Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Karel Karlík

Metoda nejmenších čtverců při nepřesných datech

Katedra aplikované matematiky

Vedoucí práce: Prof . RNDr. Karel Zimmermann, DrSc.
Studijní program: INFORMATIKA

2008

Chtěl bych poděkovat všem, kteří mi s touto prací nějak pomohli, zejména pak Prof. RNDr Zimmermannovi, DrSc za vedení práce, Prof. RNDr Jiřímu Rohnovi, DrSc za podnětné přednášky, bez kterých bych zkoumanému tématu nerozuměl a rodině za trpělivost, kterou se mnou po dobu tvorby diplomové práce měla.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 15. dubna 2008 Karel Karlík

Obsah

Název práce: Metoda nejmenších čtverců při nepřesných datech
Autor: Karel Karlík
Katedra: Katedra aplikované matematiky
Vedoucí diplomové práce: Prof. RNDr Karel Zimmermann, DrSc
e-mail vedoucího: Karel.Zimmermann@mff.cuni.cz
Abstrakt: Předmětem této diplomové práce je studium metody nejmenších čtverců s nepřesně zadanými daty. Jde především o popis množiny všech řešení a o efektivní výpočet jejího vnějšího obalu.
Klíčová slova: Metoda nejmenších čtverců, intervalová analýza

Title: Least square method with interval data
Author: Karel Karlík
Department: Department of applied mathematics
Supervisor: Prof. RNDr Karel Zimmermann, DrSc
Supervisor's e-mail address: Karel.Zimmermann@mff.cuni.cz
Abstract: The subject of this thesis is the study of the least squares method with uncertain (approximate, imprecise) input data. The focus is on the description of the set of all solutions and on the effective computation of an enclosure of said set.
Keywords: least squares method, interval analysis

Kapitola 1

Úvod

1.1 Značení

- Množinu reálných čísel označíme $\mathbb{R}$ a množinu komplexních čísel $\mathbb{C}$.
- Všechna maticová porovnání se provádějí po prvcích. Tedy, například, jsou-li matice $A, B \in \mathbb{R}^{m \times n}$:

  $$A \leq B \iff (\forall i = 1..m, j = 1..n)(a_{ij} \leq b_{ij})$$

  To znamená, že $A \geq 0$ značí nezápornou matici, $A > 0$ matici kladných prvků a $A \neq B$ že se A a B neshodují ani na jedné pozici. Nerovnost po prvcích se však obvykle nezavádí. Známou nerovnost matic budeme značit $\neg A = B$ nebo ji vyjádříme jiným opisem. Záměnám se vyhneme tím, že nerovnost po prvcích používat nebudeme.
- Stejně tak se po prvcích provádějí i operace max a min.
- Transpozici matice značíme A^T a konjugaci značíme A^H.
- Pomocí $A_{i.}$ označíme i-tý řádek matice A a pomocí $A_{.j}$ j-tý sloupec matice A.
- Vektor je považován za matici s jedním sloupcem.
- Pro matici (nebo vektor) A definujeme $A^+ = max\{A, 0\}$ a obdobně $A^- = max\{0, -A\} = (-A)^+$, kde 0 je matice (vektor) samých nul a příslušného tvaru; zjevně platí $A^+ - A^- = A$, neboť A^+ je shodná s maticí A až na to, že na místě záporných prvků ma nuly, a $-A^-$ je shodná s maticí A až na to, že má nuly na místě prvků kladných.

- Použijeme funkci $\delta(i,j) = \begin{cases} 1 & \text{když } i = j \\ 0 & \text{jinak} \end{cases}$

- I je matice identity příslušného tvaru. Jejím zobecněním je matice $I_{m,n}$, která má tvar $m \times n$ a prvky $(I_{m,n})_{ij} = \delta(i,j)$, tj. má jednotky na diagonále a nuly všude jinde.

- E je matice samých jedniček příslušného tvaru. Je-li potřeba specifikovat tvar $m \times n$, značíme E_{mn}.

- Výjimečně pouze pro vektory značíme e_j vektor, který má jedničku na j-tém místě (řádku) a nuly všude jinde, tj. j-tý sloupec matice I. Vektor e je sloupec matice E, má tedy na všech místech jedničku. Explicitní určení rozměru v tomto případě označíme $e_j^{(i)}$ pro j-tý sloupec matice I_{ii} a $e^{(i)}$ pro j-tý sloupec matice E_{ii} a $e^{(i)}$.

- Pro kapitolu 3 definujeme prostor všech ± 1-vektorů: $Y_m = \{1, -1\}^m$, tedy $Y_m = \{y \in \mathbb{R}^m : |y| = e\}$. Její mohutnost je triviálně 2^m.

- K matici A definujeme matici $sgn(A)$ opět po prvcích:

$$(sgn(A))_{ij} = sgn(A_{ij}), \text{ kde } sgn(x) = \begin{cases} +1 & \text{když } x > 0 \\ 0 & \text{když } x = 0 \\ -1 & \text{když } x < 0 \end{cases}$$

  Obdobně pro vektory. Zjevně $sgn(x) \in Y_m$ pro každý vektor $x \in \mathbb{R}^m$.

- Pro $y \in \mathbb{R}^m$ definujeme diagonální matici T_y jako

$$T_y = \mathrm{diag}(y_1, ..., y_m) = \begin{pmatrix} y_1 & 0 & \cdots & 0 \\ 0 & y_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_m \end{pmatrix}$$

- Pro $a, b \in \mathbb{R}$ označíme $[a, b]$ uzavřený interval těmito čísly ohraničený.

1.2 Intervalové matice

Definice 1 (Intervalová matice a vektor)**.** *Intervalovou maticí $A^I \subset \mathbb{R}^{m \times n}$ nazýváme množinu matic, která splňuje podmínku*

$$\exists \underline{A}, \overline{A} \in A^I : \forall A \in A^I : \underline{A} \leq A \leq \overline{A}$$

Píšeme $A^I = [\underline{A}, \overline{A}]$.

Intervalový vektor je speciální případ intervalové matice pro $n = 1$.

Upozornění: intervalová matice není matice intervalů. Je to v určitém smyslu interval v prostoru matic příslušných dimenzí. A^I je tedy množina matic:

$$A^I \in \mathcal{P}(\mathbb{R}^{m \times n})$$

kde $\mathcal{P}(X)$ značí potenční množinu množiny X.

Intervalovou matici A^I můžeme též ekvivalentně definovat pomocí takzvané centrální matice A_C a poloměru Δ:

$$\begin{aligned} A_C &= \tfrac{\overline{A}+\underline{A}}{2} \\ \Delta &= \tfrac{\overline{A}-\underline{A}}{2} \\ A^I &= [A_C - \Delta, A_C + \Delta] \end{aligned}$$

Potom totiž platí:

$$\begin{aligned} \underline{A} &= A_C - \Delta \\ \overline{A} &= A_C + \Delta \end{aligned}$$

Někdy se též používá značení $mid(A^I) = A_C$, $rad(A^I) = \Delta$.

V případě $m = n = 1$ máme $\underline{A} = (\underline{a})$ a $\overline{A} = (\overline{a})$. A^I pak splývá s reálným intervalem $[\underline{a}, \overline{a}]$, ve shodě s tím, že jednoprvková matice splývá s reálným číslem.

Definice 2 (Intervalové ohrazení)**.** *Intervalový vektor $\overline{\overline{x}}$ nazýváme intervalovým ohrazením množiny X, je-li $X \subset \overline{\overline{x}}$.*

Definice 3 (Intervalový obal)**.** *Intervalový vektor $\overline{x}$ nazýváme intervalovým obalem množiny X, je-li $\overline{x} = \bigcap\{\overline{\overline{x}} | \overline{\overline{x}}$ je intervalové ohrazení množiny $X\}$.*

Intervalový obal $\overline{x}$ je tedy (na rozdíl od intervalového ohraničení) určen jednoznačně a je roven nejmenšímu intervalovému ohraničení (ve smyslu běžného uspořádání množin na inkluzi).

1.3 Metoda nejmenších čtverců v algebře

Buď $A \in \mathbb{R}^{m \times n}$ matice, $b \in \mathbb{R}^m$ vektor. Říkáme, že $x \in \mathbb{R}^n$ řeší $Ax = b$ metodou nejmenších čtverců (píšeme x řeší $Ax = b$ m.n.č., nebo též jen $Ax = b$ m.n.č.), když

$$||Ax - b|| = min\{||Ay - b||,\ y \in \mathbb{R}^n\}$$

kde $||\cdot||$ je eukleidovská norma.

Má-li soustava $Ax = b$ řešení, pak platí:

$$x \text{ řeší } Ax = b \text{ m.n.č.} \iff Ax = b$$

Pokud má matice A lineárně nezávislé sloupce a je $n \leq m$, pak:

$$Ax = b \text{ m.n.č.} \iff A^T Ax = A^T b$$

Rovnosti $A^T Ax = A^T b$ se říká harmonická rovnice a za uvedených podmínek má vždy řešení, protože pak $A^T A$ je regulární. Z toho plyne, že $x_0 = (A^T A)^{-1} A^T b$ je řešením harmonické rovnice a tudíž je řešením $Ax = b$ metodou nejmenších čtverců.

Stojí za zmínku, že v tomto případě platí $A^\dagger = (A^T A)^{-1} A^T$, kde $A^\dagger$ je Moore-Penroseova inverze matice A (viz. např. [14]), takže $x_0 = A^\dagger b$. Moore-Penroseova inverze však není předmětem této práce.

1.4 Metoda nejmenších čtverců ve statistice

Ve statistice se metoda nejmenších čtverců používá k nalezení řešení modelu lineární regrese. Ten lze zapsat například takto:

$$b = Ax + e, A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m, n < m$$

kde A je známá matice, x je vektor neznámých parametrů, e je vektor odchylek (musí splňovat E $e = 0$, var $e = \sigma^2 I$), b je vektor známých náhodných veličin, které závisí na neznámých náhodných veličinách x.

Předpokládá se, že matice A má lineárně nezávislé sloupce, tudíž má svojí plnou hodnost a $A^T A$ je symetrická, regulární a pozitivně definitní.

Vektor x se pak odhaduje metodou nejmenších čtverců, tj. z podmínky, že $(b - Ax)^T(b - Ax)$ má být minimální. Za uvedených předpokladů je $x_0 = (A^T A)^{-1} A^T b$ řešením metodou nejmenších čtverců. Tento odhad je nestranný.

