Iterative methods: Basic relaxation techniques

- Relaxation methods: Jacobi, Gauss-Seidel, SOR
- Basic convergence results
- Optimal relaxation parameter for SOR
- See Chapter 4 of text for details.

Basic relaxation schemes
- Relaxation schemes: methods that modify one component of current approximation at a time
- Based on the decomposition $A = D - E - F$ with:
  - $D = \text{diag}(A)$, $-E = \text{strict lower part of } A$ and $-F = \text{its strict upper part}.$

Gauss-Seidel iteration for solving $Ax = b$:
- corrects $j$-th component of current approximate solution, to zero the $j$-th component of residual for $j = 1, 2, \cdots, n.$

Gauss-Seidel iteration can be expressed as:

$$ (D - E)x^{(k+1)} = Fx^{(k)} + b $$

Can also define a backward Gauss-Seidel Iteration:

$$ (D - F)x^{(k+1)} = Ex^{(k)} + b $$

and a Symmetric Gauss-Seidel Iteration: forward sweep followed by backward sweep.

Over-relaxation is based on the splitting:

$$ \omega A = (D - \omega E) - (\omega F + (1 - \omega)D) $$

→ successive overrelaxation, (SOR):

$$ (D - \omega E)x^{(k+1)} = [\omega F + (1 - \omega)D]x^{(k)} + \omega b $$

Iteration matrices
- Previous methods based on a splitting of $A$:

$$ A = M - N \rightarrow $$

$$ Mx = Nx + b \rightarrow Mx^{(k+1)} = Nx^{(k)} + b $$

$$ x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b \equiv Gx^{(k)} + f $$

Jacobi, Gauss-Seidel, SOR, & SSOR iterations are of the form

$$ G_{Jac} = D^{-1}(E + F) = I - D^{-1}A $$
$$ G_{GS} = (D - E)^{-1}F = I - (D - E)^{-1}A $$
$$ G_{SOR} = (D - \omega E)^{-1}(\omega F + (1 - \omega)D) $$

$$ = I - (\omega^{-1}D - E)^{-1}A $$
$$ G_{SSOR} = I - \omega(2 - \omega)(D - \omega F)^{-1}D(D - \omega E)^{-1}A $$
General convergence result

Consider the iteration: \( x^{(k+1)} = Gx^{(k)} + f \)

(1) Assume that \( \rho(G) < 1 \). Then \( I - G \) is non-singular and \( G \) has a fixed point. Iteration converges to a fixed point for any \( f \) and \( x^{(0)} \).

(2) If iteration converges for any \( f \) and \( x^{(0)} \) then \( \rho(G) < 1 \).

Example: Richardson’s iteration

\[ x^{(k+1)} = x^{(k)} + \alpha(b - Ax^{(k)}) \]

Assume \( \Lambda(A) \subset \mathbb{R} \). When does the iteration converge?

A few well-known results

- Jacobi and Gauss-Seidel converge for diagonal dominant matrices, i.e., matrices such that

\[ |a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, \ldots, n \]

- SOR converges for \( 0 < \omega < 2 \) for SPD matrices

- The optimal \( \omega \) is known in theory for an important class of matrices called 2-cyclic matrices or matrices with property A.

A matrix has property A if it can be (symmetrically) permuted into a \( 2 \times 2 \) block matrix whose diagonal blocks are diagonal.

\[ PAP^T = \begin{bmatrix} D_1 & E \\ E^T & D_2 \end{bmatrix} \]

Let \( A \) be a matrix which has property A. Then the eigenvalues \( \lambda \) of the SOR iteration matrix and the eigenvalues \( \mu \) of the Jacobi iteration matrix are related by

\[ (\lambda + \omega - 1)^2 = \lambda \omega^2 \mu^2 \]

The optimal \( \omega \) for matrices with property A is given by

\[ \omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(B)^2}} \]

where \( B \) is the Jacobi iteration matrix.

An observation

Introduction to Preconditioning

- The iteration \( x^{(k+1)} = Gx^{(k)} + f \) is attempting to solve \((I - G)x = f\). Since \( G \) is of the form \( G = M^{-1}(M - A) \) and \( f = M^{-1}b \), this system becomes

\[ M^{-1}Ax = M^{-1}b \]

where for SSOR, for example, we have

\[ M_{SSOR} = (D - \omega E)D^{-1}(D - \omega F) \]

referred to as the SSOR ‘preconditioning’ matrix.

In other words:

\[ \text{Relaxation iter.} \iff \text{Preconditioned Fixed Point Iter.} \]
Projection methods

- Introduction to projection-type techniques
- Sample one-dimensional Projection methods
- Some theory and interpretation –
  - See Chapter 5 of text for details.

