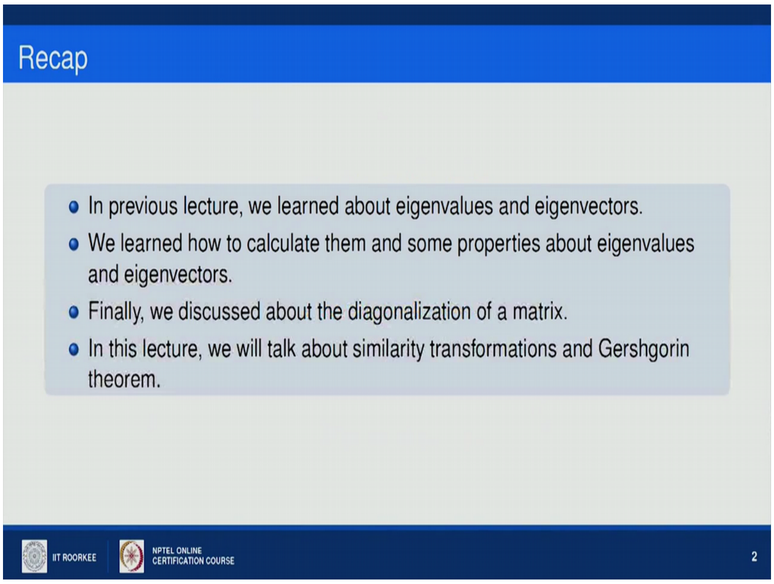
Numerical Methods
Professor Dr. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology Roorkee
Lecture 12
Similarity Transformations and Gershgorin Theorem

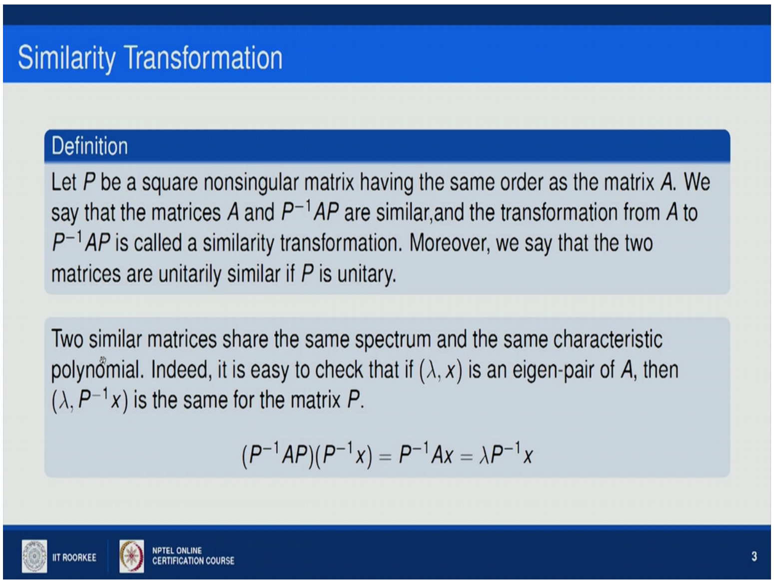
Hello everyone, so welcome to the second lecture of this module and this lecture contains some examples of similarity transformations and a beautiful results to find out the bounds on the eigenvalue of a matrix. So in the last lecture, I told you about eigenvalues and eigenvectors and in the last, I told you about diagonalization of a matrix. So in this lecture, I will generalized that concept in a more way that is, because diagonalization is a sort of transformation and it is a particular example of similarity transformations.

(Refer Slide Time: 1:10)



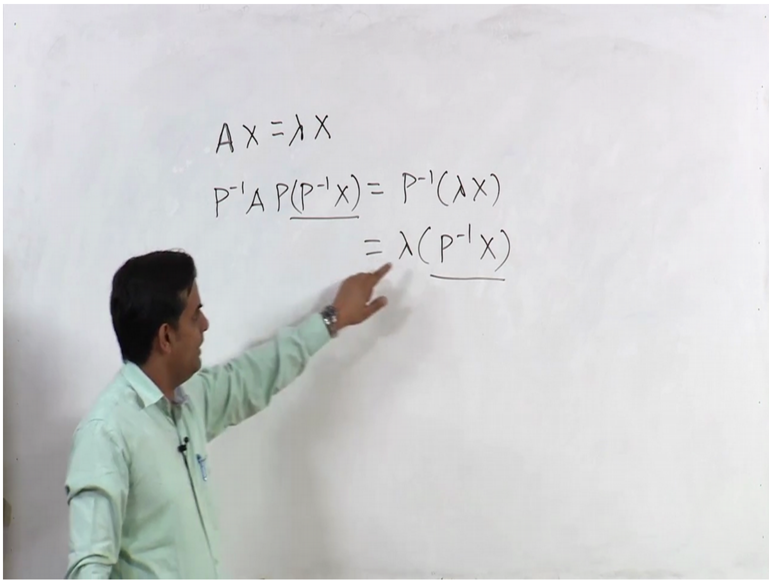
In this lecture, I will introduce few more similarity transformations and then Gershgorin Theorem for finding the bounds on eigenvalue.

