

*Numerické metódy: riešenie sústavy lineárnych rovníc
vznikajúcej pri implicitnej schéme*

Beáta Stehlíková

Finančné deriváty, FMFI UK

- Sústava tvaru $\mathbf{A}x = b$ s maticou

$$\mathbf{A} = \begin{pmatrix} 1 + 2\gamma & -\gamma & 0 & \dots & 0 \\ -\gamma & 1 + 2\gamma & -\gamma & & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 \\ \vdots & & & & \\ 0 & \dots & 0 & -\gamma & 1 + 2\gamma \end{pmatrix}$$

pre $\gamma > 0$

- Zopakujeme si postup pomocou Gauss-Seidelovej metódy
- Ukážeme jej zovšeobecnenie - SOR metóda (jej modifikácia sa bude používať pri oceňovaní amerických opcií)

Opakovanie z numerických metód

- Všeobecná schéma iteračných metód riešenia sústavy lineárnych rovníc $\mathbf{A}x = b$:
 - sústavu $\mathbf{A}x = b$ prepíšeme do tvaru $x = \mathbf{T}x + g$ a definujeme iteračnú schému $x^{(k+1)} = \mathbf{T}x^{(k)} + g$
 - ak sú všetky vlastné čísla matice \mathbf{T} menšie ako 1, tak iteračná schéma konverguje k riešeniu pre ľubovoľný štartovací bod
 - pri zápise sa často využíva rozklad matice na jej diagonálnu (\mathbf{D}), dolnú trojuholníkovú (\mathbf{L}) a hornú trojuholníkovú (\mathbf{U}) časť: $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$
- Príklady - zopakujeme si ich princíp:
 - Jacobi: $x^{(k+1)} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})x^{(k)} + \mathbf{D}^{-1}b$
 - Gauss-Seidel $x^{(k+1)} = -(\mathbf{D} + \mathbf{L})^{-1}\mathbf{U}x^{(k)} + (\mathbf{D} + \mathbf{L})^{-1}b$

