

Fundamental Theorem for Iterative methods

(1) $x^{(k+1)} = G x^{(k)} + c$

Examples

$G = I - \tau A$

$G = I - D^{-1}A$ Jacobi

Fact 3': (1) $\rho(G) \leq \|A\|$ for any matrix norm
 (2) for any $\delta > 0$ there is norm $\|\cdot\|_\delta$ s.t.
 $\|A\|_\delta \leq \rho(A) + \delta$

Fact 4: $\rho(G) < 1$ is sufficient and necessary
 for the convergence of the iteration method (1)

Proof (of necessity) Assume $\rho(G) \geq 1$. $\exists \lambda$ $|\lambda| \geq 1$ $Gy = \lambda y$
 $y \neq 0$. Then consider choosing initial guess $x^{(0)}$
 such that $e^{(0)} = x^{(0)} - x = y$

$$e^{(1)} = G e^{(0)} = G y = \lambda y$$

$$e^{(2)} = \lambda^2 y \dots e^{(n)} = \lambda^n y_0$$

$$\|e^{(n)}\| = \underbrace{|\lambda|^n}_{\geq 1} \|y_0\| \geq \|y_0\|$$

therefore the method does not converge.

0

Definition: Matrix A is said to be similar to a matrix B if there is a nonsingular matrix S such that $B = S^{-1}AS$. Obviously, similar matrices have the same eigenvalues

$$A\psi = \lambda\psi \quad (\lambda, \psi) \text{ eigenpair}$$

$$\sigma(A) = \{ \lambda : \det(A - \lambda I) = 0 \}$$

$$S^{-1}AS\varphi = \mu\varphi$$

$$AS\varphi = \mu S\varphi \quad S\varphi = \xi$$

$$A\xi = \mu\xi \quad (\mu, \xi) \text{ eigenpair}$$

$$\mu \in \sigma(A)$$

Remark: The eigenvalues of a triangular matrix are the elements of its diagonal.

Theorem (Schur). Every square matrix is unitarily similar to a triangular matrix, i.e. $\exists U$ -unitary $U^{-1} = U^T = U^*$, s.t. $A = U^*CU$ - C -upper triangular matrix.

Theorem: Every square matrix is similar to an upper triangular matrix (possibly complex) whose off-diagonal elements are arbitrarily small.

Proof: Schur ensures $A = U^* C U$, C upper triangular. Let $0 < \epsilon < 1$ and take

$$E = \text{diag}(\epsilon, \epsilon^2, \dots, \epsilon^n)$$

Then

$$C_\epsilon = E^{-1} C E \text{ has elements } c_{ij} \epsilon^{j-i} \quad j \geq i$$

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} \epsilon & & & \\ & \epsilon^2 & & \\ & & \ddots & \\ & & & \epsilon^n \end{pmatrix} = \begin{pmatrix} \epsilon c_{11} & \epsilon^2 c_{12} & \dots & \epsilon^n c_{1n} \\ 0 & \epsilon^2 c_{22} & \dots & \epsilon^n c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \epsilon^n c_{nn} \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\epsilon} & 0 & & \\ 0 & \frac{1}{\epsilon^2} & & \\ \vdots & \vdots & \ddots & \\ & & & \frac{1}{\epsilon^n} \end{pmatrix} \begin{pmatrix} \epsilon c_{11} & \epsilon^2 c_{12} & \dots & \epsilon^n c_{1n} \\ 0 & \epsilon^2 c_{22} & \dots & \epsilon^n c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \epsilon^n c_{nn} \end{pmatrix} = \begin{pmatrix} c_{11} & \epsilon c_{12} & \dots & \epsilon^{n-1} c_{1n} \\ 0 & c_{22} & \dots & \epsilon^{n-2} c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_{nn} \end{pmatrix}$$

$$\text{Then } A = \underbrace{U^* E^{-1}}_{S^{-1}} \underbrace{C_\epsilon}_{C_\epsilon} \underbrace{E U}_{S} = S^{-1} C_\epsilon S$$

$$S^{-1} = U^* E = U^* \epsilon$$

Proof (of Fact 3 second part)

$$A = S' C_{\epsilon} S$$

$$SAS^{-1} = C_{\epsilon} = D + T_{\epsilon}$$

diagonal \downarrow strictly upper Δ

$$\lambda(A) = \lambda(SAS^{-1}) = \lambda(D)$$

On the other hand

$$\#N\#2 \quad \|A\|_{\infty} = \max_j \sum_i |a_{ij}|$$

$$\|SAS^{-1}\|_{\infty} \leq \|D\|_{\infty} + \|T_{\epsilon}\|_{\infty} \leq \rho(D) + \delta$$

