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** $A^TAx = A^Tb$.

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 $\overbrace{A^TA}^{B}x = A^Tz = v$. Izračunamo

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

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 $V\overbrace{(V^Tx)}^{y} = z$. 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 $V^Tx = y$, 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*

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

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

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

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

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 $Rx = Q^Tb$. 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 $i = 1, \ldots, n-1$ **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 $k = i+1, \ldots, n-1$ **do**
    $aik = a_{ik}$;
    $a_{ik} = ca_{ik} + sa_{i+1,k}$ -|- $2\sum_{i=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$;
    $a_{i+1,k} = -s \cdot aik + ca_{i+1,k}$ -|- $2\sum_{i=1}^{n-1}(n-i-1)\ *,\ \sum_{i=1}^{n-1}(n-i-1)\ +$;
  end
end
Reši zgornje trikotni sistem $Rx = b$.

Vsota je enaka

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

Matrika Q je zgornja Hessenbergova. ■

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 $x = A^+b = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i$.