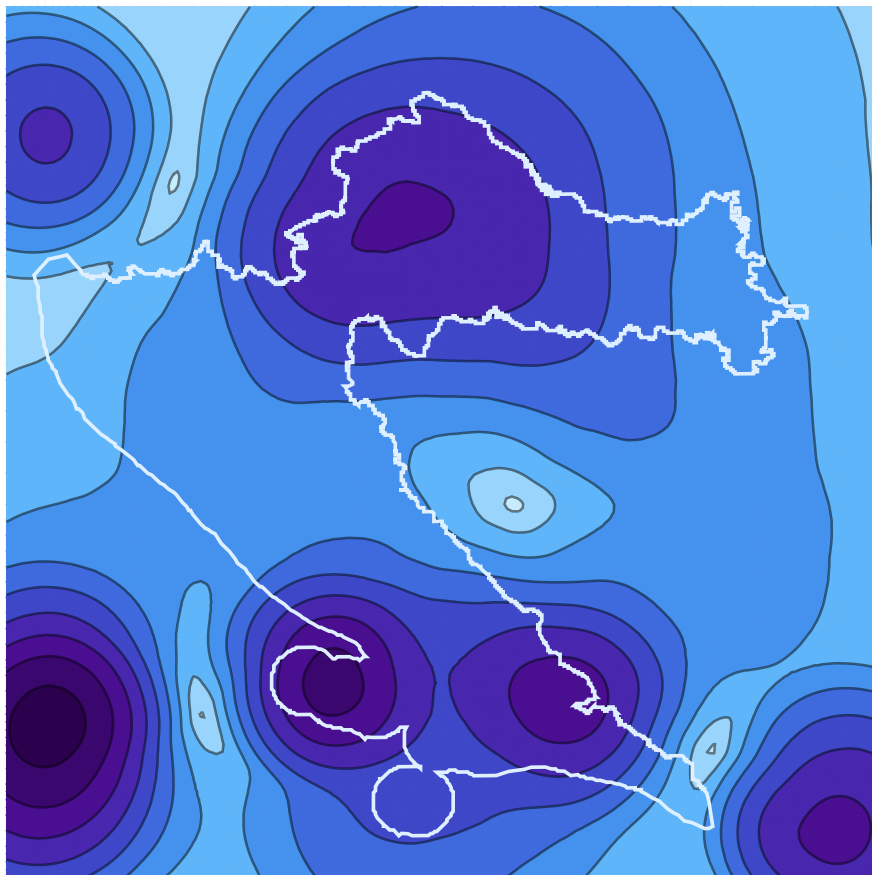


Sveučilište Josipa Jurja Strossmayera u Osijeku
Odjel za matematiku

Rudolf Scitovski

NUMERIČKA MATEMATIKA

Izmijenjeno i dopunjeno izdanje



Potresne aktivnosti u širem području Republike Hrvatske

Osijek, 2015.

Prof. dr. sc. Rudolf Scitovski, Numerička matematika

Odjel za matematiku

Sveučilište Josipa Jurja Strossmayera u Osijeku

Trg Lj. Gaja 6

HR-31 000 Osijek

e-mail: scitowsk@mathos.hr

Izdavač:

Sveučilišta Josipa Jurja Strossmayera u Osijeku – Odjel za matematiku

Recenzenti:

prof. dr. sc. Dragan Jukić, Odjel za matematiku Sveučilišta Josipa Jurja Strossmayera u Osijeku

prof. dr. sc. Miljenko Marušić, Matematički odsjek PMF-a Sveučilišta u Zagrebu

izv. prof. dr. sc. Kristian Sabo, Odjel za matematiku Sveučilišta Josipa Jurja Strossmayera u Osijeku

Lektorica:

Ivanka Ferčec, Elektrotehnički fakultet, Sveučilište Josipa Jurja Strossmayera u Osijeku

CIP zapis dostupan je u računalnom katalogu Gradske
i sveučilišne knjižnice Osijek pod brojem 140122033.

ISBN 978-953-6931-79-8

Ovaj udžbenik objavljuje se uz suglasnost Senata Sveučilišta Josipa Jurja Strossmayera u Osijeku pod brojem 11/15.

Ovaj udžbenik objavljuje se uz financijsku pomoć Ministarstva znanosti, obrazovanja i sporta Republike Hrvatske.

© Rudolf Scitovski, 1994., 2000., 2004., 2015.

Tisak: STUDIO HS Internet d.o.o. Osijek

PREDGOVOR

Ovaj tekst pisan je tako da podrazumijeva poznavanje osnova matematičke analize, linearne algebre i programiranja, a namijenjen je studentima pred-diplomskih i diplomskih studijskih programa STEM područja.

Tekst sadrži puno primjera koji mogu doprinijeti razumijevanju izložene materije. Na kraju svakog poglavlja nalaze se brojni zadaci: od sasvim jednostavnih, do onih koji mogu poslužiti kao teme seminarskih i sličnih radova. Gdje god je to moguće dana su rješenja zadataka, a kod nekih i odgovarajuće upute za rješavanje. Izrada većine zadataka povezana je s mogućnošću korištenja nekih programskih sustava (*Mathematica*, *Matlab* i sl.), kao i programa navedenih u ovom udžbeniku.

Sadržaj naveden u ovom udžbeniku, kao i priloženi programi, mogu poslužiti i u nekim praktičnim istraživanjima. U posebnim primjedbama navedene su i neke druge mogućnosti, koje se mogu pojaviti u praktičnom radu. Također, dane su i odgovarajuće upute za daljnje samostalno usavršavanje.

Novo izmijenjeno i dopunjeno izdanje ovog udžbenika korigirano je i dopunjeno u svim poglavljima. Značajnije dopune uvedene su u četvrtom poglavlju *Rješavanje nelinearnih jednadžbi*, gdje su dodana potpoglavlja o Lipschitz-neprekidnoj funkciji te dokazi konvergencije za metodu jednostavnih iteracija (odnosno Newtonovu metodu) za Lipschitz-neprekidnu (odnosno Lipschitz-derivabilnu) funkciju. Također u petom poglavlju *Aproksimacija funkcija* detaljnije se govori o ortogonalnim polinomima, a u šestom *Problemi najmanjih kvadrata* znatno više pažnje posvećeno je nelinearnoj minimizaciji.

Za sve važnije metode koje se navode u udžbeniku napisani su odgovarajući algoritmi, a na osnovi njih odgovarajući *Mathematica*-moduli koji su navedeni u devetom poglavlju (programski kodovi dostupni na <http://www.mathos.unios.hr/index.php/nastava/integrirani-nastavnicki-studij/218>). U brojnim primjerima i zadacima pokazano je korištenje tih modula. Na taj način, pored mogućnosti usvajanja osnovnih znanja iz numeričke matematike, studenti imaju mogućnost proširiti znanje programiranja u programskom sustavu *Mathematica* i također, bolje razumjeti različite numeričke metode

i pristupe. Pri tome, korištena je posljednja verzija programskog sustava *Mathematica*, za koji Sveučilište u Osijeku ima licencu.

Svi teoremi, leme, definicije, slike, tablice, primjedbe, primjeri i zadaci u tekstu imaju svoju jedinstvenu oznaku i na taj način pozivaju se u cijelom tekstu. Zbog toga su pisani velikim početnim slovom.

Zahvaljujem recenzentima prof. dr. sc. Draganu Jukiću, prof. dr. sc. Miljenku Marušiću i izv. prof. dr. sc. Kristianu Sabi i lektorici Ivanki Ferćec, koji su svojim primjedbama i prijedlozima značajno pomogli da ovaj tekst bude bolji. Posebno zahvaljujem kolegama prof. dr. sc. Šimi Ungaru, prof. dr. sc. Ninoslavu Truharu, doc. dr. sc. Zoranu Tomljanoviću, dr. sc. Ivanu Soldi, dr. sc. Danijelu Grahovcu, Ivanu Papiću, Branimiru Šajinoviću i studentici Petri Corn s Odjela za matematiku Sveučilišta u Osijeku, kao i kolegama s Matematičkog odsjeka PMF-a Sveučilišta u Zagrebu koji su na različite načine pomogli podizanju kvalitete ovog udžbenika.

Osijek, travnja, 2015.

Rudolf Scitovski

Sadržaj

1	Pogreške	1
1.1	Vrste pogrešaka	1
1.1.1	Pogreške zaokruživanja	1
1.1.2	Pogreške mjerenja	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	8
1.4	Pogreške kod izračunavanja vrijednosti funkcije	10
1.5	Inverzni problem u teoriji pogrešaka	13
1.6	Zadaci	15
2	Interpolacija. Spline interpolacija	19
2.1	Interpolacija	19
2.1.1	Lagrangeov oblik interpolacijskog polinoma	23
2.1.2	Newtonov oblik interpolacijskog polinoma	26
2.1.3	Ocjena pogreške	29
2.1.4	Nultočke Čebiševljevih polinoma kao čvorovi interpolacije	30
2.2	Spline interpolacija	32
2.2.1	Linearni interpolacijski spline	33
2.2.2	Kubični interpolacijski spline	36
2.3	Zadaci	41
3	Rješavanje sustava linearnih jednadžbi	47
3.1	Norma vektora i matrice	48
3.2	Uvjetovanost sustava linearnih jednadžbi	51

3.3	Rješavanje trokutastih sustava	53
3.4	Gaussova metoda eliminacije	55
3.5	LU-dekompozicija	59
3.6	Cholesky-dekompozicija	64
3.7	QR-dekompozicija	66
3.8	Iterativne metode	71
3.9	Dekompozicija na singularne vrijednosti	76
3.10	Zadaci	79
4	Rješavanje nelinearnih jednadžbi	85
4.1	Metoda bisekcije	87
4.2	Ocjena pogreške i brzina konvergencije	90
4.3	Metoda jednostavnih iteracija	92
4.3.1	Lipschitz-neprekidna funkcija	99
4.3.2	Metoda jednostavnih iteracija za Lipschitz-neprekidnu funkciju	102
4.4	Newtonova metoda	103
4.4.1	Modifikacije Newtonove metode	110
4.4.2	Newtonova metoda za Lipschitz-derivabilnu funkciju	113
4.5	Rješavanje sustava nelinearnih jednadžbi	115
4.5.1	Metoda jednostavnih iteracija	117
4.5.2	Newtonova metoda	119
4.5.3	Kvazi-Newtonove metode	122
4.6	Zadaci	124
5	Aproksimacija funkcija	129
5.1	Najbolja L_2 aproksimacija	130
5.1.1	Fourierovi polinomi	133
5.1.2	Ortogonalni polinomi	140
5.1.3	Čebiševljevi polinomi	145
5.2	Najbolja L_∞ aproksimacija	147
6	Problemi najmanjih kvadrata	153
6.1	Linearni problemi najmanjih kvadrata	161
6.1.1	Sustav normalnih jednadžbi za LPNK	162
6.1.2	Rješavanje LPNK pomoću QR-dekompozicije	164

6.1.3	Rješavanje LPNK pomoću dekompozicije na singularne vrijednosti	165
6.2	Nelinearni problemi najmanjih kvadrata	169
6.2.1	Gauss-Newtonova metoda	174
6.2.2	Marquardtova metoda	180
6.3	Druge mogućnosti	181
6.3.1	Princip najmanje sume apsolutnih odstupanja	181
6.3.2	Princip najmanjeg maksimalnog apsolutnog odstupanja	183
6.4	Zadaci	184
7	Numerička integracija	191
7.1	Trapezno pravilo	192
7.2	Newton-Cotesova formula	196
7.3	Simpsonovo pravilo	197
7.4	Zadaci	202
8	Numeričko rješavanje običnih diferencijalnih jednadžbi	207
8.1	Eulerova metoda	209
8.2	Metode Runge-Kutta	215
8.3	Metoda diskretizacije	220
8.4	Zadaci	222
9	Mathematica-moduli	227
9.2	Interpolacija. Spline interpolacija	227
9.2.1	Hornerov algoritam	227
9.3	Rješavanje sustava linearnih jednadžbi	227
9.3.1	Rješavanje trokutastih sustava	227
9.3.2	LU-dekompozicija	228
9.3.3	Cholesky-dekompozicija	228
9.3.4	Iterativne metode za rješavanje sustava linearnih jednadžbi	229
9.4	Rješavanje nelinearnih jednadžbi	230
9.4.1	Metoda bisekcije	230
9.4.2	Metoda jednostavnih iteracija	230
9.4.3	Newtonova metoda	231
9.5	Aproksimacija funkcija	232

9.5.1	Gram-Schmidtov postupak ortogonalizacije polinoma	232
9.7	Numerička integracija	232
9.7.1	Trapezno pravilo	232
9.7.2	Simpsonovo pravilo	232
9.8	Numeričko rješavanje običnih diferencijalnih jednažbi	233
9.8.1	Eulerova metoda	233
9.8.2	Simetrična formula	233
9.8.3	Runge-Kutta metoda	234
	Literatura	235
	Indeks	241

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 π 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 $P = 10\,000$ cm². Treba odrediti stranicu kvadrata koji će imati dvostruko manju površinu.*

Ako s x označimo stranicu traženog kvadrata, onda je $a^2 = 2x^2$, odakle je $x = a\frac{\sqrt{2}}{2}$. Broj $\sqrt{2}$ je iracionalan broj (beskonačni neperiodični 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 $P = 4970.25$ cm². Ako uzmemo $\sqrt{2} \approx 1.41421$, dobit ćemo stranicu kvadrata $x = 70.7105$ i površinu $P = 4999.97481$ cm². 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 potječe još iz antičke Grčke (vidi primjerice Brückler (2007)).

1.1.2 Pogreške mjerenja

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⁻², 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 matematički model kao njegovu aproksimaciju.

Primjer 1.2. *Prilikom proučavanja balističkih putanja u eksternoj balistici (vidi primjerice Cumin et al. (2009); Molitz i Ströbel (1963)) najvažniji problem predstavlja dobro definiranje zakona otpora zraka za balistički projektil. Kako bi se pojednostavilo postavljanje problema, 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 mogu se jako razlikovati od stvarnih. Usporedimo na primjer maksimalni domet običnog puščanog metka (10 g) i teškog topovskog projektila (150 kg) iste početne brzine $v_0 = 900$ ms⁻¹ u vakuumu i u zraku ispunjenom prostorom.*

<i>maksimalni domet projektila</i>	<i>vakuum</i>	<i>zrak</i>
<i>puščani metak</i>	82 km	3.5 km
<i>teški projektil</i>	82 km	40 km

Očigledne razlike nastale su u najvećoj mjeri zbog toga što je otpor zraka prema balističkom projektilu 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 $b \in \mathbb{N}$, $b \geq 2$ 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 i Hohmann, 2008; Gill et al., 1991; Schwarz i Köckler, 2011; Truhar, 2010).

U kompjutorskom zapisu koristi se tzv. zapis s pomičnim zarezom (*engl.: 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 preciznost (*engl.: single precision*) računala. Ako realni broj a ima više od t znamenki, računalo će ga postupkom zaokruživanja (*engl.: rounding*) ili odbacivanja (*engl.: 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 granice¹

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

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

¹Računala 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 $1 \times 10^{-38} \leq a \leq 2 \times 10^{38}$.

Primjedba 1.1. Svakom računalu pridružuje se najmanji broj $\epsilon > 0$ za koji je $1 + \epsilon > 1$. Broj ϵ nazivamo *točnost računala*, a u programskom sustavu *Mathematica* približno ga možemo izračunati na sljedeći način:

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

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

Neke je brojeve 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 δ pogreška aproksimacije.

Primjer 1.3. *Potražimo po apsolutnoj vrijednosti manji 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 $x_0 = 0$.

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 .

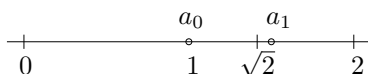
Definicija 1.1. Razliku $(a - a^*)$ između stvarne veličine a i njene aproksimacije a^* naziva se **pogreška aproksimacije**. Apsolutnu vrijednost pogreške aproksimacije naziva se **apsolutna pogreška aproksimacije** i označava s

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

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 i Scitovski, 2004, str. 82)). U Tablici 1.1 prikazana su redom prva četiri člana niza (tj. prve četiri aproksimacije broja $\sqrt{2}$), kao i odgovarajuće pogreške aproksimacija.



Slika 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).

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

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

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 $\sqrt{2}$

Spomenuti broj $\epsilon > 0$ nazivamo **granica pogreške aproksimacije**. Nejednakost (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 određena približnim brojem a^* iz intervala $[a^* - \epsilon, a^* + \epsilon]$ i da pri tome apsolutna pogreška Δa^* nije veća od ϵ .

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

<i>mjerenje</i>	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 od točnosti izmjerene veličine r . Dakle, veličinu pogreške treba promatrati u odnosu na promatranu veličinu.

Definicija 1.2. Omjer između apsolutne pogreške Δa^* i apsolutne vrijednosti veličine a ($a \neq 0$) nazivamo **relativna pogreška** δa^* 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 (%) ili u promilima (‰).

Primjedba 1.2. Kako u praksi prava vrijednost a obično nije poznata i kako je $a^* \approx a$, 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

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

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

$$\delta_{r^*} \approx \frac{\Delta_{r^*}}{|r^*|} = \frac{0.05}{2} = 0.025 = 25\text{‰}, \quad \delta_{h^*} \approx \frac{\Delta_{h^*}}{|h^*|} = \frac{0.05}{100} = 0.0005 = 0.5\text{‰}.$$

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

$$I = \int_a^b f(x) dx \text{ vrijedi}$$

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

Kao jednu od aproksimacija integrala I možemo promatrati aritmetičku sredinu

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

Pokažite da tada vrijedi

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

Odredite egzaktnu vrijednost integrala $\int_0^2 \frac{5-x}{9-x^2} dx$, te usporedite ocjene za apsolutnu i relativnu pogrešku s pravim vrijednostima.

Rješenje: $I^* = 1.1$, $\Delta I^* \leq 0.1$, $\delta I^* \leq 9\%$.

1.3 Signifikantne znamenke

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

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

gdje su $b_i \in \{0, 1, \dots, 9\}$, $b_m \neq 0$, 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 podudara 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 Δa^* 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. Aproksimativni 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 (vidi primjerice Dahlquist i Björck (2008); Demidovich i Maron (1981)):

- 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 od 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 Lagrangeovom 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.

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 π 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, pa pišemo $V^* = 38$.

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

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

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

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

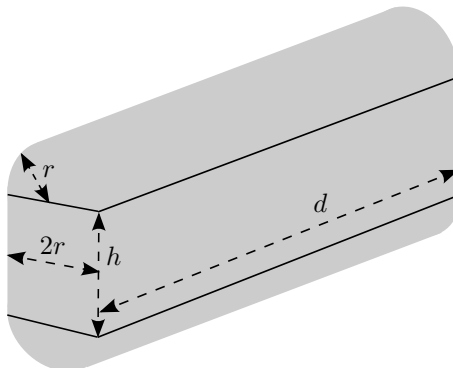
Primjer 1.12. *Kolike smiju biti apsolutne pogreške radijusa r i visine h valjka volumena $V = 12\pi \pm 0.1 \text{ m}^3$ (dozvoljena tolerancija je 100 litara!)? Neka je pri tome $r^* = 2 \text{ m}$. Broj π uzet ćemo dovoljno točno da njegova pogreška ne utječe na rezultat.*

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 10001, širine baze $2r^* = 80$ cm i visine baze $2r^* + h^* = 100$ cm. Kolika mora biti dužina d spremnika? Kolike smiju biti pogreške Δr^* , Δh^* , Δd^* , 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, $r = 40 \pm 0.08$ cm, $h = 20 \pm 0.3$ cm, $d = 151 \pm 0.5$ cm.

Primjedba 1.7. Više detalja o teoriji pogrešaka i primjenama može se vidjeti primjerice kod Demidovich i Maron (1981). Posebno o problemu kompjutorske preciznosti u numeričkim metodama može se vidjeti kod Ackleh et al. (2010); Björck i Dahlquist (2007); Deuffhard i Hohmann (2008, 2003); Gill et al. (1981, 1991); Matijević i Truhar (2012); Plato (2003); Press et al. (1992); Schatzman (2002); Schwarz i Köckler (2011); Stewart (1996); Stoer (2002); Stoer i 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 vrijednošću 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: *Modificirajte Mathematica-modul iz Primjera 1.9.*

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

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

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

p	2%	3.5%	7.5%	12%	24%	50%	120%	360%	1000%
Δp_r	0.0015	0.0046	0.02	0.05	0.19	0.73	3.21	16.44	61.21
δp_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^*)

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

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

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

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

Zadatak 1.12. Prema poučku o kosinusima, ako su za neki trokut poznate dvije stranice a, b i kut γ 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

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

Rješenje: kut γ najprije treba pretvoriti u radijane. Dobivamo:

$$\begin{aligned} \text{a) } & \Delta c^* \approx 0.0524 \text{ cm}, \quad \delta c^* \approx 0.01 \approx 1\%, \quad c^* \approx 5.4367 \text{ cm}. \\ \text{b) } & \Delta c^* \approx 0.16 \text{ m}, \quad \delta c^* \approx 0.0012 \approx 0.1\text{‰}, \quad c^* \approx 133.733 \text{ m}. \end{aligned}$$

Zadatak 1.13. Procijenite apsolutnu i relativnu pogrešku prilikom:

- zbrajanja n približnih brojeva,
- množenja n približnih brojeva,
- kvadriranja jednog približnog broja,
- potenciranja s n -tom potencijom jednog približnog broja,
- izračunavanja drugog korijena nekog približnog broja,
- izračunavanja n -tog korijena nekog približnog broja,
- 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: $\Delta a^* \approx 0.0068 \text{ cm}, \quad \delta a^* \approx 0.0033 \approx 0.3\text{‰}, \quad a^* \approx 2.03965 \text{ dm}.$

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 α 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 4.16, str. 111). Procijenite apsolutnu i relativnu pogrešku pri izračunavanju daljine gađanja ako je

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

- 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 4.16, str. 111): $X = \frac{v_0^2}{g}$. Procijenite apsolutnu i relativnu pogrešku pri izračunavanju maksimalnog dometa projektila, ako je

- a) $v_0 = 115 \pm 1 \text{ ms}^{-1}$, $g = 9.81 \pm 0.001 \text{ ms}^{-2}$.
 b) $v_0 = 900 \pm 10 \text{ ms}^{-1}$, $g = 9.81 \pm 0.001 \text{ ms}^{-2}$.

- Rješenje: a) $\Delta X^* \approx 23.58 \text{ m}$, $\delta X^* \approx 0.017 \approx 1.7\%$.
 b) $\Delta X^* \approx 1843.28 \text{ m}$, $\delta X^* \approx 0.022 \approx 2.2\%$.

Zadatak 1.19. Vrijeme leta balističkog projektila ispaljenog pod kutem α početnom brzinom v_0 na ravnom terenu u vakuumu računa se po formuli (vidi Primjer 4.16, str. 111): $T = \frac{2v_0}{g} \sin \alpha$. Procijenite apsolutnu i relativnu pogrešku pri izračunavanju vremena leta T projektila, ako je

- a) $v_0 = 115 \pm 1 \text{ ms}^{-1}$, $g = 9.81 \pm 0.001 \text{ ms}^{-2}$, $\alpha = 15^\circ \pm 0.1^\circ$
 b) $v_0 = 900 \pm 10 \text{ ms}^{-1}$, $g = 9.81 \pm 0.001 \text{ ms}^{-2}$, $\alpha = 15^\circ \pm 0.1^\circ$

- Rješenje: a) $\Delta T^* \approx 0.093 \text{ s}$, $\delta T^* \approx 0.015 \approx 1.6\%$, $T^* \approx 6.068 \text{ s}$.
 b) $\Delta T^* \approx 0.84 \text{ s}$, $\delta T^* \approx 0.018 \approx 2\%$, $T^* \approx 47.49 \text{ s}$.

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 α (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 točnima.

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 točkama: $x_0 < x_1 < \dots < x_n$. 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. Kad 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 $n + 1$ točkama (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_∞ -normi

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

ili u L_2 -normi

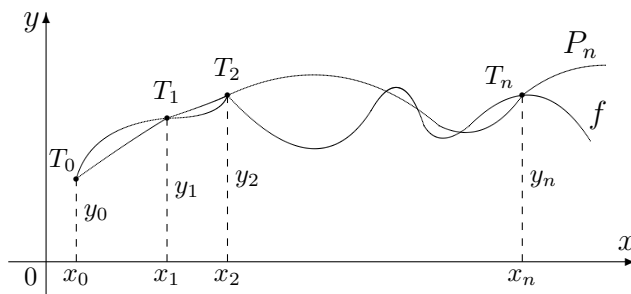
$$\|f - g\|_2 = \sqrt{\int_a^b (f(x) - g(x))^2 dx}.$$

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)

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



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 $n + 1$ jednadžbi s $(n + 1)$ -nom nepoznanicom

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

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., str. 153 (vidi također Truhar (2010)). Ako bismo 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. 51). Zbog toga ćemo se u nastavku upoznati s drugim metodama za rješavanje ovog problema.

Problem interpolacije ilustrirat ćemo primjenom programskog sustava *Mathematica*. Najprije za funkciju f definiramo n podataka sastavljenih od ekvidistantnih čvorova i vrijednosti funkcije u njima.

```
In[1]:= f[x_] := 1 - Abs[x - 1]
      a = 0; b = 2; n = 10;
      x = Table[i (b - a)/n, {i, 0, n}];
      y = f[x];
      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 *Mathematica*-modula *LinearSolve*.

```
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 Slici 2.2 korištenjem sljedećeg programa.

```

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]

```

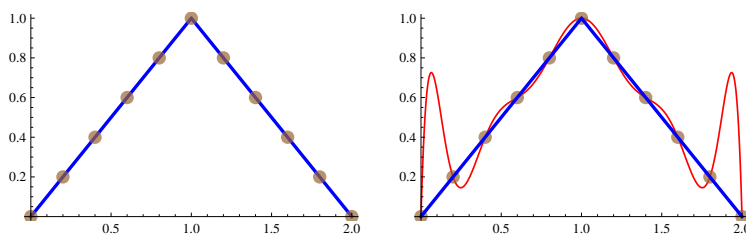
Uvjetovanost matrice sustava (2.5) u l_∞ normi dobivamo na sljedeći način.

```

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: $\text{cond} = 8.61 \times 10^8$.*



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

Vrijednost interpolacijskog polinoma P_n u proizvoljnoj točki z možemo izračunati Hornerovom shemom (vidi Jukić i Scitovski (2004)) korištenjem jednostavnog algoritma:

Algoritam 1 (Hornerov algoritam)**Input:** $a_n, \dots, a_1, a_0; z;$ {Učitati koeficijente polinoma $P_n(x) = a_n x^n + \dots + a_1 x + a_0$ i točku z u kojoj se traži vrijednost polinoma}1: $p = a_n$ 2: **for** $i = 1$ to n **do**3: $p = p z + a_{n-i}$ 4: **end for****Output:** $\{z, p\}$

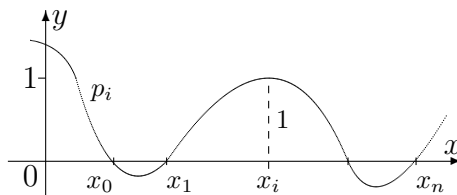
Odgovarajući Mathematica-modul `Horner[a_, z_]` može se vidjeti u poglavlju 9.2.1, str. 227.

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

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

Zapis interpolacijskog polinoma (2.9) obično nazivamo **Lagrangeovim oblikom 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 $f: \mathbb{R}_+ \rightarrow \mathbb{R}^1$, $f(x) = \ln x - 1 + \frac{1}{x}$ 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

- 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) $P_2(x) = -0.013256x^2 + 0.30111x + 0.35605$.

	$f(5.25)$	$y^* = P(5.25)$	Δy^*
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 $f: \mathbb{R}_+ \rightarrow \mathbb{R}$, $f(x) = \ln x - 1 + \frac{1}{x}$

Rezultate provjerite priloženim *Mathematica*-programom, str. 21.

Primjer 2.4. Za različite vrijednosti izlaznog kuta α i konstantnu početnu brzinu $v_0 = 115 \text{ ms}^{-1}$ primjenom formule (B.4), str. 112, izračunate su daljine gađanja X balističkog projektila u vakuumu

α°	5°	15°	25°	35°	45°
$X \text{ (m)}$	234	674	1033	1267	1348

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

¹ U cijelom tekstu koristimo sljedeće oznake:

- $\mathbb{R}_+ = \{x \in \mathbb{R}: x \geq 0\}$ – nenegativni realni brojevi;
- $\mathbb{R}_{++} = \{x \in \mathbb{R}: x > 0\}$ – pozitivni realni brojevi.

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

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

Pomoću ove funkcije za zadane daljine gađanja 500 m, 750 m i 1000 m izračunati su odgovarajući izlazni kutevi α^* , te apsolutne i relativne pogreške u odnosu na prave vrijednosti α dobivene iz formule (B.4), str. 112.

X (m)	α (°)	α^* (°)	$\Delta\alpha^*$
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. 153 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 i Scitovski, 2004, str. 47)).

Prijedimo 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 funkcija² (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 α treba odrediti tako da bude $P_2(x_2) = y_2$. Ako u (2.12) uvrstimo $x = x_2$, dobivamo

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

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

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

²Ispravno je ovu funkciju nazivati „afina funkcija”, ali zbog ustaljenog načina pisanja u tehnici i drugim područjima znanosti, u ovom udžbeniku i nadalje ćemo pisati „linearna funkcija”.

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 $f \in C^{n+1}[a, b]$ 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 i 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 ω normirani polinom $(n+1)$ -og stupnja, 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

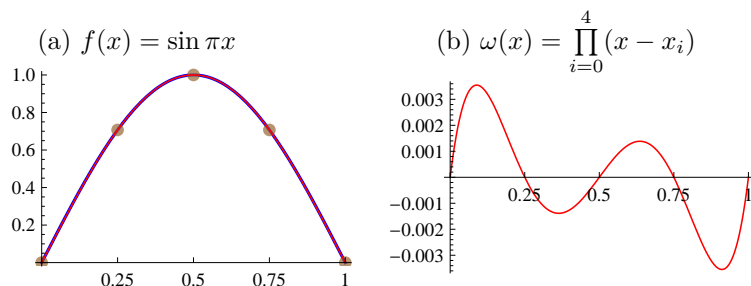
Ako označimo $M_{n+1} := \max_{x \in [a,b]} |f^{(n+1)}(x)|$, onda iz (2.15) slijedi

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

što znači da pogreška aproksimacije ovisi o ponašanju funkcije ω 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

$$P_4(x) = 2.82843x - 3.31371x(x - 0.25) - 1.83011x(x - 0.25)(x - 0.5) + 3.66022x(x - 0.25)(x - 0.5)(x - 0.75).$$

Na Slici 2.4 prikazani su odgovarajući grafovi funkcija f i ω . Iako je l_∞ pogreška 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

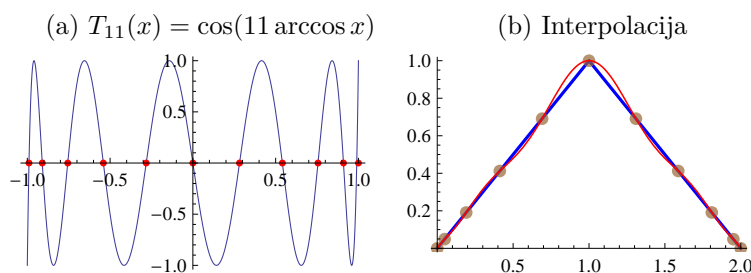
Pogreška (2.17) kod interpolacijskog polinoma ovisi o ponašanju funkcije ω . Za jednoliko raspoređene čvorove i nešto veći n pokazuje se da funkcija ω

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 ω 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. 148, 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 polinom³ proširen na interval $[a, b]$.

Na Slici 2.5a prikazan je graf Čebiševljevog polinoma T_{11} i njegove nultočke na intervalu $[-1, 1]$. Kako je $(n+1)$ -vi Čebiševljev polinom T_{n+1} zadan s



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

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

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

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

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

Za funkciju $f: [a, b] \rightarrow \mathbb{R}$ zadanu na intervalu $[a, b]$, najprije definiramo linearnu funkciju $u: [-1, 1] \rightarrow [a, b]$, $u(x) = \frac{a+b}{2} + \frac{b-a}{2}x$, a odgovarajuće nultočke u_k , $k = 0, 1, \dots, n$ Čebiševljevog polinoma T_{n+1} definiranog na $[a, b]$ dobivamo rješavajući jednadžbu $T_{n+1}(u^{-1})(x) = 0$, tj.

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

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

³Čebiševljev polinom stupnja n programom Mathematica dobiva se naredbom ChebyshevT[n,x]

Primijetite da se isti čvorovi u_k , $k = 0, 1, \dots, n$ mogu dobiti i tako da nultočke $x_k \in [-1, 1]$ preslikamo na interval $[a, b]$ pomoću linearne funkcije u . 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)

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

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

*Zadatak 2.2.*⁴ Odredite interpolacijski polinom stupnja $n = 6$ za funkciju $f: [-5, 5] \rightarrow \mathbb{R}$, $f(x) = \frac{1}{1+x^2}$ uz korištenje ekvidistantno raspoređene čvorove i uz korištenje čvorove koji su nultočke odgovarajućeg Čebiševljevog polinoma.

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); Björck i Dahlquist (2007); Ciarlet i Lions (2000); Dahlquist i Björck (1972); Demidovich i Maron (1981); Deuffhard i Hohmann (2008); Kincaid i Cheney (1996); Kress (1998); Lancaster i Šalkauskas (1986); Ortega (1990); Plato (2003); Schwarz i Köckler (2011); Stewart (1996); Stoer (2002); Stoer i Bulirsch (2002). Odgovarajuća FORTRAN – podrška može se pronaći kod Press et al. (1992).

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

⁴Klasični primjer Carl David Tolmé Runge (1856.–1927.) preuzet iz knjige Plato (2010).

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

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

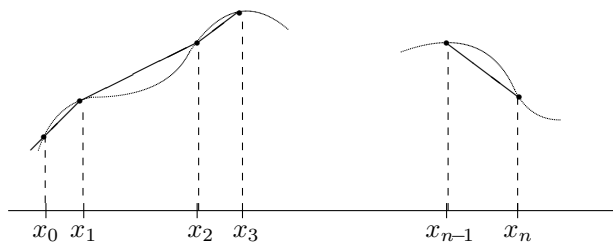
2.2.1 Linearni interpolacijski spline

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

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



Slika 2.6: Linearni interpolacijski spline

Kako je

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

funkcija φ neprekidna je 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**⁵ (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. 29, za svaki $x \in [x_{i-1}, x_i]$ postoji $\xi_i \in \langle x_{i-1}, x_i \rangle$, 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

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

pa je

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

Ako označimo $h_{max} := \max_{i=0, \dots, n} h_i$ i $M_2 := \max_{x \in [a, b]} |f''(x)|$, tada prema (2.20) vrijedi ocjena pogreške za linearni interpolacijski spline

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

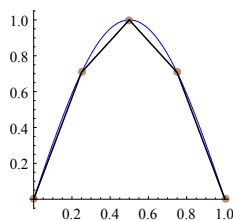
Primjer 2.7. Treba odrediti linearni interpolacijski spline za funkciju i čvorove interpolacije iz Primjera 2.6, str. 30. Dobivamo

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

Linearni interpolacijski spline možemo definirati i na drugi način - analogno kako smo u t. 2.1.1, str. 23 uveli Lagrangeov oblik interpolacijskog polinoma (vidi primjerice Lancaster i Šalkauskas (1986); Schwarz i Köckler (2011)). U tu svrhu definirajmo tzv. **hat-funkcije** p_i , $i = 1, \dots, n - 1$.

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

⁵Naziv dolazi od engleske riječi *spline*, koja označava dugu, savitljivu, drvenu letvicu.



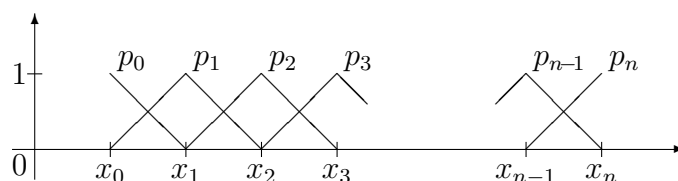
Slika 2.7: Linearni interpolacijski spline iz Primjera 2.7

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

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) na str. 23), a njihovi grafovi prikazani su na Slici 2.8.



Slika 2.8: Hat-funkcije

Sada linearni interpolacijski spline možemo definirati formulom (usporedi s (2.9), str. 24)

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

Primjedba 2.4. Ako s \mathcal{L} označimo skup svih linearnih interpolacijskih splineova 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 φ 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 $L(L(f)) = L(f)$, tj. $L^2 = L$. Dakle, linearni operator L je projektor.

Zadatak 2.3. 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 2.4. 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. 7.1, str. 192).

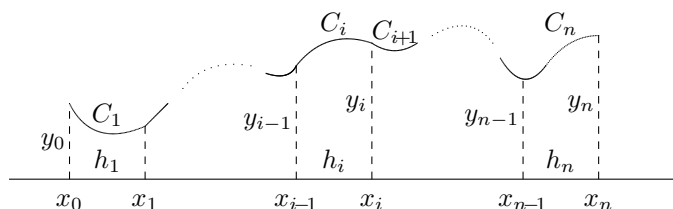
2.2.2 Kubični interpolacijski spline

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 2.9).

- (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) 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** (vidi primjerice Plato (2003); Stewart (1996); Trefethen i 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. 51), 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. Označimo

$$s_0 := C_1'''(x_0), \quad s_i := C_i'''(x_i), \quad i = 1, \dots, n.$$

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 matičnom 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)

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

Algoritam 2 (Prirodni kubični spline)

Input: (x_i, y_i) , $i = 0, 1, \dots, n$

{Učitati podatke (x_i, y_i) , $i = 0, 1, \dots, n$ }

1: 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);

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

Output: Polinomi (2.25).

Prilikom rješavanja sustava (2.31) treba iskoristiti specijalnu strukturu matrice sustava (vidi Zadatak 3.25, str. 82). 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 2.5. Izradite program koji će prema Algoritmu 2 odrediti 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 splinea 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

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

odnosno

$$\|g''\|_2 \geq \|C'''\|_2,$$

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

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

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

$$I_1 = - \sum_{i=1}^n k_i \int_{x_{i-1}}^{x_i} (C'_i(x) - g'(x)) dx = - \sum_{i=1}^n k_i [C_i(x) - g(x)]_{x_{i-1}}^{x_i} = 0,$$

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

Svojstvo (2.32) nazivamo minimalno svojstvo prirodnog kubičnog splinea. To svojstvo vrijedi i za kubični spline s drukčijim rubnim uvjetima.

Primjedba 2.8. Spline-interpolacije pojavile su se u matematičkoj literaturi šezdesetih godina prošlog stoljeća, 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 i Hohmann (2008); Lange (1999); Plato (2003); Schatzman (2002); Schwarz i Köckler (2011); Stewart (1998); Stoer (2002); Stoer i Bulirsch (2002). Programski sustav *Mathematica* omogućava izravno korištenje spline aproksimacije bilo kojeg reda razumne veličine. Odgovarajuća FORTRAN programska podrška može se naći u Press et al. (1992), kao i u *The NAG Library*. O interpolaciji funkcija jedne i više varijabli može se vidjeti primjerice kod Lancaster i Šalkauskas (1986).

2.3 Zadaci

Zadatak 2.6. Odredite determinantu sustava (2.5) (Vandermondova determinanta).

Zadatak 2.7. Odredite interpolacijski polinom za sljedeće tablično zadane funkcije

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

Rješenje:

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

Zadatak 2.8. Za tablično zadane funkcije odredite Newtonov oblik interpolacijskog polinoma te izračunajte njegovu vrijednost za $\xi = 10$.

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

Rješenje:

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

Zadatak 2.9. Zadane su vrijednosti funkcije $f(x) = 3x^2 + \pi^{-4} \ln(\pi - x)^2 + 1$ u sljedećim točkama

$$\begin{array}{c|cccc} x & 3.13 & 3.14 & 3.15 & 3.16 \\ \hline f(x) & 30.29918 & 30.44653 & 30.66939 & 30.87478 \end{array}$$

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

$$\begin{aligned} \text{Rješenje: } P_3(x) &= 483453.48 - 460723.76x + 146356.15x^2 - 15496.67x^3, \\ P_3(3.1416) &= 30.32619. \end{aligned}$$

Zadatak 2.10. Zadane su vrijednosti Γ -funkcije u sljedećim točkama

$$\begin{array}{c|cccc} x & 0.2 & 0.4 & 0.6 & 0.8 \\ \hline \Gamma(x) & 4.5910 & 2.2182 & 1.4892 & 1.1642 \end{array}$$

Odredite interpolacijski polinom, čiji će graf 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	$\Gamma(x)$	$P(x)$	ΔP
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 2.11. 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) $P(x) = -.01(6)x^3 - .1x^2 + 1.11(6)x$, $P(2) = 1.7$, $P(2.4) = 1.8736$
 (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 2.12. Za različite vrijednosti izlaznog kuta α i konstantnu početnu brzinu $v_0 = 300 \text{ ms}^{-1}$ primjenom formule (B.4) na str. 112 izračunati su dometi X balističkog projektila u vakuumu

α°	5°	10°	25°	45°
$X (m)$	1593	3138	7028	9174

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

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

X (m)	α	α^*	$\Delta\alpha^*$
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 2.13. 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:

$$a) t(h) = 0.2257 + 0.0874h + 0.2679h^2 - 0.131h^3$$

$$b) h(t) = -3.56697 + 26.7947t - 61.3122t^2 + 54.0473t^3$$

Zadatak 2.14. 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 2.15. 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 2.16. 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 2.17. 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 splinea.

Zadatak 2.18. Za zadane podatke (x_i, y_i) , $i = 0, 1, \dots, n$, odredite kvadratni interpolacijski spline i izradite odgovarajući program.

Zadatak 2.19. Zadana je periodična funkcija f perioda 2π 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 2.20. Pokažite da je dijagonalno dominantna matrica regularna.

Zadatak 2.21. 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:

$$\begin{aligned} \text{a) } C_1(x) &= 10 + \frac{43}{11}x - \frac{31}{88}x^3, \quad x \in [0, 2] \\ C_2(x) &= \frac{12}{11} + \frac{190}{11}x - \frac{147}{22}x^2 + \frac{67}{88}x^3, \quad x \in [2, 4] \\ C_3(x) &= \frac{1124}{11} - \frac{644}{11}x + \frac{135}{11}x^2 - \frac{9}{11}x^3, \quad x \in [4, 5], \\ C(1) &= 13.5568 \end{aligned}$$

$$\begin{aligned} \text{b) } C_1(x) &= 5.25 + 2.25x + 0.75x^2 - 0.25x^3, \quad x \in [1, 3] \\ C_2(x) &= -15 + 22.5x - 6x^2 + 0.5x^3, \quad x \in [3, 5] \\ C_3(x) &= 78.75 - 33.75x + 5.25x^2 - 0.25x^3, \quad x \in [5, 7] \\ C(4.5) &= 10.3125 \end{aligned}$$

Zadatak 2.22. Na osnovi podataka

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

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 2.23. Problem iz Primjera 2.4, str. 25, riješite primjenom kubičnog interpolacijskog splinea.

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

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, $\mathbf{x} \in \mathbb{R}^n$ vektor nepoznanica, a $\mathbf{b} \in \mathbb{R}^m$ 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 rješavanja.

Primjer 3.1. *Za početak pogledajmo jednostavan sustav jednadžbi*

$$\begin{aligned} 2x_1 + 400x_2 &= 200 \\ x_1 + x_2 &= 1, \end{aligned}$$

čije je rješenje

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

Simulirat ćemo rješavanje ovog sustava na računalu s 2-znamenkastom floating-point aritmetikom. Ako prvu jednadžbu pomnožimo s -0.5 i dodamo 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 $x_2^{**} = 0.497$, $x_1^{**} = 0.600$ (koristite `Mathematica`-program dostupan na: http://www.mathos.unios.hr/images/homepages/scitowsk/Num_slj.nb).

Iz ovog jednostavnog primjera vidi se da je važno znati procjenu veličine 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 $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$.

Najčešće korištene vektorske norme su

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

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

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

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 $p, q \in \{1, 2, \infty\}$ postoje konstante $\alpha, \beta > 0$ takve da vrijedi

$$\alpha \|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q \leq \beta \|\mathbf{x}\|_p, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Zadatak 3.2. Jedinična kuglina 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 sfere 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 $[a, b]$ su

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

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 **matrična 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 sve $\lambda \in \mathbb{R}$ i $\mathbf{A} \in \mathbb{R}^{n \times n}$,
- (iii) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ za sve $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$,
- (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 matricnu normu na sljedeći način

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

koju ćemo zvati **inducirana matricna norma**. Primijetite da je u (3.2) matricna norma 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

$$\|\mathbf{A}\|_1 = \max_{j=1, \dots, n} \|\mathbf{a}_j\|_1, \quad \text{gdje je } \mathbf{a}_j \text{ } j\text{-ti stupac matrice } \mathbf{A},$$

$$\|\mathbf{A}\|_2 = \sigma_1(\mathbf{A}) - \text{najveća singularna vrijednost matrice } \mathbf{A} \text{ (t. 3.9, str. 76),}$$

$$\|\mathbf{A}\|_\infty = \max_{i=1, \dots, n} \|\mathbf{a}_i^T\|_1, \quad \text{gdje je } \mathbf{a}_i \text{ } i\text{-ti redak matrice } \mathbf{A}.$$

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

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

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

Frobeniusova norma nije inducirana matricna norma, ali je kompatibilna s euklidskom 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 matricne 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 Ackleh et al. (2010); Golub i van Loan (1996); Schatzman (2002); Schwarz i Köckler (2011)):

- (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 od 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 $\text{cond}(\mathbf{A})$ puno veći od 1, kažemo da je matrica \mathbf{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 (Golub i van Loan, 1996) po formuli

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

gdje je σ_1 najveća, a σ_n najmanja singularna vrijednost matrice \mathbf{A} (vidi također t. 3.9, str. 76). Singularne vrijednosti kvadratne matrice \mathbf{A} mogu se lako izračunati Mathematica-naredbom `SingularValueList[A]`.

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 i jedinična matrica ($\text{cond}(\mathbf{D}) = 1$), a njezina determinanta je skoro nula ($\det \mathbf{D} = 10^{-20}$). S druge strane, dijagonalna matrica $\mathbf{S} = \text{diag}(10^5, 10^{-5})$ loše je 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 trokutasta matrica

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

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

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

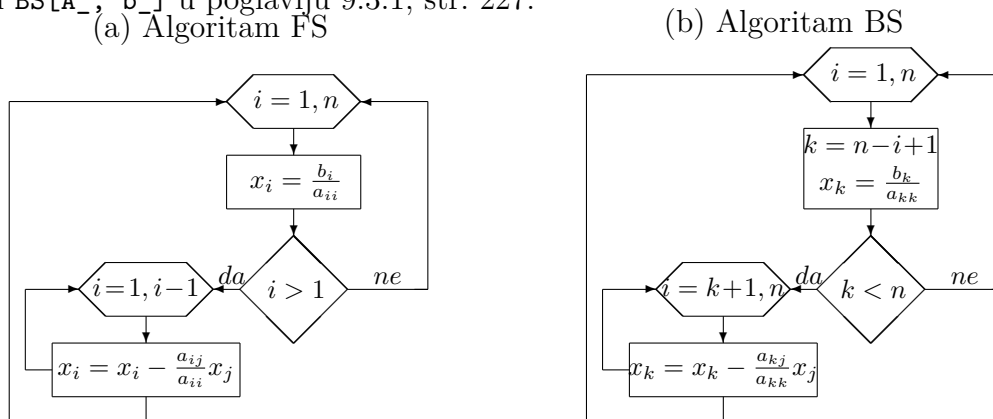
Slično, sustav linearnih jednadžbi $\mathbf{Lx} = \mathbf{b}$, gdje je $\mathbf{L} \in \mathbb{R}^{n \times n}$ donja trokutasta matrica

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

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

$$x_1 = \frac{b_1}{l_{11}}, \quad x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{k=1}^{i-1} l_{ik} x_k \right), \quad i = 2, \dots, n. \quad (3.12)$$

Na Slici 3.1 naveden je dijagram toka za rješavanje donjeg i gornjeg trokutastog sustava. Niže su navedeni algoritmi za rješavanje gornje i donje trokutastih sustava, koji su implementirani u Mathematica-modulima FS[A_, b_] i BS[A_, b_] u poglavlju 9.3.1, str. 227.



Slika 3.1: Rješavanje gornjeg i donjeg trokutastog sustava

Algoritam 3 (Supstitucija unaprijed)

Input: l_{ij} , $i, j = 1, \dots, n$; {Učitati elemente donje trokutaste matrice L }

1: Staviti: $x = 0[n]$; (nul-vektor iz \mathbb{R}^n)

2: Staviti: $x_1 = \frac{b_1}{l_{11}}$;

3: **for** $i = 2$ to n **do**

4: $x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{j=1}^{i-1} l_{ij} x_j \right)$

5: **end for**

Output: x (rješenje: vektor iz \mathbb{R}^n)

Algoritam 4 (Supstitucija unazad)**Input:** u_{ij} , $i, j = 1, \dots, n$; {Učitati elemente gornje trokutaste matrice U }1: Staviti: $x = 0[n]$; (nul-vektor iz \mathbb{R}^n)2: Staviti: $x_n = \frac{b_n}{u_{nn}}$;3: **for** $i = 1$ to $n - 1$ **do**4: $k = n - i$;5: $x_k = \frac{1}{u_{kk}} \left(b_k - \sum_{j=k+1}^n u_{kj} x_j \right)$ 6: **end for****Output:** x (rješenje: vektor iz \mathbb{R}^n)

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 rješenje¹.Budući da je matrica \mathbf{A} nesingularna, bez smanjenja općenitosti možemo pretpostaviti da je tzv. **prvi pivot-element** $a_{11} \neq 0$. Zato ćemo primjenom Gaussovih transformacija matricu \mathbf{A} svesti na oblik

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

¹Iako se Gaussovom metodom može rješavati opći sustav linearnih jednadžbi (gdje broj jednadžbi nije jednak broju nepoznanica ili matrica sustava nije regularna), mi ćemo se zadržati na ovom specijalnom slučaju, koji je i najčešći u primjenama (vidi primjerice Jukić i Scitovski (2004)).

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

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

Slično, bez smanjenja općenitosti možemo pretpostaviti da je sljedeći pivot-element $a_{22}^{(2)} \neq 0$, pa 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 $m_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$,

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

Nakon $n - 1$ ovakvih koraka (*iteracija*) matrica \mathbf{A} prijeći će u oblik gornje trokutaste matrice, gdje su svi dijagonalni elementi različiti od nule.

Cijeli postupaku moguće je provesti „bez pivotiranja” uz nešto jači uvjet na matricu \mathbf{A} , tj. sustav je moguće „bez pivotiranja” svesti na gornje trokutasti onda i samo onda ako su svi glavni minori matrice sustava \mathbf{A} različiti od nule (vidi primjerice Björck i Dahlquist (2007, 2008); Živković (2012)). Takav sustav jednadžbi lako se rješava Algoritmom BS, str. 55.

Primjer 3.3. *Opisanu Gaussovu metodu eliminacije ilustrirat ćemo na rješavanju sustava $\mathbf{Ax} = \mathbf{b}$, gdje je*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -3 \\ 4 \end{bmatrix},$$

uz primjenu sljedećeg *Mathematica* programa

