3. Relaxation Methods

May 4, 2010
Outline of This Session

- A very short recapitulation or *Where does LR make sense?*
- Fundamental properties of iterative schemes
- The Jacobi relaxation scheme
  - From a PDE's point of view
  - From an LA's point of view
  - Convergence analysis
- The Gauß-Seidel relaxation scheme
  - From a PDE's point of view
  - From an LA's point of view
  - Convergence analysis
- Relaxation Methods
3.1. Recapitulation

Properties of the $LR$ Decomposition

- What are the memory requirements?
- What means in-situ?
- What is the (general) algorithmic complexity?
- What is the complexity for PDEs discretised by local stencils?

So, where should/could we use $LR$ while we are still searching for an optimal solver, i.e. a solver with $O(n)$?
Where does $LR$ make sense?

- Sufficiently small matrices
- A multitude of right-hand sides
Where does $\mathbf{LR}$ make sense?

- Sufficiently small matrices
- A multitude of right-hand sides

\[ \partial_t u + \mathcal{L}(u) = f \]
\[ \partial_t u = f - \mathcal{L}(u) \quad \forall t \]
\[ \Omega(du)_h = M \cdot f - L \cdot u_h \quad \forall t_i \]
\[ (du)_h = \Omega^{-1} (M \cdot f - L \cdot u_h) \quad \forall t_i \]
Where does $LR$ make sense?

- Sufficiently small matrices
- A multitude of right-hand sides

\[
\begin{align*}
\partial_t u + L(u) &= f \\
\partial_t u &= f - L(u) \quad \forall t \\
\Omega(du)_h &= M \cdot f - L \cdot u_h \quad \forall t_i \\
(du)_h &= \Omega^{-1} (M \cdot f - L \cdot u_h) \quad \forall t_i 
\end{align*}
\]

- But $\Omega$ may never change, i.e. it may not depend on $t_i$.
- Video 1: Dynamic adaptivity for the heat equation
- Video 2: Fluid-structure interaction
3.2. Iterative Schemes

- Define an initial guess for the solution $u^{(0)}$.

- Define an update scheme:
  - Which entries of $u$ shall be modified?
    - All at once.
      (Jacobi or CG, e.g.)
    - One per step.
      (Gauß-Seidel, e.g.)
    - Some of them.
      (Multigrid methods or red-black Gauß-Seidel, e.g.)
  - How shall these entries be modified?
    - Stationary iteration scheme, i.e. it is always the same scheme.
    - Krylov methods.
    - Multiscale methods.

- Apply several steps of this update scheme, i.e.

\[ u^{(n)} \rightarrow u^{(n+1)} \]
Properties of Iterative Schemes

- Study the error behaviour:

\[ A \cdot u = b \iff A \cdot e = 0 \text{ with } e^{(n)} + u^{(n)} = A^{-1}b. \]

- Quantitative:

\[ \| e^{(n+1)} \| \leq p \| e^{(n)} \| q \]

- Qualitative: What does the error look like after several steps?
Stationary Schemes

- Split up matrix

\[ A \cdot u = (M + N) \cdot u = b \]
Stationary Schemes

- Split up matrix

\[ A \cdot u = (M + N) \cdot u = b \]

- Bring part of left-hand side to the right

\[ N \cdot u = b - M \cdot u \]
Stationary Schemes

- Split up matrix
  \[ A \cdot u = (M + N) \cdot u = b \]

- Bring part of left-hand side to the right
  \[ N \cdot u = b - M \cdot u \]

- Make it an iterative scheme
  \[ N \cdot u^{(n+1)} = b - M \cdot u^{(n)} \]
  \[ u^{(n+1)} = N^{-1} \left( b - M \cdot u^{(n)} \right) \]

- What should \( N \) look like?
Stationary Schemes Rewritten

• Take iterative scheme

\[
N \cdot u^{(n+1)} = b - M \cdot u^{(n)} \\

u^{(n+1)} = N^{-1} \left( b - M \cdot u^{(n)} \right)
\]