The only question we need to answer is whether $\|SAS^{-1}\|_{\infty}$ is a subordinate norm of A

$$\|A\|_{\delta} \stackrel{def}{=} \|SAS^{-1}\|_{\infty}$$

$$S'x = y$$

$$\|A\|_{\delta} = \sup \frac{\|SAS^{-1}x\|_{\infty}}{\|x\|_{\infty}} = \sup \frac{\|SAy\|_{\infty}}{\|Sy\|_{\infty}} \stackrel{def}{=} \frac{\|Ay\|_{\delta}}{\|y\|_{\delta}}$$

$$\|y\|_{\delta} = \|Sy\|_{\infty}$$

this is a vector norm

Proof (of sufficiency) Let $\rho(G) < 1$, and take $\delta = \frac{1-\rho(G)}{2} > 0$. Construct a norm $\|G\|_\delta$ such that

$$\|G\|_\delta \leq \rho(G) + \delta = \rho(G) + \frac{1-\rho(G)}{2}$$

$$\|G\|_\delta \leq \frac{1+\rho(G)}{2} < 1$$

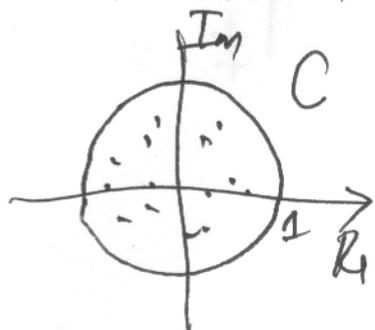
Then for the error we have

$$\|e^{(k)}\|_\delta \leq \|G\|_\delta \|e^{(k-1)}\|_\delta \leq \dots \leq \|G\|_\delta^k \|e^{(0)}\|$$

Since $\|G\|_\delta < 1$ $\|e^{(k)}\|_\delta \rightarrow 0$ when $k \rightarrow \infty$ and the iteration method converges.

Remark: Key in the proof is a construction of a proper subordinate matrix norm $\|\cdot\|_\delta$ so that $\|G\|_\delta < 1$. And we know that if this is the case then the method converges.

To study the convergence of the basic iterative methods we need to know essentially where the spectrum of the iteration matrix G is. The Fundamental Theorem says, that if $\sigma(G) \subset \{z \in \mathbb{C} : |z| < 1\}$



then we have convergence.

How we can deal with this problem?

Theorem. (Gershgorin's Th). The spectrum of a matrix $A \in \mathbb{C}^{n \times n}$ is contained in the union of the disks D_i , $i=1, \dots, n$ in the complex plane

$$D_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}, i=1, \dots, n$$

Example

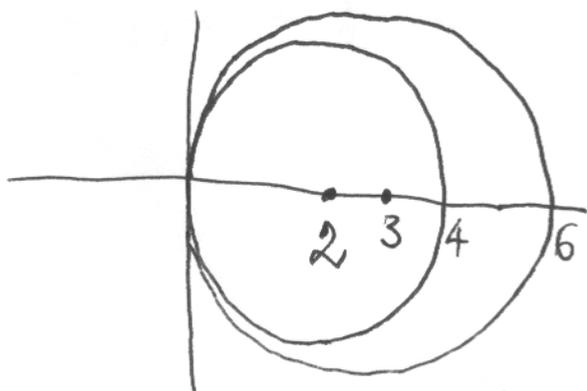
$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

$$D_1 = \{z \in \mathbb{C} : |z - 2| \leq 2\}$$

$$D_2 = \{z \in \mathbb{C} : |z - 2| \leq 2\}$$

$$D_3 = \{z \in \mathbb{C} : |z - 3| \leq 3\}$$

$$\cup D_i = D_3 = \{z \in \mathbb{C} : |z - 3| \leq 3\}$$



Richardson

$$x^{(k+1)} = \underbrace{(I - \tau A)}_G x^{(k)} + \tau b$$

$$\begin{pmatrix} \frac{1}{2} & & \\ 1 - \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix} = G$$

$$\tau = \frac{1}{4} \quad G = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

