Rest z minula PA=LU – pozrieť str. 150–153

Metóda najmenších štvorcov

Doteraz sme sa bavili o štvorcových systémoch $Ax=b$.

Vo všeobecnosti však počet rovníc (napr. výsledkov experimentov) a neznámych (fyz. veličín) nemusí byť rovnaký – zodpovedá tomu obdĺžnikový systém

$$\boxed{A_{m\times n}\, x_{n\times 1} = b_{m\times 1}} = \begin{bmatrix} a_1 \end{bmatrix} x_1 + \begin{bmatrix} a_2 \end{bmatrix} x_2 + \dots + \begin{bmatrix} a_n \end{bmatrix} x_n$$

- Alternatívne, môžeme sa pozerať na lineárne zobrazenie

$$\alpha : \mathbb{R}^n \to \mathbb{R}^m \quad \text{dané predpisom } \alpha(x) = Ax.$$

- kedy má rovnica $Ax=b$ riešenie? $\Leftrightarrow$ $\exists x \in \mathbb{R}^n : Ax = b$,

- množina $\mathcal{R}(A) = \mathrm{Im}(\alpha) = \{Ax \in \mathbb{R}^m \mid x \in \mathbb{R}^n\} = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n, Ax = y\}$

je ~~jediná~~ VP – obraz (α), resp. stĺpcový priestor A.

(range of a matrix).

Podobne máme: $\mathcal{N}(A) = \ker(\alpha) = \{x \in \mathbb{R}^n \mid Ax = 0\}$.

– jadro (α), resp. nulový priestor matice A.

ak A udáva zobrazenie α, potom matica A^T (resp. A^*) dáva zobrazenie $\alpha^* : \mathbb{R}^m \to \mathbb{R}^n$

a máme:

$$\mathcal{R}(A^T) = \mathrm{Im}(\alpha^*) = \{A^T y \in \mathbb{R}^n \mid y \in \mathbb{R}^m\}$$
$$= \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^m \; A^T y = x\}.$$
$$(y^T A = x^T)$$

→ teda ide o riadkový priestor.

$$(y_1, \dots, y_m) \begin{pmatrix} - a_1^T - \\ - a_2^T - \\ \vdots \\ - a_m^T - \end{pmatrix} = y_1 a_1^T + \dots + y_m a_m^T$$

a tiež máme ľavý nulový priestor

$$\mathcal{N}(A^T) = \ker(\alpha^*) = \{y \in \mathbb{R}^m \mid A^T y = 0\}$$

Fakty o týchto priestoroch (boli na ~~m~~ papieroch rozdaných minulý týždeň):

<u>Hodnosť:</u> $h(A) = \dim \mathcal{R}(A) = \dim \mathcal{R}(A^T) = r$

$$\dim \mathcal{N}(A) = n - r$$
$$\dim \mathcal{N}(A^T) = m - r$$

<u>Kolmosť:</u> Ak $x_1 \in \mathcal{N}(A)$ a $x_2 \in \mathcal{R}(A^T)$:

- $Ax_1 = 0$
- $\exists\, y : \; y^T A = x_2^T$
- potom $\langle x_2, x_1 \rangle = x_2^T x_1 = (y^T A) x_1 = y^T (A x_1) = y^T 0 = 0$

teda $x_2 \perp x_1$.

(podobne pre $\mathcal{N}(A)$ a $\mathcal{R}(A^*)$ v $M_{m,n}(\mathbb{C})$)

(Záver) $\Rightarrow$ $\underbrace{\mathcal{N}(A)}_{\text{nulový}} \perp \underbrace{\mathcal{R}(A^T)}_{\text{riadkový}}$

