

Úlohy nejmenších čtverců

Problémy nejmenších čtverců $Ax \approx b$

- Řešení $Ax = b$ nemusí existovat, a pokud existuje, nemusí být jednoznačné.
- Často má smysl hledat x tak, že

$$Ax \approx b.$$

V jakém smyslu má být Ax blízké vektoru b ?

- **Příklad:** A reprezentuje **matematický model**, b jsou **naměřené hodnoty**. **Cíl:** najít původní data x .
- Nepřesnosti v datech $(A, b) \rightarrow$ chceme zohlednit při volbě metody pro nalezení vhodného vektoru x .
- Interpretace $\approx$ není jednoznačná, závisí na typu problému.

Matematická formulace

- Dáno: $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$.

$n > m \rightarrow$ **přeurčená** úloha

$n < m \rightarrow$ **nedourčená** úloha.

Pro zjednodušení výkladu $\rightarrow n \geq m$.

- **Motivace 1:** Předpokládejme, že **chybami je zatížen pouze b** .

Cíl: nalézt **co nejmenší** změnu f pravé strany b tak, že x je řešením soustavy

$$Ax = b + f,$$

velikost změny f měříme euklidovskou normou vektoru.

- Jinak řečeno: hledáme x , které **minimalizuje normu rezidua**

$$f = b - Ax.$$

Problém nejmenších čtverců (LS)

Definice: Problém nejmenších čtverců

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. **Problémem nejmenších čtverců** (LS) budeme nazývat úlohu určení $x \in \mathbb{C}^m$, které minimalizuje

$$\|b - Ax\|.$$

Ekvivalentně

$$\min_f \|f\| \quad \text{za podmínky} \quad Ax = b + f.$$

- Zkratka LS pochází z anglického **least squares** (minimalizuje se odmocnina ze součtu kvadrátů prvků rezidua).
- Ve statistice bývá problém nejmenších čtverců nazýván úlohou **lineární regrese**.

Úplný problém nejmenších čtverců (TLS)

- Často nutné uvažovat chyby v b (pozorování) i v A (chyba modelu). Hledáme co nejmenší **změnu** E matice A a co nejmenší **změnu** f vektoru b tak, že x je řešením soustavy

$$(A + E)x = b + f.$$

- Velikost změny měříme **Frobeniovou normou matice** $[f, E]$.

Definice: Úplný problém nejmenších čtverců

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. **Úplným problémem nejmenších čtverců** nazveme úlohu určení $x \in \mathbb{C}^m$ tak, aby platilo

$$\min_{f, E} \|[f, E]\|_F \quad \text{za podmínky} \quad (A + E)x = b + f.$$

Total least squares, ortogonální regrese (statistika).

Jde o unitárně invariantní problémy

- Euklidovská norma vektoru i Frobeniova norma matice jsou **unitárně invariantní**.
- Přenásobíme-li například

$$Ax = b + f$$

unitární P zleva pak

$$\|Pf\| = \|f\| ,$$

tj. **velikost** hledaných změn pravé strany potřebných k dosažení rovností **nezmění**, řešení se nezmění.

- Dále jen problém nejmenších čtverců (TLS viz skriptu).

Existence řešení problému nejmenších čtverců

- $Ax = b$ má řešení $\Leftrightarrow b \in \mathcal{R}(A) = \text{span}\{a_1, \dots, a_m\}$.
- $b \notin \mathcal{R}(A) \Rightarrow$ soustava **není řešitelná** v klasickém smyslu.
- Hledáme x , které minimalizuje euklidovskou normu rezidua $\|b - Ax\| \Leftrightarrow$ hledáme nejlepší aproximaci b v $\mathcal{R}(A)$.
- Tato aproximace je dána **ortogonální projekcí b na $\mathcal{R}(A)$** $\rightarrow$ označení $b|_{\mathcal{R}(A)}$.

Věta: Existence řešení problému nejmenších čtverců

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. Vektor x je řešením problému nejmenších čtverců právě tehdy, když

$$Ax = b|_{\mathcal{R}(A)}, \quad \|b - Ax\| = \|b|_{\mathcal{N}(A^*)}\|.$$



Jednoznačnost

- Sloupce A **lineárně nezávislé** $\Rightarrow \exists!$ x tak, že $Ax = b|_{\mathcal{R}(A)}$.
- Sloupce A **lineárně závislé** $\rightarrow b|_{\mathcal{R}(A)}$ lze vyjádřit pomocí sloupců A nekonečně mnoha způsoby:
 x je řešení $\Rightarrow x + z$, $z \in \mathcal{N}(A)$, je řešení.
- Zahrnutí $z \in \mathcal{N}(A)$ do řešení problému nejmenších čtverců nepřispívá ke snížení normy rezidua $\Rightarrow$ je vhodné uvažovat **pouze řešení, která složku z $\mathcal{N}(A)$ neobsahují**.
- Následující věta $\rightarrow$ takové řešení je pouze jedno a je ze všech možných řešení LS problému minimálním v normě.

