NUMERICKÉ METODY LINEÁRNÍ ALGEBRY

Dana Černá

Katedra matematiky

Technická univerzita v Liberci

http://kma.fp.tul.cz

2024

Obsah

Kapitola 1

Řešení soustav lineárních algebraických rovnic

V této kapitole se budeme zabývat základními metodami pro řešení soustav lineárních algebraických rovnic. Řešení těchto soustav představuje základní matematický úlohu, která se často vyskytuje v mnoha různých oblastech, například ve statistice, optimalizaci, matematické analýze, numerické matematice a také v dalších disciplínách jako je fyzika, chemie a ekonomie.
Nejprve si nadefinujeme některé základní pojmy a po té se budeme věnovat základním přímým a iteračním metodám pro řešení soustav s regulární maticí.

1.1 Základní pojmy

Vektory a matice jsou charakterizovány pomocí vektorových a maticových norem. Zopakujeme si proto nejprve pojem norma a normovaný lineární prostor. V této kapitole se zaměříme na vektorový prostor nad tělesem reálných čísel, nicméně uvedené definice lze formulovat a dokázat analogicky pro množinu komplexních čísel.

Definice 1. Nezáporná funkce $\|\cdot\|$ definovaná na vektorovém prostoru V nad tělesem $\mathbb{R}$ se nazývá *norma*, jestliže platí

a) $\|x\| = 0 \Leftrightarrow x = 0$,

b) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$,

c) $\|cx\| = |c| \, \|x\| \quad \forall x \in V, \quad \forall c \in \mathbb{R}$.

Dvojici $(V, \|\cdot\|)$ nazýváme *normovaný lineární prostor*.

Z definice plyne, že $\|x\| = \|-x\|$, protože

$$\|x\| = \|(-1)\,(-x)\| = |-1| \, \|-x\| = \|-x\| \,. \tag{1.1}$$

Pokud $(V, \|\cdot\|)$ je normovaný lineární prostor, potom lze snadno ukázat, že $\rho(x,y) = \|x-y\|$ je metrika na V a říkáme, že metrika ρ je indukovaná normou $\|\cdot\|$.
V následujícím textu bude V označovat vektorový prostor, $\|\cdot\|$ bude značit normu na tomto vektorovém prostoru a metrikou budeme rozumět metriku indukovanou touto normou, pokud nebude uvedeno jinak.

Lemma 2. *Pro všechna $x, y \in V$ platí, že $\|x\| - \|y\| \leq \|x-y\|$ a $|\|x\| - \|y\|| \leq \|x-y\|$.*

Důkaz. Z trojúhelníkové nerovnosti, tj. vlastnosti b) v definici normy, plyne, že

$$\|x\| = \|x - y + y\| \leq \|x-y\| + \|y\|. \tag{1.2}$$

Tím je dokázáno první tvrzení. Vzhledem k tomu, že

$$|\|x\| - \|y\|| = \begin{cases} \|x\| - \|y\|, \text{pokud } \|x\| \geq \|y\|, \\ \|y\| - \|x\|, \text{pokud } \|y\| \geq \|x\|, \end{cases} \tag{1.3}$$

a z prvního tvrzení lemmatu plyne, že

$$\|x\| - \|y\| \leq \|x-y\|, \quad \|y\| - \|x\| \leq \|y-x\| = \|x-y\|, \tag{1.4}$$

platí $|\|x\| - \|y\|| \leq \|x-y\|$ pro všechna $x, y \in V$. □

Věta 3. *Jestliže $(V, \|\cdot\|)$ je normovaný lineární prostor, potom norma $\|\cdot\|$ je spojitá funkce na V (vzhledem k metrice indukované touto normou).*

Důkaz. Uvažujme libovolné $x \in V$. Připomeňme, že funkce f je spojitá v bodě x, jestliže ke každému $\epsilon > 0$ existuje $\delta > 0$ takové, že pro každé $y \in V$ splňující $\rho(x,y) < \delta$ platí $|f(x) - f(y)| < \epsilon$.
V našem případě $\rho(x,y) = \|x-y\|$ a $f(x) = \|x\|$. Podle předchozího lemmatu platí $|\|x\| - \|y\|| \leq \|x-y\|$. Z toho plyne, že stačí položit $\delta = \epsilon$ a dostaneme, že pokud $\|x-y\| < \delta = \epsilon$, potom $|\|x\| - \|y\|| \leq \|x-y\| \leq \epsilon$ a norma $\|\cdot\|$ je tedy spojitá. □

Definice 4. Normu definovanou na vektorovém prostoru $\mathbb{R}^n$ nazýváme *vektorová norma.*

Příkladem vektorové normy je *p–norma* definovaná pro $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ předpisem

$$\|\mathbf{x}\|_p = \begin{cases} \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} & \text{pro } p \in \mathbb{N}, \\ \max\limits_{k=1,\ldots,n} |x_k| & \text{pro } p = \infty. \end{cases} \tag{1.5}$$

Takto definovaná 2-norma se také nazývá *euklidovská norma* a tato norma reprezentuje délku vektoru. Norma $\|\cdot\|_\infty$ se také označuje jako *maximová norma*. Snadno se přesvědčíme, že takto definovaná p-norma je skutečně norma, platnost vlastností a) a c) je zřejmá a vlastnost b) pro p-normu je Minkowského nerovnost.

Definice 5. Dvě normy $\|\cdot\|_A$ a $\|\cdot\|_B$ definované na vektorovém prostoru $\mathbb{R}^n$ se nazývají *ekvivalentní*, jestliže existují kladné konstanty c a C takové, že platí

$$c\,\|\mathbf{x}\|_A \leq \|\mathbf{x}\|_B \leq C\,\|\mathbf{x}\|_A \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{1.6}$$

Věta 6. Tranzitivita norem. *Jestliže normy $\|\cdot\|_A$ a $\|\cdot\|_B$ jsou ekvivalentní a zároveň normy $\|\cdot\|_B$ a $\|\cdot\|_C$ jsou ekvivalentní, potom jsou ekvivalentní také normy $\|\cdot\|_A$ a $\|\cdot\|_C$.*

Důkaz. Z předpokladů věty plyne, že existují kladné konstanty c, C, d, D takové, že

$$c\,\|\mathbf{x}\|_B \leq \|\mathbf{x}\|_A \leq C\,\|\mathbf{x}\|_B\,, \quad d\,\|\mathbf{x}\|_B \leq \|\mathbf{x}\|_C \leq D\,\|\mathbf{x}\|_B\,, \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{1.7}$$

což implikuje

$$\|\mathbf{x}\|_C \leq D\,\|\mathbf{x}\|_B \leq \frac{D}{c}\,\|\mathbf{x}\|_A\,, \quad \|\mathbf{x}\|_A \leq C\,\|\mathbf{x}\|_B \leq \frac{C}{d}\,\|\mathbf{x}\|_C\,, \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{1.8}$$

tedy ekvivalenci norem $\|\cdot\|_A$ a $\|\cdot\|_C$. □

Vzhledem k tomu, že vektorový prostor $\mathbb{R}^n$ je konečně-dimenzionální, platí, že všechny normy definované na tomto prostoru jsou ekvivalentní.

Věta 7. *Libovolné dvě normy $\|\cdot\|_A$ a $\|\cdot\|_B$ definované na vektorovém prostoru $\mathbb{R}^n$ jsou ekvivalentní.*

Důkaz. Dokážeme, že libovolná norma $\|\cdot\|_A$ je ekvivalentní s normou $\|\cdot\|_1$. Víme, že jednotkové vektory $\mathbf{e}_k$, $k = 1, \ldots, n$, tj. vektory které mají na k-té pozici 1 a ostatní prvky nulové, tvoří bázi prostoru $\mathbb{R}^n$. Potom

$$\mathbf{x} = (x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k \mathbf{e}_k \tag{1.9}$$

a platí

$$\|\mathbf{x}\|_A \leq \sum_{k=1}^{n} |x_k|\,\|\mathbf{e}_k\|_A \leq \max_{k=1,\ldots,n} \|\mathbf{e}_k\|_A \sum_{k=1}^{n} |x_k| \leq C\,\|\mathbf{x}\|_1\,, \quad C = \max_{k=1,\ldots,n} \|\mathbf{e}_k\|_A\,. \tag{1.10}$$

Nyní uvažujme metrický prostor (V, ρ), kde ρ je metrika indukovaná 1-normou. V tomto prostoru je norma $\|\cdot\|_A$ spojitá, protože platí $|\|\mathbf{x}\|_A - \|\mathbf{y}\|_A| \leq \|\mathbf{x} - \mathbf{y}\|_A \leq C\,\|\mathbf{x} - \mathbf{y}\|_1$. Množina $B = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_1 = 1\}$ je kompaktní (vzhledem k ρ) a funkce $\|\cdot\|_A$ tedy nabývá na B minima. Označme $\mathbf{x}_{min}$ bod tohoto minima a uvažujme $\mathbf{x} \neq \mathbf{0}$. Pro $\mathbf{y} = \mathbf{x}/\,\|\mathbf{x}\|_1$ platí $\|\mathbf{y}\|_1 = 1$ a

$$\|\mathbf{x}\|_A = \|\mathbf{y}\|_A\,\|\mathbf{x}\|_1 \geq c\,\|\mathbf{x}\|_1\,, \quad c = \|\mathbf{x}_{min}\|_A\,. \tag{1.11}$$

Ze vztahů (1.10) a (1.11) plyne ekvivalence 1-normy a libovolné vektorové normy. Normy $\|\cdot\|_A$ a $\|\cdot\|_1$ jsou tedy ekvivalentní a normy $\|\cdot\|_B$ a $\|\cdot\|_1$ jsou ekvivalentní, z čehož dle předchozí věty o tranzitivitě norem plyne i ekvivalence norem $\|\cdot\|_A$ a $\|\cdot\|_B$. □

Kromě vektorových norem budeme pracovat s maticovými normami.

Definice 8. Nezáporná funkce $\|\cdot\|$ definovaná na prostoru $\mathbb{R}^{n\times n}$, tj. prostoru všech reálných matic velikosti $n \times n$, se nazývá *maticová norma*, jestliže jsou splněny vlastnosti a), b) a c) z definice normy a navíc platí

d) $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \, \|\mathbf{B}\| \quad \forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$.

V tomto textu se zaměříme na reálné matice. Pro komplexní matice je maticová norma definovaná analogicky a také většinu prezentovaných vět lze analogicky odvodit i pro komplexní matice.
Uveďme si nyní příklady maticových norem. Pro matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ budeme definovat maticovou normu $\|\cdot\|$ *generovanou vektorovou normou* $\|\cdot\|$ předpisem

$$\|\mathbf{A}\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}. \tag{1.12}$$

Maticovou a vektorovou normu v této definici často značíme stejně, neboť z kontextu je vždy zřejmé, jestli se jedná o maticovou nebo vektorovou normu. Maticovou normu generovanou p-normou nazýváme také *p-norma*.
Dále pro matici $\mathbf{A} = (a_{ij})_{i,j=1}^{n}$ definujme *Frobeniovu (Schurovu) normu* předpisem

$$\|\mathbf{A}\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}. \tag{1.13}$$

Snadno lze ověřit, že takto definované normy splňují podmínky a)-d) v definici normy, a jsou to tedy skutečně maticové normy.
Zaměřme se nyní na maticovou normu generovanou vektorovou normou. Nevýhoda výše uvedené definice je ta, že zde vystupuje supremum přes nekonečnou množinu. Následující věta definici zjednodušuje v tom smyslu, že generovanou normu lze určit také jako maximum na kompaktní množině, což je jednodušší úloha.

Věta 9. *Jestliže $\|\cdot\|$ je maticová norma generovaná vektorovou normou $\|\cdot\|$, potom pro každou matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ platí $\|\mathbf{A}\| = \max\limits_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$.*

Důkaz. Platí

$$\|\mathbf{A}\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\neq\mathbf{0}} \left\| \mathbf{A} \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| = \sup_{\|\mathbf{y}\|=1} \|\mathbf{Ay}\|, \tag{1.14}$$

protože pro $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ je

$$\|\mathbf{y}\| = \left\| \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| = \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|} = 1. \tag{1.15}$$

Vzhledem k tomu, že $f(\mathbf{x}) = \|\mathbf{Ax}\|$ je spojitá funkce a množina

$$B = \{\mathbf{x} : \|\mathbf{x}\| = 1\} \tag{1.16}$$

je kompaktní, nabývá f na B maxima a můžeme tedy supremum na pravé straně rovnice (1.14) nahradit maximem. □

Pokud budeme pracovat s vektorovými a maticovými normami, budeme často potřebovat, aby tyto normy byly v určitém vztahu, který popisuje následující definice.

Definice 10. Řekneme, že maticová norma $\|\cdot\|$ je *konzistentní* s vektorovou normou $\|\cdot\|$, jestliže platí

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n \quad \forall \mathbf{A} \in \mathbb{R}^{n\times n}. \tag{1.17}$$

Pokud navíc ke každé matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ existuje nenulový vektor $\mathbf{x}$ takový, že platí

$$\|\mathbf{Ax}\| = \|\mathbf{A}\| \|\mathbf{x}\|, \tag{1.18}$$

potom řekneme, že maticová norma $\|\cdot\|$ *vyhovuje* vektorové normě $\|\cdot\|$.

Věta 11. *Jestliže $\|\cdot\|$ je maticová norma generovaná vektorovou normou $\|\cdot\|$, potom je tato maticová norma s touto vektorovou normou konzistentní a této normě vyhovuje.*

Důkaz. Uvažujme libovolnou matici $\mathbf{A} \in \mathbb{R}^{n\times n}$. Tato matice a nulový vektor $\mathbf{x} \in \mathbb{R}^n$ vztah (1.17) splňují. Pro nenulový vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \leq \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\|, \tag{1.19}$$

tedy platí (1.17) a tato maticová norma je konzistentní s příslušnou vektorovou normou. Vzhledem k tomu, že $f(\mathbf{x}) = \|\mathbf{Ax}\|$ je spojitá funkce a množina

$$B = \{\mathbf{x} : \|\mathbf{x}\| = 1\} \tag{1.20}$$

je kompaktní, nabývá f na B maxima. Označme ho $\mathbf{x}_{max}$. Platí tedy

$$\|\mathbf{Ax}_{max}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\| = \|\mathbf{A}\| = \|\mathbf{A}\| \|\mathbf{x}_{max}\|. \tag{1.21}$$

Nalezli jsme tedy vektor $\mathbf{x} = \mathbf{x}_{max}$, pro který platí (1.18), z čehož plyne, že maticová norma generovaná normou vektoru této vektorové normě vyhovuje. □

Spektrální poloměr $\rho(\mathbf{A})$ matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je definován jako v absolutní hodnotě největší vlastní číslo matice $\mathbf{A}$, tedy

$$\rho(\mathbf{A}) := \max\{|\lambda| : \lambda \text{ je vlastní číslo matice } \mathbf{A}\}. \tag{1.22}$$

V následujícím textu budeme vždy symbolem a_{ij} značit prvek matice $\mathbf{A}$ na pozici (i, j) a analogicky symbolem x_i značit i-tou složku vektoru $\mathbf{x}$. Některé z maticových norem definovaných vztahem (1.12) můžeme vyjádřit explicitně pomocí následující věty.

Věta 12. *Pro každou matici* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *platí:*

a) $\|\mathbf{A}\|_1 = \max\limits_{j=1,\dots,n} \sum\limits_{i=1}^{n} |a_{ij}|$.

b) $\|\mathbf{A}\|_\infty = \max\limits_{i=1,\dots,n} \sum\limits_{j=1}^{n} |a_{ij}|$.

c) $\|\mathbf{A}\|_2 = \sqrt{\rho\left(\mathbf{A}^T\mathbf{A}\right)}$.

d) Je-li matice $\mathbf{A}$ *symetrická, potom* $\|\mathbf{A}\|_2 = \rho\left(\mathbf{A}\right)$.

Důkaz. a) Pro libovolný vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_1 &= \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \le \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|\,|x_j| \\ &\le \max_{j=1,\dots,n} \left(\sum_{i=1}^{n} |a_{ij}| \right) \sum_{j=1}^{n} |x_j| \le \|\mathbf{x}\|_1 \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|\,. \end{aligned} \tag{1.23}$$

Z toho plyne, že

$$\|\mathbf{A}\|_1 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \le \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|\,. \tag{1.24}$$

Označme p index, ve kterém nastává maximum, tj. pro který platí

$$\sum_{i=1}^{n} |a_{ip}| = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}| \tag{1.25}$$

a označme $\tilde{\mathbf{x}} \in \mathbb{R}^n$ jednotkový vektor takový, že $\tilde{x}_p = 1$ a $\tilde{x}_i = 0$ pro všechna $i \neq p$. Potom platí $\|\tilde{\mathbf{x}}\|_1 = 1$ a

$$\|\mathbf{A}\tilde{\mathbf{x}}\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| = \sum_{i=1}^{n} |a_{ip}| = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|\,. \tag{1.26}$$

Z toho plyne, že

$$\|\mathbf{A}\|_1 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \ge \frac{\|\mathbf{A}\tilde{\mathbf{x}}\|_1}{\|\tilde{\mathbf{x}}\|_1} \ge \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|\,. \tag{1.27}$$

Důsledkem nerovností (1.24) a (1.27) je tvrzení věty 12.
b) Důkaz tvrzení b) je analogický jako důkaz v části a). Pro libovolný vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_\infty &= \max_{i=1,\dots,n} \left| \sum_{k=1}^{n} a_{ik}x_k \right| \le \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,|x_k| \\ &\le \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}| \max_{k=1,\dots,n} |x_k| = \|\mathbf{x}\|_\infty \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,. \end{aligned} \tag{1.28}$$

Z toho plyne, že

$$\|\mathbf{A}\|_\infty = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,. \tag{1.29}$$

Označme p index, ve kterém nastává maximum, tj. pro který platí

$$\sum_{k=1}^{n} |a_{pk}| = \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,. \tag{1.30}$$

a definujme vektor $\tilde{\mathbf{x}}$ předpisem

$$\tilde{x}_k = \begin{cases} 1, & a_{pk} = 0, \\ \frac{|a_{pk}|}{a_{pk}} & a_{pk} \neq 0. \end{cases} \tag{1.31}$$

