

# Sadržaj

<b>1 Pogreške</b>	<b>1</b>
1.1 Vrste pogrešaka . . . . .	1
1.1.1 Pogreške zaokruživanja . . . . .	1
1.1.2 Pogreške nastale zbog nepreciznosti ulaznih po- dataka . . . . .	2
1.1.3 Pogreška metode . . . . .	2
1.1.4 Pogreška modela . . . . .	2
1.1.5 Strojna pogreška . . . . .	3
1.2 Apsolutna i relativna pogreška . . . . .	5
1.3 Signifikantne znamenke . . . . .	9
1.4 Pogreške kod izračunavanja vrijednosti funkcije . . . . .	11
1.5 Inverzni problem u teoriji pogrešaka . . . . .	14
1.6 Zadaci . . . . .	16
<b>2 Interpolacija. Spline interpolacija</b>	<b>21</b>
2.1 Interpolacija . . . . .	21
2.1.1 Lagrangeov oblik interpolacijskog polinoma . . . . .	25
2.1.2 Newtonov oblik interpolacijskog polinoma . . . . .	28
2.1.3 Ocjena pogreške . . . . .	30
2.1.4 Nultočke Čebiševljevih polinoma kao čvorovi in- terpolacije . . . . .	32
2.2 Spline interpolacija . . . . .	33
2.2.1 Linearni interpolacijski spline . . . . .	34
2.2.2 Kubični interpolacijski spline . . . . .	38
2.3 Zadaci . . . . .	43
<b>3 Rješavanje sustava linearnih jednadžbi</b>	<b>49</b>
3.1 Norma vektora i matrice . . . . .	50

3.2	Uvjetovanost sustava linearnih jednadžbi . . . . .	52
3.3	Rješavanje trokutastih sustava . . . . .	55
3.4	Gaussova metoda eliminacije . . . . .	57
3.5	LU-dekompozicija . . . . .	60
3.6	Cholesky - dekompozicija . . . . .	65
3.7	QR-dekompozicija . . . . .	67
3.8	Iterativne metode . . . . .	73
3.9	Dekompozicija na singularne vrijednosti . . . . .	77
3.10	Zadaci . . . . .	80
<b>4</b>	<b>Rješavanje nelinearnih jednadžbi</b>	<b>85</b>
4.1	Metoda bisekcije . . . . .	88
4.2	Ocjena pogreške i brzina konvergencije . . . . .	90
4.3	Metoda jednostavnih iteracija . . . . .	92
4.4	Newtonova metoda . . . . .	99
4.4.1	Modifikacije Newtonove metode . . . . .	104
4.5	Newtonova metoda za Lipschitz derivabilnu funkciju . . . . .	107
4.5.1	Lipschitz neprekidna funkcija . . . . .	108
4.5.2	Newtonova metoda . . . . .	110
4.6	Rješavanje sustava nelinearnih jednadžbi . . . . .	112
4.6.1	Metoda jednostavnih iteracija . . . . .	113
4.6.2	Newtonova metoda . . . . .	115
4.6.3	Quasi-Newtonove metode . . . . .	118
4.7	Zadaci . . . . .	119
<b>5</b>	<b>Aproksimacija funkcija</b>	<b>125</b>
5.1	Najbolja $L_2$ aproksimacija . . . . .	126
5.1.1	Fourierovi polinomi . . . . .	128
5.1.2	Ortogonalni polinomi . . . . .	135
5.1.3	Čebiševljevi polinomi . . . . .	138
5.2	Najbolja $L_\infty$ aproksimacija . . . . .	140
<b>6</b>	<b>Problemi najmanjih kvadrata</b>	<b>145</b>
6.1	Linearni problemi najmanjih kvadrata . . . . .	151
6.1.1	Rješavanje LPNK preko sustava normalnih jednadžbi . . . . .	152
6.1.2	Rješavanje LPNK pomoću QR-dekompozicije . . . . .	154
6.1.3	Rješavanje LPNK pomoću dekompozicije na singularne vrijednosti . . . . .	155

6.2	Nelinearni problemi najmanjih kvadrata . . . . .	159
6.2.1	Gauss-Newtonova metoda . . . . .	161
6.2.2	Marquardtova metoda . . . . .	165
6.3	Druge mogućnosti . . . . .	166
6.3.1	Princip najmanje sume apsolutnih odstupanja .	166
6.3.2	Princip najmanjeg maksimalnog apsolutnog od- stupanja . . . . .	168
6.4	Zadaci . . . . .	169
	<b>Literatura</b>	<b>175</b>
	NumMat . . . . .	175
	<b>Indeks</b>	<b>181</b>

# Poglavlje 1

## Pogreške

Vrlo često praksa nameće potrebu operiranja s približnim umjesto sa stvarnim veličinama, a umjesto stvarne fizikalne situacije, vrlo često u praksi promatramo idealiziranu sliku stvarnosti. Također, umjesto stvarnog rezultata, često se zadovoljavamo njegovom aproksimacijom. U svim takvim i sličnim situacijama moramo biti svjesni s kakvom pogreškom ulazimo u račun i što je još važnije, moramo unaprijed znati kakve će to posljedice imati na konačni rezultat.

### 1.1 Vrste pogrešaka

Općenito, pogreške u praktičnom radu i stručnim i znanstvenim istraživanjima možemo podijeliti u nekoliko skupina.

#### 1.1.1 Pogreške zaokruživanja

Često u praksi neki broj zaokružujemo na nekoliko prvih signifikantnih znamenki (vidi *Primjedbu 1.5*, str. 10), čime svjesno pravimo pogrešku. Neki puta u račun nije moguće uzeti stvarni broj (primjerice transcendentan broj  $\pi$  ili  $e$ ), a neki puta iz praktičnih razloga nije potrebno operirati s potpunim brojem (primjerice stopa rasta proizvodnje).

**Primjer 1.1.** *Zadan je kvadrat sa stranicom  $a = 100$  cm. Njegova površina je  $P = 10\,000$  cm<sup>2</sup>. Treba odrediti stranicu kvadrata koji će imati dvostruko manju površinu.*

S  $x$  označimo stranicu traženog kvadrata. Treba biti  $a^2 = 2x^2$ , odakle je  $x = a\frac{\sqrt{2}}{2}$ . Broj  $\sqrt{2}$  je iracionalan broj (beskonačni nepe-

riodični decimalni broj), pa smo prisiljeni u račun uzeti njegovu približnu vrijednost. Ako uzmemo  $\sqrt{2} \approx 1.41$ , dobivamo stranicu kvadrata  $x = 70.5$  cm, a odgovarajuća površina je  $P = 4970.25$  cm<sup>2</sup>. Ako uzmemo  $\sqrt{2} \approx 1.41421$ , dobit ćemo stranicu kvadrata  $x = 70.7105$  i površinu  $P = 4999.97481$  cm<sup>2</sup>. Koju ćemo vrijednost aproksimacije broja  $\sqrt{2}$  upotrijebiti, ovisi o tome s kakvom točnošću želimo dobiti rezultat.

Ovaj primjer pojednostavljena je verzija poznatog problema “*duplikacije kocke*”, koji potiče još iz antičke Grčke (vidi primjerice Brückler (2007))

### 1.1.2 Pogreške nastale zbog nepreciznosti ulaznih podataka

Često se u praksi koriste približne vrijednosti nekih konstanti koje utječu na veličinu pogreške konačnog rezultata. Primjerice, često za veličinu akceleracije sile teže uzimamo vrijednost  $9.81$  ms<sup>-2</sup>, iako se zna da to nije točna vrijednost, koja osim toga ovisi o zemljopisnom položaju tijela.

### 1.1.3 Pogreška metode

Različiti iterativni procesi (primjerice, rješavanje sustava linearnih ili nelinearnih jednadžbi, traženje lokalnih ekstrema funkcije, itd.) po prirodi su beskonačni iterativni procesi, koje u praksi zaustavljamo na nekom koraku (kad postignemo željenu točnost aproksimacije).

### 1.1.4 Pogreška modela

U praksi gotovo uvijek, zbog složenosti stvarnog problema, promatramo idealizirani model kao njegovu aproksimaciju.

**Primjer 1.2.** *Prilikom proučavanja balističkih putanja u eksternoj balistici (Cumin et al., 2009; Molitz and R.Ströbel, 1963) najvažniji problem predstavlja dobro definiranje zakona otpora zraka za balistički objekt. Da bi se postavljanje problema pojednostavilo, uvijek najprije promatramo balistički problem u vakuumu (dakle, balistički problem bez otpora zraka), čime se stvara idealizirana situacija u kojoj se problem lako rješava. No, rezultati koji se dobivaju ovakvom idealizacijom stvarnosti*

možu se jako razlikovati od stvarnih. Usporedimo za primjer maksimalni domet običnog puščanog metka (10 g) i teškog topovskog projektila (150 kg) iste početne brzine  $v_0 = 900 \text{ ms}^{-1}$  u vakuumu i u zraku ispunjenom prostorom.

maksimalni domet projektila	vakuum	zrak
puščani metak	82 km	3.5 km
teški projektil	82 km	40 km

Očigledne razlike nastale su u najvećoj mjeri zbog toga što je otpor zraka prema letećem objektu obrnuto proporcionalan težini projektila.

### 1.1.5 Strojna pogreška

Za potrebe numeričkih izračunavanja realan broj  $a \in \mathbb{R}$  obično zapisujemo u obliku

$$a = \pm m \times b^e, \quad (1.1)$$

gdje je  $b \in \mathbb{Z}$ ,  $b \geq 2$  baza,  $\frac{1}{b} \leq m < 1$  mantisa a  $e \in \mathbb{Z}$  eksponent realnog broja  $a$ . Ljudi uglavnom za bazu koriste  $b = 10$  (*dekadski sustav*), a kod većine računala koristi se  $b = 2$  (*binarni sustav*) ili  $b = 16$  (*heksadecimalni sustav*) (Ackleh et al., 2010; Deuffhard and Hohmann, 2008; Gill et al., 1991; Schwarz and Köckler, 2011; Truhar, 2010).

U kompjuterskom zapisu koristi se tzv. zapis s pomičnim zarezom (*floating-point representation*) broja

$$a = \pm 0.m_1m_2 \cdots m_t \times b^e, \quad (1.2)$$

gdje je mantisa  $m = 0.m_1m_2 \cdots m_t$  zapisana pomoću  $t$  nenegativnih cijelih brojeva  $0 \leq m_i < b$ ,  $m_1 \neq 0$ . Kažemo da je broj  $a$  reprezentiran u  $t$ -znamenkastoj floating-point aritmetici. Broj  $t$  je fiksna za svako pojedino računalo i predstavlja običnu točnost (*single precision*) računala. Ako realni broj  $a$  ima više od  $t$  znamenki, računalo će ga postupkom zaokruživanja (“rounding”) ili odbacivanja (“chopping”) pretvoriti u  $t$ -znamenkasti broj. Svako računalo obično ima i mogućnost izračunavanja i u *dvostrukoj preciznosti* ( $\approx 2t$  znamenki mantise). Eksponent  $e$

također ima svoje granice<sup>1</sup>

$$-L \leq e \leq U, \quad L, U \in \mathbb{Z}_+$$

Poruka “*Overflow*” znači da smo pokušali zapisati broj  $a > b^U$ , a “*underflow*” da smo pokušali zapisati broj  $a < b^{-L}$

*Primjedba 1.1.* Svakom računalu pridružuje se najmanji broj  $\epsilon > 0$  za koji je  $1 + \epsilon > 1$ . Broj  $\epsilon$  nazivamo **točnost računala**, a približno ga možemo izračunati sljedećim programom:

```
In[1]:= eps = 1;
        While[eps1 = eps + 1; eps1 > 1, eps = .5 eps]
        Print["točnost racunala = ", eps]
```

Primjerice za računalo s 2.66 GHz Intel(R) Core(TM)i5 CPU s 4GB RAM dobivamo  $1.42109 \times 10^{-14}$ .

Neke brojeve je moguće, a neke nije moguće prikazati u obliku (1.2). Primjerice, brojevi  $\sqrt{2}$ ,  $e$ ,  $\pi$  ne mogu se prikazati u tom obliku. Primijetite također, da je primjerice broj  $\frac{1}{10}$  moguće prikazati u obliku (1.2) s bazom  $b = 10$ , ali ne i s bazom  $b = 2$ .

Zaokruživanjem ili odbacivanjem decimala dobivamo tzv. **floating-point aproksimaciju**  $fl(a)$  broja  $a$ , za koju vrijedi

$$fl(a) = a(1 + \delta),$$

gdje je  $\delta$  pogreška aproksimacije.

**Primjer 1.3.** *Potražimo manji po apsolutnoj vrijednosti korijen kvadratne jednadžbe*

$$x^2 + 2px - q = 0, \quad p > 0, q > 0, \quad p \gg q.$$

Poznatom formulom iz srednje škole dobivamo

$$x_0 = -p + \sqrt{p^2 + q} \quad (*)$$

Ako računamo u floating-point aritmetici, lako se može dogoditi da je  $fl(p) = fl(\sqrt{p^2 + q})$  i da je rezultat  $x_0 = 0$ .

<sup>1</sup>Računala koja su građena prema IEEE standardima za bazu koriste  $b = 2$ . Vrijednost za  $t$  u običnoj preciznosti je 23, a u dvostrukoj preciznosti  $t = 52$ . Granice eksponenta su  $L = 126$  i  $U = 127$ , što znači da ta računala “prepoznaju” brojeve veličine  $1 \times 10^{-38} \leq a \leq 2 \times 10^{38}$

Ako (\*) napišemo u obliku

$$x_0 = \frac{q}{p + \sqrt{p^2 + q}}, \quad (**)$$

dobit ćemo znatno točniji rezultat.

Primjerice, u 4-znamenkastoj floating-point aritmetici za  $p = 0.125 \times 10^3$  i  $q = 0.125 \times 10^{-1}$  dobivamo

$$\begin{aligned} x_0 &= 0.00005 && \text{(prava vrijednost)} \\ x'_0 &= 0 && \text{(prema (*))} \\ x''_0 &= 0.00005 && \text{(prema (**))} \end{aligned}$$

Niže je naveden Mathematica-modul za pretvaranje realnog broja  $x$  u  $k$ -znamenkasti floating-point broj

```
fl[w_, k_] := Module[{sw},
  sw = MantissaExponent[w];
  sw[[1]] = Round[10^k sw[[1]] ]/10^k;
  ww = sw[[1]] 10^sw[[2]] //N
]
```

## 1.2 Apsolutna i relativna pogreška

Označimo s  $a$  stvarnu vrijednost neke poznate veličine (primjerice  $\sqrt{2}$ ) ili nepoznate veličine (primjerice korijen jednadžbe  $e^x - \sin x = 0$ ), a s  $a^*$  njezinu približnu vrijednost. Obično kažemo da je  $a^*$  aproksimacija od  $a$ .

**Definicija 1.1.** Razliku  $(a - a^*)$  između stvarne veličine  $a$  i njene aproksimacije  $a^*$  nazivamo **pogreška aproksimacije**. Apsolutnu vrijednost pogreške aproksimacije nazivamo **apsolutna pogreška aproksimacije** i označavamo

$$\Delta a^* = |a - a^*|. \quad (1.3)$$

**Primjer 1.4.** *Niz  $(a_n)$  definiran rekurzivnom formulom*

$$a_{n+1} = \frac{1}{2} \left( a_n + \frac{2}{a_n} \right), \quad a_0 = 1, \quad n = 0, 1, 2, \dots$$

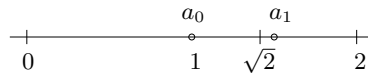


konvergira prema broju  $\sqrt{2}$  (vidi primjerice Jukić and Scitovski (2004), Primjer U1). U Tablici 1.1 prikazana su redom prva četiri člana niza (tj. prve četiri aproksimacije broja  $\sqrt{2}$ ), kao i odgovarajuće pogreške aproksimacija.

$n$	$a_n$	$a - a_n$
0	1	+0.414214
1	1.5	-0.085786
2	1.416667	-0.002453
3	1.414216	-0.000002

Tablica 1.1: Aproksimacije broja  $\sqrt{2}$

Vidi se također, da pogreške aproksimacije mogu biti pozitivni ili negativni brojevi. Apsolutna pogreška aproksimacije na brojevnom pravcu predstavlja udaljenost točke koja odgovara aproksimaciji  $a_n$  do točke koja odgovara stvarnoj vrijednosti  $a$  (vidi *Sliku 1.1*).



Slika 1.1: Aproksimacije broja  $\sqrt{2}$

U praksi stvarna vrijednost  $a$  često nije poznata, ali se zna da se pogreška aproksimacije kreće u intervalu  $[-\epsilon, \epsilon]$ , za neki  $\epsilon > 0$ . To znači da za aproksimaciju  $a^*$  vrijedi

$$|a - a^*| \leq \epsilon, \quad (1.4)$$

Spomenuti broj  $\epsilon > 0$  nazivamo **granica pogreške aproksimacije**. Nejednadžbu (1.4) možemo pisati u obliku

$$a^* - \epsilon \leq a \leq a^* + \epsilon. \quad (1.5)$$

Broj  $a^* - \epsilon$  je najmanja, a broj  $a^* + \epsilon$  najveća vrijednost koju može primiti aproksimacija broja  $a$ , pa često (posebno u tehnici) simbolički pišemo

$$a = a^* \pm \epsilon. \quad (1.6)$$

To znači da je stvarna veličina  $a$  aproksimirana brojem  $a^*$  i da pri tome apsolutna pogreška nije veća od  $\epsilon$ .

**Primjer 1.5.** *Obavljen je niz mjerenja neke veličine  $y$  i dobiveni su sljedeći rezultati*

mjerenje	1	2	3	4	5	6
$y_i$	29.2	29.3	29.25	29.28	29.24	29.26

Aritmetička sredina  $\bar{y} = 29.255$  predstavlja jednu aproksimaciju veličine  $y$ , pri čemu možemo smatrati da je granica apsolutne pogreške  $\epsilon = \max_{i=1,\dots,6} \{|y - y_i|\} = 0.055$ . Zbog toga pišemo  $y = 29.255 \pm 0.055$ .

**Primjer 1.6.** *Zadani su radijus  $r$  i visina  $h$  cilindra s odgovarajućim granicama pogreške*

$$r = 2 \pm 0.05 \text{ cm}, \quad h = 100 \pm 0.05 \text{ cm}.$$

Iako su granice pogrešaka obje veličine brojčano jednake, one nemaju isto značenje jer je veličina  $h$  relativno mnogo veća od veličine  $r$ . Točnost izmjerene veličine  $h$  cilindra relativno mnogo je veća nego točnost izmjerene veličine  $r$ . Dakle, veličinu pogreške treba promatrati u odnosu na promatranu veličinu.

**Definicija 1.2.** Omjer između apsolutne pogreške  $\Delta a^*$  i apsolutne vrijednosti veličine  $a$  ( $a \neq 0$ ) nazivamo **relativna pogreška**  $\delta a^*$  i pišemo

$$\delta a^* = \frac{\Delta a^*}{|a|}. \quad (1.7)$$

Relativna pogreška je u stvari veličina apsolutne pogreške izmjerena u odnosu na promatranu veličinu  $a$ . Zbog toga se ona često izražava u postocima (množenjem relacije (1.7) sa 100) ili u promilima (množenjem relacije (1.7) s 1000).

*Primjedba 1.2.* Kako je u praksi obično  $a^* \approx a$ , onda se relativna pogreška često izražava na sljedeći način

$$\delta a^* \approx \frac{\Delta a^*}{|a^*|}, \quad a^* \neq 0. \quad (1.8)$$

Mi ćemo nadalje relativnu pogrešku aproksimacije  $a^*$  također računati na ovaj način.

Spomenimo još da je često puta u praksi prikladno koristiti tzv. mješovitu pogrešku

$$e(a^*) = \frac{|a - a^*|}{1 + |a^*|}$$

Primijetite da je za  $|a^*| \ll 1$  mješovita pogreška  $e(a^*)$  slična apsolutnoj pogreški, a za  $|a^*| \gg 1$ , mješovita pogreška  $e(a^*)$  slična je relativnoj pogreški.

**Primjer 1.7.** U Primjeru 1.1, str. 1 relativna pogreška aproksimacije  $a^* = 1.41$  broja  $\sqrt{2}$  je

$$\delta a^* = \frac{|\sqrt{2} - 1.41|}{1.41} \approx 0.00299 \approx 3\%_0.$$

U Primjeru 1.6, str. 1 relativne pogreške veličina  $r$  i  $h$  su

$$\delta_{r^*} \approx \frac{\Delta_{r^*}}{|r^*|} = \frac{0.05}{2} = 0.025, \quad \delta_{h^*} \approx \frac{\Delta_{h^*}}{|h^*|} = \frac{0.05}{100} = 0.0005.$$

Mogli bismo reći da je veličina  $h$  zadana 50 puta pouzdanije od veličine  $r$ .

**Zadatak 1.1.** Ako je  $f : [a, b] \rightarrow \mathbb{R}$  neprekidna funkcija, onda za integral  $I = \int_a^b f(x) dx$  vrijedi

$$m(b-a) \leq I \leq M(b-a), \quad m = \min_{x \in [a, b]} f(x), \quad M = \max_{x \in [a, b]} f(x).$$

Kao jednu od aproksimacija integrala  $I$  možemo uzeti aritmetičku sredinu

$$I^* = \frac{m(b-a) + M(b-a)}{2} = (b-a) \frac{m+M}{2}.$$

Pokažite da tada vrijedi

$$\Delta I^* \leq (b-a) \frac{M-m}{2}, \quad \delta I^* \leq \frac{M-m}{M+m}.$$

Ocijenite na takav način integral  $\int_0^2 \frac{5-x}{9-x^2} dx$ , te izračunajte apsolutnu i relativnu pogrešku.

$$\text{Rješenje: } I^* = 1.1, \quad \Delta I^* \leq 0.1, \quad \delta I^* \leq 9\%$$

### 1.3 Signifikantne znamenke

Svaki pozitivni realni broj  $a$  u dekadskom sustavu možemo zapisati u obliku

$$a = b_m 10^m + b_{m-1} 10^{m-1} + \dots + b_{m-n+1} 10^{m-n+1} + b_{m-n} 10^{m-n} \dots, \quad b_m \neq 0, m \in \mathbb{Z} \quad (1.9)$$

gdje su  $b_i \in \{0, 1, \dots, 9\}$  znamenke broja  $a$ . Primjerice,

$$0.02305 = 2 \cdot 10^{-2} + 3 \cdot 10^{-3} + 0 \cdot 10^{-4} + 5 \cdot 10^{-5}.$$

Neka je

$$a^* = b_m^* 10^m + b_{m-1}^* 10^{m-1} + \dots + b_{m-n+1}^* 10^{m-n+1} + b_{m-n}^* 10^{m-n} + \dots \quad (1.10)$$

aproksimacija broja  $a$ , takva da im se podudaraju prvih  $n$  znamenki

$$b_m^* = b_m, \quad b_{m-1}^* = b_{m-1}, \quad \dots, \quad b_{m-n+1}^* = b_{m-n+1}.$$

Tada apsolutnu pogrešku  $\Delta a^*$  možemo ovako ocijeniti

$$\begin{aligned} \Delta a^* &= |a - a^*| = |(b_{m-n} - b_{m-n}^*) 10^{m-n} + (b_{m-n-1} - b_{m-n-1}^*) 10^{m-n-1} + \dots| \\ &\leq |b_{m-n} - b_{m-n}^*| 10^{m-n} + |b_{m-n-1} - b_{m-n-1}^*| 10^{m-n-1} + \dots \\ &\leq 9 \cdot 10^{m-n} + \dots = 9 \cdot 10^{m-n} (1 + 10^{-1} + \dots) \leq 9 \cdot 10^{m-n} \frac{1}{1 - \frac{1}{10}} = 10^{m-n+1}. \end{aligned}$$

Ovo je motivacija za uvođenje sljedeće definicije:

**Definicija 1.3.** Neka je

$$a^* = b_m^* 10^m + b_{m-1}^* 10^{m-1} + \dots, \quad b_m^* \neq 0, m \in \mathbb{Z}$$

aproksimacija broja  $a$  zadanog s (1.9). Kaže se da su prvih  $n$  znamenki  $b_m^*, \dots, b_{m-n+1}^*$  broja  $a^*$  **signifikantne (pouzdanе)** ako je  $n$  najveći pozitivni cijeli broj za koji vrijedi

$$\Delta a^* = |a - a^*| \leq \frac{1}{2} \times 10^{m-n+1}. \quad (1.11)$$

*Primjedba 1.3.* Aproximativni broj  $a^*$  zadan s (1.10) je signifikantan na  $k$  svojih decimala ako vrijedi

$$|a - a^*| \leq \frac{1}{2} \times 10^{-k}.$$

Naime, kako je  $n = m + k + 1$ , iz (1.11) slijedi prethodna tvrdnja.

Primjerice, ako je  $a = 0.0024357 \pm 0.3 \times 10^{-4}$ , tada je  $m = -3$  i vrijedi

$$\Delta a^* = 0.3 \times 10^{-4} < \frac{1}{2} \times 10^{-4} = \frac{1}{2} \times 10^{-3-n+1} \Rightarrow n = 2,$$

odakle zaključujemo da su samo prve dvije znamenke broja  $a^*$  signifikantne pa možemo pisati  $a^* = 0.0024$ . U ovom slučaju broj  $a^*$  ima  $k = 2 - (-3) - 1 = 4$  signifikantne decimale.

Ako je  $a = 243.5731 \pm 0.7 \times 10^{-2}$ , tada je  $m = 2$  i vrijedi

$$\Delta a^* = 0.7 \times 10^{-2} = 0.07 \times 10^{-1} < \frac{1}{2} \times 10^{-1} = \frac{1}{2} \times 10^{2-n+1} \Rightarrow n = 4,$$

odakle zaključujemo da su prve 4 znamenke broja  $a^*$  signifikantne i možemo pisati  $a^* = 243.6$ . U ovom slučaju broj  $a^*$  ima  $k = 4 - 2 - 1 = 1$  signifikantnu decimalu.

Broj 1.416667 iz *Primjera 1.4* ima prve tri, a broj 1.414216 ima prvih šest signifikantnih znamenki.

*Primjedba 1.4.* Kako je

$$|a^*| \geq a^* \geq b_m^* 10^m \geq 10^m,$$

broj signifikantnih znamenki približnog broja  $a^*$  možemo definirati i na drugi način - kao najveći pozitivni cijeli broj  $n$  za koji vrijedi

$$\delta a^* \approx \frac{|a - a^*|}{|a^*|} < \frac{1}{2} \times 10^{-n+1}. \quad (1.12)$$

Broj signifikantnih znamenki procijenjen na taj način može se od *Definicije 1.3* razlikovati najviše za jedan.

*Primjedba 1.5.* Vrlo često u praksi brojeve aproksimiramo zaokruživanjem. Pri tome treba se držati sljedećih pravila (Demidovich and Maron, 1981):

#### 1.4. POGREŠKE KOD IZRAČUNAVANJA VRIJEDNOSTI FUNKCIJE 11

- ako se iza znamenke na koju zaokružujemo broj nalazi znamenka manja od 5, znamenka na koju zaokružujemo ne mijenja se;
- ako se iza znamenke na koju zaokružujemo broj nalazi znamenka veća ili jednaka 5, znamenka na koju zaokružujemo povećava se za 1;

**Primjer 1.8.** Broj  $a = 2.351850$  treba zaokružiti na jedno, dva, tri i četiri decimalna mjesta. Redom dobivamo sljedeće aproksimacije broja  $a$ :

$$2.4, \quad 2.35, \quad 2.352, \quad 2.3519.$$

### 1.4 Pogreške kod izračunavanja vrijednosti funkcije

Zadana je realna funkcija  $n$  varijabli  $z = f(x_1, \dots, x_n)$ . Treba izračunati apsolutnu pogrešku vrijednosti funkcije u točki  $x^* = (x_1^*, \dots, x_n^*)$  ako je, u skladu s (1.6),  $x_i = x_i^* \pm \Delta x_i^*$ ,  $i = 1, \dots, n$ .

Prema teoremu o srednjoj vrijednosti je

$$\Delta z^* = |f(x_1, \dots, x_n) - f(x_1^*, \dots, x_n^*)| = \left| \sum_{i=1}^n \frac{\partial f(\tilde{x}_1, \dots, \tilde{x}_n)}{\partial x_i} (x_i - x_i^*) \right|,$$

gdje je  $\tilde{x}_i$  između  $x_i$  i  $x_i^*$ . Budući da brojevi  $\tilde{x}_i$  nisu poznati, a  $x_i^* \approx x_i$ , možemo pisati

$$\Delta z^* \leq \sum_{i=1}^n \left| \frac{\partial f(\tilde{x}_1, \dots, \tilde{x}_n)}{\partial x_i} \right| \Delta x_i^* \lesssim \sum_{i=1}^n \left| \frac{\partial f(x_1^*, \dots, x_n^*)}{\partial x_i} \right| \Delta x_i^*.$$

Zato je

$$\Delta z^* \lesssim \sum_{i=1}^n |\partial_i f^*| \Delta x_i^*, \quad \text{gdje je } \partial_i f^* := \frac{\partial f(x_1^*, \dots, x_n^*)}{\partial x_i}. \quad (1.13)$$

**Primjer 1.9.** Treba odrediti apsolutnu i relativnu pogrešku pri izračunavanju volumena kugle  $V = \frac{4}{3}r^3\pi$ , ako je radijus kugle zadan s  $r = 10.2 \pm 0.01$  cm, a broj  $\pi \approx 3.14$ .

Imamo

$$\begin{aligned} r^* &= 10.2, & \Delta r^* &= 0.01, & \pi^* &= 3.14, & \Delta \pi &= 0.0016, \\ \partial_r V(r^*, \pi^*) &= 4r^{*2}\pi^* = 4 \cdot 10.2^2 \cdot 3.14 = 1306.74, \\ \partial_\pi V(r^*, \pi^*) &= \frac{4}{3}r^{*3} = \frac{4}{3} \cdot 10.2^3 = 1414.94, \\ \Delta V^* &\approx |\partial_r V^*| \Delta r^* + |\partial_\pi V^*| \Delta \pi^* = 1306.74 \cdot 0.01 + 1414.94 \cdot 0.0016 = 15.33. \end{aligned}$$

Dakle,

$$V = \frac{4}{3}r^{*3}\pi^* \pm \Delta V^* = 4442.92 \pm 15.33 \text{ cm}^3.$$

Kako je  $\Delta V^* \approx 0.5 \times 10^{3-2+1}$ , signifikantne su samo prve dvije znamenke pa možemo zaokružiti  $V^* \approx 4400 \text{ cm}^3$ . Relativna pogreška prema (1.8) je

$$\delta V^* \approx \frac{\Delta V^*}{|V^*|} = \frac{15.33}{4442.92} = 0.00345 \approx 3.5\%.$$

Prema *Primjedbi 1.4* to bi značilo da su prve tri znamenke signifikantne, tj.  $V^* \approx 4440 \text{ cm}^3$ .

Za navedeni račun možemo koristiti kratki *Mathematica*-program:

```
In[1]:= V[r_, pi_] := 4 r^3 pi/3;
w = {r -> 10.2, pi -> 3.14};
Vz = V[r, pi] /. w
DV = .01 Abs[D[V[r, pi], r] /. w] + .0016 Abs[D[V[r, pi], pi] /. w]
dV = DV/Vz
```

**Primjer 1.10.** *Treba izračunati pogrešku pri izračunavanju površine kruga radijusa 100 m, ako uzmemo  $\pi \approx 3.14$ .*

Imamo

$$\begin{aligned} \pi^* &= 3.14, & \Delta \pi^* &= 0.0016, & r^* &= 100, & \Delta r^* &= 0, \\ \Delta P^* &\approx |\partial_r P^*| \Delta r^* + |\partial_\pi P^*| \Delta \pi^* = 0 + 10\,000 \cdot 0.0016 = 16 \text{ m}^2, \\ \delta P^* &\approx \frac{\Delta P^*}{|P^*|} = \frac{16}{31\,400} \approx 0.5\%. \end{aligned}$$

Približna površina je  $P^* = r^2 \pi^* = 31\,400 \pm 16 \text{ m}^2$ , a stvarna površina je  $P \approx 31\,415.93 \text{ m}^2$ , što se podudara s gornjom procjenom.

#### 1.4. POGREŠKE KOD IZRAČUNAVANJA VRIJEDNOSTI FUNKCIJE 13

**Primjer 1.11.** Treba procijeniti pogrešku kod izračunavanja volumena valjka ( $V = r^2\pi h$ ), ako su mu radijus  $r \approx 2$  m i visina  $h \approx 3$  m zadani s točnošću na dvije decimale, dok broj  $\pi$  smatramo egzaktnim.

Imamo

$$\begin{aligned} r^* &= 2, & h^* &= 3, & \pi^* &= 3.14159, \\ \Delta r^* &= 0.005, & \Delta h^* &= 0.005, & \Delta \pi &= 0.00000265, \\ \partial_r V^* &= 2r^*\pi^*h^* = 2 \cdot 2 \cdot 3.14159 \cdot 3 = 37.6991, \\ \partial_\pi V^* &= r^{*2}h^* = 2^2 \cdot 3 = 12, & \partial_h V^* &= r^{*2}\pi^* = 2^2 \cdot 3.14159 = 12.5664, \\ \Delta V^* &\approx |\partial_r V^*|\Delta r^* + |\partial_\pi V^*|\Delta \pi^* + |\partial_h V^*|\Delta h^* \approx 0.25 \text{ m}^3, \\ \delta V^* &\approx \frac{\Delta V^*}{|V^*|} = \frac{0.25}{37.6991} = 0.0066 \approx 7\%. \end{aligned}$$

To znači da su u broju  $V^* = 37.6991$  samo prve dvije znamenke signifikantne, tj.  $V^* = 38$ .

*Primjedba 1.6.* Prilikom izvođenja računskih operacija s približnim brojevima potrebno je unaprijed znati s kakvom točnošću možemo očekivati dobiveni rezultat. Primjenom formule (1.13) uz pretpostavku da su apsolutne pogreške ulaznih veličina malene i to je moguće procijeniti:

- Prilikom zbrajanja ili oduzimanja dva približna broja dobivamo broj čija apsolutna pogreška ne prelazi zbroj apsolutnih pogrešaka pribrojnika. Naime, ako je  $z = x + y$ , onda je

$$\Delta z^* = |z - z^*| = |x + y - x^* - y^*| \leq |x - x^*| + |y - y^*| = \Delta x^* + \Delta y^*.$$

- Prilikom množenja dva približna broja dobivamo broj čija relativna pogreška ne prelazi zbroj relativnih pogrešaka članova produkta. Naime, ako je  $z = x \cdot y$ , onda prema (1.13) vrijedi

$$\Delta z^* \approx |\partial_x z^*|\Delta x^* + |\partial_y z^*|\Delta y^* = |y^*|\Delta x^* + |x^*|\Delta y^*,$$

odakle slijedi

$$\delta z^* \approx \frac{\Delta z^*}{|z^*|} \approx \frac{|y^*|\Delta x^* + |x^*|\Delta y^*}{|x^*y^*|} = \frac{\Delta x^*}{|x^*|} + \frac{\Delta y^*}{|y^*|} = \delta x^* + \delta y^*.$$



- Prilikom dijeljenja dva približna broja dobivamo broj čija relativna pogreška ne prelazi zbroj relativnih pogrešaka članova kvocijenta. Naime, ako je  $z = x/y$ , onda prema (1.13) vrijedi

$$\Delta z^* \approx |\partial_x z^*| \Delta x^* + |\partial_y z^*| \Delta y^* = \frac{1}{|y^*|} \Delta x^* + \left| \frac{x^*}{-y^{*2}} \right| \Delta y^*,$$

odakle slijedi

$$\delta z^* \approx \frac{\Delta z^*}{|z^*|} \approx \left| \frac{y^*}{x^*} \right| \left( \frac{\Delta x^*}{|y^*|} + \frac{|x^*|}{|y^*|^2} \Delta y^* \right) = \frac{\Delta x^*}{|x^*|} + \frac{\Delta y^*}{|y^*|} \approx \delta x^* + \delta y^*.$$

*Primjedba 1.7.* U primjenama se također upotrebljava tzv. **standardna pogreška funkcije** (naročito ako je broj nezavisnih varijabli  $x_1, \dots, x_n$  nešto veći – vidi primjerice DAHLQUIST (1972)). Uz pretpostavku (vidi Benšić and Šuvak (2012)) da su sve pogreške  $x_i - x_i^*$ ,  $i = 1, \dots, n$  nezavisne i normalno distribuirane slučajne varijable s matematičkim očekivanjem 0 i standardnim devijacijama  $\Delta x_1^*, \dots, \Delta x_n^*$ , **standardna pogreška funkcije**  $z = f(x_1, \dots, x_n)$  zadana je s

$$\Delta z^* \approx \left( \sum_{i=1}^n (\partial_i f^*)^2 (\Delta x_i^*)^2 \right)^{1/2}.$$

## 1.5 Inverzni problem u teoriji pogrešaka

U prethodnom poglavlju pokazali smo kako se može procijeniti pogreška prilikom izračunavanja vrijednosti funkcije ako su vrijednosti nezavisnih varijabli približni brojevi. U praksi se obično postavlja obrnuti problem:

*s kojom točnošću moramo uzeti vrijednosti nezavisnih varijabli promatrane funkcije, tako da njezina vrijednost bude u granicama unaprijed zadane točnosti?*

Za rješavanje ovog problema koristit ćemo tzv. "princip jednakih efekata":

*pretpostavljamo da svi parcijalni diferencijali iz (1.13) imaju jednaki utjecaj na veličinu apsolutne pogreške, tj.*

$$|\partial_1 f^*| \Delta x_1^* = |\partial_2 f^*| \Delta x_2^* = \dots = |\partial_n f^*| \Delta x_n^* \quad (1.14)$$

Iz (1.13) i (1.14) lako slijedi

$$\Delta x_i^* \approx \frac{\Delta z^*}{n |\partial_i f^*|}, \quad i = 1, \dots, n. \quad (1.15)$$

**Primjer 1.12.** *Kolike smiju biti apsolutne pogreške radijusa  $r$  i visine  $h$  valjka volumena  $V = 12\pi \pm 0.1 \text{ m}^3$  (dozvoljena tolerancija je 100 litara!)? Neka je pri tome  $r^* = 2 \text{ m}$ . Broj  $\pi$  uzet ćemo dovoljno točno da njegova pogreška ne utječe na rezultat.*

Najprije iz  $12\pi = r^{*2}\pi h^*$  slijedi  $h^* = 3 \text{ m}$ . Kako je  $\partial_r V^* = 2\pi r^* h^* = 37.7$ ,  $\partial_h V^* = r^{*2}\pi = 12.57$ , iz (1.15) dobivamo

$$\Delta r^* \approx \frac{\Delta V^*}{2|\partial_r V^*|} = 0.001 \text{ m}, \quad \Delta h^* \approx \frac{\Delta V^*}{2|\partial_h V^*|} = 0.004 \text{ m}.$$

Dakle, radijus  $r$  moramo imati s točnošću do na 1 mm, a visinu do na 4 mm.

**Primjer 1.13.** *Treba načiniti spremnik za tekućinu oblika prikazanog na Slici 1.2 volumena 1000 l, širine baze  $2r^* = 80 \text{ cm}$  i visine baze  $2r^* + h^* = 100 \text{ cm}$ . Kolika mora biti dužina  $d$  spremnika? Kolike smiju biti pogreške  $\Delta r^*$ ,  $\Delta h^*$ ,  $\Delta d^*$ , ako je dozvoljena tolerancija volumena 10 litara?*

Slika 1.2 Spremnik za tekućinu

Kako je  $V^* = r^* d^* (r^* \pi + 2h^*)$ , onda je

$$d^* = \frac{V^*}{r^* (r^* \pi + 2h^*)} = \frac{1\,000\,000}{40(40\pi + 40)} \approx 151 \text{ cm}.$$

Kako je

$$\partial_r V^* = 2 d^* (r^* \pi + h^*) = 43\,990,$$

$$\partial_h V^* = 2 d^* r^* = 12\,080, \quad \partial_d V^* = r^* (r^* \pi + 2h^*) = 6\,627,$$

onda iz (1.15) dobivamo

$$\Delta r^* \approx \frac{\Delta V^*}{3 |\partial_r V^*|} = \frac{10\,000}{3 \cdot 43\,990} = 0.08 \text{ cm},$$

$$\Delta h^* \approx \frac{\Delta V^*}{3 |\partial_h V^*|} = 0.3 \text{ cm},$$

$$\Delta d^* \approx \frac{\Delta V^*}{3 |\partial_d V^*|} = 0.5 \text{ cm}.$$

Dakle,

$$r = 40 \pm 0.08 \text{ cm}, \quad h = 20 \pm 0.3 \text{ cm}, \quad d = 151 \pm 0.5 \text{ cm}.$$

*Primjedba 1.8.* Više detalja o teoriji pogrešaka i primjenama može se vidjeti primjerice kod (Demidovich and Maron, 1981). Posebno o problemu kompjuterske preciznosti u numeričkim metodama može se vidjeti kod (Ackleh et al., 2010; Deuffhard and Hohmann, 2008; Gill et al., 1981, 1991; ?; Schwarz and Köckler, 2011; Stoer and Bulirsch, 2002).

## 1.6 Zadaci

*Zadatak 1.2.* Korištenjem poznatih formula izračunajte po apsolutnoj vrijednosti manji korijen kvadratne jednadžbe  $2x^2 + 984x - 0.05 = 0$  u a) 2-znamenkastoj, b) 6-znamenkastoj, c) 7-znamenkastoj, d) 8-znamenkastoj i e) 10-znamenkastoj floating-point aritmetici. Usporedite rezultat s točnom vijednošću tog korijena:  $x_0 = 0.000050813$ . Kako treba napisati formulu da bi rješenje i u nižoj floating-point aritmetici bilo prihvatljivo? Kakva su ta rješenja?

*Uputa:* Koristite priloženi *Mathematica*-program.

*Rješenje:* a) 4, b) 0, c) 0.0001, d) 0.00005, e) 0.0000508.

*Zadatak 1.3.* U  $k$ -znamenkastoj ( $k = 3, 4, 6, 8$ ) floating-point aritmetici izračunajte vrijednost polinoma  $P(x) = -32x^3 + 96x^2 - 90x + 27$  u točki  $a = 0.6$  direktno i koristeći Hornerovu shemu. Što primjećujete?

*Zadatak 1.4.* Broj  $a^*$  je aproksimacija broja  $a$ . Za koliko znamenki broja  $a^*$  možemo reći da su signifikantne?

- a)  $a = 23.395$ ,  $a^* = 23.40$     b)  $a = 0.00275$ ,  $a^* = 0.00266$   
 c)  $a = 243317$ ,  $a^* = 243315$     d)  $a = 0.012815$ ,  $a^* = 0.0130$

- Rješenje: a)  $\Delta a^* = 0.5 \cdot 10^{-2} = 0.5 \times 10^{1-4+1}$ , 4 signifikantne znamenke  
 b)  $\Delta a^* = 0.09 \times 10^{-3} < 0.5 \times 10^{-3-1+1}$ , 1 signifikantna znamenka  
 c)  $\Delta a^* = 2 < 0.5 \times 10^1 = 0.5 \times 10^{5-5+1}$ , 5 signifikantnih znamenki  
 d)  $\Delta a^* = 0.185 \times 10^{-3} < 0.5 \times 10^{-2-2+1}$ , 2 signifikantne znamenke

Zadatak 1.5. Broj  $a$  zaokružite na jedno, dva, tri i četiri decimalna mjesta.

- a)  $a = 2.71828$    b)  $a = 0.753550$    c)  $a = 0.97965$

- Rješenje: a) 2.7, 2.72, 2.718, 2.7183  
 b) 0.8, 0.75, 0.754, 0.7536  
 c) 1, 0.98, 0.980, 0.9797

Zadatak 1.6. Zadana je godišnja kamatna stopa  $p$ . Ispodgodišnji konformni kamatnjak  $p_m$  (vidi primjerice SCITOVSKI (1993)) računa se po formuli

$$p_m = 100 \left( \sqrt[m]{1 + \frac{p}{100}} - 1 \right),$$

gdje je  $m$  broj jednakih podintervala na koje dijelimo godinu. Jedna aproksimacija konformnog kamatnjaka koja se koristi u praksi je tzv. *relativni kamatnjak*, dobiva se kao linearna aproksimacija *konformnog kamatnjaka* pomoću Taylorove formule

$$p_r = \frac{p}{m}.$$

Ako je  $m = 12$ , izračunajte apsolutnu i relativnu pogrešku aproksimacije, te relativnu pogrešku u % za različite vrijednosti godišnje kamatne stope

$$p = 2\%, 3.5\%, 5\%, 7.5\%, 12\%, 24\%, 50\%, 120\%, 360\%, 1000\%.$$

$$\text{Rješenje: } \Delta p_r = |p_m - p_r|, \quad \delta p_r = \frac{\Delta p_r}{p_m}$$

$p$	2%	3.5%	7.5%	12%	24%	50%	120%	360%	1000%
$\Delta p_r$	0.0015	0.0046	0.02	0.05	0.19	0.73	3.21	16.44	61.21
$\delta p_r$	0.0091	0.0159	0.03	0.05	0.11	0.21	0.47	1.21	2.76
%	1	1.6	3	5	11	21	47	121	276

Zadatak 1.7. Kako se ponaša apsolutna i relativna pogreška u prethodnom zadatku ako za  $m$  uzmemo 2, 4 ili 365?

*Zadatak 1.8.* Neka je  $x = 2.00 \pm 0.005$ ,  $y = 3.00 \pm 0.005$ ,  $z = 4.00 \pm 0.005$ . Procijenite pogrešku prilikom izračunavanja vrijednosti niže navedenih funkcija u točki  $(x^*, y^*, z^*)$

$$a) \quad f(x, y, z) = 3x + y - z, \quad b) \quad f(x, y, z) = x \frac{y}{z}, \quad c) \quad f(x, y) = x \sin \frac{y}{40}.$$

Rješenje: a)  $\Delta f^* \approx 0.025$ , b)  $\Delta f^* \approx 0.008$ , c)  $\Delta f^* \approx 0.003$

*Zadatak 1.9.* Procijenite apsolutnu i relativnu pogrešku pri izračunavanju površine  $P$  pravokutnika sa stranicama

$$a = 29.3 \pm 0.05 \text{ cm}, \quad b = 18.1 \pm 0.04 \text{ cm}.$$

Rješenje:  $\Delta P^* \approx 2.08 \text{ cm}^2$ ,  $\delta P^* \approx 0.0039 \approx 4\%$ .

*Zadatak 1.10.* Procijenite apsolutnu i relativnu pogrešku pri izračunavanju volumena  $V$  paralelepipeda sa stranicama

$$a = 29.3 \pm 0.05 \text{ cm}, \quad b = 18.1 \pm 0.04 \text{ cm}, \quad c = 11.2 \pm 0.03 \text{ cm}.$$

Rješenje:  $\Delta V \approx 39.17 \text{ cm}^3$ ,  $\delta V^* \approx 0.0066 \approx 7\%$ .

*Zadatak 1.11.* Procijenite apsolutnu i relativnu pogrešku pri izračunavanju volumena  $V = 2ar^2\pi^2$  torusa koji nastaje rotacijom kruga radijusa  $r$ , čije je središte za  $a$  udaljeno od centra rotacije, ako je

$$\begin{aligned} a) \quad r &= 7.5 \pm 0.05 \text{ cm}, & a &= 30 \pm 0.1 \text{ cm}. \\ b) \quad r &= 20 \pm 0.2 \text{ cm}, & a &= 100 \pm 0.5 \text{ cm}. \end{aligned}$$

Rješenje: a)  $\Delta V^* \approx 555.2 \text{ cm}^3$ ,  $\delta V^* \approx 0.017 \approx 2\%$ .  
b)  $\Delta V^* \approx 19739.2 \text{ cm}^3$ ,  $\delta V^* \approx 0.025 \approx 2.5\%$ .

*Zadatak 1.12.* Prema poučku o kosinusima, ako su za neki trokut poznate dvije stranice  $a, b$  i kut  $\gamma$  između njih, onda se treća stranica  $c$  može izračunati po formuli

$$c^2 = a^2 + b^2 - 2ab \cos \gamma.$$

Procijenite apsolutnu i relativnu pogrešku pri izračunavanju treće stranice  $c$  trokuta, ako je

$$\begin{aligned} a) \quad a &= 10.2 \pm 0.05 \text{ cm}, & b &= 7.5 \pm 0.05 \text{ cm}, & \gamma &= 31.3^\circ \pm 0.05^\circ. \\ b) \quad a &= 230.7 \pm 0.05 \text{ m}, & b &= 123.5 \pm 0.05 \text{ m}, & \gamma &= 27.4^\circ \pm 0.05^\circ. \end{aligned}$$

Rješenje: kut  $\gamma$  najprije treba pretvoriti u radijane. Dobivamo:

- a)  $\Delta c^* \approx 0.0524 \text{ cm}$ ,  $\delta c^* \approx 0.01 \approx 1\%$ ,  $c^* \approx 5.4367 \text{ cm}$ .  
 b)  $\Delta c^* \approx 0.16 \text{ m}$ ,  $\delta c^* \approx 0.0012 \approx 0.1\%$ ,  $c^* \approx 133.733 \text{ m}$ .

Zadatak 1.13. Procijenite apsolutnu i relativnu pogrešku prilikom:

- a) zbrajanja  $n$  približnih brojeva  
 b) množenja  $n$  približnih brojeva  
 c) kvadriranja jednog približnog broja  
 d) potenciranja s  $n$ -tom potencijom jednog približnog broja  
 e) izračunavanja drugog korijena nekog približnog broja  
 f) izračunavanja  $n$ -tog korijena nekog približnog broja  
 g) izračunavanja prirodnog logaritma nekog pozitivnog realnog broja

Zadatak 1.14. Treba načiniti tetraedar volumena  $V = 1 \pm 0.01 l$ . S kojom točnošću treba imati stranicu a tetraedra?

Uputa: Volumen tetraedra stranice  $a$  je  $V = \frac{a^3\sqrt{2}}{12}$ .

Rješenje:  $\Delta a^* \approx 0.0068 \text{ cm}$ ,  $\delta a^* \approx 0.0033 \approx 0.3\%$ ,  $a^* \approx 2.03965 \text{ dm}$ .

Zadatak 1.15. Ako je  $z = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$  i  $x_i = x_i^* \pm \Delta x_i^*$ , pokažite da je  $\delta z^* \approx \sum_{i=1}^n |\alpha_i| \delta x_i^*$ .

Uputa: Najprije logaritmirajte  $\ln z = \alpha_1 \ln x_1 + \cdots + \alpha_n \ln x_n$ .

Zadatak 1.16. Procijenite apsolutnu pogrešku pri izračunavanju vrijednosti funkcije  $f(x) = \sqrt{1+x^2} - x + 200$ , ako je  $x = 100 \pm 1$ .

Rješenje:  $f(x) = 200.0050 \pm 0.5 \times 10^{-4}$ .

Zadatak 1.17. Balistički projektil ispaljen u vakuumu početnom brzinom  $v_0$  pod kutem  $\alpha$  u odnosu na ravni teren, past će na udaljenosti  $d = \frac{v_0^2}{g} \sin 2\alpha$ , gdje je  $g = 9.81 \text{ ms}^{-2}$  akceleracija sile teže (vidi *Primjer 4.11*, str. 106). Procijenite apsolutnu i relativnu pogrešku pri izračunavanju daljine gađanja ako je

- a)  $v_0 = 115 \pm 1 \text{ ms}^{-1}$ ,  $\alpha = 15^\circ \pm 0.1^\circ$ .  
 b)  $v_0 = 900 \pm 10 \text{ ms}^{-1}$ ,  $\alpha = 15^\circ \pm 0.1^\circ$ .

Rješenje: a)  $\Delta d^* \approx 15.8 \text{ m}$ ,  $\delta d^* \approx 0.023 \approx 2\%$ ,  $d^* \approx 674.057 \text{ m}$ .  
 b)  $\Delta d^* \approx 1167.04 \text{ m}$ ,  $\delta d^* \approx 0.028 \approx 3\%$ ,  $d^* \approx 41\,284.4 \text{ m}$ .

*Zadatak 1.18.* Maksimalni domet balističkog projektila na ravnom terenu u vakuumu računa se po formuli (vidi *Primjer 4.11*, str. 106):  $X = \frac{v_0^2}{g}$ . Procijenite apsolutnu i relativnu pogrešku pri izračunavanju maksimalnog dometa projektila, ako je

$$\text{a) } v_0 = 115 \pm 1 \text{ ms}^{-1}, \quad g = 9.81 \pm 0.001 \text{ ms}^{-2}.$$

$$\text{b) } v_0 = 900 \pm 10 \text{ ms}^{-1}, \quad g = 9.81 \pm 0.001 \text{ ms}^{-2}.$$

$$\text{Rješenje: a) } \Delta X^* \approx 23.58 \text{ m}, \quad \delta X^* \approx 0.017 \approx 1.7\%.$$

$$\text{b) } \Delta X^* \approx 1843.28 \text{ m}, \quad \delta X^* \approx 0.022 \approx 2.2\%.$$

*Zadatak 1.19.* Vrijeme leta balističkog projektila ispaljenog pod kutem  $\alpha$  početnom brzinom  $v_0$  na ravnom terenu u vakuumu računa se po formuli (vidi *Primjer 4.11*, str. 106):  $T = \frac{2v_0}{g} \sin \alpha$ . Procijenite apsolutnu i relativnu pogrešku pri izračunavanju vremena leta  $T$  projektila, ako je

$$\text{a) } v_0 = 115 \pm 1 \text{ ms}^{-1}, \quad g = 9.81 \pm 0.001 \text{ ms}^{-2}, \quad \alpha = 15^\circ \pm 0.1^\circ$$

$$\text{b) } v_0 = 900 \pm 10 \text{ ms}^{-1}, \quad g = 9.81 \pm 0.001 \text{ ms}^{-2}, \quad \alpha = 15^\circ \pm 0.1^\circ$$

$$\text{Rješenje: a) } \Delta T^* \approx 0.093 \text{ s}, \quad \delta T^* \approx 0.015 \approx 1.6\%, \quad T^* \approx 6.068 \text{ s}.$$

$$\text{b) } \Delta T^* \approx 0.84 \text{ s}, \quad \delta T^* \approx 0.018 \approx 2\%, \quad T^* \approx 47.49 \text{ s}.$$

*Zadatak 1.20.* Balističkim projektilom početne brzine  $v_0 = 300 \text{ ms}^{-1}$  u vakuumu na ravnom terenu gađa se točka udaljena  $d = 5000 \text{ m}$ . S kojom točnošću treba imati izlazni kut  $\alpha$  (u stupnjevima), ako se tolerira apsolutna pogreška pogotka cilja od  $\Delta d^* = 50 \text{ m}$ ?

Uputa: Vidi *Zadatak 1.17*. Veličine  $v_0$  i  $g$  smatraju se točnim.

Rješenje:  $\alpha \approx 0.288 \pm 0.003$  (odnosno  $\alpha \approx 16.5^\circ \pm 0.18^\circ$ ).

## Poglavlje 2

# Interpolacija. Spline interpolacija

Često za neku funkciju  $f$  nemamo analitički izraz, ali poznajemo njenu vrijednost u konačno točaka:  $x_0 < x_1 < \dots < x_n$ . U području podataka, tj. u intervalu  $[x_0, x_n]$  treba aproksimirati funkciju  $f$  nekom jednostavnijom poznatom funkcijom  $g$ , tako da bude

$$g(x_i) = f(x_i), \quad i = 0, 1, \dots, n. \quad (2.1)$$

Problem određivanja funkcije  $g$  na osnovi zahtjeva (2.1) nazivamo **problem interpolacije**. Funkcija  $g$  obično se bira u klasi polinoma, trigonometrijskih, eksponencijalnih, racionalnih ili nekih drugih jednostavnih funkcija. Kada odredimo funkciju  $g$ , onda možemo procijeniti vrijednosti funkcije  $f$  u nekoj točki  $x$ ,  $x \neq x_i$ , tako da stavimo  $f(x) \approx g(x)$ .

### 2.1 Interpolacija

Pretpostavimo da je funkcija  $f : [a, b] \rightarrow \mathbb{R}$  neprekidna i da je njena vrijednost poznata u  $n + 1$  točaka (zvat ćemo ih **čvorovi interpolacije**)

$$a \leq x_0 < x_1 < \dots < x_n \leq b, \quad (2.2)$$

u kojima prima vrijednosti

$$y_i = f(x_i), \quad i = 0, 1, \dots, n. \quad (2.3)$$



Na osnovi podataka  $(x_i, y_i)$ ,  $i = 1, \dots, n$ , treba rekonstruirati funkciju  $f$ , tj. pronaći novu funkciju  $g : [a, b] \rightarrow \mathbb{R}$ , tako da bude  $g(x_i) = y_i$ . Pri tome pogrešku aproksimacije obično izražavamo u  $L_\infty$ -normi

$$\|f - g\|_\infty = \max_{x \in [a, b]} |f(x) - g(x)|$$

ili u  $L_2$ -normi

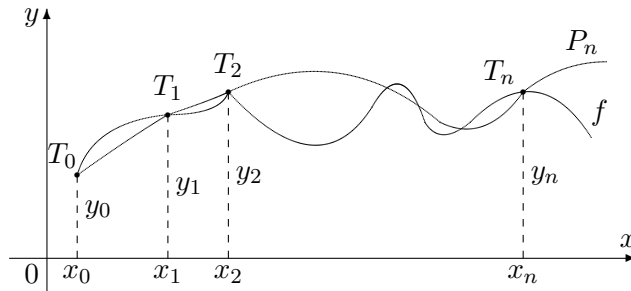
$$\|f - g\|_2 = \sqrt{\int_a^b (f(x) - g(x))^2 dx}.$$

Prijedimo na problem interpolacije polinomom. Treba pronaći interpolacijski polinom  $P_n$  stupnja  $n$

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

tako da bude (vidi Sliku 2.1)

$$P_n(x_i) = y_i, \quad i = 0, 1, \dots, n. \tag{2.4}$$



Slika 2.1: Interpolacija funkcije

Geometrijski to znači da treba pronaći polinom  $P_n$  čiji graf prolazi točkama  $T_i = (x_i, y_i)$ ,  $i = 0, 1, \dots, n$  (Slika 2.1). Koeficijente polinoma  $P_n$  mogli bismo odrediti iz uvjeta (2.4) rješavajući sustav od  $(n + 1)$  jednačine s  $(n + 1)$ -nom nepoznanicom

$$\begin{aligned} a_n x_0^n + a_{n-1} x_0^{n-1} + \dots + a_1 x_0 + a_0 &= y_0 \\ a_n x_1^n + a_{n-1} x_1^{n-1} + \dots + a_1 x_1 + a_0 &= y_1 \\ \dots & \\ a_n x_n^n + a_{n-1} x_n^{n-1} + \dots + a_1 x_n + a_0 &= y_n \end{aligned} \tag{2.5}$$

Determinanta matrice ovog sustava je poznata *Vandermondova determinanta*, koja je uz uvjet (2.2) različita od nule (vidi primjerice (Kurepa, 1967)). Zato je uz uvjet (2.2) ovaj sustav uvijek rješiv i ima jedinstveno rješenje.

Ako bismo za dane podatke  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  tražili polinom stupnja  $< n$ , interpolacijski polinom ne mora uvijek postojati (broj nepoznanica u sustavu (2.2) manji je od broja jednadžbi). Ovaj problem detaljnije ćemo razmatrati u poglavlju 6. *Problemi najmanjih kvadrata* str. ?? (vidi takodjer Truhar (2010)). Ako bi tražili polinom stupnja većeg od  $n$ , interpolacijski polinom nije jedinstven (broj nepoznanica u sustavu (2.2) veći je od broja jednadžbi).

Drugi problem je rješavanje sustava (2.5). Za veliki  $n$  i međusobno relativno bliske čvorove interpolacije tu ćemo naići na ozbiljne numeričke probleme. Matrica sustava (2.5) obično je vrlo loše uvjetovana (vidi t.3.2, str.52). Zbog toga ćemo se u nastavku upoznati s drugim metodama za rješavanje ovog problema.

