Poglavje 1

Naloge iz Matlaba

Naloga 1.1 *Z Matlabom pretvori število* -1.6 *v zapis s plavajočo vejico.*

Naloga 1.2 (Numerična stabilnost) *Preveri, poišči taka* x *in* y*, da potrdiš stabilnost/nestabilnost izračuna:*

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

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

(iii). $\sqrt{1+x} - 1$, $\frac{x}{1+\sqrt{1+x}}$ *za majhne* x*,*

(iv). $\sqrt{x^2+x} - x$, $\frac{1}{1+\sqrt{1+\frac{1}{x}}}$ *za velike* x*,*

(v). $\tan(x) - \sin(x)$, $\frac{2\sin^2(x/2)\sin(x)}{\cos(x)}$ *za majhne* x*.*

Namig, kaj se zgodi, ko računamo $1+10^{-16}$ *ali* $10^{17}+10^{34}$*? Preveri v Matlabu. Ali znaš pojasniti, zakaj ne velja* $(10^{32}+10^{16}) == 10^{32}$*?*

Naloga 1.3 *Preoblikuj naslednje izraze, tako da bo njihov izračun numerično stabilen za* $x \approx 0$*. Za podano vrednost* x *jih tudi izračunaj.*

(i). $\exp(x) - \exp(-x), x = 10^{-5}$

(ii). $1 - \cos(x), x = 10^{-5}$

(iii). $\frac{1}{\sqrt{1+x^2}-\sqrt{1-x^2}}, x = 10^{-2}$

Naloga 1.4 *Za vsak* $x_0 > -1$ *naslednje rekurzivno definirano zaporedje:*

$$x_{n+1} = 2^{n+1}\left(\sqrt{1+2^{-n}x_n} - 1\right)$$

konvergira k $\ln\left(1 + x_0 + \frac{1}{4}x_0^2\right)$*. Vendar, če v Matlabu poizkusiš z* $x_0 = 4$*,*

```
x = 4;
for n = 1 : 100
    x = 2^(n+1)*(sqrt(1+x/2^n - 1);
end;
x
```

dobiš popolnoma napačne rezultate. Pojasni problem in ga popravi.

Naloga 1.5 *Razišči gostoto predstavljivih števil.*

(i). Katero je najbližje predstavljivo število $10^{300}, 1, 10^{-30}, 10^{-300}$*?*

(ii). Pomagaj si z ukazom eps v Matlabu? Lahko si funkcijo eps tudi narišeš.

(iii). Napovej največjo pozitivno in najmajše pozitivno predstavljivo število v dvojni aritmetiki.

(iv). Preveri, če imaš prav. Ali se napoved ujema z realmin, realmax?

Naloga 1.6 *Uporabi bisekcijo za iskanje ničle enačbe* $x^3+4x^2-10=0$ *na intervalu* $[0,2]$. *Nariši graf napake (v logaritemski skali) kot funkcijo števila iteracij. Kaj opaziš?*

Naloga 1.7 *Uporabi bisekcijo za iskanje rešitve enačbe* $\tan(x)=0$, *kjer je* x *med* 1 *in* 2. *Ali dobiš pravilen rezultat? Komentiraj?*

Naloga 1.8 *Poišči rešitev enačbe* x^2-3x+2 *z bisekcijo z začetnima vrednostima* -2.4 *in* 0, *tangentno metodo za začetni približek* -2.4 *in sekantno metodo z začetnima vrednostima* -2.4 *in* 0. *Nariši graf napake (v logaritemski in navadni skali) za vse metode.*

Naloga 1.9 *Kaj opaziš, če v prejšnji nalogi namesto točke* -2.4 *uporabiš točko* 1.2.

Naloga 1.10 *Poišči rešitev enačbe* $x^3-x+3=0$ *z uporabo tangentne metode z začetno točko* 0 *in sekantne metode z začetnima točkama* 2 *in* 1. *Nariši graf napake v logaritemski skali za obe metodi. Kaj opaziš?*

Naloga 1.11 *Reši enačbo* $xe^{-x}=0$ *s sekantno metodo z začetnima točkama* 1.5 *in* 2? *Kaj opaziš? Kaj pa, če uporabiš tangentno metodo z začetnim približkom* 2?

Naloga 1.12 *Reši enačbo* $\arctan(x)=0$ *s sekantno metodo z začetnima točkama* 1.5 *in* 1.6. *Kaj se dogaja? Ali je situacija drugačna, če uporabiš tangentno metodo z začetnim približkom* 1.4.

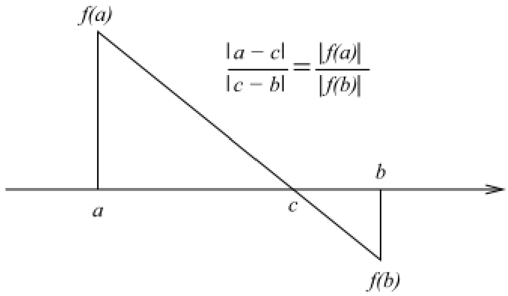
Ali lahko potegneš kakšen zaključek? Večino že pripravljenih skript lahko najdeš na strani http://www-lp.fmf.uni-lj.si/plestenjak/bor.htm izr. prof. Bora Plestenjaka.

Naloga 1.13 *Preizkusi, kako dober je približek je Stirlingovo število* $n^n e^{-n}\sqrt{2\pi n}$ *za* $n!$. *Tabeliraj* $n!$, *pripadajoče Stirlingovo število in aboslutno in relativno napako, za* $n=1,\ldots,10$. *Kaj pa, če bi rad tabeliral samo napake za* $n=1,\ldots,10000$. *Kaj se dogaja z napakami?*

Naloga 1.14 *Pokaži, da ima enačba* $f(x)=x+a\cos(x)$ *vsaj eno rešitev za vsako realno število* a. *Uporabi bisekcijo za poseben primer, ko je* $a=14$, *za začetni točki pa si izbereš* 14 *in* 16. *Ali znaš oceniti, koliko iteracij je potrebno, da bo rezultat izračunan maksimalno natančno v plavajoči vejici?*

Naloga 1.15 *Za izračun ničle prejšnje metode uporabi metodo regula-falsi, ki je podobna biskeciji. Pri metodi regula-falsi intervala ne razpolavljamo, ampak ga delimo v razmerju, da velja*

$$\frac{|a-c|}{|c-b|}=\frac{|f(a)|}{|f(b)|}.$$



Preizkusi svojo metodo na naslednjih problemih.