• Express \( M \) in terms of \( A \) and \( N \)

\[
A = M + N \Rightarrow M = A - N \\

u^{(n+1)} = N^{-1} \left( b - M \cdot u^{(n)} \right) \\
\quad = N^{-1} \left( b - (A - N) \cdot u^{(n)} \right) \\
\quad = N^{-1} \left( b - A \cdot u^{(n)} + N \cdot u^{(n)} \right) \\
\quad = N^{-1} \left( b - A \cdot u^{(n)} \right) + u^{(n)} \\
\quad = u^{(n)} + N^{-1} \text{res}^{(n)}
\]
The Residual

\[ u^{(n+1)} = u^{(n)} + N^{-1} \left( b - A \cdot u^{(n)} \right) = u^{(n)} + N^{-1} \text{res}^{(n)} \]

- The residual is something like the *error in the image*.
- The residual measures the error with respect of $A$-norm.
- The residual acts as input for the update steps.
- The better $N$ approximates $A$, the better the convergence!?
- For $N = A$, we end up with the direct solver.
3.3. Jacobi

- Method is named after German mathematician Carl Gustav Jakob Jacobi.
- Update all the entries of solution vector at once.
- Always use the same update rule: Approximate the inverse by the inverse of the diagonal.
...from a PDE’s point of view

- Have a look at the movie.
- Evaluate the residual at each point, as the residual is a (good) estimate for the error.
- Try to make the residual 0 everywhere, while we now that this induces errors at the neighbouring points.
...from an LA’s point of view

\[ u^{(n+1)} = u^{(n)} + N^{-1} \left( b - A \cdot u^{(n)} \right) = u^{(n)} + N^{-1} \text{res}^{(n)} \]

\[ u_i = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} u_j \right) \]

- Select \( N = \text{diag}(A) \), i.e. \( A = M + \text{diag}(A) = (A - \text{diag}(A)) + \text{diag}(A) \).
- Update of entry/computation of entry \( u_i^{(n+1)} \) does not depend on any other entry of \( u^{(n)} \).
- 'Better approximation' in \( u_i^{(n+1)} \) does not propagate to any \( u_j^{(n+1)} \).
- All update steps in vector \( u^{(n+1)} \) can run in parallel.
- Programming remark: We need two different arrays for \( u^{(n)} \) and \( u^{(n+1)} \).
Convergence Analysis—When does it Converge? (Part I)

- Method converges ⇔ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

$$\|e^{(n+1)}\|_2 \leq \ldots \|e^{(n)}\|_2$$
Convergence Analysis—When does it Converge? (Part I)

- Method converges $\iff$ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

$$
\| e^{(n+1)} \|_2 \leq \ldots \| e^{(n)} \|_2 \\
\| e^{(n+1)} \|_2 = \| u^{(*)} - u^{(n+1)} \|_2
$$
Convergence Analysis—When does it Converge? (Part I)

- Method converges ⇐ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(\ast)}$, and update rules:

\[
\begin{align*}
\|e^{(n+1)}\|_2 & \leq \ldots \|e^{(n)}\|_2 \\
\|e^{(n+1)}\|_2 & = \|u^{(\ast)} - u^{(n+1)}\|_2 \\
& = \|u^{(\ast)} - \left(u^{(n)} + \text{diag}^{-1}(A) \left(b - A \cdot u^{(n)}\right)\right)\|_2
\end{align*}
\]
Convergence Analysis—When does it Converge? (Part I)

- Method converges $\iff$ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

$$
\| e^{(n+1)} \|_2 \leq \ldots \| e^{(n)} \|_2 \\
\| e^{(n+1)} \|_2 = \| u^{(*)} - u^{(n+1)} \|_2 \\
= \| u^{(*)} - \left( u^{(n)} + \text{diag}^{-1}(A) \left( b - A \cdot u^{(n)} \right) \right) \|_2 \\
= \| u^{(*)} - \left( u^{(n)} + \text{diag}^{-1}(A) \left( A \cdot u^{(*)} - A \cdot u^{(n)} \right) \right) \|_2
$$
Convergence Analysis—When does it Converge? (Part I)

