Lecture 15: Recall:

Definition: Consider the system $A\vec{x} = \vec{b}$. Let $A = \underset{\text{lower}}{L} + \underset{\text{diagonal}}{D} + \underset{\text{upper}}{U}$.

If the eigenvalues of: $\alpha D^{-1}L + \frac{1}{\alpha} D^{-1}U$ $(\alpha \neq 0)$ are independent of α. Then, the matrix A is said to be <u>consistently ordered</u>.

Example of consistently ordered matrices

1. Tridiagonal matrix:

$$\begin{pmatrix} \lambda_1 & * & & 0 \\ * & \lambda_2 & * & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & * \\ 0 & & & * & \lambda_n \end{pmatrix}$$

2. Block tridiagonal matrix:

$$\begin{pmatrix} \boxed{D_1} & T_{12} & & 0 \\ T_{21} & \boxed{D_2} & T_{23} & \\ & \ddots & \ddots & \ddots \\ 0 & & & \boxed{D_p} \end{pmatrix}$$

where D_i = diagonal matrix

Theorem: [D. Young] Assume:

1. $0 < \omega < 2$
2. $M_J = N_J^{-1} P_J$ has only real eigenvalues
3. $\beta \stackrel{def}{=} \rho(M_J) < 1$
4. A is consistently ordered.

Then: $\rho(M_{SOR}) < 1$

Also, the optimal parameter ω_{opt} for fastest convergence

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \beta^2}}$$ and

$$\rho(M_{SOR,\omega_{opt}}) = \omega_{opt} - 1$$

Important example:

Consider an $n \times n$ tridiagonal matrix of the form:

$$A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} \in M_{n\times n}(\mathbb{R})$$

We want to study the convergence of Jacobi and SOR methods to solve $A\vec{x} = \vec{b}$.

In fact, for any $n \times n$ matrix of the form:

$$T_\alpha = \begin{pmatrix} \alpha & -1 & & & \\ -1 & \alpha & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & \alpha \end{pmatrix} \in M_{n\times n}(\mathbb{R})$$

, its eigenvectors are given by:

$$\vec{q}_j = \begin{pmatrix} \sin(j\theta) \\ \sin(2j\theta) \\ \vdots \\ \sin(nj\theta) \end{pmatrix}$$

w/ eigenvalue

$$\lambda_j = \alpha - 2\cos(j\theta)$$

$$\left(j = 1, 2, \ldots, n, \quad \theta = \frac{\pi}{n+1}\right)$$

Proof:

$$T_\alpha \vec{q}_j = \begin{pmatrix} \alpha \sin j\theta - \sin(2j\theta) \\ \alpha \sin(2j\theta) - \sin(j\theta) - \sin(3j\theta) \\ \alpha \sin(3j\theta) - \sin(2j\theta) - \sin(4j\theta) \\ \vdots \\ \alpha \sin(nj\theta) - \sin((n-1)j\theta) \end{pmatrix}$$

Using trigonometric formula: $\sin(a+b) + \sin(a-b) = 2\sin a \cos b$

$$\therefore \quad T_\alpha \vec{q}_j = \begin{pmatrix} \alpha \sin j\theta - 2\sin j\theta \cos j\theta \\ \alpha \sin(2j\theta) - 2\cos j\theta \sin 2j\theta \\ \vdots \\ \alpha \sin(nj\theta) - 2\cos(j\theta)\sin(nj\theta) \end{pmatrix} = \underbrace{(\alpha - 2\cos(j\theta))}_{\lambda_j} \vec{q}_j$$

Note that A is tridiagonal. Hence, it is consistently ordered.
(Condition 4 of Young's Theorem is satisfied)

Now, $M_J = -D^{-1}L - D^{-1}U = \begin{pmatrix} 0 & \frac{1}{2} & & 0 \\ \frac{1}{2} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{1}{2} \\ 0 & & \frac{1}{2} & 0 \end{pmatrix} = \frac{-1}{2} T_1$

$\therefore$ Eigenvalues of M_J are given by:

$$\lambda_j = \cos\left(j\frac{\pi}{n+1}\right) \quad \text{where } j = 1, 2, \cdots, n$$

$\therefore$ All eigenvalues of M_J are real.
(Condition 2 of Young's Theorem is satisfied)

