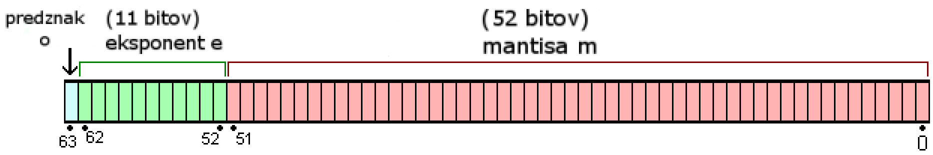
# Poglavje 1

# Plavajoča vejica




Slika 1.1: Plavajoča vejica

Zapis je oblike $(-1)^o(1+m)2^{e-1023}$, mantisa je v normalizirani obliki, eksponent je podan z zamikom. Več lahko najdete na tej strani.

**Naloga 1.1** *Zapiši naslednja števila v dvojni natančnosti:*

*(i).* 2.7185*,*

*(ii).* $-0.1$*,*

*(iii).* 255.125*.*

*Rešitev.*

(i). Število najprej pretvorimo v dvojiški zapis. Lotimo se celega dela.

$$\begin{aligned} 2 &= 1*2+0 \\ 1 &= 0*2+1 \end{aligned}$$

$2 = 10_{(2)}$. Potem pa še decimalni del.

$$\begin{aligned} 0.71875*2 &= 1+0.4375 \\ 0.4375*2 &= 0+0.875 \\ 0.875*2 &= 1+0.75 \\ 0.75*2 &= 1+0.5 \\ 0.5*2 &= 1+0 \end{aligned}$$

Torej je $2.71875 = 11.10111_{(2)}$. Ker iščemo zapis normalizirane oblike $(-1)^o(1+m)2^{e-1023}$, moramo število deliti/množiti z dva dokler ne dobimo 1.m. Torej $11.10111_{(2)} = 1.11011_{(2)} * 2$ in $e - 1023 = 1$. Dobili smo $o = 0$, $m = 110111\underbrace{0\ldots0}_{46}$, $e = 10000000000_{(2)}$.

(ii). Število pretvorimo v dvojiški zapis.

$$
\begin{aligned}
0.1 * 2 &= 0 + 0.2 \\
0.2 * 2 &= 0 + 0.4 \\
0.4 * 2 &= 0 + 0.8 \\
0.8 * 2 &= 1 + 0.6 \\
0.6 * 2 &= 1 + 0.2 \\
0.2 * 2 &= 0 + 0.4 \\
0.4 * 2 &= 0 + 0.8 \\
0.8 * 2 &= 1 + 0.6...
\end{aligned}
$$

$0.1 = 0.000\overline{1100} = 1.100\overline{1100} * 2^{-4}$. Mantisa je dolžine 52, torej moramo na ustreznem mestu odrezati. Dobimo $o = -1$, $m = 100\underbrace{1100}_{11}11010$. Iz $e - 1023 = -4$, dobimo $e = 1019$, torej je $e = 01111111011$. Zadnje števke v mantisi so take zaradi (unbiased) zaokroževanja na najbližje število, ki ima zadnji bit enak 0. Število, ki je najbližje $0. * 1100\underbrace{1}100$ in ima na označem mestu 0, je $0. * 11010$. Tukaj $*$ označuje preostale decimalke.

(iii). Število pretvorimo v dvojiški zapis. Celi del:

$$
\begin{aligned}
255 &= 2 * 127 + 1 \\
127 &= 2 * 63 + 1 \\
&\vdots \\
3 &= 2 * 1 + 1 \\
1 &= 0 * 2 + 1
\end{aligned}
$$

$255 = 11111111_{(2)}$. Še decimalni del:

$$
\begin{aligned}
0.125 * 2 &= 0 + 0.25 \\
0.25 * 2 &= 0 + 0.5 \\
0.5 * 2 &= 1 + 0
\end{aligned}
$$

$0.125 = 0.001_{(2)}$. Dobili smo $255.125 = 11111111.001_{(2)}$. Premaknemo decimalno piko, da dobimo normalizirano obliko. Dobimo $11111111.001_{(2)} = 1.1111111001 * 2^7$. Torej je $o = 0$, $m = 1111111001\underbrace{0\ldots0}_{42}$. Velja še $e - 1023 = 7$, torej je $e = 10000000110_{(2)}$.

■

# Poglavje 2

# Numerična stabilnost

Algoritem je direktno stabilen, če vrne rezultat, ki se malo razlikuje od prave vrednosti. Algoritem je direktno stabilen, če za izračunan rezultat obstajajo taki malo zmoteni podatki, da iz njih s točnim izračunom dobimo izračunano vrednost.

**Naloga 2.1** *Vrednost $z = x^2 - y^2$ računamo na dva načina:*

*(i).* $z = x^2 - y^2$

*(ii).* $z = (x - y)(x + y)$

*Analiziraj algoritma. Oceni relativno napako $\frac{|\widehat{z}-z|}{|z|}$. Ali je kateri od obeh algoritmov direktno/obratno stabilen?*

*Rešitev.*

(i). Imamo $z = x * x - y * y$, $\widehat{a} = x * x(1 + \alpha)$, $\widehat{b} = y * y(1 + \beta)$, $\widehat{z} = (\widehat{a} - \widehat{b})(1 + \gamma)$, kjer je $|\alpha|, |\beta|, |\gamma| \leq u$, $u$ je relativna natančnost. Sledi, da je

$$\widehat{z} = x^2 \overbrace{(1 + \alpha)(1 + \gamma)}^{(1+\delta_1)} - y^2 \overbrace{(1 + \beta)(1 + \gamma)}^{(1+\delta_2)}.$$

Iz ocene $1 - 2u + u^2 = (1 - u)^2 \leq (1 + \delta_1) \leq (1 + u)^2 = 1 + 2u + u^2$, pri majhnem $u$ dobimo $\delta_1, \delta_2 \leq 2 * u$. Ocenimo izraz

$$|\widehat{z} - z| = |x^2\delta_1 - y^2\delta_2| \leq x^2|\delta_1| + y^2|\delta_2| \leq 2u(x^2 + y^2).$$

Torej velja ocena:

$$\frac{|\widehat{z} - z|}{|z|} \leq 2u\frac{x^2 + y^2}{x^2 - y^2}.$$

Iz tega vidimo, da ta algoritem ni direktno stabilen, saj lahko pri $x$ in $y$, ki sta si blizu, dobimo veliko napako. Algoritem je obratno stabilen. Če definiramo $\widehat{x} = x\sqrt{(1 + \delta_1)}$ in $\widehat{y} = y\sqrt{(1 + \delta_2)}$, ki sta blizu $x$ in $y$, ter predpostavimo, da je računanje točno, dobimo zmoten izraz $\widehat{z} = \widehat{x}^2 - \widehat{y}^2$.

(ii). Oglejmo si še $z = (x - y)(x + y)$. Definirajmo $\widehat{a} = (x - y)(1 + \alpha)$, $\widehat{b} = (x + y)(1 + \beta)$ ter $\widehat{z} = \widehat{a}\widehat{b}(1 + \gamma)$, kjer je $|\alpha|, |\beta|, |\gamma| \leq u$. Izraz je enak

$$\widehat{z} = (x - y)(x + y) \overbrace{(1 + \alpha)(1 + \beta)(1 + \gamma)}^{(1+\delta)},$$

kjer s podobno oceno kot v prejšnji točki dobimo $|\delta| \leq 3u$. Torej velja ocena $|\widehat{z}-z| \leq 3u|z|$. Na koncu dobimo $\frac{|\widehat{z}-z|}{|z|} \leq 3u$. Algoritem je direktno stabilen. Prav tako je obratno stabilen, iskana zmotena $x$ in $y$ sta recimo: $\widehat{x} = x\sqrt{(1+\delta)}$, $\widehat{y} = y\sqrt{(1+\delta)}$.

■

**Naloga 2.2** *Pokaži, da za računanje vsote in produkta kompleksnih števil $x, y \in \mathbb{C}$ v premični piki velja:*

*(i).* $fl(x+y) = (x+y)(1+\delta)$, $|\delta| \leq u$

*(ii).* $fl(xy) = xy(1+\delta)$, $|\delta| \leq \sqrt{2}u(2+u)$.

*Kompleksna števila si predstavljamo kot pare realnih števil.*

*Rešitev.*

(i). Vsota $x = a + bi$ in $y = c + di$ je $x + y = (a+c) + (b+d)i$. Dobimo

$$fl(x+y) = (a+c)(1+\delta_1) + i(b+d)(1+\delta_2) = (a+c) + i(b+d) + (a+c)\delta_1 + i(b+d)\delta_2,$$

kjer je $|\delta_1|, |\delta_2| \leq u$. Po drugi strani je $fl(x+y) = (x+y)(1+\delta) = \delta(x+y) + i\delta(x+y)$. Ko te dve enakosti primerjamo, dobimo $(x+y)\delta = (a+c)\delta_1 + i(b+d)\delta_2$. Na obeh straneh izračunamo kvadrat absolutne vrednost in dobimo:

$$|x+y|^2\delta^2 = (a+c)^2\delta_1^2 + (b+d)^2\delta_2^2 \leq u^2((a+c)^2 + (b+d)^2) = u^2|x+y|^2.$$

Iz tega sledi $|\delta| \leq u$.

(ii). Poglejmo si $fl(xy) = xy(1+\delta)$. Produkt je enak $xy = (ac - bd) + i(ad + bc)$. Torej je:

$$\begin{aligned} fl(xy) = (ac(1+\delta_1) - bd(1+\delta_2))\,(1+\delta_3) + i\,(ad(1+\delta_4) + bc(1+\delta_5))\,(1+\delta_6) \\ = ac(1+\epsilon_1) - bd(1+\epsilon_2) + i\,(ad(1+\epsilon_3) + bc(1+\epsilon_4)) \\ = ac - bd + i(ad+bc) + (\epsilon_1 ac - \epsilon_2 bd) + i(ad\epsilon_3 + bc\epsilon_4). \end{aligned}$$

Torej smo dobili enakost $xy\delta = (\epsilon_1 ac - \epsilon_2 bd) + i(ad\epsilon_3 + bc\epsilon_4)$. Spet izračunamo kvadrat absolutne vrednosti na obeh straneh. Dobimo

$$|xy|^2|\delta|^2 = (ac\epsilon_1 - bd\epsilon_2)^2 + (ad\epsilon_3 + bc\epsilon_4)^2 =$$

$$(ac)^2\epsilon_1^2 + (bd)^2\epsilon_2^2 + (ad)^2\epsilon_3^2 + (bc)^2\epsilon_4^2 + \underline{-2acbd\epsilon_1\epsilon_2} + \underline{2abcd\epsilon_3\epsilon_4}.$$

Če upoštevamo še $q^2 + w^2 \geq 2qw$ in $q^2 + w^2 \geq -2qw$ v podčrtanih izrazih ter da so vsi epsiloni pod $u(2+u)$ (oz. 2u za majhen u), dobimo:

$$|xy|^2|\delta|^2 \leq 2\left((ac)^2\epsilon_1^2 + (bd)^2\epsilon_2^2 + (ad)^2\epsilon_3^2 + (bc)^2\epsilon_4^2\right)$$

$$\leq \left((ac-bd)^2 + (ad+bc)^2\right) 2\,(u(2+u))^2 = |xy|^2 2(u(2+u))^2.$$

Iz tega pa že sledi $|\delta| \leq \sqrt{2}u(2+u) \doteq 2\sqrt{2}u$.

■

# Poglavje 3

# Matrične norme

Matrična norma je preslikava $||.|| : \mathbb{C}^{n\times n} \to \mathbb{R}^+$, za katero velja

(i). $||A|| \geq 0,\ ||A|| = 0 \Leftrightarrow A = 0$,

(ii). $||\alpha A|| = |\alpha|||A||$,

(iii). $||A + B|| \leq ||A|| + ||B||$,

(iv). $||AB|| \leq ||A|| \cdot ||B||$, submultiplikativnost.

Za poljubni matriki $A$ in $B$ in poljuben $\alpha \in \mathbb{C}$.

**Naloga 3.1** *Pokaži, da je*

$$||A|| = \sup_{||x||=1} ||Ax|| = \sup_{x\neq 0} \frac{||Ax||}{||x||}$$

*matrična norma. Tukaj je $||.||$ poljubna vektorska norma.*

*Rešitev.* Očitno velja $||A|| \geq 0$. Če velja $||A|| = \sup_{||x||=1} ||Ax||$, sledi da je $Ay = 0$ za vsak $y$, torej je $A = 0$. Velja tudi

$$||\alpha A|| = \sup_{||x||=1} ||\alpha Ax|| = \sup_{||x||=1} |\alpha|||Ax|| = |\alpha| \sup_{||x||=1} ||Ax|| = |\alpha|||A||.$$

Podobno izračunamo

$$\begin{aligned} ||A + B|| = \sup_{||x||=1} ||(A + B)x|| = \sup_{||x||=1} ||Ax + Bx|| &\leq \sup_{||x||=1} (||Ax|| + ||Bx||) \\ &\leq \sup_{||x||=1} ||Ax|| + \sup_{||x||=1} ||Bx|| = ||A|| + ||B||. \end{aligned}$$

Lotimo se še zadnjega pogoja, definirajmo še $y = Bx$.

$$\begin{aligned} ||AB|| &= \sup_{||x||=1} ||(AB)x|| = \sup_{||x||=1, Bx\neq 0} \frac{||(AB)x||}{||Bx||} ||Bx|| \\ &= \sup_{||x||=1, y=Bx, y\neq 0} ||A\left(\frac{y}{||y||}\right)|| \cdot ||Bx|| \\ &\leq \sup_{||x||=1, y=Bx, y\neq 0} ||A\left(\frac{y}{||y||}\right)|| \cdot \sup_{||x||=1} ||Bx|| \\ &= \sup_{||x||=1, y=Bx, y\neq 0} ||A\frac{y}{||y||}|| \cdot ||B|| \leq \sup_{||y||=1} ||Ay|| \cdot ||B|| \end{aligned}$$

$$= ||A|| \cdot ||B||.$$

S tem smo končali. Upoštevali smo, da je supremum vsote pozitivnih števil manjši od vsote supremumov. To prav tako velja za množenje. Zadnji neenačaj dobimo, ker je supremum po večji množici kvečjemu večji. ∎

**Naloga 3.2** *Pokaži naslednje.*

*a) Dokaži, da*

$$N_\infty(A) = \max_{1\leq i,\, j\leq n} |a_{ij}|$$

*ni matrična norma.*

*b) Dokaži, da velja* $||A||_F^2 = sl(A^H A) = \sum_{i=1}^n \lambda_i(A^H A)$, *kjer so* $\lambda_i$ *lastne vrednosti* $A^H A$.

*Rešitev.* a)
Da so prve tri lastnosti matrične norme izpolnjene hitro preverimo. Matrika ima normo nič, natanko takrat ko je njen največji element po absolutni vrednosti enak 0. Če matriko množimo s skalarjem, se največji element matrike množi z absolutno vrednostjo tega skalarja. Pri tretji lastnosti uporabimo trikotniško neenakost in dejstvo, da je maksimum vsote pozitivnih števil manjši od maksimuma števil ih vsote po posameznih številih. Definirajmo matriki $A$ in $B$ takole

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad A\cdot B = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Velja $N_\infty(A) = 1,\ N_\infty(B) = 1,\ N_\infty(AB) = 2$. Veljati bi moralo $N_\infty(AB) \leq N_\infty(A)N_\infty(B)$, vendar $2 \nleq 21$. Submultiplikativnost ni izpolnjena.
b)
Za $B = A^H A$ velja $b_{ii} = \sum_{k=1}^n \overline{a_{ki}} a_{ki} = \sum_{k=1}^n |a_{ki}|^2$. Tako za sled $B$ dobimo

$$sled(A^H A) = \sum_{i=1}^n \left( \sum_{k=1}^n |a_{ki}|^2 \right) = ||A||_F^2.$$

Matrika $A^H A$ je simetrična, torej se da diagonalizirati. Podobna je matriki z lastnimi vrednostmi na diagonali. Sled podobne matrike je enaka sledi prvotne matrike, tako dobimo $||A||_F^2 = sl(A^H A) = \sum_{i=1}^n \lambda_i(A^H A)$. ∎

**Naloga 3.3** *Pokaži naslednje:*

*a)* $\frac{1}{\sqrt{n}}||A||_F \leq ||A||_2 \leq ||A||_F$

*b)* $\frac{1}{\sqrt{n}}||A||_\infty \leq ||A||_2 \leq \sqrt{n}||A||_\infty$

*c)* $\frac{1}{\sqrt{n}}||A||_1 \leq ||A||_2 \leq \sqrt{n}||A||_1$

*d)* $N_\infty(A) \leq ||A||_2 \leq nN_\infty(A)$

*Rešitev.* a)
Vemo, da velja

$$||A||_2 = \max_{i=1,\ldots,n} \sqrt{\lambda_i(A^H A)} = \sigma_1(A) \text{ in } ||A||_F^2 = \sum_{i=1}^{n} \lambda_i(A^H A).$$

Lastne vrednosti $A^H A$ so nenegativne, saj velja $< A^H Ax, x >=< \mu x, x >= \mu =< Ax, Ax >\geq 0$, $\lambda_i = \sigma_i^2$. Razporedimo jih v zaporedje $\sigma_1^2 \geq \sigma_2^2 \ldots \geq \sigma_n^2 \geq 0$. Pozitivni koreni teh lastnih vrednosti $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$ so singularne vrednosti matrike $A$. Očitno je $||A||_2 \leq ||A||_F$, saj je $||A||_2$ enaka $\sigma_1(A)$, kar je največja singularna vrednost matrike $A$. Poglejmo si $||A||_F^2 = \sum_{i=1}^n \sigma_i^2 \leq n\sigma_1^2 = n||A||_2$, kar pomeni $\frac{1}{\sqrt{n}}||A||_F \leq ||A||_2$. Neenačaj dobimo, ker je $\sigma_1(A)$ največja singularna vrednost.
b)
Za vektorske norme velja

$$||x||_\infty = \max_{1\leq i\leq n} |x_i| \leq \sqrt{|x_1|^2 + \ldots |x_n|^2} = ||x||_2 \leq \sqrt{n \max_i |x_i|^2} = \sqrt{n}||x||_\infty,$$

to je $||x||_\infty \leq ||x||_2 \leq \sqrt{n}||x||_\infty$. Razmisli zakaj neeneakosti držijo, vsi sklepi so enostavni. Iz tega dobimo

$$||A||_2 = \max_{x\neq 0} \frac{||Ax||_2}{||x||_2} \leq \max_{x\neq 0} \frac{\sqrt{n}||Ax||_\infty}{||x||_\infty} = \sqrt{n}||A||_\infty$$

in

$$||A||_2 = \max_{x\neq 0} \frac{||Ax||_2}{||x||_2} \geq \max_{x\neq 0} \frac{||Ax||_\infty}{\sqrt{n}||x||_\infty} = \frac{1}{\sqrt{n}}||A||_\infty.$$

c)
Velja $||A||_1 = ||A^H||_\infty$ in $||A||_2 = ||A^H||_2$. Iz tega že sledi neenakost.
d)
Iz a) vemo

$$||A||_2 \leq ||A||_F = \sqrt{\sum_{i,j} |a_{ij}|^2} \leq \sqrt{n^2 \max_{i,j} |a_{ij}|^2} = n \max_{i,j} |a_{ij}| = nN_\infty(A).$$

Za drugi del neenakosti upoštevamo $a_{ij} = e_i^T Ae_j$. Računajmo

$$|a_{ij}| = |e_i^T Ae_j| \leq ||e_i||_2||Ae_j||_2 = \frac{||Ae_j||_2}{||e_j||_2} \leq ||A||_2.$$

Prvi neenačaj dobimo po Cauchy-Schwartzovi neenakosti, zadnjega pa po definiciji norme $||A||_2$.

∎

**Naloga 3.4** *Izračunaj* $||\ ||_1$, $||\ ||_\infty$, $||\ ||_F||$ *za* $\begin{bmatrix} 2 & -1 & 3 \\ 5 & 4 & 1 \\ -2 & -1 & 2 \end{bmatrix}$ *in oceni* $||\ ||_2$.

*Rešitev.* Izračunajmo

$$\begin{aligned}
||A||_1 &= \max\{|2|+|5|+|-2|,\ |-1|+|4|+|-1|,\ |3|+|1|+|2|\} = 9,\\
||A||_\infty &= \max\{|2|+|-1|+|3|,\ |5|+|4|+|1|,\ |-2|+|-1|+|2|\} = 10,\\
||A||_F &= \sqrt{4+1+9+25+16+1+4+4} = \sqrt{65} \doteq 8.06226.
\end{aligned}$$

Ocenimo $||A||_2$

$$\begin{aligned}
\frac{1}{\sqrt{3}}||A||_F \le ||A||_2 \le ||A||_F &\Rightarrow 4.65475 \le ||A||_2 \le 8.06226\\
\frac{1}{\sqrt{3}}||A||_\infty \le ||A||_2 \le \sqrt{3}||A||_\infty &\Rightarrow 5,7735 \le ||A||_2 \le 17,3205\\
\frac{1}{\sqrt{3}}||A||_1 \le ||A||_2 \le \sqrt{3}||A||_1 &\Rightarrow 5,19615 \le ||A||_2 \le 15,58846\\
N_\infty(A) \le ||A||_2 \le 3N_\infty(A) &\Rightarrow 5 \le ||A||_2 \le 15.
\end{aligned}$$

Skupaj dobimo oceno $5.7735 \le ||A||_2 \le 8.06226$. Boljšo oceno dobimo, če poskusimo oceniti spektralni radij matrike

$$B = A^H A = \begin{bmatrix} 2 & 5 & -2\\ -1 & 4 & -1\\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3\\ 5 & 4 & 1\\ -2 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 33 & 20 & 7\\ 20 & 8 & -1\\ 7 & -1 & 14 \end{bmatrix}.$$

Za vsako lastno vrednost in vsako normo velja $\lambda_i(B) \le ||B|_F = 50.0899$. Vemo, da velja

$$||A||_2 \le \sqrt{||B||_F} = 7.0774.$$

Oceno navzdol dobimo, če upoštevamo $||A||_2 = ||A^T||_2$ in $||Ae_i||_2 \le ||A||_2||e_i||_2 = ||A||_2$, zadnja enakost sledi iz definicije matrične norme. Vektorji $Ae_i$ so ravno i-ti stolpci, $A^Te_i$ so i-te vrstice. Norme vrstic so $5.7446,\ 4.2426, 3.7417$. Norme stolpcev so $3.7417,\ 6.4807,\ 3$. Torej dobimo $||A||_2 \ge 6.4807$. Končna ocena je $6.4809 \le ||A||_2 \le 7.0774$. Ukaz iz Matlaba vrne $\mathtt{norm}(A) = 6.9044$. ■

**Naloga 3.5** *Izračunaj* $||A||_1$, $||A||_\infty$, $||A||_F$ *in čim natančneje oceni* $||A||_2$ *na obe strani, če je*

$$\begin{aligned}
a_{ii} &= -2i,\ i = 1,\ 2, \ldots,\ n\\
a_{i+1,i} &= a_{i,i+1} = n-i,\ i = 1,\ 2, \ldots,\ n,
\end{aligned}$$

*drugi elementi so enaki* 0.

$$A = \begin{bmatrix} -2 & n-1 & & & \\ n-1 & -4 & n-2 & & \\ & n-2 & -6 & & \\ & & \ddots & \ddots & 1\\ & & & 1 & -2n \end{bmatrix}.$$

*Upoštevaj, da velja*

$$\sum_{i=1}^{n-1} i^2 = \frac{(n-1)n(2n-1)}{6}.$$

*Rešitev.* Pri $i = 1$ dobimo za vsoto absolutnih vrednosti v stolpcu $n+1$, kar očitno ni maksimum. Izračunajmo

$$\begin{aligned} ||A||_1 &= \max_{i\neq 1}\{|-2i| + |n-i| + |n-i+1|\} = \max_{i\neq 1}\{2i + n - i + n - i + 1\} \\ &= \max_{i\neq 1}\{2n+1\} = 2n+1. \end{aligned}$$

Ker je matrika simetrična velja $||A||_1 = ||A^H||_\infty = ||A||_\infty = 2n+1$. Frobeniusova norma je

$$||A||_F^2 = \sum_{i=1}^{n}(2i)^2 + 2\sum_{i=1}^{n} i^2 = 6\sum_{i=1}^{n-1} i^2 + 4n^2 = 6\frac{(n-1)n(2n-1)}{6} + 4n^2 =$$
$$n(n-1)(2n-1) + 4n^2 = n(2n^2-3n+1) + 4n^2 = 2n^3 + n^2 + n.$$

Poglejmo si ocene za drugo normo

$$\begin{aligned} \frac{1}{\sqrt{n}}||A||_F \leq ||A||_2 \leq ||A||_F \Rightarrow \\ \frac{1}{\sqrt{n}}\sqrt{2n^3+n^2+n} \leq ||A||_2 \leq \sqrt{2n^3+n^2+n} \\ \frac{1}{\sqrt{n}}||A||_\infty \leq ||A||_2 \leq \sqrt{n}||A||_\infty \Rightarrow \frac{1}{\sqrt{n}}(2n+1) \leq ||A||_2 \leq \sqrt{n}(2n+1) \\ N_\infty(A) \leq ||A||_2 \leq nN_\infty(A) \Rightarrow 2n \leq ||A||_2 \leq 2n^2. \end{aligned}$$

Upoštevamo še, da velja $||A||_2^2 \leq ||A||_1||A||_\infty$. Iz tega dobimo

$$||A||_2^2 \leq (2n+1)^2,$$

kar pomeni $||A||_2 \leq 2n+1$. Za drugo normo smo dobili oceno

$$2n \leq ||A||_2 \leq 2n+1.$$

■

# Poglavje 4

# Reševanje linearnih sistemov

Pogojenostno število matrike, $\mathrm{cond}(A) = ||A||_2||A^{-1}||_2$ nam pove kako občutljivo je reševanje sistema. Da lahko dobimo poljubno občutljive matrike v praksi, pokaže naslednja naloga.

