boldsymbol{x}^*)\ .$$

Zbog naše pretpostavke zaključujemo da je $f(\boldsymbol{y}) \geq f(\boldsymbol{x}^*)$, za proizvoljno $\boldsymbol{y} \in \mathbb{R}^n$. Dakle, $\boldsymbol{x}^*$ je globalni minimum. ■

Uslov $\nabla f(\boldsymbol{x}) = \mathbf{0}$ je neophodan uslov za lokalni minimum u opštem slučaju. Dakle, uz konveksnost funkcije to je i dovoljan uslov.

4.1.4 Uslovna ekstremizacija

Primjer 42 ima neke sličnosti sa Primjerom 41. Naime, u oba primjera smo nalazili ekstremne vrijednosti funkcije pod restrikcijom na podskup koji je manje dimenzije. U prvom primjeru smo ekstremizirali funkciju $f(x,y) = x^2 + y^2$ sa restrikcijom na jednodimenzionalnoj elipsi $x^2 + 4y^2 = 4$. U drugom primjeru smo ekstremizirali funkciju tri varijable $M'(x,y,z) = xy + 2xz + 2yz$ sa restrikcijom na trodimenzionalnu površ $xyz = 500$.
U prvom smo primjeru problem riješili tako što smo parametrizovali elipsu, a zatim smo ekstremizirali

funkciju jedne varijable. U drugom smo izrazili z kao funkciju od x i y, a zatim smo ekstremizirali funkciju dvije varijable. U ovoj sekciji ćemo dati generalni metod za rješavanje oba ova ali i drugih sličnih problema.

U osnovnom slučaju ekstremizacija, zadata je neka (diferencijabilna) funkcija $f : \mathbb{R}^n \to \mathbb{R}$ za koju želimo naći ekstremne vrijednosti. Taj problem smo rješavali nalaženjem svih kritičnih tačaka funkcije, a onda testom druge derivacije ispitivali karakter tih tačaka. Međutim, kao što smo vidjeli u Primjeru 42, nekada treba izvršiti ekstremizaciju funkcije, pri čemu su nezavisne varijable te funkcije vezane nekim uslovom, to jest tražimo ekstremnu vrijednost funkcije $f(\boldsymbol{x})$, pod nekim uslovom na varijablu $\boldsymbol{x}$. Ovo izražavamo u sljedećoj formi:

$$\begin{aligned} f(\boldsymbol{x}) &\to ext \\ \boldsymbol{x} &\in S, \end{aligned} \tag{4.5}$$

gdje je $f : \mathbb{R}^n \to \mathbb{R}$, a $S \subset \mathbb{R}^n$.

Skup S je zadat nekom jednačinom ili sistemom jednačina, nekom nejednačinom ili sistemom nejednačina, a može biti zadat i sistemom jednačina i nejednačina. Funkciju f nazivamo *ciljna funkcija*, a skup S nazivamo *dopustivi skup*. Ovakvu vrstu ekstremizacije nazivamo *uslovna ekstremizacija*.

PRIMJER 43 : Neka treba odrediti minimum funkcije $z = x^2 + y^2$ pri uslovu $x + y = 1$, to jest rješavamo problem uslovne ekstremizacije

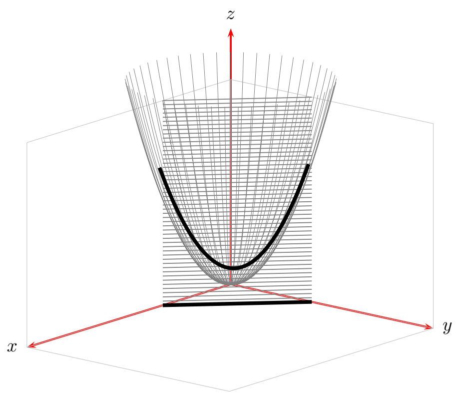
$$\begin{aligned} \min \ & x^2 + y^2 \\ & x + y = 1 \ . \end{aligned}$$

Očigledni minimum funkcije, bez uslova $x+y = 1$, je u tački $(0, 0)$ i vidimo da ta tačka ne zadovoljava uslov $x + y = 1$. Šta geometrijski predstavlja uslov u gornjem problemu?

Graf funkcije $z = x^2 + y^2$ je paraboloid, a uslov $x + y = 1$ predstavlja jednačinu ravni u $\mathbb{R}^3$. Dakle, mi tražimo minimalnu vrijednost na paraboloidu ali samo u onim tačkama u kojima se sijeku paraboloid (ciljna funkcija) i ravan (uslovna funkcija). Sa slike vidimo da se traži minimum funkcije koja predstavlja parabolu u prostoru $\mathbb{R}^3$. Zaista, koristeći uslovnu funkciju, možemo izraziti jednu varijablu, npr. $y = 1 - x$, pa stavljajući to u izraz ciljne funkcije imamo,

$$z = x^2 + (1 - x)^2 = 2x^2 - 2x + 1 \ ,$$

a ovo je zaista jednačina parabole. Sada minimum ove funkcije nalazimo kao problem ekstremizacije funkcije jedne vaerijable. $z' = 4x - 2$, pa imamo jednu stacionarnu tačku $x_0 = \frac{1}{2}$. Kako je $z'' = 4$, dakle pozitivan, to u tački x_0 funkcija ima minimum. Izračunavajući $y_0 = 1 - x_0 = \frac{1}{2}$, zaključujemo da funkcija $z = x^2 + y^2$ ima minimum u tački $(\frac{1}{2}, \frac{1}{2})$, pri uslovu $x + y = 1$.

