Višedimenziona optimizacija bez ograničenja

Zoran D. Jeličić
Mirna N. Kapetina
5. novembar 2021.

1 Uvodna razmatranja

U okviru ovog poglavlja, namera nam je da izvedemo potrebne i dovoljne uslove za koje funkcija više promenljivih ima ekstremnu (mininalnu ili maksimalnu) vrednost. Sva izvođenja su data pod pretpostavkom da promenljive x_i imaju proizvoljnu vrednost, odnosno da ne trpe ograničenja. Po svom karakteru ovo izvođenje podseća na slučaj funkcije jedne promenljive, ali su matematička izvođenja znatno složenija, pa smo ponavljanju osnovih postulata matematiče analize, posvetili uvodne paragrafe.

1.1 Razvoj funkcije više promenljivih u Tejlorov red

Posmatramo funkciju više promenljivih [1]

$$y(\mathbf{x}) = y(x_1, x_2, x_3, \ldots, x_n) \tag{1}$$

gde je očigledno vektor promenljivih stanja označen i definisan kao $\mathbf{x} = [x_1, x_2, x_3, \ldots, x_n]^T$.
Pod pretpostavkom da je funkcija (1) definisana i neprekidna zajedno sa svojih $j+1$ izvoda u okolini tačke [2] $\mathbf{x}^*$, tada je uobičajeno da se razvoj u Tejlorov red, u njegovoj opštoj formi, zapisuje kao

$$y(\mathbf{x}) = y(\mathbf{x}^*) + \frac{\Delta y}{1!} + \frac{\Delta^2 y}{2!} + \ldots + \frac{\Delta^j y}{j!} + R_{j+1} \tag{2}$$

gde je R_{j+1} ostatak razvoja u red, a ostali članovi su, [3] uz slično uopštenje na više izvode,

$$\begin{aligned}
\Delta y &= (x_1 - x_1^*)\left(\frac{\partial y}{\partial x_1}\right)^* + (x_2 - x_2^*)\left(\frac{\partial y}{\partial x_2}\right)^* + \ldots + (x_n - x_n^*)\left(\frac{\partial y}{\partial x_n}\right)^* \\
\Delta^2 y &= (x_1 - x_1^*)^2\left(\frac{\partial^2 y}{\partial x_1^2}\right)^* + 2(x_1 - x_1^*)(x_2 - x_2^*)\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^* + \\
&\quad + (x_2 - x_2^*)^2\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* + \ldots + (x_n - x_n^*)^2\left(\frac{\partial^2 y}{\partial x_n^2}\right)^*.
\end{aligned} \tag{3}$$

[1] Jasno je da ova funkcija predstavlja kriterijum optimalnosti, čiji minimum/maksimum pokušavamo da odredimo

[2] Pojam tačka se koristi, kao uobičajeni sinonim za kooridnate vektora stanja x_i u n dimenzionalnom prostoru

[3] Radi lakšeg praćenja teksta uveli smo oznaku $\left(\frac{\partial y}{\partial x_i}\right)^*$ koja označava i-ti paricijalni izvod u tački $\mathbf{x}^*$.

Odnosno u opštem obliku

$$\Delta^k y = \sum \frac{k!}{p_1!p_2!\dots p_n!}(x_1 - x_1^*)_1^p (x_2 - x_2^*)_2^p \dots (x_n - x_n^*)_n^p \left(\frac{\partial^k y}{\partial x_1^{p_1} \partial x_2^{p_2} \partial x_n^{p_n}} \right)^* \tag{4}$$

pri čemu se sumira po svim celobrojnim vrednostima p_i za koje važi $\sum p_i = k$. Svesni smo da je ova notacija na prvi pogled [4] složena i da bi u daljim izvođenjima suština bila zamenjena formom, te smo se bez gubitka na opštosti opredelili da dalju studiju nastavimo za funkciju dve promenljive, a da na kraju dobijene zakonomernosti uopštimo za funkciju sa n promenljivih. Odnosno, ako pretpostavimo da je kriterijum optimalnosti (1) funkcija dve promenljive

