Numerické metody a statistika

Radek Kučera

VŠB-TU Ostrava

2016-2017

Iterační metody řešení soustav lineárních rovnic

Formulace úlohy

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

- Předpoklad: existuje jediné řešení $\boldsymbol{x}$
- Iterační postup řešení: počítá se posloupnost vektorů $\boldsymbol{x}^{(k)}$

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}$$

Příklad

$$\begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix}$$

$$\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{pmatrix} \Longrightarrow \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{pmatrix} \Longrightarrow \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} \ldots \Longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Přímé metody

- malé soustavy ($n = 100,\ 1000$)
- plné matice (téměř všechny prvky jsou nenulové)
- dobře podmíněné matice

Iterační metody

- velké soustavy ($n > 100,\ 1000$)
- řídké matice (obsahují hodně nulových prvků, ty se neukládají)
- špatně podmíněné matice (každou iteraci lze chápat jako počáteční)
- iterační zpřesnění řešení vypočítaného přímou metodou
- iterační výpočty lze zobecnit pro složitější úlohy (s omezením)

Obecná lineární iterační metoda

- soustavu rovnic $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ převedeme do iteračního tvaru:

$$\boldsymbol{x} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{d}$$

- zvolíme počáteční aproximaci: $\boldsymbol{x}^{(0)}$
- postupně počítáme jednotlivé aproximace $\boldsymbol{x}^{(k+1)}$:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{C}\boldsymbol{x}^{(k)} + \boldsymbol{d}, \qquad k = 0, 1, 2, \ldots$$

- sledujeme přitom ukončovací kritérium:

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\| \leq \varepsilon,$$

kde ε je předepsaná tolerance

Příklad 1 (konvergentní výpočet)

$$\left.\begin{array}{rcl} 4x_1 - x_2 + 2x_3 & = & -12 \\ 2x_1 + 5x_2 + x_3 & = & 5 \\ x_1 + x_2 - 3x_3 & = & -4 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{rcl} x_1 & = & \frac{1}{4}(-12 + x_2 - 2x_3) \\ x_2 & = & \frac{1}{5}(5 - 2x_1 - x_3) \\ x_3 & = & \frac{1}{-3}(-4 - x_1 - x_2) \end{array}\right.$$

Zde je:

$$\boldsymbol{C} = \begin{pmatrix} 0 & 1/4 & -2/4 \\ -2/5 & 0 & -1/5 \\ 1/3 & 1/3 & 0 \end{pmatrix}, \qquad \boldsymbol{d} = \begin{pmatrix} -12/4 \\ 1 \\ 4/3 \end{pmatrix}$$

Iterační vzorce:

$$\begin{array}{rcl} x_1^{(k+1)} & = & \frac{1}{4}(-12 + x_2^{(k)} - 2x_3^{(k)}) \\ x_2^{(k+1)} & = & \frac{1}{5}(5 - 2x_1^{(k)} - x_3^{(k)}) \\ x_3^{(k+1)} & = & \frac{1}{-3}(-4 - x_1^{(k)} - x_2^{(k)}) \end{array}$$

Příklad 1 (konvergentní výpočet)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\lVert \boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)} \rVert_R$
0	0	0	0	—
1	−3.000	1.0000	1.3333	1.3333
2	−3.4167	1.9333	0.6667	0.9333
3	−2.8500	2.2333	0.8389	0.5667
4	−2.8611	1.9722	1.1278	0.2889
5	−3.0708	1.9189	1.0370	0.2097
6	−3.0388	2.0209	0.9494	0.1020
7	−2.9694	2.0256	0.9940	0.0694
8	−2.9906	1.9890	1.0187	0.0367
9	−3.0121	1.9925	0.9995	0.0215
10	−3.0016	2.0050	0.9935	0.0125
11	−2.9955	2.0019	1.0011	$0.0077 < \varepsilon = 10^{-2}$
	-3	2	1	

Příklad 2 (divergentní výpočet)

$$\left.\begin{array}{rcl} 4x_1 - x_2 + 2x_3 & = & -12 \\ 2x_1 + 5x_2 + x_3 & = & 5 \\ x_1 + x_2 - 3x_3 & = & -4 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{rcl} x_1 & = & -12 - 3x_1 + x_2 - 2x_3 \\ x_2 & = & 5 - 2x_1 - 4x_2 - x_3 \\ x_3 & = & -4 - x_1 - x_2 + 4x_3 \end{array}\right.$$