Potom $\|\tilde{\mathbf{x}}\|_\infty = 1$ a platí

$$\|\mathbf{A}\tilde{\mathbf{x}}\|_\infty \;=\; \max_{i=1,\dots,n} \left| \sum_{k=1}^{n} a_{ik}\tilde{x}_k \right| \geq \sum_{k=1}^{n} |a_{pk}\tilde{x}_k| = \sum_{k=1}^{n} |a_{pk}| = \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,, \tag{1.32}$$

z čehož plyne, že

$$\|\mathbf{A}\|_\infty = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \geq \frac{\|\mathbf{A}\tilde{\mathbf{x}}\|_\infty}{\|\tilde{\mathbf{x}}\|_\infty} \geq \max_{i=1,\dots,n} \sum_{k=1}^{n} |a_{ik}|\,. \tag{1.33}$$

c) Matice $\mathbf{A}^T\mathbf{A}$ je symetrická, neboť $\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\mathbf{A}$. Uvažujme libovolný vektor $\mathbf{x} \in \mathbb{R}^n$. Vzhledem k tomu, že norma je nezáporná funkce a euklidovskou normu můžeme vyjádřit pomocí skalárního součinu, dostaneme

$$0 \leq \|\mathbf{A}\mathbf{x}\|_2^2 = (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}. \tag{1.34}$$

Matice $\mathbf{A}^T\mathbf{A}$ je tedy pozitivně semidefinitní. Z toho plyne, že její vlastní čísla $\lambda_1, \dots, \lambda_n$ jsou reálná a nezáporná a existují vlastní vektory $\mathbf{u}_1, \dots, \mathbf{u}_n$ příslušné těmto vlastním číslům, které tvoří ortonormální bázi. Označme koeficienty vektoru $\mathbf{x}$ v této bázi $c_1, \dots, c_n$. Vzhledem k ortonormalitě báze platí, že $\mathbf{u}_i^T\mathbf{u}_j = 0$ pro $i \neq j$ a $\mathbf{u}_i^T\mathbf{u}_i = 1$, a proto

$$\|\mathbf{x}\|_2^2 = \mathbf{x}^T\mathbf{x} = \left(\sum_{i=1}^{n} c_i\mathbf{u}_i\right)^T \left(\sum_{j=1}^{n} c_j\mathbf{u}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_ic_j\mathbf{u}_i^T\mathbf{u}_j = \sum_{i=1}^{n} c_i^2. \tag{1.35}$$

Pro 2-normu dostaneme vztah

$$\|\mathbf{A}\mathbf{x}\|_2^2 \;=\; (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \left(\sum_{i=1}^{n} c_i\mathbf{u}_i\right)^T \mathbf{A}^T\mathbf{A} \left(\sum_{j=1}^{n} c_j\mathbf{u}_j\right) \tag{1.36}$$

$$= \sum_{i=1}^{n} c_ic_j\mathbf{u}_i^T\mathbf{A}^T\mathbf{A}\mathbf{u}_j = \sum_{i=1}^{n} c_ic_j\lambda_j\mathbf{u}_i^T\mathbf{u}_j = \sum_{i=1}^{n} c_i^2\lambda_i \tag{1.37}$$

$$\leq \rho\left(\mathbf{A}^T\mathbf{A}\right) \sum_{i=1}^{n} c_i^2 = \rho\left(\mathbf{A}^T\mathbf{A}\right) \|\mathbf{x}\|_2^2\,. \tag{1.38}$$

Maticová 2-norma tedy splňuje

$$\|\mathbf{A}\|_2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \sqrt{\rho\left(\mathbf{A}^T\mathbf{A}\right)}. \tag{1.39}$$

Pro maximální vlastní číslo $\tilde{\lambda}$ a jemu příslušný vlastní vektor $\tilde{\mathbf{x}}$ platí

$$\|\mathbf{A}\tilde{\mathbf{x}}\|_2^2 = \tilde{\mathbf{x}}^T\mathbf{A}^T\mathbf{A}\tilde{\mathbf{x}} = \tilde{\lambda}\,\|\tilde{\mathbf{x}}\|_2^2\,. \tag{1.40}$$

To implikuje nerovnost

$$\|\mathbf{A}\|_2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \geq \frac{\|\mathbf{A}\tilde{\mathbf{x}}\|_2}{\|\tilde{\mathbf{x}}\|_2} \geq \sqrt{\rho\left(\mathbf{A}^T\mathbf{A}\right)}. \tag{1.41}$$

d) Symetrická matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ má reálná vlastní čísla $\lambda_1, \ldots, \lambda_n$ a příslušné vlastní vektory $\mathbf{u}_1, \ldots, \mathbf{u}_n$ jsou lineárně nezávislé. Pro $i = 1, \ldots, n$ dostaneme

$$\mathbf{A}^T\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{A}^T\mathbf{u}_i = \lambda_i\mathbf{A}\mathbf{u}_i = \lambda_i^2\mathbf{u}_i \tag{1.42}$$

a tedy $\rho\left(\mathbf{A}\right) = \sqrt{\rho\left(\mathbf{A}^T\mathbf{A}\right)}$.

□

Norma $\|\cdot\|_2$ tedy na základě vlastností c) a d) souvisí se spektrálním poloměrem matice a proto se často nazývá *spektrální norma*. Následující věta uvádí nutnou podmínku pro to, aby určitá maticová norma vyhovovala nějaké normě vektoru. Symbol $\mathbf{I}$ označuje jednotkovou matici.

Věta 13. *Jestliže $\|\cdot\|$ je maticová norma, která vyhovuje nějaké normě vektoru, potom $\|\mathbf{I}\| = 1$.*

Důkaz. Pokud maticová norma $\|\cdot\|$ vyhovuje normě vektoru $\|\cdot\|$, potom existuje vektor $\mathbf{x} \neq \mathbf{0}$ pro který platí

$$\|\mathbf{x}\| = \|\mathbf{I}\mathbf{x}\| = \|\mathbf{I}\|\,\|\mathbf{x}\|\,. \tag{1.43}$$

Z toho plyne, že $\|\mathbf{I}\| = 1$. □

Pomocí této věty lze snadno ukázat, že Frobeniova norma nevyhovuje žádné normě vektoru. Platí totiž, že $\|\mathbf{I}\|_F = \sqrt{n} \neq 1$.

Věta 14. *Ke každé maticové normě existuje norma vektoru, která je s ní konzistentní.*

Důkaz. Uvažujme maticovou normu $\|\cdot\|$. Pro $\mathbf{x} \in \mathbb{R}^n$ definujme vektorovou normu vztahem $\|\mathbf{x}\| = \|\mathbf{G}\|$, kde $\mathbf{G} \in \mathbb{R}^{n\times n}$ je matice, jejíž první sloupec je vektor $\mathbf{x}$ a ostatní sloupce jsou nulové. Potom pro libovolnou matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ je $\mathbf{AG}$ matice, jejíž první sloupec je $\mathbf{Ax}$ a ostatní sloupce jsou nulové a platí

$$\|\mathbf{A}\mathbf{x}\| = \|\mathbf{A}\mathbf{G}\| \leq \|\mathbf{A}\|\,\|\mathbf{G}\| = \|\mathbf{A}\|\,\|\mathbf{x}\|\,. \tag{1.44}$$

Zkonstruovali jsme tedy vektorovou normu, která je konzistentní s uvažovanou maticovou normou. □

Zkonstruujeme-li vektorovou normu konzistentní s Frobeniovou normou pomocí postupu uvedeného v důkazu této věty, vyjde nám, že Frobeniova norma je konzistentní s euklidovskou normou, tj. vektorovou normou $\|\cdot\|_2$.
Důsledkem této věty je dále následující věta o odhadu spektrálního poloměru.

Věta 15. *Pro každou matici* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *a libovolnou normu matice* $\|\cdot\|$ *platí* $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$.

Důkaz. K dané maticové normě $\|\cdot\|$ existuje dle předchozí věty vektorová norma $\|\cdot\|$, která je s ní konzistentní. Uvažujme vlastní číslo λ matice $\mathbf{A}$ a vlastní vektor $\mathbf{u}$ příslušný tomuto vlastnímu číslu, tj. platí $\mathbf{Au} = \lambda\mathbf{u}$. Potom platí

$$|\lambda| \|\mathbf{u}\| = \|\lambda\mathbf{u}\| = \|\mathbf{Au}\| \leq \|\mathbf{A}\| \|\mathbf{u}\|. \tag{1.45}$$

Po vydělení tohoto vztahu $\|\mathbf{u}\|$ dostaneme $|\lambda| \leq \|\mathbf{A}\|$ a tedy $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$. □

1.2 Podmíněnost matice

Budeme se zabývat numerickým řešením soustavy m rovnic s n neznámými:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2, \\ \vdots & & & & & & \vdots & & \vdots, \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m, \end{array} \tag{1.46}$$

kde $a_{ij}, b_i \in \mathbb{R}$ pro $i = 1, \dots, m$, $i = 1, \dots, n$. Tuto soustavu můžeme zapsat ve tvaru

$$\mathbf{Ax} = \mathbf{b}, \tag{1.47}$$

kde

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Matice $\mathbf{A}$ se nazývá *matice soustavy*, vektor $\mathbf{x}$ se nazývá *vektor neznámých*, vektor $\mathbf{b}$ se nazývá *vektor pravé strany*.
Matice soustavy, která má mnohem více nulových prvků než nenulových prvků, se nazývá *řídká*. Matice, která není řídká, se nazývá *plná*.
Ještě předtím než uvedeme metody pro řešení soustav, zaměříme se na čtvercové matice a seznámíme se s problémem podmíněnosti matice. Uvažujme následující příklad.

Příklad 16. a) Soustava
$$\begin{aligned} x &+ y &&= 2 \\ x &+ 1.00000001y &&= 2.00000001 \end{aligned}$$
má řešení $x = 1$, $y = 1$.

b) Soustava
$$\begin{aligned} x &+ y &&= 2 \\ x &+ 1.00000001y &&= 2.00000002 \end{aligned}$$
má řešení $x = 0$, $y = 2$.

c) Soustava
$$\begin{aligned} x &+ y &&= 2 \\ x &+ y &&= 2 \end{aligned}$$
má nekonečně mnoho řešení $x = t$, $y = 2 - t$, $t \in \mathbb{R}$.

Vidíme, že zde velmi malé změny v koeficientech matice a vektoru pravé strany vedly k velkým změnám v řešení. Pokud je matice a vektor pravé strany zadán nepřesně, což je například v případě, že byly získány pomocí měření nebo numerických výpočtů, potom vypočtené řešení může být velmi odlišné od přesného řešení přesně zadané soustavy. Taková situace může nastat u špatně podmíněných matic.
Řekneme, že matice je *špatně podmíněná*, jestliže relativně malé změny prvků matice nebo vektoru pravé strany způsobí relativně velké změny v řešení soustavy. Řekneme, že matice je *dobře podmíněná*, jestliže relativně malé změny prvků matice a vektoru pravé strany způsobí relativně malé změny v řešení.
Nyní se budeme zabývat závislostí relativní chyby řešení soustavy na relativní chybě v zadání matice soustavy a vektoru pravé strany. K odvození vztahu popisujícímu tuto závislost je potřeba následující lemma.

Lemma 17. *Předpokládejme, že $\|\cdot\|$ je maticová norma generovaná normou vektoru $\|\cdot\|$ a $\mathbf{A} \in \mathbb{R}^{n\times n}$. Pokud $\|\mathbf{A}\| < 1$, potom matice $\mathbf{I} + \mathbf{A}$ je regulární a platí*
$$\left\|(\mathbf{I}+\mathbf{A})^{-1}\right\| \leq \frac{1}{1-\|\mathbf{A}\|}, \quad \left\|\mathbf{I}-(\mathbf{I}+\mathbf{A})^{-1}\right\| \leq \frac{\|\mathbf{A}\|}{1-\|\mathbf{A}\|}. \tag{1.48}$$

Důkaz. Nejprve dokážeme sporem, že $\mathbf{I}+\mathbf{A}$ je regulární. Předpokládejme tedy, že $\|\mathbf{A}\| < 1$ a matice $\mathbf{I}+\mathbf{A}$ je singulární. Potom soustava $(\mathbf{I}+\mathbf{A})\,\mathbf{x} = 0$ má nenulové řešení $\mathbf{x}$. Pro $\mathbf{x} \neq \mathbf{0}$ tedy platí, že
$$\mathbf{A}\mathbf{x} = -\mathbf{I}\mathbf{x} = -\mathbf{x}. \tag{1.49}$$
Z toho plyne, že $\lambda = -1$ je vlastní číslo matice $\mathbf{A}$. Z definice spektrálního poloměru a věty 15 plyne, že $|\lambda| = 1 \leq \rho(\mathbf{A}) \leq \|\mathbf{A}\|$. To je ve sporu s předpokladem $\|\mathbf{A}\| < 1$.
Označme nyní $\mathbf{B} = (\mathbf{I}+\mathbf{A})^{-1}$. Vynásobením tohoto vztahu zprava maticí $\mathbf{I}+\mathbf{A}$ dostaneme $\mathbf{B}+\mathbf{B}\mathbf{A} = \mathbf{I}$. Platí
$$\|\mathbf{B}\| = \|\mathbf{B}+\mathbf{B}\mathbf{A}-\mathbf{B}\mathbf{A}\| \leq \|\mathbf{B}+\mathbf{B}\mathbf{A}\| + \|\mathbf{B}\mathbf{A}\| \Rightarrow \|\mathbf{B}+\mathbf{B}\mathbf{A}\| \geq \|\mathbf{B}\| - \|\mathbf{B}\mathbf{A}\|, \tag{1.50}$$

a tedy

$$1 = \|\mathbf{I}\| = \|\mathbf{B} + \mathbf{BA}\| \geq \|\mathbf{B}\| - \|\mathbf{BA}\| \geq \|\mathbf{B}\| - \|\mathbf{B}\| \, \|\mathbf{A}\| \, . \tag{1.51}$$

Vydělením tohoto vztahu výrazem $1 - \|\mathbf{A}\|$ dostaneme první ze vztahů uvedených v (1.48). Dále platí, že

$$\|\mathbf{I} - \mathbf{B}\| = \|\mathbf{BA}\| \leq \|\mathbf{B}\| \, \|\mathbf{A}\| \leq \frac{\|\mathbf{A}\|}{1 - \|\mathbf{A}\|}. \tag{1.52}$$

Tím je dokázán druhý vztah v (1.48). □

Nyní již můžeme uvést větu o podmíněnosti matice.

Věta 18. *Předpokládejme, že $\|\cdot\|$ je maticová norma generovaná vektorovou normou $\|\cdot\|$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ je regulární matice a $\Delta\mathbf{A}$ je matice taková, že $\|\mathbf{A}^{-1}\Delta\mathbf{A}\| < 1$. Potom platí*

$$\frac{\left\|\mathbf{x}^{\Delta} - \mathbf{x}^{*}\right\|}{\|\mathbf{x}^{*}\|} \leq \frac{\|\mathbf{A}\| \, \|\mathbf{A}^{-1}\|}{1 - \|\mathbf{A}^{-1}\Delta\mathbf{A}\|} \left(\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \right), \tag{1.53}$$

přičemž $\mathbf{x}^{*}$ je řešení soustavy $\mathbf{Ax} = \mathbf{b}$ a $\mathbf{x}^{\Delta}$ je řešení soustavy $(\mathbf{A} + \Delta\mathbf{A})\,\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$.

Důkaz. Z předpokladu věty a předchozího lemmatu plyne, že matice $\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}$ je regulární. Matice $\mathbf{A} + \Delta\mathbf{A} = \mathbf{A}\,(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A})$ je proto také regulární. Platí

$$\begin{aligned}
\mathbf{x}^{*} - \mathbf{x}^{\Delta} &= \mathbf{x}^{*} - (\mathbf{A} + \Delta\mathbf{A})^{-1} (\mathbf{b} + \Delta\mathbf{b}) \\
&= \mathbf{x}^{*} - \left(\mathbf{A}\left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)\right)^{-1} (\mathbf{b} + \Delta\mathbf{b}) \\
&= \mathbf{x}^{*} - \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{A}^{-1} (\mathbf{b} + \Delta\mathbf{b}) \\
&= \mathbf{x}^{*} - \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{A}^{-1}\mathbf{b} + \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{A}^{-1}\Delta\mathbf{b} \\
&= \mathbf{x}^{*} - \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{x}^{*} + \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{A}^{-1}\Delta\mathbf{b} \\
&= \left(\mathbf{I} - \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1}\right) \mathbf{x}^{*} + \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1} \mathbf{A}^{-1}\Delta\mathbf{b}.
\end{aligned} \tag{1.54}$$

Z toho plyne, že

$$\left\|\mathbf{x}^{*} - \mathbf{x}^{\Delta}\right\| \leq \left\|\mathbf{I} - \left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1}\right\| \|\mathbf{x}^{*}\| + \left\|\left(\mathbf{I} + \mathbf{A}^{-1}\Delta\mathbf{A}\right)^{-1}\right\| \left\|\mathbf{A}^{-1}\right\| \|\Delta\mathbf{b}\| \, . \tag{1.55}$$

Nyní použijeme předchozí lemma a dostaneme

$$\left\|\mathbf{x}^{*} - \mathbf{x}^{\Delta}\right\| \leq \frac{\|\mathbf{A}^{-1}\Delta\mathbf{A}\|}{1 - \|\mathbf{A}^{-1}\Delta\mathbf{A}\|} \|\mathbf{x}^{*}\| + \frac{\|\mathbf{A}^{-1}\|}{1 - \|\mathbf{A}^{-1}\Delta\mathbf{A}\|} \|\Delta\mathbf{b}\| \, . \tag{1.56}$$

Ze vztahu $\mathbf{Ax}^{*} = \mathbf{b}$ dostaneme $\|\mathbf{b}\| \leq \|\mathbf{A}\| \, \|\mathbf{x}^{*}\|$ a tedy

$$\frac{1}{\|\mathbf{x}^{*}\|} \leq \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|}. \tag{1.57}$$

Vztah (1.56) vydělíme $\|\mathbf{x}^*\|$, využijeme (1.57) a dostaneme

$$\begin{aligned}\frac{\left\|\mathbf{x}^* - \mathbf{x}^{\Delta}\right\|}{\|\mathbf{x}^*\|} &\leq \frac{\|\mathbf{A}^{-1}\|\,\|\Delta\mathbf{A}\|}{1-\|\mathbf{A}^{-1}\Delta\mathbf{A}\|}\frac{\|\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\mathbf{A}^{-1}\|}{1-\|\mathbf{A}^{-1}\Delta\mathbf{A}\|}\|\Delta\mathbf{b}\|\frac{\|\mathbf{A}\|}{\|\mathbf{b}\|} \\ &= \frac{\|\mathbf{A}\|\,\|\mathbf{A}^{-1}\|}{1-\|\mathbf{A}^{-1}\Delta\mathbf{A}\|}\left(\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}\right).\end{aligned} \tag{1.58}$$

Tím je věta dokázána. □

Ve větě 18 se nejčastěji uvažuje spektrální norma $\|\cdot\|_2$. Na levé straně odhadu (1.53) se vyskytuje relativní chyba řešení soustavy, na pravé straně potom relativní chyba v zadání matice soustavy $\|\Delta\mathbf{A}\|_2 / \|\mathbf{A}\|_2$, relativní chyba v zadání vektoru pravé strany $\|\Delta\mathbf{b}\|_2 / \|\mathbf{b}\|_2$ a konstanta $\|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$, která charakterizuje změny v řešení soustavy. Tato konstanta se nazývá *číslo podmíněnosti* matice $\mathbf{A}$ a značí se $\operatorname{cond}\mathbf{A}$, tj.

$$\operatorname{cond}\mathbf{A} = \|\mathbf{A}\|_2 \left\|\mathbf{A}^{-1}\right\|_2. \tag{1.59}$$

Platí, že $\operatorname{cond}\mathbf{A} \geq 1$. Pokud je číslo podmíněnosti $\operatorname{cond}\mathbf{A}$ velké, je matice $\mathbf{A}$ špatně podmíněná. Pokud je toto číslo malé, je matice dobře podmíněná.