(i). Izračunaj π *kot ničlo enačbe* $\sin(x) = 0$ *z začetnima točkama* 2 *in* 4.

(ii). Naslednji test je rešitev enačbe $x - e^x = 0$ *za začetnim intervalom* $a = -1$ *in* $b = 0$.

(iii). Poišči vse ničle funkcije $f(x) = x + 4\cos(x)$.

Naloga 1.16 *Preizkusiš lahko tudi zgleda s predavanj s tračnico in računanjem števila* π.

Poglavje 2

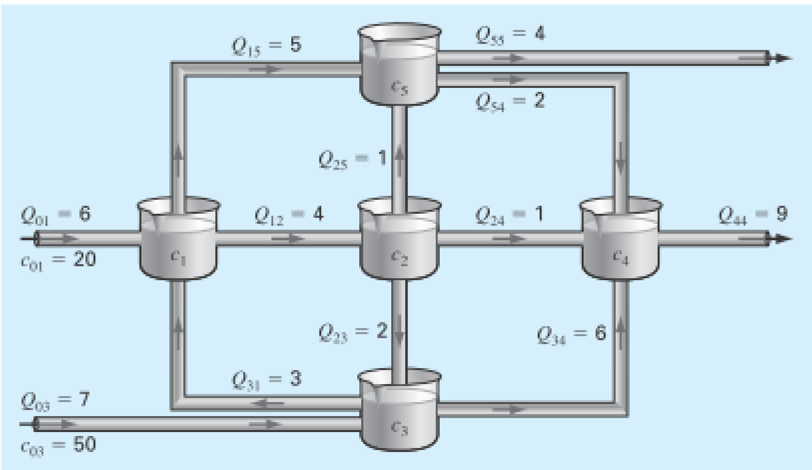
Matlab, norme, linearni sistemi

Naloga 2.1 *Z Matlabom poišči protiprimer, ki pokaže, da*

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

ni matrična norma. Pomagaj si z generiranjem naključnih matrik. Ali ti uspe poiskati celoštevilsko matriko?

Naloga 2.2 *Pet reaktorjev je povezanih s cevmi na način prikazan na spodnji sliki.*



Masni pretok skozi vsako cev izračunamo kot produkt toka (Q) *in koncentracije* (c)*. V ravnovesju sta masni pretok iz reaktorja in v reaktor enaka. Za prvi masni reaktor dobimo enačbo:*

$$Q_{01}c_{01} + Q_{31}c_3 = Q_{15}c_1 + Q_{12}c_1.$$

Zapiši preostale mankajoče enačbe in jih reši v MATLABU. Kaj lahko poveš o občutljivosti sistema?

Naloga 2.3 *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

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

Od tod dobimo

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

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

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

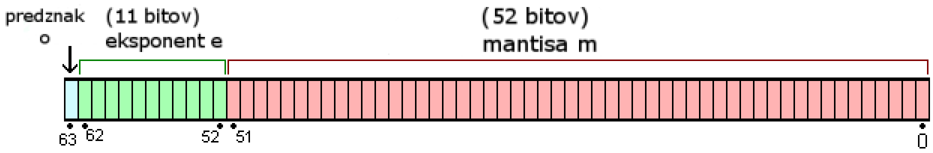
Hilbertove matrike so primer zelo občutljivih matrik. Pogojenostno število H_5 *je recimo približno* $4.766 \cdot 10^5$*, kar lahko izračunamo z ukazom cond v Matlabu. Pogojenostna števila matrik* H_n *hitro rastejo. Nariši graf pogojenostnega števila v odvisnosti od* n*. Za vsako stopnjo polinoma izračunaj še napako* E *in nariši njen graf, če aproksimiraš funkcijo* $\sin(\pi x)$*. Primerjaj dobljene rezultate z oceno s predavanj.*

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

Naloga 2.4 *Napiši funkcijo v MATLABu v kateri implemtiraš Gaussovo eliminacijo z delnim pivotiranjem (če uporabnik zahteva vrni tudi matriki* L *in* U*, ki tvorita* LU *razcep). Funkcija naj vrne še determinatno matrike in prepozna, če je sistem blizu 'singularnega' (determinanta blizu nič, tj. manj kot predpisana toleranca). A je to dober kriterij? Namig, velike digaonalne matrike. Prva vrstic funkcije naj bo* `function [x, D] = GaussPivotDet(A, b, tol),` *kjer je D determinanta in tol predpisana toleranca. Pri delnem pivotiranju je pivotna rast omejena z* 2^{n-1}*, ponavadi pa je* $O(n^{2/3})$*. Testiraj svojo funkcijo na Hilbertovih matrikah, kakšna je pivotna rast v tem primeru?*

Poglavje 3

Plavajoča vejica



Slika 3.1: 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 3.1 (Vaje) *Zapiši naslednja števila v dvojni natančnosti:*

(i). 2.71875,

(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 $2.71875 = 10.10111_{(2)}$. Ker iščemo zapis normalizirane oblike $(-1)^o(1+m)2^{e-1023}$, moramo število deliti/množiti z dva dokler ne dobimo 1.m. Torej $10.10111_{(2)} = 1.01011_{(2)} * 2$ in $e - 1023 = 1$. Dobili smo $o = 0$, $m = 010111\underbrace{0\ldots0}_{46}$, $e = 10000000000_{(2)}$.

(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 zaokroževanja. Število, ki je najbližje $0.*11001\underbrace{1}$, 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 $e = 10000000110_{(2)}$.

■

Naj bo x število in $fl(x)$ najbližje predstavljivo število. Velja

$$fl(x) = x(1+\delta) \text{ in } |\delta| \leq u,$$

kjer je u osnovna zaokrožitvena napaka. Za predstavljivo število velja $fl(x) = x$.

