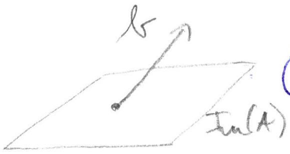
Téma 3: Řešení lin. aproximačních úloh (1)

úloha: $Ax \approx b$, $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$

$\Rightarrow$ lin. model je dán obdélníkovou maticí

chyby: modelovací, měření, ukládání dat v FPA způsobují, že $b \notin Im(A)$ (nekompatibilní problém) $\Rightarrow$ neex. přesné řešení x $\Rightarrow$ „$\approx$" v nějakém smyslu



ozn: $Im(A) = R(A)$ ... image, range A

$Ker(A) = N(A)$ ... jádro, nulový prostor A

$rank(A)$ ... hodnost A

• **chyby pouze v b, A – přesné**:

idea: najdi nejmenší korekci b tak, aby opravená úloha byla kompatibilní

$f = ?$: $f \in \mathbb{C}^n$, $\|f\|$ – malá, $Ax \overset{\text{kompatibilní}}{=} b + \underbrace{f}_{\text{oprava pozorování}}$

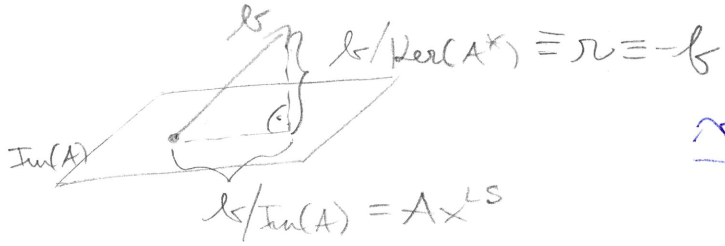
$\Rightarrow f = Ax - b =: -r$ ... residuum

$\Rightarrow$ ekvivalentně hledám $x \in \mathbb{C}^m$ tak, aby $\|b - Ax\|$ byla minimální

def: Nechť $A \in \mathbb{C}^{n \times m}$, $b \in \mathbb{C}^n$. Problém nejmenších čtverců (LS) nazveme úlohu nalezení $x^{LS} \in \mathbb{C}^m$ tak, aby

$$\min_{f \in \mathbb{C}^n} \|f\| \text{ za podmínky } Ax = b + f.$$

$$\Leftrightarrow x^{LS} = \arg\min_{x \in \mathbb{C}^m} \|b - Ax\|$$



víme: $\mathbb{C}^n = Im(A) \oplus Ker(A^*)$,
$Im(A) \perp Ker(A^*)$

$$\Rightarrow b - Ax = \underbrace{b/_{Im(A)} - Ax}_{\in Im(A)} + \underbrace{b/_{Ker(A^*)}}_{\in Ker(A^*)}$$

$$\Rightarrow \|b - Ax\|^2 = \|b/_{Im(A)} - Ax\|^2 + \|b/_{Ker(A^*)}\|^2$$

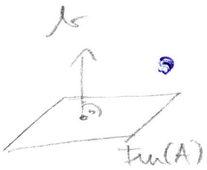
x^{LS}: $\|b - Ax^{LS}\|^2 = 0 + \|b/_{Ker(A^*)}\|^2$

existence a jednoznačnost f a x^{LS}:

- $f = -r = b/_{Ker(A^*)}$ je jednoznačně určená minimální korekce b
- x^{LS} existuje vždy a je řešením opravené soustavy $Ax^{LS} = b/_{Im(A)}$
- x^{LS} je určeno jedn. $\Leftrightarrow$ lze $b/_{Im(A)}$ reprezentovat jednoznačně v $Im(A)$ $\Leftrightarrow$ $\Leftrightarrow$ A má LN sloupce $\Leftrightarrow$ $rank(A) = m$

