

Matematika pro chemické inženýry

Drahoslava Janovská

4. Lineární algebra - pokračování.

Singulární hodnoty matic, singulární rozklad matic,
řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců, normální rovnice.

Lineární regrese

Lineární regrese, polynomiální regrese, obecný model lineární regrese.



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Povinná látka. Bude v písemkách a bude se zkoušet při ústní zkoušce
(žádné označení)

- ★ Řešené příklady k procvičení - dobrovolné
- ★ Pro studenty, kteří chtějí vědět víc. Tato látka se nebude přednášet, nebude v písemkách, nebude se zkoušet.



Obsah

1 Singulární rozklad matic

- Singulární hodnoty matic
- Singulární rozklad matic
- Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců
- Řešení normálních rovnic s využitím singulárního rozkladu
- Singulární rozvoj matic

2 Lineární regrese

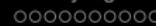
- Vyhodnocování experimentálních dat
- Náhodná veličina, distribuční funkce
- Střední hodnota a rozptyl náhodné veličiny
- Kovariance náhodných veličin

3 Základy regresní analýzy

- Základní model lineární regrese
- Ekvivalentní model
- Metoda nejmenších čtverců

4 Polynomiální regrese

5 Literatura k dalšímu studiu



Singulární hodnoty matice

Singulární hodnoty matice

$\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \neq n$... libovolná obdélníková matice

Pro obdélníkovou matici není pojem vlastního čísla definován, ale ...

$$\mathbf{A}^T \mathbf{A} \in \mathbb{R}^{n \times n} \implies \mathbf{A}^T \mathbf{A} \text{ je čtvercová}$$

$$(\mathbf{A}^T \mathbf{A})^T = \mathbf{A}^T \mathbf{A} \implies \mathbf{A}^T \mathbf{A} \text{ je symetrická}$$

$$\mathbf{x}^T (\mathbf{A}^T \mathbf{A}) \mathbf{x} = (\mathbf{Ax})^T \mathbf{Ax} = \|\mathbf{Ax}\|^2 \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \implies \mathbf{A}^T \mathbf{A} \text{ je pozitivně semidefinitní}$$

Vlastní čísla $\lambda_1, \lambda_2, \dots, \lambda_n$ matice $\mathbf{A}^T \mathbf{A}$ jsou reálná, nezáporná. Zapišme je ve tvaru $\lambda_k = \sigma_k^2$, $\sigma_k \geq 0$, $k = 1, \dots, n$. Čísla

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ se nazývají singulární hodnoty matice \mathbf{A} .

Pro největší a nejmenší singulární hodnotu matice \mathbf{A} platí:

$$\sigma_1 = \max_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}, \quad \sigma_n = \min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}.$$

Singulární hodnoty matice

Lemma: Je-li $A_{m \times n}$ reálná matici, pak vlastní čísla $n \times n$ matice $A^T A$ jsou reálná, nezáporná.

Důkaz: Necht λ je vlastní číslo $\mathbf{A}^T \mathbf{A}$, $\mathbf{v} \neq \mathbf{0}$ odpovídající vlastní vektor.

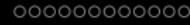
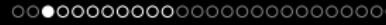
$$\begin{aligned} (\mathbf{A}^T \mathbf{A})\mathbf{v} &= \lambda \mathbf{v} \quad | \cdot \mathbf{v}^T zleva \\ \mathbf{v}^T \mathbf{A}^T \mathbf{A} \mathbf{v} &= \lambda \mathbf{v}^T \mathbf{v} \\ (\mathbf{A} \mathbf{v})^T \mathbf{A} \mathbf{v} &= \lambda \|\mathbf{v}\|^2 \quad \Rightarrow \quad \lambda = \frac{\|\mathbf{A} \mathbf{v}\|^2}{\|\mathbf{v}\|^2} \geq 0 \end{aligned}$$

Příklad Vypočtěte singulární hodnoty matice A.

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \Rightarrow \mathbf{A}^T \mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}.$$

$$\det(\mathbf{A}^T \mathbf{A} - \lambda \mathbf{E}) = 0 \Leftrightarrow \lambda^2 - 10\lambda + 16 = 0 \Leftrightarrow \lambda_1 = 8, \lambda_2 = 2.$$

Singulární hodnoty: $\sigma_1 = \sqrt{8} = 2\sqrt{2}$, $\sigma_2 = \sqrt{2}$.



Singulární rozklad matice

Věta Bud $\mathbf{A} \in \mathbb{R}^{m \times n}$ libovolná matice. Pak

- existují ortogonální matice \mathbf{U} typu $m \times m$ a ortogonální matice \mathbf{V} typu $n \times n$ takové, že matice $\mathbf{S} = \mathbf{U}^T \mathbf{A} \mathbf{V}$ typu $m \times n$ má "diagonální" tvar

$$\mathbf{S} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$\mathbf{D} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0,$$

kde $\sigma_1, \sigma_2, \dots, \sigma_r$ jsou nenulové singulární hodnoty matice \mathbf{A} a r je hodnost matice \mathbf{A} ;

- nenulové singulární hodnoty matice \mathbf{A}^T jsou rovněž čísla $\sigma_1, \sigma_2, \dots, \sigma_r$.

Rozklad $\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$... singulární rozklad matice \mathbf{A} .

Poznámka $\mathbf{S} = \mathbf{U}^T \mathbf{A} \mathbf{V}$

- sloupce matice \mathbf{U} ... m ortonormálních vlastních vektorů symetrické matice $\mathbf{A} \mathbf{A}^T$ typu $m \times m$,
- sloupce matice \mathbf{V} ... n ortonormálních vlastních vektorů symetrické matice $\mathbf{A}^T \mathbf{A}$ typu $n \times n$.

Pokračování příkladu

Vypočteme ortonormální vlastní vektory $\mathbf{v}_1, \mathbf{v}_2$ matice $\mathbf{A}^T \mathbf{A}$:

$$\lambda_1 = 8 \Rightarrow \begin{bmatrix} 5-8 & 3 \\ 3 & 5-8 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{h}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \|\mathbf{h}_1\| = \sqrt{2}.$$

Tedy $\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Obdobně vypočteme $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Zbývá vypočítat ortonormální vlastní vektory matice $\mathbf{A}\mathbf{A}^T$. K jejich výpočtu využijeme vztah

$$\mathbf{S} = \mathbf{U}^T \mathbf{A} \mathbf{V} \Rightarrow \mathbf{U} \mathbf{S} = \mathbf{A} \mathbf{V},$$

který jsme získali formálně vynásobením první rovnice maticí **U** zleva. Tuto rovnost rozepíšeme do složek:

$$\sigma_i u_i = \mathbf{A} v_i \Rightarrow u_i = \frac{\mathbf{A} v_i}{\sigma_i}, i = 1, 2 \Rightarrow$$

$$\mathbf{u}_1 = 2\sqrt{2} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = 2\sqrt{2} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix}.$$

Singulární rozklad matice

★ Vlastní čísla symetrické matice jsou reálná

Lemma Vlastní čísla symetrické matice jsou reálná a o odpovídajících vlastních vektorech můžeme také předpokládat, že jsou reálné.

Důkaz A je symetrická, λ její vlastní číslo, u odpovídající vlastní vektor. Předpokládáme, že jak vlastní čísla, tak vlastní vektory mohou být imaginární. Pak

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u} \quad \text{a také} \quad \mathbf{u}^*\mathbf{A} = \bar{\lambda}\mathbf{u}^*.$$

druhá rovnice je Hermitovsky sdružená s první. Druhou rovnici vynásobíme vektorem \mathbf{u} zprava a upravíme:

$$\mathbf{u}^*(\mathbf{A}\mathbf{u}) = \overline{\lambda} \mathbf{u}^*\mathbf{u} \quad \Rightarrow \quad \lambda \mathbf{u}^*\mathbf{u} = \overline{\lambda} \mathbf{u}^*\mathbf{u}$$

Tedy $(\lambda - \bar{\lambda})\mathbf{u}^*\mathbf{u} = 0$, $\mathbf{u} \neq 0 \Rightarrow \lambda = \bar{\lambda}$, a tedy λ je reálné.