Naloga 3.2 (Vaje) *Podan je tangens kota* $\tan(\alpha) = t$. *Poišči učinkovito metodo za izračun* $\cos(\alpha)$ *in* $\sin(\alpha)$ *brez uporabe* $\arctan$. *Poizkusi se izogniti podkoračenju pri računanju v premični piki. Kje nam podkoračenje prinese največ težav? Kaj se zgodi pri kvadriranju* $t = 2^{-600}$ *v Matlabu?*

Rešitev. Spomnimo se trigonometričnih zvez $1 + \tan^2 \alpha = \frac{1}{\cos^2 \alpha}$ in $\sin^2 \alpha + \cos^2 \alpha = 1$. Definirajmo $r = \sqrt{1 + t^2}$, $c = \frac{1}{r}$ in $s = \sqrt{1 - c^2}$. Pri računanju $\sin \alpha$ bo lahko prišlo do podkoračenja pri majhnem t. Temu se izognemo na sledeči način

$$s = \sqrt{1 - c^2} = \sqrt{1 - \frac{1}{r^2}} = \sqrt{\frac{r^2 - 1}{r^2}} = \sqrt{\frac{t^2}{1 + t^2}} = \frac{t}{r},$$

kjer se izognemo računanju t^2. ■

Naloga 3.3 (Vaje) *Izračunati želimo vrednost funkcije* $f(x) = \sqrt{1 + x}$ *pri* $x = \frac{1}{13}$, *pri čemer računamo v premični piki na pet decimalk natančno in z zaokrožanjem na najbližje predstavljivo število. Za računanje uporabimo prve tri člene razvoja* $f(x)$ *v taylorjevo vrsto. Oceni celotno napako in jo primerjaj z dejansko.*

Rešitev. Razvoj v taylorjevo vrsto je

$$\sqrt{1 + x} = \sum_{i=0}^{\infty} \binom{\frac{1}{2}}{i} x^i = 1 + \binom{\frac{1}{2}}{1} x + \binom{\frac{1}{2}}{2} x^2 + \ldots = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{6}f^3(\xi)x^3.$$

Napaka pri numeričnem računanju je sestavljena iz neodstranljive napake, napake metode in zaokrožitvene napake. Velja

$$D = D_n + D_m + D_z \text{ in seveda } |D| \leq |D_n| + |D_m| + |D_z|.$$

- **Osnovna zaokrožitvena napaka**
  Velja $u = \frac{1}{2}10^{-5} = 5 \cdot 10^{-6}$ in $\overline{x} = fl(\frac{1}{13}) = 10^{-1} \cdot 0.76923$.

- **Neodstranljiva napaka**
  Namesto z $x = \frac{1}{13}$ računamo z $\overline{x} = 0.07692$, $D_n = f(x) - f(\overline{x})$. Velja

  $$|f(x) - f(\overline{x})| = |f'(\zeta)||x - \overline{x}| \text{ za } \zeta \text{ med } x \text{ in } \overline{x}.$$

  Ocenimo odvod

  $$f'(x) = \frac{1}{2\sqrt{1 + x}} \leq \frac{1}{2}.$$

  Tako dobimo oceno za

  $$|D_n| \leq \frac{1}{2}|x - \overline{x}| \leq \frac{1}{2}u = \frac{1}{2} \cdot 5 \cdot 10^{-6}.$$

  Točna vrednost

  $$D_n = \sqrt{1 + \frac{1}{13}} - \sqrt{1 + 10^{-1}0.76923} \doteq 3.706 \cdot 10^{-8}.$$

- **Napaka metode**
  Namesto vrednosti $f(\overline{x})$ računamo $g(\overline{x})$.

  $$D_m = f(\overline{x}) - g(\overline{x}) = \sqrt{1 + \overline{x}} - (1 + \frac{1}{2}\overline{x} - \frac{1}{8}\overline{x}^2) \leq \frac{1}{6}f^{(3)}(\xi_{\overline{x}})\overline{x}^3 \leq \frac{1}{16}\overline{x}^3 \leq 2.85 \cdot 10^{-5}.$$

  Točna vrednost $D_m = 2.8444788 \cdot 10^{-6}$.

- **Zaokrožitvena napaka** nastane zaradi zaokroževanja med računanjem.
  Najprej moramo določiti vrstni red računanja, saj operacije niso več asociativne. Denimo, da najprej izračunamo produkte in nato še ostalo. Poizkusi oceniti napako. Oceni še celotno napako.

■

Poglavje 4

Numerična stabilnost

Algoritem je **direktno stabilen**, če vrne rezultat, ki se malo razlikuje od prave vrednosti. Ponavadi preverjamo relativno direktno stabilnost. Direktna absolutna in relativna napaka sta

$$|y - \widehat{y}| \text{ in } \frac{|y - \widehat{y}|}{|y|}.$$

Algoritem je **obratno stabilen**, če za izračunan rezultat obstajajo taki malo zmoteni podatki, da iz njih s točnim izračunom dobimo izračunano vrednost. Obratna absolutna napaka je najmanjši $|\Delta x|$, tako da velja $f(x + \Delta x) = \widehat{y}$. Obratna relativna napaka je $\frac{|\Delta x|}{|x|}$.

Pri obratni napaki je izračunana vrednost enaka $\widehat{y} = f(\widehat{x})$. Če je funkcija f *zvezno odvedljiva* v x, potem velja

$$|\widehat{y} - f(x)| = |f(\widehat{x}) - f(x)| \leq |f'(x)||\widehat{x} - x|.$$

Če f ni absolutno občutljiva, bo pri majhnih vrednostih $\Delta x = |\widehat{x} - x|$ napaka absolutno obratno stabilne metode majhna in metoda bo absolutno direktno stabilna. Podobno velja za relativno stabilnost in relativne napake. Več ste povedali na predavanjih.

Naloga 4.1 (Vaje) *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:

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

Ocena je smiselna. Najlažje to vidimo tako, da izberemo $\delta_1 = u$, $\delta_2 = -u$. 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: $\widehat{x} = x\sqrt{(1+\delta)}$, $\widehat{y} = y\sqrt{(1+\delta)}$.