Niže navedeni Mathematica-program rješava problem interpolacije zadane funkcije  $f$  u izabranim čvorovima rješavanjem sustava (2.5). Najprije definiramo podatke. U “komentarima” je drugi primjer funkcije i mogućnost izbora čvorova kao nultočaka Čebiševljevih polinoma.

```
In[1]:=f[x_] := 1 - Abs[x - 1]; a = 0; b = 2; n = 10;
(*f[x_] := Abs[Sin[x]]; a=0; b=2Pi;*)
x = Table[(b - a) i/n, {i, 0, n}]; y = f[x];
(* x=Table[(a+b)/2+N[.5(b-a)Cos[(n-k)Pi/n]],{k,0,n}];
y=f[x]; *)
(* cvorovi kao nultočke Cebisevljevog polinoma *)
pod = Table[{x[[i]], y[[i]]}, {i, n + 1}];
Print[TableForm[Transpose[N[pod, 2]]]]
```

Nakon toga definiramo matricu sustava prema (2.5).

```
In[2]:=mat = Table[0, {i, 0, n}, {j, 0, n}];
Do[
  Do[
    If[j < 2, mat[[i, 1]] = 1, mat[[i, j]] = N[x[[i]]^(j-1)]
    , {i, n + 1}
    , {j, n + 1}];
```

Nakon toga interpolacijski polinom dobivamo rješavanjem sustava (2.5) pomoću gotovog Mathematica-modula.

```
In[3]:=koef = LinearSolve[mat, y];
      intpol[xx_] := Sum[koef[[i]] xx^(i - 1), {i, n + 1}];
```

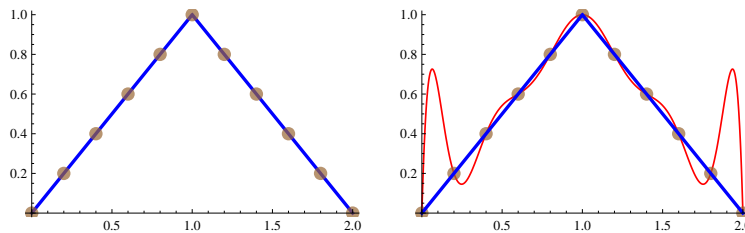
Dobiveni interpolacijski polinom i podatke prikazat ćemo na sljedeći način.

```
In[4]:=slf=Plot[f[x], {x,x[[1]],x[[n+1]]},
      PlotStyle->{Blue,Thickness[.01]}];
      slpod=ListPlot[pod,
      PlotStyle->{Brown, PointSize[.04], Opacity[.7]}];
      slp = Plot[intpol[xx], {xx,x[[1]],x[[n+1]]},
      PlotStyle -> {Red, Thickness[.005]}];
      GraphicsGrid[{{Show[slf,slpod], Show[slp,slpod,slf]}},
      ImageSize->500]
```

Uvjetovanost matrice sustava (2.5) u  $l_\infty$  normi dobivamo na sljedeći način.

```
In[4]:=cond = Norm[mat, Infinity] Norm[Inverse[mat], Infinity];
      Print["Uvjetovanost =", cond]
```

**Primjer 2.1.** *Odredimo interpolacijski polinom koji prolazi točkama  $T_i = (x_i, y_i)$ ,  $i = 1, \dots, n$ , gdje je  $n = 10$ ,  $x_i = 2\frac{i}{n}$ ,  $y_i = f(x_i)$ ,  $f: [0, 2] \rightarrow \mathbb{R}$ ,  $f(x) = 1 - |x - 1|$ . Na Slici 2.2 prikazan je graf funkcije  $f$ , točke  $T_i$  i interpolacijski polinom čiji graf prolazi točkama  $T_i$ . Primijetite da su pogreške interpolacije značajne na rubovima intervala  $[0, 2]$ . Uvjetovanost odgovarajućeg sustava jednadžbi je visoka:  $cond = 8.61 \times 10^8$*



Slika 2.2: Interpolacija funkcije  $f: [0, 2] \rightarrow \mathbb{R}$ ,  $f(x) = 1 - |x - 1|$

Vrijenost interpolacijskog polinoma  $P$  u proizvoljnoj točki  $z$  možemo izračunati Hornerovom shemom (vidi Jukić and Scitovski (2004)) korištenjem jednostavnog potprograma ( $a$  – vektor koeficijenata polinoma;  $n$  – broj koeficijenata):

```

Horner[a_,n_,z_] := Module[{s},
    p = a[[n]];
    Do[
        p = p z + a[[n-i]]
        ,{i,n-1}];
    p];

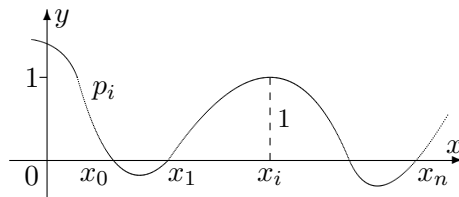
```

### 2.1.1 Lagrangeov oblik interpolacijskog polinoma

Razmotrimo najprije jedan specijalni slučaj koji se lako rješava: *treba pronaći polinom  $p_i$  stupnja  $n$  za koji vrijedi*

$$p_i(x_j) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}, \quad i, j = 0, 1, \dots, n. \quad (2.6)$$

Geometrijski to znači da treba pronaći polinom čiji graf presijeca os  $x$  u točkama  $x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ , a u točki  $x_i$  prima vrijednost 1 (Slika 2.3).



Slika 2.3: Polinom  $p_i$

Kako polinom  $p_i$  iščezava u točkama  $x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ , mora biti

$$p_i(x) = C_i(x - x_0)(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n) \quad (2.7)$$

gdje je  $C_i$  konstanta koju ćemo odrediti iz uvjeta  $p_i(x_i) = 1$

$$C_i = \frac{1}{(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}. \quad (2.8)$$

Uvrštavajući (2.8) u (2.7), dobivamo traženi polinom

$$p_i(x) = \frac{(x - x_0)(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)} = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

Prema tome, polinom  $P_n$  za koji vrijedi (2.4) uz uvjet (2.2) glasi

$$\begin{aligned} P_n(x) &= \sum_{i=0}^n y_i p_i(x) = \sum_{i=0}^n y_i \frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)} \\ &= \sum_{i=0}^n y_i \prod_{j \neq i} \frac{x-x_j}{x_i-x_j}. \end{aligned} \quad (2.9)$$

Očigledno je

$$P_n(x_j) = \sum_{i=0}^n p_i(x_j) y_i = p_j(x_j) y_j = y_j, \quad \forall j = 0, 1, \dots, n.$$

Polinom  $P_n$  obično nazivamo **Lagrangeov oblik interpolacijskog polinoma**

Primijetimo, međutim, da izračunavanje vrijednosti interpolacijskog polinoma u Lagrangeovom obliku u nekoj točki  $x \neq x_i$  zahtijeva veliki broj računskih operacija, pa time i značajno vrijeme rada računala.

**Primjer 2.2.** *Treba odrediti interpolacijski polinom čiji graf prolazi točkama  $T_0(-1, 4)$ ,  $T_1(2, 7)$ ,  $T_2(4, 29)$ .*

Prema (2.9) imamo

$$P_2(x) = \frac{(x-2)(x-4)}{(-1-2)(-1-4)} 4 + \frac{(x+1)(x-4)}{(2+1)(2-4)} 7 + \frac{(x+1)(x-2)}{(4+1)(4-2)} 29$$

ili nakon sređivanja

$$P_2(x) = 2x^2 - x + 1.$$

**Primjer 2.3.** *Treba odrediti interpolacijski polinom za funkciju  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $f(x) = \ln x - 1 + \frac{1}{x}$  i sljedeće čvorove interpolacije*

- a) 1, 2, 4, 8, 10    b) 2, 4, 8, 10  
c) 4, 8, 10        d) 2, 4, 8

*Za svaki od tih slučajeva treba izračunati  $P(5.25)$ , te apsolutnu i relativnu pogrešku ove aproksimacije.*

Prema (2.9) dobivamo (koristite priloženi program)

- a)  $P_4(x) = 0.0004455x^4 - 0.009929x^3 + 0.06339x^2 + 0.0658x - 0.11971$ ,  
b)  $P_3(x) = 0.0007619x^3 - 0.023923x^2 + 0.3438x - 0.4048$ ,  
c)  $P_2(x) = -0.007160x^2 + 0.227947x - 0.161$ ,  
d)  $P_2(x) = -0.013256x^2 + 0.30111x + 0.35605$ .

Rezultate provjerite priloženim Mathematica-programom, str. 23.

	$f(5.25)$	$y^* = P(5.25)$	$\Delta y^*$
a)	-0.848704	-0.874536	0.0258
b)	-0.848704	-0.850896	0.0022
c)	-0.848704	-0.838456	0.0102
d)	-0.848704	-0.859408	0.0107

Tablica 2.1: Interpolacija funkcije  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $f(x) = \ln x - 1 + \frac{1}{x}$

**Primjer 2.4.** Za različite vrijednosti izlaznog kuta  $\alpha$  i konstantnu početnu brzinu  $v_0 = 115 \text{ ms}^{-1}$  primjenom formule (B.4), str. 106 izračunate su daljine gađanja  $X$  balističkog projektila u vakuumu

$\alpha^\circ$	$5^\circ$	$15^\circ$	$25^\circ$	$35^\circ$	$45^\circ$
$X \text{ (m)}$	234	674	1033	1267	1348

Na osnovu ovih podataka treba odrediti polinom koji će za proizvoljnu daljinu  $X$  davati vrijednost izlaznog kuta  $\alpha$ .

Primjenom Lagrangeove interpolacijske formule (2.9) dobivamo polinom  $P_4$  s koeficijentima

$$a_0 = 57.95152, a_1 = -0.41569, a_2 = 0.00101, a_3 \approx a_4 \approx 0.$$

Pomoću ove funkcije za zadane daljine gađanja: 500 m, 750 m i 1000 m izračunati su odgovarajući izlazni kutevi  $\alpha^*$ , te apsolutne i relativne pogreške u odnosu na prave vrijednosti  $\alpha$  dobivene iz formule (B.4), str. 106.

$X \text{ (m)}$	$\alpha^\circ$	$\alpha^*(^\circ)$	$\Delta\alpha^*$
500	10.885	5.676	5.209
750	16.901	18.452	1.551
1000	23.942	24.505	0.563

Tablica 2.2: Domet projektila u ovisnosti o izlaznom kutu

*Primjedba 2.1.* Za zadane točke  $T_i(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , možemo potražiti polinom  $P_m$  stupnja nižeg od  $n$  ( $m < n$ ), za koji više neće biti ispunjen uvjet (2.4), ali čiji će graf prolaziti blizu zadanih točaka. Tako se u prethodnom primjeru pokazalo da su posljednja dva koeficijenta u interpolacijskom polinomu zanemarivo malena, pa ima smisla postavljeno problem pokušati približno riješiti pomoću polinoma 2. stupnja.

U tom slučaju aproksimacija se može odrediti tako da zahtijevamo da suma kvadrata odstupanja zadanih točaka od odgovarajućih točaka na grafu polinoma bude minimalna. Ova metoda naziva se metoda najmanjih kvadrata (vidi t.6, str. 145 i također Truhar (2010)).

### 2.1.2 Newtonov oblik interpolacijskog polinoma

Pretpostavimo da su zadane vrijednosti  $y_i = f(x_i)$ ,  $i = 0, 1, \dots, n$  neprekidne funkcije  $f : [a, b] \rightarrow \mathbb{R}$  u čvorovima  $a \leq x_0 < x_1 < \dots < x_n \leq b$ . Interpolacijski polinom tražit ćemo u obliku

$$P_n(x) = a_0 + a_1(x - x_0) + \dots + a_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}). \quad (2.10)$$

Nakon što odredimo koeficijente  $a_0, a_1, \dots, a_n$  vrijednost polinoma  $P_n$  u nekoj točki  $x \neq x_i$  možemo izračunati prema

$$P_n(x) = (\dots((a_n(x - x_{n-1}) + a_{n-1})(x - x_{n-2}) + \dots + a_1)(x - x_0) + a_0 \quad (2.11)$$

što zahtijeva znatno manje izračunavanja nego kod Lagrangeovog oblika interpolacijskog polinoma. Primijetimo da (2.11) po formi odgovara tzv. **Hornerovoj shemi** za izračunavanje vrijednosti polinoma (vidi Jukić and Scitovski (2004), str. 47).

Prijeđimo na izračunavanje koeficijenata  $a_i$  u prikazu (2.10). Ako je  $n = 1$ , onda je interpolacijski polinom čiji graf prolazi točkama  $T_0 = (x_0, y_0)$  i  $T_1 = (x_1, y_1)$  linearna funkcija (čiji je graf pravac kroz te dvije točke)

$$P_1(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0),$$

što uz oznake  $f[x_0] := y_0$ ,  $f[x_0, x_1] := \frac{y_1 - y_0}{x_1 - x_0}$ , možemo pisati

$$P_1(x) = f[x_0] + f[x_0, x_1](x - x_0).$$

Uzmimo sada  $n = 2$ . Polinom  $P_2$ , čiji graf treba prolaziti točkama  $T_0, T_1, T_2$  ima oblik

$$P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + \alpha(x - x_0)(x - x_1). \quad (2.12)$$

Očigledno je  $P_2(x_0) = y_0$  i  $P_2(x_1) = y_1$ . Parametar  $\alpha$  treba odrediti tako da bude  $P_2(x_2) = y_2$ . Ako u (2.12) uvrstimo  $x = x_2$ , dobivamo

$$\begin{aligned} \alpha &= \frac{y_2 - y_0}{(x_2 - x_1)(x_2 - x_0)} - \frac{y_1 - y_0}{(x_1 - x_0)(x_2 - x_1)} && \text{(što možemo zapisati kao)} \\ &= \frac{y_2 - y_1}{(x_2 - x_1)(x_2 - x_0)} - \frac{y_1 - y_0}{(x_1 - x_0)(x_2 - x_0)} && \text{(odnosno kao)} \\ &= \frac{1}{x_2 - x_0} \left( \frac{y_2 - y_1}{x_2 - x_1} - \frac{y_1 - y_0}{x_1 - x_0} \right). \end{aligned}$$

Ako uvedemo oznake

$$f[x_0, x_1] := \frac{y_1 - y_0}{x_1 - x_0}, \quad f[x_1, x_2] := \frac{y_2 - y_1}{x_2 - x_1}, \quad f[x_0, x_1, x_2] := \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

onda je očigledno  $\alpha = f[x_0, x_1, x_2]$ , pa polinom (2.12) glasi

$$P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1).$$

Općenito, za  $(n + 1)$  točku  $T_i(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  dobivamo

$$P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, x_1, \dots, x_n](x - x_0) \dots (x - x_{n-1}), \quad (2.13)$$

gdje brojeve

$$f[x_0, x_1, \dots, x_i] = \frac{f[x_1, \dots, x_i] - f[x_0, x_1, \dots, x_{i-1}]}{x_i - x_0}, \quad i = 1, \dots, n. \quad (2.14)$$

zovemo **podijeljene razlike**, a polinom (2.13) **Newtonov oblik interpolacijskog polinoma  $n$ -tog reda**.

**Primjer 2.5.** *Treba odrediti Newtonov oblik interpolacijskog polinoma čiji graf prolazi točkama  $T_0 = (-2, -5)$ ,  $T_1 = (2, 3)$ ,  $T_2 = (4, 211)$  i izračunati njegovu vrijednost za  $x = 1$ .*

Koeficijente  $f[x_0, x_1]$  i  $f[x_0, x_1, x_2]$  koji su potrebni u formuli (2.13) izračunat ćemo prema sljedećoj shemi:

$i$	$x_i$	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$
0	-2	-5		
1	2	3	2	
2	4	211	104	17

Prema (2.13) imamo  $P_2(x) = -5 + 2(x + 2) + 17(x + 2)(x - 2)$ . Vrijednost  $P_2(1)$  treba računati prema (2.11):

$$P_2(1) = ((17(1 - 2) + 2)(1 + 2) - 5) = -50.$$

Rezultate provjerite niže navedenim **Mathematica**-programom za izračunavanje podijeljenih razlika prema (2.14)

```
In[1]:= n=3; x={-2,2,4}; y={-5,3,211}; z=1; a = y; t = a;
Do[kk=k-1;
  Do[ii=n-i-kk;
    t[[ii]]=(y[[ii+1]]-y[[ii]])/(x[[ii+kk+1]]-x[[ii]])
```



```

, {i, n-1}];
a[[k+1]] = t[[1]]; y = t
, {k, n-1}]; a

```

```
Out[1] := {-5, 2, 17}
```

i vrijednosti polinoma u točki prema (2.11)

```

In[2] := p = a[[n]];
Do[{ii=n-i; p=p (z-x[[ii]])+a[[ii]], {i, n-1}}]; p

```

```
Out[2] := -50
```

*Zadatak 2.1.* Izvedite Newtonov oblik interpolacijskog polinoma za jednoliko raspoređene čvorove, tj. ako je  $h := x_1 - x_0 = x_2 - x_1 = \dots = x_n - x_{n-1}$ . Pokažite da u tom slučaju vrijedi

$$f[x_0, x_1, \dots, x_k] = \frac{\Delta^k y_0}{h^k k!}, \quad k = 0, 1, \dots, n,$$

gdje je:  $\Delta^1 y_i = y_{i+1} - y_i, \dots, \Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i, i = 0, 1, \dots, n-1$ .

### 2.1.3 Ocjena pogreške

**Teorem 2.1.** *Neka je  $f \in C_{[a,b]}^{n+1}$  funkcija čije vrijednosti su poznate u  $(n+1)$  točaka  $x_i, i = 0, 1, \dots, n$ ,*

$$a = x_0 < x_1 < \dots < x_n = b, \quad y_i = f(x_i), \quad i = 0, 1, \dots, n$$

*i neka je  $P_n$  odgovarajući interpolacijski polinom.*

*Tada za svaki  $\bar{x} \in [a, b]$  postoji  $\xi \in (a, b)$ , tako da je*

$$f(\bar{x}) - P_n(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega(\bar{x}), \quad \omega(\bar{x}) = (\bar{x} - x_0) \cdots (\bar{x} - x_n). \quad (2.15)$$

*Dokaz.* Za  $\bar{x} = x_i$  tvrdnja je očigledna. Za  $\bar{x} \neq x_i, i = 0, 1, \dots, n$  definirajmo pomoćnu funkciju

$$g(x) := f(x) - P_n(x) - k\omega(x), \quad (2.16)$$

gdje ćemo konstantu  $k$  odrediti tako da bude  $g(\bar{x}) = 0$ . Na taj način funkcija  $g$  imat će barem  $(n + 2)$  nultočke:  $\bar{x}, x_0, x_1, \dots, x_n$ . Prema Rolleovom teoremu (vidi Jukić and Scitovski (2004), str. 147) funkcija  $g'$  ima barem  $(n + 1)$  nultočku, funkcija  $g''$  barem  $n$  nultočaka, ... , a funkcija  $g^{(n+1)}$  ima barem jednu nultočku  $\xi \in (a, b)$ .

Primijetite da je  $\frac{d^{n+1}}{dx^{n+1}}P_n(x) = 0$ . Osim toga, kako je  $\omega$  polinom  $(n + 1)$ -og stupnja s vodećim koeficijentom 1, imamo  $\frac{d^{n+1}}{dx^{n+1}}\omega(x) = (n + 1)!$ . Zato je

$$0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - k(n + 1)!$$

odakle slijedi  $k = \frac{f^{(n+1)}(\xi)}{(n+1)!}$ . Na taj način odredili smo konstantu  $k$  tako da je  $g(\bar{x}) = 0$ . Zato iz (2.16) slijedi (2.15).  $\square$

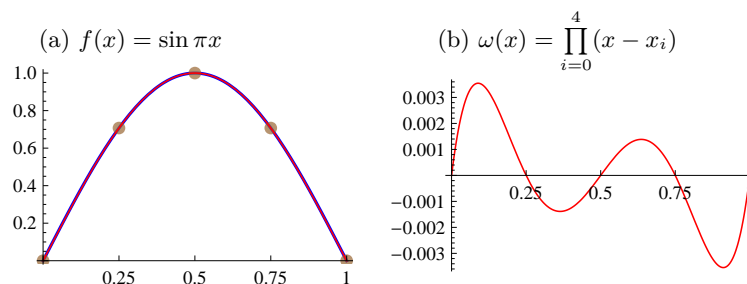
Ako označimo  $M_{n+1} := \max_{x \in [a, b]} |f^{(n+1)}(x)|$ , onda iz (2.15) slijedi

$$|f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega(x)|, \quad (2.17)$$

što znači da pogreška aproksimacije ovisi o ponašanju funkcije  $\omega$ , koju ćemo analizirati u sljedeća dva primjera.

**Primjer 2.6.** Funkciju  $f(x) = \sin(\pi x)$  na intervalu  $[0, 1]$  treba aproksimirati interpolacijskim polinomom koji će prolaziti sljedećim točkama

$x$	0	0.25	0.5	0.75	1
$f(x)$	0	0.707107	1	0.707107	0



Slika 2.4: Interpolacija funkcije  $f: [0, 1] \rightarrow \mathbb{R}$ ,  $f(x) = \sin(\pi x)$

Prema (2.13) dobivamo Newtonov oblik interpolacijskog polinoma

$$P_4(x) = 2.82843x - 3.31371x(x - .25) - 1.83011x(x - .25)(x - .5) + 3.66022x(x - .25)(x - .5)(x - .75).$$

Na Slici 2.4 prikazani su odgovarajući grafovi funkcija  $f$  i  $\omega$ . Iako je  $l_\infty$  pogreska malena

$$\|f - P_4\|_\infty = \max_{x \in [0,1]} |f(x) - P_4(x)| \leq 0.00027,$$

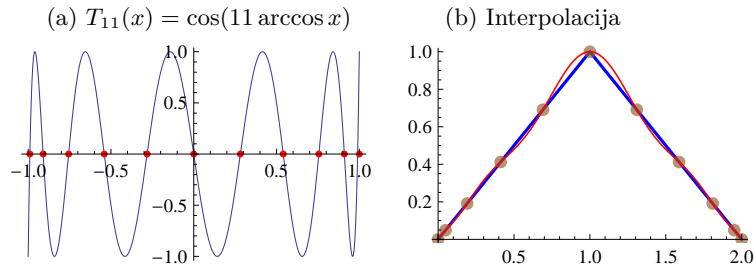
uočljiva su relativno značajnija odstupanja na rubovima promatranog intervala.

### 2.1.4 Nultočke Čebiševljevih polinoma kao čvorovi interpolacije

Pogreška (2.17) kod interpolacijskog polinoma ovisi o ponašanju funkcije  $\omega$ . Za jednoliko raspoređene čvorove i nešto veći  $n$  pokazuje se da funkcija  $\omega$  na rubovima područja interpolacije ima jake oscilacije, što ukazuje na pojavu velikih pogrešaka kod interpolacije. Čvorove interpolacije  $x_0, x_1, \dots, x_n$  normiranog polinoma  $\omega$  stupnja  $n + 1$  trebalo bi birati tako da bude

$$\max_{a \leq x \leq b} |\omega(x)| \longrightarrow \min. \quad (2.18)$$

Prema *Teoremu 5.2*, str. 141, minimum u (2.18) postići će se i biti jednak  $2^{-n}$  ako je  $\omega(x) = 2^{-n}T_{n+1}$ , gdje je  $T_{n+1}$ ,  $(n+1)$ -vi Čebiševljev polinom<sup>1</sup> proširen na interval  $[a, b]$ . Na Slici 2.5a prikazan je graf Čebiševljevog polinoma  $T_{11}$  i njegove nultočke na intervalu  $[-1, 1]$ .



Slika 2.5: Interpolacija funkcije  $f: [0, 2] \rightarrow \mathbb{R}$ ,  $f(x) = 1 - |x - 1|$  u nultočkama Čebiševljevog polinoma  $T_{11}$

Kako je  $(n + 1)$ -i Čebiševljev polinom  $T_{n+1}$  zadan s

$$T_{n+1}(x) = \cos((n + 1) \arccos x), \quad x \in [-1, 1],$$

<sup>1</sup>Čebiševljev polinom stupnja  $n$  programom *Mathematica* dobiva se naredbom `ChebyshevT[n, x]`

njegove nultočke na intervalu  $[-1, 1]$  određene su sa

$$(n+1)\arccos x = \frac{\pi}{2}(2k+1), \quad k = 0, 1, \dots, n \quad (\text{odnosno})$$

$$x_k = \cos \frac{2k+1}{n+1} \frac{\pi}{2}, \quad k = 0, 1, \dots, n.$$

Pomoću linearne funkcije  $u: [-1, 1] \rightarrow [a, b]$ ,  $u(x) = \frac{a+b}{2} + \frac{b-a}{2}x$  ove nultočke preslikat ćemo na interval  $[a, b]$

$$u_k = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{2k+1}{n+1} \frac{\pi}{2}, \quad k = 0, 1, \dots, n.$$

Brojevi  $u_0, u_1, \dots, u_n$  predstavljaju optimalan izbor čvorova interpolacije na intervalu  $[a, b]$ . One su gušće raspoređene pri rubovima intervala  $[a, b]$ . U tom slučaju interpolacijski polinom  $P_n$  imat će najmanju moguću pogrešku prema (2.17)

$$|f(x) - P_n(x)| \leq \frac{M_{n+1}}{2^n(n+1)!}, \quad x \in [a, b].$$

Na *Slici 2.5b* prikazan je interpolacijski polinom za funkciju iz *Primjera 2.1* s čvorovima u nultočkama Čebiševljevog polinoma  $T_{11}$  rastegnuto na interval  $[0, 2]$ .

Najviše zbog navedenih problema kod interpolacije šesdesetih godina prošlog stoljeća pojavile su se tzv. *spline-interpolacije*.

*Primjedba 2.2.* Za primjene je također zanimljiva tzv. Hermiteova interpolacija. U čvorovima interpolacije zadaje se vrijednost funkcije i njene derivacije (“nagib tangente”). O rješavanju ovakvih problema također se može vidjeti u niže navedenoj literaturi.

*Primjedba 2.3.* Postoji vrlo opsežna literatura o interpolaciji funkcija: (Ackleh et al., 2010; Demidovich and Maron, 1981; Deuffhard and Hohmann, 2008; Ortega, 1990; Schwarz and Köckler, 2011; Stoer, 2002; Stoer and Bulirsch, 2002). Odgovarajuća FORTRAN – podrška može se naći kod (?), a posebno u THE NAG LIBRARY.

## 2.2 Spline interpolacija

Ako je broj čvorova interpolacije velik, odgovarajući interpolacijski polinom je visokog stupnja i kao takav neuporabiv u primjenama (vidi *Teorem 2.1* i *Primjer 2.1*). Umjesto interpolacijskog polinoma, možemo pokušati interpolirati funkciju  $\varphi: [a, b] \rightarrow \mathbb{R}$ , takvu da je

$$\varphi|_{[x_{i-1}, x_i]} = \varphi_i,$$

pri čemu je  $\varphi_i : [x_{i-1}, x_i] \rightarrow \mathbb{R}$  neka jednostavnija funkcija (primjerice polinom 1. stupnja) takva da je

$$\varphi_i(x_{i-1}) = y_{i-1}, \quad \varphi_i(x_i) = y_i$$

### 2.2.1 Linearni interpolacijski spline

Pretpostavimo da je  $f : [a, b] \rightarrow \mathbb{R}$  neprekidna funkcija i da poznamo njene vrijednosti  $y_0, y_1, \dots, y_n$  u  $(n + 1)$  točaka-čvorova

$$a = x_0 < x_1 < \dots < x_n = b.$$

Funkciju  $f$  interpolirat ćemo neprekidnom po dijelovima linearnom funkcijom  $\varphi : [a, b] \rightarrow \mathbb{R}$ . Lako se vidi da takva funkcija postoji i da je jedinstvena. Za  $i = 1, \dots, n$  definirajmo

$$\begin{aligned} \varphi|_{[x_{i-1}, x_i]} &= \varphi_i, \\ \varphi_i(x) &= y_{i-1} + \frac{y_i - y_{i-1}}{x_i - x_{i-1}}(x - x_{i-1}), \quad x \in [x_{i-1}, x_i]. \end{aligned} \quad (2.19)$$

Kako je

$$\varphi_i(x_i) = y_i = \varphi_{i+1}(x_i), \quad i = 1, \dots, n - 1,$$

funkcija  $\varphi$  je neprekidna na čitavom intervalu  $[a, b]$ , a linearna na svakom podintervalu  $[x_{i-1}, x_i]$ . Kako je osim toga

$$\varphi(x_0) = \varphi_1(x_0) = y_0, \quad \varphi(x_i) = \varphi_i(x_i) = y_i, \quad i = 1, \dots, n,$$

ona interpolira funkciju  $f$  u čvorovima  $x_0, x_1, \dots, x_n$ . Nazivamo je linearni interpolacijski spline<sup>2</sup> (vidi *Sliku 2.6*). Linearni interpolacijski spline nije derivabilna funkcija u čvorovima  $x_0, x_1, \dots, x_n$  i stoga nije pogodan za daljnju uporabu.

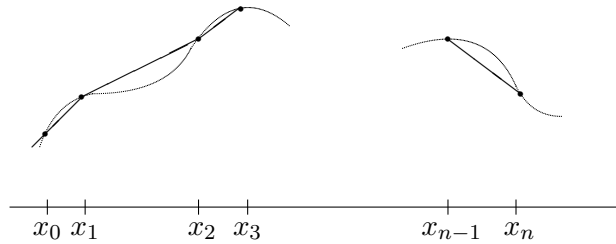
Odredimo pogrešku linearnog interpolacijskog splinea. Uz pretpostavku da je funkcija  $f \in C^2[a, b]$ , prema *Teoremu 2.1*, str. 30, za svaki  $x \in [x_{i-1}, x_i]$  postoji  $\xi_i \in (x_{i-1}, x_i)$ , takav da vrijedi

$$f(x) - \varphi_i(x) = \frac{1}{2}f''(\xi_i)(x - x_{i-1})(x - x_i), \quad i = 0, 1, \dots, n. \quad (2.20)$$

Označimo  $h_i := x_i - x_{i-1}$ . Budući da kvadratna funkcija  $x \mapsto (x - x_{i-1})(x - x_i)$  postiže minimum u točki  $x_0 = \frac{x_{i-1} + x_i}{2}$ , onda za svaki  $x \in [x_{i-1}, x_i]$  vrijedi

$$|(x - x_{i-1})(x - x_i)| \leq \left| \left( \frac{x_{i-1} + x_i}{2} - x_{i-1} \right) \left( \frac{x_{i-1} + x_i}{2} - x_i \right) \right| = \frac{h_i^2}{4},$$

<sup>2</sup>Naziv dolazi od engleske riječi *spline*, koja označava dug, savitljiv, uzak drveni klin



Slika 2.6: Linearni interpolacijski spline

pa je

$$|f(x) - \varphi_i(x)| \leq \frac{1}{2} |f''(\xi_i)| \frac{h_i^2}{4}.$$

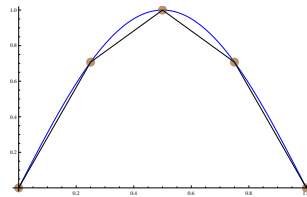
Ako označimo

$$h_{max} := \max_{i=0, \dots, n} h_i \quad M_2 := \max_{x \in [a, b]} |f''(x)|,$$

tada prema (2.20) vrijedi ocjena pogreške za linearni interpolacijski spline

$$\|f - \varphi\|_{\infty} = \max_{x \in [a, b]} |f(x) - \varphi(x)| \leq \frac{M_2}{8} h_{max}^2 \quad (2.21)$$

**Primjer 2.7.** Treba odrediti linearni interpolacijski spline za funkciju  $i$  čvorove interpolacije iz Primjera 2.6, str.31. Također, treba odrediti  $i$  pogrešku interpolacije.



Slika 2.7: Linearni interpolacijski spline iz Primjera 2.7

Dobivamo

$$\varphi(x) = \begin{cases} 2\sqrt{2}x, & 0 \leq x < \frac{1}{4} \\ \frac{\sqrt{2}}{2} + (4 - 2\sqrt{2})(x - \frac{1}{4}), & \frac{1}{4} \leq x < \frac{1}{2} \\ 1 + (2\sqrt{2} - 4)(x - \frac{1}{2}), & \frac{1}{2} \leq x < \frac{3}{4} \\ \frac{\sqrt{2}}{2} - 2\sqrt{2}(x - \frac{3}{4}), & x > \frac{3}{4} \end{cases}$$

Linearni interpolacijski spline možemo definirati i na drugi način - analogno kako smo u t.2.1.1, str. 25 uveli Lagrangeov oblik interpolacijskog polinoma. U tu svrhu definirajmo tzv. **hat-funkcije**  $p_i$ ,  $i = 1, \dots, n-1$ .

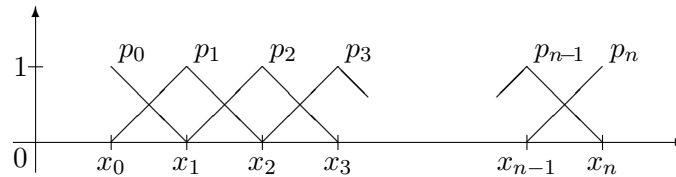
$$p_i(x) = \begin{cases} (x - x_{i-1})/(x_i - x_{i-1}), & x \in [x_{i-1}, x_i] \\ (x_{i+1} - x)/(x_{i+1} - x_i), & x \in [x_i, x_{i+1}] \\ 0, & \text{inače} \end{cases}$$

$$p_0(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_0}, & x \in [x_0, x_1] \\ 0, & \text{inače} \end{cases} \quad p_n(x) = \begin{cases} \frac{x - x_{n-1}}{x_n - x_{n-1}}, & x \in [x_{n-1}, x_n] \\ 0, & \text{inače} \end{cases}$$

Hat-funkcije  $p_i$  imaju svojstvo

$$p_i(x_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

(usporedi s (2.6) str. 25), a njihovi grafovi prikazani su na *Slici 2.8*.



Slika 2.8: Hat-funkcije

Sada linearni interpolacijski spline možemo definirati formulom

$$\varphi(x) = \sum_{i=0}^n y_i p_i(x) \quad (2.22)$$

(usporedi s (2.9), str. 26).

*Primjedba 2.4.* Ako s  $\mathcal{L}$  označimo skup svih linearnih interpolacijskih spline-ova određenih čvorovima  $a = x_0 < x_1 < \dots < x_n = b$ , onda je lako vidjeti da je  $\mathcal{L}$  vektorski potprostor na prostoru  $C[a, b]$  svih neprekidnih funkcija na intervalu  $[a, b]$ . Definirajmo linearni operator  $L: C[a, b] \rightarrow \mathcal{L}$ ,  $L(f) = \varphi$ , koji neprekidnoj funkciji  $f$  pridružuje linearni interpolacijski spline  $\varphi$  u čvorovima  $x_0, \dots, x_n$ . Specijalno, za neki linearni interpolacijski spline  $\varphi \in \mathcal{L} \subset C[a, b]$  vrijedi  $L(\varphi) = \varphi$ , pa onda i općenito za svaki  $f \in C[a, b]$  vrijedi

$$L(L(f)) = L(f), \quad \text{tj.} \quad L^2 = L.$$

Dakle, linearni operator  $L$  je projektor.

*Zadatak 2.2.* Pokažite da jednu bazu u prostoru  $\mathcal{L}$  čine hat-funkcije  $\{p_0, p_1, \dots, p_n\}$ , te da za svaki  $f \in C[a, b]$  vrijedi

$$L(f) = \sum_{i=0}^n y_i p_i,$$

gdje je  $y_i = f(x_i)$ ,  $i = 0, 1, \dots, n$ .

*Zadatak 2.3.* Izračunajte integral funkcije (2.22) u granicama od  $x_0$  do  $x_n$ . Pokažite da za ekvidistantno raspoređene čvorove dobivamo poznato generalizirano trapezno pravilo (vidi t.??, str. ??).



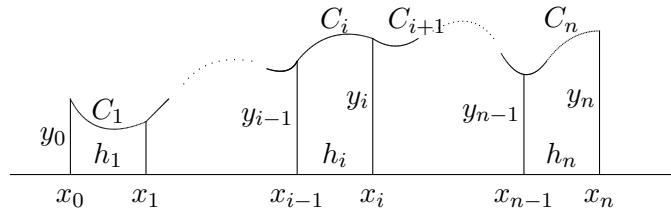
### 2.2.2 Kubični interpolacijski spline

Neprekidnu funkciju  $f: [a, b] \rightarrow \mathbb{R}$  čije vrijednosti  $y_0, y_1, \dots, y_n$  poznajemo u  $(n + 1)$  čvorova  $a = x_0 < x_1 < \dots < x_n = b$ , interpolirat ćemo funkcijom  $C: [a, b] \rightarrow \mathbb{R}$ ,

$$C(x) = C_i(x), \quad x \in [x_{i-1}, x_i], \quad i = 1, \dots, n, \quad (2.23)$$

gdje su  $C_i$ ,  $i = 1, \dots, n$ , kubični polinomi koji trebaju zadovoljavati sljedeće uvjete (usporedi *Sliku 2.7*).

- (i)  $C_i(x_{i-1}) = y_{i-1}$ ,  $i = 1, \dots, n$ ;
- (ii)  $C_i(x_i) = y_i$ ,  $i = 1, \dots, n$ ;
- (iii)  $C'_i(x_i) = C'_{i+1}(x_i)$ ,  $i = 1, \dots, n - 1$ ;
- (iv)  $C''_i(x_i) = C''_{i+1}(x_i)$ ,  $i = 1, \dots, n - 1$ .



Slika 2.9: Kubični interpolacijski spline

Uvjeti (i) – (ii) osiguravaju interpolacijske uvjete u čvorovima  $x_0, x_1, \dots, x_n$  i neprekidnost funkcije  $C$ , uvjet (iii) osigurava derivabilnost funkcije  $C$ , a uvjet (iv) osigurava da će funkcija  $C$  biti iz klase  $C^2_{[a,b]}$ . Budući da  $n$  kubičnih polinoma  $C_1, \dots, C_n$  ima ukupno  $4n$  parametara, a uvjetima (i) – (iv) određeno je  $4n - 2$  uvjeta, za jednoznačno određivanje funkcije  $C$  potrebno je postaviti još dva dodatna uvjeta. Najčešće zahtijevamo još da bude

$$C''_1(x_0) = C''_n(x_n) = 0, \quad (2.24)$$

iako je moguće tražiti i neke druge zahtjeve. Funkciju  $C$  koja zadovoljava uvjete (i)-(iv), te dodatne uvjete (2.24) nazivamo **prirodni kubični interpolacijski spline** (Plato, 2003; Stewart, 1996; Trefethen and Bau, 1997). Funkcija  $C$  je klase  $C^2[a, b]$ , koja na intervalu  $[a, b]$  interpolira funkciju  $f$  poznatu u čvorovima  $x_0, x_1, \dots, x_n$ .

*Primjedba 2.5.* Problem određivanja koeficijenata kubičnih polinoma  $C_i$  na taj način sveo se na rješavanje sustava linearnih jednačbi. Primijetimo da već za 50 čvorova treba riješiti sustav od 200 jednačbi s 200 nepoznanica. Ako je pri tome još matrica sustava loše uvjetovana (vidi t.3.2, str. 52), mogu nastupiti ozbiljni numerički problemi.

**Teorem 2.2.** *Zadani su podaci  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  koji zadovoljavaju uvjet  $x_0 < x_1 < \dots < x_n$ . Tada postoji jedinstveni prirodni kubični interpolacijski spline  $C$ , pri čemu su polinomi  $(C_i, i = 1, \dots, n)$  zadani s*

$$C_i(x) = \left( y_{i-1} - s_{i-1} \frac{h_i^2}{6} \right) + b_i(x - x_{i-1}) + \frac{s_{i-1}}{6h_i}(x_i - x)^3 + \frac{s_i}{6h_i}(x - x_{i-1})^3 \quad (2.25)$$

gdje je

$$b_i = d_i - (s_i - s_{i-1}) \frac{h_i}{6},$$

$$d_i = \frac{y_i - y_{i-1}}{h_i}, \quad h_i = x_i - x_{i-1},$$

a brojevi  $(s_i, i = 0, \dots, n)$  zadovoljavaju sustav jednačbi

$$s_{i-1}h_i + 2s_i(h_i + h_{i+1}) + s_{i+1}h_{i+1} = 6(d_{i+1} - d_i), \quad i = 1, \dots, n-1$$

$$s_0 = s_n = 0. \quad (2.26)$$

**Dokaz.** Za  $i = 1, \dots, n$  označimo

$$s_{i-1} := C_i''(x_{i-1}), \quad s_i := C_i''(x_i).$$

Primijetimo da je zbog (2.24)  $s_0 = s_n = 0$ . Kako je  $x \rightarrow C_i''(x)$  linearna funkcija čiji graf treba prolaziti točkama  $(x_{i-1}, s_{i-1})$ ,  $(x_i, s_i)$ , mora biti

$$C_i''(x) = \frac{s_{i-1}}{h_i}(x_i - x) + \frac{s_i}{h_i}(x - x_{i-1}) \quad (2.27)$$

Primitivna funkcija funkcije (2.27) je

$$C_i'(x) = b_i - \frac{s_{i-1}}{2h_i}(x_i - x)^2 + \frac{s_i}{2h_i}(x - x_{i-1})^2, \quad (2.28)$$

gdje je  $b_i$  integracijska konstanta. Nadalje, primitivna funkcija funkcije (2.28) je

$$C_i(x) = a_i + b_i x + \frac{s_{i-1}}{6h_i}(x_i - x)^3 + \frac{s_i}{6h_i}(x - x_{i-1})^3, \quad (2.29)$$

gdje je  $a_i$  integracijska konstanta. Konstante  $a_i, b_i$  odredit ćemo iz uvjeta (i)-(ii):

$$C_i(x_{i-1}) = y_{i-1}, \quad C_i(x_i) = y_i, \quad i = 1, \dots, n$$

tj. rješavanjem sustava s dvije nepoznanice  $a_i, b_i$

$$\begin{aligned} a_i + b_i x_{i-1} + s_{i-1} \frac{h_i^2}{6} &= y_{i-1}, \\ a_i + b_i x_i + s_i \frac{h_i^2}{6} &= y_i. \end{aligned}$$

Dobivamo

$$\begin{aligned} a_i &= -b_i x_{i-1} + y_{i-1} - s_{i-1} \frac{h_i^2}{6}, \\ b_i &= d_i - (s_i - s_{i-1}) \frac{h_i}{6}, \quad d_i = \frac{y_i - y_{i-1}}{h_i}. \end{aligned} \quad (2.30)$$

Zato iz (2.29) neposredno slijedi (2.25).

Još treba odrediti brojeve  $s_1, \dots, s_{n-1}$ . Njih ćemo odrediti iz uvjeta (iii). Iz (2.28) dobivamo

$$C'_i(x_i) = b_i + \frac{s_i}{2} h_i, \quad C'_{i+1}(x_i) = b_{i+1} - \frac{s_i}{2} h_{i+1},$$

pa uvjet (iii) postaje

$$d_i - (s_i - s_{i-1}) \frac{h_i}{6} + \frac{s_i}{2} h_i = d_{i+1} - (s_{i+1} - s_i) \frac{h_{i+1}}{6} - s_i \frac{h_{i+1}}{2},$$

iz čega neposredno slijedi (2.26).

Kako je  $s_0 = s_n = 0$ , (2.26) je sustav od  $(n-1)$  jednadžbe s  $(n-1)$  nepoznanica:  $s_1, \dots, s_{n-1}$ , koji možemo zapisati u matricnom obliku

$$\mathbf{H}\mathbf{s} = \mathbf{r}, \quad (2.31)$$

$$\mathbf{H} = \begin{bmatrix} 2(h_1 + h_2) & h_2 & \cdots & 0 & 0 \\ h_2 & 2(h_2 + h_3) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & 0 & \cdots & h_{n-1} & 2(h_{n-1} + h_n) \end{bmatrix},$$

$$\mathbf{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_{n-1} \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 6(d_2 - d_1) \\ \vdots \\ 6(d_n - d_{n-1}) \end{bmatrix}.$$

Primijetimo da je matrica sustava (2.31) trodijagonalna, dijagonalno dominantna i simetrična matrica s pozitivnim elementima, pa kao takva pozitivno definitna (Truhar, 2010). Zato postoji jedinstveno rješenje ovog sustava, a onda i jedinstveni prirodni kubični interpolacijski spline.  $\square$

*Primjedba 2.6.* Ako je  $f \in C^4[a, b]$  i  $h_{max} = \max_i h_i$ , onda vrijedi (Stewart, 1996)

$$\begin{aligned}\|f - C\|_\infty &\leq \frac{5}{384}h_{max}^4 \\ \|f' - C'\|_\infty &\leq \frac{1}{24}h_{max}^3 \\ \|f'' - C''\|_\infty &\leq \frac{3}{8}h_{max}^2\end{aligned}$$

---

**Algorithm 1 (SPLINE3)**


---

1. Učitati podatke  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ ;
2. Izračunati

$$h_i = x_i - x_{i-1}, \quad d_i = \frac{y_i - y_{i-1}}{h_i}, \quad i = 1, \dots, n,$$

i riješiti sustav (2.31);

3. Izračunati  $b_i = d_i - (s_i - s_{i-1})\frac{h_i}{6}$ ,  $i = 1, \dots, n$  i definirati polinome (2.25).
- 

Prilikom rješavanja sustava (2.31) treba iskoristiti specijalnu strukturu matrice sustava (vidi *Zadatak 3.24*, str. 83). Specijalno, ako su razmaci između čvorova ekvidistantni i iznose  $h$ , onda imamo

$$\mathbf{H} = \begin{bmatrix} 4 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 4 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} \frac{6}{h}(d_2 - d_1) \\ \vdots \\ \frac{6}{h}(d_n - d_{n-1}) \end{bmatrix}.$$

*Zadatak 2.4.* Izradi program koji će prema algoritmu SPLINE3 izračunavati prirodni kubični interpolacijski spline.

**Primjer 2.8.** Tablično su zadane vrijednosti funkcije

$x_i$	0	0.25	0.5	0.75	1
$y_i$	1	2	1	0	1

U ovom slučaju je  $n = 4$ , a svi  $h_i$  jednaki su 0.25. Sustav (2.31) glasi

$$\begin{bmatrix} 1 & 0.25 & 0 \\ 0.25 & 1 & 0.25 \\ 0 & 0.25 & 1 \end{bmatrix} \mathbf{s} = \begin{bmatrix} -48 \\ 0 \\ 48 \end{bmatrix}.$$

a njegovo rješenje je:  $s_1 = -48$ ,  $s_2 = 0$ ,  $s_3 = 48$ . Osim toga, iz uvjeta (2.24) imamo:  $s_0 = s_4 = 0$ . Sada iz (2.25) možemo izračunati četiri kubična polinoma  $C_i$ ,  $i = 1, 2, 3, 4$ , koji grade traženi prirodni kubični interpolacijski spline

$$C(x) = \begin{cases} C_1(x) = 1 + 6x - 32x^3, & 0 \leq x \leq 0.25 \\ C_2(x) = 18x - 48x^2 + 32x^3, & 0.25 \leq x \leq 0.5 \\ C_3(x) = 18x - 48x^2 + 32x^3, & 0.5 \leq x \leq 0.75 \\ C_4(x) = 27 - 90x + 96x^2 - 32x^3, & 0.75 \leq x \leq 1 \end{cases}$$

Ako želimo primjerice izračunati vrijednost spline-a  $C$  u točki 0.7, onda prvo moramo ustanoviti da je  $0.7 \in [0.5, 0.75]$ . Tada je  $C(0.7) = C_3(0.7) = 0.056$ .

*Primjedba 2.7.* Prirodni kubični interpolacijski spline  $C$ , koji interpolira funkciju  $f$  u čvorovima  $a = x_0 < x_1 < \dots < x_n = b$  ima svojstvo da za svaku funkciju  $g \in C^2[a, b]$ , koja također interpolira funkciju  $f$  u navedenim čvorovima, vrijedi

$$\int_a^b (g''(x))^2 dx \geq \int_a^b (C''(x))^2 dx, \quad \text{odnosno} \quad (2.32)$$

$$\|g''\|_2 \geq \|C''\|_2, \quad (2.33)$$

tj. između svih interpolirajućih funkcija prirodni kubični spline  $C$  ima najmanju  $L_2$ -normu druge derivacije.

Naime, vrijedi

$$\begin{aligned} 0 \leq \int_a^b (g''(x) - C''(x))^2 dx &= \int_a^b (g''(x))^2 dx - 2 \int_a^b g''(x)C''(x) dx + \int_a^b (C''(x))^2 dx \\ &= \int_a^b (g''(x))^2 dx - \int_a^b (C''(x))^2 dx + 2 \int_a^b C''(x) (C''(x) - g''(x)) dx. \end{aligned}$$

Pri tome, zbog  $C''(a) = C''(b) = 0$ , parcijalnom integracijom dobivamo

$$\begin{aligned} I_1 &= \int_a^b C''(x) (C''(x) - g''(x)) dx \\ &= [C''(x) (C'(x) - g'(x))]_a^b - \int_a^b C'''(x) (C'(x) - g'(x)) dx \quad [C''(a) = C''(b) = 0] \\ &= - \int_a^b C'''(x) (C'(x) - g'(x)) dx. \end{aligned}$$

Kako je  $C'''$  po dijelovima konstantna funkcija, a  $C$  i  $g$  interpolirajuće funkcije, vrijedi

$$\int_{x_i}^{x_{i+1}} C'''(C'(x) - g'(x))dx = \text{const} [C(x) - g(x)]_{x_i}^{x_{i+1}} = 0,$$

odakle slijedi  $I_1 = 0$  i tražena nejednakost (2.32).

Svojsstvo (2.32) nazivamo minimalno svojsstvo prirodnog kubičnog spline-a. To svojsstvo vrijedi i za kubični spline s drukčijim rubnim uvjetima.

*Primjedba 2.8.* Spline-interpolacije pojavile su se u matematičkoj literaturi u posljednjih tridesetak godina, a zbog mogućnosti učinkovite primjene kod rješavanja običnih i parcijalnih diferencijalnih jednadžbi i u mnogim drugim situacijama, i sada se intenzivno proučavaju. Kvalitetna i pregledna knjiga s odgovarajućom programskom podrškom (FORTRAN) iz ovog područja je de Boor (2001), ali se osnovni pojmovi i neke primjene mogu vidjeti i kod (Deuffhard and Hohmann, 2008; Ortega, 1990; Schwarz and Köckler, 2011; Stewart, 1998). Programski sustav *Mathematica* omogućava izravno korištenje spline-a bilo kojeg reda razumne veličine. Odgovarajuća FORTRAN programska podrška može se naći u (?), kao i u THE NAG LIBRARY. O interpolaciji funkcija jedne i više varijabli može se vidjeti primjerice kod (Lancaster and Šalkauskas, 1986).

## 2.3 Zadaci

*Zadatak 2.5.* Odredite determinantu sustava (2.5) (Vandermondova determinanta).

*Zadatak 2.6.* Odredite interpolacijski polinom za sljedeće tablično zadane funkcije

$$\begin{array}{ll} \text{a)} \frac{x}{y} \begin{array}{ccc|c} 2 & 5 & 10 & \\ \hline 51 & 150 & 435 & \end{array} & \text{b)} \frac{x}{y} \begin{array}{cccc|c} -1 & 0 & 1 & 3 & \\ \hline -8 & 2 & 6 & 44 & \end{array} \\ \text{c)} \frac{x}{y} \begin{array}{cccc|c} -2 & 1 & 2 & 5 & \\ \hline 34 & -2 & -10 & 62 & \end{array} & \text{d)} \frac{x}{y} \begin{array}{cccc|cc} 0 & 1 & 2 & 5 & -2 & \\ \hline -8 & -1 & 38 & 1127 & -70 & \end{array} \end{array}$$

Rješenje:

$$\begin{array}{ll} \text{a)} P_2(x) = 3x^2 + 12x + 15 & \text{b)} P_3(x) = 2x^3 - 3x^2 + 5x + 2 \\ \text{c)} P_3(x) = x^3 - 15x + 12 & \text{d)} P_4(x) = x^4 + 5x^3 - 6x^2 + 7x - 8 \end{array}$$

*Zadatak 2.7.* Za tablično zadane funkcije odredite Newtonov oblik interpolacijskog polinoma te izračunajte njegovu vrijednost za  $\xi = 10$ .

$$\begin{array}{l} \text{a) } \frac{x}{y} \left| \begin{array}{ccc} 2 & 4 & 5 \\ 91 & 171 & 220 \end{array} \right. \quad \text{b) } \frac{x}{y} \left| \begin{array}{ccc} -3 & 0 & 2 & 4 \\ -28 & 8 & 52 & 224 \end{array} \right. \\ \text{c) } \frac{x}{y} \left| \begin{array}{ccc} -2 & -1 & 0 & 1 \\ -4 & -6 & -2 & 2 \end{array} \right. \quad \text{d) } \frac{x}{y} \left| \begin{array}{ccc} -2 & 0 & 2 & 3 & 4 \\ 88 & 6 & 20 & 108 & 370 \end{array} \right. \end{array}$$

Rješenje:

$$\begin{array}{l} \text{a) } P_2(x) = 91 + 40(x-2) + 3(x-2)(x-4), \quad p_2(10) = 555 \\ \text{b) } P_3(x) = -28 + 12(x+3) + 2(x+3)x + 2(x+3)x(x-2), \quad P_3(10) = 2468 \\ \text{c) } P_3(x) = -4 - 2(x+2) + 3(x+2)(x+1) - 1(x+2)(x+1)x, \quad P_3(10) = -952 \\ \text{d) } P_4(x) = 88 - 41(x+2) + 12(x+2)x + 3(x+2)x(x-2) + \\ \quad 2(x+2)x(x-2)(x-3), \quad P_4(10) = 17356 \end{array}$$

*Zadatak 2.8.* Zadane su vrijednosti funkcije  $f(x) = 3x^2 + \pi^{-4} \ln(\pi - x)^2 + 1$  u sljedećim točkama

$$\frac{x}{f(x)} \left| \begin{array}{ccc} 3.13 & 3.14 & 3.15 & 3.16 \\ 30.29918 & 30.44653 & 30.66939 & 30.87478 \end{array} \right.$$

Odredite interpolacijski polinom, čiji će graf prolaziti ovim točkama, te izračunajte njegovu vrijednost u točki  $\xi = 3.1416$ .

Rješenje:  $P_3(3.1416) = 30.32619$ ,

$$P_3(x) = 483453.48 - 460723.76x + 146356.15x^2 - 15496.67x^3$$

*Zadatak 2.9.* Zadane su vrijednosti  $\Gamma$ -funkcije u sljedećim točkama

$$\frac{x}{\Gamma(x)} \left| \begin{array}{ccc} 0.2 & 0.4 & 0.6 & 0.8 \\ 4.5910 & 2.2182 & 1.4892 & 1.1642 \end{array} \right.$$

Odredite interpolacijski polinom, čiji graf će prolaziti ovim točkama, te izračunajte njegovu vrijednost u točkama: 0.3, 0.5, 0.7. Izračunajte apsolutnu i relativnu pogrešku ovih aproksimacija.

Rješenje: *Koeficijenti interpolacijskog polinoma su*

$$a_0 = 9.8474, \quad a_1 = -35.5573, \quad a_2 = 51.5425, \quad a_3 = -25.8292$$

$x$	$\Gamma(x)$	$P(x)$	$\Delta P$
0.3	2.9917	3.122	0.1303
0.5	1.7724	1.726	0.0464
0.7	1.2980	1.354	0.0560

*Zadatak 2.10.* Funkcija  $f$  zadana je u nekoliko točaka:  $x_0, x_1, \dots$ . Odredite interpolacijski polinom  $P$ , te izračunajte  $P(c_1)$  i  $P(c_2)$  ako je

- a)  $f(x) = x + 2/x$ ,  
 (1)  $x_0 = 1, x_1 = 2, x_2 = 2.5, c_1 = 1.5, c_2 = 1.2$ ,  
 (2)  $x_0 = 0.5, x_1 = 1, x_2 = 2, x_3 = 2.5, c_1 = 1.5, c_2 = 1.2$ ,
- b)  $f(x) = 8x2^{-x}$ ,  
 (1)  $x_0 = 0, x_1 = 1, x_2 = 2, c_1 = 1.5, c_2 = 1.3$ ,  
 (2)  $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3, c_1 = 1.5, c_2 = 1.3$ ,
- c)  $f(x) = 2 \sin(\pi x/6)$ ,  
 (1)  $x_0 = 0, x_1 = 1, x_2 = 3, c_1 = 2, c_2 = 2.4$ ,  
 (2)  $x_0 = 0, x_1 = 1, x_2 = 3, x_3 = 5, c_1 = 2, c_2 = 2.4$ ,  
 (3)  $x_0 = 0, x_1 = 1, x_2 = 3, c_1 = 4, c_2 = 3.5$ ,  
 (4)  $x_0 = 0, x_1 = 1, x_2 = 3, x_3 = 5, c_1 = 4, c_2 = 3.5$ .

Rješenje:

- a) (1)  $P(x) = 0.4x^2 - 1.2x + 3.8, P(1.5) = 2.9, P(1.2) = 2.936$ ;  
 (2)  $P(x) = -0.8x^3 + 4.8x^2 - 8.8x + 7.8, P(1.5) = 2.7, P(1.2) = 2.7696$
- b) (1)  $P(x) = -2x^2 + 6x, P(1.5) = 4.5, P(1.3) = 4.42$ ;  
 (2)  $P(x) = 0.5x^3 + 3.5x^2 + 7x, P(1.5) = 4.125, P(1.3) = 4.2835$
- c) (1)  $P(x) = -\frac{1}{6}x^2 + \frac{7}{6}x, P(2) = \frac{7}{6}, P(2.4) = 1.84$ ;  
 (2)  $P(x) = -0.01(6)x^3 - .1x^2 + 1.11(6)x, P(2) = 1.7, P(2.4) = 1.8736$   
 (3)  $P(x) = -.1(6)x^2 + 1.1(6)x, P(4) = 2, P(3.5) = 2.04167$   
 (4)  $P(x) = -.01(6)x^3 - .1x^2 + 1.11(6)x, P(4) = 1.8, P(3.5) = 1.96875$

*Zadatak 2.11.* Za različite vrijednosti izlaznog kuta  $\alpha$  i konstantnu početnu brzinu  $v_0 = 300 \text{ ms}^{-1}$  primjenom formule (B.4), str. 106 izračunati su dometi  $X$  balističkog projektila u vakuumu

$\alpha^\circ$	$5^\circ$	$10^\circ$	$25^\circ$	$45^\circ$
$X \text{ (m)}$	1593	3138	7028	9174

Na osnovi ovih podataka odredite funkciju-polinom koja će za proizvoljnu daljinu  $X$  davati vrijednost izlaznog kuta  $\alpha$ . Pomoću ove funkcije za zadane daljine: 2000 m, 3000 m, 5000 m, 7000 m, 9000 m izračunajte izlazne kuteve  $\alpha^*$ , apsolutne i relativne pogreške.



Rješenje:  $\alpha(X) = -3.2516 + 0.00669X - 0.0000011X^2 + 1.0410^{-10}X^3$ .

$X$ (m)	$\alpha$	$\alpha^*$	$\Delta\alpha^*$
2000	6.296	6.507	0.211
3000	9.543	9.613	0.070
5000	16.512	15.407	1.105
7000	24.865	24.819	0.046
9000	39.407	42.858	3.451

*Zadatak 2.12.* Olovnu kuglicu puštamo da u zraku pada s različitih visina  $h$  i mjerimo vrijeme pada  $t$

$h$ (m)	0.5	0.8	1	1.2
$t$ (s)	0.32	0.40	0.45	0.49

Odredite interpolacijski polinom

a) za funkciju  $h \mapsto t(h)$ ,      b) za funkciju  $t \mapsto h(t)$ .

Rješenje:

$$\begin{aligned} a) t(h) &= 0.2257 + 0.0874h + 0.2679h^2 - 0.131h^3 \\ b) h(t) &= -3.56697 + 26.7947t - 61.3122t^2 + 54.0473t^3 \end{aligned}$$

*Zadatak 2.13.* Olovnu kuglicu puštamo da u ulju pada s različitih visina  $h$  i mjerimo vrijeme pada  $t$

$h$ (cm)	4	8	10	16	20
$t$ (s)	2.5	5	6.25	10	12.5

Odredite interpolacijski polinom za funkciju  $h \mapsto t(h)$ .

Rješenje: a)  $t(h) = 0.625h$ .

*Zadatak 2.14.* Principom kontradikcije u dokazivanju pokažite da za dane podatke  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , koji zadovoljavaju uvjet  $x_0 < x_1 < \dots < x_n$ , postoji jedinstveni interpolacijski polinom stupnja  $\leq n$ .

*Zadatak 2.15.* Polinom  $P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$  može se napisati u obliku (Hornerova shema)

$$P_n(x) = (((((a_n x + a_{n-1})x + a_{n-2})x + a_{n-3}) \dots)x + a_1)x + a_0. \quad (*)$$

Izradite program kojim ćete učitati stupanj polinoma, sve njegove koeficijente, te za zadanu vrijednost  $x = a$  primjenom formule (\*) izračunati vrijednost polinoma u točki  $a$ .

*Zadatak 2.16.* Izradite program kojim ćete učitati  $n$  točaka  $T_i(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  i definirati linearni interpolacijski spline. Za zadani broj  $x_0 \in [x_1, x_n]$  izračunajte vrijednost linearnog interpolacijskog spline-a.

*Zadatak 2.17.* Za zadane podatke  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  odredite kvadratni interpolacijski spline i izradite odgovarajući program.

*Zadatak 2.18.* Zadana je periodična funkcija  $f$  perioda  $2\pi$  i točke  $(x_i, y_i)$ ,  $i = 0, \dots, m$ ,  $x_i = -\pi + 2\pi i/m$ ,  $y_i = f(x_i)$ . Tada postoji trigonometrijski polinom

$$T_n(x) = \frac{a_0}{2} + \sum_{j=1}^n (a_j \cos(jx) + b_j \sin(jx)), \quad (2n < m),$$

$$a_j = \frac{2}{m} \sum_{k=1}^m f(x_k) \cos(jx_k), \quad b_j = \frac{2}{m} \sum_{k=1}^m f(x_k) \sin(jx_k), \quad j = 0, \dots, n,$$

takav da je

$$\sum_{k=1}^m (f(x_k) - T_n(x_k))^2 \rightarrow \min.$$

Izradite program kojim ćete na osnovi zadanih podataka definirati trigonometrijski polinom  $T_n$ .

*Zadatak 2.19.* Pokažite da je dijagonalno dominantna matrica regularna.