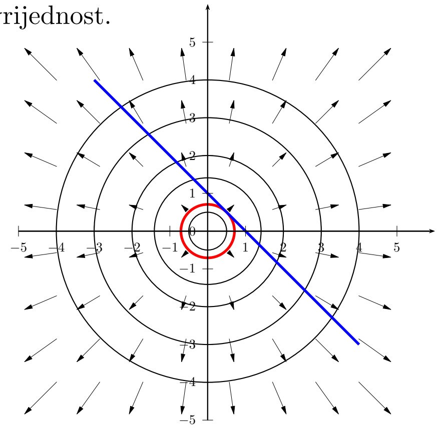


Slika 4.2: Uslovni ekstrem

Gornji primjer nam daje jedan metod za rješavanje problema uslovne ekstremizacije, ali jasno je da će primjena ovog metoda biti kudikamo složenija, za malo složenije uslovne funkcije. Zato nam je u interesu imati i neki drugi metod, a najopštiji od svih je tzv. *Lagrangeov metod*, koga ćemo sada izložiti.

Šta će biti motivacija za ovaj metod? Posmatrajmo ponovo gornji primjer i konturnu sliku grafova ciljne i uslovne funkcije (slika 4.3).

Nivo linije funkcije $z = x^2 + y^2$ predstavljaju koncentrične centralne kružnice ($x^2 + y^2 = k$), a uslovna funkcija zbog svog položaja (ortogonalna na xOy ravan), predstavljena je pravom linijom u xOy ravni. Na slici uočavamo da prava neke od konturnih linija siječe, neke nivo linije neće uopšte sjeći ali da samo jednu nivo liniju dodiruje. Strelice na slici nam pokazuju pravce rasta ciljne funkcije (gradijentni vektor u različitim tačkama), a time je onda određeno da nivo linije paraboloida bliže koordinatnom početku, odgovaraju manjim vrijednostima funkcije (u opštem slučaju ovo nije pravilo). Ovo onda znači da upravo ona nivo linija koja se dodiruje sa uslovnom funkcijom predstavlja bitan momenat. Naime, tačke na onim nivo linijama koje se ne sijeku sa uslovnom funkcijom i ne mogu biti kandidati za uslovne ekstreme, a jasno je da od momenta kada prava presječe jednu od nivo linija, sjeći će i svaku "veću" nivo liniju, pa dakle tu i nemožemo tražiti konačnu ekstremnu vrijednost.



Slika 4.3: Nivo linije funkcije $z = x^2 + y^2$ sa uslovnom funkcijom $x + y = 1$

Naravno, tražiti nivo liniju zadate površi koja će dodirivati uslovnu funkciju ne bi bio lagan posao. Zato se prisjetimo da je ugao između dvije krive koje se sijeku, jednak uglu između njihovih tangenti u presječnoj tački. Dakle, ako se dvije linije dodiruju, onda se njihove tangente u dodirnoj tački poklapaju, ili drugačije iskazano, vektori normala na tim tangentama su paralelni. Kako je gradijentni vektor upravo onaj vektor koji je ortogonalan na nivo liniju u proizvoljnoj tački, a uslov paralelnosti vektora je uslov njihove kolinearnosti, zaključujemo da mi treba da odredimo upravo one tačke $(x, y) \in \mathbb{R}^2$ u kojima vrijedi

$$\nabla(x^2 + y^2) = \lambda\nabla(x + y - 1) \; .$$

Zbog paralelnog pomjeranja, takvih vektora bi bilo beskonačno mnogo. Međutim, mi tražimo tačke na uslovnoj krivoj koje to zadovoljavaju, to jest nalazimo tačke (x, y) koje zadovoljavaju

$$\nabla(x^2 + y^2) = \lambda\nabla(x + y - 1) \;\; \text{i}$$

$$x + y - 1 = 0 \; .$$

Generalno, ako rješavamo problem

$$f(\boldsymbol{x}) \;\longrightarrow\; ext$$

$$g(\boldsymbol{x}) = 0 \; ,$$

rješenje će biti u onim tačkama $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ u kojima su zadovoljeni uslovi

$$\nabla f(\boldsymbol{x}) = \lambda\nabla g(\boldsymbol{x}) \tag{4.6}$$

$$g(\boldsymbol{x}) = 0 \; . \tag{4.7}$$

Izloženi metod se naziva *Lagrangeov metod*, a nova varijabla $\lambda \in \mathbb{R}$ koja se pojavljuje u uslovu (4.6), naziva se *Lagrangeov multiplikator*. Ako uvedemo funkciju

$$\Lambda(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x}) \ , \tag{4.8}$$

koju nazivamo *Lagrangeova funkcija* ili *lagranžijan*, nije teško uočiti da su uslovi (4.6) i (4.7), ekvivalentni uslovu

$$\nabla\Lambda(\boldsymbol{x}, \lambda) = \mathbf{0} \ . \tag{4.9}$$

Zaista, nalazeći parcijalne izvode po promjenljivima x_i $(i = 1, 2, ..., n)$ imamo

$$\frac{\partial \Lambda}{\partial x_i} = \frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} \ , \ i = 1, 2, ..., n \ .$$

Sada zbog (4.9), zaključujemo da je

$$\frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} = 0 \ ,$$

za sve $i \in \{1, 2, ..., n\}$, to jest vrijedi uslov (4.6).

Kako je

$$\frac{\partial \Lambda}{\partial \lambda} = -g(\boldsymbol{x}) \ ,$$

opet zbog (4.9) imamo

$$g(\boldsymbol{x}) = 0 \ ,$$

odnosno uslov (4.7).