■

Poglavje 5

Nelinearne enačbe

Naloga 5.1 (Vaje) *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 $f(a) = f(-0.9) = (-0.9)^3 + 1 = 0.271$, $f(b) = f(-1.2) = (-1.2)^3 + 1 = -0.728$, $\texttt{sign}(f(a)f(b)) = -1$. Zato lahko izvedemo korak bisekcije. Nadaljujmo: $c = (a + b)/2 = -1.05$, $f(c) = (-1.05)^3 + 1 = -0.157625$, $\texttt{sign}(f(a)f(c)) < 0$, 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 α 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*

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

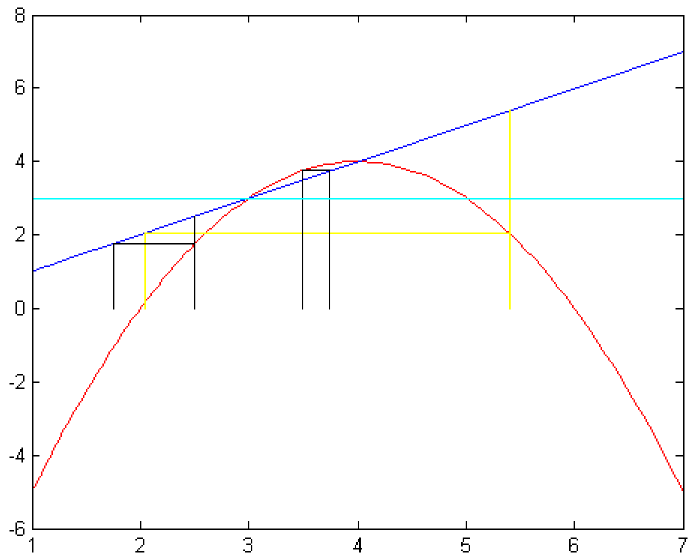
konvergira k α in velja ocena za napako

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

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 α in velja $f(\alpha) = 0$.*

Izrek 5.3 *(Banachovo skrčitveno načelo) Naj bo M poln metričen prostor in $g : M \to M$. Tedaj obstaja natanko ena negibna (fiksna) točka preslikava g, t.j. taka točka $a \in M$, da je $g(a) = a$. Če je $x_0 \in M$ poljubna točka, tedaj zaporedje $x_{r+1} = g(x_r)$ konvergira k a. Primer polnega metričnega prostora v $\mathbb{R}$ je zaprti interval.*

Naloga 5.2 (Vaje) *Za katere začetne približke je navadna iteracija za reševanje $x = g(x)$, kjer je $g(x) = -12 + 8x - x^2$, konvergentna? Kam konvergira zaporedje $x_{r+1} = g(x_r)$? Kakšen je red konvergence? Kje nam konvergenco zagotavlja izrek (Banachovo skrčitveno načelo) ? Namig: začetne približke išči na intervalu.*



Slika 5.1: Koraki iteracije

Rešitev. Iz slike preberemo, da je kandidat za območje konvergence interval $[3, 5]$. V naslednjih točkah bomo pokazali, da je ta interval res dobra izbira.

(i). Pokažimo, da velja: $x_r < 3 \Rightarrow x_{r+1} < x_r$. Za take začetne približke bomo imeli divergenco. 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: $x_0 > 5 \Rightarrow x_1 < 3$. Spet dobimo divergenco. 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

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

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 (Vaje) *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 $[0,1]$, kjer je funkcija skrčitev. Vidimo, da velja $m \leq \frac{1^4}{2} = 0.5$. Imamo $x_1 = 0.1$, $x_2 = 0.100001$ in

$$|x_2 - \alpha| \leq \frac{0.5}{1-0.5}|0.100001 - 0.1| \doteq 8 \cdot 10^{-6}.$$

Če hočemo uporabiti izrek 5.2, izberemo recimo interval $[-0.2, 0.2]$, kjer je $d = 0.2$. Na tem intervalu ocenimo $m \leq \frac{0.2^4}{2} \leq 0.0008$. Poleg tega velja tudi $|g(0) - 0| = 0.1 \leq (1 - 0.0008)0.2$. Vsi pogoji so izpolnjeni. ∎

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 5.4 *Za tangentno metodo velja naslednji izrek. Naj velja $f(\alpha) = 0$ in naj bo α m-kratna ničla. Če je $m = 1$, je konvergenca vsaj kvadratična. Če velja še $f''(\alpha) = 0$, je konvergenca kubična. Če imamo večkratno ničlo $m \geq 2$, velja $\lim_{x\to\alpha} g'(x) = 1 - 1/m$.*

Rešitev. Na vajah smo ubrali drugo pot do rešitve, ki ni bila preveč pregledna. V tej rešitvi je predstavljen preglednejši dokaz z uporabo pravila L'Hopital.

Dokazali bomo samo drugi del izrek, saj je prvi enostaven. Če je α 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 (Vaje) *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

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

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. ∎

Naloga 5.5 (Vaje) *Določi vse polinome četrte stopnje z vodilnim koeficientom 1, pri katerih se tangentna metoda veda takole:*

- *v bližini α 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 α 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

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

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 -2α je enostavna, konvergenca je vsaj kvadratična. Izračunajmo še vrednost drugega odvoda v -2α:

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

Konvergenca v -2α je kvadratična. ∎

Naloga 5.6 (Vaje) *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 $r = 0, 1, \ldots$ ∎

Poglavje 6

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 6.1 *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 $y = Bx$.

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

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

Dokaz je krajši, če upoštevamo definicijo norme in ugotovimo, da velja $||Ax|| \leq ||A||||x||-$ 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 6.2 (Vaje) *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* λ_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 številih 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 $2 \not\leq 21$. Submultiplikativnost ni izpolnjena. Alternativna možnost je, da take matrike poizkusimo poiskati v Matlabu z ukazom `randi`, ku generira celoštevilske naključne matrike
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 6.3 (Vaje) *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

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

Lastne vrednosti A^HA so nenegativne, saj velja $< A^HAx, x >=< \mu x, x >= \mu =< Ax, Ax >\geq 0$, $\lambda_i = \sigma_i^2$. 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 $a_{ij} = e_i^T A e_j$. Računajmo

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

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

Naloga 6.4 (Vaje) *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 $||A||_2$

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

Skupaj dobimo oceno $5.7735 \leq ||A||_2 \leq 8.06226$. 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 $\lambda_i(B) \leq ||B|_F = 50.0899$. Vemo, da velja

