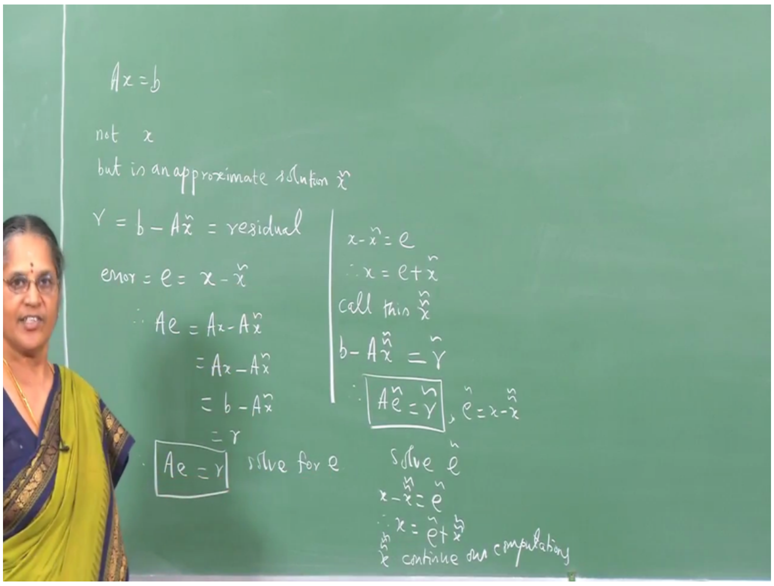
**Numerical Analysis.**
**Professor r. Usha.**
**Department of Mathematics.**
**Indian Institute of Technology, Madras.**
**Lecture-45.**
**Solution of Linear Systems of equations-8**
**Iterative Improvement Method Iterative Methods-1.**

Good morning everyone, in the last class we discussed the error analysis for the direct methods and we wanted to know whether the ideas that we understood from the error analysis, there is a possibility of improving the accuracy of the approximate solution that we obtained for a system of equations AX equal to b.

(Refer Slide Time: 0:58)




So when we solve the system of equations of the form AX is equal to b by the direct methods, then the solution that we obtained is not the exact solution X but is an approximate solution which we denote by say X tilde due the round-off errors which are incorporated at any stage of our computation. So the question now is can we improve this approximate solution that we obtained using direct methods so that we obtain the solution of the system correct to the desired degree of accuracy. So I compute the A Times X tilde where X tilde is the approximate solution that we obtained using any of the direct methods.

We expect that AX tilde is very close to the right-hand side vector b because we are solving the system AX equal to b. And X tilde being an approximate solution, there will be a difference between b and AX tilde, so I compute this error. Recall b is a vector and A is a matrix of order N cross N, X tilde is an N cross 1 vector, so this will be N cross 1 vector, b is an N cross 1 vector, so b - A tilde is an N cross 1 vector. So I compute this vector and call this as r, namely it is the residual. Now what is the error that has been incorporated? The error if I denote by e is the difference between the exact solution and the approximate solution.

So it is X - X tilde and therefore A times this error is AX - AX tilde. And so it is AX -, AX tilde is what we have computed using our approximate solution and AX is b, so it is b - AX tilde. But what is b - AX tilde, that is the residual vector that we obtain. We observe that e the equation A e is equal to r, so there is a relationship connecting the error that one step and the residual at that particular. So e at a particular step which is a vector satisfies the equation Ae is equal to r. So we again have a system of equations in which the unknown vector is e, why is it an unknown e, what is e, e is X - X tilde.

