MA633L-Numerical Analysis

Lecture 28 : Numerical Linear Algebra - Matrix Norms

Panchatcharam Mariappan[1]

[1]Associate Professor
Department of Mathematics and Statistics
IIT Tirupati, Tirupati

March 17, 2025

Recap

Iterative Schemes

The underlying principles behind iterative methods to solve $Ax = b$ are as follows:

1. Guess any $x^{(0)}$
2. Compute the residual $r^{(0)} = b - Ax^{(0)}$
3. Compute the residual norm, that is $\|r^{(0)}\|$
4. Use an algorithm to compute $x^{(1)}$ involving A, b and $x^{(0)}$
5. Recompute the residual norm $\|r^{(1)}\| = \|b - Ax^{(1)}\|$
6. For given $x^{(i)}$, iterate this process until you get the residual as zero or an equivalent condition

Norms

Definition 1 (Vector Norms)
A vector norm, usually denoted by $\|.\|$, is a function from a vector space V to the set of non-negative real numbers that obeys the following three postulates. $\|.\| : V \to \mathbb{R}_+$ such that

$$\begin{aligned} \|\mathbf{x}\| &> 0 \quad \text{if } \mathbf{x} \neq 0, \mathbf{x} \in V \\ \|\alpha \mathbf{x}\| &= |\alpha| \|\mathbf{x}\| \quad \text{if } \alpha \in \mathbb{R}, \mathbf{x} \in V \\ \|\mathbf{x} + \mathbf{y}\| &\leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \text{if } \mathbf{x}, \mathbf{y} \in V \end{aligned}$$

Matrix Norms

Definition 2 (Matrix Norm)

A matrix norm on the set of all $n \times n$ matrices, usually denoted by $\|.\|$, is a function from a $M_{n\times n}(\mathbb{R})$ to the set of non-negative real numbers that obeys the following three postulates. $\|.\| : M_{n\times n}(\mathbb{R}) \to \mathbb{R}_+$ such that

$$\|A\| \geq 0 \quad \text{(non-negativity)}$$

$$\|A\| = 0, \text{ if and only if A is a zero matrix} \quad \text{(Mapping of the Identity)}$$

$$\|\alpha A\| = |\alpha|\|A\| \quad \text{if } \alpha \in \mathbb{R} \quad \text{(Scalar multiplication)}$$

$$\|A + B\| \leq \|A\| + \|B\| \quad \text{(Triangle Inequality)}$$

$$\|AB\| \leq \|A\|\|B\| \quad \text{(Consistency)}$$

where $A, B \in M_{n\times n}(\mathbb{R})$

Matrix Norm

If $A = (a_{ij})$ is an $n \times n$ matrix and $\|.\|$ is a vector norm in $\mathbb{R}^n$, then the matrix norm is

$$\|A\| = \max_{\|x\|=1} \|Ax\| \qquad \text{(Vector Induced Norm)}$$

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| \qquad \text{(Row Sum Norm)}$$

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| \qquad \text{(Column Sum Norm)}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2} \qquad \text{(Frobenius Norm)}$$

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sigma_{max}(A)$$

Matrix Norm

- The $||A||_1$ norm is also called as the **column sum norm** as it computes the maximum absolute column sum of the matrix.
- The $||A||_\infty$ norm is also called as the **row sum norm** as it computes the maximum absolute row sum of the matrix.
- An induced matrix norm is also called as **operator norm**
- The $||A||_2$ norm is also called as the **spectral norm**
- For Symmetric matrix $||A||_1 = ||A||_\infty$ and $||A||_2 = \rho(A)$
- $||x||_2$ norm is usually called as Euclidean norm, but $||A||_2$ is not.
- If Q is orthogonal, then $||Q||_2 = 1$ and hence an orthogonal matrix is also called an isometric matrix.