Literatura: [3]

1.5 Intervalová analýza

Při řešení matematických úloh se nám z rozličných důvodů může stát, že nemáme jedno přesné zadání. Vstupní data mohou pocházet z nepřesných měření nebo odhadů. Při jejich zpracování mohlo dojít k zaokrouhlení – obzvláště náchylné je na to zpracování v počítači, kde můžeme čísla reprezentovat pouze s omezenou přesností. Nebo můžeme požadovat znalost řešení pro nějakou třídu úloh, reprezentovanou možným rozsahem nějakých parametrů.

Tento problém též úzce souvisí se studiem perturbací (např. [17]), kde chceme znát odhad na změnu řešení úlohy, pokud známe změnu zadání (perturbaci) – můžeme mít matici reprezentující stabilní stav nějakého systému a studovat chování tohoto systému při malých změnách.

Intervalová analýza přichází s tím, že (reálná) čísla obalí do (reálných) intervalů a v každém kroku výpočtu zaručuje, že skutečná hodnoty přesných řešení leží někde uvnitř těchto intervalů. Nejprimitivnějším postupem je pak přímá metoda: postupuje se jako v případě přesných dat, pouze všechny početní operace jsou nahrazeny jejich intervalovými rozšířeními. Každou funkci $f(x_1, x_2, \ldots, x_n)$ (zde budeme za funkce považovat i základní algebraické operace $+$, $-$, $*$, $/$) nahradíme rozšířením f^I definovaným jako

$$f^I([\underline{x_1}, \overline{x_1}], [\underline{x_2}, \overline{x_2}] \ldots [\underline{x_n}, \overline{x_n}]) \\ = \{f(x_1, x_2, \ldots, x_n) | \quad x_1 \in [\underline{x_1}, \overline{x_1}], x_2 \in [\underline{x_2}, \overline{x_2}] \ldots x_n \in [\underline{x_n}, \overline{x_n}]\}$$

Tato metoda má mnoho úskalí, nicméně vede alespoň k nějakým zaručeným výsledkům.

Podrobnějšímu úvodu do intervalové aritmetiky se věnuje větší počet prací, námatkou [2, 6–9, 16, 18–20].

Především [2] obsahuje pěkný přehled historie intervalové analýzy už od dob Archimeda.

S rostoucí složitostí úlohy (a počtem operací potřebných k jejímu vyřešení) však roste nepřesnost takto získaných odhadů a pro přesnější výsledky je potřeba použít metody přizpůsobené struktuře zadání. Tato práce se snaží najít metodu přizpůsobenou struktuře úlohy nejmenších čtverců.

1.6 Metoda nejmenších čtverců s nepřesnými daty

V úloze $Ax = b$ m.n.č. jsou prvky matice A a vektoru b známé pouze s určitou mírou tolerance. To může způsobit například chyba měření nebo požadavek na jedno řešení vyhovující více různým situacím. Zajíma nás, jaká všechna x musíme vzít do úvahy, pokud chceme mít jistotu, že mezi nimi bude řešení pro přesné hodnoty, nebo té situace, která skutečně nastane.

Definice 4. *Říkáme, že x řeší úlohu $A^I x = b^I$ metodou nejmenších čtverců (píšeme x řeší $A^I x = b^I$ m.n.č., nebo jen $A^I x = b^I$ m.n.č.), když*

$$\exists A \in A^I, b \in b^I : Ax = b \text{ metodou nejmenších čtverců}$$

Dále definujeme množinu všech řešení úlohy:

$$X = \{x : A^I x = b^I \text{ m.n.č.}\}$$

Tedy $X \in \mathcal{P}(\mathbb{R}^n)$, nemusí to však být intervalový vektor.

Budeme se zabývat pouze případem $n \leq m$. V případě $n > m$ nemůže matice A mít lineárně nezávislé sloupce.

Úloha je nyní najít co nejlepší intervalové ohrazení pro množinu X. Přesný popis množiny X – a dokonce už i přesný popis jejího intervalového obalu – je pro obecnou úlohu intervalové metody nejmenších čtverců NP-těžká úloha (neexistuje deterministický algoritmus, který by ji dokázal vyřešit v polynomiálním čase [1]). Spokojíme se tedy s (vnějším) intervalovým ohrazením:

$$\overline{\overline{x}} = [\underline{X}, \overline{X}] : X \subset \overline{\overline{x}}$$

Platí tedy, že $x \in X \to x \in \overline{\overline{x}}$.

1.7 Známé prostředky

Následující prostředky budou dokázané postupně v dalších kapitolách.

1 Je-li $n \leq m$ a sloupce A jsou lineárně nezávislé, pak je $A^T A$ regulární a platí
$$Ax = b \text{ m.n.č.} \iff A^T A x = A^T b$$

2 Pokud existuje $A \in A^I$ takové, že A má lineárně závislé sloupce, pak X je neomezená. Z toho vyplývá, že se můžeme omezit pouze na takové A_I, že všechny $A \in A^I$ mají lineárně nezávislé sloupce.

3 Máme metodu, která umí v polynomiálním čase najít velmi dobré intervalové ohrazení $Y^I \supset Y$ pro úlohu $A^I y = b^I,\ m = n$, kde
$$Y = \{y : \exists A \in A^I, b \in b^I : Ay = b\}$$
za podmínky, že A^I je silně regulární (podmínka silné regularity zároveň implikuje $m = n$, neboť pouze čtvercová matice může být regulární). Definice silné regularity následuje.

Definice 5 (Spektrální poloměr matice)**.** *$\rho(A)$ je spektrální poloměr matice $A \in \mathbb{R}^{m \times n}$, když*

$$\rho(A) = max\{|\lambda| \mid \exists x \in \mathbb{C}^n, x \neq 0 : Ax = \lambda x, \lambda \in \mathbb{C}\}$$

[1] alespoň pokud neplatí P=NP; této rovnosti ovšem věří málokdo, přestože problém zůstává stále otevřený

Ekvivalentně (triviální úprava):

$$\rho(A) = max\{|\lambda| \mid det(A - \lambda I) = 0, \lambda \in \mathbb{C}\}$$

V obecném případě nemáme zaručené, že $\rho(A)$ bude některé z vlastních čísel matice A.

Věta 1. *$A \geq 0 \rightarrow \exists x \geq 0, x \neq 0 : Ax = \rho(A)x$. To znamená, že je-li A nezáporná, je $\rho(A)$ vlastním číslem matice A.*

Definice 6 (Silná regularita)**.** *Intervalovou matici $A^I = [A_C - \Delta, A_C + \Delta]$ nazýváme silně regulární, platí-li $\rho(|A_C^{-1}\Delta|) < 1$*

Platí, že je-li A_C regulární a $\rho(|A_C^{-1}|\Delta) < 1$, pak každá $A \in A^I$ je regulární. Lze též ukázat, že když $max\{|A_C^{-1}|\Delta\}_{jj} \geq 1$, pak existuje $A \in A^I$ singulární.

1.8 Další postup

- $$Ax = b \text{ m.n.č.}$$
- $$\begin{aligned} A^T Ax =& \ A^T b \\ A^T(Ax - b) =& \ 0 \end{aligned}$$
- Položíme $y = Ax - b$. Pak

  $$\begin{aligned} 0x + A^T y =& \ 0 \\ Ax - Iy =& \ b \end{aligned}$$

  takže maticově

  $$\begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

- Položíme $\tilde{A} = \begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix}$. Tato matice je symetrická. Pak platí:

  Tvrzení 1. *$\tilde{A}$ je regulární $\iff$ A má lineárně nezávisle sloupce.*

- Položíme $\tilde{A}^I = \left[\begin{pmatrix} 0 & \underline{A}^T \\ \underline{A} & -I \end{pmatrix}, \begin{pmatrix} 0 & \overline{A}^T \\ \overline{A} & -I \end{pmatrix}\right]$, $\tilde{b}^I = \left[\begin{pmatrix} 0 \\ \underline{b} \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{b} \end{pmatrix}\right]$. To je intervalová matice a platí $\forall A \in A^I : \tilde{A} \in \tilde{A}^I$. Očividně je to nejmenší taková intervalová matice, protože $\underline{A}, \overline{A} \in A^I$.

- Dále postupně rozvineme popis množiny X:

$$\begin{aligned} X = & \{ \quad x : \exists A \in A^I, b \in b^T : Ax = b \text{ m.n.č.} \} \\ = & \{ \quad x : \exists A \in A^I, b \in b^T : A^T A = A^T b \} \\ = & \{ \quad x : \exists A \in A^I, b \in b^T, y \in \mathbb{R}^m : \begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \} \\ \subset & \{ \quad x : \exists \tilde{A} \in \tilde{A}^I, b \in b^I, \exists y \in \mathbb{R}^m : \tilde{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \} \\ = & \{ \quad x : \begin{pmatrix} x \\ y \end{pmatrix} = z \in \mathbb{R}^{m+n}, z \in Z = \{z : \tilde{A}^I z = \tilde{b}^I\} \} \end{aligned}$$

  Nalezneme tedy nějaké vhodné intervalové ohraničení $\bar{Z}$ pro množinu $Z = \{z : \tilde{A}^I z = \tilde{b}^I\}$. Když z prvků této množiny $\bar{Z}$ vezmeme pouze prvních n složek, dostaneme intervalové ohraničení množiny X.

- V další fázi lze využít toho, že x_i jsou spojitými funkcemi všech a_{kl}. Stačí jen aplikovat Cramerovo pravidlo, protože předpokládáme lineární nezávislost sloupců matice A, a proto máme regularitu matice $\begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix}$. Takto implicitně vyjádřenou funkci můžeme parciálně derivovat podle a_{kl} a zkoumat, jaký je vliv na x_i.

- Následně se pokusíme využít toho, že nám stačí omezit se na symetrické $A \in \tilde{A}$.

Kapitola 2

Lineární nezávislost

V této kapitole se pokusíme shrnout nějaká kritéria pro lineární nezávislost sloupců matic $A \in A^I$.

Definice 7. *Intervalová matice A^I se nazývá regulární, je-li každá $A \in A^I$ regulární. Intervalová matice A^I se nazývá singulární, existuje-li $A \in A^I$ singulární (tj. A^I je singulární právě když není regulární).*

Podle [13, 15] je už ověření regularity čtvercové A^I NP-úplný problém.