Na ovaj način smo praktično dali i opis postupka rješavanja uslovne ekstremizacije oblika

$$f(\boldsymbol{x}) \longrightarrow ext$$

$$g(\boldsymbol{x}) = 0 \ .$$

1. Prvo formiramo lagranžijan

$$\Lambda(x_1, x_2, ..., x_n, \lambda) = f(x_1, x_2, ..., x_n) - \lambda g(x_1, x_2, ..., x_n) \ ,$$

2. Određujemo $\nabla\Lambda(x_1, x_2, ..., x_n, \lambda)$.

3. Rješavamo jednačinu $\nabla\Lambda(x_1, x_2, ..., x_n, \lambda) = \mathbf{0}$, to jest sistem

$$\begin{array}{ccl} \dfrac{\partial \Lambda}{\partial x_1} & = & \dfrac{\partial f}{\partial x_1} - \lambda \dfrac{\partial g}{\partial x_1} = 0 \\ \dfrac{\partial \Lambda}{\partial x_2} & = & \dfrac{\partial f}{\partial x_2} - \lambda \dfrac{\partial g}{\partial x_2} = 0 \\ \cdots\cdots & .. & \cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial \Lambda}{\partial x_n} & = & \dfrac{\partial f}{\partial x_n} - \lambda \dfrac{\partial g}{\partial x_n} = 0 \\ \dfrac{\partial \Lambda}{\partial \lambda} & = & -g(x_1, x_2, ..., x_n) = 0 \ . \end{array}$$

Rješenja posljednjeg sistema su stacionarne tačke lagranžijana i ostaje nam još samo utvrditi karakter tih tačaka.

Primjetimo odma, da će u pronađenim stacionarnim tačkama $\boldsymbol{x}^*$, biti

$$\Lambda(\boldsymbol{x}^*, \lambda) = f(\boldsymbol{x}^*) \ ,$$

(jer je $g(\boldsymbol{x}^*) = 0$), to jest ekstremi lagranžijana ujedno su i ekstremi naše ciljne funkcije. Zato za ispitivanje karaktera tih tačaka možemo primjeniti test druge derivacije ili ispitivanjem drugog diferencijala ciljne funkcije. Naime, ako je $d^2 f(\boldsymbol{x}^*) > 0$, imamo minimum, a ako je $d^2 f(\boldsymbol{x}^*) < 0$ imamo maksimum ciljne funkcije sa zadatim uslovom. Ako je $d^2 f(\boldsymbol{x}^*) = 0$, potrebna su dodatna ispitivanja za određivanje karaktera te tačke.

Primjer 44 : Riješiti problem

$$f(x,y) = x^2 + y^2 \longrightarrow ext$$
$$x + y = 1 \ .$$

Rješenje: Kao što smo rekli, formiramo prvo lagranžijan

$$\Lambda(x,y,\lambda) = f(x,y) - \lambda g(x,y) = x^2 + y^2 - \lambda(x + y - 1) \ ,$$

gdje je sa $g(x,y) = x + y - 1$ zadata uslovna funkcija.
U drugom koraku računamo gradijent lagranžijana

$$\nabla\Lambda(x,y,\lambda) = \left(\frac{\partial\Lambda}{\partial x}, \frac{\partial\Lambda}{\partial y}, \frac{\partial\Lambda}{\partial\lambda}\right) = (2x - \lambda, 2y - \lambda, x + y - 1) \ .$$

Sada rješavamo sistem

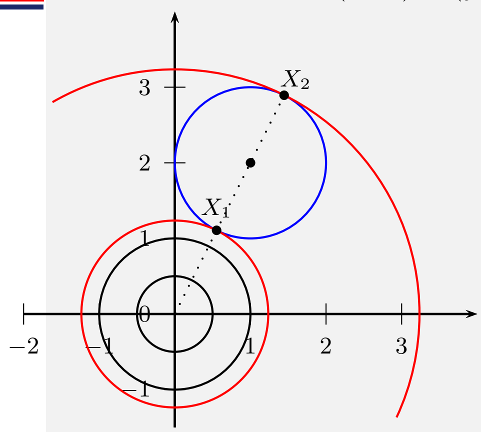
$$\begin{aligned} 2x - \lambda &= 0 \\ 2y - \lambda &= 0 \\ x + y - 1 &= 0 \ . \end{aligned}$$

Iz prve dvije jednačine sistema imamo $2x = 2y$, to jest $x = y$, pa uvrštavajući to u treću jednačinu, dobijamo $x = y = \frac{1}{2}$ i za ove vrijednosti je $\lambda = 1$. Dakle, imamo jednu stacioarnu tačku $\boldsymbol{x}_0 = \left(\frac{1}{2}, \frac{1}{2}, 1\right)$.
Posljedni korak je utvrđivanje karaktera tačke $\boldsymbol{x}_0$. Računajući druge parcijalne izvode, imamo

$$d^2 f(\boldsymbol{x}_0) = 2dx^2 + 2dy^2 \ ,$$

i vidimo da je $d^2 f(\boldsymbol{x}_0) > 0$ (kao suma kvadrata), te dakle imamo minimum funkcije f, pri uslovu g, u tački $\left(\frac{1}{2}, \frac{1}{2}\right)$, i on iznosi $f_{min} = \frac{1}{2}$.

Primjer 45 : Odrediti na kružnici $k:\ (x-1)^2 + (y-2)^2 = 1$ najbližu i najdalju tačku od koordinatnog početka.



Problem možemo riješiti jednostavno, povlačeći pravu određenu tačkama $(0,0)$ i $(1,2)$, i nalazeći njen presjek sa zadatom kružnicom.
Riješimo problem ipak na "teži" način. Razmišljajmo ovako: ako opišemo centralnu kružnicu proizvoljnog poluprečnika r, onda ona na sebi sadrži sve one tačke koje su na istom odstojanju r od koodinatnog početka.

"Naduvajmo" neku malu centralnu kružnicu, sve do momenta njenog dodira sa zadatom kručnicom k. Tačka dodira će upravo biti najbliža tačka koordinatnom početku. Ako nastavimo "naduvavanje", kružnice će sjeći kružnicu k ali tu nemamo tačaka koje su najbliže ili najdalje jer su sve one dalje od prve dodirne tačke, a naduvavanjem dobijamo sve dalje i dalje tačke. Ovo naravno vrijedi do momenta kada ponovo dobijemo kružnicu koja dodirne kružnicu k (velika crvena kružnica).