- Method converges ⇔ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

\[
\begin{align*}
\|e^{(n+1)}\|_2 & \leq \ldots \|e^{(n)}\|_2 \\
\|e^{(n+1)}\|_2 & = \|u^{(*)} - u^{(n+1)}\|_2 \\
& = \|u^{(*)} - \left(u^{(n)} + \text{diag}^{-1}(A) \left(b - A \cdot u^{(n)}\right)\right)\|_2 \\
& = \|u^{(*)} - \left(u^{(n)} + \text{diag}^{-1}(A) \left(A \cdot u^{(*)} - A \cdot u^{(n)}\right)\right)\|_2 \\
& = \left\| \left(id - \text{diag}^{-1}(A)A\right) \left(u^{(*)} - u^{(n)}\right) \right\|_2
\end{align*}
\]
Convergence Analysis—When does it Converge? (Part I)

- Method converges $\iff$ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

\[
\|e^{(n+1)}\|_2 \leq \ldots \|e^{(n)}\|_2 \\
\|e^{(n+1)}\|_2 = \|u^{(*)} - u^{(n+1)}\|_2 \\
= \|u^{(*)} - \left(u^{(n)} + \text{diag}^{-1}(A) \left(b - A \cdot u^{(n)}\right)\right)\|_2 \\
= \|u^{(*)} - \left(u^{(n)} + \text{diag}^{-1}(A) \left(A \cdot u^{(*)} - A \cdot u^{(n)}\right)\right)\|_2 \\
= \left\| \left(id - \text{diag}^{-1}(A)A\right) \left(u^{(*)} - u^{(n)}\right) \right\|_2 \\
= \left\| \left(id - \text{diag}^{-1}(A) \cdot A\right) e^{(n)} \right\|_2
\]
Convergence Analysis—When does it Converge? (Part I)

- Method converges $\Leftrightarrow$ error becomes smaller in each step.
- Error is unknown, but we can express it in terms of $A$, $u^{(n)}$, exact solution $u^{(*)}$, and update rules:

\[
\| e^{(n+1)} \|_2 \leq \ldots \| e^{(n)} \|_2 \\
\| e^{(n+1)} \|_2 = \| u^{(*)} - u^{(n+1)} \|_2 \\
= \| u^{(*)} - \left( u^{(n)} + \text{diag}^{-1}(A) \left( b - A \cdot u^{(n)} \right) \right) \|_2 \\
= \| u^{(*)} - \left( u^{(n)} + \text{diag}^{-1}(A) \left( A \cdot u^{(*)} - A \cdot u^{(n)} \right) \right) \|_2 \\
= \| \left( \text{id} - \text{diag}^{-1}(A)A \right) \left( u^{(*)} - u^{(n)} \right) \|_2 \\
= \| \left( \text{id} - \text{diag}^{-1}(A) \cdot A \right) e^{(n)} \|_2 \\
\leq \| \text{id} - \text{diag}^{-1}(A) \cdot A \|_2 \| e^{(n)} \|_2
\]
Convergence Analysis—When does it Converge? (Part II)

\[ \| e^{(n+1)} \|_2 = \| u^{(\ast)} - u^{(n+1)} \|_2 \]
\[ = \ldots \]
\[ \leq \| id - \text{diag}^{-1}(A) \cdot A \|_2 \| e^{(n)} \|_2 \]
Convergence Analysis—When does it Converge? (Part II)

\[ \| e^{(n+1)} \|_2 = \| u^{(*)} - u^{(n+1)} \|_2 \]
\[ = \ldots \]
\[ \leq \| id - diag^{-1}(A) \cdot A \|_2 \| e^{(n)} \|_2 \]
\[ = \left( \| id - diag^{-1}(A) \cdot A \|_2 \right)^n \| e^{(0)} \|_2 \]