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

Oceno navzdol dobimo, če upoštevamo $||A||_2 = ||A^T||_2$ in $||Ae_i||_2 \leq ||A||_2||e_i||_2 = ||A||_2$, zadnja enakost sledi iz definicije matrične norme. Vektorji Ae_i so ravno i-ti stolpci, $A^T e_i$ so i-te vrstice. Norme vrstic so $5.7446,\ 4.2426, 3.7417$. Norme stolpcev so $3.7417,\ 6.4807,\ 3$. Torej dobimo $||A||_2 \geq 6.4807$. Končna ocena je $6.4809 \leq ||A||_2 \leq 7.0774$. Ukaz iz Matlaba vrne `norm`$(A) = 6.9044$. V Matlabu preizkusi še, če smo tudi druge norme izračunali pravilno. ■

Naloga 6.5 *Matrična norma* $||.||_M$ *je usklajena z vektorsko normo* $||.||_V$*, če za vsako matriko* A *in vsak vektor* x *velja* $||Ax||_V \leq ||A||_M||x||_V$.

Dokaži:

a) *Za vsako matrično normo* $||.||_M$ *obstaja taka vektorska norma* $||.||_V$*, ki je z njo usklajena.*

b) *Za vsako lastno vrednost matrike* A *in za poljubno matrično normo velja ocena* $|\lambda(A)| \leq ||A||_M$.

c) *Velja* $||A||_2^2 \leq ||A||_1||A||_\infty$.

Rešitev. a)

Imamo matrično preslikavo $||.||_M$. Naj bo $\overline{M}$ preslikava, ki vektorju x priredi matriko $\overline{M}(x) = \begin{bmatrix} x & 0 & \dots & 0 \end{bmatrix}$. Definirajmo vektorsko normo $||x||_V := ||\overline{M}(x)||_M$. Tako dobimo

$$
\begin{aligned}
||Ax||_V = ||\begin{bmatrix} Ax & 0 & \dots & 0 \end{bmatrix}||_M = ||A\begin{bmatrix} x & 0 & \dots & 0 \end{bmatrix}||\\
\leq ||A||_M||\begin{bmatrix} Ax & 0 & \dots & 0 \end{bmatrix}||_M = ||A||_M||x||_V.
\end{aligned}
$$

b)
Naj bo λ lastna vrednost, $Av = \lambda v$. Po točki a) obstaja vektorska norma, ki je usklajena z matrično. Tako sledi

$$|\lambda|||v||_V = ||\lambda v||_V = ||Av||_V \leq ||A||_M ||v||_V,$$

kar pomeni, da velja $|\lambda| \leq ||A||_M$.
c)
Vemo, da je $||A||_2^2$ največja lastna vrednost matrike $A^H A$. Iz b) dobimo

$$||A||_2^2 \leq ||A^H A||_\infty \leq ||A^H||_\infty ||A||_\infty = ||A||_1 ||A||_\infty.$$

■

Naloga 6.6 *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

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

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 7

Reševanje linearnih sistemov

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

Naloga 7.1 *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

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

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 7.1 *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 7.2 *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 7.2 *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:

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

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

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

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

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

Naloga 7.3 *Sestavi učinkovit algoritem za reševanje sistema linearnih enačb*

$$\begin{bmatrix} U & -I \\ B & L \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Matrika $B = LU$ je nesingularna matrika in ima podan LU razcep. Preštej število množenj in deljenj.

Rešitev. Zmnožimo po blokih in dobimo

$$\begin{aligned} Ux - y &= a \\ L(Ux + y) &= b. \end{aligned}$$

Rešimo drugi sistem in dobimo vektor $z = Ux + y$, kar nas stane $n^2 + O(n)$ operacij. Označimo še $w = Ux - y$ in izrazimo $y = \frac{1}{2}(z - w) = \frac{1}{2}(z - a)$. Za izračun y porabimo $O(n)$ operacij. Na koncu rešimo še sistem $Ux = a + y$, za kar porabimo $n^2 + O(n)$ operacij. Skupaj smo porabili $2n^2 + O(n)$ operacij. ∎

Naloga 7.4 *Za matriko*

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

naredi:

(i). LU razcep z delnim pivotiranjem,

(ii). LU razcep s kompletnim pivotiranjem.

Rešitev. (i)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 $PA = LU$.

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

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \overset{1\updownarrow 3}{\rightarrow} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \overset{G}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1/2 & 0 & 0 \end{bmatrix} \overset{2\updownarrow 3}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = L$$

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

(ii)Algoritem poteka takole. V podmatriki $A(i:n, i:n)$ poiščemo največji element po absolutni vrednostni, recimo da je to a_{kl}. Da ta element spravimo na mesto (i,i) zamenjamo i-to vrstico in k-to vrstico ter i-ti stolpec in k-ti stolpec. To naredimo tudi za matriko L. V P menjamo samo vrstice, v Q pa samo stolpce. Normalno nadaljujemo z Gaussovo eliminacijo. Na koncu velja $PAQ = LU$.

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 0 & 1 \\ 6 & 1 & \underline{7} \end{bmatrix} \overset{1\updownarrow 3, 1\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 7 & 1 & 6 \\ 1 & 0 & 0 \\ 1 & 1 & 3 \end{bmatrix} \overset{G}{\rightarrow} \begin{bmatrix} 7 & 1 & 6 \\ 0 & -\frac{1}{7} & -\frac{6}{7} \\ 0 & \frac{6}{7} & \underline{\frac{15}{7}} \end{bmatrix} \overset{2\updownarrow 3, 2\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 7 & 6 & 1 \\ 0 & \frac{15}{7} & \frac{6}{7} \\ 0 & -\frac{6}{7} & \frac{1}{7} \end{bmatrix} \overset{G}{\rightarrow} \begin{bmatrix} 7 & 6 & 1 \\ 0 & \frac{15}{7} & \frac{6}{7} \\ 0 & 0 & \frac{1}{5} \end{bmatrix} = U$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \overset{1\updownarrow 3, 1\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \overset{G}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 1/7 & 0 & 0 \\ 1/7 & 0 & 0 \end{bmatrix} \overset{2\updownarrow 3, 2\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 1/7 & 0 & 0 \\ 1/7 & 0 & 0 \end{bmatrix} \overset{G}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 1/7 & 1 & 0 \\ 1/7 & -2/5 & 1 \end{bmatrix} = L$$

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{1\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \overset{2\leftrightarrow 3}{\rightarrow} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = Q$$

∎

Naloga 7.5 *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:

$$L = \begin{bmatrix} 1 & & & \\ l_2 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & l_n \end{bmatrix} \text{ in } U = \begin{bmatrix} u_1 & c_1 & & \\ & \ddots & \ddots & \\ & & u_{n-1} & c_{n-1} \\ & & & u_n \end{bmatrix}$$

- 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 $b \leq 2a$.