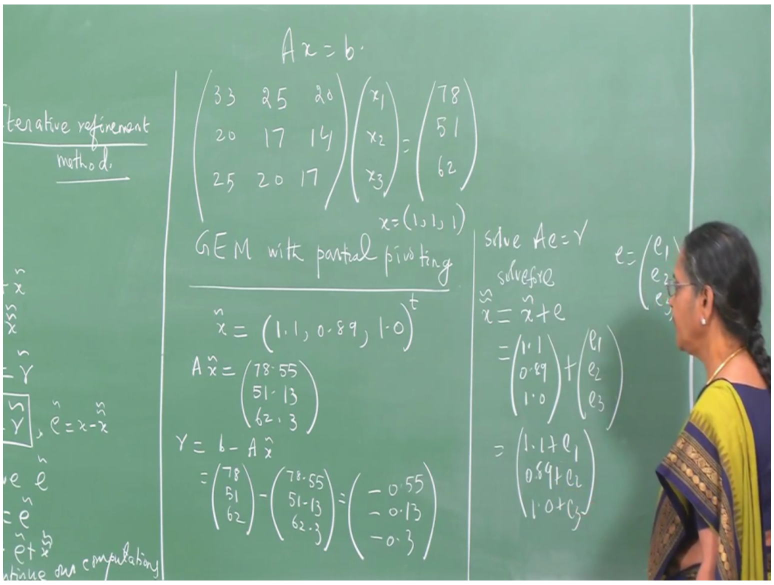
X is not known to us, it is exact solution, that is what we are trying to obtain, so e is an unknown vector, so here e is an unknown vector, so that we solve the system Ae is equal to r. Do we know the right-hand side vector r at this step? Yes, what is it, it is b - A into X tilde. So we solve the system. How do we solve? Again by any of the direct method that we have learnt and then obtained this error vector e, so once we have e, then we use the fact that the error is X - X tilde. So X - X tilde is going to be the e that we have computed and therefore X is equal to e + X tilde.

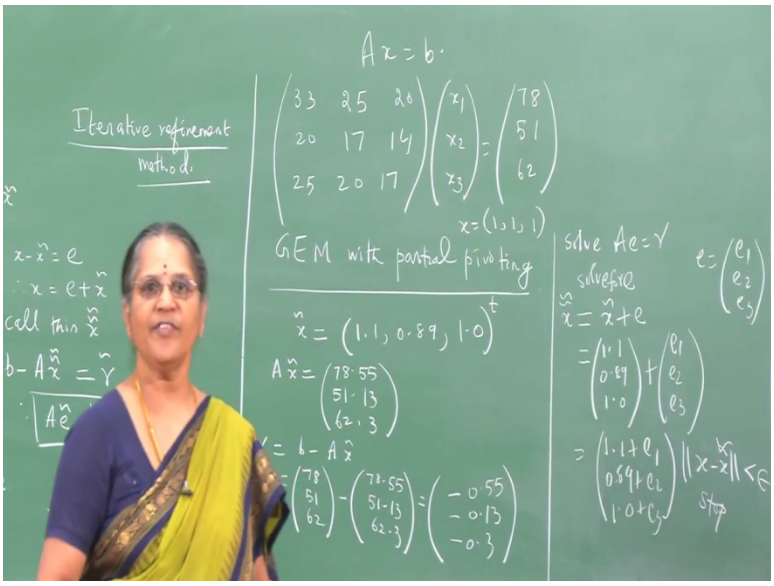
So we call this vector as X double tilde, we compute what is b - AX double tilde. If this is such that the X double tilde that we computed is within the desired degree of accuracy that we have demanded while computing our solution, then we can take X double tilde to be the solution and stop our iterations. If not, what do we do, this is the residual at this step, we call this as r tilde. And therefore we solve the new equation, namely A into e tilde is equal to r tilde. So we compute what is e tilde, what is e tilde, e tilde is X - X double tilde which was obtained as the solution at this step.

So again we have a system which satisfies the relationship A into e tilde is equal to r tilde. So we solve for this are known vector again by any of the direct methods that we know. And so we obtain X - X double tilde to be e tilde which has been obtained just now, so X is equal to e tilde + X double tilde. And we check whether the solution that we have obtained at this step is correct to the desired degree of accuracy. If so, we stop our computation, if not, we call this as X triple tilde and continue our computations till the desired degree of accuracy is achieved.

So the method described here is called the iterative refinement method, which helps us to improve the accuracy of the approximate solution that we have obtained using any of the direct methods. So this method helps us to obtain the solution of the given problem that AX equal to b correct to the desired degree of accuracy. So let us take an example and I will indicate the 1st few steps and you can complete the problem.