- Matrix \((id - diag^{-1}(A) \cdot A) =: M\) is iteration matrix.
- \(\lim_{n \to \infty} \| M \|_2^n = 0 \iff\) Jacobi method converges.
- To make the Jacobi method converge, \(M\) should be nilpotent.
- Attention: Jacobi might work for other matrices as well.
Convergence Analysis—When does it Converge? (Part II)

\[
\| e^{(n+1)} \|_2 = \| u^{(*)} - u^{(n+1)} \|_2 \\
= \ldots \\
\leq \| id - diag^{-1}(A) \cdot A \|_2 \| e^{(n)} \|_2 \\
= \left( \| id - diag^{-1}(A) \cdot A \|_2 \right)^n \| e^{(0)} \|_2
\]

- Matrix \( (id - diag^{-1}(A) \cdot A) =: M \) is iteration matrix.
- \( \lim_{n \to \infty} \| M \|_2^n = 0 \iff \text{Jacobi method converges} \).
- To make the Jacobi method converge, \( M \) should be nilpotent.
- Attention: Jacobi might work for other matrices as well.
- \( M \) nilpotent \( \iff \) spectral radius \( \rho(M) \leq 1 \).
Convergence Analysis—When does it Converge? (Part II)

\[
\| \text{e}^{(n+1)} \|_2 = \| \text{u}^{(\ast)} - \text{u}^{(n+1)} \|_2
\]
\[
= \ldots
\]
\[
\leq \| \text{id} - \text{diag}^{-1}(A) \cdot A \|_2 \| \text{e}^{(n)} \|_2
\]
\[
= \left( \| \text{id} - \text{diag}^{-1}(A) \cdot A \|_2 \right)^n \| \text{e}^{(0)} \|_2
\]

- Matrix \((\text{id} - \text{diag}^{-1}(A) \cdot A) =: M\) is iteration matrix.
- \(\lim_{n \to \infty} \|M\|_2^n = 0 \Leftrightarrow\) Jacobi method converges.
- To make the Jacobi method converge, \(M\) should be nilpotent.
- Attention: Jacobi might work for other matrices as well.
- \(M\) nilpotent \(\Leftrightarrow\) spectral radius \(\rho(M) \leq 1\).
- Proof: \(\lim_{k \to \infty} \|B^k\|^{1/k} =: \rho(B) = \max |\lambda(B)|.\)
Convergence Analysis—When does it Converge? (Part III)

\[-\Delta u = 0\]
\[u_{|_{\partial \Omega}} = 0\]
\[u^{(0)} = 1\]
Convergence Analysis—When does it Converge? (Part IV)

- Update rule for one variable is
  \[
  u_i = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} u_j \right)
  \]
  with \( b_i = 0 \),
- i.e. we make \( u_i \) the average value of its neighbours (Laplace operator).
- No entry may explode, i.e.
  \[
  |a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad \forall i
  \]
Convergence Analysis—When does it Converge? (Part V)

- Terminology: Matrix $A$ is (weak) diagonal dominant, if
  \[|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad \forall i\]

- Terminology: Matrix $A$ is strict diagonal dominant, if
  \[|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \forall i\]

- To make Jacobi converge, the stiffness matrix $A$ should be weak diagonal dominant with
  \[\exists i : |a_{ii}| > \sum_{j \neq i} |a_{ij}|\]

- Theorem: If the iteration matrix $M$ is nilpotent, and the stiffness matrix $A$ is sufficiently diagonal dominant, the Jacobi method converges.

- However: It might work otherwise as well.
Convergence Analysis—How fast does it Converge?
Convergence Analysis—How fast does it Converge?

We need something to do next week, too ...
3.4. Gauß-Seidel

- Method is named after the German mathematicians Carl Friedrich Gauss and Philipp Ludwig von Seidel.
- Other name: Liebmann method.
- Update one entry after another.
- Always use the same update rule: Approximate the inverse by the inverse of the diagonal.
...from a PDE’s point of view

- Have a look at the movie.
- Evaluate the residual in one point.
- Make the residual 0 in this point, while we now that this induces errors at the neighbouring points.
- Continue with the other points.
...from an LA’s point of view

\[ u^{(n+1)} = u^{(n)} + N^{-1} \left( b - A \cdot u^{(n)} \right) = u^{(n)} + N^{-1} \text{res}^{(n)} \]