Nyní sečteme komplexně sdružené rovnice $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ a $\bar{\mathbf{A}}\bar{\mathbf{u}} = \bar{\lambda}\bar{\mathbf{u}}$. Dostaneme

$$\mathbf{A}(u + \bar{u}) = \lambda(u + \bar{u}) \quad \Rightarrow \quad u + \bar{u} \quad \text{je vlastní vektor } \mathbf{A}$$

Kdyby $\mathbf{u} = x + iy$, pak $\bar{\mathbf{u}} = x - iy$ a $\mathbf{u} + \bar{\mathbf{u}} = 2x$, což je reálné číslo.



Příklad Vypočtěte singulární rozklad matice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Řešení Chceme vypočítat rozklad $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, kde \mathbf{U} je 2×2 ortogonální matici. \mathbf{S} má na "diagonále" singulární hodnoty. \mathbf{V} je 3×3 ortogonální matici.

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{Vlastní čísla matice } \mathbf{A}^T \mathbf{A} : \quad \lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$$

Singulární hodnoty matice \mathbf{A} : $\sigma_1 = \sqrt{3}, \sigma_2 = 1, \sigma_3 = 0$
 (řadíme podle velikosti)

Odpovídající vlastní vektory:

$$\lambda_1 = 3 : \quad (\mathbf{A}^T \mathbf{A} - 3\mathbf{E}) = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\tilde{\mathbf{v}}_1 = (1, 2, 1)^T, \quad \|\tilde{\mathbf{v}}_1\| = \sqrt{6} \quad \Rightarrow \quad \mathbf{v}_1 = \frac{1}{\sqrt{6}}(1, 2, 1)^T.$$



$$\lambda_2 = 1 : \quad (\mathbf{A}^T \mathbf{A} - \mathbf{E}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\tilde{\mathbf{v}}_2 = (1, 0, -1)^T, \quad \|\tilde{\mathbf{v}}_2\| = \sqrt{2} \quad \Rightarrow \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}}(1, 0, -1)^T.$$

$$\lambda_3 = 0 : \quad (\mathbf{A}^T \mathbf{A}) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{v}}_3 = (1, -1, 1)^T, \quad \|\tilde{\mathbf{v}}_3\| = \sqrt{3} \quad \Rightarrow \quad \mathbf{v}_3 = \frac{1}{\sqrt{3}}(1, -1, 1)^T.$$

Dostaneme

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



Rovnici $\mathbf{A} = \mathbf{USV}^T$ vynásobíme zleva maticí \mathbf{V} a dostaneme

$$\mathbf{A}\mathbf{V} \equiv \mathbf{U}\mathbf{S} \quad \Rightarrow \quad \mathbf{A}\mathbf{v}_i \equiv \sigma_i \mathbf{u}_i,$$

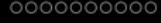
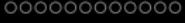
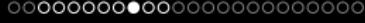
Tedy sloupce matice **U** počítáme podle vzorce

$$\mathbf{u}_i = \frac{1}{\sigma_i} \mathbf{Av}_i, \quad \text{t.j.} \quad \mathbf{u}_1 = \frac{1}{\sigma_1} \mathbf{Av}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{u}_2 = \frac{1}{\sigma_2} \mathbf{Av}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Dostaneme

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Ověřte si, že rovnost $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ skutečně platí.



Praktický výpočet singulárního rozkladu

Necht $\mathbf{A} \in \mathbb{R}^{m \times n}$ je libovolná obdélníková matice.

Je nějaký vztah mezi spektrálními rozklady matic $\mathbf{A}^T\mathbf{A}$ a $\mathbf{A}\mathbf{A}^T$?

Vypočteme spektrální rozklad matice $\mathbf{A}^T\mathbf{A}$, tedy vypočteme vlastní čísla $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ matice $\mathbf{A}^T\mathbf{A}$ a odpovídající ortonormální vlastní vektory v_1, v_2, \dots, v_n .

Pak ortonormální vlastní vektory u_1, u_2, \dots, u_m odpovídající příslušným nenulovým vlastním číslům matice $\mathbf{A}\mathbf{A}^T$ vypočteme podle jednoduchého vzorce

$$u_j = \frac{\mathbf{A}v_j}{\sigma_j}, \quad j = 1, \dots, m.$$

Není třeba počítat spektrální rozklad matice $\mathbf{A}\mathbf{A}^T$.

Singulární rozklad matice

★ Numerický výpočet

$$\mathbf{A} \in \mathbb{R}^{m \times n}, \quad m \geq n \text{ (w.l.o.g.)}$$

Golubova-Reinschova (Golubova-Kahanova) metoda: dva kroky

- **bidiagonální** matici \mathbf{A} pomocí Householderových matic zrcadlení:

$$\mathbf{A} \longrightarrow \mathbf{J}^{(0)} = \begin{pmatrix} \mathbf{J}_0 \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{J}_0 = \begin{pmatrix} x & x & & & 0 \\ & x & x & & \\ & & \ddots & \ddots & \\ & & & \ddots & x \\ 0 & & & & x \end{pmatrix}.$$

Po n redukčních krocích dostaneme (horní) dvoudiagonální matici $\mathbf{J}^{(0)}$ typu $m \times n$

$$\mathbf{J}^{(0)} = \mathbf{P}_n \mathbf{P}_{n-1} \dots \mathbf{P}_1 \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2 \dots \mathbf{Q}_{n-2},$$

$\mathbf{P}_k, \mathbf{Q}_k$ jsou Householderovy matice zrcadlení.



$\mathbf{Q} := \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{n-2}$, $\mathbf{P} := \mathbf{P}_1 \mathbf{P}_2 \cdots \mathbf{P}_n$, \mathbf{P} a \mathbf{Q} ... ortogonální,

$$\mathbf{J}^{(0)} = \mathbf{P}^T \mathbf{A} \mathbf{Q}, \quad (\mathbf{J}^{(0)})^T \mathbf{J}^{(0)} = \mathbf{J}_0^T \mathbf{J}_0 = \mathbf{Q}^T \mathbf{A}^T \mathbf{A} \mathbf{Q}.$$

Matice \mathbf{J}_0 a \mathbf{A} jsou podobné a mají tedy tytéž singulární hodnoty. Zbývá provést singulární rozklad dvoudiagonální matice \mathbf{J}_0 .

- singulární rozklad dvoudiagonální matice \mathbf{J}_0 , tj. iterační diagonalizace pomocí jisté varianty QR metody s posuny spektra s využitím vhodných Givensových matic rovinné rotace

$$\mathbf{J}_0 \longrightarrow \mathbf{J}_1 \longrightarrow \dots \longrightarrow \mathbf{D}, \quad \text{kde } \mathbf{D} \text{ je diagonální}, \quad \mathbf{J}_{k+1} = \mathbf{S}_k^T \mathbf{J}_k \mathbf{T}_k,$$

\mathbf{S}_k a \mathbf{T}_k jsou ortogonální matice. Matice \mathbf{T}_k vybíráme tak, že posloupnost třidiagonálních matic $\mathbf{M}_k = \mathbf{J}_k^T \mathbf{J}_k$ konverguje k diagonální matici, kdežto matice \mathbf{S}_k vybíráme tak, aby všechny matice \mathbf{J}_k byly ve dvoudiagonálním tvaru.