Gauss-Seidelova metóda: príklad

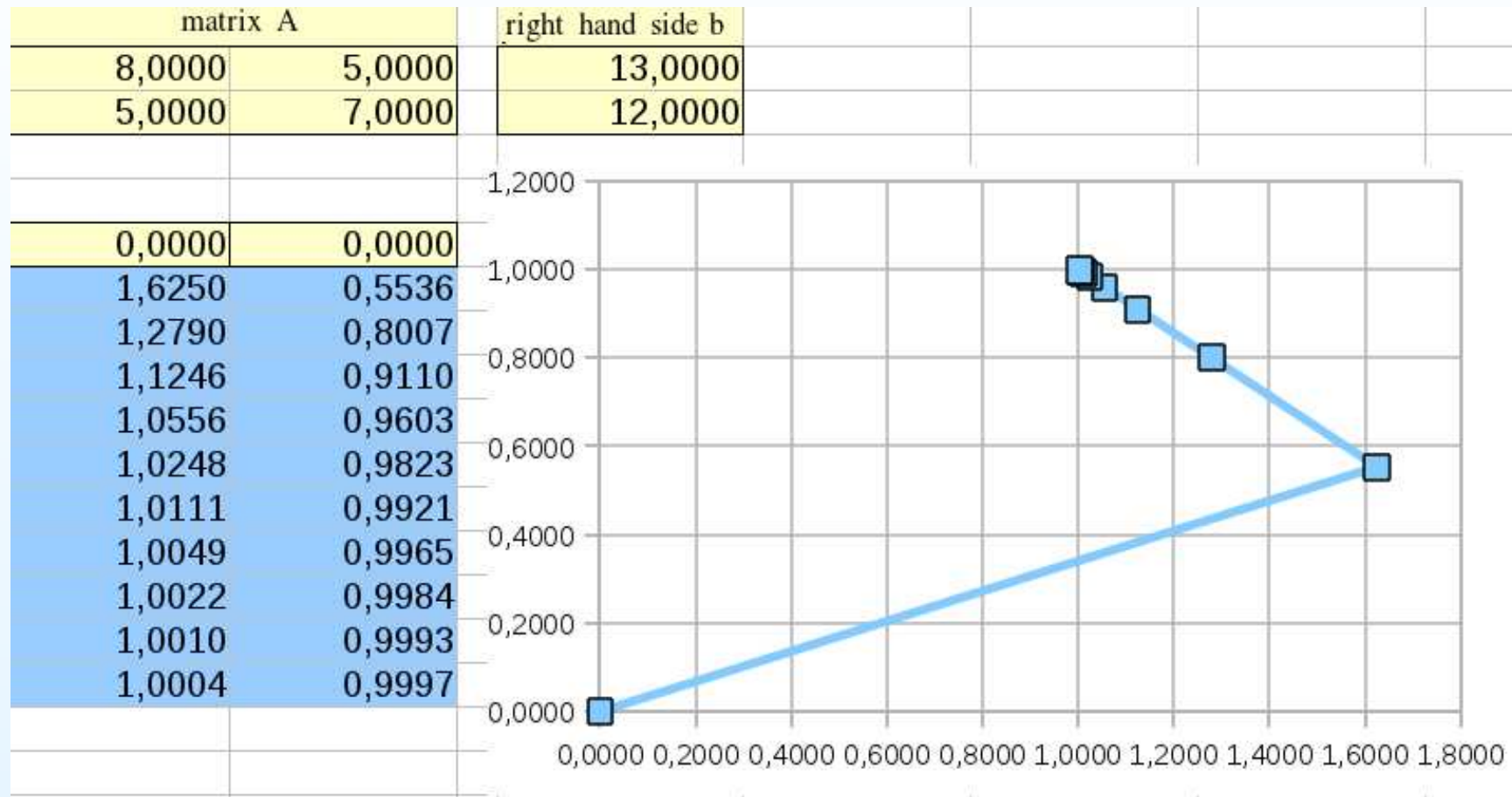
- Budeme riešiť úlohu

$$\begin{pmatrix} 8 & 5 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 12 \end{pmatrix},$$

ktorej presné riešenie je $(x_1, x_2)' = (1, 1)'$

- Použijeme Gauss-Seidelovu metódu, pričom budeme štartovať z bodu $(0, 0)'$

Gauss-Seidelova metóda: príklad



POZOROVANIE: Konverguje, ale pomaly, lebo robí príliš krátke kroky, aj keď správnym smerom \Rightarrow motivácia na vylepšenie

Vylepšenie Gauss-Seidelovej metódy

- Gauss-Seidelova metóda po zložkách:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j < i} a_{ij} x_j^{(k)} - \sum_{j > i} a_{ij} x_j^{(k-1)} \right]$$

- Pri výpočte $x_i^{(k)}$ máme
 - aktuálnu aproximáciu riešenia:

$$\left(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_i^{(k-1)}, \dots, x_n^{(k-1)} \right)$$

- rezíduum, t.j. rozdiel $b - \mathbf{A}x^{ap}$:

$$\mathbf{r}_i^{(k)} = \left(r_{1i}^{(k)}, \dots, r_{ni}^{(k)} \right)$$

Vylepšenie Gauss-Seidelovej metódy

- Gauss-Seidelov výpočet $x_i^{(k)}$ sa dá zapísať ako :

$$\begin{aligned}x_i^{(k)} &= \frac{1}{a_{ii}} \left[b_i - \sum_{j<i} a_{ij}x_j^{(k)} - \sum_{j>i} a_{ij}x_j^{(k-1)} \right] \\ &= x_i^{(k-1)} + \frac{1}{a_{ii}}r_{ii}^{(k)}\end{aligned}$$

