Numerical Methods and Computation

Prof. S.R.K. Iyengar

Department of Mathematics

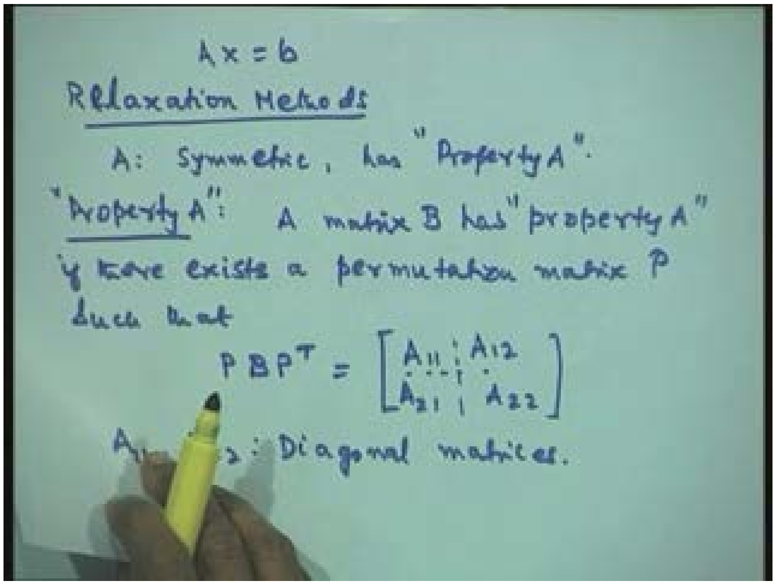
Indian Institute of Technology Delhi

Lecture No # 18

Solution of a System of Linear Algebraic Equations (Continued)

In our last lecture we derived two methods for the solution of the system of linear algebraic equations Ax is equal to b.

(Refer Slide Time: 00:04:03 min)



$Ax = b$

Relaxation Methods

A: Symmetric, has "Property A".

"Property A": A matrix B has "property A" if there exists a permutation matrix P such that

$$PBP^T = \begin{bmatrix} A_{11} & \vdots & A_{12} \\ \cdots & \vdots & \cdots \\ A_{21} & \vdots & A_{22} \end{bmatrix}$$

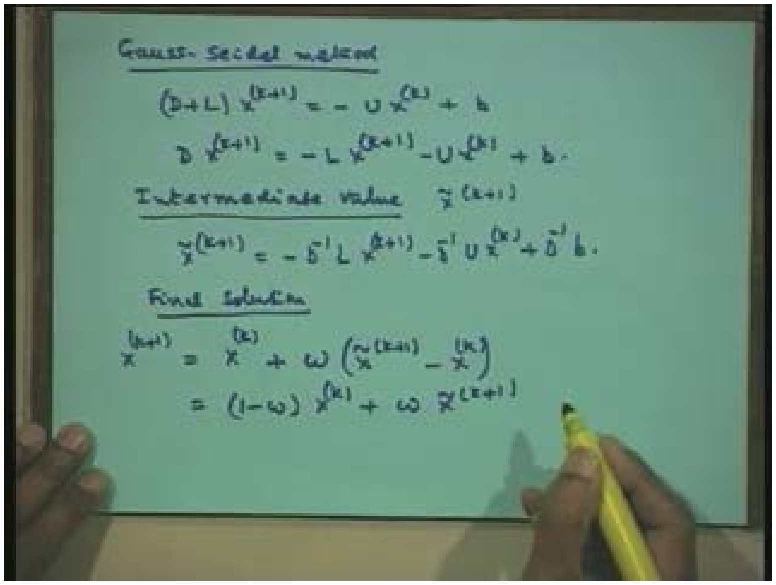
A_{11}, A_{22}: Diagonal matrices.

We have we have mentioned that the Gauss Seidel method is at least two times more faster than the Gauss Jacobi method. However in practical applications when we have thousands or lakhs of equations even the convergence of Gauss Seidel is not sufficient, because it will take too much of time. Therefore we need to develop or need to have some methods which works faster than the Gauss Seidel methods. In this direction we have words known as relaxation methods. So we will discuss what is known as the relaxation method. This is a generalization of the Gauss Seidel method. Originally the method was discovered when the matrix has special properties and that is the matrix A has properties which is symmetric and it is has what is known as property A. It's called what is known as property A. This is property A; given the co efficient matrix A, you must able to use inter changes of rows and interchanges of column, that is apply a permutation matrix similar to transformation PAP transpose similarity transformation; such that it can be partitioned as the these two corners. Top and bottom is diagonal matrices and remaining are square matrices or they are any other matrices which is not diagonal.

Let's use another name matrix B as property A. If there exists a permutation matrix P such that PBP transpose can be partitioned as A_{11} A_{12} A_{21} $A_{22,}$ where A_{11} A_{22} are diagonal matrices, these are diagonal matrices, these are diagonal matrices; what this really implies is we have pre multiplied by P post multiplied by P transpose. Therefore this gives you interchanges of rows and the corresponding interchanges of columns that is if you interchanges column two and five then it will interchange the rows two and five. So it is the corresponding interchanges of this; then I must be able to bring these matrices A_{11} A_{22} as diagonal matrices. Then the matrix is said to have property A, in fact this was done because we will see that the method requires finding the optimal value of a relaxation parameter which gives us fastest convergence for the given method. So that is why we needed this particular formulation to have the analysis of the methods. Therefore without going through whether the matrix property A or now or not even though it is symmetric we use the relaxation methods.

Now let us define what a relaxation method is. We start from the Gauss Seidel method.