```
In[1] := a={{1,1,0,3},{2,1,-1,1},{3,-1,-1,2},{-1,2,3,-1}}; b={4,1,-3,4};
n=Length[a];
Print[MatrixForm[a], MatrixForm[b]]
Do[
  Do[
    m = a[[i,1]]/a[[1,1]];
```

```

Do[a[[i,k]]=a[[i,k]]-m*a[[1,k]], {k,1,n}];
b[[i]]=b[[i]]-m*b[[1]]
,{i,1+1,n}];
Print[MatrixForm[a], MatrixForm[b] ]
,{1,n-1}]
BS[a, b]

```

Dobivamo

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 2 & 1 & -1 & 1 & 1 \\ 3 & -1 & -1 & 2 & -3 \\ -1 & 2 & 3 & -1 & 4 \end{array} \right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 0 & -1 & -1 & -5 & -7 \\ 0 & -4 & -1 & -7 & -15 \\ 0 & 3 & 3 & 2 & 8 \end{array} \right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 0 & -1 & -1 & -5 & -7 \\ 0 & 0 & 3 & 13 & 13 \\ 0 & 0 & 0 & -13 & -13 \end{array} \right]$$

$$x = (-1, 2, 0, 1)^T.$$

Primjedba 3.3. Primjenom Gaussovih transformacija matricu sustava mogli bismo svesti i na dijagonalnu matricu. U tom slučaju govorimo o **Gauss-Jordanovoj** metodi, koja ima važnu ulogu kod *simpleks-metode* za rješavanje problema linearnog programiranja (Björck i Dahlquist, 2008).

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 bez određene „strategije pivotiranja”, 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. $\mathbf{A}^T = \mathbf{A}$ i $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ za sve $\mathbf{x} \neq \mathbf{0}$. Prema poznatom Sylvesterovom kriteriju (Kurepa, 1967), simetrična matrica \mathbf{A} je pozitivno definitna onda i samo onda ako su svi njezini glavni minori² pozitivni.

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.

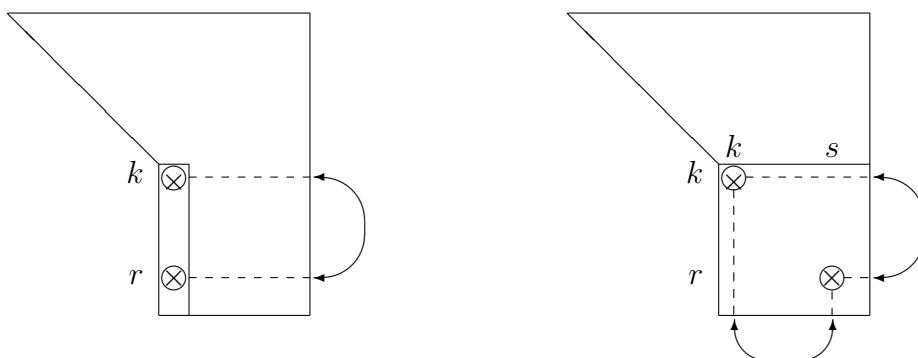
Jedan pokušaj u tom smislu je tzv. **strategija parcijalnog pivotiranja**:

²Glavni 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.

U k -toj iteraciji k -ti redak treba zamijeniti 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 3.2).



Slika 3.2: Strategija parcijalnog (lijevo) i strategija potpunog (desno) pivotiranja

Drugi pokušaj je tzv. **strategija potpunog pivotiranja**:

U k -toj iteraciji najprije pronademo 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 stupca, te k -tog i r -tog retka (vidi shemu na Slici 3.2)

Zadatak 3.4. Pokažite da je Gaussov postupak bez pivotiranja provediv onda i samo onda ako su svi glavni minori matrica sustava različiti od nule.

Primjer 3.4. Ako bismo u Primjeru 3.1, str. 47, u slučaju 2-znamenaste 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 3.5. 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 (vidi Björck i Dahlquist (2008); Golub i 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 3.6. 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 $m_{i,1} := \frac{a_{i1}}{a_{11}}$, $i = 2, \dots, n$ definirani kao u t.3.4, str. 55, 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 3.7. 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-Morriso-nova**³ 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 3.6), 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).

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

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

³J. 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.

Niže je naveden algoritam *LU-dekompozicija*, koji je implementiran u Mathematica-modulu LU[A_] u poglavlju 9.3.2, str. 228.

Algoritam 5 (LU-dekompozicija)

Input: a_{ij} , $i, j = 1, \dots, n$; {Učitati elemente matrice}

1: Staviti:

$$\begin{aligned} U &= O[n] && \text{(nul-matrica reda } n) \\ L &= I[n] && \text{(jedinična matrica reda } n) \end{aligned}$$

2: **for** $i = 1, \dots, n$ **do**

3:

$$\begin{aligned} u_{ij} &:= a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}, && j = i, \dots, n, \\ l_{ji} &:= \frac{1}{u_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk}u_{ki} \right), && j = i + 1, \dots, n. \end{aligned}$$

4: **end for**

Output: Matrice L i U

Odgovarajući Mathematica-modul može se vidjeti u poglavlju 9.3.2, str. 228.

Primjer 3.5. *Primjenom LU-dekompozicije riješit ćemo sustav linearnih jednadžbi gdje su matrica sustava \mathbf{A} i vektor slobodnih koeficijenata \mathbf{b} zadani*

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

Navedenim algoritmom 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

$$\begin{aligned} z &= (1, -1, 0)^T, \\ x &= \left(-\frac{1}{3}, \frac{1}{3}, 0\right)^T. \end{aligned}$$

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

```
In[1] := lud = LU[a];
        z = FS[lud[[1]], b]
        BS[lud[[2]], z]
```

U nekim praktičnim situacijama (primjerice Leontijev input-output model nacionalne privrede⁴) treba više puta rješavati sustav $\mathbf{Ax} = \mathbf{b}_i$, gdje je matrica sustava („matrica tehnologije”) uvijek ista, a vektor slobodnih koeficijenata („output”) se mijenja. Također se može dogoditi da treba riješiti sustave $\mathbf{Ax}_1 = \mathbf{b}_1$ i $\mathbf{Ax}_2 = \mathbf{b}_2$, gdje je \mathbf{b}_2 neka funkcija od \mathbf{x}_1 . 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.

Zadatak 3.8. Izvedite formule navedene u Algoritmu LU koristeći definiciju jednakosti dviju matrica. Specijalizirajte algoritam za slučaj trodijagonalne matrice.

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

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

Primjer 3.6. Za kvadratnu regularnu matricu $\mathbf{A} \in \mathbb{R}^{n \times n}$ vrijedi $\mathbf{AA}^{-1} = \mathbf{I}$. 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 $\mathbf{AA}^{-1} = \mathbf{I}$ možemo zapisati kao n sustava linearnih jednadžbi

$$\mathbf{Ax}_1 = \mathbf{e}_1, \quad \mathbf{Ax}_2 = \mathbf{e}_2, \quad \dots, \quad \mathbf{Ax}_n = \mathbf{e}_n.$$

⁴Wassily Leontief, *Input-Output Economics*, 2nd ed. Oxford University Press, New York, 1986.

Korištenjem LU-dekompozicije matrice \mathbf{A} te ranije sagrađenih algoritama FS i BS možemo izračunati inverznu matricu. Primjerice za matricu iz Primjera 3.5, str. 62 imamo

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

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

Zadatak 3.9. Specijalizirajte LU-dekompoziciju za slučaj trodijagonalne matrice⁵ i nakon toga primijenite algoritme FS i BS. Trodijagonalna matrica pojavljuje se primjerice kod prirodnog kubičnog splinea, str. 36.

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

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} , množeci slijeva elementarnom matricom \mathbf{M}_1 (vidi t. 3.5, str. 60) i zdesna matricom \mathbf{M}_1^T , dobivamo

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

⁵Trodijagonalna matrica je specijalni slučaj tzv. vrpčastih matrica, koje spadaju u tzv. rijetko popunjene matrice (sparse matrices).

Ponavljajući postupak kao u t. 3.5, str. 60, 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 $\mathbf{x}^* = \mathbf{L}\mathbf{z}$.

Niže je naveden algoritam Cholesky-dekompozicija, koji je implementiran u Mathematica-modulu `Cholesky[A_]` u poglavlju 9.3.3, str. 228.

Algoritam 6 (Cholesky-dekompozicija)

Input: a_{ij} , $i, j = 1, \dots, n$; {Učitati elemente matrice}

1: Staviti: $L = O[n]$ (nul-matrica reda n)

2: **for** $i = 1, \dots, n$ **do**

3:

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

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

4: **end for**

Output: Matrica L

Primjer 3.7. Cholesky dekompozicijom treba riješiti sustav $\mathbf{H}\mathbf{x} = \mathbf{b}$, gdje je \mathbf{H} Hilbertova matrica trećeg reda (str. 59), a $\mathbf{b} = (1, 1, 1)^T$.

Algoritmom Cholesky lako dobivamo

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

rješenje odgovarajućeg sustava $\mathbf{L}\mathbf{z} = \mathbf{b}$ je $\mathbf{z} = (1, \sqrt{3}, \sqrt{5})^T$, a rješenje polaznog sustava je $\mathbf{x}^* = (3, -24, 30)^T$.

3.7 QR-dekompozicija

QR-dekompozicija 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 (vidi Householder (1964)). Na taj način sustav $\mathbf{A}\mathbf{x} = \mathbf{b}$ 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 redcima ili stupcima. Mi ćemo razmotriti dovoljno općeniti slučaj, koji se javlja kod rješavanja sustava linearnih jednažbi i kod rješavanja tzv. linearnog problema najmanjih kvadrata (t. 6.1, str. 161).

Pretpostavimo 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 matrica⁶ $\mathbf{Q} \in \mathbb{R}^{m \times m}$ i gornja

⁶Ortogonalna 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 redci 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.

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 (vidi Björck (1996); Gill et al. (1991); Golub i van Loan (1996); Lawson i Hanson (1996); Stewart (1998)). Mi ćemo QR-dekompoziciju graditi primjenom Householderovih transformacija.

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 euklidske 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 $\mathbf{H}\mathbf{b} = \mathbf{a}$. \square

Primjer 3.8. Vektori $\mathbf{a} = (\frac{9}{2}, 1)^T$ i $\mathbf{b} = (3, \frac{7}{2})^T$ su različiti, ali imaju jednaku euklidsku normu, tj. zadovoljavaju (3.23). Definirat ćemo Householderov vektor \mathbf{u} , tako da bude $\mathbf{H}\mathbf{a} = \mathbf{b}$.

Stavimo

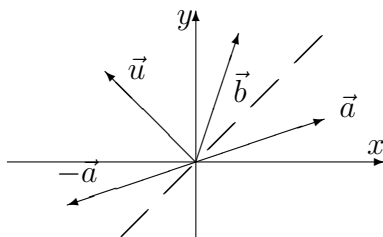
$$\mathbf{u} = \frac{\mathbf{a} - \mathbf{b}}{\|\mathbf{a} - \mathbf{b}\|} = \frac{2}{\sqrt{34}} \left(\frac{3}{2}, -\frac{5}{2} \right)^T.$$

Householderova matrica glasi

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

a za nju 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 3.3: Householderov vektor

Na Slici 3.3 vidi se da su vektori \mathbf{a} i \mathbf{b} osno simetrični na pravac kroz ishodište, koji je okomit na vektor \mathbf{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 $\|\vec{\mathbf{r}}_T\| = \|\vec{\mathbf{r}}_{T'}\|$, gdje su $\vec{\mathbf{r}}_T, \vec{\mathbf{r}}_{T'}$ radij-vektori točaka T, T' . Odredite također odgovarajuću Householderovu matricu \mathbf{H} , koja će vektor $\vec{\mathbf{r}}_T$ preslikavati u vektor $\vec{\mathbf{r}}_{T'}$.

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

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 svođenja 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} podudara, a da se posljednjih $m - k$ komponenti vektora \mathbf{b} poništava. 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

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

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 3.9. 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 3.10. *Primjenom Householderovih transformacija, matricu \mathbf{A} treba transformirati u gornju trokutastu*

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

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 $\mathbf{A} \mathbf{x} = \mathbf{b}$ (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 jednadžbe, 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 $\mathbf{A} \mathbf{x} = \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$. Pretpostavljamo da je $a_{ii} \neq 0$, $i = 1, \dots, n$.

- **Jacobijeva metoda**

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

odnosno:

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

- **Gauss-Seidelova metoda**

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

odnosno:

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

Niže je naveden Jacobijev i Gauss-Seidelov algoritam, a njihova implementacija navedena je u Mathematica-modulima `Jacobi[a_,b_,it_]` i `GS[a_,b_,it_]` u poglavlju 9.3.4, str. 229.

Algoritam 7 (Jacobi)

Input: $a_{ij}, i, j = 1, \dots, n; b_j, j = 1, \dots, n; IT$
 {Učitati elemente matrice A , vektora b i broj iteracija IT }

- 1: Staviti: $x = \xi = 0[n]$; (nul-vektor iz \mathbb{R}^n)
- 2: **for** $k = 1$ to IT **do**
- 3: **for** $i = 1$ to n **do**
- 4: $\xi_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i}^n a_{ij} x_j \right)$
- 5: **end for**
- 6: $x = \xi$
- 7: **end for**

Output: Vektor x (aproksimacija rješenja)

Algoritam 8 (Gauss-Seidel)

Input: $a_{ij}, i, j = 1, \dots, n; b_j, j = 1, \dots, n; \epsilon > 0; k = 0; IT$
 {Učitati elemente matrice A , vektora b , točnost $\epsilon > 0$ i maksimalni broj iteracija}

- 1: Staviti: $x = \xi = 1[n]$; (nenul-vektor iz \mathbb{R}^n)
- 2: **while** **do**
- 3: **for** $i = 1$ to $n - 1$ **do**
- 4: $\xi_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} \xi_j - \sum_{j=i+1}^n a_{ij} x_j \right)$
- 5: **end for**
- 6: $\|\xi - x\|_1 > \epsilon \quad \& \quad k < IT,$
- 7: $x = \xi; k=k+1;$
- 8: **end while**

Output: Vektor x (aproksimacija rješenja)

Primjer 3.11. *Jacobijevom odnosno Gauss-Seidelovom metodom riješit ćemo sustav iz Primjera 2.8, str. 40:*

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

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

U cilju ispitivanja konvergencije navedenih, pa i drugih iterativnih metoda, na Banachovom⁷ 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 vrijedi

$$d(F(\mathbf{x}), F(\mathbf{y})) \leq q d(\mathbf{x}, \mathbf{y}) \quad \text{za sve } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Primjer 3.12. Funkcija $F : [1, \infty) \rightarrow \mathbb{R}$, $F(x) = \frac{x}{2} + \frac{1}{x}$ je kontrakcija s $q = 1/2$. Naime,

$$\begin{aligned} d(F(x), F(y)) &= |F(x) - F(y)| = \left| \frac{x}{2} + \frac{1}{x} - \frac{y}{2} - \frac{1}{y} \right| = \left| \frac{1}{2}(x - y) - \frac{x-y}{xy} \right| \\ &= |x - y| \left| \frac{1}{2} - \frac{1}{xy} \right| \leq \frac{1}{2} |x - y| = \frac{1}{2} d(x, y). \end{aligned}$$

Konvergencija iterativnih metoda za rješavanje sustava linearnih jednadžbi temelji se na važnom Banachovom teoremu o fiksnoj točki (vidi Plato (2003, 2010)).

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.*

$$F(\mathbf{x}^*) = \mathbf{x}^*.$$

⁷Stefan Banach (1892.-1945.) poljski matematičar (University of Lvov) utemeljio je modernu funkcionalnu analizu i dao veliki doprinos teoriji mjere i teoriji topoloških vektorskih prostora.

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 $(\mathbf{x}^{(n)})$ 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 konvergiraće 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 ω , 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 (vidi Björck (1996); Björck i Dahlquist (2007); Gill et al. (1991); Golub i van Loan (1996); Lawson i Hanson (1996); Plato (2003); 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 σ_i^2 su svojstvene vrijednosti⁸ 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 σ_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 $\text{rang } \mathbf{A} = k \leq p$, 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) i (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 i van Loan (1996); Trefethen i Bau (1997)).

Primjer 3.13. (*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 γ , tj. $\mathbf{a}_1^T \mathbf{a}_2 = \cos \gamma$.*

⁸Kažemo 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 λ_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 $\sigma(\mathbf{A}) \subset \mathbb{R}_+$;
- $\text{Tr}(\mathbf{A}) := \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$;
- $\prod_{i=1}^n \lambda_i = \det \mathbf{A}$.

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 γ^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. Izračunajte lijeve singularne vektore $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ iz prethodnog primjera. Formirajte matrice \mathbf{U} i \mathbf{V} i provjerite 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 `Mathematica`- naredbom `SingularValueDecomposition[A]`.

Primjedba 3.9. Broj uvjetovanosti proizvoljne matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$ ranga $k < p = \min\{m, n\}$ (vidi primjerice Björck i Dahlquist (2007); Gill et al. (1991); Golub i van Loan (1996)) može se definirati na sljedeći način:

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^+\|_2 = \frac{\sigma_1}{\sigma_k}, \quad (3.32)$$

gdje su $\sigma_1 > \sigma_2 > \dots > \sigma_k > 0$ singularne vrijednosti matrice \mathbf{A} , a $\mathbf{A}^+ = \mathbf{V}\mathbf{S}^+\mathbf{U}^T$ je 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 jednažbi može se naći u Ackleh et al. (2010); Björck (1996); Björck i Dahlquist (2007, 2008); Ciarlet i Lions (2000); Dahlquist i Björck (1972); Demmel (1997); Deuffhard i Hohmann (2008); Gill et al. (1991); Golub i van Loan (1996); Kincaid i Cheney (1996); Mathews i Fink (1999); Plato (2003, 2010); Schatzman (2002); Schwarz i Köckler (2011); Stewart (1996); Stoer i Bulirsch (2002); Süli i Mayers (2006), a odgovarajuća FORTRAN-programaska podrška može se naći u Press et al. (1992), te posebno u The NAG-Library. Moguće je koristiti se i gotovim korisničkim softwareom *Mathematica* ili *Matlab* (Trefethen i Bau, 1997).

3.10 Zadaci

Zadatak 3.13. Koliko 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. Usporedite broj računskih operacija potrebnih za rješavanje sustava $\mathbf{Ax} = \mathbf{b}$ izračunavanjem inverzne matrice i korištenjem formule $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ i broj potrebnih računskih operacija za rješavanje sustava $\mathbf{Ax} = \mathbf{b}$ primjenom LU-dekompozicije.

Zadatak 3.15. Nekom od navedenih metoda riješite sustave linearnih jednažbi, 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 3.16. U 4-znamenkastoj, odnosno 6-znamenkastoj floating-point aritmetici uz primjenu neke od navedenih metoda riješite sustav linearnih jednažbi

$\mathbf{K}\mathbf{x} = \mathbf{b}$,

$$\mathbf{K} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 1 & 1 & \cdots & 1 \\ -1 & -1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Zadatak 3.17. 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 3.18. Načinimo malu promjenu u elementu a_{21} matrice \mathbf{A} iz Zadatka 3.17, 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 oko 1000 puta je veća od relativne promjene u matrici \mathbf{A} . $\text{cond}(\mathbf{A}) = 1039.52$

Zadatak 3.19. Izradite program koji će učitati red matrice n , elemente matrice \mathbf{A} , te primjenom Gaussove metode eliminacije izračunati inverznu matricu \mathbf{A}^{-1} .

Zadatak 3.20. Izradite program koji će učitati red matrice n , elemente matrice \mathbf{A} , te primjenom LU-dekompozicije izračunati inverznu matricu \mathbf{A}^{-1} .

Zadatak 3.21. Izradite program koji će Gaussovom metodom uz strategiju potpunog pivotiranja riješiti sustav $\mathbf{A}\mathbf{x} = \mathbf{b}$, te niže navedenim postupkom iterativno popraviti rješenje:

Neka su $\tilde{\mathbf{L}}$ i $\tilde{\mathbf{U}}$ neke aproksimacije matrica \mathbf{L} i \mathbf{U} , a $\tilde{\mathbf{x}}$ odgovarajuća aproksimacija rješenja sustava $\mathbf{A}\mathbf{x} = \mathbf{b}$. Označimo $\mathbf{x}^{(1)} = \tilde{\mathbf{x}}$ i računamo vektor-rezidual

$$\mathbf{r}^{(1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(1)},$$

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 dvostrukoj točnosti.

Zadatak 3.22. 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 3.23. Izradite program koji će primjenom LU-dekompozicije riješiti sustav $\mathbf{Ax} = \mathbf{b}$, te na osnovi iterativnog postupka navedenog u Zadatku 3.21 iterativno popraviti rješenje.

Zadatak 3.24. 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

$$y_{ij} = \frac{1}{a_{ii}} \left(\delta_{ij} - \sum_{k=j}^{i-1} a_{ik} y_{kj} \right), \quad i = j, \dots, n.$$

$$z_{ij} = \frac{1}{u_{ii}} \left(\delta_{ij} - \sum_{k=i+1}^j u_{ik} z_{kj} \right), \quad i = j, j-1, \dots, 1.$$

(δ_{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 3.25. Treba riješiti sustav linearnih jednadžbi $\mathbf{Ax} = \mathbf{y}$, gdje je \mathbf{A} tzv. vrpčasta matrica (vidi Björck i Dahlquist (2007, 2008); Živković (2012)).

$$\begin{bmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & c_{n-1} & a_n \end{bmatrix}$$

Matrica iz Zadatka 3.25

$$\begin{bmatrix} \times & 0 & 0 & \times & 0 & \times & 0 & \cdots \\ 0 & \times & 0 & 0 & \times & 0 & \times & \cdots \\ 0 & 0 & \times & 0 & 0 & \times & 0 & \cdots \\ \times & 0 & 0 & \times & 0 & 0 & \times & \cdots \\ 0 & \times & 0 & 0 & \times & 0 & 0 & \cdots \\ \times & 0 & \times & 0 & 0 & \times & 0 & \cdots \\ 0 & \times & 0 & \times & 0 & 0 & \times & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Matrica iz Zadatka 3.26

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 3.26. 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

$$\begin{aligned} \mathbf{a} &= (a_1, a_2, \dots, a_n)^T && \text{— elementi glavne dijagonale,} \\ \mathbf{b} &= (b_1, b_2, \dots, b_{n-4})^T && \text{— elementi 4. sporedne dij. u gornjem trokutu} \\ \mathbf{c} &= (c_1, c_2, \dots, c_{n-6})^T && \text{— elementi 6. sporedne dij. u gornjem trokutu} \\ \mathbf{d} &= (d_1, d_2, \dots, d_{n-4})^T && \text{— elementi 4. sporedne dij. u donjem trokutu} \\ \mathbf{e} &= (e_1, e_2, \dots, e_{n-4})^T && \text{— elementi 6. sporedne dij. u donjem trokutu} \end{aligned}$$

Izradite kompjutorski 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 3.27. Hilbertova matrica \mathbf{H} s elementima $h_{ij} = \frac{1}{i+j-1}$ klasični je 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 3.28. 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 3.29. 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}_∞ normi? Kako se ovo može primijeniti za komparativnu analizu kvalitete računala?

Zadatak 3.30. Cholesky dekompozicija također se može provesti i po redcima 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`-modul i ispitajte ga na prethodnim primjerima.

Poglavlje 4

Rješavanje nelinearnih jednadžbi

Promatramo realnu neprekidnu funkciju f definiranu na zatvorenom intervalu $[a, b]$. Općenito, svaki realni broj ξ , koji je rješenje jednadžbe

$$f(x) = 0, \quad (4.1)$$

nazivamo *nultočkom* funkcije f , koja geometrijski odgovara sjecištu 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

$$f(a) \cdot f(b) < 0, \quad (4.2)$$

onda (vidi (Jukić i Scitovski, 2004, 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:

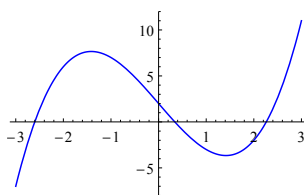
1. Separirati interval I , u kojemu funkcija ima nultočku ξ ;
2. Nekom iterativnom metodom odrediti aproksimaciju nultočke ξ s unaprijed zadanom točnošću.

Primjer 4.1. *Treba separirati realne nultočke funkcije $f(x) = x^3 - 6x + 2$.*

Graf ove funkcije lako je nacrtati (vidi Sliku 4.1). 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 se grafičkim mogućnostima računala (programski sustavi: *Mathematica*, *Matlab*). Ako jednadžbu $f(x) = 0$ preuredimo tako da na lijevu i desnu stranu dođu funkcije čije grafove možemo nacrtati

$$\varphi(x) = \psi(x),$$

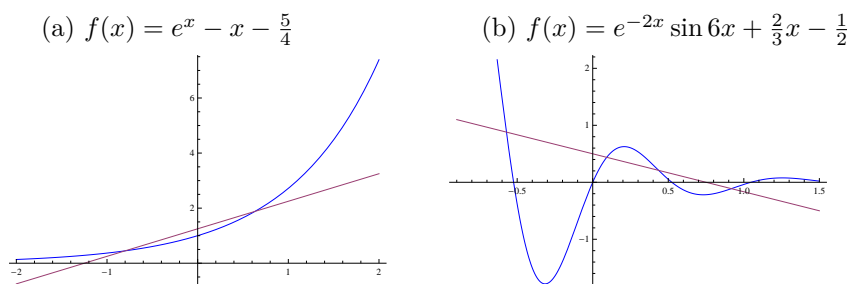
tada apscise točaka sjecišta grafova funkcija φ i ψ predstavljaju nultočke funkcije f . Na ovaj način nećemo dobiti dovoljno dobre aproksimacije nultočaka, ali tako smo u mogućnosti izolirati intervale u kojima leže nultočke.

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*

$$f(x) = e^{-2x} \sin 6x + \frac{2}{3}x - \frac{1}{2}.$$

Jednažbu $f(x) = 0$ 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 4.2b). 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 sljedeća tablica

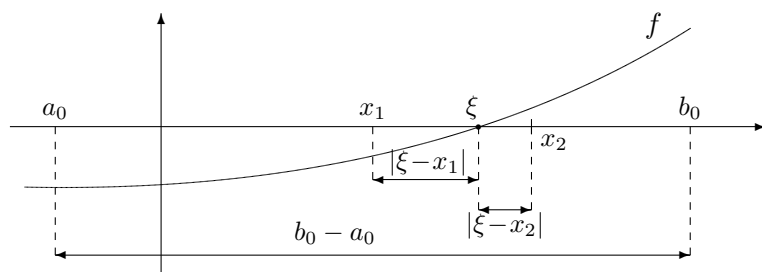
x	0	0.25	0.5	1
$\text{sgn}f(x)$	-1	+1	-1	+1

Za rješavanje jednažbe $f(x) = 0$ postoji čitav niz različitih metoda, a mi ćemo se upoznati s nekima od njih, koje se najčešće koriste u primjenama. Sve metode za rješavanje opće jednažbe $f(x) = 0$ iterativne su 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 jednažbe.

Razmotrit ćemo najprije jednostavnu metodu bisekcije (vidi primjerice Ackleh et al. (2010); Björck i Dahlquist (2007); Stoer i Bulirsch (2002)) za koju je dovoljno da je funkcija f 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



Slika 4.3: Metoda bisekcije

funkcija f prima vrijednosti suprotnog predznaka. Navedeni postupak dalje ponavljamo.

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

$$\begin{aligned}
 x_1 &= \frac{1}{2}(a_0 + b_0), & |\xi - x_1| &\leq \frac{1}{2}(b_0 - a_0), \\
 x_2 &= \begin{cases} \frac{1}{2}(a_0 + x_1), & f(a_0) \cdot f(x_1) < 0 \\ \frac{1}{2}(x_1 + b_0), & f(a_0) \cdot f(x_1) > 0 \end{cases}, & |\xi - x_2| &\leq \frac{1}{2^2}(b_0 - a_0), \\
 &\dots & &
 \end{aligned}$$

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 $f(x) = 0$, pri čemu pogrešku aproksimacije x_n možemo ocijeniti (vidi Sliku 4.3) s

$$|\xi - x_n| \leq \frac{1}{2^n}(b - a).$$

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. 85).*

Nakon 6 iteracija dobivamo $x^* = 0.3$. Rezultat izvođenja iterativnog postupka vidljiv je u Tablici 4.1.

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

Niže je naveden algoritam *Metoda bisekcije*, koji je implementiran u Mathematica-modulu `Bisekcija[f_, a_, b_, eps_, It_, P_]` u poglavlju 9.4.1, str. 230.

Algoritam 9 (Metoda bisekcije)

Input: $f: [a, b] \rightarrow \mathbb{R}$;

Input: $a, b, \epsilon > 0, IT$

{Učitati funkciju, interval, točnost i maksimalni broj iteracija;}

1: **if** $f(a) \cdot f(b) > 0$ **then**

2: $k = 0; a_0 := a; b_0 := b; x_1 = \frac{1}{2}(a_0 + b_0); \Delta = \frac{1}{2}(b_0 - a_0);$

3: **while** $\Delta > \epsilon$ **do**

4: $k = k + 1$

5: **if** $f(a_0) \cdot f(b_0) < 0$ **then**

6: $b_0 = x_1;$

7: **else**

8: $a_0 = x_1$

9: **end if**

10: $x_1 = \frac{1}{2}(a_0 + b_0); \Delta = \frac{1}{2}(b_0 - a_0);$

11: **end while**

12: **return** $\{k, x_1\},$

13: Print["Interval nije dobro izabran"]

14: **end if**

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 ξ 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 Lagrangeovom teoremu o srednjoj vrijednosti (vidi primjerice Jukić i Scitovski (2004)), imamo

$$f(\xi) - f(x_n) = (\xi - x_n)f'(c), \quad c \in I.$$

Odavde 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 ξ jednadžbe $f(x) = 0$ tražimo nekim iterativnim procesom, postavlja se pitanje kada zaustaviti taj proces. Ako pri tome tražimo da apsolutna pogreška aproksimacije x_n nultočke ξ ne bude veća od unaprijed zadano broj $\epsilon > 0$, 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 konstruirati primjer u kojemu 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 ξ (vidi Sliku 4.4a). Naime, u tom slučaju broj m_1 je malen, pa je $\frac{1}{m_1}$ velik, što znači da će zahtjev $\frac{|f(x_n)|}{m_1} < \epsilon$ moći biti ispunjen za vrlo maleni $|f(x_n)|$.

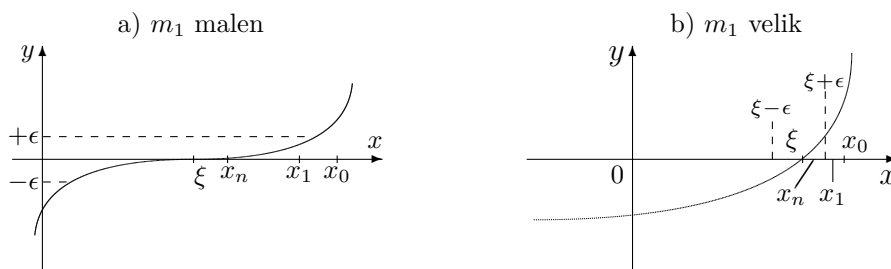
Ako je m_1 velik, $\frac{1}{m_1}$ je malen, pa prema (4.3) pogreška aproksimacije može biti malena (x_n je blizu rješenja ξ), a da vrijednost funkcije $f(x_n)$ bude relativno velika (Slika 4.4b).

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 teško odrediti konstantu m_1 , umjesto kriterija za zaustavljanje procesa $\frac{|f(x_n)|}{m_1} < \epsilon$, često se koristi

$$|x_n - x_{n-1}| < \epsilon_1 \quad \text{i} \quad |f(x_n)| < \epsilon_2, \quad (4.4)$$

gdje su ϵ_1, ϵ_2 unaprijed zadani brojevi.



Slika 4.4: Kriteriji zaustavljanja iterativnog procesa

Primjer 4.5.¹ Kao ilustraciju Primjedbe 4.1, str. 90, promatrajmo problem rješavanja sljedećih jednačbi metodom bisekcije:

$$(a) \quad f(x) = 0.001x + \frac{1}{2} + \frac{\pi}{2} + \arctg x = 0, \quad x \in [-600, -400]$$

$$(b) \quad g(x) = 1000(x - 4) - e^x = 0, \quad x \in [4, 5].$$

(a) Graf funkcije f na zadanom intervalu približno odgovara grafu prikazanom na Slici 4.4a 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 4.4b 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.

Važna karakteristika svake iterativne metode je njena brzina (red) konvergencije. To je broj koji je usko povezan s brojem potrebnih iteracija za postizanje tražene točnosti u iterativnom procesu. 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 $\Delta x_n = |\xi - x_n|$ apsolutna pogreška n -te aproksimacije. Tada, ako postoje dvije pozitivne konstante $A, r \in \mathbb{R}_+$, takve da vrijedi

$$\lim_{n \rightarrow \infty} \frac{\Delta x_{n+1}}{(\Delta x_n)^r} = A, \quad (4.5)$$

¹Ovaj primjer konstruirao je J. Muharemović, student III. godine Elektrotehničkog fakulteta Sveučilišta J.J.Strossmayera u Osijeku, 1998.

kažemo da metoda ima red konvergencije r (ili brzinu konvergencije r).

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 ξ .

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

$$\Delta x_{n+1} \leq A(\Delta x_n)^r, \quad n \geq n_0.$$

4.3 Metoda jednostavnih iteracija

Kaže se da je točka $\xi \in [a, b]$ *fiksna točka* neke funkcije $\varphi: [a, b] \rightarrow \mathbb{R}$ ako je $\varphi(\xi) = \xi$. Traženje nultočke funkcije $f: [a, b] \rightarrow \mathbb{R}$ na intervalu $[a, b]$ može se povezati s traženjem fiksne točke funkcije $\varphi: [a, b] \rightarrow \mathbb{R}$,

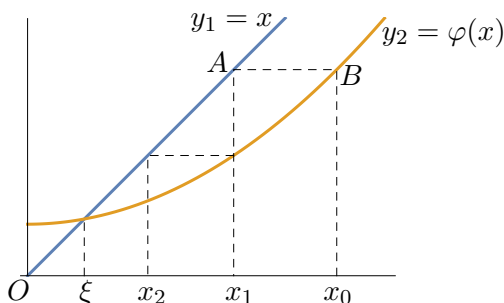
$$\varphi(x) := x - f(x), \tag{4.6}$$

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 φ , a rješavanje jednadžbe $f(x) = 0$ ekvivalentno je traženju fiksne točke funkcije φ , tj. rješavanju jednadžbe

$$x = \varphi(x). \tag{4.7}$$

Za funkciju φ 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 jednačbe $f(x) = 0$, 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 $\triangle(O, x_1, B)$ jednakokrčan. 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 pokazuje sljedeći jednostavni primjer.

Primjer 4.6. *Potražimo rješenje jednačbe $f(x) \equiv e^x - x - \frac{5}{4} = 0$ na intervalu $[0, 1]$ uz primjenu metode jednostavnih iteracija rješavanjem jednačbi*

$$\begin{aligned} x &= \varphi_1(x), & \varphi_1(x) &= e^x - \frac{5}{4}, \\ \text{ili} \\ x &= \varphi_2(x), & \varphi_2(x) &= \ln\left(x + \frac{5}{4}\right). \end{aligned}$$

Uz izbor iste početne aproksimacije $x_0 = 1$, odgovarajući nizovi funkcijskih vrijednosti $(\varphi_1(x_k))$ i $(\varphi_2(x_k))$ prikazani su u Tablici 4.2.

Kao što se može primijetiti, izbor funkcije φ_1 daje divergentan niz, a izbor funkcije φ_2 daje konvergentan niz.

Općenito, mogu se postaviti sljedeća pitanja:

k	0	1	2	3	4	5	6
$\varphi_1(x_k)$	1.4683	3.0918	20.766	1×10^9	–	–	–
$\varphi_2(x_k)$	0.8109	0.7232	0.6796	0.6573	0.6457	0.6396	0.6364

Tablica 4.2: Metoda jednostavnih iteracija

- Je li niz definiran s (4.8) konvergentan? Ako nije, je li moguće novim izborom funkcije φ 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 ξ jednadžbe (4.7) nije poznato, kako ocijeniti pogrešku pojedine aproksimacije x_n ?

Odgovor na sva ova pitanja daje sljedeći teorem².

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) $\exists q \in \langle 0, 1 \rangle$, takav da je $|\varphi'(x)| \leq q$ za svaki $x \in \langle a, b \rangle$.

Tada postoji jedinstveni $\xi \in I$ takav da je $\varphi(\xi) = \xi$. 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 ξ 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

$$|\xi - x_{n+1}| \leq q |\xi - x_n|. \quad (4.12)$$

²Ovaj teorem je specijalni slučaj općenitijeg Banachovog teorema o fiksnoj točki (vidi primjerice Björck i Dahlquist (2007); Deuffhard (1995); Deuffhard i Hohmann (2008)).

Dokaz. Kako je $x_n = \varphi(x_{n-1})$ i $x_{n+1} = \varphi(x_n)$, prema Lagrangeovom teoremu o srednjoj vrijednosti postoji realan broj $c \in \langle a, b \rangle$, takav da je

$$x_{n+1} - x_n = \varphi(x_n) - \varphi(x_{n-1}) = (x_n - x_{n-1})\varphi'(c).$$

Zato je

$$|x_{n+1} - x_n| \leq q|x_n - x_{n-1}| \quad \text{za svaki } n = 1, 2, \dots \quad (4.13)$$

Sukcesivnom primjenom nejednakosti (4.13) dobivamo

$$|x_{n+1} - x_n| \leq q|x_n - x_{n-1}| \leq \dots \leq q^n|x_1 - x_0|. \quad (4.14)$$

Konstruirajmo ovakav red

$$x_0 + (x_1 - x_0) + (x_2 - x_1) + \dots + (x_n - x_{n-1}) + \dots \quad (4.15)$$

Niz njegovih parcijalnih suma jednak je nizu (x_n) , tj. $s_n = x_n$.

Zbog (4.14) jedna majoranta ovog reda je red

$$|x_0| + |x_1 - x_0| + q|x_1 - x_0| + \dots + q^{n-1}|x_1 - x_0| + \dots \quad (4.16)$$

kojeg možemo zapisati ovako

$$|x_0| + |x_1 - x_0|(1 + q + \dots + q^{n-1} + \dots).$$

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 odgovarajući niz (s_n) parcijalnih suma reda (4.15), konvergentan. Kako je $s_n = x_n$, to znači da je i niz (x_n) konvergentan, tj. postoji realan broj ξ , 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 u (4.9) pustimo $n \rightarrow \infty$, onda zbog neprekidnosti funkcije φ imamo $\xi = \varphi(\xi)$, što znači da je ξ korijen jednačine (4.7).

Dokažimo da je ξ jedino rješenje jednačine (4.7) u intervalu I . Pretpostavimo suprotno, tj. pretpostavimo da je $\bar{\xi}$ također rješenje jednačine (4.7) i da vrijedi $\bar{\xi} = \varphi(\bar{\xi})$. Tada prema teoremu o srednjoj vrijednosti mora postojati neki $c \in \langle a, b \rangle$, takav da je

$$\bar{\xi} - \xi = \varphi(\bar{\xi}) - \varphi(\xi) = (\bar{\xi} - \xi)\varphi'(c),$$

odakle slijedi

$$(\bar{\xi} - \xi)(1 - \varphi'(c)) = 0.$$

Budući da je $1 - \varphi'(c) \neq 0$ zbog uvjeta (ii), mora biti $\bar{\xi} = \xi$. Dakle, ξ je jedino rješenje jednadžbe (4.7).

Prijedimo na dokaz 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 prethodne nejednakosti 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, zaključujemo da postoji neki $c \in \langle a, b \rangle$, takav da je

$$|\xi - x_{n+1}| = |\varphi(\xi) - \varphi(x_n)| = |(\xi - x_n)\varphi'(c)| \leq q|\xi - x_n|,$$

odakle neposredno slijedi (4.12). □

Primjedba 4.2. Ako je unaprijed zadana točnost ϵ 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 ϵ . Na taj način osigurali smo da apsolutna pogreška aproksimacije ne bude veća od ϵ .

Zadatak 4.1. Geometrijski objasnite značenja uvjeta (i) i (ii).

Primjer 4.7. 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 jednažbu napišemo u obliku

$$x = e^x - \frac{5}{4},$$

funkcija φ 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 jednažbu napisati u obliku

$$x = \ln(x + \frac{5}{4}), \quad x > -\frac{5}{4}.$$

Sada je $\varphi(x) = \ln(x + \frac{5}{4})$, a $\varphi'(x) = 1/(x + \frac{5}{4})$. Lako se vidi da je (nacrtajte graf funkcije φ')

$$\varphi'(x) \leq \frac{4}{9} < 1 \quad \text{za svaki } x \in I \quad \text{i} \quad \varphi([0, 1]) = \left[\ln \frac{5}{4}, \ln \frac{9}{4} \right] \approx [0.223, 0.811] \subset [0, 1].$$

To znači da su uvjeti Teorema 4.1 ispunjeni.

Niže je naveden algoritam *Metode jednostavnih iteracija*, koji je implementiran u `Mathematica`-modulu `MIT[fi_, a_, b_, eps_, q_, IT_]` u poglavlju 9.4.2, str. 230.

Algoritam 10 (Metoda jednostavnih iteracija)

Input: $\varphi: [a, b] \rightarrow \mathbb{R};;$

Input: $a, b, \epsilon, q, IT;$

1: Staviti $n = 0, x_n = a;$

2: **while do**

3: $n = n + 1; x_n = \varphi(x_{n-1}); \Delta x_n = \frac{q}{1-q} |x_n - x_{n-1}|;$

4: Print[$n, x_n, x_n - \varphi(x_n), \Delta x_n$];

5: $\Delta x_n > \epsilon \ \& \ n > IT,$

6: $x_{n-1} = x_n;$

7: **end while**

Output: $\{x_n, n\}$

Algoritam i odgovarajući `Mathematica`-modul `MIT[fi_, a_, b_, eps_, q_, IT_]` ilustrirat ćemo na Primjeru 4.2. Najprije definiramo funkcije f i φ

```
In[1]:= f[x_]:= Exp[x] - x - 5/4; fi[x_]:= Log[x + 5/4]
```

Izabравši interval $[0, 1]$, točnost $\epsilon = 0.005$, parametar $q = 0.8$ i maksimalni broj iteracija $IT = 20$, modul pokrećemo na sljedeći način.

