

Numerické řešení soustav lineárních rovnic

Mirko Navara

Centrum strojového vnímání, katedra kybernetiky
elektrotechnická fakulta ČVUT, Praha
<http://cmp.felk.cvut.cz/~navara>
10. 12. 2015

Úloha: Hledáme řešení soustavy n lineárních rovnic o n neznámých x_1, x_2, \dots, x_n

$$\begin{aligned} a_{1,1} x_1 + a_{1,2} x_2 + \dots + a_{1,n} x_n &= b_1 \\ a_{2,1} x_1 + a_{2,2} x_2 + \dots + a_{2,n} x_n &= b_2 \\ &\vdots \\ a_{n,1} x_1 + a_{n,2} x_2 + \dots + a_{n,n} x_n &= b_n \end{aligned}$$

Maticový tvar:

$$\mathbf{A} \mathbf{x} = \mathbf{b},$$

kde $\mathbf{A} = (a_{i,j})_{i,j=1,\dots,n}$ je (regulární) matice soustavy,
 $\mathbf{b} = (b_1, b_2, \dots, b_n)^\top$ vektor pravých stran,
 $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$ vektor neznámých.

Cramerovo pravidlo má velkou výpočetní složitost a numerické chyby.

Druhy problémů

Matice soustavy:

- plné, ne příliš velké,
- řídké, často velmi velké (mj. u kubického splinu).

Špatná podmíněnost

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$$

Malá změna koeficientů soustavy nebo pravé strany může způsobit velkou změnu řešení.
Zpětné dosazení (nepřesného) řešení \mathbf{x}_c dá **reziduum řešení**:

$$\mathbf{r} = \mathbf{b} - \mathbf{A} \mathbf{x}_c,$$

Pokud matice \mathbf{A}^{-1} má velké prvky, může být reziduum \mathbf{r} malé, i když se vektor \mathbf{x}_c podstatně liší od přesného řešení $\bar{\mathbf{x}}$.

$$\begin{aligned} \mathbf{r} &= \mathbf{A} \bar{\mathbf{x}} - \mathbf{A} \mathbf{x}_c = \mathbf{A} (\bar{\mathbf{x}} - \mathbf{x}_c), \\ \bar{\mathbf{x}} - \mathbf{x}_c &= \mathbf{A}^{-1} \mathbf{r}. \end{aligned}$$

Jsou-li prvky matice \mathbf{A}^{-1} velké, může i malá složka vektoru \mathbf{r} způsobit velký rozdíl $\bar{\mathbf{x}} - \mathbf{x}_c$.
Malé reziduum nezaručuje malou chybu řešení!

Takové soustavy nazýváme **špatně podmíněné**.

Příklad: Soustava

$$\begin{aligned} 2x + 6y &= 8 \\ 2x + 6.00001y &= 8.00001 \end{aligned}$$

má řešení $x = 1, y = 1$;

minimální změna koeficientů na soustavu

$$\begin{aligned} 2x + 6y &= 8 \\ 2x + 5.99999y &= 8.00002 \end{aligned}$$

změní řešení na $x = 10, y = -2$.

Inverzní matice k oběma soustavám mají prvky řádově 10^5 , což ukazuje na jejich špatnou podmíněnost.
Rovnice v soustavách jsou „skoro lineárně závislé“.

Zdroje chyb

- nepřesnost koeficientů soustavy a pravé strany,
- zaokrouhlovací chyby při výpočtu,
- chyby metody – nekonečný proces je nahrazen konečným počtem kroků (u iteračních metod).

Přímé metody

Po konečném počtu kroků vedou (teoreticky) k přesnému řešení.

Gaussova eliminace (GEM)

Postupné úpravy matice soustavy pomocí ekvivalentních úprav (nemění řešení soustavy) na horní trojúhelníkovou matici, ze které lze zpětným dosazením snadno získat řešení.

Rozšířená matice soustavy má prvky

$$\begin{aligned} a_{i,j}^{(0)} &= a_{i,j}, & \text{pro } i = 1, 2, \dots, n, j = 1, 2, \dots, n; \\ a_{i,n+1}^{(0)} &= b_i, & \text{pro } i = 1, 2, \dots, n. \end{aligned}$$

Soustavu

$$\begin{aligned} a_{1,1}^{(0)} x_1 + a_{1,2}^{(0)} x_2 + \dots + a_{1,n}^{(0)} x_n &= a_{1,n+1}^{(0)} \\ a_{2,1}^{(0)} x_1 + a_{2,2}^{(0)} x_2 + \dots + a_{2,n}^{(0)} x_n &= a_{2,n+1}^{(0)} \\ &\vdots \\ a_{n,1}^{(0)} x_1 + a_{n,2}^{(0)} x_2 + \dots + a_{n,n}^{(0)} x_n &= a_{n,n+1}^{(0)} \end{aligned}$$

převědeme povolenými úpravami na tvar

$$\begin{aligned} a_{1,1}^{(0)} x_1 + a_{1,2}^{(0)} x_2 + \dots + a_{1,n}^{(0)} x_n &= a_{1,n+1}^{(0)} \\ a_{2,2}^{(1)} x_2 + \dots + a_{2,n}^{(1)} x_n &= a_{2,n+1}^{(1)} \\ &\vdots \\ a_{n,n}^{(n-1)} x_n &= a_{n,n+1}^{(n-1)}, \end{aligned}$$

ze kterého zpětnou substitucí vypočítáme vektor řešení.

Pokud vyjde na diagonále nulový prvek, stačí provést záměnu řádků (resp. sloupců – v tom případě musíme **zaměnit i odpovídající složky vektoru řešení!**).

To lze, pokud je matice soustavy regulární.