(Refer Slide Time: 7:57)




So solve the system AX is equal to b, given that A is a 3 cross 3 matrix having elements as given here, multiplied by the vector X1, X2, X3 and that should be equal to 78, 51, 62, so I have a system AX is equal to b. Say apply Gauss elimination method with partial pivoting, then say I obtain the solution X tilde as 1.1, 0.89 and 1.0 transpose. When I look at the system may immediately see that X1 equal to1, X2 equal to1 and X3 equal to1 is the exact solution, since 33+25+20 is 78, 20+17+14 is 51, 25+20+17 is 62 and therefore 1, 1, 1 is the exact solution of the system.

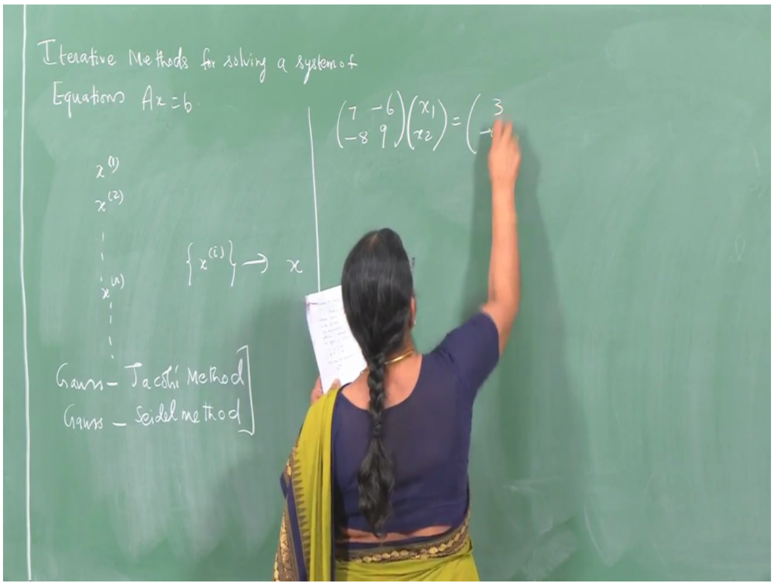
When I apply Gauss elimination method with partial pivoting, I end up with a solution which is this. So it is an approximate solution, so I must improve this approximate solution so that I go closer and closer to the exact solution which is 1, 1, 1. So what should I do according to the iterative refinement method? I must compute A into X tilde. So A is this matrix and X tilde is this vector and so when you compute A into X tilde, that turns out to be vector 78.55, 51.13, 62.3. So I compute what is the residual vector r, what is the residual vector, it is b - AX tilde. And my b is the right-hand side vector 78, 51, 62 - AX tilde is 78.55, 51.13, 62.3. So the residual vector is -0.55, -0.13 and -0.3. So this is my residual vector. So what should I do now?

I should solve the system A into e is equal to this residual vector r. You know what the matrix A is, you also know the right-hand side vector r, so you solve for e, the error vector. So once you get e, then your solution X is going to be this X tilde + e. What is X tilde, X tilde has components 1.1, 0.89 and 1. If suppose you compute your e to have components U1, U2, U3, then add this vector to X tilde, so your new X which I call as X double tilde will be 1.1+ e1, 0.89+ e2, 1.0+ e3. So this is your new solution, this is your new approximation to the solution of the system AX equal to b.

Now check whether X - X double tilde that you have computed is such that the magnitude is less than the prescribed error tolerance. If so, you stop your iterations and you say that X double tilde is your solution correct to the desired degree of accuracy. If not, compute A into X double tilde and then compute b - AX double tilde, call it the residual vector r double tilde and then continue your computations as described in the iterative refinement methods and obtain the solution and see whether that solution is current to the desired degree of accuracy.

So I leave the rest of the computations for you to complete, work out the details. I will give you some more problems on this method in the assignment sheet. So that completes our discussion on the direct methods, so we learnt some direct methods, namely the decomposition methods and Gauss elimination method, Gauss Jordan method and also we incorporated the pivoting strategy in these 2 methods. And then performed some error analysis in the previous class and then using those concepts we have now seen how we can improve the accuracy of the approximate solution that we get using direct methods when we solve a system of equations.