**Naloga 4.1** *Iščemo koeficiente polinoma* $p(x) = a_1 + a_2x + \cdots + a_nx^{n-1}$*, ki na* $[0, 1]$ *aproksimira zvezno funkcijo* $f$ *tako, da je napaka*

$$E = \int_0^1 (f(x) - p(x))^2 dx$$

*minimalna. Torej mora veljati* $\frac{\partial E}{\partial a_i} = 0$ *za* $i = 1, \ldots, n$*, kjer je*

$$\frac{\partial E}{\partial a_i} = 2\int_0^1 (p(x) - f(x))x^{i-1}dx.$$

*Definirajmo vektor*

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} \qquad in; \quad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

*Od tod dobimo*

$$F_i = \int_0^1 f(x)x^{i-1}dx = \int_0^1 \left(\sum_{j=1}^n a_jx^{j-1}x^{i-1}\right) dx = \sum_{j=1}^n a_j \int_0^1 x^{i+j-2}dx = \sum_{j=1}^n a_j \frac{1}{i+j-1}.$$

*Dobimo sistem s Hilbertovo matriko* $H_n$, $H_{ij} = \frac{1}{i+j-1}$*. Kar je sistem* $Ha = F$*. Primer za* $n = 5$ *je*

$$H_5 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix}.$$

*Hilbertove matrike so primer zelo občutljivih matrik. Pogojenostno število* $H_5$ *je recimo približno* $4.766 \cdot 10^5$*, kar lahko izračunamo z ukazom cond v Matlabu. Pogojenostna števila matrik* $H_n$ *hitro rastejo.*

*Izkaže se, da smo dobili občutljiv sistem, ker smo vzeli standardno bazo polinomov stopnje* $n$*. Za stabilno računanje moramo vzeti ortogonalno bazo polinomov.*

**Izrek 4.1** *Naslednji trditvi sta ekvivalentni*

*(i). Obstaja enolični razcep $A = LU$, kjer je $L$ spodnje trikotna matrika z enicami na diagonali, $U$ pa nesingularna zgornje trikotna matrika.*

*(ii). Vse vodilne podmatrike $A(1:k, 1:k)$ so nesingularne.*

---

***Algoritem 1**: LU razcep brez pivotiranja*

---

**for** $j = 1, \ldots, n-1$ **do**
  **for** $i = 1, \ldots, n$ **do**
    $l_{ij} = \frac{a_{ij}}{a_{jj}}$;
    **for** $k = j+1, \ldots, n$ **do**
      $a_{ik} = a_{ik} - l_{ij} a_{jk}$;
    **end**
  **end**
**end**
*Število operacij je* $\frac{2}{3}n^2 + O(n^2)$.

---

**Izrek 4.2** *Če je $A$ nesingularna, potem obstaja taka permutacijska matrika $P$, da obstaja LU razcep $PA = LU$.*

---

***Algoritem 2**: LU razcep z delnim pivotiranja*

---

$L = 0$;
$P = I$;
**for** $j = 1, \ldots, n-1$ **do**
  *Poišči* $|a_{qj}| = \max_{j \le p \le n} |a_{pj}|$;
  *Zamenjaj vrstici $q$ in $j$ v $L$, $A$ in $P$;*
  **for** $i = j, \ldots, n$ **do**
    $l_{ij} = \frac{a_{ij}}{a_{jj}}$;
    **for** $k = j+1, \ldots, n$ **do**
      $a_{ik} = a_{ik} - l_{ij} a_{jk}$;
    **end**
  **end**
**end**
*Število operacij je* $\frac{2}{3}n^3 + O(n^2)$.

---

**Naloga 4.2** *Z LU-razcepom brez pivotiranja reši sistem $Ax = b$, kjer je*

$$A = \begin{bmatrix} 2 & 3 & -4 \\ -6 & -11 & 13 \\ 4 & 0 & -8 \end{bmatrix}, \; b = \begin{bmatrix} -11 \\ 38 \\ -16 \end{bmatrix}.$$

*Zapiši matriki $L$, $U$ in vektor $y$, ki ga dobiš pri računanju.*

*Rešitev.*
Lotimo se Gaussove eliminacije:

$$\begin{bmatrix} 2 & 3 & -4 \\ -6 & -11 & 13 \\ 4 & 0 & 8 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & -4 \\ 0 & -2 & 1 \\ 0 & -6 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & -4 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix} = U$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ -3 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = L.$$

Izračunamo lahko tudi v bolj kompaktni obliki. Zgornji trikotnik končne matrike vsebuje matriko $U$, spodnji trikotnik brez diagonale pa matriko $L$ brez diagonale. Upoštevamo, da ima matrika $L$ na diagonali enice.

$$\begin{bmatrix} 2 & 3 & -4 \\ -6 & -11 & 13 \\ 4 & 0 & -8 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 3 & -4 \\ -3 & -2 & 1 \\ 2 & -6 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 3 & -4 \\ -3 & -2 & 1 \\ 2 & 3 & -3 \end{bmatrix},$$

zgornji trikotnik je $U$, spodnji trikotnik brez diagonale je matrika $L$ brez diagonale.

Rešimo sistem $Ax = b$, $L(Ux) = b$, $y = Ux$.

---

**Algoritem 3**: Prema substitucija, $Ly = b$

---

**for** $i = 1, 2, \ldots, n$ **do**
  $y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k$;
**end**

---

$$\begin{bmatrix} 1 & & \\ -3 & 1 & \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 38 \\ -16 \end{bmatrix}, \qquad \begin{aligned} y_1 &= -11 \\ y_2 &= 38 + 3 \cdot (-11) = 5 \\ y_3 &= -16 - 2 \cdot (-11) - 35 = -9 \end{aligned}$$

---

**Algoritem 4**: Obratna substitucija, $Ux = y$

---

**for** $i = n, n-1, \ldots, 1$ **do**
  $x_i = \frac{1}{u_{ii}} \left( y_i - \sum_{k=i+1}^{n} u_{ik} x_k \right)$;
**end**

---

$$\begin{bmatrix} 2 & 3 & -4 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 5 \\ -9 \end{bmatrix}, \qquad \begin{aligned} x_3 &= 3 \\ x_2 &= \tfrac{-1}{2}(5 - 3) = -1 \\ x_1 &= \tfrac{1}{2}(-11 + 3 + 12) = 2 \end{aligned}.$$

Torej dobimo $x = \begin{bmatrix} 2 & -1 & 3 \end{bmatrix}^T$. ∎

**Naloga 4.3** *Za matriko*

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 0 & 1 \\ 6 & 1 & 7 \end{bmatrix}$$

*naredi $LU$ razcep z delnim pivotiranjem.*

*Rešitev.* Algoritem poteka takole. V stolpcu $i$ poiščemo največji element, izmed $a_{ji}$, $j \geq i$ po absolutni vrednostni. Vrstico največjega elementa in $i$-to vrstico zamenjamo v $L$, $A$ in $P$. Normalno nadaljujemo z Gaussovo eliminacijo. Na koncu velja $PA = LU$.

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 0 & 1 \\ \underline{6} & 1 & 7 \end{bmatrix} \xrightarrow{1 \updownarrow 3} \begin{bmatrix} 6 & 1 & 7 \\ 0 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} \xrightarrow{G} \begin{bmatrix} 6 & 1 & 7 \\ 0 & 0 & 1 \\ 0 & \underline{1/2} & -5/2 \end{bmatrix} \xrightarrow{2 \updownarrow 3} \begin{bmatrix} 6 & 1 & 7 \\ 0 & 1/2 & -5/2 \\ 0 & 0 & 1 \end{bmatrix} = U$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{1 \updownarrow 3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{G} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1/2 & 0 & 0 \end{bmatrix} \xrightarrow{2 \updownarrow 3} \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightrightarrows \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = L$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{1 \updownarrow 3} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \xrightarrow{2 \updownarrow 3} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = P$$

∎

**Naloga 4.4** *Matrika $A$ je tridiagonalna,*

$$A = \begin{bmatrix} a_1 & c_1 & & & \\ b_1 & a_2 & c_2 & & \\ & b_3 & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ & & & b_n & a_n \end{bmatrix}.$$

- *Kakšna je oblika matrik $L$ in $U$ v $LU$ razcepu brez pivotiranja (s pivotiranjem)?*
- *Pivotna rast matrike je $R = \frac{b}{a}$, kjer je*

  $$a = \max_{1\le i,j\le n} |a_{ij}| \text{ in } b = \max_{1\le i,j\le n} b_{ij}.$$

  *Gledamo koliko se lahko poveča največji element v matriki. Dokaži, da je pri delnem pivotiranju za tridiagonalne matrike $R \le 2$.*
- *Sestavi učinkovit algoritem za izračun $LU$ razcepa tridiagonalne matrike brez pivotiranja (s pivotiranjem).*

*Rešitev.*

- Oblika matrik je naslednja:

  $$L = \begin{bmatrix} l_1 & & & \\ m_1 & l_2 & & \\ & \ddots & \ddots & \\ & & m_{n-1} & l_n \end{bmatrix} \text{ in } U = \begin{bmatrix} u_1 & p_2 & & \\ & \ddots & \ddots & \\ & & u_{n-1} & p_n \\ & & & u_n \end{bmatrix}$$

- Najprej poglejmo, kaj se zgodi s prvima dvema vrsticama. Tako bomo dobili bazo indukcije. Poglejmo si prve tri stolpce prvih dveh vrstic:

  $$\begin{matrix} a_1 & c_1 & 0 \\ b_1 & a_2 & c_2 \end{matrix}.$$

  Imamo dve možnosti. Prva je, da ne pivotiramo, torej velja $|a_1| \ge |b_1|$ in zato je $\frac{|b_1|}{|a_1|} \le 1$. Tako dobimo

  $$\begin{matrix} a_1 & c_1 & 0 \\ 0 & a_2 - \frac{b_1}{a_1}c_1 & c_2 \end{matrix}.$$

  Druga možnost je, da pivotiramo. Ko zamenjamo vrstici, dobimo

  $$\begin{matrix} b_1 & a_2 & c_2 \\ a_1 & c_1 & 0 \end{matrix}.$$

  Po koraku s pivotiranjem dobimo

  $$\begin{matrix} b_1 & a_2 & c_2 \\ 0 & c_1 - \frac{a_1}{b_1}a_2 & -c_2\frac{a_1}{b_1} \end{matrix}.$$

  Podobno kot prej velja $|b_1| < |a_1|$ in zato je $\frac{|a_1|}{|b_1|} < 1$. Dokažimo, da so diagonalni elementi manjši od $2a$, kjer je $a$ največja absolutna vrednost elementov v matriki $A$. V prvem primeru velja

  $$|a_2 - \frac{b_1}{a_1}c_1| \le |a_2| + |\frac{b_1}{a_1}||c_1| \le 2a.$$

V drugem primeru velja

$$|c_1 - \frac{a_1}{b_1}a_2| \leq |c_1| + |\frac{a_1}{b_1}||a_2| < 2a.$$

Pokažima še, da so izvendiagonalni elementi manjši ali enaki $a$. V prvem primeru velja $|a_1| \leq a$ in $|c_2| \leq a$. V drugem primeru velja $0 \leq a$ in

$$|-c_2\frac{a_1}{b_1}| \leq |c_2||\frac{a_1}{b_1}| < a.$$

Naša indukcijska predpostavka je, da je v $i$-tem koraku v $i$-ti vrstici diagonalni element manjši od $2a$, izvendiagonalni element pa manjši od $a$. Imamo sledečo situacija za elemente

$$\begin{matrix} a_1^{(i)} & c_1^{(i)} & 0 \\ b_1^{(i+1)} & a_2^{(i+1)} & c_2^{(i+1)} \end{matrix},$$

kjer je $b_1^{(i+1)} = b_i$, $a_2^{(i+1)} = a_{i+1}$ in $c_2^{(i+1)} = c_{i+1}$. Poleg tega velja še $|a_1^{(i)}| \leq 2a$ in $c_1^{(i)} \leq a$. Naredimo analogen sklep kot za bazo indukcije in s tem dokaz končamo, saj potem velja $b \leq 2a$.

- Algoritma sta prepuščena za domačo nalogo.

■

**Naloga 4.5** *Sestavi ekonomičen algoritem za izračun inverzne matrike spodnje trikotne matrike $L$ z enicami po diagonali in preštej število operacij.*

*Rešitev.* Iz pravila za izračun inverza vidimo, da je inverzna matrika zopet spodnje trikotna z enicami na diagonali. Rešujemo sistem $LY = I$, $Y$ je oblike $Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}$. Dobimo več podproblemov oblike

$$\begin{bmatrix} 1 & & & & \\ * & \ddots & & & \\ l_{j,1} & l_{j,2} \cdots l_{j,j-1} & 1 & & \\ lj+1,1 & l_{j+1,2} \cdots l_{j+1,j} & & \ddots & \\ * & * & * & * & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \\ y_{j,j} \\ \vdots \\ y_{n,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 1 \\ \\ 0 \end{bmatrix} = e_j.$$

Vemo, da je $y_{j,j} = 1$. Za $i > j$ dobimo

$$l_{i,j}y_{j,j} + l_{i,j+1}y_{j+1,j} + \cdots l_{i,i-1}y_{i-1,j} + l_{i,i}y_{i,j} = 0,$$

kar nam da $y_{ij} = -\sum_{k=j}^{i-1} l_{ik}y_{kj}$. Vseh operacij je

---

**for** $j = 1, \ldots, n$ **do**
  $y_{j,j} = 1$;
  **for** $i = j+1, \ldots, n$ **do**
    $y_{i,j} = -\sum_{k=j}^{i-1} l_{ik}y_{kj}$;
  **end**
**end**

---

$$\sum_{j=1}^{n}\sum_{i=j+1}^{n} 2(i-j) = 2\sum_{j=1}^{n}\frac{(n-j)(n-j+1)}{2} = \sum_{k=1}^{n-1} k(k+1) = \sum_{k=1}^{n-1}(k^2+k)$$

$$= \frac{(n-1)n(2n-1)}{6} + \frac{n(n-1)}{2} = \frac{n^3}{3} + O(n^2).$$

■

**Naloga 4.6** *Poišči razcep Choleskega za matriko*

$$A = \begin{bmatrix} 1 & 2 & -2 & 3 \\ 2 & 8 & -2 & 8 \\ -2 & -2 & 14 & -11 \\ 3 & 8 & -11 & 15 \end{bmatrix}.$$

*Kaj to pomeni za matriko $A$?*

*Rešitev.*

**Algoritem 5**: Razcep Choleskega – $A = V \cdot V^T,\ V = ?$

**for** $k = 1, \ldots, n$ **do**
  $v_{kk} = \sqrt{a_{kk} - \sum_{i=1}^{k-1} v_{ki}^2}$;
  **for** $j = k+1, \ldots, n$ **do**
    $v_{jk} = \frac{1}{v_{kk}}\left(a_{jk} - \sum_{i=1}^{k-1} v_{ji} \cdot v_{ki}\right)$;
  **end**
**end**

$$\begin{bmatrix} 1 & 2 & -2 & 3 \\ 2 & 8 & -2 & 8 \\ -2 & -2 & 14 & -11 \\ 3 & 8 & -11 & 15 \end{bmatrix} \rightarrow \begin{bmatrix} \sqrt{1}=1 \\ 2/1 \\ -2/1 \\ 3/1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 2 & \sqrt{8-2^2}=2 \\ -2 & 1/2(-2-(-2\cdot 2)) \\ 3 & 1/2(8-(3\cdot 2)) \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ -2 & 1 & \sqrt{14-(-2)^2-1^2}=3 \\ 3 & 1 & 1/3(-11-(3\cdot -2+1\cdot 1)) \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ -2 & 1 & 3 & 0 \\ 3 & 1 & -2 & \sqrt{15-(3^2+1^2+(-2)^2)}=3 \end{bmatrix} = V.$$

Matrika $A$ je simetrična in pozitivno definitna, saj se je razcep Choleskega izvedel. ■

# Poglavje 5

# Nelinearne enačbe

**Naloga 5.1** *Za funkcijo* $f(x) = x^3 + 1$ *naredi korak bisekcije in tangentne metode. Vzemi* $a = x_0 = -0.9$ *in* $b = -1.2$.

*Rešitev.* Korak bisekcije je $fa = f(-0.9) = (-0.9)^3 + 1 = 0.271$, $fb = f(-1.2) = (-1.2)^3 + 1 = -0.728$, $sign(fafb) = -1$. Zato lahko izvedemo korak bisekcije. Nadaljujmo: $c = (a + b)/2 = -1.05$, $fc = (-1.05)^3 + 1 = -0.157625$, $sign(fafc) < 0$, torej vzamemo $b = c$. To ponavljamo.

Za korak tangentne metode potrebujemo odvod $f'(x) = 3x^2$. Izračunajmo

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = -0.9 - \frac{(-0.9)^3 + 1}{3(-0.9)^2} \doteq 1,0115.$$

Bisekcijo in sekantno metoda uporabljam takrat, ko ne poznamo odvoda ali pa je računanje odvoda zahtevno. ■

**Izrek 5.1** *Naj bo* $\alpha$ *koren enačbe* $x = g(x)$*, naj bo* $g$ *zvezno odvedljiva na* $I = [\alpha - d, \alpha + d]$ *in naj velja* $|g'(x)| \leq m < 1$ *za vsak* $x \in I$*. Potem za vsak* $x_0 \in I$ *zaporedje*

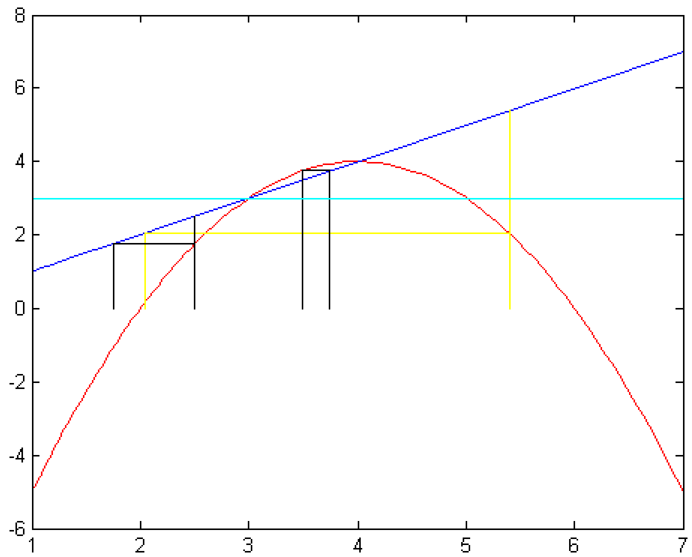
$$x_{r+1} = g(x_r),\ r = 0, 1, \cdots$$

*konvergira k* $\alpha$ *in velja ocena za napako*

$$|x_r - \alpha| \leq \frac{m}{1-m}|x_r - x_{r-1}|.$$

**Izrek 5.2** *Imamo začetno točko* $x_0$ *in interval* $I = [x_0 - d, x_0 + d]$*. Če velja* $|g(x) - g(y)| \leq m|x - y|$ *za* $m < 1$ *in* $|g(x_0) - x_0| \leq (1 - m)d$*, potem zaporedje* $x_r$ *konvergira k* $\alpha$ *in velja* $f(\alpha) = 0$.

**Naloga 5.2** *Za katere začetne približke je navadna iteracija za reševanje* $x = g(x)$*, kjer je* $g(x) = -12 + 8x - x^2$*, konvergentna? Kam konvergira zaporedje? Kakšen je red konvergence? Kje nam konvergenco zagotavlja izrek? Namig: začetne približke išči na intervalu.*




Slika 5.1: Koraki iteracije

*Rešitev.*

(i). Pokažimo, da velja: $x_r < 3 \Rightarrow x_{r+1} < x_r$. Oglejmo si $x_r - x_{r+1}$ in pokažimo, da je izraz pozitiven. Izračunajmo

$$x_r - x_{r+1} = x_r - g(x_r) = x_r - (-12 + 8x_r - x_r^2) = 12 - 7x_r + x_r^2 = (x_r - 4)(x_r - 3).$$

Zadnji izraz v enakosti je v primeru $x_r < 3$ strogo pozitiven.

(ii). Velja tudi: $x_0 > 5 \Rightarrow x_1 < 3$. Oglejmo si izraz $3 - x_1$, to nam da

$$3 - x_1 = 15 - 8x_0 + x_0^2 = (x_0 - 5)(x_0 - 3) > 0.$$

(iii). Poleg tega velja $|x_{r+1} - 4| \leq |x_r - 4|$ za $x_r \in (3, 5)$. Izračunajmo

$$|x_{r+1} - 4| = |-12 + 8x_r - x_r^2 - 4| = |-16 + 8x_r - x_r^2| = (x_r - 4)^2, \qquad (5.1)$$

ker je $x_r$ na $(3, 5)$, sledi $|x_r - 4| < 1$ in potem $|x_r - 4|^2 < |x_r - 4|$.Vidimo, da velja tudi $|x_r - 4| = (x_0 - 4)^{2^r}$, kar že zagotavlja konvergenco. Zvezo dobimo z zaporedno uporabo enačbe (5.1).

(iv). Za robne točke velja $g(3) = g(5) = 3$.

Naši možni rešitvi sta $x_1 = 4$ in $x_2 = 3$, $x_2$ dobimo samo v primeru robnih točk, kar je praktično nemogoče. Zato si oglejmo red konvergence za okolico ničle $x_1 = 4$. Spomnimo se, da velja: metoda je reda $p$, natanko takrat ko $g'(x_1) = \cdots = g^{(p-1)}(x_1) = 0$ in $g^{(p)}(x_1) \neq 0$. Odvod je $g'(x) = 8 - 2x$, torej je $g'(4) = 0$. Red je vsaj kvadratičen. Drugi odvod je enak $g''(x) = -2$. Torej je red kvadratičen.

Izrek nam konvergenco zagotavlja na intervalu, kjer velja $|g'(x)| < 1$. Torej mora veljati |8-2x|< 1, to pa je $3.5 < x < 4.5$. Izrek nam da torej le potreben pogoj, iterativna metoda pa lahko konvergira še na večjem intervalu.

∎

**Naloga 5.3** *Iščemo rešitve enačbe $f(x) = x^5 - 10x + 1 = 0$. Za iteracijsko funkcijo izberemo $g(x) = \frac{x^5+1}{10}$. Ali nam izrek zagotavlja konvergenco za $x_0 = 0$? Oceni napako drugega približka.*

*Rešitev.* Za odvod mora veljati $|g'(x)| \leq m < 1$. Kar je $|\frac{x^4}{2}| < 1$, $|x| < \sqrt[4]{2} \doteq 1.1892$. Torej imamo za $x_0$ konvergenco. Napako drugega približka bomo ocenili s pomočjo formule $|x_r - \alpha| \leq \frac{m}{1-m}|x_r - x_{r-1}|$. Oceniti moramo še $m$, izberemo si interval $[0, 0.2]$, kjer je funkcija skrčitev. Vidimo, da velja $m \leq \frac{0.2^4}{2} = 0.0008$. Imamo $x_1 = 0.1$, $x_2 = 0.100001$ in

$$|x_2 - \alpha| \leq \frac{0.0008}{1 - 0.0008}|0.100001 - 0.1| \doteq 8 \cdot 10^{-10}.$$

∎

Tangentno metodo dobimo, če za naslednjo točko v iteraciji vzamemo presečišče tangente na funkcijo $f$ z osjo $x$. Tako dobimo $x_{r+1} = g(x_r) = x_r - \frac{f(x_r)}{f'(x_r)}$.

**Izrek 5.3** *Naj velja $f(\alpha) = 0$ in naj bo $\alpha$ $m$-kratna ničla. Če je $m = 1$, je konvergenca vsaj kvadratična. Če velja še $f''(\alpha) = 0$, je konvergenca kubična. Če je $m \geq 2$, velja $\lim_{x\to\alpha} g'(x) = 1 - 1/m$.*

*Rešitev.* Dokazali bomo samo drugi del izrek, saj je prvi enostaven. Če je $\alpha$ $m$-kratna ničla funkcije $f$, potem po definiciji obstaja $l \neq 0$, da velja $\lim_{x\to\alpha} \frac{f(x)}{(x-\alpha)^m} = l$. Po L'Hopitalu dobimo $\lim_{x\to\alpha} \frac{f'(x)}{m(x-\alpha)^{m-1}} = l$, saj gresta imenovalec in števec proti 0. Najprej pokažimo, da velja $\lim_{x\to\alpha} x - \frac{f(x)}{f'(x)} = \alpha$. Zadosti bo, da dokažemo $\lim_{x\to\alpha} \frac{f(x)}{f'(x)} = 0$. Imenovalec in števec delimo z $(x-\alpha)^m$, dobimo

$$\lim_{x\to\alpha}(x-\alpha)\frac{\frac{f(x)}{(x-\alpha)^m}}{m\frac{f'(x)}{m(x-\alpha)^{m-1}}} = \lim_{x\to\alpha}(x-\alpha)\frac{l}{ml} = 0.$$

Izračunajmo še limito odvoda:

$$g'(\alpha) = \lim_{x\to\alpha} \frac{x - \frac{f(x)}{f'(x)} - g(\alpha)}{x - \alpha} = \lim_{x\to\alpha} \frac{x - \frac{f(x)}{f'(x)} - \alpha}{x - \alpha} = 1 - \lim_{x\to\alpha} \frac{f(x)}{f'(x)(x-\alpha)}.$$

Nadaljujmo

$$1 - \lim_{x\to\alpha} \frac{f(x)}{f'(x)(x-\alpha)} = 1 - \lim_{x\to\alpha} \frac{\frac{f(x)}{(x-\alpha)^m}}{m\frac{f'(x)}{m(x-\alpha)^{m-1}}} = 1 - \frac{l}{ml} = 1 - \frac{1}{m}.$$

∎

**Naloga 5.4** *Poišči red konvergence iterativne metode $x_{r+1} = g(x_r) = x_r \frac{x_r^2+3a}{3x_r^2+a}$ za iskanje korena $\sqrt{a}$, kjer $a \neq 0$. Pokaži, da je konvergentna za vsak $x_0 > 0$.*

*Rešitev.* Odvod je enak

$$g'(x) = \frac{x^2+3a}{3x^2+a} + x\frac{-6x(x^2+3a)+2x(3x^2+a)}{(3x^2+a)^2}$$
$$= \frac{3x^4+10ax^2+3a^2-16ax^2}{(3x^2+a)^2} = \frac{3(x^4-2ax^2+a^2)}{(3x^2+a)^2} = 3\frac{(x^2-a)^2}{(3x^2+a)^2},$$

velja torej $g(\sqrt{a}) = g'(\sqrt{a}) = 0$. Izračunajmo še drugi odvod

$$g''(x) = 3\frac{4x(x^2-a)(3x^2+a)^2 - 12x(x^2-a)^2(3x^2+a)}{(3x^2+a)^4}$$
$$= \frac{12x(x^2-a)(3x^2+a-3x^2+3a)}{(3x^2+a)^3} = \frac{48ax(x^2-a)}{(3x^2+a)^2}.$$

Izračunajmo še tretji odvod

$$g'''(x) = \frac{48xa}{(3x^2+a)^3}\cdot 2x + (x^2-a)\cdot\square\,|_{x=\sqrt{a}} = \frac{96a^2}{64a^3} = \frac{3}{2a} \neq 0.$$

Metoda je torej reda 3. Poglejmo si razliko $g(x_r) - \sqrt{a}$:

$$g(x_r) - \sqrt{a} = x\left(\frac{x_r^2+3a}{3x_r^2+a}\right) - \sqrt{a} = \frac{x_r^3+3ax_r-3x_r^2\sqrt{a}-a\sqrt{a}}{3x_r^2+a} = \frac{(x_r-\sqrt{a})^3}{3x_r^2+a}.$$

Tako velja

$$\frac{x_{r+1}-\sqrt{a}}{x_r-\sqrt{a}} = \frac{(x_r-\sqrt{a})^2}{3x_r^2+a}.$$

Za $0 < x_0 < \sqrt{a}$ lahko ocenimo

$$|\frac{x_0-\sqrt{a}}{\sqrt{3x_0^2+a}}| < \frac{\sqrt{a}}{\sqrt{a}} = 1.$$

Za $x_0 \geq \sqrt{a}$ pišimo $x_0 = \sqrt{a} + y_0$ in izračunajmo

$$|\frac{x_0-\sqrt{a}}{\sqrt{3x_0^2+a}}| = \left|\frac{y_0}{\sqrt{4a+3y_0^2+6\sqrt{a}y_0}}\right| < \frac{y_0}{\sqrt{3y_0^2}} = 1/\sqrt{3} < 1.$$

Tako ugotovimo, da ostanki po absolutni vrednosti strogo padajo proti 0. ∎

**Laguerrova metoda** je metoda za iskanje ničel polinoma

$$f(z) = a_0(z-\alpha_1)(z-\alpha_2)\dots(z-\alpha_n).$$

Iteracije je oblike

$$z_{r+1} = z_r - \frac{nf(z_r)}{f'(z_r) \pm \sqrt{(n-1)\left[(n-1)f'^2(z_r) - nf(z_r)f''(z_r)\right]}}.$$

Velja $\epsilon_r = x_r - \alpha$, $\epsilon_{r+1} = C\epsilon_r^3 + O(\epsilon_r^4)$, kjer je $C = \frac{(n-2)f''^2(\alpha)}{8(n-1)f'^2(\alpha)} - \frac{f'''(\alpha)}{6f'(\alpha)}$. Predznak izberemo tako, da je izraz v imenovalcu po absolutni vrednosti največji. Konvergenca je pri enostavnih ničlah vsaj kubična.

