Optimalizace

5. Metoda nejmenších čtverců

Tomáš Kroupa   Tomáš Werner

2022 LS

Fakulta elektrotechnická
ČVUT v Praze

Řešitelnost soustav lineárních rovnic

Chceme řešit soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$. Mohou nastat tyto možnosti:

Soustava má řešení

Podle Frobeniovy věty platí $\mathbf{b} \in \operatorname{rng} \mathbf{A}$.

1. Matice $\mathbf{A}$ má LN sloupce a tedy jediné řešení.
2. Matice $\mathbf{A}$ má LZ sloupce a tedy nekonečně mnoho řešení.

Soustava nemá řešení

Platí $\mathbf{b} \notin \operatorname{rng} \mathbf{A}$. Hledáme alespoň přibližné řešení.

Problém nejmenších čtverců

Problém nejmenších čtverců

Definice

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^{m}$. Problém nejmenších čtverců je úloha

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2.$$

Příklad (přeurčená soustava)

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix}$$

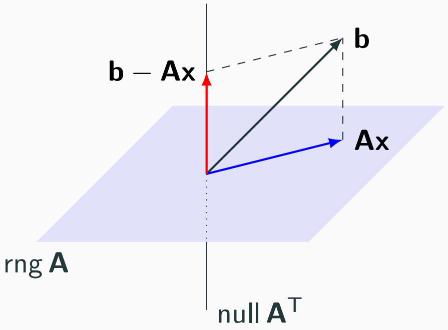
$$\min_{\mathbf{x} \in \mathbb{R}^2} (x_1 + 2x_2 - 6)^2 + (-x_1 + x_2 - 3)^2 + (x_1 + x_2 - 4)^2$$

Řešení problému nejmenších čtverců

Tvrzení

Množina všech řešení problému nejmenších čtverců je neprázdná a je rovna množině řešení soustavy normálních rovnic

$$\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{b}.$$



O matici soustavy normálních rovnic $\mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{A}^\top \mathbf{b}$

Nechť $\mathbf{A} \in \mathbb{R}^{m\times n}$. Gramova matice je matice $\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{n\times n}$.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "\mathbf{A}"] \arrow[rr, bend right, "\mathbf{A}^\top \mathbf{A}"'] & \mathbb{R}^m \arrow[r, "\mathbf{A}^\top"] & \mathbb{R}^n \end{tikzcd}$$

Tvrzení

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ platí:

$$\operatorname{null}(\mathbf{A}^\top \mathbf{A}) = \operatorname{null} \mathbf{A}$$
$$\operatorname{rng}(\mathbf{A}^\top \mathbf{A}) = \operatorname{rng} \mathbf{A}^\top$$

Navíc platí: $\mathbf{A}$ má LN sloupce $\quad\Leftrightarrow\quad$ $\mathbf{A}^\top \mathbf{A}$ regulární.

Moore-Penroseova pseudoinverze

Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ s plnou hodností definujeme pseudoinverzi

$$\mathbf{A}^+ = \begin{cases} (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T} & \operatorname{rank} \mathbf{A} = n, \\ \mathbf{A}^\mathsf{T}(\mathbf{A}\mathbf{A}^\mathsf{T})^{-1} & \operatorname{rank} \mathbf{A} = m. \end{cases}$$

- Zřejmě

$$\mathbf{A}^+ \text{ je } \quad \begin{cases} \text{levá inverze matice } \mathbf{A} & \operatorname{rank} \mathbf{A} = n, \\ \text{pravá inverze matice } \mathbf{A} & \operatorname{rank} \mathbf{A} = m. \end{cases}$$

- Je-li $\mathbf{A}$ regulární, pak $\mathbf{A}^+ = \mathbf{A}^{-1}$
- Pokud $\mathbf{A}$ nemá plnou hodnost, $\mathbf{A}^+$ se definuje pomocí SVD

Struktura řešení soustavy normálních rovnic $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$

