Finite Element Method and Computational Structural Dynamics
Prof. Manish Shrikhande
Department of Earthquake Engineering
Indian Institute of Technology, Roorkee

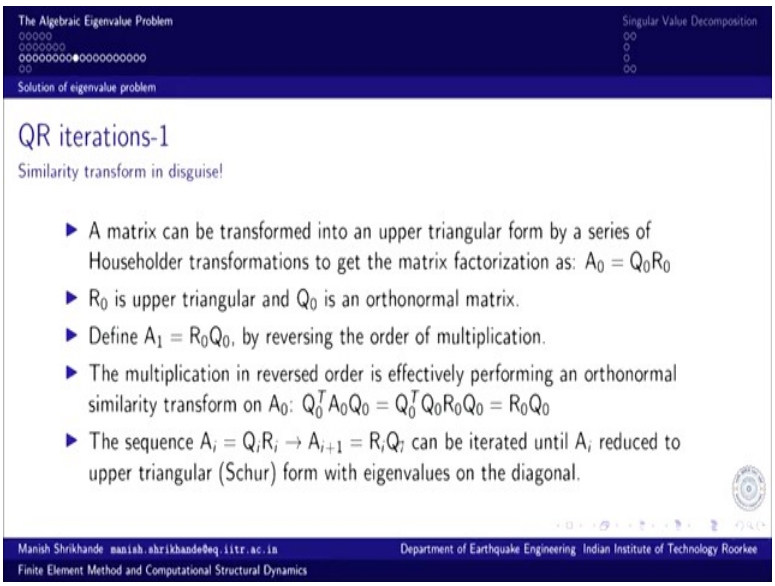
Lecture - 50
The Algebraic Eigenvalue Problem - IV

Hello friends. So, we have seen some of the methods of computing eigen values and eigen vectors, where we targeted one eigen pair at a time and then having computed one eigen pair at a time; at a time we converge to the next eigen pair by using some sort of orthogonality or purification technique or even shifting procedure and we could compute the other eigen vectors other the next eigen vectors in the spectrum.

So, now we look at some of the algorithms which are used for more large scale eigen value problem and first one very powerful and very popular technique for I mean large scale solution of eigen value problem is what we call as QR iterations. In numerical analysis or numerical techniques books, it is often referred to as HQRI or H standing for householder transformation. So, and QR refers to the factorization.

So, using householder transformation matrix, the reflection matrix we convert this matrix a into a factor of orthogonal matrix Q and a right triangular matrix. So, R because this is a matrix on the right of the matrix Q and it is it has a triangular structure.

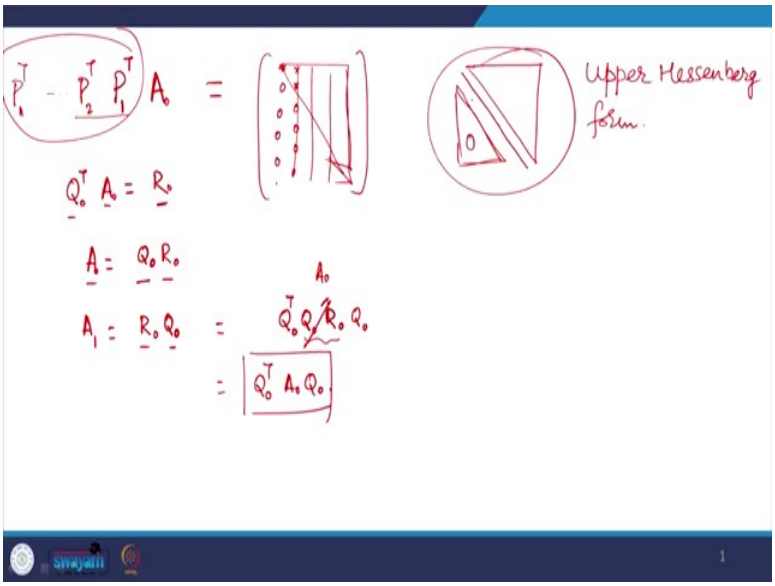
(Refer Slide Time: 02:10)



So, matrix Q is actually generated from a such that this will reduce this entire system reduces it to upper triangular matrix. So, A is converted to upper triangular form. So, a matrix can be transformed into upper triangular form by a series of householder transformation. So, its very efficient to convert matrix a given matrix to upper triangular matrix by using householder transformation and that accumulation of the transformation that sequential accumulation that is what is we call as the transformation matrix that is related to Q.