- Select \( N = L(A) \), i.e. \( A = M + L(A) = (A - L(A)) + L(A) = U(A) + L(A) \).
- Update of entry/computation of entry \( u_{i}^{(n+1)} \) does depend on preceding entries of \( u^{(n)} \).
- 'Better approximation' in \( u_{i}^{(n+1)} \) propagates immediately to \( u_{j}^{(n+1)} \) \( j > i \).
- Programming remark: we can implement it with one array for \( u^{(n+1)} \).
- Write it down as sum:
...from an LA’s point of view

\[ u^{(n+1)} = u^{(n)} + N^{-1} \left( b - A \cdot u^{(n)} \right) = u^{(n)} + N^{-1} \text{res}^{(n)} \]

• Select \( N = L(A) \), i.e. \( A = M + L(A) = (A - L(A)) + L(A) = U(A) + L(A) \).

• Update of entry/computation of entry \( u_i^{(n+1)} \) does depend on preceding entries of \( u^{(n)} \).

• ’Better approximation’ in \( u_i^{(n+1)} \) propagates immediately to \( u_j^{(n+1)} \) \( j > i \).

• Programming remark: we can implement it with one array for \( u^{(n+1)} \).

• Write it down as sum:

\[
    u_i = \frac{1}{a_{ii}} \left( b_i - \sum_{j<i} a_{ij} u_j - \sum_{j>i} a_{ij} u_j \right)
\]
Convergence Analysis—When does it Converge?

- To make a long story short: We again have to have an $A$ which is diagonal dominant.
- To make a long story short: We again have to study the iteration matrix.
Convergence Analysis—When does it Converge?

- To make a long story short: We again have to have an $A$ which is diagonal dominant.
- To make a long story short: We again have to study the iteration matrix.
- Iteration matrix $M$:

$$
\| e^{(n+1)} \|_2 = \ldots = \| (id - L^{-1} \cdot A) e^{(n)} \|_2 \\
\leq \| id - L^{-1} \cdot A \|_2 \| e^{(n)} \|_2
$$
Convergence Analysis—When does it Converge?

- To make a long story short: We again have to have an $A$ which is diagonal dominant.
- To make a long story short: We again have to study the iteration matrix.
- Iteration matrix $M$:
  \[
  \| e^{(n+1)} \|_2 = \ldots \\
  = \| (id - L^{-1} \cdot A) e^{(n)} \|_2 \\
  \leq \| id - L^{-1} \cdot A \|_2 \| e^{(n)} \|_2
  \]
- Jacobi and Gauß-Seidel seem to be in the same class of convergence, i.e. differences result from the different iteration matrices $M$.
- The convergence criteria are the same.
- Exact study will be topic of the next lecture.
3.5. Relaxation Methods

- Imagine we’ve chosen an iteration scheme that makes the solution look (almost) like an orbit.
- We’ll see that Jacobi/GS do exactly this for certain guesses (eigenvalues).
- Isn’t it a pity?
The Trivial Iteration

- What we could do: We could scale the solution after each iteration.
- The trivial iteration is $u^{(n+1)} = u^{(n)}$.
- Relaxation method: Take an iterative scheme and combine it with the trivial iteration to scale the solution:
The Trivial Iteration

- What we could do: We could scale the solution after each iteration.

- The trivial iteration is $u^{(n+1)} = u^{(n)}$.