Also, $\rho(M_J) = \cos\left(\frac{\pi}{n+1}\right) < 1$
(Condition 3 of Young's Theorem is satisfied)

$\therefore$ the optimal ω is:

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(M_J)^2}} = \frac{2}{1 + \sin\left(\frac{\pi}{n+1}\right)} \quad //$$

Another condition for the convergence (useful for analyzing SOR)

Let $A = N - P$ (N is invertible)

Iterative scheme : $N\vec{x}^{k+1} = P\vec{x}^{k} + \vec{b}$

Theorem: (Householder-John) Suppose A and $(N^* + N - A)$ are self-adjoint positive-definite matrices, then the iterative scheme converges.

Proof: Consider $M = N^{-1}P = N^{-1}(N-A) = I - N^{-1}A$.

Suffice to show that all eigenvalues of M satisfy $|\lambda| < 1$ (λ can be complex). Let λ be an eigenvalue associated to the eigenvector $\vec{x}$. Then:

$$M\vec{x} = \lambda\vec{x} \Rightarrow (I - N^{-1}A)\vec{x} = \lambda\vec{x} \Rightarrow (N-A)\vec{x} = \lambda N\vec{x}$$
$$\Rightarrow (1-\lambda)N\vec{x} = A\vec{x}$$

Note that $\lambda \neq 1$. Otherwise 0 is an eigenvalue of A. Contradiction to the fact that A is positive definite.

Multiply $\vec{x}^*$ on both sides:

$$(1-\lambda)\vec{x}^*N\vec{x} = \vec{x}^*A\vec{x} \Rightarrow \vec{x}^*N\vec{x} = \frac{1}{(1-\lambda)}\vec{x}^*A\vec{x} \quad \text{— (1)}$$

Take conjugate transpose on both sides:

$$(1-\bar{\lambda})\,\vec{x}^{*} N^{*}\vec{x} = \vec{x}^{*} A^{*} \vec{x} = \vec{x}^{*} A \vec{x}$$

$$\Rightarrow \quad \vec{x}^{*} N^{*}\vec{x} = \frac{1}{(1-\bar{\lambda})}\,\vec{x}^{*} A \vec{x} \quad \text{— (2)}$$

(1) + (2) − $\vec{x}^{*} A \vec{x}$ on both sides:

$$\vec{x}^{*}(N+N^{*}-A)\,\vec{x} = \left(\frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}} - 1\right)\vec{x}^{*} A \vec{x}$$

$$= \frac{1-|\lambda|^{2}}{|1-\lambda|^{2}}\,\vec{x}^{*} A \vec{x}$$

By assumption, A and $N+N^{*}-A$ are both positive definite.

We have: $\vec{x}^{*} A \vec{x} > 0$ and $\vec{x}^{*}(N+N^{*}-A)\,\vec{x} > 0$

Hence, $1-|\lambda|^{2} > 0$ and $|\lambda| < 1$.

$\therefore$ $\rho(M) < 1$ and the iterative scheme converges.

<u>Example:</u> Let $A = \begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix}$ be real symmetric tridiagonal matrix. Suppose A is positive-definite. Prove that the Gauss-Seidel method converges.

<u>Solution:</u> For Gauss-Seidel method,

$$N = \begin{pmatrix} \alpha_1 & & & 0 \\ \beta_1 & \alpha_2 & & \\ & \ddots & \ddots & \\ 0 & & \beta_{n-1} & \alpha_n \end{pmatrix} \quad \text{and} \quad N^* + N - A = \begin{pmatrix} \alpha_1 & & & & 0 \\ & \alpha_2 & & & \\ & & \ddots & & \\ & & & \alpha_{n-1} & \\ 0 & & & & \alpha_n \end{pmatrix}$$

Then, $N^* + N - A$ is symmetric.

Also, $\underbrace{(0, \dots, \underset{\substack{\uparrow \\ i\text{th}}}{1}, 0 \cdots 0)}_{\vec{e}_i^{\,T}} A \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\vec{e}_i} \leftarrow i\text{th} = \alpha_i > 0$. $\therefore$ $N + N^* - A$ is positive definite

($\because$ all eigenvalues are positive)

By Householder-John Theorem, Gauss-Seidel method converges.

<u>Example</u>: Suppose A is (real) self-adjoint positive-definite. Using Householder-John theorem, prove that: SOR method converges if and only if $0 < \omega < 2$.