Věta

- Má-li $\mathbf{A}$ lineárně nezávislé sloupce, existuje jediné řešení

$$\mathbf{A}^+\mathbf{b}.$$

- Má-li $\mathbf{A}$ lineárně nezávislé řádky, všechna řešení tvoří množinu

$$\mathbf{A}^+\mathbf{b} + \operatorname{null}\mathbf{A}.$$

- Je-li $\mathbf{A}$ regulární, existuje jediné řešení

$$\mathbf{A}^{-1}\mathbf{b}.$$

Řešení soustavy normálních rovnic pomocí QR rozkladu

- Chceme se vyhnout *explicitnímu* výpočtu matice $\mathbf{A}^\mathsf{T}\mathbf{A}$
- Najdeme redukovaný QR rozklad matice $\mathbf{A} = \mathbf{QR}$ s lineárně nezávislými sloupci, kde $\mathbf{Q}$ má ortonormální sloupce a $\mathbf{R}$ je regulární horní trojúhelníková:

$$\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b} \tag{1}$$

$$(\mathbf{QR})^\mathsf{T}\mathbf{QR}\mathbf{x} = (\mathbf{QR})^\mathsf{T}\mathbf{b} \tag{2}$$

$$\mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{QR}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} \tag{3}$$

$$\mathbf{R}^\mathsf{T}\mathbf{R}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} \tag{4}$$

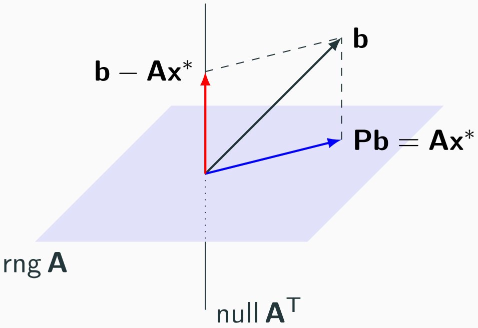
$$\mathbf{R}\mathbf{x} = \mathbf{Q}^\mathsf{T}\mathbf{b} \tag{5}$$

Ortogonální projekce na podprostor zadaný obecnou bází

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ a $\mathbf{x}^* \in \mathbb{R}^n$ je řešením problému nejmenších čtverců $\mathbf{Ax} = \mathbf{b}$.

- Ortogonální projektor $\mathbf{P}$ na $\operatorname{rng} \mathbf{A}$ splňuje $\mathbf{Pb} = \mathbf{Ax}^*$.
- Má-li $\mathbf{A}$ LN sloupce, platí $\mathbf{P} = \mathbf{AA}^+ = \mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top$.



Řešení problému nejmenších čtverců: shrnutí

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ a $\mathbf{x}^* \in \mathbb{R}^n$. Následující tvrzení jsou ekvivalentní pro problém nejmenších čtverců $\mathbf{Ax} = \mathbf{b}$.

1. Vektor $\mathbf{x}^*$ je řešením problému nejmenších čtverců.
2. $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|^2 = \|\mathbf{Ax}^* - \mathbf{b}\|^2$
3. Vektor $\mathbf{x}^*$ řeší soustavu normálních rovnic, $\mathbf{A}^\mathsf{T}\mathbf{Ax}^* = \mathbf{A}^\mathsf{T}\mathbf{b}$.
4. Platí $\mathbf{Ax}^* = \mathbf{Pb}$, kde $\mathbf{P}$ je ortogonální projektor na $\operatorname{rng} \mathbf{A}$.

Jak na to v MATLABu?

MATLAB řeší problém nejmenších čtverců $\mathbf{Ax} = \mathbf{b}$ příkazem

$$\texttt{A}\backslash\texttt{b}$$

v následujícím smyslu:

- Má-li soustava $\mathbf{Ax} = \mathbf{b}$ řešení, příkaz `A\b` najde její řešení, které má nejvýše m nulových souřadnic, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$
- Nemá-li soustava $\mathbf{Ax} = \mathbf{b}$ řešení, příkaz `A\b` najde *nějaké* řešení ve smyslu nejmenších čtverců

Vícekriteriální nejmenší čtverce

Pro $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$, $\mathbf{b}_i \in \mathbb{R}^{m_i}$ a $\mu_i \geq 0$ minimalizujeme

$$f(\mathbf{x}) = \mu_1 \|\mathbf{A}_1\mathbf{x} - \mathbf{b}_1\|^2 + \cdots + \mu_k \|\mathbf{A}_k\mathbf{x} - \mathbf{b}_k\|^2, \qquad \mathbf{x} \in \mathbb{R}^n .$$

Převod na problém nejmenších čtverců

Platí $f(\mathbf{x}) = \|\mathbf{A}'\mathbf{x} - \mathbf{b}'\|^2$, kde $\mathbf{A}' = \begin{bmatrix} \sqrt{\mu_1}\mathbf{A}_1 \\ \vdots \\ \sqrt{\mu_k}\mathbf{A}_k \end{bmatrix}, \mathbf{b}' = \begin{bmatrix} \sqrt{\mu_1}\mathbf{b}_1 \\ \vdots \\ \sqrt{\mu_k}\mathbf{b}_k \end{bmatrix}.$