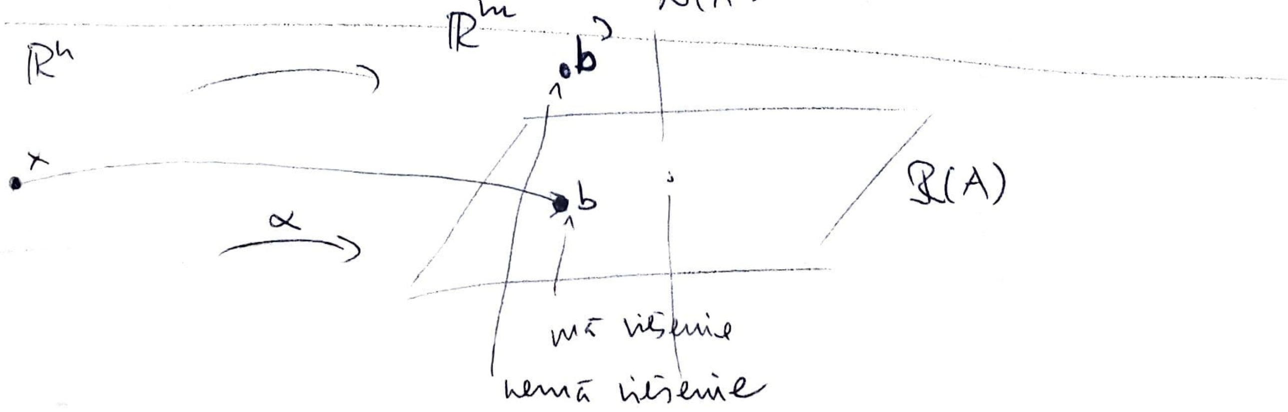
Podobne $\underbrace{\mathcal{R}(A)}_{\text{stĺpcový}} \perp \underbrace{\mathcal{N}(A^T)}_{\text{ľavý nulový}}$

Riešiteľnosť rovníc

$Ax = b$ má (klasické) riešenie $\Leftrightarrow b \in \mathcal{R}(A)$

$\Leftrightarrow b \perp \mathcal{N}(A^T)$.

Obr.



Čo robiť, ak $\vec{b} \notin \mathcal{R}(A)$?

– Ukazuje sa, že správnou odpoveďou je riešenie rovníc metódou najmenších štvorcov.

. Najprv odbočka k maticiam A^TA (a A^*A).

– Nech $A_{m\times n}$ má hodnosť r. Potom $A^T_{n\times m}A_{m\times n}$ je typu $n\times n$.

Pozorovanie : $\mathcal{N}(A^TA) \supseteq \mathcal{N}(A)$ (všeobecne $\mathcal{N}(BA)\supseteq\mathcal{N}(A)$) (ak $x\in\mathcal{N}(A)$ $Ax=0 \Rightarrow A^TAx=0$)

$\mathcal{R}(A^TA) \subseteq \mathcal{R}(A)$ $\mathcal{R}(BA)\subseteq\mathcal{R}(A)$

Tu však dostaneme rovnosť:

<u>Tvrdenie</u> : $h(A) = h(A^TA) = h(AA^T)$

. $\mathcal{R}(A^TA) = \mathcal{R}(A)$ $\qquad \mathcal{R}(AA^T) = \mathcal{R}(A^T)$

. $\mathcal{N}(A^TA) = \mathcal{N}(A)$ $\qquad \mathcal{N}(AA^T) = \mathcal{N}(A^T)$.

---

pre komplexné nahradiť A^T za A^*.

<u>Dôkaz</u> Stačí ukázať inklúziu $\mathcal{N}(A^TA) \subseteq \mathcal{N}(A)$.

Nech $A^TAx=0$. prenásobme x^T zľava:

$$0 = x^TA^TAx = (Ax)^T(Ax) = \|Ax\|^2 = 0$$

z kladnej definitnosti normy $\Rightarrow$ $Ax=0$, $x\in\mathcal{N}(A)$.

potom: $\mathcal{N}(A^TA) = \mathcal{N}(A)$, ich dimenzie sa rovnajú.