<u>Solution</u>: Note that $N_{SOR} = L + \frac{1}{\omega} D$. Now, $L = U^*$

(A is self-adjoint)

$$\therefore \quad N_{SOR} + N_{SOR}^* - A = \left(\frac{2}{\omega} - 1\right) D$$

$A =$

$\therefore$ $N_{SOR} + N_{SOR}^* - A$ is also self-adjoint positive-definite if $0 < \omega < 2$.

$\therefore$ By Householder-John theorem, SOR converges.

Eigenvalue Problem

Recall: Convergence of iterative scheme:

$N\vec{x}^{k+1} = P\vec{x}^{k} + \vec{f}$ depends on the Spectral radius $\rho(N^{-1}P)$.

$\therefore$ Need: numerical method to compute eigenvalues.

Computation of Spectral radius

Goal: Find eigenvalues with largest magnitude $\leftarrow$ Spectral radius

Two methods:

1. Power method
2. QR method

1. <u>Power method</u>

Let $A \in M_{n\times n}(\mathbb{C})$ with n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with eigenvectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$.

Let $X = \begin{pmatrix} | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \\ | & | & & | \end{pmatrix} \in M_{n\times n}(\mathbb{C})$. Then, we know:

$$AX = X \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Assuming : $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$

We'll use Power method to compute $|\lambda_1|$.

<u>Observation:</u> Start with an initial vector $\vec{x}^{(0)}$.

Consider the iterative scheme: $\vec{x}^{(k+1)} = \dfrac{A\vec{x}^{(k)}}{\|A\vec{x}^{(k)}\|_\infty}$ for $k=0,1,\ldots$

Suppose A is diagonalizable. That's, we can assume $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ form a basis for $\mathbb{C}^n$.

Take $\vec{x}^{(0)} = a_1\vec{x}_1 + a_2\vec{x}_2 + \cdots + a_n\vec{x}_n$ (assuming $a_1 \neq 0$)

Note that $A^k\vec{x}^{(0)} = a_1\lambda_1^k\left[\vec{x}_1 + \displaystyle\sum_{j=2}^{n}\frac{a_j}{a_1}\left(\frac{\lambda_j}{\lambda_1}\right)^k\vec{x}_j\right]$.

Hence, $\vec{x}^{(k)} = \dfrac{A\vec{x}^{(k-1)}}{\|A\vec{x}^{(k-1)}\|_\infty} = \dfrac{A\left(A\vec{x}^{(k-2)}/\|A\vec{x}^{(k-2)}\|_\infty\right)}{\|A\left(A\vec{x}^{(k-2)}/\|A\vec{x}^{(k-2)}\|_\infty\right)\|_\infty}$

$= \dfrac{A^2\vec{x}^{(k-2)}}{\|A^2\vec{x}^{(k-2)}\|_\infty} = \cdots = \dfrac{A^k\vec{x}^{(0)}}{\|A^k\vec{x}^{(0)}\|_\infty}$

$$\vec{x}^{(k)} = \frac{a_1 \lambda_1^k \left[\vec{x}_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k \vec{x}_j \right]}{\left\| a_1 \lambda_1^k \left[\vec{x}_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k \vec{x}_j \right] \right\|_\infty} \approx \underbrace{\frac{a_1}{|a_1|} \frac{\vec{x}_1}{\|\vec{x}_1\|_\infty} \frac{\lambda_1^k}{|\lambda_1|^k}}_{\vec{v}}$$

when k is big

($\because \left|\frac{\lambda_j}{\lambda_1}\right| < 1$ for $j = 2, \dots, n$)

Note: $\vec{v}$ is an eigenvector associated to λ_1

In fact,

$$\|A\vec{x}^{(k)}\|_\infty \longrightarrow \left\|A\left(\frac{a_1}{|a_1|} \frac{\vec{x}_1}{\|\vec{x}_1\|_\infty}\right)\right\|_\infty = \left\| \frac{a_1}{|a_1|} \lambda_1 \frac{\vec{x}_1}{\|\vec{x}_1\|_\infty} \right\|_\infty = |\lambda_1|$$

$\frac{\lambda_1^k}{|\lambda_1|^k}$ can be removed under $\|\cdot\|_\infty$