Iterative Solvers for Large Linear Systems

Part IIc: Relaxation Schemes

Andreas Meister

University of Kassel, Department of Analysis and Applied Mathematics

Outline

- Basics of Iterative Methods
- Splitting-schemes
  - Jacobi- u. Gauß-Seidel-scheme
  - Relaxation methods
- Methods for symmetric, positive definite Matrices
  - Method of steepest descent
  - Method of conjugate directions
  - CG-scheme

Outline

- Multigrid Method
  - Smoother, Prolongation, Restriction
  - Twogrid Method and Extension
- Methods for non-singular Matrices
  - GMRES
  - BiCG, CGS and BiCGSTAB
- Preconditioning
  - ILU, IC, GS, SGS, ...

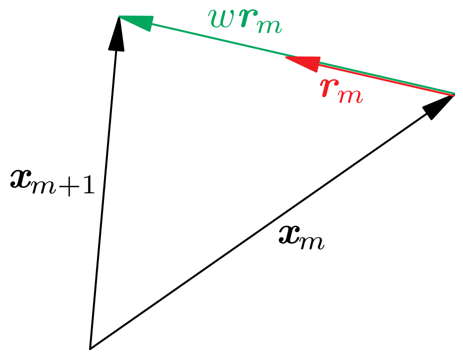
Relaxation method

Utilizing

$$\begin{aligned} x_{m+1} &= B^{-1}(B-A)x_m + B^{-1}b \\ &= x_m + \underbrace{B^{-1}(b - Ax_m)}_{=:\, r_m \text{ (correction vector)}} \end{aligned}$$

<u>Aim of the relaxation</u>

Improving the rate of convergence
by means of weighting the correction vector.



Utilizing $\omega \in \mathbb{R}^+$ (relaxation parameter) leads to

$$x_{m+1} = x_m + \omega r_m = \underbrace{(I - \omega B^{-1} A)}_{= M(\omega)} x_m + \underbrace{\omega B^{-1}}_{= N(\omega)} b$$

Optimality of the relaxation parameter: $\omega_{opt} = \arg \min_{\omega \in \mathbb{R}^+} \rho\,(M(\omega))$

Gauß-Seidel scheme:

Correction of each component

Jacobi relaxation method

Jacobi scheme:

$$
\begin{aligned}
x_{m+1} &= D^{-1}(D - A)x_m + D^{-1}b \\
&= x_m + D^{-1}(b - Ax_m)
\end{aligned}
$$

Jacobi relaxation method:

$$
\begin{aligned}
x_{m+1} &= x_m + \omega D^{-1}(b - Ax_m) \\
&= (I - \omega D^{-1}A)x_m + \omega D^{-1}b
\end{aligned}
$$

Jacobi relaxation method

Optimal relaxation parameter

Assumption:

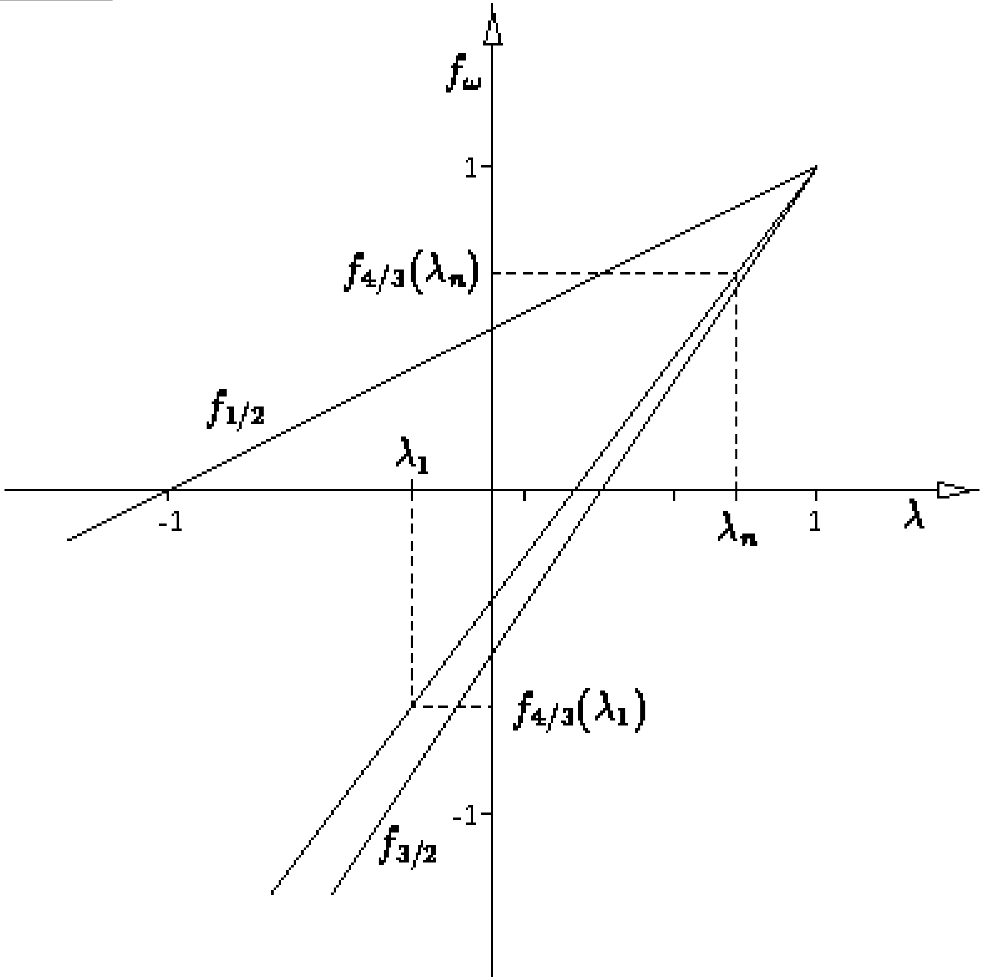
1. $M_J = D^{-1}(D - A)$ possess exclusively real eigenvalues
$$\lambda_1 \leq \cdots \leq \lambda_n.$$
2. Corresp. eigenvectors $u_1, \cdots, u_n$ are linear independent
3. It hold $\rho(M_J) < 1$.

Then:

1. Eigenvalues of the Jacobi relaxation method read:
$$\mu_i = 1 - \omega + \omega\lambda_i$$
2. Optimal relaxation parameter is given by:
$$\omega_{opt} = \frac{2}{2 - \lambda_1 - \lambda_n} > 0$$

Jacobi relaxation method

Idea to obatin $\omega_{opt.}$: $\mu_i = 1 - \omega + \omega\lambda_i, \quad \mu = f(\lambda) = 1 - \omega + \omega\lambda$



- Choose ω such that $\mu_1 = -\mu_n$.
- Improvement is only possible if $\lambda_1 \neq -\lambda_n$.

Jacobi relaxation method

Model problem:

$$A = \begin{pmatrix} 0.7 & -0.4 \\ -0.2 & 0.5 \end{pmatrix} \Longrightarrow M_J = \begin{pmatrix} 0 & 4/7 \\ 2/5 & 0 \end{pmatrix}$$

$$\Longrightarrow \qquad \lambda_{1,2} = \pm\sqrt{\frac{8}{35}}$$

$\Longrightarrow$ No improvement

Appraisement

"$\circ$" : Minor assumptions on A

"+" : Very simple calculation of matrix-vector products $\omega D^{-1} x$

"$\circ$" : often no improvement, *MGM*

SOR method (successive overrelaxation)

Basic principle:

Pointwise formulation of the Gauß-Seidel method

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

Proceeding

Weighting of the correction part

$$x_{m+1,i} = (1-\omega) x_{m,i} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_{m+1,j} - \sum_{j=i+1}^{n} a_{ij} x_{m,j} \right)$$

SOR method (successive overrelaxation)

Pointwise formulation:

$$x_{m+1,i} = (1-\omega)x_{m,i} + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_{m+1,j} - \sum_{j=i+1}^{n} a_{ij}x_{m,j}\right)$$

