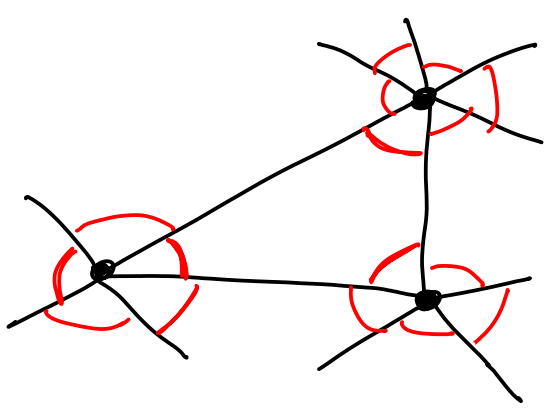
24.3.2015 – 5. cvičení

Metoda nejmenších čtverců a determinanty

Mějme soustavu $Ax=b$, řešení existuje $\Leftrightarrow b \in Im(A)$.
Když řešení existuje, umíme ho najít. Co když není?
V řadě aplikací nějaké řešení chceme najít.

Problémem se zabýval Gauss v zeměměřičství. Představme si, že chceme určit polohy bodů ze znalosti vzdáleností a úhlů.



Příklad:

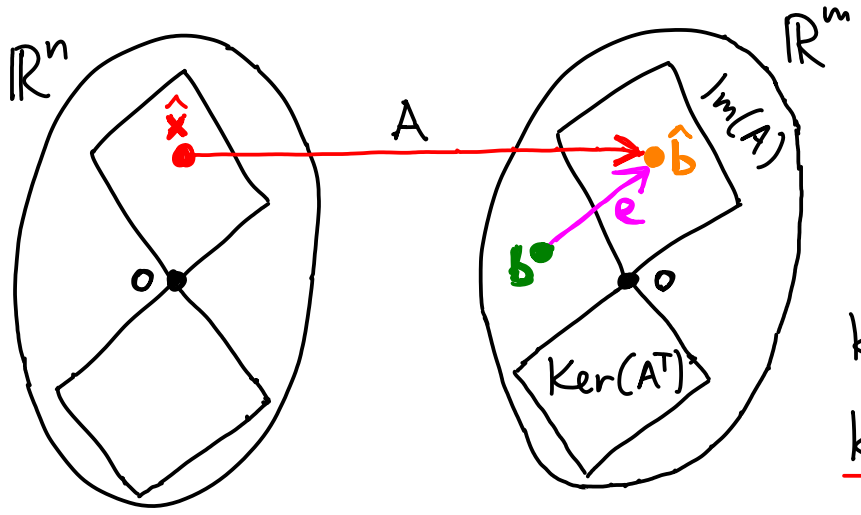
$$2x = b_1$$
$$3x = b_2$$
$$4x = b_3$$

Řešení existuje pouze tehdy, když jsou pravé strany v poměru 2:3:4.

← Pokud pravá strana pochází z měření, typicky nebude platit kvůli chybám.

více rovnic než neznámých, typicky neexistuje řešení

Mohli bychom část rovnic vynechat, ale to bychom museli předpokládat, že chyba je jen v některých, což není typické. Chceme co nejméně změnit pravou stranu, aby řešení existovalo.

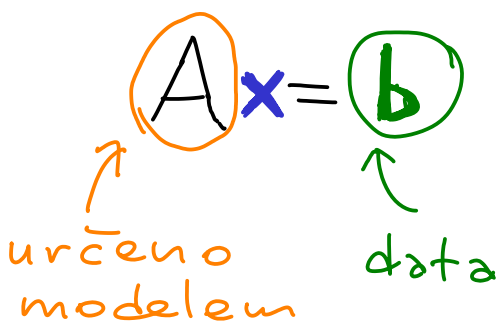


Metoda nejmenších čtverců:

$Ax=b$ .... nemá řešení
$A\hat{x}=\hat{b}$ ... chceme vyřešit, } min. součtu čtverců rozdílu

kde $\hat{b} \in Im(A)$ je nejbližší vektor k b, tedy kolmá projekce na $Im(A)$.

Ve statistice se toto nazývá lineární regrese.