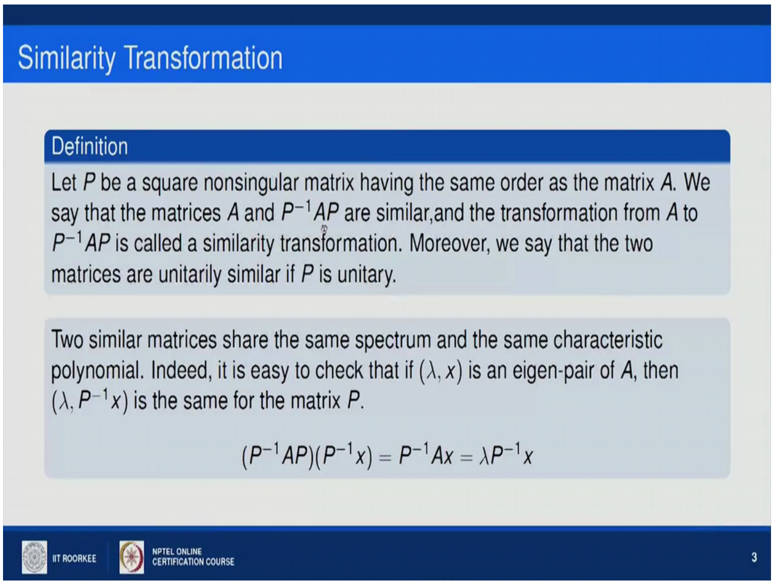
(Refer Slide Time: 1:20)



So first of all what is a similarity transformation? So let P be a square nonsingular matrix having the same order as the matrix A. So please note that P should be nonsingular matrix, it should be of the same order as A. we say that the matrices A and P inverse AP. So we are having pre-multiplication of p inverse with A and post-multiplication of P. So we say that matrix A and matrix P inverse AP are similar and the transformation from A to P inverse AP is called a similarity transformation, moreover we say that the 2 matrices are unitarily similar if P is a unitary matrix. So 2 similar matrices share the same spectrum means they contains the same eigenvalues.

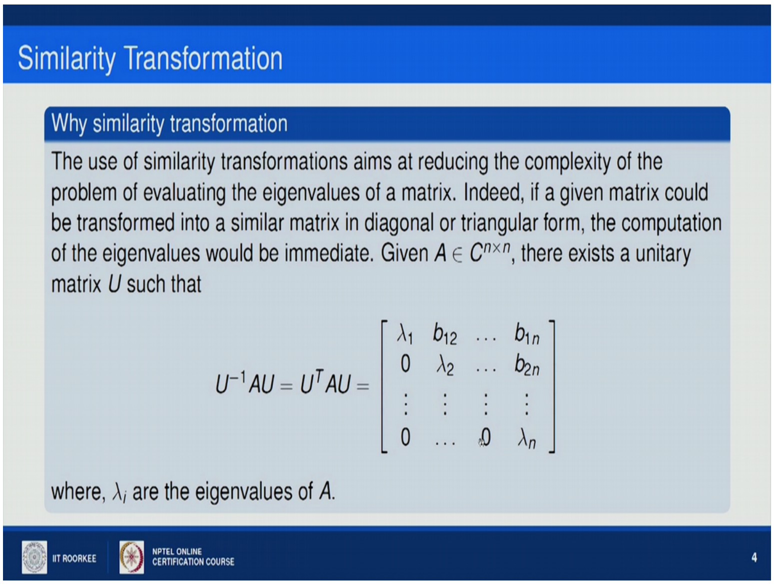
(Refer Slide Time: 2:27)





So if we can prove it here very easily. Let us say I am having a matrix A having eigenvalue lambda and eigenvector as X. So I am writing AX equals to lambda X, it means X is a eigenvector of A and lambda is an eigenvalue. Now if I write this, so what I am doing? I am pre-multiplying by P inverse in this particular equation and since here P into P inverse will be identity, so it will become P inverse into lambda X and if I take lambda, which is scalar out p inverse into X. So it tells me that, if X is the eigenvector of a corresponding to eigenvalue lambda, then the eigenvector corresponding to the same eigenvalue lambda of the similar matrix P inverse AP will be P inverse X. hence, A as well as P inverse AP are having the same values.

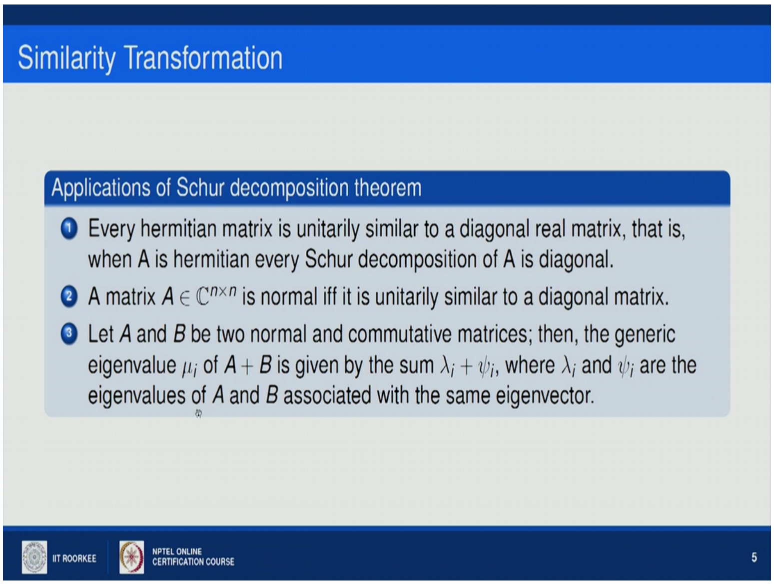
(Refer Slide Time: 3:49)