Algoritem s pivotiranjem je prepuščen za domačo nalogo.

∎

Naloga 7.6 *Dana je nesingularna matrika A reda n skupaj z inverzno matriko A^{-1}. Zapiši algoritem za izračun inverzne matrike B^{-1}, ki je*

$$B = \begin{bmatrix} A & u \\ v^T & \alpha \end{bmatrix},$$

kjer je $u, v \in \mathbb{R}^n, \alpha \in \mathbb{R}$.

Rešitev. Naj bo

$$B^{-1} = \begin{bmatrix} C & x \\ y^t & \beta \end{bmatrix}.$$

Zmnožimo po bločnih komponentah in dobimo

$$BB^{-1} = \begin{bmatrix} A & u \\ v^T & \alpha \end{bmatrix} \begin{bmatrix} C & x \\ y^t & \beta \end{bmatrix} = \begin{bmatrix} AC + uy^T & Ax + \beta u \\ v^T C + \alpha y^T & v^T x + \alpha\beta \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & 1 \end{bmatrix}.$$

Dobimo sistem

$$AC + uy^T = I \tag{7.1}$$
$$Ax + \beta u = 0 \tag{7.2}$$
$$v^T C + \alpha y^T = 0 \tag{7.3}$$
$$v^T x + \alpha\beta = 1 \tag{7.4}$$

Iz prve enačbe izrazimo C in ga vstavimo v 3 in 4 enačbo. Dobimo $C = A^{-1} - \overbrace{A^{-1}u}^{z} y^T = A^{-1} - zy^T$ in $x = -\beta z$.

Ko vstavimo v tretjo enačbo dobimo

$$v^T(A^{-1} - zy^T) + \alpha y^T$$

in

$$v^T(-\beta z) + \alpha\beta = 1.$$

Iz zadnje zveze lahko takoj izrazimo

$$\beta = \frac{1}{\alpha - v^T z}.$$

Poglejmo še tretjo enačbo, sledi

$$v^T A^{-1} - v^T z y^T + \alpha y^T = v^T A^{-1} + y^T(\alpha - v^T z) = 0.$$

Tako je

$$y^T = \frac{v^T A^{-1}}{v^T z - \alpha}.$$

Napišimo še skico algoritma

(i). $z = A^{-1}u \qquad n \times (n*, n-1+) \to n^2*, n^2 - n + |2n^2 - n$;

(ii). $\beta = \frac{1}{\alpha - v^T z} \qquad (n*, n-1+, 1-, 1/)|2n+1$;

(iii). $x = -\beta z \qquad (n*, 1-)|n+1$;

(iv). $y = -\beta(v^T A^{-1})^T = -\beta(A^{-1})^T v \qquad |2n^2 - n + 1$;

(v). $C = A^{-1} - zy^T \qquad (n^2*, n^2-)|2n^2$.

Skupaj imamo $6n^2 + n + 3$ operacij.

■

Naloga 7.7 *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.

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

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

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

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

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

Poglavje 8

Problemi najmanjših kvadratov, predoločeni sistemi

Naloga 8.1 *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 8.2 *Merili smo tir delca, ki naj bi se 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 8.3 *Naj bo $A \in \mathbb{R}^{m \times n}$, kjer je $m \geq n$ in $rang(A) = n$. Dokaži, da ima sistem*

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

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

Rešitev. 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. ■

Izrek 8.1 *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 ρ.

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 8.4 *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:

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

---

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**
    $aik = a_{ik}$;
    $a_{ik} = ca_{ik} + sa_{i+1,k}$ -|- $2\sum_{\hat{i}=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$;
    $a_{i+1,k} = -s \cdot aik + ca_{i+1,k}$ -|- $2\sum_{i=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$;
  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. ∎

Householderjevo zrcaljenje je matrika oblike $H = I - \frac{2}{w^T w} ww^T$ 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 8.5 *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

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

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

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

Naloga 8.6 *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,*
- *s pomočjo Householderjevih zrcaljenj.*

Rešitev.

- 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 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 $Q = R_{12}R_{13}R_{23}$.

■

Singularni razcep matrike $A = U\Sigma V^*$, je sestavljen iz ortogonalnih matrik U in V. Matrika Σ je diagonalna, elementi na diagonali Σ_i so ravno singularne vrednosti. S pomočjo singularnega razcepa lahko definiramo psevdoinverz matrike $A^+ = V\Sigma^+U^*$, 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 $S = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$. Rešitev predoločenega sistema $Ax = b$, lahko izrazimo kot $x = A^+b = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i$.

Poglavje 9

Lastne vrednosti in vektorji

Naloga 9.1 *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* $i = 1, \dots, n$.. *Vse lastne vrednosti matrike* A *ležijo v uniji krogov* $\cup_{i=1}^{n} C_i$.

Rešitev. Naj bo x lastni vektor in λ 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 9.2 *Naj bo* $A \in \mathbb{R}^{n\times n}$ *diagonalno dominantna, tj.*

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

Potem je A *obrnljiva.*

Rešitev. Dovolj je pokazati, da so vse lastne vrednosti različne od 0. Fiksirajmo lastno vrednost λ. 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 9.1 *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 9.2 *Velja $|\lambda| \leq ||A||_\infty$.*

Dokaza posledic sta preprosta in ju prepuščam bralcu.

