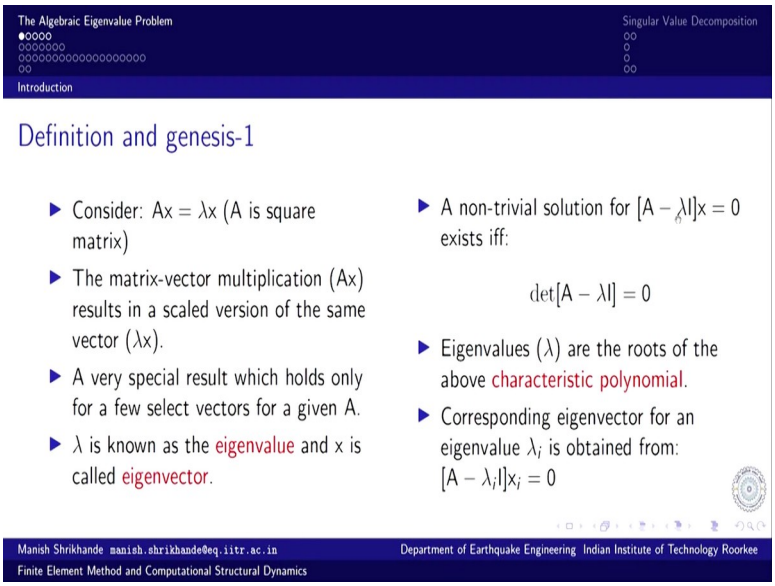
Finite Element Method and Computational Structural Dynamics
Prof. Manish Shrikhande
Department of Earthquake Engineering
Indian Institute of Technology, Roorkee

Lecture - 47
The Algebraic Eigenvalue Problem - I

Hello friends. So, today we start with a new topic very important topic which has applications in all fields of science and engineering in the importance or in the nature of prevalence how prevalent this topic is finds its usage in a different fields. It probably ranks only next to equation solver solution of simultaneous equations.

And what we are referring to is the algebraic eigenvalue problem. And singular value decomposition is a related problem which can be cast as an algebraic eigenvalue problem as well, but it is more efficient to solve in its own formulation.

(Refer Slide Time: 01:25)



So, let us define what is understood by what is meant by algebraic eigenvalue problem and what it means. So, essentially the algebraic eigenvalue problem is defined as Ax = λx where a is a^2 matrix x is of course, a vector and λ is a scalar. So, what this means is we are looking for that particular vector x which when multiplied with a matrix given matrix A results in a scaled version of its own self; that is the vector does not change its direction it only changes its magnitude.

Now, that is a very very special property. Normally, when a matrix multiplies a vector that implies a rotation in plane. So, in this particular case, we are looking for that particular vector for a given matrix which when multiplied with this matrix does not change its direction, but it only is affected it only changes its length. So, the whole idea the formal statement is matrix vector multiplication that is A times x results in a scaled version of the same vector.

A very special result which holds only for a few select vectors for a given matrix A and this does not happen for any arbitrary vector if a matrix A is given. And λ is known as the eigenvalue of this matrix A and associated with this eigenvalue is the vector x that is called the eigenvector.

So, it is this pair eigenvalue eigenvalue and eigenvector; the computation of this pair of eigenvalue and eigenvector I mean there can be several such pairs for a given matrix. So, this process of a computation of these eigenvalues and eigenvectors is what is referred to as the algebraic eigenvalue problem.

Now, a nontrivial solution for this $Ax=\lambda x$ exists only when determinant of $A-\lambda I=0$. So, that is when I take this λx to a left-hand side.

So, this is what I get $A-\lambda I x=0$. So, the solution of this when I say nontrivial; that means, I am looking for vector x which is not a null vector. So, for an non trivial solution x cannot be a null vector, then the only way this equation can hold is when this matrix becomes singular and that is when determinant of $A-\lambda I=0$. So, when we expand this determinant. So, this will be an nth degree polynomial in terms of %lambda.

So, these there will be n number of roots of these polynomial nth degree polynomial. So, those roots of this characteristic polynomial. This is called as the characteristic polynomial and roots of these characteristic polynomial are known as are referred to as the eigenvalues.