$A = [a_1 \dots a_m]$, $x^{LS} = \begin{bmatrix} x_1^{LS} \\ \vdots \\ x_m^{LS} \end{bmatrix} \Rightarrow \sum_{i=1}^{m} x_i^{LS} \cdot a_i = b/_{Im(A)}$

koeficienty lineární kombinace



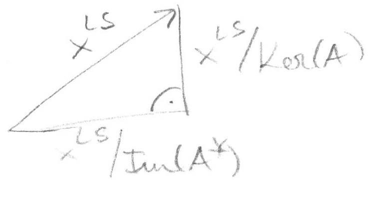
- $x^{LS} = 0 \Leftrightarrow b \perp Im(A)$ ... pozorování b není korelováno s modelem A

případ $rank(A) < m$:

nějaké LS řešení

pozorování: $z \in Ker(A) \Rightarrow A(z + x^{LS}) = Ax^{LS}$

$\Rightarrow \exists\, \infty$ mnoho x^{LS}



$$\|x^{LS}\|^2 = \|x^{LS}|_{Im(A^*)}\|^2 + \|x^{LS}|_{Ker(A)}\|^2 \geq$$

$\geq \|x^{LS}|_{Ker(A)}\|$ a $x^{LS}|_{Ker(A)}$ je také LS řešení $Ax \approx b$

důsledek: $\exists!\, x^{LS}_{min}$ řešení LS problému s minimální normou a je dáno vztahy

$$Ax^{LS}_{min} = b|_{Im(A)}, \quad x^{LS}_{min} \in Im(A^*).$$

D($\exists$): z konstrukce

D(jed.): sporem: nechť ex. x_1, x_2 splňující vztahy výše $\Rightarrow$ $Ax_1 - Ax_2 = b|_{Im(A)} - b|_{Im(A)} = 0$

$(x_1 - x_2) \in Im(A^*)$

$\Rightarrow (x_1 - x_2) \in Ker(A)$ a zároveň $(x_1 - x_2) \in Im(A^*)$

$\Rightarrow x_1 - x_2 = 0 \Rightarrow x_1 = x_2$ $\square$

množina všech LS řešení: $x^{LS}_{min} + Ker(A)$

- chyby v A i b:

idea: najdi nejmenší korekci A, b tak, aby úloha byla kompatibilní

def: Nechť $A \in \mathbb{C}^{n\times m}$, $b \in \mathbb{C}^n$. Problém úplných nejmenších čtverců (TLS) nazveme úlohu nalezení $x^{TLS} \in \mathbb{C}^m$ tak, aby

$$\min_{\substack{f \in \mathbb{C}^n \\ E \in \mathbb{C}^{n\times m}}} \|[f, E]\|_F \text{ za podm. } (A+E)x = b+f.$$

oprava modelu　　kompatibilní

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2} \; \dots \text{Frobeniova norma } A$$ (4)

pozn: • TLS problém nemusí mít řešení

př: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ kompatibilní   velikost opravy

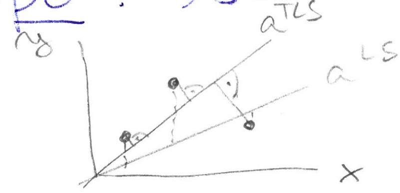
$E := \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon \end{bmatrix} \Rightarrow (A+E)x = b$, kde $x_\varepsilon = \begin{bmatrix} 1 \\ \frac{1}{\varepsilon} \end{bmatrix}$, $\|E\|_F = \varepsilon$

pokud $\varepsilon \to 0$, pak $\|E\| \to 0$, ale neex. minimum

navíc $\|x_\varepsilon\| \xrightarrow{\varepsilon \to 0} \infty$, tedy x_ε není smysluplná aproximace řešení

• analýza TLS úlohy – přes SVD a aproximaci matice $[b|A]$ maticí menší hodnosti

př: srovnání LS a TLS



$[x_i, y_i], i = 1, \dots, n$ ... měřené body

? přímka: $ax = y$, kde $a = ?$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} a \approx \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

LS (lineární regrese): opraví jen y_i

TLS (ortogonální -||-): opraví obě souřadnice

• další zobecnění:

- škálování, constraints, náhodná měření, ...

př: aproximace funkce $f(x)$

máme: přibližné hodnoty $b_i \approx f(x_i)$ v bodech, $i = 1, \dots, n$

chceme: $f(x) \approx \sum_{j=1}^{m} c_j \cdot \varphi_j(x)$, kde φ_j jsou LN fce