[4] a i svaki drugi pogled

$$y(\mathbf{x}) = y(x_1, x_2) \tag{5}$$

tada se dominantni članovi reda, na osnovu (3) i (4) mogu zapisati na sledeći način

Napomena: Prvi član reda $k = 1$, ima vrednosti parametara (p_1, p_2): $(1, 0)$ i $(0, 1)$, a drugi član reda $k = 2$ za iste parametre ima vrednosti $(2, 0)$, $(1, 1)$ i $(0, 2)$

$$\begin{aligned} \Delta y &= (x_1 - x_1^*) \left(\frac{\partial y}{\partial x_1} \right)^* + (x_2 - x_2^*) \left(\frac{\partial y}{\partial x_2} \right)^* \\ \Delta^2 y &= (x_1 - x_1^*)^2 \left(\frac{\partial^2 y}{\partial x_1^2} \right)^* + 2(x_1 - x_1^*)(x_2 - x_2^*) \left(\frac{\partial^2 y}{\partial x_1 \partial x_2} \right)^* + \\ &+ (x_2 - x_2^*)^2 \left(\frac{\partial^2 y}{\partial x_2^2} \right)^* . \end{aligned} \tag{6}$$

Prateći rezon iz jednodimenzione optimizacije, možemo uvesti smenu $x_i - x_i^* = h_i$ i pojednostaviti izraze (6) i zapisati ih u kompaktnijoj formi, koja je pogodnija za dalju analizu

$$\begin{aligned} \Delta y &= h_1 \left(\frac{\partial y}{\partial x_1} \right)^* + h_2 \left(\frac{\partial y}{\partial x_2} \right)^* \\ \Delta^2 y &= h_1^2 \left(\frac{\partial^2 y}{\partial x_1^2} \right)^* + 2h_1 h_2 \left(\frac{\partial^2 y}{\partial x_1 \partial x_2} \right)^* + h_2^2 \left(\frac{\partial^2 y}{\partial x_2^2} \right)^* . \end{aligned} \tag{7}$$

Izraz (7) ima centralnu ulogu u našoj daljoj studiji.

1.2 *Potrebni i dovoljni uslovi ekstremuma funkcije dve promenljive*

Podsećamo na slučaj funkcije jedne promenljive

$$y(x^* + h) - y(x^*) = hy'(x^*) + \frac{h^2}{2!} y''(x^*)$$

gde smo zbog *kolebanja* znaka h uveli potreban uslov da je prvi izvod u tački x^* jednak nuli, odnosno $y'(x^*) = 0$. Sličan rezon koristimo i u slučaju funkcije više promenljivih, koji se na osnovu (7) svodi na uslov da su svi parcijalni izvodi po promenljivima x_i jednaki nuli. tj.

$$\left(\frac{\partial y}{\partial x_i}\right)^* = 0 \quad \text{za} \quad i = 1, 2 \tag{8}$$

Svi kandidati za (lokalni) minimum/maksimum se nalaze među stacionarnim tačkama. Ovu činjenicu kroz navedeni intuitivni dokaz možemo formulisati na sledeći način.

Tvrđenje 1. (*Potrebni uslovi ekstremuma funkcije dve promenljive*) U slučaju da postoje parcijalni izvodi prvog reda funkcije y po svim promenljivama, onda tačka ekstremuma funkcije y u tački x^* zadovoljava sistem jednačina

$$\left(\frac{\partial y}{\partial x_i}\right)^* = 0 \quad \text{za} \quad i = 1, 2$$

odnosno uslov

$$\Delta y = 0$$

Ako tačka (x_1^*, x_2^*) zadovoljava uslove stacionarnosti [5], tada očevidno izraz (2), uz zanemarivanje članova višeg reda, postaje

$$y(\mathbf{x}) - y(\mathbf{x}^*) = \frac{\Delta^2 y}{2!}\,. \tag{9}$$

Dalje, iz izraza (9) jasno se vidi da stacionarna tačka (x_1^*, x_2^*) ima maksimalnu vrednost ako je desna strana jednačine manja od nule, odnosno minimalnu vrednost za $\Delta^2 y < 0$. Pretpostavićemo da nas interesuje, kada funkcija (5) ima maksimalnu vrednost, tj. kada je izraz

$$\Delta^2 y = h_1^2 \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* + 2h_1h_2 \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^* + h_2^2 \left(\frac{\partial^2 y}{\partial x_2^2}\right)^* < 0. \tag{10}$$