Eigenvalues

Eigenvalues and Eigenvector

$$Ax = \lambda x$$

Singular Values

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$$

The spectral radius

$$\rho(A) = \max_{1 \le i \le n} |\lambda_i|$$

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sigma_{max}(A)$$

This is the minimum norm and provides the tightest measure of size.

Condition Number

We call a matrix A as convergent if

$$\lim_{k\to\infty} (A^k)_{ij} = 0$$

for all i, j

$$\kappa(A) = \|A\| \|A^{-1}\|$$

If A is singular, it is customary to define $\kappa(A) = \infty$.

Challenges

Challenges

1. Definition of condition number involves matrix inverse, which is not easy to compute and requires more work than solving the system $Ax = b$.
2. In practice, condition number is merely estimated, using another relative expensive byproduct of the solution procedure.
3. If z is a solution of $Ax = y$, then

$$\|z\| = \|A^{-1}y\| \leq \|A^{-1}\|\|y\| \implies \frac{\|z\|}{\|y\|} \leq \|A^{-1}\|$$

4. If we can find a vector y such that $\frac{\|z\|}{\|y\|}$ is as large as possible, then we can find a reasonable estimate for $\|A^{-1}\|$.

Challenges

1. Usually, the ℓ_2 norm or spectral norm is used for the condition number as it provides the tightest measure of size.
2. Also, this norm does not require any matrix inverse, instead computation of eigenvalues. With the help of Gersghorin's theorem and Spectral radius, one can find this.
3. The condition number of a matrix in ℓ_2 norm can be defined as

$$\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}$$

where σ_{max} and σ_{min} denote the maximal and minimal singular values of A respectively.

Challenges

1. If A is normal ($A^*A = AA^*$), then

$$\kappa(A) = \frac{|\lambda_{max}(A)|}{|\lambda_{min}(A)|}$$

2. If A is unitary $A^*A = AA^* = I$, then

$$\kappa(A) = 1$$

3. If $\|.\|$ is the ℓ_∞ norm and A is a nonsingular lower or upper triangular matrix, then

$$\kappa(A) \geq \frac{\max_{1\leq i\leq n} |a_{ii}|}{\min_{1\leq i\leq n} |a_{ii}|}$$

Properties

Verify whether the following Statements are True or Not.

1. For any matrix A, $\kappa(A) \geq 1$
2. $\kappa(I) = 1$
3. $\kappa(\alpha A) = \kappa(A)$
4. $\kappa(D) = \frac{\max_i |d_i|}{\min_i |d_i|}$
5. $\kappa(AB) \leq \kappa(A)\kappa(B)$
6. If $A = A^T$, then $\kappa(A^2) = \kappa(A)^2$

Theorems

Theorem 3

If A is and $n \times n$ matrix such that $\|A\| < 1$, then $I - A$ is invertible, and

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

From this theorem, we can observe that

$$\|(I - A)^{-1}\| \leq \sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1 - \|A\|}$$

Theorems

Theorem 4

If A and B are $n \times n$ matrices such that $\|I - AB\| < 1$, then A and B are invertible. Furthermore,

$$A^{-1} = B \sum_{k=0}^{\infty} (I - AB)^k$$

and

$$B^{-1} = \sum_{k=0}^{\infty} (I - AB)^k A$$

Theorems

Proof: By previous theorem, $AB = I - (I - AB)$ is invertible and

$$(AB)^{-1} = \sum_{k=0}^{\infty} (I - AB)^k$$

$$A^{-1} = B(AB)^{-1} = B \sum_{k=0}^{\infty} (I - AB)^k$$

$$B^{-1} = (AB)^{-1} A = \sum_{k=0}^{\infty} (I - AB)^k A$$

Iterative Methods

Iterative Methods

The general algorithm for solving a system $Ax = b$ is as follows:

1. Choose a nonsingular matrix Q
2. Choose an arbitrary starting vector $x^{(0)}$
3. Generate vectors $x^{(1)}, x^{(2)}, \cdots$ recursively from the equation

$$Qx^{(k)} = (Q - A)x^{(k-1)} + b, \quad k = 1, 2, 3, \cdots \tag{1}$$

Suppose x is the solution of the system $Ax = b$ and $x^{(k)}$ converges to x, as $k \to \infty$, then

$$Qx = (Q - A)x + b$$

Note, that the system (1) should be easy to solve for $x^{(k)}$ when the right hand side is known. Also, Q should be chosen to ensure that $x^{(k)}$ converges to x, no matter, what initial vector is used and the convergence should be rapid.