$$D_{1,2} \quad |z - \frac{1}{2}| \leq \frac{1}{2}$$

$$\tau = \frac{1}{6}$$

$$|z - \frac{1}{4}| \leq \frac{3}{4}$$

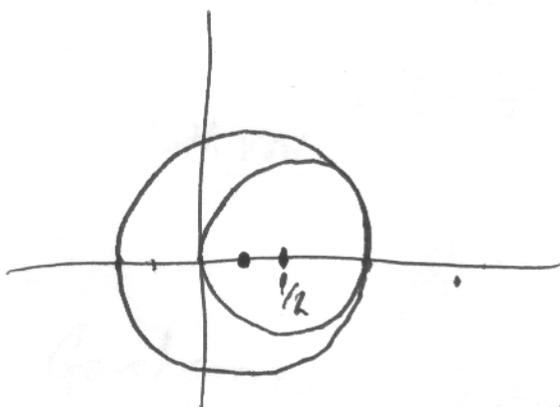
$$\lambda(G) \in D$$

$$1 - \frac{\lambda_j}{4} = \frac{1}{4} = \lambda_1$$

$$= \frac{1}{4}$$

$$1 - \frac{3 - \sqrt{3}}{4} = \lambda_2 = \frac{1}{4} - \frac{\sqrt{3}}{4}$$

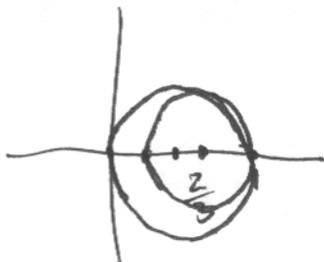
$$1 - \frac{3 + \sqrt{3}}{4} = \lambda_3 = \frac{1}{4} + \frac{\sqrt{3}}{4}$$



$$I - \frac{1}{6}A = \begin{pmatrix} \frac{2}{3} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{2}{3} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{2}{6} & \frac{1}{2} \end{pmatrix}$$

$$D_{1,2} \quad \left| \frac{2}{3} - z \right| \leq \frac{1}{3}$$

$$\left| \frac{1}{3} - z \right| \leq \frac{1}{2}$$



$$A = D - L - U$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

$$\text{Jacobi } Q = D \quad x^{(k+1)} = \underbrace{D^{-1}(L+U)}_G x^{(k)} + D^{-1}b$$

Theorem: Jacobi method is convergent for any strictly diagonally dominant matrix A .

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{strictly diagonally dominant}$$

Proof: Enough to show that $\rho(G) < 1$.

$$Gx = \lambda x \quad \lambda \text{ eigenvalue}$$

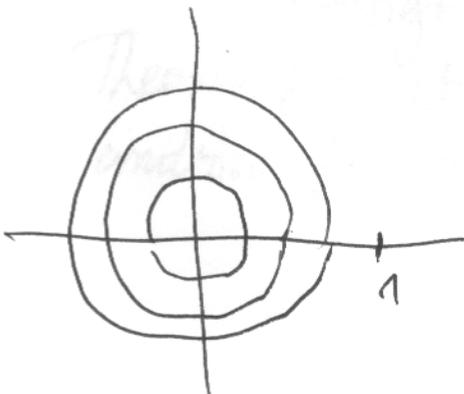
$$(G)_{i\text{-th row}} \Rightarrow \begin{vmatrix} 0 & -\frac{a_{i2}}{a_{ii}} & -\frac{a_{i3}}{a_{ii}} & \dots & -\frac{a_{in}}{a_{ii}} \\ \frac{a_{ii}}{a_{ii}} & \dots & 0 & \dots & -\frac{a_{in}}{a_{ii}} \end{vmatrix}$$

$$\text{Gershgorin } \lambda \in D = \cup D_i \subset \{ |z| < 1 \}$$

$$D_i = \left\{ z \in \mathbb{C} : |z - 0| \leq \sum_{\substack{j=1 \\ j \neq i}}^n \frac{|a_{ij}|}{|a_{ii}|} \right\}$$

$$D_i = \{ |z| \leq q < 1 \}$$

$$\rho(G) \leq q < 1$$



Proof: Let λ be an eigenvalue of $A, x \neq 0$
 be the corresponding eigenvector, i.e. $Ax = \lambda x, x \neq 0$.
 We can always normalize x so that $\|x\|_{\infty} = \max_j |x_j| = 1$.
 Then for some i we shall have $\|x\|_{\infty} = |x_i| = 1$.

Take the i -th row of $Ax = \lambda x$

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ii}x_i + \dots + a_{in}x_n = \lambda x_i$$

$$(a_{ii} - \lambda)x_i = \sum_{j \neq i} a_{ij}x_j$$

$$|a_{ii} - \lambda| \leq \sum_{j \neq i} |a_{ij}| |x_j| = \sum_{j \neq i} |a_{ij}|$$

therefore $\lambda \in D_i$

Since λ is any eigenvalue from $\sigma(A)$

$$\text{then } \sigma(A) \subset \bigcup_{i=1}^n D_i$$

The Gershgorin Localization and Fundamental Theorem for iterative methods form a sufficient condition for convergence of iteration method