1.3 Přímé metody pro řešení soustav

Nyní se budeme věnovat řešení soustav lineárních algebraických rovnic. V této kapitole budeme uvažovat pouze soustavy se čtvercovou regulární maticí s reálnými koeficienty a reálným vektorem pravé strany. Metody pro řešení takových soustav se dělí na *přímé* a *iterační*. Přímé metody jsou takové metody, které po konečně mnoha krocích dávají přesné řešení. To nastane ale pouze tehdy, když počítáme přesně. Realizujeme-li je numericky, je výsledné řešení zatíženo zaokrouhlovacími chybami. U metod iteračních konstruujeme posloupnost vektorů, která konverguje k přesnému řešení. Toto dělení je spíše tradiční, existují také metody, např. metoda sdružených gradientů, které byly odvozeny jako přímé, ale používají se jako iterační.
Přímé metody se používají především pro soustavy s malou plnou maticí a v některých speciálních případech také pro soustavy s velkou řídkou maticí. Při použití pro velké řídké matice je však třeba dbát na to, aby při použití těchto metod nedocházelo k zaplnění matice, tedy aby z velkého množství nulových prvků nevznikaly nenulové a řídká matice se nestala plnou. Iterační metody se používají spíše pro soustavy s velkými řídkými maticemi.

1.3.1 Gaussova eliminace

Uvažujme soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A}$ je čtvercová regulární matice řádu n. Po konečném počtu kroků dostaneme pomocí známých ekvivalentních úprav soustavu s horní trojúhelníkovou maticí, která je ekvivalentní s původní soustavou. Tato část algoritmu se nazývá

přímý chod Gaussovy eliminace. *Zpětný chod* Gaussovy eliminace spočívá v řešení trojúhelníkové soustavy.

Přímý chod Gaussovy eliminace:

for $k = 1, \ldots, n-1$
    for $i = k+1, \ldots, n$
        $d = \dfrac{a_{ik}}{a_{kk}}$
        for $j = k, \ldots, n$
            $a_{ij} = a_{ij} - d\, a_{kj}$
        end
        $b_i = b_i - a_{ik} b_k$
    end
end

Zpětný chod Gaussovy eliminace:

for $i = n, \ldots, 1$

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^{n} a_{ij} x_j \right)$$

end

Gaussova eliminace v této podobě je pro obecné matice numericky nestabilní. Zaokrouhlovací chyby mohou velmi negativně ovlivnit výsledek. Ke kumulaci zaokrouhlovacích chyb může dojít zejména tehdy, dělíme-li čísly, která jsou v absolutní hodnotě hodně malá. Z tohoto důvodu je Gaussova eliminační metoda prováděna s použitím pivotace, která spočívá v prohození řádků tak, abychom při výpočtu koeficientu d nedělili malým číslem.
Prvek a_{kk}, který je v k-tém kroku Gaussovy eliminace na pozici (k, k) a kterým dělíme, se nazývá *pivot*. Pokud je tento prvek malý, může dojít k velké zaokrouhlovací chybě. Proto provedeme pivotaci, tedy mezi prvky na pozicích $(k, k), \ldots, (n, k)$ nalezneme prvek, který je v absolutní hodnotě největší. Řádek s tímto prvkem vyměníme s k-tým řádkem. Tento proces se nazývá *částečná* nebo také *sloupcová pivotace*. Gaussova eliminace s *úplnou pivotací*, při které se vyměňují nejen řádky ale také sloupce, je časově náročnější a proto se tolik nepoužívá.
Gaussova eliminace matice řádu n vyžaduje celkem $\frac{2n^3}{3} + \mathcal{O}(n^2)$ operací s plovoucí řádovou čárkou.
Modifikací Gaussovy eliminace je Gaussova-Jordanova eliminace, která převádí matici soustavy pomocí elementárních úprav na diagonální matici. Celkový počet operací je v tomto případě $n^3 + \mathcal{O}(n^2)$, proto je pro velká n výhodnější použít k řešení soustav Gaussovu eliminaci než určovat inverzní matici.
Uvažujme nyní Gaussovu eliminaci pro tzv. třídiagonální matici. Třídiagonální matice je

matice, kde $a_{ij} = 0$ pro $|i - j| > 1$. Má tedy tvar

$$\mathbf{A} = \begin{pmatrix} s_1 & t_1 & 0 & 0 & \dots & 0 \\ r_2 & s_2 & t_2 & 0 & & 0 \\ 0 & r_3 & s_3 & t_3 & & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & & 0 & r_{n-1} & s_{n-1} & t_{n-1} \\ 0 & \dots & 0 & 0 & r_n & s_n \end{pmatrix}.$$

Pokud Gaussovu eliminaci provádíme tak, abychom neupravovali prvky pod diagonálou, které jsou již nulové, potom Gaussova eliminace pro třídiagonální matice vyžaduje pouze $\mathcal{O}(n)$ operací s plovoucí řádovou čárkou.
Gaussova eliminace je tedy vhodná k řešení soustav s malými plnými maticemi a některými speciálními typy řídkých matic, například třídiagonální matice. Při použití Gaussovy eliminace na obecné velké řídké matice je třeba dbát na to, aby z nulových prvků nevznikaly nenulové. Řídká matice by se pak stala plnou. Tím pádem bychom k jejímu uložení potřebovali mnohem více paměti a kromě toho by metoda mohla být velmi časově náročná. Někdy se proto provádí permutace řádků a sloupců tak, aby nenulové prvky byly na vhodných pozicích, například blízko diagonály, a při eliminace nedocházelo k zaplnění matice.

1.3.2 LU rozklad

Pokud řešíme více soustav se stejnou maticí a různými pravými stranami je výhodné rozložit matici soustavy $\mathbf{A}$ na součin $\mathbf{A} = \mathbf{LU}$, kde matice $\mathbf{L}$ je dolní trojúhelníková matice s jedničkami na diagonále a $\mathbf{U}$ je horní trojúhelníková matice. Dolní trojúhelníková matice je matice, která má nulové prvky nad hlavní diagonálou, tedy na pozicích (i, j), kde $i < j$. Horní trojúhelníkovou maticí rozumíme matici, která má nulové prvky pod hlavní diagonálou, tedy na pozicích (i, j), kde $i < j$. Matici $\mathbf{U}$ získáme pomocí Gaussovy eliminace, zatímco matice $\mathbf{L}$ obsahuje s opačným znaménkem koeficienty, kterými násobíme řádky při odečítání od jiných řádků při eliminaci. Soustavu $\mathbf{Ax} = \mathbf{b}$ můžeme zapsat ve tvaru $\mathbf{LUx} = \mathbf{b}$. Označme $\mathbf{y} = \mathbf{Ux}$. Potom řešení soustavy pomocí LU-rozkladu spočívá v řešení dvou trojúhelníkových soustav: $\mathbf{Ly} = \mathbf{b}$ a $\mathbf{Ux} = \mathbf{y}$.
Odvození LU rozkladu je založeno na Gaussově eliminaci, protože odečtení $l_{j,i}$ násobku j-tého řádku od i-tého řádku matice $\mathbf{A}$ lze reprezentovat vynásobením matice $\mathbf{A}$ maticí $\mathbf{L}_{j,i}$, která má na diagonále jedničky, na pozici (j, i) má prvek $-l_{j,i}$ a ostatní prvky má nulové. Platí tedy:

$$\mathbf{L}_{j,i}\mathbf{A} = \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ 0 & 1 & & & & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & -l_{j,i} & \dots & 1 \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{j,1} & a_{j,2} & \dots & a_{j,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{j,1} - l_{j,i}a_{i,1} & a_{j,2} - l_{j,i}a_{i,2} & \dots & a_{j,n} - l_{j,i}a_{i,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Koeficient $l_{j,i}$ je zvolen tak, aby po vynásobení maticí $\mathbf{L}_{j,i}$ byla ve výsledné matici na pozici (j, i) nula. Dostaneme

$$\mathbf{L}_{n,n-1} \dots \mathbf{L}_{3,1}\mathbf{L}_{2,1}\mathbf{A} = \mathbf{U},$$

kde $\mathbf{U}$ je horní trojúhelníková matice. Vyjádříme $\mathbf{A}$:

$$\mathbf{A} = \mathbf{L}_{2,1}^{-1}\mathbf{L}_{3,1}^{-1} \dots \mathbf{L}_{n,n-1}^{-1}\mathbf{U},$$

a položíme $\mathbf{L} = \mathbf{L}_{2,1}^{-1}\mathbf{L}_{3,1}^{-1} \dots \mathbf{L}_{n,n-1}^{-1}$. Matice $\mathbf{L}$ má tvar

$$\begin{pmatrix} 1 & 0 & \dots & & 0 \\ l_{2,1} & 1 & & & 0 \\ l_{3,1} & l_{3,2} & 1 & & \\ \vdots & & & & \\ l_{n,1} & l_{n,2} & \dots & l_{n,n-1} & 1 \end{pmatrix}$$

a splňuje $\mathbf{A} = \mathbf{LU}$.

Níže je uvedena varianta algoritmu přímého chodu Gaussovy eliminace, kde jsou koeficienty ukládány pod diagonálu. Díky tomu jsou po skončení eliminace pod diagonálou prvky matice $\mathbf{L}$. Na diagonále a nad diagonálou jsou prvky matice $\mathbf{U}$. Pomocí tohoto algoritmu tedy dostaneme LU rozklad.

LU rozklad pomocí Gaussovy eliminace
for $k = 1, \dots, n-1$

```
    for i = k + 1, ..., n
        a_ik = a_ik / a_kk
        for j = k + 1, ..., n
            a_ij = a_ij − a_ik a_kj
        end
        b_i = b_i − a_ik b_k
    end
end
```

Tato metoda pro řešení jedné soustavy je podobně časově náročná jako Gaussova eliminace, vyžaduje $\frac{2n^3}{3} + \mathcal{O}(n^2)$ operací s plovoucí řádovou čárkou. Při řešení více soustav se stejnou maticí a různými pravými stranami provádíme přímý chod Gaussovy eliminace pouze jednou a pro každou pravou stranu pak provádíme dvakrát zpětný chod, který vyžaduje pouze $\mathcal{O}(n^2)$ operací. Dochází tedy k významné redukci výpočtového času.
Kromě uvedeného postupu lze LU rozklad získat také například pomocí Doolitlova nebo Crootova algoritmu.

Pro obecné matice LU rozklad nemusí existovat. Pokud je však matice regulární, potom je vždy možné změnit pořadí řádků této matice tak, aby LU rozklad existoval.

Také LU rozklad většinou provádíme s pivotací. Potom má rozklad tvar $\mathbf{A} = \mathbf{PLU}$, kde $\mathbf{P}$ je permutační matice, tj. matice, která má v každém řádku a sloupci právě jednu jedničku, jinak pouze nuly. Vynásobení permutační matici je ekvivalentní změně pořadí řádků.

1.3.3 Choleského rozklad

Pokud je matice soustavy $\mathbf{A} \in \mathbb{R}^{n\times n}$ symetrická a pozitivně definitní, potom lze k řešení soustavy použít Choleského rozklad, který má tvar $\mathbf{A} = \mathbf{L}\mathbf{L}^T$, kde $\mathbf{L} \in \mathbb{R}^{n\times n}$ je dolní trojúhelníková matice. Choleského rozklad symetrické pozitivně definitní matice existuje, ale není určen jednoznačně.

Věta 19. *Ke každé symetrické pozitivně definitní matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ existuje dolní trojúhelníková matice $\mathbf{L} \in \mathbb{R}^{n\times n}$ taková, že platí $\mathbf{A} = \mathbf{L}\mathbf{L}^T$.*

Důkaz. Dokážeme navíc, že existuje taková matice $\mathbf{L}$ s kladnými prvky na diagonále. Budeme postupovat indukcí. Pro $n = 1$ je $\mathbf{A} = (a_{11})$ a vzhledem k tomu, že $\mathbf{A}$ je pozitivně definitní platí $a_{11} > 0$. Stačí tedy položit $l_{11} = \sqrt{a_{11}}$.
Nyní předpokládejme, že ke každé symetrické pozitivně definitní matici $\tilde{\mathbf{A}} \in \mathbb{R}^{n\times n}$ existuje dolní trojúhelníková matice $\tilde{\mathbf{L}} \in \mathbb{R}^{n\times n}$ s kladnými prvky na diagonále splňující $\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T$.
Zvolme libovolnou symetrickou pozitivně definitní matici $\mathbf{A}$ řádu $n + 1$ a zapišme ji ve tvaru

$$\mathbf{A} = \begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{a} \\ \mathbf{a}^T & a_{nn} \end{pmatrix}, \tag{1.60}$$

kde $\tilde{\mathbf{A}} \in \mathbb{R}^{n\times n}$ a $\mathbf{a} \in \mathbb{R}^n$. Potom $\tilde{\mathbf{A}}$ je pozitivně definitní a má Choleského rozklad $\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T$.

Nyní budeme hledat vektor $\mathbf{b}$ a číslo α tak, aby platilo

$$\mathbf{A} = \begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{a} \\ \mathbf{a}^T & a_{nn} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{L}} & \mathbf{0} \\ \mathbf{b}^T & \alpha \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{L}} & \mathbf{b} \\ \mathbf{0} & \alpha \end{pmatrix}^T = \begin{pmatrix} \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T & \tilde{\mathbf{L}}\mathbf{b} \\ \mathbf{b}^T\tilde{\mathbf{L}}^T & \mathbf{b}^T\mathbf{b} - \alpha^2 \end{pmatrix}. \tag{1.61}$$

Dostali jsme rovnice $\tilde{\mathbf{L}}\mathbf{b} = \mathbf{a}$ a $\mathbf{b}^T\mathbf{b} - \alpha^2 = a_{nn}$. Matice $\tilde{\mathbf{L}}$ je dolní trojúhelníková s kladnými prvky na diagonále, je tedy regulární a platí $\mathbf{b} = \tilde{\mathbf{L}}^{-1}\mathbf{a}$. Číslo $\alpha \in \mathbb{C}$ lze určit ze vztahu $\alpha^2 = \mathbf{b}^T\mathbf{b} - a_{nn}$ a tedy existuje. Zbývá ukázat, že α lze volit reálné kladné. Ze vztahu (1.61) a pozitivní definitnosti matice $\mathbf{A}$ plyne, že

$$0 < \det \mathbf{A} = \left(\det \tilde{\mathbf{L}}\right)^2 \alpha^2. \tag{1.62}$$

Přitom $\det \tilde{\mathbf{L}} > 0$, protože je tento determinant roven součinu prvků na diagonále matice $\tilde{\mathbf{L}}$ a tyto diagonální prvky jsou dle indukčního předpokladu kladné. Tím pádem $\alpha^2 > 0$ a α lze tedy skutečně volit reálné kladné. Nalezli jsme tedy reálný vektor $\mathbf{b}$ a kladné číslo α splňující (1.61) a tím jsme určili Choleského rozklad matice $\mathbf{A}$. □

Zkusme nyní určit Choleského rozklad pro symetrickou pozitivně definitní matici $\mathbf{A}$ velikosti 3×3. Součin $\mathbf{A} = \mathbf{L}\mathbf{L}^T$ rozepíšeme po složkách:

$$\begin{aligned} \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{2,1} & a_{2,2} & a_{3,2} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} &= \begin{pmatrix} l_{1,1} & 0 & 0 \\ l_{2,1} & l_{2,2} & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} \end{pmatrix} \begin{pmatrix} l_{1,1} & l_{2,1} & l_{3,1} \\ 0 & l_{2,2} & l_{3,2} \\ 0 & 0 & l_{3,3} \end{pmatrix} \\ &= \begin{pmatrix} l_{1,1}^2 & l_{1,1}l_{2,1} & l_{1,1}l_{3,1} \\ l_{1,1}l_{2,1} & l_{2,1}^2 + l_{2,2}^2 & l_{2,1}l_{3,1} + l_{2,2}l_{3,2} \\ l_{1,1}l_{3,1} & l_{2,1}l_{3,1} + l_{2,2}l_{3,2} & l_{3,1}^2 + l_{3,2}^2 + l_{3,3}^2 \end{pmatrix}. \end{aligned}$$

Porovnáním prvků matice na levé a pravé straně tohoto vztahu dostaneme vzorce pro výpočet prvků matice $\mathbf{L}$,

$$l_{1,1} = \sqrt{a_{1,1}}, \quad l_{2,1} = \frac{a_{2,1}}{l_{1,1}}, l_{3,1} = \frac{a_{3,1}}{l_{1,1}}, l_{2,2} = \sqrt{a_{2,2} - l_{2,1}^2}, l_{3,2} = \frac{a_{3,2} - l_{2,1}l_{3,1}}{l_{2,2}}, \tag{1.63}$$

a

$$l_{3,3} = \sqrt{a_{3,3} - l_{3,1}^2 - l_{3,2}^2}. \tag{1.64}$$

Analogicky lze určit vzorce pro matici $\mathbf{L}$ velikosti $n \times n$ a tím odvodit následující algoritmus.

Choleského rozklad

for $r = 1 : n$

$$l_{rr} = \left(a_{rr} - \sum_{s=1}^{r-1} l_{rs}^2 \right)^{1/2}$$

for $i = r+1 : n$

$$l_{ir} = \frac{1}{l_{rr}} \left(a_{ir} - \sum_{s=1}^{r-1} l_{rs} l_{is} \right)$$

end

end

Při řešení soustavy lineárních algebraických rovnic pomocí Choleského rozkladu postupujeme podobně jako u LU rozkladu. To znamená, že po určení matice $\mathbf{L}$ řešíme dvě soustavy s trojúhelníkovými maticemi a to $\mathbf{Ly} = \mathbf{b}$ a $\mathbf{L}^T\mathbf{x} = \mathbf{y}$. Choleského rozklad vyžaduje asi polovinu času a paměti ve srovnání s Gaussovou eliminací a LU rozkladem.