Tačnije za koje netrivijalne[6] vrednosti h_1 i h_2 je vrednost funkcije manja od nule.

$$h_1^2 \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* + 2h_1h_2 \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^* + h_2^2 \left(\frac{\partial^2 y}{\partial x_2^2}\right)^* < 0 \quad \text{za netrivijalne vrednosti} \quad h_1 \text{ i } h_2\,. \tag{11}$$

Akko važi uslov (11) tada tvrdimo da je funkcija $\Delta^2 y$ **negativno definitna** . Uobičajeno je da se u literaturi pojam definitnosti poistovećuje sa pojmovima minimuma i maksimuma, a izvođenje dovoljnih uslova

[5] Potrebne uslove ekstremuma, a tačku nazivamo stacionarana tačka

[6] netrivijalne znači, da je bar jedna vrednost h_i različita od nule

Akko je vrednost funkcije $\Delta^2 y > 0$ za netrivijalne vrednosti h_1 i h_2, tada kažemo da je funkcije Δ^2 **pozitivno definitna**
Akko je vrednost funkcije $\Delta^2 y \leq 0$ za netrivijalne vrednosti h_1 i h_2, tada kažemo da je funkcije Δ^2 **negativno semidefinitna**
Akko je vrednost funkcije $\Delta^2 y \geq 0$ za netrivijalne vrednosti h_1 i h_2, tada kažemo da je funkcije Δ^2 **pozitivno semidefinitna**
Akko je vrednost funkcije $\Delta^2 y <> 0$ za netrivijalne vrednosti h_1 i h_2, tada kažemo da je funkcije Δ^2 **nedefinitna**

se naziva ispitivanje definitnosti. Kao što ćete videti, ovo nije trivijalan problem i verovatno je najsloženiji u ovom delu kursa.

U našem konkretnom slučaju cilj nam je da nađemo dovoljne uslove pod kojima važe relacije (10) i (11), što se svodi da ove jednačine prepišemo u kompaktniju formu u kojoj će svi članovi h_1 i h_2 biti pod znakom kvadrata, što garantuje invarijantnost traženih uslova od znaka parametara h_i.

U izvođenju, koje sledi pratimo logiku, koja će nas dovesti do uslova koji garantuju negativnu (pozitivnu) definitnost, druge granične slučajeve ćemo uvesti kroz primere i nisu predmet ovog pitanja. Izraz (10) možemo razviti na sledeći način

$$\Delta^2 y = \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* \left[h_1^2 + 2h_1 h_2 \frac{\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^*}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} + h_2^2 \frac{\left(\frac{\partial^2 y}{\partial x_2^2}\right)^*}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} \right] < 0. \tag{12}$$

Prva dva člana u zagradi možemo posmatrati kao deo binomnog obrasca [7], a preostale članove ćemo posebno grupisati, tj.

[7] Prvi na kvadrat bi bio prvi član h_1^2, drugi član je dvostruki prvi puta drugi, nedostaje samo drugi član na kvadrat, koga ćemo dodati i oduzeti.

$$\Delta^2 y = \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* \left[h_1 + h_2 \frac{\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^*}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} \right]^2 + h_2^2 \left[\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* - \frac{\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{*2}}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} \right] < 0\,. \tag{13}$$

U izrazu (13) postigli smo da su svi članovi h_i pod znakom kvadrata i njihov doprinos karakteru ekstremuma nije više sporan, znači na definitnost sada utiču samo članovi, za koje ćemo pretpostaviti da su manji od 0[8].

[8] Stvarno nije teško zaključiti da negativna vrednost navedenih članova garantuje negativnu definitnost izraza (13). Može se pokazati da se i ostale moguće kombinacije u vrednosti izdvojenih članova svode na ovaj slučaj.