Naloga 9.3 *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 za območje lahko izboljšamo tako, da isto oceno naredimo za A^T in podobno matriko DAD^{-1}, kjer je D diagonalna matrika. Vsako območje je določeno z unijo krogov. Lastne vrednosti se nahajajo v preseku vseh območij. ∎

Naloga 9.4 *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 λ 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 6: 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 9.5 *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 λ_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

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

■

Algoritem 7: 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 8: *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 9.6 *Izračunaj lastne vrednosti in lastne vektorje tridiagonalne matrike dimenzije $n \times n$, kjer je $a_{i,i-1} = c$, $a_{ii} = a$, $a_{i,i+1} = b$ in $bc > 0$. Pomagaj si z ustrezno diferenčno enačbo s konstatnimi koeficienti.*

$$A = \begin{bmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{bmatrix}$$

Rešitev. Upoštevamo, da mora veljati $Ax = \lambda x$. Tako dobimo sistem enačb:

$$
\begin{aligned}
(a-\lambda)x_1 + bx_2 &= 0 && (9.1)\\
cx_{i-1} + (a-\lambda)x_i + bx_{i+1} &= 0, \quad \text{i = 2, \ldots, n-1} && (9.2)\\
cx_{n-1} + (a-\lambda)x_n &= 0 && (9.3)
\end{aligned}
$$

Zaradi lažjega računanja uvedemo še spremenljivki x_0 in x_n ter robna pogoja $x_0 = x_n = 0$. Torej rešujemo diferenčno enačbo

$$cx_{i-1} + (a-\lambda)x_i + bx_{i+1} = 0.$$

Uporabimo nastavek za homogeno rešitev $x_i = r^i$ in dobimo kvadratno enačbo za r,

$$c + (a-\lambda)r + br^2 0.$$

Enačbo rešimo, rešitvi sta

$$r_{1,2} = \frac{-(a-\lambda) \pm \sqrt{(a-\lambda)^2 - 4bc}}{2b}$$

Zapis se poenostavi, če uvedemo

$$\cos(\varphi) = \frac{\lambda - a}{2\sqrt{bc}}.$$

$$r_{1,2} = \sqrt{\frac{b}{c}}\left(\cos(\varphi) \pm \sqrt{\cos^2(\varphi) - 1} = \sqrt{\frac{b}{c}}(\cos(\varphi) \pm i\sin(\varphi)\right) = \sqrt{\frac{b}{c}}e^{\pm i\varphi}.$$

Tako dobimo nastavek za rešitev $x_i = \alpha e^{i\varphi} + \beta e^{-i\varphi}$. Upoštevamo še robne pogoje

$$
\begin{aligned}
x_0 = \alpha + \beta &= 0\\
x_{n+1} = \alpha e^{(n+1)i\varphi} + \beta e^{-(n+1)i\varphi}
\end{aligned}
$$

Iz prve enačbe dobimo $\beta = -\alpha$. Iz druge enačbe dobimo

$$\alpha\left(e^{(n+1)i\varphi} - e^{-(n+1)i\varphi}\right) = i \cdot 2\sin((n+1)\varphi).$$

Njena rešitev je

$$(n+1)\varphi = \pi k,\ k \in \mathbb{Z} \Rightarrow \varphi = \frac{k\pi}{n+1},\ k \in \mathbb{Z}.$$

Rešitev s $k = 0$ ni dobra, saj potem sledi $x_i = 0$ za vsak i. Tako dobimo

$$\varphi_k = \frac{k\pi}{n+1},\ k \in \{1, \ldots, n\}.$$

Za lastne vrednosti velja

$$\cos(\varphi) = \frac{\lambda - a}{2\sqrt{bc}} \Rightarrow \lambda_k = a + 2\sqrt{bc}\cos\left(\frac{k\pi}{n+1}\right).$$

Pripadajoči lastni vektor ima komponente

$$x_j^{(k)} = \sqrt{\frac{b}{c}}\left(e^{i\cdot j\cdot\varphi_k} - e^{-i\cdot j\cdot\varphi_k}\right).$$

Zanima nas le smer, vzamemo lahko kar

$$x_j^{(k)} = \sin\left(\frac{j\cdot k\cdot\pi}{n+1}\right).$$

■

Naloga 9.7 *Naj bo A simetrična matrika z lastnimi pari $Ax_i = \lambda_i x_i$, $i = 1, \ldots, n$. Naj bo x aproksimacija za lastni vektor x_1 in naj bo $\mu = \rho(A, x)$ Rayleighov kvocient za x. Potem sledi*

$$|\mu - \lambda_1| \leq 2||A||||x - x_1||_2^2.$$

Dokaži.

Rešitev. Matrika A je simetrična, torej lastni vektorji tvorijo ortonormirano bazo. BŠS lahko privzamemo, da velja $||x||_2 = 1$. Vektor x razvijemo po ortonormirani bazi,

$$x = \sum_{i=1}^{n} \alpha_i x_i, \text{ iz } ||x||_2 = 1 \text{ sledi } \sum_{i=1}^{n} \alpha_i^2 = 1.$$

Kratek račun pokaže, da je

$$\mu = x^T A x = \sum_{i=1}^{n} \lambda_i \alpha_i^2.$$

Tako velja

$$|\mu - \lambda_1| = \left| \sum_{i=1}^{n} \lambda_i \alpha_i^2 - \lambda_1 \sum_{i=1}^{n} \alpha_i^2 \right| = \left| \sum_{i=2}^{n} (\lambda_i - \lambda_1) \alpha_i^2 \right| \leq 2||A|| \sum_{i=2}^{n} \alpha_i^2.$$

Upoštevali smo, da velja

$$|\lambda_i - \lambda_1| \leq |\lambda_i| + |\lambda_1| \leq 2||A||.$$

Izračunajmo še

$$||x - x_1||_2^2 = (\alpha_1 - 1)^2 + \sum_{i=2}^{n} \alpha_i^2 \geq \sum_{i=2}^{n} \alpha_i^2,$$

tako dobimo

$$|\mu - \lambda_1| \leq 2||A||||x - x_1||_2^2.$$

∎

Poglavje 10

Numerična aproksimacija

Naloga 10.1 *Pokaži, da prostori zveznih funkcij niso strogo normirani za neskončno normo. Oglejmo si* $\mathcal{C}([0,1])$ *in funkcijo*

$$f(x) = \begin{cases} \frac{3}{2}x, & 0 \leq x < \frac{1}{3} \\ \frac{1}{2}, & \frac{1}{3} \leq x \leq 1 \end{cases}$$

in $g(x) = 1 - x$.

