

**Engineering Mathematics-II**  
**Professor Jitender Kumar**  
**Department of Mathematics,**  
**Indian Institute of Technology, Kharagpur.**

**Lecture 23**

**Iterative Methods for Solving System of Linear Equations (Cont.)**

So, welcome back to lectures on Engineering Mathematics 2 this is lecture number 23 and we will continue with the Iterative Methods for Solving System of Linear Equations.

(Refer Slide Time: 00:24)



And in particular, today we will be talking about the convergence of those iterative methods which we have derived in one lecture and then some numerical examples were performed in the next one. And in this lecture we will be talking about the convergence of the Jacobi method, convergence of the Gauss Seidel method.

(Refer Slide Time: 00:46)

Iterative Methods:  $x^{(k+1)} = Gx^{(k)} + Hb$

**Necessary and Sufficient Conditions:**

The iterative methods converge for any initial guess if and only if all the eigenvalues of the iteration matrix  $G$  have absolute value less than 1.

Dr. Khanna

NPTEL

Iterative Methods:  $x^{(k+1)} = Gx^{(k)} + Hb$

**Necessary and Sufficient Conditions:**

The iterative methods converge for any initial guess if and only if all the eigenvalues of the iteration matrix  $G$  have absolute value less than 1.

Dr. Khanna

NPTEL

So, in general the, all these iterative methods take the following form that the value of  $x$ , the unknown at  $k$  plus 1 th iteration is equal to  $G$ , the matrix and then the  $x$  taken from the previous iteration and  $H$  times  $b$ . So, this is the general structure we have discussed for any iterative methods for solving system of linear equations and two particular cases where we have different  $G$  and  $H$  that means the Gauss Seidel method and the Jacobi methods were discussed.

So, first we will be talking about the necessary and sufficient conditions for the convergence. So, necessary sufficient means. We will know exactly whether the numerical method will converge to the actual solution or it will not converge to the actual solution. So, the iterative method converge for any initial guess, if and only if all the eigenvalues of the iteration matrix  $G$  have absolute value less than 1.

So, as I discuss before that these iteration matrix is  $a$ , is the crucial, is the central part of a iterative method for solving system of linear equation and now we do see that if all the eigenvalues of the iteration matrix  $G$  have absolute value less than 1 in that case the iterative method will converge for any initial guess and this is if and only if.

That means if any of the initial eigenvalue in absolute value is greater than 1, then the iterative method will not converge or if it converge for any initial guess then definitely all the eigenvalues of that matrix have absolute value less than 1. So, this is if and only if condition.

(Refer Slide Time: 02:59)

The slide contains the following text:

**Iterative Methods:**  $x^{(k+1)} = Gx^{(k)} + Hb$

**Necessary and Sufficient Conditions:**

The iterative methods converge for any initial guess if and only if all the eigenvalues of the iteration matrix  $G$  have absolute value less than 1.

OR

The iterative methods converge if and only if the spectral radius (largest absolute eigenvalue) of  $G$  is less than 1, i.e.  $\rho(G) < 1$ .

The slide also features a small video inset of a man in the bottom right corner and logos for IIT Kharagpur and NPTEL at the bottom.

This result can be also stated in this way that the iterative methods converge if and only if the spectral radius. So, the spectral radius is nothing but the largest absolute value. So, if the largest absolute value is less than 1 that means all the eigenvalues are less than 1. So, these two results are actually the same. So, whether we call that the spectral radius is less than 1 or we say that all the eigenvalues of this iteration matrix  $G$  have absolute value less than 1 both are equivalent.

(Refer Slide Time: 03:39)

Lemma: Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

*Handwritten annotations:*  
- A red arrow points from  $\mathbf{0}$  to the text "Zero Matrix" written below it.  
- A red circle is drawn around the expression  $\rho(A) < 1$ .

IT Khargpur  
NPTEL

Lemma: Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

Sketch of the proof: Suppose  $A$  is diagonalizable. Then there exist a matrix  $P$  such that

$$A = PDP^{-1}$$

Where  $D$  is a diagonal matrix having the eigenvalues of  $A$  on the diagonal. Therefore

$$A^m = PD^mP^{-1}$$

*Handwritten annotations:*  
- A red circle is drawn around the expression  $A = PDP^{-1}$ .

IT Khargpur  
NPTEL

To prove this result or to see at least some idea, to get some idea of the proof of this necessary and sufficient conditions. We will first prove this result, let this  $A$  be a square matrix, then we will show that,  $A$  power  $m$ , so the product of this  $A$   $m$  times. So, if  $A$  power  $m$  and when  $m$  goes to infinity, this goes to  $0$  matrix. If and only if the spectral radius of this  $A$  that means all the eigenvalues of  $A$  have absolute value less than  $1$ .

So, if the spectral radius of  $A$  is less than  $1$  than this limit will definitely will go to the  $0$  matrix. So, this is  $0$  matrix and if this happens that means the spectral radius is less than  $1$ . This is if and only if condition in this case. So, to just see the proof of this. Suppose, so this is the assumption in this, under this restriction we are proving this but this is more general of course.

The result is more general suppose  $A$  is diagonalizable then there exist a matrix  $P$  that we know already from the linear algebra. That  $A$  can be written as  $PDP^{-1}$ , where  $P$  is invertible matrix which must exist. And this  $D$  is the diagonal matrix having the eigenvalues of  $A$  on its diagonal.

(Refer Slide Time: 05:18)

Lemma: Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

Sketch of the proof: Suppose  $A$  is diagonalizable. Then there exist a matrix  $P$  such that

$$A = PDP^{-1}$$

Where  $D$  is a diagonal matrix having the eigenvalues of  $A$  on the diagonal. Therefore

$$A^m = PD^mP^{-1} \in$$

So, that means if  $A$  is  $P, DP$  inverse then we can also find the product of  $A$  for instance  $A$  power  $m$ , and that also we have seen in the topic of linear algebra in Engineering Mathematics 1, that this product is nothing but  $P$  and the product will be just for  $D$  which is very easy because  $D$  is a diagonal matrix and this is one of the applications of this diagonalization where we can get the power of a matrix very large power we can talk about because getting  $D$  power  $m$  is trivial in this case.

(Refer Slide Time: 05:58)

**Lemma:** Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

**Sketch of the proof:** Suppose  $A$  is diagonalizable. Then there exist a matrix  $P$  such that

$$A = PDP^{-1}$$

Where  $D$  is a diagonal matrix having the eigenvalues of  $A$  on the diagonal. Therefore

$$A^m = PD^mP^{-1}$$

with  $D = \begin{bmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n^m \end{bmatrix}$

It has only the diagonal entries and the diagonal entries will get the power  $m$  at the end and we have to just  $P$  and  $P$  Inverse and we will get this product  $A$  power  $m$ . So, this  $D$  will take such a form that  $\lambda_1$  power  $m$ ,  $\lambda_2$  power  $m$  and  $\lambda_n$  power  $m$ . So, all these  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_n$  will be placed in the diagonal. So, we assume that  $A$  is diagonalizable, and these  $\lambda$ s may not be distinct. So, they may some of them may be repeated but still since it is diagonalizable, so we will have such a form for  $D$ .

(Refer Slide Time: 06:32)

**Lemma:** Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

**Sketch of the proof:** Suppose  $A$  is diagonalizable. Then there exist a matrix  $P$  such that

$$A = PDP^{-1}$$

Where  $D$  is a diagonal matrix having the eigenvalues of  $A$  on the diagonal. Therefore

$$A^m = PD^mP^{-1}$$

with  $D = \begin{bmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n^m \end{bmatrix}$

$\Rightarrow \lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff all the eigenvalues satisfy  $|\lambda_i| < 1$  ( $\rho(A) < 1$ )

And now it is clear that this  $A$  power  $m$  if  $m$  approaches to infinity because this is at the end we have  $\lambda_1$ ,  $\lambda_2$  and power goes to these eigenvalues. So, if all the eigenvalues are less than 1 or in other way around that  $\lambda_A$  is the row  $A$  is less than 1 spectral radius

is less than 1. Then all of them will go to 0 and finally we will have this matrix, 0 matrix and when we multiply so everything will become a 0 matrix.

(Refer Slide Time: 07:10)

**Lemma:** Let  $A$  be a square matrix. Then  $\lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff  $\rho(A) < 1$

**Sketch of the proof:** Suppose  $A$  is diagonalizable. Then there exist a matrix  $P$  such that

$$A = PDP^{-1}$$

Where  $D$  is a diagonal matrix having the eigenvalues of  $A$  on the diagonal. Therefore

$$A^m = PD^mP^{-1}$$

with  $D = \begin{bmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 0 & \lambda_n^m \end{bmatrix}$

$\Rightarrow \lim_{m \rightarrow \infty} A^m = \mathbf{0}$  iff all the eigenvalues satisfy  $|\lambda_i| < 1$  ( $\rho(A) < 1$ )

So, that is the condition that and say we have a, if and only if condition that if all the eigenvalues are less than 1 in absolute value, then this limit A power m will go to 0.