Iterative Methods

Note that the true solution x satisfies the equation

$$x = (I - Q^{-1}A)x + Q^{-1}b \tag{2}$$

Therefore, x is a fixed point of the mapping

$$f(x) = (I - Q^{-1}A)x + Q^{-1}b$$

From (1)

$$Qx^{(k)} = (Q - A)x^{(k-1)} + b, \quad k = 1, 2, 3, \cdots$$
$$\implies x^{(k)} = (I - Q^{-1}A)x^{(k-1)} + Q^{-1}b, \quad k = 1, 2, 3, \cdots$$

Now,

$$x^{(k)} - x = (I - Q^{-1}A)(x^{(k-1)} - x), \quad k = 1, 2, 3, \cdots \tag{3}$$

$$\implies \|x^{(k)} - x\| = \|(I - Q^{-1}A)\|\|(x^{(k-1)}) - x\|$$
$$\implies \|x^{(k)} - x\| = \|(I - Q^{-1}A)\|^{k}\|(x^{(0)}) - x\|$$

Iterative Methods

If $\|I - Q^{-1}A\| < 1$, we can conclude that

$$\lim_{k\to\infty} \|x^{(k)} - x\| = 0$$

By above theorem, it is guaranteed that if $\|I - Q^{-1}A\| < 1$, then both $Q^{-1}A$ and A are invertible.

Theorem 5

If $\|I - Q^{-1}A\| < 1$ for some matrix norm, then the sequence produced by (1) converges to the solution of $Ax = b$ for any initial vector $x^{(0)}$.

Theorem 6

If all eigenvalues of $I - Q^{-1}A$ lies in the open unit disc $|z| < 1$, then the sequence produced by (1) converges to the solution of $Ax = b$ for any initial vector $x^{(0)}$.

Iterative Methods

The above theorem implies that the spectral radius of $I - Q^{-1}A$ must be less than 1, that is,

$$\rho(I - Q^{-1}A) < 1$$

Let $r^{(k)}$ denote the residual vector obtained from $x^{(k)}$ after k iterations, then we get

$$r^{(k)} = b - Ax^{(k)}.$$

By above theorem, if $\|I - Q^{-1}A\| < 1$, then $\|r^{(k)}\| \to 0$.

Richardson/Jacobi Iteration

Richardson Iteration

$$Qx^{(k)} = (Q - A)x^{(k-1)} + b, \quad k = 1, 2, 3, \cdots$$

The Richardson iteration is obtained when $Q = I$.
So equation (1) becomes

$$x^{(k)} = (I - A)x^{(k-1)} + b = x^{(k-1)} + r^{(k-1)}$$

Richardson Iteration

Example 7

Compute the first 100 iterates on the following problem using Richardson algorithm starting with $x = (0, 0, 0)^T$

$$\begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/3 & 1 & 1/2 \\ 1/2 & 1/3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 11/18 \\ 11/18 \\ 11/18 \end{pmatrix}$$

Using the computer program, we can obtain that

$$x^{(0)} = (0, 0, 0)^T$$

$$x^{(80)} = (0.333333, 0.333333, 0.333333)^T$$

Jacobi method

In Jacobi method our choice of Q is the diagonal matrix of A.

$$x^{(k)} = (I - D^{-1}A)x^{(k-1)} + D^{-1}b$$

In particular,

$$x_i^{(k)} = -\sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k-1)} + \frac{b_i}{a_{ii}}$$

When the norm is ℓ_∞, we get that

$$\|I - D^{-1}A\|_\infty = \max_{1\leq i\leq n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right|$$

Jacobi method

Theorem 8

If A is diagonally dominant, then the sequence produced by Jacobi iteration converges to the solution of $Ax = b$ for any starting vector.