2.1 Vztah A a $\tilde{A}$

Tvrzení 2. *Jestliže existuje $A \in A^I$ taková, že A má lineárně závislé sloupce, pak množina řešení X je neomezená.*

Důkaz. Buď $A \in A^I$ s lineárně závislými sloupci a $Ax_0 = b$ m.n.č. (tedy $x_0 \in X$). Pak existuje $x_1 \neq 0 : Ax_1 = 0$. To znamená, že $A(x_0 + \alpha x_1) = Ax_0 + \alpha Ax_1 = Ax_0, \alpha \in \mathbb{R}$, a tedy že $A(x_0 + \alpha x_1) = b$ m.n.č. — množina X tedy obsahuje celou přímku $\{x_0 + \alpha x_1 \mid \alpha \in \mathbb{R}\}$ a je neomezená. □

Tvrzení 3. *A má lineárně nezávislé sloupce právě když $\tilde{A} = \begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix}$ je regulární.*

Důkaz.

- $\Leftarrow$ (sporem) nemá-li A lineárně nezávislé sloupce, pak už prvních n sloupců matice $\tilde{A}$ je lineárně závislých a $\tilde{A}$ je tedy singulární.

→ (opět sporem) měj A lineárně nezávislé sloupce a buď $\tilde{A}$ singulární, tedy existuje $x \neq 0$ takové, že $\begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix} x = 0$. Rozepíšeme $x = \begin{pmatrix} a \\ b \end{pmatrix}$, takže

$$\begin{aligned} & & A^T b &= 0 \\ Aa &- & Ib &= 0 \end{aligned}$$

Proto $Aa = Ib = b$. Nemůže být $b = 0$: A má lineárně nezávislé sloupce, takže by muselo být i $a = 0$, a to je spor s předpokladem $x \neq 0$.

Do $A^T b = 0$ dosadíme $b = Aa$ a dostaneme $A^T Aa = 0$. Přenásobíme zleva vektorem a^T, čímž dostaneme $a^T A^T Aa = \|Aa\|^2 = 0$. To je možné pouze tak, že $Aa = 0$. To je ovšem opět $b = 0$ a spor.

□

Vzhledem k této větě nám mohou stačit kritéria na ověření regularity intervalové matice $\tilde{A}^I$.

2.2 Pomocné věty

Maticové normy uvažujeme pouze submultiplikativní, to znamená, že kromě běžných podmínek

$$\begin{aligned} \|A\| &\geq 0 \\ \|A\| = 0 &\rightarrow A = 0 \\ \|\alpha A\| &= |\alpha| \cdot \|A\| \\ \|A + B\| &\leq \|A\| + \|B\| \end{aligned}$$

musí slpňovat ještě submultiplikativní podmínku

$$\|AB\| \leq \|A\| \, \|B\|$$

Běžně používané maticové normy

$$\begin{aligned} \|A\|_F &:= \sqrt{\sum_{i,j} |a_{ij}|^2} \\ \|A\|_1 &:= \max_{j=1 \rightarrow n} \sum_{i=1}^{m} |a_{ij}| \\ \|A\|_\infty &:= \max_{i=1 \rightarrow m} \sum_{j=1}^{n} |a_{ij}| \\ \|A\|_2 &:= \sqrt{\lambda_{max}(A^H A)} \end{aligned}$$

jsou všechny submultiplikativní (důkaz například [17]).

Normy $\|A\|_1$, $\|A\|_2$ a $\|A\|_\infty$ jsou zvláštní případy Hölderových maticových norem, což jsou normy tvaru $\|A\|_p := max\{\|Ax\|_p \mid \|x\|_p = 1\}$, kde $\|\cdot\|_p$ v definici je Hölderova vektorová forma; $\|A\|_F$ se nazývá Frobeniova norma; $\lambda_{max}(A)$ je největší vlastní číslo A.

Věta 2 (Schurova triangulační věta)**.** *Pro každou čtvercovou komplexní matici $A \in \mathbb{C}^{n\times n}$ existuje unitární matice U taková, že $A = UTU^H$, kde T je horní trojúhelníková matice, která má na diagonále vlastní čísla matice A v libovolném (vybraném) pořadí.*

Myšlenka důkazu: jelikož U je unitární, je $U^H U = I$, $U^H = U^{-1}$ a tedy $A = UTU^{-1}$. To znamená, že matice A a T jsou podobné a tedy mají i stejná vlastní čísla:

$$\begin{aligned} &\lambda I - A = \lambda I - UTU^{-1} = \lambda UU^{-1} - UTU^{-1} = U(\lambda I - T)U^{-1} \\ &\rightarrow \det(\lambda I - A) = \det(U)\det(\lambda I - T)\det(U^{-1}) = \det(\lambda I - T) \end{aligned}$$

Jelikož T je trojúhelníková, musí mít vlastní čísla na diagonále. Proto vlastní čísla A jsou právě všechny diagonální prvky T.

Důkaz. Dokazuje se indukcí podle n.

Pro $n = 1$: stačí položit $U = (1)$, $T = (A_{11})$.

Pro $n > 1$: předpokládáme, že věta platí až do $n-1 \geq 1$, že $\lambda_1, \lambda_2, \ldots, \lambda_n$ jsou vlastní čísla matice A ve zvoleném pořadí a že x je ten vlastní vektor příslušející λ_1, který má $\|x\|_2 = 1$.

Vektor x doplníme na unitární matici $U_1 = (xX)$. Potom

$$\begin{aligned} U_1^H A U_1 &= \begin{pmatrix} x^H \\ X^H \end{pmatrix} A\,(xX) \\ &= \begin{pmatrix} x^H Ax & x^H AX \\ X^H Ax & X^H AX \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & x^H AX \\ \lambda_1 X^H x & X^H AX \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & x^H AX \\ 0 & X^H AX \end{pmatrix} \end{aligned}$$

Poslední úprava $X^H x = 0$ platí, protože (xX) je unitární.

Matice U_1 je unitární, proto A a $U_1^H A U_1$ jsou podobné a mají stejná vlastní čísla. Vlastní číslo λ_1 už jsme použili, takže na $X^H AX$ zbývají vlastní čísla $\lambda_2, \ldots, \lambda_n$.

Podle indukčního předpokladu existuje unitární matice $\tilde{U}$ taková, že $\tilde{T} = \tilde{U}^H X^H A X \tilde{U}$ je horní trojúhelníková s diagonálními prvky $\lambda_2, \ldots, \lambda_n$. Potom

$$
\begin{aligned}
T := \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix}^H U_1^H A U_1 \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}^H \end{pmatrix} \begin{pmatrix} \lambda_1 & x^H A X \\ 0 & X^H A X \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & x^H A X \\ 0 & \tilde{U}^H X^H A X \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & x^H A X \tilde{U} \\ 0 & \tilde{U}^H X^H A X \tilde{U} \end{pmatrix}
\end{aligned}
$$

Takže položíme-li $U = U_1 \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix}$, je U (jako součin dvou unitárních matic) unitární a platí $A = UTU^H$, kde T je horní trojúhelníková a má na diagonále prvky $\lambda_1, \ldots, \lambda_n$. $\square$

Věta 3. *Pro každou matici $A \in \mathbb{C}^{n \times n}$ a reálné $\epsilon > 0$ existuje maticová norma $\|\cdot\|_*$ taková, že $\rho(A) > \|A\|_* - \epsilon$.*

Důkaz. Podle věty 2 existuje unitární matice U a horní trojúhelníková T takové, že $U^H A U = T$ a T má na diagonále vlastní čísla matice A.

Definujme pro $\tau \in \mathbb{R}^+$ diagonální matici $D_\tau := T(\tau, \tau^2, \ldots, \tau^n)$, takže $D_\tau^{-1} := T(\tau^{-1}, \tau^{-2}, \ldots, \tau^{-n})$. Podíváme se na matici $H = D_\tau T D_\tau^{-1}$.

Platí $H_{ij} = \tau^i T_{ij} \tau^{-i} = T_{ij}$. To znamená, že $H_{ii} = T_{ii}$, $H_{ij} = 0$ když $i > j$ (protože pak i $T_{ij} = 0$) a $\lim_{\tau \to \infty} H_{ij} = 0$, když $i < j$.

Proto $H = D_\tau U^H A U D_\tau^{-1}$ je horní trojúhelníková a pro dostatečně velké τ_0 je v každém jejím sloupci součet prvků nad diagonálou $\sum_{i=0}^{j-1} |T_{ij} \tau_0^{i-j}| < \epsilon$. Navíc je matice H podobná matici A a proto má na diagonále právě její vlastní čísla $\lambda_1, \ldots, \lambda_n$.

Na tuto matici použijeme normu $\|A\|_1 = \max_j \sum_i |A_{ij}|$. Podle předchozího odhadu:

$$\|D_{\tau_0} U^H A U D_{\tau_0}^{-1}\| < \max_j (|\lambda_j| + \epsilon) = \max_j |\lambda_j| + \epsilon = \rho(A) + \epsilon$$

Definujeme normu $\|A\|_* := \|D_{\tau_0} U^H A U D_{\tau_0}^{-1}\|_1$. Základní vlastnosti maticových norem splňuje triviálně z toho, že je splňuje už $\|\cdot\|_1$ a D_{τ_0} i U jsou regulární. Zbývá ověřit submultiplikativní podmínku $\|XY\|_* \leq \|X\|_* \cdot \|Y\|_*$:

$$
\begin{aligned}
\|XY\|_* &= \|D_{\tau_0} U^H X Y U D_{\tau_0}^{-1}\|_1 \\
&= \|D_{\tau_0} U^H X U D_{\tau_0}^{-1} D_{\tau_0} U^H Y U D_{\tau_0}^{-1}\|_1 \\
&\leq \|D_{\tau_0} U^H X U D_{\tau_0}^{-1}\|_1 \cdot \|D_{\tau_0} U^H Y U D_{\tau_0}^{-1}\|_1 \\
&= \|X\|_* \cdot \|Y\|_*
\end{aligned}
$$

$\square$

Věta 4. *Pro každou matici $A \in \mathbb{R}^{n \times n}$ jsou následující podmínky ekvivalentní:*

1 $\rho(A) < 1$

2 $\lim_{j\to\infty} A^j = 0$

3 $\sum_{j=0}^{\infty} A^j$ *konverguje*

Platí-li libovolné z těchto tvrzení, pak $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$.

Důkaz. Důkaz ekvivalence provedeme v cyklu $3 \to 2 \to 1 \to 3$.