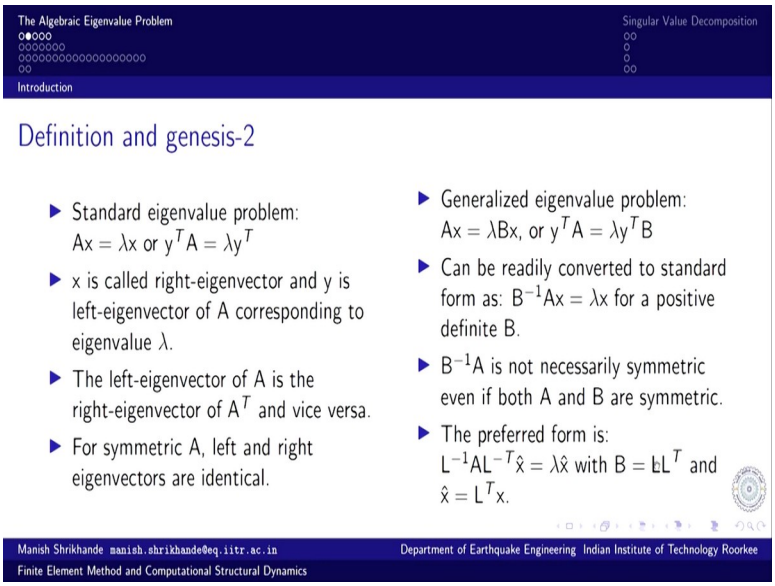
Eigenvalues are the roots of this characteristic polynomial. And once we find a particular eigenvalue, we can substitute it back in this equation and then find compute corresponding eigenvector to a degree of I mean constant of proportionality. So, we can shape we can find out one particular eigenvector for in I mean. If x is an eigen vector then, alpha times x is also an eigenvector of the matrix.

So, it can always only be determined to a constant of proportionality there is. So, eigenvector is not unique, because this matrix $A-\lambda I$ is of course, a singular matrix. So, there is no there cannot be a unique solution to this. So, there has to be a constraint additional constraint before vector x can be calculated and that constraint is by imposed by assuming some value for one of the elements or two couple of the elements depending on what is the multiplicity of this eigenvalue and then the rest of the elements of this vector can be calculated.

So, and that is why it is not possible to say that we solve for this system of equations, because this system of equations is still singular and constraints have to be imposed before we can compute the eigenvector. So, standard eigenvalue problem is referred to as $A x=\lambda x$.

So, there is only one^2 matrix a and x is the corresponding eigenvector which results after multiplication by A into a scaled version of it or we can have vector multiplying from the left side. So, there form that form becomes $y^T A=\lambda \times y^T$. So, this is also an eigenvector.

(Refer Slide Time: 07:14)



So, these are referred to as. Here, x operates on the right-hand side. So, this is referred to as the right eigenvector and here, y operates on the left side of the matrix A. So, this is referred to as the left eigenvector of matrix A. If matrix A is symmetric then; obviously, there is no difference between left eigenvector and right eigenvector they are identical.

So, x is called right eigenvector and y is called the left eigen vector and for a corresponding to a value eigenvalue specific eigenvalue λ . And these can be interchanged. I mean left eigenvector of A is the right eigenvector of A^T and vice versa.

So, that is simple mathematical operation; if we take the^T, then the result is arrived at. And for symmetric matrix A, left and right eigen vectors are identical. So, this is as far as standard eigenvalue problem what is referred to a standard eigenvalue problem is concerned.

Now, we can also have a generalised eigenvalue problem. For example, in this case instead of λx what we say is A x results into another vector which is $\lambda \times B \times x$. So, there are two matrices here. So, A x results into a scaled version of another rotation of matrix of vector x. So, $A x = \lambda B x$ or $y^T A = \lambda \times y^T B$.

So, that is generalised eigenvalue problem and; obviously, there is not I mean lot of similarities between the two generalised eigenvalue problem can be converted into standard form or standard eigenvalue problem can be considered as a special case of generalised eigenvalue problem where matrix B is identity matrix. And here, if we pre multiply by B^{-1}, then this generalised eigenvalue problem actually reduces to standard form.