(Refer Slide Time: 07:19)

The iterative methods  $x^{(k+1)} = Gx^{(k)} + Hb$  converge if and only if the spectral radius satisfies  $\rho(G) < 1$ .

**Sketch of the proof:** The error is given by  $e_k = x^{(k)} - x$

$\Rightarrow e_k = (Gx^{(k-1)} + Hb) - (Gx + Hb)$

$\Rightarrow e_k = G(x^{(k-1)} - x) \Rightarrow e_k = Ge_{k-1} \Rightarrow e_k = G^k e_0$

$e_1 = Ge_0$   
 $e_2 = Ge_1 = GGe_0$   
 $e_k = G^k e_0$

And now, using this result we will come the main result of this lecture that the iterative methods which have this general form converge if and only if the spectral radius, satisfies that means the spectral radius is less than 1. So, now we will prove this result, remember from the first lecture of this module the error is given by  $e_k$  as  $x_k$  minus  $x$ . So,  $x_k$  we denote by the approximate values after  $k$  th iteration and  $x$  is the exact value for this system  $Ax$  is equal to  $b$ .

So, having this  $e_k$  the error at the  $k$ th iteration, we have  $x_k$  by the method here by the iterative method. We can replace this by  $Gx^{(k-1)} + Hb$  because the method is  $Gx$  plus  $Hb$  and we will get this  $x_k$  plus 1. But here we have  $x_k$ . So, we have replaced here  $x_k$  minus 1. The value from the previous iteration and  $x$  since we have this set up of this iteration were done from  $x$  is equal to  $Gx$  and plus this  $Hb$ .

That means this actual  $x$  satisfies this equation and this was just a reformulation of this  $Ax$  is equal to  $b$  equation. So, here  $x$  we can replace also by  $Gx$  plus  $Hb$  and this  $Hb$  and  $Hb$  get cancelled and at the end this error  $e_k$  we can write down as  $G$  and this error at  $k$  minus 1th step. So, which is written here  $e_k$  is equal to  $Ge_{k-1}$ . So,  $e_k$  is equal to  $G$  power  $k$   $e_0$  we can get from this by induction. So, for instance this  $e_1$  is  $Ge_0$  and this  $e_2$  will be  $Ge_1$  and then  $e_1$  we can replace again by  $Ge_0$ . So, we have  $G^2 e_0$  as our  $e_2$ .

(Refer Slide Time: 09:41)

The iterative methods  $x^{(k+1)} = Gx^{(k)} + Hb$  converge if and only if the spectral radius satisfies  $\rho(G) < 1$ .

Sketch of the proof: The error is given by  $e_k = x^{(k)} - x$

$$\Rightarrow e_k = (Gx^{(k-1)} + Hb) - (Gx + Hb)$$

$$\Rightarrow e_k = G(x^{(k-1)} - x) \Rightarrow e_k = Ge_{k-1} \Rightarrow e_k = G^k e_0$$

The slide also features a video inset of a lecturer in the bottom right corner and the NPTEL logo in the bottom left corner.

So, similarly we can proceed and we can get that  $e_k$  is nothing but  $G$  power  $k$  and  $e_0$ . So, the error at  $k$ th step,  $k$ th iteration will be equal to  $G$  power  $k$  and the initial error  $e_0$ .



(Refer Slide Time: 09:51)

The iterative methods  $x^{(k+1)} = Gx^{(k)} + Hb$  converge if and only if the spectral radius satisfies  $\rho(G) < 1$ .

**Sketch of the proof:** The error is given by  $e_k = x^{(k)} - x$

$$\Rightarrow e_k = (Gx^{(k-1)} + Hb) - (Gx + Hb)$$

$$\Rightarrow e_k = G(x^{(k-1)} - x) \Rightarrow e_k = Ge_{k-1} \Rightarrow e_k = G^k e_0$$

$$\Rightarrow \lim_{k \rightarrow \infty} e_k = 0 \text{ for any } e_0 \text{ if and only if } \rho(G) < 1$$

And we have this result that this result will go to 0. As this iteration goes to infinity  $k$  goes to infinity for any  $e$  not if and only if the spectral radius is, the spectral radius of  $G$  is less than 1. So, if the spectral radius of this  $G$  is less than 1 than this error will go to 0 as,  $k$  goes to infinity and that is actually the convergence that this error goes, error between the approximate solution and the numerical solution and the exact solution goes to 0 as  $k$  goes to infinity.

(Refer Slide Time: 10:30)

The iterative methods  $x^{(k+1)} = Gx^{(k)} + Hb$  converge if and only if the spectral radius satisfies  $\rho(G) < 1$ .

**Sketch of the proof:** The error is given by  $e_k = x^{(k)} - x$

$$\Rightarrow e_k = (Gx^{(k-1)} + Hb) - (Gx + Hb)$$

$$\Rightarrow e_k = G(x^{(k-1)} - x) \Rightarrow e_k = Ge_{k-1} \Rightarrow e_k = G^k e_0$$

$$\Rightarrow \lim_{k \rightarrow \infty} e_k = 0 \text{ for any } e_0 \text{ if and only if } \rho(G) < 1$$

**Remark:** If the spectral radius of  $G$  is small, then the convergence is rapid and if the radius of  $G$  is close to unity then convergence is very slow.

So, we have proved this result if the spectral radius of  $G$  is small, that is also implication from here, if the spectral radius of  $G$  is small then the convergence is rapid because the error is written in terms of  $G$  power  $k$   $e$  naught.

So, if the spectral radius of  $G$  is small then this convergence will be faster because at the end we have seen in the previous result that this lambdas will be powered by this  $k$  and if lambdas are very small then this that matrix will go to 0 faster.

(Refer Slide Time: 11:11)

The iterative methods  $x^{(k+1)} = Gx^{(k)} + Hb$  converge if and only if the spectral radius satisfies  $\rho(G) < 1$ .

**Sketch of the proof:** The error is given by  $e_k = x^{(k)} - x$

$$\Rightarrow e_k = (Gx^{(k-1)} + Hb) - (Gx + Hb)$$

$$\Rightarrow e_k = G(x^{(k-1)} - x) \Rightarrow e_k = Ge_{k-1} \Rightarrow e_k = G^k e_0$$

$$\Rightarrow \lim_{k \rightarrow \infty} e_k = 0 \text{ for any } e_0 \text{ if and only if } \rho(G) < 1$$

**Remark:** If the spectral radius of  $G$  is small, then the convergence is rapid and if the radius of  $G$  is close to unity then convergence is very slow.

IT Khanna  
NPTEL

So, here the spectral radius of  $G$  is small, the convergence will be rapid and if the radius of  $G$  this is close to unity then the convergence will be naturally slow.

(Refer Slide Time: 11:22)

**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|, \forall \lambda \in \mathbb{R}$

IT Khanna  
NPTEL

Now, to have some more sufficient conditions, we need to go through the vector and the matrix norm again. So, the matrix norm were defined earlier as well but the direct formula was given. So, here we will go a bit more into the detail of matrix and the vector norm. So, first let me introduce you the vector norm quickly.

So, suppose we have two vectors  $x$  and  $y$  from  $\mathbb{R}^n$ . So, the norm of a vector again it is a number that measures the size of the vector or length of a vector and the properties of a vector norm are given as follows. So, the vector norm must satisfy that whenever we have a nonzero vector, the norm of this must be greater than 0. So, there are various ways or various norms we can define and they will have all these properties.

So, and the norm of  $x$  will be 0 when we are talking about the 0 vector. Another nice property is that will when we multiply by some real number to the vector. That will be equal to the absolute value of this lambda times this norm of the vector.

(Refer Slide Time: 12:36)

**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|$ ,  $\forall \lambda \in \mathbb{R}$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$

And the third property we have that this is called triangular inequality that if we add two vector here and take the norm that will be less than or equal to the sum of the norms of  $x$  and  $y$ .

(Refer Slide Time: 12:46)

**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|, \forall \lambda \in \mathbb{R}$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$

**Matrix Norm:** Let  $A, B \in \mathbb{R}^{n \times n}$ . Similar to vector norm, matrix norm also satisfies the following properties

- (i)  $\|A\| > 0$  for  $A \neq 0$  and  $\|A\| = 0$  for  $A = 0$
- (ii)  $\|\lambda A\| = |\lambda| \|A\|, \forall \lambda \in \mathbb{R}$
- (iii)  $\|A + B\| \leq \|A\| + \|B\|$

**Note:** For any vector norm, we can also define a corresponding matrix norm (called induced matrix norm) as

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

The slide also features a small video inset of a man speaking and logos for IIT Kharagpur and NPTEL at the bottom.

So, coming to the matrix norm we have a similar kind of definition or the properties of the matrix norm as well. So, suppose A B are two matrices of order this n cross n and similar to the vector norm we already discuss before the matrix norm also satisfies the following properties.

