

*XIII. Numerical methods: System of linear equations  
arising from the implicit scheme*

Beáta Stehlíková

Financial derivatives, winter term 2014/2015

Faculty of Mathematics, Physics and Informatics  
Comenius University, Bratislava

## Gauss-Seidel method: example

---

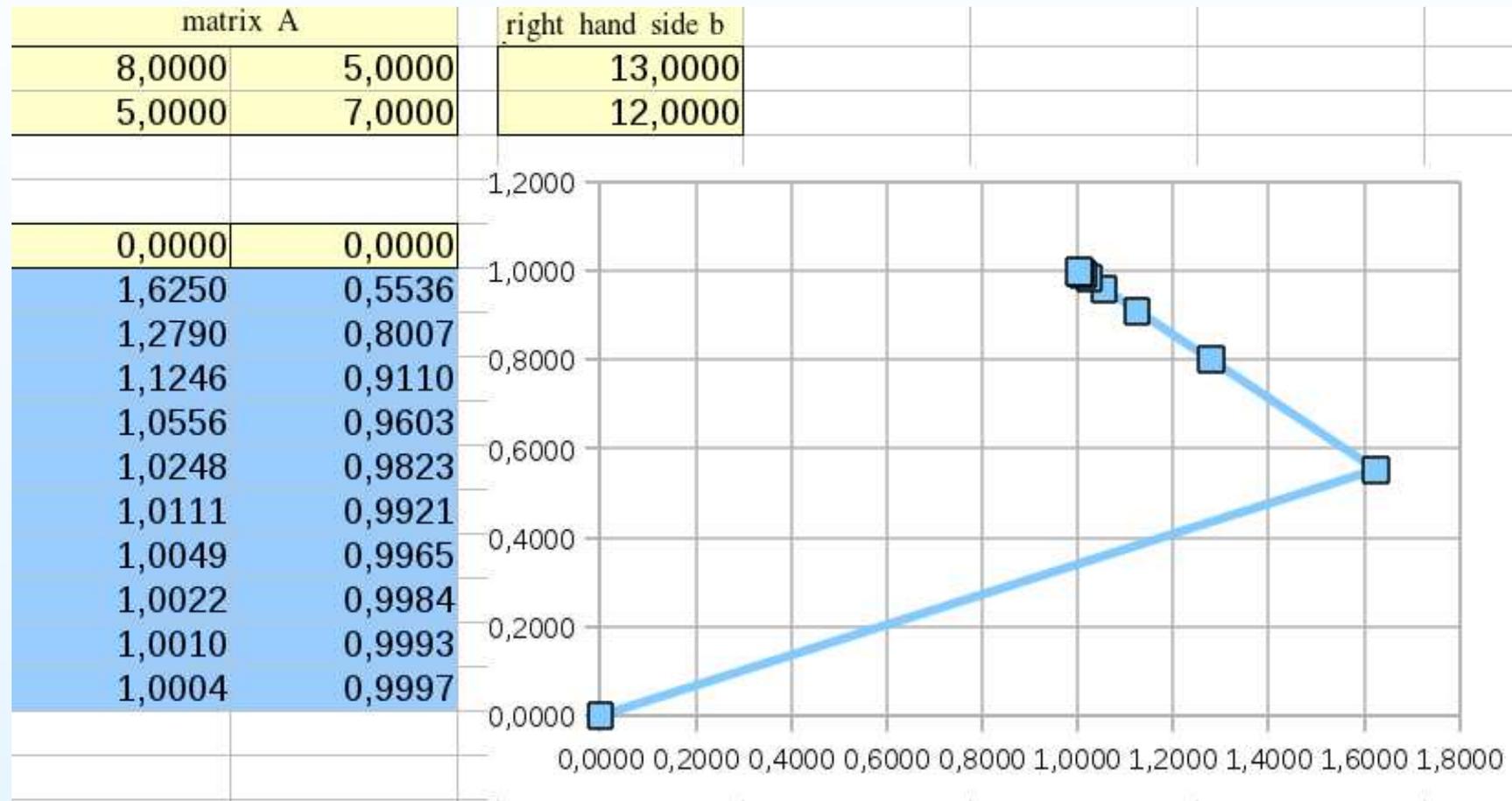
- Consider the system

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

which has the exact solution  $(x_1, x_2)' = (1, 1)'$

- We will use Gauss-Seidel method and start from  $(0, 0)'$

# Gauss-Seidel method: example



**OBSERVATION:** It converges, but slowly, because it make too short steps, although in the right direction  $\Rightarrow$  motivation for a modification of the method

