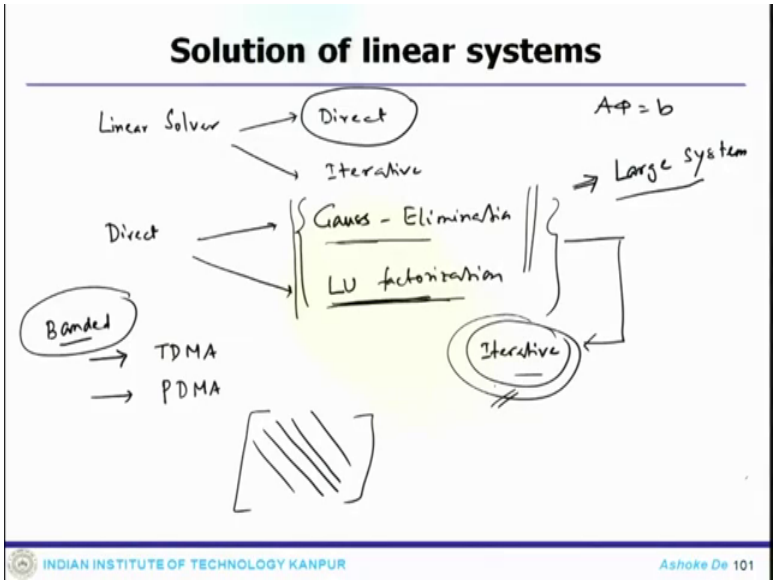
Introduction to Finite Volume Methods-II
Prof. Ashoke De
Department of Aerospace Engineering
Indian Institute of Technology, Kanpur

Lecture – 03
Linear solvers – III

Welcome back and you are continuing our discussion or the Linear solver and for this Finite Volume series. And so far what we have done till the last class actually we have looked at the direct method for solution of the linear system and under that what we have discussed? We have discussed Gauss elimination, we have discussed LU factorizations and then some special cases like tridiagonal matrix system and penta diagonal matrix system and now today we are going to start with the iterative process.

(Refer Slide Time: 00:53)



So, what summarize to that we have this linear solvers; which essentially we are solving for A phi equals to b; one is direct approach another one is iterative approach. So, this is what we have discussed in the previous class and under direct approach we have done Gauss elimination and LU factorization.

And we have looked at how many number of operations one need to do either of these two processes and what we have seen LU factorizations requires more number of operation, but it provides you some sort of a better flexibility compared to gauss elimination process. And this can be used if the coefficient matrix is not changing; it can

be also used for the different right hand side vector b. Top of that, we did some discussion on these things because these are going to fit to the or pave the way for our discussion with the iterative solvers.

So, as we go along with the lecture we would be able to see how this factorizations or elimination process that which we discussed under the direct approach would be handy for the iterative process. Third we have talked about couple of special cases that is the banded system.

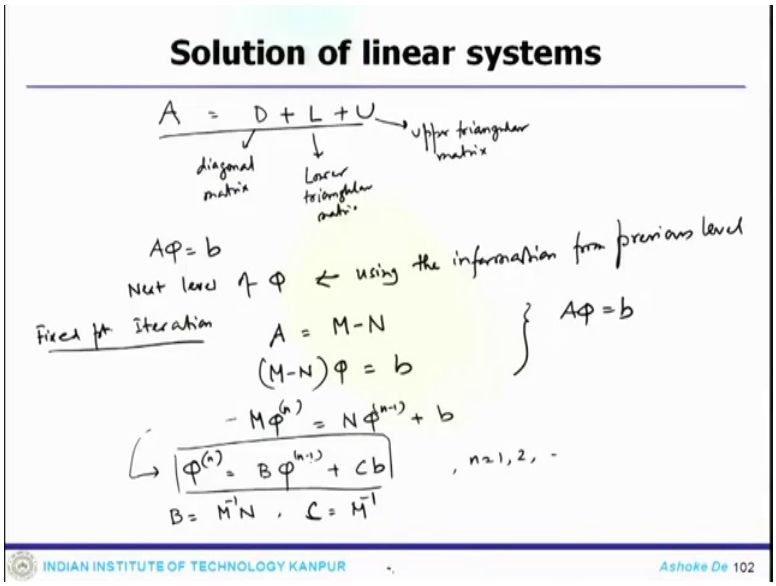
So, under the banded system what we have discussed it the TDMA which is called Thomas algorithm and tridiagonal matrix algorithm and then the one which other one PDMA; that means, the penta diagonal matrix algorithm and these are all applicable for banded system. So, one case is a kind of 3 elements, other case the 5 elements.