Algoritmus:

Pro $k = 1, 2, \dots, n - 1$, pro $i = k + 1, k + 2, \dots, n$ a $j = k + 1, k + 2, \dots, n + 1$

$$a_{i,j}^{(k)} = a_{i,j}^{(k-1)} - \frac{a_{i,k}^{(k-1)}}{a_{k,k}^{(k-1)}} a_{k,j}^{(k-1)}.$$

Pokud po provedení **přímého chodu** je nějaký diagonální prvek $a_{i,i}^{(i-1)} = 0$ (resp. $|a_{i,i}^{(i-1)}| < \varepsilon$), matice soustavy je (resp. může být) singulární.

V opačném případě použijeme **zpětnou substituci**

$$x_i = \frac{1}{a_{i,i}^{(i-1)}} \left(a_{i,n+1}^{(i-1)} - \sum_{j=i+1}^n a_{i,j}^{(i-1)} x_j \right), \quad \text{pro } i = n, n - 1, \dots, 1.$$

Výběr hlavního prvku

Pokud číslo na diagonále je v absolutní hodnotě malé, jeho malá změna vyvolá velkou změnu výsledku při dělení a rostou zaokrouhlovací chyby.

Proto v každém kroku eliminace vybereme na diagonálu koeficient s co největší absolutní hodnotou = **hlavní prvek (pivot)**.

GEM s výběrem hlavního prvku

- **úplným** – vybíráme z $(n - k)^2$ prvků zbylé čtvercové podmatice (výpočetně složité),
- **sloupcovým** – vybíráme v rámci sloupce a pouze vyměníme řádky,
- **řádkový** – vybíráme v rámci řádku a vyměníme sloupce (**i pořadí neznámých!**).

Gaussova-Jordanova redukce

GEM může pokračovat eliminací prvků nad diagonálou.

Diagonální prvky lze převést na jedničky.

Sloupec pravých stran je pak vektor řešení.

Pro jedno použití je větší složitost, ale vyplatí se, pokud máme mnoho úloh lišících se pouze pravou stranou (např. výpočet inverzní matice, kdy řešíme soustavu lineárních rovnic současně pro n pravých stran, kde vycházíme z jednotkové matice).

LU-rozklad

Označme

$$\mathbf{L}_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{2,1}}{a_{1,1}} & 1 & 0 & \dots & 0 \\ -\frac{a_{3,1}}{a_{1,1}} & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -\frac{a_{n,1}}{a_{1,1}} & 0 & \dots & 0 & 1 \end{pmatrix}$$

a vynásobme $\mathbf{L}_1 \cdot \mathbf{A}$. Dostaneme první přidruženou soustavu z GEM s vynulovaným prvním sloupcem pod diagonálou. Pokračujeme:

$$\mathbf{A}_0 = \mathbf{A}, \quad \mathbf{A}_{i+1} = \mathbf{L}_{i+1} \cdot \mathbf{A}_i \quad \text{pro } i = 0, 1, \dots, n-2,$$

kde

$$\mathbf{L}_{i+1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & & \vdots \\ & & -\frac{a_{i+2,i+1}^{(i)}}{a_{i+1,i+1}^{(i)}} & \ddots & \\ \vdots & \vdots & \vdots & & 1 & 0 \\ 0 & 0 & -\frac{a_{n,i+1}^{(i)}}{a_{i+1,i+1}^{(i)}} & \dots & 0 & 1 \end{pmatrix}.$$

Po provedení $n-1$ maticových násobení máme

$$\mathbf{L}_{n-1} \cdot \mathbf{L}_{n-2} \dots \mathbf{L}_2 \cdot \mathbf{L}_1 \cdot \mathbf{A} = \mathbf{U},$$

kde matice \mathbf{U} je horní trojúhelníková (=výsledek přímého chodu GEM) a $\bar{\mathbf{L}} = \mathbf{L}_{n-1} \cdot \mathbf{L}_{n-2} \dots \mathbf{L}_2 \cdot \mathbf{L}_1$ dolní trojúhelníková s jednotkami na diagonále. Inverzní matice $\bar{\mathbf{L}}^{-1} = \mathbf{L}$ existuje a je rovněž dolní trojúhelníková s jednotkami na diagonále.

$$\begin{aligned} \bar{\mathbf{L}} \cdot \mathbf{A} &= \mathbf{U}, \\ \mathbf{A} &= \bar{\mathbf{L}}^{-1} \cdot \mathbf{U} = \mathbf{L} \cdot \mathbf{U}. \end{aligned}$$

Původní soustavu $\mathbf{A} \mathbf{x} = \mathbf{L} \cdot \mathbf{U} \mathbf{x} = \mathbf{b}$ nahradíme dvěma soustavami s trojúhelníkovými maticemi

$$\begin{aligned} \mathbf{L} \mathbf{y} &= \mathbf{b}, \\ \mathbf{U} \mathbf{x} &= \mathbf{y}, \end{aligned}$$

(neboť $\mathbf{A} \mathbf{x} = \mathbf{L} \cdot \mathbf{U} \mathbf{x} = \mathbf{L} \mathbf{y} = \mathbf{b}$),

které řešíme zpětnou substitucí.