So, here the matrix norm will be positive whenever A is nonzero matrix and when A is a 0 matrix then we have here the matrix norm will be equal to 0 and similarly second property also that the lambda absolute value lambda can be taken out and then the norm of A and the triangular inequality also follows for matrix norm.

And there is a note here that for any vector norm. So, given a vector norm, we can also define a corresponding matrix norm which is called actually the induced matrix norm and in this way we can define that. So, if so this is A is a matrix and x is a vector. So, Ax is also a vector.

(Refer Slide Time: 13:53)

**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|$ ,  $\forall \lambda \in \mathbb{R}$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$

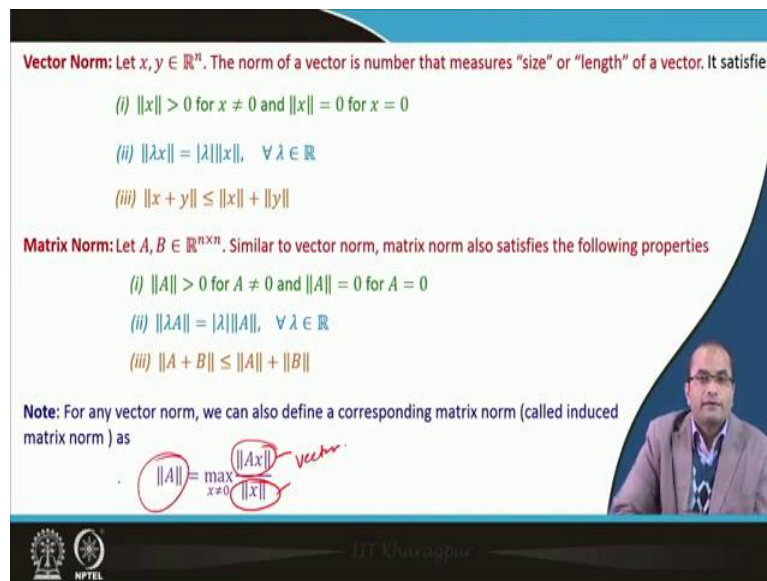
**Matrix Norm:** Let  $A, B \in \mathbb{R}^{n \times n}$ . Similar to vector norm, matrix norm also satisfies the following properties

- (i)  $\|A\| > 0$  for  $A \neq 0$  and  $\|A\| = 0$  for  $A = 0$
- (ii)  $\|\lambda A\| = |\lambda| \|A\|$ ,  $\forall \lambda \in \mathbb{R}$
- (iii)  $\|A + B\| \leq \|A\| + \|B\|$

**Note:** For any vector norm, we can also define a corresponding matrix norm (called induced matrix norm) as

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

*Handwritten note: "vector" with an arrow pointing to the numerator  $\|Ax\|$  in the formula above.*



**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|$ ,  $\forall \lambda \in \mathbb{R}$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$

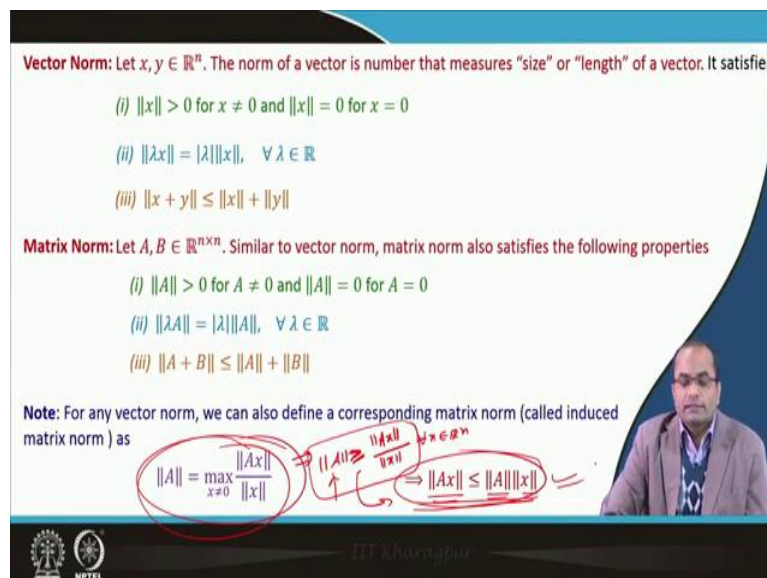
**Matrix Norm:** Let  $A, B \in \mathbb{R}^{n \times n}$ . Similar to vector norm, matrix norm also satisfies the following properties

- (i)  $\|A\| > 0$  for  $A \neq 0$  and  $\|A\| = 0$  for  $A = 0$
- (ii)  $\|\lambda A\| = |\lambda| \|A\|$ ,  $\forall \lambda \in \mathbb{R}$
- (iii)  $\|A + B\| \leq \|A\| + \|B\|$

**Note:** For any vector norm, we can also define a corresponding matrix norm (called induced matrix norm) as

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

*Handwritten notes: "vector" with an arrow pointing to the numerator  $\|Ax\|$ ; "x ∈ R^n" with an arrow pointing to the denominator  $\|x\|$ ; and the inequality  $\|Ax\| \leq \|A\| \|x\|$  with an arrow pointing from the definition to it.*



So, if this vector norm is defined here also this vector norm. So, these two are like vector norm and this A is a matrix. So, here we have the matrix norm. So, we have defined here the matrix norm equal to maximum for all these x we have taken for this Ax the norm of this vector and the x this vector norm of this x.

So, this is also a definition for the matrix norm and one can show that this matrix norm fulfill all these properties or satisfies all these properties. Having this definition there we can readily see that we have this inequality for matrix vector norm. So, here we have Ax, the vector norm there and this is the matrix norm which we have defined into this vector norm of this x.

So, how this is coming, this is trivial from here because A is defined as the maximum over this x. So, if we take A there that will be always greater than of this quantity Ax norm divide

by this x norm for any other x or any x for any x for all x in  $\mathbb{R}^n$ . This will be true because this norm of A is defined as the maximum over this ratio and we are talking about any x then there this will be always the norm will be greater than or equal and from there this inequality follows that Ax norm will be less than equal to norm A and norm x

(Refer Slide Time: 15:41)

**Vector Norm:** Let  $x, y \in \mathbb{R}^n$ . The norm of a vector is number that measures "size" or "length" of a vector. It satisfies

- (i)  $\|x\| > 0$  for  $x \neq 0$  and  $\|x\| = 0$  for  $x = 0$
- (ii)  $\|\lambda x\| = |\lambda| \|x\|, \forall \lambda \in \mathbb{R}$
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$

**Matrix Norm:** Let  $A, B \in \mathbb{R}^{n \times n}$ . Similar to vector norm, matrix norm also satisfies the following properties

- (i)  $\|A\| > 0$  for  $A \neq 0$  and  $\|A\| = 0$  for  $A = 0$
- (ii)  $\|\lambda A\| = |\lambda| \|A\|, \forall \lambda \in \mathbb{R}$
- (iii)  $\|A + B\| \leq \|A\| + \|B\|$

**Note:** For any vector norm, we can also define a corresponding matrix norm (called induced matrix norm) as

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \Rightarrow \|Ax\| \leq \|A\| \|x\|$$

Well, so having this now so we have the vector norm which satisfies these property we have matrix norm which satisfy this property and then we have also introduced this another kind of norm which is coming directly from the vector norm given a vector norm we can define also this matrix norm which is called the induced matrix norm.

(Refer Slide Time: 15:58)

**Example of Matrix Norms**

**Frobenius Norm:**  $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$

**Row Sum Norm:**  $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$

**Column Sum Norm:**  $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$

**Convergence Theorems (Iterative Methods for Solving  $Ax = b$ )** **Sufficient Conditions:**

1. If any **norm of iteration matrix G** is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

Example of Matrix Norms

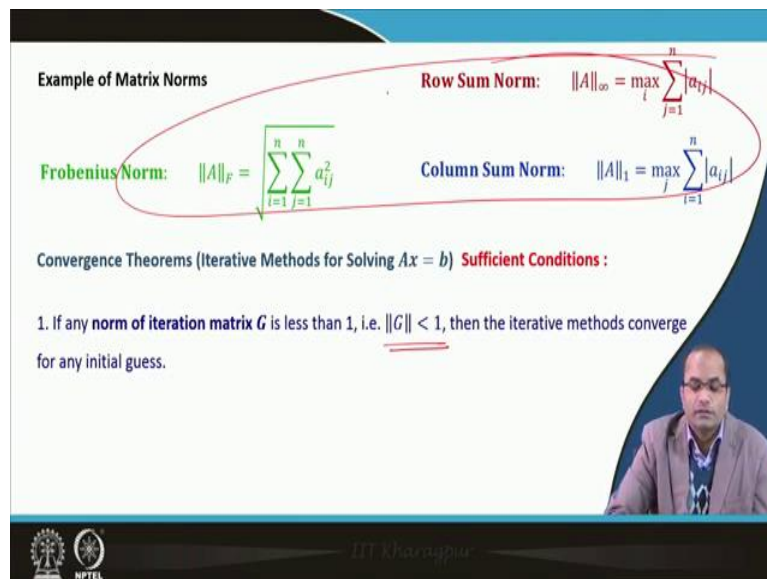
**Frobenius Norm:**  $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$

**Row Sum Norm:**  $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$

**Column Sum Norm:**  $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$

Convergence Theorems (Iterative Methods for Solving  $Ax = b$ ) **Sufficient Conditions :**

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.