So, R_0 is upper triangular and Q_0 is an orthonormal matrix and this orthonormal matrix is related to the householder transformation gradually.

(Refer Slide Time: 03:15)



So, essentially what is done is a if I multiply this householder transformation P_1, then I will get this matrix and 0 0 0 0 here and then of course, there are other rows that are there. So, it can be done like this and then once I do P_2^T , it will target this matrix and there will be elements here and then there will be 0s all along.

So, these matrices the product of these, this is what refers to. If I look at it, then $Q_0^T A$ = R because this is what is going to be converted to upper triangular matrix and if I multiply Q_0^T is of course orthogonal matrix. So, A is given as Q_0 R_0 let us say A_0, right. So, this is what the entire factorization is. A can be defined as a orthogonal matrix Q_0 multiplied by right triangular matrix R_0.

Now if I define another matrix let us say A_1 as I reverse the order, what happens R_0 Q_0 having computed QR factorization, I just interchange the order of matrix multiplication and define a new matrix. So, what is this matrix? This matrix I can write this as $Q_0 Q_0^T R_0 Q_0$ because this is a orthogonal matrix. So, $Q_0 Q_0^T$ is of course identity matrix. So, when I do this, then this is actually orthogonal transform and sorry I would. So, I define this $Q^T Q$ and $R_0 Q$ now $Q_0 R_0$ what is $Q_0 R_0$. $Q_0 R_0$ is just A_0 .

So, this essentially this by reversing the order of multiplication what I am eventually doing is implementing a similarity orthogonal similarity transform of matrix A, very efficient technique. So, this similarity transform would require multiplication. Ideally if I do implement it, then it will require multiplication of 3 n by n matrices and here I am doing the same operation by just multiplying two matrices A considerable saving in numerical effort.

So, we define new matrix by reversing the order of multiplication. So, A_0 is the factorization by series of householder transformation, we get $Q_0 R_0 A_0 = Q_0 R_0$ that is the factorization and then I define new matrix A_1 by just reversing the order of multiplication R_0 Q_0 and multiplication in reverse order is effectively performing an orthogonal similarity transform by virtue of identity that we have.

And the sequence we can keep on doing this repeatedly and we can do this $A_i = Q_i R_i$ and this new matrix will lead to R_i triangle multiplied by Q_i by interchanging the order and so on. The matrix factorization and reversed operation reverse multiplication can be continued until A is reduced until this matrix A that we have is reduced to upper triangular form completely upper triangular form.

And that will have that is called schur form with eigen values on the diagonal of the matrix and once we have the eigen values, then it can be we can go back to inverse iteration with shift parameter to compute refined eigen value and the corresponding eigen vector and generally, one iteration is good enough to compute the corresponding eigen values and eigen vectors. So, in this case by the Q R iterations, we can actually converge to the eigen values on the diagonal of the matrix. So, this diagonal matrix that we generate.

So, in practice that is the theory that in for Q R iteration, in practice we actually first consider the householder transformation to convert matrix A into upper Hessenberg form. So, upper Hessenberg form is like upper triangular structure with one sub diagonal and rest of it is all 0.

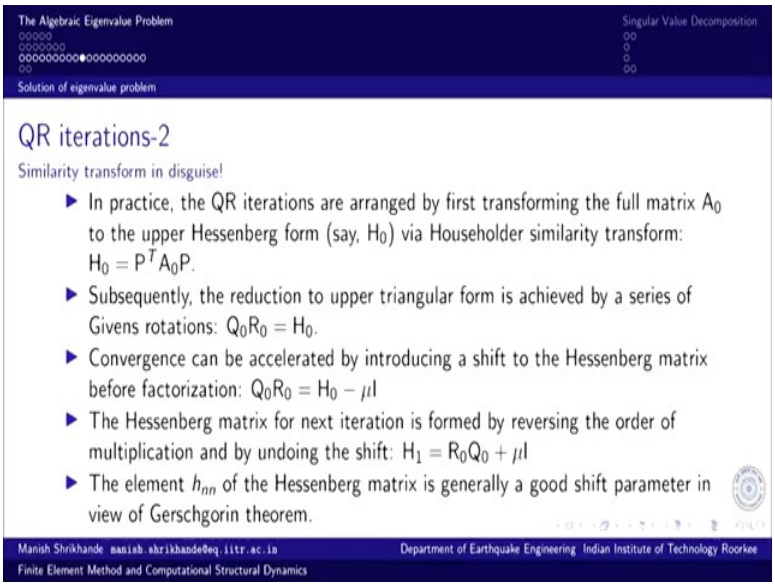
So, this is all 0. So, this structure is what is known as Upper Hessenberg Form. So, we convert the matrix instead of converting into right triangle upper triangular straight away from using householder transformation, we just use it for converting it to upper Heisenberg form. Now why do we do this upper Heisenberg form? We do this because not all eigen values are guaranteed to be real in general case.