Cijeli opisani postupak nas navodi da problem postavimo ovako: nađimo minimum i maksimum funkcije $f(x,y) = x^2 + y^2$ (to su centralne kružnice čije poluprečnike tražimo) pri uslovu $(x-1)^2 + (y-2)^2 = 1$ (na ovoj kružnici tražimo najbližu i najdalju tačku). Dakle, rješavamo

$$x^2 + y^2 \longrightarrow ext$$
$$(x-1)^2 + (y-2)^2 = 1 \ .$$

Lagranžijan problema je $\Lambda(x,y,\lambda) = x^2 + y^2 - \lambda((x-1)^2 + (y-2)^2 - 1)$, a njegov gradijent,

$\nabla\Lambda = (2x - 2\lambda(x-1), 2y - 2\lambda(y-1), (x-1)^2 + (y-2)^2 - 1)$. Rješavamo sistem

$$\begin{aligned} 2x - 2\lambda(x-1) &= 0 \\ 2y - 2\lambda(y-1) &= 0 \\ (x-1)^2 + (y-2)^2 - 1 &= 0 \ . \end{aligned}$$

Stacionarne tačke su $X_1\left(1 + \frac{1}{\sqrt{5}}, 2 + \frac{2}{\sqrt{5}}\right)$ i $X_2\left(1 - \frac{1}{\sqrt{5}}, 2 - \frac{2}{\sqrt{5}}\right)$. Prostom provjerom zaključujemo da funkcija u ovim tačkama ima najveću i najmanju vrijednost, odnosno da su to najdalja i najbliža tačka koordinatnom početku, na kružnici k.

Vratimo se na opasku "teži" način rješavanja. Postavimo problem da na proizvoljnoj liniji $g(x, y) = 0$ nađemo najbližu ili najdalju tačku od koordinatnog početka. Sada onaj "lakši" način uopšte nemožemo primjeniti, a ovaj "teži" funkcioniše. Dakle, on je univerzalnijeg karaktera i kao takav mnogo bolji način. Npr. naći na grafu funkcije $y = x^2 + x + 1$ tačku najbližu koordinatnom početku.

U prethodna dva primjera vidjeli smo kako funkcioniše Lagrangeov metod, pri čemu smo imali samo jedno oganičenje, a time i jednu dodatnu varijablu problema. Naravno da ograničenja može biti i više, međutim metod se bitno neće mijenjati. Naime, neka je zadat problem sa dva ograničenja.

$$f(x_1, x_2, ..., x_n) \longrightarrow ext$$

$$h(x_1, x_2, ..., x_n) = 0 \ ,$$

$$g(x_1, x_2, ..., x_n) = 0 \ .$$

Formiramo lagranžijan, tako da svakom ograničenju pridružimo po jedan lagrangeov multiplikator,

$$\Lambda(x_1, x_2, ..., x_n, \lambda, \mu) = f(x_1, x_2, ..., x_n) - \lambda h(x_1, x_2, ..., x_n) - \mu g(x_1, x_2, ..., x_n) \ .$$

Nalaženjem gradijenta lagranžijana, postavljamo sistem

$$\begin{aligned} \frac{\partial\Lambda}{\partial x_i} &= \frac{\partial f}{\partial x_i} - \lambda\frac{\partial g}{\partial x_i} - \mu\frac{\partial h}{\partial x_i} = 0 \ , \ i = 1, 2, ..., n \\ \frac{\partial\Lambda}{\partial\lambda} &= g(x_1, x_2, ..., x_n) = 0 \\ \frac{\partial\Lambda}{\partial\mu} &= h(x_1, x_2, ..., x_n) = 0 \ , \end{aligned}$$

čijim rješavanjem dobijamo stacionarne tačke problema. Kao i u slučaju jednog ograničenja, nekim od poznatih postupaka odredimo karakter stacionarnih tačaka.

U opštem slučaju, ako imamo k ograničenja $g_i(\boldsymbol{x}) = 0$, $i = 1, 2, \ldots, k$, postupak je isti, a lagranžijan je

$$\Lambda(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i=1}^{k} \lambda_i g_i(\boldsymbol{x}) \ , \ \boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_k) \ .$$

Primjer 46 : Riješimo problem

$$f(x, y, z) = 4y - 2z \longrightarrow ext$$

$$2x - y - z - 2 = 0 \ ,$$

$$x^2 + y^2 - 1 = 0 \ .$$

Rješenje: Za egzistenciju rješenja gornjeg problema pozivamo se na Teorem 4.0.1. Zaista, zbog drugog ograničenja, očigledno je da vrijedi $0 \leq x, y \leq 1$, a iz prvog ograničenja onda zaključujemo da je $-3 \leq z \leq 0$, pa je skup na kome tražimo ekstremne vrijednosti funkcije ograničen i zatvoren.

Lagranžijan glasi

$$\Lambda(x, y, z, \lambda, \mu) = 4y - 2z - \lambda(2x - y - z - 2) - \mu(x^2 + y^2 - 1) \ .$$

Nalazimo parcijalne izvode lagranžijana

$$\begin{aligned}
\frac{\partial \Lambda}{\partial x} &= -2\lambda - 2\mu x \\
\frac{\partial \Lambda}{\partial y} &= 4 - \lambda - 2\mu y \\
\frac{\partial \Lambda}{\partial z} &= -2 - \lambda \\
\frac{\partial \Lambda}{\partial \lambda} &= -2x + y + z + 2 \\
\frac{\partial \Lambda}{\partial \mu} &= -x^2 - y^2 + 1 \ .
\end{aligned}$$

Sistem koga rješavamo ima pet nepoznatih (x, y, z, λ, μ) i pet jednačina

$$\begin{aligned}
-2\lambda - 2\mu x &= 0 \\
4 + \lambda - 2\mu y &= 0 \\
-2 + \lambda &= 0 \\
-2x + y + z + 2 &= 0 \\
-x^2 - y^2 + 1 &= 0 \ .
\end{aligned}$$

Iz treće jednačine direktno slijedi $\lambda = 2$. Ubacujući to u prvu i drugu jednačinu, dobijamo

$$x = -\frac{2}{\mu}\ ,\ y = \frac{3}{\mu}\ .$$

Stavljajući ove rezultate u petu jednačinu, imamo

$$\frac{4}{\mu^2} + \frac{9}{\mu^2} = \frac{13}{\mu^2} = 1 \implies \mu = \pm\sqrt{13}\ .$$