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^* < 0 \quad \text{i} \quad \left[\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* - \frac{\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{*2}}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} \right] < 0 \tag{14}$$

Izraz u srednjoj zagradi može se napisati u formi racionalnog izraza odnosno, kao

$$\frac{\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* - \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{*2}}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^*} < 0 \tag{15}$$

Prvi uslov iz izraza (14), $\left(\frac{\partial^2 y}{\partial x_1^2}\right)^* < 0$ nedvosmisleno ukazuje da imenilac izraza (15) mora biti manji od nule, a da brojilac tada mora biti veći od nule, pa uslovi negativne definitnosti konačno postaju

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^* < 0 \tag{16}$$

i

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* - \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{*2} > 0 \tag{17}$$

Izrazi (16) i (17) predstavljaju dovoljne uslove da funkcija dve promenljive ima maksimum, odnosno da je funkcija (10) **negativno definitna**.
Čitaocima ostavljamo da sami izvedu uslove **pozitivne definitnosti**, koji će se bez većih problema dobiti u sledećoj formi

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^* > 0 \tag{18}$$

i

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^* \left(\frac{\partial^2 y}{\partial x_1^2}\right)^* - \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{*2} > 0 \tag{19}$$

Međutim, važno je naglasiti da stacionarna tačka može biti maksimum ili minimum kada je funkcije $\Delta^2 y$ negativno semidefinitna ili pozitivno semidefinitna respektivno [9]. Ove slučajeve ćemo spomenuti u primerima, koji slede, a čitaoce upućujemo na preporučenu literaturu za više detalja.

Uslove definitnosti (16)-(19) možemo kompaktnije zaspisati u formi *Lagranžeovog kriterijuma* [10] na sledeći način

Tvrđenje 2. (*Dovoljni uslovi ekstremuma funkcije dve promenljive*) U slučaju da postoje parcijalni izvodi drugog reda funkcije y po svim promenljivama u okolini tačke $\mathbf{x}^*$ i ako važi $\left(\frac{\partial y}{\partial x_i}\right)^* = 0$ onda tačka ekstremuma funkcije y u tački x^* zadovoljavaju sledeće uslove

Za **strogi lokalni minimum** $D_i > 0 \quad \text{za} \quad i = 1, 2$

Za **strogi lokalni maksimum** $D_1 < 0 \quad D_2 > 0$

gde su

$$D_1 = \left|\frac{\partial^2 y}{\partial x_1^2}\right|_{\mathbf{x}=\mathbf{x}^*} \qquad D_2 = \begin{vmatrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} \\ \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_2^2} \end{vmatrix}_{\mathbf{x}=\mathbf{x}^*} .$$

Navedeni dovoljni uslovi predstavljaju veoma stroge zahteve, koje

[9] Semidefinitnost se povezuje sa uslovima minimuma/maksimuma, a definitnost sa uslovima strogog minimuma/maksimuma

[10] Joseph-Louis Lagrange (25 January 1736 – 10 April 1813) italijasnko-francuski matematičar i fizičar, na ovom mestu ga prvi put spominjemo i pored koleginica Kapetine, Buljević, Mitrović, Stokanović i kolege Rapaića sigurno će obeležiti svaku lekciju do kraja kursa

funkcija mora da ispuni po pitanju ekstremuma, uslovi mogu biti i blaži, na njih ćemo samo ukazati kroz primere.

Primer 1. Uslovi ekstremuma funkcije dve promenljive

1. Posmatrajmo kriterijum optimalnosti sa dve promenljive stanja

$$y(x_1, x_2) = 3x_2^2 + x_1^4 + 2x_2^4 \, .$$

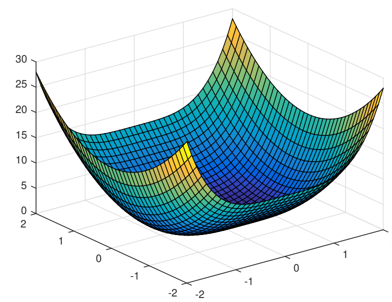
   Iz potrebnih uslova ekstrema (8) sledi

$$\begin{aligned} \frac{\partial y}{\partial x_1} &= 4x_1^3 \\ \frac{\partial y}{\partial x_1} &= 6x_2 + 8x_2^3 \, , \end{aligned}$$