*Zadatak 2.20.* Na osnovi podataka

$$\text{a) } \begin{array}{c|cccc} x_i & 0 & 2 & 4 & 5 \\ \hline y_i & 10 & 15 & 12 & 14 \end{array}, \quad \xi = 1 \quad \text{b) } \begin{array}{c|cccc} x_i & 1 & 3 & 5 & 7 \\ \hline y_i & 8 & 12 & 10 & 14 \end{array}, \quad \xi = 4.5$$

odredite prirodni kubični interpolacijski spline i izračunajte  $C(\xi)$ .

Rješenje:

$$\begin{aligned} \text{a) } C_1(x) &= 10 + \frac{43}{11}x - \frac{31}{88}x^3, \quad x \in [0, 2] \\ C_2(x) &= \frac{12}{11} + \frac{190}{11}x - \frac{147}{22}x^2 + \frac{67}{88}x^3, \quad x \in [2, 4] \\ C_3(x) &= \frac{1124}{11} - \frac{644}{11}x + \frac{135}{11}x^2 - \frac{9}{11}x^3, \quad x \in [4, 5], \\ C(1) &= 13.5568 \\ \text{b) } C_1(x) &= 5.25 + 2.25x + 0.75x^2 - 0.25x^3, \quad x \in [1, 3] \\ C_2(x) &= -15 + 22.5x - 6x^2 + 0.5x^3, \quad x \in [3, 5] \\ C_3(x) &= 78.75 - 33.75x + 5.25x^2 - 0.25x^3, \quad x \in [5, 7] \\ C(4.5) &= 10.3125 \end{aligned}$$

Zadatak 2.21. Na osnovi podataka

$$\begin{array}{c|cccccc} x_i & 0 & 1 & 2 & 3 & 4 & 5 \\ \hline y_i & 1 & 2 & 1.5 & 2 & 2.5 & 2 \end{array}, \quad \xi = 2.5,$$

odredite prirodni kubični interpolacijski spline i izračunajte  $C(\xi)$ .

Uputa: Sustav (2.31) u ovom slučaju glasi: 
$$\begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix} \mathbf{s} = \begin{bmatrix} -9 \\ 6 \\ 0 \\ -6 \end{bmatrix}, \quad i \text{ ima}$$

rješenje  $\mathbf{s} = (0, -2.8134, 2.25359, -0.200957, -1.44976, 0)^T$ .

Rješenje:

$$C_1(x) = 1 + 1.4689x - 0.4689x^3, \quad x \in [0, 1]$$

$$C_2(x) = -0.313397 + 5.409091x - 3.94019x^2 + 0.844498x^3, \quad x \in [1, 2]$$

$$C_3(x) = 9.71531 - 9.63397x + 3.58134x^2 - 0.409091x^3, \quad x \in [2, 3]$$

$$C_4(x) = 4.28947 - 4.20813x + 1.77273x^2 - 0.208134x^3, \quad x \in [3, 4]$$

$$C_5(x) = -24.4952 + 17.3804x - 3.6244x^2 + 0.241627x^3, \quad x \in [4, 5],$$

$$C(2.5) = 2.28439$$

Zadatak 2.22. Problem iz *Primjera 2.4*, str. 27 riješite primjenom kubičnog interpolacijskog spline-a. Koristite program u *Dodatku*.

Rješenje:  $C(500) = 10.54086$ ,  $C(750) = 17.2716$ ,  $C(1000) = 24.32858$ .

Zadatak 2.23. Program u *Dodatku* za izračunavanje kubičnog interpolacijskog spline-a dopunite tako da računa prvu, drugu i treću derivaciju dobivenog interpolacijskog spline-a u točki  $\xi \in [x_0, x_n]$ .

## Poglavlje 3

# Rješavanje sustava linearnih jednadžbi

Promatramo sustav linearnih jednadžbi

$$\mathbf{Ax} = \mathbf{b}, \quad (3.1)$$

gdje je  $\mathbf{A} \in \mathbb{R}^{m \times n}$  matrica,  $\mathbf{x} \in \mathbb{R}^m$  vektor nepoznanica, a  $\mathbf{b} \in \mathbb{R}^n$  vektor slobodnih koeficijenata. Osim klasičnih metoda za rješavanje sustava (3.1), kao što je Gaussova metoda, postoje metode koje se zasnivaju na različitim dekompozicijama matrice  $\mathbf{A}$ , kao što je Cholesky dekompozicija, QR dekompozicija, dekompozicija na singularne vrijednosti, itd. Posebnu klasu metoda predstavljaju tzv. iterativne metode, kao što je Gauss-Seidelova ili Jacobijeva metoda. Također, od posebnog su praktičnog značaja tzv. veliki rijetki sustavi linearnih jednadžbi za koje postoje specijalne metode za rješavanje.

**Primjer 3.1.** *Za početak pogledajmo jednostavni sustav jednadžbi*

$$2x_1 + 400x_2 = 200$$

$$x_1 + x_2 = 1$$

*čije je rješenje*

$$x_1 = 0.5025125628 \dots, \quad x_2 = 0.4974874372 \dots$$

Riješit ćemo ovaj sustav na računalu s 2-znamenkastom floating-point aritmetikom. Množeći prvu jednadžbu s  $(-0.5)$  i dodavajući drugoj, dobivamo

$$\begin{aligned} 2x_1 + 400x_2 &= 200 \\ -200x_2 &= -99 \end{aligned}$$

Iz druge jednadžbe dobivamo  $x_2^* = 0.50$ , a nakon toga iz prve,  $x_1^* = (200 - 400 \times 0.50)/2 = 0$ .

Ako bismo račun proveli u 3-znamenkastoj floating-point aritmetici, dobili bismo  $x_2^{**} = 0.497$ ,  $x_1^{**} = 0.600$ .

Iz ovog jednostavnog primjera vidi se da je važno znati procijeniti veličinu pogreške aproksimacije rješenja.

Ovisno o veličini i strukturi matrice  $\mathbf{A}$ , prilikom rješavanja sustava (3.1) u praktičnim primjenama mogu nastupiti različiti numerički problemi. Pokušat ćemo proanalizirati razloge ovim pojavama i uputiti na mogućnost dobivanja što korektnijeg rješenja u nekim za praksu važnim slučajevima.

### 3.1 Norma vektora i matrice

**Definicija 3.1.** Neka je  $X$  realni vektorski prostor. Funkciju  $\|\cdot\| : X \rightarrow [0, \infty)$ , koja svakom vektoru  $\mathbf{x} \in X$  pridružuje nenegativan realan broj (kojeg ćemo označiti s  $\|\mathbf{x}\|$ ) zovemo *norma* ako vrijedi

- (i)  $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$  (pozitivna definitnost),
- (ii)  $\|\lambda\mathbf{x}\| = |\lambda| \|\mathbf{x}\|$  za svaki  $\lambda \in \mathbb{R}$  i za svaki  $\mathbf{x} \in X$ ,
- (iii)  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  za svaki  $\mathbf{x}, \mathbf{y} \in X$  (nejednakost trokuta).

Vrijednost funkcije  $\|\cdot\|$  na nekom vektoru  $\mathbf{x} \in X$  zvat ćemo *norma vektora  $\mathbf{x}$* . U ovom poglavlju razmatrat ćemo vektorski prostor  $\mathbb{R}^n$ , čiji su elementi vektori stupci. Zato obično pišemo  $\mathbf{x} = (x_1, \dots, x_n)^T$ .

Najčešće korištene vektorske norme su

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|, \quad (l_1 \text{ norma})$$

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T \mathbf{x}} = \left( \sum_{i=1}^n x_i^2 \right)^{1/2}, \quad (l_2 \text{ ili Euklidova norma})$$

$$\|\mathbf{x}\|_\infty = \max_{i=1, \dots, n} |x_i|, \quad (l_\infty \text{ norma ili Čebiševljeva norma})$$

**Zadatak 3.1.** Pokažite da spomenute norme imaju sva svojstva navedena u *Definiciji 3.1*.

Može se pokazati da su sve tri navedene norme međusobno ekvivalentne (vidi primjerice Kurepa (1967), tj. za svaki par  $i, j \in \{1, 2, \infty\}$  postoje konstante  $\alpha, \beta > 0$  takve da vrijedi

$$\alpha \|\mathbf{x}\|_i \leq \|\mathbf{x}\|_j \leq \beta \|\mathbf{x}\|_i \quad \text{za svaki } \mathbf{x} \in \mathbb{R}^n.$$

**Zadatak 3.2.** Jedinična kuglina ljuska (sfera) sa središtem u  $\mathbf{O} \in \mathbb{R}^n$  definira se s

$$\partial K_n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}.$$

Nacrtajte jedinične kugline ljuske za  $n = 2$  i norme  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ .

*Primjedba 3.1.* Odgovarajuće norme u vektorskom prostoru  $C[a, b]$ ,  $[a, b] \subset \mathbb{R}$  svih neprekidnih funkcija na intervalu  $[a, b]$  bile bi

$$\|f\|_1 = \int_a^b |f(t)| dt, \quad \|f\|_2 = \left( \int_a^b f^2(t) dt \right)^{1/2}, \quad \|f\|_\infty = \max_{t \in [a, b]} |f(t)|$$

**Definicija 3.2.** Neka je  $\mathbb{R}^{n \times n}$  skup svih realnih matrica tipa  $(n \times n)$ . Funkciju  $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow [0, \infty)$ , koja svakoj matrici  $\mathbf{A} \in \mathbb{R}^{n \times n}$  pridružuje nenegativan realan broj (kojeg ćemo označiti s  $\|\mathbf{A}\|$ ) zovemo **norma matrice  $\mathbf{A}$**  ako vrijedi

- (i)  $\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = \mathbf{0}$ , za svaki  $\mathbf{A} \in \mathbb{R}^{n \times n}$
- (ii)  $\|\lambda \mathbf{A}\| = |\lambda| \|\mathbf{A}\|$  za svaki  $\lambda \in \mathbb{R}$  i  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,
- (iii)  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$  za svaki  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$
- (iv)  $\|\mathbf{A} \cdot \mathbf{B}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$  za sve  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$

Ako smo na  $\mathbb{R}^n$  definirali vektorsku normu  $\|\cdot\|$ , onda njoj možemo pridružiti matičnu normu na sljedeći način

$$\|\mathbf{A}\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}, \quad (3.2)$$

koju ćemo zvati **inducirana matična norma**. Primijetite da je u (3.2) norma matrica  $\mathbf{A}$  definirana kao najveći mogući omjer norme vektora

$\mathbf{Ax}$  i norme vektora  $\mathbf{x}$ ,  $\mathbf{x} \neq \mathbf{0}$ . Iz (3.2) slijedi  $\|\mathbf{A}\| \geq \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}$ , odnosno

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|. \quad (3.3)$$

Matrične norme inducirane ranije spomenutim vektorskim normama su

$$\begin{aligned} \|\mathbf{A}\|_1 &= \max_{j=1,\dots,n} \|\mathbf{a}_j\|_1, \quad \text{gdje je } \mathbf{a}_j \text{ } j\text{-ti stupac matrice } \mathbf{A} \\ \|\mathbf{A}\|_2 &= \sigma_1(\mathbf{A}) - \text{najveća singularna vrijednost matrice } \mathbf{A} \quad (\text{vidi } t.3.9, \text{ str. } 77) \\ \|\mathbf{A}\|_\infty &= \max_{i=1,\dots,n} \|\mathbf{a}_i^T\|_1, \quad \text{gdje je } \mathbf{a}_i \text{ } i\text{-ti redak matrice } \mathbf{A}. \end{aligned}$$

Ako za neku matričnu i za neku vektorsku normu vrijedi (3.3), kažemo da su one **kompatibilne** (ili konzistentne).

Osim navedenih, važna matrična norma je tzv. Frobeniusova norma matrice  $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\|\mathbf{A}\|_F = \left( \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2},$$

Frobeniusova norma nije inducirana matrična norma, ali je kompatibilna s Euklidovom vektorskom normom (vidi primjerice Truhar (2010)), tj. vrijedi

$$\|\mathbf{Ax}\|_2 \leq \|\mathbf{A}\|_F \|\mathbf{x}\|_2.$$

Ako je  $\tilde{\mathbf{x}}$  aproksimacija vektora  $\mathbf{x}$ , a  $\tilde{\mathbf{A}}$  aproksimacija matrice  $\mathbf{A}$ , onda analogno kao što smo to uradili u *t.1.2*, str. 5, možemo definirati

	Vektor	Matrica
Pogreška aproksimacije	$\Delta \mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}$	$\Delta \mathbf{A} = \tilde{\mathbf{A}} - \mathbf{A}$
Apsolutna pogreška aproksimacije	$\ \Delta \mathbf{x}\  = \ \tilde{\mathbf{x}} - \mathbf{x}\ $	$\ \Delta \mathbf{A}\  = \ \tilde{\mathbf{A}} - \mathbf{A}\ $
Relativna pogreška aproksimacije	$\frac{\ \tilde{\mathbf{x}} - \mathbf{x}\ }{\ \mathbf{x}\ }$	$\frac{\ \tilde{\mathbf{A}} - \mathbf{A}\ }{\ \mathbf{A}\ }$

### 3.2 Uvjetovanost sustava linearnih jednadžbi

Promatramo sustav jednadžbi  $\mathbf{Ax} = \mathbf{b}$  uz pretpostavku da je  $\mathbf{A}$  kvadratna regularna matrica. Tada sustav ima jedinstveno rješenje  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Možemo postaviti ovakvo pitanje:

- (i) *Koliko će se promijeniti rješenje  $\mathbf{x}$  sustava, ako se promijeni vektor slobodnih koeficijenata  $\mathbf{b}$ , a matrica  $\mathbf{A}$  ostane nepromijenjena?*

Neka je  $\mathbf{x}$  rješenje sustava  $\mathbf{Ax} = \mathbf{b}$ , a  $\tilde{\mathbf{x}} = \mathbf{x} + \Delta\mathbf{x}$  rješenje “perturbiranog sustava”  $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{b} + \Delta\mathbf{b}$ , tj.

$$\mathbf{A}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b}. \quad (3.4)$$

Iz (3.1) i (3.4) dobivamo

$$\Delta\mathbf{x} = \mathbf{A}^{-1}\Delta\mathbf{b}, \quad (3.5)$$

Korištenjem kompatibilne matrične i vektorske norme, iz (3.1) i (3.5) dobivamo

$$\|\mathbf{b}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|, \quad (3.6)$$

$$\|\Delta\mathbf{x}\| \leq \|\mathbf{A}^{-1}\|\|\Delta\mathbf{b}\|. \quad (3.7)$$

Dijeljenjem (3.7) s  $\|\mathbf{x}\|$  i korištenjem (3.6) dobivamo

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}^{-1}\| \cdot \|\mathbf{A}\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}, \quad (3.8)$$

čime smo ocijenili relativnu pogrešku vektora  $\tilde{\mathbf{x}}$  pomoću relativne pogreške vektora  $\mathbf{b}$ . Vrijedi također (vidi primjerice Golub and van Loan (1996)):

- (ii) *Ako se matrica sustava  $\mathbf{A}$  promijeni za  $\Delta\mathbf{A}$ , a vektor  $\mathbf{b}$  ostane nepromijenjen, onda relativnu pogrešku aproksimacije rješenja  $\tilde{\mathbf{x}}$  možemo ocijeniti s*

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|}{\|\tilde{\mathbf{x}}\|} \leq \|\mathbf{A}^{-1}\| \cdot \|\mathbf{A}\| \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}. \quad (3.9)$$

- (iii) *Ako se matrica sustava  $\mathbf{A}$  promijeni za  $\Delta\mathbf{A}$ , a vektor  $\mathbf{b}$  za  $\Delta\mathbf{b}$ , onda relativnu pogrešku aproksimacije rješenja  $\tilde{\mathbf{x}}$  možemo ocijeniti s*

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|(\mathbf{A} + \Delta\mathbf{A})^{-1}\|\|\mathbf{A}\| \left( \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \right). \quad (3.10)$$

U sva tri slučaja u ocjeni relativne pogreške aproksimacije rješenja pojavio se broj  $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$ . Ovaj broj zvat ćemo broj *uvjetovanosti* regularne matrice  $\mathbf{A}$  (*condition number*) i označavati s

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|.$$



Broj uvjetovanosti  $\text{cond}(\mathbf{A})$  uvijek je veći ili jednak 1. Naime,

$$1 \leq \|\mathbf{I}\| = \|\mathbf{A}\mathbf{A}^{-1}\| \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\| = \text{cond}(\mathbf{A}).$$

Ako je  $\text{cond}(\mathbf{A})$  blizu 1, kažemo da je matrica **dobro uvjetovana**, a ako je  $\text{cond}(\mathbf{A})$  mnogo veći od 1, kažemo da je matrica **A loše uvjetovana** (*ill-conditioned*).

Iz (3.8), (3.9) i (3.10) vidi se da će i male pogreške u matrici ili vektoru slobodnih koeficijenata sustava s loše uvjetovanom matricom sustava rezultirati značajnom pogreškom u rješenju sustava.

**Primjer 3.2.** *Rješenje sustava linearnih jednadžbi, gdje su matrica sustava  $\mathbf{A}$  i vektor  $\mathbf{b}$  zadani s*

$$\mathbf{A} = \begin{bmatrix} 0.234 & 0.458 \\ 0.383 & 0.750 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0.224 \\ 0.367 \end{bmatrix},$$

je  $\mathbf{x} = (-1, 1)^T$ . Promijenimo vektor  $\mathbf{b}$  za  $\Delta\mathbf{b} = (0.00009, 0.000005)^T$ .

Egzaktno rješenje sustava  $\mathbf{A}\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$  je  $\tilde{\mathbf{x}} = (-0.241774, 0.612791)^T$ , pri čemu je  $\Delta\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x} = (0.758256, -0.387209)^T$ . Izračunajmo relativne pogreške vektora  $\mathbf{b}$  i  $\mathbf{x}$  u normi  $\|\cdot\|_\infty$ .

$$\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{0.00009}{0.367} \approx 2.45 \times 10^{-4}, \quad \frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{0.758256}{1} = 0.758256.$$

Dakle, relativna pogreška rješenja je više od 3000 puta veća od relativne pogreške vektora  $\mathbf{b}$ .

Načinimo sada malu promjenu u elementu  $a_{11}$  matrice  $\mathbf{A}$ , tako da je

$$\mathbf{A} + \Delta\mathbf{A} = \begin{bmatrix} 0.233 & 0.458 \\ 0.383 & 0.750 \end{bmatrix}.$$

Egzaktno rješenje sustava  $(\mathbf{A} + \Delta\mathbf{A})\tilde{\mathbf{x}} = \mathbf{b}$  je  $\tilde{\mathbf{x}} = (0.129518, 0.423193)^T$ . Sada je

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|}{\|\tilde{\mathbf{x}}\|} = \frac{1.129518}{0.423193} \approx 2.67, \quad \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} = \frac{0.001}{1.133} \approx 0.00075,$$

pa je relativna pogreška aproksimacije rješenja opet više od 3000 puta veća od relativne pogreške u matrici  $\mathbf{A}$ .

Razloge za ove pojave treba tražiti u veličini broja uvjetovanosti matrice  $\mathbf{A}$ . Kako je

$$\mathbf{A} = \begin{bmatrix} 0.234 & 0.458 \\ 0.383 & 0.750 \end{bmatrix}, \quad \mathbf{A}^{-1} = \begin{bmatrix} 8720.93 & -5325.58 \\ -4453.49 & 2720.93 \end{bmatrix},$$

dobivamo  $\|\mathbf{A}\|_\infty = 1.133$  i  $\|\mathbf{A}^{-1}\|_\infty = 14\,046.51$ , pa je broj uvjetovanosti matrice  $\mathbf{A}$   $\text{cond}(\mathbf{A}) = 15\,914.70$ .

*Zadatak 3.3.* Provedite čitav račun iz *Primjera 3.2* za norme  $\|\cdot\|_1$  i  $\|\cdot\|_2$ .

Uputa: Broj uvjetovanosti  $\text{cond}(\mathbf{A})$  u  $\|\cdot\|_2$  može se izračunati (vidi primjerice Golub and van Loan (1996)) po formuli

$$\text{cond}(\mathbf{A}) = \frac{\sigma_1}{\sigma_n},$$

gdje je  $\sigma_1$  najveća, a  $\sigma_n$  najmanja singularna vrijednost matrice  $\mathbf{A}$  (vidi također *t.3.9*, str. 77). Singularne vrijednosti kvadratne matrice  $\mathbf{A}$  mogu se lako izračunati programom *Mathematica*- naredbom: `SingularValues[A][[2]]`.

*Primjedba 3.2.* Moglo bi se pomisliti da je veličina determinante matrice u direktnoj vezi s uvjetovanosti matrice. Da te dvije veličine nisu u direktnoj vezi vidi se već iz sljedećeg primjera. Tako je primjerice dijagonalna matrica  $\mathbf{D} = \text{diag}(10^{-10}, 10^{-10})$  perfektno uvjetovana kao jedinična matrica ( $\text{cond}(\mathbf{D}) = 1$ ), a njezina determinanta je skoro nula ( $\det \mathbf{D} = 10^{-20}$ ). S druge strane, dijagonalna matrica  $\mathbf{S} = \text{diag}(10^5, 10^{-5})$  je loše uvjetovana ( $\text{cond}(\mathbf{S}) = 10^{10}$ ), iako je  $\det \mathbf{S} = 1$ .

### 3.3 Rješavanje trokutastih sustava

Sustav linearnih jednadžbi  $\mathbf{U}\mathbf{x} = \mathbf{b}$ , gdje je  $\mathbf{U} \in \mathbb{R}^{n \times n}$  gornja<sup>1</sup> trokutasta matrica

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}, \quad u_{ii} \neq 0, i = 1, \dots, n$$

lako se rješava rješavanjem unazad (engl: *Back Substitution* (BS)):

$$x_i = \frac{1}{u_{ii}} \left( b_i - \sum_{k=i+1}^n u_{ik} x_k \right), \quad i = n, n-1, \dots, 1. \quad (3.11)$$

<sup>1</sup>engl: upper = gornji

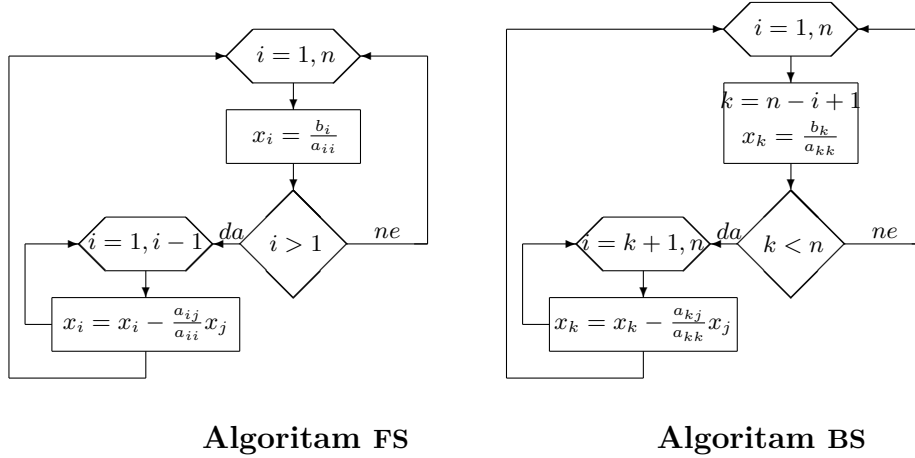
Slično, sustav linearnih jednadžbi  $\mathbf{Lx} = \mathbf{b}$ , gdje je  $\mathbf{L} \in \mathbb{R}^{n \times n}$  donja<sup>2</sup> trokutasta matrica

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}, \quad l_{ii} \neq 0, i = 1, \dots, n$$

rješava se rješavanjem unaprijed (engl: *Forward Substitution* (FS)):

$$x_i = \frac{1}{l_{ii}} \left( b_i - \sum_{k=1}^{i-1} l_{ik} x_k \right), \quad i = 1, \dots, n. \quad (3.12)$$

Niže su navedeni odgovarajući dijagrami toka, *Mathematica*-programi za rješavanje donjeg trokutastog (Algoritam FS) i gornjeg trokutastog (Algoritam BS) sustava.



```
FS[n_, a_, b_] := Module[{i, j}, x = Table[0, {i, n}];
  x[[1]] = b[[1]]/a[[1, 1]];
  Do[
    x[[i]] = (b[[i]] - Sum[a[[i, j]] * x[[j]], {j, i - 1}]) / a[[i, i]]
    , {i, 2, n}
  ]
```

```
BS[n_, a_, b_] := Module[{i, j, k}, x = Table[0, {i, n}];
  x[[n]] = b[[n]]/a[[n, n]];
```

<sup>2</sup>engl: lower = donji

```

Do [k = n-i;
  x[[k]]=(b[[k]]-Sum[a[[k,j]]*x[[j]],
  ,{j,k+1,n}])/a[[k,k]]
  ,{i,n-1}]
]

```

*Zadatak 3.4.* Specijalizirajte algoritme FS i BS za slučaj trodijagonalne matrice<sup>3</sup> (kakva se primjerice pojavljuje kod prirodnog kubičnog splinea, str. 38).

Uputa: Od cijele matrice treba “pamtiti” samo tri vektora: glavnu dijagonalu, te donju i gornju sporednu dijagonalu matrice  $\mathbf{A}$ .

### 3.4 Gaussova metoda eliminacije

Zadana je kvadratna regularna matrica  $\mathbf{A}$  i vektor slobodnih koeficijenata  $\mathbf{b}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \quad (3.13)$$

Prema *Kronecker-Capellijevom* teoremu sustav  $\mathbf{Ax} = \mathbf{b}$  je rješiv i ima jedinstveno rješenje<sup>4</sup>.

Pretpostavimo da je tzv. *pivot-element*  $a_{11} \neq 0$ . Najprije ćemo primjenom Gaussovih transformacija matricu  $\mathbf{A}$  svesti na oblik

$$\mathbf{A}^{(2)} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}, \quad \mathbf{b}^{(2)} = \begin{bmatrix} b_1 \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{bmatrix},$$

gdje je, uz oznaku  $m_{i1} = \frac{a_{i1}}{a_{11}}$ ,

$$a_{ik}^{(2)} = a_{ik} - m_{i1}a_{1k}, \quad b_i^{(2)} = b_i - m_{i1}b_1, \quad (i, k = 2, \dots, n) \quad (3.14)$$

<sup>3</sup>Trodijagonalna matrica je specijalni slučaj tzv. vrpčastih matrica, koje spadaju u tzv. rijetko popunjene matrice (sparse matrices).

<sup>4</sup>Iako se Gaussovom metodom može rješavati opći sustav linearnih jednadžbi (gdje broj jednadžbi nije jednak broju nepoznanica ili gdje matrica sustava nije regularna), mi ćemo se zadržati na ovom specijalnom slučaju, koji je i najčešći u primjenama (vidi također Jukić and Scitovski (2004)).

Uz pretpostavku da je sljedeći pivot-element  $a_{22}^{(2)} \neq 0$ , na sličan način dobivamo

$$\mathbf{A}^{(3)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}, \quad \mathbf{b}^{(3)} = \begin{bmatrix} b_1 \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(3)} \end{bmatrix},$$

gdje je, uz oznaku  $m_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$ ,

$$a_{ik}^{(3)} = a_{ik}^{(2)} - m_{i2}a_{2k}^{(2)}, \quad b_i^{(3)} = b_i^{(2)} - m_{i2}b_2^{(2)}, \quad (i, k = 3, \dots, n) \quad (3.15)$$

Uz pretpostavku da su svi pivot-elementi bili različiti od nule, nakon  $(n - 1)$  ovakvih koraka (*iteracija*) matrica  $\mathbf{A}$  prijeći će u oblik gornje trokutaste matrice, gdje su svi dijagonalni elementi različiti od nule. Takav sustav jednadžbi lako se rješava (*Algoritam BS*, str. 56).

*Primjedba 3.3.* Primjenom Gaussovih transformacija matricu sustava mogli bismo svesti i na dijagonalnu matricu. U tom slučaju govorimo o Gauss-Jordanovoj metodi, koja igra važnu ulogu kod *simplex-metode* za rješavanje problema linearnog programiranja.

Vidjeli smo da je prilikom provođenja Gaussovog postupka važno da su svi pivot-elementi:  $a_{11}, a_{22}^{(2)}, \dots$  različiti od nule. To neće biti uvijek osigurano, iako je matrica sustava regularna. Ipak, postoje dva važna slučaja, gdje će spomenuti uvjet biti ispunjen. To su slučajevi kada je matrica sustava

- dijagonalno dominantna, tj  $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$
- simetrična i pozitivno definitna, tj  $\mathbf{A}^T = \mathbf{A}$  &  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  za sve  $\mathbf{x} \neq \mathbf{0}$ . Prema poznatom Sylvesterovom kriteriju, simetrična matrica  $\mathbf{A}$  je pozitivno definitna onda i samo onda ako su svi njezini glavni minori<sup>5</sup> pozitivni.

<sup>5</sup>Glavni minori matrice  $\mathbf{A}$  su determinante  $\det \mathbf{A}_k$ ,  $k = 1, \dots, n$ , gdje je  $\mathbf{A}_k$  matrica dobivena od elemenata matrice  $\mathbf{A}$  na presjeku prvih  $k$  stupaca i prvih  $k$  redaka

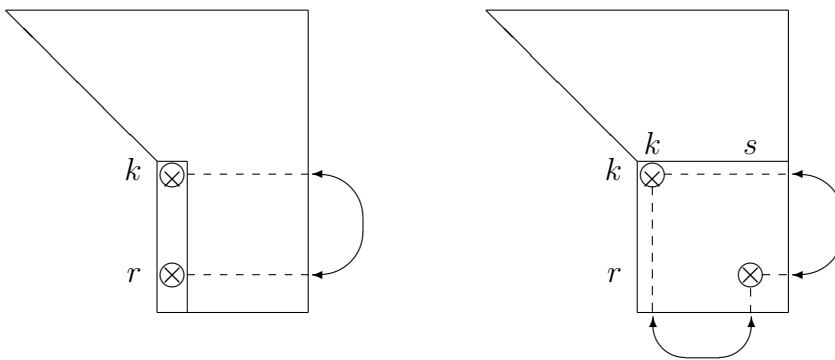
Dijagonalna dominantnost, odnosno pozitivna definitnost vrlo su jaki uvjeti na matricu  $\mathbf{A}$ . Da bi se ipak proširila klasa matrica za koje bi Gaussov postupak bio provediv, ranije opisanu Gaussovu metodu eliminacije malo ćemo modificirati.

Jedan pokušaj u tom smislu je tzv. strategija parcijalnog pivotiranja:

*u  $k$ -toj iteraciji treba zamijeniti  $k$ -ti s  $r$ -tim retkom, gdje je  $r$  najmanji indeks retka takav da je*

$$|a_{rk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$$

*(vidi shemu na Slici 3.1)*



Slika 3.1 Strategija parcijalnog (lijevo) i strategija potpunog (desno) pivotiranja

Drugi pokušaj je tzv. strategija potpunog pivotiranja:

*u  $k$ -toj iteraciji najprije pronađemo cijele brojeve  $r$  i  $s$ , za koje je*

$$|a_{rs}^{(k)}| = \max_{k \leq i, j \leq n} |a_{ij}^{(k)}|,$$

*a onda sukcesivno napravimo zamjenu  $k$ -tog i  $s$ -tog stupaca, te  $k$ -tog i  $r$ -tog retka (vidi shemu na Slici 3.1)*

**Zadatak 3.5.** Pokažite da je Gaussov postupak bez pivotiranja provediv onda i samo onda ako su svi glavni minori matrica sustava različiti od nule.

**Primjer 3.3.** *Ako bismo u Primjeru 3.1, str. 49 za slučaj 2-znamenkaste floating-point aritmetike primijenili strategiju parcijalnog pivotiranja, dobili bismo isti (loši) rezultat. Primjenom strategije potpunog pivotiranja dobili bismo “točan” rezultat:  $x_1^* = x_2^* = 0.50$ .*

*U 3-znamenkastoj floating-point aritmetici već uz primjenu strategije parcijalnog pivotiranja dobili bismo “točan” rezultat  $x_1^{**} = 0.503$ ,  $x_2^{**} = 0.498$ .*

Ovaj primjer pokazuje da se strategijom parcijalnog, a naročito strategijom potpunog pivotiranja također može postići i numerička stabilnost Gaussovog postupka.

**Zadatak 3.6.** Izradite program koji će u  $k$ -znamenkastoj floating-point aritmetici ( $k = 2, 3, \dots, 8$ ) rješavati sustav linearnih jednadžbi  $\mathbf{Ax} = \mathbf{b}$ :

- Gaussovom metodom bez pivotiranja;
- Gaussovom metodom uz strategiju parcijalnog pivotiranja;
- Gaussovom metodom uz strategiju potpunog pivotiranja;

Program ispitajte na sljedećim primjerima:

(a)

$$\begin{bmatrix} 0.0001 & 0.5 \\ 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix},$$

Egzaktno rješenje zaokruženo na 4 znamenke je  $\mathbf{x} = (0.9999, 0.9998)^T$ .

(b)

$$\begin{aligned} \epsilon x_1 &= b \\ \epsilon x_2 &= c, \\ x_1 + x_2 + x_3 &= a \end{aligned}$$

gdje je  $\epsilon, a, b, c \in \mathbb{R}$ . Egzaktno rješenje je  $x_1 = \frac{b}{\epsilon}$ ,  $x_2 = \frac{c}{\epsilon}$ ,  $x_3 = a - \frac{b}{\epsilon} - \frac{c}{\epsilon}$ .

(c)

Sustav kome je matrica sustava tzv. Hilbertova matrica  $H_{ij} = \frac{1}{i+j-1}$ ,  $i = 1, \dots, n$ , s proizvoljnim vektorom slobodnih koeficijenata.

### 3.5 LU-dekompozicija

Pretpostavimo da je  $\mathbf{A} \in \mathbb{R}^{n \times n}$  regularna kvadratna matrica, kojoj su svi glavni minori različiti od nule. Tada je na jedinstven način moguće načiniti rastav (?Golub and van Loan, 1996)

$$\mathbf{A} = \mathbf{LU}$$

gdje je  $\mathbf{L}$  donja trokutasta matrica, kojoj su na glavnoj dijagonali jedinice, a  $\mathbf{U}$  gornja trokutasta matrica, čiji dijagonalni elementi nisu nule

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix} \quad (3.16)$$

*Zadatak 3.7.* Matrica  $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$  je regularna kvadratna matrica. Pokažite da za ovu matricu ipak nije moguće načiniti LU-dekompoziciju.

Najprije ćemo pokazati da Gaussova metoda eliminacije opisana u prethodnoj točki primijenjena na kvadratnu matricu  $\mathbf{A}$ , kojoj su svi glavni minori različiti od nule generira LU-dekompoziciju te matrice.

Pretpostavimo da je  $\mathbf{A}$  kvadratna regularna matrica kojoj su svi glavni minori različiti od nule. Ako uvedemo vektore

$$\mathbf{u}_1 = (0, m_{21}, \dots, m_{n1})^T, \quad \mathbf{v}_1 = (1, 0, \dots, 0)^T,$$

i elementarnu matricu  $\mathbf{M}_1$ :

$$\mathbf{M}_1 = \mathbf{I} - \mathbf{u}_1 \mathbf{v}_1^T = \begin{bmatrix} 1 & & & \\ -m_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ -m_{n1} & 0 & \cdots & 1 \end{bmatrix},$$

gdje su elementi  $m_{i,j}$  zadani kao u t.3.4, str. 57, onda matricu  $\mathbf{A}^{(2)}$  dobivenu u prvom koraku Gaussovog postupka eliminacije možemo pisati kao

$$\mathbf{A}^{(2)} = \mathbf{M}_1 \mathbf{A},$$

Matricu  $\mathbf{A}^{(3)}$  možemo pisati kao

$$\mathbf{A}^{(3)} = \mathbf{M}_2 \mathbf{M}_1 \mathbf{A},$$

gdje je  $\mathbf{M}_2 = \mathbf{I} - \mathbf{u}_2 \mathbf{v}_2^T$  elementarna matrica s vektorima

$$\mathbf{u}_2 = (0, 0, m_{32}, \dots, m_{n2})^T, \quad \mathbf{v}_2 = (0, 1, 0, \dots, 0)^T$$



Na kraju Gaussovog postupka matrica  $\mathbf{A}$  primit će oblik gornje trokutaste matrice, koju ćemo označiti s  $\mathbf{U}$  i zapisati kao

$$\mathbf{U} := \mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A}$$

Sve matrice  $\mathbf{M}_1, \dots, \mathbf{M}_{n-1}$  su regularne donje trokutaste matrice s jedinicama na glavnoj dijagonali. Zato su takve i njihove inverzne matrice. Možemo pisati

$$\mathbf{A} = \mathbf{L}\mathbf{U} \quad \text{gdje je} \quad \mathbf{L} = \mathbf{M}_1^{-1} \mathbf{M}_2^{-1} \cdots \mathbf{M}_{n-1}^{-1}.$$

Matrica  $\mathbf{L}$  je donja trokutasta matrica s jedinicama na glavnoj dijagonali jer su sve matrice  $\mathbf{M}_k$  takvog oblika.

*Zadatak 3.8.* Pokažite da je inverzna matrica matrice  $\mathbf{M}_k = \mathbf{I} - \mathbf{u}_k \mathbf{v}_k^T$  iz gornjeg izvoda zadana s

$$\mathbf{M}_k^{-1} = \mathbf{I} + \mathbf{u}_k \mathbf{v}_k^T.$$

*Primjedba 3.4.* Općenitije može se pokazati da vrijedi tzv. SHERMAN-MORRISONOVA<sup>6</sup> lema:

Neka je  $\mathbf{A} \in \mathbb{R}^{n \times n}$  regularna matrica i  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ . Tada je matrica  $\mathbf{A} + \mathbf{u}\mathbf{v}^T$  regularna onda i samo onda ako je  $\sigma = 1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u} \neq 0$ . U tom slučaju vrijedi formula

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{1}{\sigma} \mathbf{A}^{-1} \mathbf{u}\mathbf{v}^T \mathbf{A}^{-1}.$$

*Primjedba 3.5.* Ako svi glavni minori regularne kvadratne matrice  $\mathbf{A}$  nisu različiti od nule (vidi *Primjer 3.7*), onda postoji matrica permutacija  $\mathbf{P}$ , takva da je  $\mathbf{P}\mathbf{A} = \mathbf{L}\mathbf{U}$  LU-dekompozicija matrice  $\mathbf{P}\mathbf{A}$ . Matrica  $\mathbf{P}$  pojavljuje se prilikom provođenja Gaussovog postupka uz primjenu strategije parcijalnog pivotiranja (Gill et al., 1991).

U nekim praktičnim situacijama (primjerice – analiza modela nacionalne privrede) treba više puta rješavati sustav  $\mathbf{A}\mathbf{x} = \mathbf{b}_i$ , gdje je matrica sustava (“matrica tehnologije”) uvijek ista, a vektor slobodnih koeficijenata (“outputi”) se mijenja. Također se može dogoditi da treba riješiti sustave  $\mathbf{A}\mathbf{x}_1 = \mathbf{b}_1$  i  $\mathbf{A}\mathbf{x}_2 = \mathbf{b}_2$ , gdje je  $\mathbf{b}_2$  neka funkcija od  $\mathbf{x}_1$ .

<sup>6</sup>J. SHERMAN, W. J. MORRISON, *Adjustments of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix*, Ann. Math. Stat. **20**(1949), 621.

U takvim situacijama korisno je poznavati LU-dekompoziciju matrice sustava  $\mathbf{A}=\mathbf{LU}$ .

Sustav  $\mathbf{Ax} = \mathbf{b}$  tada glasi

$$\mathbf{LUx} = \mathbf{b}, \quad (3.17)$$

Ako označimo

$$\mathbf{z} := \mathbf{Ux}, \quad (3.18)$$

onda (3.17) postaje sustav

$$\mathbf{Lz} = \mathbf{b}, \quad (3.19)$$

koji se lako rješava jer je matrica sustava  $\mathbf{L}$  donja trokutasta matrica s jedinicama na glavnoj dijagonali (*Algoritam FS*). Rješenje  $\mathbf{z}^*$  sustava (3.19) uvrstimo u (3.18), koji se sada lako rješava jer je matrica sustava  $\mathbf{U}$  gornja trokutasta matrica čiji dijagonalni elementi nisu nule (*Algoritam BS*). Rješenje  $\mathbf{x}^*$  sustava (3.18) je i rješenje polaznog sustava.

Primijetimo također da poznavajući LU-dekompoziciju matrice  $\mathbf{A}$ , prema *Binet-Cauchyjevom teoremu* (vidi primjerice Jukić and Scitovski (2004), str. 244) lako možemo izračunati determinantu matrice  $\mathbf{A}$

$$\det \mathbf{A} = u_{11}u_{22} \cdots u_{nn}.$$

Matrice  $\mathbf{L}$  i  $\mathbf{U}$  u rastavu  $\mathbf{A} = \mathbf{LU}$  možemo dobiti primjerice tzv. Croutovom metodom (Stoer, 2002; Stoer and Bulirsch, 2002):

$$\begin{aligned} (1) \quad a_{1j} &= \sum_{k=1}^1 l_{1k}u_{kj}, \quad u_{1j} := a_{1j}, \quad j = 1, \dots, n \\ (2) \quad a_{i1} &= \sum_{k=1}^1 l_{ik}u_{k1}, \quad l_{i1} := a_{i1} := \frac{a_{i1}}{u_{11}}, \quad i = 2, \dots, n \\ (3) \quad a_{2j} &= \sum_{k=1}^2 l_{2k}u_{kj}, \quad u_{2j} := a_{2j} - l_{21}u_{1j}, \quad j = 2, \dots, n \\ &\vdots \end{aligned}$$

**Algorithm 2** (LU-dekompozicija)

1. Učitati elemente matrice  $a_{ij}$ ,  $i, j = 1, \dots, n$ ;
2. Za  $i = 1, \dots, n$  izračunati

$$u_{ij} := a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}, \quad j = i, n,$$

$$l_{ji} := \frac{1}{u_{ii}} \left( a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki} \right), \quad j = i+1, \dots, n.$$

```

LU[a_, n_] := Module[{i, j, k},
  L = Table[0, {i, n}, {j, n}]; U = Table[0, {i, n}, {j, n}]; Do[L[[i, i]] = 1, {i, n}];
  Do[
    Do[
      U[[i, j]] = a[[i, j]] - Sum[L[[i, k]] U[[k, j]], {k, i-1}],
      {j, i, n}];
    Do[
      L[[j, i]] = (a[[j, i]] - Sum[L[[j, k]] U[[k, i]],
        {k, i-1}) / U[[i, i]]],
      {j, i+1, n}];
    {i, n}];
  Print["L=", MatrixForm[L], ", U=", MatrixForm[U]]
]

```

**Primjer 3.4.** Treba riješiti sustav linearnih jednadžbi gdje su matrica sustava  $\mathbf{A}$  i vektor slobodnih koeficijenata  $\mathbf{b}$ , zadani s

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 11 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Navedenim Mathematica-programom dobivamo

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 2 \end{bmatrix}$$

Sukcesivnim rješavanjem jednadžbi (3.19) i (3.18), dobivamo rješenje sustava

$$x_1 = -1/3, \quad x_2 = 1/3, \quad x_3 = 0.$$

Rješenje možemo dobiti korištenjem FS i BS modula (nakon što smo aktivirali: LU[3,A]) kako slijedi

```
FS[3, L, b]; BS[3, U, x]; x
```

*Zadatak 3.9.* Izvedite formule navedene u *Algoritmu LU* koristeći definiciju jednakosti dviju matrica. Specijalizirajte algoritam za slučaj trodijagonalne matrice

**Primjer 3.5.** Za kvadratnu regularnu matricu  $\mathbf{A} \in \mathbb{R}^{n \times n}$  vrijedi  $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ . Ako s  $\mathbf{x}_1, \dots, \mathbf{x}_n$  označimo vektore-stupce matrice  $\mathbf{A}^{-1}$ , a s:  $\mathbf{e}_1, \dots, \mathbf{e}_n$  vektore-stupce jedinične matrice  $\mathbf{I}$ , onda matričnu jednadžbu  $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$  možemo zapisati kao  $n$  sustava linearnih jednadžbi

$$\mathbf{A}\mathbf{x}_1 = \mathbf{e}_1, \quad \mathbf{A}\mathbf{x}_2 = \mathbf{e}_2, \quad \dots, \quad \mathbf{A}\mathbf{x}_n = \mathbf{e}_n,$$

Korištenjem LU-dekompozicije matrice  $\mathbf{A}$ , te ranije sagrađenih algoritama FS i BS možemo izračunati inverznu matricu. Primjerice za matricu iz *Primjera 3.4*, str. 64 imamo

```
In[1]:= n=3;
a={{1,4,7}, {2,5,8}, {3,6,11}};
LU[n,a];
inv=Table[0, {i,n}, {j,n}];
Do[
  FS[n,L,IdentityMatrix[n][[i]]];
  BS[n,U,x]; inv[[i]]=x
  ,{i,n}]
Transpose[inv]//MatrixForm
Out[2]//MatrixForm=
```

$$\begin{bmatrix} -\frac{7}{6} & \frac{1}{3} & \frac{1}{2} \\ -\frac{1}{3} & \frac{5}{3} & -1 \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix}$$

### 3.6 Cholesky - dekompozicija

Pretpostavimo da je  $\mathbf{A} \in \mathbb{R}^{n \times n}$  simetrična pozitivno definitna matrica. Tada je Gaussov postupak provediv bez strategije pivotiranja, pri čemu su svi pivot-elementi  $a_{11}, a_{22}^{(2)}, \dots$  pozitivni. Zbog simetričnosti matrice  $\mathbf{A}$ , množeći slijeva elementarnom matricom  $\mathbf{M}_1$  (vidi *t.3.5*, str. 61) i

sdesna matricom  $\mathbf{M}_1^T$ , dobivamo

$$\tilde{\mathbf{A}}^{(2)} = \mathbf{M}_1 \mathbf{A} \mathbf{M}_1^T = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

Ponavljajući postupak kao u t.3.5, str. 61, nakon  $(n - 1)$  koraka dobivamo

$$\mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{A} \mathbf{M}_1^T \mathbf{M}_2^T \cdots \mathbf{M}_{n-1}^T = \text{diag}(a_{11}, a_{22}^{(2)}, \dots, a_{nn}^{(n)}) =: \mathbf{D} \quad (3.20)$$

Kako su svi dijagonalni elementi matrice  $\mathbf{D}$  pozitivni, možemo definirati

$$\sqrt{\mathbf{D}} = \text{diag} \left( \sqrt{a_{11}}, \sqrt{a_{22}^{(2)}}, \dots, \sqrt{a_{nn}^{(n)}} \right),$$

a (3.20) zapisati kao

$$\mathbf{A} = \mathbf{M}_1^{-1} \cdots \mathbf{M}_{n-1}^{-1} \sqrt{\mathbf{D}} \sqrt{\mathbf{D}} \left( \mathbf{M}_{n-1}^T \right)^{-1} \cdots \left( \mathbf{M}_1^T \right)^{-1}.$$

Ako sada definiramo

$$\begin{aligned} \mathbf{L} &:= \mathbf{M}_1^{-1} \cdots \mathbf{M}_{n-1}^{-1} \sqrt{\mathbf{D}}, \\ \mathbf{L}^T &:= \sqrt{\mathbf{D}} \left( \mathbf{M}_{n-1}^T \right)^{-1} \cdots \left( \mathbf{M}_1^T \right)^{-1}, \end{aligned}$$

dobivamo Cholesky-dekompoziciju matrice  $\mathbf{A}$ :

$$\mathbf{A} = \mathbf{L} \mathbf{L}^T, \quad (3.21)$$

gdje je  $\mathbf{L}$  donja trokutasta matrica. Dekompoziciju (3.21) neki puta je zgodnije pisati u obliku  $\mathbf{A} = \mathbf{R}^T \mathbf{R}$ , gdje je  $\mathbf{R} = \mathbf{L}^T$  gornja trokutasta matrica. Ako je Cholesky dekompozicija matrice  $\mathbf{A}$  poznata, onda sustav  $\mathbf{A} \mathbf{x} = \mathbf{b}$ , odnosno sustav  $\mathbf{L} \mathbf{L}^T \mathbf{x} = \mathbf{b}$  rješavamo tako da najprije *Algoritmom FS* riješimo sustav  $\mathbf{L} \mathbf{z} = \mathbf{b}$ . Rješenje polaznog sustava tada je  $\mathbf{x}^* = \mathbf{L} \mathbf{z}$ .

Elemente matrice  $\mathbf{L}$  redom, stupac po stupac, možemo dobiti pomoću sljedećeg algoritma.

**Algorithm 3** (Cholesky)

1. Učitati elemente matrice  $a_{ij}$ ,  $i, j = 1, \dots, n$ ;
2. Za  $k = 1, \dots, n$  izračunati

$$l_{kk} = \left( a_{kk} - \sum_{p=1}^{k-1} l_{kp}^2 \right)^{1/2},$$

$$l_{ik} = \frac{1}{l_{kk}} \left( a_{ik} - \sum_{p=1}^{k-1} l_{ip} l_{kp} \right), \quad i = k+1, \dots, n.$$

(pri čemu navedene sume nećemo izvoditi ako je gornja granica sume manja od donje granice).

Niže je naveden odgovarajući Mathematica-modul.

```
Cholesky[a_, n_] :=
Module[{i, j, k, p}, L=Table[0, {i, n}, {j, n}];
Do[L[[k, k]] =
Sqrt[a[[k, k]] - Sum[L[[k, p]]^2, {p, k-1}]];
Do[
L[[i, k]] = (a[[i, k]] - Sum[L[[i, p]]*L[[k, p]], {p, k-1}))/L[[k, k]]
, {i, k+1, n}
, {k, n}];
Print["L=", MatrixForm[L]]
]
```

**Primjer 3.6.** Cholesky dekompozicijom treba riješiti sustav  $\mathbf{H}\mathbf{x} = \mathbf{b}$ , gdje je  $\mathbf{H}$  Hilbertova matrica 3. reda (str. 60), a  $\mathbf{b} = (1, 1, 1)^T$ .

Algoritmom Cholesky lako dobivamo

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & 0 \\ \frac{1}{3} & \frac{1}{2\sqrt{3}} & \frac{1}{6\sqrt{5}} \end{bmatrix}$$

i rješenje sustava  $\mathbf{x}^* = (3, -24, 30)^T$ .

### 3.7 QR-dekompozicija

QR-dekompozicija, koju bismo željeli sada izložiti, matricu  $\mathbf{A} \in \mathbb{R}^{m \times n}$  rastavlja na produkt  $\mathbf{Q}\mathbf{R}$ , gdje je  $\mathbf{Q}$  ortogonalna, a  $\mathbf{R}$  gornja trokutasta

matrica. Na taj način sustav  $\mathbf{Ax} = \mathbf{b}$  prelazi u trokutasti sustav

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b},$$

koji se lako rješava *Algoritmom BS*. Ono što je ovdje od izuzetnog značaja je činjenica da se pri tome broj uvjetovanosti nije promijenio, tj.

$$\text{cond}(\mathbf{A}) = \text{cond}(\mathbf{R}).$$

I to nije sve! QR-dekompozicija je provediva i u slučaju ako je matrica  $\mathbf{A}$  singularna, pa čak i ako je  $\mathbf{A}$  pravokutna matrica nepunog ranga po retcima ili stupcima. Mi ćemo razmotriti dovoljno općeniti slučaj, koji se javlja kod rješavanja sustava linearnih jednadžbi i kod rješavanja tzv. linearnog problema najmanjih kvadrata (*t.6.1*, str. 151).

Pretpostavimo dakle, da je matrica  $\mathbf{A} \in \mathbb{R}^{m \times n}$  ( $m \geq n$ ) punog ranga po stupcima ( $\text{rang } \mathbf{A} = n$ ), tada postoji ortogonalna matrica<sup>7</sup>  $\mathbf{Q} \in \mathbb{R}^{m \times m}$  i gornja trokutasta matrica  $\mathbf{R}$  s elementima na glavnoj dijagonali koji su različiti od nule, tako da je  $\mathbf{A} = \mathbf{QR}$ . Ovaj rastav nazivamo QR-dekompozicija matrice  $\mathbf{A}$ .

$$\begin{array}{c} n \\ \boxed{\mathbf{A}} \\ m \end{array} = \begin{array}{c} m \\ \boxed{\mathbf{Q}} \\ m \end{array} \begin{array}{c} n \\ \boxed{\mathbf{R}} \\ m \end{array}$$

QR-dekompozicija matrice  $\mathbf{A}$  može se dobiti na više načina. Spomenimo samo Gram-Schmidtov postupak ortogonalizacije, Givensove rotacije i Householderove transformacije (Björck, 1996; Gill et al., 1991; Golub and van Loan, 1996; Lawson and Hanson, 1996; Stewart, 1998). Mi ćemo QR-dekompoziciju graditi primjenom Householderovih transformacija.

<sup>7</sup>Ortogonalna matrica  $\mathbf{Q}$  je takva kvadratna matrica za koju vrijedi  $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ , odnosno

$$\sum_k q_{ik}q_{jk} = \delta_{ij} \quad \& \quad \sum_k q_{ki}q_{kj} = \delta_{ij},$$

tj. njeni retci i njeni stupci čine ortonormirani sustav vektora. Ona je regularna matrica za koju je  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ . Ima važno svojstvo da čuva  $l_2$  normu, tj.

$$\|\mathbf{Q}\mathbf{a}\|^2 = (\mathbf{Q}\mathbf{a})^T(\mathbf{Q}\mathbf{a}) = \mathbf{a}^T\mathbf{Q}^T\mathbf{Q}\mathbf{a} = \mathbf{a}^T\mathbf{a} = \|\mathbf{a}\|^2.$$

Ortogonalna matrica geometrijski predstavlja rotaciju ili simetriju prostora

U tu svrhu najprije ćemo uvesti pojam tzv. Householderove matrice:

$$\mathbf{H} = \mathbf{H}(\mathbf{u}) := \mathbf{I} - 2\mathbf{u}\mathbf{u}^T, \quad \mathbf{u}^T \mathbf{u} = 1, \quad \mathbf{u} \in \mathbb{R}^m, \quad \mathbf{H} \in \mathbb{R}^{m \times m}, \quad (3.22)$$

Vektor  $\mathbf{u}$  iz (3.22) nazivamo Householderov vektor. Householderova matrica je očigledno simetrična i ortogonalna. Posebno je važno sljedeće svojstvo ove matrice:

**Lema 3.1.** *Za dva vektora  $\mathbf{a}$ ,  $\mathbf{b}$  jednake Euklidove norme postoji Householderov vektor  $\mathbf{u}$ , tako da je  $\mathbf{H}\mathbf{a} = \mathbf{b}$ , odnosno, zbog simetričnosti matrice  $\mathbf{H}$ ,  $\mathbf{H}\mathbf{b} = \mathbf{a}$ .*

**Dokaz.** Ako je  $\mathbf{a} = \mathbf{b}$ , onda je  $\mathbf{H} = \mathbf{I}$ . Pretpostavimo zato da su  $\mathbf{a}$ ,  $\mathbf{b}$  dva vektora takva da je

$$\mathbf{a} \neq \mathbf{b} \quad \& \quad \|\mathbf{a}\| = \|\mathbf{b}\| \quad (3.23)$$

Ovo također znači da vektori  $\mathbf{a}, \mathbf{b}$  ne mogu biti nulvektori. Primijetimo također da zbog  $\mathbf{a} \neq \mathbf{b}$  slijedi  $\|\mathbf{a} - \mathbf{b}\| > 0$ . Zato možemo definirati Householderov vektor

$$\mathbf{u} = \frac{1}{\|\mathbf{a} - \mathbf{b}\|}(\mathbf{a} - \mathbf{b}).$$

Kako je

$$\mathbf{a} - \mathbf{H}\mathbf{b} = (\mathbf{a} - \mathbf{b}) + 2 \frac{\mathbf{a} - \mathbf{b}}{\|\mathbf{a} - \mathbf{b}\|} \frac{(\mathbf{a} - \mathbf{b})^T}{\|\mathbf{a} - \mathbf{b}\|} \mathbf{b} = \frac{\mathbf{a} - \mathbf{b}}{\|\mathbf{a} - \mathbf{b}\|^2} (\|\mathbf{a} - \mathbf{b}\|^2 + 2(\mathbf{a} - \mathbf{b})^T \mathbf{b}),$$

onda zbog

$$\|\mathbf{a} - \mathbf{b}\|^2 + 2(\mathbf{a} - \mathbf{b})^T \mathbf{b} = \mathbf{a}^T \mathbf{a} - 2\mathbf{a}^T \mathbf{b} + \mathbf{b}^T \mathbf{b} + 2\mathbf{a}^T \mathbf{b} - 2\mathbf{b}^T \mathbf{b} = \mathbf{a}^T \mathbf{a} - \mathbf{b}^T \mathbf{b},$$

vrijedi  $\mathbf{H}\mathbf{b} = \mathbf{a}$ . □

**Primjer 3.7.** *Vektori  $\mathbf{a} = (\frac{9}{2}, 1)^T$  i  $\mathbf{b} = (3, \frac{7}{2})^T$  imaju svojstvo (3.23). Definirat ćemo Householderov vektor  $\mathbf{u}$ , tako da bude  $\mathbf{H}\mathbf{a} = \mathbf{b}$ .*

Stavimo

$$\mathbf{u} = \frac{(\mathbf{a} - \mathbf{b})}{\|\mathbf{a} - \mathbf{b}\|} = \frac{2}{\sqrt{34}} \left( \frac{3}{2}, -\frac{5}{2} \right)^T.$$

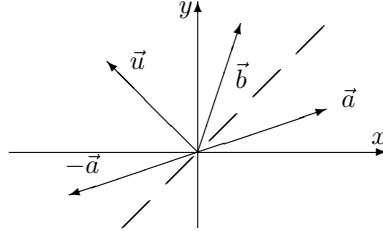
Householderova matrica glasi

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2 \frac{4}{34} \begin{bmatrix} \frac{3}{2} \\ -\frac{5}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{2} & -\frac{5}{2} \end{bmatrix} = \begin{bmatrix} \frac{8}{17} & \frac{15}{17} \\ \frac{15}{17} & -\frac{8}{17} \end{bmatrix}$$

i vrijedi

$$\mathbf{H}\mathbf{a} = \begin{bmatrix} \frac{8}{17} & \frac{15}{17} \\ \frac{15}{17} & -\frac{8}{17} \end{bmatrix} \cdot \begin{bmatrix} \frac{9}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ \frac{7}{2} \end{bmatrix} = \mathbf{b}.$$





Slika 3.2

Na *Slici 3.2* vidi se da su vektori  $a$  i  $b$  osno simetrični na pravac kroz ishodište, koji je okomit na vektor  $u$ . Dakle, Householderova matrica u geometrijskom smislu predstavlja zrcaljenje ili rotaciju prostora.

*Zadatak 3.10.* Zadana je točka  $T(4, 3)$ . Odredite točku  $T'$  na pozitivnom dijelu osi  $x$ , tako da bude  $\|\vec{r}_T\| = \|\vec{r}_{T'}\|$ , gdje su  $\vec{r}_T, \vec{r}_{T'}$  radijvektori točaka  $T, T'$ . Odredite također odgovarajuću Householderovu matricu  $\mathbf{H}$ , koja će vektor  $\vec{r}_T$  preslikavati u vektor  $\vec{r}_{T'}$ .

$$\text{Rješenje: } T'(5, 0), \quad \mathbf{H} = \frac{1}{5} \begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}.$$

*Primjedba 3.6.* Množenje nekog vektora  $\mathbf{c}$  s Householderovom matricom  $\mathbf{H}$  može se jednostavno provesti. Naime, zbog

$$\mathbf{H}\mathbf{c} = \mathbf{c} - 2(\mathbf{u}^T \mathbf{c})\mathbf{u},$$

dovoljno je napraviti skalarni produkt  $\mathbf{u}^T \mathbf{c}$ , množenje vektora sa skalarom  $2(\mathbf{u}^T \mathbf{c})\mathbf{u}$  i oduzimanje vektora  $\mathbf{c} - 2(\mathbf{u}^T \mathbf{c})\mathbf{u}$ .

U cilju svoenja matrice  $\mathbf{A}$  na gornje trokutasti oblik, njene stupce promatrat ćemo kao vektore u  $\mathbb{R}^m$ . Pretpostavimo općenito da je  $\mathbf{a} \in \mathbb{R}^m$ , da je  $k \geq 1$  i da je barem jedna od komponenti  $a_k, \dots, a_m$  vektora  $\mathbf{a}$  različita od nule. Konstruirat ćemo Householderovu matricu  $\mathbf{H}$  tako da se prvih  $(k-1)$  komponenti vektora  $\mathbf{b} := \mathbf{H}\mathbf{a}$  i vektora  $\mathbf{a}$  podudaraju, a da se posljednjih  $(m-k)$  komponenti vektora  $\mathbf{b}$  poništavaju. Dakle, treba biti

$$(\mathbf{H}\mathbf{a})_i = b_i = \begin{cases} a_i, & i = 1, \dots, k-1 \\ \pm\gamma, & i = k \\ 0, & i = k+1, \dots, m \end{cases}, \quad \gamma = \sqrt{a_k^2 + \dots + a_m^2}.$$

Pri tome ćemo predznak ( $\pm\gamma$ ) birati tako da bude suprotan predznaku od  $a_k$ . Primijetimo da je stoga  $b_k = -\text{sgn}(a_k)\gamma$ . Očigledno je  $\|\mathbf{a}\| = \|\mathbf{b}\|$ .