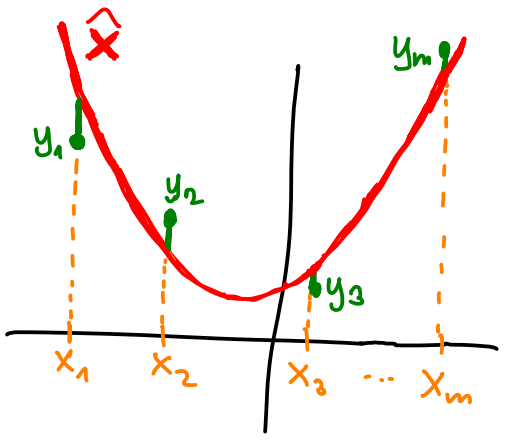
Například data jsou dvojice (x_i, y_i), model je, že leží na parabole.
Hledáme neznámé α, β, γ, aby platilo

$$\alpha x_i^2 + \beta x_i + \gamma = y_i.$$

28 Dostáváme soustavu lineárních rovnic:

$$\left(\begin{array}{ccc|c} x_1^2 & x_1 & 1 & y_1 \\ x_2^2 & x_2 & 1 & y_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_m & 1 & y_m \end{array}\right)$$

← Typicky nebude existovat přesné řešení.



Existuje i totální metoda nejmenších čtverců:

Chceme upravit $A\mathbf{x}=\mathbf{b}$ na $\hat{A}\hat{\mathbf{x}}=\hat{\mathbf{b}}$, aby se minimalizovaly rozdíly $A-\hat{A}$, $\mathbf{b}-\hat{\mathbf{b}}$. Situace, kdy nedůvěřujeme modelu. ← Matematicky složité, nebudeme řešit.

Jak vyřešit nejmenší čtverce?

$\mathbf{e}=\hat{\mathbf{b}}-\mathbf{b}$, platí, že $\mathbf{e}\perp \mathrm{Im}(A)$, neboli $\mathbf{e}\in \mathrm{Ker}(A^T)$.

$$0=A^T\mathbf{e}=A^T(\hat{\mathbf{b}}-\mathbf{b})=A^T(A\hat{\mathbf{x}}-\mathbf{b}),$$

$$\boxed{A^TA\hat{\mathbf{x}}=A^T\mathbf{b}.}$$ ← normální rovnice

Předpokládejme, že A^TA je regulární, potom

$$\hat{\mathbf{x}}=(A^TA)^{-1}A^T\mathbf{b} \quad \text{a} \quad \hat{\mathbf{b}}=\underbrace{A(A^TA)^{-1}A^T}\mathbf{b}.$$

Lze snadno ověřit, že P splňuje $P^2=P$ a $P^T=P$. ← Matice P kolmé projekce na podprostor $\mathrm{Im}(A)$.

Matice A^TA má řadu zajímavých vlastností, například je symetrická pozitivně semidefinitní.

Lemma: $\mathrm{Ker}(A^TA)=\mathrm{Ker}(A)$.

Důkaz: $\mathrm{Ker}(A^TA)\supseteq\mathrm{Ker}(A)$: $A\mathbf{x}=0 \Rightarrow A^TA\mathbf{x}=0$.

$\mathrm{Ker}(A^TA)\subseteq\mathrm{Ker}(A)$: $A^TA\mathbf{x}=0$, vynásobme zleva $\mathbf{x}^T$,

$$0=\mathbf{x}^TA^TA\mathbf{x}=(A\mathbf{x})^TA\mathbf{x}=\|A\mathbf{x}\|^2.$$

← Vektor má nulovou normu, právě když je nulový. □

Proto $\operatorname{rank}(A^TA) = \operatorname{rank}(A)$, tedy pokud má A lineárně nezávislé sloupce, je A^TA regulární a existuje $(A^TA)^{-1}$.

Lze i přes QR rozklad:

$$Ax = b \Rightarrow QRx = b$$

$$Q^TQR\hat{x} = Q^Tb$$

$$\hat{x} = R^{-1}Q^Tb$$

Numericky přesnější, protože A^TA má druhou mocninu čísla podmíněnosti A.

Typicky není čtvercová, proto neekvivalentní úprava: $A \in \mathbb{R}^{m\times n}$, $Q \in \mathbb{R}^{m\times n}$, $Q^TQ = I_n$.

Determinanty

$\det(A): \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$, multilineární alternující forma (vůči sloupcům matice A).

Kolem roku 1900 klíčový pojem lineární algebry, dnes jsou v popředí jiné pojmy, třeba vlastní čísla.
Co může říct jedno číslo o matice? Ne tolik, ale překvapivě hodně!