(Refer Slide Time: 00:08:29 min)



Gauss- Seidel method

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

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

Intermediate value $\tilde{x}^{(k+1)}$

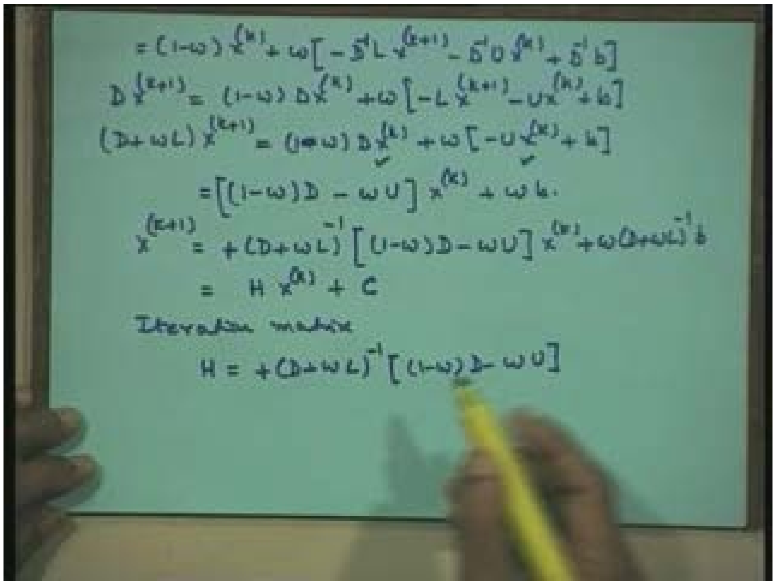
$$\tilde{x}^{(k+1)} = -D^{-1}L\,x^{(k+1)} - D^{-1}U x^{(k)} + D^{-1}b.$$

Final solution

$$x^{(k+1)} = x^{(k)} + \omega\left(\tilde{x}^{(k+1)} - x^{(k)}\right)$$
$$= (1-\omega)\,x^{(k)} + \omega\,\tilde{x}^{(k+1)}$$

We write the Gauss Seidel method as D plus L xk plus one minus U xk plus b. Now let us take the second term Lxk plus one to the right hand side. So I will write Dxk plus one is minus Lxk plus one minus Uxk plus b. Now we shall call this solution of this system as an intermediate value. So let us call it as intermediate value. Let us give a new notation to it. Let's call it is as some x tilda k plus one. Therefore x tilda k plus one is minus D inverse Lxk plus one minus D inverse Uxk plus D inverse of b. Now this intermediate value has really no meaning as it is. We shall use this intermediate value to construct the final solution and hence I mean we will not specifically say that this is going to approximate our solution or anything here. Now we shall use this intermediate value and the previous iterated value xk to build the final solution. So let us call this as a final solution. Let us first take difference between the two values; that difference between the two values will be x tilda k plus one minus xk. Therefore this will be the difference of the intermediate value and this is the value at the previous iteration. I multiply this by a parameter omega. I will call this as a relaxation parameter, the value of which we shall determine later on using the property or the condition that this iteration should converge. And the analysis of the convergence would give us what will be the omega that should be used in order that the solution of the problem can be obtained. Now we shall add this product to the previous iterate xk, this we shall say is our final solution xk plus one. Now we can use this as a correction using the parameter omega to the previous iterated value to get the current estimate. I can simplify the right hand side and write it as one minus omega into xk plus omega x tilda k plus one. Now I want to simplify this further. Therefore I shall substitute the expression for x tilda plus one from here into this particular equation.

(Refer Slide Time: 00:08:45 min)



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

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

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

$$= \left[(1-\omega) D - \omega U\right] x^{(k)} + \omega b.$$

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

$$= H x^{(k)} + C$$

Iteration matrix

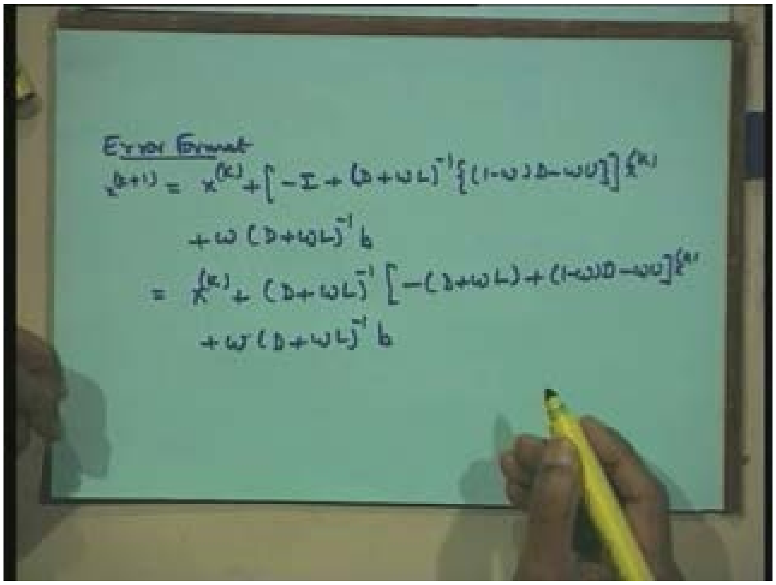
$$H = +(D+\omega L)^{-1}\left[(1-\omega) D - \omega U\right]$$

Therefore I would get here one minus omega xk plus omega into minus D inverse Lxk plus one minus D inverse Uxk plus D inverse b. Before I simplify it further I would like to pre multiply this equation by D, so that I can remove this D inverse from here. Therefore I would get Dxk plus one is one minus omega Dxk plus omega minus Lxk plus one minus Uxk plus b. I mean we have pre multiplied throughout by D, so that this will simplify and become minus L. This is minus U and this is simply b. Now this xk plus one belongs to the left hand side because it is a current iterate, it belongs to the left hand side. So let us bring it to the left hand side, then we will have D plus omega L xk plus one is equal to one minus omega one plus omega one minus omega into Dxk plus omega minus Uxk plus b. Now let us combine this xk term and this xk term. So I will have one minus omega into D minus omega into U into xk plus omega into b. Now the solution of the system is our xk plus one. Therefore let us inverse this matrix and find xk plus one. Therefore xk plus one is minus D plus omega L inverse one minus omega into D minus omega U xk plus omega D plus omega L inverse of b. Now this is our iteration method, therefore we shall write it as Hxk plus some c, where this H is iteration matrix given by this particular expression.

Therefore our iteration matrix H is equal to minus D plus omega L inverse. There is no minus sign here, there is a plus sign. There is a plus sign D plus omega L inverse one minus omega of D minus omega of U.

Now the properties of this iteration matrix, we shall study it later on. If the method is to converge then the Eigen values of this matrix has to be strictly less than one in magnitude. We are going to prove this in our later lectures. Therefore what we would like to say now here is that the behavior of this iteration matrix will tell us whether the iterations are going to converge or it’s not going to converge. However for computation purposes this particular expression is not useful for us because it is going to take a lot of computer time, therefore we would like to use the error format to derive it. Now let us write the error format of this particular method.