Metoda je rychlá a numericky stabilní. Podrobnostmi se nebudeme zabývat.

Přeuročená soustava rovnic

$$\mathbf{A} \in \mathbb{R}^{m \times n}, \quad m \geq n, \quad \mathbf{b} \in \mathbb{R}^m, \quad ? \mathbf{x} \in \mathbb{R}^n : \quad \mathbf{Ax} = \mathbf{b}.$$

Problém má více rovnic než neznámých, **soustava je přeuročená**.

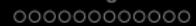
Aby takováto soustava byla řešitelná, musí **b** ležet ve sloupcovém prostoru $\mathcal{R}(\mathbf{A})$ matice **A**:

$$\mathbf{b} \in \mathcal{R}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m, \exists \mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{y}\} \subset \mathbb{R}^m,$$

t.j. **b** musí být lineární kombinací sloupců matice **A**. Složky hledaného vektoru $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ představují koeficienty této lineární kombinace:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Obvykle $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$, a tedy místo přesného řešení hledáme takové **řešení**, které je "nejblíže" k přesnému řešení.



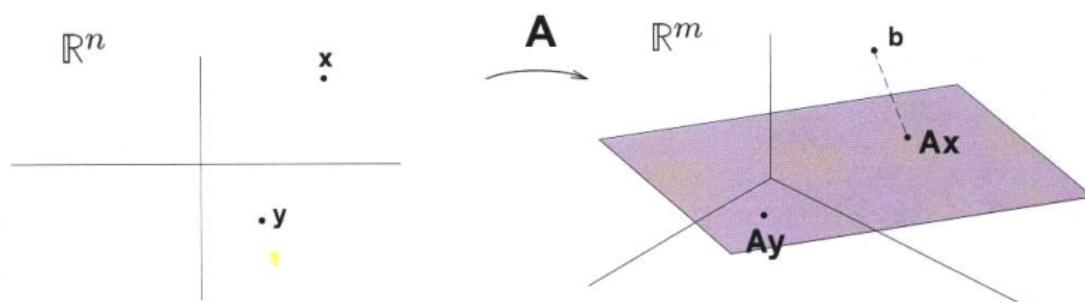
Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců

Formulace problému

Necht $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \geq n$, $\mathbf{b} \in \mathbb{R}^m$.

Hledáme řešení soustavy $\mathbf{Ax} = \mathbf{b}$ ve smyslu nejmenších čtverců, t.j. hledáme

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{x} \in \arg \min_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{Ay} - \mathbf{b}\|.$$



Řešení soustavy ve smyslu nejmenších čtverců

Normální rovnice

$\|\mathbf{x}\| = (\mathbf{x}^T \mathbf{x})^{\frac{1}{2}}$... Euklidovská norma vektoru \mathbf{x}

$E = \|\mathbf{Ax} - \mathbf{b}\|$... chyba výpočtu, které se dopustíme.

Hledáme takový bod $\bar{\mathbf{x}} \in \mathbb{R}^n$, pro který je chyba E minimální.

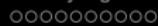
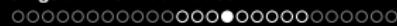
Geometricky: chyba E je vzdálenost bodů \mathbf{Ax} a \mathbf{b} . Tato vzdálenost je nejmenší, jestliže vektor $\mathbf{A}\bar{\mathbf{x}}$ je ortogonální projekcí vektoru \mathbf{b} na prostor $\mathcal{R}(\mathbf{A})$

\Rightarrow vektor chyby $\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}$ je kolmý k prostoru $\mathcal{R}(\mathbf{A})$, tj. pro každé $\mathbf{x} \in \mathbb{R}^n$ musí být $\mathbf{Ax} \in \mathcal{R}(\mathbf{A})$ kolmé k vektoru $\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}$:

$$(\mathbf{Ax})^T (\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}) = 0 \Leftrightarrow \mathbf{x}^T (\mathbf{A}^T \mathbf{A}\bar{\mathbf{x}} - \mathbf{A}^T \mathbf{b}) = 0 \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Poslední rovnice může být splněna tehdy a jen tehdy, jestliže $\bar{\mathbf{x}}$ řeší tzv. normální rovnice

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



Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců

Věta Vektor $\mathbf{x} \in \mathbb{R}^n$ je řešením soustavy $\mathbf{Ax} = \mathbf{b}$ ve smyslu nejmenších čtverců právě když \mathbf{x} řeší normální rovnice.

Navíc problém nejmenších čtverců má jediné řešení právě když hodnost $h(\mathbf{A})$ matice \mathbf{A} je maximální, tj. $h(\mathbf{A}) = n$.

Poznámka Je-li $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \geq n$, pak

$$h(\mathbf{A}) = n \iff \det(\mathbf{A}^T \mathbf{A}) \neq 0.$$

Soustava $\mathbf{Ax} = \mathbf{b}$ je tedy jednoznačně řešitelná ve smyslu nejmenších čtverců právě když matice $\mathbf{A}^T \mathbf{A}$ je regulární.

Pak z normálních rovnic dostáváme

$$\bar{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

a má-li matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ ortonormální sloupce, t.j. $\mathbf{A}^T \mathbf{A} = \mathbf{E}_n$, kde \mathbf{E}_n je jednotková matice řádu n , pak

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

Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců

★ Řešení normálních rovnic - teorie

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}, \quad \mathbf{c} := \mathbf{A}^T \mathbf{b}, \quad \mathbf{A}^T \mathbf{A} \text{ regulární} \implies \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{c}$$

Spektrální analýza matice $\mathbf{A}^T \mathbf{A} \in \mathbb{R}^{n \times n}$:

vlastní čísla $\lambda_i > 0$, vlastní vektory $\mathbf{v}_i \in \mathbb{R}^n$, $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$... báze $\mathbb{R}^n \implies$

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{v}_i, \quad \mathbf{c} = \sum_{i=1}^n \gamma_i \mathbf{v}_i$$

$$\left. \begin{aligned} \mathbf{A}^T \mathbf{A} \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i \right) &= \sum_{i=1}^n \gamma_i \mathbf{v}_i \\ \sum_{i=1}^n \alpha_i \lambda_i \mathbf{v}_i &= \sum_{i=1}^n \gamma_i \mathbf{v}_i \end{aligned} \right\} \implies \alpha_i = \frac{1}{\lambda_i} \gamma_i, \quad i = 1, \dots, n.$$

Co se skrývá za poslední rovností?

Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců



Necht vlastní čísla $\lambda_i > 0$, $i = 1, \dots, n$, matice $\mathbf{A}^T \mathbf{A}$ jsou taková, že λ_n je řádově několikrát menší než ostatní vlastní čísla:

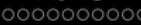
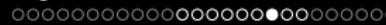
$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \gg \lambda_n > 0.$$

Matice $\mathbf{A}^T \mathbf{A}$ zobrazí jednotkovou sféru v \mathbb{R}^n na elipsoid s osami ve směrech **vlastních vektorů** \mathbf{v}_i . Délka poloosy ve směru \mathbf{v}_n je mnohem menší, než ostatní délky poloos, což znamená, že při zobrazení $\mathbf{A}^T \mathbf{A}$ je n -tá složka libovolného vektoru jednotkové délky **zanedbatelná** a elipsoid tedy leží v podstatě v \mathbb{R}^{n-1} .