Rozepsáním součinu $\mathbf{L} \cdot \mathbf{U}$ dostáváme

Algoritmus:

Pro $r = 1, 2, \dots, n$

$$u_{i,r} = a_{i,r} - \sum_{s=1}^{i-1} l_{i,s} a_{s,r} \quad \text{pro } i = 1, 2, \dots, r,$$

$$l_{i,r} = \frac{1}{u_{r,r}} \left(a_{i,r} - \sum_{s=1}^{r-1} l_{i,s} u_{s,r} \right) \quad \text{pro } i = r+1, r+2, \dots, n,$$

$$y_i = b_i - \sum_{s=1}^{i-1} l_{i,s} y_s \quad \text{pro } i = 1, 2, \dots, n,$$

$$x_i = \frac{1}{u_{i,i}} \left(y_i - \sum_{s=i+1}^n u_{i,s} x_s \right) \quad \text{pro } i = n, \dots, 2, 1.$$

Potřebujeme všechny prvky $u_{r,r} \neq 0$ (i během výpočtu) nenulové \Rightarrow výběr hlavního prvku (částečný); zmenší se tím i zaokrouhlovací chyby.

Poznámka: Celý algoritmus můžeme realizovat „na místě“, v jediné čtvercové matici. Pro překrývající se diagonální prvky použijeme $u_{r,r}$, protože na diagonále matice \mathbf{L} jsou jedničky, které nepočítáme ani neukládáme.

Poznámka: Tato metoda je zvláště vhodná pro řadu úloh lišících se pouze pravými stranami. Lze použít i výpočet inverzní matice \mathbf{A}^{-1} (ovšem s větší složitostí).

Výpočet inverzní matice

\mathbf{E} = jednotková matice

$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E} \Rightarrow \mathbf{A}$ vynásobená j -tým sloupcem matice \mathbf{A}^{-1} je rovna j -tému sloupci jednotkové matice \mathbf{E} .

$\mathbf{x}_j = j$ -tý sloupec \mathbf{A}^{-1}

$\mathbf{e}_j = j$ -tý sloupec \mathbf{E}

$$\mathbf{A} \mathbf{x}_j = \mathbf{e}_j$$

Máme soustavu rovnic, kde \mathbf{x}_j je vektor neznámých.

Jednotlivé sloupce inverzní matice \mathbf{A}^{-1} dostaneme jako řešení soustavy pro různé pravé strany – sloupce matice \mathbf{E} .

Můžeme využít GEM pro jednu matici soustavy a několik pravých stran současně; stačí „prodloužit“ cyklus pro řádky rozšířené matice soustavy typu $(n \times 2n)$, tedy $j = k+1, k+2, \dots, 2n$ (viz Gaussova-Jordanova redukce).

Použití LU-rozkladu: $\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$:

$$\mathbf{A}^{-1} = (\mathbf{L} \cdot \mathbf{U})^{-1} = \mathbf{U}^{-1} \cdot \mathbf{L}^{-1}.$$

Výpočet \mathbf{U}^{-1} a \mathbf{L}^{-1} je snadný, inverzní matice k trojúhelníkové je opět trojúhelníková.

Výpočet determinantu

Podle definice pouze pro velmi malé řádky matic.

GEM: po eliminaci jako součin prvků na diagonále:

$$\det \mathbf{A} = \pm a_{1,1}^{(0)} a_{2,2}^{(1)} a_{3,3}^{(2)} \dots a_{n,n}^{(n-1)}.$$

POZOR! Výměna řádků či sloupců (při výběru hlavního prvku) **mění znaménko** determinantu. (Stačí si pamatovat, zda počet výměn byl sudý nebo lichý.)

LU-rozklad $\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$:

$$\det \mathbf{A} = \det (\mathbf{L} \cdot \mathbf{U}) = \det \mathbf{L} \cdot \det \mathbf{U} = u_{1,1} u_{2,2} u_{3,3} \dots u_{n,n}.$$

(Jedná se o trojúhelníkové matice a \mathbf{L} má navíc na diagonále jednotky.)

Opět **nutno ošetřit znaménko** při výměnách řádků nebo sloupců.

Zpřesnění výsledků pomocí rezidua

Přesné řešení $\bar{\mathbf{x}}$ můžeme vyjádřit pomocí nepřesného řešení \mathbf{x}_c :

$$\bar{\mathbf{x}} = \mathbf{x}_c + \delta,$$

kde $\delta = (\delta_1, \delta_2, \dots, \delta_n)^T$ je **vektor oprav**.

$$\begin{aligned} \mathbf{A} \bar{\mathbf{x}} &= \mathbf{b} \\ \mathbf{A} (\mathbf{x}_c + \delta) &= \mathbf{b} \\ \mathbf{A} \mathbf{x}_c + \mathbf{A} \delta &= \mathbf{b} \\ \mathbf{A} \delta &= \mathbf{b} - \mathbf{A} \mathbf{x}_c = \mathbf{r} \end{aligned}$$

Vektor oprav je řešením stejné soustavy s pravou stranou nahrazenou reziduem \mathbf{r} .
Postup můžeme několikrát opakovat (je výhodné použít LU-rozklad) a získat další zpřesněná řešení

$$\mathbf{x}_c^{(1)}, \mathbf{x}_c^{(2)}, \mathbf{x}_c^{(3)}, \dots$$

Iterační metody

Snaží se konstruovat posloupnosti vektorů, konvergující k přesnému řešení soustavy.

Normy vektorů a matic [Handbook Lin. Alg.]

\mathbb{R}^n ... n -rozměrný aritmetický vektorový prostor

$\mathbb{R}^{n,n}$... prostor čtvercových matic řádu n

$\mathbf{o} \in \mathbb{R}^n$... nulový vektor

$\mathbf{O} \in \mathbb{R}^{n,n}$... nulová matice

Definice: Vektorová norma je zobrazení $\|\cdot\|_v: \mathbb{R}^n \rightarrow \mathbb{R}$ splňující

- $\|\mathbf{x}\|_v \geq 0$, přičemž $\|\mathbf{x}\|_v = 0 \Leftrightarrow \mathbf{x} = \mathbf{o}$,
- $\|c\mathbf{x}\|_v = |c| \|\mathbf{x}\|_v$,
- $\|\mathbf{x} + \mathbf{y}\|_v \leq \|\mathbf{x}\|_v + \|\mathbf{y}\|_v$ (trojúhelníková nerovnost).