1.4 Maticové iterační metody

Tyto metody spočívají v konstrukci posloupnosti vektorů, která je dána předpisem:

$$\mathbf{x}^{i+1} = \mathbf{B}\mathbf{x}^i + \mathbf{c}, \quad i = 0, 1, \ldots, \tag{1.65}$$

kde $\mathbf{x}^0$ je zvolený počáteční vektor. Matice $\mathbf{B}$ se nazývá *iterační matice*. Matice $\mathbf{B}$ a vektor $\mathbf{c}$ jsou zvoleny tak, že pro přesné řešení $\mathbf{x}^*$ původní soustavy platí $\mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{c}$.
Konvergenci posloupností vektorů a matic budeme chápat jako konvergenci po složkách.

Definice 20. Řekneme, že posloupnost vektorů $\left\{\mathbf{x}^k\right\}_{k=1}^{\infty}$, $\mathbf{x}^k \in \mathbb{R}^n$, konverguje k $\mathbf{x} \in \mathbb{R}^n$, jestliže pro každé $i = 1, \ldots, n$ posloupnost reálných čísel x_i^k konverguje k x_i.
Řekneme, že posloupnost matic $\left\{\mathbf{A}^{(k)}\right\}_{k=1}^{\infty}$, kde $\mathbf{A}^{(k)} \in \mathbb{R}^{m\times n}$, konverguje k $\mathbf{A} \in \mathbb{R}^{m\times n}$, jestliže pro každé $i = 1, \ldots, m$, $j = 1, \ldots, n$ posloupnost reálných čísel a_{ij}^k konverguje k a_{ij}.

V této definici je použit horní index (k) pro označení k-té matice v posloupnosti, aby bylo odlišeno indexování matic a umocnění matice $\mathbf{A}$ na k-tou, které je značeno standardně $\mathbf{A}^k$ a objevuje se v následující definici.

Definice 21. Řekneme, že matice $\mathbf{A}$ je konvergentní, jestliže posloupnost $\left\{\mathbf{A}^k\right\}_{k=1}^{\infty}$ konverguje k nulové matici.

Věta 22. Oldenburgerova
Matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je konvergentní právě tehdy, když $\rho(\mathbf{A}) < 1$.

Důkaz. Každou matici $\mathbf{A} \in \mathbb{R}^{n\times n}$ lze zapsat v Jordanově tvaru $\mathbf{A} = \mathbf{TJT}^{-1}$. Matice $\mathbf{T}$ je

regulární a matice $\mathbf{J}$ je blokově diagonální,

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{J}_{p-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_p \end{pmatrix}, \quad \mathbf{J}_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}, \tag{1.66}$$

přičemž $i = 1, \dots, p$ a $\lambda_1, \dots, \lambda_p$ jsou vlastní čísla matice $\mathbf{A}$. Potom

$$\mathbf{A}^2 = \mathbf{T}\mathbf{J}\mathbf{T}^{-1}\mathbf{T}\mathbf{J}\mathbf{T}^{-1} = \mathbf{T}\mathbf{J}^2\mathbf{T}^{-1} \tag{1.67}$$

a analogicky $\mathbf{A}^k = \mathbf{T}\mathbf{J}^k\mathbf{T}^{-1}$. Z toho plyne, že $\mathbf{A}^k$ konverguje k nulové matici právě tehdy, když $\mathbf{J}^k$ konverguje k nulové matici. Platí

$$\mathbf{J}^k = \begin{pmatrix} \mathbf{J}_1^k & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2^k & & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{J}_{p-1}^k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_p^k \end{pmatrix}, \tag{1.68}$$

kde mocniny Jordanových bloků mají tvar

$$\mathbf{J}_i^k = \begin{pmatrix} \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \binom{k}{2}\lambda_i^{k-2} & \dots & \binom{k}{n_i-1}\lambda_i^{k-n_i+1} \\ 0 & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \dots & \binom{k}{n_i-2}\lambda_i^{k-n_i+2} \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} \\ 0 & 0 & \dots & 0 & \lambda_i^k \end{pmatrix}. \tag{1.69}$$

Z toho plyne, že $\mathbf{J}_i^k$ konverguje k nulové matici právě tehdy, když $|\lambda_i| < 1$. Tím pádem matice $\mathbf{J}^k$ a tedy i matice $\mathbf{A}^k$ konvergují pro $k \to \infty$ k nulové matici právě tehdy, když $\rho(\mathbf{A}) < 1$. □

Pomocí této věty nyní snadno odvodíme nutnou a postačující podmínku konvergence maticových iteračních metod.

Věta 23. Nutná a postačující podmínka konvergence
Posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ *konverguje k přesnému řešení soustavy* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *pro libovolný počáteční vektor* $\mathbf{x}^0$ *právě tehdy, když* $\rho(\mathbf{B}) < 1$.

Důkaz. Označme $\mathbf{x}^*$ přesné řešení soustavy $\mathbf{Ax} = \mathbf{b}$. Pro chybu metody platí

$$\begin{aligned} \mathbf{x}^i - \mathbf{x}^* &= \mathbf{B}\mathbf{x}^{i-1} + \mathbf{c} - \mathbf{B}\mathbf{x}^* - \mathbf{c} = \mathbf{B}\mathbf{x}^{i-1} - \mathbf{B}\mathbf{x}^* = \mathbf{B}\left(\mathbf{x}^{i-1} - \mathbf{x}^*\right) \\ &= \mathbf{B}^2\left(\mathbf{x}^{i-2} - \mathbf{x}^*\right) = \mathbf{B}^i\left(\mathbf{x}^0 - \mathbf{x}^*\right). \end{aligned} \tag{1.70}$$

Z toho plyne, že $\{\mathbf{x}^i\}_{i=0}^{\infty}$ konverguje k $\mathbf{x}^*$ právě tehdy, když $\mathbf{B}^i$ konverguje k nulové matici, což je podle předchozí věty právě tehdy, když $\rho(\mathbf{B}) < 1$. □

Vzhledem k tomu, že spektrální poloměr matice $\mathbf{B}$ můžeme odhadnout maticovou normou této matice, platí následující věta o postačující podmínce konvergence.

Věta 24. Postačující podmínka konvergence
Posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ konverguje k přesnému řešení soustavy $\mathbf{Ax} = \mathbf{b}$ pro libovolný počáteční vektor $\mathbf{x}^0$, pokud $\|\mathbf{B}\| < 1$ pro některou z maticových norem $\|\ \|$.

Důkaz. Pokud $\|\mathbf{B}\| < 1$, potom $\rho(\mathbf{B}) < \|\mathbf{B}\| < 1$ a dle předchozí věty posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ konverguje k přesnému řešení dané soustavy □

Pokud posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ kontruovaná podle předpisu dané metody konverguje k přesnému řešení soustavy $\mathbf{Ax} = \mathbf{b}$ pro libovolný počáteční vektor $\mathbf{x}^0$, potom budeme říkat, že metoda konverguje. V dalším textu uvedeme několik konkrétních maticových iteračních metod a budeme se zabývat tím, pro jaké typy matic tyto metody konvergují.

1.4.1 Jacobiho metoda

Rozložme matici soustavy $\mathbf{A}$ na součet

$$\mathbf{A} = \mathbf{D} + \mathbf{L} + \mathbf{U}, \tag{1.71}$$

kde $\mathbf{D}$ je diagonální matice, $\mathbf{L}$ je dolní trojúhelníková matice s nulami na diagonále a $\mathbf{U}$ je horní trojúhelníková matice s nulami na diagonále. Předpokládejme, že $\mathbf{D}$ je regulární, potom platí

$$\mathbf{Ax} = \mathbf{b} \Leftrightarrow \mathbf{Dx} + (\mathbf{L} + \mathbf{U})\,\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{Dx} = -(\mathbf{L} + \mathbf{U})\,\mathbf{x} + \mathbf{b}. \tag{1.72}$$

To je dále ekvivalentní se vztahem

$$\mathbf{x} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\,\mathbf{x} + \mathbf{D}^{-1}\mathbf{b}. \tag{1.73}$$

Z poslední rovnice odvodíme předpis Jacobiho metody:

$$\mathbf{x}^{i+1} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\,\mathbf{x}^i + \mathbf{D}^{-1}\mathbf{b}, \quad i = 0, 1, \ldots. \tag{1.74}$$

Tento předpis můžeme rozepsat po složkách:

$$x_j^{i+1} = \frac{1}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^i - \sum_{k=j+1}^{n} a_{jk}x_k^i\right), \quad j = 1, \ldots, n, \quad i = 0, 1, \ldots. \tag{1.75}$$

V tomto případě je iterační matice dána vztahem $\mathbf{B} = -\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$ a posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ konverguje k přesnému řešení právě tehdy, když $\rho(-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})) < 1$. Tuto podmínku splňují například ostře diagonálně dominantní matice.

Definice 25. Matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ se nazývá *ostře diagonálně dominantní*, jestliže

$$|a_{ii}| > \sum_{j=1}^{i-1} |a_{ij}| + \sum_{j=i+1}^{n} |a_{ij}|, \quad i = 1, \ldots, n. \tag{1.76}$$

Věta 26. *Jestliže matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je ostře diagonálně dominantní, potom $\mathbf{A}$ je regulární.*

Důkaz. Provedeme důkaz sporem. Předpokládejme, že matice $\mathbf{A}$ je ostře diagonálně dominantní a singulární. Potom soustava $\mathbf{Ax} = \mathbf{0}$ má řešení $\mathbf{x} \neq \mathbf{0}$. tato soustava rozepíšeme po složkách

$$\sum_{i=1}^{n} a_{ij}x_j = a_{ii}x_i + \sum_{j\neq i} a_{ij}x_j = 0, \quad i = 1, \ldots, n. \tag{1.77}$$

Označme $x_s = \max\limits_{j=1,\ldots,n} x_j$ a položme $i = s$. Potom platí

$$|a_{ss}|\,|x_s| = |a_{ss}x_s| = \left|\sum_{j\neq s} a_{sj}x_j\right| \leq \sum_{j\neq s} |a_{sj}|\,|x_j| \leq \sum_{j\neq s} |a_{sj}|\,|x_s|. \tag{1.78}$$

Z toho plyne, že

$$|a_{ss}| \leq \sum_{j\neq s} |a_{sj}|, \tag{1.79}$$

což je ve sporu s předpokladem, že matice $\mathbf{A}$ je ostře diagonálně dominantní. □

Věta 27. *Jestliže matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je ostře diagonálně dominantní, potom Jacobiho metoda konverguje.*

Důkaz. Uvažujme vlastní číslo λ iterační matice $\mathbf{B}$ a vlastní vektor $\mathbf{x}$ příslušný tomuto vlastnímu číslu. Potom $\mathbf{Bx} = \lambda\mathbf{x}$ a tedy

$$-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\,\mathbf{x} = \lambda\mathbf{x}. \tag{1.80}$$

Z toho plyne, že

$$(\mathbf{L}+\mathbf{U})\,\mathbf{x} = -\mathbf{D}\lambda\mathbf{x}. \tag{1.81}$$

Tento vztah rozepsaný po složkách má tvar

$$\sum_{i\neq j} a_{ij}x_j = -\lambda a_{ii}x_i, \quad i = 1, \ldots, n. \tag{1.82}$$

Z toho plyne, že

$$|\lambda| \, |a_{ii}| \, |x_i| \leq \sum_{i \neq j} |a_{ij}| \, |x_j| \, . \tag{1.83}$$

Označme $|x_s| = \max\limits_{j=1,\ldots,n} |x_j|$. Pro $i = s$ dostaneme

$$|\lambda| \leq \frac{\sum\limits_{s \neq j} |a_{sj}| \, |x_j|}{|a_{ss}| \, |x_s|} \leq \frac{\sum\limits_{s \neq j} |a_{sj}| \, |x_s|}{|a_{ss}| \, |x_s|} \leq \frac{\sum\limits_{s \neq j} |a_{sj}|}{|a_{ss}|} < 1. \tag{1.84}$$

Z toho plyne, že $\rho\left(\mathbf{B}\right) < 1$ a Jacobiho metoda tedy konverguje. □

1.4.2 Gaussova-Seidelova metoda

Tato metoda se liší od předchozí tím, že při výpočtu složek vektoru $\mathbf{x}^{i+1}$ použijeme místo složek vektoru $\mathbf{x}^i$ složky vektoru $\mathbf{x}^{i+1}$, pokud již byly vypočteny.
Uvažujme matice $\mathbf{D}$, $\mathbf{L}$ a $\mathbf{U}$ jako u Jacobiho metody. Platí:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Leftrightarrow \left(\mathbf{D} + \mathbf{L}\right)\mathbf{x} + \mathbf{U}\mathbf{x} = \mathbf{b} \Leftrightarrow \left(\mathbf{D} + \mathbf{L}\right)\mathbf{x} = -\mathbf{U}\mathbf{x} + \mathbf{b}, \tag{1.85}$$

což dále je ekvivalentní se vztahem

$$\mathbf{x} = -\left(\mathbf{D} + \mathbf{L}\right)^{-1}\mathbf{U}\mathbf{x} + \left(\mathbf{D} + \mathbf{L}\right)^{-1}\mathbf{b}. \tag{1.86}$$

Z poslední rovnosti odvodíme předpis Gaussovy-Seidelovy metody:

$$\mathbf{x}^{i+1} = -\left(\mathbf{D} + \mathbf{L}\right)^{-1}\mathbf{U}\mathbf{x}^i + \left(\mathbf{D} + \mathbf{L}\right)^{-1}\mathbf{b}, \quad i = 0, 1, \ldots . \tag{1.87}$$

Tento předpis můžeme rozepsat po složkách:

$$x_j^{i+1} = \frac{1}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{i+1} - \sum_{k=j+1}^{n} a_{jk}x_k^i\right), \quad j = 1, \ldots, n, \quad i = 0, 1, \ldots . \tag{1.88}$$

V tomto případě posloupnost $\{\mathbf{x}^i\}_{i=0}^{\infty}$ konverguje k přesnému řešení právě tehdy, když

$$\rho\left(-\left(\mathbf{D} + \mathbf{L}\right)^{-1}\mathbf{U}\right) < 1. \tag{1.89}$$

Tuto podmínku splňují ostře diagonálně dominantní matice a také symetrické pozitivně definitní matice.

Věta 28. *Jestliže matice* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *je ostře diagonálně dominantní, potom Gaussova-Seidelova metoda konverguje.*

Důkaz. Uvažujme vlastní číslo λ iterační matice $\mathbf{B}$ a vlastní vektor $\mathbf{x}$ příslušný tomuto vlastnímu číslu. Potom $\mathbf{Bx} = \lambda\mathbf{x}$ a tedy

$$-\left(\mathbf{D}+\mathbf{L}\right)^{-1}\mathbf{Ux} = \lambda\mathbf{x}. \tag{1.90}$$

Z toho plyne, že

$$-\mathbf{Ux} = \lambda\left(\mathbf{D}+\mathbf{L}\right)\mathbf{x}. \tag{1.91}$$

Tento vztah rozepsaný po složkách má tvar

$$-\sum_{j=i+1}^{n} a_{ij}x_j = \lambda\sum_{j=1}^{i} a_{ij}x_j. \tag{1.92}$$

Označme $|x_s| = \max\limits_{j=1,\ldots,n} |x_j|$. Pro $i = s$ dostaneme

$$|\lambda|\left|\sum_{j=1}^{s} a_{sj}x_j\right| \leq \sum_{j=s+1}^{n} |a_{sj}|\,|x_j| \leq \sum_{j=s+1}^{n} |a_{sj}|\,|x_s|\,. \tag{1.93}$$

Budeme upravovat levou stranu této nerovnosti:

$$|\lambda|\left|\sum_{j=1}^{s} a_{sj}x_j\right| \geq |\lambda|\left(|a_{ss}x_s| - \sum_{j=1}^{s-1}|a_{sj}x_j|\right) \geq |\lambda|\left(|a_{ss}x_s| - \sum_{j=1}^{s-1}|a_{sj}|\,|x_j|\right) \tag{1.94}$$

$$\geq |\lambda|\left(|a_{ss}|\,|x_s| - \sum_{j=1}^{s-1}|a_{sj}|\,|x_s|\right). \tag{1.95}$$

Z toho plyne, že

$$|\lambda| \leq \frac{\sum\limits_{j=s+1}^{n}|a_{sj}|}{|a_{ss}| - \sum\limits_{j=1}^{s-1}|a_{sj}|}. \tag{1.96}$$

Matice $\mathbf{A}$ je ostře diagonálně dominantní, proto platí

$$|a_{ss}| - \sum_{j=1}^{s-1}|a_{sj}| > \sum_{j=s+1}^{n}|a_{sj}|\,. \tag{1.97}$$

Ve vztahu (1.96) je tedy čitatel menší než jmenovatel a proto $|\lambda| < 1$. Takže $\rho\left(\mathbf{B}\right) < 1$ a Gaussova-Seidelova metoda tedy konverguje. □

Věta 29. *Jestliže matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní, potom Gaussova-Seidelova metoda konverguje.*

Důkaz. Uvažujme vlastní číslo λ iterační matice $\mathbf{B}$ a vlastní vektor $\mathbf{x}$ příslušný tomuto vlastnímu číslu. Jak je uvedeno v předchozím důkazu, potom platí

$$-\mathbf{U}\mathbf{x} = \lambda\left(\mathbf{D}+\mathbf{L}\right)\mathbf{x}. \tag{1.98}$$

Tento vztah vynásobíme vektorem $\mathbf{x}^T$ zleva a dostaneme

$$-\mathbf{x}^T\mathbf{U}\mathbf{x} = \lambda\left(\mathbf{x}^T\mathbf{D}\mathbf{x}+\mathbf{x}^T\mathbf{L}\mathbf{x}\right). \tag{1.99}$$

Vzhledem k tomu, že $\mathbf{A}$ je symetrická, platí $\mathbf{U}=\mathbf{L}^T$. To dosadíme do předchozího vztahu a dále tento vztah umocníme na druhou. Dostaneme

$$\begin{aligned}\left(\mathbf{x}^T\mathbf{L}^T\mathbf{x}\right)^2 &= \lambda^2\left(\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)^2+2\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)+\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2\right) \\ &= \lambda^2\left(\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2+\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)\left(\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)+2\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)\right)\right) \\ &= \lambda^2\left(\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2+\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)\right).\end{aligned} \tag{1.100}$$

Matice $\mathbf{A}$ je pozitivně definitní, což znamená, že $\mathbf{x}^T\mathbf{A}\mathbf{x}>0$. Uvažujme jednotkový vektor $\mathbf{e}_i\in\mathbb{R}^n$, tedy vektor, který má na i-té pozici 1 a na ostatních pozicích 0. Potom

$$0<\mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = A_{ii}. \tag{1.101}$$

Matice $\mathbf{A}$ má tedy kladné diagonální prvky a proto

$$\mathbf{x}^T\mathbf{D}\mathbf{x} = \sum_{i=1}^{n} A_{ii}x_i^2 > 0. \tag{1.102}$$

Z toho plyne, že

$$\left(\mathbf{x}^T\mathbf{L}^T\mathbf{x}\right)^2 = \lambda^2\left(\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2+\left(\mathbf{x}^T\mathbf{D}\mathbf{x}\right)\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)\right) \geq \lambda^2\left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2. \tag{1.103}$$

Takže platí $\lambda^2<1$, a tedy $|\lambda|<1$. Z toho již plyne, že Gaussova-Seidelova metoda konverguje. □

1.4.3 SOR metoda

Název SOR metoda vychází z anglického successive overrelaxation method a tato metoda se označuje také jako superrelaxační metoda. Tuto metodu můžeme chápat jako modifikaci Gaussovy-Seidelovy metody, která se používá za účelem urychlení konvergence. V případě této metody konstruujeme posloupnost vektorů podle vzorce:

$$\mathbf{x}^{i+1} = (1-\omega)\,\mathbf{x}^i+\omega\tilde{\mathbf{x}}^{i+1}, \tag{1.104}$$

kde

$$\tilde{\mathbf{x}}^{i+1} = -\left(\mathbf{D}+\mathbf{L}\right)^{-1}\mathbf{U}\mathbf{x}^i + \left(\mathbf{D}+\mathbf{L}\right)^{-1}\mathbf{b}, \quad i = 0, 1, \ldots. \tag{1.105}$$

Parametr ω se nazývá relaxační faktor a volí se z intervalu $(0, 2)$, obvykle $\omega \in (1, 2)$. Pro symetrické pozitivně definitní matice soustavy tato volba parametru vede ke konvergenci posloupnosti $\{\mathbf{x}^i\}_{i=0}^{\infty}$ k řešení. Volbou $\omega = 1$ dostaneme Gaussovu-Seidelovu metodu. Po úpravě dostaneme

$$\mathbf{x}^{i+1} = \left(\mathbf{D}+\omega\mathbf{L}\right)^{-1}\left(-\omega\mathbf{U} + (1-\omega)\,\mathbf{D}\right)\mathbf{x}^i + \omega\left(\mathbf{D}+\omega\mathbf{L}\right)^{-1}\mathbf{b}, \quad i = 0, 1, \ldots. \tag{1.106}$$

Tento předpis můžeme rozepsat po složkách:

$$x_j^{i+1} = \frac{\omega}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{i+1} - \sum_{k=j+1}^{n} a_{jk}x_k^i\right) + (1-\omega)\,x_j^i, \tag{1.107}$$

kde $j = 1, \ldots, n$ a $i = 0, 1, \ldots$.