```
In[2]:= MIT[fi, 0., 1., .005, .8, 20]
```

Cijeli iterativni proces prikazan je u Tablici 4.3. U trećem, odnosno četvrtom stupcu prikazane su vrijednosti desne strane ocjene (4.11), odnosno ocjene (4.10). 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.

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 4.3: Metoda jednostavnih iteracija

Zadatak 4.2. Pogrešku aproksimacije alternativno možemo 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.

Primjer 4.8. 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 kojemu 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 φ 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 φ 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 ispunjeni uvjeti Teorema 4.1. Tijek odgovarajućeg iterativnog procesa prikazan je u Tablici 4.4. 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 4.4: Metoda jednostavnih iteracija

Na ovom primjeru vidi se da iako ocjena (4.10) u drugoj iteraciji ne osigurava niti 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, pa zbog toga nastojimo da ocjena pogreške ne bude previše gruba.

Primjedba 4.3. 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. 71) ili nelinearnih (t. 4.5, str. 115) jednažbi.

4.3.1 Lipschitz-neprekidna funkcija

Zahtjev derivabilnosti na funkciju φ iz Teorema 4.1 očigledno je prejak jer se u iterativnom procesu (4.8) ne koristi. Prirodni zahtjev na funkciju φ bila bi neprekidnost uz dodatnu pretpostavku ograničenosti rasta ili pada. Takva svojstva imaju tzv. „Lipschitz-neprekidne funkcije” (vidi primjerice Ackleh et al. (2010); Björck i Dahlquist (2007); Plato (2003, 2010); Stoer i Bulirsch (2002); Süli i Mayers (2006)).

Definicija 4.2. Kažemo da je funkcija $g : I \rightarrow \mathbb{R}$, $I = [a, b]$ Lipschitz-neprekidna s konstantom $L > 0$ i pišemo $g \in Lip_L(I)$ ako za svaki $x, y \in I$ vrijedi

$$|g(x) - g(y)| \leq L|x - y|. \quad (4.17)$$

Primjedba 4.4. Primijetite da je sukladno Definiciji 3.3, str. 74, Lipschitz-neprekidna funkcija s konstantom $1 > L > 0$ ujedno i kontrakcija na I .

Primjer 4.9. 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 $x \in \mathbb{R}$.

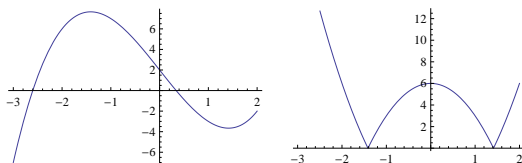
Za $x \neq y$ Lipschitzov uvjet (4.17) možemo zapisati u obliku

$$-L \leq \frac{g(x)-g(y)}{x-y} \leq L \quad \forall x, y \in [a, b],$$

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 4.5. 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 4.12). Ako je $f: [a, b] \rightarrow \mathbb{R}$ neprekidna funkcija, koja ima po dijelovima neprekidnu derivaciju, onda primjenom Lagrangeovog teorema o srednjoj vrijednosti možemo odrediti konstantu $L = \sup_{x \in [a, b]} |f'(x)|$ (vidi Primjer 4.12).

Primjer 4.10. Odredimo Lipschitzovu konstantu funkcije $f: [-3, 2] \rightarrow \mathbb{R}$, $f(x) = x^3 - 6x + 2$. Na Slici 4.6 može se vidjeti da je $L = |f'(-3)| = 21$.

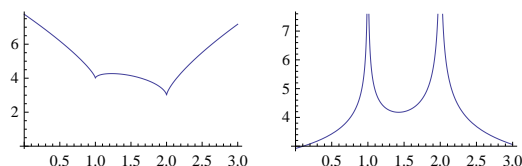


Slika 4.6: Funkcija $f(x) = x^3 - 6x + 2$ i njena derivacija $x \mapsto |f'(x)|$

Primjer 4.11. Odredimo Lipschitzovu konstantu L funkcije $f: [0, 3] \rightarrow \mathbb{R}$, $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$. Njena funkcije

$$f'(x) = \frac{4}{3\sqrt[3]{x-1}} + \frac{2}{\sqrt[3]{x-2}}$$

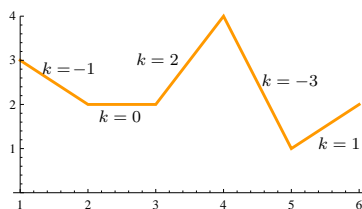
nije definirana u točkama $x \in \{1, 2\}$ (vidi Sliku 4.7), a Lipschitzova konstanta u ovom slučaju ne postoji.

Slika 4.7: Funkcija $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$ i $x \mapsto |f'(x)|$

Primjer 4.12. Lipschitzova 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 4.8, očigledno je $L = 3$.

Slika 4.8: Koeficijenti smjera po dijelovima linearne funkcije f

Primjedba 4.6. Definicija Lipschitz-neprekidne funkcije lako se može proširiti na funkcije više varijabli $g: D \rightarrow \mathbb{R}^n$, gdje je $D \subseteq \mathbb{R}^n$ podskup (vidi primjerice Ackleh et al. (2010); Dennis i Schnabel (1996); Deuffhard (1995); Deuffhard i Hohmann (2008); Ortega i Rheinboldt (1970); Plato (2003); Süli i Mayers (2006)). Kažemo da je takva funkcija *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.

4.3.2 Metoda jednostavnih iteracija za Lipschitz-neprekidnu funkciju

Ponovo razmotrimo iterativnu metodu jednostavnih iteracija

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, \dots, \quad (4.18)$$

za rješavanje jednadžbe

$$x = \varphi(x), \quad (4.19)$$

uz pretpostavku da je φ Lipschitz-neprekidna funkcija na intervalu $I = [a, b]$. Vrijedi teorem o konvergenciji sličan Teoremu 4.1.

Teorem 4.2. *Neka je $\varphi : I = [a, b] \rightarrow \mathbb{R}$ funkcija za koju vrijedi:*

- (i) $\varphi(x) \in I$ za svaki $x \in I$,
- (ii) $\varphi \in Lip_L(I)$, $0 < L < 1$.

Tada postoji jedinstveni $\xi \in I$ takav da je $\varphi(\xi) = \xi$. Osim toga, za proizvoljni $x_0 \in I$, niz definiran s (4.18) konvergira prema ξ i vrijede ovakve ocjene pogreške aproksimacije

$$|\xi - x_n| \leq \frac{L}{1-L} |x_n - x_{n-1}|, \quad (4.20)$$

$$|\xi - x_n| \leq \frac{L^n}{1-L} |x_1 - x_0|. \quad (4.21)$$

Metoda ima linearnu brzinu konvergencije, tj. vrijedi

$$|\xi - x_{n+1}| \leq L |\xi - x_n|. \quad (4.22)$$

Dokaz. Primijetimo najprije da je zbog (i), $x_n \in I$, za svaki $n = 0, 1, \dots$. Kako je $x_n = \varphi(x_{n-1})$ i $x_{n+1} = \varphi(x_n)$, koristeći (ii), dobivamo

$$|x_{n+1} - x_n| = |\varphi(x_n) - \varphi(x_{n-1})| \leq L |x_n - x_{n-1}|, \quad \forall n = 1, 2, \dots \quad (4.23)$$

Sukcesivnom primjenom nejednakosti (4.23) dobivamo

$$|x_{n+1} - x_n| \leq L |x_n - x_{n-1}| \leq \dots \leq L^n |x_1 - x_0|, \quad (4.24)$$

i također

$$\begin{aligned} |x_n - x_{n+k}| &\leq |x_n - x_{n+1}| + |x_{n+1} - x_{n+2}| + \dots + |x_{n+k-1} - x_{n+k}| \\ &\leq L (1 + L + \dots + L^{k-1}) |x_n - x_{n-1}| = L \frac{1-L^k}{1-L} |x_n - x_{n-1}| \\ &< \frac{L}{1-L} |x_n - x_{n-1}|. \end{aligned} \quad (4.25)$$

Također, koristeći (4.24) u (4.25) dobivamo

$$|x_n - x_{n+k}| \leq \frac{L^n}{1-L} |x_1 - x_0|, \quad (4.26)$$

Kako je $0 \leq L < 1$, za proizvoljni $\epsilon > 0$ postoji $n_0 \in \mathbb{N}$, takav da za $n, n+k > n_0$, $|x_n - x_{n+k}| < \epsilon$, tj. (x_n) je Cauchyjev niz u \mathbb{R} , a kako je \mathbb{R} potpun prostor (Jukić, 2012, str. 166), niz (x_n) je konvergentan.

Nadalje, zbog neprekidnosti funkcije φ iz (4.18) slijedi

$$\xi = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \varphi(x_n) = \varphi\left(\lim_{n \rightarrow \infty} x_n\right) = \varphi(\xi).$$

Dakle, niz (x_n) konvergira prema fiksnoj točki ξ funkcije φ , što znači da je ξ rješenje jednačbe (4.19).

Ako u (4.25), odnosno (4.26), pustimo $k \rightarrow \infty$, dobivamo ocjenu pogreške (4.20), odnosno (4.21).

Nadalje, dokažimo jedinstvenost fiksne točke ξ . Pretpostavimo suprotno, tj. pretpostavimo da je i $\bar{\xi} \neq \xi$ također fiksna točka preslikavanja φ . Tada vrijedi

$$|\bar{\xi} - \xi| = |\varphi(\bar{\xi}) - \varphi(\xi)| \leq L|\bar{\xi} - \xi| < |\bar{\xi} - \xi|,$$

što je kontradikcija. Dakle fiksna točka ξ je jedinstvena.

I na kraju, ocjena (4.22) slijedi iz

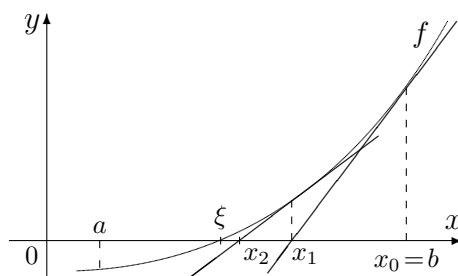
$$|\xi - x_{n+1}| = |\varphi(\xi) - \varphi(x_n)| \leq L|\xi - x_n| \quad \square$$

4.4 Newtonova metoda

Pretpostavimo da smo na neki način odredili interval $I = [a, b]$, u kojemu 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

$$f_1(x) = f(x_0) + (x - x_0)f'(x_0), \quad (4.27)$$

čiji graf je tangenta na graf funkcije f u točki $(x_0, f(x_0))$ (Slika 4.9a). Zato ćemo umjesto rješavanja jednačbe (4.1) rješavati jednačbu $f_1(x) = 0$.



Slika 4.9: Newtonova metoda tangenti

Geometrijski gledano, umjesto traženja sjecišta grafa funkcije f s osi x , tražimo sjecište tangente s osi x . To je razlog zbog kojeg ovu metodu nazivamo *metoda tangenti*. Rješenje jednadžbe $f_1(x) = 0$ označimo s x_1

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)}. \quad (4.28)$$

Ponavljajući postupak, dobivamo niz x_0, x_1, \dots zadan rekurzivnom formulom³

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots \quad (4.29)$$

Važno je primijetiti da u svakoj iteraciji konstruiramo lokalni aproksimand funkcije f i onda tražimo nultočku tog aproksimanda.

Teorem 4.3. *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 i duga derivacija funkcije f na intervalu I imaju stalan predznak.*

Tada, ako je $x_0 \in I$ izabran tako da je

$$f(x_0) \cdot f''(x_0) > 0, \quad (4.30)$$

onda niz definiran s (4.29) konvergira prema jedinstvenom rješenju ξ jednadžbe $f(x) = 0$.

Pri tome vrijedi ocjena pogreške aproksimacije

$$|\xi - x_n| \leq \frac{M_2}{2m_1} (x_n - x_{n-1})^2, \quad (4.31)$$

gdje je

$$m_1 = \min_{x \in I} |f'(x)|, \quad M_2 = \max_{x \in I} |f''(x)|, \quad (4.32)$$

³U literaturi često ovu metodu možemo naći pod imenom *Newton-Raphsonova metoda*.

a metoda ima kvadratnu brzinu konvergencije, tj. vrijedi

$$|\xi - x_{n+1}| \leq \frac{M_2}{2m_1} (\xi - x_n)^2. \quad (4.33)$$

Dokaz. Razmotrimo slučaj⁴

$$f(a) < 0, f(b) > 0, f'(x) > 0, f''(x) > 0 \text{ za svaki } x \in I.$$

Najprije primijetimo da postoji jedinstveni $\xi \in I$ takav da je $f(\xi) = 0$ i izaberimo $x_0 \in I$ tako da je ispunjeno (4.30) (primjerice možemo uzeti $x_0 = b$). Kako je po pretpostavci $f''(x_0) > 0$, onda je i $f(x_0) > 0$, pa je $x_0 > \xi$ jer je f rastuća funkcija.

Primjenom matematičke indukcije pokazat ćemo da je niz definiran s (4.29) monotonno padajući i ograničen odozdo s ξ .

- 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 ξ 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 ξ .

- 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 (4.29) slijedi

$$x_n - x_{n+1} = \frac{f(x_n)}{f'(x_n)} > 0,$$

što znači da je niz (x_n) monotonno padajući.

⁴Ostali slučajevi dokazuju se analogno.

Budući da je niz (x_n) monotono 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 je $\bar{\xi} = \lim x_n$. Ako sada u (4.29) 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 $\bar{\xi} = \xi$.

Prijedimo na dokaz ocjene (4.31). Prema Taylorovoj formuli postoji realan broj c između x_n i x_{n-1} , takav da vrijedi

$$f(x_n) = f(x_{n-1}) + f'(x_{n-1}) \cdot (x_n - x_{n-1}) + \frac{1}{2} \cdot f''(c) \cdot (x_n - x_{n-1})^2.$$

Zbog (4.29), ostaje samo

$$f(x_n) = \frac{1}{2} f''(c) (x_n - x_{n-1})^2,$$

odakle je

$$|f(x_n)| \leq \frac{1}{2} M_2 (x_n - x_{n-1})^2. \quad (4.34)$$

Iz (4.3), koristeći (4.34), dobivamo ocjenu (4.31).

Još treba dokazati kvadratnu brzinu konvergencije metode, tj. dokazati nejednakost (4.33). Prema Taylorovoj formuli postoji realan broj c između ξ i x_n , takav da vrijedi

$$0 = f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \frac{1}{2} \cdot f''(c)(\xi - x_n)^2.$$

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 (4.29), dobivamo

$$x_{n+1} - \xi = \frac{1}{2} \cdot \frac{f''(c)}{f'(x_n)} (\xi - x_n)^2,$$

odakle slijedi (4.33). □

Primjedba 4.7. 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. 9), tada je prema (4.33)

$$|\xi - x_{n+1}| \leq \frac{M_2}{2m_1}(\xi - x_n)^2 \leq 2\frac{1}{4}10^{-2k} = \frac{1}{2}10^{-2k},$$

tj. $(n + 1)$ -va aproksimacija ima $2k$ signifikantnih decimala.

Primjedba 4.8. Za Newtonovu metodu kažemo da je **lokalno konvergentna**. To znači da postoji neka okolina oko nultočke ξ u kojoj su uvjeti Teorema 4.3 ispunjeni i u kojoj možemo proizvoljno birati početnu aproksimaciju x_0 , tako da iterativni proces konvergira prema rješenju ξ .

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 4.3. Za koju će vrijednost početne aproksimacije x_0 Newtonova metoda za funkciju $f(x) = \arctg x$ proizvoditi „kruženje” ($x_1 = -x_0, x_2 = x_0, \dots$)?

Niže je naveden algoritam Newtonove metode tangenti, koji je implementiran u jednostavnom `Mathematica`-modulu `Newton[f_, a_, b_, eps_, m1_, M2_, IT_]` u poglavlju 9.4.3, str. 231.

Algoritam 11 (Newtonova metoda tangenti)

Input: $f: [a, b] \rightarrow \mathbb{R};;$

Input: $a, b, \epsilon, m_1, M_2, IT;$

1: Staviti $n = 0$

2: **if** $f[a] * f''[a] > 0$ **then**

3: $x_0 = a$

4: **else**

5: $x_0 = b$

6: **end if**

7: **while do**

8: $n = n + 1; \quad x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}; \quad \Delta x_n = \frac{|f(x_n)|}{m_1};$

9: Print[$n, x_n, f(x_n), \Delta x_n$];

10: $\Delta x_n > \epsilon$ & $n > IT$,

11: $x_{n-1} = x_n;$

12: **end while**

Output: $\{x_n, n\}$

Primjer 4.13. *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 4.10a). 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 4.5 (usporedite desne strane ocjena (4.3) i (4.31)).

n	x_n	$\frac{ f(x_n) }{m_1}$	$\frac{M_2}{2m_1}(x_n - x_{n-1})^2$	$f(x_n)$	$f'(x_n)$
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 4.5: Newtonova metoda za funkciju iz Primjera 4.13

Na prethodnom Primjeru 4.13 ilustrirat ćemo Algoritam 11 i odgovarajući Mathematica-modul `Newton[f_, a_, b_, eps_, m1_, M2_, IT_]`. Najprije definiramo funkciju f , interval $[a, b]$ u kojemu ćemo tražiti nultočku, točnost $\epsilon > 0$, parametre m_1, M_2 i maksimalni broj iteracija IT .

```
In[1]:= f[x_]:= Exp[x] + x^2 - 2;
        a=.5; b=2.; eps=.005; m1=Abs[f'[b]]; M2 = Abs[f''[a]]; IT=20;
```

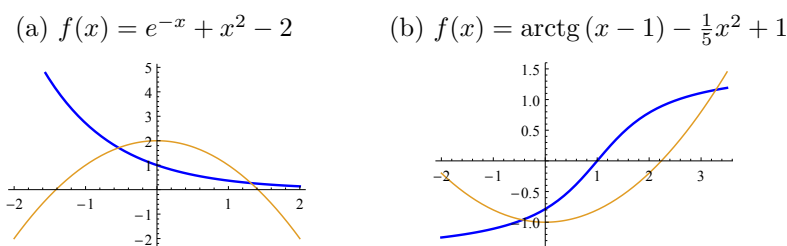
Nakon toga, modul pokrećemo na sljedeći način.

```
In[2]:= Newton[f, a, b, eps, m1, M2, IT]
```

Primjer 4.14. *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 4.10b). Mi ćemo potražiti nultočku iz intervala



Slika 4.10: Separacija nultočaka iz Primjera 4.13 i Primjera 4.14

n	x_n	(4.3)	(4.31)	$f(x_n)$	$f'(x_n)$
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 4.6: Newtonova metoda za funkciju iz Primjera 4.14

[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 4.6.

Primjer 4.15. 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?

Označimo s d duljinu spremnika, a s α kut kružnog isječka (Slika 4.11). 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

$$V_h = \frac{1}{2}r^2d(\alpha - \sin \alpha), \quad (*)$$

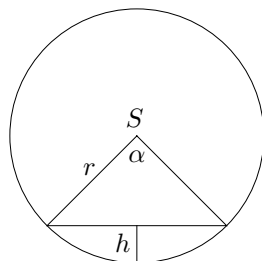
gdje je d duljina spremnika.

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 α , za koji je $V_h = V_q$, dobivamo iz jednačbe

$$\alpha - \sin \alpha - 2\pi q = 0. \quad (**)$$

Kako je $\cos \frac{\alpha}{2} = \frac{r-h}{r}$, nakon što smo iz (***) odredili α , visinu h razine nafte dobivamo iz formule

$$h = r \left(1 - \cos \frac{\alpha}{2}\right). \quad (***)$$



Slika 4.11: Razina spremnika

Primjenom Newtonove metode tangenti uz $\alpha_0 = 3$ i $\epsilon = 0.00005$, dobivamo $\alpha = 2.3099$ ($\approx 132^\circ$). Nakon toga iz (***) dobivamo $h = 59.6\text{cm}$.

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 grafa 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 4.12a)

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}, \text{ ako je } f(x_1) \neq f(x_0).$$

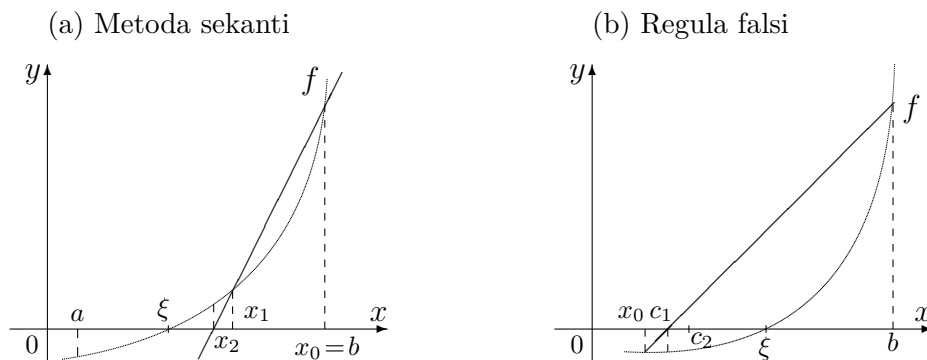
Ponavljajući postupak, dobivamo niz definiran rekurzivnom formulom

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad f(x_n) \neq f(x_{n-1}), \quad n = 1, 2, \dots \quad (4.35)$$

Ako su početne aproksimacije x_0 , x_1 birane dovoljno blizu rješenja ξ , 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 Björck i Dahlquist (2008); Schwarz i Köckler (2011); Stoer i Bulirsch (2002)).

Zadatak 4.4. Pokažite da se rekurzivna formula (4.35) može dobiti izravno iz Newtonove metode (4.29), str. 104, tako da $f'(x_n)$ aproksimiramo kvocijentom razlika

$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$



Slika 4.12: Modifikacije Newtonove metode

Regula falsi (*metoda krivih položaja*) je metoda koju također možemo smatrati jednom modifikacijom Newtonove metode, odnosno jednom varijantom metode sekanti.⁵

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 $(x_0, f(x_0))$, $(b_0, f(b_0))$ (Slika 4.12b). 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

$$\text{ako je } f(c_1)f(x_0) > 0, \quad \begin{array}{ll} x_1 = c_1; & \text{inače } x_1 = c_1 \\ b_1 = b_0 & b_1 = x_0 \end{array}$$

i povucimo pravac točkama $(x_1, f(x_1))$, $(b_1, f(b_1))$. Ponavljajući postupak dobivamo niz (x_n) koji linearnom brzinom konvergira prema jednom korijenu jednažbe $f(x) = 0$ 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 4.5. Izradite **Mathematica**-module za metodu sekanti i metodu regula falsi i testirajte ih na priloženim primjerima.

Primjer 4.16. *Gibanje balističkog projektila u vakuumu određeno je sustavom diferencijalnih jednažbi (vidi primjerice Molitz i Ströbel (1963))*

$$\frac{d^2 x(t)}{dt^2} = 0, \quad \frac{d^2 y(t)}{dt^2} = -g(y),$$

⁵Također se može pokazati (Stoer i Bulirsch, 2002) da je regula falsi specijalni slučaj metode jednostavnih iteracija.

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 , α izlazni kut, v_0 početna brzina projektila (u ms^{-1}) i g akceleracija sile teže (vidi Sliku 4.13a).

Za male visine y može se pretpostaviti da je funkcija g konstantna (približno 9.81 ms^{-2}). Rješenje sustava diferencijalnih jednadžbi možemo zapisati u parametarskom obliku

$$\begin{aligned} (a) \quad x(t) &= v_0 t \cos \alpha, \\ (b) \quad y(t) &= v_0 t \sin \alpha - \frac{1}{2} g t^2, \end{aligned} \tag{B.1}$$

ili eksplicitno (isključujući vrijeme t iz jednadžbi (B.1))

$$y = x \tan \alpha - \frac{x^2}{4h \cos^2 \alpha}, \quad h = \frac{v_0^2}{2g}. \tag{B.2}$$

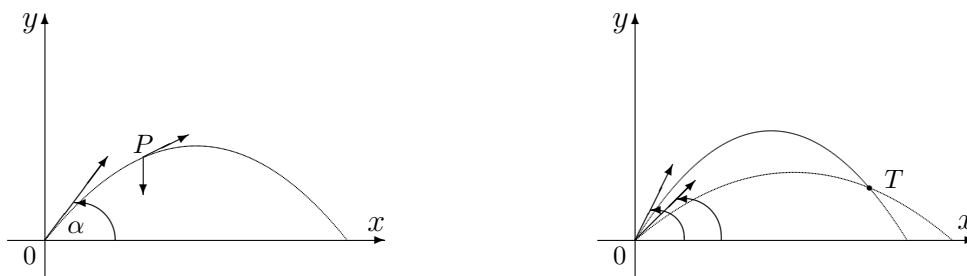
Lako se vidi da je vrijeme leta projektila (dok ne padne u točku X)

$$T = \frac{2v_0}{g} \sin \alpha, \tag{B.3}$$

uz domet

$$X = 2h \sin 2\alpha. \tag{B.4}$$

Maksimalni domet projektila postiže se za $\alpha = \frac{\pi}{4} \text{ rad}$ (45°) i iznosi $X_M = 2h$.



Slika 4.13: Kretanje balističkog projektila

Gadanje točke T

Primjer 4.17. Zadana je točka $T = (x_0, y_0)$. Treba odrediti veličinu izlaznog kuta α tako da projektil pogodi točku T (vidi Sliku 4.13b).

Kako bi projektil mogao pogoditi točku T , njene koordinate moraju zadovoljavati jednažbu

$$y_0 = x \operatorname{tg} \alpha - \frac{x_0^2}{4h \cos^2 \alpha},$$

što možemo pisati

$$y_0 \cos^2 \alpha - \frac{1}{2} x_0 \sin 2\alpha + \frac{x_0^2}{4h} = 0. \quad (*)$$

Jednažbu (*) možemo riješiti nekom od metoda navedenih u ovom poglavlju. Jednažba (*) može imati jedno ili dva rješenja (tzv. donji i gornji dijapazon kuteva). U Tablici 4.7 prikazani su dobiveni rezultati za neke karakteristične situacije (kutevi su dani u stupnjevima).

x_n	y_n	v_0	α_1	α_2	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 4.7: Gađanje točke T

4.4.2 Newtonova metoda za Lipschitz derivabilnu funkciju

Uvjet $f \in C^2[a, b]$ uz koji smo dokazali Teorem 4.3 je prejak i zahtijeva da funkcija f bude dvostruko neprekidno derivabilna na intervalu $[a, b]$. Taj uvjet može se oslabiti tako da zahtijevamo da funkcija f 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 $[a, b]$ (vidi odjeljak 4.3.1). Taj uvjet zvat ćemo *Lipschitz derivabilnost*.

Sljedeća lema daje ocjenu pogreške aproksimacije Lipschitz derivabilne funkcije ($f' \in Lip_L(I)$) afinom funkcijom $f_1(x) = f(x_0) + f'(x_0)(x - x_0)$.

Lema 4.1. *Neka je $f : I \rightarrow \mathbb{R}$ funkcija takva da je $f' \in Lip_L(I)$. Tada za proizvoljne $x_0, x \in I$ vrijedi*

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq \frac{L}{2}(x - x_0)^2. \quad (4.36)$$

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 $f' \in Lip_L(I)$ i da je $t - x_0 > 0$ za $t \in (x_0, x)$, imamo

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq L \int_{x_0}^x (t - x_0) dt = L \frac{(x - x_0)^2}{2}. \quad \square$$

Ocjena iz Leme 4.1 bliska je ocjeni iz Taylorove formule (vidi primjerice Jukić i 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 4.4. *Neka je $f: I \rightarrow \mathbb{R}$ funkcija takva da je $f' \in Lip_L(I)$. Nadalje, pretpostavimo da postoji $m_1 > 0$ tako da vrijedi $|f'(x)| \geq m_1$ za svaki $x \in I$.*

Ako postoji $\xi \in I$ takav da je $f(\xi) = 0$, onda postoji $\eta > 0$ takav da za proizvoljni $x_0 \in (\xi - \eta, \xi + \eta)$, niz generiran rekurzivnom formulom

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots \quad (4.37)$$

konvergira prema ξ kvadratnom brzinom⁶

$$|x_{n+1} - \xi| \leq \frac{L}{2m_1} (x_n - \xi)^2. \quad (4.38)$$

Dokaz. Neka je $\tau \in (0, 1)$, a $\hat{\eta}$ radijus najvećeg otvorenog simetričnog intervala oko ξ sadržanog u I i neka je $\eta = \min \left\{ \hat{\eta}, \tau \frac{2m_1}{L} \right\}$. Primijetite da to znači da $x_0 \in I$ treba birati tako da je

$$|x_0 - \xi| < \hat{\eta} \quad i \quad |x_0 - \xi| < \tau \frac{2m_1}{L},$$

⁶tj. za x_0 iz dovoljno malene okoline od ξ niz (4.37) konvergira prema ξ kvadratnom brzinom.

što može biti vrlo blizu ξ .

Najprije ćemo induktivno dokazati da vrijedi (4.38). Po pretpostavci je $|x_0 - \xi| \leq \eta$. Nadalje slijedi

$$\begin{aligned} x_1 - \xi &= x_0 - \xi - \frac{f(x_0)}{f'(x_0)} \\ &= x_0 - \xi - \frac{f(x_0) - f(\xi)}{f'(x_0)} = \frac{1}{f'(x_0)} \left(f(\xi) - f(x_0) - f'(x_0)(\xi - x_0) \right). \end{aligned}$$

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 ξ . 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

$$\lim_{n \rightarrow \infty} |x_{n+1} - \xi| = 0. \quad \square$$

4.5 Rješavanje sustava nelinearnih jednažbi

Spomenimo na kraju mogućnost generalizacije prethodno navedenih metoda na rješavanje sustava nelinearnih jednažbi

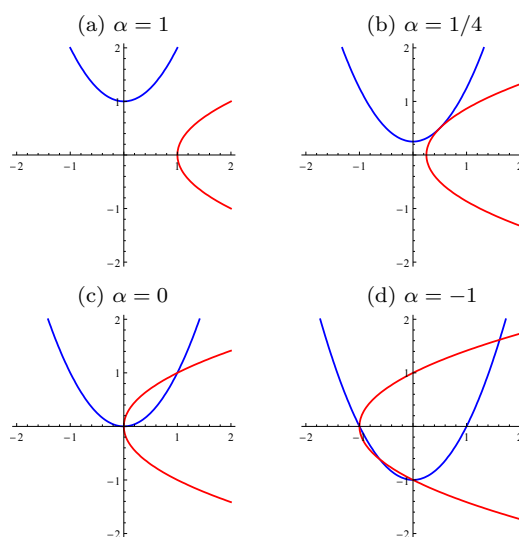
$$f_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, \dots, n, \quad (4.39)$$

odnosno u vektorskom obliku $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, 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 4.18. Promatramo sustav nelinearnih jednadžbi (Ortega i Rheinboldt, 1970) za različite vrijednosti parametra $\alpha \in [0, 1]$

$$\begin{aligned} f_1(x_1, x_2) &\equiv x_1^2 - x_2 + \alpha = 0, \\ f_2(x_1, x_2) &\equiv -x_1 + x_2^2 + \alpha = 0. \end{aligned}$$

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 4.6. Grafički ispitajte skup svih rješenja sljedećih sustava nelinearnih jednadžbi.

$$\begin{aligned} a) \quad f_1(x_1, x_2) &\equiv \frac{1}{2}x_1 \sin\left(\frac{\pi}{2}x_1\right) - x_2 = 0, & b) \quad f_1(x_1, x_2) &\equiv x_1^2 - |x_2| = 0, \\ f_2(x_1, x_2) &\equiv -x_1 + x_2^2 + 1 = 0. & f_2(x_1, x_2) &\equiv x_1^2 - x_2 = 0. \end{aligned}$$

Primjer 4.19. 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 i 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 4.20. Za dani skup točaka $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \dots, m\}$ s težinama $w_1, \dots, w_m > 0$ treba odrediti točku $c = (u, v) \in \mathbb{R}^2$ tako da suma težinskih euklidskih udaljenosti od svih točaka a^i do točke c bude minimalna. Ova točka naziva se geometrijski medijan (vidi primjerice Iyigun i Ben-Israel (2010)), a dobiva se kao rješenje sljedećeg optimizacijskog problema

$$\operatorname{argmin}_{u, v \in \mathbb{R}} F(u, v), \quad F(u, v) = \sum_{i=1}^m w_i \sqrt{(u - x_i)^2 + (v - y_i)^2}.$$

Kandidate za rješenje ovog problema (stacionarne točke funkcije F) dobit ćemo rješavanjem sljedećeg sustava nelinearnih jednadžbi

$$\frac{\partial F}{\partial u} \equiv \sum_{i=1}^m w_i \frac{u - x_i}{\sqrt{(u - x_i)^2 + (v - y_i)^2}} = 0, \quad \frac{\partial F}{\partial v} \equiv \sum_{i=1}^m w_i \frac{v - y_i}{\sqrt{(u - x_i)^2 + (v - y_i)^2}} = 0. \quad (4.40)$$

U nastavku kratko ćemo pokazati nekoliko najvažnijih metoda za rješavanje sustava nelinearnih jednadžbi i ilustrirati ih na jednostavnim primjerima.

4.5.1 Metoda jednostavnih iteracija

Analogno metodi jednostavnih iteracija (t. 4.3, str. 92) sustav (4.39) možemo zapisati u obliku

$$x_i = \varphi_i(x_1, x_2, \dots, x_n), \quad i = 1, \dots, n, \quad (4.41)$$

što upućuje na iterativni postupak metode jednostavnih iteracija (vidi primjerice Björck i Dahlquist (2007); Ortega i Rheinboldt (1970); Süli i Mayers (2006))

$$x_i^{(k+1)} = \varphi_i(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}), \quad i = 1, \dots, n, \quad k = 0, 1, \dots, \quad (4.42)$$

gdje je $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) \in \mathbb{R}^n$ početna aproksimacija. Postoje uvjeti koje moraju ispunjavati funkcije $\varphi_1, \dots, \varphi_n$ da bi iterativni proces (4.42) konvergirao prema rješenju sustava (Ortega i Rheinboldt, 1970; Süli i Mayers,

2006). Za ilustraciju, napisat ćemo sustav jednadžbi iz Primjera 4.20 u obliku (4.42)

$$u = \varphi(u, v), \quad v = \psi(u, v),$$

gdje je

$$\varphi(u, v) = \frac{\sum_{i=1}^m w_i \frac{x_i}{\sqrt{(u-x_i)^2+(v-x_i)^2}}}{\sum_{i=1}^m w_i \frac{1}{\sqrt{(u-x_i)^2+(v-x_i)^2}}}, \quad \psi(u, v) = \frac{\sum_{i=1}^m w_i \frac{y_i}{\sqrt{(u-x_i)^2+(v-x_i)^2}}}{\sum_{i=1}^m w_i \frac{1}{\sqrt{(u-x_i)^2+(v-x_i)^2}}}.$$

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 (4.40).

Ovaj iterativni proces u literaturi (vidi primjerice Iyigun i Ben-Israel (2010)) naziva se *Weiszfeldov algoritam* i ima brojne primjene (Drezner i Hamacher, 2004).

Na Slici 4.15 prikazano je 50 normalno distribuiranih podataka u okolini točke $(4, 5)$ s varijancom $\sigma^2 = 1.5$

```
data=Table[{4,5} + RandomReal[NormalDistribution[0, sig1], {2}], {i, 50}]
```

s težinama

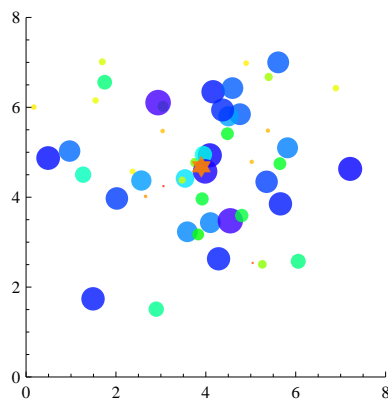
```
w=RandomReal[{0, 1}, 50];
```

Težine točaka-podataka prikazane su na slici različitim veličinama i nijansama boja. Uz izbor proizvoljne početne aproksimacije (koja se ne smije podudarati s nekom od točaka podataka!) nakon samo nekoliko iteracija Weiszfeldov algoritam daje geometrijski medijan koji je također prikazan na slici (narančasta zvjezdica).

Sustav iz Primjera 4.19 ne bi se mogao jednostavno prilagoditi za metodu jednostavnih iteracija.

Zadatak 4.7. Za dani skup točaka $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2: i = 1, \dots, m\}$ treba odrediti točku — geometrijski medijan $c \in \mathbb{R}^2$, 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.

Zadatak 4.8. Za dani skup točaka $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2: i = 1, \dots, m\}$ treba odrediti točku $c \in \mathbb{R}^2$ tako da



Slika 4.15: Geometrijski medijan

- 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?

4.5.2 Newtonova metoda

Newtonova metoda (t. 4.4, str. 103) također se može generalizirati na rješavanje sustava nelinearnih jednažbi (4.39). Najprije izaberemo početnu aproksimaciju $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \in \mathbb{R}^n$ i svaku od funkcija f_i razvijemo u Taylorov red u okolini $\mathbf{x}^{(0)}$, te odgovarajuću linearnu aproksimaciju označimo s \tilde{f}_i :

$$\tilde{f}_i(\mathbf{x}) = f_i(\mathbf{x}^{(0)}) + \sum_{j=1}^n \frac{\partial f_i(\mathbf{x}^{(0)})}{\partial x_j} (x_j - x_j^{(0)}), \quad i = 1, \dots, n. \quad (4.43)$$

Sada umjesto sustava (4.39) rješavamo sustav

$$\tilde{f}_i(\mathbf{x}) = 0, \quad i = 1, \dots, n,$$

što možemo pisati u matričnom obliku na sljedeći način

$$\mathbf{J}^{(0)} \mathbf{s} = -\mathbf{f}(\mathbf{x}^{(0)}), \quad \text{gdje je} \quad (4.44)$$

$$\mathbf{J}^{(0)} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x}^{(0)})}{\partial x_1} & \dots & \frac{\partial f_n(\mathbf{x}^{(0)})}{\partial x_n} \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} x_1 - x_1^{(0)} \\ \vdots \\ x_n - x_n^{(0)} \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}^{(0)}) = \begin{bmatrix} f_1(\mathbf{x}^{(0)}) \\ \vdots \\ f_n(\mathbf{x}^{(0)}) \end{bmatrix}.$$

Matricu \mathbf{J} nazivamo **Jacobijeva matrica** ili **Jacobijan** sustava u točki $\mathbf{x}^{(0)} \in \mathbb{R}^n$. Nova aproksimacija rješenja tada je

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)},$$

gdje je $\mathbf{s}^{(0)}$ rješenje sustava (4.44). Općenito, dobivamo iterativni postupak

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}, \quad k = 0, 1, \dots \quad (4.45)$$

gdje je $\mathbf{s}^{(k)}$ rješenje sustava

$$\mathbf{J}^{(k)} \mathbf{s} = -\mathbf{f}(\mathbf{x}^{(k)}), \quad (4.46)$$

pri čemu je $\mathbf{J}^{(k)}$ odgovarajuća Jacobijeva matrica u točki $\mathbf{x}^{(k)} \in \mathbb{R}^n$.

Primijetite da se iterativni postupak (4.45)–(4.46) uz pretpostavku regularnosti Jacobijana $\mathbf{J}^{(k)}$ može zapisati u obliku

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left(\mathbf{J}^{(k)}\right)^{-1} \mathbf{f}(\mathbf{x}^{(k)}), \quad k = 0, 1, \dots, \quad (4.47)$$

što podsjeća na Newtonov iterativni postupak (4.29), str. 104. Inače, iterativni proces (4.47) rijetko se koristi jer je postupak (4.45)–(4.46) 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]:= 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}]]
slf = 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;
```

Ponavljanjem linije In[3] dobivamo niz aproksimacija.