Řešení LS minimální v normě

Věta: Řešení LS minimální v normě

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. Potom **existuje právě jedno** řešení x problému nejmenších čtverců **minimální v normě** a je dáno vztahy

$$Ax = b|_{\mathcal{R}(A)} \quad \text{a} \quad x \in \mathcal{R}(A^*).$$



Řešení problému nejmenších čtverců

a soustava normálních rovnic

Věta

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. x je řešením problému nejmenších čtverců právě tehdy, je-li řešením **soustavy normálních rovnic**,

$$A^*Ax = A^*b.$$



Soustava normálních rovnic

Poznámky

- Má-li matice A **plnou sloupcovou hodnost**, je matice A^*A **regulární** a soustava normálních rovnic má jednoznačné řešení,

$$x = (A^*A)^{-1}A^*b,$$

neboli $x = A^\dagger b$.

- Jak uvidíme dále, má vyjádření řešení pomocí **pseudoinverze** obecnou platnost.

Přehled metod pro řešení LS

A s plnou sloupcovou hodnotí

- Předpoklad: $A \in \mathbb{C}^{n \times m}$, $n \geq m$,

$$\text{rank}(A) = m .$$

Problém nejmenších čtverců má pak **jednoznačné řešení** (nemusíme tudíž hledat řešení minimální v normě).

- Tři základní metody:
 - Řešení LS pomocí QR rozkladu.
 - Řešení LS pomocí soustavy normálních rovnic.
 - Řešení LS pomocí rozšířené soustavy rovnic.

Řešení problému nejmenších čtverců

pomocí QR rozkladu

- LS je unitárně invariantní, je-li U unitární

$$\|b - Ax\| = \|U(b - Ax)\|.$$

- Využijeme QR rozklad matice A ,

$$A = QR, \quad Q \in \mathbb{C}^{n \times n}, \quad R \in \mathbb{C}^{n \times m},$$

$$\min_x \|b - Ax\| = \min_x \|b - QRx\| = \min_x \|Q^*b - Rx\|.$$

- A má plnou sloupcovou hodnotu $\Rightarrow$

$$R = \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix},$$

kde $\hat{R} \in \mathbb{C}^{m \times m}$ je regulární horní trojúhelníková.

Řešení problému nejmenších čtverců

pomocí QR rozkladu

- Označíme-li

$$Q = [q_1, \dots, q_m | q_{m+1}, \dots, q_n] \equiv [Q_m | V_m],$$

- Hledáme řešení **ekvivalentního problému**,

$$Q^*b - Rx = \begin{bmatrix} Q_m^*b \\ V_m^*b \end{bmatrix} - \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} x = \begin{bmatrix} Q_m^*b - \hat{R}x \\ V_m^*b \end{bmatrix}.$$

- Volbou x lze ovlivnit **pouze prvních m prvků** vektoru rezidua, a hledané x je řešením soustavy

$$\hat{R}x = Q_m^*b,$$

přičemž

$$\min_x \|b - Ax\| = \|V_m^*b\|.$$

Praktická realizace

- Aplikujeme QR rozklad na rozšířenou matici $[A, b]$.
- Po aplikaci m **Householderových reflexí** dostáváme

$$H_m \dots H_1 [A, b] = \begin{bmatrix} \hat{R}_H & c_H \\ 0 & d_H \end{bmatrix}.$$

Potom $x_H = \hat{R}_H^{-1} c_H$ a $\|b - Ax_H\| = \|d_H\|$.

- Aplikací **MGS** na matici $[A, b]$ získáme obdobně

$$[A, b] = [Q_m, \tilde{q}] \begin{bmatrix} \hat{R}_G & c_G \\ 0 & d_G \end{bmatrix}.$$

Řešení opět získáme jako $x_G = \hat{R}_G^{-1} c_G$ a $\|b - Ax_G\| = d_G$.

- Pokud při užití Householderových reflexí neukládáme bázi, jsou oba postupy přibližně **stejně drahé**.
- Björck a Paige, překvapující výsledek: **zpětná chyba** výpočtu matic $\hat{R}_H$ a $\hat{R}_G$ je pro oba postupy **srovnatelná**.