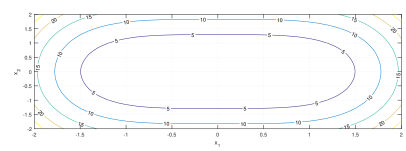
   lako dobijamo $x_1^* = x_2^* = 0$. Nastavljamo analizu dovoljnih uslova (16)-(19)

$$\begin{aligned} \frac{\partial^2 y}{\partial x_1^2} &= 12(x_1^*)^2 = 0 \\ \frac{\partial^2 y}{\partial x_1 x_2} &= 0 \\ \frac{\partial^2 y}{\partial x_2^2} &= 6 + 24(x_2^*)^2 = 6 \end{aligned}$$

   što čini da uslovi iz Tvrđenja 2 nisu zadovoljeni. Međutim, ako pogledamo da je vrednost $\Delta^2 y = 6h_2^2 \geq 0$ i izgled funkcije, slika 1, odnosno njenu projekciju na ravan x_1, x_2, slika 2, jasno vidimo da sa porastom (po apsolutnoj vrednosti) vrednosti promenljivih x_1, x_2, raste i vrednost funkcije, odnosno da je u koordinatnom početku (stacionarnoj tački) minimum kriterijuma optimalnosti.



Slika 1: Funkcija $y(x_1, x_2) = 3x_2^2 + x_1^4 + 2x_2^4$.



Slika 2: Projekcija funkcije $y(x_1, x_2) = 3x_2^2 + x_1^4 + 2x_2^4$ na ravan x_1, x_2

2. Ako je kriterijum optimalnosti, sa dve promenljive stanja dat u sledećoj formi

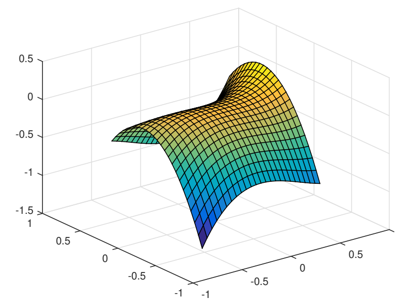
$$y(x_1, x_2) = 2x_1^2 x_2 - 2x_2^2 + x_1^3 \, ,$$

   takođe dobijamo da je stacionarna tačka $x_1^* = x_2^* = 0$, a ispitivanjem defintinosti ne možemo sa sigurnošću utvrditi karakter ekstrema, već dobijamo da je

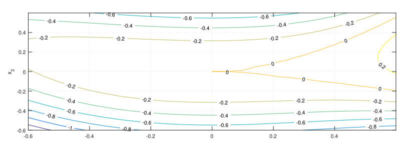
$$\Delta^2 y = -4h_2^2 \leq 0 \, .$$

Međutim, ova funkcija ima sasvim drugačiji karakter od prethodne, slika 3 .Tako na primer, ako posmatramo vrednost funkcije y u ravni x_1, x_2, slika 4, jasno vidimo da vrednost funkcije opada ili raste u zavisnosti da li se pomeramo levo ili desno od koordinatnog početka duž ose x_1. Odnosno, očevidno je da se u koordinatnom početku nalazi prevojna tačka.

Iz ova dva primera, možemo lako zaključiti da ako strogo dovoljni uslovi ekstrema Tvrđenja 2 nisu zadovoljeni, dalje ispitivanje karaktera je neophodno. Detalje o daljoj mogućoj studiji ovih problema možete naći u preporučenoj literaturi za kurs.



Slika 3: Funkcija $y(x_1, x_2) = 2x_1^2 x_2 - 2x_2^2 + x_1^3$.