Příklad:

- $\|\mathbf{x}\|_e = \sqrt{\sum_{i=1}^n x_i^2}$ euklidovská
- $\|\mathbf{x}\|_r = \max_{i=1, \dots, n} |x_i|$ maximová, Čebyševova
- $\|\mathbf{x}\|_s = \sum_{i=1}^n |x_i|$ součtová, „manhattanská“
- $\|\mathbf{x}\|_q = \left(\sum_{i=1}^n |x_i|^q\right)^{\frac{1}{q}}$, $q \geq 1$ společné zobecnění předchozích

Změna měřítka: Je-li $\|\cdot\|_v$ vektorová norma a $r > 0$, pak $\|\mathbf{x}\|_u = r \|\mathbf{x}\|_v$ je také vektorová norma.

Definice: Maticová norma je zobrazení $\|\cdot\|_M: \mathbb{R}^{n,n} \rightarrow \mathbb{R}$ splňující

- $\|\mathbf{A}\|_M \geq 0$, přičemž $\|\mathbf{A}\|_M = 0 \Leftrightarrow \mathbf{A} = \mathbf{O}$
- $\|c\mathbf{A}\|_M = |c| \|\mathbf{A}\|_M$,
- $\|\mathbf{A} + \mathbf{B}\|_M \leq \|\mathbf{A}\|_M + \|\mathbf{B}\|_M$ (trojúhelníková nerovnost),
- $\|\mathbf{A} \cdot \mathbf{B}\|_M \leq \|\mathbf{A}\|_M \|\mathbf{B}\|_M$ (Schwarzova nerovnost).

Příklad:

- $\|\mathbf{A}\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2}$ euklidovská, Frobeniova
- $\|\mathbf{A}\|_R = \max_{i=1, \dots, n} \sum_{j=1}^n |a_{i,j}|$ řádková
- $\|\mathbf{A}\|_S = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{i,j}|$ sloupcová

Co není maticová norma:

$$\|\mathbf{A}\|_M = \max_{i,j=1, \dots, n} |a_{i,j}|$$

splňuje první 3 podmínky, ale nikoli Schwarzovu nerovnost:

$$\left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\|_M = \left\| \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \right\|_M = 2 \not\leq \left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\|_M \cdot \left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\|_M = 1 \cdot 1 = 1,$$

$$\left\| \begin{pmatrix} 10^{-6} & 10^{-6} \\ 10^{-6} & 10^{-6} \end{pmatrix}^k \right\|_M = 2^k \left\| \begin{pmatrix} 10^{-6} & 10^{-6} \\ 10^{-6} & 10^{-6} \end{pmatrix} \right\|_M \rightarrow \infty \text{ pro } k \rightarrow \infty.$$

Důsledky Schwarzovy nerovnosti

Měřítka nemůže být libovolné: Dosaďme $\mathbf{B} := \mathbf{E}$ (jednotková matice):

$$\|\mathbf{A}\|_M = \|\mathbf{A} \cdot \mathbf{E}\|_M \leq \|\mathbf{A}\|_M \|\mathbf{E}\|_M \Rightarrow \|\mathbf{E}\|_M \geq 1.$$

Konvergence nekonečného součinu: Dosaďme $\mathbf{B} := \mathbf{B}^k$, $k \in \mathbb{N}$:

$$\|\mathbf{A} \cdot \mathbf{B}^k\|_M \leq \|\mathbf{A}\|_M \|\mathbf{B}\|_M^k \Rightarrow \|\mathbf{A} \cdot \mathbf{B}^k\|_M \rightarrow 0 \text{ pro } \|\mathbf{B}\|_M < 1, k \rightarrow \infty.$$

Definice: Maticová norma $\|\cdot\|_M$ je **souhlasná** s vektorovou normou $\|\cdot\|_v$, jestliže

$$\forall \mathbf{A} \in \mathbb{R}^{n,n} \quad \forall \mathbf{x} \in \mathbb{R}^n : \|\mathbf{A} \mathbf{x}\|_v \leq \|\mathbf{A}\|_M \|\mathbf{x}\|_v.$$

(Je bezrozměrná, změna měřítka vektorové normy nemá vliv.)

Změna měřítka maticové normy: Je-li $\|\cdot\|_M$ maticová norma a $r > 1$, pak $\|\mathbf{x}\|_L = r \|\mathbf{x}\|_M$ je také maticová norma (ale pro $0 < r < 1$ nikoli).

Je-li navíc $\|\cdot\|_M$ souhlasná s vektorovou normou $\|\cdot\|_v$, pak i $\|\cdot\|_L$ je souhlasná s $\|\cdot\|_v$. Můžeme mít maticové normy libovolně velké, my je však budeme chtít co nejmenší.

Věta: Ke každé vektorové normě $\|\cdot\|_v$ existuje alespoň jedna souhlasná maticová norma $\|\cdot\|_M$, a to

$$\|\mathbf{A}\|_M = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A} \mathbf{x}\|_v}{\|\mathbf{x}\|_v} = \sup_{\|\mathbf{x}\|_v=1} \|\mathbf{A} \mathbf{x}\|_v.$$

Norma $\|\cdot\|_M$ (**indukovaná** normou $\|\cdot\|_v$) splňuje $\|\mathbf{E}\|_M = 1$.