Potražimo oblik vektora  $\mathbf{u}$ . Koristeći jednakost  $|a_k| = a_k \text{sgn}(a_k)$ , dobivamo

$$\begin{aligned}\|\mathbf{a} - \mathbf{b}\|^2 &= (a_k + \text{sgn}(a_k)\gamma)^2 + a_{k+1}^2 + \cdots + a_m^2 \\ &= a_k^2 + 2\gamma a_k \text{sgn}(a_k) + \gamma^2 + a_{k+1}^2 + \cdots + a_m^2 = 2\gamma(\gamma + |a_k|).\end{aligned}$$

Zato je

$$\mathbf{u} = \frac{\mathbf{a}-\mathbf{b}}{\|\mathbf{a}-\mathbf{b}\|} = \frac{1}{\sqrt{2\gamma(\gamma+|a_k|)}} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_k + \text{sgn}(a_k)\gamma \\ a_{k+1} \\ \vdots \\ a_m \end{bmatrix}.$$

**Primjer 3.8.** Treba konstruirati Householderovu matricu  $\mathbf{H}$  tako da poništi drugu komponentu vektora  $\mathbf{a} = (4, 3)^T$ .

U ovom slučaju je  $k = 1$ ,  $\gamma = \sqrt{16+9} = 5$ ,

$$\mathbf{u} = \frac{1}{\sqrt{10(5+4)}} \begin{bmatrix} 4+5 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{90}} \begin{bmatrix} 9 \\ 3 \end{bmatrix}, \quad \mathbf{H} = -\frac{1}{5} \begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}.$$

Dobivamo:  $\mathbf{H}\mathbf{a} = (-5, 0)^T$ .

**Zadatak 3.11.** Konstruirajte Householderovu matricu  $\mathbf{H}$  tako da poništi samo treću komponentu vektora  $\mathbf{a} = (7, 4, 3)^T$ .

Rješenje:

$$k = 2, \quad \gamma = 5, \quad \mathbf{u} = \frac{1}{\sqrt{90}} \begin{bmatrix} 0 \\ 9 \\ 3 \end{bmatrix}, \quad \mathbf{H} = -\frac{1}{5} \begin{bmatrix} 5 & 0 & 0 \\ 0 & -4 & -3 \\ 0 & -3 & 4 \end{bmatrix}, \quad \mathbf{H}\mathbf{a} = \begin{bmatrix} 7 \\ -5 \\ 0 \end{bmatrix}.$$

**Primjer 3.9.** Primjenom Householderovih transformacija, matricu  $\mathbf{A}$  treba transformirati u gornju trokutastu

$$\mathbf{A} = \begin{bmatrix} 10 & 9 & 18 \\ 20 & -15 & -15 \\ 20 & -12 & 51 \end{bmatrix}.$$

72 POGLAVLJE 3. RJEŠAVANJE SUSTAVA LINEARNIH JEDNADŽBI

1. Najprije stavimo  $k = 1$  i  $\mathbf{a} = (10, 20, 20)^T$ . Dobivamo

$$\gamma = 30, \quad \mathbf{u}^{(1)} = \frac{1}{\sqrt{60(30+10)}} \begin{bmatrix} 10+30 \\ 20 \\ 20 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix},$$

$$\mathbf{H}_1 = -\frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{bmatrix}, \quad \mathbf{A}^{(1)} := \mathbf{H}_1 \mathbf{A} = \begin{bmatrix} -30 & 15 & -30 \\ 0 & -12 & -39 \\ 0 & -9 & 27 \end{bmatrix}.$$

2. Stavimo  $k = 2$ ,  $\mathbf{a} = (15, -12, -9)^T$ . Dobivamo

$$\gamma = 15, \quad \mathbf{u}^{(2)} = \frac{1}{\sqrt{30(15+12)}} \begin{bmatrix} 0 \\ -12-15 \\ -9 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 \\ -3 \\ -1 \end{bmatrix},$$

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.8 & -0.6 \\ 0 & -0.6 & 0.8 \end{bmatrix}, \quad \mathbf{A}^{(2)} := \mathbf{H}_2 \mathbf{A}^{(1)} = \begin{bmatrix} -30 & 15 & -30 \\ 0 & 15 & 15 \\ 0 & 0 & 45 \end{bmatrix}.$$

Kako je  $\mathbf{A}^{(2)} = \mathbf{H}_2 \mathbf{A}^{(1)} = \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}$ , dobivamo dekompoziciju  $\mathbf{A} = \mathbf{Q} \mathbf{R}$ , gdje je

$$\mathbf{Q} = (\mathbf{H}_1 \mathbf{H}_2) = \begin{bmatrix} -\frac{1}{3} & \frac{14}{15} & -\frac{2}{15} \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -1 & -\frac{2}{15} & \frac{11}{15} \end{bmatrix}, \quad \mathbf{R} = \mathbf{A}^{(2)}.$$

Slično kao u prethodnom primjeru, sukcesivnom primjenom Householderovih matrica  $\mathbf{H}_k$  na matricu  $\mathbf{A} \in \mathbb{R}^{m \times n}$  možemo dobiti gornju trokutastu matricu  $\mathbf{R}$

$$\mathbf{R} = \mathbf{H}_m \mathbf{H}_{m-1} \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{A},$$

iz čega zbog  $\mathbf{H}_k^{-1} = \mathbf{H}_k^T = \mathbf{H}_k$  slijedi

$$\mathbf{A} = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_m \mathbf{R}.$$

Kako je produkt ortogonalnih matrica opet ortogonalna matrica, dobivamo QR-dekompoziciju matrice  $\mathbf{A}$

$$\mathbf{A} = \mathbf{Q} \mathbf{R}, \quad \text{gdje je } \mathbf{Q} = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_{m-1} \mathbf{H}_m.$$

### 3.8 Iterativne metode

Do sada razmatrane metode za rješavanje sustava  $\mathbf{Ax} = \mathbf{b}$  (Gaussov postupak, LU–dekompozicija, Cholesky–dekompozicija, QR–dekompozicija) spadaju u tzv. direktne metode. Budući da je broj potrebnih računskih operacija kod ovih metoda reda veličine  $n^3$ , za velike matrice s puno elemenata koji iščezavaju (tzv. *large sparse matrices*) ove metode ne mogu se preporučiti. Osim toga, treba ekonomizirati i s brojem elemenata matrice za koje treba rezervirati mjesto u memoriji računala. Takve situacije javljaju se primjerice kod proučavanja električnih mreža, kod velikih ekonometrijskih modela nacionalne privrede, kod spline–interpolacija, kod rješavanja rubnih problema za obične i parcijalne diferencijalne jednačbe, itd. Upravo se iterativne metode, o kojima ćemo reći samo nekoliko osnovnih činjenica, koriste u takvim situacijama.

Spomenut ćemo dvije klasične iterativne metode za rješavanje sustava  $\mathbf{Ax} = \mathbf{b}$ ,  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Pretpostavljamo da je  $a_{ii} \neq 0$ ,  $i = 1, \dots, n$ .

#### • Jacobijeva metoda

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \dots - a_{1,n-1}x_{n-1}^{(k)} - a_{1n}x_n^{(k)}) \\ x_2^{(k+1)} &= \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} \dots - a_{2,n-1}x_{n-1}^{(k)} - a_{2n}x_n^{(k)}) \\ x_3^{(k+1)} &= \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)} \dots - a_{3,n-1}x_{n-1}^{(k)} - a_{3n}x_n^{(k)}) \\ &\vdots \\ x_n^{(k+1)} &= \frac{1}{a_{nn}}(b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - a_{n3}x_3^{(k)} \dots - a_{n,n-1}x_{n-1}^{(k)}) \end{aligned}$$

odnosno:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n \quad (3.24)$$

#### • Gauss-Seidelova metoda

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \dots - a_{1,n-1}x_{n-1}^{(k)} - a_{1n}x_n^{(k)}) \\ x_2^{(k+1)} &= \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} \dots - a_{2,n-1}x_{n-1}^{(k)} - a_{2n}x_n^{(k)}) \\ x_3^{(k+1)} &= \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} \dots - a_{3,n-1}x_{n-1}^{(k)} - a_{3n}x_n^{(k)}) \\ &\vdots \\ x_n^{(k+1)} &= \frac{1}{a_{nn}}(b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - a_{n3}x_3^{(k+1)} \dots - a_{n,n-1}x_{n-1}^{(k+1)}) \end{aligned}$$

odnosno:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n \quad (3.25)$$

**Primjer 3.10.** *Jacobijevom odnosno Gauss-Seidelovom metodom riješit ćemo sustav iz Primjera 2.8, str. 41:*

$$\begin{bmatrix} 1 & 0.25 & 0 \\ 0.25 & 1 & 0.25 \\ 0 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -48 \\ 0 \\ 48 \end{bmatrix}.$$

Jacobijeva metoda daje rješenje  $\mathbf{x}^* = (-48, 0, 48)^T$  već u prvom koraku. Gauss-Seidelova metoda također konvergira rješenju. Prvih nekoliko iteracija prikazano je u *Tablici 3.1*.

$k$	$x_1^k$	$x_2^k$	$x_3^k$
0	0	0	0
1	-48.0	12.0	45.0
2	-51.0	1.5	47.625
3	-48.375	0.1875	47.9531
4	-48.0469	0.02344	47.9941
5	-48.0059	0.00293	47.9993
6	-48.0007	0.00037	47.9999

Tablica 3.1: Gauss-Seidelova metoda

Izračunavanje je obavljeno na bazi niže navedenih Mathematica-modula.

```
Jacobi[a_,b_,n_,it_] := Module[{i,k,xs,xn}, xs=Table[0, {i,n}]; xn=xs;
Do[
Do[
xn[[i]]=(b[[i]]-Apply[Plus, ReplacePart[a[[i]],0,{i}]*xs])/a[[i,i]]
,{i, n}];
Print["It_", k, " = ", N[xn]]; xs=xn
,{k, it}
]
```

```
GS[a_,b_,n_,it_] := Module[{i,k,xs,xn}, xs = Table[0, {i,n}]; xn=xs;
```

```

Do[
  Do[
    xn[[i]]=(b[[i]] - Sum[a[[i,j]] xn[[j]], {j, i-1}]
      - Sum[a[[i,j]]*xs[[j]],{j, i+1,n}]/a[[i,i]],{i,n});
    Print["It_", k, " = ", N[xn]];
    xs = xn
    ,{k, it}]
  ]

```

U cilju ispitivanja konvergencije navedenih, pa i drugih iterativnih metoda, na Banachovom prostoru  $\mathbb{R}^n$  (snabdjevenom s jednom od normi navedenih u *t.3.1*) uvest ćemo metriku  $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty)$ ,

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

čime  $\mathbb{R}^n$  postaje i potpuni metrički prostor.

**Definicija 3.3.** Preslikavanje  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  zovemo **kontrakcija** ako postoji takav realni broj  $q < 1$ , da bude

$$d(F(\mathbf{x}), F(\mathbf{y})) < q d(\mathbf{x}, \mathbf{y}) \quad \text{za sve } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

**Primjer 3.11.** Funkcija  $F : [1, \infty) \rightarrow \mathbb{R}$ ,  $F(x) = \frac{x}{2} + \frac{1}{x}$  je kontrakcija s  $q = 1/2$ . Naime

$$\begin{aligned}
 d(F(x), F(y)) &= |F(x) - F(y)| = \left| \frac{x}{2} + \frac{1}{x} - \frac{y}{2} - \frac{1}{y} \right| = \left| \frac{1}{2}(x - y) - \frac{x-y}{xy} \right| \\
 &= |x - y| \left| \frac{1}{2} - \frac{1}{xy} \right| \leq \frac{1}{2} |x - y| = \frac{1}{2} d(x, y)
 \end{aligned}$$

**Teorem 3.1.** (Banachov teorem o fiksnoj točki) *Neka je funkcija  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  kontrakcija. Tada postoji jedinstvena fiksna točka  $\mathbf{x}^* \in \mathbb{R}^n$  funkcije  $F$ , tj.*

$$F(\mathbf{x}^*) = \mathbf{x}^*.$$

*Osim toga, za proizvoljni  $\mathbf{x}_0 \in \mathbb{R}^n$ , niz definiran rekurzivnom formulom  $\mathbf{x}_{k+1} = F(\mathbf{x}^k)$  konvergira prema  $\mathbf{x}^*$  i vrijedi formula za ocjenu pogreške*

$$d(\mathbf{x}^{k+1}, \mathbf{x}^*) \leq \frac{q}{1-q} d(\mathbf{x}^k, \mathbf{x}^{k+1}) \leq \frac{q^k}{1-q} d(\mathbf{x}_0, F(\mathbf{x}_0)), \quad k = 0, 1, \dots$$

Ako matricu sustava  $\mathbf{A}$  rastavimo na donji trokut, dijagonalu i gornji trokut,

$$\mathbf{A} = \mathbf{D}(\mathbf{L} + \mathbf{I} + \mathbf{U}), \quad \mathbf{D} = \text{diag}(a_{11}, \dots, a_{nn}),$$

dobivamo matrični oblik

- Jacobijeve metode

$$\mathbf{x}^{(k+1)} = -(\mathbf{L} + \mathbf{U})\mathbf{x}^{(k)} + \mathbf{D}^{-1}\mathbf{b}, \quad k = 0, 1, \dots$$

- Gauss-Seidelove metode

$$\mathbf{x}^{(k+1)} = -\mathbf{L}\mathbf{x}^{(k+1)} - \mathbf{U}\mathbf{x}^{(k)} + \mathbf{D}^{-1}\mathbf{b} \quad k = 0, 1, \dots$$

Sljedeća rekurzivna formula obuhvaća obje metode

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c}, \quad k = 0, 1, \dots; \quad (3.26)$$

za  $\mathbf{B} = \mathbf{B}_J := -(\mathbf{L} + \mathbf{U})$  dobivamo Jacobijevu, a za  $\mathbf{B} = \mathbf{B}_{GS} := -(\mathbf{I} + \mathbf{L})^{-1}\mathbf{U}$  dobivamo Gauss-Seidelovu metodu.

Primijetimo da je funkcija  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F(\mathbf{x}) = \mathbf{B}\mathbf{x} + \mathbf{c}$  kontrakcija onda i samo onda ako je  $\|\mathbf{B}\| = q < 1$ . Tada prema Banachovom teoremu o fiksnoj točki niz  $(x^{(n)})$  definiran s (3.26) konvergira prema jedinstvenom rješenju sustava  $\mathbf{A}\mathbf{x} = \mathbf{b}$  i vrijedi ocjena pogreške

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \frac{q}{1-q} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.$$

Kod Jacobijeve metode matrica  $\mathbf{B}_J$  ima elemente  $b_{ij} = \frac{a_{ij}}{a_{ii}}$ ,  $i \neq j$ ,  $b_{ii} = 0$ , pa je

$$\|\mathbf{B}_J\|_\infty = \max_{i=1,\dots,n} \sum_{j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|}.$$

Dakle, Jacobijeva metoda konvergira će ako je  $\mathbf{A}$  strogo dijagonalno dominantna matrica.

Slično se može pokazati da će i Gauss-Seidelova metoda konvergirati ako je  $\mathbf{A}$  strogo dijagonalno dominantna matrica.

*Primjedba 3.7.* Jednostavnom modifikacijom Gauss-Seidelove metode moguće je ubrzati konvergenciju iterativnog procesa. Proces (3.25) možemo zapisati kao

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} + r_i^{(k)}, \quad i = 1, \dots, n, \\ r_i^{(k)} &= \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right). \end{aligned} \quad (3.27)$$

Uvođenjem relaksacijskog parametra  $\omega$ , proces (3.27) možemo modificirati tako da iterativni proces

$$x_i^{(k+1)} = x_i^{(k)} + \omega r_i^{(k)}, \quad i = 1, \dots, n, \quad (3.28)$$

koji se u literaturi nalazi pod imenom Successive overrelaxation (SOR) method, brže konvergira. Može se pokazati da odgovarajuća matrica  $\mathbf{B}$  iz (3.26) u ovom slučaju glasi

$$\mathbf{B} := \mathbf{B}_\omega = (\mathbf{I} + \omega \mathbf{L})^{-1} ((\mathbf{I} - \omega \mathbf{I}) - \omega \mathbf{U}),$$

i da odgovarajući iterativni proces konvergira za  $\omega \in (0, 2)$ .

### 3.9 Dekompozicija na singularne vrijednosti

Svaku matricu  $\mathbf{A} \in \mathbb{R}^{m \times n}$  možemo zapisati kao (Björck, 1996; Gill et al., 1991; Golub and van Loan, 1996; Lawson and Hanson, 1996; Stewart, 1998)

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (3.29)$$

gdje su  $\mathbf{U} \in \mathbb{R}^{m \times m}$  i  $\mathbf{V} \in \mathbb{R}^{n \times n}$  ortogonalne matrice, a  $\mathbf{S} \in \mathbb{R}^{m \times n}$  dijagonalna matrica

$$\mathbf{S} = \text{diag}(\sigma_1, \dots, \sigma_p), \quad p = \min(m, n), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0. \quad (3.30)$$



Nenegativne brojeve  $\sigma_1, \dots, \sigma_p$  zovemo **singularne vrijednosti** matrice  $\mathbf{A}$ , a rastav (3.29) rastav na singularne vrijednosti (*Singular Value Decomposition*).

Kako je  $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{S}^T \mathbf{S} \mathbf{V}^T$ , kvadrati singularnih vrijednosti  $\sigma_i^2$  su svojstvene vrijednosti<sup>8</sup> simetrične pozitivno semidefinitne kvadratne matrice  $\mathbf{A}^T \mathbf{A}$ , a stupci matrice  $\mathbf{V}$  su odgovarajući ortonormirani svojstveni vektori. Analogno, kako je  $\mathbf{A} \mathbf{A}^T = \mathbf{U} \mathbf{S} \mathbf{S}^T \mathbf{U}^T$ , kvadrati singularnih vrijednosti  $\sigma_i^2$  također su i svojstvene vrijednosti matrice  $\mathbf{A} \mathbf{A}^T$ , a stupci matrice  $\mathbf{U}$  su odgovarajući ortonormirani svojstveni vektori.

Ako s  $\mathbf{u}_i \in \mathbb{R}^m$  i  $\mathbf{v}_i \in \mathbb{R}^n$  označimo stupce matrice  $\mathbf{U}$ , odnosno  $\mathbf{V}$ , i ako je rang  $\mathbf{A} = k \leq p$ , tada iz (3.29) slijedi

$$\mathbf{A} \mathbf{v}_i = \sigma_i \mathbf{u}_i \quad \& \quad \mathbf{A}^T \mathbf{u}_i = \sigma_i \mathbf{v}_i, \quad i = 1, \dots, p. \quad (3.31)$$

Vektore  $\mathbf{u}_i$  zovemo **lijevi**, a vektore  $\mathbf{v}_i$  **desni singularni vektori** matrice  $\mathbf{A}$ .

*Primjedba 3.8.* Singularne vrijednosti matrice  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mogu se interpretirati i kao svojstvene vrijednosti matrice  $(\mathbf{A}^T \mathbf{A})^{1/2}$

Nadalje, iz (3.29)-(3.30) lako dobivamo  $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{S}^T \mathbf{S} \mathbf{V}^T$ . Zato SVD matrice  $\mathbf{A} \in \mathbb{R}^{m \times n}$  možemo izračunati tako da najprije načinimo spektralnu dekompoziciju  $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T$  i definiramo dijagonalnu matricu  $\mathbf{S} \in \mathbb{R}^{m \times n}$ , čiji je gornji ( $n \times n$ ) blok matrica  $+\sqrt{\mathbf{D}}$ . Nakon toga matricu  $\mathbf{U}$  možemo dobiti iz jednadžbe  $\mathbf{U} \mathbf{S} = \mathbf{A} \mathbf{V}$ . Nažalost, opisani postupak nije numerički stabilan. Numerički stabilni SVD algoritmi obično se provode u dvije faze: najprije se provede ortogonalna bidijagonalizacija matrice  $\mathbf{A}$ , a nakon toga se nad dobivenom bidijagonalnom matricom provede prilagođeni SVD postupak (vidi primjerice Golub and van Loan (1996); Trefethen and Bau (1997)).

<sup>8</sup>Kažemo da je  $\lambda \in \mathbb{C}$  svojstvena vrijednost kvadratne matrice  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ako postoji  $\mathbf{x} \neq \mathbf{0}$  iz  $\mathbb{R}^n$  tako da bude  $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$ . Spomenuti vektor  $\mathbf{x}$  nazivamo svojstveni vektor matrice  $\mathbf{A}$ . Svaka kvadratna matrica  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ima  $n$  (ne nužno različitih) svojstvenih vrijednosti  $\lambda_1, \dots, \lambda_n$ . Skup  $\sigma(\mathbf{A}) = \{\lambda_1, \dots, \lambda_n\}$  naziva se spektralni radijus matrice  $\mathbf{A}$ . Ako je  $\mathbf{A}$  još i simetrična, onda su sve svojstvene vrijednosti  $\lambda_i$  realne. U tom slučaju može se definirati ortogonalna matrica  $\mathbf{V} \in \mathbb{R}^{n \times n}$  čiji su stupci ortonormirani svojstveni vektori, tako da je  $\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T$ , gdje je  $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Rastav  $\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T$  nazivamo spektralna dekompozicija matrice  $\mathbf{A}$ . Spomenimo još da vrijedi (vidi primjerice Gill et al. (1991)):

– matrica  $\mathbf{A} \in \mathbb{R}^{n \times n}$  je pozitivno definitna onda i samo onda ako je  $\sigma(\mathbf{A}) \subset \mathbb{R}_+$ ;

$$– \text{Tr}(\mathbf{A}) := \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i;$$

$$– \prod_{i=1}^n \lambda_i = \det \mathbf{A}.$$

**Primjer 3.12.** (Björck, 1996) Uzmimo primjerice da matrica  $\mathbf{A}$  ima dva stupca  $\mathbf{a}_1, \mathbf{a}_2$  koji su ujedno jedinični vektori, a kut između njih je  $\gamma$ , tj.  $\mathbf{a}_1^T \mathbf{a}_2 = \cos \gamma$ .

Tada je

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 1 & \cos \gamma \\ \cos \gamma & 1 \end{bmatrix},$$

sa svojstvenim vrijednostima  $\lambda_1 = 2 \cos^2 \frac{\gamma}{2}$ ,  $\lambda_2 = 2 \sin^2 \frac{\gamma}{2}$ . Singularne vrijednosti matrice  $\mathbf{A}$  su

$$\sigma_1 = \sqrt{2} \cos \frac{\gamma}{2}, \quad \sigma_2 = \sqrt{2} \sin \frac{\gamma}{2}.$$

Svojstveni vektori matrice  $\mathbf{A}^T \mathbf{A}$

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

ujedno su desni singularni vektori matrice  $\mathbf{A}$ .

S numeričkog aspekta, ako je  $\gamma^2$  manji od najmanjeg floating-point broja računala, onda bi zbog  $\cos \gamma \approx 1 - \frac{\gamma^2}{2} = 1$ , matrica  $\mathbf{A}^T \mathbf{A}$  imala samo jednu svojstvenu vrijednost  $\lambda_1 = 2$  i samo jednu singularnu vrijednost  $\sigma_1 = \sqrt{2}$  različitu od nule. Manja singularna vrijednost bi se izgubila.

*Zadatak 3.12.* Izračunaj lijeve singularne vektore  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  iz prethodnog primjera. Formiraj matrice  $\mathbf{U}$  i  $\mathbf{V}$  i provjeri jednakost (3.29).

Primjenom SVD možemo riješiti sustav linearnih jednadžbi  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , gdje je  $\mathbf{A} \in \mathbb{R}^{n \times n}$  kvadratna regularna matrica, koja može biti i loše uvjetovana. Korištenjem rastava (3.29) polazni sustav možemo pisati

$$\mathbf{U}\mathbf{S}\mathbf{V}^T \mathbf{x} = \mathbf{b} \implies \mathbf{S}\mathbf{V}^T \mathbf{x} = \mathbf{U}^T \mathbf{b},$$

na osnovi čega, supstitucijom  $\mathbf{z} = \mathbf{V}^T \mathbf{x}$ , polazni sustav transformiramo u dijagonalni

$$\mathbf{S}\mathbf{z} = \mathbf{U}^T \mathbf{b}.$$

SVD–dekompoziciju matrice  $\mathbf{A}$  možemo dobiti korištenjem `Mathematica`- naredbe `SingularValues[A]`.

*Primjedba 3.9.* Broj uvjetovanosti proizvoljne matrice  $\mathbf{A} \in \mathbb{R}^{m \times n}$  ranga  $k < p = \min\{m, n\}$  (Gill et al., 1991; Golub and van Loan, 1996) može se definirati na sljedeći način:

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^\dagger\|_2 = \frac{\sigma_1}{\sigma_k}, \quad (3.32)$$

gdje su  $\sigma_1 > \sigma_2 > \dots > \sigma_k > 0$  singularne vrijednosti matrice  $\mathbf{A}$ , a  $\mathbf{A}^+ = \mathbf{V}\mathbf{S}^+\mathbf{U}^T$  pseudoinverzna matrica od  $\mathbf{A}$ , pri čemu je  $\mathbf{S}^+ = \text{diag}(s_1, \dots, s_p)$ , gdje je

$$s_i = \begin{cases} \frac{1}{\sigma_i}, & \sigma_i \neq 0 \\ 0, & \sigma_i = 0 \end{cases} \quad (3.33)$$

*Primjedba 3.10.* Više o raznim numeričkim metodama za rješavanje sustava linearnih jednadžbi može se naći kod (Björck, 1996; Dahlquist and Björck, 2008; Gill et al., 1991; Golub and van Loan, 1996; Schwarz and Köckler, 2011), a odgovarajuća FORTRAN-programaska podrška može se naći kod `?`, te posebno u The NAG-Library. Mogu se također koristiti i gotovi `Mathematica` ili `Matlab` software.

### 3.10 Zadaci

*Zadatak 3.13.* Kpliko je računskih operacija potrebno da se za zadanu kvadratnu matricu  $\mathbf{A}$  reda  $n$  provede

- a) LU-dekompozicija (Gaussov postupak bez pivotiranja);  
 b) Cholesky-dekompozicija;      c) QR-dekompozicija.

*Zadatak 3.14.* Nekom od navedenih metoda riješite sustave linearnih jednadžbi, te izračunajte  $\text{cond}(\mathbf{A})$  za navedene sustave.

$$\begin{aligned} a) \quad & x_1 + 3x_2 + 9x_3 + x_4 = 10 \\ & x_1 + x_2 + x_3 + 5x_4 = 6 \\ & 2x_1 + 7x_2 + 3x_3 + 2x_4 = 15 \\ & 4x_1 + 2x_2 + x_3 + x_4 = 9 \end{aligned}$$

$$\begin{aligned} b) \quad & 100x_1 - 24x_2 + 48x_3 - 23x_4 = 39 \\ & 5x_1 + 100x_2 - 44x_3 - 31x_4 = 72 \\ & 10x_1 - 3x_2 + 100x_3 + 55x_4 = 56 \\ & -12x_1 + 7x_2 - 11x_3 + 100x_4 = 47 \end{aligned}$$

Rješenje: a)  $\mathbf{x} = \frac{1}{7}(9, 10, 3, 4)^T$ ,  $\text{cond}(\mathbf{A}) = 2.68$ .

b)  $\mathbf{x} = (0.615488, 0.955458, 0.249673, 0.504441)^T$ ,  $\text{cond}(\mathbf{A}) = 1.075$ .

*Zadatak 3.15.* U 4-znamenkastoj, odnosno 6-znamenkastoj floating-point aritmetici uz primjenu neke od navedenih metoda riješite sustav

linearnih jednadžbi  $\mathbf{K}\mathbf{x} = \mathbf{b}$ ,

$$\mathbf{K} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

*Zadatak 3.16.* Rješenje sustava linearnih jednadžbi, gdje su matrica sustava  $\mathbf{A}$  i vektor  $\mathbf{b}$  zadani s

$$\mathbf{A} = \begin{bmatrix} 0.550 & 0.423 \\ 0.484 & 0.372 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0.127 \\ 0.112 \end{bmatrix},$$

je  $\mathbf{x} = (1, -1)^T$ . Promijenimo vektor  $\mathbf{b}$  za  $\Delta\mathbf{b} = (0.00007, 0.00028)^T$ . Odredite egzaktno rješenje sustava  $\mathbf{A}\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$  i izračunajte relativne pogreške vektora  $\mathbf{b}$  i  $\mathbf{x}$ . Koliko puta je relativna promjena u rješenju veća od relativne promjene u vektoru  $\mathbf{b}$ ?

Rješenje:  $\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{0.00028}{0.127} \approx 2.2 \times 10^{-3}$ ,  $\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{0.91}{1} = 0.91$ . Relativna promjena rješenja  $\mathbf{x}$ , 413 puta je veća od relativne promjene u vektoru  $\mathbf{b}$ .

*Zadatak 3.17.* Načinimo malu promjenu u elementu  $a_{21}$  matrice  $\mathbf{A}$  iz *Zadatka 3.16*, tako da je

$$\mathbf{A} + \Delta\mathbf{A} = \begin{bmatrix} 0.550 & 0.423 \\ 0.483 & 0.372 \end{bmatrix}.$$

Odredite egzaktno rješenje sustava  $(\mathbf{A} + \Delta\mathbf{A})\tilde{\mathbf{x}} = \mathbf{b}$ . Koliko puta je relativna promjena u rješenju  $\tilde{\mathbf{x}}$  veća od relativne promjene u matrici? Koliki je  $\text{cond}(\mathbf{A})$  u  $\|\cdot\|_\infty$ ?

Rješenje: Relativna promjena u rješenju je oko 1000 puta veća od relativne promjene u matrici  $\mathbf{A}$ .  $\text{cond}(\mathbf{A}) = 1039.52$

*Zadatak 3.18.* Izradite program koji će učitati red matrice  $n$ , elemente matrice  $\mathbf{A}$ , te primjenom Gaussove metode eliminacije izračunati inverznu matricu  $\mathbf{A}^{-1}$ .

*Zadatak 3.19.* Izradite program koji će učitati red matrice  $n$ , elemente matrice  $\mathbf{A}$ , te primjenom LU-dekompozicije izračunati inverznu matricu  $\mathbf{A}^{-1}$ .

*Zadatak 3.20.* Izradite program koji će Gaussovom metodom uz strategiju potpunog pivotiranja riješiti sustav  $\mathbf{Ax} = \mathbf{b}$ , te niže navedenim postupkom iterativno popraviti rješenje:

Neka su  $\bar{\mathbf{L}}$  i  $\bar{\mathbf{U}}$  neke aproksimacije matrica  $\mathbf{L}$  i  $\mathbf{U}$ , a  $\bar{\mathbf{x}}$  odgovarajuća aproksimacija rješenja sustava  $\mathbf{Ax} = \mathbf{b}$ . Označimo  $\mathbf{x}^{(1)} = \bar{\mathbf{x}}$  i računamo vektor-rezidual

$$\mathbf{r}^{(1)} = \mathbf{b} - \mathbf{Ax}^{(1)},$$

riješimo sustav

$$\bar{\mathbf{L}}\bar{\mathbf{U}}\delta\mathbf{x}^{(1)} = \mathbf{r}^{(1)},$$

i izračunamo sljedeću aproksimaciju rješenja sustava  $\mathbf{Ax} = \mathbf{b}$  na sljedeći način

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \delta\mathbf{x}^{(1)}.$$

Navedeni postupak možemo ponoviti nekoliko puta, a preporučljivo ga je izvesti u double-precision.

*Zadatak 3.21.* Pokažite da je  $\mathbf{x}^{(2)}$  iz prethodnog zadatka točno rješenje sustava  $\mathbf{Ax} = \mathbf{b}$ , ako je  $\bar{\mathbf{L}}\bar{\mathbf{U}} = \mathbf{A}$  i ako su  $\mathbf{r}^{(1)}$  i  $\delta\mathbf{x}^{(1)}$  točno izračunati.

*Zadatak 3.22.* Izradite program koji će primjenom LU-dekompozicije riješiti sustav  $\mathbf{Ax} = \mathbf{b}$ , te na osnovi iterativnog postupka navedenog u *Zadatku 3.20* iterativno popraviti rješenje.

*Zadatak 3.23.* Neka je  $\mathbf{A} = \mathbf{LU}$ , gdje su  $\mathbf{L}$  i  $\mathbf{U}$  matrice definirane Algoritmom LU. Tada je  $\mathbf{A}^{-1} = (\mathbf{LU})^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$ . Uz oznake:  $\mathbf{Y} = \mathbf{L}^{-1}$  i  $\mathbf{Z} = \mathbf{U}^{-1}$  vrijedi:  $\mathbf{LY} = \mathbf{UZ} = \mathbf{I}$ , odnosno

$$\mathbf{Ly}^j = \mathbf{Uz}^j = \mathbf{e}^j, \quad j = 1, \dots, n.$$

Komponente vektora  $y^j$  možemo dobiti iz Algoritma FS, a komponente vektora  $z^j$  iz Algoritma BS. Pokažite da je tada

$$y_{ij} = 0, \quad i < j, \quad z_{ij} = 0, \quad i > j,$$

tj. da je  $\mathbf{L}^{-1}$  donja trokutasta, a  $\mathbf{U}^{-1}$  gornja trokutasta matrica.

Pokažite također da je

$$y_{ij} = \frac{1}{a_{ii}}(\delta_{ij} - \sum_{k=j}^{i-1} a_{ik}y_{kj}), \quad i = j, \dots, n.$$

$$z_{ij} = \frac{1}{u_{ii}}(\delta_{ij} - \sum_{k=i+1}^j u_{ik}z_{kj}), \quad i = j, j-1, \dots, 1.$$

( $\delta_{ij}$  su elementi jedinične matrice), a da elemente inverzne matrice dobivamo iz

$$a_{ij}^{-1} = \sum_{k=m}^n z_{ik}y_{kj} \quad m = \max(i, j).$$

Izradite odgovarajući program za izračunavanje inverzne matrice.

*Zadatak 3.24.* Treba riješiti sustav linearnih jednadžbi  $\mathbf{Ax} = \mathbf{y}$ , gdje je  $\mathbf{A}$  tzv. vrpčasta matrica.

$$\begin{bmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & c_{n-1} & a_n \end{bmatrix}$$

Matrica iz Zadatka 3.24

$$\begin{bmatrix} \times & 0 & 0 & \times & 0 & \times & 0 & \cdots \\ 0 & \times & 0 & 0 & \times & 0 & \times & \cdots \\ 0 & 0 & \times & 0 & 0 & \times & 0 & \cdots \\ \times & 0 & 0 & \times & 0 & 0 & \times & \cdots \\ 0 & \times & 0 & 0 & \times & 0 & 0 & \cdots \\ \times & 0 & \times & 0 & 0 & \times & 0 & \cdots \\ 0 & \times & 0 & \times & 0 & 0 & \times & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Matrica iz Zadatka 3.25

Označimo vektore

$$\mathbf{a} = (a_1, a_2, \dots, a_n)^T, \quad \mathbf{b} = (b_1, b_2, \dots, b_{n-1})^T, \quad \mathbf{c} = (c_1, c_2, \dots, c_{n-1})^T.$$

Izradite program koji će učitati red matrice  $n$ , vektore  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ , te vektor slobodnih koeficijenata  $\mathbf{y}$  i riješiti sustav primjenom a) *Gaussove metode*, c) *Cholesky dekompozicije*, d) *QR-dekompozicije*, e) *Jacobijeve metode*, f) *Gauss-Seidelove metode*.

*Uputa:* Program testirajte za slučaj:  $a_i = i$ ,  $i = 1, \dots, n$ ,  $b_i = i - 1$ ,  $i = 1, \dots, n - 1$ ,  $c_i = i$ ,  $i = 1, \dots, n - 1$ ,  $y_i = (-1)^i$ ,  $i = 1, \dots, n$ . Ako je  $n = 5$ , dobiva se  $\mathbf{x} = (-1, \frac{8}{3}, -\frac{13}{3}, \frac{14}{3}, -3)^T$ .

*Zadatak 3.25.* Treba riješiti sustav linearnih jednadžbi  $\mathbf{Ax} = \mathbf{y}$ , gdje je  $\mathbf{A}$  tzv. vrpčasta matrica, koja ima elementa različitih od nule na glavnoj dijagonali, te na četvrtoj i šestoj sporednoj dijagonali u gornjem i donjem trokutu. Označimo vektore

$$\begin{aligned} \mathbf{a} &= (a_1, a_2, \dots, a_n)^T && \text{— elementi glavne dijagonale,} \\ \mathbf{b} &= (b_1, b_2, \dots, b_{n-4})^T && \text{— elementi 4. sporedne dij. u gornjem trokutu} \\ \mathbf{c} &= (c_1, c_2, \dots, c_{n-6})^T && \text{— elementi 6. sporedne dij. u gornjem trokutu} \\ \mathbf{d} &= (d_1, d_2, \dots, d_{n-4})^T && \text{— elementi 4. sporedne dij. u donjem trokutu} \\ \mathbf{e} &= (e_1, e_2, \dots, e_{n-4})^T && \text{— elementi 6. sporedne dij. u donjem trokutu} \end{aligned}$$

Izradite kompjuterski program koji će učitati red matrice  $n$ , vektore  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ ,  $\mathbf{d}$ ,  $\mathbf{e}$ , te vektor slobodnih koeficijenata  $\mathbf{y}$  i riješiti sustav nekom od ranije spomenutih metoda.

*Zadatak 3.26.* Hilbertova matrica  $\mathbf{H}$  s elementima  $h_{ij} = \frac{1}{i+j-1}$  je klasični primjer loše uvjetovane matrice. Ispitajte neke komercijalne, kao i priložene programe, rješavajući sustav jednadžbi, tako da za matricu sustava izaberete Hilbertove matrice različitih dimenzija. Pri tome izračunavanja obavljajte u  $k$ -znamenkastoj ( $k = 2, \dots, 10$ ) floating-point aritmetici.

*Zadatak 3.27.* Pokažite da je

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & -1 \\ -2 & -4 & 5 \\ 1 & 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ -25 & -5 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 3 & -1 \\ 0 & -2.5 & 4.5 \\ 0 & 0 & 8.5 \end{bmatrix}.$$

Primjenom LU-dekompozicije riješite sustav  $\mathbf{Ax} = \mathbf{b}$ , gdje je  $\mathbf{b} = (-2, 20, 7)^T$ .

Rješenje:  $\mathbf{x} = (3, -4, 2)^T$ .

*Zadatak 3.28.* Zadan je sustav linearnih jednadžbi

$$\begin{bmatrix} 34 & 55 \\ 55 & 89 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 21 \\ 34 \end{bmatrix}$$

Koja je aproksimacija rješenja:  $\mathbf{x}_A^* = (-0.11, 0.45)^T$  ili  $\mathbf{x}_B^* = (-0.99, 1.01)^T$  bolja u  $\mathbf{l}_1$ ,  $\mathbf{l}_2$  i  $\mathbf{l}_\infty$  normi? Kako se ovo može primijeniti za komparativnu analizu kvalitete računala?

*Zadatak 3.29.* Cholesky dekompozicija također se može provesti i po retcima sljedećim algoritmom

Za  $k = 1, \dots, n$

$$l_{kj} = \frac{1}{l_{jj}} \left( a_{kj} - \sum_{p=1}^{j-1} l_{kp} l_{jp} \right), \quad l_{kk} = \left( a_{kk} - \sum_{p=1}^{k-1} l_{kp}^2 \right)^{1/2}, \quad j = 1, \dots, k-1.$$

(pri čemu navedene sume nećemo izvoditi ako je gornja granica sume manja od donje granice). Izradite odgovarajući *Mathematica*-program i ispitajte ga na prethodnim primjerima.

## Poglavlje 4

# Rješavanje nelinearnih jednadžbi

Promatramo realnu neprekidnu funkciju  $f$  definiranu na zatvorenom intervalu  $[a, b]$ . Općenito, svaki kompleksni broj  $\xi$ , koji je rješenje jednadžbe

$$f(x) = 0, \quad (4.1)$$

nazivamo nultočkom funkcije  $f$ . Mi ćemo se ograničiti na istraživanje samo realnih nultočaka funkcije  $f$ , koje geometrijski odgovaraju sjecištima grafa funkcije  $f$  s apscisom. Općenito se može dogoditi da neka funkcija ima više realnih nultočaka, da su neke višestruke ili da uopće nema realnih nultočaka.

Ako je funkcija  $f$  neprekidna na intervalu  $I = [a, b]$  i ako na rubovima intervala prima suprotne vrijednosti (tj. ako je  $f(a) \cdot f(b) < 0$ ), onda (vidi JUKIĆ (1999), str. 118), postoji barem jedna točka  $\xi \in I$ , za koju vrijedi  $f(\xi) = 0$ . Ako je, osim toga, prva derivacija  $f'$  stalnog predznaka na intervalu  $I$ , onda je to i jedina nultočka funkcije  $f$  na intervalu  $I$ . Na taj se način posao oko traženja realnog rješenja jednadžbe (4.1) svodi na dva koraka:

1. Separirati interval  $I$ , u kome funkcija ima nultočku  $\xi$ ;
2. Nekom iterativnom metodom odrediti aproksimaciju nultočke  $\xi$  s unaprijed zadanom točnošću.

U intervalu  $I = [a, b]$  nalazi se barem jedna nultočka funkcije  $f$  ako na njegovim rubovima funkcija prima vrijednosti suprotnog predznaka,



tj. ako vrijedi

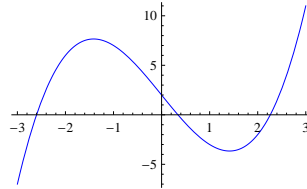
$$f(a) \cdot f(b) < 0. \quad (4.2)$$

**Primjer 4.1.** *Treba separirati realne nultočke funkcije  $f(x) = x^3 - 6x + 2$ .*

Graf ove funkcije lako je nacrtati (vidi Sliku<sup>reffig:PrNelin-1</sup>). Sada je dovoljno samo procijeniti predznak vrijednosti funkcije u nekim točkama.

$x$	-3	-1.5	0	1.5	3
$\text{sgn}f(x)$	-1	+1	+1	-1	+1

Na taj način separirali smo tri intervala:  $I_1 = [-3, -1.5]$ ,  $I_2 = [0, 1.5]$ ,  $I_3 = [1.5, 3]$ . U svakom od njih nalazi se po jedna nultočka funkcije. Budući da se u ovom slučaju radi o polinomu trećeg stupnja, ova funkcija nema drugih nultočaka.



Slika 4.1:  $f(x) = x^3 - 6x + 2$

Prilikom separiranja nultočaka funkcije možemo se koristiti i grafičkim metodama. Pri tome je naročito pogodno koristiti grafičke mogućnosti računala (programi: `Mathematica`, `Matlab`). Ako jednadžbu  $f(x) = 0$  preuredimo tako, da na lijevu i desnu stranu dođu funkcije čije grafove možemo lako nacrtati

$$\varphi(x) = \psi(x),$$

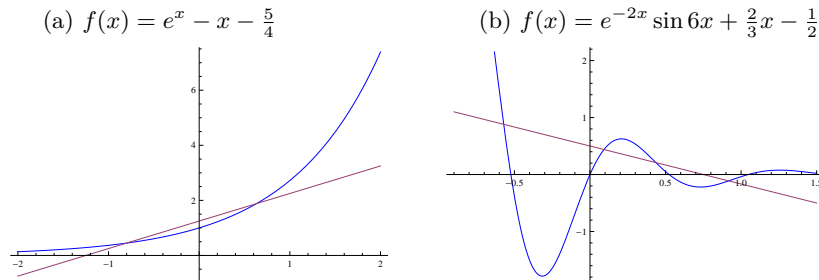
tada apscise točaka sjecišta grafova funkcija  $\varphi$  i  $\psi$  predstavljaju nultočke funkcije  $f$ . Na ovaj način nećemo dobiti dovoljno dobre aproksimacije nultočaka, ali tako smo u mogućnosti izolirati intervale u kojima leže nultočke.

**Primjer 4.2.** *Treba separirati realne nultočke funkcije*

$$f(x) = e^x - x - 5/4.$$

Jednadžbu  $f(x) = 0$  napisat ćemo u obliku:  $e^x = x + 5/4$ , te nacrtati grafove funkcija  $\varphi(x) = e^x$  i  $\psi(x) = x + 5/4$  (Slika 4.2.a). Iz slike se vidi da funkcija  $f$  ima dvije nultočke, koje se nalaze u intervalima:  $I_1 = [-1, 0]$ ,  $I_2 = [0, 1]$ , što potvrđuje i sljedeća tablica

$x$	-1	0	1
$\text{sgn} f(x)$	+1	-1	+1

Slika 4.2: Separacija nultočaka iz *Primjera 4.2* i *Primjera 4.3*

**Primjer 4.3.** *Treba separirati sve pozitivne nultočke funkcije*

$$f(x) = e^{-2x} \sin 6x + \frac{2}{3}x - \frac{1}{2}.$$

Jednadžbu  $f(x) = 0$  napisat ćemo u obliku:  $e^{-2x} \sin 6x = \frac{1}{2} - \frac{2}{3}x$ , te nacrtati grafove funkcija  $\varphi(x) = e^{-2x} \sin 6x$  i  $\psi(x) = \frac{1}{2} - \frac{2}{3}x$  (Slika 4.2.b). Iz slike se vidi da funkcija  $f$  ima tri pozitivne nultočke, koje treba tražiti u intervalima:  $I_1 = [0, 0.25]$ ,  $I_2 = [0.25, 0.5]$ ,  $I_3 = [0.5, 1]$ , što potvrđuje i slijedeća tablica

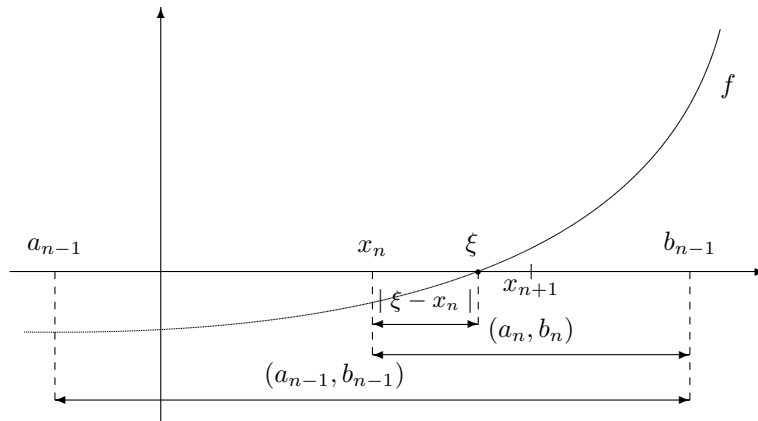
$x$	0	0.25	0.5	1
$\text{sgn} f(x)$	-1	+1	-1	+1

Za rješavanje jednadžbe  $f(x) = 0$  postoji čitav niz različitih metoda, a mi ćemo se upoznati s nekim od njih, koje se najčešće koriste u primjenama. Metode za rješavanje jednadžbe  $f(x) = 0$  uglavnom su iterativne metode u kojima se nekom rekurzivnom formulom definira niz brojeva (aproksimacija):  $x_0, x_1, \dots, x_n \dots$ , koji uz neke uvjete može konvergirati rješenju jednadžbe.

Razmotrit ćemo najprije jednostavnu metodu bisekcije (Ackleh et al., 2010; Dahlquist and Björck, 2008; Stoer and Bulirsch, 2002) za koju je dovoljno da funkcija  $f$  bude neprekidna i da na rubovima intervala  $[a, b]$  prima vrijednosti suprotnog predznaka.

## 4.1 Metoda bisekcije

Pretpostavimo da je funkcija  $f$  neprekidna na intervalu  $I = [a, b]$  i da je  $f(a) \cdot f(b) < 0$ . Ako interval  $I$  raspolovimo i ako polovište nije nultočka funkcije  $f$ , dalje promatramo onu polovinu intervala  $I$  na čijim rubovima funkcija  $f$  prima vrijednosti suprotnog predznaka. Navedeni postupak dalje ponavljamo.



Slika 4.3: Metoda bisekcije

Formalno, označimo  $a_0 := a$ ,  $b_0 := b$  i navedenim postupkom definirajmo silazni niz segmenata:  $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \dots$ , koji sadrži korijen jednadžbe  $f(x) = 0$ . Na taj način u mogućnosti smo definirati niz aproksimacija  $x_1, \dots, x_n, \dots$

$$x_1 = \frac{1}{2}(a + b), \quad |\xi - x_1| \leq \frac{1}{2}(b - a),$$

$$x_2 = \begin{cases} \frac{1}{2}(a + x_1), & f(a) \cdot f(x_1) < 0 \\ \frac{1}{2}(x_1 + b), & f(a) \cdot f(x_1) > 0 \end{cases}, \quad |\xi - x_2| \leq \frac{1}{2}|\xi - x_1| \leq \frac{1}{2^2}(b - a),$$

...

Općenito (vidi *Sliku 4.3*) niz čiji je opći član

$$x_n = \frac{1}{2}(a_{n-1} + b_{n-1}),$$

konvergira prema rješenju jednadžbe  $f(x) = 0$  linearnom brzinom jer je

$$|\xi - x_{n+1}| \leq \frac{1}{2}|\xi - x_n|,$$

pri čemu pogrešku aproksimacije  $x_n$  možemo ocijeniti (vidi *Sliku 4.4*) s

$$|\xi - x_n| \leq \frac{1}{2^n}(b - a).$$

**Algorithm 4 (BISEKCIJA)**

1. Učitati  $a, b, \epsilon > 0$ ;  
Ako je  $f(a) \cdot f(b) > 0$ , zaustaviti proces; u protivnom prijeći na sljedeći korak;
2. Staviti  $a_0 := a$ ;  $b_0 := b$  i izračunati  $x_1 = \frac{1}{2}(a_0 + b_0)$ ;
3. Ako je  $f(a_0) \cdot f(x_1) < 0$ , staviti  $b_0 := x_1$ ; u protivnom staviti  $a_0 = x_1$ ;
4. Ako je  $b_0 - a_0 > 2\epsilon$ , izračunati  $x_1 = \frac{1}{2}(a_0 + b_0)$  i prijeći na Korak 3.; U protivnom zaustaviti proces;

Niže je naveden Mathematica-program za algoritam BISEKCIJE.

```
In[1]:= f[x_] := x^3 - 6 x + 2;
a = 0; b = 1.5; eps = .01; k = 0;
If[f[a] f[b] < 0,
  xL = a; xD = b; xP = (xL + xD)/2; gr = (xD - xL)/2;
  Print["k=",k," xL=",xL," xD=",xD," xP=",xP," gr=",gr];
While[gr > eps, k = k + 1;
  If[f[xL] f[xP] < 0, xD = xP, xL = xP];
  xP = (xL + xD)/2;
  gr = (xD - xL)/2;
  Print["k=", k," xL=",xL," xD=",xD, " xP=",xP," gr=",gr]],
Print["Interval nije dobro odredjen"]
]
```

*Zadatak 4.1.* Izradi Mathematica-modul koji će pozivati funkciju  $f$ , interval  $[a, b]$  i točnost  $\epsilon > 0$ , te algoritmom BISEKCIJE davati aproksimaciju nultočke funkcije  $f$  na intervalu  $[a, b]$ .

**Primjer 4.4.** *S točnošću  $\epsilon = 0.05$  korištenjem algoritma bisekcije izračunat ćemo nultočku funkcije  $f(x) = x^3 - 6x + 2$  na intervalu  $[0, 1.5]$  (vidi Primjer 4.1 str. 86).*

Nakon 6 iteracija dobivamo  $x^* = 0.3$ . Rezultat izvođenja iterativnog postupka vidljiv je u *Tablici 4.1*.

**Primjer 4.5.**<sup>1</sup> *Kao ilustraciju Primjedbe 4.1, str. 91 promatrajmo problem rješavanja sljedećih jednadžbi metodom bisekcije:*

$$(a) \quad f(x) = 0.001x + \frac{1}{2} + \frac{\pi}{2} + \arctg x = 0, \quad x \in [-600, -400]$$

$$(b) \quad g(x) = 1000(x - 4) - e^x = 0, \quad x \in [4, 5].$$

<sup>1</sup>Ovaj primjer konstruirao je J. Muharemović, student III godine ETF-a u Osijeku, 1998.

$n$	$a_n$	$b_n$	$x_{n+1}$	$\text{sgn}f(x_{n+1})$	$(b_{n-1} + a_{n-1})/2$
0	0	1.5	0.75	-1	0.75
1	0	0.75	0.375	-1	0.375
2	0	0.375	0.1875	1	0.1875
3	0.1875	0.375	0.28125	1	0.09375
4	0.28125	0.375	0.328125	1	0.046875
5	0.328125	0.375	0.351563	-1	0.02344
6	0.328125	0.351563	0.339844	1	0.0117

Tablica 4.1: Metoda bisekcije

a) Graf funkcije  $f$  na zadanom intervalu približno odgovara grafu prikazanom na Slici 4.4.a i ima nultočku  $\xi = -501.992061$  točnu na 6 decimala. Uz  $\epsilon = .005$  (odnosno  $\epsilon = 0.00005$ ) i kriterij zaustavljanja:  $|f(x_n)| < \epsilon$  nakon 5 (odnosno 9) iteracija dobivamo aproksimaciju  $x_5 = -506.25$  (odnosno  $x_9 = -501.9531$ ). Za aproksimaciju točnu na 2 (odnosno 4) decimale bilo bi potrebno 12 (odnosno 18) iteracija.

b) Graf funkcije  $g$  na zadanom intervalu približno odgovara grafu prikazanom na Slici 4.4.b i ima nultočku  $\xi = 4.057850$  točnu na 6 decimala. Uz  $\epsilon = .005$  (odnosno  $\epsilon = 0.00005$ ) i kriterij zaustavljanja:  $|f(x_n)| < \epsilon$  nakon 16 (odnosno 18) iteracija dobivamo puno točniju aproksimaciju nego što je tražena:  $x_{16} = 4.057846$  (odnosno  $x_{18} = 4.057850$ ). Za aproksimaciju točnu na 2 (odnosno 4) decimale potrebno je samo 4 (odnosno 12) iteracija.

## 4.2 Ocjena pogreške i brzina konvergencije

Pretpostavit ćemo da je promatrana funkcija  $f$  na intervalu  $I = [a, b]$  dovoljno "glatka". tj. da je klase  $C^1[a, b]$ . Ako je  $x_n$  jedna aproksimacija nultočke  $\xi$  u intervalu  $I$  i ako je  $f$  derivabilna funkcija, takva da je  $|f'(x)| > 0$ ,  $x \in I$ , onda vrijedi ovakva ocjena apsolutne pogreške

$$|\xi - x_n| \leq \frac{|f(x_n)|}{m_1}, \quad \text{gdje je } 0 < m_1 = \min_{x \in I} |f'(x)|. \quad (4.3)$$

Naime, prema teoremu o srednjoj vrijednosti, imamo

$$f(\xi) - f(x_n) = (\xi - x_n)f'(c), \quad c \in I.$$

Oдавde je

$$|f(x_n)| = |\xi - x_n| \cdot |f'(c)| \geq m_1 \cdot |\xi - x_n|,$$

odakle zbog stroge monotonosti funkcije  $f$  slijedi ocjena (4.3).

Ako rješenje  $\xi$  jednadžbe  $f(x) = 0$  određujemo kao graničnu vrijednost jednog beskonačnog niza, postavlja se pitanje kada zaustaviti iterativni proces.

Ako želimo da apsolutna pogreška aproksimacije  $x_n$  nultočke  $\xi$  ne bude veća od  $\epsilon > 0$ , onda je prema (4.3) dovoljno ispuniti uvjet  $|f(x_n)|/m_1 < \epsilon$ . Ovo može biti jedan kriterij za zaustavljanje iterativnog procesa.

*Primjedba 4.1.* Moglo bi se pomisliti da je prirodan kriterij za zaustavljanje iterativnog procesa ispunjenje nejednakosti:  $|f(x_n)| < \epsilon$ , za neki unaprijed zadani  $\epsilon > 0$ . Međutim, lako je iskonstruirati primjer u kome će za neki  $n$  biti  $|f(x_n)| < \epsilon$ , a da ipak pogreška aproksimacije  $|\xi - x_n|$  bude velika, tj. da  $x_n$  na brojevnom pravcu bude daleko od  $\xi$  (vidi *Sliku 4.3.a*). Naime, u tom slučaju broj  $m_1$  je malen, pa je  $\frac{1}{m_1}$  velik, što znači da će biti teško ispuniti zahtjev  $\frac{|f(x_n)|}{m_1} < \epsilon$ , koji će prema (4.3) osigurati da pogreška aproksimacije  $x_n$  ne premaši  $\epsilon$ .

Ako je  $m_1$  velik,  $\frac{1}{m_1}$  je malen, pa prema (4.3) pogreška aproksimacije može biti malena ( $x_n$  je blizu rješenja  $\xi$ ), a da vrijednost funkcije  $f(x_n)$  bude relativno velika (*Slika 4.3.b*).

Ako se dogodi da je  $m_1 = 0$ , metodom bisekcije (vidi t.4.1) treba suziti interval, tako da bude  $m_1 > 0$ .

Budući da je u primjerima često puta teško odrediti konstantu  $m_1$ , umjesto kriterija za zaustavljanje procesa  $\frac{|f(x_n)|}{m_1} < \epsilon$ , često se koristi

$$|x_n - x_{n-1}| < \epsilon_1 \quad \& \quad |f(x_n)| < \epsilon_2, \quad (4.4)$$

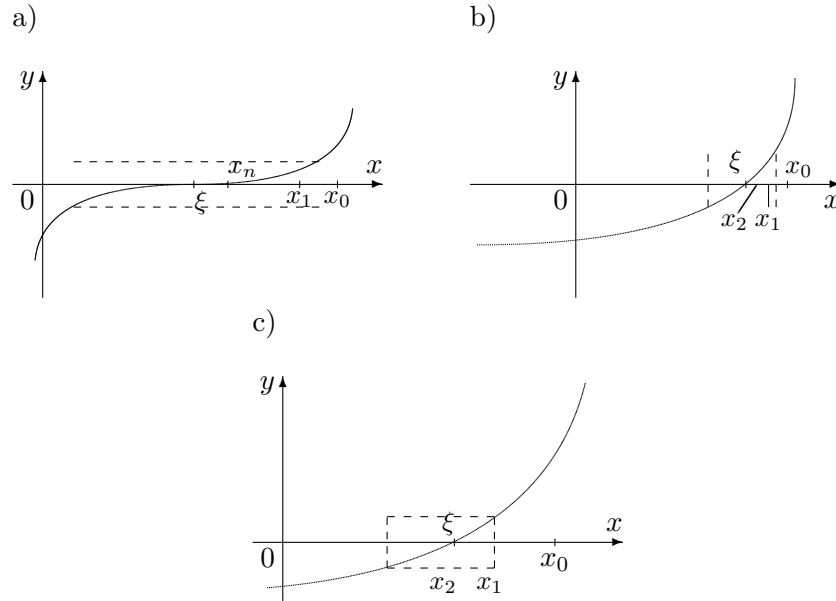
gdje su  $\epsilon_1, \epsilon_2$  unaprijed zadani brojevi.

Za neku iterativnu metodu važno je poznavati brzinu (red) konvergencije metode. U tu svrhu uvodimo sljedeću definiciju

**Definicija 4.1.** (Red konvergencije metode) Neka niz  $(x_n)$ , dobiven nekom iterativnom metodom, konvergira prema  $\xi \in \mathbb{R}$  i neka je  $e_n = \xi - x_n$  pogreška  $n$ -te aproksimacije. Tada, ako postoje dvije pozitivne konstante  $A, r \in \mathbb{R}_+$ , takve da vrijedi

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^r} = A, \quad (4.5)$$

kažemo da metoda ima red konvergencije  $r$ .



Slika 4.4: Kriteriji zaustavljanja iterativnog procesa

Specijalno, ako je  $r = 1$ , kažemo da metoda ima linearnu, a ako je  $r = 2$ , da ima kvadratnu brzinu konvergencije. Iz navedene definicije jasno je da će za veći  $r$  niz  $(x_n)$  brže konvergirati prema  $\xi$ .