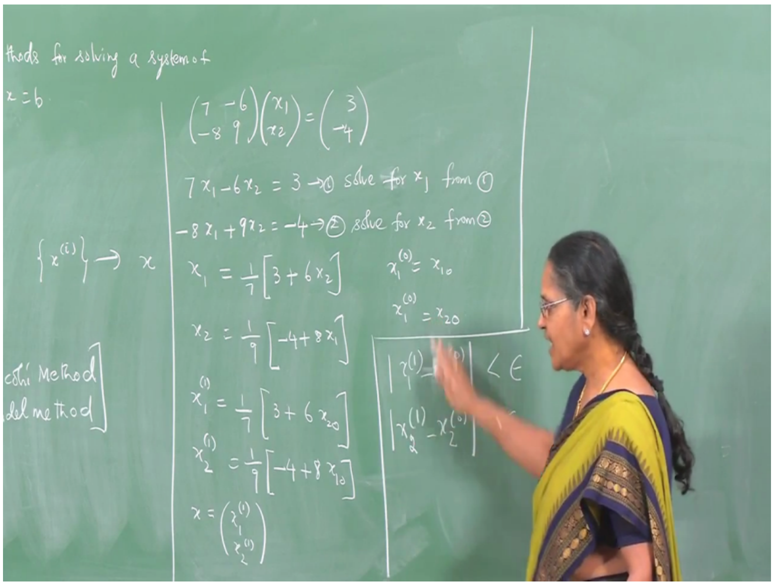
(Refer Slide Time: 14:15)




So now we move onto the iterative methods for solving the system of equations AX equal to b. So our goal now is to develop methods which are iterative in nature for solving a system of equation X equal to b. So what do these methods do? These methods generate successive iterates say X1, X2, etc. XN and so on for the solution vector X such that this sequence say XI converges to the solution X. So where do we stop these iterations? We either perform prescribed number of iterations and stop our computations and take the solution obtained at that step to be the solution of the system or we work out the details in such a way that the solution that we obtained at any stage satisfies the required accuracy demand.

So let us see some methods for solving a system of equations iteratively. One of the methods for solving this system is known as Gauss-Jacobi method, the other one that we are going to learn is called Gauss-Seidel method. So we shall describe these methods and see how one can solve a system of equations iteratively. So we shall illustrate it by taking a simple example of a system of 2 equations in 2 unknowns, the procedure is the same for a system of N equations in N unknown.

(Refer Slide Time: 15:45)