Při řešení normálních rovnic nás zajímá **inverzní zobrazení** $(\mathbf{A}^T \mathbf{A})^{-1}$. To má stejné vlastní vektory, ale vlastní čísla mají převrácenou hodnotu:

$$(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{v}_i = \frac{1}{\lambda_i} \mathbf{v}_i.$$

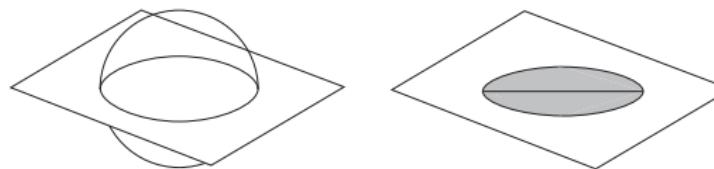
Protože $\frac{1}{\lambda_n}$ je mnohem větší než ostatní $\frac{1}{\lambda_i}$, je **příslušný elipsoid** v podstatě **jednodimenzionální**.



Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců

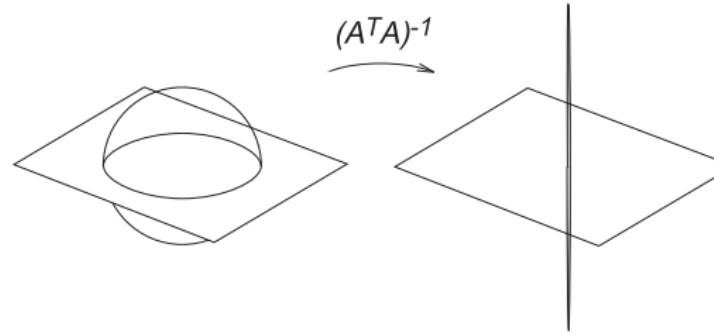


$$A^T A$$



Zobrazení jednotkové sféry maticí $A^T A$; $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \gg \lambda_n > 0$

$$(A^T A)^{-1}$$



Zobrazení jednotkové sféry maticí $(A^T A)^{-1}$; $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \gg \lambda_n > 0$

Numerické řešení normálních rovnic

Jaká je v tomto případě chyba výpočtu $E = \|\mathbf{A}^T \mathbf{A} \mathbf{x} - \mathbf{A}^T \mathbf{b}\|$?

V konečné počítačové aritmetice může být vektor $\mathbf{A}^T \mathbf{A} \mathbf{x}$ libovolně nepřesný, protože jednou ztracené cifry nelze získat zpět, a tedy všechny složky kromě poslední jsou zničené nebo úplně ztracené.

Numerické řešení $\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$:

- ★ pomocí Choleského rozkladu symetrické, pozitivně definitní matice $\mathbf{A}^T \mathbf{A}$.
Nevýhoda: museli bychom nejprve explicitně vyčíslit matici $\mathbf{A}^T \mathbf{A}$, tj. vyčíslit mnoho skalárních součinů, čímž již matice $\mathbf{A}^T \mathbf{A}$ může být zatížena značnou chybou. Vyplatí se tedy použít nějakou metodu, která nevyžaduje vytvoření matice $\mathbf{A}^T \mathbf{A}$, ale pracuje přímo s maticí \mathbf{A} .
- iterační metody.

Ztrátu platných cifer numerického výpočtu charakterizuje tzv. **číslo podmíněnosti matice**, které je v tomto případě rovno

$$\kappa(\mathbf{A}^T \mathbf{A}) = \frac{\lambda_1}{\lambda_n}.$$

Řešení soustavy lineárních rovnic ve smyslu nejmenších čtverců

Podmíněnost matice soustavy lineárních algebraických rovnic

Příklad

$$Ax = b$$

$$\begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \Rightarrow \mathbf{x} = (1, 1, 1, 1)^T$$

$$\widehat{Ax} = \widehat{b}$$

$$\begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 32, 1 \\ 22, 9 \\ 33, 1 \\ 30, 9 \end{pmatrix} \Rightarrow \widehat{\mathbf{x}} = (9, 2; -12, 6; 4, 5; -1, 1)^T$$

$$\widetilde{Ax} = b$$

$$\begin{pmatrix} 10 & 7 & 8, 1 & 7, 2 \\ 7, 08 & 5, 04 & 6 & 5 \\ 8 & 5, 98 & 9, 98 & 9 \\ 6, 99 & 4, 99 & 9 & 9, 98 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \Rightarrow \widetilde{\mathbf{x}} = (-5, 79; 12, 02; -1.57, 2.57)^T$$

Relativní chyba:

$$\varepsilon_{rel}(\widehat{\mathbf{b}}) = \frac{\|\widehat{\mathbf{b}} - \mathbf{b}\|}{\|\mathbf{b}\|} = 0, 003, \quad \varepsilon_{rel}(\widehat{\mathbf{x}}) = 8, 2$$

$$\varepsilon_{rel}(\widetilde{\mathbf{A}}) = \frac{\|\widetilde{\mathbf{A}} - \mathbf{A}\|}{\|\mathbf{A}\|} = 0, 009, \quad \varepsilon_{rel}(\widetilde{\mathbf{x}}) = 6, 64$$

Matice \mathbf{A} je symetrická, $\det(\mathbf{A}) = 1$, ale $\kappa(\mathbf{A}) = 4488$

Řešení normálních rovnic s využitím singulárního rozkladu

★ Řešení normálních rovnic s využitím singulárního rozkladu

Necht $\mathbf{A} = \mathbf{USV}^T$, $\mathbf{x} = \arg \min_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{Ay} - \mathbf{b}\| \implies \mathbf{SV}^T \mathbf{x} = \mathbf{U}^T \mathbf{b}$, kde

$\mathbf{S} = \text{diag}(\sigma_1, \dots, \sigma_p)$, $p = \min(m, n)$. Nechť hodnost matice $h(\mathbf{A}) = r < p$. Pak $\sigma_{r+1} = \dots = \sigma_p = 0$, matice \mathbf{S} je singulární a inverzní matice neexistuje.

Vynásobíme-li však druhou rovnici maticí \mathbf{S}^+ zleva,

$$\mathbf{S}^+ = \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{D}^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_r} \end{pmatrix},$$

dostaneme pro \mathbf{x} soustavu $\mathbf{V}^T \mathbf{x} = \mathbf{S}^+ \mathbf{U}^T \mathbf{b}$ s ortogonální maticí soustavy \mathbf{V}^T .

Matice \mathbf{S}^+ je tzv. **Mooreova-Penroseova pseudoinverzní matice** k matici \mathbf{S} .

Pseudoinverzemi v obecném případě se zde nebudeme zabývat.

Srovnání:

$$\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_r}, \quad \kappa(\mathbf{A}^T \mathbf{A}) = \frac{\lambda_1}{\lambda_r} = \left(\frac{\sigma_1}{\sigma_r} \right)^2 = (\kappa(\mathbf{A}))^2 \implies$$

přímým řešením normálních rovnic ztratíme **dvakrát více** platných cifer než při aplikaci SVD.

Singulární rozvoj matice

Singulární rozvoj matice

Rovnici $\mathbf{A} = \mathbf{USV}^T$ lze přepsat ve tvaru

$$\mathbf{A} = \left[\begin{array}{cccc} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \\ | & | & & | \end{array} \right] \left[\begin{array}{ccccc} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & 0 \end{array} \right] \left[\begin{array}{cccc} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{array} \right]^T =$$

$$= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

tzv. singulární rozvoj matice \mathbf{A} . Tedy

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad h(\mathbf{A}) = r, \quad h(\mathbf{u}_i \mathbf{v}_i^T) = 1 \quad \Rightarrow$$

$$\mathbf{x} \in \mathbb{R}^n \implies \mathbf{Ax} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \mathbf{x} = \sum_{i=1}^r (\mathbf{v}_i^T \mathbf{x} \sigma_i) \mathbf{u}_i \dots$$

lineární kombinace vektorů \mathbf{u}_i , $i = 1, \dots, r$.