Zde je:

$$\boldsymbol{C} = \begin{pmatrix} -3 & 1 & -2 \\ -2 & -4 & -1 \\ -1 & -1 & 4 \end{pmatrix}, \qquad \boldsymbol{d} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix}$$

Iterační vzorce:

$$\begin{array}{rcl} x_1^{(k+1)} & = & -12 - 3x_1^{(k)} + x_2^{(k)} - 2x_3^{(k)} \\ x_2^{(k+1)} & = & 5 - 2x_1^{(k)} - 4x_2^{(k)} - x_3^{(k)} \\ x_3^{(k+1)} & = & -4 - x_1^{(k)} - x_2^{(k)} + 4x_3^{(k)} \end{array}$$

Příklad 2 (divergentní výpočet)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\lVert \boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)} \rVert_R$
0	0	0	0	—
1	-12	5	-4	5
2	37	13	-13	49
3	-84	-108	-106	93
4	344	711	-236	819
5	139	-3291	-2003	205
6	286	14894	-4864	18185
7	23752	-55279	-34640	23466
	-3	2	1	

Jacobiova iterační metoda

- převod na iterační tvar podle Příkladu 1: z první rovnice vyjádříme x_1, ze druhé x_2, atd.
- obecný zápis po rovnicích: $\boldsymbol{A} = (a_{ij})$, $a_{ii} \neq 0$, $\boldsymbol{b} = (b_i)$

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

- obecný zápis maticově: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{A} = \boldsymbol{L} + \boldsymbol{D} + \boldsymbol{U}$

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

- iterační matice $\boldsymbol{C}_J = -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})$, $\boldsymbol{d}_J = \boldsymbol{D}^{-1}\boldsymbol{b}$

Gauss-Seidelova iterační metoda

Příklad 1 (pokračování)

Iterační vzorce Jacobiovy metody:

$$
\begin{aligned}
x_1^{(k+1)} &= \tfrac{1}{4}(-12 + x_2^{(k)} - 2x_3^{(k)}) \\
x_2^{(k+1)} &= \tfrac{1}{5}(5 - 2x_1^{(k)} - x_3^{(k)}) \\
x_3^{(k+1)} &= \tfrac{1}{-3}(-4 - x_1^{(k)} - x_2^{(k)})
\end{aligned}
$$

Iterační vzorce Gauss-Seidelovy metody:

$$
\begin{aligned}
x_1^{(k+1)} &= \tfrac{1}{4}(-12 + x_2^{(k)} - 2x_3^{(k)}) \\
x_2^{(k+1)} &= \tfrac{1}{5}(5 - 2x_1^{(k+1)} - x_3^{(k)}) \\
x_3^{(k+1)} &= \tfrac{1}{-3}(-4 - x_1^{(k+1)} - x_2^{(k+1)})
\end{aligned}
$$

Snahou je urychlit konvergenci iteračního výpočtu.

Příklad 1 (pokračování)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|_R$
0	0	0	0	—
1	-3.0000	2.2000	1.0667	3.0000
2	-2.9833	1.9800	0.9989	0.2200
3	-3.0044	2.0020	0.9992	0.0220
4	-2.9991	1.9998	1.0002	$0.0054 < \varepsilon = 10^{-2}$
	-3	2	1	

- obecný zápis po rovnicích: $\boldsymbol{A} = (a_{ij})$, $a_{ii} \neq 0$, $\boldsymbol{b} = (b_i)$

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

- obecný zápis maticově: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{A} = \boldsymbol{L} + \boldsymbol{D} + \boldsymbol{U}$

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

- iterační matice $\boldsymbol{C}_{GS} = -(\boldsymbol{L} + \boldsymbol{D})^{-1}\boldsymbol{U}$, $\boldsymbol{d}_{GS} = (\boldsymbol{L} + \boldsymbol{D})^{-1}\boldsymbol{b}$

Definice

Řekneme, že čtvercová matice $\boldsymbol{A} = (a_{ij})$ řádu n je ostře diagonálně dominantní, jestliže platí:

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

Věta (o konvergenci)

Jacobiova (a také Gauss-Seidelova) iterační metoda konverguje pro libovolnou počáteční aproximaci $\boldsymbol{x}^{(0)}$, jestliže matice soustavy $\boldsymbol{Ax} = \boldsymbol{b}$ je ostře diagonálně dominantní.

Poznámka

Pokud není matice soustavy ostře diagonálně dominantní, pak můžeme celou soustavu vhodně upravit.

