Cvičení 5 - zadání a řešení úloh

Základy numerické matematiky - NMNM201 Verze z 30. října 2024

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 Mooreovy–Penroseovy pseudoinverze.

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

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

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

Nápověda. *Ukázat první dvě vlastnosti je snadné, indefinitnost je obtížnější. Spočtěte výraz*

$$\begin{bmatrix} \gamma \\ y \end{bmatrix}^\top C \begin{bmatrix} \gamma \\ y \end{bmatrix}.$$

Zvolte vhodně γ a y, aby výraz měl jednou kladnou a po druhé zápornou hodnotu. Argumentujte precizně, proč lze zvolit γ a y, aby výraz byl nenulový.

Řešení. Řešme

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} \gamma \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Z druhé rovnice $A^\top\gamma = 0$. Z první rovnice postupně

$$\gamma + Ay = 0$$
$$A^\top\gamma + A^\top Ay = A^\top Ay = 0.$$

Matice $A^\top A$ je regulární, a proto $y = 0$ a dosazením do první rovnice $\gamma = 0$, čímž jsme ukázali regularitu matice C. Ta je navíc zřejmě symetrická.

Ukažme nyní indefinitnost matice C. Prostým výpočtem dostáváme

$$\begin{bmatrix} \gamma \\ y \end{bmatrix}^\top C \begin{bmatrix} \gamma \\ y \end{bmatrix} = \begin{bmatrix} \gamma \\ y \end{bmatrix}^\top \begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} \gamma \\ y \end{bmatrix} = \begin{bmatrix} \gamma \\ y \end{bmatrix}^\top \begin{bmatrix} \gamma + Ay \\ A^\top\gamma \end{bmatrix} = \gamma^\top(\gamma + Ay) + y^\top A^\top\gamma = \|\gamma\|^2 + 2\gamma^\top Ay.$$

Vektory γ a y lze zvolit tak, že znaménko tohoto výrazu je libovolné. Vskutku, pro $\gamma := Ay$ resp. $\gamma := -Ay$ máme

$$\begin{bmatrix} Ay \\ y \end{bmatrix}^\top C \begin{bmatrix} Ay \\ y \end{bmatrix} = 3\|Ay\|^2, \qquad \begin{bmatrix} -Ay \\ y \end{bmatrix}^\top C \begin{bmatrix} -Ay \\ y \end{bmatrix} = -\|Ay\|^2,$$

což je kladný resp. záporný výraz pro jisté y neboť A je nenulová matice. □

Nyní prozkoumáme souvislost problému nejmenších čtverců $\min_x \|b-Ax\|$ a soustavy normálních rovnic $A^\top Ax = A^\top b$.

Úloha 2. *Definujme $g\colon \mathbb{R}^m \to \mathbb{R}$ předpisem*

$$g(x) := \frac{1}{2}\|b - Ax\|^2.$$

Ukažte, že $\nabla g(x) = A^\top(Ax - b)$.
Nápověda. *Složky gradientu počítejte jako $\frac{\partial}{\partial x_k}\frac{1}{2}\sum_{i=1}^n (b_i - \sum_{j=1}^m a_{ij}x_j)^2$.*
Nechť $x_* \in \mathbb{R}^m$ je libovolné takové, že $A^\top(Ax_* - b) = 0$. Ukažte, že

(i) platí $g(x_* + y) - g(x_*) = \dfrac{1}{2}\|Ay\|^2$ pro každé $y \in \mathbb{R}^m$;

(ii) x_* je globální minimum g;

(iii) pokud A má lineárně nezávislé sloupce, pak x_* je jediné *globální minimum g.*

Řešení. Přímým výpočtem dostáváme

$$\begin{aligned} \frac{\partial}{\partial x_k}\frac{1}{2}\sum_{i=1}^n \left(b_i - \sum_{j=1}^m a_{ij}x_j\right)^2 &= \sum_{i=1}^n \left[\frac{\partial}{\partial x_k}\frac{1}{2}\left(b_i - \sum_{j=1}^m a_{ij}x_j\right)^2\right] \\ &= \sum_{i=1}^n \left[\left(b_i - \sum_{j=1}^m a_{ij}x_j\right)\cdot\frac{\partial}{\partial x_k}\left(b_i - \sum_{j=1}^m a_{ij}x_j\right)\right] = \sum_{i=1}^n \left[\left(b_i - \sum_{j=1}^m a_{ij}x_j\right)(-a_{ik})\right] = a_k^\top(Ax - b), \end{aligned}$$

tedy k-tou složku gradientu můžeme napsat jako skalární součin k-tého sloupce matice A a vektoru $Ax - b$; tedy maticově $\nabla g(x) = A^\top(Ax - b)$.
Snadno počítáme

$$\begin{aligned} 2g(x_* + y) - 2g(x_*) &= (b - Ax_* - Ay)^\top(b - Ax_* - Ay) - (b - Ax_*)^\top(b - Ax_*) \\ &= (-Ay)^\top(b - Ax_*) + (b - Ax_*)^\top(-Ay) + (-Ay)^\top(-Ay) \\ &= 2y^\top A^\top(Ax_* - b) + \|Ay\|^2 = \|Ay\|^2, \end{aligned}$$

kde v poslední rovnosti jsme použili $A^\top(Ax_* - b) = 0$. Máme tedy $g(z) - g(x_*) = \frac{1}{2}\|A(z - x_*)\|^2 \geq 0$ pro každé $z \in \mathbb{R}^m$, z čehož vidíme, že x_* je globální minimum. Má-li navíc A plnou sloupcovou hodnost, je $A(z - x_*) \neq 0$ kdykoliv $z \neq x_*$, a tedy x_* je jediné minimum. □

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 3. *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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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 4. *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*

$$p_0 e^{rt},$$

kde t je čas a p_0 je populace v čase $t = 0$.