**Naloga 5.5** *Izpelji metodo kubičnega korena, tako da v Laguerrovi metodi pošlješ $n \to \infty$. Z izpeljano metodo izračunaj $\sqrt[3]{a}$. Naredi tri korake za $a = 3$ in $x_0 = 1$.*

**Naloga 5.6** *Izpelji metodo $(f, f', f'')$ za iskanje enostavnih ničel, pomagaj si z razvojem inverzne funkcije v vrsto. To je posplošitev tangentne metode.*

*Rešitev.* Naj bo $f(\alpha) = 0$ in $\alpha$ enostavna ničla. Oglejmo si $y = f(x)$ v bližini $\alpha$. Ker je $\alpha$ enostavna ničla obstaja inverzna funkcija $x = f^{-1}(y) := F(y)$ in velja $\alpha = F(0)$, $x = F(f(x)) = F(y)$. Ničlo $\alpha$ razvijemo po $y = f(x)$ v Taylorjevo vrsto okoli $y = f(x)$:

$$\alpha = F(0) = F(y) - yF'(y) + \frac{y^2}{2!}F''(y) - \ldots .$$

Oglejmo si $F(y) = x$. Z odvajanjem po $x$ dobimo $F'(y)f'(x) = 1 \Rightarrow F'(y) = \frac{1}{f'(x)}$. Podobno za drugi odvod dobimo

$$F''(y)f'(x) = -\frac{f''(x)}{f'^2(x)} \Rightarrow F''(y) = -\frac{f''(x)}{f'^3(x)}.$$

Torej je

$$\alpha = x - \frac{f(x)}{f'(x)} - \frac{1}{2}f^2(x)\frac{f''(x)}{f'^3(x)} - \ldots,$$

vse višje člene zanemarimo. Dobimo iterativno formulo

$$x_{r+1} = x_r - \frac{f(x_r)}{f'(x_r)} - \frac{1}{2}f^2(x_r)\frac{f''(x_r)}{f'^3(x_r)}.$$

V bližini enostavne ničle je red metode kubičen. Dokaz je prepuščen za domačo nalogo. ■

**Naloga 5.7** *V iteracijski formuli* $x_{r+1} = \frac{A^2}{x_r^5}\left(\alpha + \frac{\beta x_r^3}{A} + \frac{\gamma x_r^6}{A^2}\right)$ *za* $\sqrt[3]{A}$ *določi* $\alpha$, $\beta$, $\gamma$, *da bo metoda vsaj kubična.*

*Rešitev.* Veljati mora $g(\sqrt[3]{A}) = \sqrt[3]{A}$, $g'(\sqrt[3]{A}) = 0$, $g''(\sqrt[3]{A}) = 0$. Imamo

$$\begin{aligned} g(x) &= \alpha A^2 x^{-5} + \beta A x^{-2} + \gamma x, \\ g'(x) &= -5\alpha A^2 x^{-6} - 2\beta A x^{-3} + \gamma, \\ g''(x) &= 30\alpha A^2 x^{-7} + 6\beta A x^{-4}. \end{aligned}$$

Dobimo sistem:

$$\begin{aligned} \alpha + \beta + \gamma &= 1, \\ -5\alpha - 2\beta + \gamma &= 0, \\ 30\alpha + 6\beta &= 0. \end{aligned}$$

Iz tretje enačbe dobimo $\beta = -5\alpha$. Vstavimo v drugo in dobimo $\gamma = -5\alpha$. Nazadnje iz prve enačbe dobimo $-9\alpha = 1$. Torej je $\alpha = -\frac{1}{9}$, $\beta = \gamma = \frac{5}{9}$. Za $A = 5$ in $x_0 = 2$ dobimo $g(x_r) = \frac{25}{9x_r^5}\left(-1 + x_r^3 + \frac{x_r^6}{5}\right)$. Navedimo nekaj začetnih približkov: $x_1 \doteq 1.71875$, $x_2 = 1.709976326$, $x_3 = 1.709975947$. ■

**Naloga 5.8** *Določi vse polinome četrte stopnje z vodilnim koeficientom 1, pri katerih se tangentna metoda veda takole:*

- *v bližini* $\alpha$ *ima linearno konvergenco,*
- *v bližini* $-\alpha$ *ima kubično konvergenco.*

*Določi še preostale ničle polinoma in red konvergence v njihovi bližini.*

*Rešitev.* Spomnimo se, da je v bližini enostavnega korena konvergenca vsaj kvadratična. Če je prvi odvod enak 0, je konvergenca kubična. V primeru večkratne ničle je konvergenca linearna. Iz tega ugotovimo, da je $\alpha$ vsaj dvakratna ničla in $-\alpha$ enostavna ničla. Polinom je oblike $p(x) = (x+\alpha)(x-\alpha)^2(x-d)$, kjer je $d$ neznana ničla. Uporabimo formulo

$$(f(x)g(x))^{(n)} = \sum_{i=0}^{n} f^{(i)}(x)g^{(n-i)}(x).$$

Dobimo

$$p''(x) = 2(x-\alpha)^2 + 4(x-\alpha)(x-d) + 4(x+\alpha)(x-d) + (x-\alpha)(x+\alpha).$$

Če hočemo, da bo konvergenca kubična, mora veljati $p''(-\alpha) = 0$. Vstavimo $-\alpha$:

$$p''(-\alpha) = 2(-\alpha-\alpha)^2 + 4(-\alpha-\alpha)(-\alpha-d) + 0 + 0 = 8\alpha(\alpha+\alpha+d) = 0,$$

iz česar dobimo $d = -2\alpha$. Ničla $-2\alpha$ je enostavna, konvergenca je vsaj kvadratična. Izračunajmo še vrednost drugega odvoda v $-2\alpha$:

$$p''(-2\alpha) = 2(-2\alpha-\alpha)^2 + 0 + 0 + (-2\alpha+\alpha)(-2\alpha-\alpha) = 30\alpha^2 \neq 0.$$

Konvergenca v $-2\alpha$ je kvadratična. ■

Newtonova metoda se uporablja za iskanje rešitev sistema $F(z) = 0$. Metoda je posplošitev tangentne metode. Označimo $\Delta z_r = z_{r+1} - z_r$. Recimo, da je $z_{r+1}$ ničla. Velja

$$0 \doteq F(z_r) + J(z_r)(z_{r+1} - z_r),$$

kjer je $J$ Jacobijeva matrika parcialnih odvodov reda 1. Dobimo

$$J(z_r)\Delta z_r = -F(z_r).$$

Pri vsakem koraku rešimo sistem in dobimo $\Delta z_r$, $z_{r+1} = z_r + \Delta z_r$.

**Naloga 5.9** *Nastavi Newtonovo metodo za reševanje sistema:*

$$\begin{aligned} f_1(x,y) &= x^3 - 3xy^2 + 1 = 0, \\ f_2(x,y) &= 3x^2y - y^3 = 0. \end{aligned}$$

*Velja* $F = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$. *Izračunajmo*

$$JF = \begin{bmatrix} f_{1x} & f_{1y} \\ f_{2x} & f_{2y} \end{bmatrix} = \begin{bmatrix} 3x^2 - 3y^2 & -6xy \\ 6xy & 3x^2 - 3y^2 \end{bmatrix}.$$

*Rešitev.* Na vsakem koraku rešimo:

$$\begin{bmatrix} 3x_r^2 - 3y_r^2 & -6x_ry_r \\ 6x_ry_r & 3x_r^2 - 3y_r^2 \end{bmatrix} \begin{bmatrix} \Delta x_r \\ \Delta y_r \end{bmatrix} = - \begin{bmatrix} x_r^3 - 3x_ry_r^2 + 1 \\ 3x_r^2y_r - y_r^3 \end{bmatrix}.$$

Izračunamo $x_{r+1} = x_r + \Delta x_r$ in $y_{r+1} = y_r + \Delta y_r$. To ponavljamo za $r = 0, 1, \ldots$ ■

**Naloga 5.10** *Pascalova matrika $P_5$ ima strukturo*

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{bmatrix}.$$

*V splošnem je $i$, $j$-ti element $P_n$ enak $\binom{i+j-2}{j-1}$. Izračunaj razcep Choleskega za $P_5$ in iz dobljenega razcepa postavi hipotezo za $P_n$ ter jo dokaži. Namig, upoštevaj zvezo med binomski koeficienti, ki jo dobiš, če razpišeš enakost $(x+y)^{a+b} = (x+y)^a(x+y)^b$.*

*Rešitev.* Za razcep Choleskega dobimo $P_5 = V_5 \cdot V_5^T$, kjer je

$$V_5 = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}.$$

Opazimo, da velja $v_{ij} = v_{i-1,j-1} + v_{i-1,j}$. Velja še $v_{i1} = v_{ii} = 1$. Kar nam da idejo za $v_{ij} = \binom{i-1}{j-1}$. Za $n = 5$ preverimo, da je to res. V splošnem moramo dokazati

$$(P_n)_{ij} = \sum_{k=1}^{j} (V_n)_{ik}(V_n)_{jk} = \sum_{k=1}^{j} \binom{i-1}{k-1}\binom{j-1}{k-1} \stackrel{?}{=} \binom{i+j-2}{j-1}.$$

Če pišemo $a = i-1$, $b = j-1$ in $k' = k-1$, dobimo

$$\sum_{k'=0}^{b} \binom{a}{k'}\binom{b}{k'}.$$

Razpišimo $(x+y)^{a+b} = (x+y)^a(x+y)^b$. Dobimo

$$\sum_{k=0}^{a+b} \binom{a+b}{k} x^k y^{a+b-k} = \left(\sum_{i=0}^{a} \binom{a}{i} x^i y^{a-i}\right)\left(\sum_{j=0}^{b} \binom{b}{j} x^{b-j} y^j\right).$$

Primerjamo koeficente pri $x^b y^a$. Dobimo

$$\binom{a+b}{b} = \sum_{i=j} \binom{a}{i}\binom{b}{j} = \sum_{k=0}^{b} \binom{a}{k}\binom{b}{k}.$$

$\blacksquare$

# Poglavje 6

# Predoločeni sistemi, ekstremi funkcij

**Naloga 6.1** *Podane so točke* $(x_1, y_1)$, $(x_2, y_2)$, $\ldots$, $(x_n, y_n)$. *Poišči funkcijo oblike* $f(x) = ae^{bx}$, *ki se bo najbolj prilegala tem točkam. Uporabi metodo najmanjših kvadratov.*

*Rešitev.* Iščemo par $(a, b)$, ki minimizira funkcijo

$$G(a,b) = \sum_{i=0}^{n} (y_i - ae^{bx_i})^2.$$

Funkcija $G$ ima minimum, ko so vsi parcialni odvodi enaki nič. Dobimo sistem

$$\begin{aligned}
\frac{\partial G}{\partial a}(a,b) &= \sum_{i=0}^{n} 2(y_i - ae^{bx_i})(-e^{bx_i}) = 0, \\
\frac{\partial G}{\partial b}(a,b) &= \sum_{i=0}^{n} 2(y_i - ae^{bx_i})(-ax_i e^{bx_i}) = 0.
\end{aligned}$$

Kar je ekvivalentno reševanju sistema

$$\begin{aligned}
f_1(a,b) &= \sum_{i=0}^{n} (y_i - ae^{bx_i})e^{bx_i} = 0, \\
f_2(a,b) &= \sum_{i=0}^{n} 2(y_i - ae^{bx_i})e^{bx_i} = 0.
\end{aligned}$$

Uporabimo Newtonovo metodo. Parcialni odvodi so enaki

$$\begin{aligned}
\frac{\partial f_1}{\partial a} &= -\sum_{i=0}^{n} e^{2bx_i} \\
\frac{\partial f_1}{\partial b} &= \sum_{i=0}^{n} x_i e^{bx_i}(y_i - 2ae^{bx_i}) \\
\frac{\partial f_2}{\partial a} &= -\sum_{i=0}^{n} x_i (e^{bx_i})^2 \\
\frac{\partial f_2}{\partial b} &= \sum_{i=0}^{n} x_i^2 e^{bx_i}(y_i - 2ae^{bx_i}).
\end{aligned}$$

Za

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial a} & \frac{\partial f_1}{\partial b} \\ \frac{\partial f_2}{\partial a} & \frac{\partial f_2}{\partial b} \end{bmatrix}, F = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

rešujemo sistem $J(z_r)\Delta z_r = -F(z_r)$. Začetni približek lahko dobimo prek linearizacije $ae^{bx} \doteq a + abx$. Druga možnost je, da poiščemo $a$, $b$, tako da gre funkcija čez dve podani točki.

■

Radi bi rešili sistem $Ax = b$, kjer je $A \in \mathbb{C}^{m \times n}$ matrika polnega ranga, $x \in \mathbb{C}^n$, $b \in \mathbb{C}$ ter $m \geq n$. Sistem v splošnem ni rešljiv, zato minimiziramo normo ostanka $||Ax - b||_2$. Izkaže se, da je tak $x$ ravno rešitev **normalnega sistema** $A^T Ax = A^T b$.

**Naloga 6.2** *Merili smo tir delca, ki se naj bi gibal po paraboli.*

| $i$ | $x_i$ | $f(x_i)$ |
|---|---|---|
| 1 | $-1$ | $\frac{11}{4}$ |
| 2 | 0 | $\frac{7}{4}$ |
| 3 | 1 | $\frac{1}{4}$ |
| 4 | 2 | $\frac{13}{4}$ |

*V bližini teh točk bi radi potegnili parabolo. Določi njene koeficiente po metodi najmanjših kvadratov. Uporabi normalni sistem in ga reši z Gaussovo eliminacijo (po metodi Choleskega).*

*Rešitev.* Parabola naj bo $a + bx + cx^2$. Dobimo

$$\begin{array}{ccccccc} a & - & b & + & c & = & \frac{11}{4} \\ a & + & 0b & + & 0c & = & \frac{7}{4} \\ a & + & b & + & c & = & \frac{1}{4} \\ a & + & 2b & + & 4c & = & \frac{13}{4} \end{array} \qquad \overbrace{\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}^{x} = \overbrace{\begin{bmatrix} \frac{11}{4} \\ \frac{7}{4} \\ \frac{1}{4} \\ \frac{13}{4} \end{bmatrix}}^{z},$$

kar je predoločen sistem $Ax = z$. Iz tega dobimo normalni sistem $\overbrace{A^T A}^{B} x = A^T z = v$. Izračunamo

$$B = A^T A = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \text{ in } v = A^T z = \begin{bmatrix} 8 \\ 4 \\ 16 \end{bmatrix}.$$

Sistem rešimo z Gaussovo eleminacijo

$$\left[\begin{array}{ccc|c} 4 & 2 & 6 & 8 \\ 2 & 6 & 8 & 4 \\ 6 & 8 & 18 & 16 \end{array}\right] \to \left[\begin{array}{ccc|c} 2 & 6 & 8 & 4 \\ 0 & -10 & -10 & 0 \\ 0 & -10 & -6 & 4 \end{array}\right] \to \left[\begin{array}{ccc|c} 2 & 6 & 8 & 4 \\ 0 & -10 & -10 & 0 \\ 0 & 0 & 4 & 4 \end{array}\right]$$

Lahko uporabimo razcep Choleskega. Najprej izračunamo razcep Choleskega matrike $B = VV^T$ in dobimo

$$V = \begin{bmatrix} 2 & & \\ 1 & \sqrt{5} & \\ 3 & \sqrt{5} & 2 \end{bmatrix}.$$

Dobimo sistem $V \overbrace{(V^T x)}^{y} = z$. Iz

$$\begin{bmatrix} 2 & & \\ 1 & \sqrt{5} & \\ 3 & \sqrt{5} & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 16 \end{bmatrix}, \quad \begin{array}{ccccccc} 2y_1 & + & 0y_2 & + & 0y_3 & = & 8 \\ y_1 & + & \sqrt{5}y_2 & + & 0y_3 & = & 4 \\ 3y_1 & + & \sqrt{5}y_2 & + & 2y_3 & = & 16 \end{array}, \text{ dobimo } \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}.$$

Rešiti moramo še sistem $V^T x = y$, kar je

$$\begin{bmatrix} 2 & 1 & 3 \\ 0 & \sqrt{5} & \sqrt{5} \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}.$$

Rešimo sistem in dobimo $a = 1$, $b = -1$, $c = 1$. Parabola je $1 - x + x^2$. ∎

**Naloga 6.3** *Sestavi ekonomičen (čim manj operacij) algoritem za računanje razcepa Choleskega, $A = VV^T$ pozitivno definitne tridiagonalne matrike. Koliko operacij porabiš za izračun $B = V^T V$.*

*Rešitev.* Matrika $A$ je oblike

$$A = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & \ddots & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix},$$

torej bo

$$A = \overbrace{\begin{bmatrix} u_1 & & & & \\ v_1 & u_2 & & & \\ & v_2 & u_3 & & \\ & & \ddots & \ddots & \\ & & & v_{n-1} & u_n \end{bmatrix}}^{V} \overbrace{\begin{bmatrix} u_1 & v_1 & & & \\ & u_2 & v_2 & & \\ & & \ddots & \ddots & \\ & & & u_{n-1} & v_{n-1} \\ & & & & u_n \end{bmatrix}}^{V^T}.$$

Ko matrike zmnožimo dobimo $u_1^2 = a_1$ in $v_i u_i = b_i$ ter $v_i^2 + u_{i+1}^2 = a_{i+1}$ za $i = 1, \ldots, n-1$. Iz tega že lahko zapišemo algoritem. Porabili smo $2n - 2$ množenj/deljenj in $n$ korenjenj. Izračunajmo

```
u_1 = sqrt(a_1);
for i = 1,...,n-1 do
    v_i = b_i/u_i;
    u_{i+1} = sqrt(a_{i+1} - v_i^2);
end
```

še $V^T V$:

$$V^T V = \begin{bmatrix} c_1 & d_1 & & & \\ d_1 & c_2 & d_2 & & \\ & d_2 & \ddots & \ddots & \\ & & \ddots & \ddots & d_{n-1} \\ & & & d_{n-1} & c_n \end{bmatrix}.$$

**for** $i = 1, \ldots, n-1$ **do**
  $c_i = u_i^2 + v_i^2$;
  $d_i = v_i u_{i+1}$;
**end**
$c_n = u_n^2$;
Porabimo $3n - 2$ množenj.

■

**Izrek 6.1** *Naj bo $A \in \mathbb{R}^{m\times n}$, kjer je $m \geq n$ in $rang(A) = n$. Potem obstaja enolični $QR$ razcep $A = QR$, kjer je $Q$ pravokotna matrika dimenzije $m \times n$ z ortogonalnimi stolpci, $R$ pa zgornje trikotna matrika s pozitivnimi diagonalnimi elementi.*

**for** $k = 1, \ldots, n$ **do**
  $q_k = a_k$;
  **for** $i = 1, \ldots, k-1$ **do**
    $r_{ik} = q_i^T a_k (CGS)$ ali $r_{ik} = q_i^T a_k (MGS)$;
    $q_k = q_k - r_{ik} q_i$;
  **end**
  $r_{kk} = ||q_k||_2$;
  $q_k = \frac{q_k}{r_{kk}}$;
**end**

Pri reševanju predoločenega sistema si lahko pomagamo s $QR$ razcepom. Boljša je različica z MGS. Dobimo, da moramo rešiti sistem $Rx = Q^T b$. Boljše je narediti razširjen $QR$ razcep:

$$Ax - b = \begin{bmatrix} A & b \end{bmatrix} \begin{bmatrix} x \\ -1 \end{bmatrix} = \begin{bmatrix} Q & q_{n+1} \end{bmatrix} \begin{bmatrix} R & z \\ 0 & \rho \end{bmatrix} \begin{bmatrix} x \\ -1 \end{bmatrix} = Q(Rx - z) - \rho q_{n+1}.$$

Iz česar dobimo, da moramo rešiti sistem $Rx = z$, maksimum pa je enak $\rho$.

**Naloga 6.4** *Poišči normalni sistem za reševanje naslednjih dveh problemov najmanjših kvadratov, kjer je $A \in \mathbb{R}^{m\times n}$, $m \geq n$ in $rang(A) = n$.*

*(i). Uteženi problem najmanjših kvadratov. Iščemo*

$$\min_x ||D(Ax - b)||_2,$$

*kjer je $D$ nesingularna diagonalna.*

*(ii). Iščemo*

$$\min_x ||Ax - b||_C,$$

*kjer je $C$ simetrična pozitivno definitna, ki generira normo*

$$||x||_C = (x^T C x)^{\frac{1}{2}}.$$

*Rešitev.* (i)
Izračunajmo

$$\min_x ||D(Ax-b)|| = \min_x ||DAx - Db||,$$

matrika $F = DA$ je spet dimenzije $m \times n$. Označimo še $Db = \widehat{b}$, tako dobimo predoločen sistem $Fx = \widehat{b}$, iz česar dobimo

$$F^T Fx = (DA)^T (DA)x = F^T \widehat{b} = (DA)^T (Db).$$

Razpisano je to enako

$$A^T D^2 Ax = A^T D^2 b$$

(ii)
Matrika $C$ je simetrična pozitivno definitna, zato zanjo obstaja razcep Choleskega $C = V^T V$, kjer $V$ zgornje trikotna. Izračunajmo

$$\min_x ||Ax-b||_C = \min_x \left((Ax-b)^T C(Ax-b)\right)^{\frac{1}{2}} = \min_x \left((Ax-b)^T V^T \overbrace{V(Ax-b)}^{y}\right)^{\frac{1}{2}},$$

uvedemo novo neznanko $y = U(Ax-b)$, dobimo

$$\min_x (y^T y)^{\frac{1}{2}} = \min_x < y, y >^{\frac{1}{2}} = \min_x ||y||_2 = \min_x ||V(Ax-b)||_2.$$

Torej iščemo

$$\min_x ||VAx - Vb||_2,$$

to je predoločen sistem

$$VAx = Vb \;\; /(VA)^T,$$

kar da

$$A^T V^T VAx = A^T V^T Vb.$$

Z upoštevanjem $C = V^T V$ dobimo

$$A^T CAx = A^T Cb.$$

■

**Naloga 6.5** *Metoda zlatega reda je metoda za iskanje minimuma funkcije. Ta metoda je podobna bisekciji za iskanje ničle funkcije. Pri bisekciji potrebujemo dve točki, za kateri velja $f(a)f(b) < 0$. Če hočemo, da bo funkcija imela minimum na intervalu $[a,c]$, mora veljati $f(b) \le f(a)$ in $f(c) \le f(a)$. Hočemo, da bo točka $b$ optimalno izbrana, dolžina intervalov naj bo čimbolj enaka. Pri bisekciji je najboljša enakomerna izbira, kjer za naslednjo točko izberemo $\frac{a+b}{2}$. Pri iskanju minimuma izberemo točko $x$ in potem poiščemo trojico zaporednih, za katere velja, da je funkcijska vrednost v srednji točki manjša kot funkcijski vrednosti v ostalih. Poišči kakšna je optimalna delitev intervala $w = \frac{b-a}{c-a}$.*

*Rešitev.* Naj bo $w = \frac{b-a}{c-a}$ relativna dolžina intervala $[a,b]$ glede na celoten interval $[a,c]$. Relativna dolžina intervala $[b,c]$ glede na $[a,c]$ je $1-w = \frac{c-b}{c-a}$. Recimo, da naslednjo točko $x$ izberemo na relativni razdalji $z$ od $b$, $z = \frac{x-b}{c-a}$. Recimo, da je $z \geq 0$. V naslednjem koraku imamo dve možnosti:

- izberemo točke $a, b, x$,
- izberemo točke $b, x, c$.