We have already seen some examples and these are again I am repeating and all these norms satisfy all these properties which we have mentioned for the norm. One was the Frobenius norm, which was defined as the sum of the square of the entries of the coefficient of A. There was Row Sum Norm which we denote by A infinity it is the, we do take the sum of the rows and then among these sums, we take the maximum one that is called the row sum norm. There is a column sum norm also which is other way around.

So, we take the sum of each column and then with the absolute value of obviously and then we look for the maximum among all these sums and that is our column sum norm. So, all these norms satisfy these properties the triangular one and the other two properties.

Now, we will be talking about the sufficient conditions though we have already discussed a necessary and sufficient conditions. So, those conditions are enough but as we have seen we have to compute the spectral radius of the iteration matrix, which is not trivial always because we have to compute the eigenvalues of the iteration matrix. So, here we will provide some sufficient conditions which could first be checked and if the given matrix satisfy the sufficient conditions then we do not have to compute the eigenvalues and we can talk about the convergence.

But if these sufficient conditions are not met for some we have to compute the spectral radius because from there we will get exactly the precisely the idea whether, the iterative method will converge or it will diverge indeed. So, these are the two results if for any norm of iteration matrix G. So, again we have to compute in this sufficient matrix at least the iteration matrix and if we check that any norm is less than 1 than the iteration matrix will converge for

any initial guess. So, as we have seen few examples of the norms here and there are many more examples of the norm.

(Refer Slide Time: 18:18)

**Example of Matrix Norms**

**Frobenius Norm:**  $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$

**Row Sum Norm:**  $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$

**Column Sum Norm:**  $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$

**Convergence Theorems (Iterative Methods for Solving  $Ax = b$ ) Sufficient Conditions :**

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

Dr. Khanna

So, if any norm here of  $G$  is less than 1 this is what, in some norms may not be less than 1 but if we check the other one it may be less than 1. So, if we find that any norm is less than 1 then we can say that the iterative method will converge this is one of the sufficient conditions we have but the sufficient conditions means if we find that norm is not less than 1 than we cannot say anything we have to actually go for the necessary and sufficient conditions that means the evaluation of the eigenvalues of this  $G$ .

(Refer Slide Time: 18:50)

**Example of Matrix Norms**

**Frobenius Norm:**  $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$

**Row Sum Norm:**  $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$

**Column Sum Norm:**  $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$

**Convergence Theorems (Iterative Methods for Solving  $Ax = b$ ) Sufficient Conditions :**

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

2. If  $A$  is strictly diagonally dominant by rows (or by columns) then the Jacobi and Gauss-Siedel methods converge for any initial guess.

Dr. Khanna

The second sufficient condition we have that if  $A$  is strictly diagonally dominant and this is the best one which we should check first because here we are talking about  $A$  the coefficient



matrix which is given directly in the system of linear equations. So, if A is strictly diagonally dominant by rows or by column then we can say the Jacobi and Gauss Siedel methods converge for any initial guess.

So, the first we will check this sufficient conditions if we have, if we observe that A is strictly diagonally dominant by rows or by column then we do not have to check for any other sufficient condition or necessary and insufficient condition because that itself is a sufficient condition for the convergence.

If this if we fails here the matrix is not the matrix A is not strictly diagonally dominant then we have to check the norm of the iteration matrix G and if we find any norm is less than 1 again we can ensure that the iteration methods will converge for any initial guess and if again we fail we are not able to check this one then we have to perhaps go for the eigenvalues of the iteration matrix where we will exactly know whether it will converge or it will diverge.

(Refer Slide Time: 20:11)

1. If any norm of iteration matrix G is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

Note that  $Gx = \lambda x \Rightarrow \|Gx\| = \|\lambda x\| = |\lambda| \|x\| \leq \|G\| \|x\|$   
 $\Rightarrow |\lambda| \|x\| \leq \|G\| \|x\|$


So, just to go through how these sufficient conditions are coming for instance this norm of reiteration matrix G is less than 1 and the iteration matrix converge for initial guess, we can take a quick look on this. So, if we note that if lambda is the eigenvalue of the matrix G then we have this relation Gx is equal to lambda x and if we take the norm both the sides we have lambda x form here and then we have Gx we have taken the norm on both the sides.


Then this property of the norm we have used that the lambda absolute value times this norm of x the right hand side this Gx we have used that inequality that this is less than G and x so this inequality is used here .

(Refer Slide Time: 21:01)

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.


Note that  $Gx = \lambda x \Rightarrow \|\lambda x\| = \|Gx\|$


$$\Rightarrow |\lambda| \|x\| \leq \|G\| \|x\|$$
$$\Rightarrow |\lambda| \leq \|G\|$$


 IIT Kharyagpur

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.


Note that  $Gx = \lambda x \Rightarrow \|\lambda x\| = \|Gx\|$


$$\Rightarrow |\lambda| \|x\| \leq \|G\| \|x\|$$
$$\Rightarrow |\lambda| \leq \|G\|$$


 IIT Kharyagpur

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

Note that  $Gx = \lambda x \Rightarrow \|\lambda x\| = \|Gx\|$

$$\Rightarrow |\lambda| \|x\| \leq \|G\| \|x\|$$
$$\Rightarrow |\lambda| \leq \|G\|$$


 IIT Kharyagpur

So, we have this result now that the absolute value lambda norm of x is less than equal to norm of G into norm of x and this implies because here we have norm of x there also we have norm of x which is positive for nonzero x. So, this can be cancelled out and then we have actually the absolute value of lambda is always less than any norm of this iteration matrix G and that exactly gives the sufficient condition because if norm is less than 1, and we know that all the eigenvalues are less than the norm then definitely all the eigenvalues are less than 1.

If norm is less than 1, if norm itself is less than 1 and we know that all the eigenvalues must be less than the norm. So, naturally all the eigenvalues will be less than 1 and then we have those necessary sufficient conditions which says that it will converge the iteration matrix the iteration scheme will converge.

(Refer Slide Time: 22:07)

1. If any norm of iteration matrix  $G$  is less than 1, i.e.  $\|G\| < 1$ , then the iterative methods converge for any initial guess.

Note that  $Gx = \lambda x \Rightarrow \|\lambda x\| = \|Gx\|$

$$\Rightarrow |\lambda| \|x\| \leq \|G\| \|x\|$$

$$\Rightarrow |\lambda| \leq \|G\|$$

$$\Rightarrow \rho(G) \leq \|G\|$$

✓ It clearly shows that if  $\|G\| < 1$  then spectral radius  $\rho(A) < 1$ .


NPTEL Dr. Khosla

So, here we have that row G is less than equal to this matrix norm and we have freedom of taking any matrix norm here. So, it clearly shows that this if the norm is less than G the norm of this matrix G is less than 1 than the spectral radius will be less than 1 and hence it follows the convergence. So, this is the sufficient condition for checking the convergence of the iteration matrix of the iterative scheme.

(Refer Slide Time: 22:37)

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.


Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L + U)x^{(k)} + D^{-1}b$



IIIT Kharagpur  
NPTEL

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L + U)x^{(k)} + D^{-1}b = -D^{-1}(A - D)x^{(k)} + D^{-1}b$



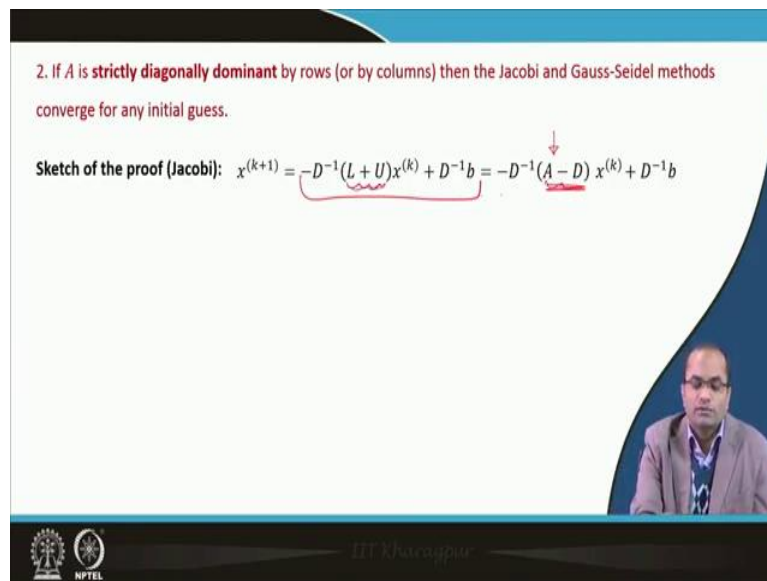
IIIT Kharagpur  
NPTEL

The second one we have discussed, that if  $A$  is strictly diagonally dominant by rows then also then also this Gauss and Jacobi methods convergence for any initial guess this we can also take a look. Here so just consider for instance the Jacobi one again we can do similar steps for the Gauss Siedel one. So, if we take the Jacobi one  $x_k$  plus 1 is equal to minus  $D$  inverse  $L$  plus  $U$   $x_k$  plus we have  $D$  inverse  $b$ . This was the scheme for written matrix form and what we have done now this right hand side part here.