Pomocí metody nejmenších čtverců najděte parametry p_0 a r tak, aby byl minimalizován výraz (3). Do předpřipraveného skriptu `USpopulation.m` *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. *V kontextu (lineárních) nejmenších čtverců je třeba, abychom byli schopni vyjádřit předpověď daného modelu tak, aby lineárně závisela na hledaných parametrech.*[1] *Toho dosáhneme zlogaritmováním původního modelu, protože* $\log(p_0 e^{rt}) = \log(p_0) + rt$*. Nyní chceme* $\log(p_0) + rt \approx \log(p(t))$ *ve smyslu nejmenšich čtverců. Na pravé straně máme* $\log(p(t))$*, kde* $p(t)$ *je naměřená velikost populace v čase* t*. Řešení ve smyslu nejmenších čtverců pak odpovídá parametrům* r *a* $\log(p_0)$*.*

Year	Population	Year	Population
1790	3,900,000	1900	76,000,000
1800	5,300,000	1910	92,000,000
1810	7,200,000	1920	105,700,000
1820	9,600,000	1930	122,800,000
1830	12,900,000	1940	131,700,000
1840	17,100,000	1950	150,700,000
1850	23,100,000	1960	179,000,000
1860	31,400,000	1970	205,000,000
1870	38,600,000	1980	226,500,000
1880	50,200,000	1990	248,700,000
1890	62,900,000		

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);
```

□

3 Aproximace dat polynomem

Uvažujme neznámou funkci f, pro niž jsou nám známy její funkční hodnoty $f(x_i)$ v bodech x_i, $i = 1, \ldots, n$. Tuto funkci se budeme snažit aproximovat polynomem p nejvýše $k$tého stupně tak, aby

$$p = \operatorname*{argmin}_{\deg(q) \leq k} \sum_{i=1}^{n} (f(x_i) - q(x_i))^2. \tag{4}$$

Úloha 5. *Formulujte* (4) *jako problém nejmenších čtverců*

$$\min_x \|Ax - b\| \tag{5}$$

pro vhodnou matici A *a vektory* b *a* x*.*

Řešení. Zřejmě můžeme psát $q(x) = \sum_{j=1}^{k+1} c_j x^{j-1}$ a tedy $q(x_i) = \sum_{j=1}^{k+1} c_j x_i^{j-1}$, přičemž úkolem je určit příslušné koeficienty c_j. Definujeme-li $a_{i,j} \equiv x_i^{j-1}$, $b_i \equiv f(x_i)$ a uspořádáme-li koeficienty c_j do vektoru c, dostáváme z (4) úlohu nejmenších čtverců

$$\min_c \|Ac - b\|,$$

přičemž $A \in \mathbb{R}^{n \times k+1}$.

Ukázání ekvivalence (4) a (5) má dva kroky:

[1] Nelineární závislost na parametrech vede na *nelineární* nejmenší čtverce, což je mimo záběr této přednášky.

(i) Rozepsáním máme

$$\operatorname*{argmin}_{\deg(q)\le k} \sum_{i=1}^{n} (f(x_i) - q(x_i))^2 = \operatorname*{argmin}_{c} \|Ac - b\|^2.$$

(ii) Z nezápornosti $\|Ac - b\|$ a ryzí monotonie funkce x^2 na $[0, \infty)$ dostáváme:

$$\operatorname*{argmin}_{c} \|Ac - b\|^2 = \operatorname*{argmin}_{c} \|Ac - b\|. \qquad \square$$

Úloha 6 (volitelná)**.** *Použijte skript* `lspoly.m`*. Pro zadané hodnoty x_i a $f(x_i)$, $i = 1, \dots, n$, a $k =$ 5 vyřešte úlohu (4) v MATLABu (či Octave) a získaný polynom p vykreslete. Zvyšujte hodnotu k a pozorujte kvalitu numerického řešení. Svá pozorování sepište (a vysvětlete, pokud to svedete).*

Nápověda. *Pro vykreslení více dat (např. polynom a původní funkci) do jednoho grafu použijte* `hold`*. Pro přesnější vykreslení zadané funkce i výsledného polynomu použijte jemnější dělení intervalu, např.* `linspace(0,5,500)`*. Pro vyhodnocování polynomu zadaného koeficienty funkci* `polyval`*.*

Řešení. Skripty `lspoly_reseni.m` a `lspoly_show.m`. Všimneme si též, že A je blokem tzv. Vandermondovy matice `https://en.wikipedia.org/wiki/Vandermonde_matrix`. Pro $k \ge 13$ je matice A velmi blízko singulární matici. $\square$