Slika 4: Projekcija funkcije $y(x_1, x_2) = 2x_1^2 x_2 - 2x_2^2 + x_1^3$ na ravan x_1, x_2

1.3 *Potrebni i dovoljni uslovi ekstremuma funkcije više promenljih*

Potrebne i dovoljne uslove funkcije više promenljvih [11] nećemo posebno izvoditi i dokazivati, već ćemo logički uopštiti studiju iz prethodnog paragrafa. Posmatramo funkciju, koja u opštem slučaju zavisi od n promenljivih

$$y(\mathbf{x}) = y(x_1, x_2, x_3, \ldots, x_n) \; . \tag{20}$$

Kao što smo pokazali razvoj u red funkcije više promenljvih, svodi se na izraz (2), pri čemu su najznačajnija prva dva člana sa desne strane jednakosti, odnosno

$$y(\mathbf{x}) - y(\mathbf{x}^*) = \frac{\Delta y}{1!} + \frac{\Delta^2 y}{2!} \; . \tag{21}$$

Potrebne uslove ekstremuma dobijamo generalizacijom Tvrđenja 1.2

Tvrđenje 3. (*Potrebni uslovi ekstremuma funkcije više promenljivih*) U slučaju da postoje parcijalni izvodi prvog reda funkcije y po svim promenljivima, onda tačka ekstremuma funkcije y u tački x^* zadovoljava sistem jednačina

$$\left(\frac{\partial y}{\partial x_i}\right)^* = 0 \quad \text{za} \quad i = 1, \ldots, n$$

odnosno uslov

$$\Delta y = 0$$

Da bi opisali složene veze iz uslova (16)-(19) u slučaju funkcije dve promenljive, uveli smo matričnu formu drugih izvoda, koja za

[11] Pod više promenljivih sada mislimo, više od dve promenljive

funkciju (20) može da se zapiše na sledeći način

$$H = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_1 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_2^2} & \frac{\partial^2 y}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_n} \\ \frac{\partial^2 y}{\partial x_1 \partial x_3} & \frac{\partial^2 y}{\partial x_2 \partial x_3} & \frac{\partial^2 y}{\partial x_3^2} & \cdots & \frac{\partial^2 y}{\partial x_3 \partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_1 \partial x_n} & \frac{\partial^2 y}{\partial x_2 \partial x_n} & \frac{\partial^2 y}{\partial x_3 \partial x_n} & \cdots & \frac{\partial^2 y}{\partial x_n^2} \end{bmatrix}. \tag{22}$$

Ova matrica (22) naziva se *Heseova*[12] matrica ili *Hesijan* i u našim daljim razmatranjima ima veoma značajnu ulogu.
Kao što smo videli na primeru funkcije dve promenljive, utvrđivanje zakonomernosti kada je kvadratna forma (7) pozitivno/negativno definitna [13] nije jednostavno i dodatno se komplikuje u slučaju kada funkcija zavisi od više od dve promenljive (3). Za ispitivanje definitnosti se po pravilu koristi Silvesterova [14] teorema.

[12] Ludwig Otto Hesse (22 April 1811 – 4 August 1874) nemački matematičar, interesatno je da je njegov mentor bio Carl Gustav Jacob Jacobi, za koga se vezuje matrica **prvih** izvoda ili Jakobijan

[13] Podsećamo da je definitnost direktno vezana za ispitivanje dovoljnih uslova ekstrema.

[14] James Joseph Sylvester (3 September 1814 – 15 March 1897) britanski matematičar.

Tvrđenje 4. (*Silvesterova teorema, uslovi definitnosti*) Simetrična kvadratna matrica $\mathbf{A} = [a_{ij}]$ dimenzije n je

pozitivno definitna akko su svi glavni minori (minori oko glavne dijagonale) matrice **A** pozitivni, odnosno

$$D_i > 0 \quad \text{za} \quad i = 1, 2, \ldots, n$$

a **negativno definitna** akko glavni minori (minori oko glavne dijagonale) matrice **A** naizmenično menjajuju znak na sledeći način

$$D_{1,3,5,7,\ldots} < 0 \quad D_{2,4,6,8,\ldots} > 0$$

gde su

$$D_1 = \begin{vmatrix} a_{11} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} \ldots \quad D_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} \ldots a_{1n} \\ a_{21} & a_{22} & a_{23} \ldots a_{2n} \\ a_{31} & a_{32} & a_{33} \ldots a_{3n} \\ \vdots & \vdots & \vdots \ddots \vdots \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nn} \end{vmatrix}$$

Konačno, lako je uočiti da je kvadratna forma funkcije više promenljivih, čiju definitnost ispitujemo (3) ustvari predstavljena *Hesijanom* (22) i da se ispitivanje definitnosti svodi na analizu glavnih minora Heseove matrice u duhu Tvrđenja 4.