Primijetimo također da iz *Definicije 4.1* slijedi da metoda ima red konvergencije  $r$  ako postoji  $A \in \mathbb{R}_+$  i  $n_0 \in \mathbb{N}$ , tako da bude

$$|\xi - x_{n+1}| \leq A|\xi - x_n|^r, \quad n \geq n_0.$$

### 4.3 Metoda jednostavnih iteracija

Traženje nultočke funkcije  $f: [a, b] \rightarrow \mathbb{R}$  na intervalu  $[a, b]$  može se povezati s traženjem fiksne točke funkcije  $\varphi: [a, b] \rightarrow \mathbb{R}$ ,

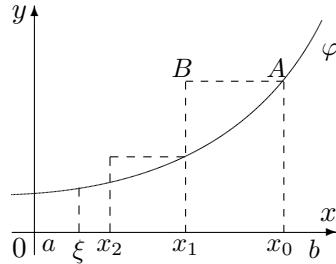
$$\varphi(x) := x - f(x). \quad (4.6)$$

na tom intervalu. Lako se vidi da je  $x_0 \in [a, b]$  nultočka funkcije  $f$  onda i samo onda ako je  $x_0$  fiksna točka funkcije  $\varphi$ , a rješavanje jednadžbe  $f(x) = 0$  ekvivalentno je traženju fiksne točke funkcije  $\varphi$ , tj. rješavanju jednadžbe

$$x = \varphi(x). \quad (4.7)$$

Za funkciju  $\varphi$  obično postoji više izbora. Primjerice, za funkciju iz *Primjera 4.2*, str. 86 za  $x > -\frac{5}{4}$  odgovarajuću jednadžbu (4.7) možemo pisati u obliku

$$x = e^x - 5/4 \quad \text{ili} \quad x = \ln(x + 5/4).$$



Slika 4.5: Metoda jednostavnih iteracija

Geometrijski gledano, umjesto rješavanja jednadžbe (4.1), tražimo sjecišta grafova funkcija  $y_1 = x$  i  $y_2 = \varphi(x)$ . Nakon što smo odredili interval  $I$  u kome se nalazi tražena nultočka, u njemu ćemo izabrati početnu aproksimaciju  $x_0$ . Sljedeću aproksimaciju  $x_1$  odredit ćemo tako da bude

$$x_1 = \varphi(x_0).$$

Na *Slici 4.5* točka  $x_1$  određena je tako da bude  $\overline{x_0A} = \overline{x_1B} = \overline{Ox_1}$ . Posljednje dvije dužine međusobno su jednake jer je trokut  $\triangle(O, x_1, B)$  jednakokratan. Ponavljajući postupak, dobivamo niz

$$x_0, x_1 = \varphi(x_0), x_2 = \varphi(x_1), \dots, x_n = \varphi(x_{n-1}), \dots \quad (4.8)$$

Ovako definirani niz može, ali i ne mora konvergirati, kao što pokazuju sljedeći jednostavni *Mathematica*-programi

$$\begin{array}{ll} \text{In}[1] := \text{phi}[x_] := e^x - \frac{5}{4} & \text{In}[1] := \text{phi}[x_] := \text{Log}[x + \frac{5}{4}] \\ x_0 = 1; & x_0 = 1; \\ \text{In}[2] := x_1 = \text{phi}[x_0] & \text{In}[2] := x_1 = \text{phi}[x_0] \\ x_0 = x_1; & x_0 = x_1; \end{array}$$

Višestruko pokretanje linije *In*[2] u lijevom programu daje divergentan niz, a višestruko pokretanje linije *In*[2] u desnom programu daje konvergentan niz.

Mogu se postaviti sljedeća pitanja:



- je li niz definiran s (4.8) konvergentan; ako nije, je li moguće novim izborom funkcije  $\varphi$  dobiti konvergentan niz?
- ako je niz (4.8) konvergentan, teži li prema rješenju jednadžbe (4.7)?
- ako je niz (4.8) konvergentan i ako teži rješenju jednadžbe (4.7), kakva je brzina (red) konvergencije?
- budući da rješenje  $\xi$  jednadžbe (4.7) nije poznato, kako ocijeniti pogrešku pojedine aproksimacije  $x_n$ ?

Odgovor na sva ova pitanja daje sljedeći teorem<sup>2</sup>.

**Teorem 4.1.** *Neka je  $\varphi : I = [a, b] \rightarrow \mathbb{R}$  neprekidno derivabilna funkcija za koju vrijedi:*

(i)  $\varphi(x) \in I$  za svaki  $x \in I$ ,

(ii)  $\exists q \in (0, 1)$ , takav da je  $|\varphi'(x)| \leq q$  za svaki  $x \in (a, b)$ .

Tada postoji jedinstveni  $\xi \in I$  takav da bude  $\varphi(\xi) = \xi$ . Osim toga, za proizvoljni  $x_0 \in I$ , niz definiran s

$$x_n = \varphi(x_{n-1}), \quad n = 1, 2, \dots \quad (4.9)$$

konvergira prema  $\xi$  i vrijede ovakve ocjene pogreške aproksimacije

$$|\xi - x_n| \leq \frac{q}{1-q} |x_n - x_{n-1}|. \quad (4.10)$$

$$|\xi - x_n| \leq \frac{q^n}{1-q} |x_1 - x_0|, \quad (4.11)$$

Metoda ima linearnu brzinu konvergencije, tj. vrijedi

$$|\xi - x_{n+1}| \leq q |\xi - x_n|. \quad (4.12)$$

*Dokaz.* Kako je  $x_n = \varphi(x_{n-1})$  i  $x_{n+1} = \varphi(x_n)$ , koristeći Lagrangeov teorem o srednjoj vrijednosti, imamo

$$x_{n+1} - x_n = \varphi(x_n) - \varphi(x_{n-1}) = (x_n - x_{n-1})\varphi'(c), \quad c \in (a, b).$$

Zato je

$$|x_{n+1} - x_n| \leq q |x_n - x_{n-1}| \quad \text{za svaki } n = 1, 2, \dots \quad (4.13)$$

<sup>2</sup>Ovaj teorem je specijalni slučaj općenitijeg Banachovog teorema o fiksnoj točki (Dahlquist and Björck, 2008; Deuffhard, 1995; Deuffhard and Hohmann, 2008)

Sukcesivnom primjenom nejednakosti (4.13) dobivamo

$$|x_{n+1} - x_n| \leq q|x_n - x_{n-1}| \leq \cdots \leq q^n|x_1 - x_0|. \quad (4.14)$$

Konstruirajmo ovakav red

$$x_0 + (x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1}) + \cdots \quad (4.15)$$

Niz njegovih  $n$ -tih parcijalnih suma odgovara nizu  $(x_n)$ , tj.  $s_n = x_n$ .

Zbog (4.14) jedna majoranta ovog reda je red

$$|x_0| + |x_1 - x_0| + q|x_1 - x_0| + \cdots + q^{n-1}|x_1 - x_0| + \cdots \quad (4.16)$$

kojeg možemo zapisati ovako

$$|x_0| + |x_1 - x_0|(1 + q + \cdots + q^{n-1} + \cdots).$$

Kako je  $q < 1$ , geometrijski red u zagradi je konvergentan, pa je i red (4.16) konvergentan. Na taj način pronašli smo jednu konvergentnu majorantu reda (4.15), što znači da je i taj red konvergentan. Po definiciji konvergencije reda, to znači da je i odgovarajući niz  $n$ -tih parcijalnih suma  $(s_n)$  reda konvergentan. Kako je  $s_n = x_n$ , to znači da je i niz  $(x_n)$  konvergentan, tj. postoji realan broj  $\xi$ , takav da je  $\xi = \lim x_n$ . Zbog uvjeta (i) čitav niz  $(x_n)$  je sadržan u  $I$ , a kako je  $I$  zatvoren skup, također je i  $\xi \in I$ .

Ako sada u (4.9) pustimo  $n \rightarrow \infty$ , onda zbog neprekidnosti funkcije  $\varphi$  imamo  $\xi = \varphi(\xi)$ , što znači da je  $\xi$  korijen jednadžbe (4.7).

Dokažimo da je  $\xi$  jedini korijen jednadžbe (4.7) u intervalu  $I$ . Pretpostavimo da je  $\bar{\xi}$  također korijen jednadžbe (4.7), tj. pretpostavimo da vrijedi  $\bar{\xi} = \varphi(\bar{\xi})$ . Tada prema teoremu o srednjoj vrijednosti mora biti

$$\bar{\xi} - \xi = \varphi(\bar{\xi}) - \varphi(\xi) = (\bar{\xi} - \xi)\varphi'(c), \quad c \in (a, b),$$

odakle je

$$(\bar{\xi} - \xi)(1 - \varphi'(c)) = 0.$$

Budući da je  $1 - \varphi'(c) \neq 0$  zbog uvjeta (ii), mora biti  $\bar{\xi} = \xi$ . Dakle,  $\xi$  je jedino rješenje jednadžbe (4.7).

Prijedimo na dokaza ocjene (4.10). Za funkciju  $f(x) = x - \varphi(x)$  zbog uvjeta (ii) vrijedi  $f'(x) = 1 - \varphi'(x) \geq 1 - q > 0$ . Zato uz korištenje teorema o srednjoj vrijednosti dobivamo

$$|x_n - \varphi(x_n)| = |f(x_n)| = |f(x_n) - f(\xi)| = |(x_n - \xi)f'(c)| \geq (1 - q)|x_n - \xi|,$$

odakle, ponovno koristeći teorem o srednjoj vrijednosti dobivamo traženu ocjenu (4.10)

$$|x_n - \xi| \leq \frac{|x_n - \varphi(x_n)|}{1-q} = \frac{|\varphi(x_{n-1}) - \varphi(x_n)|}{1-q} \leq \frac{q|x_n - x_{n-1}|}{1-q}.$$

Ocjenu (4.11) dobivamo iz (4.10) koristeći (4.14)

$$|x_n - \xi| \leq \frac{q}{1-q}|x_n - x_{n-1}| \leq \frac{q^n}{1-q}|x_1 - x_0|.$$

Još treba pokazati da metoda ima linearnu brzinu konvergencije. Koristeći  $x_{n+1} = \varphi(x_n)$  i  $\xi = \varphi(\xi)$  i teorem o srednjoj vrijednosti, dobivamo

$$|\xi - x_{n+1}| = |\varphi(\xi) - \varphi(x_n)| = |(\xi - x_n)\varphi'(c)| \leq q|\xi - x_n|.$$

odakle neposredno slijedi (4.12).  $\square$

*Primjedba 4.2.* Ako je unaprijed zadana točnost  $\epsilon$  s kojom želimo dobiti aproksimaciju rješenja jednadžbe (4.7), dovoljno je desnu stranu nejednakosti (4.11) ili desnu stranu nejednakosti (4.10) učiniti manjom od  $\epsilon$ . Na taj način osigurali smo da apsolutna pogreška aproksimacije ne bude veća od  $\epsilon$ .

*Zadatak 4.2.* Geometrijski objasnite značenja uvjeta (i), (ii).

**Primjer 4.6.** *S točnošću  $\epsilon = 0.005$  (na dvije decimale) treba pronaći pozitivnu nultočku funkcije iz Primjera 4.2. Pokazano je da se nultočka nalazi u intervalu  $I = [0, 1]$ .*

Ako odgovarajuću jednadžbu napišemo u obliku

$$x = e^x - \frac{5}{4},$$

funkcija  $\varphi$  bila bi  $\varphi(x) = e^x - \frac{5}{4}$ . Kako je  $\varphi'(x) = e^x \geq 1$  za  $x \in I$ , uvjeti *Teorema 4.1* nisu ispunjeni. Zato ćemo jednadžbu napisati u obliku

$$x = \ln(x + \frac{5}{4}), \quad x > -\frac{5}{4}.$$

Sada je  $\varphi(x) = \ln(x + \frac{5}{4})$ , a  $\varphi'(x) = 4/(4x + 5)$ . Lako se vidi (nacrtaj graf funkcije  $\varphi'$ ) da je

$$\varphi'(x) \leq \frac{4}{5} < 1 \quad \text{za svaki } x \in I \quad \text{i} \quad \varphi([0, 1]) = [\ln \frac{5}{4}, \ln \frac{9}{4}] \approx [0.223, 0.811] \subset [0, 1].$$

To znači da su uvjeti *Teorema 4.1* ispunjeni.

Niže ćemo konstruirati **Mathematica**-program za metodu jednostavnih iteracija. Definirajmo najprije funkcije  $f$  i  $\varphi$  iz prethodnog primjera.

```
In[1]:= f[x_]:= Exp[x] - x - 5/4; fi[x_]:= Log[x + 5/4]
```

Nakon toga definirat ćemo potrebne parametre i pokrenuti algoritam.

```
In[2]:= n=0; x0=1.; q=.8; eps =.005;
While[n = n + 1; x1 = fi[x0]; gr = Abs[x1 - x0] q/(1 - q);
Print["n=", n, " x", n, "=", x1, " f(x", n, ")=", f[x1],
" gr=", gr];
gr1 > eps, x0 = x1]
```

*Zadatak 4.3.* Pogrešku  $gr$  možemo alternativno definirati prema (4.11) ili uz dodatno definiranje parametra  $m_1 = \min_{x \in [a,b]} |f'(x)|$ , prema (4.3). Dopunite prethodni program tako da u svakom koraku računate pogrešku i prema ove dvije ocjene, te na taj način usporedite sve tri ocjene.

Cijeli iterativni proces prikazan je u *Tablici 4.2*. U trećem, odnosno četvrtom stupcu prikazane su vrijednosti desne strane ocjene (4.11); odnosno ocjene (4.10).

$n$	$x_n$	(4.11)	(4.10)	$f(x_n)$
0	1	-	-	0.468282
1	0.810930	0.756280	0.756280	0.189070
2	0.723157	0.605024	0.351091	0.087773
3	0.679635	0.484019	0.174090	0.043522
4	0.657331	0.387215	0.089217	0.022304
5	0.645705	0.309772	0.046504	0.011626
6	0.639591	0.247818	0.024457	0.006114
7	0.636360	0.198254	0.012922	0.003230
8	0.634649	0.158603	0.006844	0.001711
9	0.633742	0.126883	0.003630	0.000907

Tablica 4.2: Metoda jednostavnih iteracija

Vidi se da je ocjena (4.11) mnogo grublja od ocjene (4.10). Dakle, pouzdano možemo reći da apsolutna pogreška aproksimacije  $x^* = 0.63$  tražene nultočke nije veća od 0.005.

**Primjer 4.7.** *S točnošću  $\epsilon = 0.005$  (na dvije decimale) treba odrediti nultočku funkcije*

$$f(x) = x - \frac{2}{x}.$$

Kako je  $f(1) = -1$  i  $f(2) = 1$ , interval u kome leži nultočka je  $I = [1, 2]$ . Ako pripadnu jednadžbu napišemo u obliku  $x = \frac{2}{x}$  i definiramo  $\varphi(x) := \frac{2}{x}$ , proces neće

konvergirati jer je  $\varphi'(x) = -\frac{2}{x^2}$  i  $|\varphi'(x)| > 1$  na intervalu  $I$  (primjerice  $\varphi'(1) = -2$ ). Zbog toga ćemo funkciju  $\varphi$  morati drugačije definirati. Jednadžbu  $x = \frac{2}{x}$  zapisat ćemo u obliku

$$x = \frac{1}{2}\left(x + \frac{2}{x}\right).$$

Označimo  $\varphi(x) := \frac{1}{2}\left(x + \frac{2}{x}\right)$ . Lako se vidi da je derivacija funkcije  $\varphi$  na intervalu  $I$  monotono rastuća funkcija. Zato je  $q = \max\{|\varphi'(1)|, |\varphi'(2)|\} = \max\{\frac{1}{2}, \frac{1}{4}\} = \frac{1}{2}$ . Osim toga je  $\varphi([1, 2]) = [\sqrt{2}, 1.5] \subset [1, 2]$ . Time su uvjeti *Teorema 4.1* ispunjeni. Tijek odgovarajućeg iterativnog procesa prikazan je u *Tablici 4.3*. Inače, lako se vidi da je jedina nultočka promatrane funkcije broj  $\sqrt{2}$ . Zato ćemo u tablici pratiti i stvarnu apsolutnu pogrešku aproksimacija  $|\sqrt{2} - x_n|$ .

$n$	$x_n$	$ \sqrt{2} - x_n $	(4.11)	(4.10)	$f(x_n)/m_1$
0	1	0.414214	-	-	-1
1	1.500000	0.085786	0.500000	0.500000	0.166667
2	1.416667	0.002453	0.255000	0.083333	0.004902
3	1.414216	0.000002	0.125000	0.002451	0.000004

Tablica 4.3: Metoda jednostavnih iteracija

Na ovom primjeru vidi se da iako ocjena (4.10) u drugoj iteraciji ne osigurava ni jednu točnu decimalu, ipak već imamo jednu signifikantnu decimalu. U trećoj iteraciji ocjena (4.10) osigurava dvije signifikantne decimale, iako stvarna apsolutna pogreška pokazuje da smo postigli čak pet signifikantnih znamenki, od čega četiri decimalna mjesta.

Najbolju ocjenu pogreške u ovom primjeru dobili bismo korištenjem ocjene (4.3), gdje je  $m_1 = \min |f'(x)| > 1.5$ . Provjerite!

U praksi tražena nultočka obično nije poznata i zbog toga je važno da ocjena pogreške ne bude manja od stvarne pogreške, ali da ne bude previše gruba.

*Primjedba 4.3.* *Teorem 4.1* može se dokazati i uz slabije zahtjeve na funkciju. Umjesto da se traži da funkcija klase  $C^1[a, b]$  i da zadovoljava uvjete (i), (ii), može se tražiti da je *Lipschitz neprekidna* s konstantom  $0 \leq L < 1$ , tj. može se tražiti da je *kontrakcija* na  $[a, b]$  (Dahlquist and Björck, 2008; Deuffhard, 1995; Deuffhard and Hohmann, 2008).

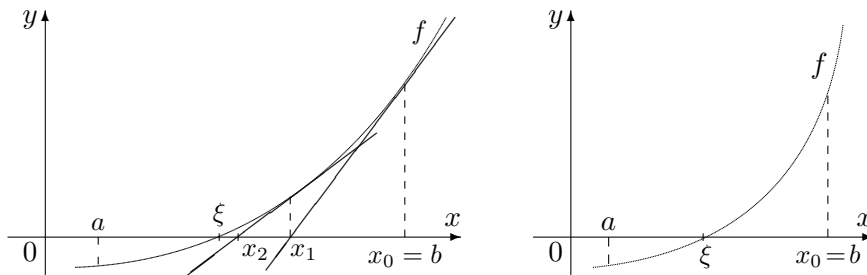
Ideja metode iteracija može se koristiti i u drugim situacijama. Tako primjerice na sličan način možemo pristupiti i rješavanju sustava linearnih (t.3.8, str. 73) ili nelinearnih (t.4.6, str. 112) jednadžbi.

## 4.4 Newtonova metoda

Pretpostavimo da smo na neki način odredili interval  $I = [a, b]$ , u kome se nalazi jedinstvena nultočka neprekidne i dovoljno “glatke” funkcije  $f$  za koju je  $f'(x) \neq 0 \forall x \in I$ . Izaberimo početnu aproksimaciju  $x_0 \in I$ , razvijmo funkciju  $f$  u Taylorov red u okolini točke  $x_0$  i zadržimo se na linearnom članu. Tako smo funkciju  $f$  u okolini točke  $x_0$  aproksimirali linearnom funkcijom

$$f_1(x) = f(x_0) + (x - x_0)f'(x_0) \quad (4.17)$$

čiji graf je tangenta na graf funkcije  $f$  u točki  $x_0$  (Slika 4.6.a).



Slika 4.6: Newtonova metoda

Izbor početne aproksimacije  $x_0$ 

Sada ćemo umjesto rješavanja jednadžbe (4.1), rješavati jednadžbu  $f_1(x) = 0$  (geometrijski gledano, umjesto traženja sjecišta grafa funkcije  $f$  s osi  $x$ , tražimo sjecište tangente s osi  $x$ ). Rješenje jednadžbe  $f_1(x) = 0$  označimo s  $x_1$

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)}. \quad (4.18)$$

Ponavljajući postupak, dobivamo niz  $x_0, x_1, \dots$  zadan rekurzivnom formulom<sup>3</sup>

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots \quad (4.19)$$

Važno je primijetiti da u svakoj iteraciji konstruiramo lokalni aproksimant funkcije  $f$  i onda tražimo nultočku lokalnog aproksimanta.

**Teorem 4.2.** *Neka funkcija  $f: I \rightarrow \mathbb{R}$  ima neprekidnu drugu derivaciju na intervalu  $I = [a, b]$ . Neka je nadalje,  $f(a) \cdot f(b) < 0$ , a prva ( $f'$ ) i druga ( $f''$ ) derivacija funkcije  $f$  na intervalu  $I$  imaju stalan predznak.*

<sup>3</sup>U literaturi često ovu metodu možemo naći pod imenom *Newton-Raphsonova metoda*.

Tada, ako je  $x_0 \in I$  izabran tako da bude

$$f(x_0) \cdot f''(x_0) > 0, \quad (4.20)$$

niz definiran s (4.19) konvergira prema jedinstvenom rješenju  $\xi$  jednadžbe  $f(x) = 0$ .

Pri tome vrijedi ocjena pogreške aproksimacije

$$|\xi - x_n| \leq \frac{M_2}{2m_1}(x_n - x_{n-1})^2, \quad (4.21)$$

gdje je

$$m_1 = \min_{x \in I} |f'(x)|, \quad M_2 = \max_{x \in I} |f''(x)|, \quad (4.22)$$

a metoda ima kvadratnu brzinu konvergencije, tj. vrijedi

$$|\xi - x_{n+1}| \leq \frac{M_2}{2m_1}(\xi - x_n)^2. \quad (4.23)$$

*Dokaz.* Razmotrimo slučaj<sup>4</sup>

$$f(a) < 0, \quad f(b) > 0, \quad f'(x) > 0, \quad f''(x) > 0 \quad \text{za svaki } x \in I.$$

Najprije primijetimo da postoji jedinstveni  $\xi \in I$ , takav da je  $f(\xi) = 0$  i izaberimo  $x_0 \in I$  tako da bude ispunjeno (4.20) (primjerice možemo uzeti  $x_0 = b$ ). Kako je po pretpostavci  $f''(x_0) > 0$ , onda je i  $f(x_0) > 0$ , pa je  $x_0 > \xi$  (vidi *Sliku 4.6.b*) jer je  $f$  rastuća funkcija.

Primjenom matematičke indukcije pokazat ćemo da je niz definiran s (4.19) monotono padajući i ograničen odozdo s  $\xi$ .

- Pretpostavimo da je  $x_n > \xi$ . Prema Taylorovoj formuli je

$$f(x) = f(x_n) + f'(x_n) \cdot (x - x_n) + \frac{1}{2} \cdot f''(c_1) \cdot (x - x_n)^2,$$

gdje je  $c_1$  između  $x$  i  $x_n$ . Specijalno, za  $x = \xi$  imamo

$$0 = f(\xi) = f(x_n) + f'(x_n) \cdot (\xi - x_n) + \frac{1}{2} \cdot f''(c_2) \cdot (\xi - x_n)^2,$$

gdje je  $c_2$  između  $\xi$  i  $x_n$ . Kako je  $f''(c_2) > 0$  i  $(\xi - x_n) \neq 0$  (zbog  $x_n > \xi$ ), onda je

$$f(x_n) + f'(x_n) \cdot (\xi - x_n) < 0.$$

Oдавde je

$$\xi < x_n - \frac{f(x_n)}{f'(x_n)} = x_{n+1},$$

pa je i  $x_{n+1} > \xi$ . Dakle niz  $(x_n)$  je odozdo ograničen sa  $\xi$ .

---

<sup>4</sup>Ostali slučajevi dokazuju se analogno.

- Kako je  $x_n > \xi$  za svaki  $n \in \mathbb{N}$ , onda zbog strogog monotonog rasta funkcije  $f$  mora biti i  $f(x_n) > f(\xi) = 0$  za svaki  $n \in \mathbb{N}$ . Osim toga po pretpostavci je i  $f'(x_n) > 0$  za svaki  $n \in \mathbb{N}$ . Zato iz (4.19) slijedi

$$x_n - x_{n+1} = \frac{f(x_n)}{f'(x_n)} > 0,$$

što znači da je niz  $(x_n)$  monotonno padajući.

Budući da je niz  $(x_n)$  monotonno padajući i ograničen odozdo, on je konvergentan i čitav se nalazi u  $I$ , pa zbog toga postoji realni broj  $\bar{\xi} \in I$ , takav da bude  $\bar{\xi} = \lim x_n$ . Ako sada u (4.19) pustimo  $n \rightarrow \infty$ , dobivamo

$$\bar{\xi} = \bar{\xi} - \frac{f(\bar{\xi})}{f'(\bar{\xi})},$$

odakle (zbog  $f'(\bar{\xi}) > 0$ ) slijedi  $f(\bar{\xi}) = 0$ . Kako je  $\bar{\xi}$  jedinstvena nultočka funkcije  $f$  na intervalu  $I$ , mora biti  $\bar{\xi} = \xi$ .

Prijeđimo na dokaz ocjene (4.21). Prema Taylorovoj formuli vrijedi

$$f(x_n) = f(x_{n-1}) + f'(x_{n-1}) \cdot (x_n - x_{n-1}) + \frac{1}{2} \cdot f''(c) \cdot (x_n - x_{n-1})^2,$$

gdje je  $c$  između  $x_n$  i  $x_{n-1}$ . Zbog (4.19), ostaje samo

$$f(x_n) = \frac{1}{2} f''(c) (x_n - x_{n-1})^2,$$

odakle je

$$|f(x_n)| \leq \frac{1}{2} M_2 (x_n - x_{n-1})^2. \quad (4.24)$$

Iz (4.3) koristeći (4.24), dobivamo ocjenu (4.21).

Još treba dokazati kvadratnu brzinu konvergencije metode, tj. dokazati nejednakost (4.23). Prema Taylorovoj formuli vrijedi

$$0 = f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \frac{1}{2} \cdot f''(c)(\xi - x_n)^2,$$

gdje je  $c$  između  $\xi$  i  $x_n$ . Dijeljenjem s  $f'(x_n)$  dobivamo

$$0 = \frac{f(x_n)}{f'(x_n)} + \xi - x_n + \frac{1}{2} \cdot \frac{f''(c)}{f'(x_n)} (\xi - x_n)^2.$$

Koristeći ovdje (4.19), dobivamo

$$x_{n+1} - \xi = \frac{1}{2} \cdot \frac{f''(c)}{f'(x_n)} (\xi - x_n)^2,$$

odakle slijedi (4.23). □



*Primjedba 4.4.* Ako je početna aproksimacija  $x_0$  izabrana u intervalu za koji vrijedi

$$\frac{M_2}{2m_1} \leq 2,$$

onda će se u svakom koraku broj signifikantnih decimala aproksimacije podvostručavati. Primjerice ako je  $n$ -ta aproksimacija imala  $k$  signifikantnih decimala, tj.  $|\xi - x_n| < \frac{1}{2}10^{-k}$  (vidi *Primjedbu 1.3*, str. 10), tada je prema (4.23)

$$|\xi - x_{n+1}| \leq \frac{M_2}{2m_1}(\xi - x_n)^2 \leq 2\frac{1}{4}10^{-2k} = \frac{1}{2}10^{-2k},$$

tj.  $(n+1)$ -va aproksimacija ima  $2k$  signifikantnih decimala.

*Primjedba 4.5.* Za Newtonovu metodu kažemo da je lokalno konvergentna. To znači da postoji neka okolina oko nultočke  $\xi$  u kojoj su uvjeti Teorema 4.2 ispunjeni i u kojoj možemo proizvoljno birati početnu aproksimaciju  $x_0$ , pa da iterativni proces konvergira prema rješenju  $\xi$ .

Primjerice, za funkciju  $f(x) = \arctg x$  ako je  $|x_0| < 1.39$ , Newtonova metoda konvergirat će prema  $\xi = 0$ , a ako je  $|x_0| > 1.40$ , Newtonova metoda će divergirati.

*Zadatak 4.4.* Za koju će vrijednost početne aproksimacije  $x_0$  Newtonova metoda za funkciju  $f(x) = \arctg x$  proizvoditi “kruženje” ( $x_1 = -x_0$ ,  $x_2 = x_0, \dots$ )?

**Primjer 4.8.** *S točnošću  $\epsilon = 0.00005$  (4 signifikantne decimale) treba odrediti pozitivnu nultočku funkcije*

$$f(x) = e^{-x} + x^2 - 2.$$

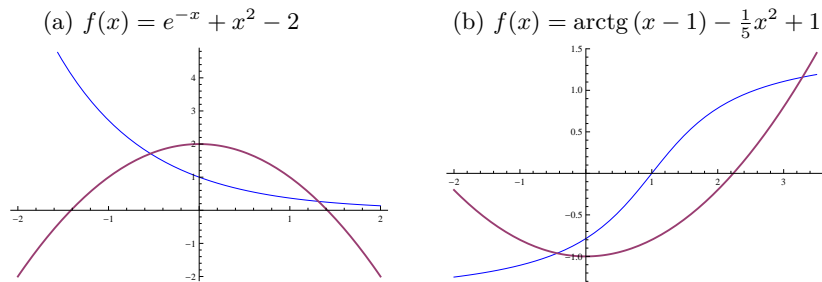
Ova funkcija ima dvije nultočke. Jedna se nalazi u intervalu  $I_1 = [-1, 0]$ , a druga u intervalu  $I_2 = [1, 2]$  (Slika 4.7.a). Mi ćemo potražiti nultočku iz intervala  $[1, 2]$ . Kako je  $f'(x) = -e^{-x} + 2x$  rastuća funkcija na  $I_2$ , bit će  $m_1 = f'(1) \geq 1.6$ , a kako je  $f''(x) = e^{-x} + 2$  padajuća funkcija na  $I_2$ , uzet ćemo  $M_2 = f''(1) \leq 2.4$ . Zbog  $f(2) \cdot f''(2) > 0$  za početnu aproksimaciju izabrat ćemo  $x_0 = 2$ .

Nakon 3 iteracije dobivamo  $x^* = 1.3160$ . Tijek iterativnog procesa prikazan je u *Tablici 4.4* ( $g_0$  je desna strana ocjene (4.3), a  $g_2$  desna strana ocjene (4.21)).

Za izvođenje iterativnog postupka može se koristiti sljedeći *Mathematica*-program

```
f[x_] := Exp[-x]+x^2-2
n=0; x0=2; m1=1.6; m2=2.4; eps=.00005;
x=x0;
While[N[Abs[f[x]]/m1]>eps,
  Print["Iteracija ",n,":  xn=",x," pogreska=",f[x]/m1];
  x=N[x-f[x]/f'[x]]; n=n+1
]
Print["Iteracija ",n,":  xn=",x," greska=",f[x]/m1];
```

$n$	$x_n$	$g_0$	$g_2$	$f(x_n)$	$f'(x_n)$
0	2	1.334585	-	2.135336	3.864665
1	1.4475	0.206462	0.228965	0.330339	2.659780
2	1.3233	0.010823	0.011569	0.017316	2.380286
3	1.3160	0.000037	0.000040	0.000060	2.363792

Tablica 4.4: Newtonova metoda za funkciju iz *Primjera 4.8*Slika 4.7: Separacija nultočaka iz *Primjera 4.8* i *Primjera 4.9*

**Primjer 4.9.** S točnošću  $\epsilon = 0.00005$  treba odrediti pozitivnu nultočku funkcije

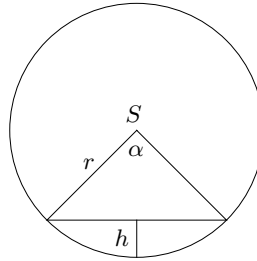
$$f(x) = \arctg(x - 1) - \frac{1}{5}x^2 + 1.$$

Ova funkcija ima dvije nultočke. Jedna se nalazi u intervalu  $I_1 = [-1, 0]$ , a druga u intervalu  $I_2 = [3, 4]$  (Slika 4.7b). Mi ćemo potražiti nultočku iz intervala  $[3, 4]$ . Kako je  $f'$  negativna i padajuća funkcija na  $I_2$ , uzet ćemo  $m_1 = |f'(3)| \approx 1$ , a kako je  $f''$  negativna i rastuća funkcija na  $I_2$ , uzet ćemo  $M_2 = |f''(3)| \approx 0.56$ . Zbog  $f(4) \cdot f''(4) > 0$  za početnu aproksimaciju izabrat ćemo  $x_0 = 4$ . Nakon 3 iteracije dobivamo  $x^* = 3.2850$ . Tijek iterativnog procesa prikazan je u *Tablici 4.5*

$n$	$x_n$	(4.3)	(4.21)	$f(x_n)$	$f'(x_n)$
0	4	0.950954	-	-0.950954	-1.500000
1	3.366031	0.095113	0.112537	-0.095113	-1.194853
2	3.286428	0.001621	0.001774	-0.001621	-1.154000
3	3.285023	0.000001	0.000001	-0.000001	-1.153272

Tablica 4.5: Newtonova metoda za funkciju iz *Primjera 4.9*

**Primjer 4.10.** Spremnik za naftu ima oblik položenog cilindra radijusa  $r = 1m$ . Kolika je visina  $h$  razine nafte ako je spremnik napunjen s  $q = \frac{1}{4}$  svog volumena?



Slika 4.8: Razina spremnika

Označimo s  $d$  duljinu spremnika, a s  $\alpha$  kut kružnog isječka (Slika 4.8). Količina (volumen) nafte zadana je formulom  $V_h = \frac{1}{2}r^2\alpha d - \frac{1}{2}r^2 \sin \alpha \cdot d$ , odnosno

$$V_h = \frac{1}{2}r^2d(\alpha - \sin \alpha). \quad (*)$$

gdje je  $d$  duljina spremnika.

S druge strane kako je volumen punog spremnika  $V = r^2\pi \cdot d$ , onda je  $q$ -ti dio njegovog volumena jednak  $V_q = r^2\pi \cdot d \cdot q$ . Kut  $\alpha$ , za koji je  $V_h = V_q$ , dobivamo iz jednadžbe

$$\alpha - \sin \alpha - 2\pi q = 0. \quad (**)$$

Kako je  $\cos \frac{\alpha}{2} = \frac{r-h}{r}$ , nakon što smo iz (\*\*\*) odredili  $\alpha$ , visinu  $h$  razine nafte dobivamo iz formule

$$h = r \left(1 - \cos \frac{\alpha}{2}\right). \quad (***)$$

Primjenom Newtonove metode tangenti uz  $\alpha_0 = 3$  i  $\epsilon = 0.00005$ , dobivamo  $\alpha = 2.3099$  ( $\approx 132^\circ$ ). Nakon toga iz (\*\*\*) dobivamo  $h = 59.6cm$ .

#### 4.4.1 Modifikacije Newtonove metode

Činjenica da u svakom koraku Newtonove metode moramo računati i vrijednost funkcije i njene derivacije, u nekim slučajevima otežava primjenu ove metode. Navest ćemo nekoliko modifikacija Newtonove metode u kojima se izbjegava izračunavanje derivacije funkcije u svakoj iteraciji.

Najvažnija modifikacija je tzv. **metoda sekanti**. U intervalu  $I = [a, b]$  izabrat ćemo dvije početne aproksimacije  $x_0$  i  $x_1$  te povući sekantu na krivulju kroz točke  $(x_0, f(x_0))$ ,  $(x_1, f(x_1))$ . Sljedeću aproksimaciju  $x_2$  dobit ćemo kao sjecište sekante s osi  $x$  (Slika 4.9a)

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}, \text{ ako je } f(x_1) \neq f(x_0).$$

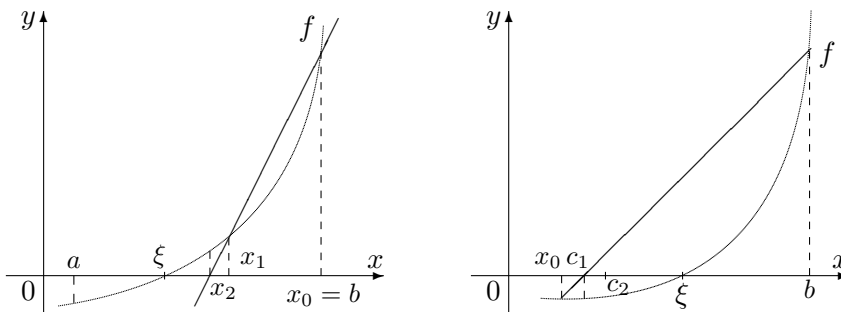
Ponavljajući postupak, dobivamo niz definiran rekurzivnom formulom

$$x_{n+1} = \frac{x_{n-1}f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad f(x_n) \neq f(x_{n-1}), \quad n = 1, 2, \dots \quad (4.25)$$

Ako su početne aproksimacije  $x_0, x_1$  birane dovoljno blizu rješenja  $\xi$ , onda uz uvjete  $f'(\xi) \neq 0$  i  $f''(\xi) \neq 0$ , metoda sekanti ima brzinu konvergencije  $\frac{1+\sqrt{5}}{2} \approx 1.618$  (vidi primjerice Dahlquist and Björck (2008); Schwarz and Köckler (2011); Stoer and Bulirsch (2002)).

*Zadatak 4.5.* Pokažite da se rekurzivna formula (4.25) može dobiti izravno iz Newtonove metode (4.19), str. 99, tako da  $f'(x_n)$  aproksimiramo podijeljenom razlikom

$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$



Slika 4.9: Metoda sekanti

Regula falsi

**Regula falsi** (*metoda krivih položaja*) je metoda, koju također možemo smatrati jednom modifikacijom Newtonove metode, odnosno jednom varijantom metode sekanti.<sup>5</sup>

Pretpostavimo da je funkcija  $f$  neprekidna na zatvorenom intervalu  $[a, b]$  i da je  $f(a) \cdot f(b) < 0$ . Označimo:  $x_0 := a, b_0 := b$  i povucimo pravac točkama  $(x_0, f(x_0)), (b_0, f(b_0))$ . On siječe os  $x$  u točki

$$c_1 = \frac{x_0 f(b_0) - b_0 f(x_0)}{f(b_0) - f(x_0)}, \quad x_0 < c_1 < b_0$$

Ako je  $f(c_1) = 0$ , nultočka je pronađena; u protivnom postupimo na sljedeći način

$$\text{ako je } \begin{array}{ll} f(c_1)f(x_0) > 0, & x_1 = c_1; \\ & b_1 = b_0 \end{array} \quad \text{inače } \begin{array}{ll} x_1 = c_1 & \\ b_1 = x_0 & \end{array}$$

<sup>5</sup>Također se može pokazati (Stoer and Bulirsch, 2002) da je regula falsi specijalni slučaj metode jednostavnih iteracija.

Ponavljajući postupak, dobivamo niz  $(x_n)$  koji linearnom brzinom konvergira prema jednom korijenu jednadžbe  $f(x) = 0$  na intervalu  $[a, b]$ . Prema izloženom vidi se da ovu metodu možemo smatrati i jednom varijantom metode bisekcije.

Za ocjenu pogreške kod svih ovih modifikacija možemo koristiti (4.3), str. 90.

**Zadatak 4.6.** Izradite *Mathematica*-programe za metodu sekani i metodu regula falsi i testirajte na priloženim primjerima.

**Primjer 4.11.** *Gibanje balističkog projektila u vakuumu određeno je sustavom diferencijalnih jednadžbi (vidi primjerice Molitz and R.Ströbel (1963))*

$$\frac{d^2x(t)}{dt^2} = 0, \quad \frac{d^2y(t)}{dt^2} = -g(y)$$

s početnim uvjetima

$$x(0) = y(0) = 0, \quad \frac{dx(0)}{dt} = v_0 \cos \alpha, \quad \frac{dy(0)}{dt} = v_0 \sin \alpha,$$

gdje su  $(x, y)$  koordinate (u metrima) proizvoljne točke  $P$  na putanji,  $t$  vrijeme (u sekundama) potrebno da projektil dođe u točku  $P$ ,  $\alpha$  izlazni kut,  $v_0$  početna brzina projektila (u  $\text{ms}^{-1}$ ) i  $g$  akceleracija sile teže (vidi Sliku 4.10a).

Za male visine  $y$  može se pretpostaviti da je funkcija  $g$  konstantna (približno  $9.81\text{ms}^{-2}$ ). Rješenje sustava diferencijalnih jednadžbi možemo zapisati u parametarskom obliku

$$\begin{aligned} \text{(a)} \quad x(t) &= v_0 t \cos \alpha \\ \text{(b)} \quad y(t) &= v_0 t \sin \alpha - \frac{1}{2}gt^2 \end{aligned} \quad (\text{B.1})$$

ili eksplicitno (isključujući vrijeme  $t$  iz jednadžbi (B.1))

$$y = x \tan \alpha - \frac{x^2}{4h \cos^2 \alpha}, \quad h = \frac{v_0^2}{2g}. \quad (\text{B.2})$$

Lako se vidi da je vrijeme leta projektila (dok ne padne u točku  $X$ )

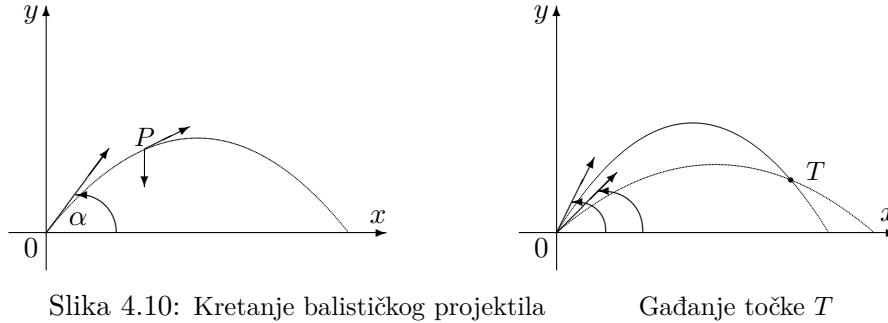
$$T = \frac{2v_0}{g} \sin \alpha, \quad (\text{B.3})$$

uz domet

$$X = 2h \sin 2\alpha. \quad (\text{B.4})$$

Maksimalni domet projektila postiže se za  $\alpha = \frac{\pi}{4} \text{ rad}$  ( $45^\circ$ ) i iznosi  $X_M = 2h$ .

**Primjer 4.12.** *Zadana je točka  $T(x_0, y_0)$ . Treba odrediti veličinu izlaznog kuta  $\alpha$  tako da projektil pogodi točku  $T$  (vidi Sliku 4.10b).*



Slika 4.10: Kretanje balističkog projektila

Gađanje točke  $T$ 

Da bi projektil mogao pogoditi točku  $T$ , njene koordinate moraju zadovoljavati jednadžbu

$$y_0 = x \operatorname{tg} \alpha - \frac{x_0^2}{4h \cos^2 \alpha},$$

što možemo pisati

$$y_0 \cos^2 \alpha - \frac{1}{2} x_0 \sin 2\alpha + \frac{x_0^2}{4h} = 0. \quad (*)$$

Jednadžbu (\*) možemo riješiti nekom od metoda navedenih u ovom poglavlju. Jednadžba (\*) može imati jedno ili dva rješenja (tzv. donji i gornji dijapazon kuteva). U *Tablici 4.6* prikazani su dobiveni rezultati za neke karakteristične situacije (kutevi su dani u stupnjevima).

$x_n$	$y_n$	$v_0$	$\alpha_1$	$\alpha_2$	max. domet (m)
1 000	10	115	24.66	65.92	1 348
50 000	500	900	19.28	71.29	82 569
300	1 000	1 200	73.33	-	146 789
500	3 000	500	-	-	25 484

Tablica 4.6: Gađanje točke  $T$ 

## 4.5 Newtonova metoda za Lipschitz derivabilnu funkciju

Uvjet  $f \in C^2[a, b]$  uz koji smo dokazali *Teorem 4.2* je prejak i zahtijeva da funkcija  $f$  bude dvostruko neprekidno derivabilna na intervalu  $[a, b]$ . Taj uvjet može se oslabiti tako da od funkcije  $f$  zahtijevamo da bude neprekidno derivabilna i da promjene (rast ili pad) njene derivacije ne budu prevelike.

Preciznije, tražit ćemo da derivacija  $f'$  funkcije  $f$  bude Lipschitz neprekidna funkcija na intervalu  $[a, b]$ . Taj uvjet zvat ćemo *Lipschitz derivabilnost*.

#### 4.5.1 Lipschitz neprekidna funkcija

Upoznajmo se najprije s pojmom “Lipschitz neprekidna funkcija”.

**Definicija 4.2.** Kažemo da je funkcija  $g : J \rightarrow \mathbb{R}$ ,  $J = (a, b)$  *Lipschitz-neprekidna s konstantom  $L$*  i pišemo  $g \in Lip_L(J)$  ako za svaki  $x, y \in J$  vrijedi

$$|g(x) - g(y)| \leq L|x - y|. \quad (4.26)$$

**Primjer 4.13.** Funkcija  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,  $g(x) = |x|$  je *Lipschitz-neprekidna s konstantom  $L = 1$*  na  $\mathbb{R}$  jer je  $\|x| - |y|| \leq 1 \cdot |x - y|$  za svaki  $x \in \mathbb{R}$ .

Za  $x \neq y$  Lipschitzov uvjet (4.26) možemo zapisati u obliku

$$-L \leq \frac{g(x) - g(y)}{x - y} \leq L \quad \forall x, y \in [a, b],$$

odakle se vidi da je relativna promjena (kvocijent diferencija) funkcije  $g$  ograničena između  $-L$  i  $L$ . Ako  $x$  fiksiramo, a  $y$  pustimo prema  $x$ , onda možemo zaključiti da ako postoji derivacija funkcije  $g$ , ona je ograničena između  $-L$  i  $L$ .

*Primjedba 4.6.* Svaka Lipschitz-neprekidna funkcija je neprekidna funkcija. Ako je  $f \in C^1[a, b]$ , onda je  $f \in Lip_L[a, b]$  (Lagrangeov teorem o srednjoj vrijednosti), ali obrat ne vrijedi (vidi *Primjer 4.16*). Ako je  $f: [a, b] \rightarrow \mathbb{R}$  neprekidna funkcija, koja ima po dijelovima neprekidnu derivaciju, onda primjenom Lagrangeovog teorema o srednjoj vrijednosti dobivamo konstantu  $L = \sup_{x \in [a, b]} |f'(x)|$  (vidi *Primjer 4.16*).

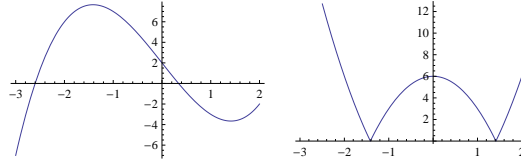
*Primjedba 4.7.* Definicija Lipschitz-neprekidne funkcije lako se može proširiti na funkciju više varijabli  $g: D \rightarrow \mathbb{R}^n$ ,  $D \subseteq \mathbb{R}^n$ . Takva funkcija je *Lipschitz-neprekidna* (zadovoljava *Lipschitzov uvjet*) ako postoji konstanta  $L > 0$  tako da za svaki  $x, y \in D$  vrijedi

$$|g(x) - g(y)| \leq L\|x - y\|,$$

gdje je  $\|\cdot\|$  neka norma.

**Primjer 4.14.** Treba odrediti Lipschitzovu konstantu funkcije  $f: [-3, 2] \rightarrow \mathbb{R}$ ,  $f(x) = x^3 - 6x + 2$ . Iz *Slike 4.11* može se vidjeti da je  $L = |f'(-3)| = 21$ .

4.5. NEWTONOVA METODA ZA LIPSCHITZ DERIVABILNU FUNKCIJU 109

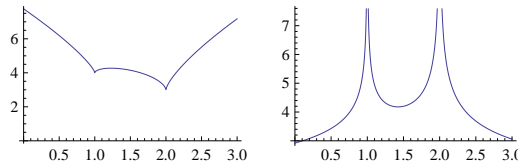


Slika 4.11: Funkcija  $f(x) = x^3 - 6x + 2$  i  $x \mapsto |f'(x)|$

**Primjer 4.15.** Treba odrediti Lipschitzovu konstantu funkcije  $f: [0, 3] \rightarrow \mathbb{R}$ ,  $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$ . Njena derivacija

$$f'(x) = \frac{4}{3\sqrt[3]{x-1}} + \frac{2}{\sqrt[3]{x-2}}$$

nije definirana u točkama iz  $\{1, 2\}$  (vidi Sliku 4.12), a Lipschitzova konstanta u ovom slučaju ne postoji.



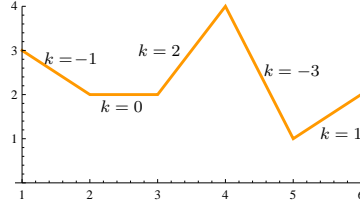
Slika 4.12: Funkcija  $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$  i  $x \mapsto |f'(x)|$

**Primjer 4.16.** Treba odrediti Lipschitzovu konstantu funkcije  $f: [1, 6] \rightarrow \mathbb{R}$ ,

$$f(x) = \begin{cases} -x + 4, & 1 \leq x \leq 2, \\ 2, & 2 \leq x \leq 3, \\ 2x - 4, & 3 \leq x \leq 4, \\ -3x + 16, & 4 \leq x \leq 5, \\ x - 4, & 5 \leq x \leq 6, \end{cases}$$

čiji graf je prikazan na Slici 4.13. Očigledno je  $L = 3$ .



Slika 4.13: Koeficijenti smjera po dijelovima linearne funkcije  $f$ 

### 4.5.2 Newtonova metoda

Sljedeća lema daje ocjenu pogreške aproksimacije Lipschitz derivabilne funkcije ( $f' \in Lip_L(J)$ ) afinom funkcijom  $f_1(x) = f(x_0) + f'(x_0)(x - x_0)$ .

**Lema 4.1.** *Neka je  $f : J \rightarrow \mathbb{R}$  funkcija, takva da je  $f' \in Lip_L(J)$ . Tada za proizvoljni  $x_0, x \in J$  vrijedi*

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq \frac{L}{2}(x - x_0)^2. \quad (4.27)$$

*Dokaz.* Iz Newton–Leibnizove formule slijedi

$$f(x) - f(x_0) = \int_{x_0}^x f'(t) dt.$$

Oduzimanjem vrijednosti  $\int_{x_0}^x f'(x_0) dt = f'(x_0)(x - x_0)$  lijevoj i desnoj strani dobivamo

$$f(x) - f(x_0) - f'(x_0)(x - x_0) = \int_{x_0}^x (f'(t) - f'(x_0)) dt.$$

Budući da je  $f' \in Lip_L(J)$  i da je  $t - x_0 > 0$  za  $t \in (x_0, x)$ , imamo

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq L \int_{x_0}^x (t - x_0) dt = L \frac{(x - x_0)^2}{2}.$$

□

Ocjena iz *Leme 4.1* bliska je ocjeni iz Taylorove formule (vidi primjerice Jukić and Scitovski (2004)), pri čemu je  $|f''(\xi)|$  ocijenjen s  $L$ . Sljedeći teorem daje uvjete uz koje Newtonova metoda za Lipschitz derivabilnu funkciju zadržava kvadratnu brzinu konvergencije.

**Teorem 4.3.** Neka je  $f: J \rightarrow \mathbb{R}$  funkcija, takva da je  $f' \in \text{Lip}_L(J)$ . Nadalje, pretpostavimo da postoji  $m_1 > 0$ , tako da vrijedi  $|f'(x)| \geq m_1$  za svaki  $x \in J$ .

Ako postoji  $\xi \in J$ , takav da je  $f(\xi) = 0$ , onda postoji  $\eta > 0$ , takav da za proizvoljno<sup>6</sup>  $x_0 \in (\xi - \eta, \xi + \eta)$ , niz generiran rekursivnom formulom

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots \quad (4.28)$$

konvergira prema  $\xi$  kvadratnom brzinom

$$|x_{n+1} - \xi| \leq \frac{L}{2m_1} (x_n - \xi)^2. \quad (4.29)$$

*Dokaz.* Neka je  $\tau \in (0, 1)$ , a  $\hat{\eta}$  radijus najvećeg otvorenog simetričnog intervala oko  $\xi$  sadržanog u  $J$  i neka je  $\eta = \min \left\{ \hat{\eta}, \tau \frac{2m_1}{L} \right\}$ . Primijetite da to znači da  $x_0 \in J$  treba birati tako da je

$$|x_0 - \xi| < \hat{\eta} \quad \& \quad |x_0 - \xi| < \tau \frac{2m_1}{L},$$

što može biti vrlo blizu  $\xi$ .

Najprije ćemo induktivno dokazati da vrijedi (4.29). Po pretpostavci je  $|x_0 - \xi| \leq \eta$ . Nadalje slijedi

$$x_1 - \xi = x_0 - \xi - \frac{f(x_0)}{f'(x_0)} = x_0 - \xi - \frac{f(x_0) - f(\xi)}{f'(x_0)} = \frac{1}{f'(x_0)} [f(\xi) - f(x_0) - f'(x_0)(\xi - x_0)].$$

Koristeći *Lemu 4.1* dobivamo

$$|x_1 - \xi| \leq \frac{1}{f'(x_0)} \frac{L}{2} (\xi - x_0)^2 \leq \frac{L}{2m_1} (\xi - x_0)^2,$$

što je baza indukcije. Slično se pokaže i korak indukcije.

Nadalje, pokažimo da je cijeli niz  $(x_n)$  sadržan u okolini  $(\xi - \eta, \xi + \eta)$ . Po pretpostavci teorema za  $x_0$  to je ispunjeno. Kako je  $|x_0 - \xi| < \eta \leq \tau \frac{2m_1}{L}$ , iz *Leme 4.1* slijedi

$$|x_1 - \xi| \leq \frac{L}{2m_1} |x_0 - \xi|^2 \leq \frac{L}{2m_1} |x_0 - \xi| \tau \frac{2m_1}{L} = \tau |x_0 - \xi| < \tau \eta < \eta.$$

Uz pretpostavku da je  $|x_n - \xi| < \eta$ , na sličan način zaključujemo da je

$$|x_{n+1} - \xi| \leq \tau |x_n - \xi| < \eta.$$

To znači da cijeli niz  $(x_n)$  ostaje u okolini  $(\xi - \eta, \xi + \eta)$ .

Pokažimo još da niz  $(x_n)$  konvergira prema  $\xi$ . Naime, iz

$$|x_{n+1} - \xi| \leq \tau |x_n - \xi| \leq \dots \leq \tau^{n+1} |x_0 - \xi| < \tau^{n+1} \eta,$$

slijedi

$$\lim_{n \rightarrow \infty} |x_{n+1} - \xi| = 0.$$

□

<sup>6</sup>tj. za  $x_0$  iz dovoljno malene okoline od  $\xi$  niz (4.28) konvergira prema  $\xi$  kvadratnom brzinom.

## 4.6 Rješavanje sustava nelinearnih jednadžbi

Spomenimo na kraju mogućnost generalizacije prethodno navedenih metoda na rješavanje sustava nelinearnih jednadžbi

$$f_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, \dots, n, \quad (4.30)$$

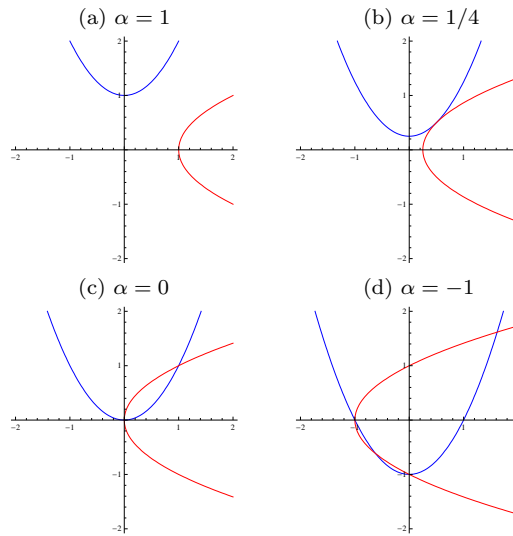
odnosno  $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ , gdje je  $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\mathbf{f}(\mathbf{x}) = (f_1(x), \dots, f_n(x))^T$ ,  $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ .

**Primjer 4.17.** Promatramo sustav nelinearnih jednadžbi (Ortega and Rheinboldt, 1970) za različite vrijednosti parametra  $\alpha \in [0, 1]$

$$f_1(x_1, x_2) \equiv x_1^2 - x_2 + \alpha = 0$$

$$f_2(x_1, x_2) \equiv -x_1 + x_2^2 + \alpha = 0$$

Kao što se može vidjeti na Slici 4.14 za različite vrijednosti parametra  $\alpha \in [0, 1]$  sustav može biti nerješiv ili može imati jedno, dva ili tri rješenja.



Slika 4.14: Grafički prikaz sustava nelinearnih jednadžbi

**Zadatak 4.7.** Grafički ispitajte skup svih rješenja sljedećih sustava nelinearnih jednadžbi.

$$a) \quad \begin{aligned} f_1(x_1, x_2) &\equiv \frac{1}{2}x_1 \sin\left(\frac{\pi}{2}x_1\right) - x_2 = 0 \\ f_2(x_1, x_2) &\equiv -x_1 + x_2^2 + 1 = 0 \end{aligned}$$

$$b) \quad \begin{aligned} f_1(x_1, x_2) &\equiv x_1^2 - |x_2| = 0 \\ f_2(x_1, x_2) &\equiv x_1^2 - x_2 = 0 \end{aligned}$$

**Primjer 4.18.** Za dani skup podataka  $(t_i, y_i)$ ,  $i = 1, \dots, m$  treba odrediti Optimalne parametre eksponencijalne funkcije  $t \mapsto b e^{ct}$  u smislu najmanjih kvadrata rješavajući optimizacijski problem (vidi primjerice Jukić and Scitovski (1997))

$$\operatorname{argmin}_{b, c \in \mathbb{R}} F(b, c), \quad F(b, c) = \sum_{i=1}^m (b e^{ct_i} - y_i)^2.$$

Kandidate za rješenje ovog problema (stacionarne točke funkcije  $F$ ) dobit ćemo rješavanjem sljedećeg sustava nelinearnih jednadžbi

$$\frac{1}{2} \frac{\partial F}{\partial b} \equiv \sum_{i=1}^m (b e^{ct_i} - y_i) e^{ct_i} = 0, \quad \frac{1}{2} \frac{\partial F}{\partial c} \equiv \sum_{i=1}^m (b e^{ct_i} - y_i) b t_i e^{ct_i} = 0.$$

**Primjer 4.19.** Za dani skup točaka  $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \dots, m\}$  treba odrediti točku  $c = (u, v) \in \mathbb{R}^2$  tako da suma euklidskih udaljenosti od svih točaka  $a^i$  do točke  $c$  bude minimalna. Ova točka naziva se geometrijski medijan (vidi primjerice Iyigun and Ben-Israel (2010)), a dobiva se kao rješenje sljedećeg optimizacijskog problema

$$\operatorname{argmin}_{u, v \in \mathbb{R}} F(u, v), \quad F(u, v) = \sum_{i=1}^m \sqrt{(u - x_i)^2 + (v - y_i)^2}.$$

Kandidate za rješenje ovog problema (stacionarne točke funkcije  $F$ ) dobit ćemo rješavanjem sljedećeg sustava nelinearnih jednadžbi

$$\frac{1}{2} \frac{\partial F}{\partial u} \equiv \sum_{i=1}^m \frac{u - x_i}{\sqrt{(u - x_i)^2 + (v - y_i)^2}} = 0, \quad \frac{1}{2} \frac{\partial F}{\partial v} \equiv \sum_{i=1}^m \frac{v - y_i}{\sqrt{(u - x_i)^2 + (v - y_i)^2}} = 0. \quad (4.31)$$

U nastavku kratko ćemo pokazati nekoliko najvažnijih metoda za rješavanje sustava nelinearnih jednadžbi i ilustrirati ih na jednostavnim primjerima.

#### 4.6.1 Metoda jednostavnih iteracija

Analogno metodi jednostavnih iteracija (t.4.3 str. 92) sustav (4.30) možemo zapisati u obliku

$$x_i = \varphi_i(x_1, x_2, \dots, x_n), \quad i = 1, \dots, n. \quad (4.32)$$

što upućuje na iterativni postupak metode jednostavnih iteracija (Ortega and Rheinboldt, 1970; Süli and Mayers, 2006)

$$x_i^{(k+1)} = \varphi_i(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}), \quad i = 1, \dots, n, \quad k = 0, 1, \dots, \quad (4.33)$$

gdje je  $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$  početna aproksimacija. Postoje uvjeti koje moraju ispunjavati funkcije  $\varphi_1, \dots, \varphi_n$  da bi iterativni proces (4.33) konvergirao prema rješenju sustava (Ortega and Rheinboldt, 1970; Süli and Mayers, 2006). Za ilustraciju napisat ćemo sustav jednadžbi iz *Primjera 4.19* u obliku (4.33)

$$u = \varphi(u, v), \quad v = \psi(u, v),$$

gdje je

$$\varphi(u, v) = \frac{\sum_{i=1}^m \frac{x_i}{\sqrt{(u-x_i)^2 + (v-x_i)^2}}}{\sum_{i=1}^m \frac{1}{\sqrt{(u-x_i)^2 + (v-x_i)^2}}}, \quad \psi(u, v) = \frac{\sum_{i=1}^m \frac{y_i}{\sqrt{(u-x_i)^2 + (v-x_i)^2}}}{\sum_{i=1}^m \frac{1}{\sqrt{(u-x_i)^2 + (v-x_i)^2}}},$$

Uz pogodan izbor početne aproksimacije  $(u_0, v_0)$  iterativni proces

$$u_{k+1} = \varphi(u_k, v_k), \quad v_{k+1} = \psi(u_k, v_k), \quad k = 0, 1, \dots$$

konvergirat će prema rješenju  $(u^*, v^*)$  sustava (4.31).

Ovaj iterativni proces u literaturi (vidi primjerice Iyigun and Ben-Israel (2010)) naziva se *Wiszfel'dov algoritam* i ima brojne primjene (Drezner and Hamacher, 2004).

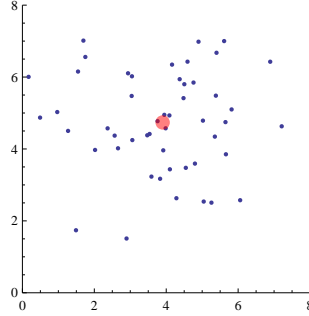
Na Slici 4.15 prikazano je 50 normalno distribuiranih podataka u okolini točke  $(4, 5)$  s varijancom  $\sigma^2 = 1.5$ :

```
Table[{4,5} + RandomReal[NormalDistribution[0, sig1], {2}], {i, 50}]
```

Uz izbor proizvoljne početne aproksimacije (koja se ne smije podudarati s nekom od točaka podataka !) nakon samo nekoliko iteracija Wiszfel'dov algoritam daje geometrijski medijan koji je također prikazan na slici.