Example 9

Compute the first 100 iterates on the following problem using Jacobi algorithm starting with $x = (0, 0, 0)^T$

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ -5 \end{pmatrix}$$

Jacobi method

$$x^{(k)} = (I - D^{-1}A)x^{(k-1)} + D^{-1}b$$

$$\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 0 \\ 1/303 & 1/3 & \\ 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ x_3^{(k-1)} \end{pmatrix} + \begin{pmatrix} 1/2 \\ 8/3 \\ -5/2 \end{pmatrix}$$

$$x^{(0)} = (0, 0, 0)^T$$

$$x^{(21)} = (2.00, 3.00, -1.00)^T$$

Gauss-Seidel and SOR

Gauss-Seidel

In Gauss-Seidel method our choice of Q is a little different. We can write A as follows:

$$A = D - L - U$$

where D is the diagonal matrix of A, L is the negative of the strictly lower triangular part of A and U is the negative of the strictly upper triangular matrix. Now, choose $Q = D - L$, that is the lower triangular part of A, then

$$Qx^{(k)} = (Q - A)x^{(k-1)} + b$$

can be written as

$$(D - L)x^{(k)} = (D - L - (D - L - U))x^{(k-1)} + b$$

$$\implies (D - L)x^{(k)} = (U)x^{(k-1)} + b$$

Gauss-Seidel

In particular,

$$x_i^{(k)} = -\sum_{\substack{j=1 \\ j<i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k)} - \sum_{\substack{j=1 \\ j>i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k-1)} + \frac{b_i}{a_{ii}}$$

Notice that Gauss-Seidel and Jacobi method is almost similar. For Jacobi method, the right-hand side values depend only on the previous iteration whereas the Gauss-Seidel method uses the updated information for its computation. However, the major drawback of Gauss-Seidel method is that it can't be parallelized. That is, in the Jacobi algorithm, the $x_i^{(k)}$ components can be computed simultaneously, whereas in the Gauss-Seidel algorithm, they must be computed serially, since the computation of $x_i^{(k)}$ depends on all $x_1^{(k)}, x_2^{(k)}, \cdots, x_{i-1}^{(k)}$. Therefore, Jacobi method is preferably used for parallel processing.

Gauss-Seidel

Theorem 10

If A is diagonally dominant, then the sequence produced by Gauss-Seidel iteration converges to the solution of $Ax = b$ for any starting vector.

Note that for both Jacobi and Gauss-Seidel method, diagonally dominant is a sufficient condition, but not necessary. There are matrices that are not diagonally dominant, but still these two methods converge.

$$A = \begin{pmatrix} 0.5 & 1 \\ 1 & 0.5 \end{pmatrix}$$

When $A = D - L - U$, the Jacobi method can also be written as

$$Dx^{(k)} = (L + U))x^{(k-1)} + b$$

SOR

In order to accelerate Gauss-Seidel method, we introduce a relaxation factor ω and obtain a new method called successive overrelaxation (SOR) method. Here we consider Q as follows:

$$Q = \frac{1}{\omega}(D - \omega L)$$

Then the algorithm is given by

$$\frac{1}{\omega}(D - \omega L)x^{(k)} = \left(\frac{1}{\omega}(D - \omega L) - D + L + U\right) x^{(k-1)} + b$$

$$\implies (D - \omega L)x^{(k)} = ((D - \omega L) - \omega(D - L - U))x^{(k-1)} + \omega b$$

$$\implies (D - \omega L)x^{(k)} = ((1 - \omega)D + \omega U)x^{(k-1)} + \omega b$$

SOR

In particular, we have

$$x_i^{(k)} = \omega \left[-\sum_{\substack{j=1 \\ j<i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k)} - \sum_{\substack{j=1 \\ j>i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k-1)} + \frac{b_i}{a_{ii}} \right] + (1-\omega) x_i^{(k-1)}$$

Here, the lower triangular part of A is chosen in such a way that each diagonal element is replaced by a_{ij}/ω.
For symmetric matrix,

$$A = D - L - L^T$$

you can apply two SOR methods in opposite directions and is called symmetric successive overrelaxation (SSOR) method (Explore it).