# Gauss-Seidel method: example

- Gauss-Seidel method - by coordinates:

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

- When computing  $x_i^{(k)}$ , we have
  - current approximation of the solution:

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

- residual, i.e., the difference  $b - \mathbf{A}x^{ap}$ :

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

# Modification of the Gauss-Seidel method

- Gauss-Seidel computation of  $x_i^{(k)}$  can be written as:

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

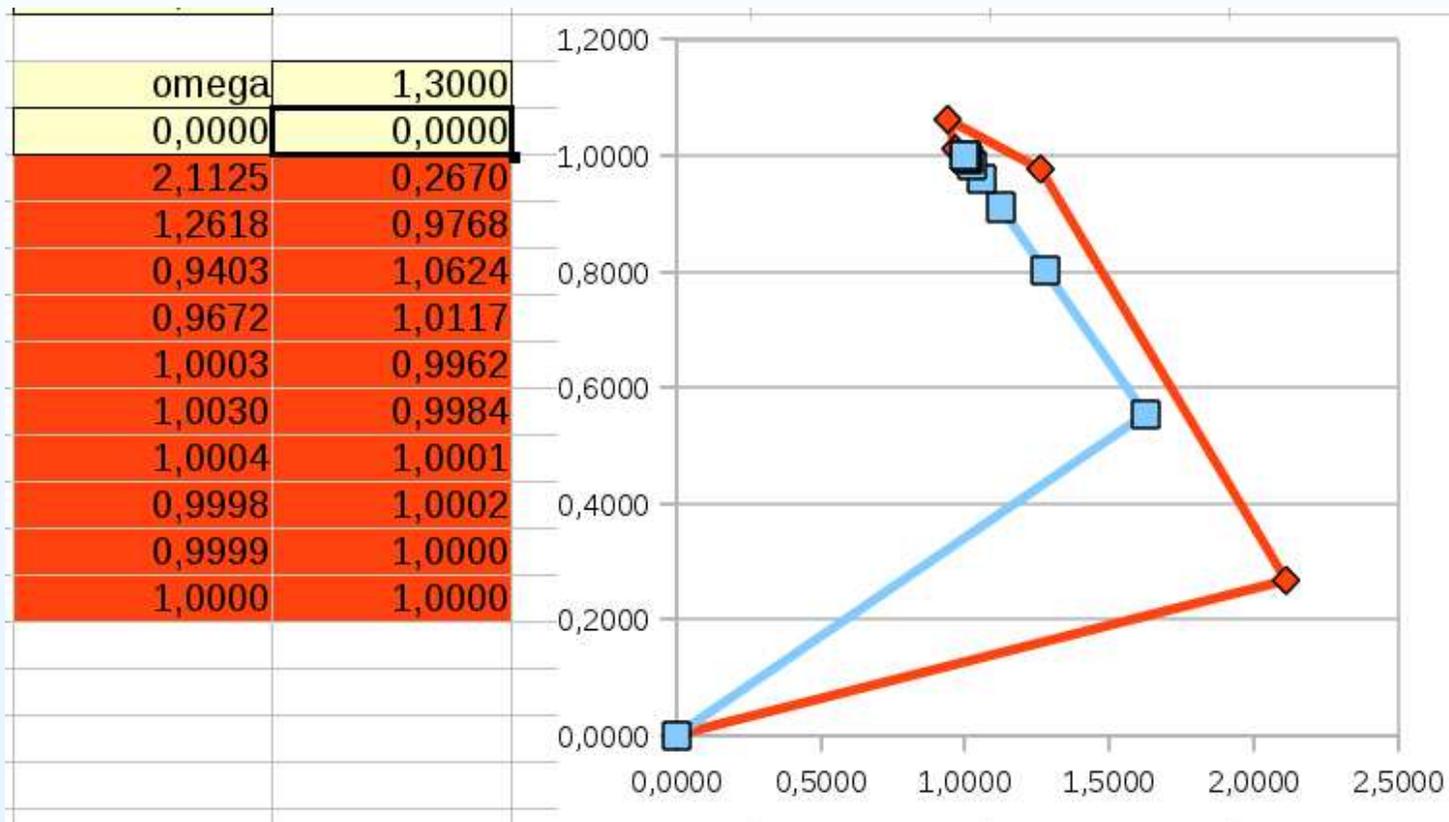
- Modification of the Gauss-Seidel method according to our earlier observation, so called successive over-relaxation (SOR) method:

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

for  $\omega > 1$  (for  $0 < \omega < 1$  it is, in fact, under-relaxation; for  $\omega = 1$  we have the original Gauss-Seidel method)

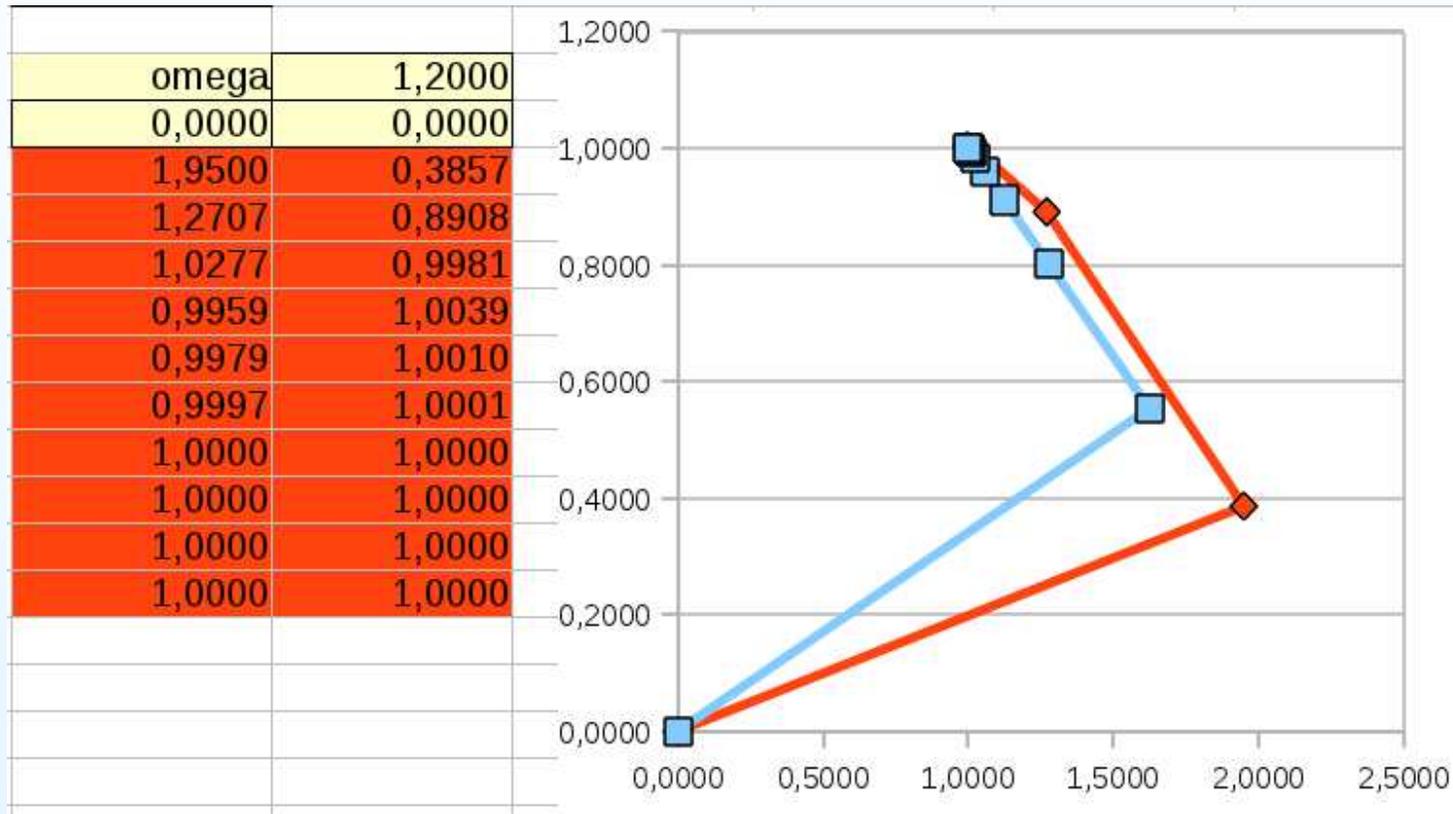
# SOR method: example

Example for the beginning of the lecture; this time we use SOR with  $\omega = 1.3$ :



# SOR method: example

Even better convergence for  $\omega = 1.2$ :



## SOR method: example

Comparison of Gauss-Seidel (left) and SOR method for  $\omega = 1.2$  (right):

		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

QUESTION:

How to choose  $\omega$ ?

What are the factors influencing the speed of convergence?

# Iteration schemes: speed of convergence

- Consider a general iteration scheme  $x^{(k+1)} = \mathbf{T}x^{(k)} + g$  with exact solution  $x^*$ , to which the scheme converges
- We have:

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

- We need to estimate the norm  $\|\mathbf{T}^k\|$
- We use spectral radius of a matrix and its properties

# Spectral radius and matrix norms

---

- Let  $\mathbf{M}$  be a square matrix
- Spectral radius of this matrix:

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

where  $\lambda_i$  are eigenvalues of  $\mathbf{M}$

- Relation of the spectral radius and matrix norms:
  - the following holds:

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

- therefore for large  $n$  we can use an approximation

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

# Iteration schemes: speed of convergence

---

- Therefore we have:

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

- Hence the spectral radius of the matrix  $\mathbf{T}$ 
  - has to be less than 1 - so that the error converges to zero and the method converges
  - should be as small as possible - to have the speed of convergence as high as possible

## Iter. schemes: speed of convergence - example

- We had the system:

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

- SOR method in matrix form  $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$ :