Sustav iz *Primjera 4.18* ne bi se mogao jednostavno prilagoditi za metodu jednostavnih iteracija.

**Zadatak 4.8.** Za dani skup točaka  $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^n : i = 1, \dots, m\}$  treba odrediti točku – geometrijski medijan  $c \in \mathbb{R}^n$  tako da suma euklidskih udaljenosti od svih točaka  $a^i$  do točke  $c$  bude minimalna. Definirajte funkciju cilja i iterativni postupak za traženje geometrijskog medijana. Izradite odgovarajući program i primjere.



Slika 4.15: Geometrijski medijan

**Zadatak 4.9.** Za dani skup točaka  $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^n : i = 1, \dots, m\}$  treba odrediti točku  $c \in \mathbb{R}^n$  tako da

- suma  $l_1$  udaljenosti od svih točaka  $a^i$  do točke  $c$  bude minimalna;
- suma kvadrata euklidskih udaljenosti od svih točaka  $a^i$  do točke  $c$  bude minimalna;

Definirajte funkcije cilja. Mogu li se ovi problemi riješiti eksplicitno ?

### 4.6.2 Newtonova metoda

Newtonova metoda (t.4.4, str. 99) također se može generalizirati na rješavanje sustava nelinearnih jednadžbi (4.30). Najprije izaberimo početnu aproksimaciju  $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$  i svaku od funkcija  $f_i$  razvijemo u Taylorov red u okolini  $\mathbf{x}^{(0)}$ , te linearnu aproksimaciju označimo s  $\tilde{f}_i$ :

$$\tilde{f}_i(\mathbf{x}) = f_i(\mathbf{x}^{(0)}) + \sum_{j=1}^n \frac{\partial f_i(\mathbf{x}^{(0)})}{\partial x_j} (x_j - x_j^{(0)}), \quad i = 1, \dots, n. \quad (4.34)$$

Sada umjesto sustava (4.30) rješavamo sustav

$$\tilde{f}_i(\mathbf{x}) = 0, \quad i = 1, \dots, n,$$

što možemo pisati u matričnom obliku na sljedeći način

$$\mathbf{J}^{(0)} \mathbf{s} = -\mathbf{f}(\mathbf{x}^{(0)}), \quad (4.35)$$

$$\mathbf{J}^{(0)} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x}^{(0)})}{\partial x_1} & \dots & \frac{\partial f_n(\mathbf{x}^{(0)})}{\partial x_n} \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} x_1 - x_1^{(0)} \\ \vdots \\ x_n - x_n^{(0)} \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}^{(0)}) = \begin{bmatrix} f_1(\mathbf{x}^{(0)}) \\ \vdots \\ f_n(\mathbf{x}^{(0)}) \end{bmatrix}.$$

Matricu  $\mathbf{J}$  nazivamo Jacobijeva matrica ili Jacobijan sustava. Nova aproksimacija rješenja tada je

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)},$$

gdje je  $\mathbf{s}^{(0)}$  rješenje sustava (4.35). Općenito, dobivamo iterativni postupak

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}, \quad k = 0, 1, \dots \quad (4.36)$$

gdje je  $\mathbf{s}^{(k)}$  rješenje sustava

$$\mathbf{J}^{(k)} \mathbf{s} = -\mathbf{f}(\mathbf{x}^{(k)}). \quad (4.37)$$

Primijetite da se iterativni postupak (4.36)–(4.37) uz pretpostavku regularnosti Jacobijana  $\mathbf{J}^{(k)}$  može zapisati u obliku

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left(\mathbf{J}^{(k)}\right)^{-1} \mathbf{f}(\mathbf{x}^{(k)}), \quad k = 0, 1, \dots, \quad (4.38)$$

što podsjeća na Newtonov iterativni postupak (4.19), str. 99. Inače, iterativni proces (4.38) rijetko se koristi jer je postupak (4.36)–(4.37) numerički stabilniji (Truhar, 2010).

Niže ćemo konstruirati odgovarajući *Mathematica*-program. Najprije definirajmo broj varijabli  $n$ , funkcije  $f_1$  i  $f_2$ , njihov Jacobijan i nacrtajmo odgovarajuće krivulje  $f_1(x_1, x_2) = 0$  i  $f_2(x_1, x_2) = 0$ .

```
In[1]:= Clear[f, J, x1, x2]; n = 2;
f1[x1_, x2_] := x2*(x1-1)-1
f2[x1_, x2_] := x1^2-x2^2-1
pom[x_] := {f1[x[1], x[2]], f2[x[1], x[2]]};
f[x] := pom[x];
J[x_] := Table[D[f[x][[i]], x[j]], {i, n}, {j, n}]
Print["f(x)=", MatrixForm[f[x]/.{x[1]->x1, x[2]->x2}],
      ", J(x)=", MatrixForm[J[x]/.{x[1]->x1, x[2]->x2}]]
s1f = ContourPlot[{f1[x1, x2]==0, f2[x1, x2]==0},
{x1,-3,3}, {x2,-3,3}, PlotRange->{-2,2}, AspectRatio->.7,
ContourStyle->{{Blue, Thickness[.005]}, {Red, Thickness[.005]}}
```

Nakon toga definirajmo početnu aproksimaciju, primjerice

```
In[2]:= x0=Table[i, {i, n}];
```

Sljedeću aproksimaciju dobivamo kako slijedi

```
In[3]:= w=Table[x[i]->x0[[i]], {i,n}];
a = J[x]/.w ;
b = -f[x]/.w;
Print["x0=", N[x0], " f0= ", -N[b]]
s0 = LinearSolve[a, b]; x1 = x0 + s0;
x0 = x1;
```

Ponavljane linije In[3] dobivamo niz aproksimacija.

**Zadatak 4.10.** Izradi `Mathematica`-modul koji će pozivati broj varijabli  $n$ , funkciju  $f$ , točnost  $\epsilon > 0$ , početnu aproksimaciju  $x_0$ , te Newtonovom metodom davati niz aproksimacija rješenja jednadžbe  $f(x) = 0$ . Program testirajte na sljedećem primjeru.

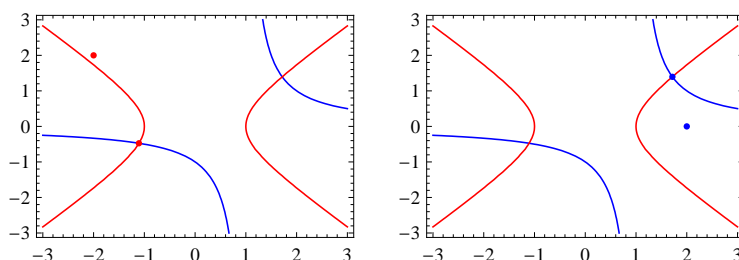
**Primjer 4.20.** Sljedeći sustav nelinearnih jednadžbi riješit ćemo Newtonovom metodom.

$$f_1(x_1, x_2) \equiv x_2(x_1 - 1) - 1 = 0, \quad f_2(x_1, x_2) \equiv x_1^2 - x_2^2 - 1 = 0.$$

U *Tablici 4.7* prikazane su iteracije Newtonove metode (zaokružene na 2 decimale) za dvije različite početne aproksimacije, a rezultati su vidljivi na *Slici 4.16*.

$n$	$x_n$	$f(x_n)$	$x_n$	$f(x_n)$
0	(-2.00, 2.00)	(-7.00, -1.00)	(2.00, 0.00)	(-1.00, 3.00)
1	(-0.75, 0.50)	(-1.88, -0.69)	(1.25, 1.00)	(-0.75, -0.44)
2	(-0.54, -0.51)	(-0.22, -0.97)	(1.86, 1.55)	(0.34, 0.08)
3	(-1.33, -0.39)	(-0.10, 0.61)	(1.73, 1.41)	(0.02, 0.00)
4	(-1.12, -0.46)	(-0.02, 0.04)	(1.72, 1.40)	(0.00, 0.00)
5	(-1.11, -0.47)	(0.00, 0.00)	-	-

Tablica 4.7: Newtonova metoda za sustav iz *Primjera 4.20*



Slika 4.16: Početna aproksimacija i rješenje sustava

Pitanje konvergencije spomenutih metoda nećemo razmatrati. Zainteresirani čitatelj to može pogledati u više knjiga iz područja numeričke analize spomenutih u *Primjedbi 4.8*, str. 119 ili u popisu referenci na kraju teksta.



### 4.6.3 Quasi-Newtonove metode

Vrlo važno mjesto među metodama za rješavanje sustava nelinearnih jednadžbi zauzimaju tzv. kvazi-Newtonove metode, koje je uveo BROYDEN (1961), a koje su nastale kao generalizacija metode sekanti (vidi t.4.4.1, str. 104):

(a) Broydenova metoda

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}, \quad k = 0, 1, \dots,$$

gdje je  $\mathbf{s}^{(k)}$  rješenje sustava

$$\mathbf{B}_k \mathbf{s} = -\mathbf{f}(\mathbf{x}^{(k)})$$

gdje su  $\mathbf{B}_k$  matrice koje uz dobar izbor početne aproksimacije  $\mathbf{B}_0$  sve više nalikuju Jacobijanu, a izračunavaju se iz rekurzivne formule:

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{(\mathbf{y}_k - \mathbf{B}_k \mathbf{s}_k) \mathbf{s}_k^T}{(\mathbf{s}_k, \mathbf{s}_k)}, \quad k = 0, 1, \dots,$$

gdje je

$$\mathbf{y}_k = \mathbf{f}(\mathbf{x}^{(k+1)}) - \mathbf{f}(\mathbf{x}^{(k)}), \quad \mathbf{s}_k = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$$

(b) Davidon-Fletcher-Powell (DFP) metoda

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{H}_k \mathbf{f}(\mathbf{x}^k), \quad k = 0, 1, \dots,$$

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{(\mathbf{y}_k, \mathbf{s}_k)} - \frac{\mathbf{H}_k \mathbf{y}_k \mathbf{y}_k^T \mathbf{H}_k}{(\mathbf{y}_k, \mathbf{H}_k \mathbf{y}_k)}, \quad k = 0, 1, \dots,$$

(c) Broyden-Fletcher-Goldfarb-Schano (BFGS) metoda

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{H}_k \mathbf{f}(\mathbf{x}^k), \quad k = 0, 1, \dots,$$

$$\mathbf{H}_{k+1} = \left( \mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{y}_k, \mathbf{s}_k} \right) \mathbf{H}_k \left( \mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{y}_k, \mathbf{s}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{y}_k, \mathbf{s}_k}, \quad k = 0, 1, \dots,$$

Primijetimo da je  $\text{rang}(\mathbf{B}_{k+1} - \mathbf{B}_k) = 1$ . Zbog toga se kaže da je Broydenova metoda metoda ranga 1. Ako je Jacobijan simetrična matrica, bolje je koristiti DFP ili BFGS metodu, u kojima su korekcije Jacobijana  $\mathbf{H}_k$  ranga 2. Može se pokazati da rekurzivne formule za generiranje korekcija  $\mathbf{H}_k$  kod DFP i BFGS metode čuvaju simetričnost i pozitivnu definitnost korekcije. Više detalja o ovim metodama može se naći kod Dennis and Schnabel (1996).

*Zadatak 4.11.* Izradite program za rješavanje sustava nelinearnih jednadžbi navedenim metodama i ispitajte ga na sljedećim primjerima

(a) (Dennis, Schnabel (1996))

$$2(x_1 + x_2)^2 + (x_1 - x_2)^2 - 8 = 0$$

$$5x_1^2 + (x_2 - 3)^2 - 9 = 0$$

$$\text{Rješenje: } \mathbf{x}_1^* = (1, 1)^T, \quad \mathbf{x}_2^* = (-1.18347, 1.58684)^T.$$

(b) (Dennis, Schnabel (1996))

$$\mathbf{f}(x_1, x_2) = (x_1^2 + x_2^2 - 2, e^{x_1-1} + x_2^3 - 2)^T = 0$$

Rješenje:  $\mathbf{x}^* = (1, 1)^T$ .

(c) (Dennis, Schnabel (1996))

$$\mathbf{f}(x_1, x_2, x_3) = (x_1, x_2^2 + x_2, e^{x_3} - 1)^T = 0$$

(d) *Rosenbrock parabolic valey*

$$\text{grad } f(x_1, x_2) = 0, \quad f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Rješenje:  $\mathbf{x}^* = (1, 1)^T$ .(e)  $\text{grad } f(x_1, x_2, x_3) = 0, \quad f(x_1, x_2, x_3) = \sum_{i=1}^{11} \left( x_3 \exp\left(-\frac{(t_i - x_1)^2}{x_2}\right) - y_i \right)^2$ ,  
gdje su brojevi  $(t_i, y_i)$ ,  $i = 1, \dots, 11$  zadani tablicom

$t_i$	0	1	2	3	4	5	6	7	8	9	10
$y_i$	.001	.01	.04	.12	.21	.25	.21	.12	.04	.01	.001

Rješenje:  $\mathbf{x}^* = (4.93, 2.62, 0.28)^T$ .

*Primjedba 4.8.* Više detalja o numeričkim metodama za rješavanje nelinearnih jednadžbi može se naći u brojnoj literaturi (Demidovich and Maron, 1981; Ortega and Rheinboldt, 1970; Schwarz and Köckler, 2011; Stoer, 2002; Stoer and Bulirsch, 2002; Süli and Mayers, 2006). FORTRAN kodovi mogu se naći u (?). O rješavanju sustava nelinearnih jednadžbi i problemima ekstrema funkcija više varijabli može se naći kod (Dennis and Schnabel, 1996; Ortega and Rheinboldt, 1970). Kvalitetna FORTRAN programska podrška nalazi se u *The NAG-Library*. Naravno, kvalitetan inženjerski “alat” za rješavanje ovakvih problema su programski sustavi *Mathematica* i *Matlab*.

## 4.7 Zadaci

*Zadatak 4.12.* Nekom iterativnom metodom riješite jednadžbu  $x^2 - 2x - 2 = 0$  s točnošću  $\epsilon = 0.5 \cdot 10^{-4}$ , tj. na 4 signifikantne decimale.

Rješenje:  $x_1 \approx 2.7321, x_2 \approx -0.7321$ .

*Zadatak 4.13.* Uz točnost  $\epsilon = 0.5 \times 10^{-6}$  odredite  $\sqrt[3]{2}$  rješavajući jednadžbu  $x^3 - 2 = 0$ .

Rješenje:  $\sqrt[3]{2} \approx 1.259921$ .

*Zadatak 4.14.* S točnošću  $\epsilon = 0.5 \cdot 10^{-4}$  riješite jednadžbu

a)  $x^3 - 2x - 2 = 0$ ,      b)  $\sin x - 5x + 1/2 = 0$ .

Rješenje: a)  $\xi \approx 1.7693$ ,      b)  $\xi \approx 0.12492$

*Zadatak 4.15.* S točnošću  $\epsilon = 0.5 \cdot 10^{-4}$  odredite sjecište grafova krivulja zadanih jednadžbama  $y = 1/(x - 1)$  i  $x^2 - y^2 = 1$ .

Rješenje:  $x_0 \approx 1.7167$ ,  $y_0 \approx 1.3953$ .

*Zadatak 4.16.* S točnošću  $\epsilon = 0.5 \cdot 10^{-4}$  odredite realne nultočke polinoma

$$P(x) = x^5 - 5x^4 - 4x^3 + 20x^2 - 5x + 25.$$

Rješenje:  $x_1 \approx -2.2361$ ,  $x_2 \approx 2.2361$ ,  $x_3 = 5$ .

*Zadatak 4.17.* S točnošću  $\epsilon = 0.5 \cdot 10^{-4}$  odredite najmanji pozitivni korijen jednadžbe  $x \sin x + 1 = 0$ .

Rješenje:  $x \approx 3.4368$ .

*Zadatak 4.18.* Jednadžba  $x^3 - x = 0$  ima tri korijena:  $-1, 0, 1$ . Ispitajte, prema kojem korijenu jednadžbe će konvergirati Newtonov iterativni proces (4.19, str. 99) ako se izabere proizvoljna početna aproksimacija  $x_0 \in \mathbb{R}$ ?

*Zadatak 4.19.* Grafičkom metodom odredite intervale u kojima funkcija  $f$  ima nultočke (DAHLQUIST (1972))

a)  $f(x) = 4 \sin x + 1 - x$ ,      b)  $f(x) = 1 - x - e^{-2x}$ ,      c)  $f(x) = (x + 1)e^{x-1}$ ,  
d)  $f(x) = x^4 - 4x^3 + 2x^2 - 8$ ,      e)  $f(x) = e^x + x^2 + x$ ,      f)  $f(x) = e^x - x^2 - 2x - 2$

Rješenje:

a)  $I_1 = [-3, -1], I_2 = [-1, 1], I_3 = [2, 3]$ ,      b)  $I_1 = [-.5, .5], I_2 = [.5, 1]$ ,  
c)  $I = [0, 1]$ ,      d)  $I_1 = [-2, 0], I_2 = [3, 4]$ ,  
e) nema nultočaka      f)  $I = [2, 4]$ .

*Zadatak 4.20.* Zadana je funkcija  $f$ . Separirajte intervale u kojima se nalaze njezine nultočke, za svaki od tih intervala odredite broj  $m_1$  iz ocjene (4.3), te metodom bisekcije ili metodom iteracija s točnošću  $\epsilon = 0.5 \times 10^{-4}$  odredite nultočke. Provjeri red konvergencije metode računajući  $|e_{n+1}|/|e_n|$  za svaki  $n$ .

a)  $f(x) = e^{-x} \sin(3x + 2) + x - 1/2$       b)  $f(x) = x \operatorname{th}(x) - x^2 + 1$   
c)  $f(x) = x^2 - 1 + \ln(x + 1)$   
d)  $f(x) = e^{-x} - 0.5 \cdot (x - 1)^2 + 1$       e)  $f(x) = \frac{\sin x}{x} - (x - 2)^3$ .

Rješenje:

a)  $I_1 = [-1, -0.4], m_1 \approx 3, x_1 \approx -0.4498, I_2 = [0, \frac{1}{2}], m_1 \approx 0.49, x_2 \approx 0.2832,$   
 $I_3 = [0.75, 1], m_1 \approx 0.79, x_3 \approx 0.9047$ .

- b)  $I_1 = [-2, -1]$ ,  $m_1 \approx 0.8$ ,  $x_1 \approx -1.5572$ ,  $I_2 = [1, 2]$ ,  $m_1 \approx 0.8$ ,  $x_2 \approx 1.5572$ .  
 c)  $I = [0, 1]$ ,  $m_1 \approx 1$ ,  $\xi \approx 0.6896$ .  
 d)  $I = [2, 3]$ ,  $m_1 \approx 1.14$ ,  $\xi \approx 2.4727$ .  
 e)  $I = [2, 3]$ ,  $m_1 \approx 0.5$ ,  $x_1 \approx 2.5879$ .

*Zadatak 4.21.* Funkcija  $f(x) = x + \ln x$  ima jednostruku realnu nultočku  $\xi$  na intervalu  $[.1, 1]$ . Koji od niže navedenih iterativnih procesa konvergira prema  $\xi$ ? Zašto?

$$\text{a) } x_{n+1} = -\ln x_n \quad \text{b) } x_{n+1} = e^{-x_n} \quad \text{c) } x_{n+1} = \frac{1}{2}(x_n + e^{-x_n})$$

*Zadatak 4.22.* Neka je  $\varphi : I = [a, b] \rightarrow \mathbb{R}$  neprekidno derivabilna funkcija, a jednačba  $x = \varphi(x)$  neka ima jedinstveno rješenje na  $I$ . Za proizvoljno izabrani  $x_0 \in I$  grafički prikažite dobivanje prve tri aproksimacije pomoću metode iteracija, ako je

- a)  $0 < \varphi'(x) < 1$ ,  $x \in I$     b)  $0 > \varphi'(x) > -1$ ,  $x \in I$   
 c)  $\varphi'(x) > 1$ ,  $x \in I$     d)  $\varphi'(x) < -1$ ,  $x \in I$ .

*Zadatak 4.23.* Neka funkcija  $f : I \rightarrow \mathbb{R}$  ima neprekidnu drugu derivaciju na intervalu  $I = [a, b]$  i neka u tom intervalu jednačba  $f(x) = 0$  ima jedinstveno rješenje  $\xi$ . Neka je nadalje,  $f(a) \cdot f(b) < 0$ , a prva ( $f'$ ) i druga ( $f''$ ) derivacija funkcije  $f$  na intervalu  $I$  imaju stalan predznak. Za izabrani  $x_0 \in I$  grafički prikažite dobivanje prve tri aproksimacije za sve četiri kombinacije predznaka prve ( $f'$ ) i druge ( $f''$ ) derivacija funkcije  $f$ :

- a) kod Newtonove metode tangenti,                      b) kod metode sekanti,  
 c) kod metode krivih položaja (regula falsi).

*Zadatak 4.24.* Rješenje jednačbe  $x = a/x$ ,  $a > 0$  je  $\xi = \pm\sqrt{a}$ . Pokažite da niz  $x_{n+1} = a/x_n$ ,  $n = 0, 1, \dots$  (formula (4.9) za funkciju  $\varphi(x) = a/x$ ) nije konvergentan ni za jedan izbor  $x_0 \in \mathbb{R}$ .

*Zadatak 4.25.* Rekurzivnom formulom zadan je niz (formula (4.19) za funkciju  $f(x) = x^2 - a$ ,  $a > 0$ )

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right), \quad x_0 > a, \quad n = 0, 1, \dots$$

Bez pozivanja na *Teorem 4.2* dokažite konvergenciju ovog niza te odredite graničnu vrijednost.

*Uputa:* Treba pokazati da je niz monotono padajući i ograničen odozdo.

*Zadatak 4.26.* Slično kao u *Primjeru 4.7*, str. 97 definirajte i analizirajte iterativne postupke za rješavanje jednačbe:

- a)  $f(x) = x^3 - a$ ,  $a \in \mathbb{R}$  (treći korijen realnog broja  $a$ ),  
 b)  $f(x) = x^k - a$ ,  $a > 0$ ,  $k \in \mathbb{N}$  ( $k$ -ti korijen pozitivnog broja  $a$ ).

*Uputa:* Koristite formulu (4.19), str. 99.

Rješenje: a)  $x_{n+1} = \frac{1}{3}(2x_n + \frac{a}{x_n^2})$ ; b)  $x_{n+1} = \frac{1}{k}((k-1)x_n + \frac{a}{x_n^{k-1}})$ .

*Zadatak 4.27.* Realni korijen jednadžbe  $x^3 - x - 4 = 0$  je  $x_0 = \sqrt[3]{2 + \frac{1}{9}\sqrt{321}} + \sqrt[3]{2 - \frac{1}{9}\sqrt{321}}$ . Izračunajte  $x_0$  s točnošću na 4 decimale iz ove formule, te također korištenjem Newtonove metode za  $x_0 = 2$ .

Rješenje:  $x_0 = 1.7963$ .

*Zadatak 4.28.* Za *Primjer 4.10*, str. 103 odredite visine svih nivoa nafte koji odgovaraju količinama 2%, 4%, 6%, ..., 100%. Izradite odgovarajuću tablicu.

*Zadatak 4.29.* Ako su u *Primjeru 4.10*, str. 103 poznate visine  $h$  razine nafte, izrazite odgovarajući volumen nafte kao funkciju od  $r$ ,  $l$ ,  $h$ . Za zadane visine razine nafte  $h : 2, 4, \dots, 100$  cm izračunajte odgovarajuće količine nafte. Izradite odgovarajuću tablicu.

Uputa: Uvrštavajući (\*\*) u (\*) dobivamo  $V$  kao funkciju od  $h$ .

*Zadatak 4.30.* Slično kao u *Primjeru 4.11*, str. 106 izvedite jednadžbe gibanja balističkog projektila uz dodatnu pretpostavku da je otpor zraka proporcionalan brzini projektila (konstantu proporcionalnosti označimo s  $c$ ). Izračunajte vrijeme leta  $T$  i domet  $X$  projektila ako je:  $\alpha = 45^\circ$ ,  $v_0 = 226.3 \text{ ms}^{-1}$ ,  $c = 10$ . Nacrtajte graf odgovarajuće balističke krivulje.

Rješenje:  $T \approx 8.7 \text{ s}$ ,  $X \approx 929 \text{ m}$ ,

$$x = cv_0 \cos \alpha \cdot (1 - e^{-t/c}), \quad y = (cv_0 \sin \alpha + gc^2) \cdot (1 - e^{-t/c}) - gct.$$

*Zadatak 4.31.* Ako je zadana početna brzina  $v_0$  balističkog projektila iz *Primjera 4.11*, str. 106, pronadite uvjet kojega moraju ispunjavati koordinate  $x_0$  i  $y_0$  točke  $T$  da bi se balističkim projektilom mogla pogoditi točka  $T$ .

Rješenje:  $4h^2 > 4hy_0 + x_0^2$ .

*Zadatak 4.32.* Kada će jednadžba (\*) iz *Primjera 4.11*, str. 106 imati jedno, a kada dva rješenja?

Rješenje: Jednadžba (\*) imat će jedno rješenje ako je  $y_0 \leq x_0$ .

*Zadatak 4.33.* Izradite program koji će lokalizirati nultočku neke funkcije  $f$ , a zatim primjenom metode bisekcije (str. 88) utvrditi dovoljno mali (zadane duljine  $\epsilon$ ) interval u kome se nalazi nultočka.

*Zadatak 4.34.* Provjerite da jednadžba:  $x - 2 \sin x = 0$  ima tri realne nultočke:  $\xi_1 = 0$ ,  $\xi_2 \in [\frac{\pi}{2}, 2]$ ,  $\xi_3 \in [-2, -\frac{\pi}{2}]$ . Pokažite da za proizvoljni  $x_0 \in [\frac{\pi}{2}, 2]$  niz definiran rekursivnom formulom

$$x_{k+1} = 2 \sin x_k, \quad k = 0, 1, \dots$$

konvergira prema  $\xi_2$ . Za koje  $x_0$  će Newtonova metoda konvergirati prema

$\xi_2$ ? Usporedite red konvergencije navedenih metoda. Izradite odgovarajuće programe.

*Zadatak 4.35.* Može li se Newtonova metoda primijeniti za rješavanje jednadžbe  $f(x) = 0$ , ako je:

$$\text{a) } f(x) = x^2 - 14x + 50; \quad \text{b) } f(x) = x^{1/3}; \quad \text{c) } f(x) = (x-3)^{1/2}, \quad x_0 = 4.$$

*Zadatak 4.36.* Pokažite da Newtonov iterativni postupak konvergira prema jedinstvenom rješenju jednadžbe  $e^{2x} + 3x + 2 = 0$  za proizvoljni  $x_0$ .

*Zadatak 4.37.* Može se pokazati da Newtonova metoda ima kvadratnu brzinu konvergencije u jednostrukoj, a linearnu u dvostrukoj nultočki. Ustanovite to na primjeru funkcije  $f(x) = x^3 - 3x + 2$  ( $\xi_1 = -2$  je jednostruka, a  $\xi_2 = 1$  dvostruka nultočka).

*Uputa:* U iterativnom procesu za svaki  $n$  računajte  $|e_{n+1}|/|e_n|^2$  za  $\xi_1 = -2$ , odnosno  $|e_{n+1}|/|e_n|$  za  $\xi_2 = 1$ .

*Zadatak 4.38.* Newtonov iterativni postupak (4.19) ima linearnu brzinu konvergencije ako je  $\xi$  dvostruka nultočka funkcije  $f$ . Pokaži da tada niz definiran s

$$x_{n+1} = x_n - 2 \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots$$

konvergira prema  $\xi$  kvadratnom brzinom konvergencije. Provjerite to na primjeru funkcije  $f(x) = x^3 - 3x + 2$ , kojoj je  $\xi = 1$  dvostruka nultočka.

*Zadatak 4.39.* Pokažite da niz definiran Newtonovim iterativnim postupkom:

- a) konvergira za  $f(x) = e^{2x} + 3x + 2$  i za proizvoljni  $x_0$ ;
- b) divergira za  $f(x) = x \cdot e^{-x}$  i  $x_0 = 2$ ;
- c) ne konvergira prema rješenju, a članovi niza se ponavljaju skoro ciklički ( $x_{k+4} \approx x_k$ ,  $k = 0, 1, \dots$ ) za  $f(x) = x^3 - x - 3$  i  $x_0 = -3$ ; ako izaberemo  $x_0 = 2$ , niz konvergira rješenju  $\xi = 1.671699881$ ;
- d) divergira u širem smislu za  $f(x) = \arctg x$  i  $x_0 = 1.45$ ; ako izaberemo  $x_0 = 0.5$ , niz konvergira prema rješenju  $\xi = 0$ .

Objasnite nastale fenomene i nacrtajte odgovarajuće slike.

*Zadatak 4.40.* Lokalni minimum neprekidne dovoljno "glatke" funkcije  $f$  može se tražiti Newtonovim iterativnim postupkom

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k), \quad k = 0, 1, \dots$$

za povoljno odabrani  $x_0$ . Izradite odgovarajući program i ispitajte ga na primjeru funkcije  $f(x) = x^3 - 2x - 5$ , koja postiže minimum u točki 0.8165.

*Zadatak 4.41.* U *Dodatku* je priložen program za traženje lokalnog minimuma funkcije  $F$ . Na početku programa unosi se analitički izraz minimizirajuće funkcije, učitavaju se granice intervala  $[XA, XB]$  u kome je lokaliziran minimum i željena točnost aproksimacije EPS. Program je napisan prema ideji izloženoj u knjizi (vidi također FORSYTHE (1974)):

R. P. BRENT, *Algorithms for minimization without derivatives*,  
Englewood Cliffs, N.J., Prentice-Hall, 1973.

Proces minimizacije obavlja se kombinacijom metode zlatnog reza i sukcesivne paraboličke aproksimacije minimizirajuće funkcije. Na početku programa ispituje se relativna preciznost računala (veličina PREC). Ispitajte navedeni program na nekoliko primjera, te rezultate komparirajte s metodom navedenom u *Zadatku 4.40*. Primijetite da navedena metoda ne zahtijeva ni prvu ni drugu derivaciju minimizirajuće funkcije.

## Poglavlje 5

# Aproksimacija funkcija

U t.2, str.21 opisan je problem *interpolacije* funkcije. Zadanu funkciju  $f$  zamijenili smo polinomom  $P$  (kao funkcijom čije se vrijednosti lako izračunavaju), tako da se funkcija  $f$  i polinom  $P$  podudaraju u zadanim točkama (čvorovima interpolacije). Sada bismo željeli istu funkciju  $f$  aproksimirati nekom jednostavnom – lako izračunljivom – funkcijom  $f^*$ , koja će danu funkciju  $f$  “dobro predstavljati” u čitavom području definicije. Pri tome ćemo kvalitetu aproksimacije  $f^*$  mjeriti “udaljenošću” funkcije  $f$  od njene aproksimacije  $f^*$

$$d(f, f^*) = \|f - f^*\|,$$

pri čemu je je norma  $\|\cdot\|$  najčešće zadana s

$$\begin{aligned} \|f\|_p &= \left( \int_a^b w(x) |f(x)|^p dx \right)^{1/p} && (L_p \text{ norma}, p \geq 1), \\ \|f\|_\infty &= \max_{a \leq x \leq b} w(x) |f(x)| && (\text{Čebiševljeva } L_\infty \text{ norma}), \end{aligned}$$

gdje je  $w(x) > 0$  tzv. težinska funkcija.

*Primjedba 5.1.* Primijetite da od svih navedenih normi jedino za  $L_2$ -normu možemo definirati skalarni produkt

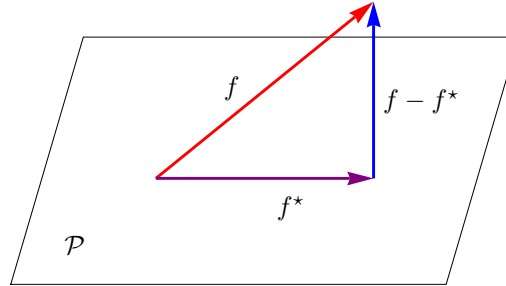
$$(f, g) = \int_a^b w(x) f(x) g(x) dx, \quad (5.1)$$

tako da je  $\|f\|_2^2 = (f, f)$ . Ako je  $w(x) = 1$ , onda  $d_2(f, f^*) = \|f - f^*\|_2$  predstavlja približnu površinu između grafova funkcija  $f$  i  $f^*$ .

Ako je  $\mathcal{P} \subset C[a, b]$  normirani potprostor nekih jednostavnih funkcija (primjerice polinoma), onda kažemo da je  $f^* \in \mathcal{P}$  najbolja aproksimacija funkcije  $f$  na potprostoru  $\mathcal{P}$  ako vrijedi

$$\|f - f^*\| \leq \|f - u\| \text{ za sve } u \in \mathcal{P}$$





Slika 5.1: Najbolja aproksimacija

Geometrijski gledano,  $f^*$  treba biti ortogonalna projekcija vektora  $f$  na potprostor  $\mathcal{P}$  (Slika 5.1). Ujedno to znači da je  $f' = f - f^*$  okomit na  $\mathcal{P}$ , tj. da vrijedi

$$(f - f^*, u) = 0 \text{ za sve } u \in \mathcal{P}. \quad (5.2)$$

Može se pokazati da ako je  $\mathcal{P}$  konačno dimenzionalni normirani vektorski potprostor, onda za svaku funkciju  $f \in C[a, b]$  postoji (barem jedna) najbolja aproksimacija  $f^* \in \mathcal{P}$ .

Posebno ćemo razmotriti problem najbolje  $L_2$  i najbolje  $L_\infty$  aproksimacije funkcije  $f$  definirane na intervalu  $[a, b]$  ili samo na konačnom skupu točaka iz tog intervala.

## 5.1 Najbolja $L_2$ aproksimacija

Pretpostavimo da je  $f \in C[a, b]$  i da je na  $C[a, b]$  definiran skalarni produkt (5.1) i odgovarajuća inducirana norma  $\|\cdot\|$ . Neka je nadalje  $\varphi_0, \varphi_1, \dots, \varphi_n \in C[a, b]$  sustav linearno nezavisnih funkcija. Promatramo problem aproksimacije funkcije  $f$  na potprostoru  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$ :

– postoji li  $a^* \in \mathbb{R}^{n+1}$ , takav da bude  $F(a^*) = \min_{a \in \mathbb{R}^{n+1}} F(a)$ , gdje je

$$F(a) = \left\| \sum_{i=0}^n a_i \varphi_i - f \right\|^2 = \int_a^b w(x) \left( \sum_{i=0}^n a_i \varphi_i(x) - f(x) \right)^2 dx? \quad (5.3)$$

Nužan uvjet egzistencije vektora  $a^* \in \mathbb{R}^{n+1}$  je  $\text{grad } F(a^*) = 0$ , tj.

$$\begin{aligned} (\varphi_0, \varphi_0)a_0^* + (\varphi_0, \varphi_1)a_1^* + \dots + (\varphi_0, \varphi_n)a_n^* &= (\varphi_0, f) \\ (\varphi_1, \varphi_0)a_0^* + (\varphi_1, \varphi_1)a_1^* + \dots + (\varphi_1, \varphi_n)a_n^* &= (\varphi_1, f) \\ \dots & \\ (\varphi_n, \varphi_0)a_0^* + (\varphi_n, \varphi_1)a_1^* + \dots + (\varphi_n, \varphi_n)a_n^* &= (\varphi_n, f) \end{aligned} \tag{5.4}$$

Sustav (5.4) nazivamo sustav normalnih jednadžbi. Matrica ovog sustava je tzv. Grammova matrica linearno nezavisnih funkcija  $\varphi_0, \varphi_1, \dots, \varphi_n$ , pa je kao takva pozitivno definitna (vidi primjerice (Kurepa, 1967)). Istovremeno ta je matrica do na pozitivnu konstantu jednaka Hessijanu funkcije  $F$ . Dakle, funkcija  $F$  na skupu  $\mathbb{R}^{n+1}$  postiže jedinstveni globalni minimum  $a^*$ , koji možemo dobiti rješavanjem sustava (5.4). Najbolja aproksimacija  $f^*$  funkcije  $f$  na potprostoru  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$  je prema tome  $f^* = a_0^*\varphi_0 + a_1^*\varphi_1 + \dots + a_n^*\varphi_n$ .

**Primjer 5.1.** Treba pronaći najbolju  $L_2$  aproksimaciju funkcije  $f: [0, 1] \rightarrow \mathbb{R}$ ,  $f(x) = e^{2x}$  na potprostoru svih polinoma stupnja  $\leq 1$  (uz  $w(x) = 1$ ).

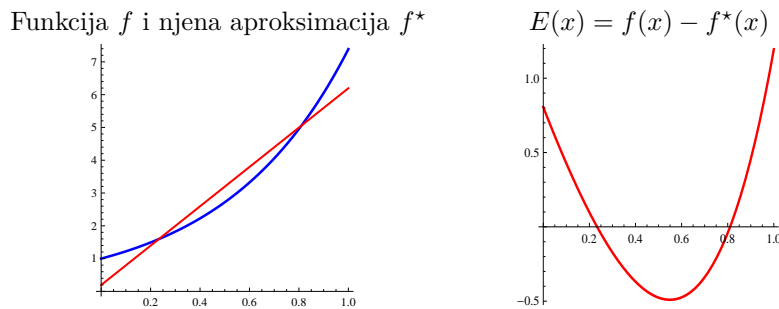
Aproksimaciju  $f^*$  tražit ćemo u obliku

$$x \mapsto a_0\varphi_0(x) + a_1\varphi_1(x), \quad \varphi_i(x) = x^i, \quad i = 0, 1.$$

Lako se vidi da su u ovom slučaju koeficijenti matrice sustava (5.4) zadani s

$$(\varphi_i, \varphi_j) = \int_0^1 \varphi_i(x)\varphi_j(x)dx = \int_0^1 x^{i+j}dx = \frac{1}{i+j+1}, \quad i, j = 0, 1,$$

a vektor slobodnih koeficijenata  $(\frac{1}{2}(e^2 - 1), \frac{1}{4}(e^2 + 1))^T = (3.19453, 2.09726)^T$ . Rješenje sustava (5.4) je  $a_0^* = 0.194528$ ,  $a_1^* = 6.0$ .



Slika 5.2: Najbolja aproksimacija funkcije  $f(x) = e^{2x}$

Na Slici 5.2 prikazani su grafovi funkcija  $f$  i  $f^*$ , te graf funkcije pogreške  $E(x) = f(x) - f^*(x)$ .

Primijetite da je matrica sustava u ovom primjeru poznata Hilbertova matrica. Ona se u literaturi obično uzima kao primjer loše uvjetovane matrice. Primjerice, za  $n = 10$ , uvjetovanost Hilbertove matrice je  $2 \times 10^7$ . Zato je traženje najbolje  $L_2$  aproksimacije rješavanjem sustava (5.4) često puta nepouzđano.

*Primjedba 5.2.* Do sustava (5.4) formalno možemo doći jednostavnije. Vektor pogreške aproksimacije  $f' = f - f^*$ , tj.

$$f'(x) = f(x) - \sum_{i=1}^m a_i \varphi_i(x).$$

je okomit na potprostor  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$  (vidi Sliku 5.1). Zato nakon skalarne množenja s  $\varphi_j$  za  $j = 0, 1, \dots, n$  dobivamo

$$0 = (f, \varphi_j) - \sum_{i=1}^m a_i (\varphi_i, \varphi_j), \quad j = 0, 1, \dots, n.$$

Ako bi funkcije  $\varphi_0, \varphi_1, \dots, \varphi_n$  činile ortogonalan sustav funkcija, tj. ako bi bilo

$$(\varphi_i, \varphi_j) = 0, \quad i \neq j \quad \|\varphi_i\| \neq 0 \quad \text{za sve } i = 0, 1, \dots, n,$$

onda bi matrica sustava (5.4) bila dijagonalna, a rješenje bi mogli eksplicitno zapisati

$$a_i^* = \frac{(f, \varphi_i)}{(\varphi_i, \varphi_i)}, \quad i = 0, 1, \dots, n. \quad (5.5)$$

Jedan takav ortogonalni sustav funkcija razmatra se u sljedećoj točki.

### 5.1.1 Fourierovi polinomi

Kao primjer ortogonalnog sustava na  $C[-\pi, \pi]$  uz  $w(x) = 1$  navodimo sustav trigonometrijskih funkcija

$$\frac{1}{2}, \sin x, \cos x, \dots, \sin nx, \cos nx. \quad (5.6)$$

Može se pokazati da vrijedi

**Lema 5.1.** *Sustav trigonometrijskih funkcija (5.6) na  $C[-\pi, \pi]$  sa skalar-nim produktom  $(f, g) = \int_{-\pi}^{\pi} f(x)g(x)dx$  je ortogonalan sustav, pri čemu su odgovarajuće norme funkcija jednake*

$$\|\frac{1}{2}\| = \sqrt{\frac{\pi}{2}}, \quad \|\sin kx\| = \|\cos kx\| = \sqrt{\pi}, \quad k = 1, \dots, n.$$

**Dokaz.** Lako se može pokazati da vrijedi

$$\begin{aligned}
 (i) \quad & \int_{-\pi}^{\pi} 1 \, dx = 2\pi \\
 (ii) \quad & \int_{-\pi}^{\pi} \sin \mu x \, dx = \int_{-\pi}^{\pi} \cos \mu x \, dx = 0, \quad \mu = 1, 2, \dots \\
 (iii) \quad & \int_{-\pi}^{\pi} \cos \mu x \cos \nu x \, dx = \frac{1}{2} \int_{-\pi}^{\pi} (\cos(\mu + \nu)x + \cos(\mu - \nu)x) \, dx = \begin{cases} 0, & \mu \neq \nu \\ \pi, & \mu = \nu \end{cases} \\
 (iv) \quad & \int_{-\pi}^{\pi} \cos \mu x \sin \nu x \, dx = \frac{1}{2} \int_{-\pi}^{\pi} (\sin(\mu + \nu)x + \sin(\mu - \nu)x) \, dx = 0 \\
 (v) \quad & \int_{-\pi}^{\pi} \sin \mu x \sin \nu x \, dx = \frac{1}{2} \int_{-\pi}^{\pi} (\cos(\mu + \nu)x - \cos(\mu - \nu)x) \, dx = \begin{cases} 0, & \mu \neq \nu \\ \pi, & \mu = \nu \end{cases}
 \end{aligned}$$

iz čega slijedi navedena tvrdnja.  $\square$

Primijetite da *Lema 5.1* ne isključuje mogućnost da sustav trigonometrijskih funkcija (5.6) bude beskonačan.

Potražimo najbolju  $L_2$  aproksimaciju funkcije  $f \in C[-\pi, \pi]$  na potprostoru s bazom (5.6). Aproksimaciju  $f^*$  funkcije  $f$  tražit ćemo u obliku trigonometrijskog polinoma stupnja  $n$

$$T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx), \quad (5.7)$$

pri čemu parametre  $a = (a_0, a_1, \dots, a_n)$ ,  $b = (b_1, \dots, b_n)$  određujemo rješavanjem sljedećeg optimizacijskog problema

$$\operatorname{argmin}_{a \in \mathbb{R}^{n+1}, b \in \mathbb{R}^n} F(a, b), \quad F(a, b) = \int_{-\pi}^{\pi} (T_n(x; a, b) - f(x))^2 \, dx. \quad (5.8)$$

Izjednačavajući sve parcijalne derivacije funkcije  $F$  s nulom ili postupajući kao u Primjedbi 5.1 i koristeći *Lemu 5.1* lako dobivamo

$$\begin{aligned}
 a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx, \quad k = 0, 1, \dots, n, \\
 b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx, \quad k = 1, \dots, n,
 \end{aligned} \quad (5.9)$$

Parametre  $a_k, b_k$  iz (5.9) zovemo **Fourierovi koeficijenti**<sup>1</sup> funkcije  $f$ , a trigonometrijski polinom  $T_n$  s Fourierovim koeficijentima (5.9) zovemo **Fourierov polinom** za funkciju  $f$  i označavamo s  $F_n$ .

<sup>1</sup>Jean Baptiste Fourier (1768 – 1830) francuski fizičar i matematičar

**Zadatak 5.1.** Pokažite da svi Fourierovi koeficijenti  $a_k$ ,  $k = 0, 1, \dots$  neparne funkcije  $f: [-\pi, \pi] \rightarrow \mathbb{R}$  iščezavaju i da svi Fourierovi koeficijenti  $b_k$ ,  $k = 1, 2, \dots$  parne funkcije  $f: [-\pi, \pi] \rightarrow \mathbb{R}$  iščezavaju.

**Primjer 5.2.** Za funkciju  $f: [-\pi, \pi] \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\pi}|x|$  treba odrediti Fourierove koeficijente i Fourierov polinom stupnja  $n \geq 1$ .

Budući da je  $f$  parna funkcija, funkcija  $x \mapsto f(x) \sin kx$  je neparna, pa svi koeficijenti  $b_1, \dots, b_n$  iščezavaju, a koeficijenti  $a_k$  su

$$a_0 = \frac{1}{2}, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx \, dx = \begin{cases} 0, & k \text{ paran} \\ -\frac{4}{k^2\pi^2} & k \text{ neparan} \end{cases}$$

Zato odgovarajući Fourierov polinom glasi

$$F_n(x) = \frac{1}{2} - \frac{4}{\pi^2} \left( \cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \dots \right)$$

**Primjer 5.3.** Za funkciju  $f: [-1, 1] \rightarrow \mathbb{R}$ ,  $f(x) = |x|$  treba odrediti Fourierove koeficijente i Fourierov polinom.

Najprije ćemo zadanu funkciju  $f: [-1, 1] \rightarrow \mathbb{R}$  “rastegnuti” na interval  $[-\pi, \pi]$  i za tako dobivenu funkciju  $\tilde{f}$  odrediti Fourierov polinom  $\tilde{F}_n$ . Nakon toga Fourierov polinom “vratit ćemo” nazad na interval  $[-1, 1]$ .

Pomoću funkcije  $u: [-\pi, \pi] \rightarrow [-1, 1]$ ,  $u(x) = \frac{1}{\pi}x$ , zadanu funkciju  $f$  transformirat ćemo u funkciju  $\tilde{f}: [-\pi, \pi] \rightarrow \mathbb{R}$ ,  $\tilde{f}(x) = (f \circ u)(x)$ , tj.

$$\tilde{f}(x) = (f(u(x))) = f\left(\frac{1}{\pi}x\right) = \frac{1}{\pi}|x|, \quad x \in [-\pi, \pi].$$

Prema prethodnom primjeru Fourierov polinom funkcije  $\tilde{f}$  glasi

$$\tilde{F}_n(x) = \frac{1}{2} - \frac{4}{\pi^2} \left( \cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \dots \right),$$

a Fourierov polinom funkcije  $f$  tada je

$$F_n(x) = (\tilde{F}_n \circ u^{-1})(x) = (\tilde{F}_n(\pi x)) = \frac{1}{2} - \frac{4}{\pi^2} \left( \cos \pi x + \frac{1}{3^2} \cos 3\pi x + \frac{1}{5^2} \cos 5\pi x + \dots \right).$$

**Zadatak 5.2.** Odredite linearnu funkciju koja interval  $[-\pi, \pi]$  preslikava na interval  $[a, b]$ . Kako glasi linearna funkcija koja interval  $[a, b]$  preslikava na interval  $[-\pi, \pi]$ ?

**Zadatak 5.3.** Pokažite da je sustav funkcija  $\varphi_i(x) = \cos ix$ ,  $i = 0, 1, \dots, n$ , ortogonalan na  $[0, \pi]$  uz  $w(x) = 1$  i da je

$$\|\varphi_0\| = \sqrt{\pi}, \quad \|\varphi_i\| = \sqrt{\frac{\pi}{2}}, \quad i = 1, \dots, n.$$

*Zadatak 5.4.* Pokažite da sustav funkcija  $\varphi_i(x) = \cos ix$ ,  $i = 0, 1, \dots, n$ , definiran na  $D = \left\{x_i = \frac{2i+1}{n+1} \frac{\pi}{2} : i = 0, 1, \dots, n\right\} \subset [0, \pi]$  čini ortogonalan sustav sa skalarnim produktom

$$(f, g) = \sum_{i=0}^n f(x_i)g(x_i),$$

i da je  $\|\varphi_0\| = \sqrt{n+1}$ ,  $\|\varphi_i\| = \sqrt{\frac{n+1}{2}}$ ,  $i = 1, \dots, n$ .

*Primjedba 5.3.* Ako su  $f, g$  ortogonalne funkcije, onda za njih vrijedi Pita-gorin teorem

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2. \quad (5.10)$$

Općenitije, za ortogonalan sustav  $\varphi_0, \varphi_1, \dots, \varphi_n$  vrijedi

$$\left\| \sum_{i=0}^n a_i \varphi_i \right\|^2 = \sum_{i=0}^n a_i^2 \|\varphi_i\|^2.$$

Nadalje, kako je  $f - f^*$  okomit na  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$  i kako je  $f^* = \sum_{i=0}^n a_i^* \varphi_i$ , onda je  $(f - f^*, f^*) = 0$  i

$$\|f - f^*\|^2 = \|f\|^2 - 2(f, f^*) + \|f^*\|^2 = \|f\|^2 - \|f^*\|^2 = \|f\|^2 - \sum_{i=0}^n (a_i^*)^2 \|\varphi_i\|^2. \quad (5.11)$$

Iz (5.11) izravno slijedi poznata Besselova nejednakost

$$\sum_{i=0}^n (a_i^*)^2 \|\varphi_i\|^2 \leq \|f\|^2. \quad (5.12)$$

*Primjedba 5.4.* Budući da je sustav funkcija iz (5.6), str. 128, ortogonalan za svaki  $n \in \mathbb{N}$ , možemo promatrati i beskonačni ortogonalni sustav funkcija

$$\frac{1}{2}, \sin x, \cos x, \dots, \sin nx, \cos nx, \dots,$$

i neku funkciju  $f$  pokušati prikazati pomoću sume reda

$$\mathcal{F}(x) := \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx, \quad (5.13)$$

gdje je prema (5.5)

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx, \quad (5.14)$$

Red (5.13) nazivamo Fourierov red, a brojeve  $a_0, a_1, b_1, \dots$  Fourierovi koeficijenti funkcije  $f$ . Kada će red (5.13) za neki  $x_0$  biti konvergentan i hoće li u tom slučaju biti  $\mathcal{F}(x_0) = f(x_0)$  određeno je *Dirichletovim teoremom* (vidi primjerice (Kurepa, 1990)).

**Teorem 5.1.** (*Dirichlet*) Neka je  $f : [-\pi, \pi] \rightarrow \mathbb{R}$  periodična funkcija temeljnog perioda  $2\pi$ , takva da

- (i) na segmentu  $[-\pi, \pi]$  ima najviše konačno mnogo prekida prve vrste,
- (ii) na segmenta  $[-\pi, \pi]$  najviše konačnomnogo puta mijenja smisao monotonosti.

Tada

- (i) Fourierov red funkcije  $f$  konvergira za svaki  $x \in \mathbb{R}$ . Neka je

$$F(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx,$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx, \quad k = 0, 1, \dots$$

- (ii) Ako je funkcija  $f$  neprekidna u točki  $x \in [-\pi, \pi]$ , onda je  $F(x) = f(x)$ ,
- (iii) Ako funkcija  $f$  ima prekid u točki  $x \in [-\pi, \pi]$ , onda je  $F(x) = \frac{1}{2}(f(x+) + f(x-))$
- (iv)  $F(-\pi) = F(\pi) = \frac{1}{2}(f(-\pi+) + f(\pi-))$

Kako je  $F(x) = f(x)$  za sve  $x \in \mathbb{R}$  osim možda za njih konačno mnogo, obično pišemo

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx,$$

i kažemo da je funkcija  $f$  razvijena u **Fourierov red**.

Kao ilustraciju pokazat ćemo da Fourierov red konvergira za svaki  $x \in \mathbb{R}$  u specijalnom slučaju ako je  $f \in C^2[-\pi, \pi]$  (vidi također (Kurepa, 1990)).

Ocijenimo najprije veličinu Fourierovih koeficijenata  $a_k, b_k$ . Primjenom metode parcijalne integracije uz  $u = f(x)$  i  $dv = \cos kx \, dx$  dobivamo

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx = \frac{1}{k\pi} [f(x) \sin kx]_{-\pi}^{\pi} - \frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) \frac{1}{k} \sin kx \, dx.$$

Prvi član iščezava, a u drugom ponovo primijenimo metodu parcijalne integracije uz  $u = f'(x)$  i  $dv = \frac{1}{k} \sin kx \, dx$ . Dobivamo

$$a_k = \frac{1}{k\pi} [f'(x) \cos kx]_{-\pi}^{\pi} - \frac{1}{k^2\pi} \int_{-\pi}^{\pi} f''(x) \cos kx \, dx.$$

Zbog neprekidnosti i periodičnosti funkcije  $f'$  prvi član iščezava.

Kako je funkcija  $f''$  neprekidna na  $[-\pi, \pi]$ , ona je i ograničena, pa postoji  $M > 0$ , tako da bude  $|f''(x)| < M$  za sve  $x \in [-\pi, \pi]$ . Zato vrijedi

$$|a_k| \leq \frac{1}{k^2\pi} \int_{-\pi}^{\pi} |f''(x)| |\cos kx| dx < \frac{2M}{k^2}.$$

Analogno dobivamo i ocjenu  $|b_k| < \frac{2M}{k^2}$ .

Na taj način za Fourierov red pronašli smo konvergentnu majorantu<sup>2</sup>

$$\frac{|a_0|}{2} + 2M \left(1 + 1 + \frac{1}{2^2} + \frac{1}{2^2} + \dots\right) = \frac{|a_0|}{2} + 4M \left(1 + \frac{1}{2^2} + \dots\right).$$

*Primjedba 5.5.* Fourierov red (5.13) može se prikazati i u kompleksnom obliku:

$$\mathcal{F}(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}, \quad \alpha_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \quad (5.15)$$

*Zadatak 5.5.* Pokažite da između Fourierovih koeficijenata  $(a_k, b_k)$  iz razvoja (5.13) i Fourierovih koeficijenata  $\alpha_k$  iz razvoja (5.15) postoji sljedeća veza

$$a_0 = 2\alpha_0, \quad a_k \cos kx + b_k \sin kx = \alpha_k e^{ikx} + \alpha_{-k} e^{-ikx}, \quad k = 1, 2, \dots$$

*Zadatak 5.6.* Neka je  $f : \mathbb{R} \rightarrow \mathbb{R}$  neprekidna periodična funkcija s temeljnim periodom  $L > 0$ .

1. Pokažite da sustav funkcija

$$\frac{1}{2}, \sin \frac{2\pi x}{L}, \cos \frac{2\pi x}{L}, \dots, \sin \frac{2n\pi x}{L}, \cos \frac{2n\pi x}{L} \quad (5.16)$$

čini ortogonalan sustav funkcija na  $C[0, L]$  uz  $w(x) = 1$ .

2. Od sustava (5.16) načinite ortonormiran sustav funkcija na  $C[0, L]$ . Odredite najbolju  $L_2$  aproksimaciju funkcije  $f \in C[0, L]$  na potprostoru razapetom dobivenim ortonormiranim funkcijama (odgovarajuće Fourierove koeficijente označite s  $a_k$  i  $b_k$ ).
3. Fourierov polinom funkcije  $f$  može se zapisati i u kompleksnom obliku

$$F_n(x) = \sum_{k=-n}^n c_k \exp\left(\frac{2k\pi ix}{L}\right) \quad (5.17)$$

Uspostavite vezu između Fourierovih koeficijenata  $a_0, a_1, \dots, a_n, b_1, \dots, b_n$  i Fourierovih koeficijenata  $c_k$  u prikazu (5.17).

---

<sup>2</sup>harmonijski red  $\sum \frac{1}{k^s}$ ,  $s > 1$ , vidi primjerice Jukić and Scitovski (2004)



Zadatak 5.7. Odredite Fourierove redove za sljedeće funkcije

$$\begin{aligned}
 a) \quad f(x) &= \begin{cases} 1, & -\frac{\pi}{2} < x < \frac{\pi}{2} \\ -1, & \frac{\pi}{2} < x < \frac{3\pi}{2} \end{cases} & b) \quad f(x) &= \begin{cases} -1, & -\pi < x < 0 \\ 1, & 0 < x < \pi \end{cases} \\
 c) \quad f(x) &= \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ -1, & \frac{1}{2} < x < \frac{3}{2} \end{cases} & d) \quad f(x) &= \begin{cases} 0, & -1 < x < 1 \\ 1, & 1 < x < 3 \\ \frac{1}{2}, & x \in \{1, 3\} \end{cases} \\
 e) \quad f(x) &= \begin{cases} x, & -1 < x < 1 \\ 0, & x = 1 \end{cases} & f) \quad f(x) &= \begin{cases} x, & -1 < x < 0 \\ 1-x, & 0 < x < 1 \\ \frac{1}{2}, & x = 0 \\ -\frac{1}{2}, & x = 1 \end{cases}
 \end{aligned}$$

Rješenje:

$$\begin{aligned}
 a) \quad \frac{4}{\pi} \left( \cos x - \frac{\cos 3x}{3} + \frac{\cos 5x}{5} - \dots \right) & \quad b) \quad \frac{4}{\pi} \left( \sin x - \frac{\sin 3x}{3} + \frac{\sin 5x}{5} - \frac{\sin 7x}{7} + \dots \right) \\
 c) \quad \frac{4}{\pi} \left( \cos \pi x - \frac{\cos 3\pi x}{3} + \frac{\cos 5\pi x}{5} - \dots \right) & \quad d) \quad \frac{4}{\pi\sqrt{2}} \left( \sin \frac{\pi x}{4} - \frac{1}{3} \sin \frac{3\pi x}{4} - \frac{1}{5} \sin \frac{5\pi x}{4} + \dots \right) \\
 e) \quad \frac{2}{\pi} \left( \frac{\sin \pi x}{1} - \frac{\sin 2\pi x}{2} + \frac{\sin 3\pi x}{3} - \dots \right) & \\
 f) \quad a_k = \frac{4}{k^2\pi^2}, \quad k = 1, 3, 5, \dots, & \quad b_k = \begin{cases} \frac{2}{k\pi}, & k = 1, 3, 5, \dots \\ 0, & k = 2, 4, 6, \dots \end{cases}
 \end{aligned}$$

**Primjer 5.4.** Iz Eulerovog identiteta  $e^{ix} = \cos x + i \sin x$  i binomne formule slijedi:

$$\begin{aligned}
 \cos nx &= \cos^n x - \binom{n}{2} \cos^{n-2} x \sin^2 x + \binom{n}{4} \cos^{n-4} x \sin^4 x - \dots \\
 \sin nx &= \binom{n}{1} \sin x \cos^{n-1} x - \binom{n}{3} \sin^3 x \cos^{n-3} x
 \end{aligned}$$

**Primjer 5.5.** Slično iz

$$\cos x = \frac{1}{2} (e^{ix} + e^{-ix}) \quad i \quad \sin x = \frac{1}{2i} (e^{ix} - e^{-ix})$$

dobivamo "Fourierov razvoj" za funkcije  $\cos^n x$ ,  $\sin^n x$   $\cos^m x \sin^n$ . Primjerice

$$\cos^3 x = \frac{1}{8} (e^{ix} + e^{-ix})^3 = \frac{1}{8} (e^{3ix} + e^{-3ix} + 3e^{ix} + 3e^{-ix}) = \frac{1}{4} (\cos 3x + 3 \cos x),$$

i općenito:

$$\cos^n x = \frac{1}{2^n} \left( \cos nx + \binom{n}{1} \cos(n-2)x + \binom{n}{2} \cos(n-4)x + \dots + \binom{n}{n} \cos(n-2n)x \right)$$

*Zadatak 5.8.* Izvedite “Fourierov razvoj” za funkciju  $\sin^n x$  posebno za parni, a posebno za neparni  $n$ .

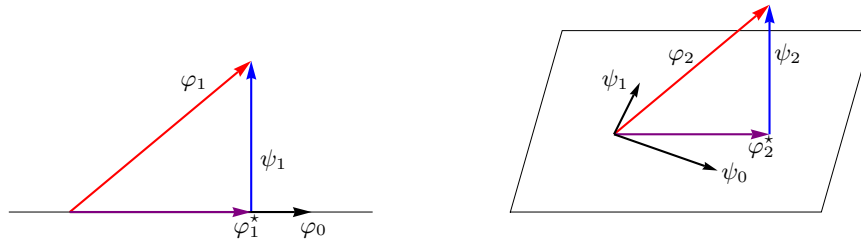
*Zadatak 5.9.* Pokažite da vrijedi:

$$\cos^3 x \sin^3 x = \frac{3}{32} \sin 2x - \frac{1}{32} \sin 6x$$

### 5.1.2 Ortogonalni polinomi

Pretpostavimo da je na vektorskom prostoru  $C[a, b]$  definiran skalarni produkt (5.1) s težinskom funkcijom  $w$ . Neka je  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$  vektorski potprostor razapet linearno nezavisnim vektorima (funkcijama)  $\varphi_0, \varphi_1, \dots, \varphi_n$ . U tom potprostoru definirat ćemo novu ortogonalnu bazu  $\psi_0, \psi_1, \dots, \psi_n$  standardnim Gram-Schmidtovim postupkom ortogonalizacije (Kurepa, 1967).