(Refer Slide Time: 00:12:45 min)



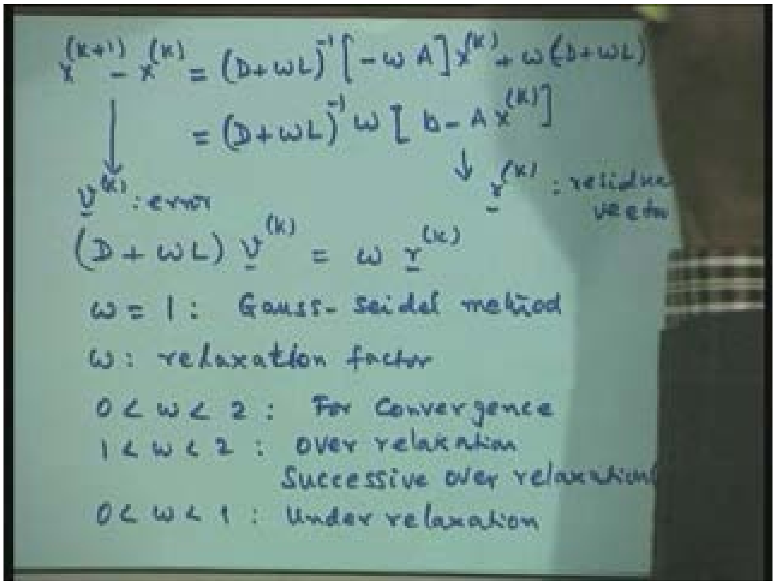
Error Format

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

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

I add and subtract xk on the right hand side, therefore I can write xk plus one is equal to xk plus minus I D plus omega L inverse one minus omega into D minus omega into Uxk plus omega D plus omega L inverse of b. Now I take D plus omega L inverse outside, so I get xk D plus omega L inverse into minus D plus omega L plus one minus omega into D minus omega Uxk plus omega D plus omega L inverse of b.

(Refer Slide Time: 00:14:12 min)



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

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

$\underset{\sim}{v}^{(k)}$: error

$\underset{\sim}{r}^{(k)}$: residue vector

$$(D + \omega L)\, \underset{\sim}{v}^{(k)} = \omega\, \underset{\sim}{r}^{(k)}$$

$\omega = 1$: Gauss- Seidel method

ω : relaxation factor

$0 < \omega < 2$: For convergence

$1 < \omega < 2$: over relaxation

Successive over relaxation

$0 < \omega < 1$: Under relaxation

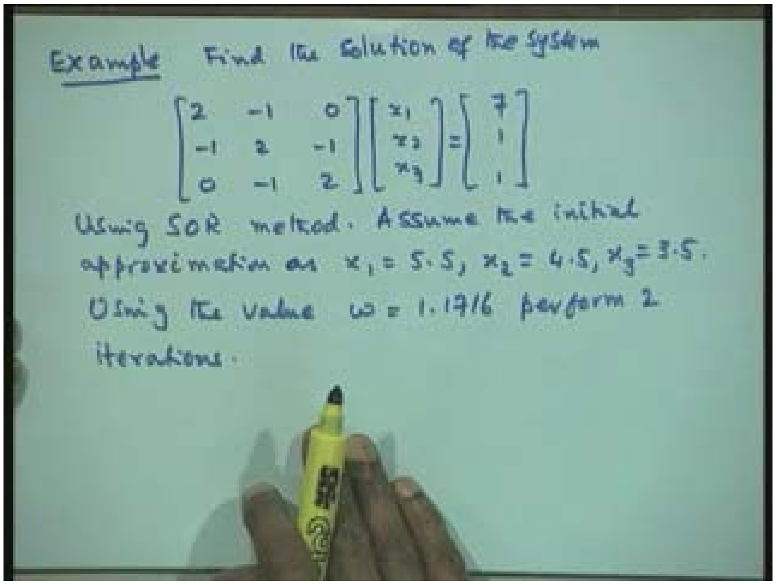
Now let us simplify it further. I can take this xk to the left hand side and write this as xk plus one minus xk. Now if you just look at this particular factor there is minus D here and there is plus D here, so minus D cancels with plus D and I have minus omega L minus omega D minus omega U. So what I have here is D plus omega L inverse of minus of omega into A. This is D and D cancels, this is minus omega L minus omega D minus omega U. So I have minus omega common, therefore D plus L plus U that is A xk plus omega D plus omega L inverse of b; therefore I can write this as D plus omega L inverse, I can take omega outside and this is b minus Axk. Therefore in the notation that we have used for the Gauss Seidel method we call this as vk and this as rk, residual vector. So this is our rk, the residual vector and this is our error in the solution vk. Therefore I can write down the method as D plus omega L vk is equal to omega rk. So this is our residual vector and this is multiplied by this omega and this D plus omega, all we have pre multiplied and this is our xk. Now you can see that this is in no way computationally expensive than Gauss Seidel except that we have put a factor over here and one more multiplication we put here. Otherwise when omega is equal to one this is same as Gauss Seidel and hence we called this as generalization of Gauss Seidel method.

Now you can see when omega is equal to one we immediately get our Gauss Seidel method here; therefore it is not computationally expensive compared to the Gauss Seidel method at all. Now the omega is called the relaxation factor. The analysis of the method requires that for convergence omega should lie between zero and two. So we should have it such that omega lies between zero and two for convergence. If we use the value greater than one that is above the Gauss Seidel, one less than omega, less than two it is called the over relaxation and it is also called a literature successive over relaxation. And this is one of the most famous and most popular methods. It is simply called SOR, Successive Over Relaxation method. If I use the value between zero and one, it is called under relaxation but it is very rarely used, so this is called the under relaxation. Now therefore this success of the over relaxation procedure would depend on the value relaxation factor that is correctly determined for the properties that I have mentioned earlier. Now symmetric and property A; it is possible for us to say what is the omega optimum value such that this has fastest convergence. It is not unusual to get at least ten times faster than Gauss Seidel in many cases. In fact if you can find the optimum relaxation factor very accurately even getting hundred times faster than Gauss Seidel is not surprising thing for this successive over relaxation.

The successive over relaxation is therefore one of the most important methods for this solution of the linear algebraic equations. There are later on some modifications to make it work better. However the even though the method was derived for the methods having property of A, later on it has been drop p plus started using or irrespective of property A, just for symmetric matrices you can use it today even you can try it for the non symmetric matrices also a relaxation

What we are really doing here is we are taking a type of the error between the Gauss Seidel and this one and then multiplying by suitable factor and adding to it in a sense by now improving upon the error in the solution and therefore it would work even for the other cases also. Now before I have a discussion on the error analysis of this, let us take a simple example for the successive over relaxation.