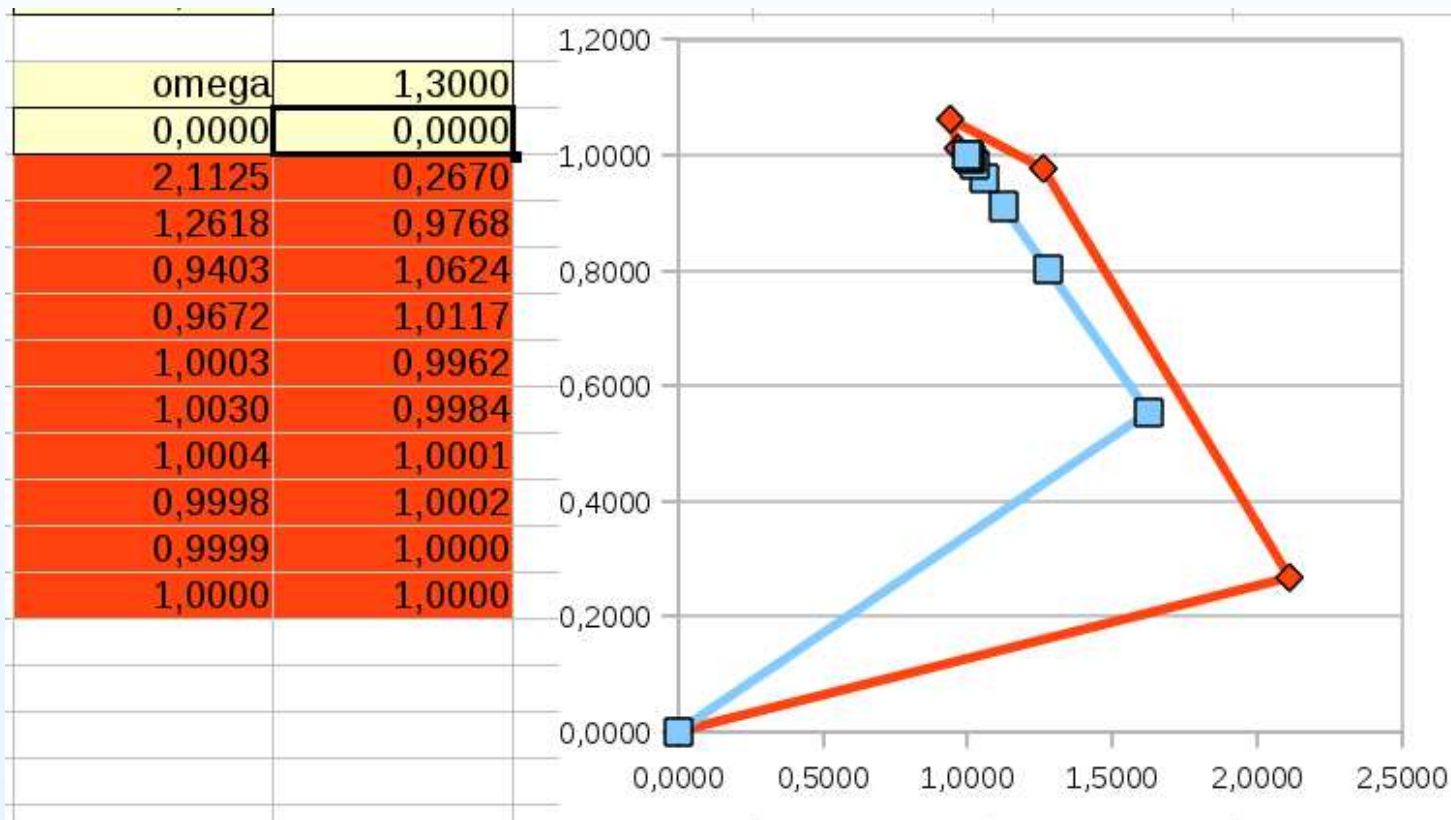
- Úprava Gauss-Seidelovej metódy v súlade s motiváciou z príkladu, tzv. successive over-relaxation (SOR) metóda:

$$x_i^{(k)} = x_i^{(k-1)} + \omega \frac{1}{a_{ii}}r_{ii}^{(k)}$$

pre $\omega > 1$ (pre $0 < \omega < 1$ hovoríme o under-relaxation; pre $\omega = 1$ dostávame pôvodnú Gauss-Seidelovu metódu)

