Cvičení 5 - zadání a řešení úloh

Základy numerické matematiky - NMNM201   Verze z 3. listopadu 2023

---

1 Problémy nejmenších čtverců

Nechť máme danou matici $A \in \mathbb{R}^{n\times m}$ a vektor $b \in \mathbb{R}^n$. V aplikacích se často stane, že řešení soustavy $Ax = b$ neexistuje nebo není jednoznačné. V takovém případě převádíme původní problém na minimalizační úlohu

$$x = \arg\min_{z\in\mathbb{R}^m} \|b - Az\|. \tag{1}$$

Úlohu lze řešit několika postupy

- řešením soustavy normálních rovnic (má-li matice A lineárně nezávislé sloupce),
- pomocí rozšířené soustavy rovnic (má-li matice A lineárně nezávislé sloupce),
- QR rozkladem (lze i v obecném případě, ale v tom případě je potřeba zobecnit QR rozklad, který jsme zaváděli pouze pro matici s lineárně nezávislými sloupci; viz skripta sekce 3.8 a 6.4),
- s využitím Moore–Penroseovy pseudoinverze.

Na přednášce jsme viděli, že pro A s lineárně nezávislými sloupci lze řešení x úlohy (1) nalézt řešením *rozšířené soustavy*

$$\underbrace{\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix}}_{C} \begin{bmatrix} \delta \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}. \tag{2}$$

Úloha 1. *Ukažte, že sedlobodová matice C v soustavě (2) je regulární, hermitovská a indefinitní. (Připomínáme předpoklad, že A má plnou sloupcovou hodnost.)*

[Nápověda: ukázat první dvě vlastnosti je snadné, indefinitnost je obtížnější. Všimněme si, že pro $v = [0, \dots, 0, \nu_{n+1}, \dots, \nu_{n+m}]^T$ je $v^*Cv = 0$. Uvažujte pak spektrální rozklad matice C a zapište v^*Cv jako součet nenulových prvků. Aby tato suma byla nulová, některé členy musí být kladné a některé záporné.]

Řešení. Řešme

$$\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix} \begin{bmatrix} \delta \\ x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Z druhé rovnice $A^*\delta = 0$. Z první rovnice postupně

$$\begin{gathered} \delta + Ax = 0 \\ A^*\delta + A^*Ax = A^*Ax = 0. \end{gathered}$$

Matice A^*A je regulární, a proto $x = 0$ a dosazením do první rovnice $\delta = 0$, čímž jsme ukázali regularitu matice C. Ta je navíc zřejmě hermitovská.

Ukažme nyní indefinitnost matice C, tj. že má kladná i záporná vlastní čísla. Nejprve si všimněme, že pro libovolný vektor ve tvaru

$$v = [\underbrace{0, \dots, 0}_{n\times}, \nu_{n+1}, \dots, \nu_{n+m}]^T$$

platí $v^*Cv = 0$. Nechť $C = S\Lambda S^*$ je spektrální rozklad matice C. Dosazením rozkladu

$$0 = v^*Cv = v^*S\Lambda S^*v = \sum_{j=1}^{n+m} \lambda_j |s_j^* v|^2,$$

kde s_j je j-tý sloupec matice S a λ_j je příslušné vlastní číslo. Pro nenulové v musí platit $|s_j^* v|$ alespoň pro nějaký index j. Protože C je regulární (to jsme ukázali výše), $\lambda_j \neq 0$ pro každé j. Aby suma mohla být nulová, musí tedy existovat kladná i záporná λ_j. □

Další způsob řešení problému nejmenších čtverců je pomocí QR rozkladu, například pomocí Householderových reflexí, jako níže. Ze cvičení o QR rozkladu víme, že je vhodné počítat QR rozklad rozšířené matice $[A \,|\, b]$. Za předpokladu plné sloupcové hodnosti matice A dostaneme

$$H_m \cdots H_1 [A \,|\, b] = \begin{bmatrix} \hat{R} & c \\ 0 & d \end{bmatrix}$$

a řešení získáme jako

$$x = \hat{R}^{-1} c, \qquad \|b - Ax\| = \|d\|.$$

Úloha 2. *Nalezněte řešení úlohy nejmenších čtverců pro*

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

Úlohu řešte Householderovými reflexemi (bez explicitní konstrukce unitární matice Q) pro rozšířenou soustavu $[A \,|\, b]$.

Řešení.

- Pomocí Householderovy reflexe H_1 vynulujeme prvky na 2. a 3. pozici v prvním sloupci a_1 matice A:

$$H_1 = I - 2\, q_1\, q_1^T,$$

kde

$$\begin{aligned} q_1 &= \frac{a_1 + \operatorname{sgn}(a_{11})\, \|a_1\|\, e_1}{\|a_1 + \operatorname{sgn}(a_1 1)\, \|a_1\|\, e_1\|} = \frac{1}{\|a_1 + \operatorname{sgn}(a_{11})\, \|a_1\|\, e_1\|} \cdot \left(\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{24}} \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Potom

$$\begin{aligned} H_1\, a_1 &= \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix}, \\ H_1\, a_2 &= (I - 2\, q_1\, q_1^T)\, a_2 = a_2 - 2\, q_1\, q_1^T\, a_2 = \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} - \frac{2}{6} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 3 \end{bmatrix}, \\ H_1\, b &= (I - 2\, q_1\, q_1^T)\, b = b - 2\, q_1\, q_1^T\, b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{6} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}. \end{aligned}$$