So, these are for banded matrix that we have discussed. So, the essentially when you talk about the iterative process; it is advantageous over direct method because here it is computationally expensive for specially every time we keep on reiterating the fact for the large system; if you have a small system then direct or iterative process may not provide you a different overhead.

But if you are talking about the iterative process I mean large scale system then this guy; the iterative process which is very often used in any CFD code and in naturally this is preferred because it has a very less computational overhead.

Now, other thing which is important is that there are different kind of or different families of iterative methods. And they can be found in literature and what we have used as a direct method; they can be they were actually introduced for the sole purpose of clarifying some fundamental processes needed for the iterative method; which I have already said these are the some sort of platform which are going to be used or invoked while talking about the iterative process.

(Refer Slide Time: 04:40)

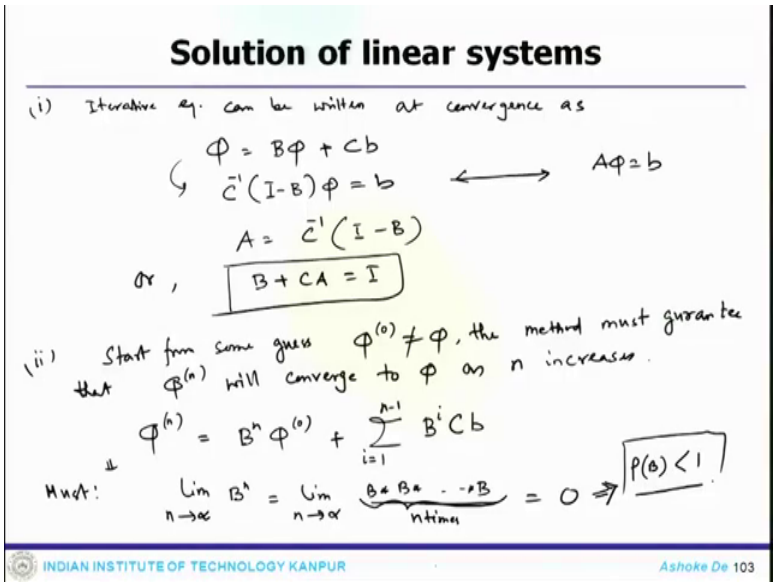


So, what we do in the iterative process? So, one of the important thing is that how we handle this coefficient matrix A. And there the thing that will decompose that in D plus L plus U; where D stands for diagonal matrix, this is diagonal matrix and this one for lower matrix or lower triangular matrix; lower triangular matrix and this is the upper triangular matrix; upper triangular matrix. So, these are the 3 different decomposition that one can do.

Now, when you solve for the A phi equals to b. So, what you are trying to find out you go in a different physical iteration and you try to find out the next level of phi using the information from previous level. So, it could be physical iteration, it could be specially and so that is what one can use to find out the; now if you talk about fixed point iteration.

In fixed point iteration; the decomposition matrix A which can be decomposed like M minus N. So, my M minus N phi that would we b because my original system is A phi equals to b. So, if we use these two information this becomes that. Now, the fixed point iteration I can write M phi n equals to N phi n minus 1 plus b; which one can rewrite like phi to the power n or the nth level phi equals to B phi n minus 1 plus C b; where n goes from 1, 2 like that and B is nothing, but M inverse N and capital C is M inverse. So, essentially from here you multiplied with M inverse; so, this guy becomes phi n equals to B phi to the power n minus 1. That means, at the nth level of phi is obtained using the matrix B and phi n minus 1

(Refer Slide Time: 07:58)



Now, before moving ahead on the different description of the iterative methods some minimum characteristics that one should know and number 1 that is the iterative equations; iterative equation can be written at convergence as phi equals to B phi plus C b. So, after rearrangement this becomes; so you do the some sort of and rearrangement, it becomes C inverse identity matrix minus B phi equals to b.

So, if you compare this one; this one if you compare with your original system then one can write A is nothing, but C inverse I minus B or alternatively one can say B plus CA is an identity matrix; so that is one of the property. So, this relation between the various matrices ensure that once the exact solution is reached all consecutive iteration will not modify it.