$3 \to 2$ Označíme $S_n := \sum_{j=0}^{n} A^j$ a $B := \lim_{n\to\infty} S_n$ (podle bodu 3 tato limita existuje). Pak též $B = \lim_{n\to\infty} S_{n+1}$, z čehož plyne

$$\lim_{j\to\infty} A^j = \lim_{j\to\infty} (S_j - S_{j-1}) = B - B = 0$$

$2 \to 1$ Buď $\lambda \in C, x \in \mathbb{C}^n$ nějaký pár vlastního čísla a vlastního vektoru matice A, tedy $Ax = \lambda x$. Podle předpokladu je $\lim_{j\to\infty} A^j = 0$ a tedy i $\lim_{j\to\infty} A^j x = \lim_{j\to\infty} \lambda^j x = 0$ Protože $x \neq 0$, musí být $\lim_{j\to\infty} \lambda^j = 0$, což je možné pouze tak, že $|\lambda| < 1$. Protože λ jsme volili libovolně, je i $\rho(A) = \max\{|\lambda|, Ax = \lambda x\} < 1$.

$1 \to 3$ Zvolme $\epsilon > 0$ dost malé, aby $\rho(A) + \epsilon < 1$. Podle věty 3 existuje maticová norma $\|\cdot\|_*$ taková, že $\|A\|_* < \rho(A) + \epsilon < 1$. Rozepíšeme nyní

$$\begin{aligned}
\|A^m + A^{m+1} + A^{m+2} + \dots\|_* &= \|A^m(I + A + A^2 + \dots)\|_* \\
&\leq \|A^m\|_* \cdot \|I + A + A^2 + \dots\|_* \\
&\leq \|A^m\|_* \cdot (\|I\|_* + \|A\|_* + \|A\|_*^2 + \dots)
\end{aligned}$$

Víme že $\|A\|_* < 1$, takže poslední závorka vpravo je součet konvergentní geometrické řady a rovná se $1/(1 - \|A\|_*)$. Máme tedy

$$\|A^m + A^{m+1} + A^{m+2} + \dots\|_* \leq \frac{\|A\|_*^m}{1 - \|A\|_*}$$

Opět proto, že $\|A\|_* < 1$, dokážeme pro každé $\epsilon > 0$ najít m dost velké, aby $\|A\|_*^m/(1 - \|A\|_*) < \epsilon$. Tím jsme ověřili Bolzano-Cauchyovu podmínku konvergence řady.

Dále ještě dokážeme $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$:

Označíme $S^m := \sum_{j=0}^{m} A^j$ a $S := \lim_{j\to\infty} S^m$. Tato limita existuje podle bodu 3.

Rozepíšeme $S^{m+1} = \sum_{j=0}^{m+1} A^j = I + AS^m$. Limitním přechodem na obou stranách pro $m \to \infty$ dostaneme $S = I + AS$, tedy $(I - A)S = I$. To znamená, že $(I - A)^{-1}$ existuje a rovná se $S = \sum_{j=0}^{\infty} A^j$. $\square$

Důsledek: je-li A nezáporná a $\rho(A) < 1$, pak $(I - A)^{-1}$ existuje a je nezáporná. Dokonce je $(I - A)^{-1} >= I$, protože se rovná $I + A + A^2 + \dots$ a všechny členy řady jsou nezáporné.

2.3 Regularita $\tilde{A}$

V následujících větách dokážeme, že když $\tilde{A}_C$ je regulární a sepktální poloměr $\rho\left(|(|\tilde{A}_C^{-1})\Delta\right) < 1$, pak $\tilde{A}$ je regulární. Protože tuto podmínku budeme potřebovat při použití věty 8 na úlohu $\tilde{A}^I z = \tilde{b}^I$, nebudeme vlastně k ověření lineární nezávislosti sloupců matic z A^I potřebovat jiné podmínky.

Nejdříve dokážeme Oettli-Pragerovu charakterizaci:

Věta 5. *Intervalová matice A^I je singulární právě tehdy, když existuje netriviální řešení soustavy nerovností $|A_C x| \le \Delta|x|$.*

Důkaz. $\leftarrow$ Máme $x \neq 0$, které řeší soustavu $|A_C x| \le \Delta|x|$. Definujeme vektory y a z a matici A:

$$
\begin{aligned}
y_i &= \begin{cases} (A_C x)_i/(\Delta|x|)_i & \text{když } (\Delta|x|)_i \neq 0 \\ 1 & \text{jinak} \end{cases} \\
z_y &= \begin{cases} 1 & \text{když } x_j \ge 0 \\ 0 & \text{jinak} \end{cases} \\
A_{ij} &= (A_C)_{ij} - y_i z_j \Delta_{ij}
\end{aligned}
$$

Je zjevné, že $|y_i| \le 0$ a $|z_i| \le 0$, takže $A \in A^I$.

Protože Δ je nazáporná matice a $|x|$ nezáporný vektor, je $(\Delta|x|)_i = 0$ možné pouze tak, že $\Delta_{ij}|x|_j = 0$ pro každé j. Po dosazení tedy máme, při použití $\text{sgn}(0) = 0$, $\text{sgn}(x_i) = z_i$ jinak:

$$A_{ij} = (A_C)_{ij} - \left(\frac{(A_C x)_i}{(\Delta|x|)_i}\right) \text{sgn}(x_i)\Delta_{ij}$$

kde podíl v závorce definujeme jako 1, když $\text{sgn}(x_i)\Delta_{ij} = 0$.

Dále máme

$$\begin{aligned}
(Ax)_{i\cdot} &= \left(\sum_j (A_c)_{ij} x_j\right) - \left(\sum_j \frac{(A_C x)_i}{(\Delta|x|)_i} (\operatorname{sgn}(x_j) x_j \Delta_{ij})\right) \\
&= (A_C x)_i - (A_C x)_i \frac{\sum_j |x_j| \Delta_{ij}}{\sum_k \Delta_{ik} |x|_k} \\
&= (A_C x)_i - (A_C x)_i \\
&= 0
\end{aligned}$$

Proto $Ax = 0$ a nalezli jsme singulární matici $A \in A^I$.

$\rightarrow$ Existuje $A \in A^I$, $x \neq 0$ takové, že $Ax = 0$. Definujme y a z stejně jako v předchozí části. Pak

$$\begin{aligned}
A_C x &= (A + T_y \Delta T_z) x = T_y \Delta |x| \\
|A_C x| &= |\ T_y \Delta |x|\ | \leq |T_y| \Delta |x| \leq \Delta |x|
\end{aligned}$$

kde T_y, T_z jsou diagonální matice s y, respektive z na diagonále.

□

Lemma 1. *Je-li $A \geq 0$, $x \geq 0$ a $Ax \geq x$, pak pro každé $n > 0$ platí $A^n x \geq x$*

Důkaz. Indukcí podle n:

Pro $n = 1$ tvrzení platí přímo z předpokladu.

Pro $n > 1$: označme $y := Ax$. Protože $y \geq x$, existuje $z \geq 0$ takové, že $y = x + z$. Z předpokladu $A \geq 0$ plyne $A^{n-1} \geq 0$, tedy i $A^{n-1} z \geq 0$. Máme teď

$$A^n x = A^{n-1} A x = A^{n-1} y = A^{n-1} x + A^{n-1} z \geq A^{n-1} x$$

a podle indukčního předpokladu $A^{n-1} x \geq x$. □

Věta 6. *Je-li A_C regulární a $\rho(|A_C^{-1}|\Delta) < 1$, pak A^I je regulární.*

Důkaz. Sporem: předpokládejme, že A^I je singulární, tedy podle věty 5 existuje $x \neq 0$ takové, že $|A_C x| \leq \Delta |x|$. Rozepíšeme

$$|x| = |A_C^{-1} A_C x| \leq |A_C^{-1}| |A_C x| \leq |A_C^{-1}| \Delta |x|$$

Máme tedy $|A_C^{-1}| \Delta |x| \geq |x|$. Jelikož součin neáporných matic je nezáporný a $|x|$ je také nezáporný vektor, je pro každé n podle tvrzení lemmatu 1 $(|A_C^{-1}|\Delta)^n |x| \geq |x|$, takže musí být $\lim_{j\to\infty} |A_C^{-1} \Delta|^j |x| \geq |x|$. Vektor $|x|$ je netriviální a nezáporný, takže nemůže být $\lim_{j\to\infty} |A_C^{-1} \Delta|^j = 0$. Ale podle věty 4 z předpokladu $\rho(|A_C^{-1}|\Delta)$ plyne, že $\lim_{j\to\infty} |A_C^{-1} \Delta|^j = 0$, takže musí být i $\lim_{j\to\infty} |A_C^{-1} \Delta|^j |x| = 0$, což je spor. □

Kapitola 3

Čtvercový systém rovností s nepřesnými daty

V této části stručně zopakujeme postup, který vede k důkazu věty o ohraničení množiny řešení čtvercové soustavy intervalových rovnic $A^I x = b^I$. Tato věta je součástí látky probírané v rámci přednášek Prof. RNDr J. Rohna a je uvedena v jeho práci [13] jako věta Hansen-Bliek-Rohn.

Podle [13] je NP-těžké najít intervalový obal (už ve velmi speciálním případě).

Poznámka: tato úloha je speciální případ úlohy nejmenších čtverců, neboť, je-li $Ax_0 = b$ pro nějaké x_0, pak platí $Ax = b \iff Ax = b$ m.n.č.: Nelze najít x takové, aby $\|Ax - b\| < 0$ (tedy x_0 řesí úlohu $Ax = b$ m.n.č.), navíc $\|Ax - b\| = 0 \iff Ax = b$. To znamená, že z NP-obtížnosti této úlohy plyne NP-obtížnost úlohy nejmenších čtverců. Podle [10] je pak NP-obtížná už nějaká přibližná varianta této úlohy.

3.1 Oettli-Pragerova věta

Připomeňmě si, že v úloze $A^I x = b^I$ uvažujeme $A^I = [A_C - \Delta, A_C + \Delta]$, $b^I = [b_c - \delta, b_c + \delta]$.

Definice 8. *Vektor $x \in \mathbb{R}^n$ se nazývá slabé řešení úlohy $A^I x = b^I$ právě tehdy, když existuje $A \in A^I$, $b \in b^I$ takové, že $Ax = b$.*

Věta 7 (Oettli-Prager)**.** *Vektor $x \in \mathbb{R}^n$ je slabé řešení úlohy $A^I x = b^I$ právě tehdy, když je řešením soustavy $|A_C x - b_c| \leq \Delta |x| + \delta$.*

Důkaz. $\rightarrow$ Předpokládáme existenci $A \in A^I$, $b \in b^I$ takových, že $Ax = b$. Dále použijeme fakt $|b - b_c| \in [-\delta, \delta]$ a $|A_C - A| \in [-\Delta, \Delta]$.