Why we need similarity transformation? Especially when we are talking about eigenvalues and eigenvectors so the use of similarity transformations aim at reducing the complexity of the problem of evaluating the eigenvalues of a matrix, for example, if a 10 by 10 matrix is given to you. So a square matrix of order 10 and will say that, okay find out the eigenvalues A. So since the order is 10 by 10, the characteristic polynomial of that matrix will be of degree 10 and it is quite difficult to calculate eigenvalues of such a matrix manually, until and unless the matrix is either a diagonal matrix or a triangular matrix, upper triangular or lower triangular.

So what we need to do? Here, our aim is to find out some similarity transformations such that we can for a given general matrix we can apply the similar transformation and we can convert it in either as a diagonal matrix or triangular matrix, such that it will be easy to find out the eigenvalue of such a matrix, for example, if A is any general matrix here. If I apply U is a unitary matrix. So if I apply U inverse AU. So as you know that a matrix shown is said to be unitary, if U inverse equals to U transpose. So I can replace this U inverse by U transpose, so U transpose AU will become a triangular matrix. So this particular results for a given matrix A, there exist a unitary matrix U such that, tis result holds is called schur lemma.

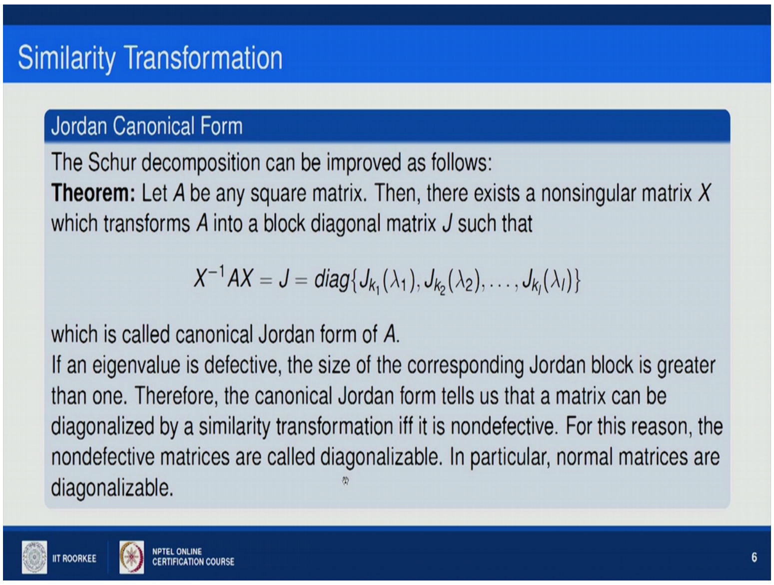
(Refer Slide Time: 5:47)



Now we are having several applications of schur decomposition lemma that is, every hermitian matrix is unitarily similar to a diagonal real matrix, means if you are having a hermitian matrix, it will be similar to a real diagonal matrix and hence, we can say, hermitian matrices are having the real eigenvalue, because at the diagonal we will be having the real entries. When A is hermitian schur decomposition of A is diagonal. A matrix A coming from a n by n matrix having the complex entries is normal if an only it is unitarily similar to a diagonal matrix.

Moreover we can say from this particular lemma that let A and B be two normal and commutative matrices then, the generic eigenvalue Mu I of A plus B is given by the some lambda I plus psi I where lambda I and psi I are the eigenvalues of A and B associated with the same eigenvector and hence, if A and B you are having the eigenvalues of A and B separately, you can find out the eigenvalues of A and B.

(Refer Slide Time: 7:18)



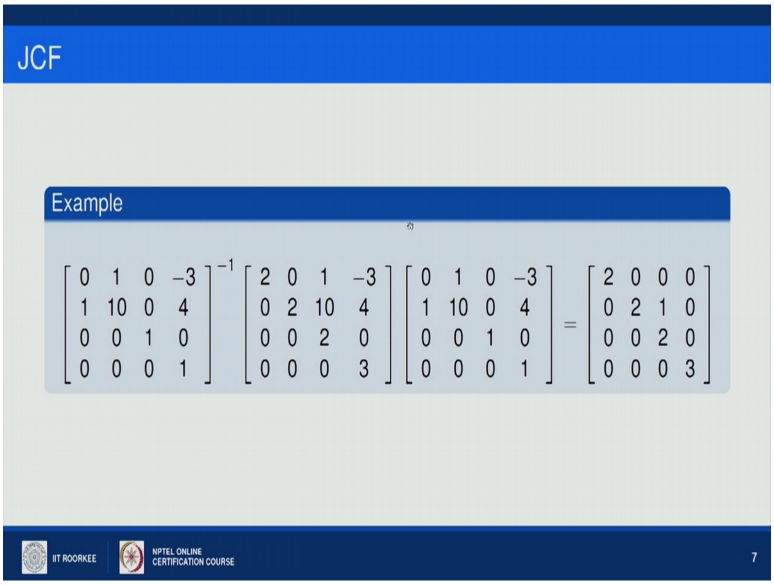
Now we can have other variants of similarity transformation, one of them is called Jordan canonical form. So what is that? Let A be any square matrix, then there exist a nonsingular matrix X, which transform A into a block diagonal matrix J such that, X inverse AX equals to J and where J will be J is a block diagonal matrix and blocks are called Jordan blocks of A corresponding to matrix A and J is called the Jordan canonical form of A.