Tvrđenje 5. (*Dovoljni uslovi ekstremuma funkcije više promenljivih*) U slučaju da postoje parcijalni izvodi drugog reda funkcije y po svim promenljivama u okolini tačke $\mathbf{x}^*$ i ako važi $\left(\frac{\partial y}{\partial x_i}\right)^* = 0$ onda tačka ekstremuma funkcije y u tački x^* zadovoljavaju sledeće uslove

Za **strogi lokalni minimum** $D_i > 0 \quad \text{za} \quad i = 1, 2, \quad , n$

Za **strogi lokalni maksimum** $D_{1,3,5,7,\ldots} < 0 \quad D_{2,4,6,8,\ldots} > 0$

gde su

$$D_1 = \left|\frac{\partial^2 y}{\partial x_1^2}\right|_{\mathbf{x}=\mathbf{x}^*} \quad D_2 = \left|\begin{matrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} \\ \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_2^2} \end{matrix}\right|_{\mathbf{x}=\mathbf{x}^*}, \ldots,$$

$$D_n = \left|\begin{matrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_1 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_2^2} & \frac{\partial^2 y}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_n} \\ \frac{\partial^2 y}{\partial x_1 \partial x_3} & \frac{\partial^2 y}{\partial x_2 \partial x_3} & \frac{\partial^2 y}{\partial x_3^2} & \cdots & \frac{\partial^2 y}{\partial x_3 \partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_1 \partial x_n} & \frac{\partial^2 y}{\partial x_2 \partial x_n} & \frac{\partial^2 y}{\partial x_3 \partial x_n} & \cdots & \frac{\partial^2 y}{\partial x_n^2} \end{matrix}\right|_{\mathbf{x}=\mathbf{x}^*}.$$

Dobro je postaviti sledeće pitanje, ako je uslov za strogi minimum/maksimum pozitivna/negativna definitnost funkcije, da li bi pod istim uslovima semidefinitnost garantovala minimum/maskimum (bez prefiksa strogi)? Odgovor je DA. Uslovi za ispitivanje semidefinitnosti u duhu Tvrđenja 5 je složeno zato u nastavku navodimo alternativni način za ispitivanje definitnosti, koji vam je dobro poznat iz teorije sistema automatskog upravljanja.

Tvrđenje 6. (*Sopstvene vrednosti i karakter ekstrema*) Neka su $\lambda_1, \lambda_2, \ldots, \lambda_n$ sopstvene vrednosti Heseove matrice H, za matricu kažemo da je

pozitivno definitna ako su sve vrednoti $\lambda_i > 0 \ldots i = 1, \ldots, n$

pozitivno semidefinitna ako su sve vrednoti $\lambda_i \geq 0 \quad i_1, \ldots, n$

negativno definitna ako su sve vrednoti $\lambda_i < 0 \quad i_1, \ldots, n$

negativno semidefinitna ako su sve vrednoti $\lambda_i \leq 0 \quad i_1, \ldots, n$

nedefinitna ukoliko sopstvene vrednosti menjaju znak

Uslov $\leq$ i $\geq$ logički podrazumeva da postoji barem jedna vrednost različita od nule.

1.4 *Linearna regresija i metod najmanjih kvadrata*

U okviru ovog poglavlja namera nam je da grubo predstavimo postupak linearne regresije uz oslonac na metod najmanjih kvadrata. Ova dva postupka predstavljaju osnovu mašinskog učenja [15] i predstavićemo ih kao optimizacioni postupak, što oni svakako jesu, bez ulaska u varijacije algoritma, koji su najčešće posledice prilagođavanja konkretnim inženjerskim problemima.