/

teda $\dim\mathcal{N}(A^TA) = n - h(A^TA)$

$\qquad \shortparallel$

$\dim\mathcal{N}(A) = n - h(A)$

Teda máme $h(A^TA) = h(A)$ $\left(\overset{\text{a aj}}{=} h(A^T) = h(AA^T)\right)$.

keďže $\dim \mathcal{R}(A^TA) = \dim \mathcal{R}(A)$, pričom

$$\mathcal{R}(A^TA) \subseteq \mathcal{R}(A) \Rightarrow \text{nutne sa rovnajú.}$$

Podobne pre $\mathcal{R}(A^T)$ a $\mathcal{R}(AA^T)$. ∎

Dôsledok pre riešenie rovníc:

Od $m \times n$ systému $Ax = b$

prejdime k $n \times n$ systému $A^TAx = A^Tb$.

normálnych rovníc

Treba si všimnúť, že normálne rovnice sú vždy konzistentné

$A^Tb \in \mathcal{R}(A^T) \overset{\text{ukázali sme}}{=} \mathcal{R}(A^TA)$, teda pravá strana A^Tb

patrí do stĺpcového priestoru matice A^TA.

Ak je $A^Tx = b$ konzistentná $\Rightarrow \exists$ čiastkové

riešenie p: a $Ap = b$ a všeobecné má tvar

$$S: p + \mathcal{N}(A)$$

~~potom aj z všeobecné tiež~~

Potom je p aj čiastkovým riešením systému

$$A^TAx = A^Tb,$$ lebo

$$A^TAp = A^Tb.$$

---

a všeobecné riešenie $S' = p + \mathcal{N}(A^TA) = p + \mathcal{N}(A) = S$.

Ak je $A\cancel{x}=b$ konzistentný a má práve jedno riešenie $(h(A)=n)$

$\Leftrightarrow N(A)=\{\vec{0}\}$ a $b\in R(A)$

v tom prípade aj $A^TAx=A^Tb$ má práve jedno riešenie $\Leftrightarrow N(A^TA)=\{\vec{0}\} \Leftrightarrow \exists (A^TA)^{-1}$

preto $x=(A^TA)^{-1}A^Tb$.

Pozor!! Ak je A obdĺžniková matica $(A^TA)^{-1} \neq A^{-1}(A^T)^{-1}$ s exponentom $^{-1}$ sa nedá ísť dovnútra!

Zhrnutie: Normálne rovnice

- Pre $m\times n$ systém $Ax=b$ máme asociovaný $n\times n$ systém normálnych rovníc $A^TAx=A^Tb$
- $A^TAx=A^Tb$ je vždy konzistentný (aj keď $Ax=b$ nie je)
- Ak $Ax=b$ je konzistentný, množiny riešení $Ax=b$ a $A^TAx=A^Tb$ sú totožné
- Ak $Ax=b$ je nekonzistentný $(b\notin R(A))$, potom ~~tieto~~ riešenia $A^TAx=A^Tb$ tvoria riešenie $Ax=b$ v najmenších štvorcoch
- $A^TAx=A^Tb$ má jednoznačné riešenie práve vtedy, keď $h(A)=n$, potom $x=(A^TA)^{-1}A^Tb$
- Ak je $Ax=b$ konzistentný s ~~pre~~ jednoznačným riešením, tak to platí aj pre $A^TAx=A^Tb$, riešenie: $x=(A^TA)^{-1}A^Tb$.

<u>Pozor!</u> ~~Najm~~ Metóda najmenších štvorcov vyzerá byť veľmi priateľská.

Ale (ako uvidíme)

$$\|A^TA\| \approx \|A\|^2$$

$$\|(A^TA)^{-1}\| \approx \|A^{-1}\|^2$$

(toto nie je pre obdĺž. definované)

teda pre $\kappa_A = \|A\|\|A^{-1}\|$ máme $\kappa_{A^TA} \approx \kappa_A^2$.