So if a matrix A is having distinct eigenvalues. So it means as I told you in the previous lecture, we will be having linearly independent eigenvectors corresponding to each distinct eigenvalue and hence, matrix will be diagonalizable. So here Jordan canonical form will become a diagonal matrix, because diagonal matrix is also called we can say it is a block diagonal matrix, moreover if this is not the case a matrix is not diagonalizable then we can write it as similar to a block diagonal matrix.

So if an eigenvalue is defective, defective means the algebraic multiplicity is not equal to geometric multiplicity. The size of the corresponding Jordan block is greater than 1. So obviously then geometric multiplicity will be less than that particular repetition of eigenvalue that is the algebraic multiplicity and hence, from this geometric multiplicity what we need to do we need to decide the block size number of block like if an eigenvalue is repeated 5 times for a matrix and it is having only 3 linearly independent eigenvectors. So what we need to say that this particular matrix is similar to a block diagonal matrix, which is having 3 Jordan blocks and total size of 3 Jordan blocks would be 5. So if I decompose 5 into 3 terms, it may be 2 plus 2 plus 1. So one Jordan block of size 2, another Jordan block of size 2 and the 3^{rd} Jordan block of size 1 or it may be 3 plus 1 plus 1 like that.

So if an eigenvalue is defective, the size of the corresponding Jordan block is greater than 1. Therefore, the Jordan form tells us that a matrix can be diagonalized by a similarity transformation, if and only if it is non-defective, for this region the non-defective matrices are called diagonalizable in particular normal matrices are diagonalizable, okay.

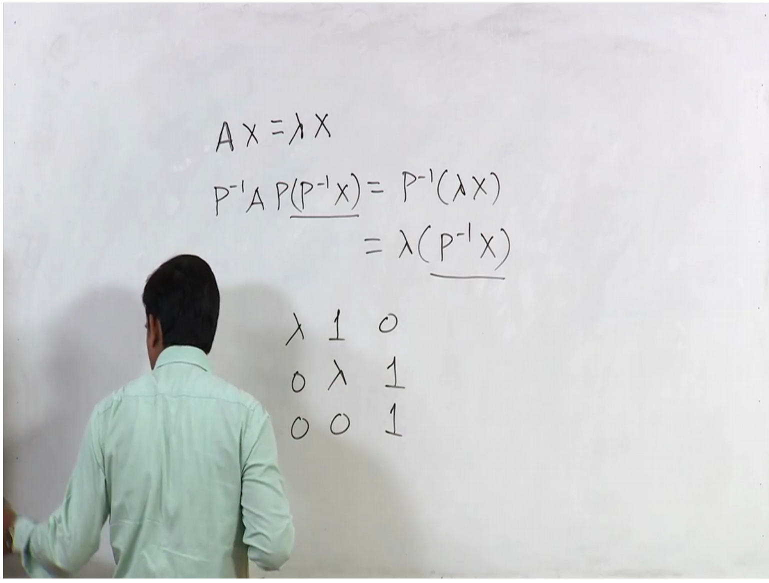
(Refer Slide Time: 10:11)



So let us take a beautiful example of Jordan canonical form and here I am having this matrix as my matrix A. so it is a 4 by 4the matrix, the first row is (2, 0,1, -3) (0, 2, 10, 4) is the second row, and the third row (0, 0, 2, 0) and in the 4^{th} row (0, 0, 0, 3) as you can notice the matrix is a an upper triangular matrix and hence, the eigenvalue of this matrix are 2, 2, 2 and 3 however; we do not know what will be the Jordan blocks corresponding to this matrix. If I find out the modern matrix P and then I calculate P inverse AP that is the similar matrix to the Jordan matrix. So A is this upper triangular matrix. Now P inverse A into P coming out as my matrix J, which is the Jordan canonical form of this matrix A so in this Jordan canonical form, you can see, this is a 1 by 1 block corresponding to eigenvalue 2.

