CHAPTER 5

Basic Iterative Methods

Solve $Ax = f$ where $A$ is large and sparse (and nonsingular). Let $A$ be split as $A = M - N$ in which $M$ is nonsingular, and solving systems of the form $Mz = r$ is much easier than solving $Ax = f$.

Iteration:

$$Mx_{k+1} = Nx_k + f$$

$$x_0 := \text{arbitrary}$$

Convergence:

$$x_{k+1} = M^{-1}Nx_k + M^{-1}f = Hx_k + g$$

but, $x = Hx + g$. Then if $\delta x_k = x_k - x$, we have $\delta x_{k+1} = H\delta x_k$, i.e.,

$$\delta x_1 = H\delta x_0$$
$$\delta x_2 = H\delta x_1 = H^2\delta x_0$$
$$\vdots$$
$$\delta x_k = H^k\delta x_0$$

or

$$\|\delta x_k\| \leq \|H^k\|\|\delta x_0\| \leq \|H\|^k\|\delta x_0\|.$$

Thus, a sufficient condition for convergence ($\lim_{k \to \infty} \|H\|^k = 0$) is $\|H\| < 1$.

Necessary condition:

Let $x_0 = 0$, then $x_{k+1} = (I + H + H^2 + \cdots + H^k)g$

$$\lim_{k \to \infty} \left(\sum_{j=0}^{k} H^j\right) \text{ converges iff } \rho(H) < 1.$$  In this case $x_k \to (I - H)^{-1}g$. Going back to

$$\|\delta x\| \leq \|H\|^k\|\delta x_0\|$$

since

$$\rho(H) \leq \|H\| \leq \rho(H) + \epsilon ; \epsilon > 0.$$  

then

$$\|\delta x_k\| \leq (\rho(H) + \epsilon)^k\|\delta x_0\|$$
If $\delta x_0$ is an eigenvector of $H$ corresponding to the eigenvalue of max. modules, i.e.,

$$H\delta x_0 = \mu \delta x_0 \ ; \ \rho(H) = |\mu|$$

\[ \therefore \delta x_k = H^k \delta x_0 = \mu^k \delta x_0 \]

$$& \| \delta x_k \| = \rho^k(H) \| \delta x_0 \| .$$

Now, to reduce the error by a factor of $10^m$, $m > 0$, we need have $\rho^k(H) \leq 10^{-m}$ or $10^m = [1/\rho(H)]^k$, i.e., $m \leq k \log_{10}[1/\rho(H)] = k \cdot R^*$ where $R^*$ is called the asymptotic rate of convergence of the iterative scheme, and

$$k \geq \frac{m}{R^*} = \frac{-m}{\log_{10} \rho(H)}$$

\[ \uparrow \]

# of iterations needed to reduce the error by a factor of $10^m$.

**Example:**

Let $\rho(H) = 0.9$ and $m = 6$, then $k \geq \frac{6}{-\log_{10}(0.9)} \simeq 130$ iterations. However, if $\rho(H) = 0.2$, then $k \geq \frac{6}{-\log_{10}(0.2)} \simeq 9$ iterations.

**Algorithm:** $x_{k+1} = M^{-1} N x_k + M^{-1} f$ but $A = M - N$, or $N = M - A$, then

$$x_{k+1} = M^{-1} (M - A) x_k + M^{-1} f$$

$$= x_k + M^{-1} (f - Ax_k)$$

$$= x_k + M^{-1} r_k$$

**Summary:**

$x :=$ arbitrary

$r := f - Ax$

while $\| r \|$ $\geq$ tolerance

Solve $M d = r$

$x := x + d$

$r := f - Ax$

end
Example:

Solve $A x = f$

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}; \quad f = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}. $$

Exact solution:

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. $$

Let $A = M - N,$

$$M = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}; \quad N = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and $x_0^T = (0, 0, 0).$ Then

$$H = M^{-1}N = \frac{1}{4} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$\therefore x_{k+1} = Hx_k + g; \quad g = M^{-1}f = \frac{1}{4} \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$

$$\lambda_k(H) = \frac{1}{4} \left(2 \cos \frac{k\pi}{4}\right) \quad k = 1, 2, 3$$

$\therefore \lambda(H) := \frac{\sqrt{2}}{4}; \quad 0; \quad -\frac{\sqrt{2}}{4}$

i.e.,

$$\rho(H) = \frac{\sqrt{2}}{4} \approx 0.3536,$$

& \quad R^* = -\log_{10}(0.3536) \approx 0.45.$$

Normally, the stopping criterion is either $\| r_k \|_2 < \text{tol1},$ or $\| r_k \|_2 / \| r_0 \|_2 \leq \text{tol2},$ where tol 1 and tol 2 are tolerances to be determined by the user. A stopping criterion based on the error in the solution is given by the following theorem.

**Theorem:** Let $A = M - N,$ with $H = M^{-1}N,$ and $\| H \| < 1.$ Thus

$$\| x - x_k \| \leq \frac{\| H \|}{1 - \| H \|} \| x_k - x_{k-1} \| \quad k = 1, 2, 3, \ldots$$
Proof:

\[ x_{k+1} = Hx_k + g; \quad g = M^{-1}f \]
\[ x_k = Hx_{k-1} + g \]

subtracting, we get

\[ x_{k+1} - x_k = H(x_k - x_{k-1}) \]

and by recursion, we have

\[ (x_{m+k+1} - x_{m+k}) = H^{m+1}(x_k - x_{k-1}) \]

Since,

\[ (x_{k+p} - x_k) = (x_{k+p} - x_{k+p-1}) + \]
\[ (x_{k+p-1} - x_{k+p-2}) + \]
\[ \vdots \]
\[ + (x_{k+1} - x_k) \]
\[ = \sum_{m=0}^{p-1} (x_{k+m+1} - x_{k+m}) \]

we have

\[ \| x_{k+p} - x_k \| \leq \sum_{m=0}^{p-1} \| H \|^{m+1} \| x_k - x_{k-1} \|. \]

But

\[ \sum_{m=0}^{p-1} \| H \|^{m+1} = \| H \| [1 + \| H \| + \| H \|^{2} + \cdots + \| H \|^{p-1}] . \]

Thus, since \( \| H \| < 1 \), we get

\[ (1- \| H \|)(1+ \| H \| + \| H \|^{2} + \cdots + \| H \|^{p-1}) = 1- \| H \|^{p} \]

i.e.,

\[ \| x_{k+p} - x_k \| \leq \frac{\| H \| (1- \| H \|^{p})}{1- \| H \|} \| x_k - x_{k-1} \| \]

as \( p \to \infty \)

\[ \| x - x_k \| \leq \frac{\| H \|}{1- \| H \|} \| x_k - x_{k-1} \|. \]
Example:

\[ A = \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} ; \ f = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \]

\[ A = M - N ; \ M = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \]

\[ N = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} \]

\[ x_0 = 0 ; \ H = M^{-1}N = \begin{pmatrix} 0 & 1/2 \\ 3/4 & 0 \end{pmatrix} \]

\[ \| H \|_\infty = \| H \|_2 \ 3/4 < 1 \]

\[ x_3 = \frac{1}{32} \begin{bmatrix} 26 \\ 23 \end{bmatrix} ; \ x_4 = \frac{1}{64} \begin{bmatrix} 55 \\ 55 \end{bmatrix} \]

\[ \therefore \ \| x - x_4 \|_\infty \leq \frac{3/4}{1 - 3/4} \ \| x_4 - x_3 \|_\infty = 3 \times \frac{1}{64} \times 9 = \frac{27}{64} \]

Since \[ x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \], then \[ \| x - x_4 \|_\infty = \frac{9}{64} \] or 1/3 of the upper bound above.

Accelerate the Iterative Scheme

\[ A = M - N \]

\[ x_{k+1} = x_k + M^{-1}r_k \] (Original scheme)

\[ r_k = f - Ax_k \].

From \((I - H)x = g, H = M^{-1}N; g = M^{-1}f\), the original scheme may also be expressed as,

\[ x_{k+1} = Hx_k + g \]

\[ = x_k + [g - (I - H)x_k] \]

\[ = x_k + r_k. \]

We can accelerate this basic scheme via

\[ x_{k+1} = x_k + \alpha \tilde{r}_k \] (**)

in which \(\alpha\) is a scalar and \(\tilde{r}_k = M^{-1}r_k\). Iteration (***) could be written as

\[ x_{k+1} = [I - \alpha G]x_k + \alpha g \]

5
where $G = (I - H)$. Necessary condition for convergence is: \( \rho(I - \alpha G) < 1 \).

**Objective:** Choose \( \alpha \) so that \( \rho(I - \alpha G) \) is minimized. Let \( \lambda(G) := [\lambda_{\min}, \lambda_{\max}] \). Let \( \lambda_k(I - \alpha G) = \mu_k \), thus \( 1 - \alpha \lambda_{\max} \leq \mu_k \leq 1 - \alpha \lambda_{\min} \). Note that if \( \lambda_{\min} < 0 \), then for some \( k \), \( \mu_k > 1 \) and the iteration diverges. Therefore, we assume that \( \lambda_{\min} > 0 \), i.e., that the basic scheme is convergent. Hence

\[
0 < \lambda_{\min} < \lambda_{\max}.
\]

Since \( |\mu_k| = |1 - \alpha \lambda_k| \), we have the st. lines

\[
|1 - \alpha \lambda_{\max}|
\]

\[
|1 - \alpha \lambda_{\min}|
\]

for \( \alpha_{\text{optimal}} \), we must have

\[
-(1 - \alpha \lambda_{\max}) = (1 - \alpha \lambda_{\min})
\]

\[
\therefore \quad \alpha_{\text{optimal}} = \frac{2}{\lambda_{\min} + \lambda_{\max}}
\]

and

\[
\rho(I - \alpha_{\text{optimal}} G) = 1 - \frac{2\lambda_{\min}}{\lambda_{\min} + \lambda_{\max}} = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\min} + \lambda_{\max}}
\]

**Diagonally Dominant Matrices**

* \( A_{(n\times n)} \) is strictly diagonal dominant if:

\[
|a_{ii}| > \sum_{\substack{j=1\atop j \neq i}}^{n} |a_{ij}|.
\]

* \( A \) is weakly diagonally dominant if:

\[
|a_{ii}| \geq \sum_{\substack{j=1\atop j \neq i}}^{n} |a_{ij}|.
\]
* A is irreducibly diagonally dominant if:

    (i) there does not exist a permutation $P$ for which

    $$ PAP^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix} $$

    where $B_{11}$ and $B_{22}$ are square matrices.

    (ii) $|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$ with strict inequality holding for at least one row $k$.

**Gerschgorin’s Theorems**

1. Let $A \in \mathbb{C}^{n \times n}$, then each eigenvalue of $A$ lies in one of the disks in the complex plane

    $$ |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|. $$

**Proof:** Consider the eigenvalue problem $Ax = \lambda x$ where, $|x_k| \geq \max_{1 \leq j \leq n} |x_j|$. Since $(\lambda I - A)x = 0$, then from the $k$-th row

    $$(\lambda - a_{kk})x_k = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{kj}x_j.$$ 

    $$|\lambda - a_{kk}| = \frac{1}{|x_k|} \left| \sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj}x_j \right|$$

    $$\leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \left| \frac{x_j}{x_k} \right|$$

    or $|\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}|.$
Example:

\[
A = \begin{bmatrix}
1 & 2 & -3 \\
1 & 5 & -1 \\
-3 & 2 & 1
\end{bmatrix}
\]

\[\lambda(A) := 0.7898 \pm 3.2885i\]

\[5.4205\]

All 3 eigenvalues are contained in the union of the 3 disks.

2. If \(k\) Gershgorin disks are disjoint from the rest, then exactly \(k\) eigenvalues lie in the union of the \(k\) disks.

Example:

\[
A = \begin{bmatrix}
1 & 1/2 & -1/2 \\
1 & 5 & -1 \\
-1/2 & 1/2 & 1
\end{bmatrix}
\]

\[\lambda(A) := 0.5, 1.5, 5.0\]
Theorem: If $A$ is strictly diagonally dominant or irreducibly diagonally dominant, then it is nonsingular.

Proof:

(a) If $A$ is strictly diagonally dominant then zero is not contained in the union of the $G$-disks.

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

(b) If $A$ is irreducibly diagonally dominant, then let $A$ be singular, i.e., there exists an $x$ s.t. $Ax = 0$.

Let $m$ be that row of $A$ for which

$$|a_{mm}| > \sum_{j=1, j \neq m}^{n} |a_{mj}|.$$ 

Let $J$ be the set of indices $J = \{k : |x_k| \geq |x_i|, i = 1, 2, ..., n, \text{ and } |x_k| > |x_j| \text{ for } 1 \leq j \leq n\}$. Clearly, $J$ is not empty, for this would imply

$$|x_1| = |x_2| = \cdots = |x_n| \neq 0$$
contradicting $|a_{mm}| > \sum_{j \neq m} |a_{mj}|$. Now, for any $k \in J$, we have

$$|a_{kk}| \leq \sum_{j \neq k} \left( \frac{1}{|x_k|} \right) |a_{kj}| |x_j|.$$

Thus, to maintain the property of diagonal dominance, we must have $a_{kj} = 0$ whenever $\frac{|x_j|}{|x_k|} < 1$ for $k \in J; j \not\in J$. But, since $A$ is irreducible, this is a contradiction. Q.E.D.

**Theorem:** If $A$ is strictly diagonally dominant, or irreducibly-diagonally dominant, then the associated Jacobi and Gauss-Seidel iterations converge for any initial iterate $x_0$.

**Proof:** Let $A = D - L - U$, where $D$ is diagonal and $L, U$ are strictly lower and upper triangular, respectively.

(a) strictly diagonally dominant case:

(i) Jacobi:

$$A = M_J - N_J$$

$$M_J = D ; \ N_J = D + L$$

$$\therefore \ H_J = M_J^{-1} N_J$$

$$= M_J^{-1} (M_J - A)$$

$$= I - M_J^{-1} A$$

$$= (I - D^{-1} A)$$

i.e., $H_J$ is of the form

$$H_J = \begin{bmatrix}
0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\
-\frac{a_{21}}{a_{22}} & 0 & \ddots & \vdots \\
\vdots & \ddots & \ddots & -\frac{a_{n-1,n}}{a_{n-1,n}} \\
-\frac{a_{n1}}{a_{nn}} & \cdots & -\frac{a_{n,n-1}}{a_{n,n}} & 0 
\end{bmatrix}.$$
\[ \rho(H_J) < 1 \]

as well as,

\[ \| H_J \|_\infty < 1 \]

\( \uparrow \) sufficient condition for convergence.

(ii) Gauss-Seidel:

\[ H_{G.S.} = (D - L)^{-1}U. \]

Consider the eigenvalue problem

\[ H_{G.S.}w = \lambda w \]

or

\[ Uw = \lambda(D - L)w. \]

Also, let \( w \) be scaled such that for some \( m \), \( |w_m| = 1 \), and \( |w_i| \leq 1 \). Then, from

\[- \sum_{j > m} a_{mj}w_j = \lambda \left[ a_{mm}w_m + \sum_{j < m} a_{mj}w_j \right] \]

i.e.,

\[ |\lambda| = \left| \sum_{j > m} a_{mj}w_j \right| / \left| a_{mm}w_m + \sum_{j < m} a_{mj}w_j \right| \]

or

\[ |\lambda| \leq \frac{\sum_{j > m} |a_{mj}|}{|a_{mm}| - \sum_{j < m} |a_{mj}|} \]

\[ = \frac{\sigma_2}{d - \sigma_1} \]

\[ = \frac{\sigma_2}{\sigma_2 + (d - \sigma_2 - \sigma_1)} \]
where,
\[ d = |a_{mm}| ; \quad \sigma_2 = \sum_{j>m} |a_{mj}| , \]
\[ \sigma_1 = \sum_{j<m} |a_{mj}| \]
and \( d > (\sigma_1 + \sigma_2) \).
Hence, \( |\lambda| < 1 \) and G.S. converges.

(a) Irreducibly diagonally dominant case:

(i) Jacobi:

From the Gerschgorin theorems, we can only show that
\[ \rho(M_J^{-1}N_J) \leq 1. \]
Let \( \rho(H_J) = 1 \), where \( H_J = M_J^{-1}N_J \) then \( (H_J - \lambda I) \) is singular with \( |\lambda| = 1 \). Hence, for \( |\lambda| = 1 \), \( M_J^{-1}(N_J - \lambda M_J) \) is singular. Since \( A = M_J - N_J \) is irreducibly diagonally dominant, then for any scalar \( \alpha \) with \( |\alpha| = 1 \), we see that \( (\alpha M_J - N_J) \) is also irreducibly diagonally dominant, i.e., \( (\alpha M_J - N_J) \) is nonsingular. Then, \( (\lambda M_J - N_J) \) is nonsingular for \( |\lambda| = 1 \Rightarrow \) contradiction if \( \lambda \) is an eigenvalue of \( H_J \). Consequently, \( \rho(H_J) < 1 \).

(ii) Gauss-Seidel:

The proof that \( \rho(H_{G.S.}) < 1 \) is identical to that of Jacobi above.

Property A:

A matrix \( A \) is 2-cyclic, or has property \( \& \) if there is a permutation \( P \) such that
\[ PAP^T = \begin{bmatrix} D_1 & E \\ F & D_2 \end{bmatrix} \]
where \( D_1 \) and \( D_2 \) are diagonal. The tridiagonal matrix
\[ A = \begin{bmatrix} 2 & -1 \\ -1 & 2 & -1 \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \]
has property $A$ with the Permutation $P$ given by,

$$P^T = [e_1, e_3, e_5, \ldots; e_2, e_4, e_6, \ldots].$$

Let $A = \begin{pmatrix} D_1 & E \\ F & F_2 \end{pmatrix}$ with $D_1, D_2$ diagonal.

(i) Jacobi:

$$H_J = \begin{pmatrix} D_1^{-1} & 0 \\ 0 & D_2^{-1} \end{pmatrix} \begin{pmatrix} 0 & -E \\ -F & 0 \end{pmatrix}$$

or

$$H_J = \begin{pmatrix} 0 & -D_1^{-1}E \\ -D_2^{-1}F & 0 \end{pmatrix}.$$ 

Now consider the eigenvalue problem

$$\begin{pmatrix} 0 & -D_1^{-1}E \\ -D_2^{-1}F & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\therefore ( +D_2^{-1}F)(D_1^{-1}E)y = \lambda^2_j y$$

for convergence we must have $\lambda^2_j < 1$.

(ii) Gauss-Seidel:

$$H_{GS} = \begin{pmatrix} D_1 & 0 \\ F & D_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -E \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{D_1} & 0 \\ \frac{1}{-D_2^{-1}F} & \frac{1}{D_2^{-1}F} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -E \end{pmatrix}$$

$$\therefore H_{GS} = \begin{pmatrix} 0 & -D_1^{-1}E \\ 0 & D_2^{-1}F D_1^{-1}E \end{pmatrix}.$$ 

Hence,

$$\lambda(H_{GS}) := 0 ; \: \lambda_{GS}(D_2^{-1}F D_1^{-1}E).$$

For convergence we must have

$$\lambda_{GS}[(D_2^{-1}F)(D_1^{-1}E)].$$

Thus,

$$\lambda_{GS} = \lambda^2_j$$

or $GS$ convergence twice as fast as Jacobi for matrices with property $A$. 

13
Positive Definite Systems

Theorem: Let $A$ be symmetric, with $A = M - N$, in which $M$ is nonsingular, and $S = MT + N$ is positive definite. Then,

$$H = M^{-1}N = I - M^{-1}A$$

is convergent (i.e., $\rho(H) < 1$) iff $A$ is positive definite.

Proof: (sufficiency). Let $\lambda, u$ be an eigenpair of $H$, i.e., $Hu = \lambda u$ or

$$(I - M^{-1}A)u = \lambda u$$

$\therefore Au = (1 - \lambda)Mu.$

Hence,

$$(1 - \lambda) = \frac{u^*Au}{u^*Mu}$$

and

$$(1 - \bar{\lambda}) = \frac{u^*Au}{u^*M^Tu}.$$  \hspace{1cm} (2)

From (2), we have

$$\frac{1}{1 - \lambda}(u^*Au) = u^*(A + N^T)u = u^*Au + u^*N^Tu$$

or,

$$(u^*Au)\left(\frac{\bar{\lambda}}{1 - \lambda}\right) = u^*N^Tu.$$  \hspace{1cm} (3)

From (1), we get

$$(u^*Au)\left(\frac{1}{1 - \lambda}\right) = u^*Mu.$$  \hspace{1cm} (4)

Adding (3) and (4), we get

$$(u^*Au)\left[\frac{1 - |\lambda|^2}{(1 - \lambda)(1 - \lambda)}\right] = u^*S^Tu.$$  

Thus, if $\lambda = \alpha + i\beta$, with $i = \sqrt{-1}$, then

$$u^*S^Tu = (u^*Au)\left[\frac{1 - |\lambda|^2}{\gamma}\right]$$

where $\gamma = (1 - \alpha)^2 + \beta^2 > 0$. As a result, if $A$ is s.p.d., and $S$ is s.p.d. then $\rho(H) < 1$. 

14
Observation 1:
Let \( A = D - L - L^T \) s.p.d.

\[ M = D; \quad e_j^T D e_j > 0 \]
\[ S = M^T + N = D + L + L^T \]
\[ \therefore \text{Jacobi converges if } S \text{ is s.p.d.} \]

Observation 2:
Gauss-Seidel
\[ M = D - L; \quad N = L^T \]
\[ S = M^T + N = D - L^T + L^T = D \quad \Rightarrow \text{s.p.d.} \]
\[ \therefore \text{Gauss-Seidel converges for s.p.d. systems.} \]

Symmetric Gauss-Seidel
Let \( A \) be s.p.d. and given by \( A = D - L - L^T \). \( D \) is diagonal, and \( L \) is strictly lower triangular.

Iteration for solving \( Ax = f \)
\[
(D - L)x_{k+1/2} = L^T x_k + f \\
(D - L^T)x_{k+1} = Lx_{k+1/2} + f.
\]

Let,
\[ \mathcal{L} = D^{-1/2} LD^{-1/2} \]
\[ y_k = D^{1/2} x_k. \]

Then the symmetric G.S. iteration is given by,
\[
(I - \mathcal{L}) y_{k+1/2} = \mathcal{L}^T y_k + D^{-1/2} f \\
(I - \mathcal{L}^T) y_{k+1} = \mathcal{L} y_{k+1/2} + D^{-1/2} f
\]
or
\[
(I - \mathcal{L}^T) y_{k+1} = \mathcal{L}(I - \mathcal{L})^{-1} \mathcal{L}^T (I - \mathcal{L})^{-T} (I - \mathcal{L}^T) y_k \\
+ [I + \mathcal{L}(I - \mathcal{L})^{-1}] D^{-1/2} f.
\]

Let \((I - \mathcal{L}^T) y_{k+1} = z_{k+1} \). Then,
\[
z_{k+1} = \mathcal{L}(I - \mathcal{L})^{-1} \mathcal{L}^T (I - \mathcal{L})^{-T} z_k + h = K z_k + h.
\]
Since $L$ is strictly lower triangular (diagonal elements are all zero), then

$$L(I - L)^{-1} = L(I + L + L^2 + \cdots + L^{n-1})$$

or

$$L(I - L)^{-1} = (I + L + L^2 + \cdots + L^{n-1})L = (I - L)^{-1}L.$$

Thus, the iteration matrix $K$ is given by,

$$K = (I - L)^{-1}LL^T(I - L)^{-T}$$

i.e., $K$ is symmetric positive semidefinite, or $\lambda(K) \geq 0$ are all real.

Also,

$$I - K = I - (I - L)^{-1}LL^T(I - L)^{-T}$$

$$= (I - L)^{-1}[(I - L)(I - L)^T - LL^T](I - L)^{-T}$$

i.e.,

$$(I - K) = (I - L)^{-1}[I - L - L^T](I - L)^{-T}$$

$$= (I - L)^{-1} \begin{bmatrix} D^{-1/2}AD^{-1/2} \end{bmatrix}(I - L)^{-T}$$

and $(I - K)$ is s.p.d., or $1 - \lambda(K) > 0$

$\therefore 0 \leq \lambda(K) < 1$. Thus, symmetric G.S. converges for s.p.d. systems.