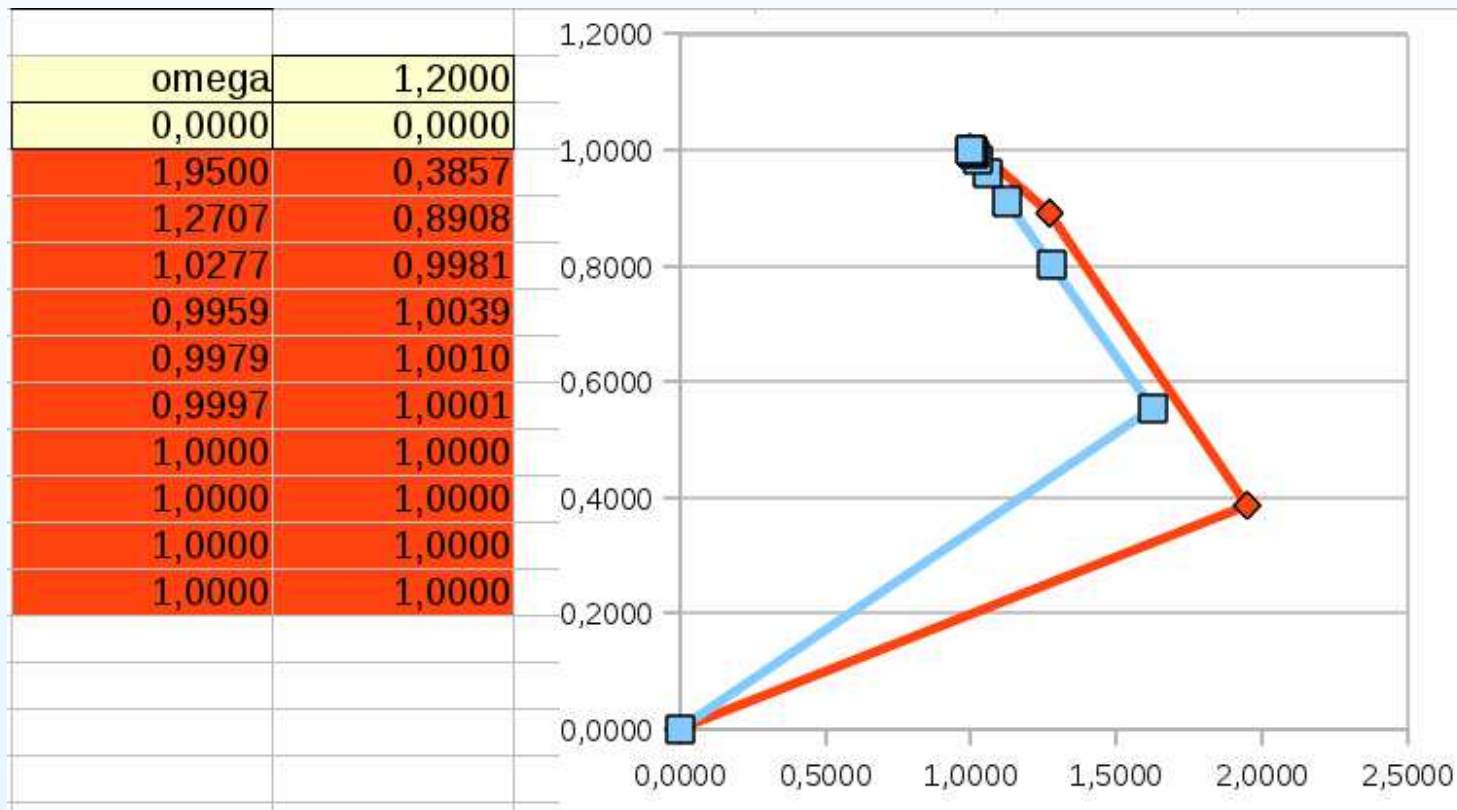
SOR metóda: príklad

Na predchádzajúci príklad použijeme SOR s $\omega = 1.3$:



SOR metóda: príklad

Ešte rýchlejšia konvergencia pre $\omega = 1.2$:



SOR metóda: príklad

Porovnanie Gauss-Seidelovej metódy (vľavo) s SOR metódou pre $\omega = 1.2$ (vpravo):

		omega	1,2000
0,0000	0,0000	0,0000	0,0000
1,6250	0,5536	1,9500	0,3857
1,2790	0,8007	1,2707	0,8908
1,1246	0,9110	1,0277	0,9981
1,0556	0,9603	0,9959	1,0039
1,0248	0,9823	0,9979	1,0010
1,0111	0,9921	0,9997	1,0001
1,0049	0,9965	1,0000	1,0000
1,0022	0,9984	1,0000	1,0000
1,0010	0,9993	1,0000	1,0000
1,0004	0,9997	1,0000	1,0000

OTÁZKY:

Ako zvoliť parameter ω ?

A od čoho vlastne závisí rýchlosť konvergencie metódy?