Přičteme nulu a rozepíšeme:

$$\begin{aligned}|A_C x - b_C| &= |(A_C - A)x + Ax - b_c| \\ &= |(A_C - A)x + b - b_c| \\ &\leq |(A_C - A)x| + |b - b_c| \\ &\leq |A_C - A||x| + \delta \\ &\leq \Delta|x| + \delta\end{aligned}$$

Tím jsme dokázali $|A_C x - b_C| \leq \Delta|x| + \delta$.

$\leftarrow$ Předpokládáme $|A_C x - b_C| \leq \Delta|x| + \delta$. Definujeme vektor z jako $\operatorname{sgn} x$ a vektor y vztahem

$$y_i := \begin{cases} \frac{(A_C x - b_C)_i}{(\Delta|x| + \delta)_i} & \text{když } (\Delta|x| + \delta)_i > 0 \\ 1 & \text{jinak} \end{cases}$$

Poznámka: není-li $(\Delta|x| + \delta)_i > 0$, musí být $(\Delta|x| + \delta)_i = 0$, protože součet nezáporných čísel $(\Delta|x| + \delta)_i$ je nezáporný.

Pro takto definované z a y platí $|x| = T_z x$, $|y_i| \leq 1$ a

$$(A_C x - b_C)_i = y_i(\Delta|x| + \delta)_i = (T_y(\Delta|x| + \delta))_i$$

To znamená, že

$$A_C x - b_C = T_y(\Delta T_z x + \delta)$$

Nyní převedeme složky s vektorem x na jednu stranu:

$$(A_C - T_y \Delta T_z)x = b_C + T_y \delta$$

Protože $|T_y \Delta T_z| \leq |T_y|\Delta|T_z| \leq \Delta$, je $A_C - T_y \Delta T_z \in A^I$. Stejně tak je $b_C + T_y\delta \in b^I$, takže jsme našli $A := A_C - T_y\Delta T_z$, $b := b_C + T_y\delta$ takové, že $A \in A^I$, $b \in b^I$ a $Ax = b$, což jsme chtěli dokázat.

$\square$

Dosazením $b^i = [0, 0]$ a přidáním podmínky $x \neq 0$ na obě strany ekvivalence dostáváme tvrzení

$$\exists A \in A^I, x \in \mathbb{R}^n, x \neq 0 : Ax = 0 \iff \exists x \in \mathbb{R}^n, x \neq 0 : |A_C x| \leq \Delta|x|$$

To je věta 5.

3.2 Intervalové ohrazení

Lemma 2. *Jsou-li x, y vektory, $x \leq y$ a A nezáporná matice, pak $Ax \leq Ab$.*

Důkaz. Z předpokladů máme $x_i \leq y_i$, tedy i $a_{ji}x_i \leq a_{ji}y_i$, protože každé $a_{ji} \geq 0$. Sečtením těchto nerovností přes i pro pevné j dostaneme j-tý řádek soustavy $Ax \leq Ay$. □

Věta 8. *Nechť $A^I = [A_C - \Delta, A_C + \Delta]$ má tvar $n \times n$ a je silně regulární ($\rho(|A_C^{-1}\Delta|) < 1$). Pak množinu X řešení úlohy $A^I x = b^I, b^I = [b_c - \delta, b_c + \delta]$ lze ohradit takto:*

$$X \subset [\min\{\underset{\sim}{x}, T_\nu \underset{\sim}{x}\}, \max\{\tilde{x}, T_\nu \tilde{x}\}]$$

kde

- min, max *jsou brány po prvcích*
- $M = (I - |A_C^{-1}|\Delta)^{-1}$
- $\mu = (M_{11}, M_{22}, \ldots, M_{nn})$ *je vektor diagonálních prvků matice M*
- $T_\nu = (2T_\mu - I)^{-1}$*; to je inverze diagnonální matice, tudíž je to také diagonální matice*
- $x_C = A_C^{-1} b_C$
- $x^* = M(|x_C| + |A_C^{-1}|\delta)$
- $\underset{\sim}{x}_i = -x^* + T_\mu(x_C + |x_C|)$
- $\tilde{x}_i = x^* + T_\mu(x_C - |x_C|)$

Důkaz. Podle věty 7 máme

$$X = \{x : |A_C x - b_C| \leq \Delta|x| + \delta\}$$

Budeme upravovat Oettli-Pragerovu podmínku:

$$|A_C x - b_C| \leq \Delta|x| + \delta$$

Vynásobíme zleva nezápornou maticí $|A_C^{-1}|$

$$|A_C^{-1}||A_C x - b_C| \leq |A_C^{-1}|(\Delta|x| + \delta)$$

Použijeme vztah $|AB| \leq |A||B|$, který platí, protože $|\sum a_i b_i| \leq \sum |a_i||b_i|$.

$$|(A_C^{-1})(A_C x - b_C)| \leq |A_C^{-1}|(\Delta|x| + \delta)$$

Na levé straně můžeme roznásobit.

$$|x - A_C^{-1}b_C| \leq |A_C^{-1}|(\Delta|x| + \delta)$$

Označíme podle tvrzení věty $x_C := A_C^{-1}b_c$

$$|x - x_C| \leq |A_C^{-1}|(\Delta|x| + \delta)$$

Levou stranu odhadneme dvakrát podle vztahů

$$a \leq |a|$$

a

$$|a| - |b| \leq ||a| - |b|| \leq |a - b|$$

Tím dostaneme

$$\begin{array}{lcl} (1) & x - x_C & \leq |x - x_C| \leq |A_C^{-1}|(\Delta|x| + \delta) \\ (2) & |x| - |x_C| & \leq |x - x_C| \leq |A_C^{-1}|(\Delta|x| + \delta) \end{array}$$

Pak pro každé $i = 1 \ldots n$ najdeme odhad pro x_i sestavením i-té nerovnosti (1) a ostatních rovnic (2).

$$\begin{array}{rcl} |x_1| - |(x_C)_1| & \leq & (|A_C^{-1}|(\Delta|x| + \delta))_1 \\ \ldots & \leq & \ldots \\ |x_{i-1}| - |(x_C)_{i-1}| & \leq & (|A_C^{-1}|(\Delta|x| + \delta))_{i-1} \\ x_i - (x_C)_i & \leq & (|A_C^{-1}|(\Delta|x| + \delta))_i \\ |x_{i+1}| - |(x_C)_{i+1}| & \leq & (|A_C^{-1}|(\Delta|x| + \delta))_{i+1} \\ \ldots & \leq & \ldots \\ |x_n| - |(x_C)_n| & \leq & (|A_C^{-1}|(\Delta|x| + \delta))_n \end{array}$$

Přejdeme k vektorům pomocí vztahů

$$\begin{pmatrix} |x_1| \\ |x_2| \\ \ldots \\ |x_{i-1}| \\ x_i \\ |x_{i+1}| \\ \ldots \\ |x_n| \end{pmatrix} = |x| + (x_i - |x_i|)e_i, \quad \begin{pmatrix} |(x_C)_1| \\ |(x_C)_2| \\ \ldots \\ |(x_C)_{i-1}| \\ (x_C)_i \\ |(x_C)_{i+1}| \\ \ldots \\ |(x_C)_n| \end{pmatrix} = |x_C| + ((x_C)_i - |(x_C)_i|)e_i$$

Pak pro každé $i = 1 \ldots n$ platí (člen s x_C převedeme na pravou stranu)

$$|x| + (x_i - |x_i|)e_i \leq |x_C| + ((x_C)_i - |x_C|_i)e_i + |A_C^{-1}|(\Delta|x| + \delta)$$

Členy s $|x|$ převedeme na levou stranu a $|x_C|$ posuneme doprava.

$$(I - |A_C^{-1}|\Delta)|x| + (x_i - |x_i|)e_i \leq ((x_C)_i - |x_C|_i)e_i + |x_C| + |A_C^{-1}|\delta$$

Označíme M a x^* podle tvrzení věty. Protože $\rho(|A_C^{-1}\Delta|) < 1$, máme podle věty 4 a jejího důsledku zaručenou existenci M a vztah $M \geq I$. Podle definic ze znění věty pak máme $\mu_i \geq 1$, $2\mu_i - 1 \geq 1$ a $0 < (T_\nu)_{ii} = 1/(2\mu_i - 1) \leq 1$.

Předchozí nerovnost přenásobíme zleva M a dostaneme

$$|x| + M(x_i - |x_i|)e_i \leq M((x_C)_i - |x_C|_i)e_i + x^*$$

Vytáhneme i-tou nerovnost

$$\begin{aligned} |x_i| + (x_i - |x_i|)(Me_i)_i &\leq x_i^* + ((x_C)_i - |x_C|_i)(Me_i)_i \\ |x_i| + (x_i - |x_i|)M_{ii} &\leq x_i^* + ((x_C)_i - |x_c|_i)M_{ii} \end{aligned}$$

Označíme $\tilde{x}_i$ podle tvrzení věty, takže $\tilde{x}_i = x_i^* + M_{ii}((x_C)_i - |x_C|_i)$ a provedeme rozbor případů podle x.

$x_i \geq 0$: $|x_i| = x_i$ a

$$\begin{aligned} x_i + (x_i - x_i)M_{ii} &\leq \tilde{x}_i \\ x_i &\leq \tilde{x}_i \end{aligned}$$

$x_i < 0$: $|x_i| = -x_i$ a

$$\begin{aligned} -x_i + (x_i + x_i)M_{ii} &\leq \tilde{x}_i \\ x_i(2M_{ii} - 1) &\leq \tilde{x}_i \\ x_i(T_\nu)_{ii}^{-1} &\leq \tilde{x}_i \\ x_i &\leq (T_\nu)_{ii}\tilde{x}_i \end{aligned}$$

Dále rozborem podle $\tilde{x}_i$:

$\tilde{x}_i < 0$: Nemůže být $x_i \geq 0$, tudíž je $x_i \leq (T_\nu)_{ii}\tilde{x}_i = \max\{\tilde{x}_i, (T_\nu)_{ii}\tilde{x}_i\}$, protože $\tilde{x}_i < 0$, $0 < (T_\nu)_{ii} \leq 1$ a tedy $\tilde{x}_i \leq (T_\nu)_{ii}\tilde{x}_i$.