1.5 Metoda sdružených gradientů

Budeme předpokládat, že matice soustavy $\mathbf{A}$ je reálná symetrická a pozitivně definitní. Metoda sdružených gradientů je založena na následující větě.

Věta 30. *Předpokládejme, že matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická a pozitivně definitní. Potom vektor $\mathbf{x}^*$ je přesným řešením soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ právě tehdy, když $\mathbf{x}^*$ je bodem minima funkce*

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b}. \tag{1.108}$$

Důkaz. Funkce F je funkce n proměnných. Určíme její stacionární bod. Platí, že

$$\nabla F(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0} \tag{1.109}$$

právě tehdy, když $\mathbf{x}$ je řešením soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$. Hessova matice je matice $\nabla^2 F(\mathbf{x}) = \mathbf{A}$, je tedy pozitivně definitní a stacionární bod je bodem minima funkce F. □

Metoda sdružených gradientů spočívá v hledání minima funkce F. To budeme hledat následujícím způsobem: Budeme konstruovat posloupnost vektorů $\mathbf{x}^i$ tak, že v i-tém kroku zvolíme směr $\mathbf{v}^i$ a určíme číslo α_i tak, aby vektor

$$\mathbf{x}^{i+1} = \mathbf{x}^i + \alpha_i\mathbf{v}^i \tag{1.110}$$

byl bodem minima funkce F na přímce $\mathbf{x}^i + t\mathbf{v}^i$, $t \in \mathbb{R}$. Určíme proto hodnotu funkce F na této přímce v závislosti na t, tedy

$$\begin{aligned} g(t) &= F\left(\mathbf{x}^i + t\mathbf{v}^i\right) = \frac{1}{2}\left(\mathbf{x}^i + t\mathbf{v}^i\right)^T\mathbf{A}\left(\mathbf{x}^i + t\mathbf{v}^i\right) + \left(\mathbf{x}^i + t\mathbf{v}^i\right)^T\mathbf{b} \qquad (1.111)\\ &= \frac{1}{2}\left(\mathbf{x}^i\right)^T\mathbf{A}\mathbf{x}^i + \frac{t}{2}\left(\mathbf{v}^i\right)^T\mathbf{A}\mathbf{x}^i + \frac{t^2}{2}\left(\mathbf{v}^i\right)^T\mathbf{A}\mathbf{v}^i + \frac{t}{2}\left(\mathbf{x}^i\right)^T\mathbf{A}\mathbf{v}^i - \left(\mathbf{x}^i\right)^T\mathbf{b} - t\left(\mathbf{v}^i\right)^T\mathbf{b}. \end{aligned}$$

Určíme stacionární bod funkce g,

$$\begin{aligned} g'(t) &= \frac{1}{2}\left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{x}^i + \frac{1}{2}\left(\mathbf{x}^i\right)^T \mathbf{A}\mathbf{v}^i + t\left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{v}^i - \left(\mathbf{v}^i\right)^T \mathbf{b} \\ &= \left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{x}^i + t\left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{v}^i - \left(\mathbf{v}^i\right)^T \mathbf{b} \end{aligned} \tag{1.112}$$

Z toho dostaneme, že $g'(t) = 0$ pro

$$t = \frac{\left(\mathbf{v}^i\right)^T \mathbf{b} - \left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{x}^i}{\left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{v}^i}. \tag{1.113}$$

Tento bod t je tedy stacionární bod. Vzhledem k tomu, že

$$g''(t) = \left(\mathbf{v}^i\right)^T \mathbf{A}\mathbf{v}^i > 0, \tag{1.114}$$

jedná se o bod minima. Symbolem $\mathbf{r}^i$ označme reziduum $\mathbf{r}^i = \mathbf{b} - \mathbf{A}\mathbf{x}^i$ a $\langle \cdot, \cdot \rangle$ budeme značit skalární součin, tj. $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T\mathbf{v}$. Potom optimální velikost kroku α_i je rovna uvedenému stacionárnímu bodu funkce g a můžeme ji zapsat ve tvaru

$$\alpha_i = \frac{\langle \mathbf{r}^i, \mathbf{v}^i \rangle}{\langle \mathbf{A}\mathbf{v}^i, \mathbf{v}^i \rangle}. \tag{1.115}$$

Volbou $\mathbf{v}^i$ dostaneme konkrétní metodu. V případě metody sdružených gradientů volíme $\mathbf{A}$-*ortogonální vektory*. To jsou vektory, pro které platí:

$$\left\langle \mathbf{A}\mathbf{v}^i, \mathbf{v}^j \right\rangle = 0, \quad i \neq j, \text{ a } \left\langle \mathbf{A}\mathbf{v}^i, \mathbf{v}^i \right\rangle \neq 0. \tag{1.116}$$

Tyto vektory lze určit pomocí Gramovy-Schmidtovy ortogonalizace vektorů reziduí. To znamená, že nejprve položíme

$$\mathbf{v}^0 = \mathbf{r}^0. \tag{1.117}$$

Nyní předpokládejme, že vektory $\mathbf{v}^0, \ldots, \mathbf{v}^{k-1}$ jsou $\mathbf{A}$-ortogonální a definujme

$$\mathbf{v}^k = \mathbf{r}^k + \sum_{j=0}^{k-1} \beta_{kj}\mathbf{v}^j. \tag{1.118}$$

Koeficienty β_{kj} je potřeba určit tak, aby tento nový vektor $\mathbf{v}^k$ byl $\mathbf{A}$-ortogonální k vektorům $\mathbf{v}^0, \ldots, \mathbf{v}^{k-1}$. To znamená, aby platilo:

$$0 = \left\langle \mathbf{v}^k, \mathbf{A}\mathbf{v}^l \right\rangle = \left\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^l \right\rangle + \sum_{j=0}^{k-1} \beta_{kj} \left\langle \mathbf{v}^j, \mathbf{A}\mathbf{v}^l \right\rangle, \quad l = 0, \ldots, k-1. \tag{1.119}$$

Protože vektory $\mathbf{v}^0, \ldots, \mathbf{v}^{k-1}$ jsou $\mathbf{A}$-ortogonální, platí

$$\left\langle \mathbf{v}^j, \mathbf{A}\mathbf{v}^l \right\rangle = 0 \tag{1.120}$$

pro $j \neq l$ a $j, l = 0, \ldots, k-1$. Z toho plyne, že

$$\beta_{kl} = -\frac{\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^l \rangle}{\langle \mathbf{v}^l, \mathbf{A}\mathbf{v}^l \rangle}. \tag{1.121}$$

Vzhledem k tomu, že v prostoru $\mathbb{R}^n$ lze určit maximálně n $\mathbf{A}$-ortogonálních vektorů, musí existovat nějaký index $N \leq n$ takový, že $\mathbf{v}^N = 0$. V následujícím textu ukážeme, že potom $\mathbf{r}^N = \mathbf{b} - \mathbf{A}\mathbf{x}^N = 0$, což znamená, že $\mathbf{x}^N$ je přesné řešení. Dále ukážeme, že u této metody jsou vektory reziduí ortogonální a že vztah (1.118) lze ještě zjednodušit. Uveďme nejprve následující lemma.

Lemma 31. *a) Pro $k, s \leq N$ platí*

$$\mathbf{r}^k = \mathbf{r}^{k-1} - \alpha_{k-1}\mathbf{A}\mathbf{v}^{k-1} = \mathbf{r}^0 - \sum_{j=0}^{k-1} \alpha_j \mathbf{A}\mathbf{v}^j. \tag{1.122}$$

b) Pro $k < s$ platí $\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^s \rangle = 0$.
c) Pro $s \in \mathbb{N}$ platí $\langle \mathbf{r}^s, \mathbf{r}^s \rangle = \langle \mathbf{r}^s, \mathbf{v}^s \rangle$.

Důkaz. a) Platí

$$\begin{aligned} \mathbf{r}^k &= \mathbf{b} - \mathbf{A}\mathbf{x}^k = \mathbf{b} - \mathbf{A}\left(\mathbf{x}^{k-1} + \alpha_{k-1}\mathbf{v}^{k-1}\right) = \mathbf{b} - \mathbf{A}\mathbf{x}^{k-1} - \alpha_{k-1}\mathbf{A}\mathbf{v}^{k-1} \\ &= \mathbf{r}^{k-1} - \alpha_{k-1}\mathbf{A}\mathbf{v}^{k-1} = \mathbf{r}^0 - \sum_{j=0}^{k-1} \alpha_j \mathbf{A}\mathbf{v}^j. \end{aligned} \tag{1.123}$$

b) Ze vztahu (1.118) a A-ortogonality vektorů $\mathbf{v}^0, \ldots \mathbf{v}^s$ plyne, že

$$\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^s \rangle = \left\langle \mathbf{v}^k - \sum_{j=0}^{k-1} \beta_{kj}\mathbf{v}^j, \mathbf{A}\mathbf{v}^s \right\rangle = 0 \quad \text{pro} \quad k < s. \tag{1.124}$$

c) Pro $s \in \mathbb{N}$ platí

$$\langle \mathbf{r}^s, \mathbf{r}^s \rangle = \left\langle \mathbf{v}^s - \sum_{j=0}^{s-1} \beta_{sj}\mathbf{v}^j, \mathbf{r}^s \right\rangle = \langle \mathbf{v}^s, \mathbf{r}^s \rangle - \sum_{j=0}^{s-1} \beta_{sj} \langle \mathbf{v}^j, \mathbf{r}^s \rangle \tag{1.125}$$

S využitím A-ortogonality vektorů $\mathbf{v}^0, \ldots \mathbf{v}^s$ a vztahu pro α_j dostaneme pro $j < s$ vztah

$$\begin{aligned} \langle \mathbf{v}^j, \mathbf{r}^s \rangle &= \left\langle \mathbf{v}^j, \mathbf{r}^0 - \sum_{l=0}^{s-1} \alpha_l \mathbf{A}\mathbf{v}^l \right\rangle = \langle \mathbf{v}^j, \mathbf{r}^0 \rangle - \sum_{l=0}^{s-1} \alpha_l \langle \mathbf{v}^j, \mathbf{A}\mathbf{v}^l \rangle \\ &= \langle \mathbf{v}^j, \mathbf{r}^0 \rangle - \alpha_j \langle \mathbf{v}^j, \mathbf{A}\mathbf{v}^j \rangle = \langle \mathbf{v}^j, \mathbf{r}^0 \rangle - \frac{\langle \mathbf{r}^j, \mathbf{v}^j \rangle}{\langle \mathbf{A}\mathbf{v}^j, \mathbf{v}^j \rangle} \langle \mathbf{v}^j, \mathbf{A}\mathbf{v}^j \rangle \\ &= \langle \mathbf{v}^j, \mathbf{r}^0 - \mathbf{r}^j \rangle = \left\langle \mathbf{v}^j, \sum_{l=0}^{j-1} \alpha_l \mathbf{A}\mathbf{v}^l \right\rangle = 0. \end{aligned} \tag{1.126}$$

Po dosazení tohoto vztahu do (1.125) již dostaneme $\langle \mathbf{r}^s, \mathbf{r}^s \rangle = \langle \mathbf{r}^s, \mathbf{v}^s \rangle$. □

Ze vztahu c) v tomto lemmatu plyne, že pokud $\mathbf{v}^N = 0$, potom je

$$\left\langle \mathbf{r}^N, \mathbf{r}^N \right\rangle = \left\langle \mathbf{r}^N, \mathbf{v}^N \right\rangle = 0 \tag{1.127}$$

a tedy $\mathbf{r}^N = 0$. To znamená, že pokud v algoritmu vyjde vektor $\mathbf{v}^N$ určující směr v následujícím kroku nulový, znamená to, že vektor $\mathbf{x}^N$ je přesným řešením dané soustavy. V následující větě ukážeme, že rezidua jsou ortogonální a že některé z koeficientů β_{kl} jsou nulové.

Věta 32. *a) Vektory* $\mathbf{r}^0, \dots \mathbf{r}^{N-1}$ *jsou ortogonální, pokud jsou nenulové.*
b) Pro $l < k-1$ *je* $\beta_{kl} = 0$ *a*

$$\beta_{k,k-1} = \frac{\left\langle \mathbf{r}^k, \mathbf{r}^k \right\rangle}{\left\langle \mathbf{r}^{k-1}, \mathbf{r}^{k-1} \right\rangle}. \tag{1.128}$$

Důkaz. a) Dokážeme, že $\left\langle \mathbf{r}^k, \mathbf{r}^s \right\rangle = 0$ pro $k < s$ a to indukcí podle s. Pro $s = 1$ s využitím předchozího lemmatu a vztahu $\mathbf{v}^0 = \mathbf{r}^0$ dostaneme

$$\begin{aligned} \left\langle \mathbf{r}^0, \mathbf{r}^1 \right\rangle &= \left\langle \mathbf{r}^0, \mathbf{r}^0 - \alpha_0 \mathbf{A}\mathbf{v}^0 \right\rangle = \left\langle \mathbf{r}^0, \mathbf{r}^0 \right\rangle - \alpha_0 \left\langle \mathbf{r}^0, \mathbf{A}\mathbf{v}^0 \right\rangle \\ &= \left\langle \mathbf{r}^0, \mathbf{r}^0 \right\rangle - \frac{\langle \mathbf{r}^0, \mathbf{v}^0 \rangle}{\langle \mathbf{v}^0, \mathbf{A}\mathbf{v}^0 \rangle} \left\langle \mathbf{r}^0, \mathbf{A}\mathbf{v}^0 \right\rangle = 0. \end{aligned} \tag{1.129}$$

Nyní předpokládejme, že platí $\left\langle \mathbf{r}^k, \mathbf{r}^s \right\rangle = 0$ pro $k < s$, a dokážeme, že potom tento vztah platí také pro $s+1$. Pro $k < s$ na základě indukčního předpokladu a předchozího lemmatu platí

$$\left\langle \mathbf{r}^k, \mathbf{r}^{s+1} \right\rangle = \left\langle \mathbf{r}^k, \mathbf{r}^s - \alpha_s \mathbf{A}\mathbf{v}^s \right\rangle = \left\langle \mathbf{r}^k, \mathbf{r}^s \right\rangle - \alpha_s \left\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^s \right\rangle = 0. \tag{1.130}$$

Pro $k = s$ platí

$$\begin{aligned} \left\langle \mathbf{r}^s, \mathbf{r}^{s+1} \right\rangle &= \langle \mathbf{r}^s, \mathbf{r}^s \rangle - \langle \mathbf{r}^s, \alpha_s \mathbf{A}\mathbf{v}^s \rangle = \langle \mathbf{r}^s, \mathbf{r}^s \rangle - \alpha_s \langle \mathbf{r}^s, \mathbf{A}\mathbf{v}^s \rangle \\ &= \langle \mathbf{r}^s, \mathbf{r}^s \rangle - \frac{\langle \mathbf{r}^s, \mathbf{v}^s \rangle}{\langle \mathbf{v}^s, \mathbf{A}\mathbf{v}^s \rangle} \langle \mathbf{r}^s, \mathbf{A}\mathbf{v}^s \rangle \\ &= \langle \mathbf{r}^s, \mathbf{r}^s \rangle - \frac{\langle \mathbf{r}^s, \mathbf{v}^s \rangle}{\langle \mathbf{v}^s, \mathbf{A}\mathbf{v}^s \rangle} \left\langle \mathbf{v}^s - \sum_{j=0}^{s-1} \beta_{sj} \mathbf{v}^j, \mathbf{A}\mathbf{v}^s \right\rangle \\ &= \langle \mathbf{r}^s, \mathbf{r}^s \rangle - \langle \mathbf{r}^s, \mathbf{v}^s \rangle = 0. \end{aligned} \tag{1.131}$$

b) Ze vztahu $\mathbf{r}^{l+1} = \mathbf{r}^l - \alpha_l \mathbf{A}\mathbf{v}^l$ vyjádříme

$$\mathbf{A}\mathbf{v}^l = \frac{\mathbf{r}^l - \mathbf{r}^{l+1}}{\alpha_l} \tag{1.132}$$

Z toho pro $l < k-1$ plyne, že

$$\beta_{kl} = -\frac{\left\langle \mathbf{r}^k, \mathbf{A}\mathbf{v}^l \right\rangle}{\left\langle \mathbf{v}^l, \mathbf{A}\mathbf{v}^l \right\rangle} = -\frac{\left\langle \mathbf{r}^k, \mathbf{r}^l \right\rangle - \left\langle \mathbf{r}^k, \mathbf{r}^{l+1} \right\rangle}{\alpha_l \left\langle \mathbf{v}^l, \mathbf{A}\mathbf{v}^l \right\rangle} = 0. \tag{1.133}$$

Dále s využitím vzorce pro α_{k-1} dostaneme

$$\begin{aligned}\beta_{k,k-1} &= -\frac{\left\langle \mathbf{r}^k, \mathbf{r}^{k-1}\right\rangle - \left\langle \mathbf{r}^k, \mathbf{r}^k\right\rangle}{\alpha_{k-1}\left\langle \mathbf{v}^{k-1}, \mathbf{A}\mathbf{v}^{k-1}\right\rangle} = \frac{\left\langle \mathbf{r}^k, \mathbf{r}^k\right\rangle}{\alpha_{k-1}\left\langle \mathbf{v}^{k-1}, \mathbf{A}\mathbf{v}^{k-1}\right\rangle} \\ &= \frac{\left\langle \mathbf{r}^k, \mathbf{r}^k\right\rangle}{\left\langle \mathbf{r}^{k-1}, \mathbf{v}^{k-1}\right\rangle} = \frac{\left\langle \mathbf{r}^k, \mathbf{r}^k\right\rangle}{\left\langle \mathbf{r}^{k-1}, \mathbf{r}^{k-1}\right\rangle}.\end{aligned} \tag{1.134}$$

□

Výše uvedený postup nyní shrneme v následujím algoritmu.