Sada imamo dva slučaja. Za $\mu = \sqrt{13}$, $x = -\frac{2}{\sqrt{13}}$ i $y = \frac{3}{\sqrt{13}}$. Iskoristimo li i četvrtu jednačinu, dobijamo $z = -2 - \frac{7}{\sqrt{13}}$. Time smo dobili prvu stacionarnu tačku

$$X_1(x, y, z, \lambda, \mu) = X_1(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}, -2 - \frac{7}{\sqrt{13}}, 2, \sqrt{13})\ .$$

Analogno, za slučaj $\mu = -\sqrt{13}$, dobijamo stacionarnu tačku

$$X_2(\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}, -2 + \frac{7}{\sqrt{13}}, 2, -\sqrt{13})\ .$$

Lahko se sada provjerava da u tački X_1 imamo maksimum

$$f_{max} = f(X_1) = 4 + \frac{26}{\sqrt{13}}\ ,$$

a minimum u tački X_2

$$f_{min} = f(X_2) = 4 - \frac{26}{\sqrt{13}}\ .$$

4.2 Još o diferenciranju funkcija

Ako nam je zadata vektorska funkcija zavisna o jednoj varijabli $\boldsymbol{r}(t) = (x(t), y(t), z(t)) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}$, tada diferenciranje ovakve funkcije izvodimo jednostavno tako što diferenciramo svaku komponentu vektora, to jest vrijedi

$$\frac{d\boldsymbol{r}(t)}{dt} \stackrel{def}{=} \left(\frac{dx(t)}{dt}, \frac{dy(t)}{dt}, \frac{dz(t)}{dt}\right) = \frac{dx(t)}{dt}\boldsymbol{i} + \frac{dy(t)}{dt}\boldsymbol{j} + \frac{dz(t)}{dt}\boldsymbol{k} \ .$$

Dakle, izvod vektora je vektor čije su komponente izvodi komponenti polaznog vektora. Primjetimo da su bazni vektori neovisni o varijabli derivacije. Primjena ovoga nam je poznata iz fizike, gdje recimo ako je $\boldsymbol{r}$ vektor položaja čestice u vremenskom trenutku t, tada $\frac{d\boldsymbol{r}}{dt}$ predstavlja brzinu čestice ($\boldsymbol{v}$), a veličina $\frac{d\boldsymbol{v}}{dt}$ predstavlja ubrzanje čestice ($\boldsymbol{a}$).
Proizvod skalara i vektora kao i skalarni i vektorski proizvod vektora diferenciraju se uobičajenim pravilima za diferenciranje proizvoda uz napomenu da se u izvodu vektorskog proizvoda treba voditi računa o redoslijedu članova.

$$\frac{d(\boldsymbol{u}\cdot\boldsymbol{v})}{dt} = \frac{d\boldsymbol{u}}{dt}\cdot\boldsymbol{v} + \boldsymbol{u}\cdot\frac{d\boldsymbol{v}}{dt} \quad , \quad \frac{d(\boldsymbol{u}\times\boldsymbol{v})}{dt} = \frac{d\boldsymbol{u}}{dt}\times\boldsymbol{v} + \boldsymbol{u}\times\frac{d\boldsymbol{v}}{dt} \ .$$

PRIMJER 47 : Neka se čestica kreće jednoliko ubrzano po kružnici. Tada vrijedi

$$\boldsymbol{r}^2 = \boldsymbol{r}\cdot\boldsymbol{r} = const \ \ \boldsymbol{i} \ \ \boldsymbol{v}^2 = \boldsymbol{v}\cdot\boldsymbol{v} = const \ .$$

Nalaženjem izvoda ovih dviju jednakosti imamo,

$$2\boldsymbol{r}\cdot\frac{d\boldsymbol{r}}{dt} = 0 \ \ \boldsymbol{i} \ \ 2\boldsymbol{v}\cdot\frac{d\boldsymbol{v}}{dt} = 0 \ .$$

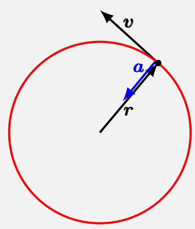
Prvu jednakost interpretiramo da je

$$\boldsymbol{r}\cdot\boldsymbol{v} = 0 \ , \tag{4.10}$$

a drugu

$$\boldsymbol{v}\cdot\boldsymbol{a} = 0 \ . \tag{4.11}$$

Jednakost (4.10) znači da su vektor položaja čestice i njen vektor brzine ortogonalni, a jednakost (4.11) da su vektor brzine i vektor ubrzanja ortogonalni. Iz ovoga zaključujemo da su vektor položaja čestice i vektor njenog ubrzanja kolinearni (paralelni) (kretanje se odvija u ravni), a to znači da je ugao između tih vektora ili 0° ili 180°. Diferenciranjem jednačine (4.10) imamo



$$\frac{d(\boldsymbol{r}\cdot\boldsymbol{v})}{dt} = \boldsymbol{r}\cdot\frac{d\boldsymbol{v}}{dt} + \boldsymbol{v}\cdot\frac{d\boldsymbol{r}}{dt} = \boldsymbol{r}\cdot\boldsymbol{a} + \boldsymbol{v}\cdot\boldsymbol{v} = 0 \ .$$

Dakle, $\boldsymbol{r}\cdot\boldsymbol{a} = -\boldsymbol{v}^2$, a kako je $\boldsymbol{r}\cdot\boldsymbol{a} = ||\boldsymbol{r}||\,||\boldsymbol{a}||\cos\angle(\boldsymbol{r},\boldsymbol{a})$, mora biti $\cos\angle(\boldsymbol{r},\boldsymbol{a}) < 0$, a to znači da je ugao između njih 180°. Dakle, vektor položaja i vektor ubrzanja čestice su paralelni i suprotno orijentisani. Dakle, ubrzanje kod jednolikog kretanja po kružnici je usmjereno prema centru kružnice. Šta više iz posljednje dvije jednakosti imamo da je veličina ubrzanja data formulom $a = \dfrac{\boldsymbol{v}^2}{r}$ $(a = ||\boldsymbol{a}||,\ r = ||\boldsymbol{r}||)$.