Background on projectors

- A projector is a linear operator that is idempotent:
  \[ P^2 = P \]

**A few properties:**

- \( P \) is a projector iff \( I - P \) is a projector
- \( x \in \text{Ran}(P) \) iff \( x = Px \) iff \( x \in \text{Null}(I - P) \)
- This means that: \( \text{Ran}(P) = \text{Null}(I - P) \)
- Any \( x \in \mathbb{R}^n \) can be written (uniquely) as \( x = x_1 + x_2 \),
  \( x_1 = Px \in \text{Ran}(P) \) \( x_2 = (I - P)x \in \text{Null}(P) \) - So:
  \[ \mathbb{R}^n = \text{Ran}(P) \oplus \text{Null}(P) \]

\[ \square \]

Prove the above properties

Background on projectors (Continued)

- The decomposition \( \mathbb{R}^n = K \oplus S \) defines a (unique) projector \( P \):
  - From \( x = x_1 + x_2 \), set \( Px = x_1 \).
  - For this \( P \): \( \text{Ran}(P) = K \) and \( \text{Null}(P) = S \).
  - Note: \( \text{dim}(K) = m \), \( \text{dim}(S) = n - m \).
  - \( \text{Pb:} \) express mapping \( x \rightarrow u = Px \) in terms of \( K, S \)
  - Note: \( u \in K \), \( x - u \in S \)
  - Express 2nd part with \( m \) constraints: let \( L = S^\perp \), then
    \[ u = Px \text{ iff } \begin{cases} u \in K \\ x - u \perp L \end{cases} \]

\[ \Rightarrow \]

Projection onto \( K \) and orthogonally to \( L \)
Projection methods

- Initial Problem: \( b - Ax = 0 \)
  Given two subspaces \( K \) and \( L \) of \( \mathbb{R}^N \) define the approximate problem:
  
  \[
  \text{Find } \tilde{x} \in K \text{ such that } b - A\tilde{x} \perp L
  \]

- Petrov-Galerkin condition
- \( m \) degrees of freedom (\( K \)) + \( m \) constraints (\( L \)) →
- a small linear system (‘projected problem’)
- This is a basic projection step. Typically a sequence of such steps are applied

Matrix representation:

Let
- \( V = [v_1, \ldots, v_m] \) a basis of \( K \) &
- \( W = [w_1, \ldots, w_m] \) a basis of \( L \)

- Write approximate solution as \( \tilde{x} = x_0 + \delta \equiv x_0 + Vy \) where \( y \in \mathbb{R}^m \). Then Petrov-Galerkin condition yields:
  
  \[
  W^T(r_0 - AVy) = 0
  \]

- Therefore,
  
  \[
  \tilde{x} = x_0 + V[W^TAV]^{-1}W^Tr_0
  \]

Remark: In practice \( W^TAV \) is known from algorithm and has a simple structure [tridiagonal, Hessenberg,..]
Prototype Projection Method

Until Convergence Do:
1. Select a pair of subspaces $K$, and $L$;
2. Choose bases: $V = [v_1, \ldots, v_m]$ for $K$ and $W = [w_1, \ldots, w_m]$ for $L$.
3. Compute:
   \[ r \leftarrow b - Ax, \]
   \[ y \leftarrow (W^T AV)^{-1} W^T r, \]
   \[ x \leftarrow x + Vy. \]

In the case $x_0 = 0$, approximate problem amounts to solving
\[ Q(b - Ax) = 0, \quad x \in K \]
or in operator form (solution is $\Pi x$)
\[ Q(b - A\Pi x) = 0 \]

Question: what accuracy can one expect?

Projection methods: Operator form representation

Let $\Pi$ = the orthogonal projector onto $K$ and $Q$ the (oblique) projector onto $K$ and orthogonally to $L$.

\[ \Pi x \in K, \quad x - \Pi x \perp K \]
\[ Qx \in K, \quad x - Qx \perp L \]

$\Pi$ and $Q$ projectors

Assumption: no vector of $K$ is $\perp$ to $L$

Let $x^*$ be the exact solution. Then
1) We cannot get better accuracy than $\| (I - \Pi) x^* \|_2$, i.e.,
\[ \| \tilde{x} - x^* \|_2 \geq \| (I - \Pi) x^* \|_2 \]
2) The residual of the exact solution for the approximate problem satisfies:
\[ \| b - QA\Pi x^* \|_2 \leq \| QA(I - \Pi) \|_2 \| (I - \Pi) x^* \|_2 \]
Two Important Particular Cases.