Rýchlosť konvergencie iteračnej schémy

- Majme iteračnú schému $x^{(k+1)} = \mathbf{T}x^{(k)} + g$ a presné riešenie x^* , ku ktorému schéma konverguje
- Platí:

$$\begin{aligned}\|x^{(k)} - x^*\| &= \|(\mathbf{T}x^{(k-1)} + g) - (\mathbf{T}x^* + g)\| \\ &= \|\mathbf{T} [x^{(k-1)} - x^*]\| \\ &= \|\mathbf{T} [(\mathbf{T}x^{(k-2)} + g) - (\mathbf{T}x^* + g)]\| \\ &= \|\mathbf{T}^2 [x^{(k-2)} - x^*]\| \\ &\quad \dots \\ &= \|\mathbf{T}^k [x^{(0)} - x^*]\| \leq \|\mathbf{T}^k\| \|x^{(0)} - x^*\|\end{aligned}$$

- Potrebujeme odhad normy $\|\mathbf{T}^k\|$
- Využije sa spektrálny polomer matice a jeho vlastnosti

Spektrálny polomer a maticové normy

- Nech \mathbf{M} je štvorcová matica
- Spektrálny polomer matice:

$$\rho(\mathbf{M}) = \max |\lambda_i|,$$

kde λ_i sú vlastné čísla matice \mathbf{M}

- Súvislosť spektrálneho polomeru s maticovými normami:
 - platí:

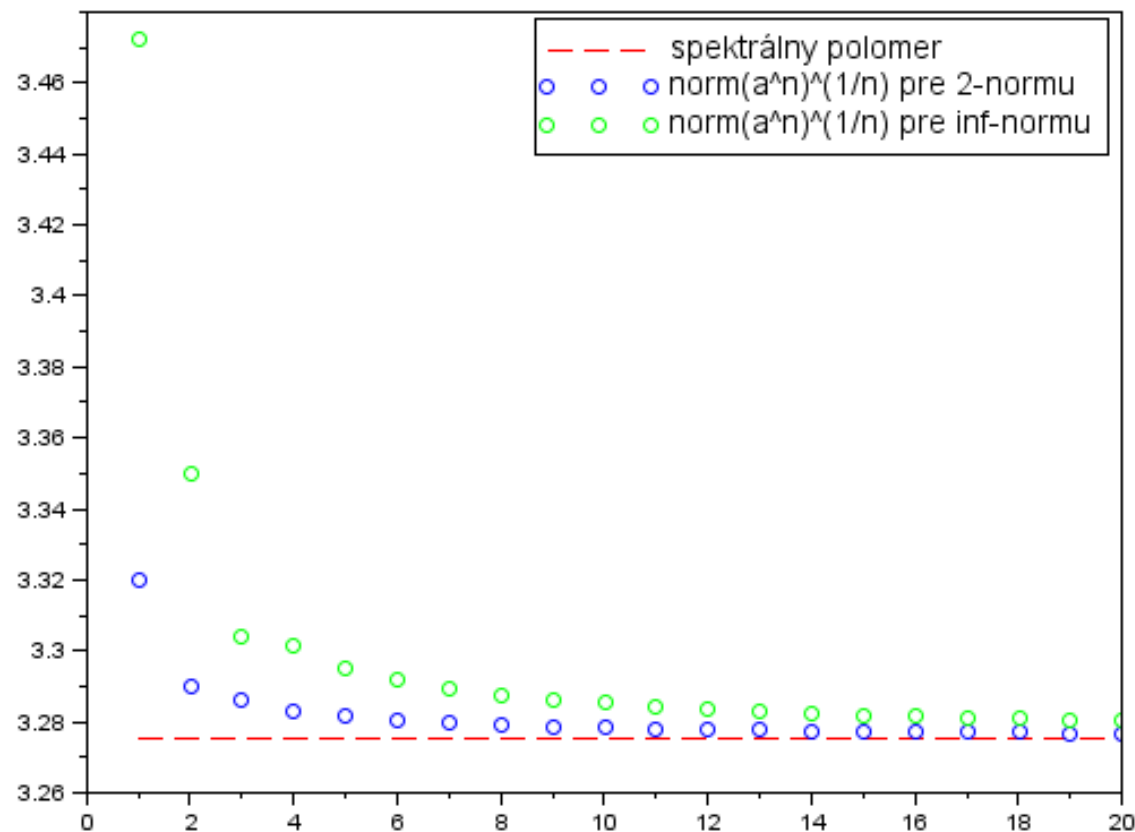
$$\lim_{n \rightarrow \infty} \|\mathbf{M}^n\|^{1/n} = \rho(\mathbf{M})$$