Hočemo, da sta intervala $[a,x]$ in $[b,c]$ enake dolžine. Veljati mora

$$\frac{x-a}{c-b} = 1 = \frac{z(c-a)+b-a}{c-b} = z\frac{c-a}{c-b} + \frac{b-a}{c-b} = \frac{z}{1-w} + \frac{w}{1-w},$$

kar je $x = a+c-b$ in $z = 1-2w$. Ker hočemo doseči optimalno izbiro, mora spet veljati $w = \frac{b-a}{x-a}$ in $w = \frac{x-b}{c-b}$. Iz prve enakosti ne dobimo nič novega, saj je $x-a = c-b$. Iz druge enakosti in prejšnjih enačb dobimo $w(1-w) = \frac{x-b}{c-b}\frac{c-b}{c-a} = \frac{x-b}{c-a} = z = 1-2w$. Rešiti moramo kvadratno enačbo $w^2 - 3w + 1 = 0$. Tako dobimo $w_{12} = \frac{3\pm\sqrt{5}}{2}$, pozitivna rešitev manjša od 1 je $w = \frac{3-\sqrt{5}}{2} \doteq 0,38196601125$. ∎

**Naloga 6.6** *(i). Naj bo $A \in \mathbb{R}^{m\times n}$, kjer je $m \geq n$ in rang$(A) = n$. Dokaži, da ima sistem*

$$\begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \end{array} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

*rešitev, ki ustreza rešitvi predoločenega sistema $Ax = b$ po metodi najmanjših kvadratov.*

*(ii). Uporabi zgornji sistem pri iterativnem izboljšanju rešitve predoločenega sistema. Pomagaj si s QR razcepom matrike $A$.*

***Algoritem 6****: Iterativno izboljšanje za obrnljivo matriko B*

$y_1 = 0$;
**while** $||r|| > eps$ **do**
  $y_0 = y_1$;
  $rr = By_0 - b$;
  *Rešimo sistem* $Bd = rr$;
  $y_1 = y_0 - d$;
**end**

*Rešitev.* (i)
Poglejmo kaj mora veljati za rešitev sistema. Po blokih dobimo $r + Ax = b$ in $A^T r = 0$. Iz prve enačbe dobimo $r = b - Ax$, kar vstavimo v drugo enačbo. Iz tega sledi

$$A^T(b - Ax) = 0 \Leftrightarrow A^T b - A^T Ax = 0 \Leftrightarrow A^T b = A^T Ax.$$

Dobili smo ravno normalni sistem.
(ii)
Upoštevajmo bločno strukturo in zapišimo spremenljivke v bločni obliki:

$$y = \begin{bmatrix} r_0 \\ x_0 \end{bmatrix}, \; d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

Za $rr$ dobimo

$$rr = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} r_0 \\ x_0 \end{bmatrix}.$$

Rešiti moramo sistem

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} d_1 & d_2 \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix}, \ \begin{bmatrix} r_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} r_0 \\ x_0 \end{bmatrix} - \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

Naredimo razširjeni $QR$ razcep matrike

$$A = \begin{bmatrix} & & \\ & Q & \\ & & \end{bmatrix} \overbrace{\begin{bmatrix} R_1 \\ \\ 0 \end{bmatrix}}^{R},$$

kjer je $Q$ ortogonalna in $R$ zgornje trikotna. Dobimo

$$\begin{bmatrix} I & QR \\ R^T Q^T & 0 \end{bmatrix} \begin{bmatrix} d_1 & d_2 \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} \Rightarrow \begin{array}{c} d_1 + QRd_2 = l_1 / Q^T \cdot \\ R^T Q^T d_1 = l_2 \end{array}.$$

Iščemo $d_1$ in $d_2$. Dobimo sistem

$$Q^T d_1 + \begin{bmatrix} R_1 \\ 0 \end{bmatrix} d_2 = Q^T l_1 \ (1), \quad \begin{bmatrix} R_1^T & 0 \end{bmatrix} Q^T d_1 = l_2 \ (2),$$

označimo še

$$Q^T d_1 = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \ Q^T l_1 = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Iz prve enačbe dobimo $R_1^T z_1 = l_2$, ko rešimo sistem dobimo $z_2$. Iz druge enačbe dobimo

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} R_1 \\ 0 \end{bmatrix} d_2 = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \text{ kar da } \begin{array}{c} z_1 + R_1 d_2 = f_1 \\ z_2 = f_2 \end{array} \Rightarrow R_1 d_2 = f_1 - z_1.$$

Dobimo še $d_2$. ■

---

$A = QR$;

$Q^T e_1 = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$;

$z_2 = f_2$;

Reši sistem $R_1^T z_1 = l_2$;

$d_1 = Q \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$;

Rešimo sistem $R_1 d_1 = f_1 - z_1$;

---

**Naloga 6.7** *Dokaži Hadamardovo neenakost za matriko $A \in \mathbb{R}^{n \times n}$:*

$$|\det(A)|^2 \leq \prod_{i=1}^{n} \left( \sum_{k=1}^{n} |a_{ki}|^2 \right).$$

*Pomagaj si s QR razcepom.*

*Rešitev.* Velja

$$\det(A) = \det(QR) = \overbrace{\det(Q)}^{\pm 1} \cdot \det(R).$$

Ker je $R$ zgornje trikotna matrika, je njena determinanta enaka produktu diagonalnih elementov. Tako dobimo

$$|\det(A)|^2 = |r_{11}|^2 |r_{22}|^2 \cdots |r_{nn}|^2.$$

Naj bosta $A^{(i)}$ in $R^{(i)}$ $i$-ta stolpca matrik $A$ in $R$, zanju velja $A^{(i)} = QR^{(i)}$. Opazimo, da velja

$$||A^{(i)}||_2^2 = \sum_{k=1}^{n} |a_{ki}|^2 = ||QR^{(i)}||_2^2 \stackrel{*}{=} ||R^{(i)}||_2^2 = \sum_{k=1}^{i} |r_{ki}|^2 \geq |r_{ii}|^2.$$

Enakost $*$ dobimo iz dejstva, da ortogonalna matrika ohranja drugo normo. Neenakosti, ki smo jih dobili zmnožimo in dobimo Hadamardovo neenakost. ■

**Naloga 6.8** *Naj bo $A \in \mathbb{R}^{m \times n}$, $m \geq n$ in $rang(A) = n$. Podana je še matrika $C \in \mathbb{R}^{p \times n}$, $p \leq n$, $rang(C) = p$. Zapiši algoritem za reševanje predoločenega sistema $Ax = b$, pri pogoju $Cx = d$.*

*Rešitev.* Iščemo $\min_{Cx=d} ||Ax - b||_2$, poglejmo si oblike matrik. Pomagajmo si z razširjenim

$Q \begin{array}{c} p \\ n-p \end{array} \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ razcepom matrike $C^T$. Rešujemo problem

$$\min_{Cx=d} ||Ax - b||_2 = \min_{CQQ^Tx=d} ||AQQ^T x - b||_2.$$

Izračunajmo $CQ = (C^T)^T Q = (QR)^T Q = R^T Q^T Q = R^T = \begin{bmatrix} R_1^T & 0 \end{bmatrix}$, uvedemo še nove neznanke $Q^T x = \begin{array}{c} p \\ n-p \end{array} \begin{bmatrix} y \\ z \end{bmatrix}$ in $AQ = \begin{array}{c} \begin{array}{cc} p & n-p \end{array} \\ \begin{bmatrix} A_1 & A_2 \end{bmatrix} \end{array}$. Nadaljujemo z

$$\min_{\begin{bmatrix} R_1^T & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = d} || \begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} - b||_2 = \min_{R_1^T y = d} ||A_1 y + A_2 z - b||_2.$$

Z $R^T y = d$ je $y$ enolično določen. Torej nam

$$\min_{R_1^T y = d} ||A_2 z - (b - \overbrace{A_1 y}^{\text{znano}})||_2$$

predstavlja predoločeni sistem $A_2 z = b - A_1 y$.

---

$C^T = QR$;
$R_1^T y = d$ (rešimo sistem, dobimo $y$);
$AQ = \begin{array}{c} \begin{array}{cc} p & n-p \end{array} \\ \begin{bmatrix} A_1 & A_2 \end{bmatrix} \end{array}$;
$\min_z ||A_2 z - (b - A_1 y)||_2$ (rešimo predoločen sistem, dobimo $z$) ;
$x = Q \begin{bmatrix} y \\ z \end{bmatrix}$;

---

■

Givensova rotacija $R_{ik}$ je matrika enaka identiteti povsod razen v $i$-ti in $k$-ti vrstici in preslika $i$-to in $k$-to komponento vektorja $x$ v vektorja $y$, ki ima $k$-to komponento enako 0.

$$R_{ik}^T([i,k],[i,k]) = \begin{bmatrix} \frac{x_i}{\sqrt{x_i^2+x_k^2}} & \frac{x_k}{\sqrt{x_i^2+x_k^2}} \\ -\frac{x_k}{\sqrt{x_i^2+x_k^2}} & \frac{x_i}{\sqrt{x_i^2+x_k^2}} \end{bmatrix}.$$

**Naloga 6.9** *Naj bo $A$ zgornja Hessenbergova matrika ($a_{ik} = 0,\ i > k+1$). Zapiši algoritem za $QR$ razcep s pomočjo Givensovih rotacij in z njegovo pomočjo reši sistem $Ax = b$. Poleg tega še preštej število operacij (korenjenja, seštevanja, množenja). Matrike $Q$ ti ni potrebno izračunati. Kakšna je oblika matrike $Q$?*

*Rešitev.* Givensova rotacija $R_{ik}$ spremeni samo $i$-to in $k$-to vrstico. Poglejmo si primer, ko je matrika dimenzije 4:

$$A = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix} \overset{R_{12}^T}{\rightarrow} \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix} \overset{R_{23}^T}{\rightarrow} \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & \times & \times \end{bmatrix} \overset{R_{34}^T}{\rightarrow} \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{bmatrix}$$

---

**for** $i = 1, \dots, n-1$ **do**
  $r = \sqrt{a_{ii}^2 + a_{i+1,i}^2}$ -|- $n-1\ \surd,\ 2n-2\ *,\ n-1\ +$ ;
  $c = \frac{a_{ii}}{r}$ -|- $n-1\ *$;
  $s = \frac{a_{i+1,i}}{r}$ -|- $n-1\ *$; $a_{ii} = r$;
  $z_1 = b_i$;
  $z_2 = b_{i+1}$;
  $b_i = cz_1 + sz_2$ -|- $2n-2\ *,\ n-1\ +$ ;
  $b_{i+1} = -sz_1 + cz_2$ -|- $2n-2\ *,\ n-1\ +$;
  **for** $k = i+1, \dots, n-1$ **do**
    $aik = a_{ik}$;
    $a_{ik} = ca_{ik} + sa_{i+1,k}$ -|- $2\sum_{i=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$ ;
    $a_{i+1,k} = -s \cdot aik + ca_{i+1,k}$ -|- $2\sum_{i=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$;
  **end**
**end**
Reši zgornje trikotni sistem $Rx = b$.

---

Vsota je enaka

$$\sum_{i=1}^{n-1}(n-i-1) = \sum_{k=1}^{n-2} k = \frac{(n-2)(n-1)}{2}.$$

Matrika $Q$ je zgornja Hessenbergova. ■

**Naloga 6.10** *Naj bo $A$ nesingularna zgornje trikotna matrika dimenzije $n \times n$. Podana sta še vektorja $u$,$v \in \mathbb{R}^n$. Zapiši čim hitrejši algoritem za $QR$ razcep matrike $B = A + uv^T$. Namig, kaj je $uv^T$, če je $u$ elementarni bazni vektor pomnožen s skalarjem, torej enak $\lambda e_i$?*

*Rešitev.* Zadosti bo, če matriko $B$ transformiramo v zgornjo Hessenbergovo matriko, saj lahko potem uporabimo prejšnjo nalogo. Če bi bil $u$ vektor iz namiga, bi imela matrika $uv^T$ neničelno samo $i$-to vrstico. Za $i = 1$ bi potem bila $B$ zgornja Hessenbergova matrika (celo zgornje trikotna). Če nam uspe $u$ transformirati z ortogonalnimi matrikami v $\lambda e_1$, tako da bo transformirana matrika A zgornja Hessenbergova, smo problem rešili. Če poizkusimo klasično, z $R_{12}^T$ uničimo 2. element, z $R_{13}^T$ uničimo 3. element, ..., z $R_{1n}^T$ uničimo $n$-ti element, se matrika $A$ transformira v polno matriko. Ta možnost ni dobra. Izkaže se, da moramo iti od spodaj navzgor. Skicirajmo to za $n = 4$.

$$\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} v^T \overset{R_{34}^T}{\longrightarrow} \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & \times & \times \end{bmatrix} + \begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \\ u_3^{(1)} \\ 0 \end{bmatrix} v^T \overset{R_{23}^T}{\longrightarrow}$$

$$\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix} + \begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \\ 0 \\ 0 \end{bmatrix} v^T \overset{R_{12}^T}{\longrightarrow} \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix} + \begin{bmatrix} u_1^{(3)} \\ 0 \\ 0 \\ 0 \end{bmatrix} v^T$$

Ker je

$$\begin{bmatrix} u_1^{(3)} \\ 0 \\ 0 \\ 0 \end{bmatrix} v^T = \begin{bmatrix} \times & \times & \times & \times \\ & & & \\ & & & \\ & & & \end{bmatrix},$$

smo dobil zgornjo Hessenbergovo matriko.

∎

Householderjevo zrcaljenje je matrika oblike $H = I - \frac{2}{w^T w} w w^T$ in predstavlja zrcaljenje prek ravnine določene z normalo $w$. Če hočemo preslikati vektor $x$ v $\pm k e_1$, lahko to naredimo s Householderjevim zrcaljenjem določenim z $w = x + sign(x_1)||x||_2 e_1$.

**Naloga 6.11** *S Householderjevimi zrcaljenji in QR razcepom reši linearni sistem*

$$\begin{array}{ccccccc} x_1 & + & 6x_2 & + & -2x_3 & = & -7 \\ 2x_1 & + & x_2 & + & -2x_3 & = & -1 \\ 2x_1 & + & 2x_2 & + & 6x_3 & = & 6 \end{array}.$$

*Rešitev.* Najprej transformiramo prvi stolpec

$$w = \begin{bmatrix} 1 + 1 \cdot (3 = ||A_1||_2) \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix},\ w^T w = 24,\ P_1 = I - \frac{1}{12} w w^T = \frac{1}{3} \begin{bmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{bmatrix},$$

$$P_1 \begin{bmatrix} 1 & 6 & -2 & -7 \\ 2 & 1 & -2 & -1 \\ 2 & 2 & 6 & 6 \end{bmatrix} = \begin{bmatrix} -3 & -4 & -2 & -1 \\ 0 & -4 & -2 & 2 \\ 0 & -3 & 6 & 9 \end{bmatrix}.$$

Izračunajmo $P_2$:

$$k = || \begin{bmatrix} -4 \\ -3 \end{bmatrix} || = 5,\ w_1 = \begin{bmatrix} -4 - 5 \\ -3 \end{bmatrix} = \begin{bmatrix} -9 \\ -3 \end{bmatrix},\ w_1^T w_1 = 90,$$

$$P_2 = I - \frac{1}{45} w_1 w_1^T = \frac{1}{5} \begin{bmatrix} -4 & -3 \\ -3 & 4 \end{bmatrix}.$$

Izračunajmo

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{-4}{5} & \frac{-3}{5} \\ 0 & \frac{-3}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} -3 & -4 & -2 & -1 \\ 0 & -4 & -2 & 2 \\ 0 & -3 & 6 & 9 \end{bmatrix} = \begin{bmatrix} -3 & -4 & -2 & -1 \\ 0 & 5 & -2 & -7 \\ 0 & 0 & 6 & 6 \end{bmatrix}.$$

Iz česar dobimo $x_3 = 1$, $x_2 = \frac{-7+2}{5} = -1$, $x_1 = \frac{-1+2-4}{3} = 1$.

Prvi stolpec je enak $z = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Givensova matrika

$$R_{1,2}^T = \begin{bmatrix} \frac{1}{\sqrt{1^2+2^2}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

uniči 2 komponento vektorja:

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \sqrt{5} \\ 0 \\ 2 \end{bmatrix} = z^{(1)}.$$

Z matriko

$$R_{1,3}^T = \begin{bmatrix} \frac{\sqrt{5}}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{bmatrix}$$

uničimo tretjo komponento $z^{(1)}$,

$$\begin{bmatrix} \frac{\sqrt{5}}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{bmatrix} \begin{bmatrix} \sqrt{5} \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}.$$

■

**Naloga 6.12** *Izračunaj QR razcep matrike*

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

*na različne načine:*

- *s pomočjo MGS,*
- *s pomočjo Givensovih rotacij.*

*Rešitev.* Delovna različica

- Najprej normiramo prvi stolpec $A$, $||a_1|| = \sqrt{1+4+1} = \sqrt{6}$. Torej je

$$q_1 = a_1/||a_1|| = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

in $r_{11} = ||a_1|| = \sqrt{6}$. V drugem koraku ortogonaliziramo drugi stolpec $A$ $a_2$ glede na $q_1$. Izračunamo $a_2 = a_2 - < q_1, a_2 > q_1$. Velja $< q_1, a_2 >= \frac{3}{\sqrt{6}}$. Tako dobimo $a_2 = \begin{bmatrix} -5/2 & 0 & 5/2 \end{bmatrix}^T$ in $r_{12} = \frac{3}{\sqrt{6}}$. Izračunamo $q_2 = a_2/||a_2|| = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}^T$ in $r_{22} = ||a_2|| = 5/\sqrt{2}$. Na koncu izračunamo še $a_3 = a_3 - < q_1, a_3 > q_1 - < q_2, a_3 > q_2$. Dobimo $r_{13} =< q_1, a_3 >= 2\sqrt{\frac{2}{3}}$, $r_{23} =< q_2, a_3 >= 0$, $r_{33} = ||a_3|| = \frac{1}{\sqrt{3}}$, $a_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T$.

- S pomočjo Givensove rotacije

$$R_{12}^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

uničimo 2. element v prvem stolpcu:

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{5} & 0 & 3/\sqrt{5} \\ 0 & \sqrt{5} & -1/\sqrt{5} \\ 1 & 3 & 1 \end{bmatrix}.$$

Z givensovo rotacijo

$$R_{13}^T = \begin{bmatrix} \sqrt{5/6} & 0 & 1/\sqrt{6} \\ 0 & 1 & 0 \\ -1/\sqrt{6} & 0 & \sqrt{5/6} \end{bmatrix}$$

uničimo 3. element v nanovo dobljeni matriki. Dobimo

$$\begin{bmatrix} \sqrt{5/6} & 0 & 1/\sqrt{6} \\ 0 & 1 & 0 \\ -1/\sqrt{6} & 0 & \sqrt{5/6} \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 & 3/\sqrt{5} \\ 0 & \sqrt{5} & -1/\sqrt{5} \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \sqrt{3/2} & \sqrt{3/2} + 1/\sqrt{6} \\ 0 & \sqrt{5} & -1/\sqrt{5} \\ 0 & \sqrt{15/2} & -\sqrt{3/10} + \sqrt{5/6} \end{bmatrix}.$$

Uporabimo še rotacijo

$$R_{23}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2/5} & \sqrt{3/5} \\ 0 & -\sqrt{3/5} & \sqrt{2/5} \end{bmatrix}$$

in dobimo isto matriko kot pri MGS, matrika $Q = R_{12}R_{13}R_{23}$.

■

**Naloga 6.13** *Naj bo $A \in \mathbb{C}^{m \times n}$ in $m \geq n$. Prevedi računanje singularnega razcepa matrike $A = U\Sigma V^H$ na računanje singularnega racepa v realnem primeru. Pokaži, da ima dobljeni realni sistem dvojne singualrne vrednosti.*

*Rešitev.* Naj bosta $u_i$ in $v_i$ $i$-ta stolpca ortogonalnih matrik $U$ in $V$. Matrika

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \\ & 0 & \\ & & \end{bmatrix}$$

ima na diagonali (realne) singularne vrednosti $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_n$. Potem velja

$$Av_i = U\Sigma \underbrace{V^H v_i}_{e_i} = U\Sigma e_i = U\sigma_1 e_i = \sigma_i u_i.$$

Vidimo, da velja $Av = \sigma u$, kjer je $u = u_i$, $v = v_i$, $\sigma = \sigma_i$. Upoštevamo, da za kompleksno matriko $A$ ter kompleksna vektorja velja:

$$\begin{aligned} A &= A_1 + iA_2, \\ u &= u_1 + iu_2, \\ v &= v_1 + iv_2. \end{aligned}$$

Ker je $\sigma$ realna dobimo

$$\begin{aligned} (A_1 + iA_2)(v_1 + iv_2) &= \sigma(u_1 + iu_2) \quad (6.1) \\ A_1v_1 - A_2v_2 + i(A_1v_2 + A_2v_1) &= \sigma u_1 + i\sigma u_2 \\ A_1v_1 - A_2v_2 &= \sigma u_1 \text{ (realni del)} \\ A_1v_2 + A_2v_1 &= \sigma u_2 \text{ (imaginarni del)}. \end{aligned}$$

Zadnji dve enačbi nam predstavljata realni bločni sistem

$$\begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} m \\ m \end{array} & \begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix} \end{array} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \sigma \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Dobljeni sistem je dvakrat večji, pokazati moramo še, da so singularne vrednosti dvojne. Pomnožimo 6.1 z $-i$ in dobimo

$$\begin{aligned} (A_1 + iA_2)(v_1 + iv_2) &= \sigma(u_1 + iu_2) \;/\cdot(-i) \\ A_1v_1 - A_2v_2 + i(A_1v_2 + A_2v_1) &= \sigma u_1 + i\sigma u_2 \;/\cdot(-i) \\ -(A_1v_1 - A_2v_2)i + A_1v_2 + A_2v_1 &= -i\sigma u_1 + \sigma u_2 \\ A_1v_2 + A_2v_1 &= \sigma u_2 \text{ (realni del)} \\ A_2v_2 - A_1v_1 &= -\sigma u_1 \text{ (imaginarni del)}. \end{aligned}$$

Torej velja

$$\begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} m \\ m \end{array} & \begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix} \end{array} \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix} = \sigma \begin{bmatrix} u_2 \\ -u_1 \end{bmatrix}.$$

Našli smo še drugi levi singularni vektor $\begin{bmatrix} u_2 \\ -u_1 \end{bmatrix}$.

■

**Naloga 6.14** *Dokaži, da za matriko $X$ z najmanjšo normo $||X||_F$, ki minimizira $||XA - I_n||_F$ velja $X = A^+$. Matika $A$ je dimenzije $m \times n$, kjer je $m \geq n$ in $rang(A) = r$.*

*Rešitev.* Naj bo podan singularni razcep matrike $A = U\Sigma V^T$ dimenzije $m \times n$ in ranga $r$. Tukaj sta $U$ in $V$ ortogonalni matriki dimenzij $m \times m$ in $n \times n$. Psevdoinverz matrike $A$ je enak $A^+ = V\Sigma^+U^T$, kjer je

$$\Sigma = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \end{array}, \quad \Sigma^+ = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{bmatrix} S^{-1} & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

in $S = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$. Lotimo se reševanja. Najprej zapišemo $X$ kot $X = VDU^T$, potem je $||X||_F = ||D||_F$ in velja

$$||XA - I||_F = ||VD\overbrace{U^TU}^{I}\Sigma V^T - I||_F \overset{V ortogonalna}{=} ||D\Sigma - I||_F.$$

Zadnji enačaj dobimo tako, da z leve pomnožimo z $V^T$, z desne pa z $V$. Matrika $V$ je ortogonalna, tako se Frobeniusova norma ne spremeni. Matriko $D$ predstavimo v bločni obliki

$$D = \begin{array}{c c} & \begin{array}{cc} r & m-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \end{array}.$$

Potem je $D\Sigma$ oblike

$$D\Sigma = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} D_{11}S & 0 \\ D_{21}S & 0 \end{bmatrix} \end{array}.$$

Če želimo, da bo $||D\Sigma - I||_F^2 = ||D_{11}S - I_r||_F^2 + ||D_{21}S||_F^2 + || - I_{n-r}||_F^2$ minimalna, moramo izbrati $D_11 = S^{-1}$ in $D_{21} = 0$. Radi bi tudi, da je $||D||_F = ||X||_F$ čim manjša, torej je očitno najbolje izbrati $D_{12} = D_{22} = 0$. Velja namreč $||D||_F^2 = ||D_{11}||_F^2 + ||D_{21}||_F^2 + ||D_{12}||_F^2 + ||D_{22}||_F^2$. Dobili smo, da mora veljat $D = \Sigma^+$, kar nam da $X = V\Sigma^+U^T = A^+$. ∎

# Poglavje 7

# Lastne vrednosti in vektorji

**Naloga 7.1** *Naj bo matrika $A \in \mathbb{R}^{n\times n}$ in $T$ njena Schurova forma. Velja zveza $A = QTQ^H$, kjer je $Q$ ortogonalna matrika in $T$ zgornje trikotna matrika. Naj bo $\lambda$ enostavna lastna vrednost matrike $A$. S pomočjo Schurove forme izračunaj pripadajoči lastni vektor.*

*Rešitev.* Za lastno vrednost in lastni vektor velja:

$$\begin{aligned} Ax = QTQ^H x &= \lambda x \\ QTQ^H x - \lambda \overbrace{QQ^H}^{I} x &= 0 \\ Q(T-\lambda I)Q^H x &= 0 \qquad /Q^H \cdot \\ (T-\lambda I)\underbrace{Q^H x}_{y} &= 0 \\ Ty &= \lambda y. \end{aligned}$$

Vektor $y$ je lastni vektor za $T$. Lastna vrednost $\lambda = t_{ii}$ je eden izmed diagonalnih elementov $T$. Oglejmo si obliko

$$(T - t_{ii}I)y = \begin{bmatrix} \lambda_1 - t_{ii} & * & * & * & * \\ & * & * & * & * \\ & & 0 & * & * \\ & & & * & * \\ & & & & \lambda_n - t_{ii} \end{bmatrix} \begin{bmatrix} y_1 \\ * \\ y_i \\ * \\ y_n \end{bmatrix} = 0.$$

Enakost si oglejmo po blokih.

$$T - \lambda I = \begin{array}{c c} & \begin{array}{ccc} i{-}1 & 1 & n{-}i \end{array} \\ \begin{array}{c} i{-}1 \\ 1 \\ n{-}i \end{array} & \begin{bmatrix} T_{11} - \lambda I & T_{12} & T_{13} \\ 0 & 0 & T_{23} \\ 0 & 0 & T_{33} - \lambda I \end{bmatrix} \end{array} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Iz tretje bločne vrstice dobimo $(T_{33} - \lambda I)y_3 = 0$, kar nam da $y_3 = 0$, saj je $(T_{33} - \lambda I)$ obrnljiva. Iz druge bločne vrstice pa dobimo $T_{23}y_3 = 0$, kar nam ne da nič novega. Iz prve bločne vrstice dobimo enačbo

$$(T_{11} - \lambda I)y_1 + T_{12}y_2 + T_{13}\overbrace{y_3}^{0} = 0,$$

iz česar dobimo

$$(T_{11} - \lambda I)y_1 = -T_{12}y_2.$$

Ker je $y_2$ skalar, lahko izberemo $y_2 = 1$. Torej je

$$y = \begin{bmatrix} -(T_{11} - \lambda)^{-1} T_{12} \\ 1 \\ 0 \end{bmatrix} = Q^H x.$$

Seveda sistem rešimo in ne računamo inverza. Na koncu dobimo $x = Qy$. ■

---

**Algoritem 7**: Potenčna metoda

---

$y = y^{(0)}$ začetni približek ;
$r = 0$ ;
**while** *premajhna natančnost* **do**
  $y^{(r+1)} = Ay^{(r)}$ ;
  $y^{(r+1)} = y^{(r+1)}/||y^{(r+1)}||_\infty$ normirana varianta;
  $r = r + 1$;
**end**

---

**Naloga 7.2** *Naj bo $A \in \mathbb{R}^{n\times n}$. Iščemo dominantne lastne vrednosti s potenčno metodo. Dokaži, da vektorji po smeri konvergirajo k dominantni lastni vrednosti. Ali lahko izluščimo lastne vektorje pripadjoče tem lastnim vrednostim? Obravnavaj naslednje primere.*

- *Dominantna lastna vrednost je ena sama, $|\lambda_1| > |\lambda_2|$.*
- *Dominantni lastni vrednosti sta dve, velja $\lambda_2 = \lambda_1$.*
- *Dominantni lastni vrednost sta dve, velja $\lambda_2 = -\lambda_1$.*

*Rešitev.* Primer $|\lambda_1| > |\lambda_2|$.
Naj bodo $x_1, x_2, \ldots, x_n$ pripadojoči lastni vektorji. Zapišimo $y^{(0)} = \sum_{i=1}^n \alpha_i x_i$. Velja $y^{(r)} = \sum_{i=1}^n \alpha_i \lambda_i^r x_i$. Tako dobimo

$$\frac{y_k^{(r+1)}}{y_k^{(r)}} = \frac{\sum_{i=1}^n \alpha_i \lambda_i^{r+1} x_{i(k)}}{\sum_{i=1}^n \alpha_i \lambda_i^r x_{i(k)}} = \frac{\alpha_1 \lambda_1 x_{1(k)} + \sum_{i=2}^n \alpha_i (\frac{\lambda_i}{\lambda_1})^r \lambda_i x_{i(k)}}{\lambda_1 x_{1(k)} + \sum_{i=2}^n \alpha_i (\frac{\lambda_i}{\lambda_1})^r x_{i(k)}}.$$

Ker je $|\frac{\lambda_i}{\lambda_1}| < 1$, velja

$$\lim_{r\to\infty} \frac{y_k^{(r+1)}}{y_k^{(r)}} = \lambda_1$$

in $y^{(r)}$ konvergira k $x_1$ po smeri.
Primer $\lambda_1 = \lambda_2$.
Podobno kot prej dobimo

$$\frac{y_k^{(r+1)}}{y_k^{(r)}} = \frac{\alpha_1 \lambda_1 x_{1(k)} + \alpha_2 \lambda_1 x_{2(k)} + \sum_{i=3}^n \alpha_i (\frac{\lambda_i}{\lambda_1})^r \lambda_i x_{i(k)}}{\alpha_1 x_{1(k)} + \alpha_2 x_{2(k)} \sum_{i=3}^n + \alpha_i (\frac{\lambda_i}{\lambda_1})^r x_{i(k)}}.$$

Torej velja

$$\lim_{r\to\infty} \frac{y_k^{(r+1)}}{y_k^{(r)}} = \lambda_1$$

in $y^{(r)}$ konvergira po smeri proti $\alpha_1 x_1 + \alpha_2 x_2$. Iz tega ne moremo dobiti lastnih vektorjev.
Primer $\lambda_2 = -\lambda_1$.
V tem primeru podzaporedje $z^{(r)} = y^{(2r)}$ konvergira po smeri k $\alpha_1 x_1 + \alpha_2 x_2$, faktor je $\lambda_1^2$. Podzaporedje $w^{(r)} = y^{(2r+1)}$ pa konvergira k $\alpha_1 x_1 - \alpha_2 x_2$. Iz $y^{(r)} \doteq \alpha x_1 + \beta x_2$ in $y^{(r+1)} \doteq \alpha\lambda_1 x_1 - \lambda_1 \beta x_2$. Iz česar dobimo

$$\begin{aligned}\lambda_1 y^{(r)} + y^{(r+1)} &= 2\alpha\lambda_1 x_1\\ \lambda_1 y^{(r)} - y^{(r+1)} &= 2\beta\lambda_1 x_2.\end{aligned}$$

■

**Naloga 7.3** *Gerschgorinov izrek.*
*Naj bo* $A \in \mathbb{C}^{n\times n}$ *in* $C_i = \{z \in \mathbb{C}_i, \quad |z - a_{ii}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}|\}$ *krog v kompleksni ravnini, za* $i = 1, \ldots, n.$. *Vse lastne vrednosti matrike* $A$ *ležijo v uniji krogov* $\cup_{i=1}^{n} C_i$.