Věta: Ke každé maticové normě $\|\cdot\|_M$ existuje alespoň jedna souhlasná vektorová norma $\|\cdot\|_v$, a to $\|\mathbf{x}\|_v = \|\mathbf{X}\|_M$, kde

$$\mathbf{X} = \begin{pmatrix} x_1 & 0 & 0 & \dots & 0 \\ x_2 & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ x_n & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Důkaz:

$$\|\mathbf{A} \mathbf{x}\|_v = \left\| \begin{pmatrix} \sum_{j=1}^n a_{1,j} x_j & 0 & \dots & 0 \\ \sum_{j=1}^n a_{2,j} x_j & 0 & \dots & 0 \\ \vdots & & & \\ \sum_{j=1}^n a_{n,j} x_j & 0 & \dots & 0 \end{pmatrix} \right\|_M = \|\mathbf{A} \cdot \mathbf{X}\|_M \leq \|\mathbf{A}\|_M \|\mathbf{X}\|_M = \|\mathbf{A}\|_M \|\mathbf{x}\|_v.$$

Příklad: Euklidovská norma matice je souhlasná s euklidovskou normou vektoru:

$$\left\| \begin{pmatrix} x_1 & 0 & 0 & \dots & 0 \\ x_2 & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ x_n & 0 & 0 & \dots & 0 \end{pmatrix} \right\|_E^2 = \sum_{i=1}^n x_i^2 = \|\mathbf{x}\|_e^2.$$

Věta: Následující dvojice norem (vektorová, maticová) jsou souhlasné:

$$\|\cdot\|_r, \|\cdot\|_R,$$

$$\|\cdot\|_s, \|\cdot\|_S,$$

$$\|\cdot\|_e, \|\cdot\|_E.$$

Navíc

maticová norma $\|\cdot\|_R$ je indukována vektorovou normou $\|\cdot\|_r$,

maticová norma $\|\cdot\|_S$ je indukována vektorovou normou $\|\cdot\|_s$, ale

maticová norma $\|\cdot\|_E$ **není** indukována vektorovou normou $\|\cdot\|_e$.

Definice: Posloupnost vektorů $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$ **konverguje k vektoru** $\mathbf{x} \in \mathbb{R}^n$, jestliže

$$\lim_{k \rightarrow \infty} x_i^{(k)} = x_i \quad \text{pro } i = 1, 2, \dots, n.$$

Věta: Posloupnost vektorů $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$ konverguje k vektoru $\mathbf{x} \in \mathbb{R}^n$, právě když

$$\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\|_v = 0,$$

kde za normu můžeme zvolit libovolnou z výše uvedených vektorových norem.

(Konvergence nezáleží na volbě normy.)

Definice: Posloupnost matic $\mathbf{A}^{(0)} = (a_{i,j}^{(0)})_{i,j=1}^n$, $\mathbf{A}^{(1)} = (a_{i,j}^{(1)})_{i,j=1}^n$, $\mathbf{A}^{(2)} = (a_{i,j}^{(2)})_{i,j=1}^n, \dots$ **konverguje k matici** $\mathbf{A} = (a_{i,j})_{i,j=1}^n$, jestliže

$$\lim_{k \rightarrow \infty} a_{i,j}^{(k)} = a_{i,j} \quad \text{pro každé } i, j = 1, 2, \dots, n.$$

($\mathbf{A}^{(k)}$... k -tý člen posloupnosti)

Definice: Matice \mathbf{B} je **konvergentní**, jestliže posloupnost matic $\mathbf{B}, \mathbf{B}^2, \mathbf{B}^3, \mathbf{B}^4, \dots$ konverguje k nulové matici. V opačném případě řekneme, že matice \mathbf{B} je **divergentní**.

(\mathbf{B}^k ... k -tá mocnina matice \mathbf{B})

Vlastní čísla a spektrální poloměr

Definice: Číslo $\lambda \in \mathbb{C}$ je **vlastní (charakteristické) číslo** matice $\mathbf{A} \in \mathbb{R}^{n,n}$, jestliže existuje nenulový vektor $\mathbf{x} \in \mathbb{R}^n$ (**vlastní (charakteristický) vektor**) splňující

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

Věta: Číslo λ je vlastní číslo matice \mathbf{A} , právě když $\det(\mathbf{A} - \lambda \mathbf{E}) = 0$.

Důkaz: Je-li λ vlastní číslo matice \mathbf{A} , pak $\exists \mathbf{x} \neq \mathbf{o} : \lambda \mathbf{x} = \mathbf{A} \mathbf{x}$, tedy $\mathbf{A} \mathbf{x} - \lambda \mathbf{x} = \mathbf{o}$, $(\mathbf{A} - \lambda \mathbf{E}) \mathbf{x} = \mathbf{o}$.

Nechť $\det(\mathbf{A} - \lambda \mathbf{E}) \neq 0$, tedy matice $\mathbf{A} - \lambda \mathbf{E}$ je regulární. Homogenní soustava rovnic s regulární maticí soustavy má pouze triviální řešení \Rightarrow spor.

Naopak z předpokladu $\det(\mathbf{A} - \lambda \mathbf{E}) = 0$ přímo plyne, že matice $\mathbf{A} - \lambda \mathbf{E}$ je singulární a $\exists \mathbf{x} \neq \mathbf{o} : (\mathbf{A} - \lambda \mathbf{E}) \mathbf{x} = \mathbf{o}$. Tedy $\mathbf{A} \mathbf{x} - \lambda \mathbf{x} = \mathbf{o}$, $\lambda \mathbf{x} = \mathbf{A} \mathbf{x}$.

Výpočet vlastních čísel

Pro matice malých řádů:

- Vypočteme $\det(\mathbf{A} - \lambda \mathbf{E})$, což je polynom v proměnné λ stupně n (**charakteristický polynom matice \mathbf{A}**).
- Vlastní čísla λ_i jsou kořeny charakteristického polynomu matice \mathbf{A} (rovnice $\det(\mathbf{A} - \lambda \mathbf{E}) = 0$ je **charakteristická rovnice**).