Zadatak 4.9. Izradite *Mathematica*-modul koji će pozivati broj varijabli n , funkciju f , točnost $\epsilon > 0$, početnu aproksimaciju x_0 , te Newtonovom metodom davati niz aproksimacija rješenja jednažbe $f(x) = 0$. Program testirajte na sljedećem primjeru.

Primjer 4.21. *Sljedeći sustav nelinearnih jednažbi riješit ćemo Newtonovom metodom.*

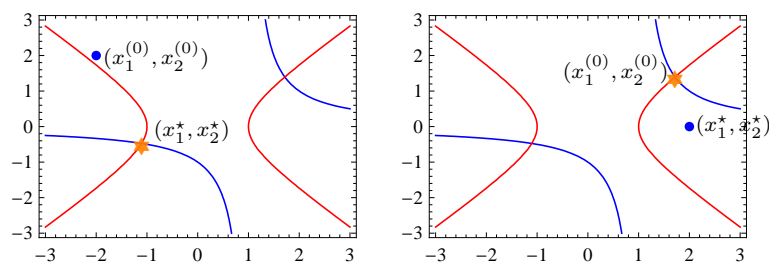
$$f_1(x_1, x_2) \equiv x_2(x_1 - 1) - 1 = 0, \quad f_2(x_1, x_2) \equiv x_1^2 - x_2^2 - 1 = 0.$$

U Tablici 4.8 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 4.8: Newtonova metoda za sustav iz Primjera 4.21

Pitanje konvergencije spomenutih metoda nećemo razmatrati. Zainteresirani čitatelj to može pogledati u više knjiga iz područja numeričke analize koje su navedene u Primjedbi 5.10, str. 151 ili u popisu literature na kraju teksta.



Slika 4.16: Početna aproksimacija i rješenje sustava

4.5.3 Kvazi-Newtonove metode

Vrlo važno mjesto među metodama za rješavanje sustava nelinearnih jednadžbi zauzimaju tzv. kvazi-Newtonove metode, koje je 1965. godine uveo C. G. Broyden⁷ (vidi primjerice Ackleh et al. (2010)), a koje su nastale kao generalizacija metode sekanti (vidi t. 4.4.1, str. 110):

(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-Powellova (DFP) metoda

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{H}_k \mathbf{f}(\mathbf{x}^k), \quad k = 0, 1, \dots,$$

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{(\mathbf{y}_k, \mathbf{s}_k)} - \frac{\mathbf{H}_k \mathbf{y}_k \mathbf{y}_k^T \mathbf{H}_k}{(\mathbf{y}_k, \mathbf{H}_k \mathbf{y}_k)}, \quad k = 0, 1, \dots$$

⁷C.G.Broyden, *A class of methods for solving nonlinear simultaneous equations*, Mathematics of Computation **19**(1965), 577–593.

(c) Broyden-Fletcher-Goldfarb-Schanoova (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 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 i Schnabel, 1996).

Zadatak 4.10. Izradite program za rješavanje sustava nelinearnih jednažbi navedenim metodama i ispitajte ga na sljedećim primjerima

(a) (Dennis i Schnabel, 1996)

$$2(x_1 + x_2)^2 + (x_1 - x_2)^2 - 8 = 0$$

$$5x_1^2 + (x_2 - 3)^2 - 9 = 0$$

Rješenje: $\mathbf{x}_1^* = (1, 1)^T$, $\mathbf{x}_2^* = (-1.18347, 1.58684)^T$.

(b) (Dennis i 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 i 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 valley*

$$\text{grad } f(x_1, x_2) = 0, \quad f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Rješenje: $\mathbf{x}^* = (1, 1)^T$.

(e) $\text{grad } f(x_1, x_2, x_3) = 0$, $f(x_1, x_2, x_3) = \sum_{i=1}^{11} \left(x_3 \exp\left(-\frac{(t_i - x_1)^2}{x_2}\right) - y_i \right)^2$,
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	0.001	0.01	0.04	0.12	0.21	0.25	0.21	0.12	0.04	0.01	0.001

Rješenje: $\mathbf{x}^* = (4.93, 2.62, 0.28)^T$.

Primjedba 4.9. Više detalja o numeričkim metodama za rješavanje nelinearnih jednačbi može se naći u brojnoj literaturi: Ackleh et al. (2010); Demidovich i Maron (1981); Dennis i Schnabel (1996); Kress (1998); Lange (1999); Ortega i Rheinboldt (1970); Ortega (1990); Plato (2003); Schwarz i Köckler (2011); Stoer (2002); Stoer i Bulirsch (2002); Süli i Mayers (2006)). FORTRAN kodovi mogu se naći u Press et al. (1992). O rješavanju sustava nelinearnih jednačbi i problemima ekstrema funkcija više varijabli može se naći kod Ackleh et al. (2010); Dennis i Schnabel (1996); Ortega i 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**.

4.6 Zadaci

Zadatak 4.11. Nekom iterativnom metodom riješite jednačbu $x^2 - 2x - 2 = 0$ 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 4.12. Uz točnost $\epsilon = 0.5 \times 10^{-6}$ odredite približnu vrijednost od $\sqrt[3]{2}$ rješavajući jednačbu $x^3 - 2 = 0$.

Rješenje: $\sqrt[3]{2} \approx 1.259921$.

Zadatak 4.13. S točnošću $\epsilon = 0.5 \cdot 10^{-4}$ riješite jednačbe

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 4.14. S točnošću $\epsilon = 0.5 \cdot 10^{-4}$ odredite sjecište grafova krivulja zadanih jednačbama $y = 1/(x - 1)$ i $x^2 - y^2 = 1$.

Rješenje: $x_0 \approx 1.7167$, $y_0 \approx 1.3953$.

Zadatak 4.15. 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 4.16. S točnošću $\epsilon = 0.5 \cdot 10^{-4}$ odredite najmanji pozitivni korijen jednačbe $x \sin x + 1 = 0$.

Rješenje: $x \approx 3.4368$.

Zadatak 4.17. Jednačba $x^3 - x = 0$ ima tri rješenja $-1, 0, 1$. Ispitajte prema kojem će rješenju jednačbe konvergirati Newtonov iterativni proces (4.29, str. 104) ako se izabere proizvoljna početna aproksimacija $x_0 \in \mathbb{R}$? Odredite intervale konvergencije.

Zadatak 4.18. Grafičkom metodom odredite intervale u kojima funkcija f ima nultočke (Björck i Dahlquist, 2007; Dahlquist i Björck, 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 4.19. 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. Provjerite 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 4.20. Funkcija $f(x) = x + \ln x$ ima jednostruku realnu nultočku ξ na intervalu $[0.1, 1]$. Koji od niže navedenih iterativnih procesa konvergira prema ξ ? Zašto?

- a) $x_{n+1} = -\ln x_n$ b) $x_{n+1} = e^{-x_n}$ c) $x_{n+1} = \frac{1}{2}(x_n + e^{-x_n})$

Zadatak 4.21. Neka je $\varphi : I = [a, b] \rightarrow \mathbb{R}$ neprekidno derivabilna funkcija, a jednažba $x = \varphi(x)$ 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 4.22. Neka funkcija $f : I \rightarrow \mathbb{R}$ ima neprekidnu drugu derivaciju na intervalu $I = [a, b]$ i neka u tom intervalu jednažba $f(x) = 0$ ima jedinstveno rješenje ξ . Neka je nadalje, $f(a) \cdot f(b) < 0$, a prva i druga 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 i druge derivacije funkcije f :

- a) kod Newtonove metode tangenti, b) kod metode sekanti,
 c) kod metode krivih položaja (regula falsi).

Zadatak 4.23. Rješenje jednadžbe $x = a/x$, $a > 0$ 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 4.24. Rekurzivnom formulom zadan je niz (formula (4.29)) 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 4.25. Slično kao u Primjeru 4.8, str. 98, definirajte i analizirajte iterativne postupke za rješavanje jednadžbe:

- a) $f(x) \equiv x^3 - a = 0$, $a \in \mathbb{R}$, (treći korijen realnog broja a),
 b) $f(x) \equiv x^k - a = 0$, $a > 0$, $k \in \mathbb{N}$, (k -ti korijen pozitivnog broja a).

Uputa: Koristite formulu (4.29), str. 104.

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 4.26. 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 4.27. Za Primjer 4.15, str. 109, odredite visine svih nivoa nafte koji odgovaraju količinama 2%, 4%, 6%, ..., 100%. Izradite odgovarajuću tablicu.

Zadatak 4.28. Ako su u Primjeru 4.15, str. 109, 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 4.29. Slično kao u Primjeru 4.16, str. 111, 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: $T \approx 8.7$ s, $X \approx 929$ 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 4.30. Ako je zadana početna brzina v_0 balističkog projektila iz Primjera 4.16, str. 111, 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 4.31. Kada će jednadžba (*) iz Primjera 4.16, str. 111, imati jedno, a kada dva rješenja?

Rješenje: Jednadžba (*) imat će jedno rješenje ako je $y_0 \leq x_0$.

Zadatak 4.32. Izradite program koji će lokalizirati nultočku neke funkcije f , a zatim primjenom metode bisekcije (str. 87) utvrditi dovoljno mali (zadane duljine 2ϵ) interval u kome se nalazi nultočka.

Zadatak 4.33. Provjerite da jednadžba: $x - 2 \sin x = 0$ ima tri realna rješenja: $\xi_1 = 0$, $\xi_1 \in [\frac{\pi}{2}, 2]$, $\xi_3 \in [-2, -\frac{\pi}{2}]$. Pokažite da za proizvoljni $x_0 \in [\frac{\pi}{2}, 2]$ niz definiran rekurzivnom formulom

$$x_{k+1} = 2 \sin x_k, \quad k = 0, 1, \dots$$

konvergira prema ξ_2 . Za koje x_0 će Newtonova metoda konvergirati prema ξ_2 ? Usporedite red konvergencije navedenih metoda. Izradite odgovarajuće programe.

Zadatak 4.34. 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 4.35. Pokažite da Newtonov iterativni postupak konvergira prema jedinstvenom rješenju jednadžbe $e^{2x} + 3x + 2 = 0$ za proizvoljni x_0 .

Zadatak 4.36. 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 4.37. Newtonov iterativni postupak (4.29) ima linearnu brzinu konvergencije ako je ξ 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 ξ kvadratnom brzinom konvergencije. Provjerite to na primjeru funkcije $f(x) = x^3 - 3x + 2$, kojoj je $\xi = 1$ dvostruka nultočka.

Zadatak 4.38. 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 pojave i nacrtajte odgovarajuće slike.

Zadatak 4.39. 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 $x^* = 0.8165$.

Zadatak 4.40. U članku R. P. Brent, *Algorithms for minimization without derivatives*, Englewood Cliffs, N. J., Prentice-Hall, 1973, predložena je metoda za traženje lokalnog minimuma neprekidne funkcije jedne realne varijable. Proces minimizacije odvija se kombinacijom metode zlatnog reza i sukcesivne paraboličke aproksimacije minimizirajuće funkcije.

Ispitajte navedeni program na nekoliko primjera, te rezultate usporedite s metodom navedenom u Zadatku 4.39. Primijetite da navedena metoda ne zahtijeva ni prvu ni drugu derivaciju minimizirajuće funkcije.

Poglavlje 5

Aproksimacija funkcija

U t. 2, str. 19, opisan je problem interpolacije funkcije čije su vrijednosti poznate na diskretnom skupu točaka. Takvu funkciju aproksimirali smo polinomom P (kao funkcijom čije se vrijednosti lako izračunavaju), tako da se funkcija i polinom podudaraju na zadanom diskretnom skupu 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, kvalitetu aproksimacije f^* mjerit ćemo „udaljenošću” funkcije f do njene aproksimacije f^*

$$d(f, f^*) = \|f - f^*\|,$$

pri čemu 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.

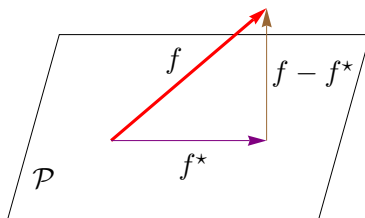
Primjedba 5.1. Primijetite da od svih navedenih normi jedino za L_2 -normu možemo definirati skalarni produkt

$$(f, g) = \int_a^b w(x) f(x) g(x) dx, \tag{5.1}$$

tako da je $\|f\|_2^2 = (f, f)$.

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

$$\|f - f^*\| \leq \|f - u\|, \quad \forall u \in \mathcal{P}.$$



Slika 5.1: Najbolja aproksimacija funkcije f na potprostoru \mathcal{P}

Geometrijski gledano, f^* treba biti ortogonalna projekcija vektora f na potprostor \mathcal{P} (vidi Sliku 5.1). To ujedno znači da je vektor $f - f^*$ okomit na \mathcal{P} , tj. da vrijedi

$$(f - f^*, u) = 0, \quad \forall u \in \mathcal{P}. \quad (5.2)$$

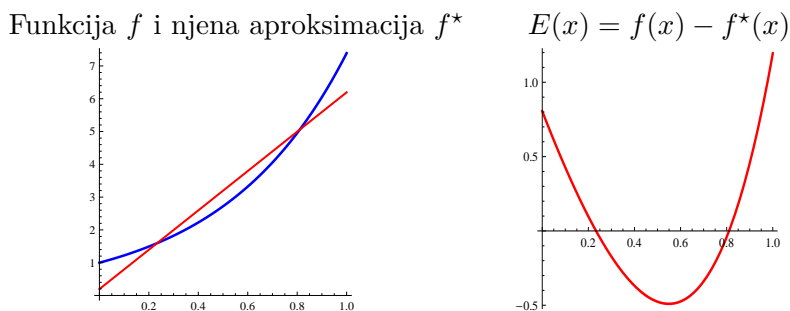
Može se pokazati da ako je \mathcal{P} konačno dimenzionalni normirani vektorski potprostor, onda za svaku funkciju $f \in C[a, b]$ postoji (barem jedna) najbolja aproksimacija $f^* \in \mathcal{P}$.

Posebno ćemo razmotriti problem najbolje L_2 i najbolje L_∞ 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. Odrediti najbolju aproksimaciju funkcije f na potprostoru $L(\varphi_0, \varphi_1, \dots, \varphi_n)$ znači pronaći $\mathbf{a}^* \in \mathbb{R}^{n+1}$, takav da je $F(\mathbf{a}^*) = \min_{\mathbf{a} \in \mathbb{R}^{n+1}} F(\mathbf{a})$, gdje je

$$F(\mathbf{a}) = \left\| \sum_{i=0}^n a_i \varphi_i - f \right\|^2 = \int_a^b w(x) \left(\sum_{i=0}^n a_i \varphi_i(x) - f(x) \right)^2 dx. \quad (5.3)$$

Slika 5.2: Najbolja aproksimacija funkcije $f(x) = e^{2x}$

a vektor slobodnih koeficijenata $\left(\frac{1}{2}(e^2 - 1), \frac{1}{4}(e^2 + 1)\right)^T = (3.19, 2.097)^T$. Rješenje sustava (5.4) je $a_0^* = 0.195$, $a_1^* = 6.0$.

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 2×10^7 . Zato je traženje najbolje L_2 aproksimacije rješavanjem sustava (5.4) često puta nepouzđano.

Primjedba 5.4. Do sustava normalnih jednadžbi (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 φ_j za $j = 0, 1, \dots, n$, dobivamo sustav (5.4)

$$0 = (f, \varphi_j) - \sum_{i=1}^m a_i (\varphi_i, \varphi_j), \quad j = 0, 1, \dots, n.$$

Ako bi funkcije $\varphi_0, \varphi_1, \dots, \varphi_n$ činile ortogonalan sustav funkcija, tj. ako bi bilo

$$(\varphi_i, \varphi_j) = 0, \quad i \neq j \quad \text{i} \quad \|\varphi_i\| \neq 0 \quad \text{za sve } i = 0, 1, \dots, n,$$

onda bi matrica sustava (5.4) bila dijagonalna, a rješenje bismo mogli eksplicitno zapisati

$$a_i^* = \frac{(f, \varphi_i)}{(\varphi_i, \varphi_i)}, \quad i = 0, 1, \dots, n. \quad (5.6)$$

¹Jørgen Pedersen Gram (1850.–1916.) danski matematičar

Jedan takav ortogonalni sustav funkcija razmatra se u sljedećoj točki.

5.1.1 Fourierovi polinomi

Kao primjer ortogonalnog sustava na $C[-\pi, \pi]$ uz težinsku funkciju $w(x) = 1$ navodimo sustav trigonometrijskih funkcija

$$\frac{1}{2}, \sin x, \cos x, \dots, \sin nx, \cos nx. \quad (5.7)$$

Lema 5.1. *Sustav trigonometrijskih funkcija (5.7) 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 (5.7) bude beskonačan.

Potražimo najbolju L_2 aproksimaciju funkcije $f \in C[-\pi, \pi]$ na potprostoru s bazom (5.7). Aproksimaciju f^* funkcije f tražit ćemo u obliku trigonometrijskog polinoma stupnja n

$$T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx), \quad (5.8)$$

pri čemu parametre $\mathbf{a} = (a_0, a_1, \dots, a_n)^T \in \mathbb{R}^{n+1}$, $\mathbf{b} = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ određujemo rješavanjem sljedećeg optimizacijskog problema

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^{n+1}, \mathbf{b} \in \mathbb{R}^n} F(\mathbf{a}, \mathbf{b}), \quad F(\mathbf{a}, \mathbf{b}) = \int_{-\pi}^{\pi} (T_n(x; \mathbf{a}, \mathbf{b}) - f(x))^2 dx. \quad (5.9)$$

Izjednačavajući sve parcijalne derivacije funkcije F s nulom ili postupajući kao u Primjedbii 5.1 i koristeći Lemu 5.1 lako dobivamo

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, & k &= 0, 1, \dots, n, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx, & k &= 1, \dots, n, \end{aligned} \quad (5.10)$$

Parametre a_k, b_k iz (5.10) zovemo **Fourierovi koeficijenti**² funkcije f , a trigonometrijski polinom T_n s Fourierovim koeficijentima (5.10) zovemo **Fourierov polinom** za funkciju f i označavamo ga s F_n (vidi primjerice Kurepa (1990); Lange (1999); Plato (2003); Schatzman (2002)).

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

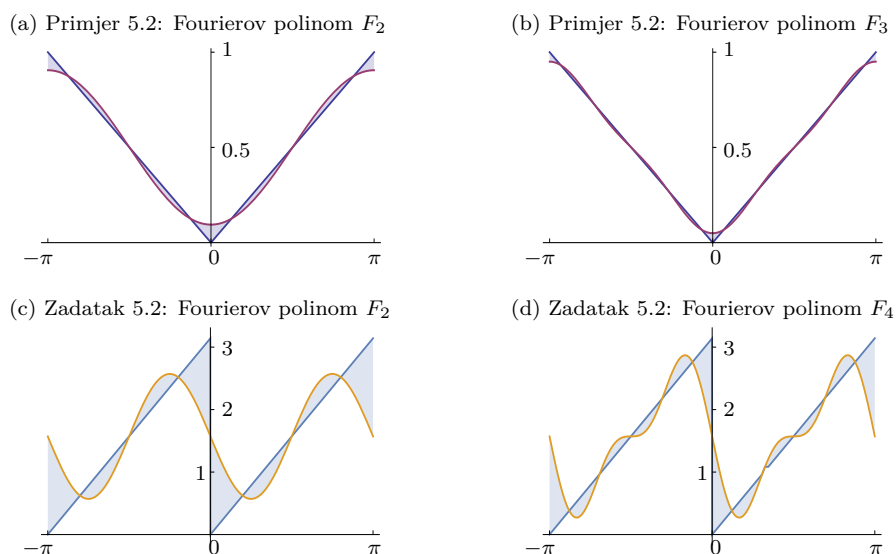
$$a_0 = 1, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx dx = \begin{cases} 0, & k \text{ paran} \\ -\frac{4}{k^2\pi^2} & k \text{ neparan} \end{cases}$$

Zato odgovarajući Fourierov polinom glasi (vidi Sliku 5.3a-b)

$$F_n(x) = \frac{1}{2} - \frac{4}{\pi^2} \left(\cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \dots + \frac{1}{n^2} \cos nx \right).$$

Zadatak 5.2. Zadana je funkciju $f: [-\pi, \pi] \rightarrow \mathbb{R}$, $f(x) = \begin{cases} x + \pi, & x \in [-\pi, 0) \\ x, & x \in [0, \pi] \end{cases}$. Ispitajte njenu parnost i korištenjem niže navedenog **Mathematica**-programa odredite Fourierove koeficijente i Fourierov polinom stupnja $n \geq 1$. Na Slici 5.3c-d prikazan je graf funkcije f i Fourierovi polinomi F_2 i F_4 .

²Jean Baptiste Fourier (1768.-1830.) francuski fizičar i matematičar.



Slika 5.3: Fourierovi polinomi

```

In[1] := n = 4;
f[x_] := If[x < 0, x + Pi, x]
a0 = (1/Pi) Integrate[f[x], {x, -Pi, Pi}];
a = Table[(1/Pi) Integrate[f[x] Cos[k x], {x, -Pi, Pi}], {k, n}];
b = Table[(1/Pi) Integrate[f[x] Sin[k x], {x, -Pi, Pi}], {k, n}];
Print["a0=", a0, " a=", MatrixForm[a], " b=", MatrixForm[b]]
F[x_] := a0/2 + Sum[a[[k]] Cos[k x] + b[[k]] Sin[k x], {k, n}]
Plot[{f[x], F[x]}, {x, -Pi, Pi}, Filling -> {1 -> {2}}]

```

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

$$\tilde{F}_n(x) = \frac{1}{2} - \frac{4}{\pi^2} \left(\cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \cdots + \frac{1}{n^2} \cos nx \right),$$

a Fourierov polinom funkcije f tada je

$$\begin{aligned} F_n(x) &= (\tilde{F}_n \circ u^{-1})(x) = (\tilde{F}_n(\pi x)) \\ &= \frac{1}{2} - \frac{4}{\pi^2} \left(\cos \pi x + \frac{1}{3^2} \cos 3\pi x + \frac{1}{5^2} \cos 5\pi x + \dots \right) + \frac{1}{n^2} \cos n\pi x. \end{aligned}$$

Zadatak 5.3. 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 5.4. 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 5.5. Pokažite da sustav funkcija $\varphi_i(x) = \cos ix$, $i = 0, 1, \dots, n$, definiran na $D = \{x_i = \frac{2i+1}{n+1} \frac{\pi}{2} : i = 0, 1, \dots, n\} \subset [0, \pi]$ 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 5.5. Ako su f, g ortogonalne funkcije, onda za njih vrijedi Pitagorin teorem

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2. \quad (5.11)$$

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

$$\|f - f^*\|^2 = \|f\|^2 - 2(f, f^*) + \|f^*\|^2 = \|f\|^2 - \|f^*\|^2 = \|f\|^2 - \sum_{i=0}^n (a_i^*)^2 \|\varphi_i\|^2. \quad (5.12)$$

Iz (5.12) izravno slijedi poznata Besselova nejednakost

$$\sum_{i=0}^n (a_i^*)^2 \|\varphi_i\|^2 \leq \|f\|^2. \quad (5.13)$$

Primjedba 5.6. Budući da je sustav funkcija iz (5.7), str. 133, 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

$$F(x) := \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx, \quad (5.14)$$

gdje je prema (5.6)

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx. \quad (5.15)$$

Red (5.14) nazivamo **Fourierov red**, a brojeve (5.15) **Fourierovi koeficijenti** funkcije f . Kada će red (5.14) za neki x_0 biti konvergentan i hoće li u tom slučaju biti $\mathcal{F}(x_0) = f(x_0)$ određeno je *Dirichletovim teoremom* (Kurepa, 1990).

Teorem 5.1. (Dirichlet) *Neka je $f : [-\pi, \pi] \rightarrow \mathbb{R}$ periodična funkcija temeljnog perioda 2π , takva da*

- (i) *na segmentu $[-\pi, \pi]$ ima najviše konačno mnogo prekida prve vrste,*
- (ii) *na segmenta $[-\pi, \pi]$ najviše konačno mnogo 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 $f \in C^2[-\pi, \pi]$ (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 majorantu³

$$\frac{|a_0|}{2} + 2M \left(1 + 1 + \frac{1}{2^2} + \frac{1}{2^2} + \dots \right) = \frac{|a_0|}{2} + 4M \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots \right).$$

Primjedba 5.7. Fourierov red (5.14) može se prikazati i u kompleksnom obliku:

$$F(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}, \quad \alpha_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \quad (5.16)$$

Zadatak 5.6. Pokažite da između Fourierovih koeficijenata $\{a_k, b_k\}$ iz razvoja (5.14) i Fourierovih koeficijenata α_k iz razvoja (5.16) postoji sljedeća veza

$$a_0 = 2\alpha_0, \quad a_k \cos kx + b_k \sin kx = \alpha_k e^{ikx} + \alpha_{-k} e^{-ikx}, \quad k = 1, 2, \dots$$

³Hiperharmonijski red $\sum \frac{1}{k^s}$, $s > 1$, vidi primjerice Jukić i Scitovski (2004)

Zadatak 5.7. Neka je $f : \mathbb{R} \rightarrow \mathbb{R}$ neprekidna periodična funkcija s temeljnim periodom $T > 0$.

1. Pokažite da sustav funkcija

$$\frac{1}{2}, \sin \frac{2\pi x}{T}, \cos \frac{2\pi x}{T}, \dots, \sin \frac{2n\pi x}{T}, \cos \frac{2n\pi x}{T} \quad (5.17)$$

čini ortogonalan sustav funkcija na $C[0, T]$ uz $w(x) = 1$.

2. Od sustava (5.17) načinite ortonormiran sustav funkcija na $C[0, T]$. Odredite najbolju L_2 aproksimaciju funkcije $f \in C[0, T]$ 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

$$F_n(x) = \sum_{k=-n}^n c_k \exp\left(\frac{2k\pi ix}{T}\right) \quad (5.18)$$

Uspostavite vezu između Fourierovih koeficijenata $a_0, a_1, \dots, a_n, b_1, \dots, b_n$ i Fourierovih koeficijenata c_k u prikazu (5.18).

Zadatak 5.8. Odredite Fourierove redove za sljedeće funkcije

$$a) \quad f(x) = \begin{cases} 1, & -\frac{\pi}{2} < x < \frac{\pi}{2} \\ -1, & \frac{\pi}{2} < x < \frac{3\pi}{2} \end{cases} \quad b) \quad f(x) = \begin{cases} -1, & -\pi < x < 0 \\ 1, & 0 < x < \pi \end{cases}$$

$$c) \quad f(x) = \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ -1, & \frac{1}{2} < x < \frac{3}{2} \end{cases} \quad d) \quad f(x) = \begin{cases} 0, & -1 < x < 1 \\ 1, & 1 < x < 3 \\ \frac{1}{2}, & x \in \{1, 3\} \end{cases}$$

$$e) \quad f(x) = \begin{cases} x, & -1 < x < 1 \\ 0, & x = 1 \end{cases} \quad f) \quad f(x) = \begin{cases} x, & -1 < x < 0 \\ 1-x, & 0 < x < 1 \\ \frac{1}{2}, & x = 0 \\ -\frac{1}{2}, & x = 1 \end{cases}$$

Rješenje:

$$a) \quad \frac{4}{\pi} \left(\cos x - \frac{\cos 3x}{3} + \frac{\cos 5x}{5} - \dots \right) \quad b) \quad \frac{4}{\pi} \left(\sin x - \frac{\sin 3x}{3} + \frac{\sin 5x}{5} - \frac{\sin 7x}{7} + \dots \right)$$

$$c) \quad \frac{4}{\pi} \left(\cos \pi x - \frac{\cos 3\pi x}{3} + \frac{\cos 5\pi x}{5} - \dots \right) \quad d) \quad \frac{4}{\pi\sqrt{2}} \left(\sin \frac{\pi x}{4} - \frac{1}{3} \sin \frac{3\pi x}{4} - \frac{1}{5} \sin \frac{5\pi x}{4} + \dots \right)$$

$$e) \quad \frac{2}{\pi} \left(\frac{\sin \pi x}{1} - \frac{\sin 2\pi x}{2} + \frac{\sin 3\pi x}{3} - \dots \right)$$

$$f) \quad a_k = \frac{4}{k^2\pi^2}, \quad k = 1, 3, 5, \dots, \quad b_k = \begin{cases} \frac{2}{k\pi}, & k = 1, 3, 5, \dots \\ 0, & k = 2, 4, 6, \dots \end{cases}$$

Primjer 5.4. Iz Eulerovog identiteta $e^{ix} = \cos x + i \sin x$ i binomne formule slijedi:

$$\begin{aligned}\cos nx &= \cos^n x - \binom{n}{2} \cos^{n-2} x \sin^2 x + \binom{n}{4} \cos^{n-4} x \sin^4 x - \dots, \\ \sin nx &= \binom{n}{1} \sin x \cos^{n-1} x - \binom{n}{3} \sin^3 x \cos^{n-3} x + \dots.\end{aligned}$$

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 $\cos^n x$, $\sin^n x$, $\cos^m x \sin^n x$. 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 5.9. Odredite Fourierov red za funkciju $\sin^n x$ posebno za parni, a posebno za neparni n .

Zadatak 5.10. Pokažite da vrijedi:

$$\cos^3 x \sin^3 x = \frac{3}{32} \sin 2x - \frac{1}{32} \sin 6x.$$

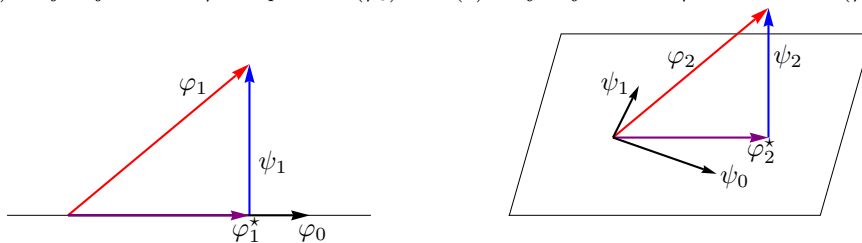
5.1.2 Ortogonalni polinomi

Pretpostavimo da je na vektorskom prostoru $C[a, b]$ definiran skalarni produkt (5.1) s težinskom funkcijom $w: [a, b] \rightarrow \mathbb{R}_+$. 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).

Stavimo $\psi_0 := \varphi_0$. Neka je nadalje φ_1^* ortogonalna projekcija vektora φ_1 u smjeru vektora φ_0 i neka je $\psi_1 = \varphi_1 - \varphi_1^*$. Tada postoji $\alpha \in \mathbb{R}$ takav da je (vidi Sliku 5.4a) $\varphi_1^* = \alpha \varphi_0$, tj.

$$\psi_1 = \varphi_1 - \alpha \varphi_0.$$

- (a) Projekcija vektora φ_1 na pravac $L(\varphi_0)$ (b) Projekcija vektora φ_2 na ravninu $L(\varphi_0, \varphi_1)$



Slika 5.4: Gram-Schmidtov postupak ortogonalizacije

Množeći skalarno ovu jednakost s vektorom φ_0 dobivamo

$$0 = (\psi_1, \varphi_0) = (\varphi_1, \varphi_0) - \alpha(\varphi_0, \varphi_0),$$

iz čega dobivamo

$$\alpha = \frac{(\varphi_1, \varphi_0)}{(\varphi_0, \varphi_0)}, \quad \text{i} \quad \psi_1 = \varphi_1 - \frac{(\varphi_1, \varphi_0)}{(\varphi_0, \varphi_0)}\varphi_0. \quad (5.19)$$

Nadalje, neka je φ_2^* ortogonalna projekcija vektora φ_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 5.4b) $\varphi_2^* = \alpha\psi_0 + \beta\psi_1$ i

$$\psi_2 = \varphi_2 - \alpha\psi_0 - \beta\psi_1.$$

Množeći skalarno redom ovu jednakost s ψ_0, ψ_1 dobivamo

$$\begin{aligned} 0 &= (\psi_2, \psi_0) = (\varphi_2, \psi_0) - \alpha(\psi_0, \psi_0), \\ 0 &= (\psi_2, \psi_1) = (\varphi_2, \psi_1) - \beta(\psi_1, \psi_1), \end{aligned}$$

iz čega dobivamo

$$\alpha = \frac{(\varphi_2, \psi_0)}{(\psi_0, \psi_0)}, \quad \beta = \frac{(\varphi_2, \psi_1)}{(\psi_1, \psi_1)}, \quad \text{i} \quad \psi_2 = \varphi_2 - \frac{(\varphi_2, \psi_0)}{(\psi_0, \psi_0)}\psi_0 - \frac{(\varphi_2, \psi_1)}{(\psi_1, \psi_1)}\psi_1. \quad (5.20)$$

Ponavljajući postupak dobivamo

$$\begin{aligned} \psi_0 &= \varphi_0, \\ \psi_1 &= \varphi_1 - \frac{(\varphi_1, \psi_0)}{(\psi_0, \psi_0)}\psi_0, \\ &\vdots \\ \psi_n &= \varphi_n - \sum_{j=0}^{n-1} \frac{(\varphi_n, \psi_j)}{(\psi_j, \psi_j)}\psi_j. \end{aligned} \quad (5.21)$$

Budući da je ψ_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 5.8. 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

$$f^* = \sum_{k=0}^n \alpha_k \psi_k \quad \text{i} \quad f' = f - f^* = f - \sum_{k=0}^n \alpha_k \psi_k.$$

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

$$\alpha_i = \frac{(f, \psi_i)}{(\psi_i, \psi_i)}, \quad \text{odnosno} \quad f^* = \sum_{k=0}^n \frac{(f, \psi_k)}{(\psi_k, \psi_k)} \psi_k. \quad (5.22)$$

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, ψ_k) Fourierovi koeficijenti.

Niže je naveden algoritam *Gram-Schmidtov postupak ortogonalizacije* linearno nezavisnih vektora (funkcija) $\varphi_0, \varphi_1, \dots, \varphi_n$, uz težinsku funkciju w , koji je implementiran u jednostavnom `Mathematica`-modulu `GSP[w_, a_, b_, n_]` u poglavlju 9.5.1, str.232.

Algoritam 12 (Gram-Schmidtov postupak ortogonalizacije)

Input: $w: [a, b] \rightarrow \mathbb{R}; \{w(t) > 0, \forall t \in [a, b]\};$

Input: $a, b, n;$

1: Staviti $\varphi_i(x) = x^i, i = 0, 1, \dots, n$

2: $\psi_0 = \varphi_0;$

3: **for** $i = 1$ to n **do**

4: $\psi_i = \varphi_i - \sum_{j=0}^{i-1} \frac{(\varphi_i, \psi_j)}{(\psi_j, \psi_j)} \psi_j$

5: **end for**

Output: $\{\psi_0, \psi_1, \dots, \psi_n\}$

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 primjenom Gram-Schmidtovog postupka ortogonalizacije.

Traženi ortogonalni sustav dobit ćemo primjenom Algoritma 12 i odgovarajućeg Mathematica-modul `GSP[w_, a_, b_, n_]`. Najprije definiramo težinsku funkciju w , interval $[a, b]$ i broj n ortogonalnih polinoma.

```
In[1] := w[x_] := 1;
        a=-1; b=1.; n=5;
```

Nakon toga, pokretanjem modula `GSP[w_, a_, b_, n_]` dobivamo prvih 6 ortogonalnih polinoma

```
In[2] := GSP[w, a, b, n]
```

$$\begin{aligned} \psi_0(x) &= 1, & \psi_1(x) &= x, & \psi_2(x) &= x^2 - \frac{1}{3}, \\ \psi_3(x) &= x^3 - \frac{3}{5}x, & \psi_4(x) &= x^4 - \frac{6}{7}x^2 + \frac{3}{35}, & \psi_5(x) &= x^5 - \frac{10}{9}x^3 + \frac{5}{21}x. \end{aligned} \quad (5.23)$$

Polinomi (5.23) se do na konstantu podudaraju s poznatim Legendreovim polinomima, koji se obično definiraju kao

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n \in \mathbb{N}. \quad (5.24)$$

Zadatak 5.11. Pokažite da za Legendreove polinome (5.24) 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 za funkciju $\varphi(x) = x^2 - 1$ vrijedi $\varphi(-1) = \varphi(1) = 0$ i koristite Rolleov teorem.

c) $P_0(x) = 1$, $P_1(x) = x$, a za $n = 1, 2, \dots$ vrijedi rekurzivna formula:

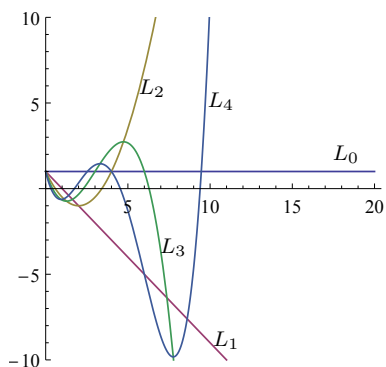
$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x).$$

- d) Primjenom *Mathematica*- naredbe `LegendreP[n,x]` odredite nekoliko Legendreovih polinoma i nacrtajte njihove grafove.

Zadatak 5.12. Koristeći Legendreove polinome odredite najbolju L_2 aproksimaciju funkcije $f(x) = e^{2x}$, $x \in [0, 1]$, na potprostoru polinoma stupnja ≤ 1 . Rezultate usporedite s Primjerom 5.1.

Uputa: Koristeći transformaciju (5.25) 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 (5.6).

Zadatak 5.13. Prethodno opisan Gram-Schmidtov postupak i pripadni Algoritam 12 daje sustav ortogonalnih polinoma, pri čemu svi dobiveni ortogonalni polinomi nemaju jednaku normu. Pokušajte modificirati algoritam tako da svi dobiveni ortogonalni polinomi imaju jednaku normu.



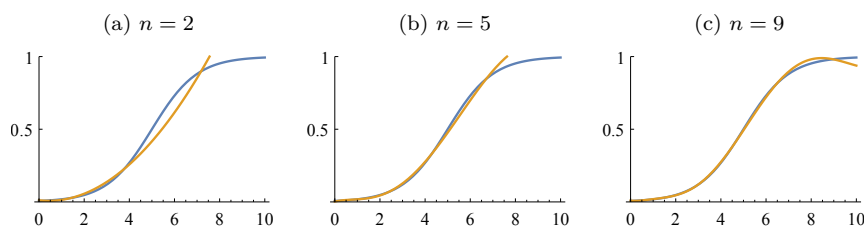
Slika 5.5: Grafovi Laguerrovih polinoma

Primjer 5.7. Sustav ortogonalnih polinoma na unitarnom vektorskom prostoru $C[0, +\infty)$ sa standardnim skalarnim produktom i težinskom funkcijom $w(x) = e^{-x}$ čine poznati Laguerreovi⁴ polinomi. Niže je navedeno prvih pet Laguerrovih polinoma, a njihovi grafovi prikazani su na Slici 5.5. Neka njihova svojstva mogu se vidjeti kod Corn i Scitovski (2014).

$$\begin{aligned} L_0(x) &= 1, & L_1(x) &= -x + 1, & L_3(x) &= \frac{1}{3!}(-x^3 + 9x^2 - 18x + 6), \\ L_2(x) &= \frac{1}{2!}(x^2 - 4x + 2) & L_4(x) &= \frac{1}{4!}(x^4 - 16x^3 + 72x^2 - 96x + 24). \end{aligned}$$

⁴Edmond Laguerre (1834.-1886.), francuski matematičar.

Zadatak 5.14. Nacrtajte graf funkcije $f: [0, +\infty) \rightarrow \mathbb{R}$, $f(x) = \frac{1}{1+e^{-x}}$. Nadite najbolje aproksimacije ove funkcije pomoću Laguerreovih polinoma stupnja $n \leq 10$. Grafovi ovih aproksimacija za $n = 2, 5, 9$ prikazani su na Slici 5.6. Nacrtajte ostale.



Slika 5.6: Laguerrove aproksimacije funkcije $f(x) = \frac{1}{1+e^{-x}}$

Primjedba 5.9. Uz pretpostavku da je poznat ortogonalan sustav $\psi_0, \psi_1, \dots, \psi_n$ na intervalu $[-1, 1]$, jednostavnom transformacijom

$$\gamma: [a, b] \rightarrow [-1, 1], \quad \gamma(x) = \frac{2x}{b-a} - \frac{a+b}{b-a}, \quad (5.25)$$

možemo dobiti sustav funkcija $\tilde{\psi}_0, \tilde{\psi}_1, \dots, \tilde{\psi}_n$, $\tilde{\psi}_i = \psi_i \circ \gamma$, $\tilde{\psi}_i: [a, b] \rightarrow \mathbb{R}$, koji je ortogonalan na intervalu $[a, b]$.

5.1.3 Čebiševljevi polinomi

Neka je (\cdot, \cdot) skalarni produkt definiran na $C[-1, 1]$ uz težinsku funkciju $w: [-1, 1] \rightarrow \mathbb{R}_+$, $w(x) = \frac{1}{\sqrt{1-x^2}}$. Sustav ortogonalnih funkcija dobiven Gram-Schmidtovim postupkom ortogonalizacije funkcija $\varphi_i(x) = x^i$, $i = 0, 1, \dots$, nazivamo **Čebiševljevi polinomi**⁵ (vidi primjerice Björck i Dahlquist (2007); Stewart (1998); Plato (2003, 2010)).