(Refer Slide Time: 00:20:03 min)



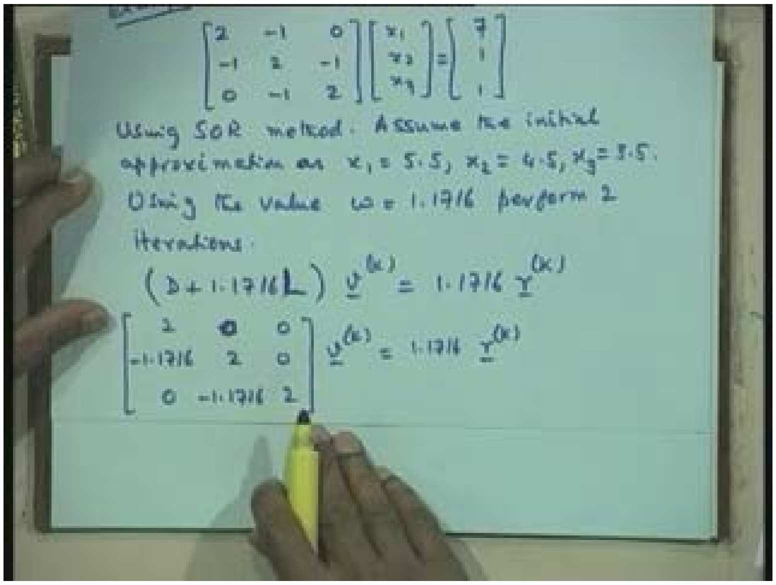
Example Find the solution of the system

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

Using SOR method. Assume the initial approximation as $x_1 = 5.5$, $x_2 = 4.5$, $x_3 = 3.5$. Using the value $\omega = 1.1716$ perform 2 iterations.

So let us take an example, so let's call it as - find the solution of the system of equations. Let me take a simple example, 2, -1, 0; -1, 2, -1; 0, -1, 2 and x_1, x_2, x_3; 7, 1, 1. Using SOR method assumes the initial approximation as x_1 one is equal to 5.5, x_2 is equal to 4.5, x_3 is equal to 3.5. Using the value will give the value of omega here. Using the value, omega is 1.1716; perform two iterations. Now here we are given the initial approximation. We have also given the omega value and we need to iterations of this as SOR and we shall see later on how we are going to obtain this value of omega.

(Refer Slide Time: 00:22:06 min)



$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

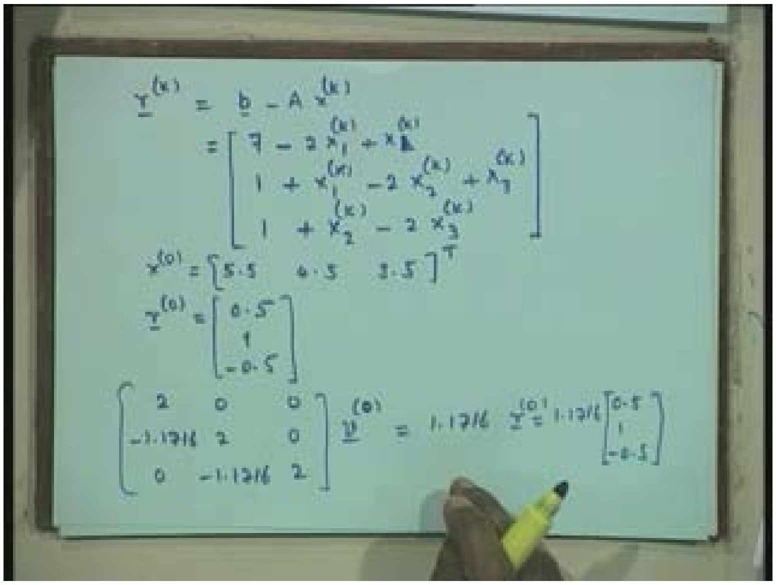
Using SOR method. Assume the initial approximation as $x_1 = 5.5$, $x_2 = 4.5$, $x_3 = 3.5$. Using the value $\omega = 1.1716$ perform 2 iterations.

$$(D + 1.1716 L)\, \underline{v}^{(k)} = 1.1716\, \underline{r}^{(k)}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ -1.1716 & 2 & 0 \\ 0 & -1.1716 & 2 \end{bmatrix} \underline{v}^{(k)} = 1.1716\, \underline{r}^{(k)}$$

Now let us use this error format of writing the system; that is our D plus omega L D plus omega L into vk is equal to omega 1.1716 rk. So D is the same, L is this part; so I am multiplying by omega, so I can write down 2, -1, 0 omega into this minus 1.1716. Therefore 0, - 1.1716, 2; this is our D plus omega times L. This is our diagonal part. This is the part omega into lower triangular part, that is D plus omega L, vk is 1.1716. Now let’s write down rk in the right, let’s write down what is rk in the next step.

(Refer Slide Time: 00:23:25 min)



$$\underline{r}^{(k)} = \underline{b} - A\,\underline{x}^{(k)}$$

$$= \begin{bmatrix} 7 - 2x_1^{(k)} + x_2^{(k)} \\ 1 + x_1^{(k)} - 2x_2^{(k)} + x_3^{(k)} \\ 1 + x_2^{(k)} - 2x_3^{(k)} \end{bmatrix}$$

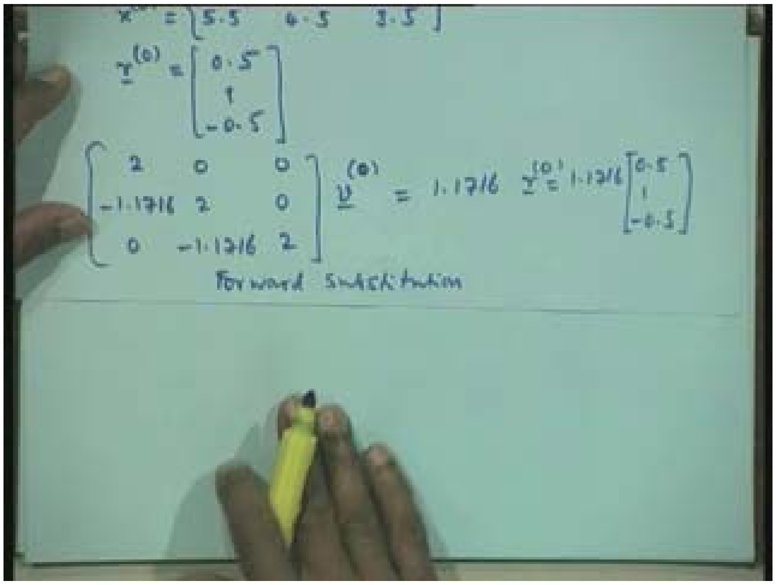
$$x^{(0)} = [5.5 \quad 4.5 \quad 3.5]^T$$

$$\underline{r}^{(0)} = \begin{bmatrix} 0.5 \\ 1 \\ -0.5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ -1.1716 & 2 & 0 \\ 0 & -1.1716 & 2 \end{bmatrix} \underline{v}^{(0)} = 1.1716\ \underline{r}^{(0)} = 1.1716 \begin{bmatrix} 0.5 \\ 1 \\ -0.5 \end{bmatrix}$$