$x := x_i \Rightarrow f(x_i) \approx \sum_{j=1}^{m} c_j \cdot \varphi_j(x_i)$, $i = 1, \dots, n$

LS: $\min_{c_1, \dots, c_m \in \mathbb{R}} \sum_{i=1}^{n} \left(b_i - \sum c_j \cdot \varphi_j(x_i) \right)^2$

ω_i – váhy dle přesnosti b_i

3.1 Metody řešení problému LS

⑤

případ rank(A) = m: $\exists!\, x^{LS}$, $A = {}_{n}\overset{m}{\boxed{}}$, $n \geq m$

pozorování: $\|b - Ax\| = \|U^*(b - Ax)\| \;\; \forall U \in \mathbb{C}^{n\times n}$ unitární

$\Rightarrow$ LS (i TLS) je unitárně invariantní,

$\Rightarrow$ A, b: unit. tranformace dat nemění řešení

- **řešení pomocí QR rozkladu:**

$$A = QR = [\overbrace{Q_m}^{m}, \overbrace{\tilde{Q}_m}^{n-m}] \begin{array}{l} \overbrace{\begin{bmatrix} R_m \\ 0 \end{bmatrix}}^{m} \begin{array}{l} \}\, m \\ \}\, n-m \end{array} \end{array} = Q_m R_m$$

ekonomický rozklad

platí: 1, $R_m = \boxed{\diagdown} \in \mathbb{C}^{m\times m}$ je regulární

D: triv.

2, $Q_m = [q_1, \dots, q_m]$, $\operatorname{span}\{q_1, \dots, q_m\} = \operatorname{Im}(A)$

3, $\tilde{Q}_m = [q_{m+1}, \dots, q_n]$, $\operatorname{span}\{q_{m+1}, \dots, q_n\} = \operatorname{Ker}(A^*)$

$$Ax \approx b \;/\, Q^* \cdot \iff Rx \approx Q^* b$$

$$\begin{array}{r} m\,\{ \\ n-m\,\{ \end{array} \begin{bmatrix} R_m \\ \hline 0 \end{bmatrix} x \approx \begin{bmatrix} Q_m^* b \\ \hline \tilde{Q}_m^* b \end{bmatrix} \begin{array}{l} \}\, m \\ \}\, n-m \end{array}$$

koeficienty b v Im(A)

koef. b v Ker(A*)

$\Rightarrow$ $\boxed{R_m x^{LS} = Q_m^* b}$ ... zpětnou substitucí

$\|b - Ax^{LS}\| = \|\tilde{Q}_m^* b\|$ ... velikost residua

implicitní postup výpočtu: Q nekonstruuji.

$$n\{[\overbrace{A}^{m} | \overbrace{b}^{1}] \quad \xrightarrow[\text{NEBO Giv. rotace}]{\text{Hous. reflexe}} \quad \underbrace{H_{m+1} \cdots H_2 H_1}_{\equiv\, Q^*} \cdot [A|b] = \left[\begin{array}{c|c} \overbrace{R_m}^{m} & \overbrace{Q_m^* b}^{1} \\ \hline 0 & \tilde{Q}_m^* b \end{array}\right] \begin{array}{l} \}\, n \\ \}\, n-m \end{array}$$

postup výpočtu přes CGS (MGS):

→ GS počítá ekon. rozklad $A = Q_m R_m$

→ pustíme GS na matici $[A|b]$

$$[A|b] = [\overbrace{Q_m}^{m}|\overbrace{q}^{1}] \cdot \begin{array}{c} \\ \begin{bmatrix} R_m & \overset{\text{OG koef. } b \text{ proti } q_1,\dots,q_m}{b_1} \\ 0 & b_2 \end{bmatrix} \begin{array}{l} \}m \\ \}1 \end{array} \\ \underbrace{\quad}_{m}\ \underbrace{\ }_{1} \end{array}$$

... co je b_1, b_2?