Ovi polinomi mogu se također definirati pomoću eksplicitne formule

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, \dots \quad (5.26)$$

Korištenjem jednostavne relacije

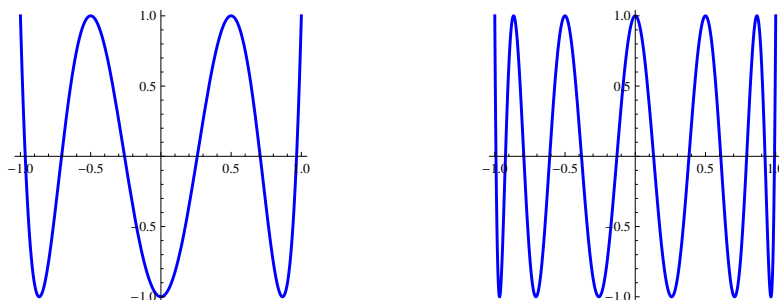
$$\cos(n+1)\varphi + \cos(n-1)\varphi = 2 \cos \varphi \cos n\varphi, \quad n \in \mathbb{N},$$

⁵Oznaka „T” dolazi od njemačke transkripcije prezimena Čebišev: *Tschebyscheff*.

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 π) odakle neposredno slijedi ortogonalnost Čebiševljevih polinoma. Čebiševljevi polinomi u programskom sustavu *Mathematica* pozivaju se s `ChebyshevT[n, x]`.



Slika 5.7: Čebiševljevi polinomi T_6 i T_{12}

Čebiševljevi polinomi imaju sljedeća svojstva:

- $|T_n(x)| \leq 1$, $x \in [-1, 1]$, $n = 0, 1, \dots$;
- Čebiševljev polinom T_n na intervalu $[-1, 1]$ ima n različitih nultočaka

$$x_k = \cos\left(\frac{2k-1}{n} \frac{\pi}{2}\right), \quad k = 1, \dots, n;$$

- 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;

- $T_0(x) = 1$, $T_1(x) = x$, 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 5.15. 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 5.16. (Björck i Dahlquist, 2007) Odredite najbolju aproksimaciju funkcije $x \mapsto \cos x$ na potprostoru razapetom Čebiševljevim polinomima stupnja $\leq n$, za $n = 2, 4, 6$. Kolika je pogreška aproksimacije u L_∞ -normi?

Zadatak 5.17. 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, ξ_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_∞ 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 $w(x) = 1$ (vidi primjerice Stewart (1998)). Treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{(a_0, \dots, a_n) \in \mathbb{R}^{n+1}} F(a_0, \dots, a_n), \quad \text{gdje je} \quad (5.27)$$

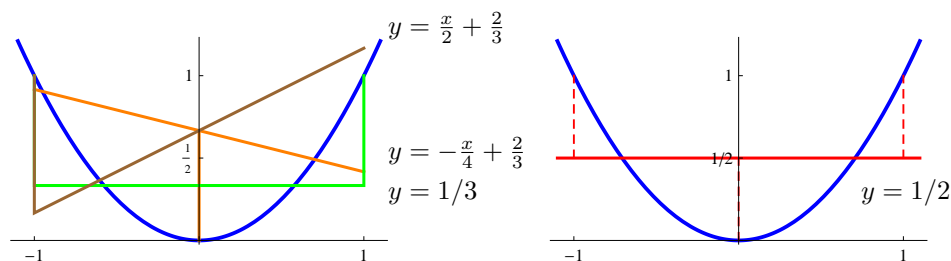
$$F(a_0, \dots, a_n) = \left\| \sum_{i=0}^n a_i \varphi_i - f \right\|_\infty = \max_{a \leq x \leq b} \left| \sum_{i=0}^n a_i \varphi_i(x) - f(x) \right|.$$

Ako je $\mathbf{a}^* = (a_0^*, \dots, a_n^*)^T \in \mathbb{R}^{n+1}$ rješenje problema (5.27), onda je funkcija $f^* = \sum_{i=0}^n a_i^* \varphi_i$ najbolja L_∞ 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_∞ 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 5.8. Za danu funkciju $f(x) = x^2$, $x \in [-1, 1]$, treba pronaći najbolju L_∞ aproksimaciju na prostoru polinoma stupnja ≤ 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 5.8: Linearna L_∞ aproksimacija kvadratne funkcije $f(x) = x^2$

Ako s \mathcal{P}_1 označimo skup svih polinoma stupnja ≤ 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_∞ aproksimacija kvadratne funkcije $f(x) = x^2$ na potprostoru svih polinoma prvog stupnja, funkcija $f^*(x) = \frac{1}{2}$. Primijetite (vidi Sliku 5.8) da se 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_∞ aproksimaciju na prostoru \mathcal{P}_{n-1} polinoma stupnja $\leq n - 1$, tj. treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{P \in \mathcal{P}_{n-1}} \|x^n - P(x)\|_\infty. \quad (5.28)$$

Rješenje daje sljedeći teorem.

Teorem 5.2. *Između svih normiranih polinoma⁶ stupnja $\leq n$ polinom $2^{1-n}T_n$, gdje je T_n n -ti Čebiševljev polinom, ima najmanju L_∞ normu na intervalu $[-1, 1]$, koja iznosi 2^{1-n} .*

⁶Kažemo da je polinom normiran ako mu je koeficijent uz najvišu potenciju jednak 1.

Drugim riječima,

$$\operatorname{argmin}_{P \in \mathcal{P}_n} \|P\|_\infty = 2^{1-n} T_n \quad \text{i} \quad \min_{P \in \mathcal{P}_n} \|P\|_\infty = \|2^{1-n} T_n\|_\infty = 2^{1-n}.$$

Zato rješenje optimizacijskog problema (5.28) 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 5.8 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.

Dokaz. Uočimo najprije da je L_∞ norma Čebiševljevog polinoma T_n jednaka $\|T_n\|_\infty = 1$ (svojstvo a), str. 146). Osim toga, njegov koeficijent uz najvišu potenciju jednak je 2^{n-1} (svojstvo d), str. 146), 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

$$-2^{1-n} \stackrel{(1)}{<} p_n(x) \stackrel{(2)}{<} 2^{1-n}, \quad \forall x \in [-1, 1]. \quad (5.29)$$

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. 146) 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 „polinoma pogreške” $E_n(x) := p_n(x) - 2^{1-n} T_n(x)$. U tu svrhu iz nejednakosti (5.29) za $x = \xi_0, \xi_1, \dots, \xi_n$ računamo:

$$\begin{aligned}
x = \xi_0 : p_n(\xi_0) &\stackrel{(2)}{<} 2^{1-n} \quad T_n(\xi_0) \stackrel{=}{=} 1 & 2^{1-n}T_n(\xi_0) &\implies E_n(\xi_0) < 0 \\
x = \xi_1 : p_n(\xi_1) &\stackrel{(1)}{>} 2^{1-n} \quad T_n(\xi_1) \stackrel{=}{=} -1 & 2^{1-n}T_n(\xi_1) &\implies E_n(\xi_1) > 0 \\
x = \xi_2 : p_n(\xi_2) &\stackrel{(2)}{<} 2^{1-n} \quad T_n(\xi_2) \stackrel{=}{=} 1 & 2^{1-n}T_n(\xi_2) &\implies E_n(\xi_2) < 0 \\
&\dots
\end{aligned}$$

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.31). Iz Teorema 2.1, str. 29, 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

$$\|f - P_n\|_\infty \leq \frac{M_{n+1}}{(n+1)!} \|\omega\|_\infty, \quad (5.30)$$

gdje je ω normirani polinom stupnja $n + 1$. Prema Teoremu 5.2, desna strana u (5.30) bit će najmanja (najbolja ocjena!) ako umjesto polinoma ω 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 (5.30) prelazi u

$$\|f - P_n\|_\infty \leq \frac{M_{n+1}}{2^n(n+1)!}, \quad (5.31)$$

a kao čvorove interpolacije moraju se uzeti nultočke Čebiševljevog polinoma T_{n+1} proširenog na interval $[a, b]$.

Zadatak 5.18. Odredite najbolju L_∞ 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 generalizira sljedeći važni teorem⁷:

⁷*Chebyshev's Equioscillation Theorem* (vidi primjerice Ackleh et al. (2010); Cheney (1998); Björck i Dahlquist (2008))

Teorem 5.3. *Neka je $f \in C[a, b]$. Polinom p^* stupnja $\leq n$ je najbolja L_∞ aproksimacija funkcije f na $[a, b]$ onda i samo onda ako postoje barem $n + 2$ točke,*

$$a \leq x_0 < x_1 < \cdots < x_{n+1} \leq b$$

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 5.19. Primjenom Teorema 5.3 odredite najbolju L_∞ aproksimacija funkcije $f(x) = \sin x$ na $[0, \frac{\pi}{2}]$ u klasi polinoma prvog stupnja.

Rješenje: $p_1^*(x) = 0.105 + \frac{2x}{\pi}$.

Primjedba 5.10. Više detalja o numeričkim metodama o aproksimaciji funkcija može se naći u brojnoj literaturi (vidi primjerice Ackleh et al. (2010); Plato (2003); Schatzman (2002); Schwarz i Köckler (2011); Stewart (1998)). FORTRAN kodovi mogu se naći u Press et al. (1992). Kvalitetna FORTRAN programska podrška nalazi se u **The NAG-Library**. Naravno, kvalitetan inženjerski „alat” za rješavanje ovakvih problema su programski sustavi **Mathematica** i **Matlab**.

Poglavlje 6

Problemi najmanjih kvadrata

Problem najbolje l_2 -aproximacije funkcije koja je zadana na konačnom skupu točaka obično se u literaturi naziva diskretni problem najmanjih kvadrata¹ (usporedite Primjedb 5.2 na str. 131).

Najprije ćemo na nekoliko tipičnih primjera pokazati neke probleme iz primjena, koji se rješavaju metodom najmanjih kvadrata. Još neke 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 reprezentant α^* veličine A tako da sve izmjerene vrijednosti y_1, y_2, \dots, y_m budu „što bliže” reprezentantu α^* .*



Slika 6.1: Aritmetička sredina podataka

¹Princip 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 objavio algebarski postupak metode najmanjih kvadrata.

Pri tome pojam „što bliže” shvaćamo u smislu zahtjeva da suma kvadrata odstupanja svih mjerenja od nekog α bude minimalna, tj. rješavamo sljedeći optimizacijski problem

$$\alpha^* = \operatorname{argmin}_{\alpha \in \mathbb{R}} F_{LS}(\alpha), \quad F_{LS}(\alpha) = \sum_{i=1}^m (y_i - \alpha)^2. \quad (6.1)$$

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 α^* 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 do mjerenja y_1, \dots, y_m najmanja.

Zadatak 6.1. Neka su y_1, \dots, y_m , $m \geq 1$, podaci s težinama $w_i > 0$. Pokažite da je $\alpha^* = \frac{1}{W} \sum_{i=1}^m w_i y_i$ jedinstveno rješenje optimizacijskog problema

$$\operatorname{argmin}_{\alpha \in \mathbb{R}} F_{LS}(\alpha), \quad F_{LS}(\alpha) = \sum_{i=1}^m w_i (y_i - \alpha)^2,$$

gdje je $W = \sum_{i=1}^m w_i$. Broj α^* nazivamo *težinska aritmetička sredina* podataka y_1, \dots, y_m .

Primjedba 6.1. Problem (6.1) često zovemo i problem diskretne l_2 aproksimacije za razliku od L_2 aproksimacije funkcije u t. 5, str. 129.

Problem traženja dobre aproksimacije veličine α^* mogli bismo postaviti i tako da umjesto (6.1) promatramo problem

$$\alpha^* \in \operatorname{argmin}_{\alpha \in \mathbb{R}} F_1(\alpha), \quad F_1(\alpha) = \sum_{i=1}^m |y_i - \alpha|, \quad (6.2)$$

ili

$$\alpha^* \in \operatorname{argmin}_{\alpha \in \mathbb{R}} F_\infty(\alpha), \quad F_\infty(\alpha) = \max_{i=1, \dots, m} |y_i - \alpha|. \quad (6.3)$$

Problem (6.2) naziva se problem najbolje diskretne l_1 aproksimacije. Može se pokazati da je njegovo rješenje medijan podataka y_1, \dots, y_m

$$\alpha^* = \underset{i=1, \dots, m}{\text{med}} y_i.$$

Definicija, svojstva i izračunavanje medijana podataka može se vidjeti u Ben-šić i Šuvak (2012); Cupec et al. (2009); Sabo i Scitovski (2008); Sabo et al. (2011); Vazler et al. (2012). Ocem diskretnih l_1 aproksimacija smatra se hrvatski znanstvenik Josip Ruđer Bošković².

Problem (6.3) naziva se problem najbolje diskretne l_∞ aproksimacije. Može se pokazati da je njegovo rješenje

$$\alpha^* = \frac{1}{2} (y_{\min} + y_{\max}), \quad y_{\min} = \min_{i=1, \dots, m} y_i, \quad y_{\max} = \max_{i=1, \dots, m} y_i.$$

Ocem diskretnih l_∞ aproksimacija smatra se ruski matematičar P. L. Čebišev.

Primjer 6.2. *Zadani su podaci*

$$\begin{array}{c|cccc} x_i & -4 & -2 & 2 & 4 \\ \hline y_i & 8 & 6 & 3 & 5 \end{array}$$

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

$$y_i \approx a_1 + a_2x_i, \quad i = 1, \dots, 4,$$

pri čemu „ \approx ” znači minimalno kvadratno odstupanje. Drugim riječima, definirat ćemo odstupanja

$$r_i = y_i - a_1 - a_2x_i, \quad i = 1, \dots, 4$$

² Princip najmanje sume apsolutnih odstupanja pripisuje se hrvatskom znanstveniku Josipu Ruđeru Boškoviću (1711.–1787.), koji je ovaj princip iznio još 1757. godine u radu Boscovich (1757). Dugo je vremena ovaj princip 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).

i pokušati riješiti optimizacijski problem

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(a_1, a_2), \quad F_{LS}(a_1, a_2) = \sum_{i=1}^m (y_i - a_1 - a_2 x_i)^2 = \sum_{i=1}^m r_i^2 = \|\mathbf{r}\|_2^2 \quad (6.4)$$

gdje je \mathbf{r} vektor s komponentama r_i . Problem (6.4) također je jedan problem najmanjih kvadrata koji se lako rješava. Za brojeve $\bar{x} = \sum_{i=1}^m x_i$ i $\bar{y} = \sum_{i=1}^m y_i$ prema (6.1) i Zadatku 6.7, str. 184, vrijedi

$$\sum_{i=1}^m (y_i - a_1 - a_2 x_i)^2 = \sum_{i=1}^m ((y_i - a_2 x_i) - a_1)^2 \geq \sum_{i=1}^m ((y_i - a_2 x_i) - (\bar{y} - a_2 \bar{x}))^2,$$

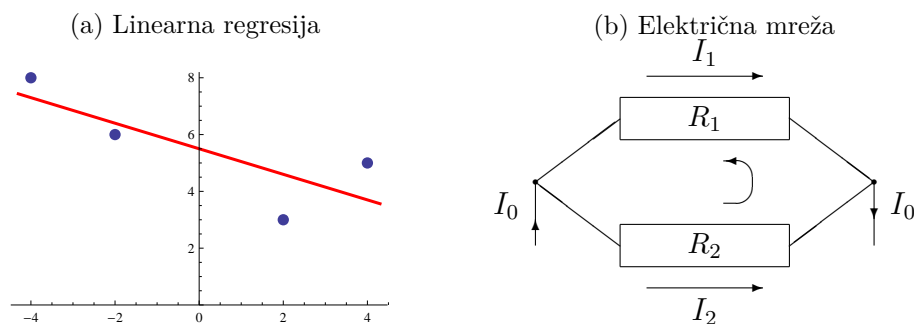
gdje je $a_1 = \bar{y} - a_2 \bar{x}$. To znači da najbolji pravac u smislu najmanjih kvadrata prolazi *centroidom* podataka (\bar{x}, \bar{y}) . Nadalje, sukladno Zadatku 6.1, str. 154, dobivamo

$$\begin{aligned} \sum_{i=1}^m ((y_i - \bar{y}) - a_2(x_i - \bar{x}))^2 &= \sum_{x_i = \bar{x}} (y_i - \bar{y})^2 + \sum_{x_i \neq \bar{x}} (x_i - \bar{x})^2 \left(\frac{y_i - \bar{y}}{x_i - \bar{x}} - a_2\right)^2 \\ &\geq \sum_{x_i = \bar{x}} (y_i - \bar{y})^2 + \sum_{x_i \neq \bar{x}} (x_i - \bar{x})^2 \left(\frac{y_i - \bar{y}}{x_i - \bar{x}} - a_2^*\right)^2, \end{aligned}$$

gdje je

$$a_2^* = \frac{1}{\sum_{x_k \neq \bar{x}} (x_k - \bar{x})^2} \sum_{x_i \neq \bar{x}} (x_i - \bar{x})^2 \frac{y_i - \bar{y}}{x_i - \bar{x}} = \frac{1}{\sum_{x_k \neq \bar{x}} (x_k - \bar{x})^2} \sum_{x_i \neq \bar{x}} (x_i - \bar{x})(y_i - \bar{y}),$$

$$a_1^* = \bar{y} - a_2^* \bar{x}.$$



Slika 6.2: Problemi najmanjih kvadrata

Zadatak 6.2. Pokažite da se u prethodnim izrazima uvjeti $x_k \neq \bar{x}$ te $x_i \neq \bar{x}$ mogu izostaviti ako postoje dva različita $x_i \neq x_j$. Što se može zaključiti u slučaju $x_1 = \dots = x_n$?

Primjedba 6.2. Problem bismo mogli postaviti i tako da umjesto (6.4) promatramo problem

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_1(a_1, a_2), \quad F_1(a_1, a_2) = \sum_{i=1}^m |y_i - a_1 - a_2 x_i| = \|\mathbf{r}\|_1 \quad (6.5)$$

ili

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_\infty(a_1, a_2), \quad F_\infty(a_1, a_2) = \max_{i=1, \dots, m} |y_i - a_1 - a_2 x_i| = \|\mathbf{r}\|_\infty \quad (6.6)$$

Inače, funkcije F_1 i F_∞ nisu derivabilne, pa su zato problemi (6.5) i (6.6) znatno složeniji i zbog toga ćemo ih samo kratko predstaviti u t.6.3, str. 181.

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 preodređenom 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

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(a_1, a_2), \quad F_{LS}(a_1, a_2) = \|\mathbf{J}\mathbf{a} - \mathbf{y}\|_2^2 = \sum_{i=1}^m (a_1 + a_2 x_i - y_i)^2.$$

Primijetite da će za optimalne vrijednosti parametara $a_1 = 5.5$, $a_2 = -0.45$, suma kvadrata odstupanja lijevih od desnih strana jednakosti 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:

Pravilo čvora: Za svaki čvor mreže vrijedi $\sum_k I_k = 0$. Pri tome, dolazeće struje smatraju se pozitivnima, a odlazeće negativnima;

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 pozitivnima, a struje u suprotnom smjeru, negativnima.

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

$$\begin{aligned} I_0 - I_1 - I_2 &= 0 \\ -I_0 + I_1 + I_2 &= 0 \\ -I_1 R_1 + I_2 R_2 &= 0 \end{aligned}$$

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 primljen 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 α_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

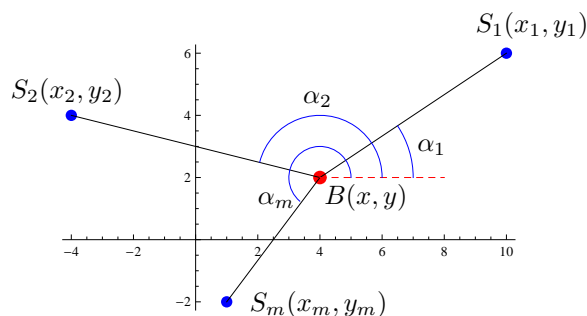
$$\alpha_i = \arctg \frac{y-y_i}{x-x_i}, \quad i = 1, \dots, m. \quad (*)$$

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

$$\operatorname{argmin}_{(a_1, a_2) \in \mathbb{R}^2} F_{LS}(x, y), \quad F_{LS}(x, y) = \sum_{i=1}^m (\arctg \frac{y-y_i}{x-x_i} - \alpha_i)^2.$$

Pretpostavimo općenito da zavisna varijabla y ovisi o nezavisnoj varijabli x po funkcionalnom zakonu

$$y = f(x; \mathbf{a}), \quad \mathbf{a} \in \mathbb{R}^n,$$



Slika 6.3: Traženje pozicije broda

gdje je $\mathbf{a} = (a_1, \dots, a_n)^T$ vektor parametara. 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

$$r_i(\mathbf{a}) = f(x_i; \mathbf{a}) - y_i. \quad (6.7)$$

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

$$F(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{a}) = \frac{1}{2} \|\mathbf{r}(\mathbf{a})\|_2^2, \quad \mathbf{a} \in \mathbb{R}^n, \quad (6.8)$$

postizhe globalni minimum, tj. treba riješiti sljedeći problem globalne optimizacije

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}). \quad (6.9)$$

Funkciju f nazivamo model-funkcija, 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).

Primjedba 6.3. Ako želimo da prilikom procjene optimalnih parametara a_1^*, \dots, a_n^* suma relativnih odstupanja stvarnih od teorijskih vrijednosti bude minimalna, onda ćemo umjesto minimiziranja funkcije (6.8), minimizirati funkciju

$$G(a) = \frac{1}{2} \sum_{i=1}^m \left(\frac{f(x_i; a) - y_i}{y_i} \right)^2.$$

Ako želimo pronaći stacionarne točke minimizirajuće funkcije (6.7), odnosno (6.10), potrebno je gradijent funkcije F izjednačiti s nulom. Tako dobivamo sustav od n jednadžbi s n nepoznanica a_1, \dots, a_n

$$\mathbf{J}^T \mathbf{r} = \mathbf{0} \quad \text{ili} \quad \begin{cases} \sum_{i=1}^m r_i \frac{\partial r_i}{\partial a_1} = 0, \\ \vdots \\ \sum_{i=1}^m r_i \frac{\partial r_i}{\partial a_n} = 0. \end{cases} \quad (6.14)$$

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 model-funkcija f linearna u svim parametrima a_1, \dots, a_n . Tada govorimo o linearnom problemu najmanjih kvadrata (LPNK).

U suprotnom, ako je model-funkcija f nelinearna u barem jednom od parametara a_1, \dots, a_n , govorimo o nelinearnom problemu najmanjih kvadrata (NPNK). U tom slučaju je (6.14) sustav nelinearnih jednadžbi.

Primijetimo da su primjeri 6.1, 6.2 i 6.3. linearni problemi, dok je Primjer 6.4 nelinearan problem najmanjih kvadrata.

6.1 Linearni problemi najmanjih kvadrata

Pretpostavimo da je model-funkcija f linearna u svim parametrima a_1, \dots, a_n i da je općenito oblika

$$f(x; a_1, \dots, a_n) = a_1 \varphi_1(x) + \dots + a_n \varphi_n(x), \quad x \in \mathbb{R} \quad (6.15)$$

gdje su φ_i neprekidne funkcije. Primjerice, kada je model-funkcija polinom, onda su funkcije φ_i potencije $\varphi_i(x) = x^{i-1}$, $i = 1, \dots, n$.

Jacobijeva matrica za model-funkciju (6.15) ne ovisi o vektoru parametara \mathbf{a} . Naime, kako u ovom slučaju (6.7) glasi

$$r_i(\mathbf{a}) = a_1 \varphi_1(x_i) + \dots + a_n \varphi_n(x_i) - y_i,$$

imamo

$$\mathbf{J}_{ij} = \frac{\partial r_i}{\partial a_j} = \varphi_j(x_i), \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Hessijan funkcije F također je bitno jednostavniji i također ne ovisi o vektoru parametara \mathbf{a}

$$\mathbf{H}_F = \mathbf{J}^T \mathbf{J}. \quad (6.16)$$

Vektor odstupanja \mathbf{r} s komponentama $r_i = a_1\varphi_1(x_i) + \dots + a_n\varphi_n(x_i) - y_i$ možemo pisati

$$\mathbf{r} = \mathbf{J}\mathbf{a} - \mathbf{y}, \quad (6.17)$$

gdje je $\mathbf{y} = (y_1, \dots, y_m)^T$. Treba riješiti sljedeći optimizacijski problem

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad \text{gdje je} \quad F(\mathbf{a}) = \frac{1}{2} \|\mathbf{r}(\mathbf{a})\|^2 = \frac{1}{2} \|\mathbf{J}\mathbf{a} - \mathbf{y}\|^2. \quad (6.18)$$

Često puta se u literaturi LPNK (6.18) označava kao

$$\mathbf{J}\mathbf{a} \simeq \mathbf{y}. \quad (6.19)$$

Primjerice, za podatke i model-funkciju 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 Sustav normalnih jednadžbi za LPNK

Stacionarne točke funkcije F dobit ćemo rješavanjem jednadžbe (6.11)

$$\operatorname{grad} F = \mathbf{J}^T \mathbf{r} = 0. \quad (6.20)$$

Hessian \mathbf{H}_F ne ovisi o vektoru \mathbf{a} i zadan je formulom (6.16). Sljedeća lema pokazuje kada će Hessian \mathbf{H}_F biti pozitivno definitan.

Lema 6.1. *Neka je $\mathbf{J} \in \mathbb{R}^{m \times n}$, $m > n$. Matrica $\mathbf{J}^T \mathbf{J}$ pozitivno je 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 bismo 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 $\operatorname{rang} \mathbf{J} = n$, 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

$$(\mathbf{H}_F \mathbf{a}, \mathbf{a}) = (\mathbf{J}^T \mathbf{J} \mathbf{a}, \mathbf{a}) = \mathbf{a}^T (\mathbf{J}^T \mathbf{J} \mathbf{a}) = (\mathbf{J}\mathbf{a})^T (\mathbf{J}\mathbf{a}) = \|\mathbf{J}\mathbf{a}\|^2 > 0. \quad \square$$

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 (6.17) uvrstiti (6.20). Dobivamo jednadžbu

$$\mathbf{J}^T \mathbf{J} \mathbf{a} - \mathbf{J}^T \mathbf{y} = 0 \quad (6.21)$$

Lako se vidi da je rješenje ove jednadžbe

$$\mathbf{a}^* = \mathbf{J}^+ \mathbf{y}, \quad \text{gdje je } \mathbf{J}^+ = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T.$$

Jednadžbu (6.21) nazivamo sustav normalnih jednadžbi, a matricu \mathbf{J}^+ pseudo-inverzna matrica ili Moore-Penroseov generalizirani inverz³ matrice \mathbf{J} .

Primjedba 6.4. U slučaju da je model-funkcija (6.15) polinom n -tog stupnja, a broj podataka m jednak $n + 1$, rješavanjem odgovarajućeg LPNK dobivamo *interpolacijski polinom*.

Primjedba 6.5. 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 (6.21), 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. 78) 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 LPNK rješavali preko sustava normalnih jednadžbi na računalu sa 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 $\mathbf{a}^* = (2.000001, 0)^T$.

³Više o svojstvima i značenju Moore-Penroseovog generaliziranog inverza može se vidjeti primjerice kod Gill et al. (1991); Golub i van Loan (1996); Lawson i 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 (6.18) potražiti ćemo na sljedeći način. Kako je $\mathbf{J} \in \mathbb{R}^{m \times n}$, najprije ćemo algoritmom opisanim u t.3.7, str. 66, 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

$$\mathbf{r} = \mathbf{Q}\mathbf{R}\mathbf{a} - \mathbf{y}. \quad (6.22)$$

Budući da ortogonalna matrica čuva normu, vrijedi

$$\|\mathbf{r}\| = \|\mathbf{Q}\mathbf{R}\mathbf{a} - \mathbf{y}\| = \|\mathbf{Q}(\mathbf{R}\mathbf{a} - \mathbf{Q}^T\mathbf{y})\| = \|\mathbf{R}\mathbf{a} - \mathbf{Q}^T\mathbf{y}\|, \quad (6.23)$$

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 $(\mathbf{Q}^T\mathbf{y})_{m-n}$, su 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

$$\mathbf{R}_n\mathbf{a}^* = (\mathbf{Q}^T\mathbf{y})_n, \quad (6.24)$$

pri čemu je

$$F(\mathbf{a}^*) = \frac{1}{2}\|\mathbf{r}(\mathbf{a}^*)\|^2 = \frac{1}{2}\|(\mathbf{Q}^T\mathbf{y})_{m-n}\|_{(m-n)}^2. \quad (6.25)$$

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 (6.24)

$$\begin{bmatrix} -2. & 0. \\ 0 & 6.32456 \end{bmatrix} \begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} -11. \\ -2.84605 \end{bmatrix},$$

odakle dobivamo $\mathbf{a}^* = (5.5, -0.45)^T$ i $F(\mathbf{a}^*) = 1.1068$.

Primjer 6.8. U Primjeru 2.4, str. 25, 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. 26.

X (m)	α (°)	α^* (°)	$\Delta\alpha^*$
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. 162, Hesssijan 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 se rastavom matrice na singularne vrijednosti (SVD).

Promatrajmo ponovo problem minimizacije (6.18) 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 tim da je $\text{rang}(\mathbf{J}) = k \leq n < m$. Neka je $\mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ dekompozicija matrice \mathbf{J} na singularne vrijednosti, gdje su \mathbf{U} , \mathbf{S} i \mathbf{V} matrice definirane u t. 3.9, str. 76. 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

$$\mathbf{S}_k \mathbf{z}^* = (\mathbf{U}^T \mathbf{y})_k. \quad (6.26)$$

Na taj način određeno je prvih k komponenti vektora \mathbf{z}^* ,

$$z_i^* = \frac{u_i^T \mathbf{y}}{\sigma_i}, \quad i = 1, \dots, k,$$

gdje su $\mathbf{u}_1, \dots, \mathbf{u}_k$ prvih k stupaca matrice \mathbf{U} . Preostalih $n - k$ komponenta z_{k+1}^*, \dots, z_n^* vektora \mathbf{z}^* biramo tako da rješenje \mathbf{a}^* ima minimalnu euklidsku normu. Dakle,

$$\mathbf{a}^* = \mathbf{V}\mathbf{z}^* = \sum_{i=1}^k \frac{u_i^T \mathbf{y}}{\sigma_i} \mathbf{v}_i, \quad \|\mathbf{a}^*\| = \min_{\mathbf{J}\mathbf{a} - \mathbf{y} = \mathbf{r}} \|\mathbf{a}\|, \quad (6.27)$$

gdje su $\mathbf{v}_1, \dots, \mathbf{v}_k$ prvih k stupaca matrice \mathbf{V} . Minimalna vrijednost funkcije F je

$$F(\mathbf{a}^*) = \frac{1}{2} \|(\mathbf{U}^T \mathbf{y})_{m-k}\|_{(m-k)}^2 = \sum_{i=k+1}^m (\mathbf{u}_i^T \mathbf{y})^2. \quad (6.28)$$

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

$$\mathbf{J} = \begin{bmatrix} -0.707107 & -0.707107 \\ -0.707107 & 0.707107 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -0.707107 & -0.707107 \\ -0.707107 & 0.707107 \end{bmatrix}.$$

Primjenom formule (6.27) dobivamo rješenje (koje ima minimalnu l_2 normu!)

$$\mathbf{a}^* = -\frac{1}{2} \left((-0.707107, -0.707107) \cdot (1, 2)^T \right) \cdot (-0.707107, -0.707107)^T = (0.75, 0.75)^T.$$

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 $\mathbf{a}^* = (0.75, 0.75)^T$ i to je rješenje postavljenog LPNK.

Primjedba 6.6. Ako je rang $\mathbf{A} = k$, onda dijagonalnu matricu \mathbf{S} iz SVD rastava (3.29), str. 76 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 σ_i , a sve ostale nule. Zato SVD matrice \mathbf{A} možemo pisati

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U} \left(\sum_{i=1}^k \mathbf{S}_i \right) \mathbf{V}^T = \sum_{i=1}^k \mathbf{U}\mathbf{S}_i\mathbf{V}^T = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \quad (6.29)$$

Najbolja l_2 aproksimacija ranga $\rho < k$ matrice \mathbf{A} tada je (vidi primjerice Björck (1996))

$$\tilde{\mathbf{A}} = \sum_{i=1}^{\rho} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{pri čemu je } \|\mathbf{A} - \tilde{\mathbf{A}}\|_2 = \sigma_{\rho+1}. \quad (6.30)$$

Zadatak 6.3. Kako izgleda rastav (6.29) za matricu \mathbf{A} iz Primjera 3.13, str. 77?

Ova činjenica može se upotrijebiti u svrhu kompresije slike (vidi primjerice Demmel (1997); Trefethen i 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 300×300 s elementima iz intervala $[0, 1]$, pri čemu 0 znači potpuno crno, a 1 potpuno bijelo.

(a) Original



(b) 5% podataka



(c) 10% podataka



(d) 20% podataka



Slika 6.4: Kompresija slike

```
In[1]:= SetDirectory[NotebookDirectory[]];
slika = Import["Kocka.pdf", ColorSpace -> "Graylevel"][[1]]
Print["Dimenzije slike: ", slika[[2]]]
podaci = ImageData[slika];
```

Uzimajući izabrani dio slike, definirajmo matricu \mathbf{A} i prikazimo 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]
```

Aproksimacija s prvih ρ , ($\rho = 15, 30, 60$) 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 (6.29) dobivamo slike aproksimacije originala prikazane na Slikama 6.4 b–d.

6.2 Nelinearni problemi najmanjih kvadrata

Ako model-funkcija 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** (vidi Dennis i Schnabel (1996); Hendrix i Tóth (2010); Ortega i Rheinboldt (1970)).

Radi ilustracije, razmotrimo jednodimenzionalni problem minimizacije. Pretpostavimo da je $F \in C^2[a, b]$ i da se u točki ξ 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)

$$k_1(x) = F(x_0) + F'(x_0)(x - x_0) + \frac{1}{2}F''(x_0)(x - x_0)^2$$

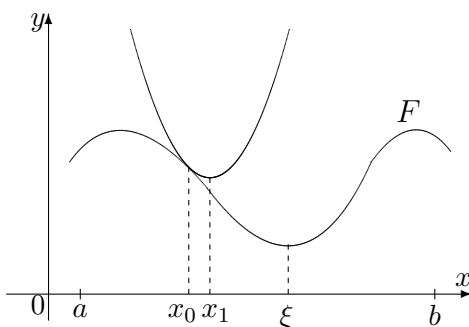
Sljedeću aproksimaciju x_1 tražene točke ξ birat ćemo tako da odredimo točku u kojoj se postiže minimum kvadratne funkcije k_1

$$x_1 = x_0 - \frac{F'(x_0)}{F''(x_0)}.$$

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 i Schnabel (1996)) konvergira prema ξ . Ovo je tzv. **obična Newtonova metoda minimizacije**. Primijetite



Slika 6.5: Newtonova metoda minimizacije

da se ona podudara s Newtonovom metodom (t. 4.4, str. 103) 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 i Schnabel (1996); Ortega i 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, umjesto izračunavanja inverzne matrice $\left(\mathbf{H}_F(\mathbf{a}^k)\right)^{-1}$, u svakom koraku riješit ćemo odgovarajući sustav linearnih jednadžbi, a prethodni iterativni postupak modificirat ćemo na sljedeći način:

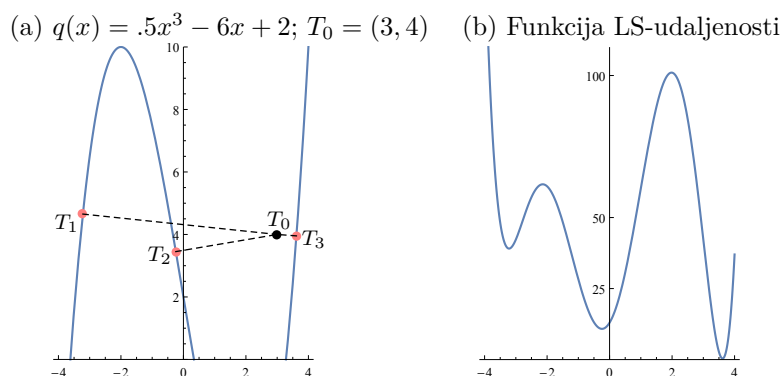
$$\mathbf{a}^{k+1} = \mathbf{a}^k + \mathbf{s}^k, \quad \mathbf{a}^0 \in \mathbb{R}^n, \quad (6.31)$$

gdje je \mathbf{s}^k vektor smjera kretanja iz točke \mathbf{a}^k u točku \mathbf{a}^{k+1} , a dobiva se kao rješenje jednadžbe

$$\mathbf{H}_F(\mathbf{a}^k) \mathbf{s} = -\text{grad } F(\mathbf{a}^k). \quad (6.32)$$

Algoritam 13 (Newtonova metoda minimizacije)**Input:** $F: \mathbb{R}^n \rightarrow \mathbb{R}$;**Input:** $a^0 \in \mathbb{R}^n$; $\epsilon > 0$; $k = 0$;

{Učitati funkciju, početnu aproksimaciju, točnost i maksimalni broj iteracija;}

1: **while** $\|\text{grad } F(\mathbf{a}^k)\| > \epsilon$ **do**2: Riješi sustav $\mathbf{H}_F(\mathbf{a}^k) \mathbf{s} = -\text{grad } F(\mathbf{a}^k)$ i rješenje označi sa \mathbf{s}^k ;3: Stavi $\mathbf{a}^{k+1} = \mathbf{a}^k + \mathbf{s}^k$; $k = k + 1$;4: **end while****Output:** $\{k, \mathbf{a}^{k+1}\}$,Slika 6.6: Euklidska udaljenost točke T_0 do kubne parabole q

Primjer 6.10. Treba izračunati LS-udaljenost točke $T_0 = (3, 4)$ do kubne parabole $q(x) = 0.5x^3 - 6x + 2$ (vidi Sliku 6.6a). LS-udaljenost točke $T_0 = (x_0, y_0)$ do neke točke $T = (x, q(x))$ na grafu funkcije q jednaka je

$$F(x) = d_2(T_0, T) = (x - x_0)^2 + (q(x) - y_0)^2, \quad (6.33)$$

pa je određivanje udaljenosti točke T_0 do grafa funkcije q problem određivanja globalnog minimuma funkcije F na intervalu $[a, b]$ (vidi Sliku 6.6b). Newtonovom metodom minimizacije potražimo lokalni minimum funkcije (6.33) na intervalu $[-1, 1]$ uz odabir početne aproksimacije $x_0 = 0.6$.

Najprije definiramo funkcije q i f , odaberemo početnu aproksimaciju x_0 i ispišemo $\{x_0, F(x_0), |F'(x_0)|\}$:

```
In[1] := q[x_] := x^3/2 - 6 x + 2; u0 = 3; v0 = 4;
        F[x_] := (x - u0)^2 + (q[x] - v0)^2
        x0 = .6; {x0, F[x0], Abs[F'[x0]]}
```

Nakon toga sukcesivne ponavljamo:

```
In[4] := s = -F'[x0]/F''[x0];
        x1 = x0 + s; Print["{x1,F1,F'1}= ", {x1, F[x1], Abs[F'[x1]]}]
        x0 = x1;
```

Na taj način dobivamo niza aproksimacija u kojima funkcija F prima padajuću niz vrijednosti (vidi Tablicu 6.2). Tako dobivamo točku $(-0.24, 10.82)$ lokalnog minimuma funkcije F , čija apscisa odgovara apscisi točke T_2 na Slici 6.6a, a ordinata kvadratu udaljenosti $\|T_0 - T_2\|$.

k	x_k	$F(x_k)$	$ F'(x_k) $
0	0.6	35.92	58.88
1	-0.72	18.34	29.64
2	-0.09	11.64	11.02
3	-0.24	10.82	0.09

Tablica 6.2: Newtonova metoda minimizacije

Međutim, uz početnu aproksimaciju $x_0 = 0.8$, gdje je $F(x_0) = 47.66$, u sljedećoj aproksimaciji dobivamo: $x_1 = -2.07$, i $F(x_0) = 61.56$, dakle vrijednost funkcije ne opada.

Zato ćemo u Algoritam 13 uključiti izračunavanje duljine koraka u smjeru vektora \mathbf{s}^k . Na taj način osigurat ćemo opadanje vrijednosti minimizirajuće funkcije i konvergenciju metode (više o tome može se vidjeti kod Dahlquist i Björck (2008); Scitovski et al. (2014); Stoer i Bulirsch (2002)).

Algoritam 14 (Newton-korak)

Input: $F: \mathbb{R}^n \rightarrow \mathbb{R}$; $s \in \mathbb{R}^n$;

Input: $\mathbf{a}^k \in \mathbb{R}^n$; $\tau \in \langle 0, \frac{1}{2} \rangle$; $\nu \in \langle 0, 1 \rangle$; $\lambda = 1$;

{Učitati funkciju, k -tu aproksimaciju i parametre τ, ν ;}

1: Riješiti sustav: $\mathbf{H}_F(\mathbf{a}^k) \mathbf{s} = -\text{grad } F(\mathbf{a}^k)$ i rješenje označiti sa \mathbf{s}^k ;

2: **while** $F(\mathbf{a}^k + \lambda \mathbf{s}^k) - F(\mathbf{a}^k) \geq \tau \lambda (\text{grad } F(\mathbf{a}^k), \mathbf{s}^k)$, **do**

3: $\lambda = \lambda \cdot \nu$

4: **end while**

Output: λ ,

Nakon toga u Algoritmu 13 u Koraku 3 stavimo

$$\mathbf{a}^{k+1} = \mathbf{a}^k + \lambda \mathbf{s}^k, \quad k = k + 1,$$

a sukcesivni dio *Mathematica*-programa možemo ovako prilagoditi:

```
In[4]:= s = -F'[x0]/F''[x0]; tau = .25; nu = .5; lam = 1;
While[F[x0 + lam s] - F[x0] >= tau lam F'[x0] s,
Print["{lam,F0,Flam}= ", {lam, F[x0], F[x0 + lam s]}];
lam = lam nu
];
Print["{lam, flam}= ", {lam, F[x0 + lam s]}]
x1 = x0 + lam s;
Print["{x1,F1,F'1}= ", {x1, F[x1], Abs[F'[x1]]}]
x0 = x1;
```

k	x_k	$F(x_k)$	$ F'(x_k) $	λ_k
0	0.8	47.66	61.56	0.5
1	-0.64	16.14	25.63	1
2	-0.16	11.07	6.10	1
3	-0.24	10.82	0.05	—

Tablica 6.3: Newtonova metoda minimizacije s regulacijom koraka

Tada Newtonov iterativni proces uz regulaciju duljine koraka s početnom aproksimacijom $x_0 = 0.8$ daje niz aproksimacija u kojima funkcija F postiže opadajuće vrijednosti (vidi Tablicu 6.3). Odgovarajući *Mathematica*-modul za Newtonovu metodu s regulacijom duljine koraka je `FindMinimum`.

Zadatak 6.4. Pronađite i preostale dvije točke T_1, T_3 sa Slike 6.6a, te odredite sve euklidske udaljenosti $d(T_0, T_i)$, $i = 1, 2, 3$.

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 i Schnabel, 1996). Ako je Hessijan funkcije F u točki \mathbf{a}^k pozitivno definitna matrica, može se pokazati da je uz zahtjev $\text{grad } F(\mathbf{a}^k) \neq \mathbf{0}$, uvjet

$$F(\mathbf{a}^k + \lambda \mathbf{s}^k) - F(\mathbf{a}^k) \geq \tau \lambda (\text{grad } F(\mathbf{a}^k), \mathbf{s}^k), \quad (6.34)$$

uvijek moguće ispuniti. Ako je $\text{grad } F(\mathbf{a}^k) = \mathbf{0}$, onda je \mathbf{a}^k stacionarna točka (pa i mogući kandidat za lokalni minimum) funkcije F .

Zadatak 6.5. Pomoću Newtonove metode pronađite lokalni minimum funkcije $F: \mathbb{R}^2 \rightarrow \mathbb{R}$, $F(x_1, x_2) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1)$.

Rješenje: $x_1^* = 0.5$, $x_2^* = -1$, $F(x_1^*, x_2^*) = 0$.

Zadatak 6.6. Za dani skup točaka $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2: i = 1, \dots, m\}$ treba odrediti točke $c_1, c_2 \in \mathbb{R}^2$ u kojima se postiže minimum funkcije

$$F(c_1, c_2) = \sum_{i=1}^m \min\{\|c_1 - a^i\|^2, \|c_2 - a^i\|^2\}.$$

Nacrtajte `ContourPlot` ove funkcije. Može li se za rješavanje ovog problema upotrijebiti Newtonov metoda? Više o ovom problemu može se vidjeti primjerice u Morales-Esteban et al. (2014); Sabo et al. (2010); Sabo i Scitovski (2015); Scitovski i Sabo (2014).

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

$$F(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m (f(x_i; \mathbf{a}) - y_i)^2 = \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{a}) = \frac{1}{2} \|r(\mathbf{a})\|_2^2, \quad \mathbf{a} \in \mathbb{R}^n, \quad (6.35)$$

gdje je $r_i(\mathbf{a}) = f(x_i; \mathbf{a}) - y_i$, $i = 1, \dots, m$.

Prilikom korištenja Newtonove metode potrebno je poznavati sve prve i druge parcijalne derivacije minimizirajuće funkcije. Budući da je u našem slučaju minimizirajuća funkcija (6.35) specijalnog oblika, Newtonovu metodu moguće je prilagoditi i pojednostavniti.

Uz pretpostavku da je u blizini rješenja \mathbf{a}^* član $\mathbf{J}^T(\mathbf{a}) \mathbf{J}(\mathbf{a})$ dominantan u izrazu (6.12), Hessijan \mathbf{H}_F funkcije F možemo aproksimirati s

$$\mathbf{H}_F(\mathbf{a}) \approx \mathbf{J}^T(\mathbf{a}) \mathbf{J}(\mathbf{a}). \quad (6.36)$$

Gradijent minimizirajuće funkcije zadan je s (6.11). Na taj način u Newtonovom iterativnom postupku (6.31)–(6.32) vektor smjera kretanja \mathbf{s}^k dobivamo rješavajući sustav linearnih jednadžbi

$$\mathbf{J}^T(\mathbf{a}^k) \mathbf{J}(\mathbf{a}^k) \mathbf{s}^k = -\mathbf{J}^T(\mathbf{a}^k) \mathbf{r}(\mathbf{a}^k). \quad (6.37)$$

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 (6.37) ima jedinstveno rješenje, a uvjet (6.34) iz kojeg biramo duljinu koraka λ_k uvijek se može ispuniti.

Sustav (6.37) možemo shvatiti kao sustav normalnih jednadžbi za LPNK

$$\mathbf{J}(\mathbf{a}^k) \mathbf{s}^k \simeq -\mathbf{r}(\mathbf{a}^k). \quad (6.38)$$

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. 164).

Primjer 6.11. Zadani su podaci (x_i, y_i) , $i = 1, \dots, m$, gdje je

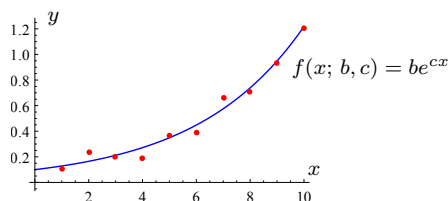
$$x_i = i, \quad y_i = 0.1e^{0.25x_i} + \epsilon_i, \quad \text{gdje je } \epsilon_i \sim N[0, \sigma^2].$$

Treba pronaći optimalne parametre b^* , c^* model-funkcije $f(x; b, c) = be^{cx}$.

Optimalne parametre eksponencijalne model-funkcije potražiti ćemo minimizacijom funkcije (6.35) uz primjenu Gauss-Newtonove metode opisane u Algoritmu 13 pri čemu se koriste specijalni izrazi za gradijent i Hessijan minimizirajuće funkcije.

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, ImageSize -> 200]
```



Izaberimo početnu aproksimaciju $b_0 = 1$, $c_0 = 0.1$ i sukcesivnim pokretanjem niže navedenog programa


```
In[2]:= b = b0; c = c0;
       jac = Table[{f[x[[i]], 1, c], x[[i]] f[x[[i]], b, c]}, {i, m}];
       q = Transpose[jac].(f[x, b, c] - y);
       sol = LinearSolve[Transpose[jac].jac, -q];
       {b0, c0} = {b, c} + sol;
```

dobivamo niz aproksimacija (**grad** označava $\| \text{grad } F(b_k, c_k) \|$):

```
It_0: b = 1          c = 0.1          F(b,c)= 17.3937
It_1: b = 0.02760   c = 0.13948       F(b,c)= 2.94035      grad= 164.786
It_2: b = 0.06252   c = 0.97136       F(b,c)= 1.24512*10^6 grad= 13.3947
It_3: b = 0.00121   c = 0.96954       F(b,c)= 388.789     grad= 2.34097*10^7
It_4: b = 0.00123   c = 0.87385       F(b,c)= 47.5995     grad= 345607.
It_5: b = 0.00270   c = 0.66903       F(b,c)= 1.38512     grad= 46829.8
```

Primjećujemo da je proces vrlo nestabilan: niz funkcijskih vrijednosti ($F(b_k, c_k)$), kao ni norma gradijenta minimizirajuće funkcije monotono ne opada. Zato ćemo u svakoj iteraciji prema Algoritmu 14 odrediti duljinu koraka. Ponovo izaberimo početnu aproksimaciju $b_0 = 1, c_0 = 0.1$, izračunajmo $F_0 = F[x, y, b_0, c_0]$ i izaberimo $\tau = 0.5$. Sukcesivnim pokretanjem niže navedenog programa

```
In[3]:= b = b0; c = c0; lam = 1;
       jac = Table[{f[x[[i]], 1, c], x[[i]] f[x[[i]], b, c]}, {i, m}];
       While[q = Transpose[jac].(f[x, b, c] - y);
            sol = LinearSolve[Transpose[jac].jac, -q];
            {b0, c0} = {b, c} + lam sol; F1 = F[x, y, b0, c0];
            F1 - F0 > -tau lam Plus @@ (sol q), lam = lam/2.
       ]
```

dobivamo:

```
It 0: b = 0.5        c = 0.1          F = 1.74005
It 1: b = 0.26380   c = 0.13948       F = 0.26895   lam=0.5 grad=20.5445
It 2: b = 0.06252   c = 0.22653       F = 0.88786   lam=1   grad= 2.18007
It 3: b = 0.09735   c = 0.27137       F = 0.14935   lam=1   grad=16.6483
It 4: b = 0.09794   c = 0.25302       F = 0.02558   lam=1   grad=10.6933
It 5: b = 0.09879   c = 0.25045       F = 0.02484   lam=1   grad= 0.68737
It 6: b = 0.09880   c = 0.25044       F = 0.02484   lam=1   grad= 0.00309
```

Konačno, cijeli program konstruiramo tako da unaprijed zadamo traženu točnost $\epsilon > 0$ i iterativni postupak ponavljamo tako dugo dok $\| \text{grad } F(b_k, c_k) \|$ ne padne ispod ϵ .

```
In[4]:= b0=.5; c0=.1; F0=F[x,y,b0,c0]; tau=.5; eps=.005; k=0;
       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].(f[x, b, c] - y);
```

```

sol = LinearSolve[Transpose[jac].jac, -q];
{b0, c0} = {b, c} + lam sol; F1 = F[x, y, b0, c0];
F1 - F0 > -tau lam Plus @@ (sol q), lam = lam/2.
];
norm = Norm[q]; norm > eps,
]

```

Za $\epsilon = 0.005$ na kraju petlje dobivamo

```
It 6: b = 0.09880 c = 0.25044 F = 0.02483 lam=1 grad= 0.00309
```

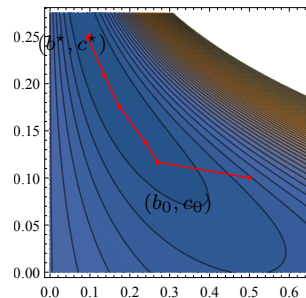
Za procjenu optimalnih parametara matematičkog modela možemo se koristiti Mathematica-modul `FindMinimum`. Također je potrebno zadati početnu aproksimaciju. Prilikom korištenja ovog modula moguće je pratiti iterativni proces (vidi Sliku 6.7).

```

In[5]:= G[b_, c_] := Plus @@ ((y - f[x, b, c])^2);
pts = Reap[
FindMinimum[G[b, c], {{b,.5}, {c,.1}},
StepMonitor -> Sow[{b, c}]]
][[2,1]];
pts = Join[{{.5, .1}}, pts];
ContourPlot[G[b, c], {b,0,1}, {c,0,.3},
Epilog -> {Red, Line[pts], Point[pts]}
]

```

Na taj način dobivamo optimalne parametre $b^* = 0.0988025$, $c^* = 0.250437$ i $F(b^*, c^*) = 0.0248365$.



Slika 6.7: Praćenje iterativnog procesa modula `FindMinimum`

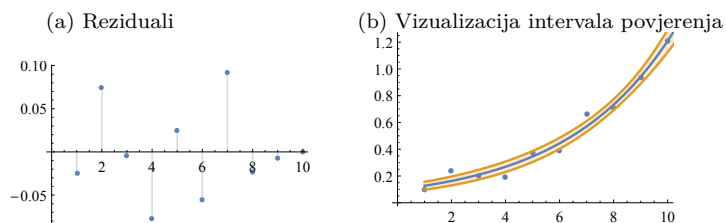
Optimalne parametre matematičkog modela možemo procijeniti i korištenjem specijalno prilagođenog Mathematica-modula `NonlinearModelFit`

```

In[5]:= Clear[b, c, x]
nlm = NonlinearModelFit[pod, b Exp[c x], {b, c}, x]

```

U ovom slučaju, dodatno naredbom `nlm["FitResiduals"]`, možemo dobiti rezidualne (Slika 6.8a), a naredbom `nlm["ParameterConfidenceIntervals", ConfidenceLevel->.9]` možemo odrediti primjerice 90 % intervale povjerenja za parametre (vidi Benšić i Scitovski (1996)): $[0.07292, 0.12468]$, $[0.22039, 0.28048]$ (vidi također Sliku 6.8b).



Slika 6.8: Dodatne mogućnosti modula `NonlinearModelFit`

Primjer 6.12. *Promatramo problem iz Primjera 6.5, str. 158. Podaci o koordinatama točaka $\mathbf{S}_i(x_i, y_i)$, $i = 1, 4$, kao i kutevima α_i , $i = 1, 4$ (u stupnjevima i radijanima prenešenim u interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$), dani su u Tablici 6.4.*

Za ovaj problem najmanjih kvadrata vektor parametara je $\mathbf{a} = (x, y)^T$, komponente r_i ($i = 1, 4$) vektora reziduala su

$$r_i = f(x_i, y_i; x, y) - \alpha_i, \quad \text{gdje je } f(x_i, y_i; x, y) = \arctg \frac{y - y_i}{x - x_i}, \quad (\text{a})$$

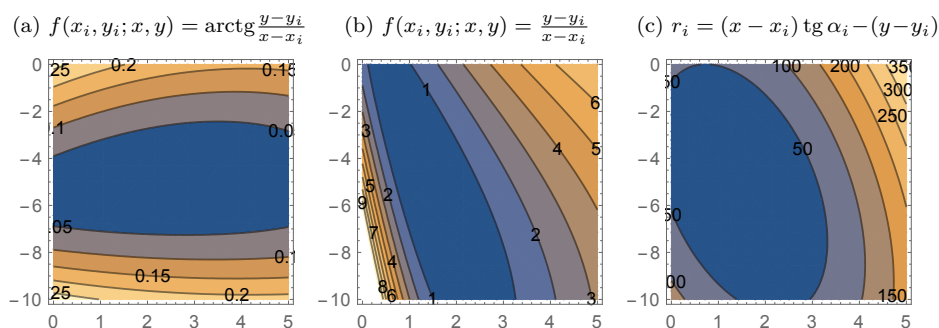
a elementi Jacobijeve matrice su

$$\mathbf{J}_{i1} = \frac{y_i - y}{(x - x_i)^2 + (y - y_i)^2}, \quad \mathbf{J}_{i2} = \frac{x - x_i}{(x - x_i)^2 + (y - y_i)^2}, \quad i = 1, 4.$$

Iteracija	Podaci			Procjena na bazi podataka			
	x_i	y_i	$\alpha_i(^{\circ})$	$\alpha_i(rad)$	$\alpha_i^*(rad)$	$\alpha_i^*(^{\circ})$	r_i
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 6.4: Traženje pozicije broda

Problem ćemo riješiti Gauss-Newtonovom metodom primjenom ranije spomenutih programa. Uz početnu aproksimaciju $\mathbf{a}^0 = (2, -3)^T$ i točnost $\epsilon = 0.00001$,



Slika 6.9: ContourPlot funkcije (6.35) za različite model-funkcije f

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 6.5: Empirijski podaci i procijenjene vrijednosti

nakon 3 iteracije dobivene su koordinate tražene točke $\mathbf{B}_1 = (2.705558, -4.972401)$. Tijek iterativnog procesa prikazan je u Tablici 6.5. Proces se zaustavlja kada se postigne $\|\mathbf{g}^k\| < \epsilon$.

U Tablici 6.4 mogu se također vidjeti veličine „teorijskih kuteva” $\bar{\alpha}_i$ (u radijanima i stupnjevima), pojedini reziduali (odstupanja) i relativna odstupanja u postocima.

Na Slici 6.9a 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

$$r_i = f(x_i, y_i; x, y) - \operatorname{tg} \alpha_i, \quad \text{gdje je} \quad f(x_i, y_i; x, y) = \frac{y - y_i}{x - x_i}. \quad (\text{b})$$

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 6.9b

Ako komponente $r_i, i = 1, \dots, 4$, vektora reziduala zadamo s

$$r_i = (x - x_i) \operatorname{tg} \alpha_i - (y - y_i), \quad (\text{c})$$

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 6.9c) sada su koncentrične elipse.

Koji ćemo pristup: (a), (b) ili (c) primijeniti ovisi o drugim zahtjevima i dodatnoj analizi.

6.2.2 Marquardtova metoda

U Gauss-Newtonovom iterativnom postupku mogu nastupiti problemi ako je rang $\mathbf{J}(\mathbf{a}^k) < n$ za neki k . Naime, tada je matrica $\mathbf{J}^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k)$ singularna. U tom slučaju, Hessijan $\mathbf{H}(\mathbf{a}^k)$ umjesto s (6.36), aproksimirat ćemo s

$$\mathbf{H}_F(\mathbf{a}) = \mathbf{J}(\mathbf{a})^T\mathbf{J}(\mathbf{a}) + \mu\mathbf{I}, \quad \mu > 0, \quad (6.39)$$

gdje je \mathbf{I} jedinična matrica. Vektor smjera kretanja bit će rješenje jednadžbe

$$(\mathbf{J}^T(\mathbf{a}^k)\mathbf{J}(\mathbf{a}^k) + \mu_k\mathbf{I})\mathbf{s}^k = -\mathbf{J}^T(\mathbf{a}^k)\mathbf{r}(\mathbf{a}^k). \quad (6.40)$$

Sustav (6.40) možemo shvatiti kao sustav normalnih jednadžbi za LPNK

$$\begin{bmatrix} \mathbf{J}(\mathbf{a}^k) \\ \sqrt{\mu_k}\mathbf{I} \end{bmatrix} \mathbf{s}^k \simeq \begin{bmatrix} -\mathbf{r}(\mathbf{a}^k) \\ 0 \end{bmatrix}, \quad (6.41)$$

Strategija biranja parametra μ sastoji se u tome da uvijek bude ispunjen uvjet (6.34), a da se pri tome nastojimo zadržati na što manjoj vrijednosti parametra μ . Naime, u tom slučaju vektor smjera kretanja blizak je vektoru smjera kretanja kod Gauss-Newtonove metode, što osigurava bržu konvergenciju procesa. Kada $\mu \rightarrow +\infty$, vektor smjera kretanja približava se smjeru antigradijenta, što također osigurava konvergenciju procesa, ali znatno sporiju (Schwarz i Köckler, 2011).

Primjedba 6.7. 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, rezultati dobiveni logaritmiranjem ili nekim drugim postupkom (Scitovski, 1993), a u slučaju $n = 2$ (model-funkcija ima dva nezavisna parametra) mogu se iskoristiti grafičke mogućnosti računala (programski sustavi: *Mathematica*, *MATLAB*).

Primjedba 6.8. Više o problemima najmanjih kvadrata, njihovom numeričkom rješavanju i praktičnim primjenama može se naći kod Björck i Dahlquist (2007); Gill et al. (1991, 1981); Golub i van Loan (1996); Schatzman (2002); Scitovski (1993);

Scitovski et al. (2014). Spomenimo da se programska podrška navedena kod Press et al. (1992) nažalost, ne može uspješno primijeniti na 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`-modul `NonlinearModelFit`.

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 (vidi Cheney (1998); Schöbel (1999); Watson (1980)), kao što smo spomenuli u Primjedbi 6.2, str. 157.

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 (*engl.: Least Absolute Deviations (LAD)*), tj. treba riješiti optimizacijski problem (vidi primjerice Cheney (1998); Sabo i Scitovski (2008); Sabo et al. (2011); Schöbel (1999))

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \sum_{i=1}^m |f(x_i; \mathbf{a}) - y_i|. \quad (6.42)$$

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 (vidi Cadzow (2002); Watson (1980)).

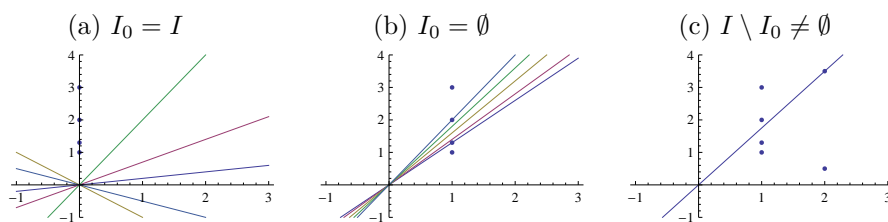
Princip ćemo samo ilustrirati na jednom jednostavnom primjeru (Sabo et al., 2011) procjene parametara linearne model-funkcije $f(x; \alpha) = \alpha x$ iz Hookeovog⁴ zakona o rastezanju elastičnog tijela.

⁴Robert Hooke (1635.–1703.) engleski fizičar – bio je tajnik Royal Society.

Uz oznake $I := \{1, \dots, m\}$ i $I_0 := \{i \in I : x_i = 0\}$, naš problem prelazi u sljedeći optimizacijski problem

$$\operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{i \in I} |y_i - \alpha x_i| = \sum_{i \in I_0} |y_i| + \operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{i \in I \setminus I_0} |x_i| \left| \frac{y_i}{x_i} - \alpha \right|, \quad (6.43)$$

koji je u literaturi poznat kao *problem težinskog medijana* (vidi Gurwitz (1990); Rousseeuw i Leroy (2003); Sabo i Scitovski (2008); Vazler et al. (2012)). Kod rješavanja problema (6.43) razlikujemo tri slučaja:



Slika 6.10: Rješenja optimizacijskog problema (6.43)

(a) Ako je $I_0 = I$, svaki $\alpha \in \mathbb{R}$ rješenje je problema (6.43) (Slika 6.10a);

(b) Ako je $I_0 = \emptyset$, pri čemu je $x_1 = \dots = x_m \neq 0$, problem (6.43) 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 (6.43) (Slika 6.10b);

(c) Ako je $I \setminus I_0 \neq \emptyset$, pri čemu $0 < x_1 < x_m$, rješenje problema (6.43) je $\alpha^* = \operatorname{med}\left(|x_i|, \frac{y_i}{x_i}\right)$ (Slika 6.10c). Veličina $\operatorname{med}\left(|x_i|, \frac{y_i}{x_i}\right)$ je težinski medijan podataka $\frac{y_i}{x_i}$ s težinama $|x_i|$ (vidi Sabo i Scitovski (2008)).

Ako je model-funkcija opća linearna funkcija $f(x; \alpha, \beta) = \alpha x + \beta$, problem (6.42) 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 (vidi Sabo i Scitovski (2008); Schöbel (1999); Watson (1980), a posebno *Metodu dvije točke* u radu Sabo i Scitovski (2008)).

Problem određivanja najbolje ravnine u prostoru na osnovi poznatih točaka-podataka $T_i = (x_i, y_i, z_i)$, $i = 1, \dots, m$, 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 (vidi Fischler i Bolles (1981); Schöbel (1999); Watson (1980), a posebno *Metodu tri točke* u radu Cupec et al. (2009)).

Općenito se problem (6.42) 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 preodređenog 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

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \|\mathbf{z} - \mathbf{X}\mathbf{a}\|_1 = \sum_{i=1}^m |z_i - \mathbf{a}^T \mathbf{x}_i|, \quad (6.44)$$

gdje je $\mathbf{X} \in \mathbb{R}^{m \times n}$ matrica punog ranga čiji su redci vektori \mathbf{x}_i (Sabo et al., 2011).

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 (vidi primjerice Cheney (1998))

$$\operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} F(\mathbf{a}), \quad F(\mathbf{a}) = \max_{i=1, \dots, m} |f(x_i; \mathbf{a}) - y_i|. \quad (6.45)$$

Ako uvedemo vektor reziduala \mathbf{r} s komponentama

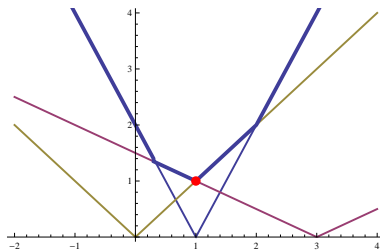
$$r_i = f(x_i; \mathbf{a}) - y_i, \quad i = 1, \dots, m,$$

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_∞ aproksimacije (vidi Cadzow (2002); Watson (1980)).

Princip ćemo samo ilustrirati na jednostavnom primjeru procjene parametara linearne model-funkcije $f(x; \alpha) = \alpha x$ iz Hookeovog zakona.



Slika 6.11: Najbolje l_∞ rješenje

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 $f(x; \alpha) = \alpha x$ prema principu najmanjeg maksimalnog apsolutnog odstupanja. Optimizacijski problem (6.45) 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 6.11).

6.4 Zadaci

Zadatak 6.7. Neka su $x, y \in \mathbb{R}^m$, $m \geq 1$ i $\alpha, \beta \in \mathbb{R}$. Pokažite da vrijedi

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y),$$

gdje je $A(x) = \frac{1}{m} \sum_{i=1}^m x_i$ aritmetička sredina podataka x_1, \dots, x_m .

Zadatak 6.8. Neka su $x, y \in \mathbb{R}^m$, $w \in \mathbb{R}_{++}^m$, $m \geq 1$, i $\alpha, \beta \in \mathbb{R}$. Pokažite da vrijedi

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y),$$

gdje je $A(x) = \frac{1}{W} \sum_{i=1}^m w_i x_i$, a $W = \sum_{i=1}^m w_i$ je težinska aritmetička sredina podataka x_1, \dots, x_m .

Zadatak 6.9. Neka su $x, y \in \mathbb{R}^m$, $m \geq 1$, i $\alpha, \beta \in \mathbb{R}$. Pokažite da ne vrijedi

$$\text{med}(\alpha x + \beta y) = \alpha \text{med}(x) + \beta \text{med}(y),$$

gdje je $\text{med}(x) = \text{med}_{i=1, \dots, m} x_i$ medijan podataka x_1, \dots, x_m .

Zadatak 6.10. U 4-znamenkastoj floating-point aritmetici preko sustava normalnih jednadžbi riješite

$$\begin{aligned} \epsilon x_1 &= 20 \\ \epsilon x_2 &= 30 \\ \epsilon x_3 &= 40 \\ x_1 + x_2 + x_3 &= 10 \end{aligned}$$

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 6.11. 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 6.12. Pokušajte riješiti Zadatak 6.10 i Zadatak 6.11 rješavanjem sustava normalnih jednadžbi pomoću Cholesky dekompozicije.

Zadatak 6.13. Olovnu kuglicu puštamo da pada u vakuumu s 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 6.14. Umjesto interpolacijskog polinoma u Zadatku 2.12, str. 43, primjenom metode najmanjih kvadrata odredite odgovarajući polinom 2. stupnja.

Rješenje: $\alpha(X) = 6.810254 - 0.00454X + 610^{-7}X^2$.

$X(\text{m})$	α	α^*	$\Delta\alpha^*$
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 6.15. Umjesto interpolacijskog polinoma u Zadatku 2.13, str. 44, 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 6.16. Mjerenjem je ustanovljeno da električni otpor R bakrene žice ovisi o temperaturi t na sljedeći način

t ($^{\circ}\text{C}$)	19.1	25	30.1	36	40	45.1	50
R (Ω)	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 6.17. Zadana je model-funkcija $f(x; a_1, a_2) = a_1 + a_2x$ 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 6.18. Ako je \mathbf{A} kvadratna regularna matrica, pokažite da se tada pseudo-inverzna matrica \mathbf{A}^+ podudara s \mathbf{A}^{-1} .

Zadatak 6.19. 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 6.20. Neka je \mathbf{H} Householderova matrica s Householderovim vektorom \mathbf{u} . Pokažite da vrijedi

a) ako je \mathbf{s} , vektor okomit na vektor \mathbf{u} , tada je $\mathbf{H}\mathbf{s} = \mathbf{s}$;

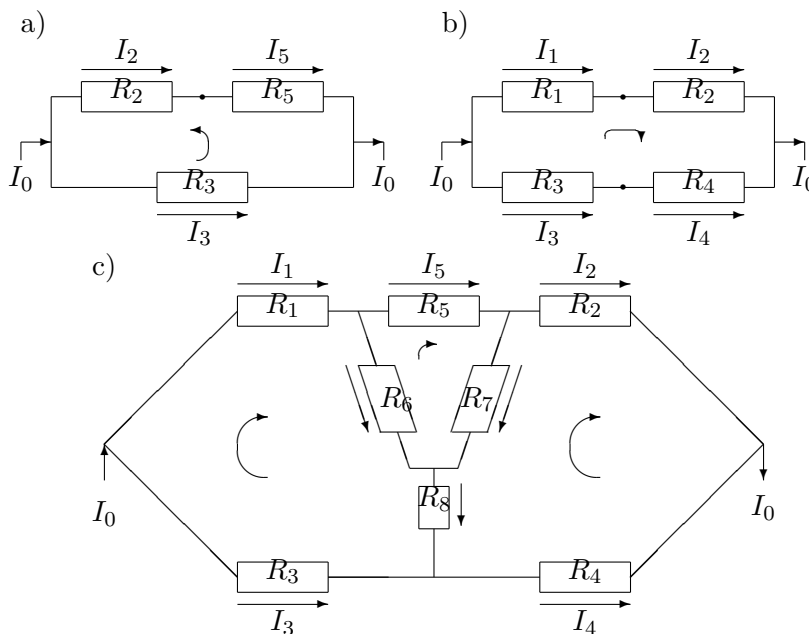
b) ako je \mathbf{z} , vektor paralelan s vektorom \mathbf{u} , tada je $\mathbf{H}\mathbf{z} = -\mathbf{z}$;

- c) 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 6.21. Za mreže prikazane na slikama a), b) i c) izračunajte veličine struja ako je

- a) $I_0 = 5 \text{ A}$, $R_1 = 50 \Omega$, $R_2 = 10 \Omega$, $R_3 = 5 \Omega$
 b) $I_0 = 10 \text{ A}$, $R_1 = 100 \Omega$, $R_2 = 10 \Omega$, $R_3 = 20 \Omega$, $R_4 = 30 \Omega$
 c) $I_0 = 10 \text{ A}$, $R_1 = 5 \Omega$, $R_2 = 50 \Omega$, $R_3 = 5 \Omega$, $R_4 = 50 \Omega$,
 $R_5 = 10 \Omega$, $R_6 = 12 \Omega$, $R_7 = 20 \Omega$, $R_8 = 30 \Omega$



Rješenje:

- a) $I_1 = 0.38462 \text{ A}$, $I_2 = 0.38462 \text{ A}$, $I_3 = 4.61539 \text{ A}$
 b) $I_1 = 3.125 \text{ A}$, $I_2 = 3.125 \text{ A}$, $I_3 = 6.875 \text{ A}$, $I_4 = 6.875 \text{ A}$
 c) $I_1 = 4.485 \text{ A}$, $I_2 = 4.6996 \text{ A}$, $I_3 = 5.515 \text{ A}$, $I_4 = 5.3 \text{ A}$
 $I_5 = 3.519 \text{ A}$, $I_6 = 0.9657 \text{ A}$, $I_7 = 1.1803 \text{ A}$, $I_8 = 0.2146 \text{ A}$

Zadatak 6.22. 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 6.23. 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 6.24. 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. 158, odnosno Primjeru 6.12, str. 178. Izračunajte koordinate točke B prema sve tri varijante navedene u Primjeru 6.12.

Rješenje:

Zadatak 6.25. 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, pokažite 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 6.26. Kako se problem procjene parametara u niže navedenim model-funkcijama može linearizirati (Scitovski, 1993)

$$\begin{array}{lll} a) y = \frac{a}{x} + b, & b) y = \frac{a}{x+c}, & c) y = \frac{1}{ax+b}, \\ d) y = \frac{x}{a+bx}, & e) y = a \ln x + b, & f) y = a \exp(bx), \\ g) y = \frac{1}{(ax+b)^2}, & h) y = ax \exp(-bx), & i) y = \frac{A}{1+b \exp(-cx)} \end{array}$$

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 6.27. Newtonovom metodom minimizacije odredite lokalne minimume funkcije iz Zadatka 4.15, str. 124.

Zadatak 6.28. 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.

Zadatak 6.29. 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.

$$\text{a) } F(x, y) = (ax + by + c)^2, \quad a, b, c \in \mathbb{R}.$$

$$\text{b) } F(x, y) = (a_1x + b_1y + c_1)^2 + (a_2x + b_2y + c_2)^2, \quad a_i, b_i \in \mathbb{R}.$$

Kako se navedeni problemi mogu generalizirati?

Uputa: Pokažite da su glavni minori Hesseove matrice funkcije F

$$\text{a) } \Delta_1 = 2a^2, \quad \Delta_2 = 0;$$

$$\text{b) } \Delta_1 = 2(a_1^2 + a_2^2), \quad \Delta_2 = 2(a_1b_2 + a_2b_1)^2.$$

Zadatak 6.30. Riješite sljedeće LPNK

$$\begin{array}{ll} \text{a) } ax + by + c \simeq 0 & \text{b) } a_1x + b_1y + c_1 \simeq 0 \\ & a_2x + b_2y + c_2 \simeq 0 \end{array}$$

Rješenje: a) $x = \frac{ac}{a^2+b^2}$, $y = \frac{bc}{a^2+b^2}$, $ab \neq 0$.

Poglavlje 7

Numerička integracija

Ako je $f : [a, b] \rightarrow \mathbb{R}$ neprekidna funkcija, a G njena primitivna funkcija, onda se Riemannov integral na intervalu $[a, b]$ može izračunati primjenom Newton-Leibnizove formule:

$$I = \int_a^b f(x) dx = G(b) - G(a).$$

Međutim, u praksi se najčešće pojavljuju takve situacije, gdje nije moguće primijeniti ovu formulu. Može se dogoditi:

- da primitivnu funkciju G nije moguće dobiti elementarnim metodama (vidi Primjer 7.5 i zadatke 7.3, 7.4 i 7.5)
- da je podintegralna funkcija poznata samo u nekoliko (konačno) točaka (kao u Primjeru 7.2 i Zadatku 7.6).

Da bismo ipak aproksimativno (približno) izračunali vrijednost integrala I , podintegralnu funkciju f interpolirat ćemo (vidi t. 2, str. 19) nekom jednostavnijom funkcijom φ i na taj način dobiti aproksimaciju integrala I

$$I^* = \int_a^b \varphi(x) dx.$$

Pri tome, aproksimirajuća funkcija φ treba biti takva da za zadanu točnost $\epsilon > 0$, bude:

$$\Delta I^* = |I - I^*| < \epsilon.$$

Uz pretpostavku poznavanja funkcije f u $n + 1$ točaka $x_0, \dots, x_n \in [a, b]$, za funkciju φ možemo primjerice uzeti Lagrangeov interpolacijski polinom (t. 2.1.1), str. 23.

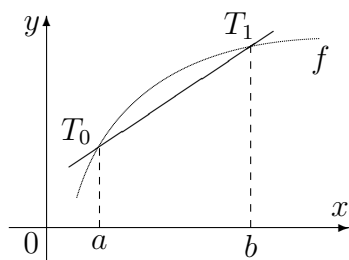
7.1 Trapezno pravilo

Funkciju $f : [a, b] \rightarrow \mathbb{R}$ interpolirat ćemo interpolacijskim polinomom P_1 , stupnja 1 koristeći čvorove interpolacije

$$x_0 = a, \quad x_1 = b.$$

Graf funkcije P_1 pravac je koji prolazi točkama $T_0 = (a, f(a))$, $T_1 = (b, f(b))$,

$$P_1(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$



Slika 7.1: Trapezno pravilo

Lako dobivamo

$$I^* = \int_a^b P_1(x) dx = \frac{b-a}{2} (f(a) + f(b)). \quad (7.1)$$

Geometrijski, I^* predstavlja površinu trapeza sa stranicama $f(a)$ i $f(b)$ i visinom $h = b - a$. Apsolutna pogreška ΔI^* predstavlja veličinu površine između pravca koji prolazi točkama T_0, T_1 i grafa funkcije f (vidi Sliku 7.1). Vrijedi sljedeći teorem.

Teorem 7.1. *Neka je $f \in C^2[a, b]$. Tada postoji $c \in \langle a, b \rangle$, tako da je:*

$$\int_a^b f(x) dx = \frac{b-a}{2} (f(a) + f(b)) - \frac{(b-a)^3}{12} f''(c). \quad (7.2)$$

Dokaz. Prema Teoremu 2.1, str. 29, postoji $\xi(x) \in \langle a, b \rangle$ (ovisan je o x !) tako da je

$$E := I - I^* = \int_a^b (f(x) - P_1(x)) dx = \int_a^b \frac{f''(\xi(x))}{2} (x-a)(x-b) dx.$$

Kako je $(x-a)(x-b) \leq 0$ za sve $x \in [a, b]$, koristeći **poopćeni teorem o srednjoj vrijednosti integralnog računa**¹, zaključujemo da postoji $c \in \langle a, b \rangle$ takav da je

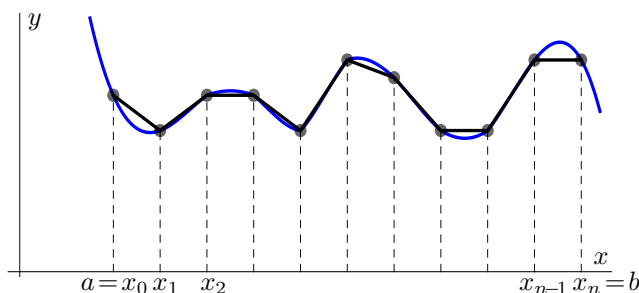
$$E = I - I^* = \frac{1}{2}f''(c) \int_a^b (x-a)(x-b) dx.$$

Uvođenjem supstitucije $x = a + (b-a)t$ dobivamo

$$E = I - I^* = \frac{1}{2}f''(c)(b-a)^3 \int_0^1 t(t-1) dt = -\frac{(b-a)^3}{12}f''(c). \quad \square$$

Ako je interval integracije $[a, b]$ velik, i pogreška E bit će velika. U cilju postizanja bolje aproksimacije I^* integrala I , interval $[a, b]$ podijelit ćemo na više podintervala i onda na svakom od njih primijeniti trapezno pravilo (7.2). Pretpostavimo zato da funkciju $f \in C^2[a, b]$ poznajemo u $n+1$ ekvidistantno raspoređenih točaka $x_0, \dots, x_n \in [a, b]$,

$$x_1 - x_0 = \dots = x_n - x_{n-1} =: h \quad \& \quad x_0 = a, \quad x_n = b.$$



Slika 7.2: Generalizirano trapezno pravilo

Očigledno, točke x_0, \dots, x_n dijele interval $[a, b]$ na n jednakih podintervala duljine $h = \frac{b-a}{n}$. Označimo

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

¹**Teorem.** Ako je $f \in C[a, b]$, a g integrabilna funkcija stalnog predznaka na $[a, b]$ (ili je $g(x) \geq 0$ ili je $g(x) \leq 0$ za sve $x \in [a, b]$), onda postoji $c \in [a, b]$, takav da vrijedi (vidi primjerice (Blanuša, 1973, II. dio))

$$\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx$$

Na svakom podintervalu primijenit ćemo trapezno pravilo (7.2). Tako za interval $[x_{i-1}, x_i]$ dobivamo:

$$\int_{x_{i-1}}^{x_i} f(x) dx = \frac{h}{2} (y_{i-1} + y_i) - \frac{h^3}{12} f''(c_i) \text{ za neki } c_i \in \langle x_{i-1}, x_i \rangle.$$

Cijeli integral I bit će:

$$I = \int_a^b f(x) dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) dx = \frac{h}{2} (y_0 + 2y_1 + \cdots + 2y_{n-1} + y_n) - \frac{h^3}{12} \sum_{i=1}^n f''(c_i). \quad (7.3)$$

Zbog neprekidnosti funkcije f'' na intervalu $\langle a, b \rangle$ postoji $c \in \langle a, b \rangle$, takav da je

$$\frac{1}{n} \sum_{i=1}^n f''(c_i) = f''(c). \quad (7.4)$$

Uvrštavajući (7.4) u (7.3), dobivamo generalizirano trapezno pravilo:

$$I = I^* + E_n,$$

gdje je:

$$I^* = \frac{h}{2} (y_0 + 2y_1 + \cdots + 2y_{n-1} + y_n), \quad (7.5)$$

$$E_n = -\frac{b-a}{12} h^2 f''(c). \quad (7.6)$$

Primjedba 7.1. Ako je zadana točnost $\epsilon > 0$ s kojom treba izračunati integral I i ako označimo:

$$M_2 := \max_{x \in [a, b]} |f''(x)|, \quad (7.7)$$

onda iz (7.6) možemo ocijeniti apsolutnu pogrešku ΔI^* aproksimacije I^* :

$$\Delta I^* = |E_n| \leq \frac{b-a}{12} h^2 M_2 < \epsilon. \quad (7.8)$$

Iz (7.8) tada lako dobivamo broj n podintervala na koji treba podijeliti interval $[a, b]$, kako bismo primjenom generaliziranog trapeznog pravila postigli točnost ϵ

$$n > (b-a) \sqrt{\frac{M_2}{\epsilon} \cdot \frac{b-a}{12}}. \quad (7.9)$$

Niže je naveden algoritam *Trapezno pravilo* implementiran u `Mathematica`-modulu `Trapez[f_, a_, b_, M2_, eps_]` u poglavlju 9.7.1, str. 232, koji primjenom generalizirane trapezne formule izračunava približnu vrijednost integrala funkcije f .

Algoritam 15 (Trapezno pravilo)**Input:** $f: [a, b] \rightarrow \mathbb{R}$;**Input:** $a, b, M2, \epsilon > 0$; $\{M2 = \max_{x \in [a, b]} |f''(x)|\}$ 1: Izračunati: $n = \left\lceil (b - a) \sqrt{\frac{M2(b-a)}{12\epsilon}} \right\rceil + 1$;2: Staviti: $h = (b - a)/n$;3: Izračunati: $Int = \frac{h}{2} \left(f(a) + f(b) + 2 \sum_{i=1}^{n-1} f(a + i * h) \right)$;**Output:** $\{n, Int\}$ **Primjer 7.1.** Znamo da je:

$$I = \int_0^1 \frac{4}{1+x^2} dx = \pi.$$

Izračunat ćemo približnu vrijednost ovog integrala primjenom generaliziranog trapeznog pravila i usporediti rezultate.

Kako je

$$f''(x) = \frac{8(3x^2 - 1)}{(1 + x^2)^3},$$

lako dobivamo $M_2 = |f''(0)| = 8$ (vidi Sliku 7.6a, str. 200). U Tablici 7.1 prikazani su rezultati izračunavanja dobiveni za različite vrijednosti točnosti ϵ primjenom sljedećeg Mathematica-programa.

```
f[x_] := 4/(1 + x^2);
a=0; b=1; M2=8; eps0=5;
Do[
  eps = eps0/10.^i;
  Trapez[f, a, b, M2, eps];
  Print["E=", Abs[int - Pi] ]
, {i, 6}]
```

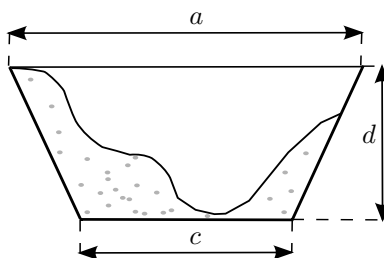
Primjer 7.2. Mjerenjem dubina poprečnog presjeka kanala širine $a = 10$ m na svakih metar, dobiveni su ovakvi podaci

x_i (m)	0	1	2	3	4	5	6	7	8	9	10
y_i (m)	0	0.1	0.5	1.2	1.8	2.3	2.1	2.5	2.1	1.5	0.9

Kanal treba tako prokopati da dobije oblik pravilnog trapeza sa stranicama $a = 10$ m, $c = 6$ m i dubine $d = 3$ m. Treba procijeniti količinu otkopa koji će u tu svrhu biti potrebno uraditi na jednom dužnom metru kanala.

ϵ	Broj točnih decimala	n	I	ΔI
0.5	0	2	3	0.041593
0.05	1	4	3.1	0.010416
0.005	2	12	3.14	0.001157
0.0005	3	37	3.141	0.000122
0.00005	4	116	3.1415	0.000012
0.000005	5	366	3.14159	0.000001

Tablica 7.1: Aproksimacija integrala iz Primjera 7.1 uz primjenu trapeznog pravila i različite točnosti $\epsilon > 0$



Slika 7.3: Presjek kanala

Površinu poprečnog presjeka ΔP zemlje u kanalu koju treba iskopati dobit ćemo primjenom trapezne formule (7.5)

$$\Delta P = 24 - 14.55 = 9.45 \text{ m}^2,$$

a količina iskopa na jednom dužnom metru kanala je $\Delta V = 9.45 \text{ m}^3$.

Zašto u ovom primjeru ne bi bilo dobro najprije odrediti interpolacijski polinom P_{10} i primijeniti formulu $\Delta P = \frac{a+c}{2}d - \int_0^{10} P_{10}(x) dx$? Provjerite biste li tako dobili rezultat $\Delta P = 6.48 \text{ m}^2$.

7.2 Newton-Cotesova formula

U cilju bolje aproksimacije funkcije f , pa time i bolje aproksimacije integrala I , interval $[a, b]$ mogli bismo podijeliti na n jednakih podintervala jednoliko raspoređenim čvorovima $x_i = a + i \frac{b-a}{n}$, a funkciju f tada aproksimirati

interpolacijskim polinomom n -tog reda u Lagrangeovom obliku

$$P_n(x) = \sum_{k=0}^n f(x_k) p_k(x), \quad p_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x-x_i}{x_k-x_i}.$$

Približnu vrijednost I^* integrala I mogli bismo dobiti kao

$$I^* = \int_a^b P_n(x) dx = (b-a) \sum_{k=0}^n \left(\frac{1}{b-a} \int_a^b p_k(x) dx \right) f(x_k) = (b-a) \sum_{k=0}^n \omega_k f(x_k),$$

gdje su $\omega_k = \frac{1}{b-a} \int_a^b p_k(x) dx$ težine. Uz supstituciju $x = a + (b-a)t$ one postaju jednostavnije

$$\omega_k = \frac{1}{b-a} \int_a^b p_k(x) dx = \frac{1}{b-a} \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x-x_i}{x_k-x_i} dx = \int_0^1 \prod_{\substack{i=0 \\ i \neq k}}^n \frac{nt-i}{k-i} dt.$$

Tako dobivamo poznatu kvadraturnu ili Newton-Cotesovu formulu n -tog reda za aproksimaciju integrala I :

$$\int_a^b P_n(x) dx = (b-a) \sum_{k=0}^n \omega_k f(x_k), \quad \omega_k = \int_0^1 \prod_{\substack{i=0 \\ i \neq k}}^n \frac{nt-i}{k-i} dt \quad (7.10)$$

Zadatak 7.1. Pokažite da su ω_k , $k = 0, 1, \dots, n$, iz (7.10) racionalni brojevi sa svojstvom $\sum_{k=0}^n \omega_k = 1$.