- preto sa pre veľké n dá použiť aproximácia

$$\|\mathbf{M}^n\|^{1/n} \sim \rho(\mathbf{M}) \Rightarrow \|\mathbf{M}^n\| \sim \rho(\mathbf{M})^n$$

Spektrálny polomer a maticové normy

Numerický príklad:



Rýchlosť konv. iteračnej schémy: pokračovanie

- Máme teda:

$$\|x^{(k)} - x^*\| \leq \|\mathbf{T}^k\| \|x^{(0)} - x^*\| \sim \rho(\mathbf{T})^k \|x^{(0)} - x^*\|$$

- Spektrálny polomer iteračnej matice \mathbf{T} teda
 - musí byť menší ako 1 - aby chyba konvergovala k nule, t.j. aby metóda konvergovala
 - mal by byť čo najmenší - aby bola konvergencia čo najrýchlejšia

Rýchlosť konv. iteračnej schémy: príklad

- Riešili sme sústavu:

$$\begin{pmatrix} 8 & 5 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 12 \end{pmatrix},$$

- SOR metóda v maticovom zápise:

$$x^{(k+1)} = (\mathbf{D} + \omega\mathbf{L})^{-1}[(1 - \omega)\mathbf{D} - \omega\mathbf{U}]x^{(k)} + \omega(\mathbf{D} + \omega\mathbf{L})^{-1}b$$

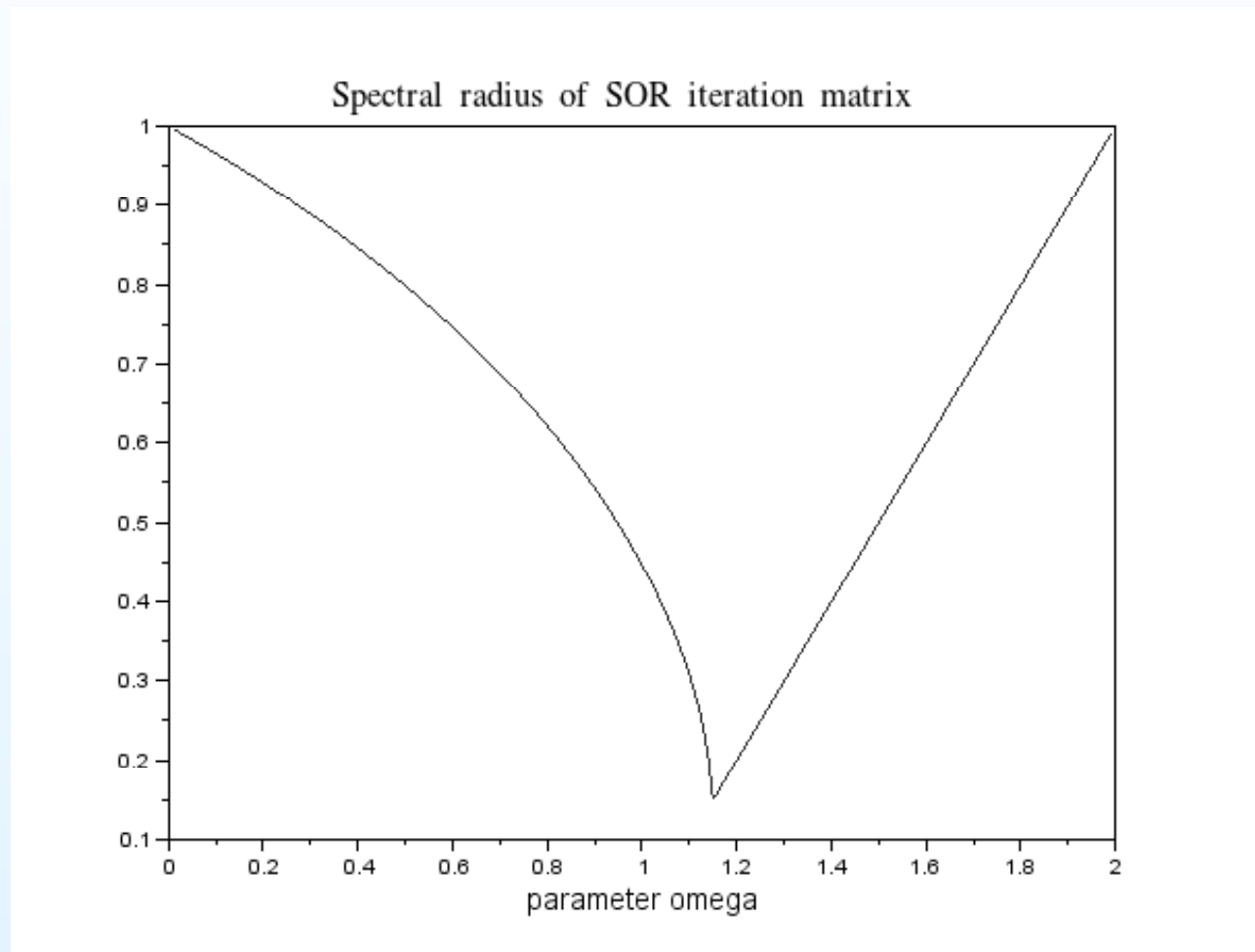
- V našom prípade teda iteračná matica je