Jak definovat determinant? ← překvapivě těžké zvolit správný přístup

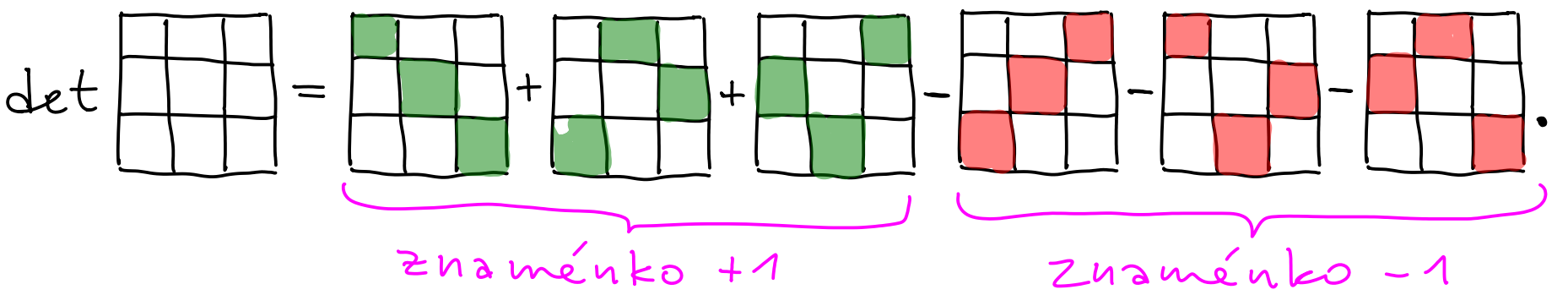
1. Vlastnostmi. ← jako úkol
2. Geometricky. ← příště
3. Formulí v koeficientech:

$$\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

($\operatorname{sgn}(\sigma) \in \{\pm 1\}$)

není jasné, proč zrovna tato definice je zajímavá

Příklad 3×3:



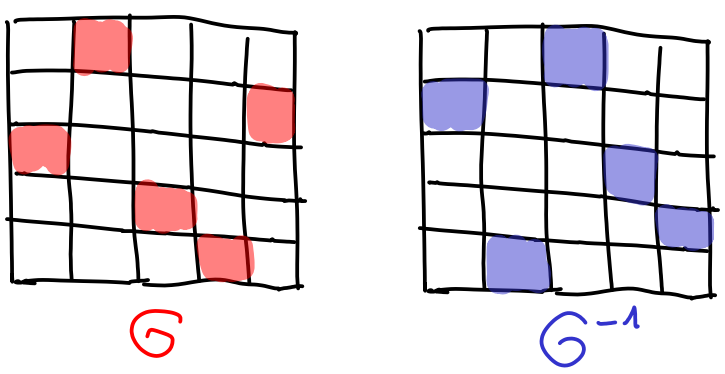
Chceme odvodit zajímavé vlastnosti det(A):

1) $\det(cA) = c^n \cdot \det(A)$. ← každý člen $\prod$ je vynásoben c

2) Vynásobení libovolného řádku c vynásobí c i det.
↖ právě jeden člen $\prod$ je vynásoben c

30 3) $\det(A^T) = \det(A)$. ← Vlastnosti řádků platí i pro sloupce.

Důkaz: Inverzní permutace násobí přes indexy zrcadlově podle diagonály a má stejné znaménko.



Proto člen sumy $\det(A)$ pro G je roven členu $\det(A^T)$ pro G^{-1}, tedy $\det(A) = \det(A^T)$. □

4) Nulový řádek implikuje nulový determinant.

5) Determinant je alternující:
Prohození dvou řádků i a j mění znaménko.

Důkaz: Pro každou permutaci G dostaneme stejné členy Π po prohození řádků při složení G s transpozicí (i,j). Protože složení mění znaménka, má determinant má opačné znaménko. □

6) Dva stejné řádky $\Rightarrow$ nulový determinant. ← stačí prohodit

7) Trojúhelníková matice T má $\det(T) = \Pi t_{i,i}$.

Každá jiná permutace násobí přes prvek pod diagonálou, což je nula.

8) Determinant je lineární v každém řádku.

$$\det\begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & a_i + b_i & — \\ & \vdots & \\ — & a_n & — \end{pmatrix} = \det\begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & a_i & — \\ & \vdots & \\ — & a_n & — \end{pmatrix} + \det\begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & b_i & — \\ & \vdots & \\ — & a_n & — \end{pmatrix}.$$

Důkaz: Každý součin obsahuje jeden člen $a_i + b_i$, lze rozdělit na dva determinanty. □

9) $\det(A) \neq 0 \Leftrightarrow A$ je regulární.

nemění nenulovost

Důkaz: Převod do odstupňovaného tvaru a 7). □