Zadatak 7.2. Primjenom Newton-Cotesove formule izvedite trapeznu formulu (7.2).

7.3 Simpsonovo pravilo

Specijalno, iz Newton-Cotesove formule (7.10) za $n = 2$ dobivamo poznato **Simpsonovo pravilo**. Primijetite da je u ovom slučaju: $x_0 = a$, $x_1 = \frac{a+b}{2}$, $x_2 = b$. Najprije ćemo izračunati težine

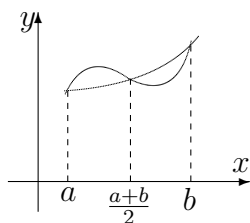
$$\omega_0 = \frac{1}{2} \int_0^1 (2t-1)(2t-2) dt = \frac{1}{6}$$

$$\omega_1 = - \int_0^1 2t(2t-2) dt = \frac{2}{3}$$

$$\omega_2 = \frac{1}{2} \int_0^1 2t(2t-1) dt = \frac{1}{6}.$$

Nakon toga lako dobivamo Simpsonovo pravilo (vidi također Sliku 7.4).

$$I^* = \int_a^b P_2(x) dx = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right). \quad (7.11)$$



Slika 7.4: Simpsonovo pravilo

Može se također pokazati (vidi Stoer (2002); Stewart (1996)) da je pogreška aproksimacije jednaka

$$E = I - I^* = -\frac{1}{90} \left(\frac{b-a}{2} \right)^5 f^{(4)}(c) \text{ za neki } c \in \langle a, b \rangle. \quad (7.12)$$

Zadatak 7.3. Neka je $f \in C^4[a, b]$. Pokažite da tada postoji $c \in \langle a, b \rangle$, tako da je:

$$E = \int_a^b (f(x) - P_2(x)) dx = -\frac{1}{90} \left(\frac{b-a}{2} \right)^5 f^{(4)}(c),$$

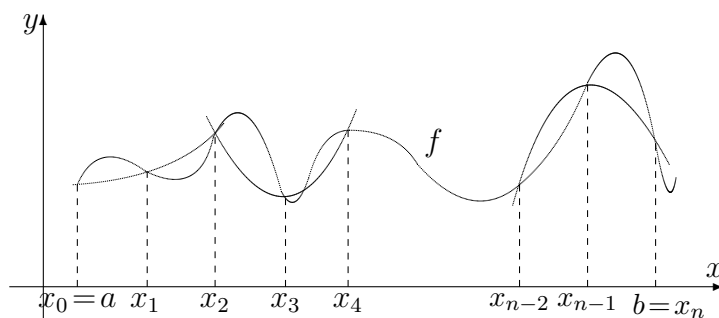
gdje je P_2 interpolacijski polinom koji prolazi točkama: $T_0 = (a, f(a))$, $T_1 = \left(\frac{a+b}{2}, f\left(\frac{a+b}{2}\right) \right)$, $T_2 = (b, f(b))$.

Ako je interval integracije $[a, b]$ velik, i pogreška E bit će velika. U cilju postizanja bolje aproksimacije I^* integrala I , interval $[a, b]$ podijelit ćemo na parni broj $n = 2m$ podintervala duljine $h = \frac{b-a}{n}$ čvorovima $x_i = a + ih$, $i = 0, 1, \dots, n$. Uz oznaku $y_i = f(x_i)$, $i = 0, 1, \dots, n$, redom na po dva podintervala primijenit ćemo Simpsonovo pravilo (7.11) (vidi također Sliku 7.5). Lako se može pokazati da vrijedi generalizirano Simpsonovo pravilo:

$$I = I^* + E_n, \quad (7.13)$$

$$I^* = \frac{h}{3} ((y_0 + y_{2m}) + 4(y_1 + \dots + y_{2m-1}) + 2(y_2 + \dots + y_{2m-2})), \quad (7.14)$$

$$E_n = -\frac{b-a}{180} h^4 f^{(4)}(c) \text{ za neki } c \in \langle a, b \rangle. \quad (7.15)$$



Slika 7.5: Generalizirano Simpsonovo pravilo

Primjedba 7.2. Ako je zadana točnost $\epsilon > 0$ s kojom treba izračunati integral I i ako označimo

$$M_4 = \max_{x \in [a, b]} |f^{(4)}(x)|, \quad (7.16)$$

onda iz (7.15) možemo ocijeniti apsolutnu pogrešku ΔI^* aproksimacije I^*

$$\Delta I^* = |E_n| \leq \frac{b-a}{180} h^4 M_4 < \epsilon. \quad (7.17)$$

Iz (7.17) tada lako dobivamo broj podintervala na koji treba podijeliti interval $[a, b]$, kako bi se primjenom generaliziranog Simpsonovog pravila postigla točnost ϵ

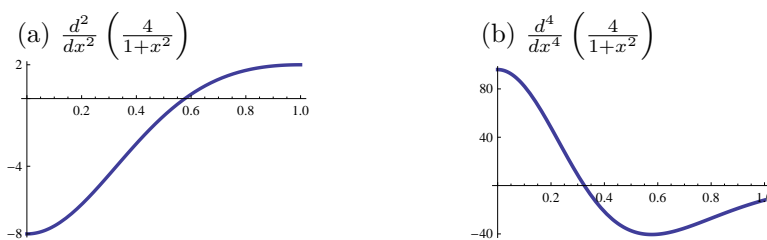
$$n > (b-a) \sqrt[4]{\frac{M_4}{\epsilon} \cdot \frac{b-a}{180}}. \quad (7.18)$$

Niže je naveden algoritam *Simpsonovo pravilo* koji je implementiran u modulu `Simpson[f_, a_, b_, M4_, eps_]` u poglavlju 9.7.2, str. 232. Primjenom generalizirane Simpsonove formule algoritam izračunava približnu vrijednost integrala funkcije f .

Algoritam 16 (Simpsonovo pravilo)**Input:** $f: [a, b] \rightarrow \mathbb{R}$;**Input:** $a, b, M_2, \epsilon > 0$; $\{M_4 = \max_{x \in [a, b]} |f^{(4)}(x)|\}$ 1: Izračunati: $n = \left\lceil (b-a) \sqrt[4]{\frac{M_4}{\epsilon} \cdot \frac{b-a}{180}} \right\rceil + 1$;2: **if** n neparan **then**3: $n = n + 1$;4: **end if**5: Staviti: $h = (b-a)/n$; $n_1 = \frac{n}{2}$

6: Izračunati:

$$Int = \frac{h}{3} \left(f(a) + f(b) + 4 \sum_{i=1}^{n_1} f(a + (2i-1)h) + 2 \sum_{i=1}^{n_1-1} f(a + 2ih) \right);$$

Output: $\{n, Int\}$ Slika 7.6: Druga i četvrta derivacija funkcije $f(x) = \frac{4}{1+x^2}$

Primjer 7.3. Integral iz Primjera 7.1 približno ćemo izračunati primjenom generaliziranog Simpsonovog pravila.

Kako je $f^{(4)}(x) = \frac{96(5x^4 - 10x^2 + 1)}{(x^2 + 1)^5}$, vrijedi (vidi Sliku 7.6b): $M_4 = |f^{(4)}(0)| = 96$.

Primjenom niže navedenog Mathematica-programa u Tablici 7.2 prikazani su rezultati izračunavanja za različite točnosti $\epsilon > 0$.

```
In[1]:=f[x_]:=4/(1+x^2);
a=0; b=1; M4=96; eps0=.5;
Do[
  eps=eps0/10.^i;
  int=Simpson[f, a, b, M4, eps];
  Print["E=", Abs[int - Pi] ]
, {i, 6}]
```

ϵ	Broj točnih znamenki	n	I	ΔI
0.05	2	2	3.1	0.0082
0.005	3	4	3.14	0.0000
0.0005	4	6	3.142	0
0.00005	5	8	3.1416	0
0.000005	6	12	3.14159	0

Tablica 7.2: Aproximacija integrala iz Primjera 7.1 uz primjenu Simpsonovog pravila i različite točnosti $\epsilon > 0$

Zadatak 7.4. Izvedite Newton-Cotesove formule reda $n = 3$ i $n = 4$, te izradite odgovarajuće `Mathematica`-module.

Primjer 7.4. *Riješit ćemo problem postavljen u Primjeru 7.2 primjenom generaliziranog Simpsonovog pravila. Dobivamo: $I^* = 14.77$, odnosno: $\Delta P = 24 - 14.77 = 9.234 \text{ m}^3$.*

Primjer 7.5. *Prilikom određivanja zakrivljenosti zavoja ceste pojavljuju se tzv. Fresnelovi integrali² (Schwarz i Köckler, 2011)*

$$I = \int_0^s \cos\left(\frac{\pi}{2}t^2\right) dt, \quad J = \int_0^s \sin\left(\frac{\pi}{2}t^2\right) dt, \quad s > 0.$$

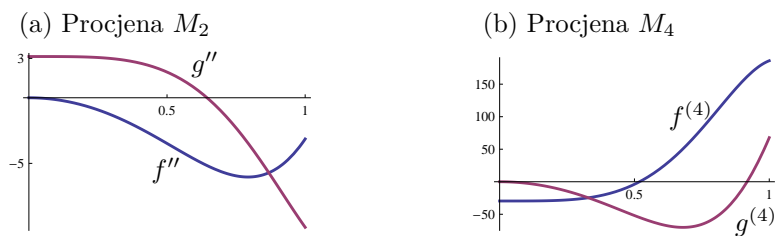
Budući da ne postoje elementarne primitivne funkcije ovih podintegralnih funkcija, vrijednosti ovih integrala odredit ćemo numeričkom integracijom. Za $s = 0.5$ i $s = 1$ izračunat ćemo ove integrale primjenom generaliziranog trapeznog i Simpsonovog pravila uz točnost $\epsilon = 0.00005$ (4 korektne decimale). Da bismo odredili konstante M_2 i M_4 za navedene podintegralne funkcije, nacrtat ćemo grafove njihovih drugih (Slika 7.7a) i četvrtih (Slika 7.7b) derivacija na intervalu $[0, 1]$. Tako dobivamo

$f(t) = \cos(\frac{\pi}{2}t^2)$	M_2	M_4	$g(t) = \sin(\frac{\pi}{2}t^2)$	M_2	M_4
$s = 0.5$	3.5	29.6	$s = 0.5$	3	52
$s = 1$	6	186	$s = 1$	10	70

Za $s = 0.5$ dobivamo: $I = 0.4923$, $J = 0.0647$, a za $s = 1$: $I = 0.7799$, $J = 0.4382$

U niže navedenoj tablici naveden je broj podintervala na koji se mora podijeliti interval integracije da bi se primjenom trapeznog, odnosno Simpsonovog pravila, postigla tražena točnost integrala.

²Augustin-Jean Fresnel (1788.-1827.), francuski inženjer i fizičar



Slika 7.7: Druge i četvrte derivacije funkcija $f(t) = \cos(\frac{\pi}{2}t^2)$ i $g(t) = \sin(\frac{\pi}{2}t^2)$ iz Primjera 7.5

$(s = 0.5)$	Trapezno pravilo	Simpsonovo pravilo	$(s = 0.5)$	Trapezno pravilo	Simpsonovo pravilo
I	28	4	I	101	12
J	26	4	J	130	10

Primjedba 7.3. Više o numeričkim metodama integracije i njenim teorijskim i praktičnim aspektima može se naći kod Ackleh et al. (2010); Dahlquist i Björck (2008); Deuffhard i Hohmann (2008); Schatzman (2002); Schwarz i Köckler (2011); Stoer (2002); Stoer i Bulirsch (2002). Posebno, tu se mogu pronaći i metode koje se zasnivaju na Gaussovima kvadraturnim formulama, a koje u ovom udžbeniku nisu korištene. Programaska podrška FORTRAN naći će se kod Press et al. (1992), a može se koristiti i gotovi korisnički software kao primjerice, *Mathematica* ili *Matlab*.

7.4 Zadaci

Zadatak 7.5. Primjenom generaliziranog trapeznog pravila za zadanu podintegralnu funkciju f i interval integracije $[a, b]$ uz točnost $\epsilon = 0.00005$ odredite konstantu M_2 , broj podintervala n na koji treba podijeliti interval integracije $[a, b]$, te vrijednost aproksimacije integrala I .

- a) $f(x) = \sin^2 x, [0, \pi/4]$; b) $f(x) = \cos^2 x, [0, \pi/4]$;
 c) $f(x) = \sin x - \ln x + e^x, [0.2, 1]$; d) $f(x) = \sqrt{1 + 2x}, [0, 1]$;
 e) $f(x) = x^2(1 - x^2)^{-1/2}, [0, 0.5]$; f) $f(x) = \frac{\sin x}{x}, [\pi/4, \pi/2]$;
 g) $g(x) = \frac{5-x}{9-x^2}, [0, 2]$.

Uputa: prilikom određivanja konstante M_2 može se koristiti kompjutorska grafika, a za izračunavanje integrala navedeni *Mathematica*-modul.

Rješenje:

- a) $M_2 = 2$, $n = 41$, $I^* = 0.1427$.
 b) $M_2 = 2$, $n = 41$, $I^* = 0.6427$.
 c) $M_2 = f''(0.2) = 26$, $n = 149$, $I^* = 2.4148$.
 d) $M_2 = |f''(0)| = 1$, $n = 41$, $I^* = 1.3987$.
 e) $M_2 = |f''(0.5)| = 5$, $n = 33$, $I^* = 0.04532$.
 f) $M_2 = |f''(\pi/4)| \simeq 0.3$, $n = 16$, $I^* = 0.6118$.
 g) $M_2 = |f''(2)| \simeq 0.6$, $n = 90$, $I^* = 1.0473$.

Zadatak 7.6. Primjere iz Zadatka 7.5 riješite primjenom generaliziranog Simpsonovog pravila.

Rješenje:

- a) $M_4 = |f^{(4)}(0)| = 8$, $n = 6$, $I^* = 0.1427$.
 b) $M_4 = |f^{(4)}(0)| = 8$, $n = 6$, $I^* = 0.6427$.
 c) $M_4 = f^{(4)}(0.2) = 3752$, $n = 20$, $I^* = 2.4148$.
 d) $M_4 = |f^{(4)}(0)| = 15$, $n = 8$, $I^* = 1.3987$.
 e) $M_4 = |f^{(4)}(0.5)| \simeq 120$, $n = 6$, $I^* = 0.0453$.
 f) $M_4 = |f^{(4)}(\pi/4)| \simeq 0.2$, $n = 2$, $I^* = 0.6118$.
 g) $M_4 = |f^{(4)}(2)| \simeq 8$, $n = 14$, $I^* = 1.0473$.

Zadatak 7.7. Duljina elipse $\frac{x^2}{a^2} + y = 1$ ($a > 0$) poznati je eliptički integral

$$I = 4 \int_0^{\pi/2} \sqrt{1 - (1 - a^2) \sin^2 t} dt.$$

Za $a \neq 1$ ne postoji elementarna primitivna funkcija ove podintegralne funkcije. Za $a = 0.5$ i $a = 1.5$ izračunajte vrijednost ovog integrala primjenom generaliziranog trapeznog i Simpsonovog pravila uz $\epsilon = 0.00005$.

Rješenje: Za $a = 0.5$ dobivamo:

- a) *Trapezno pravilo*: $M_2 = f''(\pi/2) = 6$, $n = 197$, $I^* = 4.8442$;
 b) *Simpsonovo pravilo*: $M_4 = |f^{(4)}(\pi/2)| = 78$, $n = 18$, $I^* = 0.8442$;

Zadatak 7.8. Za podintegralnu funkciju u nepravom integralu

$$I(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{t^2}{2}) dt$$

ne postoji elementarna primitivna funkcija. Za $x = 0.5, 1, 5$ primjenom trapeznog ili Simpsonovog pravila približno izračunajte vrijednost ovog integrala s točnošću $\epsilon = 0.00005$.

Uputa: pokažite najprije da vrijedi $I(x) = 0.5 + \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-\frac{t^2}{2}) dt$.

Rješenje: $I^*(0.5) = 0.6914$, $I^*(1) = 0.8414$, $I^*(5) = 1$. Pri tome za sva tri intervala integracije imamo:

$$M_2 = |f''(0)| \simeq 0.4, \quad M_4 = |f^{(4)}(0)| \simeq 1.2.$$

Potreban broj podintervala (n) vidi se u sljedećoj tablici.

n	$\alpha = 5^\circ$	$\alpha = 10^\circ$	$\alpha = 45^\circ$
Trapezno pravilo	10	26	289
Simpsonovo pravilo	2	4	26

Zadatak 7.9. Matematičko njihalo s kutem otklona α ima period $T = 4K\sqrt{\frac{l}{g}}$ gdje je

$$K = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1 - \sin^2 \frac{\alpha}{2} \sin^2 \phi}}.$$



Elementarnim metodama ne može se odrediti primitivna funkcija za ovu podintegralnu funkciju. Za $\alpha = 5^\circ, 10^\circ, 45^\circ$ primjenom trapeznog ili Simpsonovog pravila približno izračunajte vrijednost ovog integrala s točnošću $\epsilon = 0.00005$.

Rješenje:

- a) $\alpha = 5^\circ$, $M_2 = |f''(0)| \approx 0.0075$, $M_4 = |f^{(4)}(0)| \approx 0.03$, $K^* = 1.56783$.
 b) $\alpha = 10^\circ$, $M_2 = |f''(0)| \approx 0.03$, $M_4 = |f^{(4)}(0)| \approx 0.13$, $K^* = 1.55915$.
 c) $\alpha = 45^\circ$, $M_2 \approx 0.5$, $M_4 \approx 4.25$, $K^* = 1.41574$.

Potreban broj podintervala na koji treba dijeliti interval integracije prikazan je u sljedećoj tablici.

	$\alpha = 5^\circ$	$\alpha = 10^\circ$	$\alpha = 45^\circ$
Trapezno pravilo	7	14	57
Simpsonovo pravilo	4	4	10

Zadatak 7.10. Širina rijeke je 20 m. Mjerenjem dubina na svaka 2 m njenog poprečnog presjeka dobiveni su sljedeći podaci:

$x_i(\text{m})$	0	2	4	6	8	10	12	14	16	18	20
$dubina(\text{m})$	0.2	0.5	0.9	1.1	1.3	1.7	2.1	1.5	1.1	0.6	0.2

Primjenom trapeznog i Simpsonovog pravila izračunajte površinu poprečnog presjeka rijeke.

Uputa: Postupite slično kao u Primjeru 7.2, str. 195, te iskoristite navedene *Mathematica*-module.

Rješenje:

- a) *Trapezno pravilo* : $S^* = 22 \text{ m}^2$
- b) *Simpsonovo pravilo* : $S^* = 21.9 \text{ m}^2$

Zadatak 7.11. Izvedite Newton-Cotesovu formulu za $n = 3$. Primijenite dobiveni rezultat kod numeričke integracije.

Rješenje: $I^* = \frac{b-a}{8} \left(f(a) + 3f\left(\frac{2a+b}{3}\right) + 3f\left(\frac{a+2b}{3}\right) + f(b) \right)$

Poglavlje 8

Numeričko rješavanje običnih diferencijalnih jednažbi

Prilikom rješavanja diferencijalnih jednažbi u praktičnim primjenama vrlo često nailazimo na slučajeve koje ne možemo elementarno riješiti ili bi elementarno rješenje bilo previše složeno. Tada problem možemo pokušati riješiti numeričkim metodama. Primjenom numeričkih metoda za rješavanje diferencijalnih jednažbi s početnim ili rubnim uvjetom dobivamo tablično zadanu funkciju, pri čemu korak u vrijednosti nezavisne varijable možemo dovoljno fino izabrati. Kontinuiranu funkciju kao aproksimaciju rješenja možemo dobiti kao interpolacijski spline (t. 2.2, str. 32).

Promatramo sljedeći problem

Za zadanu funkciju $f(x, y)$ treba naći funkciju $y(x)$, $x \in [x_0, b]$, koja zadovoljava diferencijalnu jednažbu 1. reda

$$\frac{dy}{dx} = f(x, y), \quad (8.1)$$

uz početni uvjet

$$y(x_0) = y_0. \quad (8.2)$$

Ovaj problem nazivamo **Cauchyjev problem** ili **inicijalni problem**¹. U nekim slučajevima (primjerice, ako je funkcija f linearna) ovaj problem može se riješiti elementarno (kao u Primjeru 8.2, str. 210). Međutim, u većini praktičnih problema prisiljeni smo tražiti aproksimativno rješenje. Štoviše, u praktičnim

¹Engl.: initial value problem, njem.: Anfangswertaufgabe

istraživanjima obično se javlja problem rješavanja sustava diferencijalnih jednadžbi

$$\frac{dy_i}{dx} = f_i(x, y_1, \dots, y_m), \quad i = 1, \dots, m,$$

s početnim uvjetima $y_i(x_0) = c_i$, $i = 1, \dots, m$, ili u vektorskom obliku

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}), \quad \mathbf{y}(x_0) = \mathbf{c}, \quad (8.3)$$

gdje je $\mathbf{y} = (y_1, \dots, y_m)^T$, $\mathbf{c} = (c_1, \dots, c_m)^T$, $\mathbf{f} = (f_1, \dots, f_m)^T$.

Primjedba 8.1. Diferencijalna jednadžba višeg reda može se svesti na sustav diferencijalnih jednadžbi 1. reda. Primjerice, diferencijalna jednadžba 3. reda

$$y''' = g(x, y, y', y''),$$

s početnim uvjetima $y(0) = c_1$, $y'(0) = c_2$, $y''(0) = c_3$, uz supstitucije

$$z_1 = y, \quad z_2 = y', \quad z_3 = y'',$$

prelazi u sustav diferencijalnih jednadžbi 1. reda

$$\begin{aligned} z_1' &= z_2, & z_1(0) &= c_1 \\ z_2' &= z_3, & z_2(0) &= c_2 \\ z_3' &= g(x, z_1, z_2, z_3), & z_3(0) &= c_3. \end{aligned}$$

Zato ima smisla posebnu pozornost posvetiti rješavanju Cauchyjevog problema za linearnu diferencijalnu jednadžbu (8.1)–(8.2), odnosno za sustav linearnih diferencijalnih jednadžbi (8.3).

Zadatak 8.1. Putanja balističkog projektila u vakuumu opisana je sustavom diferencijalnih jednadžbi navedenom u Primjeru 4.16, str. 111. Napišite ovaj sustav u obliku Cauchyjevog problema za sustav linearnih diferencijalnih jednadžbi i riješite ga.

Primjer 8.1. (*Kosi hitac u zraku ispunjenom prostoru – vidi Björck i Dahlquist (2007)*) *Pretpostavljamo da je otpor zraka proporcionalan kvadratu trenutne brzine kretanja projektila. Tada putanju balističkog projektila u zraku ispunjenom prostoru možemo opisati sljedećim Cauchyjevim problemom (\dot{x} označava prvu, a \ddot{x} drugu derivaciju puta u smjeru x po vremenu)*

$$\begin{aligned} m\ddot{x} &= -\gamma\dot{x}\sqrt{\dot{x}^2 + \dot{y}^2}, & x(0) &= 0, & \dot{x}(0) &= z_0 \cos \varphi, \\ m\ddot{y} &= -mg - \gamma\dot{y}\sqrt{\dot{x}^2 + \dot{y}^2}, & y(0) &= 0, & \dot{y}(0) &= z_0 \sin \varphi, \end{aligned}$$

gdje je m masa, φ izlazni kut, z_0 početna brzina projektila, a $\gamma = C\delta S$ (vidi Primjer 8.7, str. 217).

Uz oznake

$$z = \frac{\gamma}{m} \sqrt{\dot{x}^2 + \dot{y}^2}, \quad u := \dot{x}, \quad v := \dot{y},$$

prethodni sustav postaje

$$\begin{aligned} \ddot{x} &= -z u, \\ \ddot{y} &= -g - z v. \end{aligned}$$

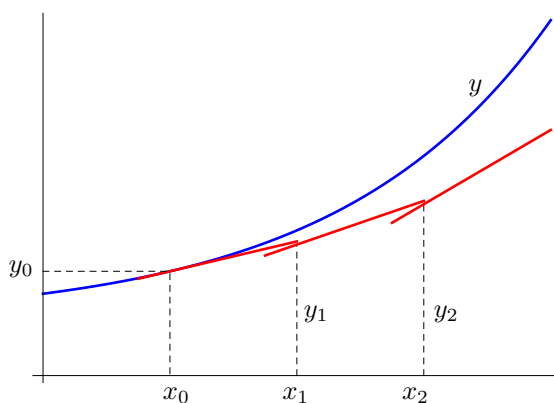
Na taj način dobivamo Cauchyjev problem s četiri linearne diferencijalne jednačbe s odgovarajućim početnim uvjetima

$$\begin{aligned} \frac{dx}{dt} &= u, & \frac{du}{dt} &= -z u, \\ \frac{dy}{dt} &= v, & \frac{dv}{dt} &= -g - z v, \\ x(0) &= 0, \quad y(0) = 0, & u(0) &= z_0 \cos \varphi, \quad v(0) = z_0 \sin \varphi, \end{aligned}$$

koji ćemo rješavati u Primjeru 8.4, str. 212.

8.1 Eulerova metoda

Promatramo problem (8.1)–(8.2). Kako je zbog (8.1) u točki x_0 poznata



Slika 8.1: Eulerova metoda

derivacija funkcije $x \mapsto y(x)$

$$y'(x_0) = f(x_0, y_0),$$

onda funkciju y u okolini točke x_0 možemo aproksimirati linearnim aproksimantom, čiji je graf tangenta (vidi Sliku 8.1)

$$l(x) = y_0 + f(x_0, y_0)(x - x_0). \quad (8.4)$$

Za maleni h , točka $x_1 = x_0 + h$ je bliska točki x_0 , pa ćemo vrijednost tražene funkcije y u točki x_1 aproksimirati vrijednošću linearnog aproksimanta (8.4)

$$y_1 \equiv l(x_1) = y_0 + f(x_0, y_0)h.$$

Ponavljajući postupak, dobivamo Eulerov iterativni proces

$$y_{k+1} = y_k + h f(x_k, y_k), \quad k = 0, 1, \dots \quad (8.5)$$

koji daje aproksimaciju funkcije y u točkama

$$x_i = x_0 + i h, \quad i = 0, 1, \dots$$

Niže je naveden algoritam *Eulerova metoda* implementiran u Mathematica-modulu `Euler[f_, a_, b_, y0_, h_]` u poglavlju 9.8.1, str. 233, koji primjenom Eulerove metode daje diskretno rješenje Cauchyjevog problema.

Algoritam 17 (Eulerova metoda)

Input: $f: [a, b] \rightarrow \mathbb{R}$;

Input: a, b, y_0, h ; {Cauchyjev problem $y' = f(x, y)$, $y(a) = y_0$ rješava se na intervalu $[a, b]$ po koraku $h > 0$ };

1: Staviti: $x = a$; $y = y_0$; `tab={{a, y0}}`;

2: **while** $x < b$ **do**

3: $y = y + f(x, y)$; $x = x + h$;

4: `tab = Append[tab, {x, y}]`;

5: **end while**

Output: `tab`

Eulerova metoda² ne upotrebljava se za rješavanje praktičnih problema. Ako duljina koraka h nije jako malena, pogreške se mogu brzo akumulirati, a ako odaberemo potrebno maleni korak h , broj iteracija postaje nerazumno velik (vidi Primjer 8.2). Ipak, ova metoda ima važno teorijsko značenje, jer se druge, mnogo efikasnije metode, zasnivaju na ovoj ideji.

Primjer 8.2. *Treba riješiti sljedeći Cauchyjev problem na intervalu $[0, 1]$.*

$$y' = x + y, \quad y(0) = 1,$$

²Ovu metodu neki puta naći ćemo u literaturi pod imenom **metoda tangenti** ili Cauchyjeva metoda jer je Cauchy otprilike 100 godina kasnije (1828.) dokazao konvergenciju ove metode.

Rješenje problema je funkcija $y = 2e^x - x - 1$, a potražiti ćemo ga također i pomoću Eulerove metode uz korake $h = 0.2$ i $h = 0.1$. Dobiveni rezultati, kao i usporedba sa stvarnim vrijednostima prikazani su u Tablici 8.1.

x_k	$y(x_k)$	h=0.2		h=0.1	
		y_k	$y(x_k) - y_k$	y_k	$y(x_k) - y_k$
0	1	1	0	1	0
0.1	1.11034			1.11	-0.00034
0.2	1.24281	1.24	-0.00281	1.241	-0.00181
0.3	1.39972			1.3951	-0.00462
0.4	1.58365	1.568	-0.01565	1.57461	-0.00904
0.5	1.79744			1.78207	-0.01537
0.6	2.04424	2.0016	-0.04264	2.02028	-0.02396
0.7	2.32751			2.29231	-0.03520
0.8	2.65108	2.56192	-0.08916	2.60154	-0.04954
0.9	3.01921			2.95169	-0.06752
1	3.43656	3.2743	-0.16226	3.34686	-0.08970

Tablica 8.1: Rješavanje inicijalnog problema $y' = x + y$, $y(0) = 1$

Primjer 8.3. *Promatrajmo jednostavni Cauchyjev problem*

$$y' = y, \quad y(0) = 1.$$

U ovom slučaju Eulerov iterativni postupak (8.5) postaje

$$y_{k+1} = y_k(1 + h), \quad k = 0, 1, \dots$$

ili općenito, za neki $n \in \mathbb{N}$ imamo

$$y_n = y_{n-1}(1 + h) = y_{n-1}(1 + h)^2 = \dots = y_0(1 + h)^n = (1 + h)^n.$$

Da bismo izračunali vrijednost tražene funkcije y u nekoj fiksnoj točki $x > 0$, u intervalu $[0, x]$, odredit ćemo n jednoliko raspoređenih točaka x_i s razmakom $h = \frac{x}{n}$. Primjenom Eulerove metode dobivamo

$$y(x) \approx y_n = \left(1 + \frac{x}{n}\right)^n. \quad (*)$$

Kada pustimo $n \rightarrow \infty$ (tj. $h \rightarrow 0$), dobivamo egzaktno rješenje $y(x) = \exp(x)$, što znači da Eulerova metoda konvergira.

U svrhu ispitivanja brzine konvergencije Eulerove metode u ovom slučaju raspišimo (*) pomoću Taylorove formule

$$y(x) \approx \exp\left(n \ln\left(1 + \frac{x}{n}\right)\right) = \exp\left(x - \frac{x^2}{2n} + \dots\right) \approx e^x e^{-\frac{x^2}{2n}}.$$

Tako dobivamo približnu pogrešku aproksimacije u fiksnoj točki x

$$\Delta y(x) = |y(x) - e^x| \approx e^x \left(1 - e^{-\frac{x^2}{2n}}\right).$$

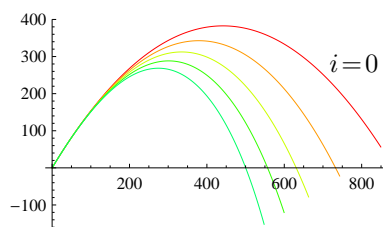
Tako primjerice (koristite niže navedeni *Mathematica*-program), ako želimo vrijednost funkcije y u točki $x = 1$ dobiti s jednom točnom decimalom ($\epsilon = 0.05$), potrebno je staviti $n = 27$ (odnosno $h \approx 0.04$). Za 5 točnih decimala ($\epsilon = 0.5 \times 10^{-5}$) treba staviti $n = 271\,828$ ($h \approx 0.000004$).

```
In[1]:=Clear[n]
      eps = .05; x = 1;
      nn = NSolve[eps == Exp[x] (1 - Exp[-x^2/(2 n)]), n]
      h = x/Floor[n /. nn] // N
```

Primjer 8.4. *Primjenom Eulerove metode riješit ćemo problem gibanja projektila u zraku ispunjenom prostoru opisanom u Primjeru 8.1, str. 208. Dobivamo*

$$\begin{aligned} x_{n+1} &= x_n + h u_n, & u_{n+1} &= u_n - h z_n u_n, \\ y_{n+1} &= y_n + h v_n, & v_{n+1} &= v_n - h(g + z_n v_n), \\ z_n &:= \frac{\gamma}{m} \sqrt{u_n^2 + v_n^2}. \end{aligned}$$

```
In[1]:= n = 1700;
      Euler[h_, ff_] := Module[{g = 9.81, gam = .25 10^(-3) ff,
      x = 0, y = 0, u = 100 Cos[Pi/3], v = 100 Sin[Pi/3]},
      tab = {{x, y, u, v}} // N;
      Do[
      x = x + h u; y = y + h v; z = gam Sqrt[u^2 + v^2];
      u = u - h z u; v = v - h (g + z v);
      tab = Append[tab, {x, y, u, v}]
      ,{i, n}]; tab]
In[2]:=Do[
      ee = Euler[.01, i];
      xx = ee[[All, 1]]; yy = ee[[All, 2]];
      s1[i+1] = ListPlot[Table[{xx[[i]], yy[[i]]},{i,n}],
      PlotStyle->{Hue[.1 i]}, Joined->True, ImageSize->200]{i,0,4}
In[3]:=Show[s1[1], s1[2], s1[3], s1[4], s1[5]]
```

Slika 8.2: Kosi hitac u zraku ispunjenom prostorom s form-faktorom $i = 0, \dots, 4$

Na Slici 8.2 prikazana je putanja balističkog projektila za različite vrijednosti form-faktora.

Primjedba 8.2. Kod Eulerove metode sljedeću aproksimaciju funkcije uvijek izračunavamo samo na bazi jedne prethodne aproksimacije. Zato kažemo da je to jednokoračna metoda.

Eulerova metoda može se poboljšati pomoću tzv. **simetrične formule** ako (8.5) modificiramo na sljedeći način

$$\frac{y_{k+1} - y_k}{h} = \frac{1}{2} (f(x_k, y_k) + f(x_{k+1}, y_{k+1})), \quad y_0 = y(x_0), \quad k = 0, 1, 2, \dots \quad (8.6)$$

Primjerice, za $k = 1$ dobivamo

$$\begin{aligned} y_1 - y_0 &= \frac{1}{2}h (f(x_0, y_0) + f(x_1, y_1)), \\ y_1 - \frac{1}{2}hf(x_1, y_1) &= F_0, \quad F_0 = y_0 + \frac{1}{2}hf(x_0, y_0), \end{aligned}$$

što znači da treba riješiti nelinearnu jednačbu da bi se dobila vrijednost y_1 .

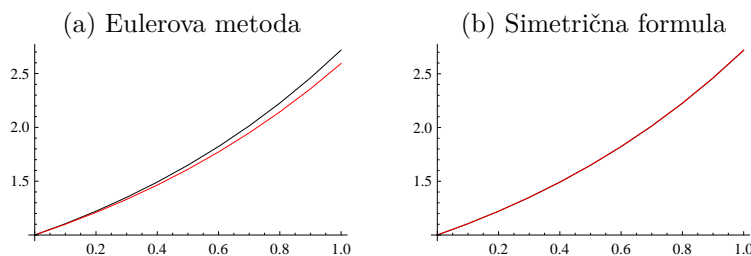
Općenito, (8.6) možemo zapisati kao

$$y_{k+1} - \frac{1}{2}hf(x_{k+1}, y_{k+1}) = F_k, \quad F_k = y_k + \frac{1}{2}hf(x_k, y_k). \quad (8.7)$$

Uz poznavanje y_k , sljedeća vrijednost y_{k+1} dobiva se rješavanjem nelinearne jednačbe (8.7). Budući da je y_{k+1} zadan implicitno, kažemo da je simetrična formula implicitna metoda. Ona daje bolje rezultate od Eulerove metode. Niže je naveden algoritam *Simetrična formula* implementiran u Mathematica-modulu `Sim[f_, a_, b_, y0_, h_]` u poglavlju 9.8.2, str. 233, koji primjenom Simetrične formule daje diskretno rješenje Cauchyjevog problema.

Algoritam 18 (Simetrična formula)**Input:** $f: [a, b] \rightarrow \mathbb{R}$;**Input:** a, b, y_0, h ; {Cauchyjev problem $y' = f(x, y)$, $y(a) = y_0$ rješava se na intervalu $[a, b]$ po koraku $h > 0$ };1: Staviti: $x = a$; $y = y_0$; $\text{tab} = \{\{a, y_0\}\}$;2: **while** $x < b$ **do**3: $x = x + h$; $F = y + .5hf(x - h, y)$;4: Po y riješiti jednadžbu: $\zeta - .5hf(x, y) = F$ na intervalu $[a, b]$;
 {Označimo rješenje s y_1 };5: $\text{tab} = \text{Append}[\text{tab}, \{x, y_1\}]$;6: **end while****Output:** tab

Na Slici 8.3 prikazano je rješenje Cauchyjevog problema iz Primjera 8.3 dobiveno Eulerovom metodom i primjenom simetrične formule.



Slika 8.3: Numeričko rješavanje Cauchyjevog problema iz Primjera 8.3

Zadatak 8.2. Sljedeći Cauchyjev problem

$$\frac{dy}{dt} = -\frac{1}{2}y^3, \quad y(1) = 1,$$

riješite primjenom Eulerove metode i simetrične formule.

Rješenje: $y(t) = \frac{1}{\sqrt{t}}$

Osim navedenih, postoje i brojne druge metode za numeričko rješavanje Cauchyjevog problema kao primjerice, metode Runge-Kutta, višekoračne metode (predictor-corrector metoda, Milneova metoda, Adamsova metoda), Richardsonove ekstrapolacije (specijalno Bulirsch-Stoerova metoda), itd. (vidi primjerice Ackleh et al. (2010); Björck i Dahlquist (2007); Ortega (1990); Samarskij i Gulin (1989); Stoer i Bulirsch (2002)). Malo detaljnije upoznat ćemo se samo s familijom tzv. Runge-Kutta metoda.

8.2 Metode Runge-Kutta

Promatramo Cauchyjev problem (8.1)–(8.2)

$$y' = f(x, y), \quad y(x_0) = y_0.$$

Posebno važno mjesto za rješavanje ovog problema u primjenama zauzima dobro poznata Runge-Kutta (RK) metoda³. Pretpostavimo da poznajemo aproksimaciju y_n tražene funkcije $x \mapsto y(x)$ u točki x_n . Želimo odrediti $(n+1)$ -vu aproksimaciju y_{n+1} u točki $x_n + h$. U tu svrhu na intervalu $(x_n, x_n + h)$ u nekoliko strateških točaka aproksimirat ćemo vrijednost funkcije $x \mapsto f(x, y(x))$, te pomoću njih što bolje aproksimirati razliku $y_{n+1} - y_n$. Egzakti izvod i analiza ovih metoda izlazi iz okvira ovog udžbenika, a može se naći primjerice kod Schwarz i Köckler (2011); Stoer (2002).

Najjednostavniji primjer iz familije RK metoda je tzv. Heunova metoda

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{2}(k_1 + k_2), \\ k_1 &= hf(x_n, y_n), \\ k_2 &= hf(x_n + h, y_n + k_1). \end{aligned} \tag{8.8}$$

Primjedba 8.3. Kako je

$$y(x_{n+1}) - y(x_n) = \int_{x_n}^{x_n+h} \frac{dy}{dx} dx = \int_{x_n}^{x_n+h} f(x, y(x)) dx, \tag{8.9}$$

ako pretpostavimo da je f samo funkcija od x , Heunova metoda (8.8) odgovara trapeznom pravilu (7.5), str. 194, koja ima pogrešku reda veličine $O(h^2)$. Primijetimo također da je za svaku aproksimaciju y_n potrebno dva puta izračunavati vrijednost funkcije f .

RK metoda definirana je s

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ k_1 &= hf(x_n, y_n), \\ k_2 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}), \\ k_3 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}), \\ k_4 &= hf(x_n + h, y_n + k_3). \end{aligned} \tag{8.10}$$

³Ideju je prvi izložio C. Runge u radu *Über die numerische Auflösung von Differentialgleichungen*, *Mathematische Annalen* **46**(1895), 167–178, a kasnije razvio W. Kutta u radu *Beitrag zur näherungsweise Integration von Differentialgleichungen*, *Zeitschrift für Mathematik und Physik* **46**(1901), 435–453.

Primjedba 8.4. Ako pretpostavimo da je f samo funkcija od x , onda iz (8.9) možemo pokazati da RK-metoda (8.10) odgovara Simpsonovoj formuli (7.14), str. 198, uz zamjenu $h \mapsto \frac{h}{2}$. Sjetimo se da Simpsonova formula ima pogrešku reda veličine $O(h^2)$, što se prenosi i na RK metodu (8.10) i u općem slučaju — kada je f funkcija od x i od y . Primijetimo također da je kod RK metode (8.10) za svaku aproksimaciju y_n potrebno četiri puta izračunavati vrijednost funkcije f .

Niže je naveden algoritam *Runge-Kutta* implementiran u *Mathematica*-modulu `RK[f_, a_, b_, y0_, h_]` u poglavlju 9.8.3, str. 234, koji primjenom Runge-Kutta metode daje diskretno rješenje Cauchyjevog problema.

Algoritam 19 (Runge-Kutta metoda)

Input: $f: [a, b] \rightarrow \mathbb{R}$;

Input: a, b, y_0, h ; {Cauchyjev problem $y' = f(x, y)$, $y(a) = y_0$ rješava se na intervalu $[a, b]$ po koraku $h > 0$ };

1: Staviti: $x = a$; $y = y_0$; `tab={{a, y0}}`;

2: **while** $x < b$ **do**

3: $k_1 = hf(x, y)$; $k_2 = hf(x + \frac{h}{2}, y + \frac{k_1}{2})$;

4: $k_3 = hf(x + \frac{h}{2}, y + \frac{k_2}{2})$; $k_4 = hf(x + h, y + k_3)$;

5: $y = y + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$; `x=x+h`;

6: `tab = Append[tab, {x, y}]`;

7: **end while**

Output: `tab`

Primjer 8.5. Na primjeru Cauchyjevog problema

$$y' = 30(\sin x - y), \quad y(0) = 0,$$

čije je egzaktno rješenje lako dobiti (linearna diferencijalna jednažba!),

$$y(x) = \frac{30}{901} (30 \sin x - \cos x + e^{-30x}),$$

na intervalu $[0, 1]$ usporedit ćemo efikasnost Eulerove i RK metode (8.10).