CGS (krok $m+1$): $\tilde{q} := b - \overbrace{Q_m Q_m^* b}^{b/\mathrm{Im}(A)} \in b/\mathrm{Im}(A^*)$

$q = \tilde{q}/\|\tilde{q}\|$

$$\Rightarrow b_1 = Q_m^* b = \begin{bmatrix} q_1^* b \\ \vdots \\ q_m^* b \end{bmatrix}, \quad b_2 = \|\tilde{q}\| = \|b/\mathrm{Im}(A^*)\| = \|b - Ax^{LS}\|$$

shrnutí: $[A|b] \xrightarrow[MGS]{CGS} [Q_m|q] \cdot \begin{bmatrix} R_m & Q_m^* b \\ 0 & \|b - Ax^{LS}\| \end{bmatrix}$

+ zpětná substituce $R_m x^{LS} = Q_m^* b$

vlastnosti: $\varkappa(A) = \dfrac{\sigma_1}{\sigma_m} = \varkappa(R_m)$ (singulární čísla A) $\Rightarrow$ je-li $\varkappa(A) \ll \dfrac{1}{\varepsilon_{mach}}$, pak R_m je numericky regulární a výpočet x^{LS} přes QR je spolehlivý.

- normální rovnice:

$Ax \approx b \quad / A^*$

$\underbrace{A^*A}_{\text{HPD, regulární}} x = A^* b$ .. normální rovnice

platí: Je-li $\mathrm{rank}(A) = m$, pak x^{LS} je řešení normální rovnice $A^*Ax = A^*b$.

D: $A^*Ax = A^*b \overset{A=QR}{\iff} R^*Q^*QRx = R^*Q^*b \iff R_m^* R_m x = R_m^* Q_m^* b \quad / R_m^{-*} \cdot \iff R_m x = Q_m^* b$

řešení norm. rce:

- sestavení (A^*A) je pro $n \gg m$ a m-malé, neboť dostanu malou soustavu rovnic
- iterační metody - používám jen násobení $(A^*A)z = A^*.(A.z)$
  - ex. speciální it. metody pro HPD matice
- přímé metody - Choleského rozklad $\equiv$ $\equiv$ spec. varianta GE pro HPD matice

vlastnosti: norm. rce mají mnohem větší podmíněnost

$$\varkappa(A^*A) = \frac{\sigma_1(A^*A)}{\sigma_m(A^*A)} = \frac{\sigma_1^2(A)}{\sigma_m^2(A)} = \varkappa^2(A)$$

$\Rightarrow$ spolehlivé, je-li $\varkappa^2(A) \ll \frac{1}{\varepsilon_{mach}}$

- rozšířená matice:

$$\underbrace{\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix}}_{\text{sedlobodová matice: regul., herm., indefinitní}} \begin{bmatrix} r \\ x \end{bmatrix} \overset{(*)}{=} \begin{bmatrix} b \\ 0 \end{bmatrix}, \quad r = b - Ax$$

platí: Je-li $\operatorname{rank}(A) = m$, pak x^{LS} je řešení sedlobodové soustavy výše.

D: $(*) \iff \begin{matrix} r + Ax = b \\ A^*r = 0 \end{matrix} \iff \begin{matrix} r = b - Ax \\ A^*r = 0 \end{matrix} \iff$

$\iff A^*(b - Ax) = 0 \iff A^*Ax = A^*b$ □

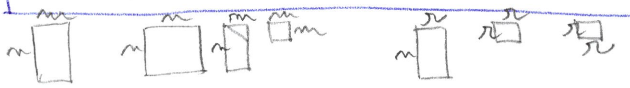
případ $\mathrm{rank}(A) < m$: $\exists!\, x^{LS}_{min}$

$r := \mathrm{rank}(A)$

- řešení pomocí SVD:

‖ $A = U\Sigma V^* = U_r \Sigma_r V_r^*$, U, V - unitární

ekonomický SVD



$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$

víme: U_r má ve sloupcích bázi $\mathrm{Im}(A)$

V_r — || — bázi $\mathrm{Im}(A^*)$

odvození: ozn.: x^{LS} - lib. LS řešení

$A x^{LS} = b|_{R(A)}$ OG projektor na $\mathrm{Im}(A)$

$U_r \Sigma_r V_r^* x^{LS} = \overbrace{U_r U_r^*}^{\text{OG projektor na Im}(A)} b \quad / U_r^* \cdot$