So, if they are a complex eigen values in some cases, then those complex eigen values must occur in pair and those complex eigen values would occur would appear I mean in this real matrix, they would occur as two by two sub blocks on this upper Heisenberg form. So, this upper Heisenberg form is capable to also represent the possibility of complex eigen values being there in the matrix.

So, of course for symmetric matrices they are now we do not have complex eigen values, but for a general unsymmetric real matrix. We can have eigen values that are complex and for that purpose instead of looking at upper triangular factor, we look at matrix transformation householder transformation to reduce it to upper Heisenberg form and when we do this for Q R iteration.

Subsequently if there are any complex eigen values, they would appear as 2 by 2 sub blocks diagonals on this upper Heisenberg form and the eigen values would be obtained from the from solving the 2 by 2 eigen value problem and the complex conjugate eigen values can be recovered from there.

(Refer Slide Time: 11:29)



So, subsequently we try to once we have this upper Heisenberg form, then there is only theoretically there is only one off diagonal term in each column that needs to be eliminated. So, the QR factorization is achieved by using givens rotation subsequent givens rotation. So, we have we accumulate givens rotations and convert this H upper Heisenberg form into a QR factorization and then reverse the order to get the next upper Heisenberg form.

Now, convergence can be accelerated by introducing a shift to the a Heisenberg form before factorization. So, instead of H_0 we can factorize $H_0 - \mu I$ where μ is the estimate of an eigen value. Now, Heisenberg matrix for the next iteration is obtained by reversing the order of multiplication and by undoing the shift in the that was implemented in the first iteration.

And then once we have this H_1, then again $Q_0 R_0$ will again be computed by having new shift parameter depending on what is the new H_1. So, this shift parameter is generally taken as the n n element of Heisenberg matrix and it is generally good shift parameter in view of Gershgorin theorem because there is only one off diagonal term.

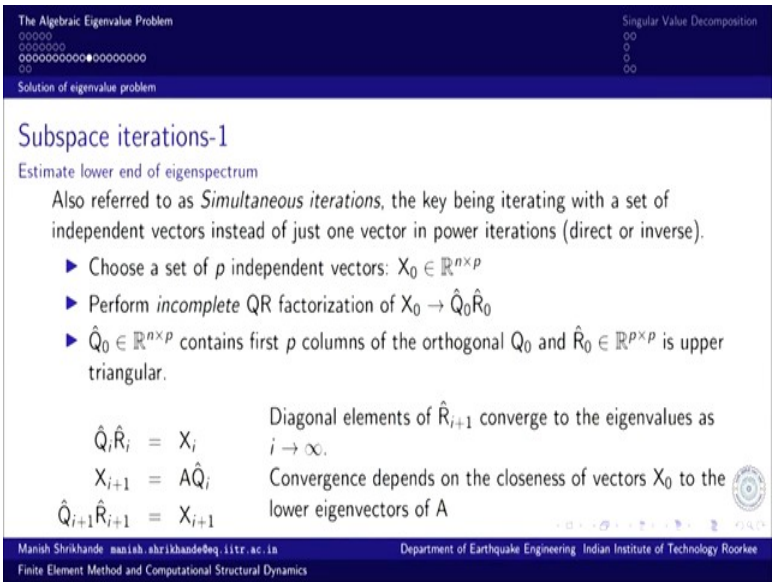
So, there is an eigen value which lies in between h_nn and + or - the off diagonal term. So, h_nn is a generally a good estimate of the of an eigen value and this shift parameter it is a good choice for shift parameter, but of course this is only for when eigen value is real for complex eigen value. The procedure becomes little more complex in the sense

that we now look at need to look at 2 by 2 sub diagonal 2 by 2 block diagonal elements and implement that in the shift parameter.