Celkem tedy dostaneme

$$H_1\,[A\,|\,b] = \left[\begin{array}{cc|c} -3 & -2 & -1 \\ 0 & 4 & -1 \\ 0 & 3 & 0 \end{array}\right],$$

kde označíme $\tilde{a}_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ a $\tilde{b} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

- Druhou Householderovou reflexí budeme chtít změnit pouze prvky v druhém a třetím řádku matice $H_1\,[A\,|\,b]$. Bude mít proto tvar

$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\widetilde{H}_2} \end{bmatrix}$$

a $\widetilde{H}_2 \in \mathbb{R}^{2\times 2}$ bude matice Householderovy reflexe, která vynuluje prvek na druhé pozici v $\tilde{a}_2$. Tedy

$$\widetilde{H}_2 = I - 2\,q_2\,q_2^T,$$

kde

$$\begin{aligned} q_2 &= \frac{\tilde{a}_2 + \operatorname{sgn}(\tilde{a}_{21})\,\|\tilde{a}_2\|\,e_1}{\|\tilde{a}_2 + \operatorname{sgn}(\tilde{a}_{21})\,\|\tilde{a}_2\|\,e_1\|} = \frac{1}{\|\tilde{a}_2 + \operatorname{sgn}(\tilde{a}_{21})\,\|\tilde{a}_2\|\,e_1\|} \cdot \left(\begin{bmatrix} 4 \\ 3 \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \end{bmatrix}\right) \\ &= \frac{1}{\sqrt{90}} \begin{bmatrix} 9 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}. \end{aligned}$$

Potom

$$\begin{aligned} \widetilde{H}_2\,\tilde{a}_2 &= \begin{bmatrix} -5 \\ 0 \end{bmatrix}, \\ \widetilde{H}_2\,\tilde{b} &= (I - 2\,q_2\,q_2^T)\,\tilde{b} = \tilde{b} - 2\,q_2\,q_2^T\,\tilde{b} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} - \frac{2}{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 0 \end{bmatrix} - \frac{3}{5} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}. \end{aligned}$$

Celkem dostaneme

$$H_2\,H_1\,[A\,|\,b] = \left[\begin{array}{cc|c} -3 & -2 & -1 \\ 0 & -5 & 4/5 \\ \hline 0 & 0 & 3/5 \end{array}\right],$$

kde pro poslední prvek ve třetím sloupci platí

$$\|b - A\,x^{LS}\| = 3/5.$$

- Řešení x^{LS} pak získáme zpětnou substitucí ze systému

$$\begin{bmatrix} -3 & -2 \\ 0 & -5 \end{bmatrix} x^{LS} = \begin{bmatrix} -1 \\ 4/5 \end{bmatrix}.$$

$\square$

2 Aproximace měřených dat

Metoda nejmenších čtverců se často používá v případech, kdy chceme naměřená data (tedy neznámou funkci) reprezentovat pomocí nějaké funkce vybrané z prostoru dobře uchopitelných funkcí (například polynomy nebo goniometrické funkce). Funkci pak vybíráme tak, aby se minimalizoval výraz

$$\sum_{i=1}^{n}(y_i - g(x_i))^2, \tag{3}$$

kde y_i jsou hodnoty naměřené v bodech x_i a $g(x)$ je funkce z daného prostoru funkcí, která výraz (3) minimalizuje.

Úloha 3. *V tabulce Table 1.1 jsou uvedena data ze sčítání lidu v USA mezi lety 1790 a 1990. Růst populace budeme aproximovat pomocí modelu exponenciálního růstu funkcí*

$$x_0 e^{rt},$$

kde t je čas.

Pomocí metody nejmenších čtverců najděte parametry x_0 a r tak, aby byl minimalizován výraz (3). Do předpřipraveného skriptu `USpopulation` *doplňte část, kde se počítá řešení LS problému pomocí QR rozkladu. Porovnejte předpověď populace danou modelem pro rok 2020 se skutečným stavem (331,5 miliónu).*

[Nápověda: Je vhodné pracovat s funkcí exponenciálního růstu ve tvaru $\log(x_0 e^{rt}) = \log(x_0) + rt$. Snažíme se tedy aproximovat $\log(y)$, kde y je naměřená velikost populace, a řešení LS odpovídá parametrům r a $\log(x_0)$.]

Year	Population
1790	3,900,000
1800	5,300,000
1810	7,200,000
1820	9,600,000
1830	12,900,000
1840	17,100,000
1850	23,100,000
1860	31,400,000
1870	38,600,000
1880	50,200,000
1890	62,900,000

Year	Population
1900	76,000,000
1910	92,000,000
1920	105,700,000
1930	122,800,000
1940	131,700,000
1950	150,700,000
1960	179,000,000
1970	205,000,000
1980	226,500,000
1990	248,700,000

Table 1.1 US census data from 1790 to 1990.

Řešení.
```
Ab = [A b];
[Q,R] = qr(Ab);
x_ls = (R(1:2,1:2))\R(1:2,3);
```
□