*Rešitev.* Naj bo $x$ lastni vektor in $\lambda$ pripadajoča lastna vrednost. Poglejmo si enakost po vrsticah.

$$A(i,:)x = \sum_{j=1}^{n} a_{ij}x_j = \lambda x_i.$$

Kar je ekvivalentno

$$\sum_{j=1,j\neq i}^{n} a_{ij}x_j = (\lambda - a_{ii})x_i.$$

Uporabimo absolutno vrednost in ocenimo:

$$|\lambda - a_{ii}||x_i| = |\sum_{j=1,j\neq i}^{n} a_{ij}x_j| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}||x_j| \leq ||x||_\infty \sum_{j=1,j\neq i}^{n} |a_{ij}|.$$

Naj bo $|x_k| = ||x||_\infty$. Če postavimo $i = k$ dobimo,

$$|\lambda - a_{kk}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}|.$$

Kar pomeni $\lambda \in C_k$. Iz tega že sledi, da vsaka lastna vrednost leži v uniji krogov. ■

**Naloga 7.4** *Naj bo* $A \in \mathbb{R}^{n\times n}$ *diagonalno dominantna, tj.*

$$|a_{ii}| > \sum_{j=1,j\neq i}^{n} |a_{ij}| \quad \textit{za} \quad \textit{vsak} \quad i.$$

*Potem je* $A$ *obrnljiva.*

*Rešitev.* Dovolj je pokazati, da so vse lastne vrednosti različne od 0. Fiksirajmo lastno vrednost $\lambda$. Po izreku leži v vsaj enem krogu. Torej velja

$$|\lambda - a_{ii}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}|.$$

Ali je lahko $\lambda = 0$. Če bi bila, bi veljalo $|a_{ii}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}| < |a_{ii}|$. Kar je protislovje. Matrika je res obrnljiva. ■

**Posledica 7.1** *Naj bo $A \in \mathbb{R}^{n\times n}$ simetrična, vse lastne vrednosti so realne. Vsak lastna vrednost leži v enem od intervalov*

$$[a_{ii} - \sum_{j=1,j\neq i}^{n} |a_{ij}|, a_{ii} + \sum_{j=1,j\neq i}^{n} |a_{ij}|].$$

**Posledica 7.2** *Velja $|\lambda| \leq ||A||_\infty$.*

Dokaza posledic sta preprosta in ju prepuščam bralcu.

**Naloga 7.5** *Določi območje v katerem se nahajajo lastne vrednosti matrike*

$$A = \begin{bmatrix} 10 & 4 & 1 \\ 1 & 10 & 0.5 \\ 1.5 & -3 & 20 \end{bmatrix}.$$

*Uporabi Gerschgorinov izrek.*

*Rešitev.* Iz prve vrstice dobimo $|\lambda - 10| \leq 5$. Iz druge vrstice dobimo $|\lambda - 10| \leq 1.5$. Iz tretje vrstice dobimo $|\lambda - 20| \leq 4.5$. Oceno lahko izboljšamo tako, da isto oceno naredimo za podobno matriko $DAD^{-1}$, kjer je $D$ diagonalna matrika. ∎

Simetrične matrike.

**Algoritem 8**: Rayleighova iteracija

$\widetilde{z_0} \neq 0$;
$z_0 = \frac{1}{||z_0||_\infty}\widetilde{z_0}$;
**for** $k = 0, 1, 2, \ldots$ **do**
    $\sigma_k := \rho(z_k, A) = \frac{z_k^T A z_k}{z_k^t z_k}$;
    reši $(A - \sigma_k I)\widetilde{z_{k+1}} = z_k$;
    $z_{k+1} = \frac{1}{||z_{k+1}||_\infty}\widetilde{z_{k+1}}$ - v praksi;
    $z_{k+1} = \frac{1}{||z_{k+1}||_2}\widetilde{z_{k+1}}$ - za naslednjo naloga;
**end**

**Algoritem 9**: $QR$ iteracija z enojnim premikom

$A_0 = A$;
**for** $k = 0, 1, 2, \ldots$ **do**
    $\eta_k = (A_k)_{n,n}$;
    $A_k - \eta_k I = Q_k R_k$;
    $A_{k+1} = R_k Q_k + \eta_k I$;
**end**

**Naloga 7.6** *Če vzamemo pri Rayleighovi iteraciji za začetni vektor $z_0 = e_n$, potem velja $\sigma_k = \eta_k$ za vsak $k$.*

*Rešitev.* Trditev bomo dokazali z indukcijo. Za $k=0$ dobimo:

$$\sigma_0 = \rho(e_n, A) = (e_n^T A e_n)/e_n^T e_n = (A)_{n,n} = (A_0)_{(n,n)} = (A)_{n,n}.$$

Poiščimo še $z_1$. Upoštevajmo, da je $\widetilde{z_1}$ rešitev sistema $(A-\sigma_0 I)\widetilde{z_1} = e_n$. Iz prvega koraka $QR$ iteracije dobimo $A_0 - \sigma_0 I = Q_0 R_0$, kar je $(A_0-\sigma_0 I)Q_0 = R_0^T$. Iz tega sledi $(A_0-\sigma_0 I)Q_0 e_n = R_0^T e_n = (R_0)_n^T = * e_n$. Potem je $z_1 = Q_0 e_0$. Velja še $A_1 = R_0 Q_0 + \eta_0 I = Q_0^T (A_0 - \eta_0 I) Q_0 + \eta_0 I = Q_0^T A_0 Q_0 = Q_0^T A Q_0$. Naj bo naša indukcijska predpostavka $\sigma_k = \eta_k$, $z_k = Q_0 Q_1 \dots Q_{k-1} e_n$ in $A_k = Q_{k-1}^T A_{k-1} Q_{k-1}$.
Velja $(A-\sigma_k I)\widetilde{z_{k+1}} = z_k$ in $(A_k - \sigma_k I) Q_k e_n = R_k^T e_n = * e_n$. Upoštevamo še

$$A_0 = Q_0 \dots Q_{k-1} A_k Q_{k-1}^T \dots Q_0^T.$$

Iz $(A-\sigma_k I)\widetilde{z_{k+1}} = z_k$ izračunamo

$$(Q_0 \dots Q_{k-1} A_k Q_{k-1}^T \dots Q_0^T - \sigma_k I)\widetilde{z_{k+1}} = Q_0 \dots Q_{k-1} e_n,$$

iz česar dobimo

$$(A_k - \sigma_k I) Q_{k-1}^T \dots Q_0^T \widetilde{z_{k+1}} = e_n.$$

Iz enakosti dobljene iz $QR$ iteracije sklepamo

$$Q_{k-1}^T \dots Q_0^T \widetilde{z_{k+1}} = Q_k e_n,$$

kar nam da

$$z_{k+1} = Q_0 \dots Q_k e_n.$$

Zdaj lahko izračunamo

$$\sigma_{k+1} = \rho(z_{k+1}, A) = e_n^T \underbrace{Q_k^T \dots Q_0 A Q_0^T \dots Q_k^T}_{A_{k+1}} e_n = (A_{k+1})_{n,n} = \eta_{k+1}.$$

■

**Naloga 7.7** *Dokaži, da je* $\det(I + xy^T) = 1 + y^T x$.

*Rešitev.* Najprej povejmo elegantno rešitev. Definirajmo matriki

$$\begin{bmatrix} 1 & -y^T \\ x & I \end{bmatrix} \quad \text{in} \quad \begin{bmatrix} 1 & 0 \\ -x & I \end{bmatrix}.$$

Izračunamo

$$AB = \begin{bmatrix} 1 + y^T x & -y^T \\ 0 & I \end{bmatrix} \quad \text{in} \quad BA = \begin{bmatrix} 1 & -y^T \\ 0 & 1 + x^T y \end{bmatrix}.$$

Velja $\det(AB) = 1 + y^T x = \det(BA) = \det(I + x^T y)$.
Druga možnost je, da uporabimo operacije na vrsticah in stolpcih, ki ohranjajo determinanto.

$$I + xy^T = \begin{bmatrix} x_1 y_1 + 1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 + 1 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_n & x_n y_2 & \cdots & x_n y_n \end{bmatrix}.$$

Prvo vrstico pomnožimo z $\frac{x_i}{x_1}$ ter odštejemo od $i$-te vrstice. Dobimo matriko

$$\begin{bmatrix} 1+x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ -\frac{x_2}{x_1} & 1 & & \\ \vdots & & \ddots & \\ -\frac{x_n}{x_1} & & & 1 \end{bmatrix}.$$

Zdaj pomnožimo $j$-ti stolpec z $\frac{x_j}{x_1}$ in ga prištejemo prvemo stolpcu, dobimo

$$\begin{bmatrix} 1+\sum_{i=1}^n x_iy_i & x_1y_2 & \cdots & x_1y_n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

S tem smo končali. ■

To izkoristimo pri reševanju sekularne enačbe. Dobimo, da moramo rešiti enačbo

$$\det(D+\rho uu^T-\lambda I)=\det((D-\lambda I)(I+\rho(D-\lambda I)^{-1}uu^T)),$$

kar je ekvivalentno reševanju sekularne enačbe

$$\det(I+\rho(D-\lambda I)^{-1}uu^T)=1+\rho\sum_{i=1}^n\frac{u_i^2}{d_i-\lambda}=0.$$

Če je $\alpha$ lastna vrednost $D+\rho uu^T$, je $(D-\alpha I)^{-1}u$ ustrezni lastni vektor.

**Naloga 7.8** *Reševanje sekularne enačbe. Iščemo ničle funkcije*

$$f(\lambda)=1+\rho\sum_{i=1}^n\frac{u_i^2}{d_i-\lambda}.$$

*Njen odvod je*

$$f'(\lambda)=\rho\sum_{i=1}^n\frac{u_i^2}{(d_i-\lambda)^2}.$$

*Za $\rho>0$ je $f'>0$ in funkcija ima asimptoto* 1.

*Recimo, da iščemo rešitev na intervalu $(d_i,d_{i+1})$, kjer je začetni približek $x_r$. Navadne tangentne metode ne moremo porabiti, saj so ničle zelo blizu polov, zato bi potrebovali zelo dober začetni približek. Namesto aproksimacije funkcije s tangento raje uporabimo preprosto racionalno funkcijo, ki se prilega funkciji $f$ na tem intervalu. Poiščemo racionalno funkcijo oblike*

$$h(\lambda)=\frac{c_1}{d_i-\lambda}+\frac{c_2}{d_{i+1}-\lambda}+c_3,$$

*za katero velja $h(x_r)=f(x_r)$ in $f'(x_r)=h'(x_r)$. Zaradi stabilnosti razdelimo $f$ na dva dela kot*

$$f(\lambda)=1+\rho\sum_{k=1}^i\frac{u_k^2}{d_k-\lambda}+\rho\sum_{k=i+1}^n\frac{u_k^2}{d_k-\lambda}=:1+\psi_1(\lambda)+\psi_2(\lambda).$$

*To naredimo tako, da v vsoti $\psi_1(\lambda)$ oziroma $\psi_2(\lambda)$ seštevamo enako predznačene člene. Sedaj določimo $c_1,c_1'$ tako da za*

$$h_1(\lambda)=\frac{c_1}{d_i-\lambda}+c_1'$$

*velja* $h_1(x_r) = \psi_1(x_r)$ *in* $h'(x_r) = \psi_1'(x_r)$. *Podobno določimo tudi* $c_2$ *in* $c_2'$ *tako, da za*

$$h_2(\lambda) = \frac{c_2}{d_{i+1} - \lambda} + c_2'$$

*velja* $h_2(x_r) = \psi_2(x_r)$ *in* $h_2'(x_r) = \psi_2'(x_r)$. *Sedaj je* $h(\lambda) = 1 + h_1(\lambda) + h_2(\lambda)$ *iskana racionalna funkcijo. Enačba* $h(\lambda) = 0$ *ima dve rešitvi, za nov približek* $x_{r+1}$ *vzamemo tisto, ki leži na intervalu* $(d_{i+1}, d_i)$. *Oglejmo si primer v Mathematici.*

**Naloga 7.9** *Poišči vektor* $v$, *tako da za*

$$T = \begin{bmatrix} a_1 & b_1 & & 0 \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ 0 & & b_{n-1} & a_n \end{bmatrix}$$

*in*

$$T_1 = \begin{bmatrix} a_1 & b_1 & & 0 \\ b_1 & \ddots & \ddots & \\ & \ddots & a_{k-1} & b_{k-1} \\ 0 & & b_{k-1} & a_k - b_k \end{bmatrix}, \quad T_2 = \begin{bmatrix} a_{k+1} - b_k & b_{k+1} & & 0 \\ b_{k+1} & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ 0 & & b_{n-1} & a_n \end{bmatrix}$$

*velja*

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} + b_k v v^T.$$

*Rešitev.* V originalni matriki $T$ sta se spremenili le vrstici in stolpca $k$ in $k+1$. Veljati mora

$$\begin{bmatrix} a_k & b_k \\ b_k & a_{k+1} \end{bmatrix} = \begin{bmatrix} a_k - b_k & 0 \\ 0 & a_{k+1} - b_k \end{bmatrix} + b_k z z^T.$$

Torej mora veljati $z = [1, 1]^T$. Tako je potem $v = \begin{array}{c} \begin{array}{ccccccc} & & k & k+1 & & & \end{array} \\ \begin{bmatrix} 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \end{bmatrix}^T \end{array}$. ■

**Izrek 7.1** *Courant-Fisherjev minimax izrek.*
*Za simetrično matriko* $A \in \mathbb{R}^{n \times n}$ *velja*

$$\lambda_k(A) = \min_{R \subseteq \mathbb{R}^n, \dim(R) = n-k+1} \max_{x \in R, x \neq 0} \rho(x, A) = \max_{S \subseteq \mathbb{R}^n, \dim(S) = k} \min_{x \in S, x \neq 0} \rho(x, A).$$

**Izrek 7.2** *Cauchyjev izrek o prepletanju.*
*Če je* $A$ *simerična matrika in* $A_r$ *glavna* $r \times r$ *podmatrika matrike* $A$, *potem velja*

$$\lambda_{r+1}(A_{r+1}) \leq \lambda_r(A_r) \leq \lambda_r(A_{r+1}) \ldots \leq \lambda_2(A_{r+1}) \leq \lambda_1(A_r) \leq \lambda_1(A_{r+1}.$$

**Naloga 7.10** *Naj bosta* $A, E \in \mathbb{R}^{n \times n}$ *simetrični, da velja*

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A) \qquad \text{in} \qquad \lambda_1(E) \geq \lambda_2(E) \geq \ldots \geq \lambda_n(E).$$

*Dokaži, da potem velja*

$$\lambda_k(A) + \lambda_n(E) \leq \lambda_k(A + E) \leq \lambda_k(A) + \lambda_1(E).$$

*Uporabi Courant-Fisherjev minimax izrek.*

*Rešitev.* Velja

$$\lambda_k(A) = \min_{R\subseteq\mathbb{R}^n,\dim(R)=n-k+1} \max_{x\in R, x\neq 0} \rho(x,A),$$

kjer je $\rho(x, A+E) = \frac{x^T(A+E)x}{x^Tx} = \rho(x,A) + \rho(X,E)$. Dobimo

$$\lambda_k(A+E) = \min_{R\subseteq\mathbb{R}^n,\dim(R)=n-k+1} \max_{x\in R, x\neq 0} (\rho(x,A)+\rho(x,E)) \leq$$

$$\leq \min_{R\subseteq\mathbb{R}^n,\dim(R)=n-k+1} (\max_{x\in R, x\neq 0} \rho(x,A)+\rho(x,E)).$$

Velja tudi

$$\lambda_n(E) \leq \rho(x,E) \leq \lambda_1(E).$$

Če vzamemo $\max \rho(x,E)$ po celem prostoru, dobimo kvečjemu več. Torej velja $\lambda_k(A+E) \leq \lambda_k(A) + \lambda_1(E)$.

Za drugi del neenakosti lahko uporabimo drugi del minimax izreka. Lahko pa uporabimo pravkar dokazano neenakost za $-A$ in $-E$. To nam da

$$\lambda_k(-A-E) = -\lambda_{n-k+1}(A+E) \leq \lambda_k(-A) + \lambda_1(-E) = -(\lambda_{n-k+1}(A) + \lambda_n(E)).$$

Če neenakost pomnožimo z $-1$, smo končali. ■

**Naloga 7.11** *Naj bo* $\rho(x,A) = \frac{x^TAx}{x^Tx}$ *Rayleighov kvocient simetrične matrike A. Dokaži, da velja*

*(i). Vrednost* $\beta = \rho(x,A)$ *minimizira* $\min_\beta ||Ax - \beta x||_2$.

*(ii). Poišči* $\lambda$ *in* $\mu$, *ki minimizira* $\min_{\lambda,\mu} ||Ax - \lambda Bx - \mu Cx||_2$.

*Vse matrike so simetrične.*

*Rešitev.*

(i). Problem najlažje rešimo, če prepoznamo, da gre v bistvu za predoločen sistem

$$x\beta = Ax = b.$$

Tako dobimo normalni sistem $x^Tx\beta = x^TAx$, kar nam že da željeno rešitev.

Rešitev lahko dobimo tudi z iskanjem maksimuma skalarne funkcije $f(\beta) = ||Ax - \beta x||_2$. Lahko pa pokažemo tudi, da velja $x \perp Ax - \rho(x,A)x$.

(ii). Problem spet poizkusimo prevesti na reševanje predoločenega sistema.

$$\begin{aligned} Ax &= \lambda Bx + \mu Cx \\ \begin{bmatrix} Bx & Cx \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} &= Ax \quad \text{predoločen sistem} \\ \begin{bmatrix} (Bx)^T \\ (Cx)^T \end{bmatrix} \begin{bmatrix} Bx & Cx \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} &= \begin{bmatrix} (Bx)^T \\ (Cx)^T \end{bmatrix} Ax. \end{aligned}$$

Rešitev dobimo tako, da rešimo sistem

$$\begin{bmatrix} <Bx,Bx> & <Bx,Cx> \\ <Cx,Bx> & <Cx,Cx> \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} <Bx,Ax> \\ <Cx,Ax> \end{bmatrix}.$$

■

**Naloga 7.12** *Za simetrično tridiagonalno matriko*

$$A = \begin{bmatrix} d_1 & u_1 & & & & \\ u_1 & d_2 & u_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & u_{n-2} & d_{n-1} & u_{n-1} \\ & & & & u_{n-1} & d_n \end{bmatrix}$$

*definiramo $p_i(\lambda)$ kot karakteristični polinom glavne podmatrike $i \times i$. Pokaži, da velja naslednja zveza*

$$p_i(\lambda) = (d_i - \lambda)p_{i-1}(\lambda) - u_{i-1}^2 p_{i-2}(\lambda), \text{ za } i = 2, \ldots, n,$$

*kjer je $p_0(\lambda) = 1$ in $p_1(\lambda) = d_1 - \lambda$.*

*Rešitev.* Trditev očitno velja za matrike dimenzije 1 in 2. Naš indukcijska predpostavka bo, da trditev velja za vse matrike do dimnezije $n - 1$. Oglejmo si, kaj dobimo ko razvijemo

$$\begin{vmatrix} d_1 - \lambda & u_1 & & & & \\ u_1 & d_2 - \lambda & u_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & u_{n-2} & d_{n-1} - \lambda & u_{n-1} \\ & & & & u_{n-1} & d_n - \lambda \end{vmatrix}$$

po zadnjem stolpcu. Dobimo ravno

$$\det(A_{n-1} - \lambda I)(d_n - \lambda) - u_{n-1} \begin{vmatrix} d_1 - \lambda & u_1 & & & & \\ u_1 & d_2 - \lambda & u_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & u_{n-3} & \\ & & & u_{n-3} & d_{n-2} - \lambda & u_{n-2} \\ & & & & & u_{n-1} \end{vmatrix}.$$

Iz česar pa že dobimo rekurzivno zvezo. ∎

**Naloga 7.13** ***Sturmovo zaporedje in bisekcija.***
*Naj bo $A$ simetrična tridiagonalna irreducibilna matrika, noben diagonalni element ni enak 0. Potem je število ničel $p_n$ (lastnih vrednosti $A$) na intervalu $[a, \infty)$ enako številu ujemanj predznaka v zaporedju*

$$\overbrace{p_0(a), p_1(a), \ldots, p_{n-1}(a), p_n(a)}^{P(a)}.$$

*Če velja $p_{k+1}(a) = 0$, to štejemo kot ujemanje predznaka z $p_k(a)$. Na vajah smo definirali malo drugače. Število ničel na intervalu $[a, b]$ je enako število sprememb predznaka v zaporedju $P(b)$ - število sprememb predznaka v zaporedju $P(a)$. Če se v sredini zaporedju nahaja 0, jo iz zaporedja izpustimo. To je dualna definicija k ujemanjem predznaka.*
*Podana je matrika*

$$T = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

*Polinomi so enaki* $p_0(\lambda) = 1, p_1(\lambda) = 2 - \lambda, p_2(\lambda) = (2-\lambda)^2 - 1$ *in*

$$p_3(\lambda) = (2-\lambda)^3 - 2(2-\lambda).$$

*Na tej matriki si oglejmo idejo bisekcije s Sturmovim zaporedjem.*

*Rešitev.* Začnimo z dovolj velikim začetnim intervalom. Vse lastne vrednosti zagotovo ležijo na intervalu $[a,b] = [-||T||_\infty, ||T||_\infty] = [-4,4]$. Zaporedje $P(4)$ je enako $1,-2,3,-4$ zaporedje $P(-4)$ je enako $1,6,35,204$. Tako dobimo, da so na intervalu 3 ničle. Nadaljujemo z $[0,4]$. Dobimo zaporedje $P(0)$ $1,2,3,4$ Torej so na intervalu $[0,4]$ 3 ničle. Spet razpolovimo interval in dobimo zaporedje $P(2)$ $1,0,-1,0$. Na intrevalu $[2,4]$ so tako 2 ničli. Nadaljujemo ... ■

Na vsakem koraku klasične Jacobijeve metode uničujemo izvendiagonalne elemente simetrične matrike $A$.

**Algoritem 10**: Jacobijeva metoda

Ponavljaj: Poišči $j,k$, tako da je $a_{jk}$ največji izvendiagonalni element po $|\dots|$.
Izberi Jacobijev rotacijo $J$, ki uniči $a_{jk}$. $J$ je ortogonalna matrika, ki spremeni le $j$-to in $k$-to vrstico. $A' = J^T AJ$.

**Naloga 7.14** *Pokaži, da klasična Jacobijeva metoda konvergira vsaj linearno. Najprej pokaži, da velja*

$$off^2(B) = off^2(A) - 2a_{jk}^2.$$

*Kjer je* $off$ *koren vsote kvadratov elementov matrike brez diagonale in* $B = J^T AJ$.