(Refer Slide Time: 23:18)

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$



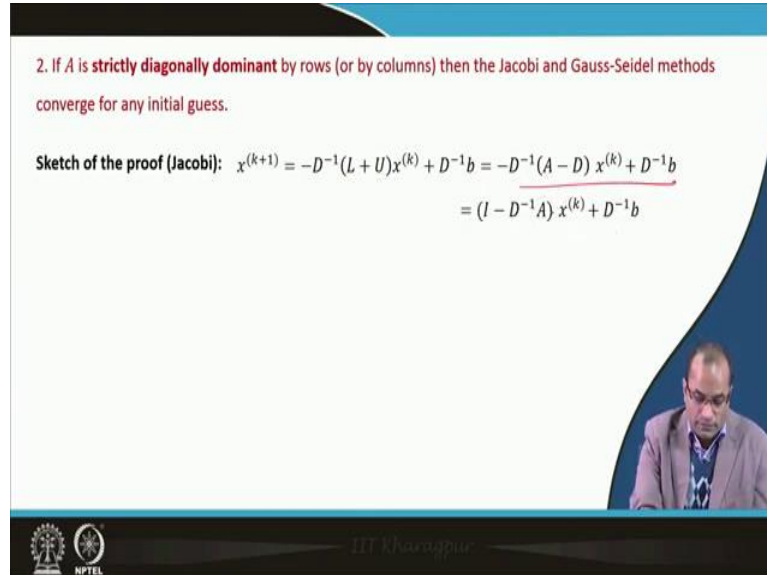
NPTEL IIT Kharagpur

We have written this L plus U because L plus U is the lower triangular matrix plus the upper triangular matrix. If we add the two it is nothing but A minus D, we can write this L plus U as A minus D. A is the complete matrix and if we remove the diagonal entries we put instead of this diagonal entries at 0 now. So, in A that is actually L plus U or A minus D here.

(Refer Slide Time: 23:45)

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.


Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$





NPTEL IIT Kharagpur

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = \underline{-D^{-1}(A-D)} x^{(k)} + D^{-1}b$   
 $= \underline{(I - D^{-1}A)} x^{(k)} + D^{-1}b$



IIT Kharagpur


So, having this now we can again here take this D minus D and here also we have D. So, that becomes I there with minus minus plus, minus this D inverse A this product there and xk plus D inverse b.

(Refer Slide Time: 24:05)



2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

Sketch of the proof (Jacobi):  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D) x^{(k)} + D^{-1}b$   
 $= \underline{(I - D^{-1}A)} x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1$  (Row sum norm)



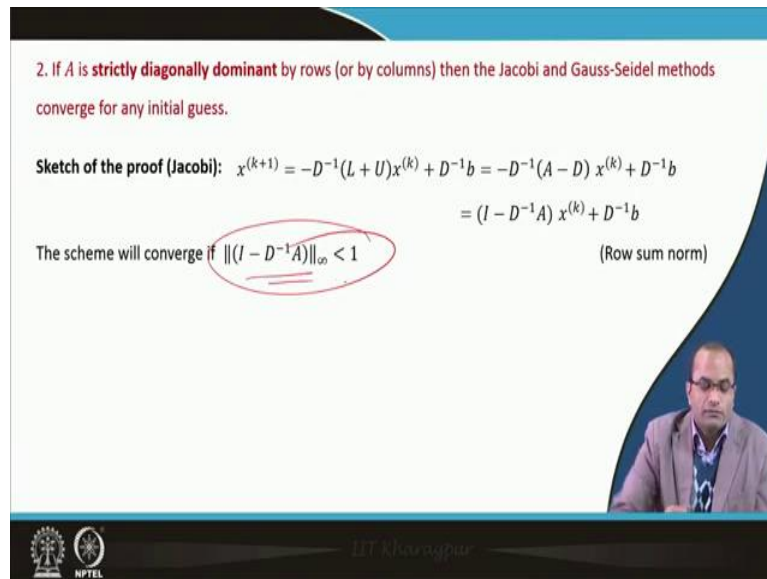
IIT Kharagpur

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

**Sketch of the proof (Jacobi):**  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1$  (Row sum norm)



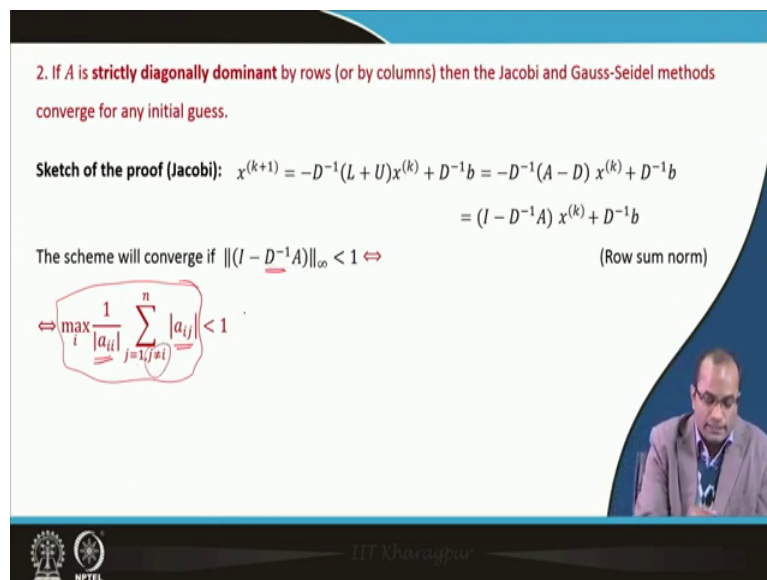
So, we have this iteration matrix here. So, the scheme will converge we know the sufficient condition which we have discussed just before that any norm is less than 1. So, the scheme will converge if this norm for instance we have taken the infinity norm which is the row sum norm if this row sum norm is less than 1 then we know that it is sufficient condition for the convergence meaning the Jacobi method will converge.

(Refer Slide Time: 24:37)

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

**Sketch of the proof (Jacobi):**  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1 \Leftrightarrow$  (Row sum norm)

$$\Leftrightarrow \max_i \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1$$


2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

**Sketch of the proof (Jacobi):**  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1 \Leftrightarrow$  (Row sum norm)

$\Leftrightarrow \max_i \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1$

So, this norm less than 1 meaning, if and only if so this is the definition of the, definition of this norm. So, 1 over a ii and j goes from 1 to n and then we have aij. So, because of this D inverse we have this 1 over a ii term and since i minus. So, that will take care when j is not equal to i. So, this is less than 1. This is what given here that this row some norm so this is exactly, we are taking the maximum because this is the row sum norm maximum for each i we are computing this sum here and that should be less than 1.

(Refer Slide Time: 25:25)

2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

**Sketch of the proof (Jacobi):**  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1 \Leftrightarrow$  (Row sum norm)

$\Leftrightarrow \max_i \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1 \Leftrightarrow \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1, \forall i$



2. If  $A$  is **strictly diagonally dominant** by rows (or by columns) then the Jacobi and Gauss-Seidel methods converge for any initial guess.

**Sketch of the proof (Jacobi):**  $x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + D^{-1}b = -D^{-1}(A-D)x^{(k)} + D^{-1}b$   
 $= (I - D^{-1}A)x^{(k)} + D^{-1}b$

The scheme will converge if  $\|(I - D^{-1}A)\|_{\infty} < 1 \Leftrightarrow$  (Row sum norm)

$$\Leftrightarrow \max_i \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1 \Leftrightarrow \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1, \forall i \Leftrightarrow \sum_{j=1, j \neq i}^n |a_{ij}| < |a_{ii}|, \forall i$$

If  $A$  is diagonally dominant by rows then  $\|(I - D^{-1}A)\|_{\infty} < 1$  and hence Jacobi method will converge. Similarly one can prove if  $A$  is diagonally dominant by columns. Convergence of Gauss-Seidel also follows similar steps.

So, if this is true that for each  $i$  this is less than 1, then we have a that for all  $i$  this should be true because of the maximum is less than 1 than for each row this should be true here that this should be less than 1 and that implies that this summation here  $a_{ij}$  is strictly less than  $a_{ij}$  and this is precisely the condition what we call is strictly diagonally dominant.

So, if  $A$  is strictly diagonally dominant by row that means here we have seen that this norm will be less than 1 and from the previous result we know that the method will converge and similarly we can prove if  $A$  is diagonally dominant by columns in that case we have to talk about this 1 norm and then the convergence of the Gauss Siedel can be also proved similarly.