Matrix formulation:

$$\underbrace{(I+\omega D^{-1}L)}_{=D^{-1}(D+\omega L)} x_{m+1} = \underbrace{\Big((1-\omega)I - \omega D^{-1}R\Big)}_{=D^{-1}((1-\omega)D-\omega R)} x_m + \omega D^{-1}b$$

$$\Rightarrow \quad (D+\omega L)x_{m+1} = ((1-\omega)D - \omega R)\,x_m + \omega b$$

$$\Rightarrow \quad x_{m+1} = \underbrace{(D+\omega L)^{-1}((1-\omega)D - \omega R)}_{=M_{GS}(\omega)} x_m + \underbrace{\omega(D+\omega L)^{-1}}_{N_{GS}(\omega)} b$$

SOR method (successive overrelaxation)

Restrictions on the relaxation parameter:

The method is divergent for all $\omega \notin (0,2)$.

Reason:

$$
\begin{aligned}
\prod_{i=1}^{n} \lambda_i &= \det M_{GS}(\omega) \\
&= \det(D+\omega L)^{-1} \cdot \det((1-\omega)D - \omega R) \\
&= \det D^{-1} \cdot \det((1-\omega)D) \\
&= (1-\omega)^n
\end{aligned}
$$

$$
\Longrightarrow \qquad \underbrace{\max_{i=1,\ldots,n} |\lambda_i|}_{=\rho(M_{GS}(\omega))} \geq |1-\omega|
$$

SOR method (successive overrelaxation)

Optimal relaxation parameter

Assumptions:

1. A is consistently ordered.
2. Eigenvalues $\lambda_1, \ldots, \lambda_n$ of M_J are real.
3. $\rho := \rho(M_J) < 1$

Then:

1. The method is convergent for all $\omega \in (0, 2)$
2. $\omega_{opt} = \dfrac{2}{1 + \sqrt{1 - \rho^2}}\,, \; \rho(M_{GS}(\omega_{opt})) = \omega_{opt} - 1$
3. $\lambda_i = \dfrac{\mu_i + \omega - 1}{\omega \mu_i^{1/2}}$ with the eigenvalue μ_i of $M_{GS}(\omega)$

SOR method (successive overrelaxation)

Definition: Consistently ordered

The matrix $A = D + L + R$ with a non-singular diagonal part D is called consistently ordered, if the eigenvalues of

$$C(\alpha) = -(\alpha D^{-1} L + \alpha^{-1} D^{-1} R) \quad , \quad \alpha \in \mathbb{C} \backslash \{0\}$$

are independent of α.

SOR method (successive overrelaxation)

Examples:

- Each 2×2 matrix is consistently ordered.
- Each tridiagonal matrix

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ c_2 & a_2 & b_2 & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & b_{n-1} \\ & & & c_n & a_n \end{pmatrix}$$