$\tilde{x}_i \geq 0$ a $x_i \geq 0$: $x_i \leq \tilde{x}_i = \max\{\tilde{x}_i, (T_\nu)_{ii}\tilde{x}_i\}$, protože $\tilde{x}_i \geq 0$, $0 < (T_\nu)_{ii} \leq 1$ a tedy $\tilde{x}_i \geq (T_\nu)_{ii}\tilde{x}_i$.

$\tilde{x}_i \geq 0$ a $x_i < 0$: $x_i < 0 \leq \max\{\tilde{x}_i, (T_\nu)_{ii}\tilde{x}_i\}$.

Dohromady tedy máme horní mez pro x_i:

$$x_i \leq \max\{\tilde{x}_i, (T_\nu)_{ii}\tilde{x}_i\}$$

což je přesně i-tá složka horní meze ohrazení z tvrzení věty.

Obrácený odhad dostaneme ze symetrie. Protože $Ax = b \rightarrow -Ax = -b$, platí

$$A^I(x) = [b_C - \delta, b_C + \delta] \iff A^I(-x) = [-b_C - \delta, -b_C + \delta]$$

Když zopakujeme předchozí důkaz s dosazením $-x$ místo x a $-b_C$ místo b_C, dostaneme všude $-x_C$ místo x_C. Matice M se nezmění. Veličina x^* se nezmění, protože $M(|-x_C|+|A_C^{-1}|\delta = M(|x_C|+|A_C^{-1}|\delta$. Při určování horního odhadu pro $-x_i$ pak místo $\tilde{x}_i$ dostaneme $-\underset{\sim}{x}_i$:

$$\begin{aligned} -\underset{\sim}{x}_i &:= x_i^* + M_{ii}((-x_C)_i - |-x_C|_i) \\ \underset{\sim}{x}_i &= -x_i^* + M_{ii}((x_C)_i + |x_C|_i) \end{aligned}$$

Dále máme $-x_i \leq \max\{-\underset{\sim}{x}_i, -(T_\nu)_{ii}\underset{\sim}{x}_i\}$, takže:

$$x_i \geq \min\{\underset{\sim}{x}_i, (T_\nu)_{ii}\underset{\sim}{x}_i\}$$

Dohromady jsme tedy dostali

$$x_i \in [\min\{\underset{\sim}{x}_i, (T_\nu)_{ii}\underset{\sim}{x}_i\}, \max\{\tilde{x}_i, (T_\nu)_{ii}\tilde{x}_i\}],\ i = 1, \ldots, n$$

$\square$

Kapitola 4

První aproximace

Definice 9. *Pro* $A \in \mathbb{R}^{m\times n}$ *definujeme sloupcový prostor matice* $R(A) = \{Ax, x \in \mathbb{R}^n\}$ *a jádro* $N(A) = \{x \in \mathbb{R}^n, Ax = 0\}$

Definice 10. *Je-li* M *podporostor nějakého skalárního prostoru* V*, označíme jeho komplementární podprostor* $M^{\perp}$*.*

Komplementární podprostor lze získat například tak, že vezmeme libovolnou ortonormální bázi $u_1, \dots, u_r$ podprostoru M, rozšíříme ji na ortonormální bázi $u_1, \dots, u_n$ prostoru V a za bázi podprostoru $M^{\perp}$ vezmeme $u_{r+1}, \dots, u_n$. Tato definice nezávisí na konkrétní volbě báze.

Triviálně pak platí, že pro každé $x \in M, y \in M^{\perp}$ je $x^T y = 0$. Dále je vidět, že pro každé $x \in V$ existuje právě jedno $x' \in M$ a právě jedno $x'' \in M^{\perp}$ takové, že $x = x' + x''$ — stačí x vyjádřit v bázi $u_1, \dots, u_n$ a rozdělit na část v bázi $u_1, \dots, u_r$ a na část v bázi $u_{r+1}, \dots, u_n$. Stejný argument ukazuje, že $x'^T x'' = 0$.

Protože $a = b + c \;\&\; a = b + d \to c = d$, máme navíc

Tvrzení 4. *Je-li* $a \in V$*,* $a = b + c$ *a* $b \in M$*, pak* $c \in M^{\perp}$*.*

Věta 9 (Closest point theorem)**.** *Buď* M *podprostor skalárního (lineárního) prostoru* V*,* $b \in V$*,* $u_1, \dots, u_r$ *libovolná ortonormální báze* M*. Definujeme* $P_M := \sum_{i=1}^{r} u_i u_i^T$*.* P_M *se nazývá ortonormální projekce na* M*. Potom existuje právě jedno* $p \in M : \|p - b\|_2 = \min_{p\in M} \|p - b\|_2$*. Toto* p *je rovno* $P_M b$*.*

Pozn.: zjevně $P_M b \in M$, neboť se jedná o lineární kombinaci vektorů báze M.

Důkaz. Existence: buď $p = P_M b$, pak:

- Pro každé $m \in M$ je $p - m \in M$.

- $b-p=(I-P_M)b \in M^{\perp}$ podle tvrzení 4, neboť $(I-P_M)b+P_Mb=b\in V$ a $P_Mb\in M$
- Tedy $(p-m)^T(b-p)=0$.
- Z toho dále plyne

$$\begin{aligned}\|b-m\|_2^2 &= \|b-p+p-m\|_2^2\\ &= \|b-p\|_2^2+2(b-p)^T(p-m)+\|p-m\|_2^2\\ &\geq \|p-m\|_2^2\end{aligned}$$

Protože $m\in M$ jsme zvolili libovolně, máme

$$\min_{m\in M}\|b-m\|_2 \geq \|p-m\|_2$$

a rovnost nastává, protože $p\in M$ a tudíž m vlevo může nabýt i hodnoty p.

Jedinečnost: buď $m\in M$ a $\|b-m\|_2=\|b-p\|_2$. Protože opět platí $(b-p)^T(p-m)=0$ (vektory z navzájem kolmých podprostorů) a navíc $\|b-m\|_2=\|b-p\|_2$, máme

$$\begin{aligned}\|b-m\|_2^2 &= \|b-p+p-m\|_2^2\\ &= \|b-p\|_2^2+2(b-p)^T(p-m)+\|p-m\|_2^2\\ &= \|b-m\|_2^2+\|p-m\|_2^2\end{aligned}$$

a musí být $\|p-m\|_2^2=0$. To už znamená, že $p=m$. $\square$

Je vidět, že $R(P_M)=M$, protože $P_Mb\in M$ pro každé b a z věty 9 plyne, že $P_Mm=m$ pro každé $m\in M$.

Tvrzení 5. $R(A)^{\perp}=N(P_{R(A)})$

Důkaz. Především platí $0\in R(A)$, protože $A0=0$, a dále $0\in N(P_{R(A)})$, protože $P_{R(A)}0=0$. Dále

$$\begin{aligned}N(P_{R(A)}) &= \{y \mid P_{R(A)}y=0\}\\ y &= P_{R(A)}y+(I-P_{R(A)})y\\ P_{R(A)}y &\in R(A)\\ (I-P_{R(A)})y &\in R(A)^{\perp} \text{ podle tvrzení 4}\\ y\in N(P_{R(A)}) &\rightarrow y=(I-P_{R(A)})y\in R(A)^{\perp}\end{aligned}$$

Tím je dokázáno $N(P_{R(A)})\subset R(A)^{\perp}$.

Není-li $y \in N(P_{R(A)})$, a $y \neq 0$: vektor y má nenulovou část ortogonálního rozkladu v množině $R(A)$, protože $P_{R(A)}y \neq 0$. Nemůže tedy existovat ortogonální rozklad $y = 0 + y$, $0 \in R(A)$, $y \in R(A)^\perp$. Tím je dokázáno $R(A)^\perp \subset N(P_{R(A)})$ □

Tvrzení 6. $R(A)^\perp = N(A^T)$

Důkaz. Aby $x \in R(A)^\perp$, musí být vyjádřitelné v nějaké bázi prostoru $R(A)^\perp$. Ta je ortogonálním doplňkem nějaké ortogonální báze prostoru $R(A)$. Báze prostoru $R(A)$ generuje právě stejný prostor jako sloupce matice A. Proto musí být x ortogonální na všechny sloupce matice A, tedy $x^T A_{\cdot j}, j = 1 \ldots n$. To je dohromady $x^T A = 0^T$, po transpozici $A^T x = 0$, což je právě tehdy, když $x \in N(A^T)$. □

Věta 10. *Následující je ekvivalentní:*

1 x řeší $A^T A x = Ab$

2 $Ax = P_{R(A)}b$ (P_M je definováno stejně jako ve větě 9)

3 $\|Ax - b\|_2 = \min_{y \in \mathbb{R}^n} \|Ay - b\|_2$

Důkaz. $1 \leftrightarrow 2$: Nejdříve použijeme $Ax \in R(A)$, takže $P_{R(A)}Ax = Ax$.

$$\begin{aligned} Ax = P_{R(A)}b &\iff P_{R(A)}Ax = P_{R(A)}b \\ &\iff P_{R(A)}(Ax - b) = 0 \\ &\iff Ax - b \in N(P_{R(A)}) \end{aligned}$$

Podle tvrzení 5 a 6 je $N(P_{R(A)}) = R(A)^\perp = N(A^T)$. Pokračujeme teď

$$\begin{aligned} Ax = P_{R(A)}b &\iff Ax - b \in N(A^T) \\ &\iff A^T(Ax - b) = 0 \\ &\iff A^T Ax = A^T b \end{aligned}$$

což bylo dokázat.

$2 \leftrightarrow 3$: Chceme dokázat, že

$$Ax = P_{R(A)}b \iff \|Ax - b\|_2 = \min_{y \in \mathbb{R}^n} \|Ay - b\|_2$$

Přepíšeme $p = Ax$, dostaneme

$$p = P_{R(A)}b \iff \|p - b\|_2 = \min_{p \in R(A)} \|p - b\|_2$$

To je znění věty 9 pro $M = R(A), V = \mathbb{R}^n$. □

Předchozí věty dokazují, že úloha $Ax = b$ m.n.č. má vždy řešení a že $Ax = b$ m.n.č. $\iff A^TAx = A^Tb$.