*Rešitev.* Spomnimo se, da se z ortogonalno transformacijo Frobeniusova norma matrike ohranja. Torej sta Frobeniusovi normi matrik $A$ in $B$ enaki. Frobeniusova norma je ravno koren vsote kvadratov elementov matrike. Torej dobimo

$$off(B)^2 = ||B||_F^2 - \sum_{i=1}^n b_{ii}^2 = ||A||_F^2 - \sum_{i=1}^n b_{ii}^2.$$

Upoštevamo, da so se spremenila le $j$-ti in $k$-ti stolpec ter vrstica. Torej sta se spremenila le $j$-ti in $k$-ti diagonalni element. Velja

$$\begin{bmatrix} b_{jj} & b_{jk} \\ b_{kj} & b_{kk} \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_{jj} & a_{jk} \\ a_{kj} & a_{kk} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Ker gre za ortogonalno transformacijo, velja enakost Frobeniusovih norm

$$a_{jj}^2 + a_{kk}^2 + 2a_{jk}^2 = b_{jj}^2 + b_{kk}^2 + 2b_{jk}^2.$$

Nadaljujmo

$$off^2(B) = ||A||_F^2 - \sum_{i=1, i\neq j,k}^n -b_{jj}^2 - b_{kk}^2 = off^2(A) - 2a_{jk}^2.$$

Dokažimo še, da je

$$off(B) \le \sqrt{(1 - \frac{1}{N})} off(A),$$

kjer je $N = \frac{n(n-1)}{2}$. Vemo, da je

$$off^2(A) = \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} a_{ij}^2 \leq 2N a_{jk}^2,$$

saj je bil $a_{jk}$ izbran kot maksimalen izvendiagonalni element. Iz tega dobimo oceno

$$2a_{jk}^2 \geq \frac{off^2(A)}{N} \geq 0.$$

Kar nam že, da želeno.

Po $r$ korakih torej velja

$$off(B) \leq (1 - \frac{1}{N})^{\frac{r}{2}}.$$

Konvergenca je res vsaj linearna.

■

# Poglavje 8

# Iterativno reševanje linearnih sistemov

**Naloga 8.1** *Pokaži naslednjo trditev. Če Jacobijeva metodo konvergira za sistem* $Ax = b$, *potem konvergira tudi za* $A^T x = d$.

*Rešitev.* Spomnimo se iteracijske matrike Jacobijeve metode za reševanje linearnih sistemov. Rešujemo sistem $Ax = b$. Matriko $A$ razdelimo na spodnji trikotnik brez diagonale $L$, diagonalo $D$ in zgornji trikotnik brez diagonale $U$. Torej iz $(D + L + U)x = b$ dobimo iteracijo $x^{(1)} = -D^{-1}(L + U)x^{(0)} + D^{-1}b$. Iteracijska matrika je $R_J = -D^{-1}(L + U)$, metoda konvergira, natanko takrat ko je $\rho(R_J) < 1$. Pokazali bomo, da velja $\rho(R_J)(A) = \rho(R_J)(A^T)$. Izračunajmo $R_J(A^T) = -D^{-1}(L^T + U^T)$ in transponirajmo. Dobimo $R_J(A^T)^T = -(L+U)D^{-1}$. Če izrazimo $-(L + U)$ iz obeh iteracijskih matrik, dobimo

$$R_J(A^T)^T D = DR_J(A), \quad R_J(A) = D^{-1}R_J(A^T)D.$$

Torej sta $R_J(A^T)^T$ in $R_J(A)$ podobni matriki, tako imata enak spektralni radij. Prav tako pa imata enak spektralni radij transponirani matriki. Torej je $\rho(R_J(A)) = \rho(R_J(A^T))$. ∎

**Posledica 8.1** *Če je* $A$ *po stolpcih diagonalno dominantna matrika, potem Jacobijeva metoda konvergira.*

*Rešitev.* Iz izreka iz predavanj vemo, da je to res za matriko, ki je diagonalno dominantna po vrsticah. Taka pa je matrika $A^T$, saj je $A$ diagonalno dominantna po stolpcih. ∎

---

**Algoritem 11**: Jacobijeva iteracija

---

$x^{(0)}$ začetni približek;
Ponavljaj

$$x_{k+1}^{(r+1)} = \frac{1}{a_{kk}} \left( b_k - \sum_{i=1, i \neq k}^{n} a_{ki} x_i^{(r)} \right)$$

Krajše $R_J = -D^{-1}(L + U), \quad c = D^{-1}b$;
$x^{(r+1)} = R_J x^{(r)} + c$;

---

**Algoritem 12**: Gauss-Seidlova metoda

$x^{(0)}$ začetni približek;
Ponavljaj

$$x_{k+1}^{(r+1)} = \frac{1}{a_{kk}}\left(b_k - \sum_{i=1}^{k-1} a_{ki}x_i^{(r+1)} - \sum_{i=k+1}^{n} a_{ki}x_i^{(r)}\right)$$

Krajše $R_{GS} = -(L+D)^{-1}U, \quad c = (L+D)^{-1}b$;
$x^{(r+1)} = R_{GS}x^{(r)} + c$;

**Algoritem 13**: SOR metoda

$x^{(0)}$ začetni približek;
Izberi $\omega \in (0,2)$;
Ponavljaj

$$x_{k+1}^{(r+1)} = (1-\omega)x_k^{(r)} + \omega\frac{1}{a_{kk}}\left(b_k - \sum_{i=1}^{k-1} a_{ki}x_i^{(r+1)} - \sum_{i=k+1}^{n} a_{ki}x_i^{(r)}\right)$$

Krajše $R_{SOR} = -(\omega L + D)^{-1}((1-\omega)D - \omega U)$;

**Naloga 8.2** *Preuredi sistem enačb*

$$\begin{array}{ccccccc} 2x_1 & + & 4x_2 & + & 3x_3 & = & 1 \\ x_1 & + & x_2 & + & 4x_3 & = & 4 \\ 4x_1 & + & x_2 & & & = & 3 \end{array},$$

*tako da bo Jacobijeva metoda zagotovo konvergirala. Izračunaj prvi približek z začetnim vektorjem* $x = \begin{bmatrix}1.5 & -0.5 & 1\end{bmatrix}^T$.

*Rešitev.* Matrične vrstice preuredimo tako, da bo nova matrika diagonalno dominantna po stolpcih. Rešujemo ekvivalenten problem

$$\begin{array}{ccccccc} 4x_1 & + & x_2 & & & = & 3 \\ 2x_1 & + & 4x_2 & + & 3x_3 & = & 1 \\ x_1 & + & x_2 & + & 4x_3 & = & 4 \end{array}.$$

Nova matrika je diagonalno dominantna po stolpcih, velja $|4| > |2| + |1|$, $|4| > |1| + |1|$ ter $|4| > |3| + |0|$. Izračunajmo

$$\begin{aligned} x_1^{(1)} &= \frac{1}{4}(3 - x_2^{(0)} - x_3^{(0)}) = \frac{1}{4}(3+0.5) = \frac{7}{8} = 0.875 \\ x_2^{(1)} &= \frac{1}{4}(1 - 2x_1^{(0)} - 3x_3^{(0)}) = \frac{1}{4}(1 - 2\frac{3}{2} - 3) = -\frac{5}{4} = -1.25 \\ x_3^{(1)} &= \frac{1}{4}(4 - x_1^{(0)} - x_2^{(0)}) = \frac{1}{4}(4 - \frac{3}{2} + \frac{1}{2}) = \frac{3}{4} = 0.75. \end{aligned}$$

■

**Naloga 8.3** *Naj bo $||R|| < 1$ in poleg tega je matrična norma usklajena z vektorsko. Dokaži, da velja*

*(i).*

$$||x^{(r+1)} - x^{(r)}|| \leq ||R||||x^{(r)} - x^{(r-1)}||$$

*(ii).*

$$||x^{(r)} - x^*|| \leq \frac{||R||}{1 - ||R||}||x^{(r)} - x^{(r-1)}||$$

*(iii).*

$$||x^{(r)} - x^*|| \leq \frac{||R||^{(r)}}{1 - ||R||}||x^{(1)} - x^{(0)}||,$$

*kjer je $x^*$ točna rešitev. Velja še $x^{(r+1)} = Rx^{(r)} + c$.*

*Rešitev.*

(i). Izračunajmo

$$||x^{(r+1)} - x^{(r)}|| = ||Rx^{(r)} + c - Rx^{(r-1)} - c|| =$$
$$||Rx^{(r)} - Rx^{(r-1)}|| \leq ||R||||x^{(r)} - x^{(r-1)}||.$$

(ii). Velja

$$||x^{(r)} - x^*|| = ||x^{(r)} - x^{(r+1)} + x^{(r+1)} - x^{(r+2)} + \ldots + x^*|| \leq$$
$$||x^{(r)} - x^{(r+1)}|| + ||x^{(r+1)} - x^{(r+2)}|| + \ldots \leq$$
$$||R||||x^{(r)} - x^{(r-1)}|| + ||R||^2||||x^{(r+1)} - x^{(r)}|| + \ldots \leq$$
$$\left(||R|| + ||R||^2 + \ldots\right)||||x^{(r)} - x^{(r-1)}|| =$$
$$||R||\left(1 + ||R|| + ||R||^2 + \ldots\right)||||x^{(r)} - x^{(r-1)}|| = ||R||\frac{1}{1 - ||R||}||x^{(r)} - x^{(r+1)}||.$$

Vse dobimo z zaporedno uporabo (i).

(iii). Iz (ii) dobimo

$$||x^{(r)} - x^*|| \leq \frac{||R||}{1 - ||R||}||x^{(r)} - x^{(r-1)}||$$
$$\leq \frac{||R||^2}{1 - ||R||}||x^{(r-1)} - x^{(r-2)}|| \leq \ldots \leq \frac{||R||^r}{1 - ||R||}||x^{(1)} - x^{(0)}||.$$

■

**Naloga 8.4** *Izračunaj prva dva približka sistema*

$$\begin{array}{rcrcrcl} 10x_1 & - & 2x_2 & + & 4x_3 & = & 8 \\ -x_1 & + & 8x_2 & + & 2x_3 & = & 8 \\ x_1 & - & 3x_2 & + & 10x_3 & = & 5 \end{array},$$

*po Jacobijevi in Gauss-Seidlovi metodi z začetnim približkom $\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$. Oceni koliko korakov iteracije potebuješ, da bo veljalo $||x^{(r)} - x^*||_\infty \leq 10^{-5}$.*

*Rešitev.* Matrika je enaka

$$A = \begin{bmatrix} 10 & -2 & 4 \\ -1 & 8 & 2 \\ 1 & -3 & 10 \end{bmatrix}.$$

Vektorja sta

$$b = \begin{bmatrix} 8 \\ 8 \\ 5 \end{bmatrix}, \quad x^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Velja ocena

$$||x^{(r)} - x^*|| \leq \frac{||R||^{(r)}}{1 - ||R||} ||x^{(1)} - x^{(0)}||,$$

oziroma

$$||x^{(r)} - x^*|| \leq \frac{||R||^{(r-1)}}{1 - ||R||} ||x^{(2)} - x^{(1)}||.$$

To smo dokazali.

Najprej izračunajmo približka za Jacobijevo metodo

$$x_{k+1}^{(r+1)} = \frac{1}{a_{kk}} \left( b_k - \sum_{i=1, i \neq k}^{n} a_{ki} x_i^{(r)} \right).$$

$$\begin{aligned} x_1^{(1)} &= \frac{1}{10}(8 + 2x_2^{(0)} - 4x_3^{(0)}) = \frac{1}{10}(8 + 2) = 1 \\ x_2^{(1)} &= \frac{1}{8}(8 + x_1^{(0)} - 2x_3^{(0)}) = \frac{1}{8}(8 + 1) = 1.125 \\ x_3^{(1)} &= \frac{1}{10}(5 - x_1^{(0)} + 3x_2^{(0)}) = \frac{1}{10}(5 - 1 + 3) = 0.7 \end{aligned}$$

Dobili smo $x^{(1)} = \begin{bmatrix} 1 & 1.125 & 0.7 \end{bmatrix}^T$.

$$\begin{aligned} x_1^{(2)} &= \frac{1}{10}(8 + 2x_2^{(1)} - 4x_3^{(1)}) = \frac{1}{10}(8 + 2 \cdot 1.125 - 4 \cdot 0.7) = 0.745 \\ x_2^{(2)} &= \frac{1}{8}(8 + x_1^{(1)} - 2x_3^{(1)}) = \frac{1}{8}(8 + 1 - 2 \cdot 0.7) = 0.95 \\ x_3^{(2)} &= \frac{1}{10}(5 - x_1^{(1)} + 3x_2^{(1)}) = \frac{1}{10}(5 - 1 + 3 \cdot 1.125) = 0.7375 \end{aligned}$$

Dobili smo $x^{(2)} = \begin{bmatrix} 0.745 & 0.95 & 0.7375 \end{bmatrix}^T$. Za oceno rabimo še $||R_J||_\infty$. Izračunajmo

$$R_J = -D^{-1}(L + U) = -\begin{bmatrix} \frac{1}{10} & & \\ & \frac{1}{8} & \\ & & \frac{1}{10} \end{bmatrix} \begin{bmatrix} 0 & -2 & 4 \\ -1 & 0 & 2 \\ 1 & -3 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & -\frac{1}{5} & \frac{2}{5} \\ -\frac{1}{8} & 0 & \frac{1}{4} \\ \frac{1}{10} & -\frac{3}{10} & 0 \end{bmatrix}.$$

Neskončna norma je ravno največja vsota absolutnih vrednosti po vrsticah, $||R_J||_\infty = \frac{1}{5} + \frac{2}{5} = 0.6$. Izračunajmo

$$||x^{(2)} - x^{(1)}||_\infty = || \begin{bmatrix} -0.255 \\ -0.175 \\ 0.0375 \end{bmatrix} ||_\infty = 0.225.$$

Upoštevajmo oceno in dobimo

$$||x^{(r)} - x^*||_\infty \leq \frac{0.6^{r-1}}{1-0.6}||x^{(2)} - x^{(1)}||_\infty = \frac{0.6^{r-1}}{0.4}0.225 \leq 10^{-5}.$$

Tako dobimo, da mora veljati

$$0.6^{r-1} \leq \frac{0.4 \cdot 10^{-5}}{0.225}.$$

Z antilogaritmiranjem dobimo $r \geq 22.411 \doteq 23$.

Izračunajmo še približka za Gauss-Seidlovo metodo

$$x_{k+1}^{(r+1)} = \frac{1}{a_{kk}} \left( b_k - \sum_{i=1}^{k-1} a_{ki} x_i^{(r+1)} - \sum_{i=k+1}^{n} a_{ki} x_i^{(r)} \right)$$

$$\begin{aligned}
x_1^{(1)} &= \frac{1}{10}(8 + 2x_2^{(0)} - 4x_3^{(0)}) = \frac{1}{10}(8+2) = 1 \\
x_2^{(1)} &= \frac{1}{8}(8 + x_1^{(1)} - 2x_3^{(0)}) = \frac{1}{8}(8+1) = 1.125 \\
x_3^{(1)} &= \frac{1}{10}(5 - x_1^{(1)} + 3x_2^{(1)}) = \frac{1}{10}(5 - 1 + 3 \cdot 1.125) = 0.7375
\end{aligned}$$

Dobili smo $x^{(1)} = \begin{bmatrix} 1 & 1.125 & 0.7375 \end{bmatrix}^T$.

$$\begin{aligned}
x_1^{(1)} &= \frac{1}{10}(8 + 2x_2^{(1)} - 4x_3^{(1)}) = \frac{1}{10}(8 + 2 \cdot 1.125 - 4 \cdot 0.7375) = 0.73 \\
x_2^{(1)} &= \frac{1}{8}(8 + x_1^{(2)} - 2x_3^{(1)}) = \frac{1}{8}(8 + 0.73 - 2 \cdot 0.7375) = 0.906875 \\
x_3^{(1)} &= \frac{1}{10}(5 - x_1^{(2)} + 3x_2^{(2)}) = \frac{1}{10}(5 - 0.73 + 3 \cdot 0.906875) = 0.6990625
\end{aligned}$$

Dobili smo $x^{(2)} = \begin{bmatrix} 0.73 & 0.906875 & 0.6990625 \end{bmatrix}^T$. Za oceno rabimo še $||R_{GS}||_\infty$. Izračunajmo

$$R_{GS} = -(D+L)^{-1}U = -\begin{bmatrix} 10 & 0 & 0 \\ -1 & 8 & 0 \\ 1 & -3 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -2 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0.2 & -0.4 \\ 0 & 0.025 & -0.3 \\ 0 & -0.0125 & -0.05 \end{bmatrix}.$$

Tako dobimo $||R_{GS}||_\infty = 0.6$, $||x^{(2)} - x^{(1)}||_\infty = 0.27$. Hočemo, da velja

$$||x^{(r)} - x^*||_\infty \leq \frac{0.6^{r-1}}{1-0.6}||x^{(2)} - x^{(1)}||_\infty = \frac{0.6^{r-1}}{0.4} \cdot 0.27.$$

Podobno kot prej z antilogaritmiranjem dobimo $r \geq 22.768 \doteq 23$. ∎

**Naloga 8.5** *Naj bo $B$ kvadratna matrika dimenzije $n \times n$ s singularnimi vrednostmi in vektorji $Bv_i = \sigma_i u_i$, za $i = 1, \ldots, n$. Lastne vrednosti matrike* $\begin{array}{cc} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} 0 & B^T \\ B & 0 \end{bmatrix} \end{array}$ *so $\pm\sigma_i$, ustrezni lastni vektorji pa $\frac{1}{\sqrt{2}} \begin{bmatrix} v_i \\ \pm u_i \end{bmatrix}$, $i = 1, \ldots, n$.*

*Rešitev.* Naj bo $B = U\Sigma V^T$, potem je $B^T = V\Sigma U^T$. Torej velja tudi $B^T u_i = \sigma_i v_i$, saj sta matriki $U$ in $V$ ortogonalni. Izračunajmo

$$\begin{bmatrix} 0 & B^T \\ B & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} v_i \\ \pm u_i \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \pm B^T u_i \\ B v_i \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \pm\sigma_i v_i \\ \sigma_i u_i \end{bmatrix} = \pm\sigma_i \frac{1}{\sqrt{2}} \begin{bmatrix} v_i \\ \pm u_i \end{bmatrix}.$$

Torej velja

$$B \begin{bmatrix} \frac{1}{\sqrt{2}} U & \frac{1}{\sqrt{2}} U \\ \frac{1}{\sqrt{2}} V & \frac{1}{\sqrt{2}} - V \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} U & \frac{1}{\sqrt{2}} U \\ \frac{1}{\sqrt{2}} V & \frac{1}{\sqrt{2}} - V \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}.$$

■

**Naloga 8.6** *Naj bo $\Gamma$ $k$-regularen graf z $n$ vozlišči. Potem je vektor*

$$\mathbf{1} = {}_n \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix},$$

*dimenzije $n$, lastni vektor matrike sosednosti grafa. Ustrezna pripadajoča lastna vrednost je $k$.*

*Rešitev.* Graf je $k$-regularen, torej ima vsako vozlišče $k$ sosedov. V vsaki vrstici matrike sosednosti je torej natanko $k$ enk, drugod pa so ničle. Če matrično vrstico pomnnožimo z vektorjem samih enk, dobimo natanko $k$. Neničelne produkte dobimo samo, ko zmnožimo dve enki. Torej je $\mathbf{1}$ res lastni vektor z lastno vrednostjo $k$. ■

**Naloga 8.7** *Poišči Jacobijevo rotacijo $J$, ki transformira matriko $A = \begin{bmatrix} a_1 & b \\ b & a_2 \end{bmatrix}$ v diagonalno matriko. Matrika $J$ naj bo ortogonalna, veljati mora*

$$JAJ^T = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

*Rešitev.* Matrika $J$ naj bo $J = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$. Izračunajmo

$$JAJ^T = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_1 & b \\ b & a_2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} ca_1 + sb & cb + sa_2 \\ -sa_1 + cb & -sb + sa_2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} =$$

$$\begin{bmatrix} c^2 a_1 + scb + scb + s^2 a_2 & -csa_1 - s^2 b + c^2 b + csa_2 \\ -csa_1 - s^2 b + c^2 b + csa_2 & s^2 a_1 - scb - scb + sca_2 \end{bmatrix}.$$

Veljati mora

$$-csa_1 - s^2 b + c^2 b + csa_2 = 0,$$

ker je matrika ortogonalna velja še $c^2 + s^2 = 1$. Velja $cs(a_2 - a_1) = -b(c^2 - s^2)$. Označimo $D = a_2 - a_1$. Iz tega s kvadriranjem dobimo

$$c^2 s^2 D^2 = b^2 (c^2 - s^2)^2.$$

Upoštevamo še, da velja $s^2 = 1 - c^2$. Tako dobimo

$$c^2 (1 - c^2) D^2 = b^2 (2c^2 - 1)^2.$$

Uvedemo še novo neznanko $C = c^2$. Dobimo kvadratno enačbo

$$C^2(D^2 + 4b^2) + C(-4b^2 - D^2) + b^2 = 0.$$

Označimo še $E = (D^2 + 4b^2)$. Rešujemo torej $C^2E - CE + b^2 = 0$. Tako dobimo rešitev

$$C = \frac{E \pm \sqrt{E^2 - 4Eb^2}}{2E} = \frac{1 \pm \sqrt{1 - 4b^2/E}}{2}.$$

Tako dobimo $c = \sqrt{C}$ in $s = \sqrt{1 - C}$. ∎

**Naloga 8.8** *Dokaži, da Jacobijeva in Gauss-Seidlova iterativna metoda konvergirata za matriko* $A \in \mathbb{R}^{n\times n}$

$$A = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}.$$

*Rešitev.* Spmonimo se, da je iteracijska matrika za Jacobijevo metodo

$$R_J = -D^{-1}(L + U) = \begin{bmatrix} 0 & \frac{1}{2} & & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & & \\ & \frac{1}{2} & 0 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2} & 0 & \frac{1}{2} \\ & & & & \frac{1}{2} & 0 \end{bmatrix}.$$

Pokazati moramo, da velja $\rho(R_J) < 1$. Kar pomeni, da mora za vsako lastno vrednost veljati $|\lambda| < 1|$. Poglejmo si matriko

$$R_J - \lambda I = \begin{bmatrix} -\lambda & \frac{1}{2} & & & & \\ \frac{1}{2} & -\lambda & \frac{1}{2} & & & \\ & \frac{1}{2} & -\lambda & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2} & -\lambda & \frac{1}{2} \\ & & & & \frac{1}{2} & -\lambda \end{bmatrix}.$$

Če velja $|\lambda| > 1$, je matrika $R_J - \lambda I$ strogo diagonalno dominantna in tako obrnljiva. Pokazati moramo še, da $R_J$ nima lastne vrednosti 1 ali $-1$.

Izračunajmo determinatno

$$\det(R_J - I) = \begin{vmatrix} -1 & \frac{1}{2} & & & & \\ \frac{1}{2} & -1 & \frac{1}{2} & & & \\ & \frac{1}{2} & -1 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2} & -1 & \frac{1}{2} \\ & & & & \frac{1}{2} & -1 \end{vmatrix} = \frac{1}{(-2)^n}\det(A).$$

Uporabili bomo indukcijo in upoštevali, da lahko uporabimo izračun determinante kot v Sturmovem zaporedju. Ugotovimo, da velja $\det(A_n) = n + 1$. Uporabimo

$$p_i(\lambda) = (d_i - \lambda)p_{i-1}(\lambda) - u_{i-1}^2 p_{i-2}(\lambda), \text{ za } i = 2, \ldots, n,$$

torej velja

$$\det(A_i) = 2\det(A_{i-1}) - (-1)^2 \det(A_{i-2}) = 2i - (i - 1) = i + 1.$$

Kar pomeni, da je matrika $A$ simetrična in pozitivno definitna. Podobno dokažemo, da velja $\det(R_J+I_n) = \begin{vmatrix} 1 & \frac{1}{2} & & & & \\ \frac{1}{2} & 1 & \frac{1}{2} & & & \\ & \frac{1}{2} & 1 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2} & 1 & \frac{1}{2} \\ & & & & \frac{1}{2} & 1 \end{vmatrix} = \frac{1}{(-2)^n}\det(A-4I) = \frac{1}{(-2)^n}(-1)^n(n+1)$. Upoštevamo, da velja

$$\det(A_i - 4I_i) = (2 - 4)(-1)^{i-1}i - (-1)^2(-1)^{i-2}(i - 1) = (-1)^i(2i - (i - 1)) = (-1)^i(i + 1).$$

■

# Poglavje 9

# Polinomska interpolacija

Polinom $\mathcal{L}_{n,i}(x) = \Pi_{j=0,j\neq i}^{n} \frac{x-x_j}{x_i-x_j}$ je Lagrangeev bazni polinom. Naj bo $I_n(x)$ polinom, ki se s funkcijo $f$ ujema v točkah $x_0, \ldots, x_n$. Potem je to ravno polinom

$$I_n(x) = \sum_{i=0}^{n} f(x_i)\mathcal{L}_{n,i}(x).$$

Če definiramo $\omega(x) = \Pi_{i=0}^{n}(x - x_i)$, lahko zapišemo

$$\mathcal{L}_{n,i}(x) = \frac{\omega(x)}{(x - x_i)\omega'(x_i)}.$$

Za ostanek $n + 1$-krat odvedljive funkcije na $[a, b]$, velja

$$f(x) - I_n(x) = \frac{f^{(n+1)}(\zeta)}{(n+1)!}\omega(x),$$

za nek $\zeta \in [a, b]$. Veljati mora, da je $x_i \in [a, b]$.

**Naloga 9.1** *Dokaži, da velja* $\sum_{i=0}^{n} \mathcal{L}_{n,i}(x) = 1$.

*Rešitev.* Če za $f$ vzamemo kar konstanto 1, iz enačbe za ostanek dobimo željeno enakost. Odvod konstante je enak 0. ■

Lagrangeva oblika interpolacijskega polinoma ni najprimernejša za konstrukcijo in računanje vrednosti interpolacijskega polinoma. Porabimo namreč veliko število operacij, poleg tega mora biti stopnja polinoma vnaprej določena. Boljše so zaporedne linerne interpolacije in tudi deljene in končne diference.

**Nevilleova shema**

Naj bo $I_{i,\ldots,i+k}$ interpolacijski polinom, ki se ujema z $f$ v točkah $x_i, \ldots, x_{i+k}$. Potem velja

$$I_{i,\ldots,i+k}(x) = \frac{1}{x_{i+k} - x_i} \begin{vmatrix} x - x_i & I_{i,\ldots,i+k-1}(x) \\ x - x_{i+k} & I_{i+1,\ldots,i+k}(x) \end{vmatrix}.$$

Primer Nevillove sheme za 4 točke:

$$\begin{array}{c|c|c|ccc}
x_i & x - x_i & y_i & & & \\
\hline
x_0 & x - x_0 & y_0 & & & \\
 & & & I_{0,1} & & \\
x_1 & x - x_1 & y_1 & & I_{012} & \\
 & & & I_{1,2} & & I_{0123} \\
x_2 & x - x_2 & y_2 & & I_{123} & \\
 & & & I_{2,3} & & \\
x_3 & x - x_3 & y_3 & & &
\end{array}.$$

**Naloga 9.2** *Poišči vrednosti interpolacijskega polinoma v točki* $x = 1$ *za podatke* $x_0 = 3, x_1 = 2, x_2 = 5$ *in* $f(x_0) = 2, f(x_1) = 4, f(x_2) = 0$.