$\Sigma_r V_r^* x^{LS} = U_r^* b$

‖ $V_r^* x^{LS} \overset{(*)}{=} \Sigma_r^{-1} \cdot U_r^* b$ ... koef. x^{LS} v $\mathrm{Im}(A^*)$

x^{LS} je min. v normě $\Leftrightarrow x^{LS} \in \mathrm{Im}(A^*)$

$\Leftrightarrow \exists y \in \mathbb{C}^r: x^{LS} = V_r \cdot y$ ... dosadím do $(*)$

určení y: $\overbrace{V_r^* V_r}^{=I} y = \Sigma_r^{-1} U_r^* b \Rightarrow$

$\Rightarrow$ ‖ $x^{LS}_{min} = V_r \cdot \Sigma_r^{-1} U_r^* b$ ... je LS řešení s min. normou

def: Matici $A^+ := V_r \Sigma_r^{-1} U_r^*$ nazveme

‖ pseudoinverze matice A (Moore-Penroseova)

spec: A - regul. $\Rightarrow A^+ = A^{-1}$

vlastnosti: výpočet přes SVD je zpětně stabilní (OG transf.), ale stojí $O(n^3)$

• **řešení pomocí QR rozkladu**: (čtení navíc)

(obecný postup)

$\exists P \in \mathbb{C}^{n\times n}$ permutační: $AP = n\Big\{\big[\overbrace{\tilde{A}}^{r} \,\big|\, \overbrace{\hat{A}}^{n-r}\big]$, $\operatorname{rank}(\tilde{A}) = r$ ... plný

ekonom. QR

$$\Rightarrow AP = n\Big\{\big[\overbrace{Q_1}^{r} \,\big|\, \overbrace{Q_2}^{n-r}\big] \cdot \begin{array}{c} \begin{bmatrix} \overbrace{R_1}^{r} & \overbrace{R_2}^{n-r} \\ 0 & 0 \end{bmatrix} \begin{array}{l} \}r \\ \}n-r \end{array} \end{array} = Q_1 \cdot \big[R_1 \,\big|\, R_2\big]$$

... počítá se QR rozklad se sloupcovou pivotací (např. přes MGS s pivotací)

potom: $Ax \approx b \iff (AP)\underbrace{(P^T x)}_{=:\hat{x}} \approx b$

$$\big[AP \,\big|\, b\big] \xrightarrow{MGS} \big[Q_1 \,\big|\, Q_2\big]^* \big[AP \,\big|\, b\big] = \left[\begin{array}{cc|c} R_1 & R_2 & Q_1^* b \\ \hline 0 & 0 & Q_2^* b \end{array}\right] \begin{array}{l} \}r \\ \}n-r \end{array}$$

$$\iff \begin{bmatrix} \overbrace{R_1}^{r} & \overbrace{R_2}^{n-r} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \begin{array}{l} \}r \\ \}n-r \end{array} = \begin{bmatrix} Q_1^* b \\ Q_2^* b \end{bmatrix} \iff \hat{x} \text{ není jedn. a}$$

splňuje $\underbrace{R_1 \hat{x}_1 + R_2 \hat{x}_2}_{=\hat{x}} = Q_1^* b \iff \hat{x} = \begin{bmatrix} R_1^{-1}(Q_1^* b - R_2 w) \\ w \end{bmatrix}$

lze volit $\Rightarrow$ vol. 0

shrnutí: $\boxed{x^{LS}_{min} = P \cdot \begin{bmatrix} R_1^{-1} Q_1^* b \\ 0 \end{bmatrix}}$ je LS řeš. s nejmenší normou

$$\boxed{\|r\| = \|b - Ax^{LS}\| = \|Q_2^* b\|} \text{ .. residuum}$$

vlastnosti: - jako QR pro $\operatorname{rank}(A) = n$

- pivotujeme za běhu (jako u GE s pivotací)