4.2.1 Divergencija i rotor

Kao što smo to već ranije vidjeli, gradijent skalarnog polja ∇f nije ništa drugo do množenje gradijenta (vektora $\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$) skalarom (skalarna funkcija $f(x_1, x_2, \ldots, x_n)$). Nabla operator još nazivamo i operator diferenciranja jer očigledno on na određen način daje informacije o "izvodu" funkcije (rezultat njegovog djelovanja na skalarnu funkciju je vektor parcijalnih izvoda te funkcije).

Međutim, nabla operatorom možemo djelovati i na drugi vektor, to jest možemo ga množiti i drugim vektorom. Pri tome smo zainteresovani kako za vektorske osobine, tako i za diferencijalne osobine tog množenja. Za vektore imamo skalarni proizvod i vektorski proizvod, tako da nas to vodi ka dva nova pojma u radu sa nabla operatorom.

Definicija 4.2.1. *Neka je $\boldsymbol{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ vektorsko polje u $\mathbb{R}^3$, to jest vektorska funkcija vektorske promjenljive. Divergencija vektorskog polja $\boldsymbol{F}$ je funkcija zadata sa*

$$\operatorname{div}(\boldsymbol{F}) = \nabla \cdot \boldsymbol{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \ .$$

Primijetimo da je divergencija vektorskog polja definisana skalarnim proizvodom gradijenta i vektorskog polja. Takođe primjetimo da je divergencija vektorskog polja skalarno polje.

PRIMJER 48 : Neka je dato vektorsko polje $\boldsymbol{F}(x,y,z) = (x^2y^2, y^2z^2, x^2z^2) = x^2y^2\boldsymbol{i} + y^2z^2\boldsymbol{j} + x^2z^2\boldsymbol{k}$. Tada je

$$\begin{aligned}\operatorname{div}(\boldsymbol{F}) = \nabla \cdot \boldsymbol{F} &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot (x^2y^2, y^2z^2, x^2z^2)\\ &= \frac{\partial (x^2y^2)}{\partial x} + \frac{\partial (y^2z^2)}{\partial y} + \frac{\partial (x^2z^2)}{\partial z} = 2xy^2 + 2yz^2 + 2x^2z \ .\end{aligned}$$

Divergencija je operator koji mjeri intenzitet izvora ili ponora vektorskog polja u datoj tački. Naprimjer, za vektorsko polje koje pokazuje brzinu širenja zraka kada se on zagrijava, divergencija polja brzine imala bi pozitivnu vrijednost, jer se zrak širi. Da se zrak hladi i skuplja, divergencija bi bila negativna. Lokalno gledano, divergencija mjeri prostornu promjenu vektorskog polja. Divergencija bi se mogla opisati i kao mjera promjene u gustoći.

PRIMJER 49 : Neka je $\boldsymbol{r} = (x,y,z)$ vektor položaja proizvoljne tačke u $\mathbb{R}^3$. Tada je

$$\nabla \cdot \boldsymbol{r} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot (x,y,z) = 3 \ .$$

Neka se jedan električni naboj (monopol) nalazi u koordinatnom početku. Neka se električna sila tog naboja prostire radijalno i ovisna je direktno proporcionalno o položaju tačke i obrnuto proporcionalno sa kubom rastojanja od naboja. Dakle, $\boldsymbol{F} = \frac{\boldsymbol{r}}{r^3}$, gdje je $\boldsymbol{r} = (x,y,z)$ vektor položaja proizvoljne tačke, a $r = ||\boldsymbol{r}|| = \sqrt{x^2+y^2+z^2}$. Izračunajmo divergenciju ovog vektorskog polja.

$$\begin{aligned}\operatorname{div}(\boldsymbol{F}) = \nabla \cdot \boldsymbol{F} = \nabla \cdot \left(\frac{\boldsymbol{r}}{r^3}\right) &= \frac{\partial}{\partial x}\left(\frac{x}{r^3}\right) + \frac{\partial}{\partial y}\left(\frac{y}{r^3}\right) + \frac{\partial}{\partial z}\left(\frac{z}{r^3}\right)\\ &= \frac{3}{r^3} + x\frac{\partial}{\partial x}\left(\frac{1}{r^3}\right) + y\frac{\partial}{\partial y}\left(\frac{1}{r^3}\right) + z\frac{\partial}{\partial z}\left(\frac{1}{r^3}\right)\\ &= \frac{3}{r^3} + x\frac{x}{r}\left(-\frac{3}{r^4}\right) + y\frac{y}{r}\left(-\frac{3}{r^4}\right) + z\frac{z}{r}\left(-\frac{3}{r^4}\right)\\ &= \frac{3}{r^3} + \frac{x^2+y^2+z^2}{r}\left(-\frac{3}{r^4}\right) = \frac{3}{r^3} + \frac{r^2}{r}\left(-\frac{3}{r^4}\right) = 0 \ .\end{aligned}$$

Sve ovo izvedeno vrijedi za $\boldsymbol{r} \neq (0,0,0)$ jer je tačka gdje se nalazi naboj singularna tačka.
Na osnovu ovog razmatranja se dobija Coulombov potencijal $V(r) = -\frac{q}{4\pi\varepsilon_0 r}$, sa električnim poljem $\boldsymbol{E} = -\nabla V = \frac{q\boldsymbol{r}}{4\pi\varepsilon_0 r^2}$.

Osnovne osobine divergencije iskazujemo narednim tvrđenjem.