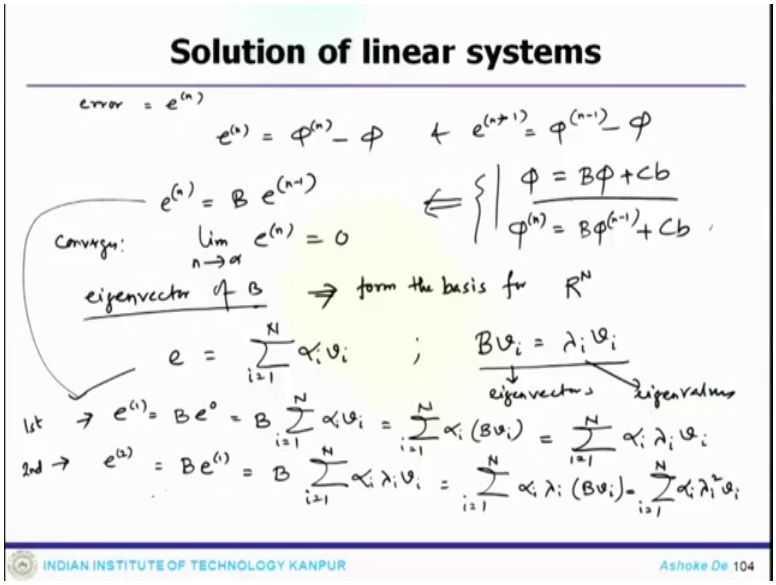
Now, the second point is that you start from some guess; let say phi 0 which is not equals to phi. The method must guarantee that phi n will converge to phi as n increases.

Since phi n can be expressed in term of phi 0; then one can write that phi n equals to B n; phi 0 plus summation of i 0 equals to n minus 1; B to the power i C b. So, in this particular expression; we must satisfy this. So, this is must, we must satisfy this limiting condition; if n tends to infinity, B to the power n must be B into B dot dot dot B divided by; so this is n times must be 0; so B must satisfy that.

So, what does it imply when you have this kind of complete I mean condition. This implies that the spectral radius of B, this is in term which we have already discussed in our proceeding lectures the under the properties of linear system. The spectral radius of B is less than 1; so, this is the condition that it satisfy.

Now, this condition guaranties that the iterative method itself is correct it. And it is also robust enough to any error which inserted into the solution vector of B. Now, one can look at slightly in a different inside can be obtained by defining some sort of an error.

(Refer Slide Time: 12:00)



And now if you define this kind of error in the solution; then one can say error at nth level is must be phi n minus phi. So, theoretically this error should be minimized or going towards 0; when the solution converges.

Now, if you use these things with your original system of equation; then one can write that e to the power n equals to, so, previously what we have obtained is that phi equals to B phi plus C b. So, now using this and you write this expression also you have got this phi to the power n equals to B to the power n minus 1 plus C b.

So, using these two; one can write that error could be written as B e to the power n minus 1. So, when it converges what happened? The limit condition n tends to infinity the error must be 0. So, in order to take this expression to some meaningful conclusion; one has to look at the eigenvectors of B. So, the eigenvectors of B are assumed to be complete and

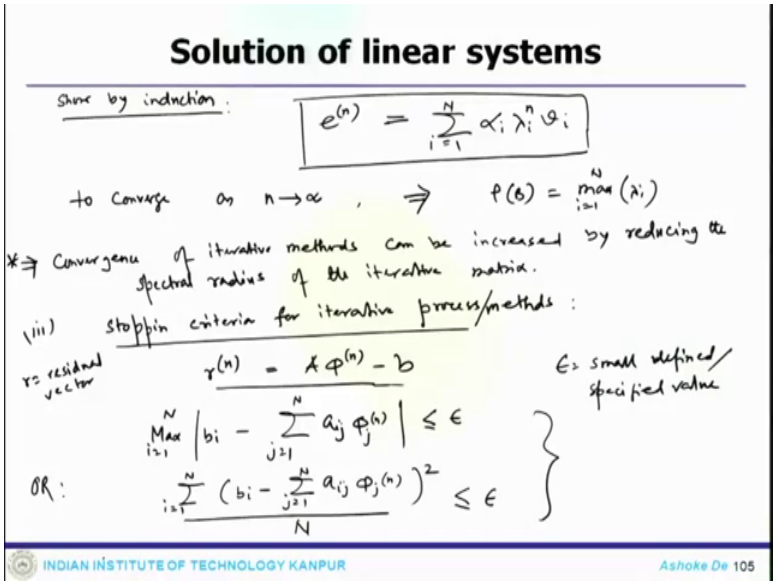
to form a full set; meaning they form the basis for, so they essentially form the basis for R to the power N.