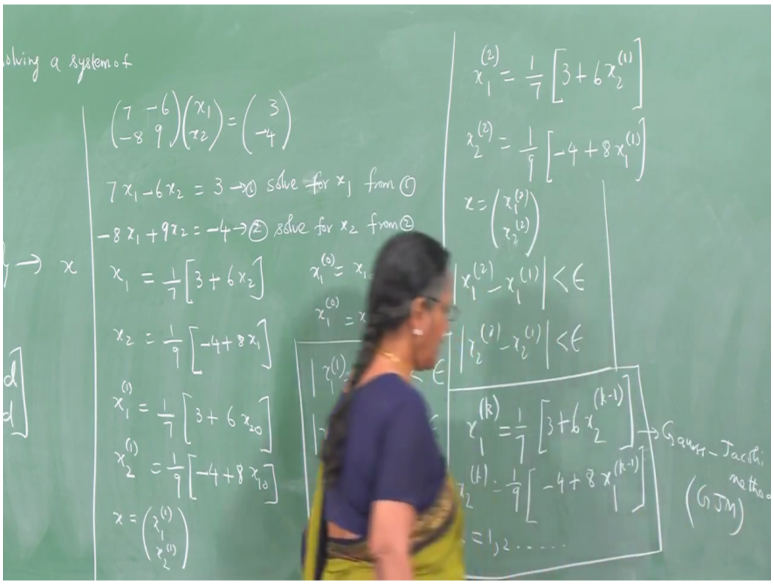



Suppose saying we are given the system of equations 7, -6, -8, 9, X1, X2 equal to 3, -4. If I write out this equation, then it is 7 X1 -6 X2 is 3, -8 X1 + 9 X2 is -4. So the procedure for solving this system using Gauss-Jacobi method is as follows. I take the 1st equation to be associated with the 1st unknown, namely X1 and I solve for X1 from the 1st equation. So solve for X1 from equation 1 and similarly I will associate the 2nd unknown X2 with the 2nd equation and use the 2nd equation and solve for X2 using the 2nd equation. Then in that case I get X1 to be equal to 1 by 7, 3+6 X2, then X2 is 1 by 9 into -4+8 X1. This is the 1st step that I do when I want to apply Gauss-Jacobi method.

Now since it is an iterative method, I must start with an initial approximation. So initially I shall take my solution, right, as say X1 0 and X2 0. And use that solution in the right-hand side and compute what is the solution that I get for the unknown variable. So when I do that, then I get X1 to be 1 by 7 into 3+6 into X2 0 because my initial approximation for X2 is X2 0, this is what I guess for X2. So when I substitute for X2 0, I have a value for X1, so I say that this is the solution that I obtain of the 1st step, namely X1 1. Now I go to the 2nd equation, then that is 1 by 9 into -4+8 times X1 0. So I have a number now which gives me X2, so I call this as X2 1.

And therefore at the 1st I have my solution for the unknown vector X, namely it is X1 1 and X2 1 and now I continue my computations. Before continuing my iterations, I check whether the absolute value of the solution that I obtained for X1 at the 1st step - the solution that I assumed for X1 at the initial step is less than the prescribed error tolerance and I do that also for the 2nd variable, namely X2 1 - X2 0, its absolute value is less than epsilon. If both these conditions are satisfied, then I can stop my computations and take X1 1 and X2 1 to be the solution of the system correct to the desired degree of accuracy.

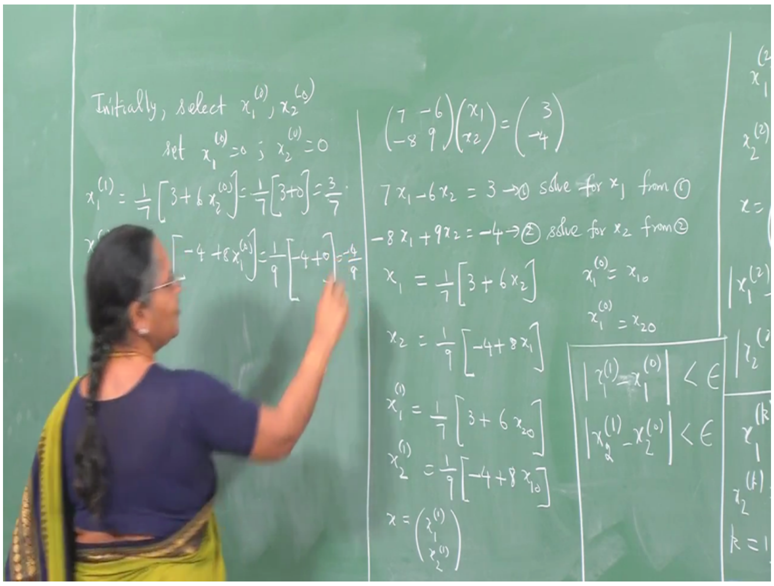
(Refer Slide Time: 20:04)