Má-li $\mathbf{A}'$ LN sloupce, existuje jediné optimální řešení

$$\mathbf{x}^* = (\mu_1\mathbf{A}_1^\top\mathbf{A}_1 + \cdots + \mu_k\mathbf{A}_k^\top\mathbf{A}_k)^{-1}(\mu_1\mathbf{A}_1^\top\mathbf{b}_1 + \cdots + \mu_k\mathbf{A}_k^\top\mathbf{b}_k).$$

Tichonovova regularizace problému nejmenších čtverců

Pro zadanou váhu $\mu > 0$ minimalizujeme

$$f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \mu\|\mathbf{x}\|^2, \qquad \mathbf{x} \in \mathbb{R}^n .$$

Převod na problém nejmenších čtverců

Platí $f(\mathbf{x}) = \|\mathbf{A}'\mathbf{x} - \mathbf{b}'\|^2$, kde $\mathbf{A}' = \begin{bmatrix} \mathbf{A} \\ \sqrt{\mu}\mathbf{I} \end{bmatrix}, \mathbf{b}' = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$.

Matice $\mathbf{A}'$ má LN sloupce a proto má optimální řešení tvar

$$\mathbf{x}^* = (\mathbf{A}^\mathsf{T}\mathbf{A} + \mu\mathbf{I})^{-1}\mathbf{A}^\mathsf{T}\mathbf{b}$$

Jiné odvození řešení problému nejmenších čtverců

- Účelová funkce

$$f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$$

  je konvexní kvadratická a tudíž spojitě diferencovatelná
- Minimum tak lze nalézt pomocí podmínek optimality známých z matematické analýzy
- Existuje jediné globální optimum $\mathbf{x}^*$ a to splňuje podmínku

$$\nabla f(\mathbf{x}^*) = \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x}^* - \mathbf{A}^\mathsf{T}\mathbf{b} = \mathbf{0}$$

Řešení soustavy s nejmenší normou

Vybíráme řešení nedourčené soustavy

Definice

Předpokládejme, že má soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ alespoň jedno řešení.

Řešení s nejmenší normou je řešení úlohy

$$\min \|\mathbf{x}\|^2 \qquad \text{za podmínky } \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n$$

Příklad

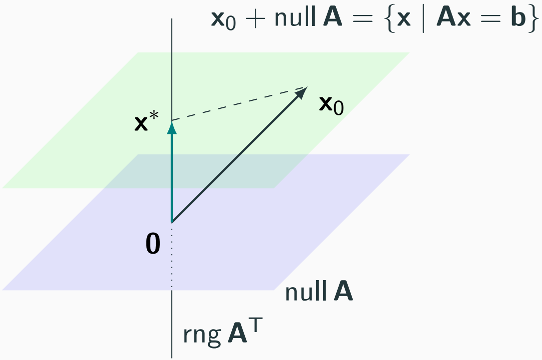
$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\min\ x_1^2 + x_2^2 + x_3^2 \qquad \text{za podmínky } \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in \mathbb{R}^3$$

Řešení s nejmenší normou nedourčené soustavy $\mathbf{Ax} = \mathbf{b}$

Tvrzení

Existuje právě jedno řešení $\mathbf{x}^*$ s nejmenší normou. Pokud má matice $\mathbf{A}$ LN řádky, platí $\mathbf{x}^* = \mathbf{A}^\mathsf{T}(\mathbf{A}\mathbf{A}^\mathsf{T})^{-1}\mathbf{b} = \mathbf{A}^+\mathbf{b}$.



Řešení s nejmenší normou soustavy normálních rovnic

Sjednocující formulace

$$\min \|\mathbf{x}\|^2 \qquad \text{za podmínky } \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b},\ \ \mathbf{x} \in \mathbb{R}^n$$

Úloha má jediné optimální řešení $\mathbf{x}^*$.

1. Pokud soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ *nemá řešení*, vektor $\mathbf{x}^*$ je řešením problému nejmenších čtverců s nejmenší normou
2. Má-li soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ *jediné řešení*, je to vektor $\mathbf{x}^*$
3. Má-li soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ *nekonečně mnoho řešení*, vektor $\mathbf{x}^*$ je jejím řešením s nejmenší normou.