Set 8: Iterative Methods for Solving Equations: Part 2

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 1
Fall 2012

Examples for Jacobi and Gauss-Seidel

$$A_0 = \begin{pmatrix} 3 & 7 & -1 \\ 7 & 4 & 1 \\ -1 & 1 & 2 \end{pmatrix} \quad A_1 = \begin{pmatrix} 3 & 0 & 4 \\ 7 & 4 & 2 \\ -1 & -1 & 2 \end{pmatrix} \quad A_2 = \begin{pmatrix} -3 & 3 & -6 \\ -4 & 7 & -8 \\ 5 & 7 & -9 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 4 & 1 & 1 \\ 2 & -9 & 0 \\ 0 & -8 & -6 \end{pmatrix} \quad A_4 = \begin{pmatrix} 7 & 6 & 9 \\ 4 & 5 & -4 \\ -7 & -3 & 8 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} 6 & -2 & 0 \\ -1 & 2 & -1 \\ 0 & -6/5 & 1 \end{pmatrix} \quad A_6 = \begin{pmatrix} 5 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3/2 & 1 \end{pmatrix}$$

Examples for Jacobi and Gauss-Seidel

$$A_7 = \begin{pmatrix} 4 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 4 \end{pmatrix}$$

Examples for Jacobi and Gauss-Seidel

$$A_8 = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

Solution x is all 1's and $x_0 = 0$. Iterated until $\|x_k - x\|_2 / \|x\|_2 \leq 10^{-6}$.

A	n	$\rho(G_J)$	$\|G_J\|_2$	k_J	$\rho(G_{GS})$	$\|G_{GS}\|_2$	k_{GS}
A_0	3	2.15	2.42	$--$	4.73	5.79	$--$
A_1	3	1.23	1.94	$--$	0.25	2.28	12
A_2	3	0.81	2.49	55	1.11	2.69	$--$
A_3	3	0.44	1.36	18	.019	0.38	5
A_4	3	0.64	1.73	33	0.77	2.55	56
A_5	3	0.88	1.25	105	0.77	0.83	52
A_6	3	0.92	1.51	173	0.85	0.92	84
A_7	7	0.46	0.46	18	0.21	0.32	11
A_8	7	0.92	0.92	174	0.85	0.89	88

Examples for Jacobi and Gauss-Seidel

Consider:

- $\rho(G)$ vs. $\|G\|$
- Jacobi behavior vs. Gauss-Seidel behavior
- n vs k

$\therefore$ We need:

- acceleration/preconditioning to keep $k << n$
- algebraic properties of classes of matrices where Jacobi and Gauss-Seidel (and accelerated methods) have predictable behavior
- practical characteristics that can serve as sufficient conditions

Convergence Analysis via Eigenvalues of G

Let $M = I$ to get accelerated simple Richardson's method. What is $\rho(G) = \rho(I - \alpha A)$? Assume $\alpha > 0$ and G has real eigenvalues, e.g., A is symmetric.

$$Ax = \lambda x \rightarrow x - \alpha Ax = x - \alpha\lambda x \rightarrow (I - \alpha A)x = (1 - \alpha\lambda)x$$

$$\therefore Gx = \mu x$$

$$\alpha\lambda_{min} \leq \alpha\lambda_{max} \rightarrow -\alpha\lambda_{max} \leq -\alpha\lambda_{min}$$

$$1 - \alpha\lambda_{max} \leq \mu \leq 1 - \alpha\lambda_{min}$$

(Note this includes sign.)

Convergence Conditions

Lemma 8.1. *Suppose A has real eigenvalues, λ_i.*

If $\lambda_{min} < 0$ and $\lambda_{max} > 0$, i.e., A is indefinite then at least one eigenvalue of G is greater than 1 in magnitude. In particular,

$$\mu_{max} = 1 - \alpha\lambda_{min} \geq 1$$

and, therefore, for any α, there exists x_0 such that Richardson's method diverges.