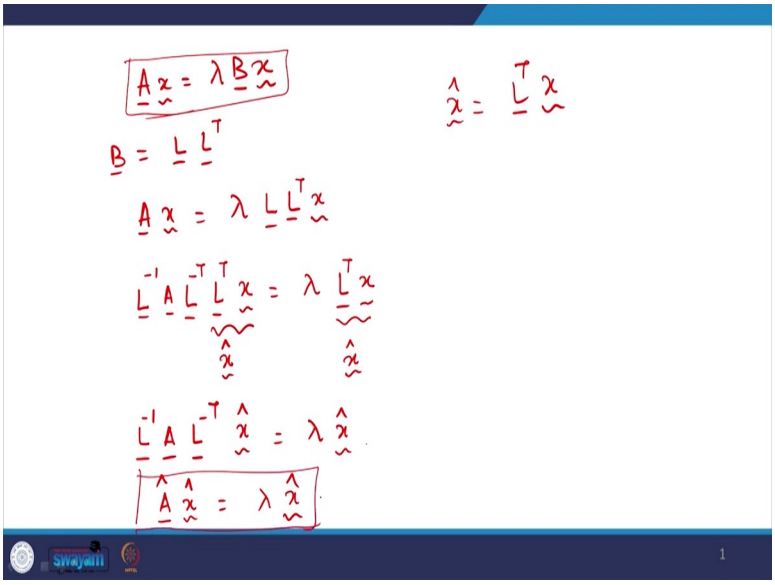
So, it can be readily converted to standard form as $B^{-1} A x = \lambda x$ for a positive definite matrix B. Now, the trouble with this is $B^{-1} A$ is not necessarily symmetric, even if both A and B are symmetric.

Now, the reason for symmetry is very very important. It is not just limited to the storage, because if it is if A and B are symmetric, then we can get away or we can do all the computation by storing only about half of the matrix. We do not really need to store other half. I mean lower diagonal lower triangular portion of the matrix right.

So, that is that results in considerable saving in the storage space and resulting into more compact storage, but if we pre multiply by B^{-1} , then this symmetry may be lost; even if A and B are both symmetric, $B^{-1} A$ is not guaranteed to be symmetric. So, that is another reason why pre multiplication by inverse is not really a good idea and not really required as well.

The preferred form is to ah. Since B is considered to be positive definite and it is symmetric. So, its decomposition is going to be LL^T a Cholesky decomposition. So, once we have this, then this B^{-1} can be easily converted into this particular form. So, this entire algebraic equation, this standard eigenvalue I mean generalised eigenvalue problem can be referred to as.

(Refer Slide Time: 11:30)



Let me just say $Ax = \lambda Bx$. So, if $B = LL^T$, then I can write this as. So, this I can write it as right. So, L^{-1} is easy to compute that you can be computed using simultaneous equation. I mean the triangular system of equations can be solved for a multiple right-hand size which = identity matrix and inverse of this triangular matrix L can be easily computed. So, that is what we arrive here, I mean L^{-1} .

So, and this is of course, identity matrix. So, we define assume this to be like this. So, we define another vector $\hat{x}$ which is given by $L^T x$ and that defines new problem $\hat{x} = \lambda \times \hat{x}$. And if I write this as $\hat{A}\hat{x}\lambda\hat{x}$. So, this becomes the standard eigenvalue problem which has the same eigenvalue. Eigenvalue is same as the original problem. Eigenvectors are related to the original eigenvectors through $L^T x = \hat{x}$.

So, once we compute $\hat{x}$, we can recover the eigenvector corresponding to the original problem from this simple relation that $\hat{x} = L^T x$ and solving this system of equations we can compute the vector x desired vector x. So, this is how the conversion transference

conversion of generalised eigenvalue problem into standard eigenvalue problem format is actually implemented in practice. It is never done as $B^{-1}A$.

So, that is what we are referring to here. So, preferred form is $L^{-1}AL^{i}\hat{x}=\lambda\hat{x}$. So, that becomes alternate form I mean how a generalised eigenvalue problem can be converted into a standard form. So, why is it so important, why is eigenvalue problem so important or so why prevalent so prevalent in a different fields.