(Refer Slide Time: 26:19)

**Example:** Consider the following system of equations

$$\begin{cases} 5x + y + 2z = 13 \\ x + 3y + z = 12 \\ -x + 2y + 4z = 8 \end{cases}$$

Discuss the convergence of the Jacobi and Gauss-Seidel methods?

The coefficient matrix is strictly diagonally dominant by rows and hence both the methods will converge for any initial guess.

$A = \begin{bmatrix} 5 & 1 & 2 \\ 1 & 3 & 1 \\ -1 & 2 & 4 \end{bmatrix}$


$5 > 1+2$   
 $3 > 1+1$   
 $4 > |-1|+2$

**Example:** Consider the following system of equations


$$\begin{aligned}5x + y + 2z &= 13 \\x + 3y + z &= 12 \\-x + 2y + 4z &= 8\end{aligned}$$

Discuss the convergence of the Jacobi and Gauss-Seidel methods?

The coefficient matrix is strictly diagonally dominant by rows and hence both the methods will converge for any initial guess.



IT Khanna

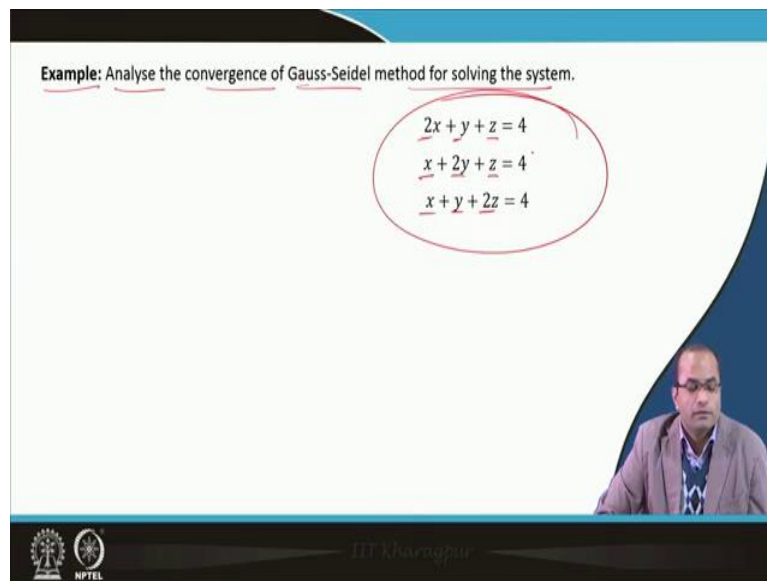


Now, we will go through quickly some of the examples we will analyze that what are the possibilities for the convergence for instance if we talk about the system which we have solved already in previous lecture and we want to discuss only the convergence of this two methods. So, if looking at this matrix A the coefficient matrix we have 5, 1, 2 then we have then we have 1, 3, 1 then we have minus 1, 2 and 4.

This is the coefficient matrix A, it clearly one can see that it is diagonally dominant by rows because this 5 is strictly greater than 1 plus 2 in the second rows says that 3 is greater than 1 plus 1 this 4, third one says that 4 is greater than again minus 1 absolute value, that is 1 and plus 2. So, here is 3 and here is 4. So, it is a strictly diagonally dominant and hence both the methods will converge, the Gauss Jacobi method and Gauss Siedel method will also converge.

(Refer Slide Time: 27:24)

Example: Analyse the convergence of Gauss-Seidel method for solving the system.

$$\begin{aligned} 2x + y + z &= 4 \\ x + 2y + z &= 4 \\ x + y + 2z &= 4 \end{aligned}$$


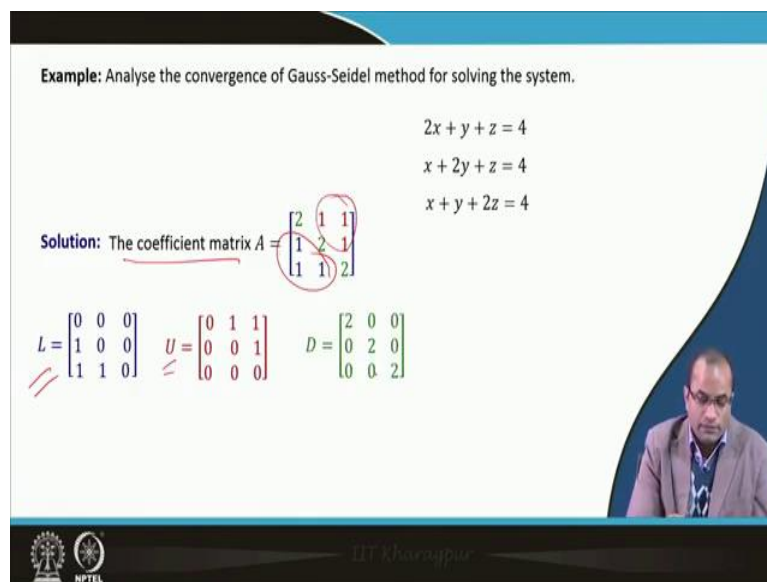
Now here we analyze the convergence of Gauss Seidel method for solving this system of equation. So, if we take a look at this system where we have coefficient matrix 2, 1, 1 and 1, 2, 1 and 1, 1, 2. So it is not diagonally dominant neither by rows or nor by column. So, we do not have that diagonal dominance here. So, then we have to either... We have to construct the iteration matrix in that case if you do not see that the matrix is diagonally dominant.

(Refer Slide Time: 28:02)

Example: Analyse the convergence of Gauss-Seidel method for solving the system.

$$\begin{aligned} 2x + y + z &= 4 \\ x + 2y + z &= 4 \\ x + y + 2z &= 4 \end{aligned}$$

Solution: The coefficient matrix  $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$


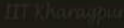
$$L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad U = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$


**Example:** Analyse the convergence of Gauss-Seidel method for solving the system.

$$\begin{aligned} 2x + y + z &= 4 \\ x + 2y + z &= 4 \\ x + y + 2z &= 4 \end{aligned}$$

**Solution:** The coefficient matrix  $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

$$L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad U = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$


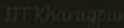
$$G_{GS} = -(L+D)^{-1}U = -\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$



**Example:** Analyse the convergence of Gauss-Seidel method for solving the system.

$$\begin{aligned} 2x + y + z &= 4 \\ x + 2y + z &= 4 \\ x + y + 2z &= 4 \end{aligned}$$

**Solution:** The coefficient matrix  $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

$$L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad U = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$G_{GS} = -(L+D)^{-1}U = -\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = -\frac{1}{8} \begin{bmatrix} 4 & 0 & 0 \\ -2 & 4 & 0 \\ -1 & -2 & 4 \end{bmatrix}$$



So, in that case we have our coefficient matrix is here, we have the lower triangular matrix we have the upper triangular matrix and then we have the diagonal matrix taking these diagonal entries as 2, 2, 2 and the rest 0 there. Then Gauss Siedel we have we know the iteration matrix that is L plus D inverse U. So, we can compute that we need to compute this inverse here and then we have to multiply, finally we got this matrix which is a iteration matrix.

(Refer Slide Time: 28:36)

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\}$$

IT Khargpur  
NPTEL

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$

IT Khargpur  
NPTEL

So, having this iteration matrix, for the Gauss Seidel method what we observe let us compute the one norm that is the column sum norm. So, here we have 0 there and if we add this so we have here 1 then we 2 there. So, 3 and then 4, 7. So, 7 by 8 the second one and from the third one if we check this is 9 by 8. So, the maximum of this when we take is 9 by 8 which is greater than 1 so does not work actually for the convergence we it is not less than 1. If would have been less than 1, than we could have declared here that the Gauss Seidel method will converge.

(Refer Slide Time: 29:22)

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\Rightarrow \|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$
$$\Rightarrow \|G_{GS}\|_\infty = \max\{1, 1/2, 1/2\}$$


DT Khosla  
NPTEL

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$
$$\|G_{GS}\|_\infty = \max\{1, 1/2, 1/2\} = 1$$


DT Khosla  
NPTEL

Now, let us check the other one at least these standard norms we can check. So, for the infinity one, the row sum norm that means here we have the sum 1 there here we have sum half and here we have sum again half. So, in this case we are getting this 1 and again we cannot did not declare anything because, if it is less than 1 then only we can say that the method will converge.

(Refer Slide Time: 29:46)

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$
$$\|G_{GS}\|_\infty = \max\{1, 1/2, 1/2\} = 1$$
$$\|G_{GS}\|_F = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} + \frac{9}{64}} = \sqrt{\frac{32+8+10}{64}} = \sqrt{\frac{50}{64}}$$



NPTEL IIT Kharyagpur

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$
$$\|G_{GS}\|_\infty = \max\{1, 1/2, 1/2\} = 1$$
$$\|G_{GS}\|_F = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} + \frac{9}{64}} = \sqrt{\frac{32+8+10}{64}} = \sqrt{\frac{50}{64}} < 1$$


NPTEL IIT Kharyagpur

$$G_{GS} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}$$
$$\|G_{GS}\|_1 = \max\{0, 7/8, 9/8\} > 1$$
$$\|G_{GS}\|_\infty = \max\{1, 1/2, 1/2\} = 1$$
$$\|G_{GS}\|_F = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} + \frac{9}{64}} = \sqrt{\frac{32+8+10}{64}} = \sqrt{\frac{50}{64}} < 1$$

⇒ The Gauss-Siedel method will converge.