Rezultati, prikazani u Tablici 8.2, dobiveni su korištenjem *Mathematica*-modula `RK[f_, a_, b_, y0_, h_]` u poglavlju 9.8.3, str. 234.

x_k	$y(x_k)$	Eulerova metoda		Runge-Kutta metoda	
		y_k	$y(x_k) - y_k$	y_k	$y(x_k) - y_k$
0	0	0	0	0	0
0.1	0.068	0.299	-0.231	0.112	-0.044
0.2	0.165	-0.002	0.168	0.228	-0.062
0.3	0.263	0.892	-0.629	0.349	-0.085
0.4	0.358	-0.616	0.975	0.475	-0.116
0.5	0.449	2.671	-2.222	0.609	-0.159
0.6	0.536	-3.649	4.186	0.754	-0.218
0.7	0.618	9.232	-8.614	0.916	-0.298
0.8	0.693	-16.313	17.006	1.102	-0.408
0.9	0.761	34.976	-34.214	1.321	-0.560
1	0.822	-67.428	68.250	1.590	-0.768

Tablica 8.2: Cauchyjev problem $y' = 30(\sin x - y)$, $y(0) = 0$, $x \in [0, 1]$

Primjer 8.6. *Primjenom RK-modula riješen je Cauchyjev problem*

$$y' = xy^2 + 1, \quad y(0) = 0,$$

na intervalu $[0, 1]$, a rezultati su prikazani u Tablici 8.3.

x_k	y_k	t_k	v_k	$v(t_k)$
0	0	0	0	0
0.1	0.1	2	18.568	18.573
0.2	0.2	4	32.193	32.211
0.3	0.302	6	40.093	40.123
0.4	0.406	8	44.070	44.1
0.5	0.516	10	45.933	45.955
0.6	0.634	12	46.766	46.79
0.7	0.766	14	47.152	47.16
0.8	0.92	16	47.319	47.323
0.9	1.107	18	47.391	47.394
1	1.350	20	47.424	47.425

Tablica: 8.3 $y' = xy^2 + 1$, $y(0) = 0$

Tablica: 8.4 $\dot{v} = g - \frac{\gamma}{m}v^2$, $v(0) = 0$

Primjer 8.7. *Promatramo slobodni pad tijela mase m u zrakom ispunjenom prostoru.*

Uz pretpostavku da je otpor zraka proporcionalan kvadratu brzine, vrijedi jednačba gibanja

$$m\ddot{s} = mg - \gamma\dot{s}^2 \tag{*}$$

gdje je $g = 9.81 \text{ ms}^{-2}$, $\gamma = C \delta S$, gdje je S površina poprečnog presjeka tijela (m^2), δ specifična gustoća zraka (kg m^{-3}), a C konstanta bez dimenzije koja ovisi o obliku tijela — tzv. form-faktor (vidi Molitz i Ströbel (1963)). Veličina δ u jednadžbi (*) je funkcija temperature (u K) i visine s , ali uz pretpostavku da se radi o malim visinama (nekoliko stotina metara), možemo pretpostaviti da je ona konstantna i jednaka svojoj vrijednosti na razini mora $\delta_0 = 1.22 \text{ kg m}^{-3}$.

Uz oznake $v = \dot{s}$ i $\frac{dv}{dt} = \ddot{s}$, jednadžba (*) glasi

$$\frac{dv}{dt} = g - \frac{\gamma}{m}v^2,$$

čije rješenje je

$$v = b \operatorname{th}(abt), \quad (**)$$

gdje je $a = \gamma/m$, $b^2 = mg/\gamma$. Primijetimo da za $t \rightarrow \infty$ brzina v teži konačnoj veličini b . Za $m = 70 \text{ kp}$, $S = 0.25 \text{ m}^2$ i $C = 1$ primjenom RK metode riješit ćemo jednadžbu (*), na intervalu $[0, 20]$ (pad tijela pratimo prvih 20 sekundi) uz korak u metodi $h = 2$. Dobiveni rezultati u Tablici 8.4 uspoređeni su s egzaktnim vrijednostima iz jednadžbe (**).

Primjedba 8.5. Runge-Kuttovom metodom možemo rješavati i sustave diferencijalnih jednadžbi. Postupak je analogan iterativnom procesu (8.10). Primjerice, za

$$y' = f(x, y, z), \quad z' = g(x, y, z), \quad y(x_0) = y_0, \quad z(x_0) = z_0,$$

imamo

$$\begin{aligned} y_1 &= y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ z_1 &= z_0 + \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4), \end{aligned} \quad (8.11)$$

gdje je

$$\begin{aligned} k_1 &= hf(x_0, y_0, z_0) & m_1 &= hg(x_0, y_0, z_0) \\ k_2 &= hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{m_1}{2}\right), & m_2 &= hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{m_1}{2}\right), \\ k_3 &= hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{m_2}{2}\right), & m_3 &= hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{m_2}{2}\right), \\ k_4 &= hf(x_0 + h, y_0 + k_3, z_0 + m_3), & m_4 &= hg(x_0 + h, y_0 + k_3, z_0 + m_3). \end{aligned}$$

Primjerice, za funkcije $f(x, y, z) = x + z$, $g(x, y, z) = -x$ i početni uvjet $y(0) = z(0) = 1$ na intervalu $[0, 1]$ dobivamo vrijednosti funkcija $x \mapsto y(x)$ i $x \mapsto z(x)$

prikazane u Tablici 8.5.

x_k	y_k	z_k
0	1	1
0.1	1.104	0.995
0.2	1.218	0.98
0.3	1.340	0.955
0.4	1.469	0.92
0.5	1.604	0.875
0.6	1.744	0.82
0.7	1.887	0.755
0.8	2.034	0.68
0.9	2.183	0.595
1	2.333	0.5

Tablica 8.5

x_k	y_k	y'_k
0	1	0
0.2	1.020	0.204
0.4	1.083	0.433
0.6	1.197	0.718
0.8	1.377	1.101
1.0	1.648	1.648
1.2	2.054	2.465
1.4	2.664	3.729
1.6	3.596	5.753
1.8	5.051	9.093
2	7.386	14.772

Tablica 8.6

Primjedba 8.6. Diferencijalna jednačba višeg reda rješava se svodenjem na sustav diferencijalnih jednačbi 1. reda. Tako primjerice Cauchyjev problem

$$y'' = g(x, y, y'), \quad y(x_0) = \alpha, \quad y'(x_0) = \beta,$$

svodimo na sustav

$$\begin{aligned} y' &= z, & y(x_0) &= y_0, \\ z' &= g(x, y, z), & z(x_0) &= z_0, \end{aligned}$$

koji možemo riješiti RK metodom

$$\begin{aligned} y_1 &= y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ z_1 &= z_0 + \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4), \end{aligned}$$

$$\begin{aligned} k_1 &= h z_0 & m_1 &= h g(x_0, y_0, z_0) \\ k_2 &= h \left(z_0 + \frac{m_1}{2} \right), & m_2 &= h g \left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{m_1}{2} \right), \\ k_3 &= h \left(z_0 + \frac{m_2}{2} \right), & m_3 &= h g \left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{m_2}{2} \right), \\ k_4 &= h (z_0 + m_3), & m_4 &= h g(x_0 + h, y_0 + k_3, z_0 + m_3). \end{aligned}$$

Primjer 8.8. Treba riješiti Cauchyjev problem

$$y'' - xy' - y = 0, \quad y(0) = 1, \quad y'(0) = 0.$$

Problem ćemo pretransformirati na sustav diferencijalnih jednadžbi 1. reda

$$\begin{aligned}y' &= z, & y(0) &= 1, \\z' &= xz + y, & z(0) &= 0,\end{aligned}$$

a rješenje potražiti na intervalu $[0, 2]$ primjenom RK metode s korakom $h = 0.2$ (vidi Tablicu 8.6).

Primjer 8.9. *Tzv. Lotka-Volterrin biološki model grabežljivac-plijen opisan je kao sljedeći Cauchyjev problem (vidi primjerice Murray (2002))*

$$\begin{aligned}\frac{dx}{dt} &= x(p_1 - p_2y), & x(0) &= 1, \\ \frac{dy}{dt} &= y(p_3x - p_4), & y(0) &= 0.3.\end{aligned}\tag{8.12}$$

Uz vrijednosti parametara

$$p_1 = 0.77069, \quad p_2 = p_3 = 2.1268, \quad p_4 = 1.984,$$

primjenom RK metode dobivamo funkcije x i y tablično prikazane u Tablici 8.3 i na Slici 8.4a. Funkcije $t \mapsto x(t)$ i $t \mapsto y(t)$ su vremenske periodične funkcije.

t	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5
x	1.0	1.052	1.060	1.013	0.934	0.860	0.819	0.8198	0.854	0.915
y	0.3	0.332	0.381	0.427	0.446	0.428	0.386	0.341	0.307	0.291
t	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
x	0.986	1.044	1.062	1.026	0.950	0.873	0.824	0.816	0.844	0.901
y	0.296	0.324	0.370	0.419	0.445	0.434	0.396	0.350	0.312	0.293

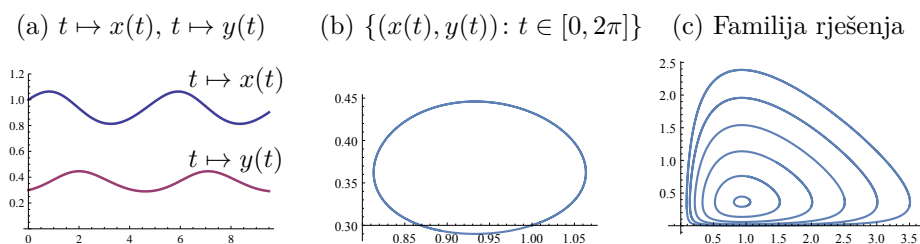
Tablica 8.3: Podaci za Lotka-Volterrin model grabežljivac-plijen

Na Slici 8.4b prikazano je parametarsko rješenje Cauchyjevog problema (8.12) za $t \in [0, 2\pi]$, a na Slici 8.4c prikazana je familija rješenja Cauchyjevog problema (8.12) uz $x(0) \in \{1, 1.5, 2, 2.5, 3, 3.5\}$.

8.3 Metoda diskretizacije

Osim inicijalnih problema koje smo do sada razmatrali, rješavanje praktičnih problema često vodi na rješavanje tzv. rubnih problema⁴.

⁴Engl.: two point boundary value problem, njem.: Randwertaufgabe



Slika 8.4: Lotka-Volterrin model grabežljivac - plijen

Za danu funkciju $f(x, y, y')$ treba odrediti funkciju $x \mapsto y(x)$, $x \in [a, b]$, koja zadovoljava diferencijalnu jednačbu 2. reda

$$y'' = f(x, y, y'), \quad (8.13)$$

uz tzv. rubne uvjete

$$y(a) = \alpha, \quad y(b) = \beta, \quad (a < b). \quad (8.14)$$

Postoji više metoda za rješavanje rubnih problema (vidi primjerice Matthews i Fink (1999); Samarskij i Gulin (1989); Stoer (2002)). Mi ćemo ovdje samo kratko opisati tzv. metodu diskretizacije.

Interval $[a, b]$ podijelit ćemo na n jednakih dijelova stavljajući

$$h = \frac{b-a}{n}, \quad x_i = a + ih.$$

Prvu i drugu derivaciju funkcije y možemo aproksimirati s

$$y'(x_k) \approx \frac{y_{k+1} - y_{k-1}}{2h}, \quad y''(x_k) \approx \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2}.$$

Na taj način rubni problem (8.13)–(8.14) svodi se na rješavanje sustava tzv. diferencijalnih jednačbi

$$y_{k+1} - 2y_k + y_{k-1} = h^2 f_k, \quad k = 1, 2, \dots, n-1,$$

$$y_0 = \alpha, \quad y_n = \beta,$$

gdje je

$$f_k = f\left(x_k, y_k, \frac{y_{k+1} - y_{k-1}}{2h}\right).$$

ili u matričnom obliku

$$\mathbf{A} \mathbf{y} = h^2 \mathbf{F}(\mathbf{y}) - \mathbf{r}, \quad (8.15)$$

gdje je

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}, \mathbf{F}(\mathbf{y}) = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix}, \mathbf{r} = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ \beta \end{bmatrix}.$$

Ako je funkcija f linearna, sustav (8.15) je sustav linearnih jednadžbi i možemo ga pokušati riješiti nekom metodom navedenom u t. 3, str. 47 (posebno vidi Zadatak 3.25, str. 82). Ako je funkcija f nelinearna, onda treba koristiti neku iterativnu metodu za rješavanje sustava nelinearnih jednadžbi (t. 4.5, str. 115).

Primjedba 8.7. Postoji opsežna literatura o numeričkom rješavanju običnih diferencijalnih jednadžbi (vidi primjerice Ackleh et al. (2010); Björck i Dahlquist (2007); Collatz (1966); Deuffhard i Hohmann (2008); Kincaid i Cheney (1996); Kress (1998); Ortega (1990); Plato (2003); Samarskij i Gulin (1989); Schatzman (2002); Stoer i Bulirsch (2002); Süli i Mayers (2006)).

Za ovo područje numeričke matematike postoji vrlo dobra **Fortran**, **Matlab**, **Mathematica** programska podrška.

8.4 Zadaci

Zadatak 8.3. Numerički riješite diferencijalnu jednadžbu s početnim uvjetom na intervalu $[a, b]$. Numerički dobiveno rješenje usporedite s egzaktnim, te nacrtajte pripadne grafove.

- a) $y' + \frac{1}{x}y = \frac{\sin x}{x}, \quad y(\frac{\pi}{2}) = 1, \quad x \in [\frac{\pi}{2}, \pi],$
- b) $y' = 3x^2y, \quad y(0) = 2, \quad x \in [0, 1],$
- c) $x + yy' = 0, \quad y(0) = 2, \quad x \in [0, 2],$
- d) $2xy' = y, \quad y(1) = 2, \quad x \in [1, 3],$
- e) $y' + y = 0, \quad y(0) = 1, \quad x \in [0, 0.6],$
- f) $y' = y_2, \quad y(0) = 0.25, \quad x \in [0, 3],$
- g) $y' = 1 + xy + y^2, \quad y(0) = 0, \quad x \in [0, 2]$

Rješenje:

- a) $y = -\frac{\cos x}{x} + \frac{\pi}{2x},$ b) $y = 2 \exp(x^3),$ c) $x^2 + y^2 = 4,$ d) $y^2 = 4x,$
- e) $y = e^{-x},$ f) $y = \frac{1}{4-x},$ g)

Zadatak 8.4. Primjenom Runge-Kutta metode riješite Cauchyjev problem

$$y' = 100(\sin x - y), \quad y(0) = 0,$$

na intervalu $[0, \pi/2]$ za različite veličine koraka $h = 0.1, 0.01, 0.001$. Dobivene rezultate usporedite s egzaktnim rješenjem

$$y = \frac{1}{1.0001}(\sin x - 0.01 \cos x + 0.01e^{-100x}).$$

Nacrtajte odgovarajuće grafove egzaktnih rješenja i aproksimacija rješenja.

Zadatak 8.5. Numerički riješite Cauchyjev problem na intervalu $I = [0, 1]$

a) $y'' = \sqrt{1 + y'^2}, \quad y(0) = 1 \quad y'(0) = 0,$

b) $y'' - xy' + y = 0, \quad y(0) = 0, \quad y'(0) = 1.$

Rješenje: a) $y = \operatorname{ch} x$, b)

Zadatak 8.6. Numerički riješite sustav diferencijalnih jednačbi

$$\begin{array}{ll} \text{a) } y' = -7y + z, & y(0) = 0, \\ z' = -2y - 5z, & z(0) = 1, \\ x \in [0, \pi] & \end{array} \quad \begin{array}{ll} \text{b) } y' = z, & y(1) = 1 \\ z' = -y, & z(1) = 1 \\ x \in [1, 2] & \end{array}$$

Zadatak 8.7. Primjenom Runge-Kutta metode riješite Cauchyjev problem

a) $y'' + 1001y' + 1000y = 0, \quad y(0) = 0, \quad y'(0) = 1, \quad x \in [0, 0.2],$

b) $y'' + y = x, \quad y(0) = 1, \quad y'(0) = 2, \quad x \in [0, 1],$

za različite veličine koraka $h = 0.01, 0.005, 0.001$. Dobivene rezultate usporedite s egzaktnim rješenjem, te nacrtajte odgovarajuće grafove egzaktnih rješenja i aproksimacija rješenja.

Rješenje: a) $y = \frac{e^{-x} - e^{-1000x}}{999}$, b) $y = \cos x - \sin x + x$.

Zadatak 8.8. Rješenje rubnog problema

$$y'' + y = x, \quad y(0) = 1, \quad y\left(\frac{\pi}{2}\right) = \frac{\pi}{2} - 1,$$

je funkcija

$$y = \cos x - \sin x + x.$$

Riješite ovaj problem metodom diskretizacije koja je spomenuta u t. 8.3, str. 220, uzevši $m = 5$ ili $m = 10$, te usporedite rješenja s egzaktnim rješenjem. Nacrtajte grafove egzaktnih rješenja i aproksimacija rješenja.

Zadatak 8.9. Za koju vrijednost parametra λ sljedeći rubni problem ima netrivialno rješenje ($y \neq 0$)

$$y'' + \lambda y = 0, \quad y(0) = 0, \quad y(1) = 0.$$

Rješenje: $\lambda = n^2\pi^2$, $n = 1, 2, \dots$

Zadatak 8.10. Zadatak 8.9 riješite metodom diskretizacije uz $n = 3$.

Rješenje:

$$\lambda = 9 \text{ (prava vrijednost } \pi^2 = 9.8696),$$

$$\lambda = 27 \text{ (prava vrijednost } 4\pi^2 = 39.48).$$

Zadatak 8.11. S podacima iz Primjera 8.7, str. 217, odredite put s kao funkciju vremena t na intervalu vremena $[0, 20]$.

Uputa: riješite diferencijalnu jednadžbu $\frac{ds}{dt} = b \operatorname{th}(abt)$.

Zadatak 8.12. S podacima iz Primjera 8.7, str. 217 odredite brzinu v kao funkciju puta s na intervalu $[0, 800]$.

$$\text{Rješenje: } v = \sqrt{\frac{mg}{\gamma}} \sqrt{1 - \exp(-\frac{2\gamma}{m})}.$$

Zadatak 8.13. Pokažite da se Runge-Kuttova metoda podudara sa Simpsonovom formulom ako funkcija f ne ovisi o y .

Zadatak 8.14. Izradite kompjutorski program za metodu diskretizacije (t. 8.3, str. 220) koristeći LU-dekompoziciju matrice.

Zadatak 8.15. Slično kao u (8.10), str. 215, definirajte Runge-Kutta metodu za rješavanje diferencijalne jednadžbe 3. reda iz Primjedbe 8.1, str. 208. Izradite odgovarajući kompjutorski program.

Zadatak 8.16. Riješite Cauchyjev problem zadan s

- a) $x' = x - xy, \quad x(0) = 4,$
 $y' = -y + xy, \quad y(0) = 1, \quad t \in [0, 8], \quad h = 0.1$
- b) $x' = -3x - 2y - 2xy^2, \quad x(0) = 0.8,$
 $y' = 2x - y + 2y^3, \quad y(0) = 0.6, \quad t \in [0, 4], \quad h = 0.1$
- c) $x' = y^2 - x^2, \quad x(0) = 2,$
 $y' = 2xy, \quad y(0) = 0.1, \quad t \in [0, 1.5], \quad h = 0.05$
- d) $x' = x^2 - y^2, \quad x(0) = 2,$
 $y' = 2xy, \quad y(0) = 0.6, \quad t \in [0, 1.6], \quad h = 0.02$
- e) $x' = 1 - y, \quad x(0) = -1.2,$
 $y' = x^2 - y^2, \quad y(0) = 0, \quad t \in [0, 5], \quad h = 0.1$
- f) $x' = x^3 - 2xy^2, \quad x(0) = 1,$
 $y' = 2x^2y - y^3, \quad y(0) = 0.2, \quad t \in [0, 2], \quad h = 0.025$

Zadatak 8.17. Za sustave diferencijalnih jednačbi iz Zadatka 8.16, nađite fazno rješenje $y = y(x; c)$, te nacrtajte odgovarajući pramen (familiju) krivulja.

Zadatak 8.18. Riješite sljedeće rubne probleme za $h = 0.2, 0.1, 0.05$

$$a) \quad x'' = \frac{2t}{1+t^2}x'(t) - \frac{2}{1+t^2}x(t) + 1, \quad x(0) = 1.25, \quad x(4) = -0.95,$$

$$b) \quad x'' = -\frac{2}{t}x'(t) + \frac{2}{t^2}x(t) + \frac{10 \cos(\ln t)}{t^2}, \quad x(1) = 1, \quad x(3) = -1,$$

$$c) \quad x'' + \frac{1}{t}x'(t) + \frac{16}{t^2}x(t) = \frac{1}{t^2}, \quad x(1) = 0.75, \quad x(7) = 0.3,$$

$$d) \quad x'' + \frac{2}{t}x'(t) - \frac{2}{t^2}x(t) = \frac{\sin t}{t^2}, \quad x(1) = -0.02, \quad x(6) = 0.02.$$

Poglavlje 9

Mathematica-moduli

9.2 Interpolacija. Spline interpolacija

9.2.1 Hornerov algoritam

Nakon unosa vektora koeficijenata $a = (a_n, \dots, a_1, a_0)$ polinoma $P_n(x) = a_n x^n + \dots + a_1 x + a_0$ i realnog broja z , modul daje vrijednost polinoma P_n u točki z (vidi primjerice Jukić i Scitovski (2004)).

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

9.3 Rješavanje sustava linearnih jednadžbi

9.3.1 Rješavanje trokutastih sustava

Niže su navedeni *Mathematica*-moduli za rješavanje trokutastih sustava (vidi poglavlje 3.3, str. 53). Nakon unosa donje trokutaste matrice \mathbf{A} i vektora \mathbf{b} modul **FS** daje rješavanje donjeg trokutastog sustava $\mathbf{Ax} = \mathbf{b}$ prema (3.12). Slično, nakon unosa gornje trokutaste matrice \mathbf{A} i vektora \mathbf{b} modul **BS** daje rješavanje donjeg trokutastog sustava $\mathbf{Ax} = \mathbf{b}$ prema (3.11).

```
FS[A_, b_] := Module[{x, n=Length[A]},
```

```

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}]; x

BS[A_, b_] := Module[{x, n=Length[A]},
  x = Table[0, {i, n}];
  x[[n]] = b[[n]]/A[[n, n]];
  Do[
    k = n - i;
    x[[k]]=(b[[k]]-Sum[A[[k,j]] x[[j]], {j,k+1,n}))/A[[k,k]]
    ,{i,n-1}]; x
  ]

```

9.3.2 LU-dekompozicija

Niže je naveden *Mathematica*-modul za LU-dekompoziciju regularne kvadratne matrice \mathbf{A} , kojoj su svi glavni minori različiti od nule (vidi poglavlje 3.5, str. 59). Nakon unosa elemenata matrice \mathbf{A} , modul daje donje trokutastu \mathbf{L} i gornje trokutastu \mathbf{U} matricu, tako da je $\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$.

```

LU[A_] := Module[{L,U}, n=Length[A];
L=IdentityMatrix[n]; U=Table[0,{i,n},{j,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]];
{L,U}
]

```

9.3.3 Cholesky-dekompozicija

Niže je naveden *Mathematica*-modul za Cholesky-dekompoziciju simetrične pozitivno definitne matrice \mathbf{A} (vidi poglavlje 3.6, str. 64). Nakon unosa elemenata pozitivno definitne simetrične matrice \mathbf{A} , modul daje donje trokutastu matricu \mathbf{L} , takvu da je $\mathbf{A} = \mathbf{L}^T \mathbf{L}$.

```

Cholesky[A_]:= Module[{L, n=Length[A]},
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]];L
]

```

9.3.4 Iterativne metode za rješavanje sustava linearnih jednadžbi

Niže je naveden Mathematica-modul za rješavanje sustava linearnih jednadžbi najpoznatijim iterativnim metodama (vidi poglavlje 3.8, str. 71).

Jacobijev algoritam

Nakon unosa elemenata matrice sustava \mathbf{A} i vektora slobodnih koeficijenata \mathbf{b} , modul daje aproksimaciju rješenja dobivenu nakon unaprijed zadanog broja iteracija zadanih s (3.24).

```

Jacobi[a_,b_,it_] := Module[{n=Length[b], xs, xn},
xs = xn = Table[0, {i, n}];
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}
]

```

Gauss-Seidelov algoritam

Nakon unosa elemenata matrice sustava \mathbf{A} , vektora slobodnih koeficijenata \mathbf{b} , točnosti $\epsilon > 0$, sukladno (3.25) modul daje aproksimaciju rješenja do na unaprijed zadanu točnost $\epsilon > 0$, pri čemu se ne smije premašiti unaprijed zadani maksimalni broj iteracija.

```

GS[a_,b_,eps_] := Module[{k=0, xs, xn, n=Length[b], it=25},
xs = xn = Table[1, {i,n}];
While[

```

```

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}];
  Norm[xn-xs,1] > eps && k < it, xs=xn; k=k+1];
  Print["it_", k, ": ", N[xn],",", ",Norm[xn-xs,1]]
]

```

9.4 Rješavanje nelinearnih jednadžbi

9.4.1 Metoda bisekcije

Niže je naveden Mathematica-modul za metodu bisekcije za rješavanje jednadžbe $f(x) = 0$ (vidi poglavlje 4.1, str. 87). Nakon unosa funkcije, intervala, točnosti i maksimalnog broja iteracija, modul najprije ispituje postoji li u intervalu nultočka funkcije; Ako postoji, iterativni proces se nastavlja tako dugo dok se ne postigne tražena točnost ili dok se ne premaši maksimalni broj iteracija. Ako se za vrijednost parametra P unese vrijednost 1, ispisivat će se svaka iteracija.

```

Bisekcija[f_,a_,b_,eps_,It_,P_] :=Module[{k=0, xL, xD, xP, gr},
  If[f[a] f[b] < 0,
  xL=a; xD=b; xP=(xL+xD)/2; gr=(xD-xL)/2;
  If[P==1,
    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;
    If[P==1,
      Print["k=",k," xL=",xL," xD=",xD," xP=",xP," gr=",gr]]
  ],
  Print["Interval nije dobro odredjen"]
]; {k, xP}
]

```

9.4.2 Metoda jednostavnih iteracija

Niže je naveden Mathematica-modul za metodu jednostavnih iteracija za rješavanje jednadžbe $f(x) = 0$ (vidi poglavlje 4.3, str. 92). Nakon unosa

funkcije, intervala $[a, b]$, točnosti $\epsilon > 0$, parametra $q < 1$ i maksimalnog broja iteracija IT , modul kao početnu aproksimaciju bira lijevi rub a intervala $[a, b]$, određuje sljedeću aproksimaciju i ispisuje vrijednosti svih parametara. Iterativni postupak se nastavlja tako dugo dok je apsolutna pogreška aproksimacije veća od ϵ i dok je broj iteracija manji od IT .

```
MIT[fi_,a_,b_,eps_,q_,IT_] := Module[{n = 0, x0 = a, gr1},
While[
  n=n+1; x1=fi[x0]; gr1=Abs[x1-x0]*q/(1-q);
  Print["n=",n," x",n,"=",x1," f(x",n,")=" ,x1-fi[x1],
    " gr1=",gr1];
  gr1 > eps && n < IT, x0 = x1
  ]; {x1, n}
]
```

9.4.3 Newtonova metoda

Niže je naveden Mathematica-modul za Newtonovu metodu tangenti za rješavanje jednadžbe $f(x) = 0$ (vidi poglavlje 4.4, str. 103). Nakon unosa funkcije, intervala $[a, b]$, točnosti $\epsilon > 0$, parametara m_1, M_2 i maksimalnog broja iteracija IT , modul kao početnu aproksimaciju bira lijevi rub a intervala $[a, b]$ ako je $a \cdot f''(a) > 0$. U protivnom, za početnu aproksimaciju uzima se desni rub b intervala $[a, b]$. Nakon toga određuje se sljedeća aproksimacija i ispisuje vrijednosti svih parametara. Iterativni postupak se nastavlja tako dugo dok je apsolutna pogreška aproksimacije veća od ϵ i dok je broj iteracija manji od IT .

```
Newton[f_,a_,b_,eps_,m1_,M2_,IT_] := Module[{n=0, x0, x1, gr},
If[f[a] f''[a] > 0, x0 = a, x0 = b ];
While[n = n + 1; x1 = x0 - f[x0]/f'[x0]; gr = Abs[f[x1]]/m1;
  Print["n=",n," x",n,"=",x1," f(x",n,")=" ,f[x1],
    " gr=",gr , " gr1=",Abs[f[x1]]/m1];
  gr > eps && n < IT, x0 = x1
  ]; {x1, n}
]
```


9.5 Aproksimacija funkcija

9.5.1 Gram-Schmidtov postupak ortogonalizacije polinoma

Niže je naveden Mathematica-modul za Gram-Schmidtov postupak ortogonalizacije polinoma (vidi poglavlje 5.1.2, str. 140). Nakon unosa težinske funkcije w , intervala $[a, b]$ i traženog broja n ortogonalnih polinoma, modul sukcesivno izračunava ortogonalne polinome.

```
GSP[w_, a_, b_, n_] := Module[{fi, psi, ortpol},
fi[i_, x_] := x^(i-1);
Do[
  psi[i, t] = fi[i, t] -
  Sum[
    Integrate[w[t] fi[i, t] psi[j, t], {t, a, b}] psi[j, t]/
    Integrate[w[t] psi[j, t]^2, {t, a, b}], {j, i-1}
  ], {i, n}];
ortpol = Table[{i-1, Simplify[psi[i, t]]}, {i, n}
]
```

9.7 Numerička integracija

9.7.1 Trapezno pravilo

Niže je naveden Mathematica-modul za izračunavanje približne vrijednosti integrala uz primjenu generaliziranog trapeznog pravila (vidi poglavlje 7.1, str. 192). Nakon unosa funkcije f , intervala $[a, b]$, broja $M_2 = \max_{x \in [a, b]} |f''(x)|$ i zahtijevane točnosti $\epsilon > 0$, modul daje približnu vrijednost integrala s točnošću $\epsilon > 0$.

```
Trapez[f_, a_, b_, M2_, eps_] := Module[{h, int},
n = Floor[(b - a) Sqrt[M2 (b - a)/(12 eps)]] + 1;
h = (b - a)/n;
int = (h/2) (f[a] + f[b] + 2 Sum[f[a + i h], {i, n - 1}]) // N
]
```

9.7.2 Simpsonovo pravilo

Niže je naveden Mathematica-modul za izračunavanje približne vrijednosti integrala uz primjenu generaliziranog Simpsonovog pravila (vidi poglavlje 7.3,

str. 197). Nakon unosa funkcije f , intervala $[a, b]$, broja $M_4 = \max_{x \in [a, b]} |f^{(4)}(x)|$ i zahtijevane točnosti $\epsilon > 0$, modul daje približnu vrijednost integrala s točnošću $\epsilon > 0$.

```
Simpson[f_, a_, b_, M4_, eps_] := Module[{h, n, n1, int},
n = Floor[(b - a) (M4 (b - a)/(180 eps))^(1/4)] + 1;
If[OddQ[n], n = n + 1];
h = N[(b - a)/n]; n1 = n/2;
int = (h/3)(f[a] + f[b] + 4 Sum[f[a + (2 i-1) h], {i, n1}]
+ 2 Sum[f[a + 2 i h], {i, n1-1}])
(* Print["n=",n," I=",int] *)
]
```

9.8 Numeričko rješavanje običnih diferencijalnih jednažbi

9.8.1 Eulerova metoda

Niže je naveden Mathematica-modul za rješavanje Cauchyjevog problema $y' = f(x, y)$, $y(a) = y_0$, na intervalu $[a, b]$ po koraku $h > 0$ (vidi poglavlje 8.1, str. 209). Nakon unosa funkcije f , intervala $[a, b]$, vrijednosti tražene funkcije y u točki $x = a$ i koraka $h > 0$, modul daje diskretno rješenje Cauchyjevog problema u obliku tablice (x_i, y_i) , $i = 1, \dots, n$, gdje je $n = \lfloor \frac{b-a}{h} \rfloor$.

```
Euler[f_, a_, b_, y0_, h_] := Module[{x = a, y = y0, tab = {{a, y0}}},
While[x < b,
y = y + f[x, y] h; x = x + h; tab = Append[tab, {x, y}];
]; tab
]
```

9.8.2 Simetrična formula

Niže je naveden Mathematica-modul za rješavanje Cauchyjevog problema $y' = f(x, y)$, $y(a) = y_0$, na intervalu $[a, b]$ po koraku $h > 0$ (vidi poglavlje 8.1, str. 209). Nakon unosa funkcije f , intervala $[a, b]$, vrijednosti tražene funkcije y u točki $x = a$ i koraka $h > 0$, modul daje diskretno rješenje Cauchyjevog problema u obliku tablice (x_i, y_i) , $i = 1, \dots, n$, gdje je $n = \lfloor \frac{b-a}{h} \rfloor$.

```

Sim[f_, a_, b_, y0_, h_] := Module[{x = a, y = y0, tab = {{a, y0}}, sol},
While[x < b,
  x = x + h; F = y + .5 h f[x - h, y];
  sol = NSolve[yy - .5*h*f[x, yy] - F == 0, yy];
  y = yy /. sol[[1]];
  tab = Append[tab, {x, y}]
];
tab]

```

9.8.3 Runge-Kutta metoda

Niže je naveden *Mathematica*-modul za rješavanje Cauchyjevog problema $y' = f(x, y)$, $y(a) = y_0$, na intervalu $[a, b]$ po koraku $h > 0$ (vidi poglavlje 8.10, str. 215). Nakon unosa funkcije f , intervala $[a, b]$, vrijednosti tražene funkcije y u točki $x = a$ i koraka $h > 0$, modul daje diskretno rješenje Cauchyjevog problema u obliku tablice (x_i, y_i) , $i = 1, \dots, n$, gdje je $n = \lfloor \frac{b-a}{h} \rfloor$.

```

RK[f, a_, b_, y0_, h_] :=
Module[{x = a, y = y0, k1, k2, k3, k4, tab = {{a, y0}}, sol},
While[x < b,
  k1 = h f[x, y];          k2 = h f[x + h/2, y + k1/2];
  k3 = h f[x + h/2, y + k2/2]; k4 = h f[x + h, y + k3];
  y = y + (1/6) (k1 + 2 k2 + 2 k3 + k4);  x = x + h;
  tab = Append[tab, {x, y}]
]; tab
]

```

Literatura

- 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.
- M. BENŠIĆ, R. SCITOVSKI, *Određivanje intervala povjerenja za neke specijalne nelinearne regresije*, U: T. HUNJAK, L. MARTIĆ, L. NERALIĆ, ur., *Proceedings of the 6th International Conference on Operational Research*, Croatian Operational Research Society, Zagreb, 1996, 87–92.
- 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, SIAM, Philadelphia, 2007.
- A. BJÖRCK, G. DAHLQUIST, *Numerical Methods in Scientific Computing*, Volume II, SAIM, 2008.
- 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 Commentarrii, 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_∞ norm approximate solutions to an overdetermined system of linear equations*, Digital Signal Processing, **12**(2002) 524–560.
- E. W. CHENEY, *Introduction to Approximation Theory*, AMS, 1998.
- 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.
- P. CORN, R. SCITOVSKI, *Jedan problem globalne optimizacije*, Osječki matematički list, **14**(2014) 1–9.
- 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*, SIAM, Philadelphia, 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 1,2 Eine algorithmisch orientierte Einführung*, W. de Gruyter, Berlin, 2008.
- Y. DODGE, editor, *Statistical data analysis based on the L_1 -norm and related methods*, Proceedings of The Third International Conference on Statistical Data Analysis Based on the L_1 -norm and Related Methods. Elsevier, Neuchâtel, 1997, 1997.
- Z. DREZNER, H. W. HAMACHER, *Facility Location: Applications and Theory*, Springer, 2004.
- 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.
- E. M. T. HENDRIX, B. G. TÓTH, *Introduction to Nonlinear and Global Optimization*, Springer, 2010.
- 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.

- F. JARRE, J. STOER, *Optimierung*, Springer Verlag, Berlin Heidelberg, 2004., Springer Verlag, Berlin Heidelberg, 2003.
- D. JUKIĆ, *Mjera i integral*, Sveučilište J. J. Strossmayera u Osijeku, Odjel za matematiku, 2012.
- 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. ŠALKAUSKAS, *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.
- A. MORALES-ESTEBAN, F. MARTÍNEZ-ÁLVAREZ, S. SCITOVSKI, R. SCITOVSKI, *A fast partitioning algorithm using adaptive Mahalanobis clustering with application to seismic zoning*, Computers & Geosciences, **73**(2014) 132–141.

- J. D. MURRAY, *Mathematical Biology: I. An Introduction*, Springer, 2002.
- 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.
- R. PLATO, *Numerische Mathematik kompakt, 4. Auflage*, Vieweg + Teubner, 2010.
- 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, *Grupiranje podataka: klasteri*, Osječki matematički list, **10**(2010) 149–176.
- 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.
- K. SABO, R. SCITOVSKI, *An approach to cluster separability in a partition*, Information Sciences, **305**(2015) 208–218.
- A. A. SAMARSKIJ, A. V. GULIN, *Numeričke metode (na ruskom)*, Nauka, Moskva, 1989.
- M. SCHATZMAN, *Numerical Analysis - a mathematical introduction*, Oxford University Press, Oxford, 2002.
- 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, N. TRUHAR, Z. TOMLJANOVIĆ, *Metode optimizacije*, Odjel za matematiku, Sveučilište u Osijeku, 2014.
- R. SCITOVSKI, K. SABO, *Analysis of the k -means algorithm in the case of data points occurring on the border of two or more clusters*, Knowledge-Based Systems, **57**(2014) 1–7.
- 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.
- J. STOER, R. BULIRSCH, *Numerische Mathematik 2*, Springer, 2005.
- L. N. TREFETHEN, D. BAU, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
- N. TRUHAR, *Numerička linearna algebra*, Odjel za matematiku, Sveučilište u Osijeku, 2010.
- I. VAZLER, K. SABO, R. SCITOVSKI, *Weighted median of the data in solving least absolute deviations problems*, Communications in Statistics - Theory and Methods, **41**(2012) 1455–1465.
- T. ŽIVKOVIĆ, *LU-dekompozicija trodijagonalne i cikličke trodijagonalne matrice*, Osječki matematički list, **12**(2012) 153–168.
- G. A. WATSON, *Approximation Theory and Numerical Methods.*, John Wiley & Sons, Chichester, 1980.

Kazalo

- Čebiševljevi polinomi, 31
- Čvorovi interpolacije, 19
- Čebiševljevi polinomi, 145
- Algoritam
 - Back Substitution, 63, 66
 - Bisekcije, 89
 - Eulerov, 210
 - Eulerova metoda, 210, 233
 - Forward Substitution, 63
 - Gauss-Newtonov, 175
 - Gauss-Seidel, 73
 - Gauss-Seidelov, 229
 - Gram - Schmidtov, 142
 - Gram-Schmidtov, 232
 - Hornerov, 22, 29, 227
 - invertiranje matrice, 64
 - Jacobijev, 73, 229
 - Metoda jednostavnih iteracija, 97
 - Runge-Kutta metoda, 234
 - Runge-Kutta metoda, 216
 - Simetrična formula, 214, 233
 - Simpsonovo pravilo, 200, 232, 233
 - Supstitucija unaprijed, 55
 - Supstitucija unazad, 55
 - Trapezno pravilo, 232
 - trapezno pravilo, 195
- Aproksimacija
 - funkcije, 130
 - matrice, 50, 167
 - najbolja, 130
 - najbolja L_2 aproksimacija, 130
 - najbolja l_2 aproksimacija, 167
 - najbolja L_∞ aproksimacija, 147
 - vektora, 50
- Aritmetička sredina, 154
- Balistika, 112
- Besselova nejednakost, 136
- Brzina konvergencije, 91
- Cauchyjev problem, 207
- Dekompozicija
 - Cholesky, 64, 65, 83, 228
 - LU, 62, 228
 - LU-dekompozicija, 59
 - na singularne vrijednosti, 76, 165
 - QR-dekompozicija, 66, 164
- Diferencijalna jednačba
 - obična, 207
- Floating-point aritmetika, 3, 48
- Gaussova metoda eliminacije, 55, 60
- Givensove rotacije, 67
- Gram-Schmidtov postupak ortogonalizacije, 140
- Hat-funkcija, 34
- Hesijan, 160
- Hornerova shema, 27, 29, 44
- Householderove transformacije, 67

- Interpolacija, 19
- Interpolacijski polinom, 20
 - Lagrangeov oblik, 23
 - Newtonov oblik, 26
 - ocjena pogreške, 29
- Jacobijan, 160
- Kontrakcija, 74
- Lipschitz
 - derivabilna funkcija, 113
 - neprekidna funkcija, 99
- Matrica
 - determinanta matrice, 63
 - dijagonalno dominantna, 38, 57, 76
 - elementarna, 60
 - glavni minori, 57
 - Gramova, 131
 - Hilbertova, 59, 66, 132
 - Householderova, 67
 - inverzna, 63
 - permutacija, 61
 - pozitivno definitna, 57
 - pseudoinverzna, 163
 - rijetko popunjena, 64, 71
 - trodiagonalna, 38, 64
 - vrpčasta, 82
- Metoda
 - bisekcije, 87, 230
 - Brentova, 128
 - Croutova, 61
 - diskretizacije, 220
 - Eulerova, 209
 - Gauss-Jordanova, 57
 - Gauss-Newtonova, 174
 - Gauss-Seidelova, 72
 - Gaussova metoda eliminacije, 55, 60
 - Heunova, 215
 - Jacobijeva, 72
 - jednostavnih iteracija, 92, 230
 - Marquardtova, 180
 - Newtonova, 103, 169, 231
 - regula falsi, 111
 - Runge-Kutta, 215
 - sekanti, 110
 - simetrična formula, 213
 - SOR, 76
- Minimizacija
 - jednodimenzionalna, 128, 169
 - višedimenzionalna, 170
- Model
 - Lotka–Volterra, 220
- Nelinearna jednadžba, 85
- Newton-Cotesova formula, 196
- Newtonova
 - metoda minimizacije, 171
- Norma, 48
 - l_1 norma, 49
 - Čebiševljeva L_∞ norma, 129
 - Čebiševljeva l_∞ norma, 49
 - Euklidska l_2 norma, 49
 - Frobeniusova, 50
 - matrična, 49
 - matrična inducirana, 50
 - vektorska, 48
- Nultočka funkcije, 86
- Ocjena pogreške, 90, 96, 104
- Ortogonalan sustav, 132
 - Fourierov sustav, 137
- Ortogonalni polinomi, 140
 - Čebiševljevi polinomi, 145

- Legendreovi polinomi, 143
- Pivot-element, 55
- Podijeljene razlike, 28
- Pogreška
 - apsolutna, 5
 - inverzni problem, 13
 - metode, 2
 - modela, 2
 - relativna, 6
 - strojna, 3
 - zaokruživanja, 1, 10
- Pouzdanе znamenke, 9
- Problem najmanjih kvadrata, 153
 - linearni, 155, 161
 - nelinearni, 158, 169
- Separacija nultočki, 85
- Sherman-Morrisonova lema, 61
- Signifikantne znamenke, 8
- Simpsonovo pravilo, 197
 - generalizirano, 198
- Singularne vrijednosti matrice, 53, 77
- Singularni vektori matrice, 77
- Slobodni pad, 217
- Spline, 32
 - kubični interpolacijski, 36
 - linearni interpolacijski, 33
 - ocjena pogreške, 34
- Strategija
 - parcijalnog pivotiranja, 57
 - potpunog pivotiranja, 58
- Sustav
 - trokutasti, 227
- Sustav linearnih jednadžbi
 - diferencijskih, 221
- Sustav linearnih jednadžbi, 47
 - iterativne metode, 71
 - perturbirani, 51
 - rješavanje unazad, 53
 - supstitucija unaprijed, 54
 - trokutasti sustav, 53
- Sustav nelinearnih jednadžbi, 115
 - BFGS metoda, 123
 - Broydenova metoda, 122
 - DFP metoda, 122
 - kvazi-Newtonova metoda, 122
 - metoda jednostavnih iteracija, 117
 - Newtonova metoda, 119
- Sustav normalnih jednadžbi, 131, 162
- Svojstvene vrijednosti matrice, 77
- Svojstveni vektori matrice, 77
- Sylvesterov kriterij, 57
- Težine podataka, 160
- Teorem
 - Banachov o fiksnoj točki, 94, 102
 - Banachov o fiksnoj točki, 74
 - Chebyshev's Equioscillatio, 150
 - Pitagorin, 136
- Trapezno pravilo, 36, 192, 215
 - generalizirano, 194
- Uvjetovanost, 51
 - broj uvjetovanosti, 52, 78
- Vandermondova determinanta, 21