$$\mathbf{T} = \begin{pmatrix} 8 & 0 \\ 5\omega & 7 \end{pmatrix}^{-1} \begin{pmatrix} 8(1 - \omega) & -5\omega \\ 0 & 7(1 - \omega) \end{pmatrix}$$

a potrebujeme závislosť maximálnej absolútnej hodnoty vlastného čísla od ω

Rýchlosť konv. iteračnej schémy: príklad

Výsledok:



Použili sme: $\omega = 1$ (Gauss-Seidel), $\omega = 1.3$, $\omega = 1.2$

Rýchlosť konvergenencie SOR schémy

Niekoľko užitočných tvrdení (bez dôkazov):

1. [Kahan] Nech $a_{ii} \neq 0$. Potom $\rho(\mathbf{T}_\omega) \geq |\omega - 1|$.
2. [Ostrowski-Reich] Nech A je kladne definitná a nech $0 < \omega < 2$. Potom SOR metóda konverguje pre ľubovoľný štartovací bod
3. Nech A je kladne definitná a trojdiagonálna. Potom $\rho(\mathbf{T}_{gs}) = \rho(\mathbf{T}_j)^2$ a optimálna voľba parametra ω pre SOR metódu je

$$\omega = \frac{2}{1 + \sqrt{1 - [\rho(\mathbf{T}_j)]^2}}.$$

Pri tejto voľbe platí: $\rho(\mathbf{T}_\omega) = \omega - 1$.

Označenie:

$\mathbf{T}_j, \mathbf{T}_{gs}, \mathbf{T}_\omega$ - iteračné matice Jacobiho, Gauss-Seidelovej a SOR metódy.

Rýchlosť konvergenzie SOR schémy

Dôsledky:

- Z tvrdenia 1 vyplýva, že pre $\omega \notin (0, 2)$ je $\rho(\mathbf{T}_\omega) \geq 1$. Podmienka $\omega \in (0, 2)$ je teda nutnou podmienkou konvergenzie metódy.
- Tvrdenie 2 udáva triedu matíc, pre ktorú je podmienka $\omega \in (0, 2)$ aj postačujúcou podmienkou konvergenzie.

Tvrdenie 3 pre náš príklad

Matica našej sústavy je kladne definitná a trojdiagonálna

$$\begin{pmatrix} 8 & 5 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 12 \end{pmatrix}$$

JACOBIHO METÓDA:

- iteračná matica:

$$\mathbf{T}_j = -\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U}) = -\begin{pmatrix} 8 & 0 \\ 0 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 5 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{5}{8} \\ -\frac{5}{8} & 0 \end{pmatrix}$$

- vlastné čísla matice \mathbf{T}_j : $\lambda_1 = \frac{5}{2\sqrt{14}}$, $\lambda_2 = -\frac{5}{2\sqrt{14}}$
- spektrálny polomer: $\rho(\mathbf{T}_j) = \frac{5}{2\sqrt{14}}$