So, that procedure is a little bit involved, but that can also be taken care of in numerical operations. So, this is work hours and householder QR iterations. They generally work very well and they need to be followed up with one cycle of inverse iteration with shift parameter to compute a really very good very accurate eigen value and corresponding eigen vector. So, this QR iteration, they will provide us with the eigen values of good approximation of the eigen values along the diagonal of the triangular matrix that we eventually get by repeated iterations. Now we look at what is known as the sub space iterations.

Sub space iterations is an extension of power iteration. It is also referred to as simultaneous iteration and just as in power iteration or inverse iteration, we dealt with one vector one trial vector in sub space iteration. We actually deal with we actually use a set of independent vectors for iterations.

(Refer Slide Time: 15:28)



So, instead of just one vector in power iterations direct or inverse, we choose a set of p number of independent vectors in real number space and then perform incomplete QR factorization. So, X_0 there are p number of vectors in this X_0 matrix. So, it is n by p matrix. So, this n by p matrix we can perform incomplete QR factorization in the sense that this is Q_0 is of course n by n and R_0 is n by p matrix.

So, R_0 will be upper triangular p by p and rest of the rows n - p rows. They are all null null matrix entirely null matrix. So, Q_0 that contains first p columns. I mean $\hat{Q}$ contains the first p columns of the orthogonal q matrix and $\hat{R}$ is p by p. That is upper triangular matrix upper triangular I mean rest of the things n - p rows are all 0s. So, Q_0 is the orthogonal n by n matrix and hat Q is first p columns of this orthogonal matrix. So, eventually we perform these iterations.

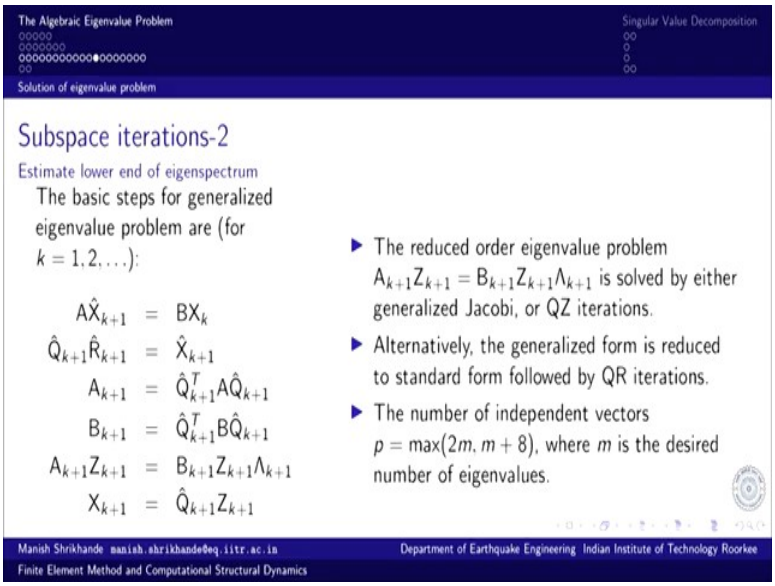
So, X_i is given as this decomposition Q_i R_i and then, we define new set of trial vectors X_{i+1} next iteration as matrix A operating on this incomplete orthogonal matrix Q hat i and once I obtain this new set of eigen trial vectors X_{i+1} . I again decompose them into factorization Q_{i+1} multiplied by R_{i+1} hat and again go back to the system.

And diagonal elements of $\hat{R_{i+1}}$ they converge to eigen values as I tends to infinity as iterations progress. The diagonal elements of this triangular matrix R they would converge to the eigen values and once we have the eigen values, then again the one cycle of inverse iteration will lead to a very accurate eigen value and corresponding eigen vector or we can just substitute in the this eigen value into the governing eigen value problem statement and compute the corresponding eigen vector directly.

So, convergence depends on closeness of vectors X_0 to the lower eigenvectors of A. So, this choice of vectors that we have if they are in some way representative of the lower end or the eigen vectors corresponding to smaller eigen values of matrix A, then the convergence is going to be very very rapid.

So, the basic steps I mean for generalized eigen value problem what we discussed here is for these standard eigen value problem $Ax = \lambda x$ for generalized eigen value problem things become little more involved in the sense that there are two matrices and the trial vectors need to be computed as a solution first.

(Refer Slide Time: 19:27)



So, for starting with X_k trial vector, we find what is B times X_k and then, find what is the new trial vector $\hat{X_{k+1}}$ going to be by solution of this system of equations and once we find this $\hat{X_{k+1}}$, we try we compute its incomplete Q R factorization and this incomplete Q R factorization, we impose similarity transform on A.