Príklad (DÚ)

$$A = \begin{pmatrix} 3 & 6 \\ 1 & 2{,}01 \end{pmatrix} \qquad b = \begin{pmatrix} 9 \\ 3{,}01 \end{pmatrix}$$

riešiť $Ax = b$ v 3-číslicovej aritmetike

$$A^TA = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix} \qquad A^Tb = \begin{pmatrix} 30 \\ 60{,}1 \end{pmatrix}$$

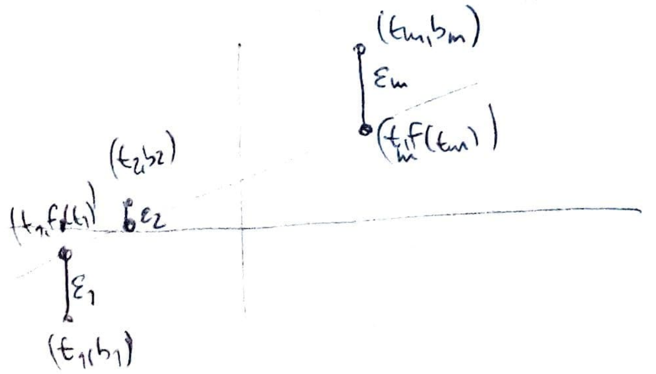
→ prax: normálnym rovniciam sa numerika vyhýba.

↳ metódu najmenších štvorcov (ktorá je ~~d~~ veľmi dôležitá) treba vylepšiť. → QR rozklad

<u>Prečo najmenšie štvorce?</u>

pozorovania v čase (t_i)

Dáta: $D = \{(t_1,b_1),(t_2,b_2),\dots,(t_m,b_m)\}$ – chceme nájsť lineárnu krivku – priamku, ktorá im bude najlepšie zodpovedať.



→ $f(t) = \alpha + \beta t$

→ chybový vektor $\varepsilon_i = |f(t_i) - b_i|$
$= |\alpha + \beta t_i - b_i|$

nájsť α, β také, aby
$$\sum \varepsilon_i^2 = \sum_{i=1}^{m} (\alpha + \beta t_i - b_i)^2$$ bolo minimálne.

4,4

→ minimum → treba parciálne derivácie a kritický bod

$$0 = \frac{\partial\left(\sum_{i=1}^{m}(\alpha+\beta t_i-b_i)^2\right)}{\partial\alpha} = 2\sum_{i=1}^{m}(\alpha+\beta t_i-b_i)$$

$$0 = \frac{\partial\left(\sum_{i=1}^{m}(\alpha+\beta t_i-b_i)^2\right)}{\partial\beta} = 2\sum_{i=1}^{m}(\alpha+\beta_i-b_i)\,t_i$$

→ to dá sústavu dvoch rovníc o dvoch neznámych:

$$\left(\sum_{i=1}^{m}1\right)\alpha + \left(\sum_{i=1}^{m}t_i\right)\beta = \sum_{i=1}^{m}b_i$$

$$\left(\sum_{i=1}^{m}t_i\right)\alpha + \left(\sum_{i=1}^{m}t_i^2\right)\beta = \sum_{i=1}^{m}t_ib_i$$

Resp. pre $x=\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$, $A=\begin{pmatrix}1 & t_1\\ 1 & t_2\\ \vdots & \vdots\\ 1 & t_m\end{pmatrix}$, $b=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}$

tieto rovnice zodpovedajú presne $A^TA\,x = A^Tb$. – normálne rovnice pre systém $Ax=b$.