Algoritmus metody sdružených gradientů:
Je dáno: $\mathbf{x}^0$, $\mathbf{A}$ symetrická pozitivně definitní, $\mathbf{b}$.
$\mathbf{r}^0 = \mathbf{b} - \mathbf{A}\mathbf{x}^0$
$\mathbf{v}^0 = \mathbf{r}^0$
for $i = 0, 1, 2, 3, \ldots$
$\alpha_i = \frac{\left\langle \mathbf{r}^i, \mathbf{v}^i\right\rangle}{\left\langle \mathbf{v}^i, \mathbf{A}\mathbf{v}^i\right\rangle}$
$\mathbf{x}^{i+1} = \mathbf{x}^i + \alpha_i \mathbf{v}^i$
$\mathbf{r}^{i+1} = \mathbf{r}^i - \alpha_i \mathbf{A}\mathbf{v}^i$
$\beta_i = \frac{\left\langle \mathbf{r}^{i+1}, \mathbf{r}^{i+1}\right\rangle}{\left\langle \mathbf{r}^i, \mathbf{r}^i\right\rangle}$
$\mathbf{v}^{i+1} = \mathbf{r}^{i+1} + \beta_i \mathbf{v}^i$
end

Jak již bylo zmíněno, počítáme-li bez zaokrouhlovacích chyb, vede tato metoda po n nebo méně krocích k přesnému řešení. Jedná se tedy o přímou metodu. Jelikož je však řád matice n často velké číslo, používá se tato metoda jako iterační, neboť dává řešení s požadovanou přesností mnohem dříve než po n iteracích.
Rychlost konvergence metody sdružených gradientů je charakterizována odhadem

$$\left\|\mathbf{x}^k - \mathbf{x}^*\right\|_{\mathbf{A}} \leq 2\left(\frac{\sqrt{\kappa_2(\mathbf{A})} - 1}{\sqrt{\kappa_2(\mathbf{A})} + 1}\right)^k \left\|\mathbf{x}^0 - \mathbf{x}^*\right\|_A, \tag{1.135}$$

kde $\|\cdot\|_{\mathbf{A}}$ je energetická norma

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\langle \mathbf{A}\mathbf{x}, \mathbf{x}\rangle} \tag{1.136}$$

a κ_2 označuje podmíněnost vzhledem k $\|\cdot\|_2$ normě

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \left\|\mathbf{A}^{-1}\right\|_2. \tag{1.137}$$

Vidíme, že rychlost konvergence metody sdružených gradientů závisí na čísle podmíněnosti matice soustavy. Konvergenci můžeme urychlit pomocí *předpodmínění*. To znamená, že místo úlohy $\mathbf{A}\mathbf{x} = \mathbf{b}$ budeme řešit ekvivalentní úlohu

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}\mathbf{P}^T\mathbf{x} = \mathbf{P}^{-1}\mathbf{b}, \tag{1.138}$$

tedy

$$\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}, \tag{1.139}$$

kde

$$\tilde{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}, \quad \tilde{\mathbf{x}} = \mathbf{P}^{T}\mathbf{x}, \quad \tilde{\mathbf{b}} = \mathbf{P}^{-1}\mathbf{b}. \tag{1.140}$$

Matice $\mathbf{P}$ se nazývá *předpodmiňovač* a volíme ji tak, aby byla regulární a aby číslo podmíněnosti matice $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}$ bylo menší než číslo podmíněnosti matice $\mathbf{A}$. Mezi nejjednodušší volbu patří

$$\mathbf{P} = \begin{pmatrix} \sqrt{a_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{a_{22}} & 0 & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \sqrt{a_{nn}} \end{pmatrix}. \tag{1.141}$$

Tento předpodmiňovač ovšem vede ke snížení čísla podmíněnosti pouze u některých typů matic. Obecně volba vhodného předpodmiňovače závisí na vlastnostech dané matice.

Kapitola 2

Výpočet vlastních čísel a vlastních vektorů matic

Volba metody pro nalezení vlastních čísel závisí na tom, zda řešíme *částečný problém vlastních čísel*, tedy zda potřebujeme nalézt největší nebo několik největších vlastních čísel, nebo *úplný problém vlastních čísel*, tedy zda chceme nalézt všechna vlastní čísla.
Nejprve si zopakujeme základní pojmy z lineární algebry týkající se problematiky vlastních čísel. Číslo λ se nazývá *vlastní číslo* matice $\mathbf{A} \in \mathbb{R}^{n\times n}$, jestliže existuje nenulový vektor $\mathbf{x}$ takový, že platí

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{2.1}$$

Vektor $\mathbf{x}$ se potom nazývá *vlastní vektor* matice $\mathbf{A}$.
Platí

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow \mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \Leftrightarrow (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}, \tag{2.2}$$

kde $\mathbf{I}$ označuje jednotkovou matici. Homogenní soustava má netriviální řešení právě tehdy, když je determinant matice soustavy roven nule. Vlastní čísla matice $\mathbf{A}$ proto musí splňovat podmínku

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{2.3}$$

Polynom

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) \tag{2.4}$$

se nazývá *charakteristický polynom*. *Algebraická násobnost* vlastního čísla je jeho násobnost jako kořene charakteristického polynomu. Vlastní čísla matice $\mathbf{A}$ jsou tedy kořeny jejího charakteristického polynomu. Výpočet vlastních čísel jako kořenů charakteristického polynomu se příliš nepoužívá, protože je výpočetně náročný a navíc numericky nestabilní, neboť koeficienty charakteristického polynomu jsou citlivé na malé změny v prvcích matice. Spíše se používá postup opačný a to převedení problému nalezení kořenů polynomu na problém nalézt vlastní čísla matice.
Základní věta algebry říká, že polynom stupně n má právě n kořenů, pokud každý kořen počítáme tolikrát, kolik je jeho násobnost. Každá matice řádu n má tedy právě n vlastních čísel, pokud je počítáme v jejich násobnosti.
Ze vztahu (2.3) ihned plyne následující věta.

Věta 33. *Vlastní čísla trojúhelníkové matice jsou její diagonální prvky.*

Řekneme, že čtvercová matice $\mathbf{A}$ je *podobná* matici $\mathbf{B}$, jestliže existuje regulární matice $\mathbf{T}$ taková, že platí

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}. \tag{2.5}$$

Věta 34. *Podobné matice mají stejná vlastní čísla včetně jejich násobnosti.*

Mnoho metod pro výpočet vlastních čísel matice je založeno na myšlence postupně transformovat matici na matice jí podobné a nalézt podobnou matici, která je trojúhelníková. Potom diagonální prvky podobné matice budou vlastní čísla původní matice.

2.1 Mocninná metoda

Mocninná metoda slouží k výpočtu dominantního vlastního čísla. Předpokládejme, že matice $\mathbf{A}$ je reálná čtvercová matice řádu n jejíž vlastní čísla λ_i splňují

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n| \tag{2.6}$$

a předpokládejme, že matice $\mathbf{A}$ má n lineárně nezávislých vlastních vektorů $\mathbf{v}^1, \ldots, \mathbf{v}^n$. Zvolme $\mathbf{x}^0$ a konstruujme posloupnost

$$\mathbf{x}^{i+1} = \mathbf{A}\mathbf{x}^i, \quad i = 1, 2, \ldots. \tag{2.7}$$

Protože platí

$$\mathbf{x}^{i+1} = \mathbf{A}\mathbf{x}^i = \mathbf{A}^2\mathbf{x}^{i-1} = \ldots = \mathbf{A}^{i+1}\mathbf{x}^0, \quad i = 1, 2, \ldots, \tag{2.8}$$

nazývá se tato metoda *mocninná metoda*. Vlastní vektory tvoří bázi prostoru $\mathbb{R}^n$, můžeme tedy $\mathbf{x}^0$ vyjádřit v této bázi:

$$\mathbf{x}^0 = c_1\mathbf{v}^1 + \ldots + c_n\mathbf{v}^n. \tag{2.9}$$

Z toho plyne, že

$$\begin{aligned} \mathbf{x}^i &= \mathbf{A}^i\mathbf{x}^0 = c_1\mathbf{A}^i\mathbf{v}^1 + \ldots + c_n\mathbf{A}^i\mathbf{v}^n && (2.10)\\ &= c_1\lambda_1^i\mathbf{v}^1 + \ldots + c_n\lambda_n^i\mathbf{v}^n && (2.11)\\ &= c_1\lambda_1^i\left(\mathbf{v}^1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^i\mathbf{v}^2 \ldots + \frac{c_n}{c_1}\left(\frac{\lambda_n}{\lambda_1}\right)^i\mathbf{v}^n\right), && (2.12) \end{aligned}$$

za předpokladu $c_1 \neq 0$. Pro $i \to \infty$ se $\mathbf{x}^i$ blíží násobku vlastního vektoru $\mathbf{v}^1$ a platí $\mathbf{x}^{i+1} \approx \lambda_1\mathbf{x}^i$. Aproximaci dominantního vlastního čísla λ_1^i vypočteme porovnáním složek vektorů $\mathbf{x}^i$ a $\mathbf{x}^{i+1}$, například porovnáním v absolutní hodnotě největších složek:

$$\lambda_1^i = \frac{\max_{1\leq j\leq n}\left|\mathbf{x}_j^{i+1}\right|}{\max_{1\leq j\leq n}\left|\mathbf{x}_j^i\right|} \tag{2.13}$$

nebo pomocí podílu

$$\lambda_1^i = \frac{\left(\mathbf{x}^k\right)^T \mathbf{A}\mathbf{x}^k}{\left(\mathbf{x}^k\right)^T \mathbf{x}^k}. \tag{2.14}$$

Potom λ_1^i pro $i \to \infty$ konverguje k dominantnímu vlastnímu číslu.

Algoritmus mocninné metody:
zvol $\mathbf{x}^0$
for $k = 0, 1, 2, \ldots$
$\quad \mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k$
$\quad \mathbf{x}^{k+1} = \frac{\mathbf{x}^{k+1}}{\left\|\mathbf{x}^{k+1}\right\|}$
$\quad \lambda^{k+1} = \left(\mathbf{x}^{k+1}\right)^T \mathbf{A}\mathbf{x}^{k+1}$
end

Rychlost konvergence závisí na podílu $\frac{\lambda_2}{\lambda_1}$. Pokud se absolutní hodnoty vlastních čísel λ_1 a λ_2 příliš neliší, potom je tato konvergence velmi pomalá. Pokud je $|\lambda_1|$ o hodně větší než $|\lambda_2|$, je mocninná metoda efektivní metodou k výpočtu dominantního vlastního čísla.

2.1.1 Inverzní mocninná metoda

Pomocí inverzní mocninné metody můžeme nalézt v absolutní hodnotě nejmenší vlastní číslo matice. Předpokládejme, že matice $\mathbf{A}$ je reálná čtvercová matice řádu n jejíž vlastní čísla λ_i splňují

$$|\lambda_1| > |\lambda_2| \geq \ldots > |\lambda_n| > 0, \tag{2.15}$$

a předpokládejme, že matice $\mathbf{A}$ má n lineárně nezávislých vlastních vektorů $\mathbf{v}^1, \ldots, \mathbf{v}^n$. Potom matice $\mathbf{A}^{-1}$ má vlastní čísla $\frac{1}{\lambda_i}$, $i = 1, \ldots, n$, která splňují

$$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \geq \ldots \geq \left|\frac{1}{\lambda_1}\right| > 0, \tag{2.16}$$

a stejné vlastní vektory jako matice $\mathbf{A}$. Mocninnou metodou tedy můžeme určit dominantní vlastní číslo $\frac{1}{\lambda_n}$ matice $\mathbf{A}^{-1}$ a tedy v absolutní hodnotě nejmenší vlastní číslo λ_n matice $\mathbf{A}$.

2.1.2 Inverzní mocninná metoda se spektrálním posunem

Inverzní mocninná metoda konverguje poměrně rychle, pokud absolutní hodnota vlastního čísla λ_n je výrazně menší než absolutní hodnota ostatních vlastních čísel. Toho můžeme využít k urychlení konvergence mocninné metody. Známe-li dobrý odhad μ vlastního čísla λ_n, potom mocninná metoda aplikovaná na matici $\mathbf{A} - \mu\mathbf{I}$ bude rychle konvergovat, neboť nejmenší vlastní číslo matice $\mathbf{A} - \mu\mathbf{I}$ je rovno $\lambda_n - \mu$, což je blízko nule. Inverzní mocninná metoda aplikovaná na $\mathbf{A} - \mu\mathbf{I}$ se nazývá *inverzní mocninná metoda se spektrálním posunem.* Používá se také v případě, že známe aproximaci vlastního čísla a chceme vypočítat vlastní vektor příslušný tomuto vlastnímu číslu.

Algoritmus inverzní mocninné metody se spektrálním posunem:
zvol $\mathbf{x}_0$
for $k = 1, 2, 3, \ldots$
$\quad$ řeš soustavu $(\mathbf{A} - \mu\mathbf{I})\,\mathbf{x}^k = \mathbf{x}^{k-1}$
$\quad \mathbf{x}^k = \frac{\mathbf{x}^k}{\|\mathbf{x}^k\|}$
$\quad \lambda_n^k = \left(\mathbf{x}^k\right)^T \mathbf{A}\mathbf{x}^k$
end

V každé iteraci řešíme soustavu $(\mathbf{A} - \mu\mathbf{I})\,\mathbf{x}^k = \mathbf{x}^{k-1}$. Protože matice soustavy je stále stejná a mění se pouze pravá strana, je výhodné použít k řešení soustavy LU rozklad. Inverzní mocninná metoda konverguje lineárně.
Konvergenci můžeme ještě urychlit, pokud v každé iteraci bude posun μ roven aktuálnímu odhadu vlastního čísla λ_n^k. Tato metoda se nazývá *metoda Rayleighových podílů*. Je známo, že konverguje kvadraticky, ale LU rozklad musíme provést v každé iteraci. Je-li matice $\mathbf{A}$ symetrická, potom dokonce konverguje kubicky.

2.2 QR algoritmus

Jak jsme se již zmínili, podobné matice mají stejná vlastní čísla. Mnoho algoritmů pro výpočet vlastních čísel matice je proto založeno na převodu matice na matici s ní podobnou, která má jednodušší tvar. Pokud bude výsledná matice například horní trojúhelníková, potom její vlastní čísla budou prvky na diagonále.
Jedním takovým algoritmem je *QR algoritmus*. Tento algoritmus používá rozklad matice na součin ortonormální matice $\mathbf{Q}$ a horní trojúhelníkové matice $\mathbf{R}$, kterému říkáme *QR rozklad*. Připomeňme, že matice $\mathbf{Q}$ se nazývá ortonormální, jestliže platí $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$.

Věta 35. *Každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$, $m \geq n$, lze rozložit na součin $\mathbf{A} = \mathbf{QR}$, kde $\mathbf{Q} \in \mathbb{R}^{m\times m}$ je ortonormální matice a $\mathbf{R} \in \mathbb{R}^{m\times n}$ je horní trojúhelníková matice.*

QR algoritmus:
$\mathbf{A}_0 = \mathbf{A}$
for $k = 1, 2, 3, \ldots$
$\quad$ proveď QR rozklad matice $\mathbf{A}_{k-1}$: $\mathbf{A}_{k-1} = \mathbf{Q}_k\mathbf{R}_k$
$\quad \mathbf{A}_k = \mathbf{R}_k\mathbf{Q}_k$
end

Ukážeme, že matice $\mathbf{A}_k$ je podobná matici $\mathbf{A}$:

$$\begin{aligned}
\mathbf{A}_k &= \mathbf{R}_k\mathbf{Q}_k = \mathbf{Q}_k^T\mathbf{Q}_k\mathbf{R}_k\mathbf{Q}_k \\
&= \mathbf{Q}_k^T\mathbf{A}_{k-1}\mathbf{Q}_k \\
&= \mathbf{Q}_k^T\mathbf{Q}_{k-1}^T\ldots\mathbf{Q}_1^T\mathbf{A}_{k-1}\mathbf{Q}_1\mathbf{Q}_2\ldots\mathbf{Q}_k \\
&= \left(\mathbf{Q}_1\mathbf{Q}_2\ldots\mathbf{Q}_k\right)^T\mathbf{A}_0\mathbf{Q}_1\mathbf{Q}_2\ldots\mathbf{Q}_k \\
&= \mathbf{Q}^T\mathbf{A}_0\mathbf{Q} = \mathbf{Q}^T\mathbf{A}\mathbf{Q},
\end{aligned} \tag{2.17}$$

kde $\mathbf{Q} = \mathbf{Q}_1\mathbf{Q}_2 \dots \mathbf{Q}_k$ je ortonormální, protože je součinem ortonormálních matic.
Lze ukázat, že matice $\mathbf{A}_k$ konvergují k matici $\mathbf{T}$, která je horní trojúhelníková. Z této matice určíme aproximace vlastních čísel matice $\mathbf{A}$. Vlastní vektory příslušné těmto vlastním číslům můžeme určit například pomocí inverzní metody s posunem, kde za posun volíme vypočtenou aproximaci vlastního čísla.
Nevýhodou QR algoritmu v této podobě může být jeho pomalá konvergence. Pro poddiagonální prvky matic $\mathbf{A}_k = \left(a_{ij}^k\right)_{i,j=1}^n$, $k \in \mathbb{N}$, platí

$$a_{ij}^k \leq \frac{|\lambda_i|}{|\lambda_j|} a_{ij}^{k-1}, \quad k \in \mathbb{N}, \tag{2.18}$$

za předpokladu, že pro vlastní čísla matice $\mathbf{A}$ platí

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|. \tag{2.19}$$

Pokud mají nějaká dvě vlastní čísla přibližně stejnou absolutní hodnotu, potom je konvergence poddiagonálních prvků k nule velmi pomalá. Konvergenci můžeme urychlit podobně jako v případě inverzní mocninné metody zavedením posunů.