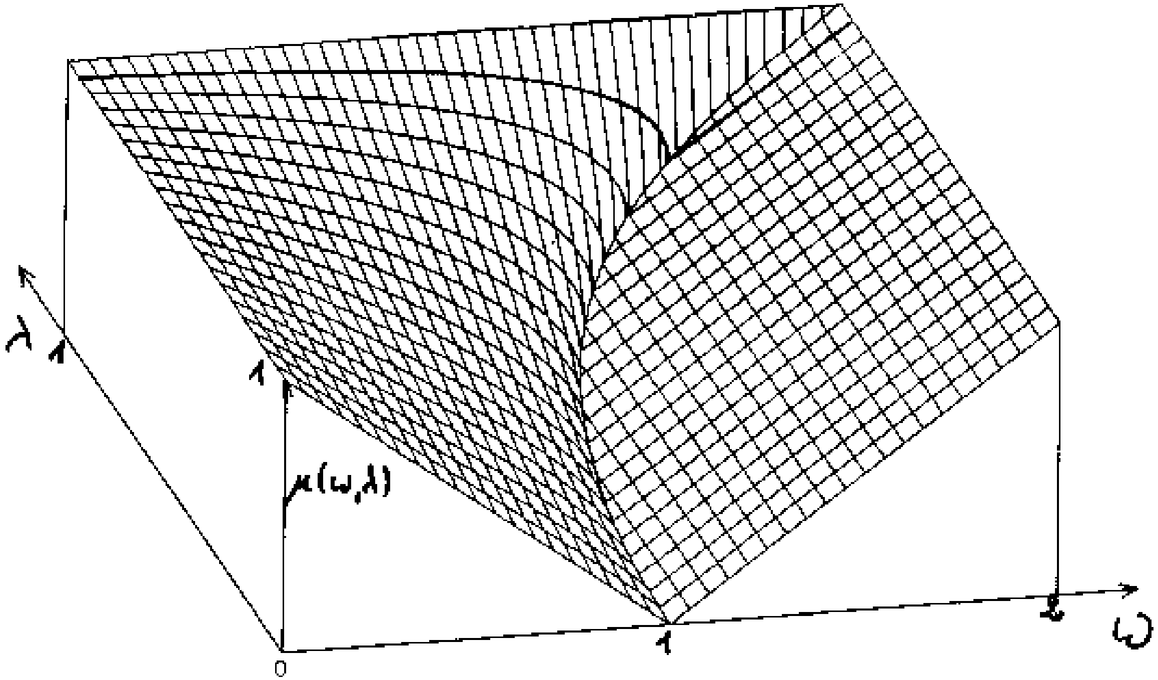
  is consistently ordered.
- Each matrix of the form

$$A = \begin{pmatrix} I & A_{12} \\ A_{21} & I \end{pmatrix}$$

  is consistently ordered.

"Distribution of the eigenvalues "of the relaxation method

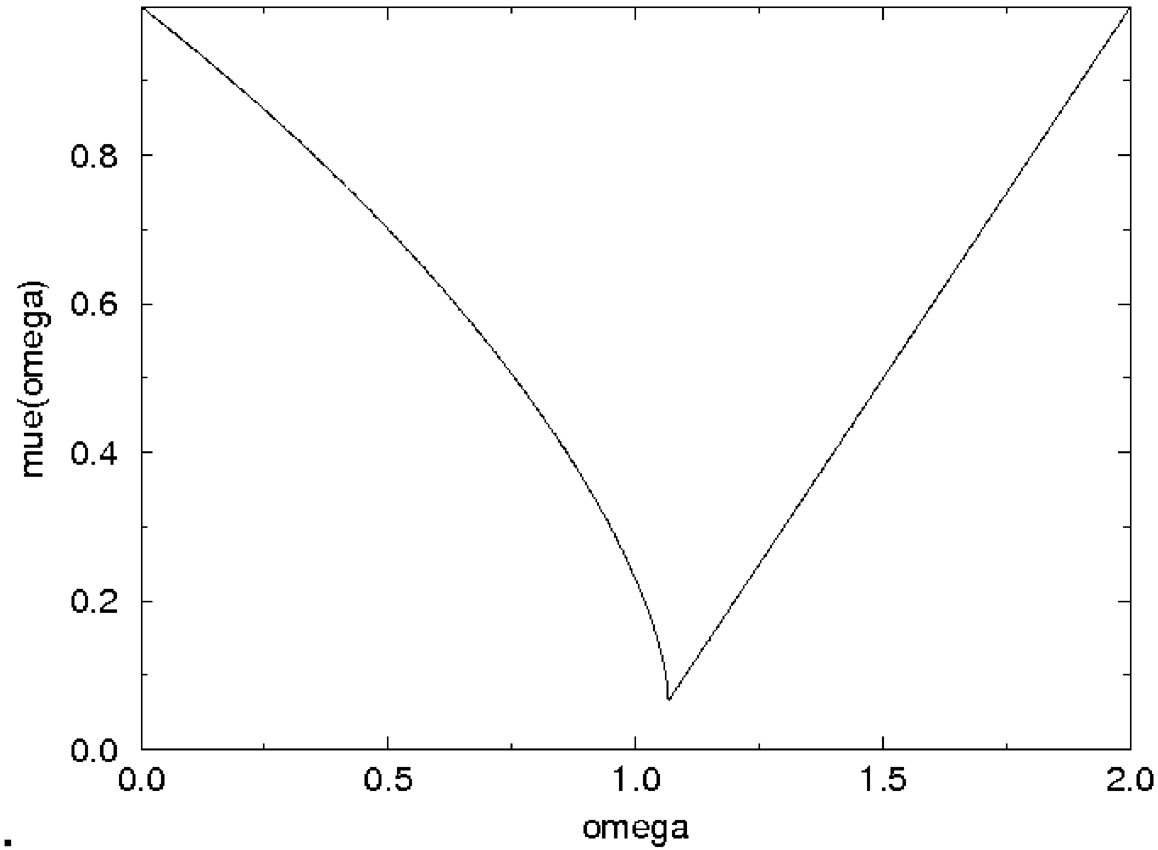
$$\mu(\omega,\lambda) = \begin{cases} \frac{1}{2}\lambda^2\omega^2 - (\omega - 1) + \lambda\omega\sqrt{\frac{1}{4}\lambda^2\omega^2 - (\omega - 1)} & , \ 0 < \omega < \omega^\star(\lambda) \\ \omega - 1 & , \ \omega^\star(\lambda) \le \omega < 2 \end{cases}$$