★ Komprese dat

Aplikace : Komprese dat

$\mathbf{A} \in \mathbb{R}^{m \times n}$... obsahuje naměřená data. Hledáme approximaci této matice maticí \mathbf{B} takovou, že $h(\mathbf{B}) = k < \min(m, n)$ a matice \mathbf{B} zachytí nejdůležitější informace obsažené v datech, tj. v matici \mathbf{A} . Kdybychom například chtěli, aby hodnota matice \mathbf{B} byla rovna jedné, položíme $\mathbf{B} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$.

Pozor! Různá volba k ovlivní kvalitu získaných výsledků.

★ Komprese dat

Příklad Chceme **digitalizovat fotografii** tak, že obraz nahradíme maticí 24×24 pixelů. Prvky matice jsou bud 0 (černá buňka) nebo 1 (bílá buňka). S přesností na 4 desetinná místa získáme 16 nenulových singulárních hodnot, všechny ostatní jsou s touto přesností 0:

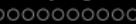
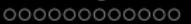
9,5403	6,6288	5,6369	3,4756	2,7385	2,2023	1,5835	1,5566
1,4207	1,2006	0,9905	0,9258	0,7479	0,6744	0,6122	0,4698

Hledáme k tak, aby relativní chyba obrazu nebyla větší než 10, relativní chyba

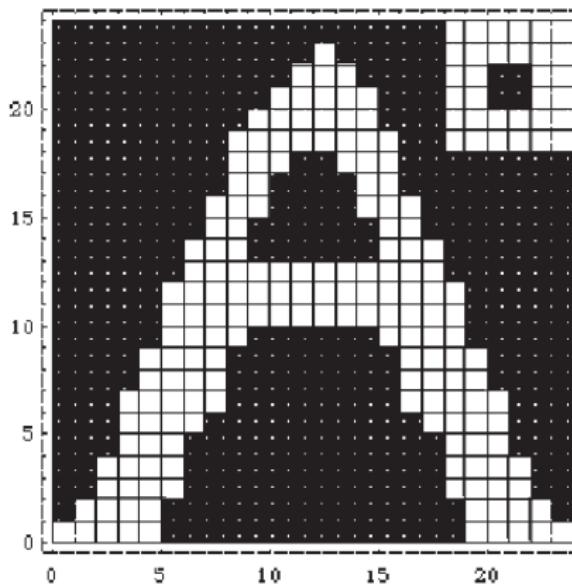
$$e(k) = 1 - \sqrt{\frac{\sum_{i=1}^k \sigma_i^2}{\sum_{i=1}^{16} \sigma_i^2}}$$

Konkrétně: $e(2) = 0,18$, $e(3) = 0,09 \implies$ tři členy singulárního rozvoje matice \mathbf{A} , \implies

$$\mathbf{B} = \sum_{i=1}^3 \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad h(\mathbf{B}) = 3$$

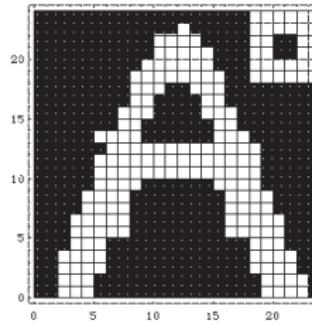
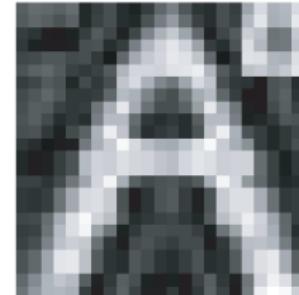
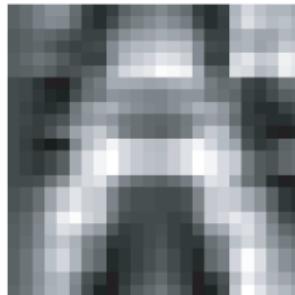


Singulární rozvoj matice



Originální fotografie

Singulární rozvoj matice



$k = 3; k = 5; k = 5$, ale prvky \mathbf{B} jsou zaokrouhleny na 0 nebo 1

Lineární regrese

Vyhodnocování experimentálních dat

- Při řešení chemicko-inženýrských problémů jsme obvykle schopni odvodit model procesu nebo děje probíhajícího v zařízení, ale...
 - Často nejsme schopni určit numerické hodnoty parametrů, které v modelu vystupují

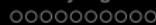
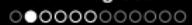
Z matematického hlediska mohou být modely různé povahy: lineární nebo nelineární algebraické rovnice, obyčejné diferenciální rovnice nebo parciální diferenciální rovnice.

Zajímá nás závislost modelu na množině bezrozměrných parametrů.

Předpokládáme závislost ve tvaru

$$y = f(\mathbf{x}, \mathbf{a}), \quad (1)$$

kde $\mathbf{x} = (x_1, x_2, \dots, x_n)$ je vektor nezávisle proměnných, $\mathbf{a} = (a_1, \dots, a_P)$ je vektor parametrů, jejichž hodnoty je třeba určit, a y je závisle proměnná (odezva). Předpokládáme, že nezávisle proměnné x_i se měří s prakticky zanedbatelnou chybou ve srovnání s chybou měření závisle proměnné y .



Náhodná veličina, distribuční funkce

★ Náhodná veličina

Náhodná veličina X je reálná funkce, definovaná na množině všech elementárních jevů Ω s hodnotami v \mathbb{R} . Je to tedy funkce, která přiřazuje každému náhodnému pokusu $\omega \in \Omega$ reálné číslo $X(\omega)$.

Náhodné veličiny dělíme na **diskrétní** – jejich obor hodnot je konečná nebo spočetná množina, a na **spojité** – jejich obor hodnot je nespočetná množina.

Pravděpodobnost, se kterou náhodná proměnná nabývá určité hodnoty nebo je obsažena v určitém intervalu hodnot, se nazývá **rozdělení pravděpodobnosti**. Základní možnost, jak popsat pravděpodobnostní rozdělení náhodné veličiny X , je určit její distribuční funkci.

Distribuční funkcí náhodné veličiny X nazveme reálnou funkci $F(x)$, $x \in \mathbb{R}$, definovanou vztahem

$$F(x) = P(X \leq x).$$

Tedy hodnota distribuční funkce v bodě x je pravděpodobnost toho, že náhodná veličina X bude mít hodnotu menší nebo rovnou tomuto x .

★ Příklad

Příklad: Hraní rulety

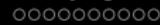
Sledujme dva hody kuličkou a pozorujme, kolikrát přitom padne lichá.

Diskrétní náhodná veličina X má za hodnoty počet kol, v nichž padlo liché číslo. Prostor elementárních jevů tohoto náhodného pokusu (označíme S sudou a L lichou) má čtyři prvky:

$$\omega_1 = SS, \quad \omega_2 = SL, \quad \omega_3 = LS, \quad \omega_4 = LL$$

a náhodná veličina má tyto hodnoty:

$$X(\omega_1) = 0, \quad X(\omega_2) = X(\omega_3) = 1, \quad X(\omega_4) = 2.$$



★ Vlastnosti distribuční funkce

Vlastnosti distribuční funkce:

- $0 \leq F(x) \leq 1$,
- je neklesající,
- je zprava spojitá,
- má konečně nebo nejvýše spočetně mnoho bodů nespojitosti
- $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$.

★ Pravděpodobnostní funkce

Diskrétní náhodná veličina X má konečně nebo nejvýše spočetně mnoho hodnot $\{x_1, x_2, \dots, x_n, \dots\}$. Pravděpodobnosti, že tyto hodnoty náhodná veličina nabude, jsou kladné, tj. $P(X = x_i) > 0$ a platí pro ně

$$\sum_{x_i} P(X = x_i) = 1.$$

Ríkáme také, že náhodná veličina X ma rozdelení diskrétního typu.