Příklad

Daná soustava:

$$\begin{aligned} -8x_1 + 2x_2 + x_3 &= -1 \\ -2x_1 + x_2 - 3x_3 &= -9 \\ x_1 + x_2 + 3x_3 &= 12 \end{aligned}$$

Upravená soustava, která má stejné řešení a ostře diagonální matici:

$$\begin{aligned} -8x_1 + 2x_2 + x_3 &= -1 \\ -x_1 + 2x_2 &= 3 \\ x_1 + x_2 + 3x_3 &= 12 \end{aligned}$$

Konvergence iteračních metod

Definice

Nechť $\boldsymbol{A}$ je čtvercová matice řádu n. Číslo λ (obecně komplexní), pro které má soustava rovnic

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$$

nenulové řešení, se nazývá vlastní číslo matice $\boldsymbol{A}$. Odpovídající nenulové řešení $\boldsymbol{v} \in \mathbb{R}^n$ se nazývá vlastní vektor matice $\boldsymbol{A}$.

Příklad

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

- tři vlastní čísla: $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$
- tři lineárně nezávislé vlastní vektory:

$$\boldsymbol{v}_1 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Závěr

- vlastní čísla jsou kořeny charakteristického polynomu:

$$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I})$$

- odpovídající vlastní vektory lze určit řešením singulárních soustav:

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$$

- každá matice řádu n má právě n vlastních čísel (mohou být komplexní a násobná)
- počet lineárně nezávislých vlastních vektorů je nejvýše n

Příklad pro matice $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Věta

Vlastní čísla trojúhelníkové matice jsou její diagonální prvky.

Věta

Nechť λ, $\boldsymbol{v}$ jsou vlastní číslo a odpovídající vlastní vektor matice $\boldsymbol{A}$, $c \in \mathbb{R}$, $k \in \mathbb{N}$. Pak $c\lambda^k$ je vlastní číslo matice $c\boldsymbol{A}^k$ odpovídající vlastnímu vektoru $\boldsymbol{v}$.

$$c\boldsymbol{A}^k\boldsymbol{v} = c\lambda\boldsymbol{A}^{k-1}\boldsymbol{v} = \cdots = c\lambda^k\boldsymbol{v}$$

Věta

Nechť $\boldsymbol{A}$ je reálná a symetrická čtvercová matice. Platí:

1. všechna vlastní čísla jsou reálná
2. všechny vlastní vektory odpovídající různým vlastním číslům jsou ortogonální
3. k-násobnému vlastnímu číslu odpovídá k lineárně nezávislých vlastních vektorů, které lze zvolit tak, aby byly ortogonální

Používáme komplexně sdružená čísla a vektory: $\overline{\lambda}$, $\overline{\boldsymbol{v}}$

1.

$$\left.\begin{array}{lll} \boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} & \Longrightarrow & \overline{\boldsymbol{v}}^\top\boldsymbol{A}\boldsymbol{v} = \lambda\overline{\boldsymbol{v}}^\top\boldsymbol{v} \\ \boldsymbol{A}\overline{\boldsymbol{v}} = \overline{\lambda}\overline{\boldsymbol{v}} & \Longrightarrow & \boldsymbol{v}^\top\boldsymbol{A}\overline{\boldsymbol{v}} = \overline{\lambda}\boldsymbol{v}^\top\overline{\boldsymbol{v}} \end{array}\right\} \text{odečteme}$$

$$0 = (\lambda - \overline{\lambda}) \cdot \overline{\boldsymbol{v}}^\top\boldsymbol{v} \quad \Longrightarrow \quad \lambda = \overline{\lambda} \quad \Longrightarrow \quad \lambda_{Im} = 0$$

2.

$$\left.\begin{array}{lll} \boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} & \Longrightarrow & \boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{w}^\top\boldsymbol{v} \\ \boldsymbol{A}\boldsymbol{w} = \mu\boldsymbol{w} & \Longrightarrow & \boldsymbol{v}^\top\boldsymbol{A}\boldsymbol{w} = \mu\boldsymbol{v}^\top\boldsymbol{w} \end{array}\right\} \text{odečteme}$$

$$0 = (\lambda - \mu) \cdot \boldsymbol{w}^\top\boldsymbol{v} \quad \Longrightarrow \quad \boldsymbol{w}^\top\boldsymbol{v} = 0$$

Důsledek

Vlastní vektory u reálné symetrické matice lze zvolit tak, že tvoří v $\mathbb{R}^n$ bázi (která je dokonce ortogonální).