Poznámka: Pro reálnou matici mohou vycházet „vlastní čísla“ komplexní. (O komplexních vlastních číslech můžeme hovořit v komplexním vektorovém prostoru, nikoli v reálném, kde není násobení komplexním číslem definováno.)

Definice: Spektrální poloměr matice $\mathbf{A} \in \mathbb{R}^{n,n}$ je číslo

$$\varrho(\mathbf{A}) = \max_{i=1,2,\dots,n} |\lambda_i|,$$

kde $\lambda_i, i = 1, 2, \dots, n$, jsou vlastní čísla matice \mathbf{A} .

Věta: Každá maticová norma $\|\cdot\|_M$ splňuje

$$\|\mathbf{A}\|_M \geq \varrho(\mathbf{A}).$$

Důkaz: Nechť $\varrho(\mathbf{A}) = |\lambda|$,

λ = vlastní číslo s největší absolutní hodnotou,

\mathbf{y} = vlastní vektor příslušný λ , $\lambda \mathbf{y} = \mathbf{A} \mathbf{y}$.

Maticová norma $\|\cdot\|_M$ je souhlasná s vektorovou normou $\|\cdot\|_v$, pro kterou platí

$$\|\mathbf{A}\|_M = \sup_{\mathbf{x} \neq \mathbf{o}} \frac{\|\mathbf{A} \mathbf{x}\|_v}{\|\mathbf{x}\|_v} \geq \frac{\|\mathbf{A} \mathbf{y}\|_v}{\|\mathbf{y}\|_v} = \frac{|\lambda| \|\mathbf{y}\|_v}{\|\mathbf{y}\|_v} = |\lambda| = \varrho(\mathbf{A}).$$

Věta: Matice \mathbf{B} je konvergentní $\iff \varrho(\mathbf{B}) < 1$.

Věta: Postačující podmínka konvergentnosti matice \mathbf{B} :

- $\|\mathbf{B}\|_M < 1$ pro **nějakou** maticovou normu, neboli
- lineární zobrazení reprezentované maticí \mathbf{B} je kontraktivní (vzhledem k nějaké vektorové normě, souhlasné s maticovou normou $\|\cdot\|_M$).

Maticové iterační metody

Hledáme posloupnost vektorů $\mathbf{x}^{(k)} \in \mathbb{R}^n$ splňující $\mathbf{x}^{(k)} \rightarrow \bar{\mathbf{x}}$, tj. $\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|_v \rightarrow 0$.
Použijeme rekurentní vzorec

$$\mathbf{x}^{(k+1)} = F_k(\mathbf{x}^{(k)}, \mathbf{x}^{(k-1)}, \dots, \mathbf{x}^{(k-m)}).$$

Omezíme se na **lineární jednobodové stacionární maticové iterační metody**, tj. F_k nezávisí na k , závisí jen na $\mathbf{x}^{(k)}$, a to „lineárně“ (správně afinně),

$$\mathbf{x}^{(k+1)} = \mathbf{B} \mathbf{x}^{(k)} + \mathbf{c}.$$

Např.

$$\begin{aligned} \mathbf{b} &= \mathbf{A} \mathbf{x}, \\ \mathbf{x} + \mathbf{b} &= \mathbf{x} + \mathbf{A} \mathbf{x} = (\mathbf{E} + \mathbf{A}) \mathbf{x}, \\ \mathbf{x} &= (\mathbf{E} + \mathbf{A}) \mathbf{x} - \mathbf{b}, \\ \mathbf{x}^{(k+1)} &= \underbrace{(\mathbf{E} + \mathbf{A})}_{\mathbf{B}} \mathbf{x}^{(k)} - \underbrace{\mathbf{b}}_{\mathbf{c}}. \end{aligned}$$

Vektor chyby:

$$\begin{aligned} \varepsilon^{(k)} &= \mathbf{x}^{(k)} - \bar{\mathbf{x}} \quad (\text{předpokládáme } \neq \mathbf{o}) \\ \mathbf{x}^{(k)} &= \mathbf{B} \mathbf{x}^{(k-1)} + \mathbf{c} \\ \bar{\mathbf{x}} &= \mathbf{B} \bar{\mathbf{x}} + \mathbf{c} \\ \mathbf{x}^{(k)} - \bar{\mathbf{x}} &= \mathbf{B}(\mathbf{x}^{(k-1)} - \bar{\mathbf{x}}) = \mathbf{B}^2(\mathbf{x}^{(k-2)} - \bar{\mathbf{x}}) = \dots = \mathbf{B}^k(\mathbf{x}^{(0)} - \bar{\mathbf{x}}) \\ \varepsilon^{(k)} &= \mathbf{B} \varepsilon^{(k-1)} = \dots = \mathbf{B}^k \varepsilon^{(0)} \\ \lim_{k \rightarrow \infty} \varepsilon^{(k)} = \mathbf{o} &\Leftrightarrow \lim_{k \rightarrow \infty} \mathbf{B}^k = \mathbf{O}. \end{aligned}$$

Věta: Nutná a postačující podmínka konvergence iterační metody tvaru $\mathbf{x}^{(k+1)} = \mathbf{B} \mathbf{x}^{(k)} + \mathbf{c}$ je $\rho(\mathbf{B}) < 1$. Pak pro vzdálenost od limity $\bar{\mathbf{x}}$ platí

$$\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|_v \leq \frac{\rho(\mathbf{B})}{1 - \rho(\mathbf{B})} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v.$$

(Počáteční odhad i vektorová norma mohou být libovolné.)