Now the value of rk (that is the residual vector) is b minus Axk, which is b minus Ax k. So I can write down b minus Axk from here; b is this, entire A is this, so as this goes to the right hand side, I will have this as seven minus two times x one plus x two at the previous iterate xk one. Then I am taking this to the right hand side. So plus x_1 of k, two times x_2 of k plus x_3 of k. So I have taken these three values of the right hand side and third the component is one; this is plus x two of k two times x three of k. Now we are given the initial approximation as our vector x_0 is given as 5.5, 4.5 and 3.5. So we are given this as our initial approximation. So I can just substitute it in rk. This is the value, so I can substitute x_1 is equal to 5.5, x_2 is 4.5 and x_3 is equal to 3.5. Then I would get r_0. I will give you the value that I obtained here, this is a very simple thing, 0.5, -0.5. So I just multiplied this, and this gives you minus eleven and this gives four point five; this gives you four point five and that is eleven point five, minus eleven. So point five; similarly I get the other values of r. Once the value of r is obtained, i substitute it in this. In this particular step, so the next step I would have here is 2, 0, 0; -1.1716, 2, 0; 0, -1, 1.1716 and the last element is 2; v_0 is equal to 1.1716, r_0 is 1.1716, 0.5, 1, -0.5.

(Refer Slide Time: 00:26:23 min)



$$x^{(0)} = [5.5 \quad 4.5 \quad 3.5]$$

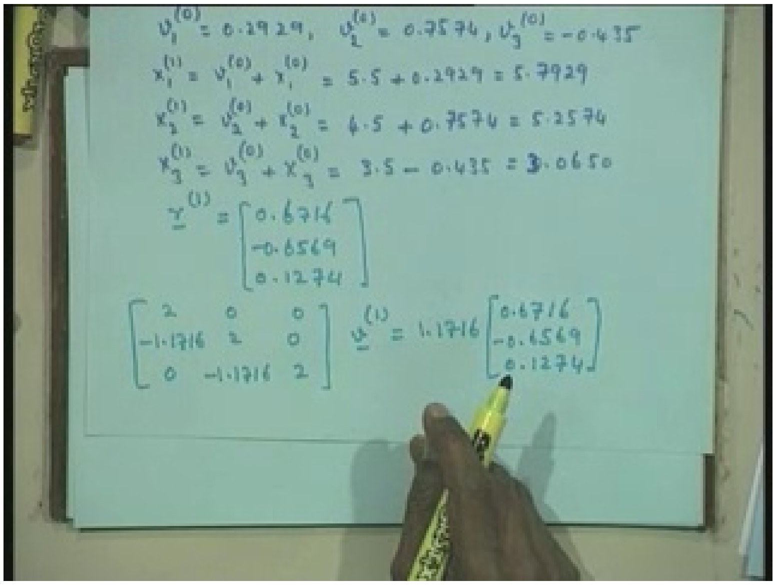
$$\underline{r}^{(0)} = \begin{bmatrix} 0.5 \\ 1 \\ -0.5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ -1.1716 & 2 & 0 \\ 0 & -1.1716 & 2 \end{bmatrix} \underline{v}^{(0)} = 1.1716 \ \underline{r}^{(0)} = 1.1716 \begin{bmatrix} 0.5 \\ 1 \\ -0.5 \end{bmatrix}$$

Forward substitution

Now this is a forward substitution method. So I can just find out the first component of this v_1 as this is divided by two and then use this value to get the value of v_2. Then I can get the value of $e_{3,}$ so by forward substitution I can get the solution for this problem.

(Refer Slide Time: 00:26:47 min)



$v_1^{(0)} = 0.2929, \quad v_2^{(0)} = 0.7574, \quad v_3^{(0)} = -0.435$

$x_1^{(1)} = v_1^{(0)} + x_1^{(0)} = 5.5 + 0.2929 = 5.7929$

$x_2^{(1)} = v_2^{(0)} + x_2^{(0)} = 4.5 + 0.7574 = 5.2574$

$x_3^{(1)} = v_3^{(0)} + x_3^{(0)} = 3.5 - 0.435 = 3.0650$

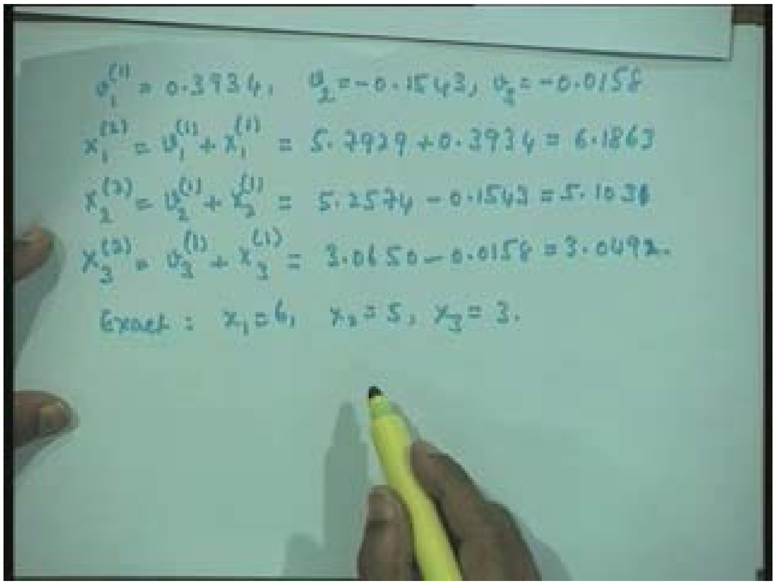
$$\underline{r}^{(1)} = \begin{bmatrix} 0.6716 \\ -0.6569 \\ 0.1274 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ -1.1716 & 2 & 0 \\ 0 & -1.1716 & 2 \end{bmatrix} \underline{v}^{(1)} = 1.1716 \begin{bmatrix} 0.6716 \\ -0.6569 \\ 0.1274 \end{bmatrix}$$