Teorem 4.2.1. *Neka su f skalarno polje i $\boldsymbol{F}$ i $\boldsymbol{G}$ vektorska polja, sve diferencijabilne funkcije i $a, b \in \mathbb{R}$. Tada vrijedi:*

1. *Ako je $\boldsymbol{F}$ konstantno vektorsko polje, tada je $\operatorname{div}(\boldsymbol{F}) = 0$.*
2. $\operatorname{div}(a\boldsymbol{F} + b\boldsymbol{G}) = a\operatorname{div}(\boldsymbol{F}) + b\operatorname{div}(\boldsymbol{G})$.
3. $\operatorname{div}(f\boldsymbol{F}) = \nabla f \cdot \boldsymbol{F} + f\operatorname{div}(\boldsymbol{F})$.

Definicija 4.2.2. *Neka je $\boldsymbol{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ vektorsko polje u $\mathbb{R}^3$, to jest vektorska funkcija vektorske promjenljive. Rotor vektorskog polja $\boldsymbol{F}$ je funkcija data sa*

$$\operatorname{rot}(\boldsymbol{F}) = \nabla \times \boldsymbol{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) .$$

Primijetimo da je rotor vektorskog polja ustvari vektorski proizvod gradijenta i vektorskog polja i da je on opet vektorsko polje.

Primjer 50 : Za vektorsko polje zadato sa $\boldsymbol{F}(x,y,z) = (x-y, x+y, z)$, njegova divergencija je skalar

$$\operatorname{div}(\boldsymbol{F}) = \nabla \cdot \boldsymbol{F} = \frac{\partial(x-y)}{\partial x} + \frac{\partial(x+y)}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3.$$

a njegov rotor je vektor

$$\operatorname{rot}(\boldsymbol{F}) = \nabla \times \boldsymbol{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x-y & x+y & z \end{vmatrix} = 0\boldsymbol{i} + 0\boldsymbol{j} + (1+1)\boldsymbol{k} = 2\boldsymbol{k} = (0,0,2) .$$

Rotor vektorskog polja ima smisla definisati samo u Euklidskom prostoru $\mathbb{R}^3$ (šta bi bila četvrta vrsta u determinanti?), za razliku od gradijenta i divergencije koje možemo posmatrati na proizvoljnom n-dimenzionalnom euklidskom prostoru, $n \in \mathbb{N}$. Osnovne osobine rotora iskazujemo narednim tvrđenjem.

Teorem 4.2.2. *Neka su f skalarno polje i $\boldsymbol{F}$ i $\boldsymbol{G}$ vektorska polja, sve diferencijabline funkcije i $a, b \in \mathbb{R}$. Tada vrijedi:*

1. *Ako je $\boldsymbol{F}$ konstantno vektorsko polje, tada je* $\operatorname{rot}(\boldsymbol{F}) = 0$.
2. $\operatorname{rot}(a\boldsymbol{F} + b\boldsymbol{G}) = a\operatorname{rot}(\boldsymbol{F}) + b\operatorname{rot}(\boldsymbol{G})$.
3. $\operatorname{rot}(f\boldsymbol{F}) = \nabla f \times \boldsymbol{F} + f\operatorname{rot}(\boldsymbol{F})$.

Neke zajedničke osobine divergencije i rotora iskazane su narednom tvrdnjom.

Teorem 4.2.3. *Neka su f skalarno polje i $\boldsymbol{F}$ i $\boldsymbol{G}$ vektorska polja, sve diferencijabline funkcije. Tada vrijedi:*

1. $\operatorname{div}(\boldsymbol{F} \times \boldsymbol{G}) = \boldsymbol{G} \cdot (\nabla \times \boldsymbol{F}) - \boldsymbol{F} \cdot (\nabla \times \boldsymbol{G}) = \boldsymbol{G} \cdot \operatorname{rot}(\boldsymbol{F}) - \boldsymbol{F} \cdot \operatorname{rot}(\boldsymbol{G})$.
2. $\operatorname{rot}(\boldsymbol{F} \times \boldsymbol{G}) = \boldsymbol{F} \cdot \operatorname{div}(\boldsymbol{G}) - \boldsymbol{G} \cdot \operatorname{div}(\boldsymbol{F}) + (\boldsymbol{G} \cdot \nabla)\boldsymbol{F} - (\boldsymbol{F} \cdot \nabla)\boldsymbol{G}$.
3. $\operatorname{rot}(\nabla f) = 0$.
4. $\operatorname{div}(\operatorname{rot}(\boldsymbol{F})) = 0$.

Neka je $f \in C^2(\mathbb{R}^3)$ skalarno polje. Tada je ∇f vektorsko polje, pa na njega možemo primijeniti operatore divergencije i rotora. Rotor vektorskog polja ∇f je

$$\begin{aligned}\operatorname{rot}(\nabla f) = \nabla \times \nabla f &= \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} \\ &= \boldsymbol{i}\left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}\right) + \boldsymbol{j}\left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}\right) + \boldsymbol{k}\left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}\right) = 0\end{aligned}$$

Sa ovim smo praktično dokazali osobinu 3. gornjeg teorema.

Pretpostavimo da nam je poznato vektorsko polje $\boldsymbol{F}(x,y,z)$ u $\mathbb{R}^3$ i želimo da odredimo da li postoji skalarna funkcija φ takva da je $\boldsymbol{F} = \nabla\varphi$. Odgovor na ovo pitanje dobijamo izračunavanjem rotora vektorskog polja $\boldsymbol{F}$. Naime, ako je $\mathrm{rot}(\boldsymbol{F}) \neq 0$, tada ne postoji skalarna funkcija φ takva da je $\boldsymbol{F} = \nabla\varphi$. Ovo zaključujemo iz 3. prethodnog teorema jer mora biti $\mathrm{rot}(\nabla\varphi) = 0$. Dakle, u prostoru glatkih funkcija koje su definisane nad povezanim domenom (oblast “bez rupa”), potreban i dovoljan uslov da postoji skalarna funkcija φ takva da je $\boldsymbol{F} = \nabla\varphi$ jeste da je $\mathrm{rot}(\boldsymbol{F}) = 0$.