[15] *machine learning*

Uspostavljanje funkcionalne zavisnosti između dva izmerena obeležja je problem regresije, i ukoliko je ta zavisnost linearna, postupak se očigledno naziva **linearna regresija**. Posmatraćemo dvodimenzionalni uzorak (x_i, y_i) gde je x_i nezavisna ili prediktorska promenljiva, a y_i se naziva zavisna promenljiva ili odziv (ishod). Vezu između obeležja smo pretpostavili u lineranoj formi [16]

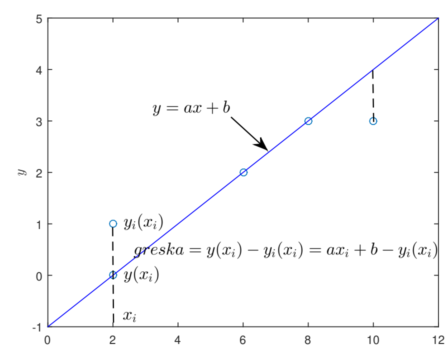
[16] Postupci regresije zavise od pretpostavljnog profila veze između nezavisnih i zavisnih promenljivih, mada suštinski u pozadini svih njih se nalaze slični optimizacioni problem, čije rešavanje zavisi od složenosti pretpostavljene funkcionalne veze.

$$y = ax + b \tag{23}$$

gde su a i b nepoznati parametri, koje želimo da odredimo u optimizacionoj proceduri, pri čemu parametar a predstavlja nagib krive, a parametar b je mesto preseka sa ordinatom slika 5 [17].
Pretpostavili smo kriterijum optimalnosti u kvadratnoj formi, gde želimo da minimizujemo grešku predviđanja (reziduala) [18], odnosno

[17] Sigurni smo da su se čitaoci upoznali sa problemom linearne regresije, makar intutitvno, primenjujući Hukov zakon i računajući Jangov modul elastičnosti.

$$I = \sum_{i=1}^{n} (y - y_i)^2 \,. \tag{24}$$



Slika 5: Ilustracija linearne regresije

[18] Odatle i ime optimizaciong problema, *metod najmanjih kvadrata*

Respektujući pretpostavljeni profil funkcionalne veze (23), kriterijum optimalnosti (24) sada postaje

$$I = \sum_{i=1}^{n} (ax_i + b - y_i)^2 \,. \tag{25}$$

Kao što smo ranije napomenuli, cilj nam je da odredimo optimalne vrednosti parametara a i b tako da je kriterijum optimalnosti (25) u minimumu. Na osnovu potrebnih uslova ekstrema Tvrđenje 3, sledi

$$\begin{aligned}
\frac{\partial I}{\partial a} &= \sum_{i=1}^{n} 2\left(ax_i + b - y_i\right) x_i = \sum_{i=1}^{n} 2\left(ax_i^2 + bx_i - y_i x_i\right) = 0 \\
\frac{\partial I}{\partial b} &= \sum_{i=1}^{n} 2\left(ax_i + b - y_i\right) = 0 \,.
\end{aligned} \tag{26}$$

Linearni sistem jednačina (26) može se napisti i u matričnoj formi

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix} \tag{27}$$

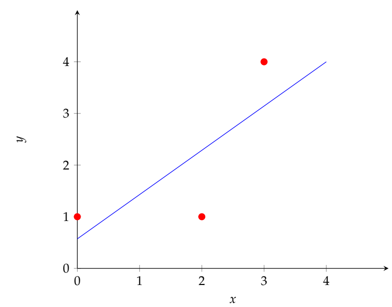
Ukoliko smo u stanju da rešimo [19] sistem jednačina (27) i dobi-

[19] Rešavanje ovog matričnog sistema jednačina je po pravilu problem sam za sebe i spada u domen linearne algebre.

jamo optimalne vrednosti parametara a i b, formalno smo završili postupak linearne regresije.

.

Čitaocima za vežbu ostavljamo da samostalno reše problem linearne regresije ukoliko su se merenjem dobila sledeća obeležja $(0,1)$, $(2,1)$ i $(3,4)$.



Slika 6: Funkcija dobijena linearnom regresijom je $y = \frac{6}{7}x + \frac{4}{7}$