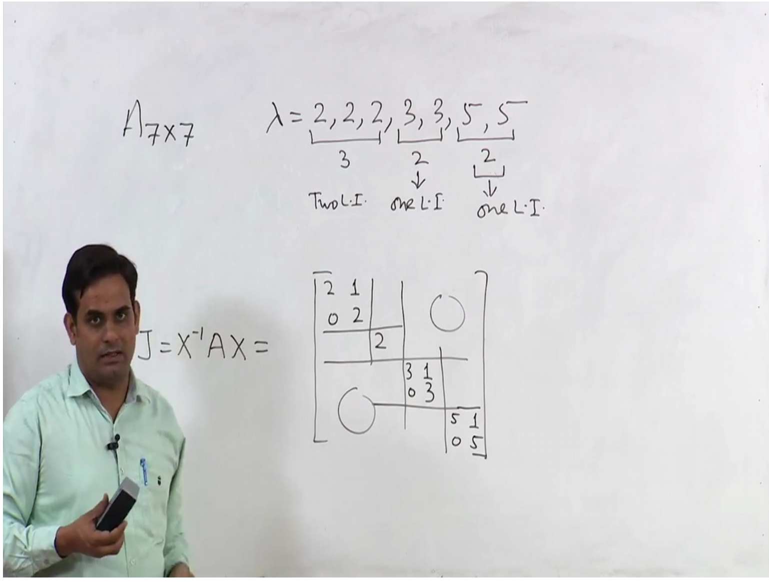
Now this is a 2 by 2 block corresponding to eigenvalue 2 and this is 1 by 1 block corresponding to eigenvalue 3. Since, the algebraic multiplicity of 3 is 1. So hence we are shure that only there will be 1 block or one Jordan block for 3, because there will be only one linearly independent eigenvectors, if you solve find out the eigenvectors of this matrix corresponding lambda equals to 2. What you will find? The number of linearly independent eigenvectors for this will come out as 2. So it means the Jordan blocks will be 2 total 2 Jordan blocks, 1 Jordan block of size 1, another Jordan block of size 2, because we have to factorize 3 in 2 terms as the sum of 2 number. So obviously it will be 1 and 2. So this is the Jordan canonical form of this matrix A and hence these two matrices are similar matrices and this transformation is a similar transformation.

(Refer Slide Time: 12:34)



So if I want to write a Jordan block of size 3, so it will be something like that (lambda 1 0) (0 lambda 1) (0 0 1).

(Refer Slide Time: 12:57)

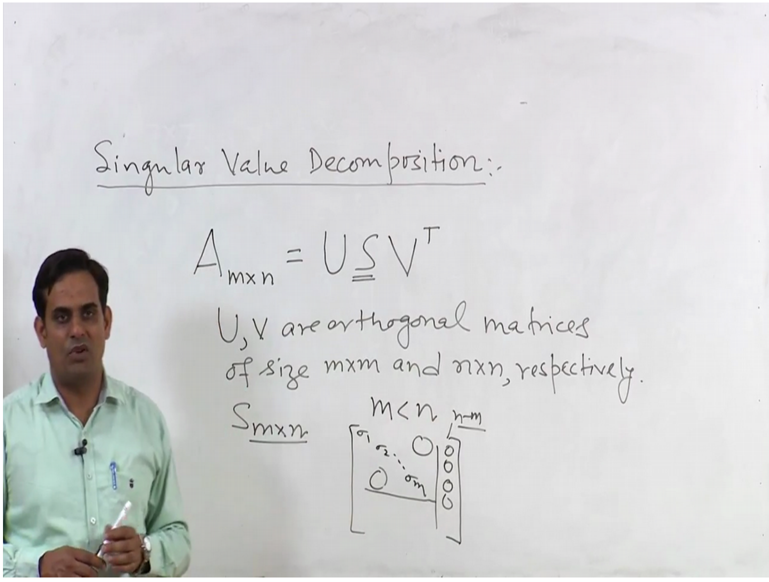


So how we can write the Jordan canonical form of a matrix, suppose I am having a 7 by 7 matric, which is having so A is a 7 by 7 matrix which is having eigenvalue as 2,2,2,3,3,5,5. So here algebraic multiplicity of 3 is 2 and algebraic multiplicity of 5 is 2. Let we are having only 2 LI eigenvector corresponding to lambda equals to 2 that is the geometric multiplicity of 2 is 2. Let us say I am having only 1 linearly independent eigenvector corresponding to 3 and the 1 linearly independent eigenvector corresponding to lambda equals to 5. So hence all the eigenvalues are defective here.

Now what will be the Jordan canonical form of this particular matrix? So here J will be I need to find out an invertible matrix X such that j equals to X inverse AX. So here you can see algebraic multiplicity is 2. So total I will be having 3 by 3 size reserve for this eigenvalue out of which I am having only 2 linearly independent eigenvector. So it means I will be having only 2 blocks. So how I can decompose 3 into 2 it will be 2 plus 1 or 1 plus 2. So 2 plus 1 means one Jordan block of size 2 and another Jordan block of size 1.

Similarly, here I am having only one linearly independent eigenvector corresponding to lambda equals to 3. So here I will be having 1 block of size 2, geometric multiplicity is giving me the number of blocks corresponding to that particular eigenvalue. So it will be 3, 1, 0, 3 and finally I am having lambda equals to 5 again only one block. So 5, 1, 0, 5 and rest of the entries will be zero. So this will be Jordan canonical form corresponding to this matrix A.

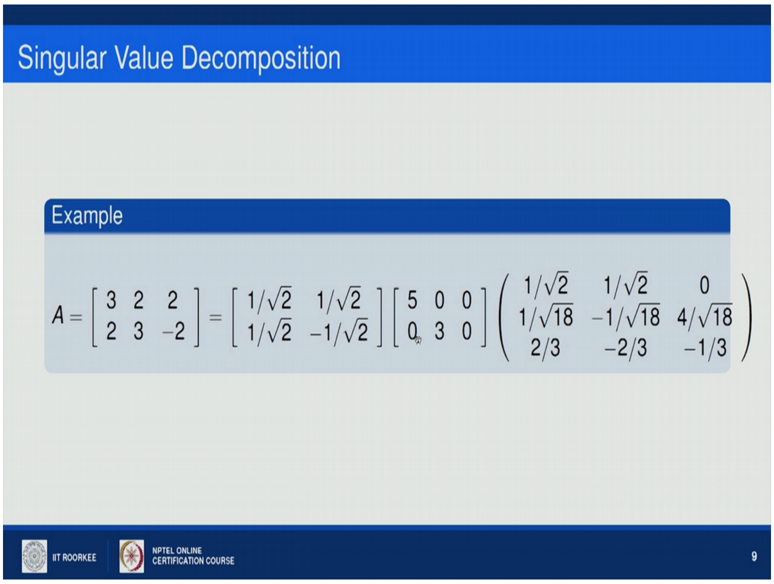
(Refer Slide Time: 15:39)

