Lecture 6: Stationary iterative methods

School of Mathematical Sciences, Xiamen University

1. Splitting and stationary iterative method

Definition 1

A *splitting* of $\mathbf{A} \in \mathbb{C}^{n\times n}$ is a decomposition $\mathbf{A} = \mathbf{M} - \mathbf{K}$, with $\mathbf{M}$ nonsingular.

Remark 2

A splitting yields a stationary iterative method as follows. The equation

$$\mathbf{A}\mathbf{x} = (\mathbf{M} - \mathbf{K})\mathbf{x} = \mathbf{b}$$

implies

$$\mathbf{x} = \mathbf{M}^{-1}\mathbf{K}\mathbf{x} + \mathbf{M}^{-1}\mathbf{b} := \mathbf{R}\mathbf{x} + \mathbf{c}.$$

Given a starting vector $\mathbf{x}^{(0)}$*, we obtain a stationary iterative method*

$$\mathbf{x}^{(m)} = \mathbf{R}\mathbf{x}^{(m-1)} + \mathbf{c}, \quad m = 1, 2, \ldots$$

We note that $\mathbf{R} = \mathbf{M}^{-1}\mathbf{K} = \mathbf{M}^{-1}(\mathbf{M} - \mathbf{A}) = \mathbf{I} - \mathbf{M}^{-1}\mathbf{A}$.

2. Convergence criterion

Definition 3

The spectral radius of a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is $\rho(\mathbf{A}) = \max\limits_{\lambda \in \Lambda(\mathbf{A})} |\lambda|$.

Exercise. If $\mathbf{A}$ is singular and $\mathbf{A} = \mathbf{M} - \mathbf{K}$ with $\mathbf{M}$ nonsingular, then $\rho(\mathbf{M}^{-1}\mathbf{K}) \geq 1$.

Proposition 4

Let $\|\cdot\|$ denote a matrix norm on $\mathbb{C}^{n\times n}$ induced by a vector norm on $\mathbb{C}^n$. For any $\mathbf{A} \in \mathbb{C}^{n\times n}$, we have $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$.

Lemma 5

For any given $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\varepsilon > 0$ there exists an induced matrix norm $\|\cdot\|_\star$ such that

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|_\star \leq \rho(\mathbf{A}) + \varepsilon.$$

The norm $\|\cdot\|_\star$ depends on both $\mathbf{A}$ and ε.

Proof.

Let $\mathbf{A} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}$ be a Jordan form of $\mathbf{A}$. Let

$$\mathbf{D}_\varepsilon = \mathrm{diag}\{1, \varepsilon, \varepsilon^2, \cdots, \varepsilon^{n-1}\}.$$

Now for all $\mathbf{x} \in \mathbb{C}^n$ and for all $\mathbf{B} \in \mathbb{C}^{n\times n}$, define the vector norm

$$\|\mathbf{x}\|_\star := \|(\mathbf{S}\mathbf{D}_\varepsilon)^{-1}\mathbf{x}\|_\infty$$

and the corresponding induced matrix norm

$$\begin{aligned}
\|\mathbf{B}\|_\star : &= \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{B}\mathbf{x}\|_\star}{\|\mathbf{x}\|_\star} = \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq\mathbf{0}} \frac{\|(\mathbf{S}\mathbf{D}_\varepsilon)^{-1}\mathbf{B}\mathbf{x}\|_\infty}{\|(\mathbf{S}\mathbf{D}_\varepsilon)^{-1}\mathbf{x}\|_\infty} \\
&= \sup_{\mathbf{y}\in\mathbb{C}^n, \mathbf{y}\neq\mathbf{0}} \frac{\|(\mathbf{S}\mathbf{D}_\varepsilon)^{-1}\mathbf{B}(\mathbf{S}\mathbf{D}_\varepsilon)\mathbf{y}\|_\infty}{\|\mathbf{y}\|_\infty} \\
&= \|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{B}\mathbf{S}\mathbf{D}_\varepsilon\|_\infty.
\end{aligned}$$

The statement follows from $\|\mathbf{A}\|_\star = \|\mathbf{D}_\varepsilon^{-1}\mathbf{J}\mathbf{D}_\varepsilon\|_\infty \leq \rho(\mathbf{A}) + \varepsilon$. □

Theorem 6

The iteration $\mathbf{x}^{(m)} = \mathbf{R}\mathbf{x}^{(m-1)} + \mathbf{c}$ *converges to the solution of* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *for all starting vectors* $\mathbf{x}^{(0)}$ *if and only if* $\rho(\mathbf{R}) < 1$.