- Relaxation method: Take an iterative scheme and combine it with the trivial iteration to scale the solution:

$$
\begin{align*}
  u^{(n+1)} &= u^{(n)} + N^{-1} \cdot \omega \in \mathbb{R} \\
  u^{(n+1)} &= u^{(n)} \cdot 1 - \omega \in \mathbb{R}
\end{align*}
$$
The Trivial Iteration

- What we could do: We could scale the solution after each iteration.
- The trivial iteration is \( u^{(n+1)} = u^{(n)} \).
- Relaxation method: Take an iterative scheme and combine it with the trivial iteration to scale the solution:

\[
\begin{align*}
u^{(n+1)} &= u^{(n)} + N^{-1} \text{res}^{(n)} \cdot \omega \in \mathbb{R} \\
u^{(n+1)} &= u^{(n)} \cdot 1 - \omega \in \mathbb{R}
\end{align*}
\]

\[
\begin{align*}
u^{(n+1)} &= u^{(n)} + \omega N^{-1} \text{res}^{(n)}
\end{align*}
\]
The Trivial Iteration

- What we could do: We could scale the solution after each iteration.
- The trivial iteration is $u^{(n+1)} = u^{(n)}$.
- Relaxation method: Take an iterative scheme and combine it with the trivial iteration to scale the solution:

$$u^{(n+1)} = u^{(n)} + N^{-1} \text{res}^{(n)} \cdot \omega \in \mathbb{R}$$
$$u^{(n+1)} = u^{(n)} \cdot 1 - \omega \in \mathbb{R}$$

$$u^{(n+1)} = u^{(n)} + \omega N^{-1} \text{res}^{(n)}$$

- Convergence:
The Trivial Iteration

- What we could do: We could scale the solution after each iteration.
- The trivial iteration is $u^{(n+1)} = u^{(n)}$.
- Relaxation method: Take an iterative scheme and combine it with the trivial iteration to scale the solution:

$$
\begin{align*}
    u^{(n+1)} &= u^{(n)} + N^{-1} \text{res}^{(n)} \cdot \omega \in \mathbb{R} \\
    u^{(n+1)} &= u^{(n)} \cdot 1 - \omega \in \mathbb{R}
\end{align*}
$$

$$
\begin{align*}
    u^{(n+1)} &= u^{(n)} + \omega N^{-1} \text{res}^{(n)}
\end{align*}
$$

- Convergence: Still depends on the spectral radius $\rho(M)$. 

3. Relaxation Methods
Scientific Computing II, Tobias Weinzierl
Relaxation Variants

\[ u^{(n+1)} = u^{(n)} + \omega N^{-1} \text{res}^{(n)} \]

- Richardson iteration, i.e. \( N = \theta \text{id} \).
Relaxation Variants

\[ u^{(n+1)} = u^{(n)} + \omega N^{-1} \text{res}^{(n)} \]

- Richardson iteration, i.e. \( N = \theta id \).
  - Usually not used.
  - Consequently, relaxation not used.
- Jacobi method, i.e. \( N = \text{diag}(A) \).
Relaxation Variants

\[ u^{(n+1)} = u^{(n)} + \omega N^{-1} \text{res}^{(n)} \]

- Richardson iteration, i.e. \( N = \theta \text{id} \).
  - Usually not used.
  - Consequently, relaxation not used.
- Jacobi method, i.e. \( N = \text{diag}(A) \).
  - \( \omega \in (0, 1) \).
  - Successive Underrelaxation.
  - Why does \( \omega = 0 \) not work?
  - Why does \( \omega > 1 \) not work? (start with oscillation).
- Gauß-Seidel method, i.e. \( N = L(A) \).
Relaxation Variants

\[ u^{(n+1)} = u^{(n)} + \omega N^{-1} \text{res}^{(n)} \]

- Richardson iteration, i.e. \( N = \theta \text{id} \).
  - Usually not used.
  - Consequently, relaxation not used.

- Jacobi method, i.e. \( N = \text{diag}(A) \).
  - \( \omega \in (0, 1) \).
  - Successive Underrelaxation.
  - Why does \( \omega = 0 \) not work?
  - Why does \( \omega > 1 \) not work? (start with oscillation).

- Gauß-Seidel method, i.e. \( N = L(A) \).
  - \( \omega \in (0, 1) \Rightarrow \) Successive Underrelaxation (SUR).
  - \( \omega \in (1, 2) \Rightarrow \) Successive Overrelaxation (SOR).
  - Why does \( \omega = 0 \) not work?
  - What is \( \omega = 1 \)?
  - Why does \( \omega \geq 2 \) not work? (start with oscillation).