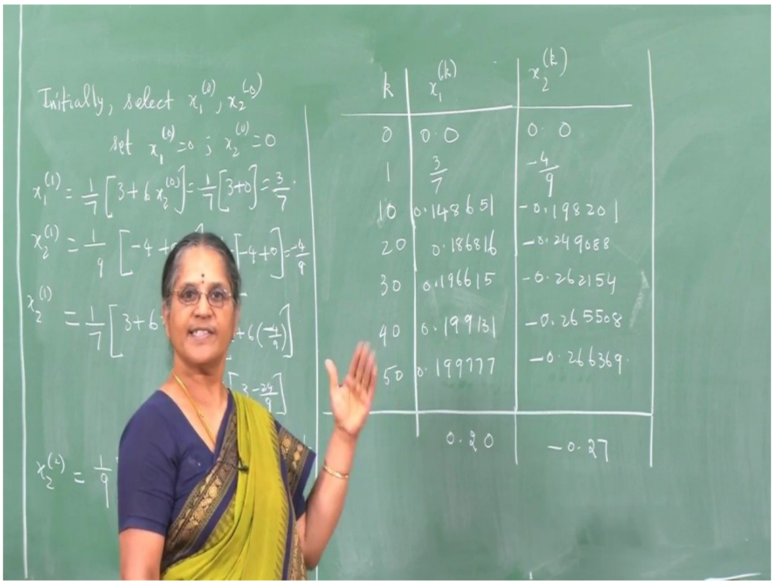



If that does not happen, then what should I do? I must use the current solution that I have for X1 and X2, namely X1 1, X2 1 and compute the new solution. What is the new solution, it is going to be X1 but now I obtain it at the 2nd step and that is going to be 1 by 7 into 3+6 into the variable X2 has the value X2 1 which we have computed at this step. Similarly, X2 at the 2nd step is 1 by 9 into -4+8 times X1 1. So I have the solution available at the 2nd step, namely X1 2, X2 2. So I again check whether the absolute value of X1 2 - X1 1 is less than prescribed error tolerance and X2 2 - X2 1 is less than the prescribed error tolerance. If so I stop the computations and take X1 2, X2 2 to give me the solution for X1 and X2 which saw the given system of equations, if not I continue my iterations like this.

So what is it that we do in Gauss-Jacobi method? We make use of the solution obtained at the previous that for the unknown variables, say X1, X2 and use them on the right-hand side whenever they occur and compute the solution of the current step, that is what we do in the Gauss-Jacobi method. Let us apply this method to the given system and arrive at the solution. At the Kth step, X1 K is given by 1 by 7 into 3+ X times X2, what is the value of X2 that I use here, that is available to me in the previous step. So it is X2 K -1. What about X2 K, it is 1 by 9 into -4+8 times X1 available at the K -1 step. So this is valid for K equal to1, 2, 3, etc.

(Refer Slide Time: 22:59)