One thing is we have already seen several instances all through this course we have seen the importance of knowing the basis defining the basis of an n dimensional space or vector space. So, knowing the proper basis to expand or to analyse any vector or to solve any problem is very very important.

Now, finding a suitable basis is not a very easy task except for constructing the basis for eulidean ah euclidean vector spaces, where the unit vectors are the basis. So, 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 and so on.

So, eulidean unit vectors they form they are very easy to construct and they can be generated, but while they are easy to generate, they are not really very convenient or to simplify the problem and many a times eigenvectors that we have, because these are derived from the operating matrix of the problem.

So, they often offer considerable simplicity to the problem being posed. So, it is possible to transform the problem into a much simpler form by using by using the transformation of variables where eigenvectors of the matrix they serve as the basis.

Now, why is it? So, convenient or why the eigenvectors form a convenient basis. It is, because of the very interesting property that eigenvectors form and orthogonal set of base vectors or they are mutually orthogonal to each other.

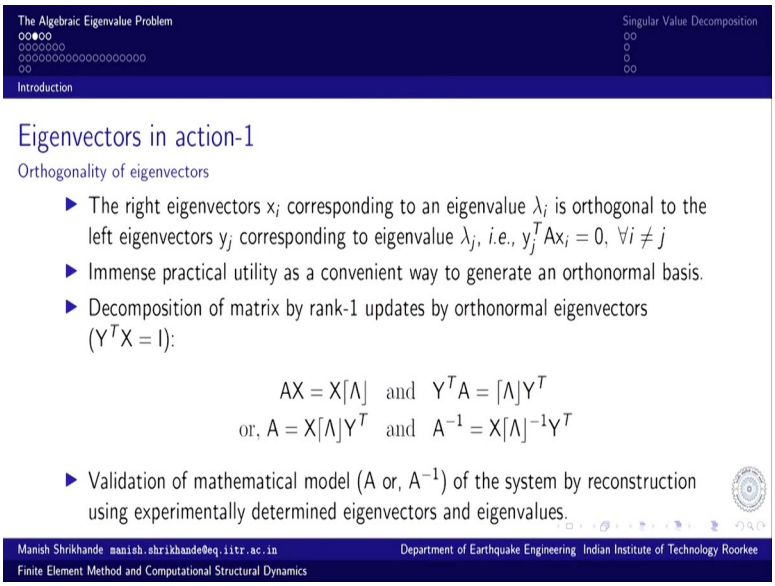
So, if I compute. I mean there are as I said there are different eigenvectors corresponding to different eigenvalues. So, one matrix can have several eigenvalues. Those are the roots of the characteristic polynomial. So, for each eigenvalue, there would be a corresponding eigenvector.

Now, what we are trying to say here is as long as the eigenvalues are distinct that is two eigenvalues that I am looking at they are different from each other they are distinct, then the eigenvectors are orthogonal to each other and that is a very very powerful property.

And orthogonality is always desirable that always leads to a very efficient transformation system, because for an orthogonal matrix we do not need to compute the inverse, its transpose its own inverse.

So, that we have ah discussed earlier when the when discussing the change of coordinate system or change of basis. So, the transformation between one basis to another basis if it is related through an orthogonal matrix or orthogonal set of column vectors then, it becomes the moving from one basis to another basis becomes very very efficient calculation and very very stable numerical operation. So, let us look at this orthogonality of eigenvectors in little detail.

(Refer Slide Time: 18:40)



So, right eigenvectors x_i corresponding to an eigenvalue λ_i is orthogonal to the left eigenvector at y_j corresponding to the eigenvalue λ_j ; these two are different λ_i is different from λ_j so; that means, $y_j^T A x_i = 0$ and of course, it also means that $y_j^T x_i$ that should also be 0. If this is 0, then the other way round is also possible for all i not equal to j.

Now, for symmetric matrix A, this is actually this reduces to the same set of vectors x_i is same as y_i left eigenvectors are same as right eigenvectors, but this is the general statement if A is un symmetric or then, this is the general rule that left eigenvectors are orthogonal to right eigenvectors. And this is has immense practical utility as a convenient basis to generate an orthonormal basis set of basis vectors.