1. \( L = K \)
   - When \( A \) is SPD then \( \| x^* - \tilde{x} \|_A = \min_{z \in K} \| x^* - z \|_A \).
   - Class of Galerkin or Orthogonal projection methods
   - Important member of this class: Conjugate Gradient (CG) method

2. \( L = AK \)
   - In this case \( \| b - A\tilde{x} \|_2 = \min_{z \in K} \| b - Az \|_2 \)
   - Class of Minimal Residual Methods: CR, GCR, ORTHOMIN, GMRES, CGNR, ...

One-dimensional projection processes

\[ K = \text{span}\{d\} \]
\[ L = \text{span}\{e\} \]

Then \( \tilde{x} = x + \alpha d \). Condition \( r - A\delta \perp e \) yields
\[ \alpha = \frac{(r,e)}{(Ad,e)} \]

Three popular choices:
1. Steepest descent
2. Minimal residual iteration
3. Residual norm steepest descent

Convergence based on the Kantorovitch inequality: Let \( B \) be an SPD matrix, \( \lambda_{\text{max}}, \lambda_{\text{min}} \) its largest and smallest eigenvalues. Then,
\[
\frac{(Bx,x)(B^{-1}x,x)}{(x,x)^2} \leq \frac{(\lambda_{\text{max}} + \lambda_{\text{min}})^2}{4 \lambda_{\text{max}} \lambda_{\text{min}}}, \quad \forall x \neq 0.
\]

This helps establish the convergence result

Let \( A \) an SPD matrix. Then, the \( A \)-norms of the error vectors \( d_k = x^* - x_k \) generated by steepest descent satisfy:
\[
\| d_{k+1} \|_A \leq \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\lambda_{\text{max}} + \lambda_{\text{min}}} \| d_k \|_A
\]

Algorithm converges for any initial guess \( x_0 \).
Proof: Observe \( \|d_{k+1}\|_A^2 = (Ad_{k+1},d_{k+1}) = (r_{k+1},d_{k+1}) \)

- by substitution, \( \|d_{k+1}\|_A^2 = (r_{k+1},d_k - \alpha_k r_k) \)

- By construction \( r_{k+1} \perp r_k \) so we get \( \|d_{k+1}\|_A^2 = (r_{k+1},d_k) \).

Now:

\[
\|d_{k+1}\|_A^2 = (r_k - \alpha_k Ar_k, d_k) = (r_k, A^{-1}r_k) - \alpha_k(r_k, r_k) = \|d_k\|_A^2 \left( 1 - \frac{(r_k,r_k)}{(r_k,Ar_k)} \times \frac{(r_k,r_k)}{(r_k,A^{-1}r_k)} \right).
\]

Result follows by applying the Kantorovich inequality. \( \square \)

2. Minimal residual iteration.

A positive definite \((A + A^T)\) is SPD. Take at each step \( d = r \) and \( e = Ar \).

Iteration:

\[
\begin{align*}
r &\leftarrow b - Ax, \\
\alpha &\leftarrow (Ar, r)/(Ar, Ar) \\
x &\leftarrow x + \alpha r
\end{align*}
\]

- Each step minimizes \( f(x) = \|b - Ax\|_2^2 \) in direction \( r \).

- Converges under the condition that \((A + A^T)\) is SPD.

As is formulated, the above algorithm would require 2 'matvecs' at each step. Reformulate it so that only one matvec is required.

Convergence

Let \( A \) be a real positive definite matrix, and let

\[
\mu = \lambda_{\min}(A + A^T)/2, \quad \sigma = \|A\|_2.
\]

Then the residual vectors generated by the Min. Res. Algorithm satisfy:

\[
\|r_{k+1}\|_2 \leq \left( 1 - \frac{\mu^2}{\sigma^2} \right)^{1/2} \|r_k\|_2
\]

- In this case Min. Res. converges for any initial guess \( x_0 \).
3. Residual norm steepest descent.

A is arbitrary (nonsingular). Take at each step $d = A^T r$ and $e = Ad$.

\[
\begin{align*}
    r &\leftarrow b - Ax, \\
    d &\leftarrow A^T r \\
\end{align*}
\]

Iteration:
\[
\begin{align*}
    \alpha &\leftarrow \frac{\|d\|^2_2}{\|Ad\|^2_2} \\
    x &\leftarrow x + \alpha d \\
\end{align*}
\]

▷ Each step minimizes $f(x) = \|b - Ax\|^2_2$ in direction $-\nabla f$.

▷ Important Note: equivalent to usual steepest descent applied to normal equations $A^T Ax = A^T b$.

▷ Converges under the condition that $A$ is nonsingular.