Věta: Postačující podmínka konvergence iterační metody tvaru $\mathbf{x}^{(k+1)} = \mathbf{B} \mathbf{x}^{(k)} + \mathbf{c}$ je $\|\mathbf{B}\|_M < 1$ (pro **nějakou** maticovou normu).

Jacobiova iterační metoda (JIM)

Vyjádříme matici \mathbf{A} ve tvaru

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

kde \mathbf{D} je diagonální, \mathbf{L} ostře dolní trojúhelníková a \mathbf{U} ostře horní trojúhelníková.

$$\begin{aligned} \mathbf{A} \mathbf{x} &= (\mathbf{D} + \mathbf{L} + \mathbf{U}) \mathbf{x} = \mathbf{D} \mathbf{x} + (\mathbf{L} + \mathbf{U}) \mathbf{x} = \mathbf{b}, \\ \mathbf{D} \mathbf{x} &= -(\mathbf{L} + \mathbf{U}) \mathbf{x} + \mathbf{b}, \\ \mathbf{x} &= \mathbf{D}^{-1} (-(\mathbf{L} + \mathbf{U}) \mathbf{x} + \mathbf{b}), \end{aligned}$$

volíme

$$\mathbf{x}^{(k+1)} = \underbrace{-\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})}_{\mathbf{B}_{\text{JIM}}} \mathbf{x}^{(k)} + \underbrace{\mathbf{D}^{-1} \mathbf{b}}_{\mathbf{c}_{\text{JIM}}}.$$

Předpokládáme, že hlavní diagonála matice \mathbf{A} neobsahuje žádný nulový prvek (toho lze dosáhnout výměnou řádků nebo sloupců).

Chceme, aby na diagonále byly „velké“ prvky.

Po složkách:

$$\begin{aligned} x_1 &= -\frac{1}{a_{1,1}} (a_{1,2} x_2 + a_{1,3} x_3 + \dots + a_{1,n} x_n) + \frac{b_1}{a_{1,1}} \\ x_2 &= -\frac{1}{a_{2,2}} (a_{2,1} x_1 + a_{2,3} x_3 + \dots + a_{2,n} x_n) + \frac{b_2}{a_{2,2}} \\ &\vdots \\ x_n &= -\frac{1}{a_{n,n}} (a_{n,1} x_1 + a_{n,2} x_2 + \dots + a_{n,n-1} x_{n-1}) + \frac{b_n}{a_{n,n}} \\ x_i^{(k+1)} &= -\frac{1}{a_{i,i}} \left(\sum_{j=1}^{i-1} a_{i,j} x_j^{(k)} + \sum_{j=i+1}^n a_{i,j} x_j^{(k)} \right) + \frac{b_i}{a_{i,i}}, \quad i = 1, 2, \dots, n. \end{aligned}$$

Podmínka ukončení: $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v < \varepsilon$.

Gaussova-Seidelova iterační metoda (GSM)

Každá už vypočtená složka vektoru $\mathbf{x}^{(k+1)}$ se ihned použije v dalším výpočtu.

$$x_i^{(k+1)} = -\frac{1}{a_{i,i}} \left(\sum_{j=1}^{i-1} a_{i,j} x_j^{(k+1)} + \sum_{j=i+1}^n a_{i,j} x_j^{(k)} \right) + \frac{b_i}{a_{i,i}}, \quad i = 1, 2, \dots, n.$$

Pro realizaci výpočtu stačí pouze jeden vektor \mathbf{x} .

Maticový tvar:

$$\begin{aligned} (\mathbf{D} + \mathbf{L}) \mathbf{x} &= -\mathbf{U} \mathbf{x} + \mathbf{b}, \\ \mathbf{x} &= -(\mathbf{D} + \mathbf{L})^{-1} (\mathbf{U} \mathbf{x} - \mathbf{b}), \\ \mathbf{x}^{(k+1)} &= \underbrace{-(\mathbf{D} + \mathbf{L})^{-1} \mathbf{U}}_{\mathbf{B}_{\text{GSM}}} \mathbf{x}^{(k)} + \underbrace{(\mathbf{D} + \mathbf{L})^{-1} \mathbf{b}}_{\mathbf{c}_{\text{GSM}}}. \end{aligned}$$

Věta: JIM (resp. GSM) konverguje pro libovolnou počáteční iteraci právě tehdy, když $\rho(\mathbf{B}_{\text{JIM}}) < 1$ (resp. $\rho(\mathbf{B}_{\text{GSM}}) < 1$).

Pokud $\rho(\mathbf{B}_{\text{JIM}}) = \rho(\mathbf{B}_{\text{GSM}}) = 0$, dostaneme teoreticky přesné řešení v konečném počtu kroků.

Může se stát, že konverguje pouze jedna z těchto metod (nebo žádná).

Problém: určení vlastních čísel matice.

Věta: Necht $\|\mathbf{B}_{\text{JIM}}\|_M < 1$ (resp. $\|\mathbf{B}_{\text{GSM}}\|_M < 1$), pak JIM (resp. GSM) konverguje a pro příslušné iterační posloupnosti získané z libovolné počáteční iterace $\mathbf{x}^{(0)} \in \mathbb{R}^n$ platí odhad

$$\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|_v \leq \frac{\|\mathbf{B}\|_M}{1 - \|\mathbf{B}\|_M} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v,$$

kde \mathbf{B} je příslušná iterační matice (normy matice a vektorů můžeme volit, ale musí být souhlasné).

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

$$|a_{i,i}| > \sum_{j=1, j \neq i}^n |a_{i,j}| \quad \text{pro každé } i = 1, 2, \dots, n.$$

Poznámka: Pro transponovanou matici dostaneme jinou podmínku, ale stejně užitečnou.