Example: $\lambda_{min} = -1$ and $\lambda_{max} = 5$

$$\mu_{min} = 1 - 5\alpha \leq \cdots \leq \mu_i \leq \cdots \leq 1 + \alpha = \mu_{max}$$

Convergence Conditions

Lemma 8.2. *Suppose A has real eigenvalues, λ_i, with $\lambda_{min} > 0$ and $\alpha > 0$ then convergence is guaranteed if*

$$\alpha < \frac{2}{\lambda_{max}}$$

Proof. We must have

$$-1 < 1 - \alpha\lambda_{max} \leq 1 - \alpha\lambda_{min} < 1$$

$\alpha\lambda_{min} > 0 \rightarrow 1 - \alpha\lambda_{min} < 1$. We must therefore have

$$1 - \alpha\lambda_{max} > -1$$

$$\therefore \alpha < \frac{2}{\lambda_{max}}$$

$\square$

Optimal Parameter

Spectral radius is

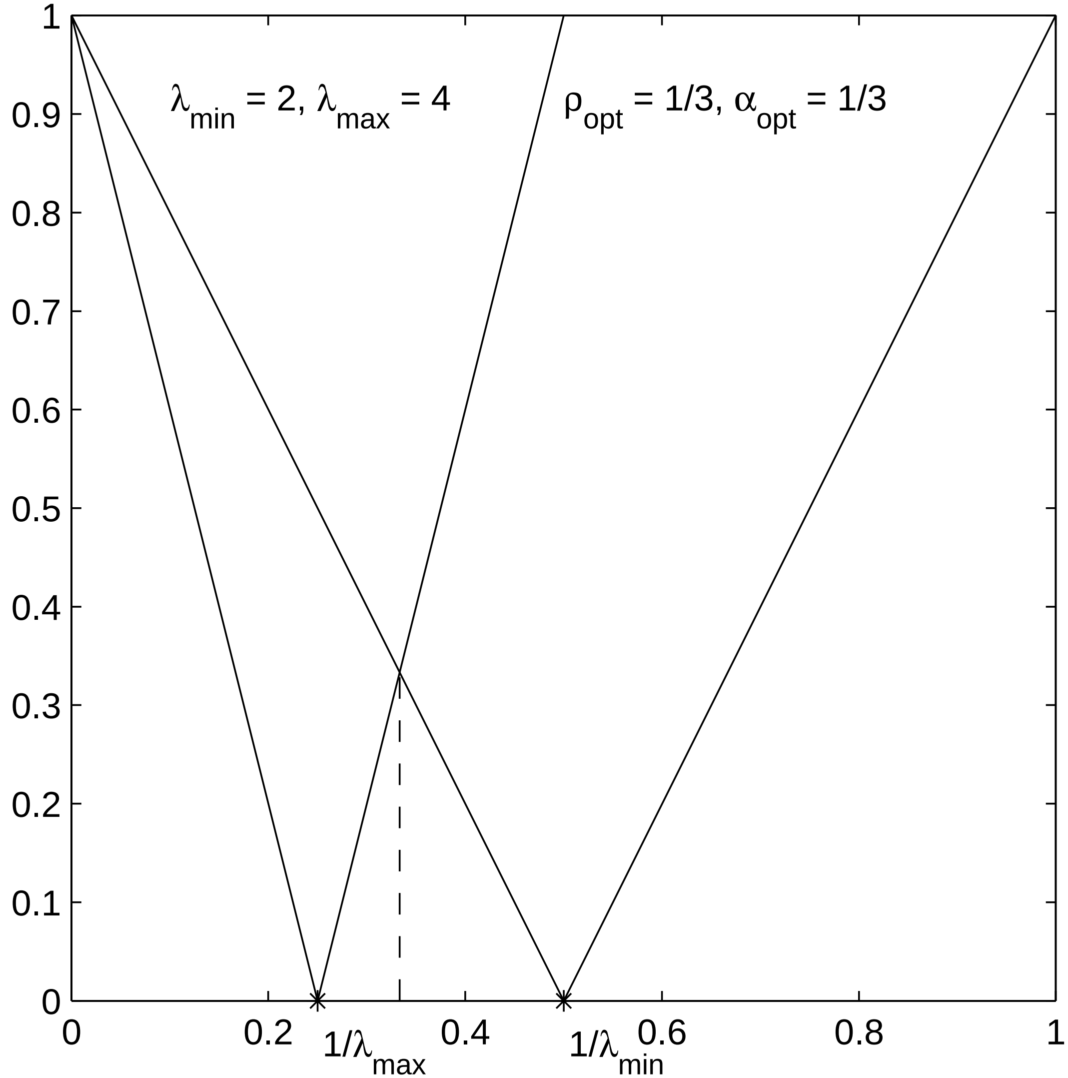
$$\rho(G(\alpha)) = max\left\{|1 - \alpha\lambda_{max}|, |1 - \alpha\lambda_{min}|\right\}$$

this has its minimum at

$$\alpha_{opt} = \frac{2}{\lambda_{max} + \lambda_{min}} = \frac{2}{\lambda_{max}}\left\{\frac{1}{1 + \frac{\lambda_{min}}{\lambda_{max}}}\right\} = \frac{2}{\lambda_{max}}\left\{\frac{1}{1+\gamma}\right\}$$

Substituting this into the spectral radius

$$\rho_{opt} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} = \frac{1-\gamma}{1+\gamma}$$



$\lambda_{min} = 2$, $\lambda_{max} = 4$
$\rho_{opt} = 1/3$, $\alpha_{opt} = 1/3$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.2
$1/\lambda_{max}$
0.4
$1/\lambda_{min}$
0.6
0.8
1

Observations

- The method does not converge for all matrices.
- Eigenvalue estimates needed to approximate α.
- As $\gamma \to 1$ the optimal convergence rate improves, i.e., $\rho \to 0$
- As $\gamma \to 0$ the optimal convergence rate slows, i.e., $\rho \to 1$
- In general, the more eccentric the spectrum the slower the convergence.
- The text has a bit more general discussion of this topic.

Convergence Theory for SOR

- It is possible to analyze the convergence of SOR and determine the optimal relaxation parameter.
- The basic problem is to identify characteristics of matrices that allow the analysis eigenvalues of the iteration matrices for SOR and Jacobi.

$$G_J = D^{-1}(L+U)$$

$$G_{SOR} = (D - \omega L)^{-1} \left[\omega U + (1-\omega) D\right]$$

- The early work of Young on matrices with "Property A" and consistent orderings has been generalized by Varga via p-cyclic matrices and $T-$matrices.
- See text, Hageman and Young or Varga books.

Matrix Properties of Interest

Types:

- Algebraic:
  - symmetric positive definite matrices
  - diagonally dominant matrices : strictly or irreducibly
  - M matrices
- Structural:
  - block tridiagonal matrices
  - Property A

These were analyzed for the classical methods because matrices from finite difference approximations to PDEs tend to have one or more of these structures.

Symmetric Positive Definite Matrices

Theorem 8.3. *Suppose $A \in \mathbb{R}^{n \times n}$ and $A = A^T$. A is positive definite if and only if one of the following is satisfied:*

- $\forall x \in \mathbb{R}^n, \;\; x \neq 0 \rightarrow x^T A x > 0$
- *If λ is an eigenvalues of A then $\lambda > 0$. (symmetry implies $\lambda \in \mathbb{R}$)*
- $\exists H \in \mathbb{R}^{n \times n}$ *such that H is nonsingular and $A = H^T H$.*

Symmetric Positive Definite Matrices

- A symmetric positive definite matrix is nonsingular.
- If the strict inequalities are replaced by $\geq$ the matrix is semidefinite (and singular).
- A matrix need not be symmetric to be positive definite. There are more general definitions for nonsymmetric and nonhermitian.

Convergence and Positive Definiteness

Theorem 8.4. *Suppose* $A \in \mathbb{R}^{n \times n}$ *is symmetric positive definite and it has a splitting* $A = M - N$.

If $(M + M^T) - A$ *is symmetric positive definite then* M^{-1} *exists and the iterative method defined by the splitting has*

- $\rho(G) \leq \|G\|_A < 1$ *(convergent)*
- $\|e^{(k+1)}\|_A < \|e^{(k)}\|_A$ *(monotonic)*

Convergence and Positive Definiteness

Corollary 8.5. *If, additionally,* M *and* $2M - A$ *in Theorem 8.4 are both symmetric positive definite then*

- $\rho(G) = \|G\|_M = \|G\|_A < 1$ *(convergent)*
- $\|e^{(k+1)}\|_A < \|e^{(k)}\|_A$ *(monotonic)*
- $\|e^{(k+1)}\|_M < \|e^{(k)}\|_M$ *(monotonic)*

Convergence and Positive Definiteness

Theorem 8.6. *If A is symmetric positive definite then for any $0 < \omega < 2$ SOR and SSOR converge for any x_0.*

Note. Setting $\omega = 1$ yields Gauss-Seidel and Symmetric Gauss-Seidel in Theorem 8.6.

Theorem 8.7. *If A is symmetric positive definite then for any $0 < \omega < \mu$ JOR converges for any x_0, where $\mu = 2/\rho(D^{-1}A)$ and $D = diag(A)$.*

Note. Note that A symmetric positive definite does not guarantee Jacobi converges.

Diagonal Dominance

Definition 8.1. A matrix A is (by columns)

- weakly diagonally dominant if, for $1 \le j \le n$,
  $|\alpha_{jj}| \ge \sum_{i=1,\ i \ne j}^{n} |\alpha_{ij}|$
- strictly diagonally dominant if, for $1 \le j \le n$,
  $|\alpha_{jj}| > \sum_{i=1,\ i \ne j}^{n} |\alpha_{ij}|$
- irreducibly diagonally dominant if, for $1 \le j \le n$,
  $|\alpha_{jj}| \ge \sum_{i=1,\ i \ne j}^{n} |\alpha_{ij}|$, A is irreducible, and strict inequality for at least one j.

Note. "By rows" forms of dominance are also easily created analogously, e.g., see textbook.

Eigenvalue Locations

Pages 184–186 of the textbook contain very useful definitions and theorems related to locating eigenvalues in regions of the complex plane. They include:

- A theorem giving a bounding box for the eigenvalues of $A \in \mathbb{C}^{n \times n}$.
- The definition and a simple sufficient condition for A to be irreducible.
- The definition of Gershgorin row and column circles.
- The first two Gershgorin theorems that relate the location of eigenvalues of a matrix to its row and column circles.
- The third Gershgorin theorem that relates the eigenvalues of an irreducible matrix to the row circles.

Convergence

Theorem 8.8. *If A is strictly or irreducibly diagonally dominant then A is nonsingular and the associated Jacobi and Gauss-Seidel iterations converge for any x_0.*

If A is strictly diagonally dominant then SOR converges for any $0 < \omega \leq 1$.

Convergence

For Jacobi, strict diagonal dominance is easily seen to guarantee convergence. $G_J = D^{-1}(L+U)$ and strict diagonal dominance implies that the i-th row of G_J has elements

$$\gamma_{ii} = 0$$

$$\gamma_{ij} = \frac{\alpha_{ij}}{\alpha_{ii}}$$

$$\|e_i^T G_j\|_1 = \sum_{j=1, j\neq i}^{n} \frac{|\alpha_{ij}|}{|\alpha_{ii}|} = \frac{1}{|\alpha_{ii}|}\left(\sum_{j=1, j\neq i}^{n} |\alpha_{ij}|\right) < 1$$

$$\|G_J\|_\infty < 1$$

M Matrices

Definition 8.2. A matrix $A \in \mathbb{R}^{n \times n}$ is an M-matrix if the following three conditions hold:

1. $\alpha_{ii} > 0$ for $1 \leq i \leq n$
2. $\alpha_{ij} \leq 0$ for $i \neq j$
3. $A^{-1} \geq 0$, i.e., all elements are nonnegative.