So if I do that I would get $v_1 0$ as 0.2929. I have v_2 of zero as 0.7574; v_3 of zero is – 0.435. This is the solution of this system by just forward substitution. Now I generate my new value, the computed value which will be equal to; x_1 of one is equal to v_1 of zero plus x_1 of zero. So this is the increment, therefore I will write this as the x_1 of zero plus v_1 of zero and x_1 of zero was given as 5.5 and I have this as 0.2929, so I will have 5.7929. Similarly the second component x_2 of one is v_2 of zero plus x_2 of zero. Now the initial approximation is 4.5 plus 0.7574 and that gives you 5.2574. I have added these two to give 5.2574 and the third component is x_3 of one is v_3 of zero plus x_3 of zero and this value was given as 3.5 minus 0.435 that is 3.060 and this is -4.5.

Once I have obtained this value to go to the next iterate I need again this residual vector. So I go back and substitute this in the residual vector and get the next solution for this one. So if I substitute that I would get this particular value, so I get r_1, which is as same. By substituting this value of the residual vector I get this as 0.6716, -0.6569, and 0.1274. Now I need to solve this system again. So I will again have to solve this system 2, 0, 0; -1.1716, 2, 0; 0, -1.1716, 2 and v_1 is 1.1716 into r_1. So I could now write this as 0.6716, -0.6590 and 0.1274.

(Refer Slide Time: 00:30:25 min)



$v_1^{(1)} = 0.3934, \quad v_2 = -0.1543, \quad v_3 = -0.0158$

$x_1^{(2)} = v_1^{(1)} + x_1^{(1)} = 5.7929 + 0.3934 = 6.1863$

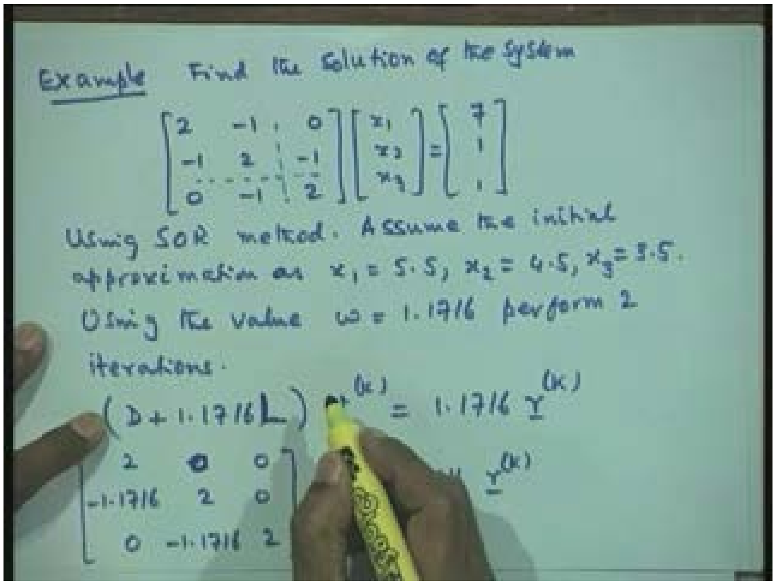
$x_2^{(2)} = v_2^{(1)} + x_2^{(1)} = 5.2574 - 0.1543 = 5.1031$

$x_3^{(2)} = v_3^{(1)} + x_3^{(1)} = 3.0650 - 0.0158 = 3.0492.$

Exact : $x_1 = 6, \quad x_2 = 5, \quad x_3 = 3.$

Now we have to solve this equation again by forward substitution and I get the solution of this as v_1 of one, that's at this stage it is v_1 of one is 0.3934. I get v_2 is -0.1543, v_3 three is -0.0158, that is from just computation of this value again by forward substitution. Now I will update my value of x, that is x_1 of two is equal to v_1 of one plus x_1 of one. Now let's go back to this slide. So the value of x_1 was 5.7929 plus 0.3934 that is 6.1863. Then x_2 of two is v_2 of one plus x_2 of one, the value of previous iterate was 5.2574. So I will have 5.2574, -0.1543 that is 5.1030, here it is 31 and lastly x_3 of two is v_3 of one plus x_3 of one the value of x_3 is 3.0653, 0.065 and the value of v_3 is -0.0158 that is 3.0492. Now we have completed two iterations and the exact solution for this problem is x_1 is 6, x_2 is 5, x_3 is equal to 3. Now the interesting observation that we would like to do here is that the convergence is so fast here. Look at this particular value $r_{0,}$ this is the error in satisfying our given equations, that is 0.51, - 0.51. When we have gone to the first step the error in satisfying the equations was 0.6716, -0.569. Now the solution has now moved faster and now the solution has come almost close to the exact solution the second iteration. Now if you find the error r_2, you will find that this value is very small. Therefore the convergence compared to the Gauss Seidel is much faster. In fact for this particular problem let us take about three or four decimal places accuracy; Gauss Seidel will require almost a hundred iterations to get that accuracy, whereas you will find in SOR that the required is much less almost a factor or so less than the required number of iterations. So that's why the successive over relaxation method is very fast compared to the other methods but the only hitch here is that we have to find suitable value of omega, which gives the fastest convergence.

(Refer Slide Time: 00:34:21 min)



Example Find the solution of the system

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

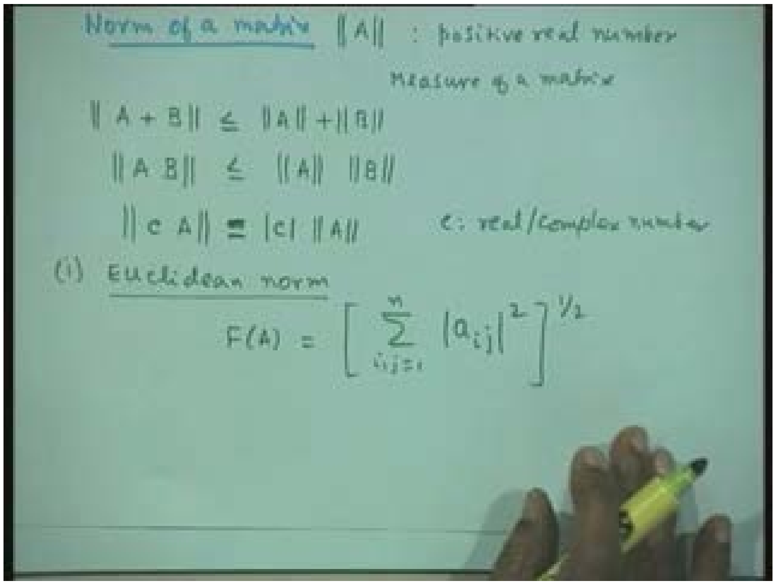
Using SOR method. Assume the initial approximation as $x_1 = 5.5$, $x_2 = 4.5$, $x_3 = 3.5$. Using the value $\omega = 1.1716$ perform 2 iterations.