*Rešitev.*

| $x_i$ | $x - x_i$ | $y_i$ | | |
|---|---|---|---|---|
| $x_0 = 3$ | $x - x_0 = -2$ | $y_0 = 2$ | | |
| | | | $I_{0,1} = 6$ | |
| $x_1 = 2$ | $x - x_1 = -1$ | $y_1 = 4$ | | $I_{012} = \frac{20}{3}$ . |
| | | | $I_{1,2} = \frac{16}{3}$ | |
| $x_2 = 5$ | $x - x_2 = -4$ | $y_2 = 0$ | | |

Do vrednosti smo prišči z naslednjim postopkom

$$I_{01}(x) = \frac{1}{x_1 - x_0} \begin{vmatrix} x - x_0 & I_0(x) \\ x - x_1 & I_1(x) \end{vmatrix} = -1 \begin{vmatrix} -2 & 2 \\ -1 & 4 \end{vmatrix} = 6,$$

$$I_{12}(x) = \frac{1}{x_2 - x_1} \begin{vmatrix} x - x_1 & I_1(x) \\ x - x_2 & I_2(x) \end{vmatrix} = \frac{1}{3} \begin{vmatrix} -1 & 4 \\ -4 & 0 \end{vmatrix} = \frac{16}{3},$$

$$I_{012}(x) = \frac{1}{x_2 - x_0} \begin{vmatrix} x - x_0 & I_{01}(x) \\ x - x_2 & I_{12}(x) \end{vmatrix} = \frac{1}{2} \begin{vmatrix} -2 & 6 \\ -4 & \frac{16}{3} \end{vmatrix} = \frac{20}{3}.$$

■

**Deljene diference**

Deljena diferenca $[\mathbf{x_0}, \ldots, \mathbf{x_k}]\mathbf{f}$ je vodilni koeficient interpolacijskega polinoma stopnje $k$, ki se ujema z $f$ v paroma različnih točkah $x_0, x_1, \ldots, x_k$. Velja še

$$I_n(x) = [x_0]f + \sum_{i=1}^{n} (x - x_0) \cdots (x - x_{i-1})[x_0, x_1, \ldots, x_i].$$

Za k-krat zvezno odvedljivo funkcijo velja

$$[x_0, x_1, \ldots, x_k]f = \frac{1}{k!} f^{(k)}(\zeta).$$

Za $n + 1$-krat odvedljivo funkcijo $f$ velja

$$f(x) - I_n(x) = (x - x_0) \ldots (x - x_n)[x_0, x_1, \ldots, x_n, x]f.$$

Vrednosti deljenih diferenc najlažje izračunamo s pomočjo rekurzivne formule:

$$[x_i, x_{i+1}, \ldots, x_{i+k}]f = \begin{cases} \frac{f^{(k)}(x_i)}{k!} & \text{za } x_i = \ldots = x_{i+k} \\ \frac{[x_i, \ldots, x_{r-1}, x_{r+1}, \ldots, x_{i+k}]f - [x_i, \ldots, x_{s-1}, x_{s+1}, \ldots, x_{i+k}]f}{x_s - x_r} & \text{za } x_s \neq x_r. \end{cases}$$

**Naloga 9.3** *Naj bo* $f$ $C^1$*-funkcija in naj bodo* $x_0 < x_1 < \ldots < x_n$. *Poišči polinom* $p$ *stopnje* $2n+1$, *ki se bo z dano funkcijo* $f$ *ujemal dvakratno, to je v vrednostih in odvodih. Poišči polinom* $H_n(x)$, *tako da bo veljalo* $H_n(x_i) = f(x_i)$ *in* $H_n'(x_i) = f(x_i)$. *Poleg tega naj velja* $A_i(x_j) = \delta_{ij}$, $A_i'(x_j) = 0$, $B_i(x_j) = 0$ *in* $B_i'(x_j) = \delta_{ij}$. *Polinoma poizkusi poiskati sam. Če ti ne uspe, lahko dokažeš, da temu zadoščata kar nastavka*

$$A_i(x) = \left(1 - 2\mathcal{L}_{in}'(x_i)(x - x_i)\right) \mathcal{L}_{in}^2(x)$$

*in*

$$B_i(x) = (x - x_i)\mathcal{L}_{in}^2.$$

*Rešitev.* Očitno velja $B_i(x_j) = 0$, saj je $\mathcal{L}_{in}(x_j) = \delta_{ij}$. Izračunajmo

$$B_i'(x) = \mathcal{L}_{in}^2(x) + 2(x - x_j)\mathcal{L}_{in}(x)\mathcal{L}_{in}'(x).$$

Torej velja $B_i'(x_j) = \delta_{ij}^2 = \delta_{ij}$. Velja tudi

$$A_i(x_j) = (1 - 2\mathcal{L}_{in}'(x_i)(x_j - x_i))\mathcal{L}_{in}^2(x_j) = ((1 - 2\mathcal{L}_{in}'(x_j)\delta_{ij}(x_j - x_i))\delta_{ij}^2 = \delta_{ij}.$$

Izračunajmo še

$$A_i'(x) = -2\mathcal{L}_{in}'(x_i)\mathcal{L}_{in}^2(x) + (1 - 2\mathcal{L}_{in}'(x_i)(x - x_i))2\mathcal{L}_{in}(x)\mathcal{L}_{in}'(x).$$

Torej je $A_i'(x_j) = 0$ za $i \neq j$, saj je $\mathcal{L}_{in}(x_j) = 0$. Za $j = i$ dobimo

$$A'(x_i) = -2\mathcal{L}_{in}'(x_i) + 2\mathcal{L}_{in}'(x_i) = 0.$$

■

**Naloga 9.4** *Dana je funkcija* $f(x) = \frac{4}{1+x}$.

*(i). Preko deljenih diferenc poišči interpolacijski polinom stopnje* 5, *za katerega velja:*

$$p(0) = f(0), \quad p'(0) = f'(0), \quad p''(0) = f''(0), \quad p(1) = f(1), \quad p'(1) = f'(1)$$

*in* $p''(1) = f''(1)$. *Izračunaj njegovo vrednost v* $x = \frac{1}{2}$.

*(ii). Čimbolj oceni napako* $\max |f(x) - p(x)|$ *za* $x \in [0, 1]$.

*Rešitev.* (i)
Izračunajmo

$$f(x) = \frac{4}{1+x}, \quad f'(x) = \frac{-4}{(1+x)^2}, \quad f''(x) = \frac{8}{(1+x)^3}.$$

Tako dobimo

$$f(0) = 4,\ f'(0) = -4\ , f''(0) = 8,\ f(1) = 2,\ f'(1) = -1,\ f''(1) = 1.$$

Uporabimo Newtonovo obliko interpolacijskega polinima

$$p(x) = [x_0]f + \sum_{i=1}^{n}(x - x_0)\dots(x - x_{i-1})[x_0, x_1, \dots, x_i]f.$$

Upoštevamo, da velja $[x_0, x_0]f = \frac{f^{(1)}(x_0)}{1!}$, $[x_0, x_0, x_0]f = \frac{f^{(2)}(x_0)}{2!}$. Izračunajmo deljene diference:

$$\begin{array}{c|cccccc}
\hline
\underline{0} & 4 & & & & & \\
 & & -4 & & & & \\
0 & 4 & & \underline{4} & & & \\
 & & -4 & & \underline{-2} & & \\
0 & 4 & & \underline{2} & & 1 & \\
 & & -2 & & -1 & & -\frac{1}{2} \\
\underline{1} & 2 & & 1 & & \frac{1}{2} & \\
 & & -1 & & -\frac{1}{2} & & \\
1 & 2 & & \frac{1}{2} & & & \\
 & & -1 & & & & \\
1 & 2 & & & & &
\end{array}.$$

Tako dobimo interpolacijski polinom

$$p(x) = 4 + (-4x) + 4x^2 + (-2)x^3 + x^3(x-1) + (-\frac{1}{2}x^3(x-1)^2).$$

Vrednost polinoma izračunamo po Hornerju:

$$p(x) = 4 + x\left(-4 + x\left(4 + x\left(-2 + (x-1)\left(1 + (x-1) - \frac{1}{2}\right)\right)\right)\right).$$

Tako dobimo

$$p(\frac{1}{2}) = 4 + \frac{1}{2}(-4 + \frac{1}{2}(4 + \frac{1}{2}(-2 - \frac{1}{2}(-\frac{1}{2}))))) = \frac{171}{64}.$$

(ii)
Ocenimo še napako. Vemo, da velja

$$|f(x) - p(x)| = |x^3(x-1)^3[0,0,0,1,1,1,x]f| = |x^3(x-1)^3| \cdot |\frac{f^{(6)}(\zeta)}{6!}|.$$

Šesti odvod funkcije je enak

$$f^{(6)}(x) = \frac{4 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot (-1)^6}{(1+x)^7}.$$

Torej velja $f^{(6)}(\zeta) \leq 4 \cdot 6!$. Ocenimo še

$$|x^3(x-1)^3| = |x(x-1)|^3 \leq (\frac{1}{2}\frac{1}{2})^3 = \frac{1}{4^3}.$$

Dobili smo oceno

$$|f(x) - p(x)| \leq \frac{1}{4^3}\frac{4 \cdot 6!}{6!} = \frac{1}{16}.$$

■

**Naloga 9.5** *Izračunaj deljeno diferenco*

$$[x_0, x_1, \ldots, x_n]\left(\frac{x}{1+x}\right),$$

*kjer so točke različne, tj.* $x_j \neq x_i$ *za* $i \neq j$*.*

*Rešitev.* Označimo $f(x) = \frac{x}{1+x}$, Nalogo rešimo z indukcijo. Najprej poglejmo, kaj dobimo za bazo indukcije in še en primer.

$$[x_0, x_1]f = \frac{[x_1]\frac{x}{1+x} - [x_0]\frac{x}{1+x}}{x_1 - x_0} = \frac{\frac{x_1}{1+x_1} - \frac{x_0}{1+x_0}}{x_1 - x_0} = \frac{x_1 + x_0x_1 - x_0 - x_0x_1}{(x_1 - x_0)(1+x_1)(1+x_0)}$$

$$= \frac{1}{(1+x_0)(1+x_1)}.$$

$$[x_0, x_1, x_2]f = \frac{[x_0, x_1]\frac{x}{1+x} - [x_1, x_2]\frac{x}{1+x}}{x_1 - x_3} = \frac{\frac{1}{(1+x_0)(1+x_1)} - \frac{1}{(1+x_1)(1+x_2)}}{x_0 - x_2} =$$

$$\frac{1 + x_2 - 1 - x_0}{(1+x_0)(1+x_1)(1+x_2))(x_0 - x_0)} = \frac{-1}{(1+x_0)(1+x_1)(1+x_2)}.$$

Naša indukcijska predpostavka bo torej

$$[x_0, x_1, \ldots, x_{n-1}]\frac{x}{1+x} = \frac{(-1)^n}{(1+x_0)(1+x_1)\ldots(1+x_{n-1})}.$$

Tako dobimo

$$[x_0, x_1, \ldots, x_n]f = \frac{[x_0, x_1, \ldots, x_{n-1}]f - [x_1, \ldots, x_n]f}{x_0 - x_n} =$$

$$\frac{(-1)^n \frac{1}{(1+x_0)(1+x_1)\ldots(1+x_n)} - (-1)^n \frac{1}{(1+x_0)(1+x_1)\ldots(1+x_n)}}{x_0 - x_n} =$$

$$\frac{(-1)^n(1+x_n-1-x_0}{(1+x_0)(1+x_1)\ldots(1+x_n)(x_0-x_n)} = \frac{(-1)^{n+1}}{(1+x_0)(1+x_1)\ldots(1+x_n)}.$$

■

**Naloga 9.6** *Izračunaj napako interpolacije za* $f(x) = \sin(x)$ *na* $[0, \frac{\pi}{6}]$ *s točkami* $0, \frac{\pi}{12}, \frac{\pi}{6}$ *splošno in pri Hermitovi interpolaciji, kjer zahtevamo še ujemanje odvodov.*

*Rešitev.* Vemo, da velja

$$|f(x) - p_3(x)| = |(x-x_0)(x-x_1)(x-x_2)[x_0, x_1, x_2, x]f| =$$

$$|(x-x_0)(x-x_1)(x-x_2)\frac{f^{(3)}(\zeta)}{3!}|.$$

Pišimo še $h = \frac{\pi}{12}$. Tako velja $x_i = x_0 + i \cdot h$. Izračunati moramo $||x(x-h)(x-2h)||_\infty$ na $[0, \frac{\pi}{6}]$. Izračunajmo odvod

$$(x(x-h)(x-2h))' = x(x-h) + (x-h)(x-2h) + x(x-2h) =$$

$$3x^2 - 6hx + 2h^2 = 3(x-h)^2 - h^2.$$

Ničli sta torej $x_{1,2} = h(1 \pm \frac{\sqrt{3}}{3})$. Samo ena leži na našem intervalu. Makismum je torej

$$|(h + \frac{\sqrt{3}}{3}h)(\frac{\sqrt{3}}{3}h)(\frac{\sqrt{3}}{3} - h)| = |h^3\frac{\sqrt{3}}{3}(1 - \frac{1}{3})| = \frac{2}{9}\sqrt{3}h^3.$$

Torej je

$$|f(x) - p_3(x)| \leq \frac{2}{9}\sqrt{3}h^3\frac{f^{(3)}(\zeta)}{3!} = \frac{\sqrt{3}}{27}(\frac{\pi}{12})^3 \doteq 1.1510^{-3}.$$

V primeru Hermitove interpolacije dobimo

$$|f(x) - p_6(x)| = |(x-x_0)^2(x-x_1)^2(x-x_2)^2[x_0, x_0, x_1, x_1, x_2, x_2, x]f| =$$

$$|x^2(x-h)^2(x-2h)^2\frac{f^{(6)}(\zeta)}{3!}|.$$

Naredimo iste ocene kot prej in dobimo

$$|f(x) - p_6(x)| \leq (\frac{2}{9}\sqrt{3}h^3)^2\frac{f^{(6)}(\zeta)}{5!} = (\frac{\pi}{12})^8\frac{1}{9720} \doteq 3.3 \cdot 10^{-8}.$$

Upoštevamo še, da velja

$$||\sin^{(6)}||_{\infty,[0,\pi/6]} = \frac{1}{2}.$$

■

**Naloga 9.7** *Dane so točke* $x_0, x_1, \ldots, x_n$ *in ustrezne vrednosti* $y_i = f(x_i)$*. Kako bi izračunali* $f^{-1}(y)$ *v dani točki* $y$*? Funkcija* $f$ *mora biti monotona.*

1. možnost
Skozi točke $(x_i, y_i)$ napeljemo interpolacijski polinom $p(x)$. Približek $x$ za $f^{(-1)}(y)$ dobimo kot rešitev enačbe $p(x) = y$.
2. možnost
Zamenjamo vlogi $x$ in $y$. Skozi točke $(y_i, x_i)$ napeljemo interpolacijski polinom $p(y)$. Primera z enakimi $y$ nimamo zaradi monotonosti. Približek $x$ za $f^{-1}(y)$ dobimo kot vrednost $p(y)$ v točki $y$. Ta postopek ponavljamo iterativno. Če poznamo funkcijo $f$, potem za približek $x$ izračunamo $f(x)$ in točko $(x, f(x))$ zamenjamo z najbolj oddaljeno točko $(x_i, y_i)$ gledano po $y$. Red konvergence je vedno med 1 in 2. Ko gre $n \to \infty$, gre red proti 2.

**Naloga 9.8** *Naredi dva koraka inverzne interpolacije za izračun* $\arcsin(\frac{1}{3})$*, kjer so dane točke* $x_0 = 0, x_1 = \frac{\pi}{6}, x_2 = \frac{\pi}{2}$.

*Rešitev.* 1. korak
Za ta primer je $f(x) = \sin(x)$. Izračunamo še $\sin(0) = 0, \sin(\pi/6) = 1/2, \sin(\pi/2) = 1$. Torej imamo $x_0 = 0, y_0 = 0, x_1 = \pi/6, y_1 = 1/2, x_2 = \pi/2, y_2 = 1$. V prvem koraku zamenjamo vlogi $x$ in $y$ in izračunamo interpolacijski polinom v $y$ in izračunamo $p(1/3)$.
2. korak
Pogledamo vrednost $f(x_3)$ in odvržemo najbolj oddaljeno točko po koordinati $y$. ■

## 9.1 Numerično odvajanje

Naj bo $f \in C^1(I)$. Za približek odvoda funkcije vzamemo kar odvod interpolacijskega polinoma. Velja namreč
$$f(x) = I_n(x) + (x - x_0)\ldots(x - x_n)[x_0, x_1, \ldots, x_n, x]f.$$
Z odvajanjem enakosti dobimo
$$f'(x) = I_n'(x) + \underbrace{\omega'(x)[x_0, x_1, \ldots, x_n, x]f + \omega(x)([x_0, x_1, \ldots, x_n, x]f))'}_{\text{napaka}}.$$
Za približek vzamemo kar odvod interpolacijskega polinoma $I_n'(x)$.

**Naloga 9.9** *Preko deljenih diferenc izrazi formulo za odvod* $f'(x_0)$ *s točkami* $f(x_0),\ f(x_1),\ f(x_2)$*. Točke so ekvidistantne.*

*Rešitev.* Velja
$$f(x) = \overbrace{f(x_0) + (x - x_0)[x_0, x_1]f + (x - x_0)(x - x_1)[x_0, x_1, x_2]f}^{p_2(x)} + \underbrace{(x - x_0)(x - x_1)(x - x_2)}_{\omega(x)}[x_0, x_1, x_2, x]f.$$
Izračunajmo odvod v $x = x_0$.
$$f'(x_0) = ([x_0, x_1]f + (2x - x_0 - x_1)[x_0, x_1, x_2]f)|_{x=x_0} + \omega'(x)[x_0, x_1, x_2, x]f|_{x=x_0}.$$
$$f'(x_0) = [x_0, x_1]f + (-h)[x_0, x_1, x_2]f + (-h)(-2h)\frac{f^{(3)}(\zeta}{6}$$
$$f'(x_0) = \frac{f(x_1) - f(x_0)}{h} - h\frac{\frac{f(x_2)-f(x_1)}{h} - \frac{f(x_1)-f(x_0)}{h}}{x_2 - x_0} + 2h^2\frac{f^{(3)}(\zeta)}{6}.$$

$$f'(x_0) = \frac{2f(x_1) - 2f(x_0) - f(x_2) + f(x_1) - f(x_0)}{2h} + h^2 \frac{f^{(3)}(\zeta)}{3}$$
$$f'(x_0) = \frac{-3f(x_0) + 4f(x_1) - f(x_2)}{2h} + \frac{h^3}{3} f^{(3)}(\zeta).$$

■

**Naloga 9.10** *Izpelji formulo za numerično odvajanje*

$$f'(x_2) = Af(x_0) + Bf(x_1) + Cf(x_2) + Df(x_3) + Ef(x_4) + Ff^{(m)}(\zeta).$$

*Točke $x_i$ so ekvidistantne.*

*Rešitev.* Uporabimo metodo nedoločenih koeficientov. Koeficiente $A, B, C, D, E$ določimo tako, da bo formula točna za polinome čim višjih stopenj.

Izberemo primerno bazo $1, x - x_2, (x - x_2)^2, (x - x_2)^3, \ldots$. Vstavljamo dokler enačbe niso linearno odvisne.

$$1 \quad : \quad 0 = A + B + C + D + E \tag{9.1}$$
$$x - x_2 \quad : \quad 1 = (-2h)A + (-h)B + hD + 2hE \tag{9.2}$$
$$(x - x_2)^2 \quad : \quad 0 = 4h^2A + h^2B + h^2D + 4h^2E \tag{9.3}$$
$$(x - x_2)^3 \quad : \quad 0 = -8h^3A - h^3B + h^3D + 8h^3E \tag{9.4}$$
$$(x - x_2)^4 \quad : \quad 0 = 16h^4A + h^4B + h^4D + 16h^4E \tag{9.5}$$

Dobimo sistem enačb

$$0 = A + B + C + D + E \tag{9.6}$$
$$\frac{1}{h} = -2A - B + D + 2E \tag{9.7}$$
$$0 = 4A + B + D + 4E \tag{9.8}$$
$$0 = -8A - B + D + 8E \tag{9.9}$$
$$0 = 16A + B + D + 16E. \tag{9.10}$$

Iz tretje in pete enačbe dobimo $0 = 12A + 12E$, kar je $A = -E$. It druge in četrte enačbe dobimo $\frac{1}{h} = 6A - 6E = 12A$, kar nam da $A = \frac{1}{12h}$. Iz prve in tretje enačbe dobimo $C = 0$. Potem iz prve enačbe sledi $B = -D$. Iz četrte enačbe pa potem $D = -8E = \frac{2}{3h}$. Iz tega dobimo

$$f'(x_2) = \frac{1}{12h}(f(x_0) - 8f(x_1) + 8f(x_3) - f(x_1)) + Ff^{(m)}(\zeta).$$

Za ostanek dobimo

$$(x - x_2)^5 \quad : \quad 0 = \frac{1}{12h}(-32h^5 + 8h^5 + -32h^5) + F5!, \tag{9.11}$$

saj nastavek ni več točen za polinome stopnje 5. Tako dobimo $4h^4 = 5!F$, kar je $F = \frac{h^4}{30}$. ■

## 9.2 Numerično integriranje

**Naloga 9.11** *Naj bodo točke ekvidistantne. Z metodo nedoločenih koeficientov izpelji Simpsonovo pravilo*

$$\int_{x_0}^{x_2} f(x)dx = \frac{h}{3}\left(f(x_0) + 4f(x_1) + f(x_2)\right) + Rf.$$

*Zapiši tudi sestavljeno pravilo in določi ostanek.*

*Rešitev.* Uporabimo metodo nedoločenih koeficentov. Določimo $A, B, C$ tako, da bo formula

$$\int_{x_0}^{x_2} f(x)dx = Af(x_0) + Bf(x_1) + Cf(x_2) + Df^{(m)}(\zeta)$$

točna za polinome čim višjih stopenj. Za bazo polinomov vzamemo $1, (x - x_0), (x - x_0)^2, \dots$. Ko te polinome po vrsti vstavimo v nastavek, dobimo enačbe:

$$\begin{aligned} 2h &= A + B + C \\ \frac{(2h)^2}{2} &= Bh + 2Ch \\ \frac{(2h)^3}{3} &= Bh^2 + 4Ch^2 \\ &. \end{aligned}$$

Tako dobimo sistem enačb:

$$2h = A + B + C \tag{9.12}$$
$$2h = B + 2C \tag{9.13}$$
$$\frac{8h}{3} = B + 4C \tag{9.14}$$

Upoštevali smo

$$\int_{x_0}^{x_2} (x - x_0)^n = \frac{(x_2 - x_0)^{n+1}}{n} = \frac{(2h)^{n+1}}{n}.$$

Če odštejemo drugo od tretje dobimo $C = \frac{h}{3}$. Odštejemo še tretjo do prve in dobimo $A = C = \frac{h}{3}$. Nazadnje dobimo še $B = \frac{4h}{3}$. Tako smo dobili

$$\int_{x_0}^{x_2} = \frac{h}{3}\left(f(x_0) + 4f(x_1) + f(x_2)\right) + Df^{(m)}(\zeta).$$

Določiti moramo še ostanek. Vstavimo še naslednje bazne polinome in poglejmo, če je formula zanje točna. Za $(x - x_0)^3$ dobimo, da je formula točna, za $(x - x_0)^4$ pa ne več. Torej je $m = 4$.

$$\begin{aligned} \frac{(2h)^4}{4} &= \frac{h}{3}(4h^3 + 8h^3) = 4h^4 \\ \frac{(2h)^5}{5} &= \frac{h}{3}(4h^4 + 16h^4) + 24D \end{aligned}$$

Tako dobimo $24D = \frac{32h^5}{5} - \frac{20h^5}{3}$, torej je $D = -\frac{1}{90}h^5$. Ostanek je $Rf = -\frac{1}{90}h^5 f^{(4)}(\zeta)$.

Izpeljimo še sestavljeno pravilo. Število delilnih točk mora biti $2n + 1$. Tako je $a = x_0$, $b = x_{2n}$ in $h = \frac{b-a}{2n}$. Tako dobimo

$$\int_{x_0}^{x_{2n}} f(x)dx = \sum_{i=0}^{n-1} \int_{x_{2i}}^{x_{2i+2}} f(x)dx = \sum_{i=0}^{n-1} \frac{h}{3}\left(f(x_{2_i}) + 4f(x_{2i+1}) + f(x_{2i+2})\right).$$

Ostanek je enak

$$Rf = \sum_{i=0}^{n-1} -\frac{1}{90}h^5 f^{(4)}(\zeta_i) = -\frac{1}{90}h^5 n f^{(4)}(\eta) = -\frac{1}{90}h^4 \frac{b-a}{2} f^{(4)}(\eta).$$

∎

**Naloga 9.12** *Integral* $\int_0^1 e^{-x^2}dx$ *računamo s sestavljenim trapeznim pravilom. Kolikšen naj bo* $h$, *da bo napaka metode manjša od* $10^{(-6)}$.