Funkce $p(x) = P(X = x)$ se nazývá **pravděpodobnostní funkce** diskrétní náhodné veličiny X . Je definovaná pouze na oboru hodnot $\{x_1, x_2, \dots, x_n, \dots\}$ náhodné veličiny X a platí

$$p(x_i) > 0, \quad \sum_{x_i} p(x_i) = 1.$$

Pravděpodobnostní funkce umožňuje určit distribuční funkci F náhodné veličiny X :

$$F(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p(x_i), \quad -\infty < x < \infty.$$



Tedy distribuční funkce diskrétní náhodné veličiny je

- nespojitá v bodech x_1, x_2, \dots
 - v každém x_i má skok velikosti $P(X = x_i)$.
 - v intervalech $(x_i; x_{i+1})$ je vždy konstantní.

Spojité náhodné veličiny X má **rozdělení pravděpodobnosti spojitého typu**, existuje-li nezáporná reálná funkce $f(x)$ taková, že distribuční funkci $F(x)$ lze vyjádřit ve tvaru

$$F(x) = \int_{-\infty}^{\infty} f(t)dt, \quad -\infty < x < \infty.$$

Funkce $f(x)$ je definovaná pro všechna reálná x a nazývá se **hustota pravděpodobnosti** náhodné veličiny X . Lze ukázat, že takto definovaná distribuční funkce F spojité náhodné veličiny je spojitá pro všechna x a ve všech bodech, v nichž má derivaci, je $f(x) = F'(x)$.



Střední hodnota a rozptyl náhodné veličiny

★ Číselné charakteristiky náhodných veličin

Střední hodnota náhodné veličiny X je charakteristikou její polohy (myšleno ve smyslu polohy hodnot veličiny X na číselné ose), je to jistá **průměrná hodnota**, kolem níž náhodná veličina náhodně kolísá.

Je-li X náhodná veličina s diskrétním rozdělením pravděpodobnosti, která nabývá hodnoty x_1, x_2, \dots a má pravděpodobnostní funkci $p(x)$, pak její **střední hodnota** je číslo

$$E(X) = \sum_{x_i} x_i \cdot p(x_i).$$

Je-li X náhodná veličina se spojitým rozdělením pravděpodobnosti a s hustotu $f(x)$, je její **střední hodnota** číslo

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx.$$



Rozptyl se také značí σ^2 . Odmocnině z rozptylu se říká **směrodatná odchylka**. Tedy

$$\sigma = \sqrt{D(X)}.$$

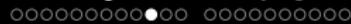
Je-li σ malé, nabývá náhodná veličina s velkou pravděpodobností hodnot, velmi blízkých své střední hodnotě $E(X)$.

Rozptyl se často počítá podle vzorce

$$D(X) = E(X^2) - (E(X))^2, \quad \text{tedy}$$

pro veličinu s diskrétním rozdělením
$$D(X) = \sum_{x_i} x_i^2 p(x_i) - (E(X))^2,$$

pro veličinu se spojitým rozdělením
$$D(X) = \int_{-\infty}^{x_i} x^2 f(x) dx - (E(X))^2.$$

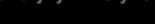


★ Dvojice náhodných veličin

Uvažujme **dvojici náhodných veličin**, někdy říkáme, že tvoří **náhodný vektor (X, Y)** nebo také **dvourozměrnou náhodnou veličinu**. **Distribuční funkcí** náhodné veličiny (X, Y) nazveme reálnou funkci $F(x,y)$ definovanou v \mathbb{R}^2 vztahem

$$F(x,y) = P(X \leq x, Y \leq y).$$

Střední hodnotou dvourozměrné náhodné veličiny (X, Y) budeme nazývat dvojici $(E(X), E(Y))$.



Kovariance náhodných veličin

★ Kovariance náhodných veličin

Kromě číselných charakteristik jednotlivých náhodných veličin X a Y jsou důležité číselné charakteristiky, které vyjadřují jejich vzájemnou souvislost:
Kovariance náhodných veličin X a Y je číslo

$$\text{cov}(X, Y) = E([X - E(X)] \cdot [Y - E(Y)]) \quad \Rightarrow$$

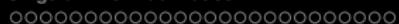
$$\text{cov}(X, X) = E([X - E(X)]^2) = D(X).$$

K výpočtu kovariance slouží nejčastěji vzorec

$$\text{cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y), \quad \text{kde} \quad E(X \cdot Y) = \sum_{x_i} \sum_{y_j} x_i y_j p(x_i, y_j).$$

Kovariance je mírou lineární závislosti veličin X a Y . Pro zhodnocení takové závislosti je ale většinou vhodnější tzv. **korelační koeficient** $\varrho(X, Y)$,

$$\varrho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}}.$$



Kovariance náhodných veličin

★ Kovarianční matice

Pro korelační koeficient vždy platí $-1 \leq \varrho \leq 1$. Například závisí-li Y na X lineárně, tj. $Y = aX + b$, platí:

- ① je-li grafem této závislosti rostoucí přímka, čili $a > 0$, je $\varrho = 1$,
- ② je-li grafem této závislosti klesající přímka, čili $a < 0$, je $\varrho = -1$

Je-li kovariance X a Y rovna nule, je také jejich korelační koeficient roven nule a takové veličiny nazýváme **nekorelované**. Pro takové veličiny pak platí

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

Matice

$$\begin{bmatrix} \text{cov}(X, X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & \text{cov}(Y, Y) \end{bmatrix} = \begin{bmatrix} D(X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & D(Y) \end{bmatrix}$$

se nazývá **kovarianční matice** dvojice náhodných veličin X a Y .

Kovarianční matice je analogie k jednorozměrnému rozptylu náhodné veličiny X .

Regresní funkce

Funkci $\eta(x) = E(Y(x))$ definovanou na definičním oboru $A \subset \mathbb{R}$ proměnné x nazveme **regresní funkcí**. Regresí rozumíme závislost střední hodnoty náhodné veličiny $Y(x)$ na veličině x .

Předpokládáme, že známe tvar regresní funkce, a na základě náhodného výběru odhadujeme její neznámé parametry:

Vybereme n hodnot x_j , $j = 1, \dots, n$, $x_j \in A$, nezávisle proměnné. Pro každé x_j napozorujeme (naměříme) realizaci (hodnotu) y_j náhodné veličiny Y_j :

$$x_j, j = 1, \dots, n, \in A \quad \longrightarrow \quad y_j = Y(x_j).$$

Získané dvojice hodnot $(x_1, y_1), \dots, (x_n, y_n)$ nám poslouží k odhadu neznámých parametrů regresní funkce.

Speciální případ: **Jednoduchá lineární regresní funkce** má tvar

$$\eta(x) = \beta_1 + \beta_2 x,$$

kde β_1, β_2 jsou parametry této funkce, jejichž hodnoty hledáme. Grafem jednoduché lineární regresní funkce je přímka se směrnicí β_2 , říkáme, že jede o **přímkovou regresi**.

Základní model lineární regrese

Základní model lineární regrese

Model lineární regrese by měl splňovat:

1. Regresní funkce $\eta(x)$ je lineární funkcí tvaru

$$\eta(x) = \sum_{k=1}^p \beta_k f_k(x),$$

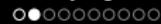
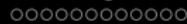
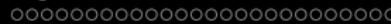
kde $f_k(x)$ jsou známé funkce a β_k , $k = 1, \dots, p$, neznámé parametry. Funkce η je lineární vzhledem k parametrům.

- 2.** Hodnotě x_i je přiřazena náhodná veličina Y_i , pro kterou platí

$$E(Y_j) = \eta(x_j), \quad D(Y_j) = \sigma^2, \quad j = 1, \dots, n,$$

Druhá rovnice znamená, že rozptyl náhodné veličiny Y nezávisí na x_j a je tedy konstantní, což může např. znamenat, že všechny realizace y_1, \dots, y_n náhodných veličin Y_1, \dots, Y_n jsou naměřeny se stejnou přesností.