Řešení problému nejmenších čtverců

pomocí soustavy normálních rovnic

- A má plnou sloupcovou hodnost $\rightarrow A^*A$ je regulární, hermitovská, pozitivně definitní $\Rightarrow$ řešení x je jednoznačné $\rightarrow$ **Choleského rozklad** či **metoda sdružených gradientů**.
- $\text{rank}(A) = m \rightarrow$ řešení soustavy normálních rovnic je ekvivalentní s řešením soustavy vzniklé z QR rozkladu

$$\begin{aligned} A^*Ax &= A^*b && \iff R^*Rx = R^*Q^*b \\ &&& \iff \hat{R}^*\hat{R}x = \hat{R}^*Q_m^*b \\ &&& \iff \hat{R}x = Q_m^*b. \end{aligned}$$

$\hat{R}^*\hat{R}$ je Choleského rozklad matice A^*A .

- Výpočet pomocí soustavy normálních rovnic může být **výhodný** např. tehdy, když je $n \gg m$ a m je malé.
- Pokud A velká $\rightarrow$ **iterační metody**, násobení A^*Av realizujeme postupně jako $A^*(Av)$.

Řešení problému nejmenších čtverců

pomocí soustavy normálních rovnic

- Nevýhoda $\rightarrow$ nárůst čísla podmíněnosti,

$$\kappa(A) = \frac{\sigma_1(A)}{\sigma_m(A)},$$

pak pro matici soustavy normálních rovnic platí

$$\kappa(A^*A) = \frac{\sigma_1(A^*A)}{\sigma_m(A^*A)} = \left(\frac{\sigma_1(A)}{\sigma_m(A)} \right)^2 = \kappa^2(A).$$

- Možné **numerické problémy**, výpočet pomocí soustavy normálních rovnic bezpečný v případě

$$\kappa^2(A) \ll \frac{1}{\varepsilon}.$$

Řešení problému nejmenších čtverců

pomocí rozšířené soustavy rovnic

- $\text{rank}(A) = m \rightarrow x$ je jednoznačné řešení $Ax = b|_{\mathcal{R}(A)}$ a platí

$$\delta \equiv b - Ax = b - b|_{\mathcal{R}(A)} \in \mathcal{N}(A^*) \quad \Rightarrow \quad A^* \delta = 0.$$

- $\delta = b - Ax$ lze zapsat maticově

$$\begin{bmatrix} I & A \end{bmatrix} \begin{bmatrix} \delta \\ x \end{bmatrix} = b.$$

- Dohromady

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

máme **rozšířenou soustavu rovnic**, získali jsme soustavu se čtvercovou hermitovskou maticí dimenze $(m + n) \times (m + n)$, tzv. **sedlobodová** matice.

Řešení problému nejmenších čtverců

pomocí rozšířené soustavy rovnic

Lemma

Uvažujme tzv. *sedlobodovou matici*

$$C = \begin{bmatrix} B & A \\ A^* & 0 \end{bmatrix},$$

kde $B \in \mathbb{C}^{n \times n}$ je hermitovská pozitivně definitní a $A \in \mathbb{C}^{n \times m}$, $n \geq m$, má plnou sloupcovou hodnotu. Potom je C je **regulární, hermitovská a indefinitní**.



Řešení problému nejmenších čtverců

pomocí rozšířené soustavy rovnic - shrnutí

Věta

Nechť $A \in \mathbb{C}^{n \times m}$, $n \geq m$, $b \in \mathbb{C}^n$ a nechť $A = m$. Řešení x problému nejmenších čtverců lze nalézt řešením rozšířené soustavy

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

se sedlobodovou maticí $\rightarrow$ je regulární, hermitovská a indefinitní.

- Soustavy s indefinitními maticemi mohou být velmi **obtížně řešitelné**. I přesto může být tento postup výpočtu v některých případech úspěšný.

Řešení problému nejmenších čtverců

obecný případ

Singulární rozklad A umožní nalézt **řešení minimální v normě** pro obecný problém nejmenších čtverců.

Věta

Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$, $A = U_r \Sigma_r V_r^*$ je ekonomický singulární rozklad matice A . Pak

$$x = V_r \Sigma_r^{-1} U_r^* b = A^\dagger b$$

představuje řešení problému nejmenších čtverců minimální v normě.



Cvičení

6.2 Necht má matice A plnou sloupcovou hodnost m , $A = QR$ je její QR rozklad a $\hat{R}$ je regulární horní trojúhelníkový blok matice R . Dokažte

$$\sigma_1(\hat{R}) = \sigma_1(A) \quad \text{a} \quad \sigma_m(\hat{R}) = \sigma_m(A).$$

6.3 Ukažte, že sedlobodová matice

$$C = \begin{bmatrix} B & A \\ A^* & 0 \end{bmatrix},$$

kde $B \in \mathbb{C}^{n \times n}$ je hermitovská pozitivně definitní a $A \in \mathbb{C}^{n \times m}$, $n \geq m$, je regulární.