Tvrdenie 3 pre náš príklad

GAUSS-SEIDELOVA METÓDA:

- iteračná matica:

$$\mathbf{T}_{gs} = -(\mathbf{D} + \mathbf{L})^{-1}\mathbf{U} = - \begin{pmatrix} 8 & 0 \\ 5 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{5}{8} \\ 0 & \frac{25}{56} \end{pmatrix}$$

- vlastné čísla matice \mathbf{T}_{gs} : $\lambda_1 = 0, \lambda_2 = \frac{25}{56}$
- spektrálny polomer: $\rho(\mathbf{T}_{gs}) = \frac{25}{56}$

NAOZAJ PLATÍ:

$$\begin{aligned} \rho(\mathbf{T}_{gs}) &= \rho(\mathbf{T}_j)^2 \\ \frac{25}{56} &= \left(\frac{5}{2\sqrt{14}} \right)^2 \end{aligned}$$

Tvrdenie 3 pre náš príklad

SOR METÓDA:

- Optimálna hodnota parametr ω podľa tvrdenia 3:

$$\omega = \frac{2}{1 + \sqrt{1 - [\rho(\mathbf{T}_j)]^2}} = \frac{2}{1 + \sqrt{1 - \frac{25}{56}}} = \frac{2\sqrt{56}}{\sqrt{56} + \sqrt{31}} \approx 1.147$$

- Príslušná hodnota spektrálneho polomeru iteračnej matice podľa tvrdenia 3:

$$\rho(\mathbf{T}_\omega) = \omega - 1 = \frac{\sqrt{56} - \sqrt{31}}{\sqrt{56} + \sqrt{31}} \approx 0.147$$

Tvrdenie 3 pre náš príklad

- Vypočítame spektrálny polomer pre dané ω priamo:

```
omega:2*sqrt(56)/(sqrt(56)+sqrt(31)) $
```

```
pom1:invert(matrix([8,0],[5*omega,7])) $
```

```
pom2:matrix([8*(1-omega),-5*omega],[0,7*(1-omega)]) $
```

```
t:pom1.pom2;
```

$$\begin{bmatrix} 1 - \frac{4\sqrt{14}}{\sqrt{31} + 2\sqrt{14}} & -\frac{5\sqrt{14}}{2(\sqrt{31} + 2\sqrt{14})} \\ -\frac{40\left(1 - \frac{4\sqrt{14}}{\sqrt{31} + 2\sqrt{14}}\right)}{\sqrt{14}(\sqrt{31} + 2\sqrt{14})} - \frac{4\sqrt{14}}{\sqrt{31} + 2\sqrt{14}} + \frac{100}{(\sqrt{31} + 2\sqrt{14})^2} + 1 & \end{bmatrix}$$

```
eigenvalues(t);
```

$$\left[\left[\frac{25}{4\sqrt{14}\sqrt{31} + 87} \right], [2] \right]$$

Dostávame teda: $\rho(\mathbf{T}_\omega) = \frac{25}{4\sqrt{14}\sqrt{31} + 87}$

Tvrdenie 3 pre náš príklad

- Zhoduje sa to s výsledkom z tvrdenia 3:

```
rho:25/(4*sqrt(14)*sqrt(31)+87) $  
float(rho);  
.1467733350400546  
  
ratsimp(rho-((sqrt(56)-sqrt(31))/(sqrt(56)+sqrt(31))));  
0
```

(najprv numericky, potom zhoda dvoch presných čísel)

Poznámky k použitým príkazom v softvéri **wxMaxima**:

\$ za príkazom: výsledok sa nevypíše

float: numerická hodnota

ratsimp: zjednodušenie výrazu

eigenvalues: vráti vlastné čísla matice a ich násobnosti, napr. $[[1/3, 1/2], [1, 2]]$

by znamenalo vlastné číslo $1/3$ s násobnosťou 1 a vlastné číslo $1/2$ s násobnosťou 2

Tvrdenie 3 pre náš príklad

- Získanú optimálnu hodnotu $\omega \approx 1.147$ a príslušnú hodnotu $\rho(\mathbf{T}_\omega) \approx 0.147$ porovnáme s priamym výpočtom zo str. 17:

