Classical Numerical Analysis, Chapter 06

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

Chapter 06, Part 2 of 3
Linear Iterative Methods

Relaxation Methods

Relaxation Methods

Let us consider one last classical method, namely the relaxation method. Recall that Gauss–Seidel reads

$$\mathsf{L}\mathbf{x}_{k+1} + \mathsf{D}\mathbf{x}_{k+1} + \mathsf{U}\mathbf{x}_k = \mathbf{f},$$

where $\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}$, with the usual assumptions and notation. We will weight the contribution of the diagonal D by introducing the parameter $\omega > 0$:

$$\mathsf{L}\mathbf{x}_{k+1} + \omega^{-1}\mathsf{D}\mathbf{x}_{k+1} + (1 - \omega^{-1})\mathsf{D}\mathbf{x}_k + \mathsf{U}\mathbf{x}_k = \mathbf{f}.$$

In doing that we obtain the method

$$(\mathsf{L} + \omega^{-1}\mathsf{D})\mathbf{x}_{k+1} = \left((\omega^{-1} - 1)\mathsf{D} - \mathsf{U}\right)\mathbf{x}_k + \mathbf{f}.$$

The Iterator and Error Transfer Matrices

The relaxation method is defined via

$$(\mathsf{L} + \omega^{-1}\mathsf{D})\mathbf{x}_{k+1} = \left((\omega^{-1} - 1)\mathsf{D} - \mathsf{U}\right)\mathbf{x}_k + \mathbf{f}.$$

If we choose $\omega > 1$, the method is termed a (successive) *over–relaxation* (SOR) method; if $0 < \omega < 1$ the method is called an *under relaxation* method.

The iterator matrix is clearly

$$\mathsf{B}_\omega = \mathsf{L} + \omega^{-1}\mathsf{D}.$$

The error transfer matrix is

$$\mathsf{T}_\omega = \mathsf{I}_n - \mathsf{B}_\omega^{-1}\mathsf{A} = (\mathsf{L} + \omega^{-1}\mathsf{D})^{-1}\left((\omega^{-1} - 1)\mathsf{D} - \mathsf{U}\right).$$

Theorem (convergence)

Let $A \in \mathbb{C}^{n\times n}$ have non–zero diagonal entries. A necessary condition for convergence of the relaxation method is that $\omega \in (0, 2)$.

Proof.

Since $A \in \mathbb{C}^{n\times n}$ has non–zero diagonal entries, the relaxation method is well defined. We know that a necessary and sufficient condition for convergence is $\rho(T_\omega) < 1$. Since the eigenvalues are roots of the characteristic polynomial,

$$\chi_T(\lambda) = \mathbf{det}(T_\omega - \lambda I_n) = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i),$$

then setting $\lambda = 0$ we obtain $\chi_T(0) = \mathbf{det}(T_\omega) = \prod_{i=1}^{n} \lambda_i$. Therefore, if we have that $|\mathbf{det}(T_\omega)| \geq 1$ this means that there must be at least one eigenvalue that satisfies $|\lambda_i| \geq 1$ and the method cannot converge.

Proof, Cont.

Consequently, a necessary condition for unconditional convergence — that is convergence for any starting point — is that

$$|\det(\mathsf{T}_\omega)| < 1.$$

In the case of the relaxation method we have,

$$\begin{aligned}
\det(\mathsf{T}_\omega) &= \frac{\det((\omega^{-1}-1)\mathsf{D}-\mathsf{U})}{\det(\mathsf{L}+\omega^{-1}\mathsf{D})} \\
&= \frac{\prod_{i=1}^n(\omega^{-1}-1)d_{i,i}}{\prod_{i=1}^n \omega^{-1}d_{i,i}} \\
&= \frac{(\omega^{-1}-1)^n}{\omega^{-n}} \\
&= \omega^n(\omega^{-1}-1)^n \\
&= (1-\omega)^n.
\end{aligned}$$

If $\omega \notin (0,2)$, then $|\det(\mathsf{T}_\omega)| \geq 1$. In other words, if $\omega \notin (0,2)$ the method cannot converge unconditionally, meaning $\omega \in (0,2)$ is a necessary condition for unconditional convergence. □

Theorem (convergence)

Let $A \in \mathbb{C}^{n\times n}$ *be HPD and* $\omega \in (0, 2)$. *Then the relaxation method converges. In particular, the Gauss–Seidel method converges, since* $T_{\omega=1} = T_{\mathrm{GS}}$.

Proof.

Recall that

$$T_\omega = I_n - B_\omega^{-1} A = B_\omega^{-1}\left((\omega^{-1} - 1)D - U\right)$$

and

$$I_n - T = B_\omega^{-1} A.$$

Now suppose that $(\lambda, \mathbf{w})$ is an eigenpair of T_ω set

$$\mathbf{y} = (I_n - T)\mathbf{w} = (1 - \lambda)\mathbf{w}.$$

Proof, Cont.

The previous computation implies that $B_\omega \mathbf{y} = A\mathbf{w}$. For this reason,

$$
\begin{aligned}
(B_\omega - A)\mathbf{y} &= (B_\omega - A)B_\omega^{-1}A\mathbf{w} \\
&= (B_\omega B_\omega^{-1}A - AB_\omega^{-1}A)\mathbf{w} \\
&= (A - AB_\omega^{-1}A)\mathbf{w} \\
&= A(I_n - B_\omega^{-1}A)\mathbf{w} \\
&= AT_\omega \mathbf{w} \\
&= \lambda A\mathbf{w}.
\end{aligned}
$$

Taking inner product with $\mathbf{y}$, and using the fact that A is Hermitian, we find

$$(B_\omega \mathbf{y}, \mathbf{y})_2 = (A\mathbf{w}, \mathbf{y})_2 = (1 - \bar{\lambda})(\mathbf{w}, A\mathbf{w})_2$$

which implies, using the explicit form of B_ω that

$$(L\mathbf{y}, \mathbf{y})_2 + \omega^{-1}(D\mathbf{y}, \mathbf{y}_2) = (1 - \bar{\lambda})(\mathbf{w}, A\mathbf{w})_2. \tag{1}$$

Proof, Cont.

Similarly,

$$(\mathbf{y}, (\mathsf{B}_\omega - \mathsf{A})\mathbf{y})_2 = \bar{\lambda}(\mathbf{y}, \mathsf{A}\mathbf{w})_2 = \bar{\lambda}(1-\lambda)(\mathbf{w}, \mathsf{A}\mathbf{w})_2,$$

which implies

$$(\omega^{-1} - 1)(\mathsf{D}\mathbf{y}, \mathbf{y})_2 - (\mathbf{y}, \mathsf{U}\mathbf{y})_2 = \bar{\lambda}(1-\lambda)(\mathbf{w}, \mathsf{A}\mathbf{w}), \qquad (2)$$

Adding (1) and (2) — and observing that, since A is Hermitian $(\mathsf{L}\mathbf{y}, \mathbf{y})_2 = (\mathbf{y}, \mathsf{U}\mathbf{y})_2$ — we obtain

$$(2\omega^{-1} - 1)(\mathsf{D}\mathbf{y}, \mathbf{y})_2 = (1 - |\lambda|^2)(\mathbf{w}, \mathsf{A}\mathbf{w})_2.$$

Recall that, since A is HPD so is its diagonal D. The expression on the left is positive, provided $\omega \in (0, 2)$. This means that

$$1 - |\lambda|^2 = \frac{(2\omega^{-1} - 1)(\mathsf{D}\mathbf{y}, \mathbf{y})_2}{(\mathbf{w}, \mathsf{A}\mathbf{w})_2} > 0,$$

which implies $|\lambda| < 1$. It follows that $\rho(\mathsf{T}_\omega) < 1$. □

The Householder–John Criterion

Theorem (Householder–John criterion[1])

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is non–singular and Hermitian and* $\mathsf{B} \in \mathbb{C}^{n\times n}$ *is non–singular. Assume that*

$$\mathsf{Q} = \mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}$$

is HPD. Then the two–layer stationary linear iteration method with error transfer matrix $\mathsf{T} = \mathsf{I}_n - \mathsf{B}^{-1}\mathsf{A}$ *converges unconditionally iff* A *is HPD.*

Proof.

For this proof, we use the standard spectral theory.

Suppose that A is HPD. Let $(\lambda, \mathbf{w})$ be an arbitrary eigenpair of T. We want to show that $|\lambda| < 1$. Using the definition of T, we observe that

$$(1-\lambda)\mathsf{B}\mathbf{w} = \mathsf{A}\mathbf{w}.$$

It follows that $\lambda \neq 1$. Otherwise, A would be singular.

Proof, Cont.

Thus,

$$\mathbf{w}^{\mathrm{H}} \mathsf{B} \mathbf{w} = \frac{1}{1-\lambda} \mathbf{w}^{\mathrm{H}} \mathsf{A} \mathbf{w}.$$

Taking the conjugate transpose of this equation, we have

$$\mathbf{w}^{\mathrm{H}} \mathsf{B}^{\mathrm{H}} \mathbf{w} = \frac{1}{1-\bar{\lambda}} \mathbf{w}^{\mathrm{H}} \mathsf{A} \mathbf{w},$$

using the fact that A is Hermitian. Combining the last two equations,

$$\mathbf{w}^{\mathrm{H}} \left(\mathsf{B}^{\mathrm{H}} + \mathsf{B} - \mathsf{A} \right) \mathbf{w} = \left(\frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}} - 1 \right) \mathbf{w}^{\mathrm{H}} \mathsf{A} \mathbf{w} = \frac{1-|\lambda|^2}{|1-\lambda|^2} \mathbf{w}^{\mathrm{H}} \mathsf{A} \mathbf{w}.$$

Since

$$\mathbf{w}^{\mathrm{H}} \mathsf{A} \mathbf{w} > 0, \qquad \mathbf{w}^{\mathrm{H}} \left(\mathsf{B}^{\mathrm{H}} + \mathsf{B} - \mathsf{A} \right) \mathbf{w} > 0,$$

it must be that

$$\frac{1-|\lambda|^2}{|1-\lambda|^2} > 0.$$

Proof, Cont.

We conclude that

$$|\lambda| < 1.$$

Using our previous calculations, but only assuming that A is non singular and Hermitian, if

$$\frac{1 - |\lambda|^2}{|1 - \lambda|^2} > 0$$

and

$$\mathbf{w}^{\mathsf{H}} \left(\mathsf{B}^{\mathsf{H}} + \mathsf{B} - \mathsf{A} \right) \mathbf{w} > 0,$$

it is easy to see that

$$\mathbf{w}^{\mathsf{H}} \mathsf{A} \mathbf{w} > 0,$$

for every eigenvector $\mathbf{w} \in \mathbb{C}^n_\star$. Unfortunately, this is not enough to prove that A is HPD. There is a bit more work to do... (Homework exercise.) □

Symmetrization and Symmetric Relaxation

Symmetrized Methods

When the coefficient matrix, A, is symmetric, it is often desirable the iterator matrix, B, is as well. For the standard Gauss–Seidel method, in particular, this is not the case. However, there is a simple way to symmetrize the iterator.

Definition (symmetrized method)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ be invertible and $\mathbf{f} \in \mathbb{C}^n$. Suppose that $\mathbf{x} = \mathsf{A}^{-1}\mathbf{f}$ and consider the stationary two–layer method

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathsf{B}^{-1}(\mathbf{f} - \mathsf{A}\mathbf{x}_k) \tag{3}$$

defined by the invertible iterator matrix $\mathsf{B} \in \mathbb{C}^{n\times n}$. The **symmetrized stationary two–layer method** is defined as follows

$$\mathbf{x}_{k+\frac{1}{2}} = \mathbf{x}_k + \mathsf{B}^{-1}(\mathbf{f} - \mathsf{A}\mathbf{x}_k), \quad \mathbf{x}_{k+1} = \mathbf{x}_{k+\frac{1}{2}} + \mathsf{B}^{-\mathsf{H}}\left(\mathbf{f} - \mathsf{A}\mathbf{x}_{k+\frac{1}{2}}\right). \tag{4}$$

Lemma (standard form)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ *be invertible,* $\mathbf{f} \in \mathbb{C}^{n}$*, and* $\mathbf{x} = \mathsf{A}^{-1}\mathbf{f}$*. Suppose that* $\mathsf{B} \in \mathbb{C}^{n\times n}$ *is invertible, and consider the symmetrized stationary two–layer method* (4)*. Then*

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathsf{C}_{\mathrm{S}}\left(\mathbf{f} - \mathsf{A}\mathbf{x}_k\right),$$

where

$$\mathsf{C}_{\mathrm{S}} = \mathsf{B}^{-\mathsf{H}}\left(\mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}\right)\mathsf{B}^{-1}.$$

If A *is Hermitian, then* C_{S} *is as well.*

Definition (symmetric relaxation)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ be invertible with nonzero diagonal entries, and consider the relaxation method with iterator

$$\mathsf{B}_\omega = \mathsf{L} + \omega^{-1}\mathsf{D}, \quad \omega > 0,$$

where $\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}$ is the standard splitting of A into lower triangular, diagonal, and upper triangular parts, respectively. Note that B_ω is invertible. The **symmetric relaxation method** is the symmetrized stationary two–layer method with respect to B_ω, that is

$$\mathsf{C}_{\omega,\mathrm{S}} = \mathsf{B}_\omega^{-\mathsf{H}} \left(\mathsf{B}_\omega + \mathsf{B}_\omega^{\mathsf{H}} - \mathsf{A}\right) \mathsf{B}_\omega^{-1}.$$

When $\omega = 1$, the method is called the **symmetric Gauss–Seidel method**.

Notice that, if A is Hermitian, then $\mathsf{C}_{\omega,\mathrm{S}}$ is as well, and, in particular,

$$\mathsf{C}_{\omega,\mathrm{S}} = \mathsf{B}_\omega^{-\mathsf{H}} \left(\mathsf{L} + \omega^{-1}\mathsf{D} + \mathsf{U} + \omega^{-1}\mathsf{D} - \mathsf{A}\right) \mathsf{B}_\omega^{-1} = \mathsf{B}_\omega^{-\mathsf{H}} \left((2\omega^{-1} - 1)\mathsf{D}\right) \mathsf{B}_\omega^{-1}.$$

Theorem (convergence)

Let $\mathsf{A} \in \mathbb{C}^{n \times n}$ *be HPD and* $\omega \in (0, 2)$. *Then the symmetric relaxation method converges. In particular, the symmetric Gauss–Seidel method, obtained by setting* $\omega = 1$, *converges.*

Proof.

Since A is HPD, it has positive diagonal entries. Using this fact and the fact that $0 < \omega < 2$, it follows that $\mathsf{C}_{\omega,\mathrm{S}}$ is invertible. Then

$$\mathsf{C}_{\omega,\mathrm{S}}^{-1} (\mathbf{x}_{k+1} - \mathbf{x}_k) = \mathbf{f} - \mathsf{A}\mathbf{x}_k .$$

In other words, the iterator matrix for the symmetric relaxation method is precisely

$$\mathsf{C}_{\omega,\mathrm{S}}^{-1} = \left(\frac{2}{\omega} - 1 \right)^{-1} \mathsf{B}_\omega \mathsf{D}^{-1} \mathsf{B}_\omega^{\mathsf{H}},$$

One can show that

$$\mathsf{Q} = \mathsf{C}_{\omega,\mathrm{S}}^{-1} + \mathsf{C}_{\omega,\mathrm{S}}^{-\mathsf{H}} - \mathsf{A}$$

is HPD, since A is HPD. Applying the Householder–John criterion, we prove that the method converges. The details are left for the reader. □

Convergence in Energy Norm

Energy Methods and the Energy Norm

In this section, we provide a powerful alternative method for proving convergence, called the energy method. Recall that, if A is HPD we can define the so–called *energy norm* of the matrix by

$$\|\mathbf{x}\|_{A}^{2} = (\mathbf{x}, \mathbf{x})_{A} = (A\mathbf{x}, \mathbf{x})_{2}.$$

Lemma (positive definite)

Suppose that $\mathsf{Q} \in \mathbb{C}^{n\times n}$ *is positive definite in the sense that*

$$\Re\left((\mathsf{Q}\mathbf{y},\mathbf{y})_2\right) > 0, \quad \forall \mathbf{y} \in \mathbb{C}^n_\star,$$

but Q *is not necessarily Hermitian. Then*

$$\|\mathbf{w}\|_{\mathsf{Q}} = \sqrt{\Re\left((\mathsf{Q}\mathbf{w},\mathbf{w})_2\right)}, \quad \forall \mathbf{w} \in \mathbb{C}^n,$$

defines a norm.

Proof.

Suppose that Q is not Hermitian, to avoid the simple case. Then

$$\mathsf{Q} = \mathsf{Q}_H + \mathsf{Q}_A,$$

where

$$\mathsf{Q}_H = \frac{1}{2}\left(\mathsf{Q} + \mathsf{Q}^{\mathsf{H}}\right), \qquad \mathsf{Q}_A = \frac{1}{2}\left(\mathsf{Q} - \mathsf{Q}^{\mathsf{H}}\right),$$

are the Hermitian and anti-Hermitian parts, respectively. Observe that $\mathsf{Q}_H^{\mathsf{H}} = \mathsf{Q}_H$ and $\mathsf{Q}_A^{\mathsf{H}} = -\mathsf{Q}_A$.

Proof, Cont.

It follows that $(Q_A\mathbf{y},\mathbf{y})_2$ is purely imaginary for any $\mathbf{y}\in\mathbb{C}^n$, because

$$\overline{(Q_A\mathbf{y},\mathbf{y})_2} = \overline{\mathbf{y}^{\mathsf{H}}Q_A\mathbf{y}} = \mathbf{y}^{\mathsf{H}}Q_A^{\mathsf{H}}\mathbf{y} = -\mathbf{y}^{\mathsf{H}}Q_A\mathbf{y} = -(Q_A\mathbf{y},\mathbf{y})_2.$$

Therefore, for all $\mathbf{y}\in\mathbb{C}^n_\star$,

$$0 < \Re\left((Q\mathbf{y},\mathbf{y})_2\right) = \Re\left((Q_H\mathbf{y},\mathbf{y})_2\right) + \Re\left((Q_A\mathbf{y},\mathbf{y})_2\right) = (Q_H\mathbf{y},\mathbf{y})_2\,,$$

since $(Q_H\mathbf{y},\mathbf{y})_2$ is real. Therefore, Q_H is HPD. Since

$$\|\mathbf{w}\|_{Q_H} = \sqrt{(Q_H\mathbf{w},\mathbf{w})_2}, \quad \text{for all} \quad \mathbf{w}\in\mathbb{C}^n,$$

defines a norm, the result follows. □

Theorem (convergence in energy)

Let A *be HPD. Suppose that* $\mathsf{B} \in \mathbb{C}^{n \times n}$ *is the invertible iterator describing a two–layer stationary linear iteration method. If* $\mathsf{Q} = \mathsf{B} - \frac{1}{2}\mathsf{A}$ *is positive definite in the sense that*

$$\Re\left((\mathsf{Q}\mathbf{y}, \mathbf{y})_2\right) > 0, \quad \forall\, \mathbf{y} \in \mathbb{C}^n_\star,$$

but is not necessarily Hermitian, then method (3) *converges.*

Proof.

Define the error $\mathbf{e}_k = \mathbf{x} - \mathbf{x}_k$, as usual, and notice that for any stationary two–layer linear iteration method we have

$$\mathsf{B}\left(\mathbf{e}_{k+1} - \mathbf{e}_k\right) + \mathsf{A}\mathbf{e}_k = \mathbf{0},$$

which is equivalent to

$$\mathsf{B}\mathbf{q}_{k+1} + \mathsf{A}\mathbf{e}_k = \mathbf{0},$$

where $\mathbf{q}_{k+1} = \mathbf{e}_{k+1} - \mathbf{e}_k$.

Proof, Cont.

Taking the inner product of this identity with $\mathbf{q}_{k+1}$ and using the fact that

$$\mathbf{e}_k = \tfrac{1}{2}(\mathbf{e}_{k+1} + \mathbf{e}_k) - \tfrac{1}{2}(\mathbf{e}_{k+1} - \mathbf{e}_k)$$

we obtain

$$\begin{aligned}
0 &= (\mathsf{B}\mathbf{q}_{k+1}, \mathbf{q}_{k+1})_2 + (\mathsf{A}\mathbf{e}_k, \mathbf{q}_{k+1})_2 \\
&= \left(\left(\mathsf{B} - \frac{1}{2}\mathsf{A}\right)\mathbf{q}_{k+1}, \mathbf{q}_{k+1}\right)_2 + \frac{1}{2}(\mathsf{A}\mathbf{e}_{k+1}, \mathbf{e}_{k+1})_2 - \frac{1}{2}(\mathsf{A}\mathbf{e}_k, \mathbf{e}_k)_2 \\
&\quad + \mathrm{i}\Im\left((\mathsf{A}\mathbf{e}_k, \mathbf{e}_{k+1})_2\right) \\
&= (\mathsf{Q}\mathbf{q}_{k+1}, \mathbf{q}_{k+1})_2 + \frac{1}{2}\|\mathbf{e}_{k+1}\|_{\mathsf{A}}^2 - \frac{1}{2}\|\mathbf{e}_k\|_{\mathsf{A}}^2 + \mathrm{i}\Im\left((\mathsf{A}\mathbf{e}_k, \mathbf{e}_{k+1})_2\right).
\end{aligned}$$

Consequently,

$$\Re\left((\mathsf{Q}\mathbf{q}_{k+1}, \mathbf{q}_{k+1})_2\right) + \frac{1}{2}\|\mathbf{e}_{k+1}\|_{\mathsf{A}}^2 = \frac{1}{2}\|\mathbf{e}_k\|_{\mathsf{A}}^2 \implies \frac{1}{2}\|\mathbf{e}_{k+1}\|_{\mathsf{A}}^2 + \|\mathbf{q}_{k+1}\|_{\mathsf{Q}}^2 = \frac{1}{2}\|\mathbf{e}_k\|_{\mathsf{A}}^2.$$

Proof, Cont.

This allows us to conclude that the sequence $\|\mathbf{e}_k\|_{\mathsf{A}}$ is non increasing. Since it is bounded below, by the monotone convergence theorem, it must have a limit, $\lim_{k\to\infty}\|\mathbf{e}_k\|_{\mathsf{A}} = \alpha$, say. Passing to the limit in this identity then tells us that

$$\|\mathbf{q}_{k+1}\|_{\mathsf{Q}} \to 0, \quad \text{as } k \to \infty.$$

But, since A is invertible,

$$\mathbf{e}_k = -\mathsf{A}^{-1}\mathsf{B}\mathbf{q}_{k+1}$$

and this implies $\mathbf{e}_k \to \mathbf{0}$. Therefore, the method converges. □

Convergence of Richardson's Method via Energy

For Richardson's method we have $\mathsf{Q} = \frac{1}{\alpha}\mathsf{I}_n - \frac{1}{2}\mathsf{A}$, $\alpha > 0$, and

$$(\mathsf{Q}\mathbf{x}, \mathbf{x})_2 = \frac{1}{\alpha}\|\mathbf{x}\|_2^2 - \frac{1}{2}(\mathsf{A}\mathbf{x}, \mathbf{x})_2 \geq \|\mathbf{x}\|_2^2 \left(\frac{1}{\alpha} - \frac{\|\mathsf{A}\|_2}{2} \right).$$

A sufficient condition for Q to be HPD, and, therefore, for the method to converge, is that

$$0 < \alpha < \frac{2}{\|\mathsf{A}\|_2}.$$

But, since A is HPD, $\|\mathsf{A}\|_2 = \lambda_n$, the largest eigenvalue. If $0 < \alpha\lambda_n < 2$, Richardson's method converges via the energy method.

Convergence of the Relaxation Method via Energy

In the relaxation method, the iterator is

$$\mathsf{B}_\omega = \mathsf{L} + \frac{1}{\omega}\mathsf{D},$$

where $\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}$ is the standard matrix splitting and $\omega > 0$. Therefore

$$\mathsf{Q}_\omega = \mathsf{B}_\omega - \frac{1}{2}\mathsf{A} = \frac{1}{\omega}\mathsf{D} + \mathsf{L} - \frac{1}{2}(\mathsf{L} + \mathsf{D} + \mathsf{L}^{\mathsf{H}}) = \left(\frac{1}{\omega} - \frac{1}{2}\right)\mathsf{D} + \frac{1}{2}\left(\mathsf{L} - \mathsf{L}^{\mathsf{H}}\right),$$

and

$$\begin{aligned}\Re\left((\mathsf{Q}_\omega \mathbf{x}, \mathbf{x})_2\right) &= \left(\frac{1}{\omega} - \frac{1}{2}\right)(\mathsf{D}\mathbf{x}, \mathbf{x})_2 + \Re\left(\frac{1}{2}((\mathsf{L} - \mathsf{L}^{\mathsf{H}})\mathbf{x}, \mathbf{x})_2\right) \\ &= \left(\frac{1}{\omega} - \frac{1}{2}\right)(\mathsf{D}\mathbf{x}, \mathbf{x})_2.,\end{aligned}$$

where we used the fact that $\frac{1}{2}((\mathsf{L} - \mathsf{L}^{\mathsf{H}})$ is anti–Hermitian. If $\frac{1}{\omega} - \frac{1}{2} > 0$ or, equivalently, $\omega < 2$, Q_ω is positive definite — though not Hermitian — and the relaxation method converges.

Theorem (Householder–John)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is non singular and Hermitian and* $\mathsf{B} \in \mathbb{C}^{n\times n}$ *is non singular. Assume that*

$$\mathsf{Q} = \mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}$$

is HPD. Then the two–layer stationary linear iteration method with error transfer matrix $\mathsf{T} = \mathsf{I}_n - \mathsf{B}^{-1}\mathsf{A}$ *converges unconditionally iff* A *is HPD.*

Proof.

Let us prove one direction and save the other for a homework exercise.

Suppose that A is HPD. Recall that $\|\cdot\|_{\mathsf{A}}$ defines a norm on $\mathbb{C}^n$. The error equation is precisely

$$\mathbf{e}_{k+1} = \left(\mathsf{I} - \mathsf{B}^{-1}\mathsf{A}\right)\mathbf{e}_k .$$

Proof, Cont.