- 3.** Hodnoty nezávisle proměnné x_j , $j = 1, \dots, n$, nejsou všechny stejné.



4. Matice $F = (f_{ij})$, kde $f_{ij} = f_i(x_j)$, $i = 1, \dots, p$, $j = 1, \dots, n$. má hodnost p . Poznamenejme, že počet n dvojic (x_j, y_j) musí být větší než počet neznámých parametrů p , přesněji, mělo by platit $n - p > 2$.
5. Náhodné veličiny Y_1, \dots, Y_n jsou nekorelované, t.j.

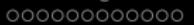
$$\text{cov}(Y_i, Y_j) = 0, \quad i, j = 1, \dots, n, \quad i \neq j.$$

Maticově zapsáno

$$C_y = \sigma^2 E_n,$$

kde E_n je jednotková matice řádu n , C_y je matice kovariance veličin Y_1, \dots, Y_n .

★ Příklad Pro regresní přímku, tj. regresní funkci tvaru $\eta(x) = \alpha + \beta x$ je počet neznámých parametrů $p = 2$ a $\beta_1 = \alpha$, $f_1 = 1$, $\beta_2 = \beta$, $f_2 = x$.



Ekvivalentní model

★ Ekvivalentní model

Popsaný model v ekvivalentním tvaru

$$\mathbf{Y}_j = \eta(x_j) + \varepsilon_j = \sum_{k=1}^p \beta_k f_{kj} + \varepsilon_j, \quad j = 1, \dots, n, \quad (2)$$

kde hodnoty x_1, \dots, x_n jsou hodnotami nenáhodné proměnné, hodnoty $f_{kj} = f_k(x_j)$ splňují podmínu 3. modelu. Pro náhodné chyby $\varepsilon_j, j = 1, \dots, n$, a pro kovarianční matici C_ε náhodného vektoru $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ platí

$$E(\varepsilon_j) = 0, \quad j = 1, \dots, n, \quad C_\varepsilon = \sigma^2 E_n = C_y.$$

Rovnici (2) lze zapsat maticově

$$\vec{\mathbf{Y}} = \mathcal{F}^\top \vec{\beta} + \vec{\varepsilon}.$$

Metoda nejmenších čtverců

Metoda nejmenších čtverců

Odhady neznámých parametrů β_1, \dots, β_p v popsaném modelu lineární regrese budeme hledat **metodou nejmenších čtverců**. Označme tyto odhady b_1, \dots, b_p , což jsou výběrové funkce náhodného výběru Y_1, \dots, Y_n . Minimalizujeme součet čtverců odchylek napozorovaných hodnot y_j od středních hodnot $\eta_j = \eta(x_j)$, tedy součet čtverců

$$Q(\beta_1, \dots, \beta_p) = \sum_{j=1}^n (y_j - \eta_j)^2 = \sum_{j=1}^n \left(y_j - \sum_{k=1}^p \beta_k f_{kj} \right)^2.$$

Odhady b_1, \dots, b_p tedy najdeme jako řešení soustavy rovnic

$$\frac{\partial Q}{\partial \beta_k} = 0, \quad k = 1, \dots, p.$$

Tato soustava se nazývá **soustavou normálních rovnic**.

Metoda nejmenších čtverců

Soustavu pro hledané odhady b_1, \dots, b_p zapíšeme v přehledném tvaru

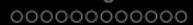
$$\begin{aligned} b_1 S_{11} + b_2 S_{12} + \cdots + b_p S_{1p} &= S_{1y} \\ b_1 S_{21} + b_2 S_{22} + \cdots + b_p S_{2p} &= S_{2y} \\ &\vdots \\ b_1 S_{p1} + b_2 S_{p2} + \cdots + b_p S_{pp} &= S_{py}, \end{aligned}$$

kde

$$S_{ki} = \sum_{j=1}^n f_{kj} f_{ij}, \quad i, k = 1, \dots, p,$$

$$S_{ky} = \sum_{j=1}^n f_{kj} y_j, \quad k = 1, \dots, p.$$

Zřejmě $S_{ik} = S_{ki}$ pro $i, k = 1, \dots, p$.



Metoda nejmenších čtverců

Maticově:

Je-li $\vec{y} = (y_1, \dots, y_n)^T$, $\vec{b} = (b_1, \dots, b_p)^T$, pak **normální rovnice** lze zapsat ve tvaru

$$FF^T \vec{b} = F \vec{y}. \quad (3)$$

Podle předpokladu je $h(F) = p$, pak také $h(FF^T) = p$ a FF^T je typu $p \times p$, regulární \implies existuje $(FF^T)^{-1}$, a tedy z rovnice (3) lze vyjádřit vektor \vec{b} :

$$\vec{b} = (FF^T)^{-1} F \vec{y},$$

vektor \vec{b} je jednoznačně určen a jeho jednotlivé složky jsou lineárními kombinacemi hodnot y_1, \dots, y_n .

Pozor! Výpočet je extrémně numericky nestabilní, viz přednáška "Lineární algebra".



Metoda nejmenších čtverců

★ Příklad

Necht **regresní přímka** prochází počátkem, $\eta(x) = ax$, pak $f_{ij} = x$ a $\beta_1 = a$. Označme $\vec{x} = (x_1, \dots, x_n)^T$, $F = (x_1, \dots, x_n)$. Pak

$$FF^T = (x_1, \dots, x_n) \cdot (x_1, \dots, x_n)^T = \sum_{j=1}^n x_j^2 \implies (FF^T)^{-1} = \frac{1}{\sum_{j=1}^n x_j^2}.$$

Odhad parametru a je

$$a = (FF^T)^{-1}F\vec{y} = \left(\frac{1}{\sum_{j=1}^n x_j^2} \right) \vec{x}^T \vec{y} = \frac{\sum_{j=1}^n x_j y_j}{\sum_{j=1}^n x_j^2}.$$

Metoda nejmenších čtverců

★ Příklad

Růst ledových krystalů

Ledové krystaly byly uloženy do boxu s konstantní teplotou -5°C . Cílem bylo analyzovat růst krystalů jako funkci času. V následující tabulce jsou uvedena naměřená data, y je délka krystalů v mikronech, x je čas v sekundách. Uvedena jsou i opakovaná měření.

S využitím přímkové regrese $y = \beta_0 + \beta_1 x$ určete parametry β_0 a β_1 .

$x[s]$	$y[mm]$	$x[s]$	$y[mm]$
50	19	125	28
60	20, 21	130	31, 32
70	17, 22	135	34, 25
80	25, 28	140	26, 33
90	21, 25, 31	145	31
95	25	150	36, 33
100	30, 29, 33	155	41, 33
105	35, 32	160	40, 30, 37
110	30, 28, 30	165	32
115	31, 36, 30	170	35
120	36, 25, 28	180	38

★ Řešení

Experimentální data zapíšeme maticově

$$y = \begin{bmatrix} 19 \\ 20 \\ 21 \\ 17 \\ \vdots \\ 35 \\ 38 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 50 \\ 1 & 60 \\ 1 & 60 \\ 1 & 70 \\ \vdots & \vdots \\ 1 & 170 \\ 1 & 180 \end{bmatrix},$$

kde $y \in \mathbb{R}^n$ je vektor a $X \in \mathbb{R}^{n \times p}$ matice. Celkový počet měřicích bodů je $n = 43$ a p reprezentuje počet parametrů, v tomto případě dva: β_0, β_1 . První sloupec matice X má v každém řádku jen 1. Počet navzájem různých měřicích bodů je $m = 22$. Protože se mnoho experimentů opakuje, je m podstatně menší, než celkový počet měřicích bodů n . Označme n_i počet měření pro každé x_i , $i = 1, \dots, m$, přičemž $n = \sum_{i=1}^m n_i$. Parametry modelu můžeme spočítat následovně:

$$\mathbf{b} = [\beta_0, \beta_1] = (X^T X)^{-1} X^T y = \begin{bmatrix} 14.19 \\ 0.1346 \end{bmatrix}.$$

Získali jsme model $y = 14.19 + 0.1346 x$.