QR algoritmus s posuny:

$\mathbf{A}_0 = \mathbf{A}$
for $k = 1, 2, 3, \ldots$
    urči posun μ_k
    proveď QR rozklad $\mathbf{A}_{k-1} - \mu_k\mathbf{I} = \mathbf{Q}_k\mathbf{R}_k$
    $\mathbf{A}_k = \mathbf{R}_k\mathbf{Q}_k + \mu_k\mathbf{I}$
end

Potom platí

$$a_{ij}^k \leq \frac{|\lambda_i - \mu_k|}{|\lambda_j - \mu_k|} a_{ij}^{k-1}, \quad k \in \mathbb{N}, \tag{2.20}$$

Za parametr μ_k můžeme zvolit například prvek, který je v pravém dolním rohu matice $\mathbf{A}_{k-1}$, tedy prvek a_{nn}^{k-1}. Tato volba se nazývá *Rayleigheův posun.*
Další nevýhodou je vysoká výpočetní náročnost QR rozkladu. QR rozklad plné matice řádu n vyžaduje $\mathcal{O}(n^3)$ operací. Je-li matice $\mathbf{A}_k$ v horním Hessenbergově tvaru, potom i v dalších iteracích dostaneme matice v horním Hessenbergově tvaru a QR rozklad takové matice vyžaduje $\mathcal{O}(n^2)$ operací. Proto nejprve převedeme matici na horní Hessenbergův tvar, což vyžaduje $\mathcal{O}(n^3)$ operací a potom provedeme QR algoritmus, ve kterém budeme v každém kroku potřebovat pouze $\mathcal{O}(n^2)$ operací. Je-li matice symetrická, potom její převod na horní Hessenbergův tvar vede na třídiagonální matici. To je výhodné, neboť QR rozklad třídiagonální matice vyžaduje pouze $\mathcal{O}(n)$ operací.

2.2.1 Výpočet QR rozkladu pomocí Householderovy transformace

QR rozklad lze provést pomocí několika algoritmů, například pomocí Householderových transformací, Givensovy rovinné rotace nebo Grammovy-Schmidtovy ortogonalize. Nejprve

si uvedeme QR rozklad pomocí Householderovy transformace.
Definujme *Householderovu transformaci* matice $\mathbf{A}$ vztahem

$$\mathbf{A}_k = \mathbf{H}_k \mathbf{A}_{k-1}, \quad k = 1, \ldots, n-1, \tag{2.21}$$

kde $\mathbf{A}_0 = \mathbf{A}$ a *Householderovy matice* budeme pro $k = 1, \ldots, n-1$ definovat

$$\mathbf{H}_k = \mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T, \tag{2.22}$$

kde $\mathbf{w}_k$ je vektor jednotkové délky. Potom $\mathbf{H}_k$ je ortonormální, neboť

$$\begin{aligned}
\mathbf{H}_k^T \mathbf{H}_k &= \left(\mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T\right)^T \left(\mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T\right) && (2.23)\\
&= \mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T + 4\mathbf{w}^k \left(\mathbf{w}^k\right)^T \mathbf{w}^k \left(\mathbf{w}^k\right)^T && (2.24)\\
&= \mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T + 4\mathbf{w}^k \left(\mathbf{w}^k\right)^T && (2.25)\\
&= \mathbf{I}. && (2.26)
\end{aligned}$$

Vektory $\mathbf{w}^k$ volíme

$$\mathbf{w}^k = \frac{\mathbf{z}^k}{\left\|\mathbf{z}^k\right\|_2}, \tag{2.27}$$

a složky vektoru $\mathbf{z}^k$ jsou definovány

$$\mathbf{z}_j^k = \begin{cases} 0, & j < k, \\ a_{kk}^{k-1} - \operatorname{sign}\left(a_{kk}\right)\left(\sum_{l=k}^{m} a_{lk}^2\right)^{1/2}, & j = k, \\ a_{jk}^{k-1}, & j > k. \end{cases} \tag{2.28}$$

Potom v prvních k sloupcích bude mít matice $\mathbf{A}_k$ pod diagonálou nuly a matice

$$\mathbf{R} = \mathbf{A}_n = \mathbf{H}_n \ldots \mathbf{H}_2 \mathbf{H}_1 \mathbf{A} \tag{2.29}$$

je horní trojúhelníková a matice

$$\mathbf{Q} = \mathbf{H}_1 \mathbf{H}_2 \ldots \mathbf{H}_n \tag{2.30}$$

je ortonormální, neboť je součinem ortonormálních matic. A plati $\mathbf{A} = \mathbf{QR}$.

2.2.2 Výpočet QR rozkladu pomocí Givensovy rovinné rotace

Definujme transformaci:

$$\mathbf{A}_k = \mathbf{G}_{k-1} \mathbf{A}_{k-1}, \quad \mathbf{A}_0 = \mathbf{A}, \tag{2.31}$$

kde $\mathbf{G}_{k-1} = \mathbf{G}_{i,j,\phi}^{k-1}$ je *Givensova matice rovinné rotace*. To je matice, která má na diagonále jedničky, kromě pozic (i,i) a (j,j), kde je $\cos\phi$, na pozici (i,j) má prvek $\sin\phi$, na pozici

(j,i) má prvek $-\sin\phi$ a ostatní prvky jsou nulové, tj.

$$\mathbf{G}_{k-1} = \mathbf{G}_{i,j,\phi}^{k-1} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos\phi & & & \sin\phi & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & -\sin\phi & & & \cos\phi & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}. \tag{2.32}$$

Snadno se ověří, že $\mathbf{G}_{k-1}^T\mathbf{G}_{k-1} = \mathbf{I}$. Matice $\mathbf{G}_{k-1}$ je tedy ortonormální. Za indexy i, j volíme postupně $j = 1, \ldots, n-1$, $i = j+1, \ldots, n$. Úhel ϕ volíme tak, aby platilo

$$\cos\phi = \frac{a_{jj}^{k-1}}{\sqrt{\left(a_{jj}^{k-1}\right)^2 + \left(a_{ij}^{k-1}\right)^2}}, \quad \sin\phi = \frac{a_{ij}^{k-1}}{\sqrt{\left(a_{jj}^{k-1}\right)^2 + \left(a_{ij}^{k-1}\right)^2}}. \tag{2.33}$$

Potom vynásobením matice $\mathbf{A}_{k-1}$ maticí $\mathbf{G}_{k-1} = \mathbf{G}_{i,j,\phi}^{k-1}$ vynulujeme prvek na pozici (i,j) a navíc ostatní již vynulované prvky zůstávají nulové. Tímto způsobem tedy vynulujeme všechny nenulové prvky pod diagonálou a máme

$$\mathbf{R} = \mathbf{G}_{m-1} \ldots \mathbf{G}_0 \mathbf{A} = \mathbf{Q}^T\mathbf{A}, \quad \mathbf{Q} = \mathbf{G}_0 \ldots \mathbf{G}_{m-1}, \tag{2.34}$$

kde $m = \frac{(n-1)n}{2}$ označuje počet poddiagonálních prvků. Potom

$$\mathbf{A} = \mathbf{QR}, \tag{2.35}$$

kde $\mathbf{R}$ je horní trojúhelníková a $\mathbf{Q}$ je ortonormální, protože je součinem ortonormálních matic.
Pokud a_{ij}^{k-1} je již nulový, potom matice $\mathbf{G}_{k-1} = \mathbf{G}_{i,j,\phi}^{k-1}$ je rovna jednotkové matici a provádíme tedy jen tolik kroků kolik je nenulových poddiagonálních prvků. Proto je vhodné tuto metodu použít pro QR rozklad matice, která má mnoho prvků pod diagonálou nulových.

2.2.3 Převod na horní Hessenbergův tvar

Nechť $\mathbf{A}$ je matice řádu n. Tuto matici převedeme na horní Hessenbergův tvar pomocí Householderových matic:

$$\mathbf{A}_k = \mathbf{P}_k\mathbf{A}_{k-1}\mathbf{P}_k, \quad k = 1, \ldots n-2, \tag{2.36}$$

kde $\mathbf{A}_0 = \mathbf{A}$ a *Householderovy matice* budeme pro $k = 1, \ldots, n-1$ definovat

$$\mathbf{P}_k = \mathbf{I} - 2\mathbf{w}^k \left(\mathbf{w}^k\right)^T, \tag{2.37}$$

kde $\mathbf{w}_k$ je vektor jednotkové délky. Vektory $\mathbf{w}^k$ volíme

$$\mathbf{w}^k = \frac{\mathbf{z}^k}{\|\mathbf{z}^k\|_2}, \tag{2.38}$$

a složky vektoru $\mathbf{z}^k$ jsou definovány

$$\mathbf{z}_j^k = \begin{cases} 0, & j \leq k, \\ a_{k+1,k}^{k-1} - \operatorname{sign}\left(a_{k+1,k}^{k-1}\right) \left(\sum_{l=k+1}^{n} \left(a_{lk}^{k-1}\right)^2\right)^{1/2}, & j = k+1, \\ a_{jk}^{k-1}, & j > k+1. \end{cases} \tag{2.39}$$

Potom je matice $\mathbf{A}_{n-2}$ v horním Hessenbergově tvaru a matice

$$\mathbf{Q} = \mathbf{P}_1 \mathbf{P}_2 \ldots \mathbf{P}_{n-2} \tag{2.40}$$

je ortonormální, protože je součinem ortonormálních matic. Platí $\mathbf{A}_{n-2} = \mathbf{Q}^T \mathbf{A} \mathbf{Q}$, matice $\mathbf{A}_2$ je tedy podobná s maticí $\mathbf{A}$ a tyto matice mají stejná vlastní čísla.

Kapitola 3

Řešení soustav s obdélníkovými maticemi

V první kapitole jsme se zabývali řešením soustav lineárních algebraických rovnic se čtvercovou regulární maticí, nyní se budeme zabývat obecnějšími soustavami s obdélníkovými maticemi. Budeme řešit soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{b} \in \mathbb{R}^m$. Matici

$$\mathbf{A_b} = \Big(\mathbf{A} \mid \mathbf{b}\Big) \tag{3.1}$$

nazýváme *rozšířená matice soustavy*. Symbolem $h(\mathbf{A})$ budeme značit *hodnost matice* $\mathbf{A}$, tedy počet lineárně nezávislých řádků matice $\mathbf{A}$. Z lineární algebry je známa následující věta.

Věta 36. Frobeniova věta
Předpokládejme, že $\mathbf{A} \in \mathbb{R}^{m \times n}$ *a* $\mathbf{b} \in \mathbb{R}^m$.
a) Je-li $h(\mathbf{A}) < h(\mathbf{A_b})$*, potom soustava* $\mathbf{Ax} = \mathbf{b}$ *nemá řešení.*
b) Je-li $h(\mathbf{A}) = h(\mathbf{A_b})$ *a* $h(\mathbf{A}) = n$*, potom soustava* $\mathbf{Ax} = \mathbf{b}$ *má právě jedno řešení.*
c) Je-li $h(\mathbf{A}) = h(\mathbf{A_b})$ *a* $h(\mathbf{A}) < n$*, potom soustava* $\mathbf{Ax} = \mathbf{b}$ *má nekonečně mnoho řešení.*

Zatímco v první kapitole jsme se zaměřili na soustavy s regulární maticí, které mají právě jedno řešení, nyní se zaměříme na soustavy, které mají nekonečně mnoho řešení a kromě toho také na takzvané nejlepší řešení soustavy ve smyslu nejmenších čtverců v případě, že daná soustava nemá přesné řešení.

3.1 Singulární rozklad

K řešení soustav s obdélníkovými maticemi můžeme použít singulární rozklad matice soustavy.

Věta 37. Věta o singulárním rozkladu.
Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ *lze rozložit na součin*

$$\mathbf{A} = \mathbf{USV}^T, \tag{3.2}$$

kde $\mathbf{U} \in \mathbb{R}^{m\times m}$ *a* $\mathbf{V} \in \mathbb{R}^{n\times n}$ *jsou ortonormální matice a matice* $\mathbf{S} \in \mathbb{R}^{m\times n}$ *je diagonální s nezápornými diagonálními prvky* $\sigma_1, \sigma_2, \ldots, \sigma_k$, $k = \min(m, n)$.

Důkaz. Předpokládejme, že $m \leq n$. Potom matice $\mathbf{A}^T\mathbf{A}$ je pozitivně semidefinitní, neboť

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|_2^2 \geq 0. \tag{3.3}$$

Z toho plyne, že vlastní čísla matice $\mathbf{A}^T\mathbf{A}$ jsou nezáporná a vlastní vektory ortogonální. Jordanův rozklad matice má tvar $\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^T$, kde $\mathbf{J}$ je diagonální matice s vlastními čísly $\lambda_1, \ldots, \lambda_n$ na diagonále a $\mathbf{V}$ je ortonormální matice, jejíž sloupce jsou vlastní vektory matice $\mathbf{A}^T\mathbf{A}$. Potom platí $(\mathbf{A}\mathbf{V})^T\mathbf{A}\mathbf{V} = \mathbf{J}$. Označme $(\mathbf{A}\mathbf{V})_i$ i-tý sloupec matice $\mathbf{A}\mathbf{V}$ a předpokládejme, že vlastní čísla jsou uspořádána sestupně a počet kladných vlastních čísel je r. Potom platí, že $(\mathbf{A}\mathbf{V})_i^T(\mathbf{A}\mathbf{V})_i = \lambda_i$. Pro λ_i kladné definujme

$$\mathbf{w}_i = \frac{(\mathbf{A}\mathbf{V})_i}{\sqrt{\lambda_i}}, \quad i = 1, \ldots, r. \tag{3.4}$$

Potom $\mathbf{w}_i^T\mathbf{w}_i = 1$ a $\mathbf{w}_i^T\mathbf{w}_j = 0$ pro $i \neq j$. To znamená, že vektory $\mathbf{w}_i$ jsou ortonormální. Tyto vektory $\mathbf{w}_1, \ldots, \mathbf{w}_r$ doplníme vektory $\mathbf{w}_{r+1}, \ldots, \mathbf{w}_n$ na ortonormální bázi. Definujme nyní matici $\mathbf{U}$ tak, že její sloupce jsou postupně vektory $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Dále definujme matici $\mathbf{S}$ o rozměrech $m \times n$ tak, že její diagonální prvky budou postupně $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}$ a ostatní prvky jsou nulové. Pro $i = 1, \ldots, r$ je i-tý sloupec matice $\mathbf{A}\mathbf{V}$ roven $\sqrt{\lambda_i}\mathbf{w}_i$, což je také i-tý sloupec matice $\mathbf{U}\mathbf{S}$. Pro $i = r, \ldots, n$ jsou i-té sloupce matice $\mathbf{A}\mathbf{V}$ i matice $\mathbf{U}\mathbf{S}$ nulové. Z toho plyne, že $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{S}$ a tedy $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$.
Pokud $m > n$, potom aplikujeme analogický postup na matici $\mathbf{A}\mathbf{A}^T$. □

Z důkazu uvedené věty plyne, že singulární rozklad není určen jednoznačně, čísla $\sigma_1, \ldots, \sigma_n$ mohou být na diagonále matice $\mathbf{S}$ v různém pořadí. Často se ovšem singulární rozklad určuje tak jako v uvedeném důkazu, kde jsou singulární čísla uspořádáno sestupně.
Diagonální prvky $\sigma_1, \sigma_2, \ldots, \sigma_k$ matice $\mathbf{S}$ se nazývají *singulární čísla* matice $\mathbf{A}$. Dále i-tý sloupec matice $\mathbf{U}$ se nazývá i-tý *levý singulární vektor*, i-tý sloupec matice $\mathbf{V}$ se nazývá i-tý *pravý singulární vektor*. Zřejmě platí

$$\mathbf{u}_i^T\mathbf{A} = \sigma_i\mathbf{v}_i^T, \quad \mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i, \quad i = 1, 2, \ldots, k. \tag{3.5}$$

Singulární rozklad má v lineární algebře mnohé aplikace. Některé z nich uvádíme níže.

a) **Výpočet pseudoinverzní matice**

   V dalším textu budeme definovat pseudoinverzní matice a ukážeme, že singulární rozklad můžeme použít k výpočtu těchto matic.

b) **Určení hodnosti matice**

   Dále můžeme singulární rozklad můžeme použít k výpočtu hodnosti matice, neboť hodnost matice je rovna počtu nenulových singulárních čísel.

c) **Výpočet spektrální normy matice**

Kromě toho se singulární rozklad používá k určení spektrální normy matice. Spektrální norma byla definována pro čtvercové matice. Pro obecnou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ je spektrální norma definována analogicky:

$$\|\mathbf{A}\|_2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}. \tag{3.6}$$

Spektrální norma matice $\mathbf{A}$ je rovna jejímu maximálnímu singulárnímu číslu,

$$\|\mathbf{A}\|_2 = \sigma_{\max}, \quad \sigma_{\max} = \max_{i=1,\ldots,k} \sigma_i. \tag{3.7}$$

d) **Výpočet čísla podmíněnosti matice**

Z toho dále plyne vztah pro výpočet čísla podmíněnosti matice vzhledem ke spektrální normě

$$\mathrm{cond}\mathbf{A} = \frac{\sigma_{\max}}{\sigma_{\min}}, \tag{3.8}$$

přičemž $\sigma_{\min}$ označuje minimální nenulové singulární číslo matice $\mathbf{A}$. Číslo podmíněnosti bylo definováno pro čtvercové regulární matice, v dalším textu bude tato definice zobecněna a číslo podmíněnosti bude definováno pro všechny matice. Zde uvedený vztah platí pro všechny nenulové matice.

e) **Řešení soustav homogenních rovnic**

Vektory $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_k$ tvoří bázi nulového prostoru matice $\mathbf{A}$, singulární rozklad lze tedy použít pro výpočet řešení homogenní soustavy lineárních algebraických rovnic.

f) **Určení báze prostoru $\mathcal{R}(\mathbf{A})$**

Vektory $\mathbf{u}_1, \ldots, \mathbf{u}_r$ tvoří bázi prostoru

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m : \exists \mathbf{x} \in \mathbb{R}^n \text{ takové, že } \mathbf{A}\mathbf{x} = \mathbf{y}\}. \tag{3.9}$$

Metody pro numerický výpočet singulárního rozkladu se podobají metodám pro výpočet vlastních čísel matice, spočívají v převedení na matici v jednodušším tvaru a následném rozkladu této jednodušší matice. Například funkce DGESVD, která je součástí knihovny LAPACK, nejprve matici převede na bidiagonální matici pomocí Householderovy transformace a tato bidiagonální matice je potom rozložena pomocí QR algoritmu.