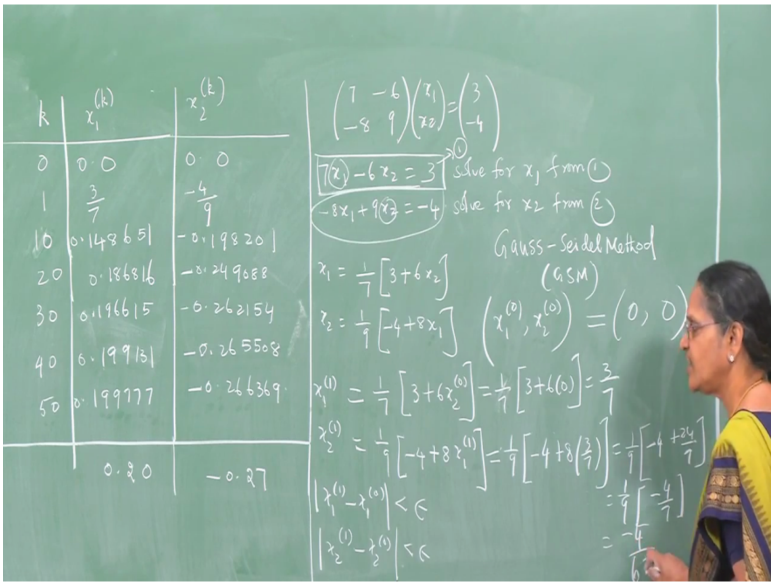


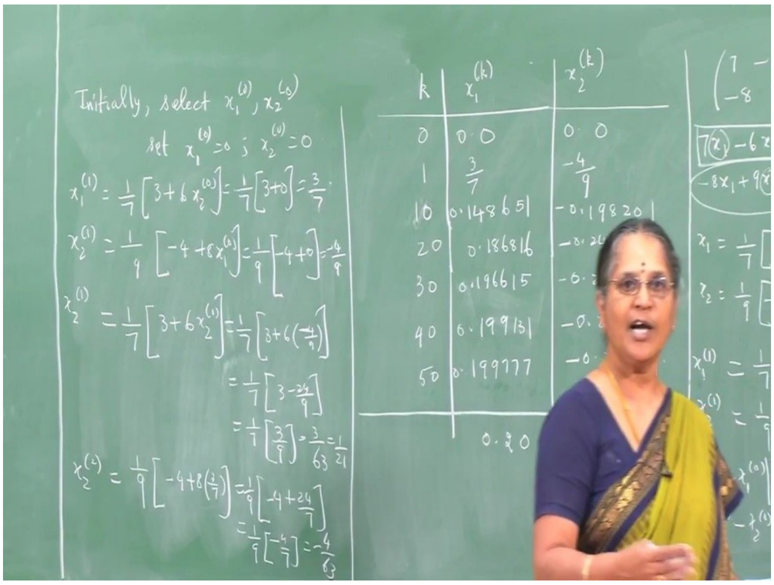




Then I start with an initial approximation for X1 0, X2 0, so this is what is known as Gauss-Jacobi method. So let us apply this technique and solve this system. So we initially select X1 0 and X2 0. So we simply set them to be 0. And then use that, so what is X1 1, it is going to be 1 by 7, 3+6 into X2 0, so it is 1 by 7, 3+0, so 3 by 7. And what is X2 1, it is 1 by 9 into -4+8 times X1 0. So it is 1 by 9 into -4+8 into 0 and so it is -4 by 9. And suppose say we are asked to do 50 iterations and produce the solution, so we have listed in the table the solutions up to 50 iterations that we have performed and see that X1 at this step turns out to be 0.199777 and X2 that we have obtained at the $50^{th}$ iteration as -0.266369.

So the $2^{nd}$ variable X2 has been determined correct to, say 3 decimal, correct to 2 decimal accuracy and the solution is 0.27 correct up to 2 decimal accuracy. And the solution for X1 is obtained in this case again correct to 2 decimal accuracy and that turns out to be 0.20. So at the end of 50 iterations, the solutions have been obtained correct to 2 decimal places using Gauss-Jacobi method. So at this stage we shall see what is Gauss-Seidel method for the same system of equations.

(Refer Slide Time: 25:25)




Let us look at the equations which are given to us, namely 7, -6, -8, 9 is the coefficient matrix, the unknown vector is having 2 components X1, X2, the right-hand side vector is 3, -4. So as before then we write down the system of equations, it is 7 X1 -6 X2 is 3 and -8 X1 +9 X2 is -4. As before I associate my 1st equation with the 1st unknown, namely X1 and I solve for X1 from the 1st equation and associate the 2nd equation with the 2nd unknown X2 and solve for X2 from the 2nd equation. So what do we get, X1 is going to be 1 by 7 into 3+6 X2 and X2 is 1 by 9 into -4+8 X1.

Now this is an iterative procedure that I would like to describe, namely it is Gauss-Seidel method. So as before I have to start with an initial approximation for these 2 variables X1 and X2. So suppose you have chosen X1 0, X2 0 as the initial approximation, then the solution at the 1st step, namely the 1st approximation for X1 is given by 1 by 7 into 3+6 times X2 0, because I have information about X2, I have guessed that the solution is X2 0. Now I move over to the next iteration, so X2 1, which is 1 and into -4+8 times, at this stage I see that I have X1 appearing here but X1 has already been obtained at this step, namely, the 1st equation gives me X1.

