Lineární algebra pro pokročilé

Martin Černý, Peter Zeman

kam.mff.cuni.cz/∼cerny
cerny@kam.mff.cuni.cz

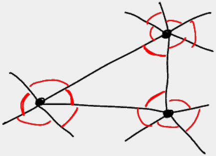
March 8, 2021

Metoda nejmenších čtverců

- $A\mathbf{x} = \mathbf{b}$
- soustava má řešení $\iff \mathbf{b} \in Im(A)$
- soustava nemá řešení $\iff \mathbf{b} \notin Im(A)$
- **Cíl:** Najít řešení, které neexistuje...

Proč řešit něco, co nejde řešit?

Trochu historie: Zeměměřičství a Carl Friedrich Gauss

- **Vstup:** vzdálenosti mezi body a úhly
  - s chybami odpovídajícími měření před 200 lety
- **Výstup:** poloha bodů
- 
- více rovnic než neznámých (*přeurčený systém* rovnic)
- obvykle řešení neexistuje

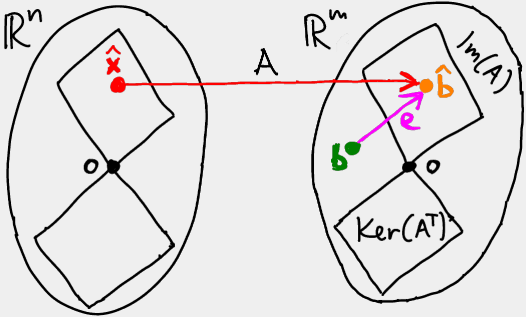
Proč řešit něco, co nejde řešit 2?

- Příklad: přeurčený systém o 1 neznámé:
  - $2x = b_1$
  - $3x = b_2$
  - $4x = b_3$
- Řešení existuje $\iff$ pravé strany v poměru $2 : 3 : 4$
  - pokud data pochází z měření $\implies$ nebude zachováno

Co s tím?

1. Vypustit některé rovnice
   - Které?
   - Předpoklad, že chyba jenom v některých
   - $\implies$ neodpovídá chybám měření
2. Změnit ***co nejméně*** pravou stranu $\mathbf{b}$, aby řešení existovalo
   - $\implies$ ***Metoda nejmenších čtverců***

- $A\mathbf{x} = \mathbf{b}$ ... nemá řešení
- $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ ... chceme vyřešit
  - $\hat{\mathbf{b}}$ je nejbližší k $\mathbf{b}$
  - $\implies$ kolmá projekce $\mathbf{b}$ do $Im(A)$



-

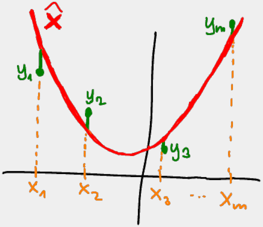
Lineární regrese = Metoda nejmenších čtverců

Lineární regrese: $A\mathbf{x} = \mathbf{b}$

- A ... určeno modelem
- $\mathbf{b}$ ... data
- Příklad:
  - (x_i, y_i) ... data
  - α, β, γ ... body leží na parabole $\alpha x_i^2 + \beta x_i + \gamma = y_i$
  - $\left(\begin{array}{ccc|c} x_1^2 & x_1 & 1 & y_1 \\ x_2^2 & x_2 & 1 & y_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_m & 1 & y_m \end{array}\right)$
  - typicky nebude existovat přesné řešení

Lineární regrese: $A\mathbf{x} = \mathbf{b}$

- A ... určeno modelem
- $\mathbf{b}$ ... data
- Příklad:
  - (x_i, y_i) ... data
  - α, β, γ ... body leží na parabole $\alpha x_i^2 + \beta x_i + \gamma = y_i$
  - $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$



  -

TOTÁLNÍ METODA NEJMENŠÍCH ČTVERCŮ $\hat{A}\hat{\mathbf{x}} = \hat{b}$

- $A\mathbf{x} = \mathbf{b} \rightarrow \hat{A}\hat{\mathbf{x}} = \hat{\mathbf{b}}$
- minimalizujeme $\hat{A} - A$ a $\mathbf{b} - \hat{\mathbf{b}}$
- Použití: když nedůvěřujeme modelu
  - matematicky složité
  - nebudeme dělat

Jak vyřešit nejmenší čtverce?

- $\mathbf{e} = \hat{\mathbf{b}} - \mathbf{b}$
- $\mathbf{e} \perp Im(A)$
  - $e \in Ker(A^T)$
- $\mathbf{o} = A^T\mathbf{e} = A^T(\hat{\mathbf{b}} - \mathbf{b}) = A^T(A\hat{\mathbf{x}} - \mathbf{b})$
- $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ ... **normální rovnice**
- pokud A^TA regulární:
  1. $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$
  2. $\hat{\mathbf{b}} = A(A^TA)^{-1}A^T\mathbf{b}$

  - $P = A(A^TA)^{-1}A^T$ ... matice kolmé projekce na podprostor $Im(A)$
  - snadno lze ověřit, že $P^T = P$ a $P^2 = P$

Jádro A^TA a A

$Ker(A^TA) = Ker(A)$

Důkaz:

1. $Ker(A^TA) \supseteq Ker(A)$
   - $A\mathbf{x} = \mathbf{o} \implies A^TA\mathbf{x} = o$
2. $Ker(A^TA) \subseteq Ker(A)$
   - $A^TA\mathbf{x} = \mathbf{o}$
   - vynásobíme zleva $\mathbf{x}^T$:
     - $\mathbf{o} = \mathbf{x}^TA^TA\mathbf{x} = (A\mathbf{x})^TA\mathbf{x} = \|A\mathbf{x}\|^2$
     - $\|A\mathbf{x}\|^2 = o \iff A\mathbf{x} = o$

Kdy nastane, že $A^T A$ je regulární?

Jádro $A^T A$ a A

$Ker(A^T A) = Ker(A)$

Důkaz:

1. $Ker(A^T A) \supseteq Ker(A)$
   - $A\mathbf{x} = \mathbf{o} \implies A^T A\mathbf{x} = o$
2. $Ker(A^T A) \subseteq Ker(A)$
   - $A^T A\mathbf{x} = \mathbf{o}$
   - vynásobíme zleva $\mathbf{x}^T$:
     - $\mathbf{o} = \mathbf{x}^T A^T A\mathbf{x} = (A\mathbf{x})^T A\mathbf{x} = \|A\mathbf{x}\|^2$
     - $\|A\mathbf{x}\|^2 = o \iff A\mathbf{x} = o$

Důsledek pro hodnosti matic

$rank(A^T A) = rank(A)$

- A má lineárně nezávislé sloupce $\implies A^T A$ regulární $\implies (A^T A)^{-1}$

Metoda nejmenších čtverců přes QR rozklad

- $A\mathbf{x} = \mathbf{b} \implies QR\mathbf{x} = \mathbf{b}$
- $Q^T QR\hat{\mathbf{x}} = Q^T\mathbf{b}$
  - není ekvivalentní úprava
  - Q není typicky čtvercová
  - $A \in \mathbb{R}^{m\times n}, Q \in \mathbb{R}^{m\times n}, Q^T Q = I_n$
- $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$
- Tento postup numericky přesnější
  - IDEA:
  - ***číslo podmíněnosti*** $\sim$ numerické vlastnosti
  - čím větší, tím horší vlastnosti
  - $A^T A$ má drunou mocninu čísla podmíněnosti matice A