Rešitev. Izračunajmo še:

$$g(x) = f(1-x) \begin{cases} \frac{1}{2}, & 0 \leq x < \frac{2}{3} \\ \frac{3}{2} - \frac{3}{2}x, & \frac{2}{3} \leq x \leq 1 \end{cases}$$

Tako velja:

$$f(x) + g(x) = \begin{cases} \frac{3}{2}x + \frac{1}{2}, & 0 \leq x < \frac{1}{3} \\ 1, & \frac{1}{3} \leq x \leq \frac{2}{3} \\ 2 - \frac{3}{2}x, & 1 \leq x \leq 1 \end{cases}$$

Očitno velja

$$||f+g|| = \max_{0 \leq x \leq 1} |f(x) + g(x)| = 1 = \overbrace{\max_{0 \leq x \leq 1} |f(x)|}^{\frac{1}{2}} + \overbrace{\max_{0 \leq x \leq 1} |g(x)|}^{\frac{1}{2}} = 1$$

Funkcija f očitno ni večkratnik funkcije g. ■

Naloga 10.2 *Vektorski prostor* $X = \mathbb{R}^3$ *je opremljen z neskončno normo. Podan je vektor* $f = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}^T$ *in iščemo element najboljše aproksimacije v prostoru* $\mathcal{L}\left(\{(1,0,0),(0,1,0)\}\right)$.

(i). Ali je prostor strogo normiran.

(ii). Ali za f obstaja enoličen element najboljše aproksimacije?

Rešitev.

(i). Prostor ni strogo normiran, saj bi to zagotavljalo enoličnost elementa najboljše aproksimacije.

(ii). Za f ne obstaja enoličen element najboljše aproksimacije. Oglejmo si vektorja $g_1 = \begin{bmatrix} 1 & 3 & 0 \end{bmatrix}^T$ in $g_2 = \begin{bmatrix} 0 & 3 & 0 \end{bmatrix}^T$. Očitno velja $||f - g_1||_\infty = ||f - g_2||_\infty = 2$.

■

Naloga 10.3 *Danih je n točk v ravnini. Poišči premico, ki po metodi najmanjših kvadratov aproksimira točke. Preveri formulo za točke* $(1,2),(2,3),(3,5),(4,8)$.

Rešitev. Zapišimo ciljno funkcijo

$$f(a,b)=\sum_{i=1}^{n}(ax_i+b-y_i)^2$$

. Iščemo za katera a in b je dosežen minimum.

Izračunajmo parcialna odvoda in rešimo sistem

$$\frac{\partial f}{\partial a}=2\sum_{i=1}^{n}x_i(ax_i+b-y_i)=0$$
$$\frac{\partial f}{\partial b}=2\sum_{i=1}^{n}(ax_i+b-y_i)=0$$

Iz česar dobimo linearni sistem

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

∎

Naloga 10.4 *Dan je interval $[-1,1]$ in funkcija $f(x)=e^x$. Poišči premico, ki f najbolje aproksimira po zvezni metodi najmanjših kvadratov. Torej v normi, ki jo porodi skalarni produkt:*

$$<f,g>:=\int_{-1}^{1}f(x)g(x)dx.$$

Rešitev. Minimizirati moramo

$$H(a,b)=\int_{-1}^{1}(e^x-ax-b)^2.$$

Velja

$$\frac{\partial H}{\partial a}=\int_{-1}^{1}2(e^x-ax-b)(-x)dx\rightarrow\quad x-e^x-a\frac{x^3}{3}-b\frac{x^2}{2}\Big|_{-1}^{1}=\quad e+e^{-1}-e+e^{-1}-a\frac{2}{3}=0$$
$$\frac{\partial H}{\partial b}=\int_{-1}^{1}2(e^x-ax-b)(-1)dx\rightarrow\quad e^x-a\frac{x^2}{2}-bx\Big|_{-1}^{1}=\quad e-e^{-1}-2b=\quad 0$$

∎

Iz česar dobimo $a=\frac{3}{e}$ in $b=\frac{-1+e^2}{2e}$.

Naloga 10.5 *Podan je prostor funkcij z diskretnim skalarnim produktom*

$$<f,g>:=f(1)g(1)+2f(2)g(2)+2f(3)g(3)+2f(4)g(4)+f(5)g(5)$$

Poišči prve tri ortogonalne polinome z uporabo Gramm-Schmidtove ortogonalizacije. Preizkusi še rekurzivno tričleno rekurzijo za izračun ortogonalne baze. Aproksimiraj funkcijo $f(x)=2\cos^2(\frac{\pi x}{4})$ po metodi najmanjših kvadratov s parabolo.

Rešitev. Začnemo z bazo prostora $1, x, x^2$. Izračunamo $< 1, 1 >= 8$. Kar pomeni, da po normiranju dobimo $f_1(x) = \frac{1}{\sqrt{8}}$. Izračunamo $< \frac{1}{\sqrt{8}}, x >= 6\sqrt{2}$. Odštejemo projekcijo na prvi vektor $h_2(x) = x - 6\sqrt{2}/\sqrt{8} = x - 3$. Dobljeno moramo še normirati, tako dobimo $< h_2, h_2 >= 12$ in $f_2(x) = h_2(x)/\sqrt{12} = \frac{x-3}{\sqrt{12}}$.

Izračunamo še $< f_1, x^2 >= 21\sqrt{2}$ in $< f_2, x^2 >= 12\sqrt{3}$ in tako dobimo

$$h_3(x) = x^2 - 21\sqrt{2}f_1(x) - 12\sqrt{3}f_2(x) = x^2 - 6x + \frac{15}{2}.$$

Izračunamo še $< h_3, h_3 >= 18$ in tako dobimo $f_3(x) = \frac{1}{\sqrt{18}}(x^2 - 6x + \frac{15}{2})$.

Najboljša aproksimacija za funkcija $f(x)$ je torej ravno ortogonalna projekcija na prostor razpet z $1, x, x^2$. Ker poznamo njegovo ortogonalno bazo, je potrebno izračunati le skalarne produkte $< f, f_1 >, < f, f_2 >$ in $< f, f_3 >$. ∎

Poglavje 11

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 11.1 *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 11.2 *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 11.3 *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*

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

in

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

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 11.4 *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 11.5 *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 11.6 *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 11.7 *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 11.8 *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. ■