Now, this being the case then e can be expressed as a linear combination of the n eigenvectors v of b. So, B has n eigenvectors; so one can write error is some sort of i goes to N alpha i small v i; where each of the eigenvectors satisfy B v i equals to lambda i; v i. Here B are the eigen vectors and lambda are the eigen values; so, one can now write this things.

Now, once you combine with this one e to the power n of the nth level error is equals to B e to the power n minus 1; then combining these two, one can say my error at the first level is B e to the power of 0; which is B summation of i equals to 1 to N alpha i v i which is i equals to 1 to N alpha i B v i which also can be written as i 1 to N alpha i lambda i v i.

Now, this is for first iteration; so second iteration if you move; that means, the second iteration level, the error is e 2; which is B e 1 equals to B i equals to 1 to N alpha i; lambda i; v i equals to i equals to 1 to N; alpha i, lambda i B small v i; which is summation of i equals to 1 to N; alpha i lambda i square v i.

(Refer Slide Time: 16:55)



Now, this can be continued like third iteration, four iteration and so on and one can finally, obtain or show by induction is that or the finding the spectral radius; you express

this one for the nth level, this is summation of i equals to 1 to N; alpha i, lambda i to the power n small v i. So, that is what one can write for n iterative process. So, this means this small n stands for the number of iteration requires for this particular process to be or method to be converged.

So, the for the iterative procedure to converge as n approaches to our infinity or rather one can think about for a large values of n; this particular process converges and then error gets minimized. All the eigen values should be less than 1; if any one of them is greater than 1, then the error will tend to infinity.

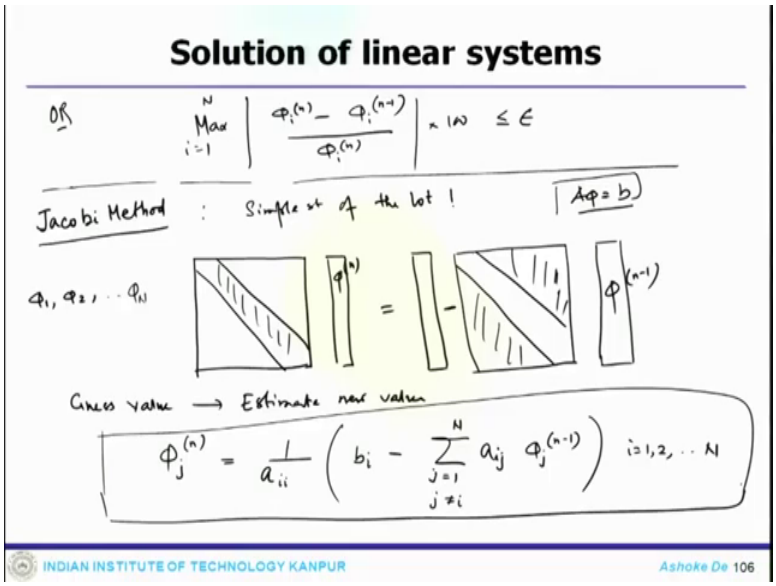
So, which means my spectral radius of B should be less than 1. Because lambda corresponds to the eigen values for that matrix B. Now, we essentially the convergence of iterative methods can be increased by reducing the spectral radius of the iterative matrix. So, B is here the iterative matrix; so this essentially the heart of any iterative process.

Now the third important point is the some sort of an stopping criteria is very much necessary for iterative process. So, that brings the stopping criteria for iterative process or iterative methods. So, which dictates or tells you where to stop or when to stop; so very often people use that based on the variation of the norm of the residual error. Like the residual error defined as r to the power n; r is the residual vector and this error comes as A phi to the power n minus b.

So, one criteria is to find the maximum residual in the domain and to require its value to become less than some threshold value or user defined values small value epsilon. So, for example one can do like that the maximum of i goes to N bi minus summation of j 1 goes to N; aij phi j which less than epsilon. So, epsilon is a small defined or specified value.

So, that is one or one can see or one can find out the alternatively the root mean square residual might be smaller than some epsilon value. So, in that case it could be i 1 to N bi minus summation of j equals to 1 to N aij; phi j square divided by N less than equals to epsilon. So, either of these criteria can be used for the convergence limit.

(Refer Slide Time: 21:59)



So, another possible criteria is for the maximum normalize difference between two consecutive iterations; which can be dropped below the epsilon which means or alternatively max i goes to N; phi i n minus phi i n minus 1 divided by phi i n; the percentage below some value.