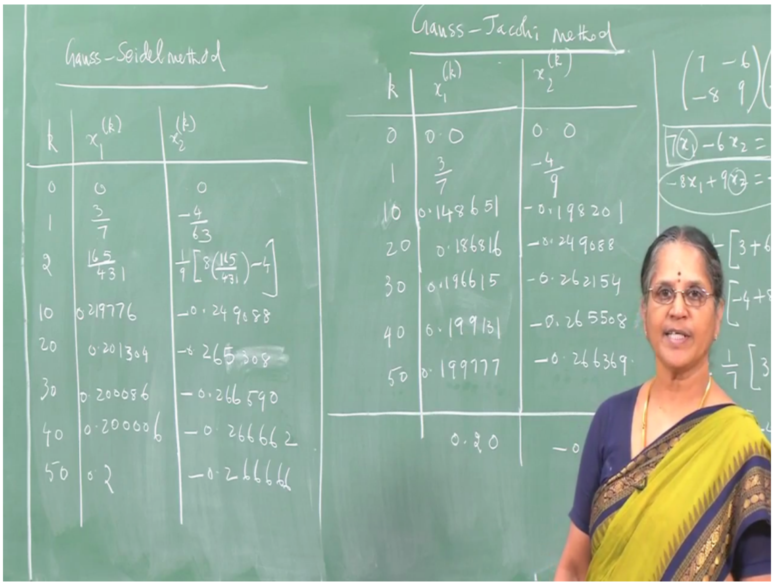
So why not I use the currently available solution for X1. So I take X1 1 for X1 here and then use that to compute the solution X2 1. I recall in Gauss-Jacobi method, we did not use the currently available value for any unknown on the right-hand side. We had any iteration, we always used the solution that was available in the previous step in the previous iteration. Whereas in Gauss-Seidel method, whenever it is possible for us to make use of the currently available solution, it is made use of on the right-hand side to compute the solution at that particular step from the left-hand side. So X2 1 is computed as 1 by 9 into -4+8 times, substitute for X1, the currently available value which you have computed just now.

And then now you have to check what is the difference between X1 1 - X1 0 in absolute value and X2 1 - X2 0 in absolute value. If they satisfy the desired condition for accuracy, then stop your computations and call X1 1, X2 1 as the solution of the problem, otherwise continue your computations till the desired degree of accuracy is obtained. So let us work out the details for this problem. So if I take the initial approximation as 0, 0, + I would like to compare my solutions with Gauss-Jacobi method. So at the 1st step what will I get, I get 1 by 7 into 3+6 times X2 0 is 0, so it is 3 by 7, that is what we got for X1 1 using Gauss-Jacobi method also.

So now I work out X2 1 and I see that it is 1 by 9 into -4+ 8 times, although initially I have approximated it to be 0, in the previous step, right, I would like to make use of the solution that is available to me for X1 currently. Namely I already know that X1 at this step which is 3 by 7, so I would like to make use of this and write down the solution as one by 9 into -4+24 by 7, so it is 1 by 9 into -28+24 by 7, so it is -4 by 63. So you observe that the solution for X2 even at the 1st step, right, coincides with the solution for X2 that was obtained at the 2nd step using Gauss-Jacobi method.

And so we expect faster convergence using Gauss-Seidel method because it makes use of the currently available solution which is computed at the previous step for an unknown on the right-hand side to compute the solution at this particular step. So we carry out our computations, say we perform 50 iterations and see what our solution is using Gauss-Seidel method. So we present the details in the following table.

(Refer Slide Time: 31:40)




Let us now look at the table for the solutions that we have computed for using Gauss-Seidel method we have performed 50 iterations. And if we still require, say 2 decimal place accuracy for our solution, we observe that this has already been obtained at the 20th iteration. So correct to 2 decimal places, the solution is 0.20 for X1 and 0.27 for X2, whereas for Gauss-Jacobi method, we could achieve that accuracy only at the end of 50 iterations. So Gauss-Seidel method and Gauss-Jacobi method, both generate a sequence of iterates for the solution of the system of equation AX equal to b which converge to the actual solution of the problem.

Gauss-Seidel method converges must faster than Gauss-Jacobi method, the reason being it takes into account the currently available solution at any particular step for the unknowns which have been already computed, as a result the method converges faster. So let us take another example where we have 3 equations into the unknowns and solve the system by both the methods and compare the solutions.