SOR

Theorem 11

If A is symmetric, positive definite matrix and $0 < \omega < 2$, then the SOR method converges for any starting vector $x^{(0)}$.

Example 12

Compute the first 100 iterates on the following problem using Richardson algorithm starting with $x = (0,0,0)^T$ and $\omega = 1.1$

$x = (0,0,0)^T$

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ -5 \end{pmatrix}$$

$$x^{(0)} = (0,0,0)^T$$

$$x^{(7)} = (2.00, 3.00, -1.00)^T$$

JOR

The other iterative scheme can be improve by the introduction of auxiliary equation and an acceleration parameter ω was follows:

$$Qz^{(k)} = (Q - A)x^{(k-1)} + b$$

$$x^{(k)} = \omega z^{(k)} + (1 - \omega)x^{(k-1)}$$

or

$$x^{(k)} = \omega[(I - Q^{-1}A)x^{(k-1)} + Q^{-1}b] + (1 - \omega)x^{(k-1)}$$

JOR

- When $\omega = 1$, this reduces to basic iterative methods.
- When $1 < \omega < 2$, the rate of convergence may be improved, which is called overrelaxation.
- When $Q = D$, we have the Jacobi overrelaxation (JOR) method

$$x^{(k)} = \omega[(I - D^{-1}A)x^{(k-1)} + D^{-1}b] + (1 - \omega)x^{(k-1)}$$

Theorem 13

If A is symmetric, positive definite matrix and $0 < \omega < \frac{2}{\rho(D^{-1}A)}$, then the JOR method converges for any starting vector $x^{(0)}$.

Stationary Iterative Methods

- Iterative methods for $Ax = b$ is called stationary iterative methods if it can be written as

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

  with constant R

- This iteration converges to the solution x if and only if $\rho(G) < 1$.
- A splitting is a decomposition $A = M - K$ with nonsingular M
- Stationary iterative method from splitting

$$Ax = Mx - Kx = b \implies Mx = Kx + b \implies x = M^{-1}Kx + M^{-1}b = Gx + c$$

- Find a splitting $A = M - K$ such that $M^{-1}Kx$ and $M^{-1}b$ are easy to compute and $\rho(M^{-1}K)$ is small.

Stationary Iterative Methods

- When $M = I$, $\rho(M^{-1}K) = \rho(I - A) =$ is not small
- When $M = A$, $K = 0, \rho(M^{-1}K) = 0$, but expensive to compute M^{-1}
- Split $A = D - L - U$
- Jacobi method $M = D, K = L + U$
- Gauss-Seidel, $M = D - L, K = U$

Theorems on Convergence of Iterative Methods

Convergence

Let us find necessary and sufficient condition for the convergence of the iterative method

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

For example, as per our splitting in Jacobi, Gauss-Seidel, Richardson and SOR, we can write the splitting as

$$x^{(k)} = Gx^{(k-1)} + c, \quad k = 1, 2, 3, \cdots \tag{4}$$

Convergence

Theorem 14

The iteration (4) converges to $(I - G)^{-1}c$ if and only if $\rho(G) < 1$.

Proof for Sufficient Part: Suppose $\rho(G) < 1$.