Věta: Necht matice $\mathbf{A} \in \mathbb{R}^{n,n}$ je ostře diagonálně dominantní, pak JIM i GSM konverguje pro libovolnou počáteční iteraci.

Poznámka: Ostře diagonálně dominantní je např. matice soustavy rovnic pro koeficienty kubického splinu, kdy navíc matice soustavy je řídká, takže složitost jedné iterace je úměrná n .

Definice: Matice $\mathbf{A} \in \mathbb{R}^{n,n}$ se nazývá **pozitivně definitní**, jestliže pro každý nenulový vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0.$$

Věta: Necht matice $\mathbf{A} \in R^{n,n}$ je symetrická a pozitivně definitní, pak GSM konverguje.

Poznámka: Symetrická a pozitivně definitní je např. matice soustavy normálních rovnic u metody nejmenších čtverců.

Superrelaxační metoda (SOR – Successive OverRelaxation method)

Konvergenci urychlí jakákoliv modifikace, která vede ke zmenšení spektrálního poloměru matice \mathbf{B} .

$$\begin{aligned} \mathbf{D} \mathbf{x} &= \mathbf{D} \mathbf{x} + \omega \overbrace{(-\mathbf{A} \mathbf{x} + \mathbf{b})}^{\mathbf{o}} = \\ &= \mathbf{D} \mathbf{x} + \omega ((-\mathbf{L} - \mathbf{D} - \mathbf{U}) \mathbf{x} + \mathbf{b}), \\ \mathbf{D} \mathbf{x} + \omega \mathbf{L} \mathbf{x} &= (1 - \omega) \mathbf{D} \mathbf{x} - \omega \mathbf{U} \mathbf{x} + \omega \mathbf{b}, \quad (\text{GSM: } \omega := 1) \\ (\mathbf{D} + \omega \mathbf{L}) \mathbf{x} &= [(1 - \omega) \mathbf{D} - \omega \mathbf{U}] \mathbf{x} + \omega \mathbf{b}, \\ \mathbf{x} &= (\mathbf{D} + \omega \mathbf{L})^{-1} [(1 - \omega) \mathbf{D} - \omega \mathbf{U}] \mathbf{x} + \omega \mathbf{b} = \\ &= (\mathbf{D} + \omega \mathbf{L})^{-1} [(1 - \omega) \mathbf{D} - \omega \mathbf{U}] \mathbf{x} + \omega (\mathbf{D} + \omega \mathbf{L})^{-1} \mathbf{b}, \\ \mathbf{x}^{(k+1)} &= \underbrace{(\mathbf{D} + \omega \mathbf{L})^{-1} [(1 - \omega) \mathbf{D} - \omega \mathbf{U}]}_{\mathbf{B}_\omega} \mathbf{x}^{(k)} + \underbrace{\omega (\mathbf{D} + \omega \mathbf{L})^{-1} \mathbf{b}}_{\mathbf{c}_\omega}, \end{aligned}$$

kde ω je **relaxační faktor**. (Pro $\omega = 1$ dostáváme GSM.) Po složkách:

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

Věta: Metoda SOR konverguje pro libovolnou počáteční iteraci právě tehdy, když $\rho(\mathbf{B}_\omega) < 1$.

Věta: (Ostrowského) Necht \mathbf{A} je symetrická matice s kladnými prvky na diagonále. Pak platí $\rho(\mathbf{B}_\omega) < 1$ právě tehdy, když matice \mathbf{A} je pozitivně definitní a $0 < \omega < 2$.

Poznámka: Často je lepší relaxační faktor nadhodnotit.

Jaký postup volit?

Poznámka: Složitost řešení soustavy n lineárních rovnic o n neznámých s jedním vektorem pravých stran pomocí GEM a dalších přímých metod je úměrná n^3 . Je-li matice soustavy trojúhelníková (speciálně při zpětném chodu GEM a v řešení soustavy při známém LU-rozkladu), vychází složitost úměrná n^2 . Rovněž vynásobení vektoru pravých stran známou inverzní maticí má složitost úměrnou n^2 . Je-li matice soustavy diagonální, je složitost řešení úměrná n .

U iteračních metod má v obecném případě jeden krok složitost úměrnou n^2 . Počet kroků závisí na požadované přesnosti a rychlosti konvergence (a ta na spektrálním poloměru iterační matice).

Úloha:

$$\mathbf{A} \mathbf{x} = \mathbf{B}$$

Poznámka: Matice \mathbf{B} zde reprezentuje pravé strany soustavy, nikoli iterační matici.

- \mathbf{A} plná (málo nulových prvků), \mathbf{B} málo sloupců: GEM nebo její modifikace
- \mathbf{A} plná, \mathbf{B} mnoho sloupců: výpočet \mathbf{A}^{-1} a řešení maticovým násobením nebo LU-rozklad
- \mathbf{A} řídká (mnoho nulových prvků) nepravidelně: GEM s výběrem hlavního prvku
- \mathbf{A} řídká pravidelně: iterační metody JIM, GSM, SOR

*

Literatura

[Navara, Němeček] Navara, M., Němeček, A.: *Numerické metody*. Skriptum ČVUT, Praha, 2005.

[Num. Recipes] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: *Numerical Recipes (The Art of Scientific Computing)*. Cambridge University Press, Cambridge, 1986.

[Num. Recipes] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: *Numerical Recipes (The Art of Scientific Computing)*. 2nd edition, Cambridge University Press, Cambridge, 1992.

[Handbook Lin. Alg.] Hogben, L. (ed.): *Handbook of Linear Algebra*. Chapman & Hall/CRC, Boca Raton/London/New York, 2007.