3.2 Řešení soustav s obdélníkovou maticí

Nyní se zaměříme na řešení soustav s obdélníkovou maticí. Nejprve zobecníme pojem inverzní matice a zavedeme takzvanou pseudoinverzní matici.

Definice 38. Řekneme, že matice $\mathbf{A}^+ \in \mathbb{R}^{n\times m}$ je *pseudoinverzní* k matici $\mathbf{A} \in \mathbb{R}^{m\times n}$, jestliže jsou splněny podmínky:

a) $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$,

b) $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$,

c) $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$,

d) $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$.

Matice $\mathbf{A}^+$ se také nazývá *Mooreova-Penroseova pseudoinverzní matice.*

Pokud je matice $\mathbf{A}$ regulární, potom inverzní matice k $\mathbf{A}$ splňuje všechny podmínky uvedené v této definici a pseudoinverzní matice $\mathbf{A}^+$ k regulární matici $\mathbf{A}$ je tedy rovna $\mathbf{A}^{-1}$. Následující věta se zabývá existencí a jednoznačností pseudoinverzní matice.

Věta 39. *Ke každé matici* $\mathbf{A}$ *existuje právě jedna matice, která je k ní pseudoinverzní.*

Důkaz. Uvažujme matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ a její singulární rozklad $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ a definujme $\mathbf{X} = \mathbf{V}\mathbf{S}^+\mathbf{U}^T$, kde $\mathbf{S}^+$ je matice pseudoinverzní k $\mathbf{S}$. Pokud $\sigma_1, \ldots, \sigma_r$ jsou nenulové diagonální prvky matice $\mathbf{S}$ a prvek σ_k je na pozici (k, k), potom $\mathbf{S}^+$ je diagonální matice velikosti $n \times m$, která má na pozici (k, k) prvek $1/\sigma_k$ pro $i = 1, \ldots, r$ a ostatní prvky jsou nulové. Nejprve se všimněme, že platí

$$\mathbf{S}\mathbf{S}^+ = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{S}^+\mathbf{S} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{3.10}$$

kde $\mathbf{I}$ je jednotková matice velikosti $r \times r$ a $\mathbf{0}$ jsou nulové matice. Na základě toho se již snadno ověří, že $\mathbf{S}^+$ splňuje podmínky a)-d) z předchozí definice a je tedy skutečně pseudoinverní k $\mathbf{S}$.
Ukážeme, že matice $\mathbf{X}$ je pseudoinverzní k matici $\mathbf{A}$, to je splňuje vlastnosti a)-d) z předchozí definice.

a) $\mathbf{A}\mathbf{X}\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^+\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U}\mathbf{S}\mathbf{S}^+\mathbf{S}\mathbf{V}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{A}$.

b) $\mathbf{X}\mathbf{A}\mathbf{X} = \mathbf{V}\mathbf{S}^+\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^+\mathbf{U}^T = \mathbf{V}\mathbf{S}^+\mathbf{S}\mathbf{S}^+\mathbf{U}^T = \mathbf{V}\mathbf{S}\mathbf{U}^T = \mathbf{X}$.

c) $(\mathbf{A}\mathbf{X})^T = \left(\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^+\mathbf{U}^T\right)^T = \left(\mathbf{U}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{U}^T\right)^T = \mathbf{U}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{U}^T = \mathbf{A}\mathbf{X}$.

d) $(\mathbf{X}\mathbf{A})^T = \left(\mathbf{V}\mathbf{S}^+\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T\right)^T = \left(\mathbf{V}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T\right)^T = \mathbf{V}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T = \mathbf{X}\mathbf{A}$.

Všechny vlastnosti z definice jsou tedy splněny a ke každé matici tedy existuje matice k ní pseudoinverní.
Zbývá ukázat, že je tato matice určena jednoznačně. Předpokládejme, že matice $\mathbf{X}$ a $\mathbf{Y}$ jsou pseudoinverzní k $\mathbf{A}$. Potom platí:

$$\begin{aligned}\mathbf{X} &= \mathbf{XAX} = \mathbf{X}\left(\mathbf{AX}\right)^T = \mathbf{XX}^T\mathbf{A}^T = \mathbf{XX}^T\mathbf{A}^T\mathbf{Y}^T\mathbf{A}^T = \mathbf{XY}^T\mathbf{A}^T \\ &= \mathbf{X}\left(\mathbf{AY}\right)^T = \mathbf{XAY} = \mathbf{XAYAY} = \mathbf{XA}\left(\mathbf{YA}\right)^T\mathbf{Y} = \mathbf{XAA}^T\mathbf{Y}^T\mathbf{Y} \\ &= \left(\mathbf{XA}\right)^T\mathbf{A}^T\mathbf{Y}^T\mathbf{Y} = \mathbf{A}^T\mathbf{X}^T\mathbf{A}^T\mathbf{Y}^T\mathbf{Y} = \mathbf{A}^T\mathbf{Y}^T\mathbf{Y} = \mathbf{YAY} = \mathbf{Y}.\end{aligned} \tag{3.11}$$

Z toho již plyne, že pseudoinverzní matice k matici $\mathbf{A}$ je určena jednoznačně. □

Pomocí pseudoinverzní matice můžeme zobecnit definici čísla podmíněnosti obdélníkové matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ vzhledem ke spektrální normě následovně:

$$\operatorname{cond}\left(\mathbf{A}\right) = \left\|\mathbf{A}\right\|_2 \left\|\mathbf{A}^+\right\|_2. \tag{3.12}$$

Důkaz věty nám zároveň dává návod, jak určit pseudoinverzní matici k dané matici pomocí singulárního rozkladu. Pokud singulární rozklad matice $\mathbf{A}$ má tvar $\mathbf{A} = \mathbf{USV}^T$, potom pseudoinverzní matice k matici $\mathbf{A}$ je dána vztahem $\mathbf{A}^+ = \mathbf{VS}^+\mathbf{U}^T$. V následující větě uvedeme další možnost, jak určit pseudoinverní matici, a to v případě, že matice soustavy má více řádků než sloupců.

Věta 40. *Jestliže $\mathbf{A} \in \mathbb{R}^{m\times n}$, přičemž $m > n$ a $h\left(\mathbf{A}\right) = n$, potom matice $\mathbf{A}^T\mathbf{A}$ je regulární a $\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$.*

Důkaz. Nejprve ukážeme, že je matice $\mathbf{A}^T\mathbf{A}$ regulární. K tomu použijeme důkaz sporem. Předpokládejme tedy, že $\mathbf{A}^T\mathbf{A}$ je singulární. Potom existuje nenulový vektor $\mathbf{x}$ takový, že $\mathbf{A}^T\mathbf{Ax} = \mathbf{0}$ a proto platí

$$\mathbf{0} = \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} = \left(\mathbf{Ax}\right)^T\mathbf{Ax} = \left\|\mathbf{Ax}\right\|_2^2. \tag{3.13}$$

Z toho vyplývá, že $\mathbf{Ax} = \mathbf{0}$. To je ovšem ve sporu s předpokladem, že $h\left(\mathbf{A}\right) = n$, neboť tento předpoklad dle Frobeniovy věty zaručuje, že soustava $\mathbf{Ax} = \mathbf{0}$ má právě jedno a to nulové řešení.
Nyní ověříme, že $\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ splňuje podmínky a)-d) z definice pseudoinverzní matice.
a) $\mathbf{AA}^+\mathbf{A} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{A} = \mathbf{A}$.
b) $\mathbf{A}^+\mathbf{AA}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \mathbf{A}^+$.
c) $\left(\mathbf{AA}^+\right)^T = \left(\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\right)^T = \mathbf{A}\left(\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\right)^T\mathbf{A}^T = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \mathbf{AA}^+$.
d) $\left(\mathbf{A}^+\mathbf{A}\right)^T = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{A} = \mathbf{I}^T = \mathbf{I} = \mathbf{A}^+\mathbf{A}$ □

Věta 41. *Jestliže $\mathbf{A} \in \mathbb{R}^{m\times n}$, přičemž $m < n$ a $h\left(\mathbf{A}\right) = m$, potom matice $\mathbf{AA}^T$ je regulární a $\mathbf{A}^+ = \mathbf{A}^T\left(\mathbf{AA}^T\right)^{-1}$.*

Důkaz. Důkaz je analogický důkazu předchozí věty. □

Nyní se již zaměříme na řešení soustavy lineárních algebraických rovnic s obecnou maticí. Pokud taková soustava nemá řešení, je potřeba v některých aplikacích určit vektor, pro který je euklidovská norma rezidua minimální. Takový vektor nazýváme nejlepší řešení soustavy $\mathbf{Ax} = \mathbf{b}$ ve smyslu nejmenších čtverců.

Definice 42. Jestliže soustava $\mathbf{Ax} = \mathbf{b}$ nemá řešení, potom vektor $\tilde{\mathbf{x}}$ nazýváme *nejlepší řešení ve smyslu nejmenších čtverců* soustavy $\mathbf{Ax} = \mathbf{b}$, pokud

$$\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_2 \leq \|\mathbf{b} - \mathbf{Ax}\|_2 \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{3.14}$$

Řešení soustavy lineárních algebraických rovnic s obdélníkovou maticí můžeme vyjádřit pomocí pseudoinverzní matice.

Věta 43. *Předpokládejme, že $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{b} \in \mathbb{R}^m$.*
a) Jestliže $m = n$ a $\mathbf{A}$ je regulární, potom $\mathbf{A}^+ = \mathbf{A}^{-1}$ a soustava $\mathbf{Ax} = \mathbf{b}$ má právě jedno řešení $\mathbf{x} = \mathbf{A}^+\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}$.
b) Jestliže $m < n$ a $h(\mathbf{A}) = m$, potom soustava $\mathbf{Ax} = \mathbf{b}$ má nekonečně mnoho řešení a vektor $\mathbf{x} = \mathbf{A}^+\mathbf{b}$ je jedno z těchto řešení.
c) Jestliže $m > n$ a $h(\mathbf{A}) = n$, potom soustava $\mathbf{Ax} = \mathbf{b}$ nemá řešení a vektor $\tilde{\mathbf{x}} = \mathbf{A}^+\mathbf{b}$ je nejlepší řešení ve smyslu nejmenších čtverců této soustavy.

Důkaz. a) Toto tvrzení plyne z Frobeniovy věty a definice pseudoinverzní matice.
b) Počet řešení plyne z Frobeniovy věty. Dle předchozí věty máme

$$\mathbf{x} = \mathbf{A}^+\mathbf{b} = \mathbf{A}^T\left(\mathbf{AA}^T\right)^{-1}\mathbf{b}.$$

Tento vztah vynásobíme maticí $\mathbf{A}$ zleva a obdržíme

$$\mathbf{Ax} = \mathbf{AA}^T\left(\mathbf{AA}^T\right)^{-1}\mathbf{b} = \mathbf{b}. \tag{3.15}$$

Vektor $\mathbf{x} = \mathbf{A}^+\mathbf{b}$ je tedy řešením soustavy $\mathbf{A}x = \mathbf{b}$.
c) Počet řešení opět plyne z Frobeniovy věty. Z věty 40 vyplývá, že

$$\tilde{\mathbf{x}} = \mathbf{A}^+\mathbf{b} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}. \tag{3.16}$$

Tento vztah vynásobíme maticí $\mathbf{A}^T\mathbf{A}$ zleva a dostaneme

$$\mathbf{A}^T\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}^T\mathbf{b}. \tag{3.17}$$

Dále platí

$$\begin{aligned}
\|\mathbf{b} - \mathbf{Ax}\|_2^2 &= \langle \mathbf{b} - \mathbf{Ax}, \mathbf{b} - \mathbf{Ax}\rangle \\
&= \langle \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}} + \mathbf{A}\tilde{\mathbf{x}} - \mathbf{Ax}, \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}} + \mathbf{A}\tilde{\mathbf{x}} - \mathbf{Ax}\rangle \\
&= \langle \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}, \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\rangle + \langle \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}, \mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x})\rangle \\
&+ \langle \mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x}), \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\rangle + \langle \mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x}), \mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x})\rangle .
\end{aligned} \tag{3.18}$$

S použitím (3.17) odvodíme, že platí

$$\langle \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}, \mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x}) \rangle = \left\langle \mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{A}\tilde{\mathbf{x}}, (\tilde{\mathbf{x}} - \mathbf{x}) \right\rangle = 0. \tag{3.19}$$

Dosazením (3.19) do (3.18), dostaneme

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 = \|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_2^2 + \|\mathbf{A}(\tilde{\mathbf{x}} - \mathbf{x})\|_2^2 \geq \|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_2^2 . \tag{3.20}$$

□

Nejlepší řešení ve smyslu nejmenších čtverců určité soustavy můžeme tedy určit pomocí pseudoinverní matice, kterou umíme vypočítat pomocí singulárního rozkladu nebo pomocí vztahu z věty 40. Další možností se zabývá následující věta.

Věta 44. *Uvažujme soustavu* $\mathbf{A}\mathbf{x} = \mathbf{b}$*, kde* $\mathbf{A} \in \mathbb{R}^{m\times n}$*,* $\mathbf{b} \in \mathbb{R}^m$*,* $m > n$ *a* $h(\mathbf{A}) = n$*. Potom soustava*

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b} \tag{3.21}$$

má právě jedno řešení $\tilde{\mathbf{x}}$ *a toto řešení* $\tilde{\mathbf{x}}$ *je řešením soustavy* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *ve smyslu nejmenších čtverců.*

Důkaz. Tato věta vyplývá z důkazu tvrzení c) předchozí věty. □

Soustava (3.21) se nazývá *soustava normálních rovnic*. Řešit původní soustavu pomocí této věty může být vhodné zejména pokud počet proměnných je výrazně menší než počet rovnic, protože soustava (3.21) je potom výrazně menší než původní soustava. Matice této soustavy je symetrická pozitivně definitní a k jejímu řešení je tedy vhodné použít Choleského rozklad. Nevýhodou tohoto postupu je to, že číslo podmíněnosti matice soustavy normálních rovnic je rovno kvadrátu čísla podmíněnosti matice původní soustavy:

$$\operatorname{cond}\left(\mathbf{A}^T\mathbf{A}\right) = \left(\operatorname{cond}(\mathbf{A})\right)^2 . \tag{3.22}$$

Pro soustavy se špatně podmíněnou maticí je tedy řešení soustavy normálních rovnic špatně podmíněná úloha.

3.3 Řešení soustav pomocí QR rozkladu

Další metoda řešení soustavy ve smyslu nejmenších čtverců je založena na QR rozkladu matice soustavy. Připomeňme, že podle Věty 35 lze každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$, $m \geq n$, rozložit na součin $\mathbf{A} = \mathbf{Q}\mathbf{R}$, kde $\mathbf{Q} \in \mathbb{R}^{m\times m}$ je ortogonální matice a $\mathbf{R} \in \mathbb{R}^{m\times n}$ je horní trojúhelníková matice.

Budeme předpokládat, že matice má plnou sloupcovou hodnost, to znamená $h(\mathbf{A}) = n$. Platí

$$\begin{aligned} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 &= \|\mathbf{b} - \mathbf{Q}\mathbf{R}\mathbf{x}\|_2^2 = (\mathbf{b} - \mathbf{Q}\mathbf{R}\mathbf{x})^T (\mathbf{b} - \mathbf{Q}\mathbf{R}\mathbf{x}) \\ &= (\mathbf{b} - \mathbf{Q}\mathbf{R}\mathbf{x})^T \mathbf{Q}\mathbf{Q}^T (\mathbf{b} - \mathbf{Q}\mathbf{R}\mathbf{x}) \\ &= \left(\mathbf{Q}^T\mathbf{b} - \mathbf{Q}^T\mathbf{Q}\mathbf{R}\mathbf{x}\right)^T \left(\mathbf{Q}^T\mathbf{b} - \mathbf{Q}^T\mathbf{Q}\mathbf{R}\mathbf{x}\right) \\ &= \left\|\mathbf{Q}^T\mathbf{b} - \mathbf{R}\mathbf{x}\right\|_2^2. \end{aligned} \tag{3.23}$$

Označme nyní

$$\mathbf{Q} = \left(\mathbf{Q}_1 \mid \mathbf{Q}_2\right), \quad \mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix}, \tag{3.24}$$

kde $\mathbf{Q}_1 \in \mathbb{R}^{m\times n}$, $\mathbf{Q}_2 \in \mathbb{R}^{m\times(m-n)}$, $\mathbf{R}_1 \in \mathbb{R}^{n\times n}$ a $\mathbf{R}_2 \in \mathbb{R}^{(m-n)\times n}$ je nulová matice. Potom dostaneme

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 = \left\| \begin{pmatrix} \mathbf{Q}_1^T\mathbf{b} - \mathbf{R}_1\mathbf{x} \\ \mathbf{Q}_2^T\mathbf{b} \end{pmatrix} \right\|_2. \tag{3.25}$$

Z toho vyplývá, že $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$ nabývá minima v bodě $\mathbf{x}^*$, který splňuje

$$\mathbf{R}_1\mathbf{x}^* = \mathbf{Q}_1^T\mathbf{b}. \tag{3.26}$$

Algoritmus metody:

1. Proveď QR rozklad $\mathbf{A} = \mathbf{Q}\mathbf{R} = \left(\mathbf{Q}_1 \mid \mathbf{Q}_2\right) \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix}$

2. Řeš trojúhelníkovou soustavu $\mathbf{R}_1\mathbf{x} = \mathbf{Q}_1^T\mathbf{b}$. Řešení $\mathbf{x}^*$ této soustavy je řešením soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ ve smyslu nejmenších čtverců.

Literatura

[1] M. Fiedler: Speciální matice a jejich použití v numerický matematice. SNTL 1981.

[2] Ch. W. Ueberhuber: Numerical Computation 1: Methods, Software and Analysis. Springer 2013.

[3] Ch. W. Ueberhuber: Numerical Computation 2: Methods, Software and Analysis. Springer 2013.

[4] E. Vitásek: Numerické metody. SNTL 1987.