- (a) Projekcija vektora  $\varphi_1$  na pravac  $L(\varphi_0)$       (b) Projekcija vektora  $\varphi_2$  na ravninu  $L(\varphi_0, \varphi_1)$



Slika 5.3: Gram-Schmidtov postupak ortogonalizacije

Neka je  $\varphi_1^*$  projekcija vektora  $\varphi_1$  u smjeru vektora  $\varphi_0$  i neka je  $\psi_1 = \varphi_1 - \varphi_1^*$ . Tada postoji  $\alpha \in \mathbb{R}$  takav da je (vidi Sliku 5.3a)

$$\psi_1 = \varphi_1 - \alpha\varphi_0.$$

Množeći skalarno ovu jednadžbu vektorom  $\varphi_0$  dobivamo

$$0 = (\psi_1, \varphi_0) = (\varphi_1, \varphi_0) - \alpha(\varphi_0, \varphi_0),$$

iz čega uz oznaku  $\psi_0 := \varphi_0$  dobivamo

$$\alpha = \frac{(\varphi_1, \psi_0)}{(\psi_0, \psi_0)}, \quad \text{i} \quad \psi_1 = \varphi_1 - \frac{(\varphi_1, \psi_0)}{(\psi_0, \psi_0)}\psi_0. \quad (5.18)$$

Nadalje, neka je  $\varphi_2^*$  projekcija vektora  $\varphi_2$  na ravninu  $L(\psi_0, \psi_1)$  i neka je  $\psi_2 = \varphi_2 - \varphi_2^*$ . Tada postoje  $\alpha, \beta \in \mathbb{R}$  takvi da je (vidi Sliku 5.3b)

$$\psi_2 = \varphi_2 - \lambda\psi_0 - \mu\psi_1.$$

Množeći skalarno redom ovu jednadžbu s  $\psi_0, \psi_1$  dobivamo

$$\begin{aligned} 0 &= (\psi_2, \psi_0) = (\varphi_2, \psi_0) - \lambda(\psi_0, \psi_0), \\ 0 &= (\psi_2, \psi_1) = (\varphi_2, \psi_1) - \mu(\psi_1, \psi_1), \end{aligned}$$

iz čega dobivamo

$$\lambda = \frac{(\varphi_2, \psi_0)}{(\psi_0, \psi_0)}, \quad \mu = \frac{(\varphi_2, \psi_1)}{(\psi_1, \psi_1)}, \quad \text{i} \quad \psi_2 = \varphi_2 - \frac{(\varphi_2, \psi_0)}{(\psi_0, \psi_0)}\psi_0 - \frac{(\varphi_2, \psi_1)}{(\psi_1, \psi_1)}\psi_1. \quad (5.19)$$

Ponavljajući postupak dobivamo

$$\begin{aligned} \psi_0 &= \varphi_0, \\ \psi_1 &= \varphi_1 - \frac{(\varphi_1, \psi_0)}{(\psi_0, \psi_0)}\psi_0, \\ &\vdots \\ \psi_n &= \varphi_n - \sum_{j=0}^{n-1} \frac{(\varphi_n, \psi_j)}{(\psi_j, \psi_j)}\psi_j. \end{aligned} \quad (5.20)$$

Budući da je  $\psi_k$  linearna kombinacija vektora  $\varphi_0, \varphi_1, \dots, \varphi_k$ , za svaki  $k = 0, 1, \dots, n$ , vrijedi  $L(\varphi_0, \varphi_1, \dots, \varphi_n) = L(\psi_0, \psi_1, \dots, \psi_n)$ . Zato najbolju aproksimaciju funkcije  $f$  možemo tražiti kao linearnu kombinaciju ortogonalnih funkcija  $\psi_0, \psi_1, \dots, \psi_n$ .

*Primjedba 5.6.* Neka je  $\mathcal{P} \subset C[a, b]$  potprostor na kome je definirana ortogonalna baza  $\psi_0, \psi_1, \dots, \psi_n$ , a  $f^*$  ortogonalna projekcija funkcije  $f \in C[a, b]$  na  $\mathcal{P}$ . Tada je

$$f^* = \sum_{k=0}^n \alpha_k \psi_k \quad \& \quad f' = f - f^* = f - \sum_{k=0}^n \alpha_k \psi_k.$$

Zbog ortogonalnosti funkcija  $\psi_0, \psi_1, \dots, \psi_n$  imamo

$$0 = (f', \psi_i) = (f, \psi_i) - \alpha_i(\psi_i, \psi_i), \quad i = 0, 1, \dots, n$$

iz čega slijedi slijedi

$$\alpha_i = \frac{(f, \psi_i)}{(\psi_i, \psi_i)}, \quad \text{odnosno} \quad f^* = \sum_{k=0}^n \frac{(f, \psi_k)}{(\psi_k, \psi_k)}\psi_k \quad (5.21)$$

Ako je baza  $\psi_0, \psi_1, \dots, \psi_n$  ortonormirana, onda je

$$f^* = \sum_{k=0}^n (f, \psi_k)\psi_k,$$

Fourierova aproksimacijom funkcije  $f$  na potprostoru  $\mathcal{P}$  s ortonormiranom bazom  $\psi_0, \psi_1, \dots, \psi_n$ , pri čemu su  $(f, \psi_k)$  Fourierovi koeficijenti.

**Primjer 5.6.** Neka je  $(\cdot, \cdot)$  skalarni produkt definiran na  $C[-1, 1]$  uz  $w(x) = 1$ . Sustav funkcija  $\varphi_i(x) = x^i$ ,  $i = 0, 1, \dots, 5$  ortogonalizirat ćemo prema Gram-Schmidtovom postupku (5.20).

Niže navedenim *Mathematica*-programom dobivamo

$$\begin{aligned} \psi_0(x) &= 1, & \psi_1(x) &= x, & \psi_2(x) &= x^2 - \frac{1}{3}, \\ \psi_3(x) &= x^3 - \frac{3}{5}x, & \psi_4(x) &= x^4 - \frac{6}{7}x^2 + \frac{3}{35}, & \psi_5(x) &= x^5 - \frac{10}{9}x^3 + \frac{5}{21}x. \end{aligned} \quad (5.22)$$

(\* Gram - Schmidtov postupak ortogonalizacije \*)

```
fi[i_, x_] := x^(i-1);
om[t_] := 1; a=-1; b=1; n=6;
Do[
  psi[i,t]=fi[i,t] -
    Sum[Integrate[om[t] fi[i,t] psi[j,t],{t,a,b}] psi[j,t]/
      Integrate[om[t] psi[j,t]^2, {t,a,b}], {j,i-1}];
  ,{i,n}]
ortpol = Table[{i-1, Simplify[psi[i,t]]}, {i,n}];
Print[TableForm[ortpol] ]
```

Polinomi (5.22) se do na konstantu podudaraju s poznatim Legendreovim polinomima, koji se obično definiraju kao

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n \in \mathbb{N}. \quad (5.23)$$

*Zadatak 5.10.* Pokažite da za Legendreove polinome (5.23) vrijedi

a) Svojstvo ortogonalnosti

$$\int_{-1}^1 P_m(x) P_n(x) dx = \begin{cases} 0, & m \neq n \\ \frac{2}{2n+1}, & m = n, \quad m, n = 0, 1, \dots \end{cases}$$

b) Legendreov polinom  $P_n$ ,  $n \geq 1$ , u intervalu  $(-1, 1)$  ima  $n$  jednostrukih nultočaka

Uputa: primijetite da je  $(x^2 - 1)/_{x=-1,1} = 0$  i koristite Rolleov teorem.

c)  $P_0(x) = 1$ ,  $P_1(x) = x$ , a za  $n = 1, 2, \dots$  vrijedi rekurzivna formula:

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x).$$

- d) Primjenom *Mathematica*-naredbe `LegendreP[n, x]` odredite nekoliko Legendreovih polinoma i nacrtajte njihove grafove.

*Primjedba 5.7.* Uz pretpostavku da je poznat ortogonalan sustav  $\psi_0, \psi_1, \dots, \psi_n$  na intervalu  $[-1, 1]$ , jednostavnom transformacijom

$$\gamma : [a, b] \longrightarrow [-1, 1], \quad \gamma(x) = \frac{2x}{b-a} - \frac{a+b}{b-a}, \quad (5.24)$$

možemo dobiti sustav funkcija  $(\tilde{\psi}_i)$ ,  $\tilde{\psi}_i = \psi_i \circ \gamma$ ,  $\tilde{\psi}_i : [a, b] \rightarrow \mathbb{R}$ , koji je ortogonalan na  $[a, b]$ .

*Zadatak 5.11.* Koristeći Legendreove polinome odrediti najbolju  $L_2$  aproksimaciju funkcije  $f(x) = e^{2x}$ ,  $x \in [0, 1]$  na potprostoru polinoma stupnja  $\leq 1$ . Rezultate usporedite s *Primjerom 5.1*.

*Uputa:* Koristeći transformaciju (5.24) na Legendreovim polinomima  $P_0$  i  $P_1$  dobivamo:  $\tilde{\psi}_0(x) = 1$ ,  $\tilde{\psi}_1(x) = 2x - 1$ , pa je  $f^* = a_0^* \cdot 1 + a_1^*(2x - 1)$ , gdje se  $a_0^*$  i  $a_1^*$  računaju prema (5.5).

### 5.1.3 Čebiševljevi polinomi

Sustav ortogonalnih polinoma na  $[-1, 1]$  s težinskom funkcijom  $w(x) = \frac{1}{\sqrt{1-x^2}}$  nazivaju se Čebiševljevi polinomi<sup>3</sup>. Mogu se definirati pomoću eksplisite formule

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, \dots \quad (5.25)$$

Korištenjem jednostavne relacije

$$\cos(n+1)\varphi + \cos(n-1)\varphi = 2 \cos \varphi \cos n\varphi, \quad n \in \mathbb{N},$$

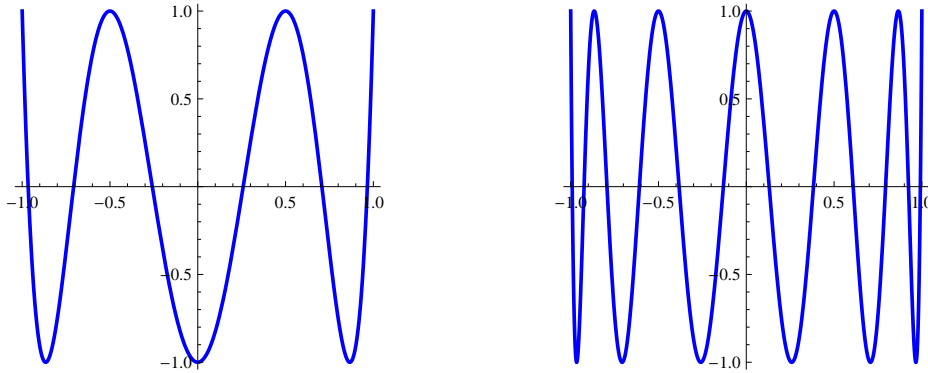
induktivno zaključujemo da su  $T_n$  polinomi stupnja  $n$ . Nadalje, koristeći činjenicu da je  $\varphi \mapsto \cos i\varphi \cos j\varphi$  parna funkcija uz supstituciju  $\varphi = \arccos x$ , dobivamo

$$\int_{-1}^1 T_i(x)T_j(x) \frac{dx}{\sqrt{1-x^2}} = - \int_{\pi}^0 \cos i\varphi \cos j\varphi d\varphi = \frac{1}{2} \int_{-\pi}^{\pi} \cos i\varphi \cos j\varphi d\varphi = \frac{\pi}{2} \delta_{ij},$$

(osim u slučaju  $i = j = 0$ , kada je rezultat  $\pi$ ) odakle neposredno slijedi ortogonalnost Čebiševljevih polinoma. Čebiševljevi polinomi u programskom sustavu *Mathematica* pozivaju se s `ChebyshevT[n, x]`.

Čebiševljevi polinomi imaju sljedeća svojstva:

<sup>3</sup>Oznaka polinoma "T" dolazi od njemačke transkripcije prezimena Čebišev: **Tschebscheff**

Slika 5.4: Čebiševljevi polinomi  $T_6$  i  $T_{12}$ 

- a)  $|T_n(x)| \leq 1$ ,  $x \in [-1, 1]$ ,  $n = 0, 1, \dots$ ;  
 b) Čebiševljev polinom  $T_n$  na intervalu  $[-1, 1]$  ima  $n$  različitih nultočaka

$$x_k = \cos\left(\frac{2k-1}{n}\frac{\pi}{2}\right), \quad k = 1, \dots, n;$$

- c) Čebiševljev polinom  $T_n$  na intervalu  $[-1, 1]$  ima  $n + 1$  različitih stacionarnih točaka

$$\xi_k = \cos\frac{k\pi}{n}, \quad k = 0, 1, \dots, n,$$

u kojima naizmjenično postiže vrijednosti globalnog minimuma i maksimuma;

- d)  $T_0(x) = 1$ ,  $T_1(x) = x$ , a za  $n = 1, 2, \dots$  vrijedi rekurzivna formula:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \dots$$

**Zadatak 5.12.** Izračunajte  $\|T_n\|$ ,  $n = 0, 1, \dots$ , te odgovarajuće koeficijente  $a_k^*$ ,  $k = 0, 1, \dots$ , najbolje  $L_2$  aproksimacije funkcije.

Rješenje:  $\|T_0\| = \sqrt{\pi}$ ,  $\|T_k\| = \sqrt{\frac{\pi}{2}}$ ,  $k = 0, 1, \dots$ ,  $a_k^* = \frac{2}{\pi} \int_0^\pi f(\varphi) \cos k\varphi d\varphi$ ,  $k = 0, 1, \dots$

**Zadatak 5.13.** Nultočke, kao i točke u kojima Čebiševljev polinom naizmjenično postiže globalne minimume i maksimume nisu jednoliko razmještene na intervalu  $[-1, 1]$ . One su gušće raspoređene pri rubovima intervala, a rjeđe oko središta. Označite točke  $x_k$ ,  $\xi_k$ ,  $k = 0, 1, \dots, n$  za  $n = 10$ , tako da na trigonometrijskoj kružnici najprije označite točke koje odgovaraju realnim brojevima  $\frac{2k-1}{n}\frac{\pi}{2}$  i  $\frac{k\pi}{n}$ ,  $k = 0, 1, \dots, n$ .

## 5.2 Najbolja $L_\infty$ aproksimacija

Razmotrimo sada problem najbolje aproksimacije neprekidne funkcije  $f \in C[a, b]$  na potprostoru  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$  s normom  $\|\cdot\|_\infty$  i težinskom funkcijom  $w(x) = 1$ . Treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{(a_0, \dots, a_n) \in \mathbb{R}^{n+1}} F(a_0, \dots, a_n), \quad \text{gdje je} \quad (5.26)$$

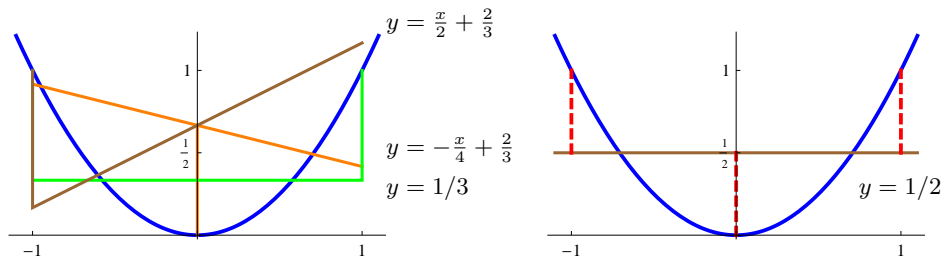
$$F(a_0, \dots, a_n) = \left\| \sum_{i=0}^n a_i \varphi_i - f \right\|_\infty = \max_{a \leq x \leq b} \left| \sum_{i=0}^n a_i \varphi_i(x) - f(x) \right|.$$

Ako je  $a^* = (a_0^*, \dots, a_n^*)^T \in \mathbb{R}^{n+1}$  rješenje problema (5.26), onda je funkcija  $f^* = \sum_{i=0}^n a_i^* \varphi_i$  najbolja  $L_\infty$  aproksimacija funkcije  $f$  na potprostoru  $L(\varphi_0, \varphi_1, \dots, \varphi_n)$ .

Zadržat ćemo se u okvirima jednog specijalnog slučaja: tražit ćemo najbolju  $L_\infty$  aproksimaciju funkcije  $f$  na prostoru polinoma stupnja  $\leq n$ , tj. na prostoru razapetom funkcijama  $\varphi_i(x) = x^i$ ,  $i = 0, 1, \dots, n$ .

**Primjer 5.7.** Za danu funkciju  $f(x) = x^2$ ,  $x \in [-1, 1]$  treba pronaći najbolju  $L_\infty$  aproksimaciju na prostoru polinoma stupnja  $\leq 1$ , tj. treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{a_0, a_1 \in \mathbb{R}} F(a_0, a_1), \quad F(a_0, a_1) = \max_{-1 \leq x \leq 1} |x^2 - (a_0 + a_1 x)|.$$



Slika 5.5: Linearna  $L_\infty$  aproksimacija kvadratne funkcije  $f(x) = x^2$

Ako s  $\mathcal{P}_1$  označimo skup svih polinoma stupnja  $\leq 1$ , onda odgovarajući optimizacijski problem možemo zapisati kao

$$\operatorname{argmin}_{P \in \mathcal{P}_1} \|x^2 - P(x)\|_\infty.$$

Budući da funkcija  $F$  nije derivabilna na  $[-1, 1]$ , ne možemo upotrijebiti svoje predznanje o istraživanju ekstrema derivabilnih funkcija. Za primjer izaberimo nekoliko polinoma  $P_i \in \mathcal{P}_1$  i pogledajmo pogreške  $\|x^2 - P_i\|_\infty$ .

Pokazat će se (vidi *Teorem 5.2*) da se najbolja aproksimacija dobije za  $a_0^* = \frac{1}{2}$ ,  $a_1^* = 0$ , tj. da je najbolja  $L_\infty$  aproksimacija kvadratne funkcije  $f(x) = x^2$  na potprostoru svih polinoma prvog stupnja, funkcija  $f^*(x) = \frac{1}{2}$ . Primijetite (vidi Sliku 5.4) da su u tom slučaju maksimalna odstupanja postižu u točkama  $-1, 0, 1$ , da su po apsolutnoj vrijednosti jednaka  $\frac{1}{2}$  i da po predznaku alterniraju.

Pokušajmo sada riješiti općenitiji problem.

*Za funkciju  $f(x) = x^n$ ,  $x \in [-1, 1]$  treba pronaći najbolju  $L_\infty$  aproksimaciju na prostoru  $\mathcal{P}_{n-1}$  polinoma stupnja  $\leq n - 1$ , tj. treba riješiti sljedeći optimizacijski problem*

$$\operatorname{argmin}_{P \in \mathcal{P}_{n-1}} \|x^n - P(x)\|_\infty. \quad (5.27)$$

Rješenje daje sljedeći teorem.

**Teorem 5.2.** *Između svih normiranih polinoma<sup>4</sup> stupnja  $\leq n$  polinom  $2^{1-n}T_n$ , gdje je  $T_n$   $n$ -ti Čebiševljev polinom, ima najmanju  $L_\infty$  normu na intervalu  $[-1, 1]$ , koja iznosi  $2^{1-n}$ .*

Drugim riječima,

$$\operatorname{argmin}_{P \in \mathcal{P}_n} \|P\|_\infty = 2^{1-n}T_n \quad \& \quad \min_{P \in \mathcal{P}_n} \|P\|_\infty = \|2^{1-n}T_n\|_\infty = 2^{1-n}.$$

Zato rješenje optimizacijskog problema (5.27) dobivamo iz

$$\begin{aligned} x^n - P(x) = 2^{1-n}T_n(x) &\implies P(x) = x^n - 2^{1-n}T_n(x), \\ \min_{P \in \mathcal{P}_{n-1}} \|x^n - P(x)\|_\infty &= 2^{1-n}. \end{aligned}$$

Specijalno, rješenje optimizacijskog problema iz *Primjera 5.7* je  $P_1(x) = x^2 - 2^{1-2}T_2(x)$ , gdje je  $T_2(x) = 2x^2 - 1$ , odnosno rješenje je  $P_1(x) = \frac{1}{2}$ , pri čemu je

$$\min_{P \in \mathcal{P}_1} \|x^2 - P(x)\|_\infty = 2^{1-2} = \frac{1}{2}.$$

Prijedimo sada na dokaz *Teorema 5.2*.

---

<sup>4</sup>Kažemo da je polinom normiran ako mu je koeficijent uz najvišu potenciju jednak 1



*Dokaz.* Uočimo najprije da je  $L_\infty$  norma Čebiševljevog polinoma  $T_n$  jednaka  $\|T_n\|_\infty = 1$  (svojstvo *a*), str. 139). Osim toga, njegov koeficijent uz najvišu potenciju jednak je  $2^{n-1}$  (svojstvo *d*), str. 139), pa je polinom  $2^{1-n}T_n$  normiran.

Dokaz tvrdnje teorema provest ćemo principom kontradikcije. Pretpostavimo da postoji neki drugi normirani polinom  $p_n$ , takav da je

$$\|p_n\|_\infty < \|2^{1-n}T_n(x)\|_\infty,$$

odnosno, takav da je  $\|p_n\|_\infty < 2^{1-n}$ , tj.  $\max_{x \in [-1,1]} |p_n(x)| < 2^{1-n}$ . Odavde slijedi

$$|p_n(x)| < 2^{1-n}, \quad \forall x \in [-1, 1],$$

što možemo pisati

$$-2^{1-n} \stackrel{(1)}{<} p_n(x) \stackrel{(2)}{<} 2^{1-n}, \quad \forall x \in [-1, 1]. \quad (5.28)$$

Neka su  $\xi_0, \xi_1, \dots, \xi_n$  točke u kojima Čebiševljev polinom  $T_n$  sukcesivno prima ekstremne vrijednosti  $-1$  ili  $1$  (vidi svojstvo *c*) str. 139) i neka je  $T_n(\xi_0) = 1$  (dokaz bi tekao slično i u slučaju  $T_n(\xi_0) = -1$ ). U tim točkama razmotrimo ponašanje “polinom pogreške”  $E_n(x) := p_n(x) - 2^{1-n}T_n(x)$ . U tu svrhu iz nejednakosti (5.28) za  $x = \xi_0, \xi_1, \dots, \xi_n$  računamo:

$$\begin{aligned} x = \xi_0 : p_n(\xi_0) &\stackrel{(2)}{<} 2^{1-n} \quad T_n(\xi_0) \stackrel{=}{=} 1 \quad 2^{1-n}T_n(\xi_0) \implies E_n(\xi_0) < 0 \\ x = \xi_1 : p_n(\xi_1) &\stackrel{(1)}{>} 2^{1-n} \quad T_n(\xi_1) \stackrel{=}{=} -1 \quad 2^{1-n}T_n(\xi_1) \implies E_n(\xi_1) > 0 \\ x = \xi_2 : p_n(\xi_2) &\stackrel{(2)}{<} 2^{1-n} \quad T_n(\xi_2) \stackrel{=}{=} 1 \quad 2^{1-n}T_n(\xi_2) \implies E_n(\xi_2) < 0 \\ &\dots \end{aligned}$$

To bi značilo da “polinom pogreške”  $E_n$ ,  $n$  puta mijenja predznak na intervalu  $[-1, 1]$ , tj. ima  $n$  različitih nultočaka. Budući da su polinomi  $p_n$  i  $2^{1-n}T_n$  normirani, stupanj polinoma  $E_n$  je  $(n-1)$ . Zbog toga je  $E_n \equiv 0$ , tj.  $p_n = 2^{1-n}T_n$ , što se protivi polaznoj pretpostavci.  $\square$

Ovo važno svojstvo Čebiševljevih polinoma već smo iskoristili u *t.2.1.3* prilikom optimalnog izbora čvorova interpolacije funkcije (str. 32). Iz *Teorema 2.1*, str. 30, uz oznaku  $M_{n+1} := \max_{x \in [a,b]} |f^{(n+1)}(x)|$  slijedi

$$|f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega(x)|, \quad \forall x \in [a, b]$$

odnosno

$$\|f - P_n\|_\infty \leq \frac{M_{n+1}}{(n+1)!} \|\omega\|_\infty, \quad (5.29)$$

gdje je  $\omega$  normirani polinom stupnja  $(n+1)$ . Prema *Teoremu 5.2* desna strana u (5.29) bit će najmanja (najbolja ocjena!) ako umjesto polinoma  $\omega$  stavimo polinom  $2^{-n}T_{n+1}$ , gdje je  $T_{n+1}$   $(n+1)$ -vi Čebiševljev polinom proširen na interval  $[a, b]$ . U tom slučaju ocjena (5.29) prelazi u

$$\|f - P_n\|_\infty \leq \frac{M_{n+1}}{2^n(n+1)!}, \quad (5.30)$$

a kao čvorove interpolacije mora se uzeti nultočke Čebiševljevog polinoma  $T_{n+1}$  proširenog na interval  $[a, b]$ .

*Zadatak 5.14.* Odredite najbolju  $L_\infty$  aproksimaciju funkcije  $f(x) = x^n$ ,  $x \in [-1, 1]$  na prostoru polinoma stupnja  $\leq n-1$ .

*Odgovor:* Prema *Teoremu 5.2* najbolja aproksimacija je funkcija  $f^*(x) = x^n - 2^{1-n}T_n(x)$ .

Prethodne tvrdnje mogu se generalizirati kroz sljedeći važni teorem (Ackleh et al., 2010; Cheney, 1966; Dahlquist and Björck, 2008).

**Teorem 5.3.** (Chebyshev's Equioscillation Theorem) *Neka je  $f \in C[a, b]$ . Polinom  $p^*$  stupnja  $\leq n$  je najbolja  $L_\infty$  aproksimacija funkcije  $f$  na  $[a, b]$  onda i samo onda ako postoje barem  $n+2$  točke,*

$$a \leq x_0 < x_1 < \dots < x_{n+1} \leq b$$

*u kojima funkcija pogreške  $E_n = p_n^* - f$  prima maksimalna odstupanja  $\|p_n^* - f\|_\infty$  s alternirajućim predznakom, tj.*

$$E(x_{i+1}) = -E(x_i), \quad i = 0, 1, \dots, n.$$

*Zadatak 5.15.* Primjenom *Chebyshev's Equioscillation Theorem* odredite najbolju  $L_\infty$  aproksimaciju funkcije  $f(x) = \sin x$  na  $[0, \frac{\pi}{2}]$  u klasi polinoma prvog stupnja.

*Rješenje:*  $p_1^*(x) = 0.105 + \frac{2t}{\pi}$ .



## Poglavlje 6

# Problemi najmanjih kvadrata

Problem najbolje  $l_2$ -aproksimacije funkcije koja je zadana na konačnom skupu točaka obično se u literaturi naziva problem najmanjih kvadrata<sup>1</sup>.

Najprije ćemo na nekoliko tipičnih primjera pokazati neke probleme iz primjena, koji se rješavaju metodom najmanjih kvadrata. Još neke druge primjene mogu se vidjeti kod (Scitovski, 1993).

**Primjer 6.1.** *Pretpostavimo da su  $y_1, y_2, \dots, y_m$  rezultati mjerenja neke veličine  $A$ . Treba odrediti aproksimaciju  $\alpha^*$  veličine  $A$  tako da sve izmjerene vrijednosti  $y_1, y_2, \dots, y_m$  budu “što bliže” aproksimaciji  $\alpha^*$*



Slika 6.1: Aritmetička sredina podataka

Pri tome pojam “što bliže” shvaćamo u smislu zahtjeva da suma kvadrata (En. Least Squares) odstupanja svih mjerenja od aproksimacije  $\alpha$  bude mi-

<sup>1</sup>Princip najmanjih kvadrata postavio je Carl Friedrich Gauss 1795. godine prilikom izučavanja kretanja nebeskih tijela, što je objavio 1809 godine u radu “*Teoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*”, (Perthes and Besser, Hamburg). Treba također reći da je francuski matematičar Legendre 1805. prvi publicirao algebarsku proceduru metode najmanjih kvadrata

nimalna, tj. da bude

$$F_{LS}(\alpha^*) = \min_{\alpha \in \mathbb{R}} F_2(\alpha), \quad F_{LS}(\alpha) = \sum_{i=1}^m (\alpha - y_i)^2 \quad (6.1)$$

Primijetite da je  $F_{LS}(\alpha) = \|\alpha \cdot \mathbf{1} - \mathbf{y}\|^2$  kvadrat euklidske norme razlike  $(\alpha \cdot \mathbf{1} - \mathbf{y})$ , gdje je  $\mathbf{y} \in \mathbb{R}^m$  vektor podataka, a  $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^m$ . Ovaj princip nazivamo princip najmanjih kvadrata ili češće metoda najmanjih kvadrata, a problem određivanja aproksimacije  $\alpha^*$  na osnovi principa najmanjih kvadrata, nazivamo problem najmanjih kvadrata. Problem najmanjih kvadrata je specijalni problem ekstrema bez ograničenja. Lako se vidi da je rješenje problema (6.1) aritmetička sredina podataka

$$\alpha^* = \frac{1}{m} \sum_{i=1}^m y_i$$

Dakle, aritmetička sredina mjerenja  $y_1, \dots, y_m$  je takva veličina koja ima svojstvo da je suma kvadrata odstupanja te veličine od mjerenja  $y_1, \dots, y_m$  najmanja.

**Primjer 6.2.** *Zadani su podaci*

$x_i$	-4	-2	2	4
$y_i$	8	6	3	5

Treba pronaći linearnu funkciju  $f(x) = a_1 + a_2x$  tako da njezin graf prolazi što bliže danim točkama  $T_i = (x_i, y_i)$ ,  $i = 1, 2, 3, 4$  (vidi Sliku 6.2a).

Jasno je da bi nam već prve dvije točke  $T_1$  i  $T_2$  bile dovoljne za određivanje tražene funkcije. Ali što je onda s preostalim točkama – podacima? Zato ćemo pokušati parametre  $a_1, a_2$ , odrediti tako da bude

$$y_i \approx a_1 + a_2x_i, \quad i = 1, \dots, 4, \quad (*)$$

pri čemu će nam  $y_i \approx y_j$  u (\*) značiti minimalno kvadratno odstupanje. Drugim riječima, definirat ćemo odstupanja

$$r_i = a_1 + a_2x_i - y_i, \quad i = 1, \dots, 4$$

i pokušati riješiti optimizacijski problem

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(a_1, a_2), \quad F_{LS}(a_1, a_2) = \sum_{i=1}^n (a_1 + a_2x_i - y_i)^2 = \sum_{i=1}^n r_i^2 = \|\mathbf{r}\|_2^2 \quad (**)$$

gdje je  $\mathbf{r}$  vektor s komponentama  $r_i$ . Odgovarajući problem (\*\*) također je jedan problem najmanjih kvadrata.

*Primjedba 6.1.* Problem  $(**)$  često zovemo i problem diskretne  $l_2$  aproksimacije za razliku od  $L_2$  aproksimacije funkcije u *t.5, str.125*. Problem bismo mogli postaviti i tako da umjesto  $(**)$  promatramo problem

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_1(a_1, a_2), \quad F_1(a_1, a_2) = \sum_{i=1}^n |a_1 + a_2 x_i - y_i| = \|\mathbf{r}\|_1 \quad (***)$$

ili

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_\infty(a_1, a_2), \quad F_\infty(a_1, a_2) = \max_{i=1, \dots, m} |a_1 + a_2 x_i - y_i| = \|\mathbf{r}\|_\infty \quad (***)$$

Problem  $(***)$  naziva se problem najbolje diskretne  $l_1$  aproksimacije, a problem  $(****)$  problem najbolje diskretne  $l_\infty$  aproksimacije. Ocem diskretnih  $l_1$  aproksimacija smatra se hrvatski znanstvenik *R. Bošković*, a ocem diskretnih  $l_\infty$  aproksimacija ruski matematičar *P. L. Čebišev*. Inače funkcije  $F_1$  i  $F_\infty$  nisu derivabilne, pa su zato problemi  $(***)$  i  $(****)$  znatno složeniji i zbog toga ćemo ih samo kratko predstaviti u *t.6.3, str.166*.

**Primjer 6.3.** *Promatramo sustav linearnih jednadžbi  $\mathbf{J}\mathbf{a} = \mathbf{y}$ , gdje je*

$$\mathbf{J} = \begin{bmatrix} 1 & -4 \\ 1 & -2 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 8 \\ 6 \\ 3 \\ 5 \end{bmatrix}.$$

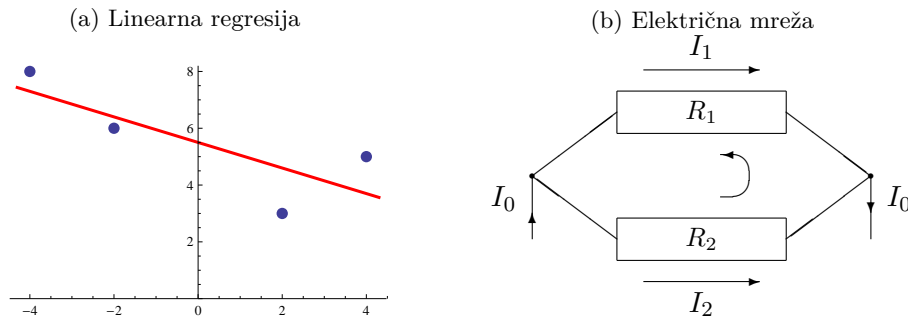
Ovaj sustav nije rješiv jer vektor  $\mathbf{y}$  ne leži u potprostoru  $L(\mathbf{j}_1, \mathbf{j}_2)$ , gdje su  $\mathbf{j}_1, \mathbf{j}_2$  stupci matrice  $\mathbf{J}$ . Kažemo da se radi o *predređenom sustavu*. Potražiti ćemo najbolju  $l_2$  aproksimaciju vektora  $\mathbf{y}$  u potprostoru  $L(\mathbf{j}_1, \mathbf{j}_2)$ , tj. riješiti ćemo sljedeći optimizacijski problem

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(a_1, a_2), \quad F_{LS}(a_1, a_2) = \|\mathbf{J}\mathbf{a} - \mathbf{y}\|_2^2 = \sum_{i=1}^n (a_1 + a_2 x_i - y_i)^2.$$

Primijetite da će za optimalne vrijednosti parametara  $a_1 = 5.5$ ,  $a_2 = -0.45$  suma kvadrata odstupanja lijevih od desnih strana jednadžbi biti minimalna. I u ovom slučaju radi se o jednom problemu najmanjih kvadrata.

Iz navedenog primjera jasno je da i problemu rješavanja sustava linearnih jednadžbi također možemo pristupiti kao jednom problemu najmanjih kvadrata.

**Primjer 6.4.** *U električnoj mreži bez izvora, sastavljenoj od otpora  $R_k$ , struje  $I_k$  izračunavaju se pomoću Kirchhoffovih zakona:*



Slika 6.2: Problemi najmanjih kvadrata

**Pravilo čvora:** Za svaki čvor mreže vrijedi  $\sum_k I_k = 0$ . Pri tome, dolazeće struje smatraju se pozitivnim, a odlazeće negativnim;

**Pravilo petlje:** Za svaku zatvorenu petlju (konturu) vrijedi  $\sum_k U_k = \sum_k I_k R_k = 0$ . Pri tome, ako konturu obilazimo u nekom smjeru, onda se struje u tom smjeru smatraju pozitivnim, a struje u suprotnom smjeru, negativnim.

Za primjer na Slici 6.2b i zadane otpore  $R_1, R_2$  i ulaznu struju  $I_0$  treba izračunati struje  $I_1$  i  $I_2$ . Imamo

$$\begin{array}{rcccc} I_0 & -I_1 & -I_2 & = & 0 \\ -I_0 & +I_1 & +I_2 & = & 0 \\ & -I_1 R_1 & +I_2 R_2 & = & 0 \end{array}$$

Na taj način obično dobivamo više jednadžbi nego nepoznanica, a zbog prisutnosti pogrešaka u mjerenjima, takav sustav je često preodređen i nema rješenje, pa se obično zadovoljavamo najboljom  $l_2$  aproksimacijom rješenja.

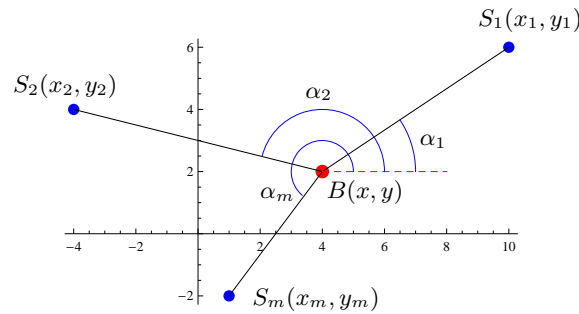
**Primjer 6.5.** SOS - poziv u pomoć s nekog broda primiljen je na  $m$  ( $m \geq 2$ ) radio stanica. Poznate su koordinate položaja tih radiostanica:  $S_i(x_i, y_i)$ ,  $i = 1, \dots, m$ , kao i kutevi:  $\alpha_i$ ,  $i = 1, \dots, m$ , pod kojima su stanice primile poziv s broda (vidi Sliku 6.3) Treba odrediti koordinate  $x, y$  položaja broda koji šalje poziv u pomoć. Zbog jednostavnosti nećemo uzimati u obzir zakrivljenost Zemljine površine.

Očigledno vrijedi

$$\alpha_i = \arctg \left( \frac{y - y_i}{x - x_i} \right), \quad i = 1, \dots, m. \quad (*)$$

Kada bi svi izmjereni kutevi bili potpuno točni, sustav (\*) imao bi jedinstveno rješenje. Budući da se to u praksi ne može očekivati, onda ćemo rješenje sustava (\*) potražiti u smislu najmanjih kvadrata

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(x, y), \quad F_{LS}(x, y) = \sum_{i=1}^m \left[ \operatorname{arctg} \left( \frac{y-y_i}{x-x_i} \right) - \alpha_i \right]^2.$$



Slika 6.3: Traženje pozicije broda

Pretpostavimo općenito da zavisna varijabla  $y$  ovisi o nezavisnoj varijabli  $x$  po funkcionalnom zakonu

$$y = f(x; \mathbf{a}), \quad \mathbf{a} \in \mathbb{R}^n,$$

gdje je  $\mathbf{a} = (a_1, \dots, a_n)^T$  vektor parametar. Osim toga, neka su zadani podaci

$$(x_i, y_i), \quad i = 1, \dots, m, \quad m \geq n.$$

Nadalje, neka je  $\mathbf{r}$  vektor odstupanja (reziduala) s komponentama

$$r_i(\mathbf{a}) = f(x_i; \mathbf{a}) - y_i. \quad (6.2)$$

Treba pronaći točku  $\mathbf{a}^* = (a_1^*, \dots, a_n^*) \in \mathbb{R}^n$ , u kojoj funkcija  $F: \mathbb{R}^n \rightarrow \mathbb{R}$ , definirana s

$$F(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{a}) = \frac{1}{2} \|\mathbf{r}(\mathbf{a})\|_2^2, \quad \mathbf{a} \in \mathbb{R}^n, \quad (6.3)$$

postizuje globalni minimum, tje treba riješiti sljedeći problem globalne optimizacije

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}). \quad (6.4)$$

Funkciju  $f$  nazivamo funkcija-model, vektor  $\mathbf{r}$  vektor odstupanja ili reziduala, a brojeve  $a_1^*, \dots, a_n^*$  optimalni parametri. Problem određivanja optimalnih parametara  $a_1^*, \dots, a_n^*$  nazivamo problem najmanjih kvadrata (u daljnjem tekstu: PNK).





gdje su  $\mathbf{H}_k$  matrica reda  $n$  s elementima

$$(\mathbf{H}_k)_{ij} = \frac{\partial^2 r_k}{\partial a_i \partial a_j}, \quad (6.8)$$

Ako želimo pronaći stacionarne točke minimizirajuće funkcije (6.2), odnosno (6.5), potrebno je gradijent funkcije  $F$  izjednačiti s nulom. Tako dobivamo sustav od  $n$  jednadžbi s  $n$  nepoznanica  $a_1, \dots, a_n$

$$\mathbf{J}^T \mathbf{r} = \mathbf{0} \quad \text{ili} \quad \begin{cases} \sum_{i=1}^m r_i \frac{\partial r_i}{\partial a_1} = 0, \\ \vdots \\ \sum_{i=1}^m r_i \frac{\partial r_i}{\partial a_n} = 0. \end{cases} \quad (6.9)$$

Ovaj sustav jednadžbi bit će linearan ako su sve funkcije  $r_i = r_i(a_1, \dots, a_n)$  linearne u svakom od argumenata  $a_1, \dots, a_n$ . To će se dogoditi onda ako je funkcija-model  $f$  linearan u svim parametrima  $a_1, \dots, a_n$ . Tada govorimo o linearnom problemu najmanjih kvadrata (LPNK).

U suprotnom, ako je funkcija-model  $f$  nelinearna u barem jednom od parametara:  $a_1, \dots, a_n$ , govorimo o nelinearnom problemu najmanjih kvadrata (NPNK). U tom slučaju je (6.9) sustav nelinearnih jednadžbi.

Primijetimo da su *Primjeri 6.1, 6.2. i 6.3.* linearni problemi, a *Primjer 6.4.* nelinearni problem najmanjih kvadrata.

## 6.1 Linearni problemi najmanjih kvadrata

Pretpostavimo da je funkcija-model  $f$  linearan u svim parametrima  $a_1, \dots, a_n$  i da je općenito oblika

$$f(x; a_1, \dots, a_n) = a_1 \varphi_1(x) + \dots + a_n \varphi_n(x), \quad x \in \mathbb{R} \quad (6.10)$$

gdje su  $\varphi_i$  neprekidne funkcije. Primjerice za polinom-model, funkcije  $\varphi_i$  su potencije:  $\varphi_i(x) = x^{i-1}$ ,  $i = 1, \dots, n$ .

Jacobijeva matrica za funkciju-model (6.10) ne ovisi o vektoru parametara  $\mathbf{a}$ . Naime, kako u ovom slučaju (6.2) glasi

$$r_i(\mathbf{a}) = a_1 \varphi_1(x) + \dots + a_n \varphi_n(x) - y_i,$$

imamo

$$\mathbf{J}_{ij} = \frac{\partial r_i}{\partial a_j} = \varphi_j(x_i), \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Hesijan funkcije  $F$  također je bitno jednostavniji i također ne ovisi o vektoru parametara  $\mathbf{a}$

$$\mathbf{H}_F = \mathbf{J}^T \mathbf{J}. \quad (6.11)$$

Vektor odstupanja  $\mathbf{r}$  s komponentama  $r_i = a_1\varphi_1(x_i) + \cdots + a_n\varphi_n(x_i) - y_i$  možemo pisati

$$\mathbf{r} = \mathbf{J}\mathbf{a} - \mathbf{y}, \quad (6.12)$$

gdje je  $\mathbf{y} = (y_1, \dots, y_m)^T$ . Treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad \text{gdje je} \quad F(\mathbf{a}) = \frac{1}{2} \|\mathbf{r}(\mathbf{a})\|^2 = \frac{1}{2} \|\mathbf{J}\mathbf{a} - \mathbf{y}\|^2. \quad (6.13)$$

Često puta se u literaturi LPNK (6.13) označava kao

$$\mathbf{J}\mathbf{a} \simeq \mathbf{y}. \quad (6.14)$$

Primjerice, za podatke i funkciju-model iz *Primjera 6.2* imamo:  $\varphi_1(x) = 1$ ,  $\varphi_2(x) = x$  i

$$\mathbf{J} = \begin{bmatrix} 1 & -4 \\ 1 & -2 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}, \quad \mathbf{H}_F = \mathbf{J}^T \mathbf{J} = \begin{bmatrix} 4 & 0 \\ 0 & 40 \end{bmatrix}.$$

### 6.1.1 Rješavanje LPNK preko sustava normalnih jednadžbi

Stacionarne točke funkcije  $F$  dobit ćemo rješavanjem jednadžbe (6.6)

$$\operatorname{grad} F = \mathbf{J}^T \mathbf{r} = 0. \quad (6.15)$$

Hesijan  $\mathbf{H}_F$  ne ovisi o vektoru  $\mathbf{a}$  i zadan je formulom (6.11). Sljedeća lema pokazuje kada će Hessijan  $\mathbf{H}_F$  biti pozitivno definitan

**Lema 6.1.** *Neka je  $\mathbf{J} \in \mathbb{R}^{m \times n}$ ,  $m > n$ . Matrica  $\mathbf{J}^T \mathbf{J}$  je pozitivno definitna onda i samo onda ako je  $\mathbf{J}$  punog ranga po stupcima ( $\operatorname{rang} \mathbf{J} = n$ ).*

**Dokaz.** (*Nužnost*) Neka je  $\mathbf{J}^T \mathbf{J}$  pozitivno definitna. Pretpostavimo da je  $\operatorname{rang} \mathbf{J} < n$ , tj. da su stupci od  $\mathbf{J}$  linearno zavisni. Tada bi za neki  $\mathbf{a}_0 \neq \mathbf{0}$  imali  $\mathbf{J}\mathbf{a}_0 = \mathbf{0}$ , odakle bi slijedilo  $\mathbf{a}_0^T \mathbf{J}^T \mathbf{J} \mathbf{a}_0 = 0$ , što bi značilo da  $\mathbf{J}^T \mathbf{J}$  nije pozitivno definitna.

(Dovoljnost) Pretpostavimo da je rang  $\mathbf{J} = n$ , tj. da su stupci od  $\mathbf{J}$  linearno nezavisni. Tada za  $\mathbf{a} \neq \mathbf{0} \Rightarrow \mathbf{J}\mathbf{a} \neq \mathbf{0}$ , pa je

$$(\mathbf{H}_F \mathbf{a}, \mathbf{a}) = (\mathbf{J}^T \mathbf{J} \mathbf{a}, \mathbf{a}) = \mathbf{a}^T (\mathbf{J}^T \mathbf{J} \mathbf{a}) = (\mathbf{J} \mathbf{a})^T (\mathbf{J} \mathbf{a}) = \|\mathbf{J} \mathbf{a}\|^2 > 0.$$

□

Dakle, LPNK, za koji je pripadna Jacobijeva matrica punog ranga (što je u praksi gotovo uvijek ispunjeno), rješiv je i postoji jedinstveno rješenje, koje možemo dobiti tako da pronađemo kritične točke funkcije  $F$ . Zato ćemo (6.12) uvrstiti (6.15). Dobivamo jednadžbu

$$\mathbf{J}^T \mathbf{J} \mathbf{a} - \mathbf{J}^T \mathbf{y} = 0 \quad (6.16)$$

Lako se vidi da je rješenje ove jednadžbe

$$\mathbf{a}^* = \mathbf{J}^+ \mathbf{y}, \quad \text{gdje je} \quad \mathbf{J}^+ = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T.$$

Jednadžbu (6.16) nazivamo sustav normalnih jednadžbi, a matricu  $\mathbf{J}^+$  pseudoinverzna matrica ili Moore-Penroseov generalizirani inverz<sup>2</sup> matrice  $\mathbf{J}$ .

*Primjedba 6.3.* U slučaju ako je funkcija-model (6.10) polinom  $n$ -tog stupnja, a broj podataka  $m$  jednak  $(n + 1)$ , rješavanjem odgovarajućeg LPNK dobivamo *interpolacijski polinom*.

*Primjedba 6.4.* Kako je  $\text{cond}(\mathbf{J}^T \mathbf{J}) = \text{cond}^2(\mathbf{J})$  (vidi primjerice Gill et al. (1991)), matrica  $\mathbf{J}^T \mathbf{J}$  bit će vrlo loše uvjetovana ako je matrica  $\mathbf{J}$  loše uvjetovana, a rješenje LPNK, dobiveno rješavanjem sustava normalnih jednadžbi (6.16), bit će vrlo nepouzdana.

**Primjer 6.6.** *Zadano je*

$$\mathbf{J} = \begin{bmatrix} 1 & 1 \\ 10^{-3} & 0 \\ 0 & 10^{-3} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 2 \\ 10^{-3} \\ 10^{-3} \end{bmatrix}$$

Može se pokazati (vidi *Primjedbu 3.9*, str. 79) da je  $\text{cond}(\mathbf{J}) \approx 1.4 \times 10^3$  i da je rješenje odgovarajućeg LPNK  $\mathbf{a}^* = (1, 1)^T$  uz  $\|\mathbf{r}(\mathbf{a}^*)\| = 0$ .

Ako bismo ovaj LPNK rješavali preko sustava normalnih jednadžbi na računaru s 6-znamenkastom floating-point aritmetikom, matrica  $fl(\mathbf{J}^T \mathbf{J}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$  bila bi singularna. U 7-znamenkastoj floating-point aritmetici rješenje bi bilo  $\mathbf{a}^* = (2.000001, 0)^T$ .

<sup>2</sup>Više o svojstvima i značenju Moore-Penroseovog generaliziranog inverza može se vidjeti primjerice kod (Gill et al., 1991; Golub and van Loan, 1996; Lawson and Hanson, 1996)

### 6.1.2 Rješavanje LPNK pomoću QR-dekompozicije

Minimum funkcije  $F(\mathbf{a}) = \frac{1}{2}\|\mathbf{J}\mathbf{a} - \mathbf{y}\|^2$  iz (6.13) potražiti ćemo na sljedeći način. Kako je  $\mathbf{J} \in \mathbb{R}^{m \times n}$ , najprije ćemo algoritmom opisanim u t.3.7, str. 67, načiniti QR-dekompoziciju  $\mathbf{J} = \mathbf{Q}\mathbf{R}$ , gdje je  $\mathbf{Q} \in \mathbb{R}^{m \times m}$  ortogonalna, a  $\mathbf{R} \in \mathbb{R}^{m \times n}$  gornja trokutasta matrica. Vektor odstupanja  $\mathbf{r} = \mathbf{J}\mathbf{a} - \mathbf{y}$ , sada možemo ovako pisati

$$\mathbf{r} = \mathbf{Q}\mathbf{R}\mathbf{a} - \mathbf{y}. \quad (6.17)$$

Budući da ortogonalna matrica čuva normu, vrijedi

$$\|\mathbf{r}\| = \|\mathbf{Q}\mathbf{R}\mathbf{a} - \mathbf{y}\| = \|\mathbf{Q}(\mathbf{R}\mathbf{a} - \mathbf{Q}^T\mathbf{y})\| = \|\mathbf{R}\mathbf{a} - \mathbf{Q}^T\mathbf{y}\|, \quad (6.18)$$

pa je

$$\begin{aligned} 2F(\mathbf{a}) = \|\mathbf{r}\|^2 = \|\mathbf{R}\mathbf{a} - \mathbf{Q}^T\mathbf{y}\|^2 &= \|\mathbf{R}_n\mathbf{a} - (\mathbf{Q}^T\mathbf{y})_n\|_{(n)}^2 + \|(\mathbf{Q}^T\mathbf{y})_{m-n}\|_{(m-n)}^2 \\ &\geq \|(\mathbf{Q}^T\mathbf{y})_{m-n}\|_{(m-n)}^2, \end{aligned}$$

gdje je  $\mathbf{R}_n$  matrica sastavljena od prvih  $n$  redaka matrice  $\mathbf{R}$ , a  $(\mathbf{Q}^T\mathbf{y})_n$ , odnosno  $(\mathbf{Q}^T\mathbf{y})_{m-n}$  vektori sastavljeni od prvih  $n$ , odnosno posljednjih  $m-n$  komponenti originalnih vektora. Indeksi  $(n)$  i  $(m-n)$  označavaju odgovarajuće vektorske  $l_2$  norme u  $\mathbb{R}^n$ , odnosno  $\mathbb{R}^{m-n}$ . Zato se minimum funkcije  $F$  postiže na vektoru  $\mathbf{a}^* \in \mathbb{R}^n$ , za koji je

$$\mathbf{R}_n\mathbf{a}^* = (\mathbf{Q}^T\mathbf{y})_n. \quad (6.19)$$

pri čemu je

$$F(\mathbf{a}^*) = \frac{1}{2}\|\mathbf{r}(\mathbf{a}^*)\|^2 = \frac{1}{2}\|(\mathbf{Q}^T\mathbf{y})_{m-n}\|_{(m-n)}^2. \quad (6.20)$$

**Primjer 6.7.** LPNK iz Primjera 6.2

$$\mathbf{J}\mathbf{a} = \begin{bmatrix} 1 & 4 \\ 1 & 2 \\ 1 & -2 \\ 1 & -4 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \simeq \mathbf{y} = \begin{bmatrix} 8 \\ 6 \\ 3 \\ 5 \end{bmatrix}.$$

*riješit ćemo primjenom QR-dekompozicije.*

Primjenom *Mathematica*-naredbe `QRDecomposition[J]` dobivamo sustav (6.19)

$$\begin{bmatrix} -2. & 0. \\ 0 & 6.32456 \end{bmatrix} \begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} -11. \\ -2.84605 \end{bmatrix},$$

odakle dobivamo  $\mathbf{a}^* = (5.5, -0.45)^T$  i  $F(\mathbf{a}^*) = 1.1068$ .

**Primjer 6.8.** U Primjeru 2.4, str. 27, pokazalo se da su posljednja dva koeficijenta u interpolacijskom polinomu zanemarivo malena, pa ima smisla potražiti polinom stupnja 2, koji će u smislu najmanjih kvadrata najbolje aproksimirati zadane podatke i dati funkcionalnu zavisnost  $\alpha = P_2(X)$ .

Primjenom QR-dekompozicije za rješavanje odgovarajućeg LPNK dobivamo

$$a_0 = 5.0467, a_1 = -0.0033, a_2 = 0.000023.$$

Usporedite rezultate prikazane u *Tablici 6.1* s *Tablicom 2.2*, str. 27.

$X$ (m)	$\alpha(^{\circ})$	$\alpha^*(^{\circ})$	$\Delta\alpha^*$
500	10.885	9.14	1.745
750	16.901	15.49	1.411
1000	23.942	24.71	0.768

Tablica 6.1: Domet projektila u ovisnosti o izlaznom kutu

### 6.1.3 Rješavanje LPNK pomoću dekompozicije na singularne vrijednosti

Ako Jacobijeva matrica  $\mathbf{J}$  nije punog ranga, prema *Lemi 6.1*, str. 152, Hesijan  $H_F$  odgovarajućeg LPNK  $\mathbf{J}\mathbf{a} \simeq \mathbf{y}$  nije pozitivno definitan. LPNK u tom slučaju ima beskonačno mnogo rješenja. Od svih tih rješenja potražiti ćemo ono koje ima najmanju normu. U tom slučaju i u slučaju kada je Jacobijan  $\mathbf{J}$  loše uvjetovana matrica koristit ćemo rastav matrice na singularne vrijednosti (SVD).

Promatrajmo ponovo problem minimizacije (6.13) funkcije

$$F(\mathbf{a}) = \frac{1}{2} \|\mathbf{r}(\mathbf{a})\|^2, \quad \text{gdje je } \mathbf{r} = \mathbf{J}\mathbf{a} - \mathbf{y},$$

s time da je  $\text{rang}(\mathbf{J}) = k \leq n < m$ . Neka je  $\mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{V}^T$  dekompozicije matrica  $\mathbf{J}$  na singularne vrijednosti, gdje su  $\mathbf{U}$ ,  $\mathbf{S}$  i  $\mathbf{V}$  matrice definirane u *t.3.9*, str. 77. Budući da ortogonalna matrica čuva normu, vrijedi

$$\|\mathbf{r}\| = \|\mathbf{J}\mathbf{a} - \mathbf{y}\| = \|\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{a} - \mathbf{y}\| = \|\mathbf{S}\mathbf{V}^T\mathbf{a} - \mathbf{U}^T\mathbf{y}\|.$$

Zato, uz oznaku  $\mathbf{z} = \mathbf{V}^T\mathbf{a}$ , imamo

$$\begin{aligned} 2F(\mathbf{a}) = \|\mathbf{r}\|^2 = \|\mathbf{S}\mathbf{z} - \mathbf{U}^T\mathbf{y}\|^2 &= \|\mathbf{S}_k\mathbf{z} - (\mathbf{U}^T\mathbf{y})_k\|_{(k)}^2 + \|(\mathbf{U}^T\mathbf{y})_{m-k}\|_{(m-k)}^2 \\ &\geq \|(\mathbf{U}^T\mathbf{y})_{m-k}\|_{m-k}^2, \end{aligned}$$

gdje je  $\mathbf{S}_k$  matrica sastavljena od prvih  $k$  redaka dijagonalne matrice  $\mathbf{S}$ , a  $(\mathbf{U}^T \mathbf{y})_k$ , odnosno  $(\mathbf{U}^T \mathbf{y})_{m-k}$  vektori sastavljeni od prvih  $k$ , odnosno posljednjih  $m-k$  komponenti originalnih vektora. Indeksi  $(k)$  i  $(m-k)$  označavaju odgovarajuće vektorske  $l_2$  norme u  $\mathbb{R}^k$ , odnosno  $\mathbb{R}^{m-k}$ . Zato se minimum funkcije  $F$  postiže na vektoru  $\mathbf{z}^* \in \mathbb{R}^n$ , za koji je

$$\mathbf{S}_k \mathbf{z}^* = (\mathbf{U}^T \mathbf{y})_k. \quad (6.21)$$

Na taj način određeno je prvih  $k$  komponenti vektora  $\mathbf{z}^*$ ,

$$z_i^* = \frac{u_i^T y}{\sigma_i}, \quad i = 1, \dots, k,$$

gdje su  $\mathbf{u}_1, \dots, \mathbf{u}_k$  prvih  $k$  stupaca matrice  $\mathbf{U}$ . Preostalih  $(n-k)$  komponenti  $z_{k+1}^*, \dots, z_n^*$  vektora  $\mathbf{z}^*$  biramo tako da rješenje  $\mathbf{a}^*$  ima minimalnu euklidsku normu. Dakle,

$$\mathbf{a}^* = \mathbf{V} \mathbf{z}^* = \sum_{i=1}^k \frac{u_i^T y}{\sigma_i} \mathbf{v}_i, \quad \|\mathbf{a}^*\| = \min_{\mathbf{J} \mathbf{a} = \mathbf{y}} \|\mathbf{a}\|, \quad (6.22)$$

gdje su  $\mathbf{v}_1, \dots, \mathbf{v}_k$  prvih  $k$  stupaca matrice  $\mathbf{V}$ . Minimalna vrijednost funkcije  $F$  je

$$F(\mathbf{a}^*) = \frac{1}{2} \|(\mathbf{U}^T \mathbf{y})_{m-k}\|_{(m-k)}^2 = \sum_{i=k+1}^m (\mathbf{u}_i^T \mathbf{y})^2. \quad (6.23)$$

**Primjer 6.9.** Promatramo LPNK  $\mathbf{J} \mathbf{a} \simeq \mathbf{y}$ , gdje je

$$\mathbf{J} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Budući da je rang  $\mathbf{J} = k = 1$ , LPNK ima beskonačno mnogo rješenja. SVD ove matrice glasi

$$\mathbf{J} = \begin{bmatrix} -.707107 & -.707107 \\ -.707107 & .707107 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -.707107 & -.707107 \\ -.707107 & .707107 \end{bmatrix}.$$

Primjenom formule (6.22) dobivamo rješenje (koje ima minimalnu  $l_2$  normu!)

$$\mathbf{a}^* = -\frac{1}{2} ((-.707107, -.707107) \cdot (1, 2)^T) \cdot (-.707107, -.707107)^T = (.75, .75)^T.$$

Isti rezultat možemo dobiti i izravno, tražeći minimum funkcije

$$F(a_1, a_2) = \frac{1}{2}(a_1 + a_2 - 1)^2 + \frac{1}{2}(a_1 + a_2 - 2)^2$$

Lako se vidi da ova funkcija postiže globalni minimum  $F^* = 0.25$  na beskonačno mnogo vektora  $\mathbf{a}^* = (a_1^*, a_2^*)^T$ , čije komponente  $a_1, a_2$  zadovoljavaju jednadžbu

pravca  $\eta = \frac{3}{2} - \xi$ . Među njima vektor najmanje  $l_2$  norme dobit ćemo tako da potražimo minimum funkcije

$$\xi \mapsto \xi^2 + \left(\frac{3}{2} - \xi\right)^2.$$

On se postiže za  $\xi = \frac{3}{4}$ . Dakle, vektor najmanje  $l_2$  norme je  $\mathbf{a}^* = (0.74, 0.75)^T$  i to je rješenje postavljenog LPNK.

*Primjedba 6.5.* Ako je rang  $\mathbf{A} = k$ , onda dijagonalnu matricu  $\mathbf{S}$  iz SVD rastava (3.29), str. 77 možemo pisati kao  $\mathbf{S} = \sum_{i=1}^k \mathbf{S}_i$ , gdje je  $\mathbf{S}_i$  dijagonalna matrica koja na mjestu  $(i, i)$  ima element  $\sigma_i$ , a sve ostale nule. Zato SVD matrice  $\mathbf{A}$  možemo pisati

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U} \left( \sum_{i=1}^k \mathbf{S}_i \right) \mathbf{V}^T = \sum_{i=1}^k \mathbf{U}\mathbf{S}_i\mathbf{V}^T = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \quad (6.24)$$

Najbolja  $l_2$  aproksimacija ranga  $\rho < k$  matrice  $\mathbf{A}$ , tada je (vidi primjerice Björck (1996))

$$\tilde{\mathbf{A}} = \sum_{i=1}^{\rho} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{pri čemu je } \|\mathbf{A} - \tilde{\mathbf{A}}\|_2 = \sigma_{\rho+1}. \quad (6.25)$$

*Zadatak 6.1.* Kako izgleda rastav (6.24) za matricu  $\mathbf{A}$  iz *Primjera 3.12*, str. 79?