Nyní můžeme zopakovat postup z části 1.8 až po vyjádření množiny řešení X pomocí Z:

$$X \subset \{x : \begin{pmatrix} x \\ y \end{pmatrix} = z \in \mathbb{R}^{m_n}, z \in Z = \{z : \tilde{A}^I z = b^I\}\}$$

kde

$$A^I = \left[\begin{pmatrix} 0 & \underline{A}^T \\ \underline{A} & -I \end{pmatrix}, \begin{pmatrix} 0 & \overline{A}^T \\ \overline{A} & -I \end{pmatrix} \right]$$

Pomocí věty 8 najdeme ohrazení $\tilde{Z}$ pro množinu Z. Předpokládáme-li lineární nezávislost sloupců pro všechny matice $A \in A^I$, stačí nám už jen silná regularita $\tilde{A}^I$, tedy podmínka

$$\rho \left(\left| \begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix}^{-1} \right| \begin{pmatrix} 0 & \Delta^T \\ \Delta & 0 \end{pmatrix} \right) < 1$$

Kapitola 5

Další zpřesnění

5.1 Využití jiných metod

Různé metody mohou dát řešení, která se netriviálně překrývají. Protože každá metoda zaručuje, že její výstup množinu řešení X obsahuje, je X obsažena i v průniku řešení všech metod, které použijeme.

Všimneme si, že podle postupu z části 1.8 je řešení $A^I x = b^I$ m.n.č. ekvivalentní „symetrickému" řešení $\tilde{A}^I x = \tilde{b}^I$:

$$\begin{aligned} X &= \{x : \exists A \in A^I, b \in b^T, y \in \mathbb{R}^m : \begin{pmatrix} 0 & A^T \\ A & -I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \} \\ &= \{x : \exists \tilde{A} \in \tilde{A}^I, A = A^T, \exists b \in b^I, \exists y \in \mathbb{R}^m : \tilde{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \} \end{aligned}$$

Dostaneme tedy lepší ohrazení, pokud se omezíme ne symetrické $\tilde{A}$.

Následující přehled shrnuje některé metody, které se věnují nalezení nějakého intervalového ohrazení množiny

$$X_{SYM} := \{x : \exists A \in A^I, A = A^T, \exists b \in b^I : Ax = b\}$$

a tedy se dají použít k nalezení potenciálně lepšího intervalového ohrazení množiny X.

Práce [1] využívá intervalovou variantu Fourier-Motzkinovy eleminace a konstruktivně poskytuje popis množiny X_{SYM} pomocí soustavy lineárních a kvadratických nerovností. tento popis je přesný, ale velikost systému (a tedy i čas potřebný k řešení) je exponenciální.

Práce [4] se zmiňuje o intervalové variantě Choleského metody, která ale experimentálně nevychází nejlépe. Dále se zmiňuje o metodě Hansen-Bliek-Rohn-Ning-Kearfott, která je založená na vlastnostech H-matic (viz. například [12]). Také zavádí několik metod založených na QR-faktorizaci a intervalových Householderových transformacích.

Kniha [17] se zabývá přímo intervalovou úlohou nejmenších čtverců, ale jako výsledek dává pouze odhad pro normu $\|x_C - x\|$, když $x \in X$.

Všechny tyto metody mají společnou tu vlastnost, že když dostaneme nějaké ohrazení pro množinu X, které není neomezené, máme zaručeno, že všechny matice $A \in A^I$ mají lineárně nezávislé sloupce.

5.2 Využití parciálních derivací bez symetrie

Podívejme se na čtvercový systém $Ax = b$ s předpokladem, že A' je regulární na nějakém malém okolí A, tj. že existuje $\epsilon > 0$ takové, že $[A_C - \epsilon E, A_C + \epsilon E]$ je regulární intervalová matice.

Protože x_i lze podle Cramerova pravidla vyjádřit z koeficientů matice A pouze pomocí základních aritmetických operací, je x_i jako funkce všech a_{kl} spojitá.

Vyjádříme nyní parcialní derivovace $\frac{\partial x_i}{\partial a_{kl}}$.

Všechny prvky matice A považujeme za nezávislé, takže $\frac{\partial}{\partial a_{kl}} a_{kl} = 1$, $\frac{\partial}{\partial a_{kl}} a_{ij} = 0$ jinak.

Z k-tého řádku $Ax = b$ dostaneme

$$\begin{aligned}
\frac{\partial}{\partial a_{kl}} \sum_{j=1}^{n} (a_{kj} x_j) &= \frac{\partial}{\partial a_{kl}} b_k \\
\sum_{j=1}^{n} \left(\frac{\partial a_{kj}}{\partial a_{kl}} x_j + a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
\frac{\partial a_{kl}}{\partial a_{kl}} x_l + \sum_{j=1}^{n} \left(a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
x_l + \sum_{j=1}^{n} \left(a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0
\end{aligned}$$

a ze všech ostatních řádků i dostaneme

$$\begin{aligned}
\frac{\partial}{\partial a_{kl}} \sum_{j=1}^{n} (a_{ij} x_j) &= \frac{\partial}{\partial a_{kl}} b_i \\
\sum_{j=1}^{n} \left(\frac{\partial a_{ij}}{\partial a_{kl}} x_j + a_{ij} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
\sum_{j=1}^{n} \left(a_{ij} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0
\end{aligned}$$

Ve vektorové formě pak dohromady máme

$$A\nabla_{a_{kl}}x = -x_l e_k$$

kde

$$\nabla_{a_{kl}}x = \begin{pmatrix} \frac{\partial x_1}{\partial a_{kl}} \\ \vdots \\ \frac{\partial x_n}{\partial a_{kl}} \end{pmatrix}$$

Položme nyní $B := A^{-1}$, takže $\nabla_{a_{kl}}x = -x_l B e_k$. Vybráním l-tého řádku dostaneme

$$\frac{\partial x_l}{\partial a_{kl}} = -x_l B_{lk}$$

Zkonstruujeme intervalovou matici B^I podle věty 8 a vztahů

$$\begin{aligned} A^I B_{\cdot 1} &= e_1 \\ &\dots \\ A^I B_{\cdot i} &= e_i \\ &\dots \\ A^I B_{\cdot n} &= e_n \end{aligned}$$

kde výsledné ohrazení X_i pro $B_{\cdot i}$ budeme brát jako i-tý sloupec B^I v tom smyslu, že $\underline{X_i} = \underline{B}_{\cdot i}$ a $\overline{X_i} = \overline{B}_{\cdot i}$. Potom pro každé $A \in A^I$ existuje $B \in B^I$ takové, že $AB = I$. Pro x takové, že existuje $A \in A^I$, $b \in b^I$ tak, že $AX = b$ pak máme

$$\frac{\partial x_l}{\partial a_{kl}} \in \left[\min\{-x_l\underline{B}_{lk}, -x_l\overline{B}_{lk}\}, \max\{-x_l\underline{B}_{lk}, -x_l\overline{B}_{lk}\}\right]$$

Není-li $0 \in [\underline{B}_{lk}, \overline{B}_{lk}]$, znamená to, že $x_l(a_{kl})$ je monotónní pro $x > 0$ a monotónní pro $x < 0$. To znamená, že může nabývat pouze extrémů $x_l(\underline{a}_{kl})$, $x_l(\overline{a}_{kl})$ a $x_l(a_{kl} = 0$. Pokud navíc z předchozích výpočtů ohrazení víme buďto $x_l > 0$ nebo $x_l < 0$, víme i ve kterém případě nastává který extrém a můžeme vynechat možnost $x_l = 0$.

Můžeme nyní dočasně nahradit intervalovou matici A^I intervalovými maticemi F^I, G^I které budou mít $\underline{F}_{kl} = \overline{F}_{kl} = \underline{A}_{kl}$, $\underline{G}_{kl} = \overline{G}_{kl} = \overline{A}_{kl}$ a ostatní prvky shodné s A^I.

Zopakujeme postup z kapitoly 4 a ze získaných ohrazení vezmeme část odpovídající x_l. Tím jsme potenciálně zlepšili už získané ohrazení.

Výsledné ohrazení bude v l-té složce průnikem l-tých složek takto vzniklých ohrazení pro všechna k, pro která $0 \notin [\underline{B}_{lk}, \overline{B}_{lk}]$ a původního ohrazení z kapitoly 3.

Nakonec můžeme vektor x (a jeho ohrazení) opět rozložit na původní x a y (a jejich ohrazení).

Tuto metodu bohužel nemůžeme snadno rozšířit na postupnou eliminaci nepřesností, protože nemáme zaručené, že různá x_i, x_j budou nabývat extrémů současně. Protože však $x_i(a_{11}, \ldots, a_{nn})$ musí nabývat globálního extrému pouze tam, kde nabývají extrému všechny $x_i(a_{ij})$, můžeme jejich hledání sloučit, tedy pro dané k vyzkoušet všechny kombinace

$$\begin{aligned} \underline{F}_{kl} &= \overline{F}_{kl} \in \{\underline{A}_{kl}, \overline{A}_{kl}\} \text{ kde } 0 \notin [\underline{B}_{lk}, \overline{B}_{lk}] \\ \underline{F}_{ij} &= \underline{A}_{ij}, \overline{F}_{ij} = \overline{A}_{ij} \text{ jinak} \end{aligned}$$

a ze vzniklých ohrazení vzít pro celkové ohrazení v každé složce minimum a maximum.

Pokud pro některé k, l víme, který extrém x_l nastává pro kterou hodnotu a_{kl}, můžeme tuto úlohu dále rozdělit na výpočet ohrazení pro maximální x_l a výpočet ohrazení pro minimální x_l. Při těchto výpočtech neuvažujeme kombinace, o kterých víme, že v nich příslušný extrém nemůže nastat.

5.3 Využití parciálních derivací se symetrií

Uvažujeme-li symetrii, máme $\frac{\partial a_{kl}}{\partial a_{kl}} = 1$, $\frac{\partial a_{lk}}{\partial a_{kl}} = 1$, $\frac{\partial a_{ij}}{\partial a_{kl}} = 0$ jinak.

Pro $k = l$ dostaneme stejný výsledek jako v předchozí části.