NPTEL IIT Kharyagpur

So, let us check the Frobenius norm. So, if we check the Frobenius norm in this case then it is coming as  $\sqrt{50}$  by  $\sqrt{64}$  square root which is less than 1 which is less than 1 and hence now we can say that for this norm this is less than 1. So, at least there is a norm whose value is less than 1 and that is already sufficient because the sufficient condition we have seen in the proof that if norm is less than 1, then all the spectral radius will be also less than 1 and hence it will converge. Therefore the Gauss Seidel method will converge.

(Refer Slide Time: 30:25)

A comparative Study of Jacobi and Gauss-Seidel Method

Consider the following system of linear equations

$$\begin{cases} 5x + y + 2z = 13 \\ x + 3y + z = 12 \\ -x + 2y + 4z = 8 \end{cases}$$

Exact Solution:  $[1.6364 \quad 3.1818 \quad 0.8182]$

Spectral Radius of the Iterative Methods:

Gauss-Seidel Method  $\rho_{GS}$ :  $0.1667$

Jacobi Method  $\rho_J$ :  $0.4860$

A comparative Study of Jacobi and Gauss-Seidel Method

Consider the following system of linear equations

$$\begin{cases} 5x + y + 2z = 13 \\ x + 3y + z = 12 \\ -x + 2y + 4z = 8 \end{cases}$$

Exact Solution:  $[1.6364 \quad 3.1818 \quad 0.8182]$

Spectral Radius of the Iterative Methods:

Gauss-Seidel Method  $\rho_{GS}$ :  $0.1667$

Jacobi Method  $\rho_J$ :  $0.4860$

Now, we will do a quick comparative study of the Gauss Seidel and the Jacobi method by taking the same system which we have solved already in the previous lecture. So, here the exact solution also one can get which is 1.6364, 3.188 and 0.8182. We have also checked the spectral radius for this analysis and the spectral radius for the Gauss Seidel is 0.1667 and for the Jacobi it is 0.4860.



So the spectral radius says that both the methods will converge because the spectral radius are less than 1. What also we see that in the Jacobi Method it is 0.4 and while in the Gauss Seidel it is 0.1 and so on. So, it is smaller for the Gauss Seidel and for Jacobi it is large. So here also they indicate that the Gauss Seidel the convergence will be faster than the Jacobi method at least in this example we can say looking at the spectral radius.

(Refer Slide Time: 31:33)

| Iteration | Jacobi Method        |
|-----------|----------------------|
| 1         | 2.0000 3.3333 1.7500 |
| 2         | 1.2333 2.7500 0.8333 |
| 3         | 1.7167 3.3111 0.9333 |
| 4         | 1.5644 3.1167 0.7736 |
| 5         | 1.6672 3.2206 0.8328 |
| 6         | 1.6228 3.1667 0.8065 |
| ⋮         |                      |
| 13        | 1.6365 3.1819 0.8182 |
| 14        | 1.6363 3.1818 0.8182 |
| 15        | 1.6364 3.1818 0.8182 |
| 16        | 1.6364 3.1818 0.8182 |

Exact Solution:  $[1.6364 \ 3.1818 \ 0.8182]$   $\rho_{GS} = 0.1667$

Iterative Method:  $x^{(k+1)} = Gx^{(k)} + Hb$   $\rho_J = 0.4860$

Initial Guess =  $[1 \ 1 \ 1]$

So, now let us compute here with the initial guess here 1, 1, 1 and what do we see that if we apply Jacobi method the iterations then we are getting here after 14 steps 14 iterations that the values are no more changing here. So 1.6364 which actually the exact solution up to this 4 digit 4 decimal. Here also we have 1, 6, 3, 4. So, there is no change further. So, naturally the convergence has occurred already now after this fourth place number may be changing but at least here we do see that we can stop the iterations here.

(Refer Slide Time: 32:19)

Exact Solution: [1.6364 3.1818 0.8182]      $\rho_{GS} = 0.1667$   
 $\rho_J = 0.4860$

Iterative Method:  $x^{(k+1)} = Gx^{(k)} + Hb$

Initial Guess = [1 1 1]

| Iteration | Jacobi Method               |
|-----------|-----------------------------|
| 1         | 2.0000 3.3333 1.7500        |
| 2         | 1.2333 2.7500 0.8333        |
| 3         | 1.7167 3.3111 0.9333        |
| 4         | 1.5644 3.1167 0.7736        |
| 5         | 1.6672 3.2206 0.8328        |
| 6         | 1.6228 3.1667 0.8065        |
| ⋮         |                             |
| 13        | 1.6365 3.1819 0.8182        |
| 14        | 1.6363 3.1818 0.8182        |
| 15        | <b>1.6364 3.1818 0.8182</b> |
| 16        | 1.6364 3.1818 0.8182        |

| Iteration | Gauss-Seidel Method         |
|-----------|-----------------------------|
| 1         | 2.0000 3.0000 1.0000        |
| 2         | 1.6000 3.1333 0.8333        |
| 3         | 1.6400 3.1756 0.8222        |
| 4         | 1.6360 3.1806 0.8187        |
| 5         | 1.6364 3.1816 0.8183        |
| 6         | <b>1.6364 3.1818 0.8182</b> |
| 7         | 1.6364 3.1818 0.8182        |

After 15 th iteration we got this value and if we apply the Gauss Seidel method for the same initial guess for the same condition what we observe that after 6 iteration itself after 5 iteration we got the same value. So, which we can see here that this Gauss Seidel is much faster than the Jacobi method which was also coming from our analysis because the radius the spectral radius of the Jacobi is smaller of Gauss Seidel is smaller than the Jacobi method.

(Refer Slide Time: 32:52)

Exact Solution: [1.6364 3.1818 0.8182]      $\rho_{GS} = 0.1667$   
 $\rho_J = 0.4860$

Iterative Method:  $x^{(k+1)} = Gx^{(k)} + Hb$

Initial Guess = [0 0 0]

| Iteration | Jacobi Method               |
|-----------|-----------------------------|
| 1         | 2.6000 4.0000 2.0000        |
| 2         | 1.0000 2.4667 0.6500        |
| 3         | 1.8467 3.4500 1.0167        |
| 4         | 1.5033 3.0456 0.7367        |
| 5         | 1.6962 3.2533 0.8531        |
| 6         | 1.6081 3.1502 0.7974        |
| ⋮         |                             |
| 15        | 1.6364 3.1819 0.8182        |
| 16        | 1.6363 3.1818 0.8182        |
| 17        | <b>1.6364 3.1818 0.8182</b> |
| 18        | 1.6364 3.1818 0.8182        |

| Iteration | Gauss-Seidel Method         |
|-----------|-----------------------------|
| 1         | 2.6000 3.1333 1.0833        |
| 2         | 1.5400 3.1256 0.8222        |
| 3         | 1.6460 3.1773 0.8229        |
| 4         | 1.6354 3.1806 0.8186        |
| 5         | 1.6365 3.1817 0.8183        |
| 6         | <b>1.6364 3.1818 0.8182</b> |
| 7         | 1.6364 3.1818 0.8182        |

So, this for initial guess 1, 1, 1. If we change here the initial guess instead of 1, 1 we take 0, 0 nothing much will happen now. Only few more iterations for the Jacobi method and only 6 iterations for the Gauss Seidel method we will get the convergence of the result. So, they are matching up to this 4 decimal places.

(Refer Slide Time: 33:17)

REMARK

Gauss Seidel as by construction seems to be faster than Jacobi method. However this is not true in general.

There are examples where Jacobi converges faster than Gauss Seidel.

If we change the order of the equations the iterative methods may not be convergent.

For example, consider the following system of linear equations

$$\begin{cases} x + 3y + z = 12 \\ -x + 2y + 4z = 8 \\ 5x + y + 2z = 13 \end{cases}$$

The coefficient matrix is:

$$\begin{bmatrix} 1 & 3 & 1 \\ -1 & 2 & 4 \\ 5 & 1 & 2 \end{bmatrix}$$

IIT Kharagpur  
NPTEL

Well so, the remark here the Gauss Seidel as by construction we have seen that because we are using there the recently evaluated value seems to be faster than the Jacobi method which is in general true but this is not true however for all the cases. There are examples where the Jacobi converges faster than the Gauss Seidel method and another remark here that if we change the order of equations the iterative method may not be convergent.


So, it is not that it matters that in which order we have taken the equations because like the earlier system if we change just the order now. So, for instance if we have taken this order now first of all now our coefficient matrix is no more is no more diagonally dominant. So, that is exactly the point here. So, earlier we were able to say it is a diagonally dominant matrix and then definitely it will converge. But now we do see that, we do not have this diagonally dominance character of this matrix hence anything can happen and which is indeed true.