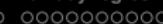
★ Nejlepší nestranný odhad lineární parametrické funkce

Úloha Hledáme nejlepší nestranný odhad lineární funkce parametrů $\vec{\beta} = (\beta_1, \dots, \beta_p)^T$. Uvažujme parametrickou funkci

$$\gamma = \sum_{k=1}^p c_k \beta_k = \vec{c}^T \cdot \vec{\beta},$$

kde $\vec{c} = (c_1, \dots, c_p)^T$ je známý nenulový vektor ($\vec{c} \neq 0$).

Tvrzení Nejlepším nestranným lineárním odhadem parametrické funkce $\vec{c}^T \cdot \vec{\beta}$ je výběrová funkce (statistika) $g = \vec{c}^T \cdot \vec{b}$, kde \vec{b} je řešením normálních rovnic. $E(g) = \gamma$ a "nejlepší" znamená, že rozptyl $D(g)$ je minimální ve třídě nestranných odhadů.



Polynomiální regrese

Regresní model, který je lineární v parametrech, ale popisuje nelineární závislost mezi proměnnými.

$$y = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n,$$

kde b_i jsou neznámé parametry, které chceme určit, n je stupeň polynomu.

- Model obsahuje pouze jednu **nezávisle proměnnou**, která se v něm vyskytuje **vždy ve všech mocninách od 1 do n** .
- Speciálním případem je **přímka - polynom 1. stupně**.
- **Multikolinearita** - přibližná rovnoběžnost sloupcových vektorů v matici **X**
- **Špatná podmíněnost matici $\mathbf{X}^T \mathbf{X}$** . Protože některá vlastní čísla jsou blízká nule, nelze provést inverzi. Připomeňme číslo podmíněnosti

$$\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} \quad \text{nebo také} \quad \kappa = \frac{\sigma_{\max}}{\sigma_{\min}}.$$

Poznámka: Hodnota $\kappa > 1000$ znamená silnou kolinearitu.

Lineární regresní model - regresní parabola

Uvažujme model regresní paraboly tvaru

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + e_i, \quad i = 1, \dots, n,$$

kde neznámé parametry jsou $\beta_1, \beta_2, \beta_3$. Maticově

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}, \quad \mathbf{X}^T \mathbf{X} = \begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{bmatrix},$$

$$\mathbf{X}^T \mathbf{y} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 y_i \end{bmatrix}.$$

Parametry modelu vypočteme následovně

$$\mathbf{b} = [\beta_1, \beta_2, \beta_3] = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Příklad U automobilu Trabant se měřila spotřeba paliva v litrech na 100 km (**y**) v závislosti na jeho rychlosti (**X**).

	Rychlosť	40	50	60	70	80	90	100
Spotřeba		6,1	5,8	6,0	6,5	6,8	8,1	10,0

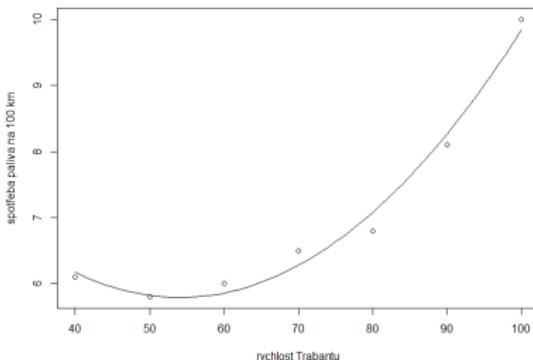
$$\mathbf{X} = \begin{bmatrix} 1 & 40 & 1600 \\ 1 & 50 & 2500 \\ 1 & 60 & 3600 \\ 1 & 70 & 4900 \\ 1 & 80 & 6400 \\ 1 & 90 & 8100 \\ 1 & 100 & 10000 \end{bmatrix}, \quad \mathbf{X}^T \mathbf{X} = \begin{bmatrix} 7 & 490 & 37100 \\ 490 & 37100 & 2989000 \\ 37100 & 2989000 & 252350000 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} 6.1 \\ 5.8 \\ 6.0 \\ 6.5 \\ 6.8 \\ 8.1 \\ 10.0 \end{bmatrix} \quad \mathbf{X}^T \mathbf{y} = [49.3, 3622.0, 286840]$$

$$\mathbf{b} = [11.392857, -0.20726191, 0.0019166667]$$

Parabolická regresní funkce má tvar

$$\mathbf{y} = 11,392857 - 0,207262x + 0,001917x^2$$



Závislost spotřeby paliva na rychlosti Trabantu.

Poznámka Vícenásobná lineární regrese Pokud měřené hodnoty \mathbf{y} závisejí na více faktorech, tedy jsou-li hodnoty \mathbf{y} lineárně závislé na několika nezávisle proměnných \mathbf{x}_i , není vždy možné provést takovou sérii pokusů, při nichž by se měnila pouze hodnota jediné nezávisle proměnné a ostatní proměnné by zůstaly konstantní. V tomto případě lze experimentální výsledky zpracovat pomocí vícenásobné regrese:

$$\mathbf{y} = b_0 + b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \cdots + b_k \mathbf{x}_k ,$$

kde b_i , $i = 1, \dots, k$, je k regresních parametrů. Je-li $k = 1$ dostaneme jednoduchou lineární regresi.



Volba regresní funkce

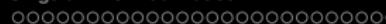
Některé typy lineárních regresních funkcí:

- přímková regrese $y = \beta_1 + \beta_2 x$,
- hyperbolická regrese $y = \beta_1 + \frac{\beta_2}{x}$,
- logaritmická regrese $y = \beta_1 + \beta_2 \ln X$,
- parabolická regrese $y = \beta_1 + \beta_2 x + \beta_3 x^2$,
- polynomiální regrese $y = \beta_1 + \beta_2 x + \cdots + \beta_p x^p$.

Poději se budeme věnovat také nelineární regresi.

Některé typy nelineárních regresních funkcí:

- exponenciální regrese $y = \beta_1 \beta_2^x$,
- mocninná regrese $y = \beta_1 x^{\beta_2}$.



Literatura k dalšímu studiu

- D. R. Cox, Christl A. Donnelly: Principles of Applied Statistics. Cambridge University Press, 2011.
- M. David: Polynomická regrese, <http://ach.upol.cz/user-files/intranet/13-pokrocilaregresy-2017-1505723153.pdf>
- Kubíček M., Dubcová M., Janovská D.: Numerické metody a algoritmy, VŠCHT Praha, 2005 (second edition).
- Harvey Motulsky, Arthur Christopoulos: Fitting Models to Biological Data using Linear and Nonlinear Regression. A practical guide to curve fitting. 2003, GraphPad Software Inc. San Diego CA, www.graphpad.com.
- J. Neubauer: Regresní analýza - Statistika II, <https://k101.unob.cz/neubauer/pdf/regresey1.pdf>
- J. Pavlík a kol.: Aplikovaná statistika. VŠCHT Praha, 2005.
- Anders Rasmuson, Bengt Andersson, Louise Olsson, Ronnie Andersson: Mathematical Modeling in Chemical Engineering. Cambridge University Press, 2014.