And decomposition of matrix by rank one updates by orthonormal eigen vectors. So, because $y^T x = I$. So, if I collect all eigenvalues and eigenvectors in a common form. So, A x, x is the column of eigenvectors and λ this is the diagonal matrix of all eigenvalues. So, this is the in essence statement of all n number of eigen values and corresponding eigenvectors. So, this is for right eigenvectors and this is for the left eigenvector.

So, $y^T A = \lambda y^T$ or if we can post multiply y^T , then it becomes $x \lambda y^T = A$, because $x y^T = I$.

And then we have this A^{-1} can be easily converted to, because A^{-1} of matrix A can be easily converted by this matrix, because λ already involves, because x and y they are orthogonal with respect to each other. So, their inverses can be easily computed they are given by respective transpose matrices and inverse of a λ that is a diagonal matrix. So, that is also easily computed.

So, this computation of inverse by using eigenvalue expansion is a very very simple rank one update. So, this actually implies if we expand this in terms of summation of these vector products, then each term. There are n number of terms here, and each term actually contributes to one rank building up one rank of the matrix A complete matrix. So, A if A is full rank matrix n with rank n, then n number of terms each of these terms will actually add one rank to the system.

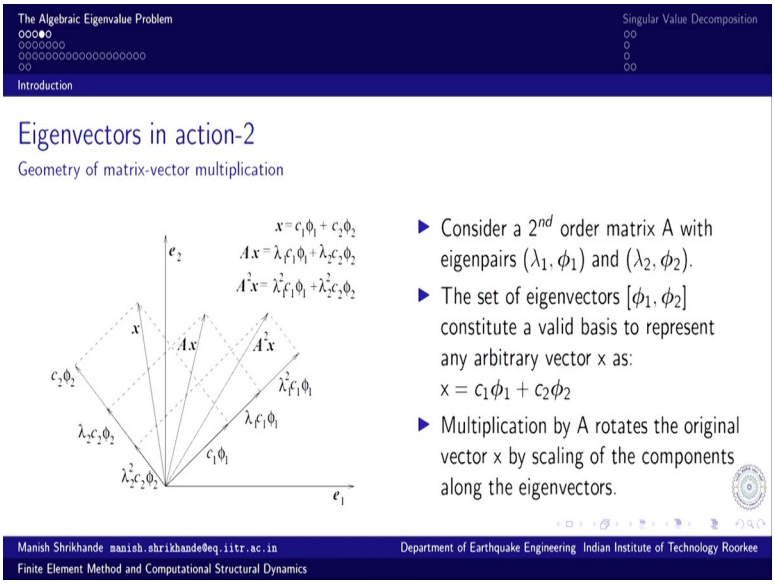
And this validation of mathematical model A or A^{-1} of the system by reconstruction by using experimentally determined eigenvectors and eigenvalue. So, this is a very very useful tool for reconstruction of the matrix. For example, system flexibility matrix can be a structural flexibility can be constructed by observing some eigenvalues and eigenvectors from the vibration data of this structure.

And using those flexibility matrices, we can draw useful inferences about the health of the structure. So, these are all applications that are possible. And hence, these tricks I mean the eigenvalues and eigenvectors the importance of eigen values and eigenvectors in constituting or in giving the orthonormal basis is very very crucial in simplifying the problems and even finding new insights into the problem.

So, now, how this actually works or what is what happens when we multiply eigen matrix with a vector. And in last lecture while discussing conjugate gradients, I had discussed that ah in Krylov subspace we have the matrix we have the sub space or the Krylov subspace elements of this Krylov sub space are given by r_0 , then $A r_0$ and $A^2 r_0$, $A^3 r_0$ and so on.

So, higher powers of a matrix A are involved and that is what gives us new directions new directions with each iteration. And there we said that for higher powers of A, it is possible that the vectors may begin to lose their independence and that we will see.

(Refer Slide Time: 24:40)



Now, why that will happen why that can happen. So, what happens is if x is the original vector and it is multiplied by A. So, it is rotated like this. And let us say %phi_1 and %phi_2, there are two eigenvectors right. So, of matrix A.