M Matrices

A more practical sufficient condition:

Lemma 8.9. *If* $A \in \mathbb{R}^{n \times n}$ *and*

1. $\alpha_{ii} > 0$ *for* $1 \leq i \leq n$
2. $\alpha_{ij} \leq 0$ *for* $i \neq j$
3. A *is strictly diagonally dominant*

then A *is an M-matrix.*

M Matrices

Another sufficient condition:

Lemma 8.10. *If A can be written $A = I - P$ where $P \geq 0$ and $\rho(P) < 1$ then A is an M-matrix.*

- The proof follows from the theorem about $(I - B)^{-1}$ used when discussing convergent matrices .
- This can easily be modified to give a necessary and sufficient condition.
- There are many other interesting characterizations of M-matrices. The textbook discusses some.

Convergence and M-matrices

Theorem 8.11. *If A is such that*

1. $\alpha_{ii} > 0, \quad 1 \le i \le n$
2. $\alpha_{ij} \le 0$ *for* $i \neq j$

then A is an M-matrix if and only if Jacobi converges, i.e., $\rho(I - D^{-1}A) < 1$ *where* $D = diag(\alpha_{11}, \ldots, \alpha_{nn})$

Convergence and M-matrices

Definition 8.3. $A = M - N$ is a regular splitting if

- M is nonsingular
- $M^{-1} \geq 0$
- and $N \geq 0$

Convergence and M-matrices

Theorem 8.12. *If $A = M - N$ is a regular splitting and A is an M-matrix then $\rho(M^{-1}N) < 1$.*

- Regular splitings of M-matrices generate convergent linear fixed point methods.
- How do you generate a regular splitting of an M-matrix?

Convergence and M-matrices

- If A is an M-matrix and the splitting $A = M - N$ simply puts the diagonal elements and selected off-diagonal elements into M and the rest of the off-diagonal elements into N then M is an M-matrix and the splitting is regular.
- Easy to see if A is a strictly diagonally dominant M-matrix.
- This does not say anything about which choice of elements for inclusion in M results in an acceptable convergence rate.
- The more elements that are included the more expensive the system solve required on each step becomes.

Convergence and Structure

Theorem 8.13. *If $A \in \mathbb{R}$ is nonsingular and either tridiagonal or block tridiagonal then*

$$\rho(G_{GS}) = \rho^2(G_J)$$

Under the assumptions of Theorem 8.13

- Gauss-Seidel and Jacobi either both diverge or converge.
- If A is additionally symmetric positive definite then Jacobi converges since Gauss-Seidel converges for any symmetric positive definite matrix.
- Additional assumptions about Property A and consistent ordering can be invoked to analyze SOR. (see text)

Examples for Jacobi and Gauss-Seidel

$$A_7 = \begin{pmatrix} 4 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 4 \end{pmatrix}$$

Examples for Jacobi and Gauss-Seidel

- A_7 is tridiagonal $\rightarrow$ behaviors of Jacobi and Gauss-Seidel are the same (but not the rates).
- A_7 is symmetric positive definite $\rightarrow$ Gauss-Seidel converges
- A_7 is strictly diagonally dominant $\rightarrow$ Jacobi and Gauss-Seidel converge
- A_7 is irreducibly diagonally dominant $\rightarrow$ Jacobi and Gauss-Seidel converge
- A_7 is an M-matrix $\rightarrow$ Jacobi converges

Examples for Jacobi and Gauss-Seidel

$$A_8 = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

Examples for Jacobi and Gauss-Seidel

- A_8 is tridiagonal $\rightarrow$ behaviors of Jacobi and Gauss-Seidel are the same (but not the rates).
- A_8 is symmetric positive definite $\rightarrow$ Gauss-Seidel converges
- A_8 is irreducibly diagonally dominant $\rightarrow$ Jacobi and Gauss-Seidel converge
- A_8 is an M-matrix $\rightarrow$ Jacobi converges