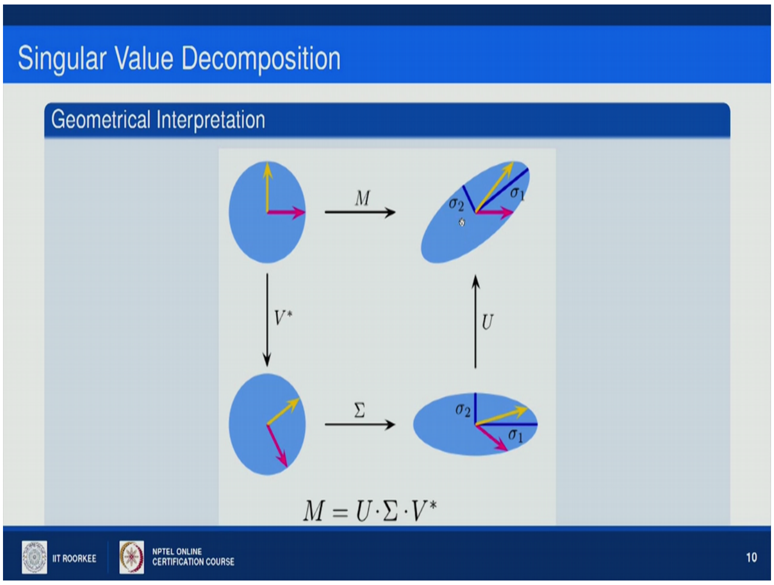


Next example of similarity transformation is singular value decomposition and it is different from previous examples, because this decomposition holds or this similarity transformation hold for rectangular matrices also not like Jordan canonical form or schur decomposition theorem, which is applicable only to the square matrices. So here it is say that any matrix A of size m by n can be written as the product of 3 matrices U S and V transpose, where U and V are orthogonal matrices of size m by n and n by n respectively and S is a matrix of size m by n in which all the up diagonal entries are zero. Here since it is of size m by n, so here I am saying up diagonal means a bit odd because for a rectangular matrix how you will decide the diagonal. Here my if m is less than n then what will happen, the matrix will be like this I will be having a m by n matrix a square matrix, which is diagonal let us say like this sigma1, sigma2, sigma m are the diagonal entries rest of the thing will be 0. So this is m by m matrix and then what I will be having? I am having n minus m number of columns to make it m by n matrix.

Similarly if m is greater than n then I will be having n by n square matrix which is diagonal matrix and then a minus n number of rows will be appended in the bottom of this matrix. Hence and the sigma1, sigma2, sigma m are called singular values of a matrix A and they are the square root of the eigenvalues of A transpose or A transpose A. The eigen vectors of A into A transpose will be the columns of U and the eigenvectors of a transpose A will be the columns of V and hence in this way we can achieve this decomposition.

(Refer Slide Time: 19:13)

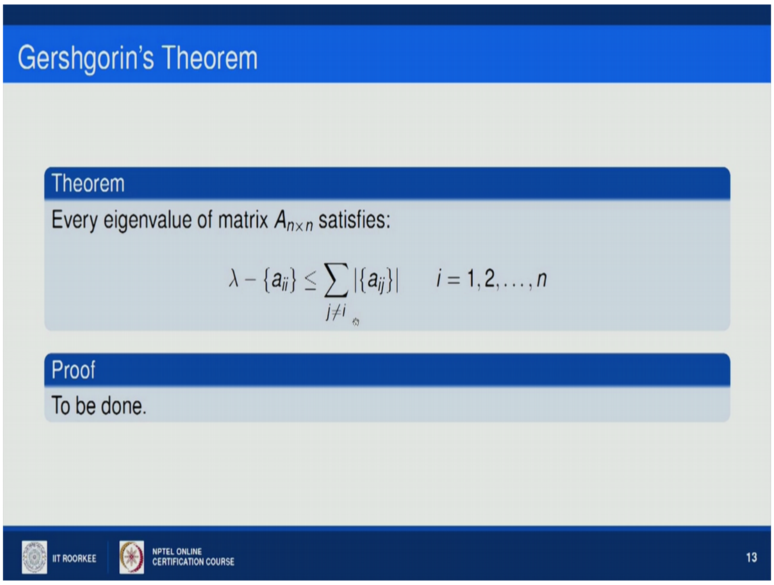




So this is an example of singular value decomposition. Here I am having a 3 by 2 2 by 3 matrix, which is first row is 3, 2, 2,second row is 2,3,-2 and this is the singular value decomposition of matrix A. So this is matrix U S and V transpose geometrically I can say like this. so I am having a circle which is transform to this ellipse by a transformation M which is a matrix in terms of singular value decomposition, it will be like that first I am applying V star on it that is V transpose. So what will happen, it will rotate orientation will change, because it is an orthogonal matrix. Hence, it is a rotation matrix then what will happen then I will apply this, it is a diagonal matrix, so what will happen, it will change the scale and it will deform this particular shape. So circle will become ellipse and finally U will noted the ellipse.