So, $\hat{Q_{k+1}^T} A \hat{Q_{k+1}} = A_{k+1}$ and similarly $\hat{Q_{k+1}^T} B \hat{Q_{k+1}} = B_{k+1}$. Now you look at it these are all small smaller size I mean A_{k+1} and B_{k+1}, they are in general smaller in size than A and B because these are incomplete factorization.

So, while A and B are matrices of size n by n A_{k+1} and B_{k+1} would be matrices of size p by p which is generally very small compared to n by n and then this small system of eigen values, eigen problem is solved by suitable technique whichever technique we may adopt. So, this smaller eigen value problem.

Now, this small eigen value problem it is what we need to appreciate that we are looking at scaling down by several orders of magnitude. While it is common in case of typical finite element problem that A and B would be of the order of several millions in the size while these matrices A_{k+1} and B_{k+1}, these are we are referring to only of size a few 100 by 100.

So, that is a tremendous saving or reduction in problem size and this eigen value problem this is the reduced order eigen value problem which can be solved and once we can solve this problem, then we find new set of eigen vectors as $\hat{Q}$ times the eigen vectors of this smaller eigen value problem, right and then again the entire cycle begins from this step again.

So, again we multiply B with this X_{k+1} . Find out what is $\hat{X_{k+1}}$ by solution of simultaneous equation and repeat these process again. So, reduced orders eigen value problem that we had been looking at, they can be solved either by generalized Jacobi method.

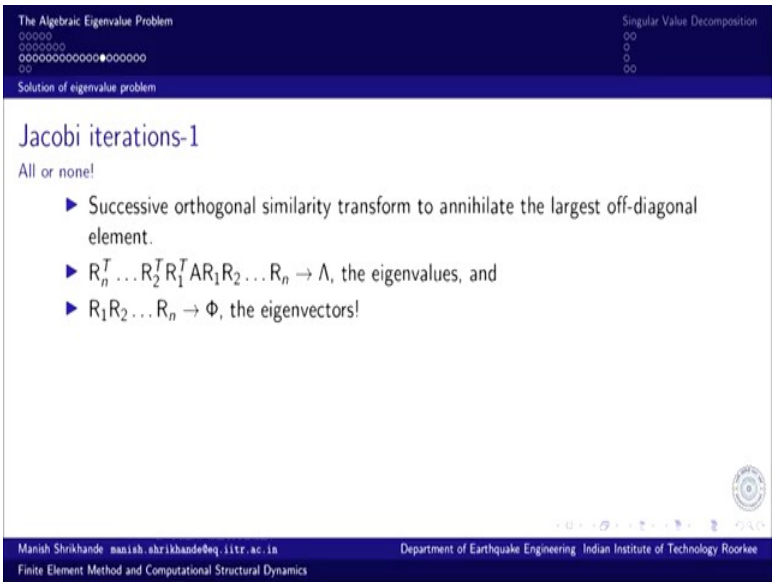
We will discuss that generalized Jacobi which is actually series of givens rotations, so that the matrix is nudge to diagonal form or QZ iterations again similar to QR iterations, but QR iterations refer to standard eigen value problem, QZ iterations refer to generalized eigen value problem where we perform decompositions on both the matrices. So, alternatively the generalized form is reduced to a standard form. As we discussed earlier matrix B can be converted to Cholesky factorization LL^T and then the problem can be transformed into standard eigen value form.

And then it can be solved by using QR iterations. The guidance of how many numbers that we need to take the vectors is given by maximum of 2 m or m + 8, where m is the number of desired eigen values. So, as I said the size the matrix A and B can be of several millions of degrees of freedom and accordingly, there are going to be millions of eigen values, but not all these eigen values are required.

So, we can work with very few eigen values and may be only a few hundred. So, those 100 eigen values. So, if I need 100 eigen values, then the number of eigen the columns independent vectors that I need to generate for these using subspace iteration is only going to be twice m. So, 200 vectors that I need for completing the problem and the size of the problem that I will need to solve eigen value problem would only be 200 by 200 instead of several millions.

So, that is how the sub space iterations are work arranged and then Jacobi iterations is as I said it is successive orthogonal similarity transform by using givens rotation to annihilate the largest off diagonal element.