*Rešitev.* Trapezno pravilo je $\int_{x_0}^{x_1} f(x)dx = \frac{h}{2}(f(x_0)+f(x_1)) - \frac{h^3}{12}f^{(2)}(\zeta)$. Za $h = \frac{b-a}{n}$ in $x_i = a+ih$ dobimo

$$\int_a^b f(x)dx = \sum_{i=0}^{n-1}\int_{x_i}^{x_{i+1}} f(x)dx = \sum_{i=0}^{n-1}\left(\frac{h}{2}\left(f(x_i)+f(x_{i+1})\right)+R_if\right),$$

kjer je

$$Rf = \sum_{i=0}^{n-1} R_if = \sum_{i=0}^{n-1} -\frac{h^3}{12}f^2(\zeta_i) = -\frac{h^3}{12}nf^2(\eta) = -\frac{h^2}{12}(b-a)f^2(\eta).$$

Naše sestavljeno pravilo je

$$\int_a^b f(x)dx = \frac{h}{2}\left(f(x_0)+f(x_n))+2\sum_{i=1}^{n-1}f(x_i)\right) - \frac{h^2}{12}(b-a)f^{(2)}(\eta).$$

Drugi odvod funkcije je enak

$$f''(x) = (-2xe^{-x^2})' = 4x^2e^{-x^2} - 2e^{-x^2} = 2e^{-x^2}(2x^2-1).$$

Maksimum $e^{-x^2}$ je dosežen v $x=0$. Funkcija $2x^2-1$ na intervalu $[0,1]$ zavzame vrednosti $[-1,1]$. Torej je maksimum $|f''(x)|$ dosežen v $x=0$ in je enak 2. Torej mora veljati

$$|Rf| \leq \frac{h^2}{12}(1-0)2 < 10^{-6}.$$

Tako dobimo $h^2 < 6\cdot 10^{-6}$ in $h < \sqrt{6}\cdot 10^{-3}$. Velja še $nh=1$. Torej mora veljati $n > 408.24$, kar je $n \geq 409$. ■

Naj bo funkcija neskončnokrat zvezno odvedljiva. Oznaka

$$T_h(f) = \frac{h}{2}(f(x_0)+2f(x_1)+2f(x_2)+\ldots+2f(x_{n-1})+f(x_n))$$

predstavlja sestavljeno trapezno pravilo za ekvidistantne točke na intervalu $[a,b]$. Velja

$$I(f) = \int_{x_0}^{x_n} f(x)dx = T_h(f) + a_{1,0}h^2 + a_{2,0}h^4 + a_{3,0}h^6 + \ldots = T_h(f) + \sum_{j=1}^{m} a_{j,0}h^{2j} + Rf.$$

Zapišimo enakosti še za $\frac{h}{2}$ in $\frac{h}{4}$.

$$\begin{aligned} I(f) &= T_h(f) + a_{1,0}h^2 + a_{2,0}h^4 + a_{3,0}h^6 + \ldots \\ I(f) &= T_{\frac{h}{2}}(f) + a_{1,0}(\frac{h}{2})^2 + a_{2,0}(\frac{h}{2})^4 + a_{3,0}(\frac{h}{2})^6 + \ldots \end{aligned}$$

Če drugo enačbo pomnožimo s 4 in odštejemo prvo, dobimo

$$4I(f) - I(f) = 4T_{\frac{h}{2}} - T_h(f) + \tilde{a}_{2,1}h^4 + \tilde{a}_{3,1}h^6 + \ldots,$$

kar nam da

$$I(f) = \frac{4T_{\frac{h}{2}}(f) - T_h(f)}{3} + a_{2,1}h^4 + a_{3,1}h^6 + \ldots.$$

Tako smo dobili nov približek za integral

$$T^{(1)}_{\frac{h}{2}}(f) = \frac{4T_{\frac{h}{2}}(f) - T_h(f)}{3}.$$

Tako dobimo shemo oblike

$$T^{(j)}_{\frac{h}{2^k}}(f) = \frac{4^j T^{(j-1)}_{\frac{h}{2^k}} - T^{(j-1)}_{\frac{h}{2^{k-1}}}}{4^j - 1}.$$

Kjer računamo $T_{\frac{h}{2}}$ na sledeči način

$$T_{\frac{h}{2}} = \frac{1}{2}T_h(f) + \frac{h}{2}(f(\frac{x_0 + x_1}{2}) + \ldots + f(\frac{x_{n-1} + x_n}{2})).$$

**Naloga 9.13** *Z Rombergovo ekstrapolacijo (sestavljene) trapezne metode izračunaj integral* $\int_0^{\frac{\pi}{2}} \sin(x)dx$. *Za začetni* $h$ *vzemi* $h = \frac{\pi}{4}$. *Naredi dva koraka Rombergove ekstrapolacije.*

| $h'$ | | | |
|---|---|---|---|
| $\frac{\pi}{4}$ | $T_h(f)$ | | |
| $\frac{\pi}{8}$ | $T_{\frac{h}{2}}(f)$ | $T^{(1)}_{\frac{h}{2}}(f)$ | |
| $\frac{\pi}{16}$ | $T_{\frac{h}{4}}(f)$ | $T^{(1)}_{\frac{h}{4}}(f)$ | $T^{(2)}_{\frac{h}{4}}(f)$ |

*Rešitev.* Nalogo rešimo z direktno uporabo sestavljene trapezne metode in formule za Rombergovo ekstrapolacijo.

$$T_h(f) = \frac{\pi}{8}\left(\sin(0) + \sin(\frac{\pi}{2})\right) + \frac{\pi}{4}\sin(\frac{\pi}{4}) \doteq 0.948059$$

$$T_{h/2}(f) = \frac{1}{2}T_h(f) + \frac{\pi}{16}\left(\sin(\frac{\pi}{8}) + \sin(\frac{3\pi}{8})\right) \doteq 0.987116$$

$$T_{h/4}(f) = \frac{1}{2}T_{h/2}(f) + \frac{\pi}{16}\left(\sin(\frac{\pi}{16}) + \sin(\frac{3\pi}{16}) + \sin(\frac{5\pi}{16}) + \sin(\frac{3\pi}{8})\right) \doteq 0.996785$$

$$T^{(1)}_{\frac{h}{2}}(f) = \frac{4T_{h/2} - T_h(f)}{4-1} \doteq 1.000135$$

$$T^{(1)}_{\frac{h}{4}}(f) = \frac{4T_{h/4} - T_{h/2}(f)}{4-1} \doteq 1.000008$$

$$T^{(2)}_{\frac{h}{4}}(f) = \frac{16T^{(1)}_{\frac{h}{4}}(f) - T^{(1)}_{\frac{h}{2}}(f)}{16-1} \doteq 0.9999995$$

∎

**Naloga 9.14** *Izpelji Gaussovo integracijsko pravilo na dveh točkah.*

$$\int_{-1}^{1} f(x)dx = Af(x_0) + Bf(x_1) + Rf.$$

V splošnem je Gaussovo integracijsko pravilo

$$\int_a^b = \sum_{i=0}^{n} A_i f(x_i) + Rf$$

na vozlih $x_0$, $x_1$, ..., $x_n$ točno za polinome stopnje $2n+1$. Pri ostanku je torej $m \geq 2n+2$. Vozle določimo tako, da je $\omega(x)$ pravokoten na polinome stopnje manjše ali enake $n$. Veljati mora $\int \omega(x)\rho(x)x^i = 0$, za $i = 1, \ldots, n$. Tukaj je $\omega(x) = (x - x_0)\ldots(x - x_n)$, $\rho(x)$ pa pozitivna integrabilna funkcija. Koeficiente potem določimo z metodo nedoločenih koeficientov.

*Rešitev.* Določiti moramo vozla $x_0$ in $x_1$, tako da bo veljalo $\int_{-1}^{1}\omega(x)dx = 0$ in $\int_{-1}^{1}\omega(x)xdx = 0$. Tako bo formula točna za polinom stopnje $\leq 3$. Najprej določimo uteži, potem pa recimo z metodo nedoločenih koeficientov določimo še $A$ in $B$. Izračunajmo integrale.

$$\int_{-1}^{1}\omega(x) = \int_{-1}^{1}(x-x_0)(x-x_1)dx = \int_{-1}^{1}x^2-(x_0+x_1)x+x_0x_1dx$$

$$\frac{x^3}{3}-(x_0+x_1)\frac{x^2}{2}+x_0x_1x\Big|_{-1}^{1} = \frac{2}{3}+2x_0x_1$$

$$\int_{-1}^{1}x\omega(x)f(x)dx = \int_{-1}^{1}x^3-(x_0+x_1)x^2+x_0x_1xdx =$$

$$\frac{x^4}{4}-(x_0+x_1)\frac{x^3}{3}+x_0x_1\frac{x^2}{2}\Big|_{-1}^{1} = -2(x_0+x_1)$$

Dobimo enačbi

$$2+\frac{2}{3}x_0x_1 = 0 \quad \text{in} -2(x_0+x_1) = 0.$$

Tako dobimo $x_0x_1 = -\frac{1}{3}$ in $x_1 = -x_0$. To nam da $x_0 = -\frac{1}{\sqrt{3}}$ in $x_1 = \frac{1}{\sqrt{3}}$, če naj velja $x_1 \geq x_0$. Z metodo nedoločenih koeficentov določimo še $A$ in $B$ in ostanek.

$$\begin{aligned}\int_{-1}^{1}1dx &= 2 = A+B\\ \int_{-1}^{1}xdx &= 0 = -A\frac{1}{\sqrt{3}}+B\frac{1}{\sqrt{3}}\end{aligned}$$

Iz enačb $A+B=2$ in $A-B=0$ dobimo $A=B=1$. Določimo še ostanek. Vemo, da je formula točna za vsaj polinome stopnje 3. Torej je $m=4$ in $(x^4)^{(4)} = 4! = 24$. Izračunajmo

$$\int_{-1}^{1}x^4dx = \frac{2}{5} = \frac{1}{9}+\frac{1}{9}+C24,$$

torej je $C = \frac{1}{135}$. Naša formula je

$$\int_{-1}^{1}f(x)dx = f(-\frac{1}{\sqrt{3}})+f(\frac{1}{\sqrt{3}})+\frac{1}{135}f^{(4)}(\zeta).$$

■

**Naloga 9.15** *Izračunaj izlimitirani integral* $\int_0^{2h}\frac{f(x)}{\sqrt{x}}dx$ *na dva načina:*

*(i). S substitucijo odstrani singularnost in uporabi Simpsonovo pravilo.*

*(ii). Izpelji formulo*

$$\int_0^{2h}\frac{f(x)}{\sqrt{x}}dx = Af(0)+Bf(h)+Cf(2h)+Df^{(m)}(\zeta)$$

*z metodo nedoločenih koeficientov, tako da bo formula točna za polinome čim večjih stopenj.*

*Preizkusi oba načina na* $\int_0^{0.2}\frac{e^{-x}}{\sqrt{x}}dx$.

*Rešitev.*

(i). Uvedemo substitucijo $t = \sqrt{x}$, $dx = 2tdt$ in na integralu uporabimo Simpsonovo pravilo.

$$\int_0^{2h} \frac{f(x)}{\sqrt{x}} dx = \int_0^{\sqrt{2h}} \frac{f(t^2)}{t} 2tdt = 2\int_0^{\sqrt{2h}} f(t^2)dt$$

Simpsonovo pravilo je

$$\int_{x_0}^{x_2} g(x)dx = \frac{h}{3}\left(g(x_0) + 4g(x_1) + g(x_2)\right) - \frac{1}{90}h^5 g^{(4)}(\zeta).$$

To pravilo uporabimo za $g(t) = f(t^2)$, $x_0 = 0$, $x_1 = \sqrt{2h}/2$, $x_2 = \sqrt{2h}$ in zamenjamo $h$ z $\sqrt{2h}/2$. Naš integral je enak

$$\int_0^{2h} \frac{f(x)}{\sqrt{x}} dx = \frac{\sqrt{2h}}{6}\Big(f(0) + 4f(\frac{h}{2}) + f(2h)\Big) - \frac{1}{90}(\sqrt{2h}/2)^5 g^{(4)}(\zeta).$$

(ii). Za bazo vzamemo 1, $x$, $x^2$, … Izračunamo integrale

$$\begin{aligned}
\int_0^{2h} x^{\frac{-1}{2}} &= 2x^{\frac{1}{2}}\Big|_0^{2h} = 2\sqrt{2h} \\
\int_0^{2h} x^{\frac{1}{2}} &= \frac{2}{3}x^{\frac{3}{2}}\Big|_0^{2h} = \frac{2}{3}\sqrt{2h}2h \\
\int_0^{2h} x^{\frac{3}{2}} &= \frac{2}{5}x^{\frac{5}{2}}\Big|_0^{2h} = \frac{2}{5}\sqrt{2h}(2h)^2
\end{aligned}$$

Tako dobimo enačbe

$$\begin{aligned}
2\sqrt{2h} &= A + B + C \\
\frac{2}{3}\sqrt{2h}2h &= Bh + 2Ch \\
\frac{2}{5}\sqrt{2h}4h^2 &= Bh^2 + 4Ch^2.
\end{aligned}$$

Ko rešimo sistem enačb dobimo $A = \frac{12}{15}\sqrt{2h}$, $B = \frac{16}{15}\sqrt{2h}$, $C = \frac{2}{15}\sqrt{2h}$. Določimo še napako, $f(x) = x^3$,

$$\int_0^{2h} x^{\frac{7}{2}} = \frac{2}{7}8h^3\sqrt{2h} = \frac{2}{15}\sqrt{2h}(6f(0) + 8f(h) + f(2h)) + D6.$$

Kar je

$$\frac{2}{7}8h^3\sqrt{2h} = \frac{2}{15}\sqrt{2h}16h^3 + 6D.$$

Tako dobimo $D = \frac{8}{315}h^3\sqrt{2h}$.

Preizkus obeh metod na konkretnem primeru je prepuščen bralcu.

■

**Naloga 9.16** *Izpelji Gaussovo kvadraturno pravilo*

$$\int_{-\infty}^{\infty} e^{-|x|} f(x)dx = A_0 f(x_0) + A_1 f(x_1) + A_2 f(x_2) + R(f).$$

*Upoštevaj* $\int_0^\infty e^{-x}x^n dx = n!$.

**Naloga 9.17** ***Naloga ni bila narejena na vajah***
*Diferencialno enačbo* $y' = -50y + 100$, $y(0) = y_0$, *ki ima točno rešitev* $y(x) = (y_0 - 2)e^{-50x} + 2$ *rešujemo z*

- *eksplicitno Eulerjevo metodo* $y_{n+1} = y_n + hf(x_n, y_n)$,
- *implicitno Eulerjevo metodo* $y_{n+1} = y_n + hf(x_{n+1}, y_{n+1})$.

*V obeh primerih izpelji splošno formulo za* $y_n$ *in določi vrednosti za* $h$, *pri katerih* $y_n$ *konvergira proti točni rešitvi, ko gre* $x$ *proti* $\infty$.

*Rešitev.*

- Imamo enačbo $y' = f(x, y) = -50y + 100$, $y(0) = y_0$. Tako dobimo zvezo

  $$y_{n+1} = y_n + hf(x_n, y_n) = y_n + h(-50y_n + 100) = y_n(1 - 50h) + 100h.$$

  Kar je diferenčna enačba z začetnim pogojem $y_0$. Najprej rešimo homogeni del $y_{n+1} = y_n(1 - 50h)$. Splošni nastavek je $y_n = A\alpha^n$. Tako dobimo $\alpha = (1 - 50h)$. Rabimo še partikularno rešitev, ki jo kar uganemo. Zlahka se prepričamo, da je to $y_n = 2$. Tako dobimo skupno rešitev $y_n = A(1 - 50h)^n + 2$. Iz začetnega pogoja dobimo $A = y_0 - 2$. Rešitev je $y_n = (y_0 - 2)(1 - 50h)^n + 2$. Točna rešitev $y(x) = (y_0 - 2)e^{-50x} + 2$ ima limito 2, ko gre $x \to \infty$. To mora veljati tudi za numerično rešitev, zato mora biti $|1 - 50h| < 1$. Ker je $h > 0$, mora veljati $0 < h < \frac{1}{50}$.

- Imamo enačbo $y' = f(x, y) = -50y + 100$, $y(0) = y_0$. Tako dobimo zvezo

  $$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1}) = y_n + h(-50y_{n+1} + 100).$$

  Kar je diferenčna enačba z začetnim pogojem $y_0$,

  $$y_{n+1} = \frac{y_n + 100h}{1 + 50h} = \frac{y_n}{1 + 50h} + \frac{100h}{1 + 50h}.$$

  Najprej rešimo homogeni del. Splošni nastavek je $y_n = A\alpha^n$. Tako dobimo $\alpha = \frac{1}{1+50h}$. Rabimo še partikularno rešitev, ki jo kar uganemo. Zlahka se prepričamo, da je to $y_n = 2$. Tako dobimo skupno rešitev $y_n = A\frac{1}{(1+50h)^n} + 2$. Iz začetnega pogoja dobimo $A = y_0 - 2$. Rešitev je $y_n = (y_0 - 2)\frac{1}{(1+50h)^n} + 2$. Točna rešitev $y(x) = (y_0 - 2)e^{-50x} + 2$ ima limito 2, ko gre $x \to \infty$. To mora veljati tudi za numerično rešitev, zato mora biti $|\frac{1}{1+50h}| < 1$. Ker je $h > 0$, je neenakost zmeraj izpolnjena.

Enačba $y' = \lambda y$, kjer je $Re(\lambda) < 0$ spada med toge probleme. Zanje je značilno, da rešitev hitro pada, ko gre $x \to \infty$. ■

**Naloga 9.18** *Dana je diferencialna enačba* $y' = x^2 - y^2$, $y(x_0) = y_0$.

*(i). Izpelji trapezno formulo* $y_{n+1} = y_n + \frac{h}{2}\left(f(x_n, y_n) + f(x_{n+1}, y_{n+1})\right)$.

*(ii). S trapezno formulo izračunaj približek* $y_1$.

*Rešitev.*

(i). Rešujemo diferencialno enačbo $y' = f(x,y)$. To enačbo integriramo in za izračun integrala uporabimo trapezno pravilo.

$$y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} y'(x)dx = \int_{x_n}^{x_{n+1}} f(x,y(x))dx = \frac{h}{2}\left(f(x_n,y_n) + f(x_{n+1},y_{n+1})\right).$$

Tako dobimo

$$y_{n+1} = y_n + \frac{h}{2}\left(f(x_n,y_n) + f(x_{n+1},y_{n+1})\right).$$

(ii). V zvezo vstavimo $f(x,y) = x^2 - y^2$. Tako dobimo

$$y_1 = y_0 + \frac{h}{2}\left(x_0^2 - y_0^2 + x_1^2 - y_1^2\right).$$

Iz česar dobimo kvadratno enačba za $y_1$

$$\frac{h}{2}y_1^2 + y_1 - \left(y_0 + \frac{h}{2}(x_0^2 - y_0^2 + x_1^2)\right).$$

Kvadratno enačbo rešimo in dobimo

$$y_1 = \frac{-1 \pm \sqrt{1 + 2hy_0 + h^2(x_0^2 - y_0^2 + x_1^2)}}{h}.$$

Izbrati moramo še pravo rešitev. To bo tista, ki ima limito, ko gre $h \to 0$ enako $y_0$. Ker je $\sqrt{1 + 2hy_0 + h^2(x_0^2 - y_0^2 + x_1^2)} = 1 + hy_0 + O(h^2)$, je to rešitev s plusom.

■

**Naloga 9.19** *Podana je Runge-Kutta metoda*

$$\begin{aligned}
k_1 &= hf(x_n, y_n) \\
k_2 &= hf(x_n + h/2, y_n + k_1/2) \\
k_3 &= hf(x_n + h/2, y_n - k_1 + 2k_2) \\
y_{n+1} &= y_n + \frac{1}{6}(k_1 + 4k_2 + k_3)
\end{aligned}$$

*Rešujemo problem $y' = \lambda y$, $y(0) = y_0$. Zapiši eksplicitno formulo za $y_n$. Kako velik je lahko $h$, da se bo pri $\lambda = -4$ rešitev $y_n$ obnašala kot prava rešitev, ko gre $n \to \infty$.*

*Rešitev.* V našem primeru je $f(x,y) = \lambda y$. Izračunamo

$$\begin{aligned}
k_1 &= h\lambda y_n \\
k_2 &= h\lambda(y_n + k_1/2) = h\lambda(y_n + 1/2h\lambda y_n) = y_n(h\lambda + 1/2h^2\lambda^2) \\
k_3 &= h\lambda(y_n - k_1 + 2k_2) = y_n(h\lambda + h^2\lambda^2 + h^3\lambda^3) \\
y_{n+1} &= y_n + \frac{1}{6}(k_1 + 4k_2 + k_3) = y_n \underbrace{(1 + h\lambda + 1/2h^2\lambda^2 + 1/6h^3\lambda^3)}_{\phi(h\lambda)}
\end{aligned}$$

Točna rešitev je enaka $y(x) = y_0 e^{\lambda x}$. Da se bo numerična rešitev obnašala kot točna rešitev mora veljati $|\phi(h\lambda)| < 1$. Za primer $\lambda = -4$ dobimo

$$\phi(h \cdot -4) = 1 - 4h + 1/216h^2 + 1/6h^3(-64) = 1 - 4h + 8h^2 - \frac{32}{3}h^3.$$

Zanima nas torej, kdaj je

$$|1 - 4h + 8h^2 - \frac{32}{3}h^3| \leq 1.$$

Z nekaj truda ugotovimo, da mora veljati $0 < h < 0.6281$.

Metoda je A-stabilna, če je za vsak $\lambda$, kjer $Re(\lambda) < 0$, in vsak $h > 0$ stabilna pri reševanju enačbe $y' = \lambda y$. Kar pomeni, da pada proti 0, ko gre $n \to \infty$. ∎

**Naloga 9.20** ***Naloga ni bila narejena na vajah***

*Dane je diferencialna enačba drugega reda*

$$y'' - y'y^2 + y = 0,$$

*z začetnima pogojema $y(0) = 1$ in $y'(0) = 0$. Preuredi diferencialno enačbo na sistem diferencialnih enačb prvega reda, uporabi Runge-Kutta metodo 4. reda ter določi numerično rešitev v točki $y_1$ za $h = 0.2$.*

$$\begin{aligned}
k_1 &= hf(x_n, y_n) \\
k_2 &= hf(x_n + h/2, y_n + k_1/2) \\
k_3 &= hf(x_n + h/2, y_n + k_2/2) \\
k_3 &= hf(x_n + h, y_n + k_3) \\
y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)
\end{aligned}$$

*Rešitev.* Uvedemo novo spremenljivko $y' = z$. Kar nam da $y'' = z' = y'y^2 - y = zy^2 - y$. Tako dobimo sistem

$$\begin{bmatrix} y \\ z \end{bmatrix}' = \begin{bmatrix} z \\ zy^2 - y \end{bmatrix}.$$

$$Y(x) = \begin{bmatrix} y(x) \\ z(x) \end{bmatrix}, \quad Y' = F(x, Y) = \begin{bmatrix} z \\ zy^2 - y \end{bmatrix}$$

Naj bo $\mathbf{k} = \begin{bmatrix} k \\ l \end{bmatrix}$.

$$\begin{bmatrix} k_1 \\ l_1 \end{bmatrix} = hF(x_0, \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}) = h \begin{bmatrix} z_0 \\ z_0 y_0^2 - y_0 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.2 \end{bmatrix}$$

$$\begin{bmatrix} k_2 \\ l_2 \end{bmatrix} = hF(x_0 + h/2, \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}) + 1/2 \begin{bmatrix} k_1 \\ l_1 \end{bmatrix}) =$$

$$h \begin{bmatrix} z_0 + 1/2 l_1 \\ (z_0 + 1/2 l_1)(y_0 + 1/2 k_1)^2 - (y_0 + 1/2 k_1) \end{bmatrix} = \begin{bmatrix} -0.02 \\ -0.22 \end{bmatrix}$$

$$\begin{bmatrix} k_3 \\ l_3 \end{bmatrix} = hF(x_0 + h/2, \begin{bmatrix} y_0 + 1/2 k_2 \\ z_0 + 1/2 l_2 \end{bmatrix}) =$$

$$h \begin{bmatrix} z_0 + l_3 \\ (z_0 + l_3)(y_0 + k_3)^2 - (y_0 + k_3) \end{bmatrix} = \begin{bmatrix} -0.022 \\ -0.219562 \end{bmatrix}$$

$$\begin{bmatrix} k_4 \\ l_4 \end{bmatrix} = hF(x_0 + h/2, \begin{bmatrix} y_0 + k_3 \\ z_0 + l_3 \end{bmatrix} = h \begin{bmatrix} z_3 + l_3 \\ (z_0 + l_3)(y_0 + k_3)^2 - (y_0 + k_3) \end{bmatrix}$$

$$\begin{bmatrix} k_4 \\ l_4 \end{bmatrix} = \begin{bmatrix} -0.0439124 \\ -0.237602 \end{bmatrix}$$

Na koncu dobimo

$$y_1 = y_0 + \frac{1}{2}(k_1 + 2k_2 + 2k_3 + k_4) = 0.978681$$

in

$$z_1 = z_0 + \frac{1}{2}(l_1 + 2l_2 + 2l_3 + l_4) = -0.216453.$$

■

**Naloga 9.21** *Poišči red formule*

$$y_{n+1} - y_n = \frac{h}{12}(5y'_{n+1} + 8y'_n - y'_{n-1}) + O(h^{p+1}).$$

*Ali je metoda ničelno stabilna?*

*Rešitev.* Preverimo za katere polinome $y = x^i$ je formula točna. Za $y = 1$ je formula očitno točna. Za $y = x$ dobimo

$$x_n + h - x_n = \frac{h}{12}12 = h.$$

Formula je točna. Za $y = x^2$ dobimo

$$(x_n + h)^2 - x_n^2 = 2x_nh + h^2 = \frac{h}{12}(10(x_n + h)16x_n - 2(x_n - h) = 2x_n + h^2.$$

Za $y = x^3$ dobimo

$$(x_n + h)^3 - x_n^3 = \frac{h}{12}(15(x_n + h)^2 + 24x_n^2 - 3(x_n - h)^2),$$

torej je formula spet točna. Ker mora biti formula točna za vsak $x_n$ in vsak $h$, lahko izberemo $x_n = 0$ Tako za $y = x^4$ dobimo

$$h^4 = \frac{h}{12}(20h^3 - 4h^3) + O(h^4) = \frac{4}{3}h^4 + O(h^4).$$

Formula je torej reda 3.

Formula je tudi ničelno stabilna. Vstavimo $f(x, y) = 0$. Tako dobimo $y_{n+1} - y_n = 0$. Polinom $\sigma(\zeta) = \zeta^2 - \zeta$ ima enostavno ničlo 1, torej je ničelna stabilnost zagotovljena.

Formula je ničelno stabilna, če za enostavno ničlo $\sigma$ velja $|\sigma| \leq 1$. Za večkratno ničlo pa mora veljati $|\zeta| < 1$.

■

**Naloga 9.22** ***Naloga ni bila narejena na vajah***
*Robni problem $y''(x) + 2y'(x) - 2y(x) = 3$, $y(0) = 0$, $y(1) = 2$ numerično rešujemo z diferenčno metodo. Pri tem prve in druge odvode nadomestimo s simetričnimi diferencami skozi tri sosednje točke. Zapiši dobljeni sistem linearnih enačb in ga reši v primeru $h = 1/3$.*

*Rešitev.* V točki $x_i$ se enačba spremeni v

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + 2\frac{y_{i+1} - y_{i-1}}{2h} - 2y_i = 3.$$

Tako dobimo enačbo

$$y_{i+1}(1+h) + y_i(2+2h^2) + y_{i-1}(1-h) = 3h^2.$$

Če je $h = \frac{1}{3}$ dobimo sistem

$$y_2(1+\frac{1}{3}) - y_1(2+2\frac{1}{9}) + y_0\frac{2}{3} = \frac{1}{3}$$

$$y_3(1+\frac{1}{3}) - y_2(2+2\frac{1}{9}) + y_1\frac{2}{3} = \frac{1}{3}.$$

Sistem rešimo in dobimo $y_1 = \frac{24}{41}$, $y_2 = \frac{201}{164}$. ■