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

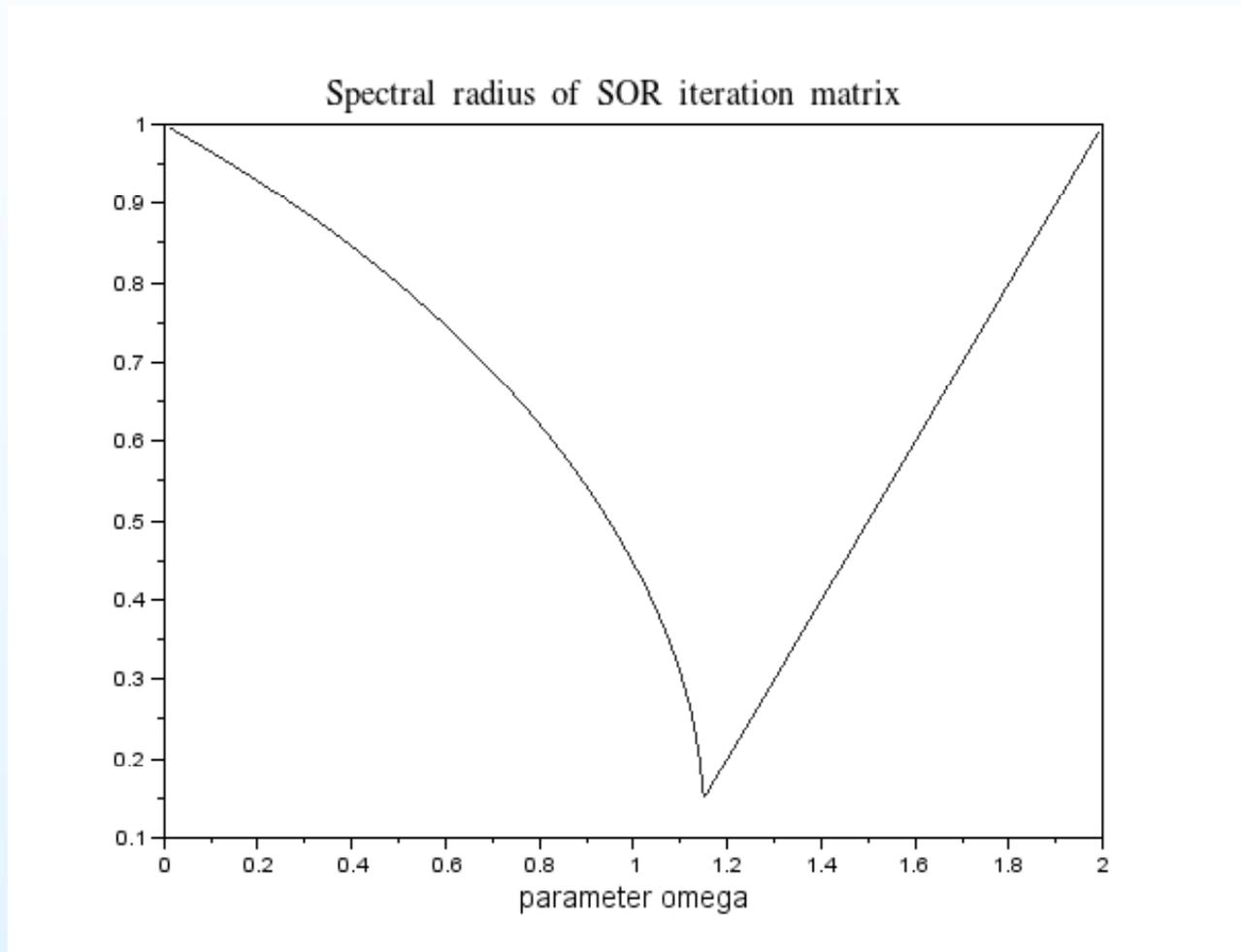
- Hence, in our case the iteration matrix is

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

and we need the dependence of the maximal absolute value of an eigenvalue on  $\omega$

# Iter. schemes: speed of convergence - example

Result:



We have used:  $\omega = 1$  (Gauss-Seidel),  $\omega = 1.3$ ,  $\omega = 1.2$

# Iteration schemes: speed of convergence

Some useful theorems (without proofs):

1. [Kahan] Let  $a_{ii} \neq 0$ . Then  $\rho(\mathbf{T}_\omega) \geq |\omega - 1|$ .
2. [Ostrowski-Reich] Let  $A$  be positive definite and let  $0 < \omega < 2$ . Then the SOR converges for any initial point.
3. Let  $A$  be a positive definite tridiagonal matrix. Then  $\rho(\mathbf{T}_{gs}) = \rho(\mathbf{T}_j)^2$  and the optimal choice of parameter  $\omega$  for the SOR method is

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

For this choice, we have:  $\rho(\mathbf{T}_\omega) = \omega - 1$ .

Notation:

$\mathbf{T}_j, \mathbf{T}_{gs}, \mathbf{T}_\omega$  - iteration matrices of Jacobi, Gauss-Seidel and SOR methods.

With the decomposition  $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$  we express  $\mathbf{T}_j$  as  $-\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$

# SOR method: speed of convergence

---

## Corollaries:

- From Theorem 1: for  $\omega \notin (0, 2)$  we have  $\rho(\mathbf{T}_\omega) \geq 1$ . Condition  $\omega \in (0, 2)$  is therefore a necessary condition for convergence.
- Theorem 2 gives a class of matrices, for which the condition  $\omega \in (0, 2)$  is also a sufficient condition for convergence.