(Refer Slide Time: 34:33)

Exact Solution:  $[1.6364 \quad 3.1818 \quad 0.8182]$  Initial Guess =  $[1 \ 1 \ 1]$

| Iteration | Jacobi Method                               | Gauss-Seidel Method                        |
|-----------|---|--|
| 1         | 8.0000 2.5000 3.5000                        | 8.0000 6.0000 -16.5000                     |
| 2         | 1.0000 1.0000 -14.7500                      | 10.5000 42.2500 -40.8750                   |
| 3         | 23.7500 34.0000 3.5000                      | -73.8750 48.8125 166.7813                  |
| 4         | -93.5000 8.8750 -69.8750                    | -301.2188 -480.1719 999.6328               |
| ⋮         |   |  |
| 10        | $10^4 \times [-2.5925 \ -0.6066 \ -2.9968]$ | $10^6 \times [0.5066 \ -1.4169 \ -0.5580]$ |

$\rho_{GS} = 3.8730$   
 $\rho_J = 2.7233$



That if we check now the spectral radius for this new system the system is saying the exact solution is same for this system. But we have change the order of the equation third equation was placed at the first place and so on but now the spectral radius of this iteration matrix has changed and now we have this 3.87 and 2.72 for the Jacobi.

Which we can see in the iterations also the Jacobi method we have these large numbers after few iterations here also we have the very large numbers and they show the diverging behaviour indeed for Gauss Seidel since this spectral radius was very high larger than the Jacobi one and we are here also getting large number after tenth iteration larger than indeed of the Jacobi method. So, here itself we can see that the diverging behaviour of the numerical scheme just by changing the order of the equations.

(Refer Slide Time: 35:34)

**Remark:** Without analyzing iteration matrix, it is difficult to conclude that which of the methods converges faster.

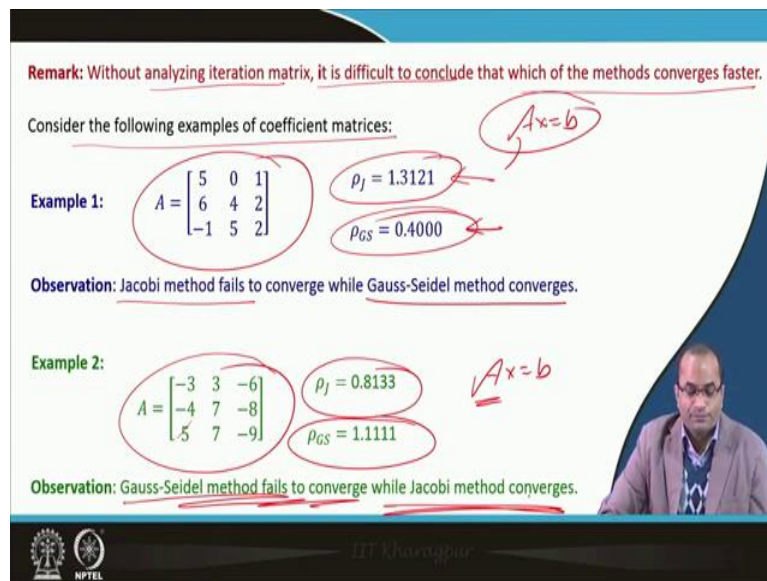
Consider the following examples of coefficient matrices:

**Example 1:**  $A = \begin{bmatrix} 5 & 0 & 1 \\ 6 & 4 & 2 \\ -1 & 5 & 2 \end{bmatrix}$   $\rho_J = 1.3121$   $\rho_{GS} = 0.4000$   $Ax=b$

**Observation:** Jacobi method fails to converge while Gauss-Seidel method converges.

**Example 2:**  $A = \begin{bmatrix} -3 & 3 & -6 \\ -4 & 7 & -8 \\ 5 & 7 & -9 \end{bmatrix}$   $\rho_J = 0.8133$   $\rho_{GS} = 1.1111$   $Ax=b$

**Observation:** Gauss-Seidel method fails to converge while Jacobi method converges.



So, without analysing iteration matrix it is difficult to conclude looking at just the coefficient matrix that which method converges faster we have to compute this iteration matrix which is crucial. Once we have iteration matrix we have the spectral radius than we can tell many things about that method. So, consider following examples for instance of the coefficient matrix. So our system is always  $Ax$  is equals to  $b$  and we are talking about  $A$ .

So, if  $A$  is this one if we compute the spectral radius of this Jacobi it is the 1.312 and for Gauss Seidel it is 0.400 and they indicate that the Jacobi method fails here because the spectral radius is greater than 1 and the Gauss Seidel method converges because the spectral radius is less than 1.

Whereas in this example what do we see now, for this  $A$  spectral radius of Jacobi is 0.8 and the spectral radius of Gauss Seidel is 1.1. So, in this case when we form system of equation with this matrix what we will happen the Gauss Seidel method will fail to converge whereas the Jacobi method converges. So, it is other way around.

(Refer Slide Time: 36:55)

**Example 3:**

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 2 & -9 & 0 \\ 0 & -8 & -6 \end{bmatrix}$$

$\rho_J = 0.4438$   
 $\rho_{GS} = 0.0185$

**Observation:** Jacobi method is more slowly convergent than Gauss-Seidel.

**Example 4:**

$$A = \begin{bmatrix} 7 & 6 & 9 \\ 4 & 5 & -4 \\ -7 & -3 & 8 \end{bmatrix}$$

$\rho_J = 0.6411$   
 $\rho_{GS} = 0.7746$

**Observation:** Gauss-Seidel method is more slowly convergent than Jacobi.

The slide also features the NPTEL logo and the name 'Dr. Dharmapala' at the bottom.

Not only this if we take this coefficient matrix for instance the row J for the Jacobi method the spectral radius is 0.44 whereas for the Gauss Seidel it is 0.01. So, which in tells that both the methods will converge but the Gauss Seidel will converge faster and if we take this example for instance where the spectral radius of this Jacobi 0.64 and for Gauss Seidel it is 0.77. So, it tells that the Jacobi will...the convergence in with the Jacobi method will be faster than the Gauss Seidel method.

So, here also written it is Gauss Seidel method is more slowly convergent than the Jacobi method. So, what we have seen in these four examples in one Jacobi converges, other one Gauss Seidel converges, in one the convergence of the Gauss Seidel is faster whereas in other one convergence of the Jacobi method is faster.

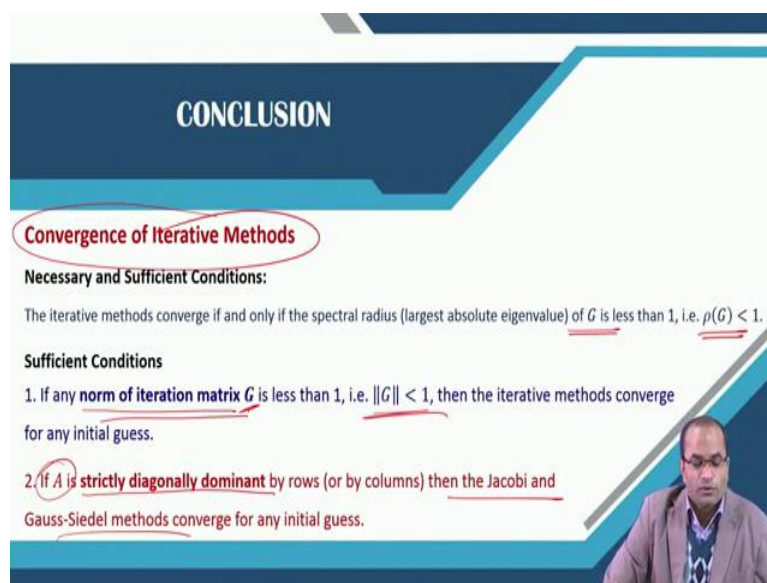
So, in general we cannot tell looking at the coefficient matrix that weather it will converge or not convergence can be seen just if we observe that it is diagonally dominant otherwise we have to compute the spectral radius of the iteration matrix and then we can analyse the speed of the convergence etc.

(Refer Slide Time: 38:22)



So, these are the references used for preparing this lecture.

(Refer Slide Time: 38:28)



And just to conclude that we have discussed a very important topic here that is convergence of the iterative method and the necessary and sufficient conditions we have discussed and seen we seen the proof also and it depends on that we have to compute the spectral radius of this iteration matrix and if the spectral radius is less than 1 than the method will converge, if spectral radius is greater than 1 the method will not converge.

The sufficient conditions for quick check we have also seen that if any norm we can find for this iteration matrix is less than 1 then we can declare that the iteration matrix will converge for any initial guess and for directly looking at A if you find that it is strictly diagonally

dominant by rows then the Jacobi and the Gauss Siedel method will converge for any initial guess. So, that is all for this lecture and I thank you very much for your attention.