So, if there are two eigenvectors, then they constitute the basis, because they are orthonormal vectors. So, I can use them as a basis vector to expand the any any arbitrary

vector. So, vector x here, can be expressed as $C_1 \times \phi_1 + C_2 \times \phi_2$; ϕ_1 and ϕ_2 are two eigenvectors of matrix A corresponding to the eigenvalues λ_1 and λ_2 .

So, $A \times x$ will refer to $A \times C_1 \phi_1 + A \times C_2 \phi_2$ and $A \times \phi_1$ is of course, equal to $\lambda_1 \phi_1$ that is the by definition of eigenvalue problem and $A \times \phi_2$ is going to be $\lambda_2 \phi_2$. So, that is what we have. So, $A x = \lambda_1 C_1 \phi_1 + \lambda_2 C_2 \phi_2$. Now, this is what we look at it.

So, $\lambda A x$. So, this is vector $C_1 \phi_1$ this is what makes it. So, vector x can be decomposed as a combination of resultant of this $C_1 \phi_1$ and $C_2 \phi_2$. So, this is what vector x is representing and resultant of this $C_2 \phi_2$ and $C_1 \phi_1$. Now, after multiplying with matrix A it results in $\lambda_1 C_1 \phi_1$. So, this contribution get scaled by λ_1 and this component gets scaled by λ_2.

So, $\lambda_2 C_2 \phi_2$ resultant is A x and if I multiply with A again, then it becomes $A^2 x$ which is resulted by $\lambda^2 C_1 \phi_1$, because this will again be equal to $\lambda_1 C_1 A \phi_1$. So, that will be equal to $\lambda_1 \phi_1$. So, that will be $\lambda_1 \times \lambda_1 = \lambda^2$. So, $\lambda^2 C_1 \phi_1$. And similarly here, it becomes $\lambda_2^2 C_2 \phi_2$.

So, gradually as we keep on increasing multiplying with A repeatedly, what will happen is; this matrix or this vector A raised to the power n let us say some higher power of ah matrix A. This vector will be aligned will tend to align with ϕ_1 . The components of ϕ_2 will be subsumed. I am of course, assuming that λ_1 is greater than 1 and λ_2 is smaller than 1 which will happen some of the cases I mean relative speaking.

So, if λ_1 is much larger than λ_2 , then gradually this multiplication will align with the dominant eigenvalue and corresponding eigenvector. So, this is what it means.

So, what will happen is for higher powers of A, this will all these vectors will tend to be very close to the eigenvector corresponding to the dominant eigenvalue of matrix A and that is why the vectors in the Krylov subspace begin to lose their independence for higher powers of A and that is why it is required to make them orthogonal by using gram Schmidt orthogonalization after a while.

So, that this is basic idea of the matrix vector multiplication by successive multiplication of the matrix. We will we can nudge a vector to align with its dominant eigenvector. So, we consider a second order matrix or 2 by 2 matrix and the eigen pairs are $\lambda_1 \phi_1$ and $\lambda_2 \phi_2$.

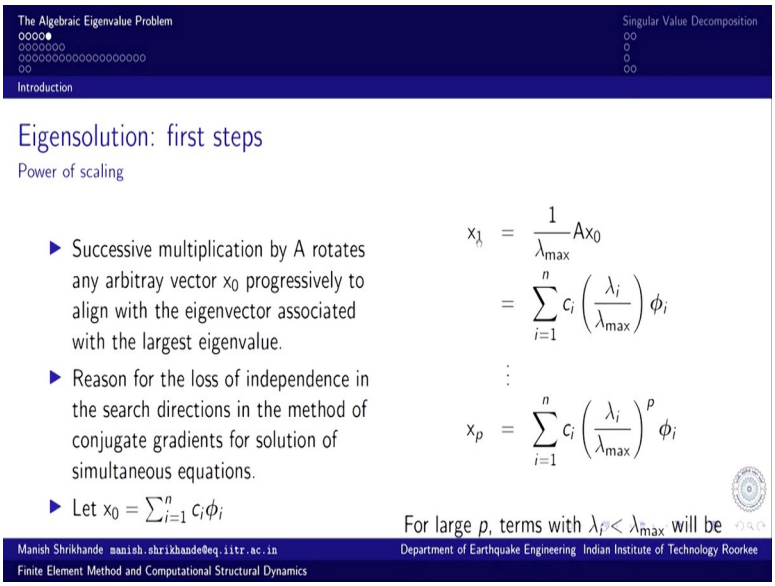
The set of eigen pairs eigenvectors $\phi_1 \phi_2$, they constitute a valid basis to represent any arbitrary vector x as a linear combination. So, that is why we can write $x = C_1 \phi_1 + C_2 \phi_2$. And multiplication by A rotates the original vector x by scaling of the components along the eigenvectors by I mean by scaling will be by the corresponding eigenvalue. So, ϕ_1 is scaled by ϕ_1 component is scaled by λ_1 , ϕ_2 component is scaled by λ_2 .