Ova činjenica može se upotrijebiti u svrhu kompresije slike (vidi primjerice Demmel (1997); Trefethen and Bau (1997)). Kao test primjer može se uzeti neka od "TestImage" iz programskog sustava *Mathematica* ili se može uzeti vlastiti test primjer.

Koristit ćemo crno-bijelu sliku dvorišta Odjela za matematiku Sveučilišta u Osijeku (vidi *Sliku 6.4a*), koja je zapisana je u obliku matrice veličine  $300 \times 300$  s elementima iz intervala  $[0, 1]$ , pri čemu 0 znači potpuno crno, a 1 potpuno bijelo.

```
In[1] := SetDirectory[NotebookDirectory[]];
        slika = Import["Kocka.pdf", ColorSpace -> "Graylevel"][[1]]
        Print["Dimenzije slike: ", slika[[2]]]
        podaci = ImageData[slika];
```

Uzimajući izabrani dio slike, definiramo matricu  $A$  i prikažimo sliku.



```

In[2]:= n=300; A=Table[0, {i, n}, {j, n}];
Do[
  Do[
    A[[i, j]] = podaci[[i + 15, j + 100, 1]]
      , {i, n}], {j, n}]
Image[A]

```

(a) Original



(b) 5% podataka



(c) 10% podataka



(d) 20% podataka



Slika 6.4: Kompresija slike

Aproksimacija s prvih  $\rho$ , ( $\rho = 15, 30, 60$ ) singularnih vrijednosti dobiva se na sljedeći način.

```

In[3]:= {u, sig, v} = SingularValueDecomposition[A, 30];
A1 = u.sig.Transpose[v];

```

Image [A1]

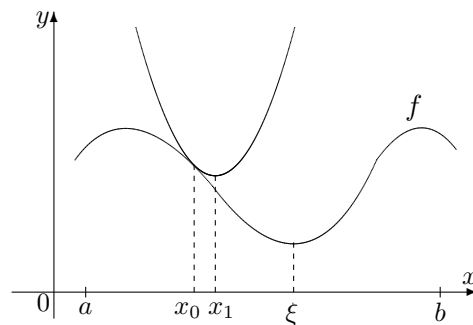
Aproksimirajući matricu  $A$  koristeći redom prvih  $\rho = 15, 30, 60$  singularnih vrijednosti u formuli (6.24) dobivamo slike aproksimacije originala prikazane na *Slikama 6.4 b-d*.

## 6.2 Nelinearni problemi najmanjih kvadrata

Ako funkcija-model  $f$  nije linearna u barem jednom od svojih parametara  $a_1, \dots, a_n$ , govorimo o nelinearnom problemu najmanjih kvadrata (NPNK). Ovaj problem rješavat ćemo kao nelinearni problem minimuma bez ograničenja. Najpoznatija metoda za rješavanje ovakvih problema je tzv. **Newtonova metoda** (Dennis and Schnabel, 1996; Ortega and Rheinboldt, 1970).

Radi ilustracije razmotrimo jednodimenzionalni problem minimizacije. Pretpostavimo da je  $f \in C^2[a, b]$  i da se u točki  $\xi$  postiže jedinstveni globalni minimum funkcije  $f$ . Izaberimo  $x_0 \in (a, b)$  i na osnovi Taylorove formule funkciju  $f$  u okolini točke  $x_0$  aproksimiramo kvadratnom funkcijom (vidi *Sliku 6.5*)

$$k_1(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2$$



Slika 6.5: Newtonova metoda minimizacije

Sljedeću aproksimaciju  $x_1$  tražene točke  $\xi$  birat ćemo tako da odredimo minimum kvadratne funkcije  $k_1$

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)}.$$

Ponavljajući postupak dobivamo niz  $x_0, x_1, \dots, x_n, \dots$  zadan rekurzivnom formulom

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}, \quad n = 0, 1, \dots$$

koji uz neke uvjete (vidi primjerice Dennis and Schnabel (1996)) konvergira prema  $\xi$ . Ovo je tzv. obična Newtonova metoda minimizacije. Primijetite da se ona podudara s Newtonovom metodom (t.4.4, str. 99) za rješavanje jednadžbe  $f'(x) = 0$ .

Analogno prethodnom, lokalni minimum dva puta neprekidno derivabilne funkcije  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  mogli bismo potražiti iterativnim postupkom (vidi primjerice Dennis and Schnabel (1996); Ortega and Rheinboldt (1970))

$$\mathbf{a}^{k+1} = \mathbf{a}^k - \left( \mathbf{H}_F(\mathbf{a}^k) \right)^{-1} \text{grad } F(\mathbf{a}^k), \quad \mathbf{a}^0 \in \mathbb{R}^n,$$

gdje je  $\mathbf{H}_F(\mathbf{a}^k)$  Hessijan funkcije  $F$  u točki  $\mathbf{a}^k$ . U cilju povećanja numeričke stabilnosti i osiguranja konvergencije, navedeni iterativni postupak modifikira se na sljedeći način:

$$\mathbf{a}^{k+1} = \mathbf{a}^k - \lambda_k \mathbf{s}^k, \quad \mathbf{a}^0 \in \mathbb{R}^n, \quad (6.26)$$

gdje je  $\mathbf{s}^k$  tzv. vektor smjera spusta, a dobiva se kao rješenje jednadžbe

$$\mathbf{H}_F(\mathbf{a}^k) \mathbf{s} = \text{grad } F(\mathbf{a}^k). \quad (6.27)$$

Broj  $\lambda_k \in (0, 1]$  je duljina koraka u smjeru vektora spusta koji je određen sljedećim algoritmom

---

**Algorithm 5** (Duljine koraka)

---

1. izabрати brojeve  $\tau \in (0, \frac{1}{2})$ ,  $\nu \in (0, 1)$ ;
2. Staviti  $\lambda = 1$ ;
3. Provjeriti nejednakost

$$F(\mathbf{a}^k - \lambda \mathbf{s}^k) - F(\mathbf{a}^k) < \tau \lambda (\mathbf{g}^k, \mathbf{s}^k), \quad (6.28)$$

gdje je  $\mathbf{g}^k = \text{grad } F(\mathbf{a}^k)$ , a  $(\mathbf{g}^k, \mathbf{s}^k)$  uobičajeni skalarni produkt;

4. Ako je nejednakost (6.28) ispunjena, staviti  $\lambda_k = \lambda$ ; Inače smanjiti duljinu koraka stavljajući  $\lambda := \lambda \nu$  i prijeći na Korak 3.
- 

Algoritam se ponavlja tako dugo dok se ne ispuni uvjet (6.28) jer na taj način osiguravamo smanjenje vrijednosti minimizirajuće funkcije ( $F(\mathbf{a}^{k+1}) <$

$F(\mathbf{a}^k)$ ) i konvergenciju niza aproksimacija  $(\mathbf{a}^k)$  prema lokalnom minimumu funkcije  $F$ . Iterativni postupak (6.26)-(6.27) nazivamo Newtonova metoda minimizacije

Uz neke uvjete na funkciju  $F$  i odgovarajući izbor početne aproksimacije  $\mathbf{a}^0$  Newtonova metoda konvergira prema lokalnom minimumu  $\mathbf{a}^*$  funkcije  $F$  (Dennis and Schnabel, 1996). Ako je Hessijan funkcije  $F$  u točki  $\mathbf{a}^k$  pozitivno definitna matrica, može se pokazati da je za  $\mathbf{g}^k \neq \mathbf{0}$  uvjet (6.28) uvijek moguće ispuniti. Ako je  $\mathbf{g}^k = \mathbf{0}$ , onda je  $\mathbf{a}^k$  stacionarna točka (pa i mogući kandidat za lokalni minimum) funkcije  $F$ .

*Zadatak 6.2.* Pomoću Newtonove metode pronađite lokalni minimum funkcije

$$f(x_1, x_2) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1).$$

$$\text{Rješenje: } x_1^* = 0.5, \quad x_2^* = -1, \quad f(x_1^*, x_2^*) = 0.$$

### 6.2.1 Gauss-Newtonova metoda

Za dane podatke  $(x_i, y_i)$ ,  $i = 1, \dots, m$ ,  $m \geq n$  i model-funkciju  $x \mapsto f(x; \mathbf{a})$  nelinearnu u vektoru parametara  $\mathbf{a} \in \mathbb{R}^n$  treba pronaći globalni minimum funkcije

$$F(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m (f(x_i; \mathbf{a}) - y_i)^2 = \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{a}) = \frac{1}{2} \|r(\mathbf{a})\|_2^2, \quad \mathbf{a} \in \mathbb{R}^n. \quad (6.29)$$

Prilikom korištenja Newtonove metode potrebno je poznavati sve prve i druge parcijalne derivacije minimizirajuće funkcije. Budući da je u našem slučaju minimizirajuća funkcija (6.29) specijalnog oblika, Newtonovu metodu moguće je prilagoditi i pojednostaviti.

Uz pretpostavku da je u blizini rješenja  $\mathbf{a}^*$  član  $\mathbf{J}^T(\mathbf{a})\mathbf{J}(\mathbf{a})$  dominantan u izrazu (6.7), Hessijan  $\mathbf{H}_F$  funkcije  $F$  možemo aproksimirati s

$$\mathbf{H}_F(\mathbf{a}) \approx \mathbf{J}^T(\mathbf{a})\mathbf{J}(\mathbf{a}). \quad (6.30)$$

Gradijent minimizirajuće funkcije zadan je s (6.6). Na taj način u Newtonovom iterativnom postupku (6.26)-(6.27) vektor smjera spusta  $\mathbf{s}^k$  dobivamo rješavajući sustav linearnih jednadžbi

$$\mathbf{J}^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k)\mathbf{s}^k = \mathbf{J}^T(\mathbf{a}^k)\mathbf{r}(\mathbf{a}^k). \quad (6.31)$$

Uz pretpostavku da je rang  $\mathbf{J}(\mathbf{a}^k) = n$ , matrica  $\mathbf{J}^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k)$  je pozitivno definitna, sustav (6.31) ima jedinstveno rješenje, a uvjet (6.28) iz kojeg biramo duljinu koraka  $\lambda_k$  uvijek se može ispuniti.

Sustav (6.31) možemo shvatiti kao sustav normalnih jednadžbi za LPNK

$$\mathbf{J}(\mathbf{a}^k) \mathbf{s}^k \simeq \mathbf{r}(\mathbf{a}^k). \quad (6.32)$$

Na taj način rješavanje NPNK Gauss-Newtonovom metodom svodi se na rješavanje niza linearnih problema najmanjih kvadrata, koje je najbolje rješavati primjenom QR-dekompozicije (vidi t.6.1.2, str. 154).

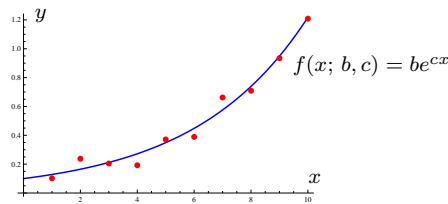
**Primjer 6.10.** Zadani su podaci  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , gdje je

$$x_i = i, \quad y_i = 0.1e^{0.25x_i} + \epsilon_i, \quad \text{gdje je } \epsilon_i \sim N[0, \sigma^2].$$

Treba pronaći optimalne parametre  $b^*$ ,  $c^*$  funkcije-modela  $f(x; b, c) = be^{cx}$ .

Niže navodimo odgovarajući *Mathematica*-program i dobivene rezultate. Najprije definiramo, model-funkciju i podatke i crtamo odgovarajuću sliku

```
In[1]:= m=10; b=.1; c=.25; SeedRandom[17];
f[x_,b_,c_] := b Exp[c x];
slf=Plot[f[x,b,c], {x,0,m}, PlotStyle->{Blue,Thickness[.005]}];
F[x_,y_,b_,c_] := Plus@@((y - f[x, b, c])^2);
x = Table[i, {i, m}];
y=Table[f[i,b,c]+Random[NormalDistribution[0,.05]], {i,m}];
pod = Table[{x[[i]], y[[i]]}, {i, m}];
slpod=ListPlot[pod, PlotStyle->{Red, PointSize[.02]}],
PlotRange->{0,1.5}];
Show[slf, slpod]
```



Nakon toga primjenom Gauss-Newtonove metode odredimo optimalne parametre model funkcije minimizacijom odgovarajuće funkcije (6.29)

```
In[2]:= b0=.5; c0=.1; F0=F[x,y,b0,c0]; tau=.5; eps=.005; k=0;
Print["It ", 0, ": b = ", b0, " c = ", c0, " F = ", F0]
While[b = b0; c = c0; lam = 1; k = k + 1;
jac = Table[{f[x[[i]],1,c], x[[i]] f[x[[i]],b,c]}, {i,m}];
While[q = -Transpose[jac].(y - f[x, b, c]);
p = -Inverse[Transpose[jac].jac].q;
{b0, c0} = {b, c} + lam p; F1 = F[x, y, b0, c0];
```

```

F1 - F0 > -tau lam Plus@@(p q), lam = lam/2.
];
norm = Sqrt[Abs[Plus@@(q^2)]]; norm > eps,
Print["It ",k,": b = ",b0," c = ",c0," F = ", F1,
      " lam=", lam, " grad= ", norm]
];

```

Iterativni proces teče kako slijedi:

```

It 0: b = 0.5 c = 0.1 F = 1.74005
It 1: b = 0.263802 c = 0.139479 F = 0.268953 lam=0.5 grad=20.5445
It 2: b = 0.0625254 c = 0.226529 F = 0.887862 lam=1 grad=2.18007
It 3: b = 0.0973459 c = 0.271372 F = 0.149347 lam=1 grad=16.6483
It 4: b = 0.0979424 c = 0.253021 F = 0.0255789 lam=1 grad=10.6933
It 5: b = 0.0987964 c = 0.250451 F = 0.0248364 lam=1 grad=0.687365

```

**Primjer 6.11.** Promatramo problem iz Primjera 6.5, str. 148. Podaci o koordinatama točaka  $\mathbf{S}_i(x_i, y_i)$ ,  $i = 1, 4$ , kao i kutevima  $\alpha_i$ ,  $i = 1, 4$  (u stup-njevima i radijanima prenešenim u interval  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ) vide se u Tablici 6.2.

Za ovaj problem najmanjih kvadrata vektor parametara je  $\mathbf{a} = (x, y)^T$ , kompo-nente  $r_i$  ( $i = 1, 4$ ) vektora reziduala su

$$r_i = f(x_i, y_i; x, y) - \alpha_i, \quad \text{gdje je } f(x_i, y_i; x, y) = \arctg \frac{y-y_i}{x-x_i}, \quad (a)$$

a elementi Jacobijeve matrice su

$$\mathbf{J}_{i1} = \frac{y_i - y}{(x - x_i)^2 + (y - y_i)^2}, \quad \mathbf{J}_{i2} = \frac{x - x_i}{(x - x_i)^2 + (y - y_i)^2}, \quad i = 1, 4.$$

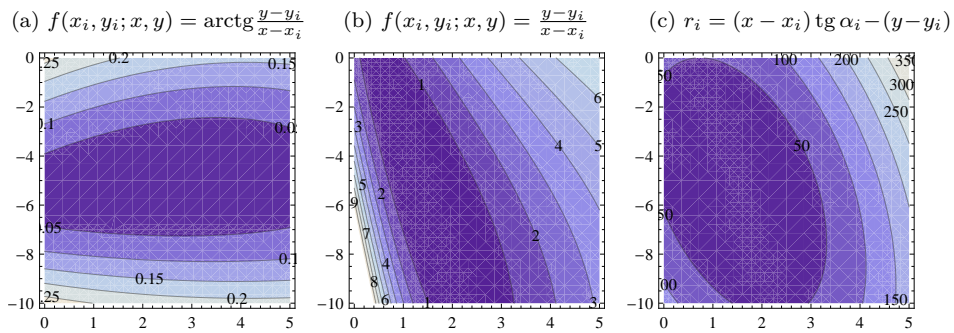
Iteracija $x_i$	Podaci			Procjena na bazi podataka			
	$y_i$	$\alpha_i(^{\circ})$	$\alpha_i(rad)$	$\alpha_i(\bar{rad})$	$\alpha_i(\bar{^{\circ}})$	$r_i$	Odst. u %
13	7	50	0.8727	0.8606	49.3	0.0121	0.2%
-2	14	100	-1.396	-1.328	103.9	-0.0685	0.5 %
-7	-7	190	0.1745	0.2059	191.8	-0.0314	0.4 %
15	-11	330	-0.5236	-0.4558	333.9	-0.0678	0.6 %

Tablica 6.2: Traženje pozicije broda

Problem ćemo riješiti Gauss-Newtonovom metodom primjenom ranije spome-nutih programa. Uz početnu aproksimaciju  $\mathbf{a}^0 = (2, -3)^T$  i točnost  $\epsilon = 0.00001$ , nakon 3 iteracije dobivene su koordinate tražene točke  $\mathbf{B}_1 = (2.705558, -4.972401)$ . Tijek iterativnog procesa prikazan je u Tablici 6.3. Proces se zaustavlja kada se postigne  $\|\mathbf{g}^k\| < \epsilon$ .

$k$	$x^k$	$y^k$	$\ \mathbf{g}^k\ $	$F(x_k, y_k)$
0	2	-3	0.033842	0.040778
1	2.62135	-5.052535	0.001222	0.005276
2	2.696991	-4.973894	0.000049	0.005214
3	2.705558	-4.972401	0.000004	0.005214

Tablica 6.3: Empirijski podaci i procijenjene vrijednosti

Slika 6.6: ContourPlot funkcije (6.29) za različite model funkcije  $f$ 

U *Tablici 6.2* mogu se također vidjeti veličine “teoretskih kuteva”  $\bar{\alpha}_i$  (u radijanima i stupnjevima), pojedini reziduali (odstupanja) i relativna odstupanja u %.

Na *Slici 6.6.a* prikazane su tzv. izo-krivulje (ContourPlot) odgovarajuće minimizirajuće funkcije  $F$ , pri čemu je model funkcija zadana s (a).

Isti problem mogli bismo riješiti i tako da komponente  $(r_i, i = 1, \dots, 4)$  vektora reziduala zadamo s

$$r_i = f(x_i, y_i; x, y) - \operatorname{tg} \alpha_i, \quad \text{gdje je} \quad f(x_i, y_i; x, y) = \frac{y-y_i}{x-x_i}. \quad (b)$$

Tada se minimum postiže u točki  $\mathbf{B}_2 = (1.44266, -5.49173)$ , pri čemu dobivamo vrijednost funkcije  $F^* = 0.0208275$ , a izo-krivulje odgovarajuće minimizirajuće funkcije vide se na *Slici 6.6.b*

Ako komponente  $r_i (i = 1, 4)$  vektora reziduala zadamo s

$$r_i = (x - x_i) \operatorname{tg} \alpha_i - (y - y_i), \quad (c)$$

tada se radi o LPNK. Minimum se postiže u točki  $\mathbf{B}_3 = (1.49565, -5.3101)$ , a vrijednost funkcije  $F$  je  $F^* = 3.35062$ . Izo-krivulje odgovarajuće minimizirajuće funkcije (vidi *Sliku 6.6.c*) sada su koncentrične elipse.

Koji pristup: (a), (b) ili (c) ćemo primijeniti ovisi o drugim zahtjevima i dodatnoj analizi.

### 6.2.2 Marquardtova metoda

U Gauss-Newtonovom iterativnom postupku mogu nastupiti problemi onda ako je rang  $J(\mathbf{a}^k) < n$  za neki  $k$ . Naime, tada je matrica  $J^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k)$  singularna. U tom slučaju, Hessijan  $\mathbf{H}(\mathbf{a}^k)$  umjesto s (6.30), aproksimirat ćemo s

$$\mathbf{H}_F(\mathbf{a}) = \mathbf{J}(\mathbf{a})^T\mathbf{J}(\mathbf{a}) + \mu\mathbf{I}, \quad \mu > 0, \quad (6.33)$$

gdje je  $\mathbf{I}$  jedinična matrica. Vektor smjera spusta bit će rješenje jednadžbe

$$(\mathbf{J}^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k) + \mu_k\mathbf{I})\mathbf{s}^k = \mathbf{J}^T(\mathbf{a}^k)\mathbf{r}(\mathbf{a}^k). \quad (6.34)$$

Sustav (6.34) možemo shvatiti kao sustav normalnih jednadžbi za LPNK

$$\begin{bmatrix} \mathbf{J}(\mathbf{a}^k) \\ \sqrt{\mu_k}\mathbf{I} \end{bmatrix} \mathbf{s}^k \simeq \begin{bmatrix} \mathbf{r}(\mathbf{a}^k) \\ 0 \end{bmatrix}, \quad (6.35)$$

Strategija biranja parametra  $\mu$  sastoji se u tome da uvijek bude ispunjen uvjet (6.28), a da se pri tome nastojimo zadržati na što manjoj vrijednosti parametra  $\mu$ . Naime, u tom slučaju vektor smjera spusta blizak je vektoru smjera spusta kod Gauss-Newtonove metode, što osigurava bržu konvergenciju procesa. Kada  $\mu \rightarrow +\infty$ , vektor smjera spusta približava se smjeru antigradijenta, što također osigurava konvergenciju procesa, ali znatno sporiju (Schwarz and Köckler, 2011).

*Primjedba 6.6.* Kod gotovo svih iterativnih metoda, pa tako i kod navedenih metoda za rješavanje *NPNK*, važan je izbor kvalitetne početne aproksimacije. Pri tome, mogu se koristiti neki uvjeti iz teorema o egzistenciji rješenja, mogu se koristiti rezultati dobiveni logaritmiranjem ili nekim drugim postupkom (Scitovski, 1993), a u slučaju  $n = 2$  (funkcija-model ima dva nezavisna parametra) mogu se iskoristiti grafičke mogućnosti računala (programski paketi: *Mathematica*, *MATLAB* i sl.)

*Primjedba 6.7.* Više o problemima najmanjih kvadrata, njihovom numeričkom rješavanju i praktičnim primjenama može se naći kod Björck and Dahlquist (2007); Gill et al. (1991, 1981); Golub and van Loan (1996); Scitovski (1993). Spomenimo da se programska podrška navedena kod (Press et al., 1992) nažalost, ne može uspješno koristiti za sve probleme najmanjih kvadrata, koji se mogu javiti u primjenama. Programi za rješavanje nelinearnih problema najmanjih kvadrata ovdje su napisani tako da se u svakom koraku rješava sustav normalnih jednadžbi, što ozbiljno umanjuje stabilnost numeričkog procesa. Za praktičnu primjenu bolje je preporučiti *The Nag-Library* ili *Mathematica Package: Statistics'NonlinearFit'*.



### 6.3 Druge mogućnosti

Optimalne parametre model funkcije možemo potražiti na osnovi zadanih podataka korištenjem i nekih drugih principa, kao što je princip najmanje sume apsolutnih odstupanja ili princip najmanjeg maksimalnog apsolutnog odstupanja (Schöbel, 1999; Watson, 1980), kao što smo spomenuli u *Primjedbi 6.1, str.147*.

#### 6.3.1 Princip najmanje sume apsolutnih odstupanja

Pretpostavimo da su zadani podaci  $(x_i, y_i)$ ,  $i = 1, \dots, m$  na osnovi kojih treba procijeniti vektor optimalnih parametara  $\mathbf{a}^* \in \mathbb{R}^n$ ,  $n \ll m$  model funkcije  $y = f(x; \mathbf{a})$  primjenom principa najmanjih apsolutnih odstupanja<sup>3</sup>, tj. treba riješiti optimizacijski problem

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \sum_{i=1}^m |f(x_i; \mathbf{a}) - y_i|. \quad (6.36)$$

Ako uvedemo vektor reziduala  $\mathbf{r}$  kao u prethodnoj točki *t.6.3.2*, onda funkciju  $F$  možemo zapisati kao

$$F(\mathbf{a}) = \|\mathbf{r}(\mathbf{a})\|_1.$$

Zato se u literaturi ovaj princip traženja optimalnog parametra  $\mathbf{a}^* \in \mathbb{R}^n$  model funkcije  $y = f(x; \mathbf{a})$  naziva još i traženje (procjena) optimalnih parametara kao najbolje  $l_1$  aproksimacije (Cadzow, 2002; Watson, 1980).

Princip ćemo samo ilustrirati na jednom jednostavnom primjeru procjene parametara linearne model funkcije  $f(x; \alpha) = \alpha x$  iz Hookeovog zakona o rastezanju elastičnog tijela<sup>4</sup>.

Uz oznake  $I := \{1, \dots, m\}$  i  $I_0 := \{i \in I : x_i = 0\}$ , naš problem prelazi u sljedeći optimizacijski problem

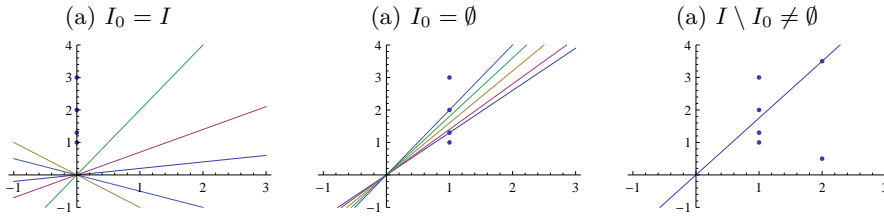
$$\operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{i \in I} |y_i - \alpha x_i| = \sum_{i \in I_0} |y_i| + \operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{i \in I \setminus I_0} |x_i| \left| \frac{y_i}{x_i} - \alpha \right|, \quad (6.37)$$

koji je u literaturi poznat kao *težinski medijan problem* (Gurwitz, 1990; Rousseeuw and Leroy, 2003; Sabo and Scitovski, 2008). Kod rješavanja problema (6.37) razlikujemo tri slučaja:

- (a) Ako je  $I_0 = I$ , svaki  $\alpha \in \mathbb{R}$  je rješenje problema (6.37) (Slika 6.7a);

<sup>3</sup>En. Least Absolute Deviations (LAD).

<sup>4</sup>Robert Hooke (1635 - 1703) engleski fizičar – bio je tajnik Royal Society.



Slika 6.7: Rješenja optimizacijskog problema (6.37)

(b) Ako je  $I_0 = \emptyset$ , pri čemu je  $x_1 = \dots = x_m \neq 0$ , problem (6.37) postaje

$$\operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{i \in I} |y_i - \alpha x_i| = |x_1| \min_{\alpha \in \mathbb{R}} \sum_{i \in I} \left| \frac{y_i}{x_1} - \alpha \right|,$$

a broj  $\alpha^* = \operatorname{med} \left( \frac{y_i}{x_1} \right) = \frac{1}{x_1} \operatorname{med} y_i$  je rješenje problema (6.37) (Slika 6.7b);

(c) Ako je  $I \setminus I_0 \neq \emptyset$ , pri čemu  $0 < x_1 < x_m$ , rješenje problema (6.37) je  $\alpha^* = \operatorname{med} \left( |x_i|, \frac{y_i}{x_i} \right)$  (Slika 6.7c).

Ako je model funkcija opća linearna funkcija  $f(x; \alpha, \beta) = \alpha x + \beta$ , problem (6.36) je složeniji i ne može se direktno riješiti. U literaturi se mogu pronaći različite iterativne metode koje se zasnivaju na činjenici da postoji najbolja linearna funkcija oblika  $f(x; \alpha, \beta) = \alpha x + \beta$ , čiji graf prolazi kroz barem dvije točke podataka (Sabo and Scitovski, 2008; Schöbel, 1999; Watson, 1980) (vidi primjerice *Metodu dviju točaka* u radu Sabo and Scitovski (2008)).

Problem određivanja najbolje ravnine u prostoru na osnovi poznatih točaka-podataka  $T_i = (x_i, y_i, z_1)$ ,  $i = 1, \dots, m$  u prostoru, a uz primjenu LAD-principa koristi se primjerice u robotici (Cupec, 2005). U literaturi se mogu pronaći različite iterativne metode koje se zasnivaju na činjenici da postoji najbolja LAD-ravnina koja prolazi kroz barem tri točke podataka (Fischler and Bolles, 1981; Schöbel, 1999; Watson, 1980) (vidi primjerice *Metodu tri točke* u radu Cupec et al. (2009)).

Općenito se problem (6.36) može promatrati za linearnu funkciju  $f(\mathbf{x}; \mathbf{a}) = \mathbf{a}^T \mathbf{x}$  čiji graf zovemo hiperravnina. Ovaj problem ekvivalentan je problemu rješavanja predodređenog sustava linearnih jednadžbi u smislu da suma apsolutnih odstupanja lijevih od desnih strana jednadžbi bude minimalna, tj. za dane podatke  $(\mathbf{x}_i, z_i)$ ,  $\mathbf{x}_i = (x_1^{(i)}, \dots, x_n^{(i)})^T \in \mathbb{R}^n$ ,  $z_i \in \mathbb{R}$ ,  $i = 1, \dots, m$ , promatra se ovakav optimizacijski problem

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \|\mathbf{z} - \mathbf{X}\mathbf{a}\|_1 = \sum_{i=1}^m |z_i - \mathbf{a}^T \mathbf{x}_i|, \quad (6.38)$$

gdje je  $\mathbf{X} \in \mathbb{R}^{m \times n}$  matrica punog ranga čiji su retci vektori  $\mathbf{x}_i$  (Sabo et al., 2011).

*Primjedba 6.8.* Princip najmanje sume apsolutnih odstupanja pripisuje se hrvatskom znanstveniku J.R.Boškoviću<sup>5</sup>, koji je ovaj princip iznio još 1757. godine u radu Boscovich (1757). Dugo vremena ovaj princip je zapostavljan u odnosu na Gaussov princip najmanjih kvadrata zbog složenosti računskih procesa. Tek dolaskom modernih računala ovaj princip ponovo je zauzeo važno mjesto u znanstvenim istraživanjima, posebno zbog svojstva svoje robusnosti: ovaj princip, za razliku od Gaussovog principa najmanjih kvadrata, u značajnoj mjeri ignorira jako stršeće podatke (“outliers”) u skupu podataka. U Švicarskom gradu Neuchâtelu još uvijek se redovito održavaju znanstvene konferencije posvećene  $l_1$  metodama i primjenama, a na naslovnici zbornika radova nalazi se slika hrvatske novčanice s likom J.R.Boškovića (Dodge, 1997).

### 6.3.2 Princip najmanjeg maksimalnog apsolutnog odstupanja

Pretpostavimo da su zadani podaci  $(x_i, y_i)$ ,  $i = 1, \dots, m$  na osnovi kojih treba procijeniti vektor optimalnih parametara  $\mathbf{a}^* \in \mathbb{R}^n$ ,  $n \ll m$  model funkcije  $y = f(x; \mathbf{a})$  primjenom principa traženja najmanjeg maksimalnog apsolutnog odstupanja, tj. treba riješiti optimizacijski problem

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \max_{i=1, \dots, m} |f(x_i; \mathbf{a}) - y_i|. \quad (6.39)$$

Ako uvedemo vektor reziduala  $\mathbf{r}$  s komponentama

$$r_i = f(x_i; \mathbf{a}) - y_i, \quad i = 1, \dots, m,$$

onda funkciju  $F$  možemo zapisati kao

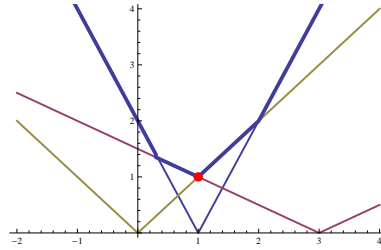
$$F(\mathbf{a}) = \|\mathbf{r}(\mathbf{a})\|_\infty.$$

Zato se u literaturi ovaj princip traženja optimalnog parametra  $\mathbf{a}^* \in \mathbb{R}^n$  model funkcije  $y = f(x; \mathbf{a})$  naziva još i traženje (procjena) optimalnih parametara kao najbolje  $l_\infty$  aproksimacije (Cadzow, 2002; Watson, 1980).

Princip ćemo samo ilustrirati na jednostavnom primjeru procjene parametara linearne model funkcije  $f(x; \alpha) = \alpha x$  iz Hookeovog zakona.

Pretpostavimo da su poznati rezultati mjerenja  $(x_i, y_i)$ ,  $i = 1, \dots, m$  na osnovi kojih treba procijeniti parametar  $\alpha \in \mathbb{R}$  linearne model-funkcije

<sup>5</sup>Josip Ruder Bošković (1711-1787)

Slika 6.8: Najbolje  $l_\infty$  rješenje

$f(x; \alpha) = \alpha x$  prema principu najmanjeg maksimalnog apsolutnog odstupanja. Optimizacijski problem (6.39) u ovom slučaju glasi

$$\operatorname{argmin}_{\alpha \in \mathbb{R}^n} F(\alpha), \quad F(\alpha) = \|\mathbf{r}(\alpha)\|_\infty = \max\{|y_1 - \alpha x_1|, \dots, |y_m - \alpha x_m|\}.$$

Primjerice, za podatke  $(2, 2)$ ,  $(\frac{1}{2}, \frac{3}{2})$ ,  $(-1, 0)$  imamo

$$F(\alpha) = \max\{|2 - 2\alpha|, |\frac{3}{2} - \frac{1}{2}\alpha|, |\alpha|\}.$$

Minimum funkcije  $F$  postiže se u točki  $\alpha^* = 1$  i iznosi  $F(\alpha^*) = 1$  (vidi Sliku 6.8).

## 6.4 Zadaci

*Zadatak 6.3.* U 4-znamenkastoj floating-point aritmetici preko sustava normalnih jednadžbi riješite

$$\begin{array}{rcl} \epsilon x_1 & & = 20 \\ & \epsilon x_2 & = 30 \\ & & \epsilon x_3 = 40 \\ x_1 + x_2 + x_3 & = & 10 \end{array}$$

a)  $\epsilon = 0.1$       b)  $\epsilon = 0.01$       c)  $\epsilon = 0.001$

Egzaktno rješenje je

$$\mathbf{x} = \left( \frac{10(-3+\epsilon+2\epsilon^2)}{3\epsilon+\epsilon^3}, \frac{10(1+3\epsilon)}{3+\epsilon^2}, \frac{10(3+\epsilon+4\epsilon^2)}{3\epsilon+\epsilon^3} \right)^T$$

*Zadatak 6.4.* Riješite LPNK  $\mathbf{J}\mathbf{a} \simeq \mathbf{y}$ , gdje je

$$\mathbf{J}^T = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad \mathbf{y}^T = [1, 2, 3, 1, 2, 1].$$

Rješenje:  $\mathbf{a} = \left(3, \frac{7}{4}, \frac{5}{4}\right)^T$ .

*Zadatak 6.5.* Pokušajte riješiti *Zadatak 6.3* i *Zadatak 6.4* rješavanjem sustava normalnih jednažbi pomoću Cholesky dekompozicije.

*Zadatak 6.6.* Olovnu kuglicu puštamo da pada u vakuumu sa različitih visina  $h$  i mjerimo vrijeme dok kuglica ne padne na tlo

$h(m)$	7.5	10	15	20
$t(s)$	1.23	1.44	1.75	2.01

Poznat je zakon gibanja:  $h = \frac{1}{2}at^2$ . Primjenom metode najmanjih kvadrata procijenite veličinu parametra  $a$ .

Rješenje:  $a = 9.8$

*Zadatak 6.7.* Umjesto interpolacijskog polinoma u *Zadatku 2.11*, str. 45, primjenom metode najmanjih kvadrata odredite odgovarajući polinom 2. stupnja.

Rješenje:  $\alpha(X) = 6.810254 - 0.00454X + 610^{-7}X^2$ .

$X(m)$	$\alpha$	$\alpha^*$	$\Delta\alpha^*$
2000	6.296	6.507	0.211
3000	9.543	9.613	0.070
5000	16.512	15.407	1.105
7000	24.865	24.819	0.046
9000	39.407	42.858	3.451

*Zadatak 6.8.* Umjesto interpolacijskog polinoma u *Zadatku 2.12*, str. 46, za funkciju  $t \mapsto h(t)$  primjenom metode najmanjih kvadrata odredite polinom 2.stupnja.

Rješenje:  $h(t) = -0.060599 + 0.23098t + 4.7645t^2$

*Zadatak 6.9.* Mjerenjem je ustanovljeno da električni otpor  $R$  bakrene žice ovisi o temperaturi  $t$  na sljedeći način

$t^\circ C$	19.1	25	30.1	36	40	45.1	50
$R(\Omega)$	76.3	77.8	79.75	80.8	82.35	83.9	85.1

Uz pretpostavku da se radi o linearnoj ovisnosti, odredite parametre  $a$  i  $b$  linearne funkcije  $R(t) = a + bt$ .

Rješenje:  $R(t) = 70.76228 + 0.288072t$

*Zadatak 6.10.* Zadana je funkcija model  $f(x; a_1, a_2) = a_1 + a_2x$  i podaci  $(x_i, y_i)$ ,  $i = 1, \dots, n$ . Pokažite da odgovarajući PNK ima jedinstveno rješenje ako je  $x_i \neq x_j$  za barem jedan par indeksa  $i, j$ .

*Zadatak 6.11.* Ako je  $\mathbf{A}$  kvadratna regularna matrica, pokažite da se tada pseudoinverzna matrica  $\mathbf{A}^+$  podudara s  $\mathbf{A}^{-1}$ .

*Zadatak 6.12.* Za vektor  $\mathbf{x} = (-4, 1, 4, 2, 5, 1, 1)^T$  odredite Householderov vektor  $\mathbf{u}_k$ ,  $k = 1, \dots, 7$ , tako da je samo  $k$ -ta komponenta transformiranog vektora  $\mathbf{H}\mathbf{x}$  različita od nule.

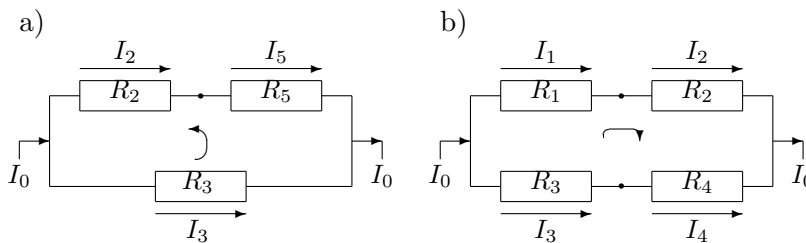
*Zadatak 6.13.* Neka je  $\mathbf{H}$  Householderova matricu s Householderovim vektorom  $\mathbf{u}$ . Pokažite da vrijedi

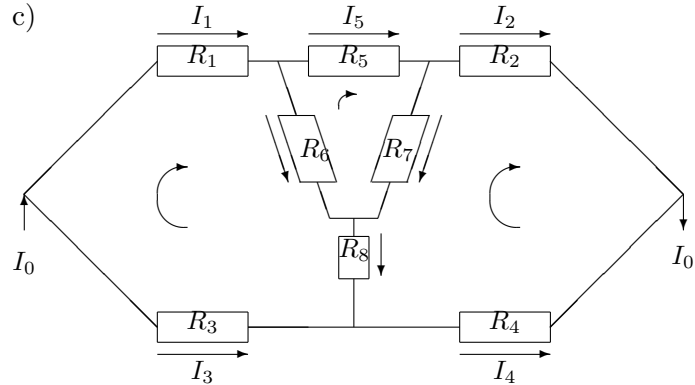
- ako je  $\mathbf{s}$ , vektor okomit na vektor  $\mathbf{u}$ , tada je  $\mathbf{H}\mathbf{s} = \mathbf{s}$ ;
- ako je  $\mathbf{z}$ , vektor paralelan s vektorom  $\mathbf{u}$ , tada je  $\mathbf{H}\mathbf{z} = -\mathbf{z}$ ;
- ako je  $\mathbf{x} = \mathbf{s} + \mathbf{z}$ , gdje je  $\mathbf{s}$  vektor okomit na vektor  $\mathbf{u}$ , a  $\mathbf{z}$  vektor paralelan s vektorom  $\mathbf{u}$ , tada je  $\mathbf{H}\mathbf{x} = \mathbf{s} - \mathbf{z}$ .

Kako se zove geometrijska transformacija koja vektor  $\mathbf{a}$  prevodi u  $\mathbf{H}\mathbf{a}$  ?

*Zadatak 6.14.* Za mreže prikazane na slikama a), b) i c) izračunajte veličine struja ako je

- $I_0 = 5A$ ,  $R_1 = 50\Omega$ ,  $R_2 = 10\Omega$ ,  $R_3 = 5\Omega$
- $I_0 = 10A$ ,  $R_1 = 100\Omega$ ,  $R_2 = 10\Omega$ ,  $R_3 = 20\Omega$ ,  $R_4 = 30\Omega$
- $I_0 = 10A$ ,  $R_1 = 5\Omega$ ,  $R_2 = 50\Omega$ ,  $R_3 = 5\Omega$ ,  $R_4 = 50\Omega$ ,  
 $R_5 = 10\Omega$ ,  $R_6 = 12\Omega$ ,  $R_7 = 20\Omega$ ,  $R_8 = 30\Omega$





Rješenje:

- a)  $I_1 = 0.38462A$ ,  $I_2 = 0.38462A$ ,  $I_3 = 4.61539A$   
 b)  $I_1 = 3.125A$ ,  $I_2 = 3.125A$ ,  $I_3 = 6.875A$ ,  $I_4 = 6.875A$   
 c)  $I_1 = 4.485A$ ,  $I_2 = 4.6996A$ ,  $I_3 = 5.515A$ ,  $I_4 = 5.3A$   
 $I_5 = 3.519A$ ,  $I_6 = 0.9657A$ ,  $I_7 = 1.1803A$ ,  $I_8 = 0.2146A$

Zadatak 6.15. Pokažite da je

$$\mathbf{Q} = \begin{bmatrix} \cos(-60^\circ) & \sin(-60^\circ) \\ -\sin(-60^\circ) & \cos(-60^\circ) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

ortogonalna matrica. Također, pokažite da je  $\mathbf{Q}^T(1, \sqrt{3})^T = (2, 0)^T$ . Kako se zove geometrijska transformacija koja vektor  $\mathbf{a}$  prevodi u vektor  $\mathbf{Q}^T \mathbf{a}$ ?

Zadatak 6.16. Pokažite da je

$$\mathbf{Q} = \begin{bmatrix} \cos 30^\circ & \sin 30^\circ \\ \sin 30^\circ & -\cos 30^\circ \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}$$

ortogonalna simetrična matrica. Također, pokažite da je  $\mathbf{Q}^T(1, \sqrt{3})^T = (\sqrt{3}, -1)^T$ . Kako zovemo geometrijsku transformaciju koja vektor  $\mathbf{a}$  prevodi u  $\mathbf{Q}^T \mathbf{a}$ ?

Zadatak 6.17. Zadane su točke:  $S_1(8, 6)$ ,  $S_2(-4, 5)$ ,  $S_3(1, -3)$  i kutevi  $\alpha_1 = 42^\circ$ ,  $\alpha_2 = 158^\circ$ ,  $\alpha_3 = 248^\circ$  pod kojima se “vidi” tražena točka  $B$  kao u *Primjeru 6.5*, str. 148, odnosno *Primjeru 6.11*, str. 163. Izračunajte koordinate točke  $B$  prema sve tri varijante navedene u *Primjeru 6.11*.

Rješenje:

*Zadatak 6.18.* Zadani su podaci  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , na osnovi kojih su metodom najmanjih kvadrata određeni parametri  $k$  i  $l$  linearne regresije  $f(x) = ax + b$ .

a) Pokažite da vrijedi

$$a = \frac{1}{D} (m \sum x_k y_k - \sum x_k \sum y_k),$$

$$b = \frac{1}{D} (\sum x_k^2 \sum y_k - \sum x_k \sum x_k y_k), \quad D = m \sum x_k - (\sum x_k)^2,$$

b) Ako je:  $\bar{x} = \frac{1}{m} \sum x_k$ ,  $\bar{y} = \frac{1}{m} \sum y_k$ , pokažite da točka  $(\bar{x}, \bar{y})$  leži na pravcu  $y = ax + b$ . Također, pokaži da vrijedi

$$a = \frac{1}{D} \sum (x_k - \bar{x})(y_k - \bar{y}), \quad b = \bar{y} - a\bar{x}, \quad D = \sum (x_k - \bar{x})^2.$$

c) Kako glase formule za  $a$  i  $b$  ako središte koordinatnog sustava postavimo u točku  $(\bar{x}, \bar{y})$  ?

*Zadatak 6.19.* Kako se problem procjene parametara u niže navedenim funkcijama-modelima može linearizirati (vidi SCITOVSKI (1993))

$$a) y = \frac{a}{x} + b, \quad b) y = \frac{a}{x+c}, \quad c) y = \frac{1}{ax+b},$$

$$d) y = \frac{x}{a+bx}, \quad e) y = a \ln x + b, \quad f) y = a \exp(bx),$$

$$g) y = \frac{1}{(ax+b)^2}, \quad h) y = ax \exp(-bx), \quad i) y = \frac{A}{(1+b \exp(-cx))}$$

Za proizvoljno izabrane podatke, parametre u modelima procijenite kao NPNK i kao LPNK dobiven linearizacijom. Usporedite dobivene parametre te nacrtajte odgovarajuće izo-krivulje. Također, grafički prikažite podatke i dobiveni regresijski model.

*Zadatak 6.20.* Newtonovom metode minimizacije odredite lokalne minimume funkcije iz *Zadatka 4.16*, str. 120.

*Zadatak 6.21.* Pokažite da funkcija  $F(x, y) = (x - y)^2(x + y)^2$ ,  $x, y \in \mathbb{R}$  postiže svoj globalni minimum  $F^* = 0$  u beskonačno mnogo točaka. Odredite geometrijska mjesta tih točaka. Graf funkcije  $F$  prikazan je na naslovnoj stranici ovog udžbenika.

*Zadatak 6.22.* Analizirajte problem minimuma funkcije  $F$ . Uz koje uvjete je njezin Hessijan pozitivno definitan, a uz koje pozitivno semidefinitan? Kakve to posljedice ima na egzistenciju lokalnog i globalnog minimuma? Nacrtajte graf funkcije  $F$  i njene izo-krivulje.

$$a) F(x, y) = (ax + by + c)^2, \quad a, b, c \in \mathbb{R}.$$



$$b) F(x, y) = (a_1x + b_1y + c_1)^2 + (a_2x + b_2y + c_2)^2, \quad a_i, b_i \in \mathbb{R}$$

Kako se mogu generalizirati navedeni problemi?

Uputa: Pokažite da su glavni minori Hesseove matrice funkcije  $F$

$$a) \Delta_1 = 2a^2, \quad \Delta_2 = 0;$$

$$b) \Delta_1 = 2(a_1^2 + a_2^2), \quad \Delta_2 = 2(a_1b_2 + a_2b_1)^2.$$

*Zadatak 6.23.* Riješite sljedeće LPNK

$$a) ax + by + c \simeq 0 \qquad b) \begin{array}{l} a_1x + b_1y + c_1 \simeq 0 \\ a_2x + b_2y + c_2 \simeq 0 \end{array}$$

$$\text{Rješenje: } a) x = \frac{ac}{a^2+b^2}, \quad y = \frac{bc}{a^2+b^2}, \quad ab \neq 0.$$

# Bibliografija

- A. S. ACKLEH, E. J. ALLEN, R. B. KEARFOTT, P. SESHAIYER, *Classical and Modern Numerical Analysis, Theory, Methods and Practice*, Chapman and Hall/CRC, 2010.
- F. S. ACTON, *Numerical Methods That Work*, MAA, Washington, 1990.
- K. ATKINSON, *An Introduction to Numerical Analysis*, John Wiley & Sons, Canada, 1998.
- M. BENŠIĆ, N. ŠUVAK, *Primijenjena statistika*, Odjel za matematiku,, 2012.
- A. BJÖRCK, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.
- A. BJÖRCK, G. DAHLQUIST, *Numerical Methods in Scientific Computing (Volume I)*, volume II, SIAM, Philadelphia, 2007.
- D. BLANUŠA, *Viša matematika*, Tehnička knjiga, Zagreb, 1973.
- R. J. BOSCOVICH, *De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis, ac habentur plura eius ex exemplaria etiam sensorum impressa*, Bononienci Scientiarum et Artium Znstituto Atque Academia Commentarii, 4(1757) 353–396.
- F. M. BRÜCKLER, *Povijest matematike I*, Odjel za matematiku, Sveučilište u Osijeku, 2007.
- J. A. CADZOW, *Minimum  $l_1$ ,  $l_2$  and  $l_\infty$  norm approximate solutions to an overdetermined system of linear equations*, Digital Signal Processing, 12(2002) 524–560.
- E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, NY, 1966.

- P. CIARLET, J. LIONS, *Handbook of Numerical Analysis - Vol. I - XII*, Elsevier, Amsterdam, 2000.
- L. COLLATZ, *Numerical Treatment of Differential Equations*, Springer-Verlag, Berlin, Heidelberg, New York, 1966.
- J. CUMIN, B. GRIZELJ, R. SCITOVSKI, *Numerical solving of ballistic flight equations for big bore air rifle*, *Technical Gazette*, **16**(2009) 41–46.
- R. CUPEC, *Scene reconstruction and free space representation for biped walking robots*, Ph.D. thesis, TU München, 2005.
- R. CUPEC, R. GRBIĆ, K. SABO, R. SCITOVSKI, *Three points method for searching the best least absolute deviations plane*, *Applied Mathematics and Computation*, **215**(2009) 983–994.
- G. DAHLQUIST, A. BJÖRCK, *Numerische Methoden*, R. Oldenbourg Verlag, München, Wien, 1972.
- G. DAHLQUIST, A. BJÖRCK, *Numerical Methods in Scientific Computing (Volume II)*, volume 1, SAIM, 2008.
- C. DE BOOR, *A Practical Guide to Splines*, Springer Verlag, New York, 2001.
- B. P. DEMIDOVICH, I. MARON, *Computational Mathematics*, MIR, Moscow, 1981.
- W. J. DEMMEL, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
- J. E. DENNIS, R. B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.
- P. DEUFLHARD, *Numerical Analysis: A First Course in Scientific Computation*, Walter de Gruyter, Berlin, 1995.
- P. DEUFLHARD, A. HOHMANN, *Numerical Analysis in Modern Scientific Computing*, Springer - Verlag, New York, 2003.
- P. DEUFLHARD, A. HOHMANN, *Numerische Mathematik 1, Eine algorithmisch orientierte Einführung*, W. de Gruyter, Berlin, 2008.

- Y. DODGE, editor, *Statistical data analysis based on the L1-norm and related methods*, Proceedings of The Third International Conference on Statistical Data Analysis Based on the L1-norm and Related Methods. Elsevier, Neuchâtel, 1997, 1997.
- Z. DREZNER, H. W. HAMACHER, *Facility Location: Applications and Theory*, Springer, 2004.
- I. DUFF, G. WATSON, *The State of the Art in Numerical Analysis*, IMA, Oxford, 1997.
- M. A. FISCHLER, R. C. BOLLES, *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*, Graphics Image Process, **24**(1981) 381–395.
- P. E. GILL, W. MURRAY, M. H. WRIGHT, *Practical Optimization*, Academic Press, 1981.
- P. E. GILL, W. MURRAY, M. H. WRIGHT, *Numerical Linear Algebra and Optimization, Vol.1.*, Addison-Wesley Publishing Company, New York, 1991.
- G. H. GOLUB, C. F. VAN LOAN, *Matrix Computations*, The J.Hopkins University Press, Baltimore and London, 1996.
- C. GURWITZ, *Weighted median algorithms for  $l_1$  approximation*, BIT, **30**(1990) 301–310.
- A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Dover Publications, New York, 1964.
- C. IYIGUN, A. BEN-ISRAEL, *A generalized weiszfeld method for the multi-facility location problem*, Operations Research Letters, **38**(2010) 207–214.
- D. JUKIĆ, R. SCITOVSKI, *Existence of optimal solution for exponential model by least squares*, J. Comput. Appl. Math., **78**(1997) 317–328.
- D. JUKIĆ, R. SCITOVSKI, *Matematika I*, Odjel za matematiku, Sveučilište u Osijeku, 2004.
- D. KINCAID, W. CHENEY, *Numerical Analysis*, Brooks/Cole Publishing Company, 1996.
- R. KRESS, *Numerical Analysis*, Springer-Verlag, Berlin, 1998.

- S. KUREPA, *Konačno dimenzionalni vektorski prostori i primjene*, Tehnička knjiga, Zagreb, 1967.
- S. KUREPA, *Matematička analiza II*, Tehnička knjiga, Zagreb, 1990.
- P. LANCASTER, K. ŠALKASKAS, *Curve and Surface Fitting*, Academic Press, London, 1986.
- K. LANGE, *Numerical Analysis for Statisticians*, Springer-Verlag, New York, 1999.
- C. L. LAWSON, R. J. HANSON, *Solving Least Squares Problems*, SIAM, Philadelphia, 1996.
- J. H. MATHEWS, K. D. FINK, *Numerical Methods Using MATLAB*, Prentice Hall, 1999.
- D. MATIJEVIĆ, N. TRUHAR, *Uvod u računarstvo*, Odjel za matematiku, Sveučilište u Osijeku, 2012.
- H. MOLITZ, R. STRÖBEL, *Äussere Ballistik.*, Springer, 1963.
- J. ORTEGA, *Numerical Analysis: a Second Course*, SIAM, Philadelphia, 1990.
- J. M. ORTEGA, W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- R. PLATO, *Concise Numerical Mathematics*, American Mathematical Society, Providence, 2003.
- W. H. PRESS, B. P. FLANNERY, S. A. TEUKOLSKY, W. T. VETTERLING, *Numerical Recipes*, Cambridge University Press, Cambridge, 1992.
- P. J. ROUSSEEUW, A. M. LEROY, *Robust Regression and Outlier Detection*, Wiley, New York, 2003.
- K. SABO, R. SCITOVSKI, *The best least absolute deviations line – properties and two efficient methods*, ANZIAM Journal, **50**(2008) 185–198.
- K. SABO, R. SCITOVSKI, I. VAZLER, *Searching for a best LAD-solution of an overdetermined system of linear equations motivated by searching for a best LAD-hyperplane on the basis of given data*, Journal of Optimization Theory and Applications, **149**(2011) 293–314.

- A. A. SAMARSKIJ, A. V. GULIN, *Numeričke metode (na ruskom)*, Nauka, Moskva, 1989.
- M. SCHATZMAN, *Numerical Analysis - a mathematical introduction*, Oxford University Press, Oxford, xxxx.
- A. SCHÖBEL, *Locating Lines and Hyperplanes: Theory and Algorithms*, Springer Verlag, Berlin, 1999.
- H. R. SCHWARZ, N. KÖCKLER, *Numerische Mathematik*, Springer Verlag, Berlin, 2011.
- R. SCITOVSKI, *Problemi najmanjih kvadrata. Financijska matematika*, Ekonomski fakultet, Elektrotehnički fakultet, Sveučilište u Osijeku, 1993.
- R. SCITOVSKI, *Numerička matematika*, Odjel za matematiku, Sveučilište u Osijeku, 2004.
- R. SCITOVSKI, N. TRUHAR, Z. TOMLJANOVIĆ, *Metode optimizacije*, Odjel za matematiku, Sveučilište u Osijeku, 2014.
- E. SÜLI, D. MAYERS, *An Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, 2006.
- G. W. STEWART, *Afternotes on Numerical Analysis*, SIAM, Philadelphia, 1996.
- G. W. STEWART, *Afternotes goes to Graduate School*, SIAM, Philadelphia, 1998.
- J. STOER, *Numerische Mathematik I, 8. Ed.*, Springer-Verlag, Berlin, 2002.
- J. STOER, R. BULIRSCH, *Introduction to Numerical Analysis, 2nd Ed.*, Springer Verlag, New York, 2002.
- L. N. TREFETHEN, D. BAU, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
- N. TRUHAR, *Numerička linearna algebra*, Odjel za matematiku, Sveučilište u Osijeku, 2010.
- C. UEERHUBER, *Numerical Computation*, Springer-Verlag, Berlin, 1997.
- G. A. WATSON, *Approximation Theory and Numerical Methods.*, John Wiley & Sons, Chichester, 1980.

C. J. ZAROWSKI, *An Introduction to Numerical Analysis for Electrical and Computer Engineers*, Wiley, New Jersey, 2004.

# Indeks

- Čebiševljevi polinomi, 32
- Čvorovi interpolacije, 21
- Čebiševljevi polinomi, 138
- Algoritam
  - Back Substitution, 56, 63, 66, 68
  - bisekcije, 89
  - Cholesky, 67
  - duljina koraka, 160
  - Forward Substitution, 56, 63, 66
  - Gauss-Newtonov, 162
  - Gauss-Seidelov, 75
  - Gram - Schmidtov, 137
  - Hornerov, 30
  - invertiranje matrice, 65
  - Jacobijev, 74
  - LU, 64
  - Newtonova metoda, 102
  - spline, 41
- Aproksimacija
  - funkcije, 126
  - matrice, 52, 157
  - najbolja, 126
  - najbolja  $L_2$  aproksimacija, 126
  - najbolja  $l_2$  aproksimacija, 157
  - najbolja  $L_\infty$  aproksimacija, 140
  - vektora, 52
- Aritmetička sredina, 146
- Balistika, 106
- Besselova nejednakost, 131
- Brzina konvergencije, 91
- Dekompozicija
  - Cholesky, 65, 84
  - LU-dekompozicija, 60
  - na singularne vrijednosti, 77, 155
  - QR-dekompozicija, 67, 154
- Floating-point aritmetika, 3, 49
- Gaussova metoda eliminacije, 57, 61
- Givensove rotacije, 68
- Gram-Schmidtov postupak ortogonalizacije, 135
- Hat-funkcija, 36
- Hesijan, 150
- Hornerova shema, 28, 30, 46
- Householderove transformacije, 68
- Interpolacija, 21
- Interpolacijski polinom, 22
  - Lagrangeov oblik, 25
  - Newtonov oblik, 28
  - ocjena pogreške, 30
- Jacobijan, 150
- Kontrakcija, 75
- Matrica
  - determinanta matrice, 63
  - dijagonalno dominantna, 40, 58, 77
  - elementarna, 61
  - glavni minori, 59



- Gramova, 127
- Hilbertova, 60, 67, 127
- Householderova, 69
- inverzna, 65
- permutacija, 62
- pozitivno definitna, 58
- pseudoinverzna, 153
- rijetko popunjena, 57, 73
- trodijagonalna, 40, 57
- vrpčasta, 83
- Metoda
  - bisekcije, 88
  - Brentova, 124
  - Croutova, 63
  - Gauss–Jordanova, 58
  - Gauss–Newtonova, 161
  - Gauss–Seidelova, 73
  - Gaussova metoda eliminacije, 57, 61
  - Jacobijeva, 73
  - jednostavnih iteracija, 92
  - Marquardtova, 165
  - Newtonova, 99, 159
  - regula falsi, 105
  - sekanti, 104
  - SOR, 77
- Minimizacija
  - jednodimenzionalna, 124, 159
  - višedimenzionalna, 160
- Nelinearna jednadžba, 85
- Norma, 50
  - $l_1$  norma, 50
  - Čebiševljeva  $L_\infty$  norma, 125
  - Čebiševljeva  $l_\infty$  norma, 50
  - Euklidova  $l_2$  norma, 50
  - Frobeniusova, 52
  - matrična, 51
  - matrična inducirana, 52
  - vektorska, 50
- Nultočka funkcije, 86
- Ocjena pogreške, 90, 96, 100
- Ortogonalan sustav, 128
  - Fourierov sustav, 131
- Ortogonalni polinomi, 135
  - Čebiševljevi polinomi, 138
  - Legendreovi polinomimi, 137
- Pivot-element, 57
- Podijeljene razlike, 29
- Pogreška
  - apsolutna, 5
  - inverzni problem, 14
  - metode, 2
  - modela, 2
  - relativna, 7
  - standardna, 14
  - strojna, 3
  - zaokruživanja, 1, 10
- Pouzdana znamenke, 9
- Problem najmanjih kvadrata, 145
  - linearni, 146, 151
  - nelinearni, 148, 159
- Separacija nultočki, 86
- Sherman–Morrisonova lema, 62
- Signifikantne znamenke, 9
- Singularne vrijednosti matrice, 55, 78
- Singularni vektori matrice, 78
- Spline, 33
  - kubični interpolacijski, 38
  - linearni interpolacijski, 34
  - ocjena pogreške, 35
- Strategija
  - parcijalnog pivotiranja, 59
  - potpunog pivotiranja, 59
- Sustav linearnih jednadžbi, 49
  - iterativne metode, 73
  - perturbirani, 53
  - rješavanje unaprijed, 56

- rješavanje unazad, 55
- trokutasti sustav, 55
- Sustav nelinearnih jednadžbi, 112
  - BFGS metoda, 118
  - Broydenova metoda, 118
  - DFP metoda, 118
  - kvazi-Newtonova metoda, 118
  - metoda jednostavnih iteracija, 114
  - Newtonova metoda, 115
- Sustav normalnih jednadžbi, 127, 152
- Svojstvene vrijednosti matrice, 78
- Svojstveni vektori matrice, 78
- Sylvesterov kriterij, 59
  
- Težine podataka, 150
- Teorem
  - Banachov o fiksnoj točki, 94
  - Banachov o fiksnoj točki, 76
  - Chebyshev's Equioscillatio, 143
  - Pitagorin, 131
- Trapezno pravilo, 37
  
- Uvjetovanost, 52
  - broj uvjetovanosti, 53, 79
  
- Vandermondova determinanta, 23