Then

$$
\begin{aligned}
\|\mathbf{e}_{k+1}\|_{\mathsf{A}}^2 &= \left(\left(\mathsf{I} - \mathsf{B}^{-1}\mathsf{A}\right)\mathbf{e}_k\right)^{\mathsf{H}} \mathsf{A}\left(\left(\mathsf{I} - \mathsf{B}^{-1}\mathsf{A}\right)\mathbf{e}_k\right) \\
&= \left({\mathbf{e}_k}^{\mathsf{H}}\left(\mathsf{I} - \mathsf{A}\mathsf{B}^{-\mathsf{H}}\right)\right)\mathsf{A}\left(\left(\mathsf{I} - \mathsf{B}^{-1}\mathsf{A}\right)\mathbf{e}_k\right) \\
&= \left({\mathbf{e}_k}^{\mathsf{H}} - {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathsf{B}^{-\mathsf{H}}\right)\left(\mathsf{A}\mathbf{e}_k - \mathsf{A}\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k\right) \\
&= {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathbf{e}_k - {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathsf{B}^{-\mathsf{H}}\mathsf{A}\mathbf{e}_k - {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k + {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathsf{B}^{-\mathsf{H}}\mathsf{A}\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k \\
&= \|\mathbf{e}_k\|_{\mathsf{A}}^2 - {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\left(\mathsf{B}^{-\mathsf{H}} + \mathsf{B}^{-1} - \mathsf{B}^{-\mathsf{H}}\mathsf{A}\mathsf{B}^{-1}\right)\mathsf{A}\mathbf{e}_k \\
&= \|\mathbf{e}_k\|_{\mathsf{A}}^2 - {\mathbf{e}_k}^{\mathsf{H}}\mathsf{A}\mathsf{B}^{-\mathsf{H}}\left(\mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}\right)\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k \\
&= \|\mathbf{e}_k\|_{\mathsf{A}}^2 - \left(\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k\right)^{\mathsf{H}}\left(\mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}\right)\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k.
\end{aligned}
$$

Since $\mathsf{B} + \mathsf{B}^{\mathsf{H}} - \mathsf{A}$ is HPD and $\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k \neq \mathbf{0}$, in general, it follows that

$$\|\mathbf{e}_{k+1}\|_{\mathsf{A}}^2 + \left\|\mathsf{B}^{-1}\mathsf{A}\mathbf{e}_k\right\|_{\mathsf{Q}}^2 = \|\mathbf{e}_k\|_{\mathsf{A}}^2,$$

and $\|\mathbf{e}_k\|_{\mathsf{A}}$ is a decreasing sequence.

Proof, Cont.

Therefore, by the monotone convergence theorem, $\|\mathbf{e}_k\|_A$ converges, that is, there is some $\alpha \in [0, \infty)$, such that

$$\lim_{k\to\infty} \|\mathbf{e}_k\|_A = \alpha = \lim_{k\to\infty} \|\mathbf{e}_{k+1}\|_A .$$

This implies

$$\lim_{k\to\infty} \left\|B^{-1}A\mathbf{e}_k\right\|_Q = 0,$$

which, in turn, implies that $\mathbf{e}_k \to \mathbf{0}$ as $k \to \infty$. $\square$

Convergence of the Relaxation Method, Re-revisited

Let us provide yet another proof of convergence of the relaxation method. Suppose $\mathsf{A} \in \mathbb{C}^{n\times n}$ is HPD. We once again recall that the iterator matrix for the relaxation method is

$$\mathsf{B}_\omega = \mathsf{L} + \frac{1}{\omega}\mathsf{D},$$

where $\mathsf{A} = \mathsf{L} + \mathsf{D} + \mathsf{U}$ is the standard matrix splitting and $\omega > 0$. Therefore

$$\mathsf{Q}_\omega = \mathsf{B}_\omega + \mathsf{B}_\omega^{\mathsf{H}} - \mathsf{A} = \frac{2}{\omega}\mathsf{D} + \mathsf{L} + \mathsf{L}^{\mathsf{H}} - (\mathsf{L} + \mathsf{D} + \mathsf{L}^{\mathsf{H}}) = \left(\frac{2}{\omega} - 1\right)\mathsf{D}.$$

If $0 < \omega < 2$, Q_ω is HPD and the method converges.