## Theorem 3 and our example

The matrix of our system is positive definite and tridiagonal:

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

JACOBI METHOD:

- iteration matrix:

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

- eigenvalues of  $\mathbf{T}_j$ :  $\lambda_1 = \frac{5}{2\sqrt{14}}$ ,  $\lambda_2 = -\frac{5}{2\sqrt{14}}$
- spectral radius:  $\rho(\mathbf{T}_j) = \frac{5}{2\sqrt{14}}$

## Theorem 3 and our example

### GAUSS-SEIDEL METHOD:

- iteration matrix:

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

- eigenvalues of  $\mathbf{T}_{gs}$ :  $\lambda_1 = 0, \lambda_2 = \frac{25}{56}$
- spectral radius:  $\rho(\mathbf{T}_{gs}) = \frac{25}{56}$

INDEED WE HAVE:

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

## Theorem 3 and our example

---

### SOR METHOD:

- Optimal value of parameter  $\omega$  according to Theorem 3:

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

- Corresponding value of spectral radius of the iteration matrix according to Theorem 3:

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

## Theorem 3 and our example

- We compute the spectral radius for the given  $\omega$  directly:

```
(%i1) omega:2*sqrt(56)/(sqrt(56)+sqrt(31)) $  
  
m1:invert(matrix([8,0],[5*omega,7])) $  
m2:matrix([8*(1-omega),-5*omega],[0,7*(1-omega)]) $  
t:m1.m2;  
  
(%o4) 
$$\begin{bmatrix} 1 - \frac{4\sqrt{14}}{\sqrt{31}+2\sqrt{14}} & -\frac{5\sqrt{14}}{2(\sqrt{31}+2\sqrt{14})} \\ -\frac{40\left(1 - \frac{4\sqrt{14}}{\sqrt{31}+2\sqrt{14}}\right)}{\sqrt{14}(\sqrt{31}+2\sqrt{14})} - \frac{4\sqrt{14}}{\sqrt{31}+2\sqrt{14}} + \frac{100}{(\sqrt{31}+2\sqrt{14})^2} + 1 & \end{bmatrix}$$
  
  
(%i5) eigenvalues(t);  
(%o5) 
$$\left[ \left[ \frac{25}{4\sqrt{14}\sqrt{31}+87} \right], [2] \right]$$

```

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

## Theorem 3 and our example

- This agrees with the result given in Theorem 3:

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

(firstly numerically, then equality of two exact numbers)

Remarks on **wxMaxima** commands:

\$ at the end: the result of the computation is not printed

float: numerical value

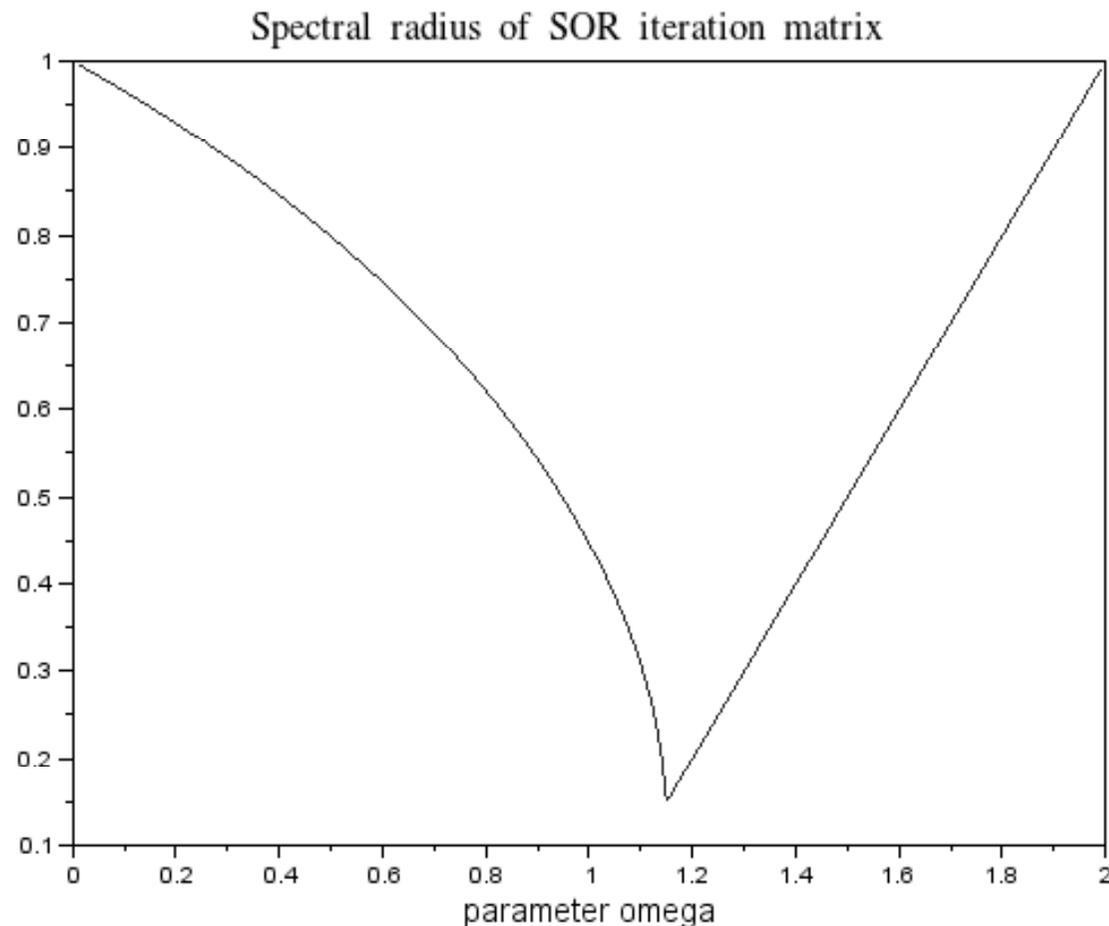
ratsimp: simplifies the expression

eigenvalues: returns eigenvalues and their multiplicity, e.g.,  $[[1/3, 1/2], [1, 2]]$

means eigenvalue  $1/3$  with multiplicity 1 and eigenvalue with multiplicity 2

## Theorem 3 and our example

- Comparison of the optimal value  $\omega \approx 1.147$  which we obtained and the corresponding value of  $\rho(\mathbf{T}_\omega) \approx 0.147$  with the application of the formula from p. 17:



## Exercise

---

- Repeat for another tridiagonal positive definite matrix:
  - computation of the iterations
  - spectral radius, the optimal choice of the parameter  $\omega$

# SOR method and pricing derivatives

---

- EXERCISES SESSION:  
SOR method for solving the system of linear equations arising in the implicit scheme
  - practical implementation of a given  $\omega$
  - choice of  $\omega$  - numerical computation of the Jacobi matrix spectrum, formulation of the theorem about its spectrum and spectral radius
  - optimal  $\omega$  for different mesh grids - observations
- HOMEWORK 1:  
Show that this matrix is positive definite.  
(Hence the convergence theorems, which we have just seen, apply to our problem.)

# Direct methods for solving systems of lin. eq.

---

- Direct methods for systems of linear equations - Gaussian elimination method and its modifications, LU decomposition
- Tridiagonal, diagonally dominant matrix  $\Rightarrow$  a convenient way of solving the system is the LU decomposition
- HOMEWORK 2 - LU decomposition method
  - What is the LU decomposition? How it can be used to solve a system of linear equations?
  - Show that the uniqueness of the decomposition follows from the matrix being diagonally dominant.
  - Why is this convenient, when solving a system with a tridiagonal matrix? (you do not need to memorize the formulae, but you need to show the principle of the computation on an example and explain why we do not obtain zeros in denominators of fractions.)

References for HW2: e.g., [Ševčovič, Stehlíková, Mikula]