[Bunse, Bunse-Gerstner: Numerische Lineare Algebra]

Relaxation scheme

Spectral radius versus relaxation parameter



Rule of thumb:

- Better to choose ω a bit larger than ω_{opt} instead of a bit smaller than ω_{opt}.

Model problem

$$Ax = b \text{ with } A = \begin{pmatrix} 0.7 & -0.4 \\ -0.2 & 0.5 \end{pmatrix}, \quad b = \begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}$$

1. A is consistently ordered.
2. The eigenvalues of $M_J = -D^{-1}(L+R)$ are $\lambda_{1,2} = \pm\sqrt{\frac{8}{35}} \in \mathbb{R}$
3. $\rho(M_J) < 1$

$$\Longrightarrow \qquad \omega_{opt} = \frac{2}{1+\sqrt{1-\frac{8}{35}}} \approx 1.06479$$

$$\rho(M_{GS}(\omega_{opt})) = \omega_{opt} - 1 \approx 0.06479 \approx 0.25^2 \approx \rho(M_{GS})^2$$

Expectation:

Approximately half as much iterations as the Gauß-Seidel method to reach the same accuracy

SOR method (successive overrelaxation)

SOR method				
m	$x_{m,1}$	$x_{m,2}$	$\varepsilon_m := \|x_m - x^*\|_\infty$	$\varepsilon_m / \varepsilon_{m-1}$
0	2.100000e+01	-1.900000e+01	2.000000e+01	
1	-1.246473e+01	-3.439090e+00	1.346473e+01	6.732366e-01
2	-8.286241e-01	5.087570e-01	1.828624e+00	1.358084e-01
3	8.195743e-01	9.549801e-01	1.804257e-01	9.866748e-02
4	9.842969e-01	9.962285e-01	1.570309e-02	8.703354e-02
5	9.987226e-01	9.997003e-01	1.277401e-03	8.134709e-02
6	9.999004e-01	9.999770e-01	9.960642e-05	7.797587e-02
7	9.999925e-01	9.999983e-01	7.544695e-06	7.574507e-02
8	9.999994e-01	9.999999e-01	5.595127e-07	7.415974e-02
9	1.000000e-00	1.000000e-00	4.083051e-08	7.297514e-02
10	1.000000e-00	1.000000e-00	2.942099e-09	7.205638e-02
11	1.000000e-00	1.000000e-00	2.098393e-10	7.132298e-02
12	1.000000e-00	1.000000e-00	1.484068e-11	7.072406e-02
13	1.000000e-00	1.000000e-00	1.042055e-12	7.021612e-02
14	1.000000e-00	1.000000e-00	7.260859e-14	6.967824e-02
15	1.000000e-00	1.000000e-00	4.884981e-15	6.727829e-02