Primjer 51 : Neka je $\mathrm{div}(\boldsymbol{F}) = 0$. Tada se rješenje ove jednačine može tražiti u obliku $\boldsymbol{F} = \nabla \times f$, gdje je $f = f(\boldsymbol{r})$ ($\boldsymbol{r} = (x,y,z)$) vektorski potencijal. Zaista, tada bi imali

$$\begin{aligned}
\mathrm{div}(\boldsymbol{F}) = \nabla\cdot(\nabla\times f) &= \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\cdot\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z},\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x},\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\\
&= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\\
&= \frac{\partial^2 R}{\partial y\partial x}-\frac{\partial^2 Q}{\partial z\partial x}+\frac{\partial^2 P}{\partial z\partial y}-\frac{\partial^2 R}{\partial x\partial y}+\frac{\partial^2 Q}{\partial x\partial z}-\frac{\partial^2 P}{\partial y\partial z} = 0\ .
\end{aligned}$$

Gornje rezonovanje se neće promjeniti i da smo vektorskom potencijalu dodali gradijent neke skalarne funkcije, a to nam govori da vektorski potencijal koji određuje funkciju $\boldsymbol{F}$ nije jednoznačan.

Divergencija vektorskog polja $\nabla\varphi$ je

$$\begin{aligned}
\mathrm{div}(\nabla\varphi) = \nabla\cdot\nabla\varphi &= \frac{\partial}{\partial x}\left(\frac{\partial\varphi}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial\varphi}{\partial y}\right)+\frac{\partial}{\partial z}\left(\frac{\partial\varphi}{\partial z}\right)\\
&= \frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2}
\end{aligned}$$

Ovaj operator se naziva Laplaceov operator ili laplasijan i označava se sa Δ (čitamo ”delta”). Dakle,

$$\Delta = \mathrm{div}\nabla = \nabla\cdot\nabla = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\ .$$

Ovaj operator ima brojne primjene u mnogim oblastima matematike, te u fizici, elektrostatici, kvantnoj mehanici, obradi snimaka, itd. Nazvan je po francuskom matematičaru Pjeru Simonu Laplaceu.

Primjer 52 : Laplasijan skalarnog polja $f(x,y,z) = x^2+y^2+z^2$ jeste

$$\Delta f = \frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2} = 6\ .$$

Iskažimo još neke veze divergencije, rotora i laplasijana.

Teorem 4.2.4. *Neka su f i g skalarna polja i $\boldsymbol{F}$ vektorsko polje, sve diferencijabline funkcije. Tada vrijedi:*

1. $\mathrm{div}(f\nabla g) = \nabla f\cdot\nabla g + f\Delta g$.
2. $\mathrm{rot}(f\nabla g) = \nabla f\times\nabla g$.
3. $\mathrm{rot}(\mathrm{rot}(\boldsymbol{F})) = \nabla\mathrm{div}(\boldsymbol{F}) - \Delta\boldsymbol{F}$.

Za kraj navedimo jednu konkretnu implementaciju rotora i divergencije. Neka je $\boldsymbol{E}$ električno polje (vektorsko polje) i $\boldsymbol{B}$ magnetsko polje (vektorsko polje). Elektromagnetski fenomeni u vakuumu opisani su poznatim Maxwellovim jednačinama, koje u diferencijalnoj formi glase:

$$\begin{aligned}
\nabla \cdot \boldsymbol{E} &= \mathrm{div}(\boldsymbol{E}) = \frac{\rho}{\varepsilon_0} \\
\nabla \cdot \boldsymbol{B} &= \mathrm{div}(\boldsymbol{B}) = 0 \\
\nabla \times \boldsymbol{E} &= \mathrm{rot}(\boldsymbol{E}) = -\frac{\partial \boldsymbol{B}}{\partial t} \\
\nabla \times \boldsymbol{B} &= \mathrm{rot}(\boldsymbol{B}) = \mu_0 \boldsymbol{J} + \mu_0 \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}
\end{aligned}$$

gdje su
ρ - gustoća električnog naboja
ε_0 - dielektrična konstanta vakuuma
μ_0 - permeabilnost vakuuma
$\boldsymbol{J}$ - gustoća električne struje.

Prva Maxwellova jednačina poznata je kao Gaussov zakon, koji govori da je električni naboj izvor (ili ponor) električnog polja. Ukupni električni tok kroz zatvorenu površinu proporcionalan je količini električnog naboja koji se nalazi unutar zapremine te površi. Ako unutar te zatvorene površi nema električnog naboja (ili je količina pozitivnog jednaka količini negativnog električnog naboja), ukupni električni tok kroz tu zatvorenu površ je nula. Ovo ipak ne znači da u toj zapremini nema električnog polja, već samo da ukupni tok iščezava. Dakle, ako nema električnog naboja u toj posmatranoj zapremini, koliko silnica električnog polja ulazi kroz površ koja opisuje zapreminu, toliko silnica negdje i izlazi iz te iste zatvorene površi.

Druga Maxwellova jednačina slična je prvoj (u situaciji u kojoj ne postoji naboj) i ona opisuje magnetsko polje. Ova jednačina govori da ne postoji "magnetski naboj" (magnetski monopol), to jest ne postoji izvor magnetskog polja iz kojega bi proizlazio magnetski tok različit od nule. U svakoj tački prostora, količina silnica magnetskog polja koja ulazi u tu tačku jednaka je količini silnica koje izlaze iz te tačke. Silnice magnetskog polja nemaju izvora (ili ponora), pa zbog toga ukupni magnetski tok kroz zatvorenu površ uvijek iščezava. To vrijedi i za izvore magnetskog polja, stoga je svaki izvor magnetskog polja barem dipol.

Treća jednačina je poznata kao Faradayjev zakon indukcije: Svaka promjena magnetnog polja stvara električno polje.

Četvrta jednačina je prošireni Ampereov zakon: Oko provodnika kojim teče struja indukuje se magnetsko polje, ali i svako promjenjivo električno polje indukuje magnetnko polje.