→

riešenie: $x=(A^TA)^{-1}A^Tb$ (pre 2x2 maticu $\begin{pmatrix}a & b\\ c & d\end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$)

$$= \frac{1}{m\sum t_i^2-(\sum t_i)^2}\begin{pmatrix}\sum t_i^2 & -\sum t_i\\ -\sum t_i & m\end{pmatrix}\begin{pmatrix}\sum b_i\\ \sum t_ib_i\end{pmatrix} =$$

$$= \frac{1}{m\sum t_i^2-(\sum t_i)^2}\begin{pmatrix}\sum t_i^2\sum b_i - \sum t_i\sum t_ib_i\\ m\sum t_ib_i - \sum t_i\sum b_i\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix},$$

Pričom chyba je daná ako:

$$\sum_{i=1}^{m}\varepsilon_i^2 = \sum_{i=1}^{m}(\alpha+\beta t-b_i)^2 = (Ax-b)^T(Ax-b)$$

Metóda najmenších štvorcov

Pre A typu $m \times n$ a $b \in \mathbb{R}^m$ označme $\varepsilon = \varepsilon(x) = Ax - b$

~~Problém~~ najmenších štvorcov je nájsť x v ktorom sa nadobúda minimum

$$\sum_{i=1}^{m} \varepsilon_i^2 = \varepsilon^T \varepsilon = (Ax-b)^T (Ax-b)$$

- vektor, kde sa takéto minimum nadobúda sa nazýva **riešením v najmenších štvorcoch**
- množina riešení v najmenších štvorcoch zodpovedá množine riešení normálnych rovníc $A^T A x = A^T b$
- Ak je rieš. v najmenších štvorcoch jednoznačné, je dané $x = (A^T A)^{-1} A^T b$
- (klasické) Ak je $Ax = b$ konzistentný, riešenie $Ax = b$ je rovnaké ako ~~r~~riešenie v najmenších štvorcoch a $\varepsilon = \vec{0}$ (dĺžka je nulová),

---

Dôkaz ① Ak x minimalizuje $\varepsilon^T \varepsilon$ $\Rightarrow$ x je riešením norm. rovníc.

Rozpíšme: $$\sum_{i=1}^{m} \varepsilon_i^2 = \varepsilon^T \varepsilon = (Ax-b)^T (Ax-b) = x^T A^T A x - 2x^T A^T b + b^T b$$

musíme nájsť minimum

$$f(x_1, \dots, x_n) = x^T A^T A x - 2x^T A^T b + b^T b$$

$\rightarrow$ parciálne derivovať podľa x_i.

dá sa ukázať (du) $$\frac{\partial f}{\partial x_i} = \frac{\partial x^T}{\partial x_i} A^T A x + x^T A^T A \frac{\partial x}{\partial x_i} - 2 \frac{\partial x^T}{\partial x_i} A^T b$$

$$\frac{\partial x}{\partial x_i} = e_i \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \qquad \frac{\partial x}{\partial x_i} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

teda: $$\frac{\partial f}{\partial x_i} = e_i^T A^T A x + x^T A^T A e_i - 2 e_i^T A^T b = 2 e_i^T A^T A x - 2 e_i^T A^T b$$

(lebo $A^T A$ je symetrická)

$e_i^T A^T = (A^T)_{i*}$ ← je i-ty riadok A^T

4.5

máme n rovníc $(A^T)_{i*} Ax = (A^T)_{i*} b$

spojením:

$$A^T A x = A^T b.$$

Analýza dáva iba nutnú podmienku → x je minimum ⇒ x je krit. bod ⇔ x rieši $A^T A x = A^T b$.

Chceme ukázať, že každé riešenie $A^T A x = b x$ zodpovedá minimu.

Nech z je riešením $A^T A x = A^T b$.

$z^T A^T b - z^T A^T b$

Potom $\underbrace{f(z)}_{\text{hodnota chyby v z je:}} = (Az-b)^T(Az-b) = z^T A^T A z - 2z^T A^T b + b^T b$

$= b^T b - z^T A^T b.$

pre iné $y = z + u$ máme:

$$f(y) = (Ay-b)^T(Ay-b) = (Az+Au-b)^T(Az+Au-b)$$

$$= (Az-b)^T(Az-b) - 2u^T A^T(Az-b) + u^T A^T A u$$

$$= f(z) + \|Au\|^2$$

→ t.j. $f(y) \geq f(z)$, rovnosť práve vtedy, keď $Au = \vec{0}$ ⇔ $u \in N(A) = N(A^T A)$