SOR method (successive overrelaxation)

Model problem:

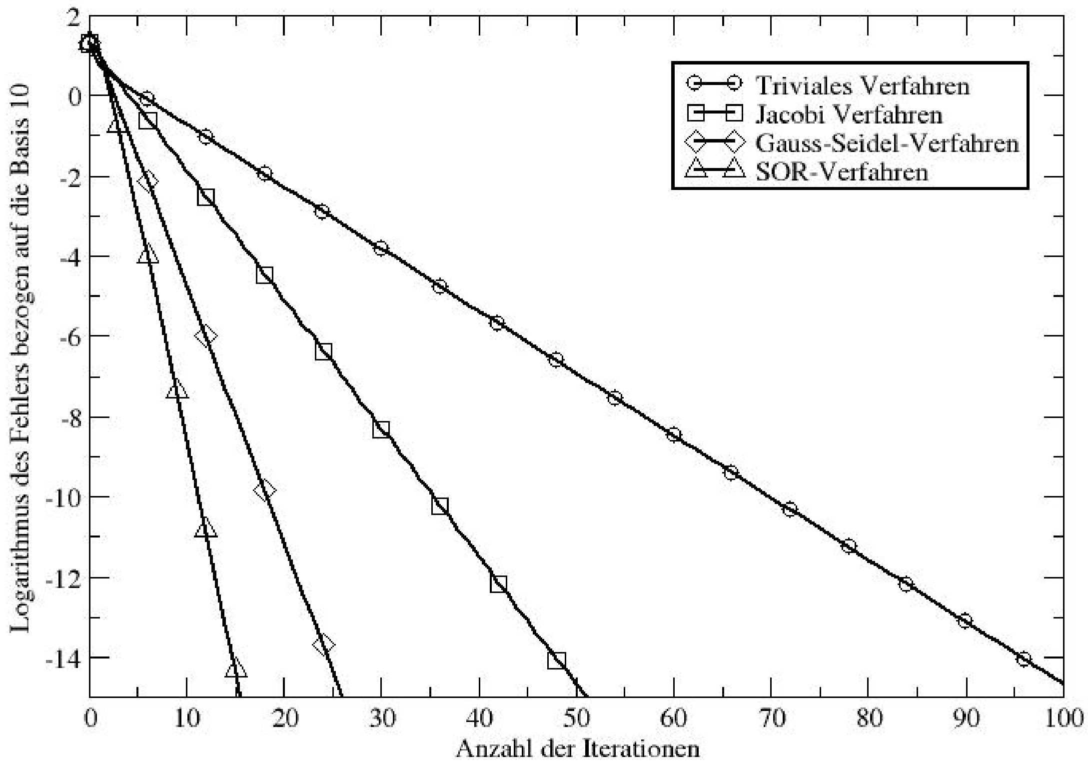


Abbildung: Convergence history $\log_{10} \varepsilon_m$ of the SOR method

SOR method (successive overrelaxation)

<u>Appraisement:</u>

"$\circ$"	:	Minor assumptions on the matrix A
		($a_{ii} \neq 0$ für $i = 1, \ldots, n$)
"+"	:	Simple calculation of matrix-vector products $(D + \omega L)^{-1} x$
"++"	:	Fast convergence behaviour

Summary:

Splitting methods:

- Easy to derive
- Simple to implement

Rate of convergence:

- Is determined by $\rho(M) = \rho(B^{-1}(B - A))$.
- Rule of thumb: Choose the approximation B as close as possible w.r.t. A to obtain a good convergence behaviour.

Summary:

Typs of Splitting methods

- Trivial scheme: $B = I$ (bad rate of convergence)
- Jacobi method : $B = D$
- Gauß-Seidel method : $B = D + L$
- Relaxation methods
  - Weighting of the correction vector
  - Optimizing $\rho(M(\omega))$