Výpočet vlastních čísel metodou LU-rozkladu

Definice

Čtvercové matice $\boldsymbol{A}$, $\boldsymbol{B}$ jsou podobné jestliže existuje regulární matice $\boldsymbol{C}$ taková, že platí:

$$\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$$

Věta

Podobné matice mají stejná vlastní čísla.

$$\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} \implies \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{C}^{-1}\boldsymbol{v} = \lambda\boldsymbol{C}^{-1}\boldsymbol{v} \implies \boldsymbol{B}\boldsymbol{w} = \lambda\boldsymbol{w},\ \ \boldsymbol{w} = \boldsymbol{C}^{-1}\boldsymbol{v}$$

Princip metody

LU-rozklad násobení

$$A_0 = A: \quad A_{k-1} = L_k U_k, \quad A_k = U_k L_k, \quad k = 1, 2, \ldots$$

- předokládejme konvergenci posloupností matic:

$$\{A_k\} \to \overline{A}, \quad \{L_k\} \to \overline{L}, \quad \{U_k\} \to \overline{U}$$

- všechny matice posloupnosti $\{A_k\}$ jsou podobné

$$A_k = U_k L_k = L_k^{-1} L_k U_k L_k = L_k^{-1} A_{k-1} L_k$$

- všechny matice posloupnosti $\{U_k\}$ jsou trojúhelníkové, a proto také $\overline{U}$ je trojúhelníková
- za určitých předpokladů je $\overline{A} = \overline{U}$
- vlastní čísla matice A jsou diagonální prvky $\overline{U}$

Věta o konvergenci obecné iterační metody

Zjistíme jak se chová iterační chyba: $\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}$

$$\left.\begin{array}{lcl} \boldsymbol{x} & = & \boldsymbol{C}\boldsymbol{x} + \boldsymbol{d} \\ \boldsymbol{x}^{(k+1)} & = & \boldsymbol{C}\boldsymbol{x}^{(k)} + \boldsymbol{d} \end{array}\right\} \begin{array}{ll} \text{odečteme:} & \boldsymbol{e}^{(k+1)} = \boldsymbol{C}\boldsymbol{e}^{(k)} \\ \text{odtud:} & \boldsymbol{e}^{(k)} = \boldsymbol{C}^k \boldsymbol{e}^{(0)} \end{array}$$

Předpoklad: $\boldsymbol{C}$ má vlastní čísla $\lambda_1, \dots, \lambda_n$ a vlastní vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$, které tvoří bázi v $\mathbb{R}^n$.

$$\begin{aligned} \boldsymbol{e}^{(0)} &= c_1 \boldsymbol{v}_1 + \dots + c_n \boldsymbol{v}_n \\ \boldsymbol{e}^{(k)} &= c_1 \boldsymbol{C}^k \boldsymbol{v}_1 + \dots + c_n \boldsymbol{C}^k \boldsymbol{v}_n = c_1 \lambda_1^k \boldsymbol{v}_1 + \dots + c_n \lambda_n^k \boldsymbol{v}_n \end{aligned}$$

Dostáváme:

$$\lim_{k\to\infty} \boldsymbol{e}^{(k)} = \boldsymbol{0} \iff \lim_{k\to\infty} \lambda_i^k = 0 \quad \forall i$$

což nastane, jestliže

$$|\lambda_i| < 1 \quad \forall i$$

Věta

Obecná lineární iterační metoda konverguje pro každou počáteční aproximaci $\boldsymbol{x}^{(0)}$ právě, když všechna vlastní čísla iterační matice $\boldsymbol{C}$ jsou v absolutní hodnotě menší než 1.

Věta

Obecná lineární iterační metoda konverguje pro každou počáteční aproximaci $\boldsymbol{x}^{(0)}$, jestliže $\|\boldsymbol{C}\| < 1$.

$$|\lambda| \cdot \|\boldsymbol{v}\| = \|\lambda \boldsymbol{v}\| = \|\boldsymbol{C}\boldsymbol{v}\| \leq \|\boldsymbol{C}\| \cdot \|\boldsymbol{v}\| \quad \Longrightarrow \quad |\lambda| \leq \|\boldsymbol{C}\| < 1$$

Důsledek: konvergence Jakobiovy metody

Použijeme řádkovou normu. Podmínka $\|\boldsymbol{C}_J\|_R < 1$, kde $\boldsymbol{C}_J = -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})$, je ekvivalentní tomu, že matice $\boldsymbol{A}$ je ostře diagonálně dominantní.

Souhrn

- obecná lineární iterační metoda
- Jacobiova a Gauss-Seidelova iterační metoda
- důkaz konvergence
- vlastní čísla a vlastní vektory matic
- výpočet vlastních čísel metoda LU-rozkladu

Zdroj

- skripta Numerické metody
- str. 66–91