So geometrically it is a stage of 3, 2 rotations and one deformation. So here we have seen some similar transformations and from them what we can say we can apply those transformations to the given matrix and we can say that this matrix is similar to sum of the diagonal matrix or triangular matrix and hence, it is a similarity transformation. So both original matrix and diagonal matrix will be having the same spectrum and hence, the same eigenvalue and it is easy to find out the eigenvalues of a diagonal matrix. We will see some methods on the basis of this similarity transformations in this module in next lectures; however before that let me introduce a very beautiful results proposed around 1930 by a Russian mathematician jurs Gershgorin and it is called Gershgorin disc theorem or gershgorin circle theorem.

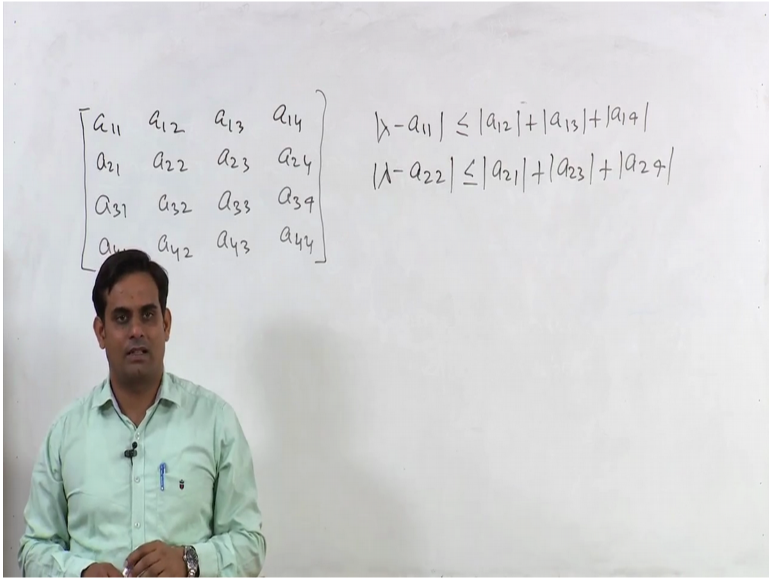
(Refer Slide Time: 23:08)



So basically this theorem tells us about gives a bound on the eigenvalues, suppose I am giving you a 4 by 4 matrix and I will say okay tell me the eigenvalue of this matrix, just by looking on the matrix I cannot say about the eigenvalues, one of the thing is if I can at the diagonal elements and hence I can find out the trace and I will say okay trace is 5. So some of the eigenvalues will be, but if trace is 5 still we cannot say anything about eigenvalues, if it is a 4 by 4 matrix, it may happens 2 of the matrix are 2 of the eigenvalues are quite high and 2 are having let us say 1 is 100, another one is 105 and rest 2 are minus 100, -100. So that sum will be 5 and trace equals to trace is 5 or it may be eigenvalues are (0,0,1,4) or it may be (0,0,0,5). So I cannot get any idea or any guess about the eigenvalues just by having the trace. So how to get some idea of the eigenvalues just by looking on the matrix this particular theorem tells us about it.

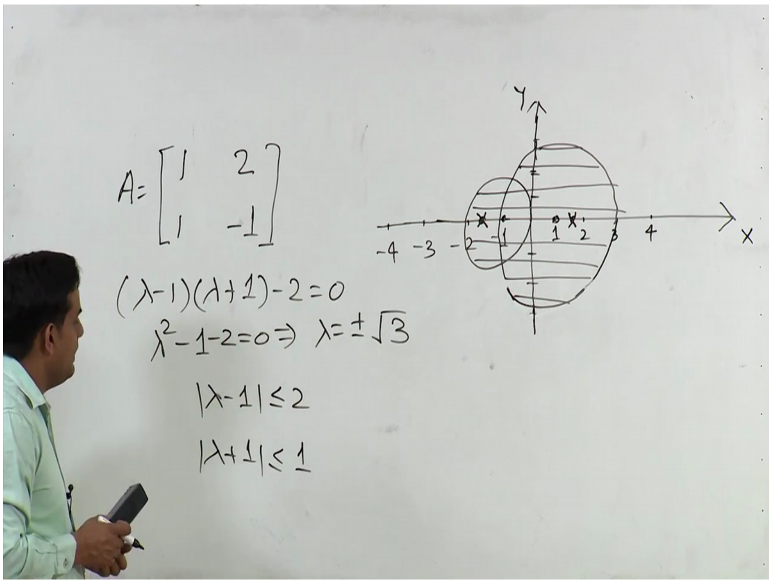
So this theorem tells that every eigenvalue of matrix A which is of square matrix of order n satisfies this particular inequality that is if lambda is an eigenvalue lambda minus aii, so aii is the diagonal element in ith row will be less than equals to sum of a flu sum of all the elements in that ith row except the diagonal element.