Claim: The iteration (4) converges to $(I - G)^{-1}c$

Then there exists a matrix norm $\|.\|$ such that $\|G\| < 1$. Hence,

$$x^{(k)} = G^k x^{(0)} + \sum_{j=0}^{k-1} G^j c, \quad k = 1, 2, 3, \cdots$$

The first term goes to zero as $k \to \infty$ since

$$\|G^k x^{(0)}\| \leq \|G\|^k \|x^{(0)}\|,$$

Hence the claim follows.

Convergence

Proof for Necessary Part: Suppose the iteration (4) converges to $(I-G)^{-1}c$.
<u>Claim:</u> $\rho(G) < 1$
Suppose $\rho(G) \geq 1$. Let λ, u be an eigenpair of G with $|\lambda| \geq 1$. Let $x^{(0)} = 0$ and $c = u$, then we obtain

$$x^{(k)} = \sum_{j=0}^{k-1} G^j u = \sum_{j=0}^{k-1} \lambda^j u = \begin{cases} ku & \lambda = 1 \\ \frac{1-\lambda^k}{1-\lambda} u & \lambda \neq 1 \end{cases}$$

Therefore, the iteration (4) does not converges when $\rho(G) \geq 1$. Hence the claim follows.

Convergence

Corollary 1
The iteration

$$x^{(k)} = (I - Q^{-1}A)x^{(k-1)} + Q^{-1}b, \quad k = 1, 2, 3, \cdots \tag{5}$$

converges to $Ax = b$ for any initial guesses $x^{(0)}$ if and only if $\rho(I - Q^{-1}A) < 1$.

Convergence of Gauss-Seidel

Theorem 15
If A is diagonally dominant then the Gauss-Seidel method converges for any initial guess $x^{(0)}$

By above corollary, if we prove $\rho(I - Q^{-1}A) < 1$, the theorem follows. Let λ be an eigenvalue and x be the corresponding eigenvector of $I - Q^{-1}A$ with $\|x\|_\infty = 1$.

$$(I - Q^{-1}A)x = \lambda x \implies (Q - A)x = \lambda Qx$$

$$-\sum_{j=i+1}^{n} a_{ij}x_j = \lambda \sum_{j=1}^{i} a_{ij}x_j, 1 \leq i \leq n$$

Convergence of Gauss-Seidel

$$\lambda a_{ii}x_i = -\lambda \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j, 1 \leq i \leq n$$

Now, pick the index i such that $|x_i| = 1$, then

$$|\lambda||a_{ii}| = \left| -\lambda \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j \right| \leq |\lambda| \sum_{j=1}^{i-1} |a_{ij}| + \sum_{j=i+1}^{n} |a_{ij}|$$

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

Convergence of Gauss-Seidel

$$|\lambda|\left(|a_{ii}| - \sum_{j=1}^{i-1}|a_{ij}|\right) \leq \sum_{j=i+1}^{n}|a_{ij}| \implies |\lambda| \leq \frac{\displaystyle\sum_{j=i+1}^{n}|a_{ij}|}{|a_{ii}| - \displaystyle\sum_{j=1}^{i-1}|a_{ij}|}$$

Since

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

Since this holds for all eigenvalues of $I - Q^{-1}A$, we obtain $\rho(I - Q^{-1}A) < 1$.

Convergence of Gauss-Seidel

Theorem 16
If $\|I - Q^{-1}A\| < 1$ for some matrix norm, then the sequence produced by (5) converges to the solution of $Ax = b$ for any initial vector $x^{(0)}$

Proof: We know that for any $n \times n$ matrix $\rho(A) \leq \|A\|$ for any natural norm.

$$\|I - Q^{-1}A\| < 1 \implies \rho(I - Q^{-1}A) \leq \|I - Q^{-1}A\| < 1$$

Thanks

Doubts and Suggestions

panch.m@iittp.ac.in

MA633L-Numerical Analysis

Lecture 28 : Numerical Linear Algebra - Matrix Norms

Panchatcharam Mariappan[1]

[1]Associate Professor
Department of Mathematics and Statistics
IIT Tirupati, Tirupati

March 17, 2025