So, these are the points for any iterative process or method to have or rather the heart of this process which needs to be satisfied. Now said that we will now move on to the discussion of different iterative methods; so, the one first one to start with is the Jacobi method. So, this is again probably the simplest of the lot and; one can solve the linear system by Jacobi method to get the solution done.

Now, if you see or graphically if someone represent this; then I can define this one like; these are my process. So, which will be like this and so you have some bandedness; so this is your phi n; then this minus, this is phi n minus 1; so that is what a graphical representation shows like that, that how you get this things.

And the equation that we are solving is A phi equals to b. So, you consider this equation and says that if the diagonal elements are non zero; then the first equation can be used to solve is phi 1 and then the second one is phi 2 and so on.

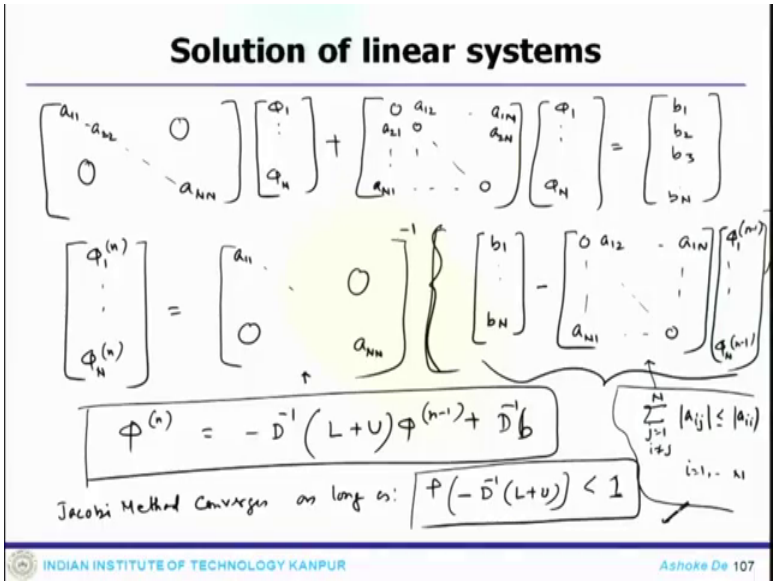
So, the solution process actually starts by assigning some guessed value to the unknown value of phi and this gauss values are used to calculate the estimate. So, essentially you

start within gauss value; then you estimate the new value and then you move on to the iterative process to get and converge solution and this one you keep on doing till you get the solution is converge.

So, essentially first level you assume some gauss value and obtain the other values in the domain. And then using those value; you correct it and check whether the solution is converged not. So, one can write this in a compact form that at level nth level the j value would be 1 by a ii b i minus j equals to 1 to N; where j not equals to i a ij; phi j n minus 1 where i goes from 1, 2 to N.

So, this is how for any i; one can find out this process.

(Refer Slide Time: 26:42)



Now once you get this generic equation; then you can actually form some sort of an matrix. Like, I can have a 1 1, like a 2 2 and so on; is in a N N. So, this as a diagonal matrix and the rest of the elements are 0; which is multiplied with phi 1 to phi N plus, I can have 0 on the diagonal and I can have like a 12 so on a 1 N; here a 21; a 2 N so on a N 1 like that; which is also multiplied with phi 1 to phi N that is b 1, b 2, b 3 dot dot b N.

Now, once you solve for phi N; you get phi 1 n to phi N at n equals to; this guy a 11 or the diagonal element inverse; multiplied with multiplied with b 1 to b N minus the; this equation which is 0 in the diagonals. This will be 1 2 to a 1N and this would be a N1 to like that into 1; n minus 1 dot dot phi N; N minus 1; so that is the term.

So, the complete matrix is multiplied with this factor. So, this is nothing, but so if one has to write this one can see it is a diagonal matrix; this only the diagonal elements are 0. So, if you write in more compact form or in the matrix form; so one can write phi to the power n is minus D inverse; L plus U phi n minus 1 plus D inverse b; so that is for any quantity.

Now, this particular method or the Jacobi method; it converges as long as it satisfied the spectral radius of minus D inverse; L plus U less than 1. So, as long as this condition is satisfied, this particular method converges nicely. So, this is a condition for large class of matrices including diagonally dominant ones; this coefficients actually satisfy that.

So, they need to satisfy I mean this condition is that summation of j 1 to N; i not equals to j A ij less that equals to a ii; where i goes from 1, 2 to N; so, this is the condition that has to be satisfied. So, we will stop here today and we will take from here in the follow up lectures.

Thank you.