Pro $k \neq l$ z k-tého řádku $Ax = b$ dostaneme

$$\begin{aligned} \frac{\partial}{\partial a_{kl}} \sum_{j=1}^{n} (a_{kj} x_j) &= \frac{\partial}{\partial a_{kl}} b_k \\ \sum_{j=1}^{n} \left(\frac{\partial a_{kj}}{\partial a_{kl}} x_j + a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\ \frac{\partial a_{kl}}{\partial a_{kl}} x_l + \sum_{j=1}^{n} \left(a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\ x_l + \sum_{j=1}^{n} \left(a_{kj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \end{aligned}$$

z l-tého řádku

$$\begin{aligned}
\frac{\partial}{\partial a_{kl}} \sum_{j=1}^{n} (a_{lj} x_j) &= \frac{\partial}{\partial a_{kl}} b_l \\
\sum_{j=1}^{n} \left(\frac{\partial a_{lj}}{\partial a_{kl}} x_j + a_{lj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
\frac{\partial a_{lk}}{\partial a_{kl}} x_k + \sum_{j=1}^{n} \left(a_{lj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
x_k + \sum_{j=1}^{n} \left(a_{lj} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0
\end{aligned}$$

a ze všech ostatních řádků i

$$\begin{aligned}
\frac{\partial}{\partial a_{kl}} \sum_{j=1}^{n} (a_{ij} x_j) &= \frac{\partial}{\partial a_{kl}} b_i \\
\sum_{j=1}^{n} \left(\frac{\partial a_{ij}}{\partial a_{kl}} x_j + a_{ij} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0 \\
\sum_{j=1}^{n} \left(a_{ij} \frac{\partial x_j}{\partial a_{kl}} \right) &= 0
\end{aligned}$$

Ve vektorové formě a se stejnou definicí B a B^I pak dohromady dostaneme

$$\begin{aligned}
A \nabla_{a_{kl}} x &= -x_l e_k - x_k e_l \\
\nabla_{a_{kl}} x &= -x_l B e_k - x_k B e_l
\end{aligned}$$

Vytažením l-tého a k-tého řádku dostaneme diferenciální soustavu

$$\begin{aligned}
\frac{\partial}{\partial a_{kl}} x_l &= -x_l B_{lk} - x_k B_{ll} \\
\frac{\partial}{\partial a_{kl}} x_k &= -x_l B_{kk} - x_k B_{kl}
\end{aligned}$$

Pokud máme z předchozích výpočtů vhodné odhady pro x_l, x_k, B_{ll}, B_{lk}, B_{kl} a B_{kk}, můžeme teoreticky dostat nové informace o monotonii $x_l(a_{kl})$ nebo monotonii $x_k(a_{kl})$ a dále postupovat jako v předchozí části.

Můžeme si zde také všimnout, že jsme se omezili na symetrické $A \in A^I$, a proto i jejich inverze $B \in B^I$ nám stačí brát symetrické — to znamená $B_{lk} = B_{kl}$.

Kapitola 6

Příklad

Použijeme stejný příklad, který byl použitý v práci [4].

$$A = \left[\begin{pmatrix} 0,1 & 0,9 \\ 8,9 & 0,4 \\ 0,9 & 6,9 \end{pmatrix}, \begin{pmatrix} 0,3 & 1,1 \\ 9,1 & 0,6 \\ 1,1 & 7,1 \end{pmatrix}\right]$$

$$b = \left[\begin{pmatrix} 0,8 \\ -0,2 \\ 1,8 \end{pmatrix}, \begin{pmatrix} 1,2 \\ 0,2 \\ 2,2 \end{pmatrix}\right]$$

Tedy

$$A_C = \begin{pmatrix} 0,2 & 1,0 \\ 9,0 & 0,5 \\ 1,0 & 7,0 \end{pmatrix}, \quad \Delta = \begin{pmatrix} 0,1 & 0,1 \\ 0,1 & 0,1 \\ 0,1 & 0,1 \end{pmatrix}$$

$$b_C = \begin{pmatrix} 1,0 \\ 0,0 \\ 2,0 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0,2 \\ 0,2 \\ 0,2 \end{pmatrix}$$

Práce [4] uvádí celkem 7 různých řešení, jejichž průnik je roven ohrazení

$$X_1 = \left[\begin{pmatrix} -0,0558 \\ 0,2579 \end{pmatrix}, \begin{pmatrix} 0,0232 \\ 0,3485 \end{pmatrix}\right]$$

Postupem podle kapitoly 4 dostáváme:

$$\tilde{A}_C = \begin{pmatrix} 0 & 0 & 0,2 & 9,0 & 1,0 \\ 0 & 0 & 1,0 & 0,5 & 7,0 \\ 0,2 & 1,0 & -1 & 0 & 0 \\ 9,0 & 0,5 & 0 & -1 & 0 \\ 1,0 & 7,0 & 0 & 0 & -1 \end{pmatrix}, \quad \tilde{\Delta} = \begin{pmatrix} 0 & 0 & 0,1 & 0,1 & 0,1 \\ 0 & 0 & 0,1 & 0,1 & 0,1 \\ 0,1 & 0,1 & 0 & 0 & 0 \\ 0,1 & 0,1 & 0 & 0 & 0 \\ 0,1 & 0,1 & 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{b} = \begin{pmatrix} 0 \\ 0 \\ 1,0 \\ 0,0 \\ 2,0 \end{pmatrix}, \quad \tilde{\delta} = \begin{pmatrix} 0 \\ 0 \\ 0,2 \\ 0,2 \\ 0,2 \end{pmatrix}$$

$$M = \begin{pmatrix} 1,01262533 & 0,01262533 & 0,00165285 & 0,00165285 & 0,00165285 \\ 0,01860471 & 1,01860471 & 0,00249901 & 0,00249901 & 0,00249901 \\ 0,11638913 & 0,11638913 & 1,00257230 & 0,00257230 & 0,00257230 \\ 0,00250460 & 0,00250460 & 0,01321621 & 1,01321621 & 0,01321621 \\ 0,01858355 & 0,01858355 & 0,01544154 & 0,01544154 & 1,01544154 \end{pmatrix}$$

$$x_c = \begin{pmatrix} -0.01629608442 \\ 0.3023017749 \\ -0.700957442 \\ 0.004486127629 \\ 0.09981633974 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0.04690012726 \\ 0.3474509669 \\ 0.9728884069 \\ 0.02093576899 \\ 0.1553385726 \end{pmatrix}$$

$$T_\mu = \begin{pmatrix} 1.01262533 & 0 & 0 & 0 & 0 \\ 0 & 1.01860471 & 0 & 0 & 0 \\ 0 & 0 & 1.00257230 & 0 & 0 \\ 0 & 0 & 0 & 1.01321621 & 0 \\ 0 & 0 & 0 & 0 & 1.01544154 \end{pmatrix}$$

$$T_\nu = \begin{pmatrix} 0.975371228 & 0 & 0 & 0 & 0 \\ 0 & 0.964125452 & 0 & 0 & 0 \\ 0 & 0 & 0.994881735 & 0 & 0 \\ 0 & 0 & 0 & 0.97424827 & 0 \\ 0 & 0 & 0 & 0 & 0.97004212 \end{pmatrix}$$

$$\underset{\sim}{x} = \begin{pmatrix} -0.04690012726 \\ 0.2684010563 \\ -0.9728884069 \\ -0.01184493454 \\ 0.04737674213 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 0.01389647145 \\ 0.3474509669 \\ -0.4326326205 \\ 0.02093576899 \\ 0.1553385726 \end{pmatrix}$$

$$T_\nu \underset{\sim}{x} = \begin{pmatrix} -0.04574503473 \\ 0.2587722898 \\ -0.9679089059 \\ -0.01153990694 \\ 0.04595743534 \end{pmatrix}, \quad T_\nu \tilde{x} = \begin{pmatrix} 0.01355421843 \\ 0.3349863206 \\ -0.4304182919 \\ 0.02039663663 \\ 0.1506849581 \end{pmatrix}$$

Zajímají nás první dvě složky vektorů $\min\{\underset{\sim}{x}, T_\nu \underset{\sim}{x}\}$ a $\max\{\tilde{x}, T_\nu \tilde{x}\}$. Tímto způsobem získáme ohrazení

$$X \subset \left[\begin{pmatrix} -0.04690012726 \\ 0.2587722898 \end{pmatrix}, \begin{pmatrix} 0.01389647145 \\ 0.3474509669 \end{pmatrix} \right]$$

Literatura

[1] Alefeld Götz, Kreinovich Vladik, and Mayer Günter, *On the solution sets of particular classes of linear interval systems*, Journal of Computational and Applied Mathematics **152** (2003), 1–15.

[2] Alefeld Götz and Mayer Günter, *Interval analysis: theory and applications*, Journal of Computational and Applied Mathematics **121** (2000), 421–464.

[3] Anděl Jiří, *Matematická statistika*, SNTL/ALFA, 1978.

[4] Bentbib A. H., *Solving the full rank interval least-square problem*, Applied Numerical Mathematics **41** (2002), 283–294.

[5] Dhompongsa Sompong, Kreinovich Vladik, and Hung T. Nguyen, *Interval Mathematics: Algebraic aspects*, 2001.

[6] Dwyer Paul S., *Linear Computations* (Walter A. Shewhart, ed.), John Wiley & Sons, Inc., 1951.

[7] Hayes Brian, *Lucid Interval*, American Scientist **91/6** (2003), 484–488.

[8] Kearfott R.B., *Interval Computations: Introductions, Uses and Resources*, Department of Mathematics, University of Southwest Lousiana.

[9] Young Rosalind Cecily, *The Algebra of Many-Valued Quantities*, Matematische Annalen, 1931, pp. 261–290.

[10] Loreto Aline B., Toscani Laira V., Robeiro Leila, Dalcídio Cláudio M., Leal Liara S., Longpré Luc, and Kreinovich Vladik, *If an Exact Interval Computation Problem Is NP-Hard, then the Approximate Problem Is Also NP-Hard: A Meta-Result*, 2004.

[11] Meyer Carl D., *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[12] Neumarier Arnold, *A simple derivation of the Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations*, Institut für Mathematik, Universität Wien, 1998.

[13] Rohn J., *Systems of interval linear equations and inequalities (rectangular case)*, Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07, Prague 8, 2002.

[14] ______, *Přehled některých důležitých vět z teorie matic*, Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07, Prague 8, 2003.

[15] Rohn J. and Rex Georg, *Sufficient Conditions for Regularity and Singularity of Interval Matrices*, SIAM Journal on Matrix Analysis and Applications **20** (1998), 437–445.

[16] Shayer Sidney, *Interval Arithmetic with some Applications for Digital Computers*, Lockheed Missiles & Space Company, Lockheed Palo Alto Research Laboratory, 1965.

[17] Stewart G. W. and Sun Ji-guang, *Matrix Perturbation Theory* (Reinboldt Werner and Siewiorek Daniel, eds.), Academic Press, Inc., 1990.

[18] Stolfi Jorge and de Figueiredo Luiz Henrique, *Self-Validated Numerical Methods and Applications*, 1997.

[19] Sunaga Teruo, *Theory of and Interval Algebra and its Application to Numerical Analysis*, Raig Memoirs **2** (1958).

[20] Warmus M., *Calculus of Approximations*, Bulletin de l'académié Polonaise des sciences **4/5** (1956), 253–259.