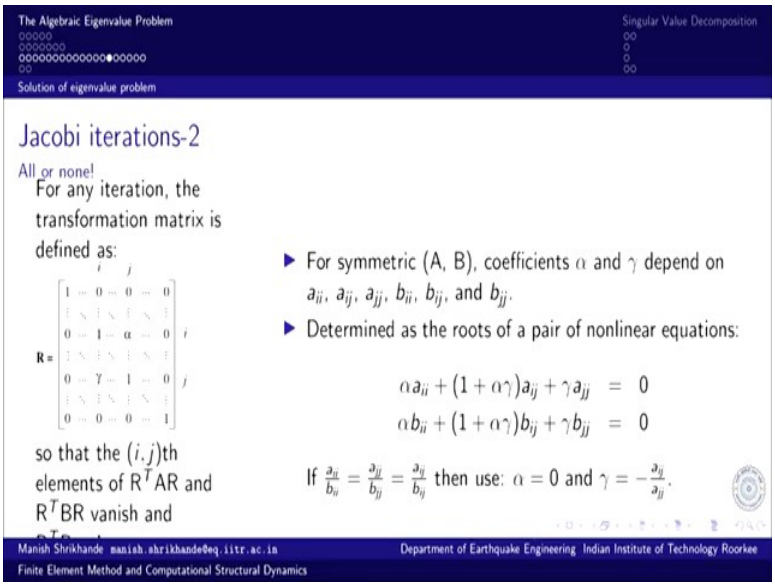
(Refer Slide Time: 25:13)



And that is what we do and we do it from both sides until the matrix A converges to the diagonal matrix which are the eigen values and interesting part is accumulation of these product matrices orthogonal matrices then converges to the eigen vectors and that in the sense is the Jacobi iteration. So, it is essentially a sequence of givens rotations and we can find out what is the iteration and how to arrange for that and for generalized eigen value problem, there can be generalized Jacobi transformation again using similar rotation matrices on both the terms both the matrices A and B.

And such that the they are progressively nudged towards diagonal, both of the matrices and eigen values are obtained as the ratios of the diagonal corresponding diagonal elements and that in a sense is Jacobi iteration and the problem is Jacobi iterations they will provide us complete eigen spectrum as on in one go. So, it is all are none. So, either we get the complete eigen values and eigen vectors or we do not get any.

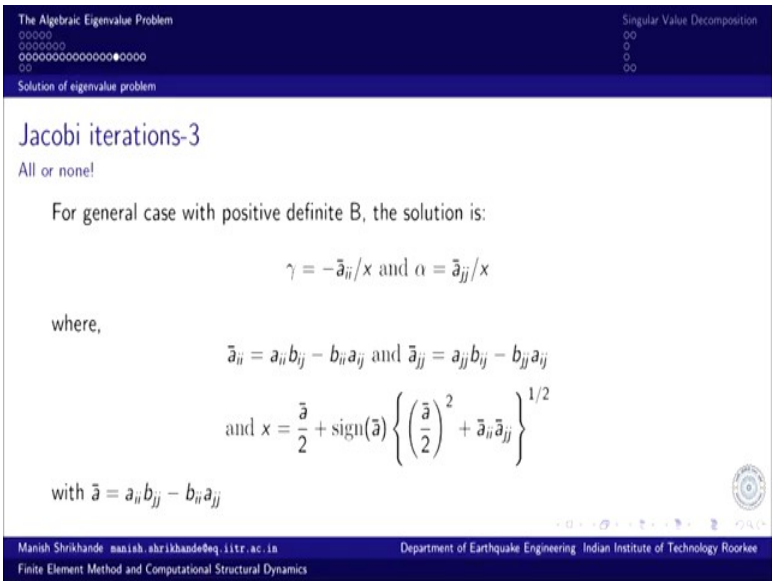
(Refer Slide Time: 26:32)