Proof.

For all $\mathbf{x}^{(0)}$, *we have* $\mathbf{x}^{(m)} - \mathbf{x} = \mathbf{R}(\mathbf{x}^{(m-1)} - \mathbf{x}) = \cdots = \mathbf{R}^m(\mathbf{x}^{(0)} - \mathbf{x})$.

If $\rho(\mathbf{R}) \geq 1$, *choose* $\mathbf{x}^{(0)} - \mathbf{x}$ *to be an eigenvector of* $\mathbf{R}$ *with eigenvalue* λ *where* $|\lambda| = \rho(\mathbf{R})$. *Then* $\mathbf{x}^{(m)} - \mathbf{x} = \lambda^m(\mathbf{x}^{(0)} - \mathbf{x})$ *will not approach* $\mathbf{0}$.

If $\rho(\mathbf{R}) < 1$, *by Lemma 5 there exists an induced matrix norm* $\|\cdot\|_\star$ *such that* $\|\mathbf{R}\|_\star < 1$, *then we have* $\|\mathbf{x}^{(m)} - \mathbf{x}\|_\star \leq \|\mathbf{R}\|_\star^m \|\mathbf{x}^{(0)} - \mathbf{x}\|_\star \to 0$ *for all* $\mathbf{x}^{(0)}$. □

Remark 7

The goal is to choose a splitting $\mathbf{A} = \mathbf{M} - \mathbf{K}$ *so that both*

(1) $\mathbf{R}\mathbf{v} = \mathbf{M}^{-1}\mathbf{K}\mathbf{v}$ *and* $\mathbf{c} = \mathbf{M}^{-1}\mathbf{b}$ *are easy to evaluate, and*

(2) $\rho(\mathbf{R})$ *is small* (< 1).

- (1) and (2) are conflicting goals, and need to be balanced.
- If $\Lambda(\mathbf{R}) \subset (-\rho(\mathbf{R}), \rho(\mathbf{R}))$, then Chebyshev acceleration technique can be used. See Demmel's book ANLA, section 6.5.6.

3. Classical stationary iterative methods

- Let $\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}$, where

  $\mathbf{D}$ is the diagonal matrix with diagonal entries $d_{ii} = a_{ii}$,

  $-\mathbf{L}$ is the strictly lower triangular part of $\mathbf{A}$,

  $-\mathbf{U}$ is the strictly upper triangular part of $\mathbf{A}$.

- Assume that $\mathbf{A}$ has no zero diagonal entries. We will introduce

  (1) Jacobi's method,

  (2) Gauss–Seidel method,

  (3) Successive overrelaxation: SOR(ω),

  (4) Symmetric successive overrelaxation: SSOR(ω).

3.1. Jacobi's method

- The splitting is

$$\mathbf{A} = \mathbf{D} - (\mathbf{L} + \mathbf{U}),$$

and the corresponding

$$\mathbf{R} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}) = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A} \quad \text{and} \quad \mathbf{c} = \mathbf{D}^{-1}\mathbf{b}.$$

Algorithm 1: Jacobi's method

for $j = 1$ **to** n

$$x_j^{(m+1)} = \frac{1}{a_{jj}} \left(b_j - \sum_{k \neq j} a_{jk} x_k^{(m)} \right)$$

end

Theorem 8

If $\mathbf{A}$ is Hermitian and $a_{ii} > 0$ for all i, then Jacobi's method converges for all starting vectors if and only if $\mathbf{A}$ and $2\mathbf{D} - \mathbf{A}$ are both HPD.

3.2. Gauss–Seidel method

- The splitting is

$$\mathbf{A} = (\mathbf{D} - \mathbf{L}) - \mathbf{U},$$

and the corresponding

$$\mathbf{R} = (\mathbf{D} - \mathbf{L})^{-1}\mathbf{U},$$

and

$$\mathbf{c} = (\mathbf{D} - \mathbf{L})^{-1}\mathbf{b}.$$

Algorithm 2: Gauss–Seidel method

---

for $j = 1$ **to** n

$$x_j^{(m+1)} = \frac{1}{a_{jj}} \left(b_j - \sum_{k=1}^{j-1} a_{jk} x_k^{(m+1)} - \sum_{k=j+1}^{n} a_{jk} x_k^{(m)} \right)$$

end

3.3. Successive overrelaxation: SOR(ω), $\omega \in \mathbb{R}$

- The splitting is $\omega\mathbf{A} = (\mathbf{D} - \omega\mathbf{L}) - ((1-\omega)\mathbf{D} + \omega\mathbf{U})$, and the corresponding
$$\mathbf{R} = (\mathbf{D} - \omega\mathbf{L})^{-1}((1-\omega)\mathbf{D} + \omega\mathbf{U}),$$
and
$$\mathbf{c} = \omega(\mathbf{D} - \omega\mathbf{L})^{-1}\mathbf{b}.$$
- $\omega = 1$: Gauss–Seidel method
- $0 < \omega < 2$: Necessary in some sense (see Theorem 13)
- Optimal ω:

Algorithm 3: SOR(ω), here ω is the relaxation parameter

---

for $j = 1$ **to** n

$$x_j^{(m+1)} = (1-\omega)x_j^{(m)} + \frac{\omega}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{(m+1)} - \sum_{k=j+1}^{n} a_{jk}x_k^{(m)}\right)$$

end

3.4. Symmetric successive overrelaxation: SSOR(ω), $\omega \in \mathbb{R}$

- This method uses two splittings:

$$
\begin{aligned}
\omega \mathbf{A} &= (\mathbf{D} - \omega \mathbf{L}) - ((1-\omega)\mathbf{D} + \omega \mathbf{U}) \\
&= (\mathbf{D} - \omega \mathbf{U}) - ((1-\omega)\mathbf{D} + \omega \mathbf{L}),
\end{aligned}
$$

and the corresponding

$$
\mathbf{R} = (\mathbf{D} - \omega \mathbf{U})^{-1}((1-\omega)\mathbf{D} + \omega \mathbf{L})(\mathbf{D} - \omega \mathbf{L})^{-1}((1-\omega)\mathbf{D} + \omega \mathbf{U}),
$$

$$
\mathbf{c} = \omega(2-\omega)(\mathbf{D} - \omega \mathbf{U})^{-1}\mathbf{D}(\mathbf{D} - \omega \mathbf{L})^{-1}\mathbf{b}.
$$

Algorithm 4: SSOR(ω)

for $j = 1$ **to** n

$$
x_j^{(m+1/2)} = (1-\omega)x_j^{(m)} + \frac{\omega}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{(m+1/2)} - \sum_{k=j+1}^{n} a_{jk}x_k^{(m)}\right)
$$

end

for $j = n$ **to** 1

$$
x_j^{(m+1)} = (1-\omega)x_j^{(m+1/2)} + \frac{\omega}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{(m+1/2)} - \sum_{k=j+1}^{n} a_{jk}x_k^{(m+1)}\right)
$$

end

3.5. Convergence (see Demmel's book ANLA, section 6.5.5)

Definition 9

$\mathbf{A}$ is an irreducible matrix if there is no permutation matrix such that

$$\mathbf{P}\mathbf{A}\mathbf{P}^\top = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}.$$

Definition 10

A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is weakly row diagonally dominant if for all i,

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

with strict inequality at least once. A matrix $\mathbf{A}$ is strictly row diagonally dominant if for all i:

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

Theorem 11

If $\mathbf{A}$ *is strictly row diagonally dominant, then both Jacobi's and Gauss–Seidel methods converge. In fact* $\|\mathbf{R}_{\mathrm{GS}}\|_\infty \le \|\mathbf{R}_{\mathrm{J}}\|_\infty < 1$.

Theorem 12

If $\mathbf{A}$ *is irreducible and weakly row diagonally dominant, then both Jacobi's and Gauss–Seidel methods converge, and* $\rho(\mathbf{R}_{\mathrm{GS}}) < \rho(\mathbf{R}_{\mathrm{J}}) < 1$.

Theorem 13

For any matrix $\mathbf{A}$*, it holds* $\rho(\mathbf{R}_{\mathrm{SOR}(\omega)}) \ge |\omega - 1|$. *Therefore* $0 < \omega < 2$ *is required for the convergence of* $\mathrm{SOR}(\omega)$ *for all starting vectors.*

Theorem 14

If $\mathbf{A}$ *is Hermitian positive definite, then* $\rho(\mathbf{R}_{\mathrm{SOR}(\omega)}) < 1$ *for all* $0 < \omega < 2$*, i.e.,* $\mathrm{SOR}(\omega)$ *converges for all* $0 < \omega < 2$. *Gauss–Seidel* $(\mathrm{SOR}(1))$ *converges for Hermitian positive definite* $\mathbf{A}$.