And this actually hints at the first algorithm for computing eigenvalues and eigenvectors. We just keep on multiplying whatever, we start with the vector any vector it could be any vector we multiply with A.

Again, this new vector that we get we multiply with A new vector we get we multiply with A. Eventually, it will align with the dominant eigenvector as we see here, in this graphical picture graphical. I mean gradually x we started x somewhere here and the first dominant eigenvector is aligned here.

So, this dominant eigenvector. So, this vector x can be nudged towards dominant eigenvector by successive multiplication of matrix A and that suggests us the first algorithm to solve for to find out to compute the eigen spectrum the complete computation complete representation of eigenvalues with eigen vectors that is referred to as the eigen spectrum. So, this is what we refer to as the power law or power of scaling.

(Refer Slide Time: 31:21)



So, successive multiplication by a rotates any arbitrary vector x_0 progressively to align with the eigenvector associated with the largest eigenvalue. Now, the only thing that we need to do is we need to scale them after each multiplication after each ah matrix multiplication so that we can detect the convergence; otherwise, it will keep on increasing by a factor of λ and it may not be easy to detect the convergence.

And as I said this is the loss this is the reason for loss of independence in search direction in the method of conjugate gradients for solution of simultaneous equations. So, let us assume x_0 as a a linear combination of basis vectors or eigenvectors as the basis.

So, $\sum_{i=1}^{n} C_i \phi_i$ n number of terms. So, when I start with I mean starting with x_0 as the trial vector; first vector x_1. If I multiply with a x_0 and then, divided by some let us say some number. Let us say it is I call it as a λ_{max}. So, that is the maximum eigenvalue, it can be some any number some.

So, if I scale it by this number. So, $\frac{\lambda_i}{\lambda_{max}}$. So, each term would be scaled by this $\frac{\lambda_i}{\lambda_{max}}$; λ_{max} is the largest eigenvalue. So, if λ_{max} is largest eigenvalue, then is; obviously, going to be less than unity. And there will be only one term which is going to be unity that will be $\frac{\lambda_{max}}{\lambda_{max}}$.

So, whichever eigenvector that λ_{max} corresponds to that term will have coefficient of 1. All others will have coefficients. I mean this ratio of eigenvalues less than one. And for this repeated multiplication of matrix A this ratio will; obviously, diminish with positive powers with every iteration and any ratio any number which is smaller than 1 less than 1. And if we keep on repeatedly multiplying it its value will; obviously, diminish it will approach 0.

So, for large p, terms with λ_i . So, eigenvalues ith eigenvalues λ eigenvalues with smaller than λ_{max} will be annihilated. They will be suppressed and eventually, we will be left with only one term that is corresponding to λ_{max} and that is why the power iterations this is called power iterations.

So, the by virtue of the power of multiplication repeated multiplication, any trial vector starting with any trial vector we can converge to the eigenvectors associated with dominant eigenvalue the largest eigenvalue largest eigenvalue in magnitude.

So, this is the first algorithm, but we computed the eigenvalues and eigenvectors I mean the dominant eigenvalue and eigenvector corresponding to dominant eigenvalue, but what next. This is just one of the n number of eigenvalues and eigenvectors and quite often, it is not the highest eigenvalue and eigenvector that is of interest rather it is the other way round.

So, lower end of this spectrum is of more interest for engineering analysis. So, how do we compute that. How do we go about with that computation of rest of the eigen spectrum we will discuss in our next lecture.

Thank you.