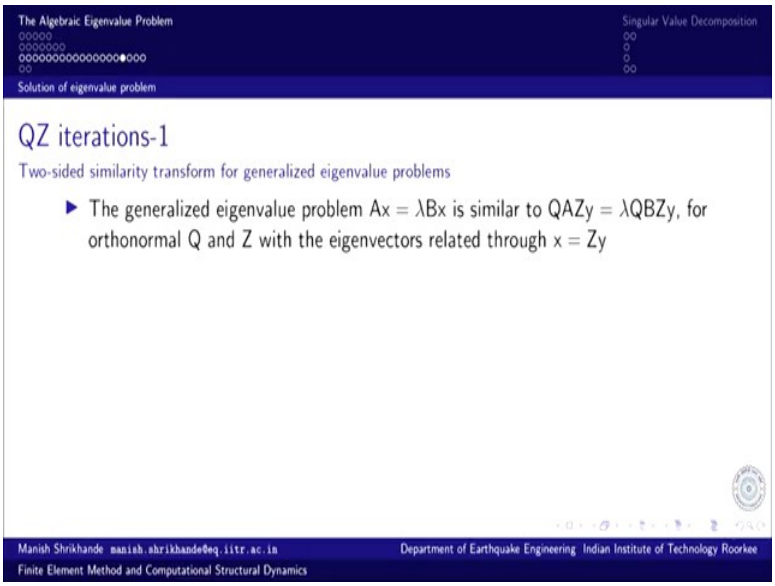
So, that is the problem and for this is the similar arrangement for Gibbons rotation, but for generalized eigen value problem there are there is standard algorithm that can be arranged and coefficient rotation matrices or orthogonal transformation matrices can be defined as modification of the identity matrix.

You can see that there are two elements which differentiate, which make this rotation matrix different from the identity matrix and these coefficients are evaluated from the solution of these equations.

(Refer Slide Time: 27:14)



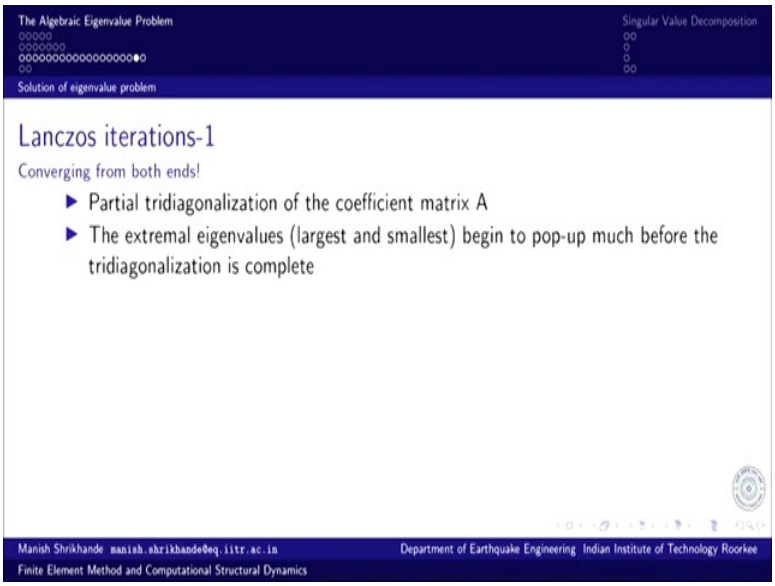
(Refer Slide Time: 27:30)



So, eventually this is what we have for definite positive definite Bt. The solution can be obtained for these non-linear equations and similarly QZ iterations, they are similarity transforms two sided similarity transform similar to QR iterations, but these are applicable to generalized eigen value problem. So, this advantage is these QZ iterations. They are robust numerically even if matrix B is near singular, all other algorithms that we have discussed for generalized eigen problem, Jacobi rotations etcetera they work only when matrix B is positive definite.

But QZ iterations would work even if the matrix B is nearly singular or not a very well conditioned. So, simultaneous reduction of A and B to upper triangular form and that is what we call as generalized schur form and then the eigen values are obtained from the ratio of diagonal terms of the respective diagonal elements and once we have these eigen values, then the corresponding inverse iteration can be computed and associated eigen vectors can be generated.

(Refer Slide Time: 28:44)



So, with that we stop the discussion and Lanczos iterations are again related to Krylov subspace that we had encountered in the discussion of Conjugate Gradient method and they are essentially related to converting it to tridiagonal form and then using the similar orthogonal transformation and then it begins to converge from both ends.

We get the eigen values from the lower end of this spectrum as well as from the higher end of this spectrum and both extreme eigen values can be obtained gradually and then it moves to the center. So, in some applications Lanczos iterations are quite useful, but for most of the cases we will find that QR iterations and sub space iterations, they more than serve the needs of the computation that we need.

So, with that we end our discussion on eigen value problem. Of course, this is a very vast subject and I just skimmed at the surface of this and hopefully that should allow you to explore this particular subject in more detail on your own. In next lecture, we start with Integration of Equations of Motion in Time.

Thank you.