$$(D + 1.1716\,L)\,x^{(k)} = 1.1716\,\underline{r}^{(k)}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ -1.1716 & 2 & 0 \\ 0 & -1.1716 & 2 \end{bmatrix} \quad \underline{r}^{(k)}$$

One observation that we should also make here is that when I gave this particular example, which matches exactly over requirements, if I look at this and see its partition, we can apply the permutation matrix such that these are thrown to other way and then we will be able to get the diagonal matrix here and the diagonal matrix here; or you could have another partitions also and then try to attempt that particular way. However as I said it’s not necessary nowadays, we can use even if it is a symmetric matrix. Now in order to do the analysis of this even though I will not go deep into how to obtain the omega, or how to obtain for successive over relaxation but we would like to discuss how the Gauss Seidel or the Jacobi or SOR convergence and how do you define by a value that this is the rate of convergence of this particular method; and thereby we can compare the various methods that this is the convergence of the method, therefore it is faster than the other methods. Now to do that I need the background of your matrixes. Let us just revise what we have done earlier and little bit more of the some of the concept about the property of the matrixes. One important property we will need is the norm of a matrix, so let us define the norm of matrix which you might have already done in the course but let’s revise it.

(Refer Slide Time: 00:35:48 min)



Norm of a matrix $\|A\|$: positive real number

Measure of a matrix

$$\|A + B\| \leq \|A\| + \|B\|$$

$$\|A B\| \leq \|A\| \, \|B\|$$

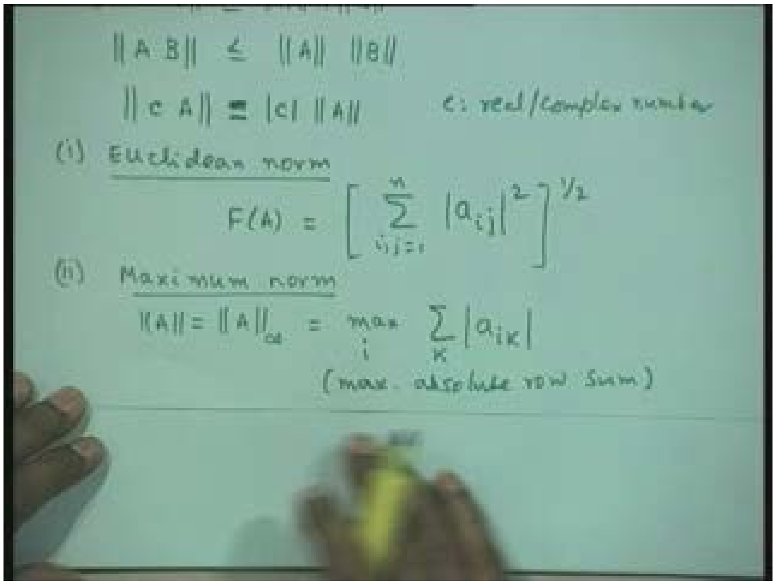
$$\|c A\| = |c| \, \|A\| \qquad c\text{: real/complex number}$$

(i) Euclidean norm

$$F(A) = \left[\sum_{i,j=1}^{n} |a_{ij}|^2 \right]^{1/2}$$

So I would like to define what's known as norm of a matrix. If given a matrix A, I will denote the norm of the matrix by this, this is a positive real number. This is a positive real number in a sense that it is a measure of a matrix. So it is in a sense actually that we can call this as a measure of a matrix; so we can say given a matrix, this is the norm of this matrix; given another matrix you can say norm of a matrix is this. So you can compare them in terms of norms of the matrixes; so in a sense we can call it as a measure of a matrix. Now if I have two matrixes A and B, then norm of this is your triangular in equality holds here; that is norm of A plus norm of B and if I take the product of two matrixes norm of A into B, this is also norm of A into norm of B. And if I have a complex, any number multiplying, say complex number C into A norm will be where we are talking of C is a real or a complex number. C is a real or a complex number, then this will be equal to magnitude of C into norm of A. This means that if C is a positive or a negative real number or C is complex, then this equality holds. Now let us see what the measures that we can give for this are. The first one norm is called the Euclidean norm. We are talking of a general matrix means real or complex matrix. Therefore what I will do is I will take the magnitude of all the elements, square them, add up and then whatever number that you would get, take the square root of that; that will be called the Euclidean norm. In this you know it is the sort of an Euclidean norm. If you take the distances of points, there is nothing but the distance between the two points is the length. It is the measure, it is Euclidean norm. So that is actually a representation of this in the matrixes. So this is the nothing but I take magnitude of aij square, then summation over i is equal to j one to n, all of them. Then I take the square root of this and this is called the Euclidean norm. It is usually represented by F of A. So it's a very simple thing that we just take the magnitude of all the elements of the matrix, square them, sum them up and then take the square root of that, a different norm is used in different circumstances.

(Refer Slide Time: 00:39:26 min)



$\|A\,B\| \le \|A\|\;\|B\|$

$\|c\,A\| = |c|\;\|A\|$ c: real/complex number

(i) Euclidean norm

$$F(A) = \left[\sum_{i,j=1}^{n} |a_{ij}|^2 \right]^{1/2}$$

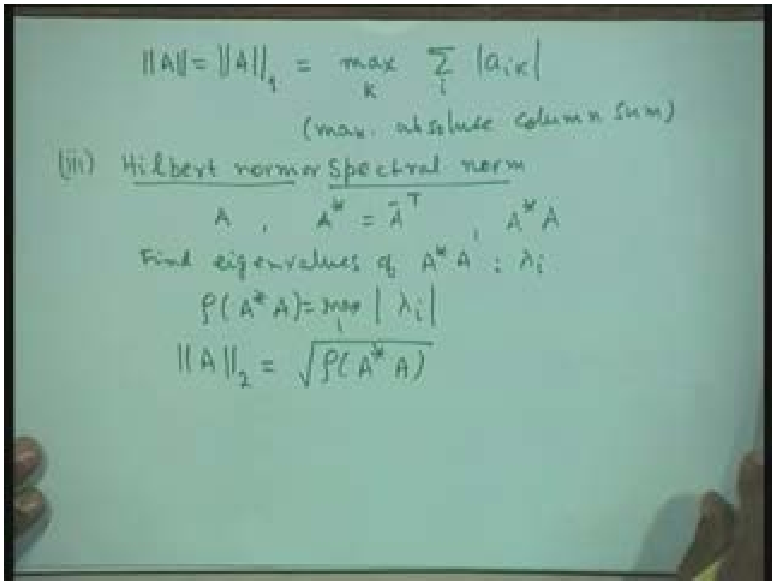
(ii) Maximum norm

$$\|A\| = \|A\|_\infty = \max_{i} \sum_{k} |a_{ik}|$$

(max. absolute row sum)

So let us have another norm called the maximum norm. It is denoted by norm of A with a suffix as infinity. I take any row, take the absolute values of all the elements, sum them up, there will be n such row sums. So I will take the largest of them, that is the maximum absolute row; sum is a measure and that is called the maximum norm. So that is denoted by this; that means this is maximum with respect to row I, summation with respect to k; of all aik, that is these are all the magnitudes of the elements and this is maximum with respect to I; that is row and I am summing up with respect to k. So that is all the values in the row are added in magnitude and then I maximize it with respect to all the rows and then this will be the maximum norm. So I can call this is as maximum absolute row sum.

(Refer Slide Time: 00:40:58 min)



$$\|A\| = \|A\|_1 = \max_{k} \sum_{i} |a_{ik}|$$

(max. absolute column sum)

(iii) Hilbert norm or Spectral norm

A , $A^* = \bar{A}^T$, A^*A

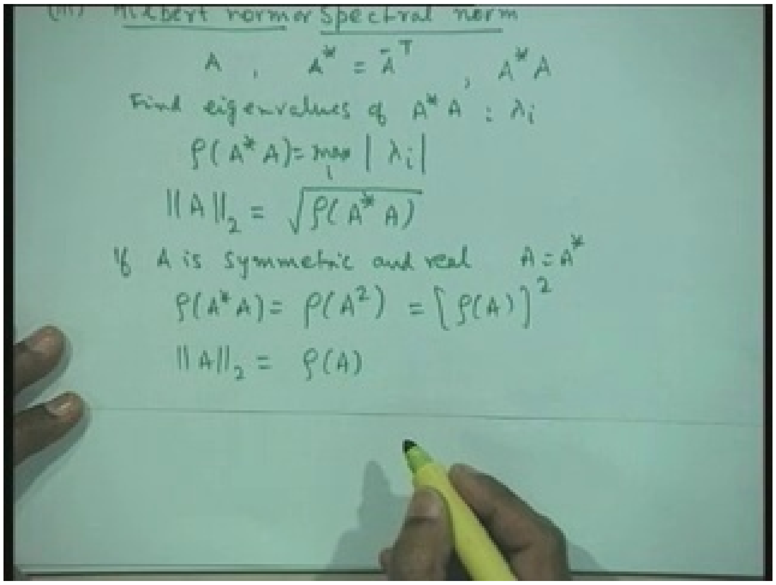
Find eigenvalues of A^*A : λ_i

$\rho(A^*A) = \max_{i} |\lambda_i|$

$\|A\|_2 = \sqrt{\rho(A^*A)}$

If I can take maximum absolute row sum, I should be able to take the maximum absolute column sum also. Therefore that would also give me another measure which I would call under the maximum norm only; but with a different notation the notation that is used is this; that is maximum with respect to k, summation with respect to I, aik. So now I am taking the maximum with respect to column and now I am adding the rows. Therefore this is the maximum absolute column sum and the third norm that we need is the one that is we use very often called the Hilbert norm or it is also called spectral norm or spectral norm. More often we use it as a spectral norm. What we do here is, given a matrix A, I would find its conjugate transpose and we defined it as Astar; that is your conjugate transpose. Then I form the matrix A star A, then I will find the Eigen values of A star A. We have defined the largest Eigen value of a matrix as its spectral radius, so I will find the largest Eigen values of this. Therefore I will take this as this, let's call it as lambda. I find Eigen values of this, that is lambda I; the maximum Eigen value is the spectral radius of this.

(Refer Slide Time: 00:43:27 min)



(iii) Hilbert norm or Spectral norm

$A \,,\quad A^* = \bar{A}^T \,,\quad A^*A$

Find eigenvalues of A^*A : λ_i

$\rho(A^*A) = \max_i |\lambda_i|$

$\|A\|_2 = \sqrt{\rho(A^*A)}$

If A is symmetric and real $\quad A = A^*$

$\rho(A^*A) = \rho(A^2) = [\rho(A)]^2$

$\|A\|_2 = \rho(A)$

The Hilbert norm says that I can define the norm as square root of spectral radius of A star A. That means I find the matrix A star A, find its Eigen value, find its spectral radius, square root of that will be the Hilbert norm. Suppose A is symmetric and real; then look at this definition, A star is conjugate transpose, it's real, therefore it is same as this called symmetric. Therefore A is equal to A transpose, that means A is equal to A star. Therefore that means this is equal to A, is equal to A star by definition and whereas this spectral radius of A star A will become spectral radius of A square A which is equal to A star; therefore it will simply be spectral radius of A square. Now we know that if lambda i is an Eigen value of A, then lambda i square is the Eigen values of A square. Therefore spectral radiuses of A square of that Eigen values of A to the power of m is lambda i to the power of m; lambda is the Eigen values of A. Therefore in this case this spectral norm will be simply spectral radius of A. Therefore if you have a symmetric matrix and real we are very happy because we can just find the Eigen values of that matrix. The largest Eigen values in magnitude will be simply the spectral norm. Now all these norms that we have defined are all independent norms; the Euclidian norms or the maximum norm or the maximum absolute column sum norm or the Hilbert norm, they are all independent definitions. Therefore if you are given a matrix and if you want to find the norm and if you find all of them, since all are independent you can take the smallest of them as a required norm. Suppose say given a matrix A, what is the largest value of norm this matrix has got. If you are finding the norm of all the four, the smallest of that will be the required bound; that will give the norm of that particular matrix.