(Refer Slide Time: 23:44)



So, how to do it? Basically, so if I am having a 4 by 4 matrix. So Gershgorin theorem tells us that the eigenvalues will be like from the first row I am saying that lambda minus a11 will be a14. So absolute value of lambda minus a11 less than equals to absolute sum of rest of the entries from the first row and since, eigenvalues are coming from the field of complex numbers. So I need to find out this particular inequality is giving me a disc in the complex plane, which is having center at a11 and there this sum. Similarly, second row will tells second row is giving me another disc and similarly I will get another form the third row and the last one from fourth row and hence, gershgorin theorem tells us that all the eigenvalues will lie in the union of upon all disc or all Gershgorin disc corresponding to that particular matrix, okay.

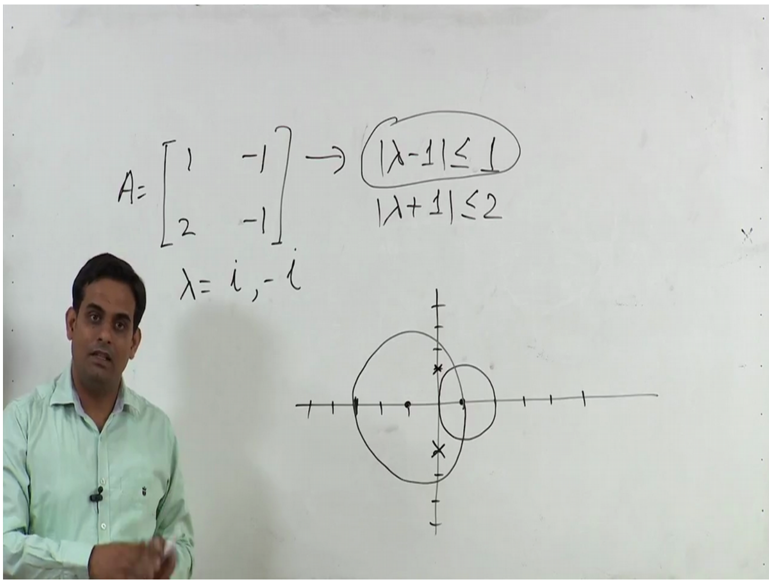
(Refer Slide Time: 26:04)



So let us take some example of gershgorin disc, so whether the eigenvalues are coming inside the gershgorin disc that is the union of all disc or not. So let us take a 2 by 2 example, a matrix A which si given a (1, 2, 1,-1). Now if I find out the eigenvalue of this matrix then the characteristic polynomial will be lambda minus 1 into lambda plus 1 minus 2 equals to zero. So this will be lambda square minus 1 minus 2 equals to zero. So, lambda equals to plus minus root 3 now if I plot the disk of this matrix according to gershgorin theorem. So let us say, this is my X and Y that is the complex plane. So the imaginary axis and real axis let us say 1,2,3,4. Similarly here, -1,-2,-3,-4 and 1real, 2, 3 that is the i, 2i,3i i,2i, 3i.

Now from the first line what I am getting that is, the first disc will be lambda minus 1 less than equals to 2. It means the center is at 1 and radius is 2. So center is 1, radius is 2. So it will be this disc okay. So ya so this is the center of the disc if I plot the another disc will be lambda plus 1less than equals to 1, that is the lambda plus 1 less than equals to 1. So the second, the center of second disc at minus 1 and radius is 1. So it means it will start from here and this will be like this. Hence, this region will be the union of these two disc. Now the eigenvalue is root 3, root 3 will comes somewhere here and minus root 3 will come somewhere here. So these are the 2 eigenvalues and here we can say the eigenvalues lie in the union of gershgorin disc. Here, eigenvalues are real.

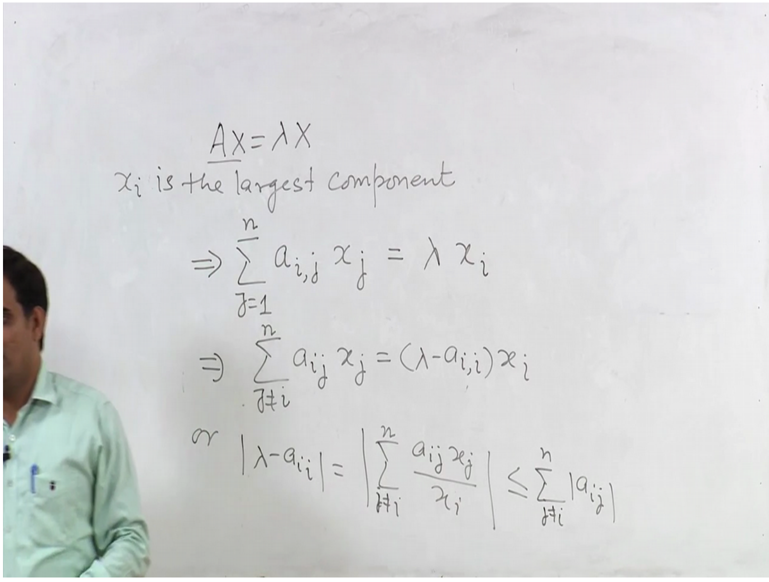
(Refer Slide Time: 29:22)



Let us take another example, where eigenvalues are where eigenvalues are imaginary eigenvalues. So again for sake of simplicity, let me take a simple matrix that is a 2 by 2 matrix, (1, -1, 2, -1). Now eigenvalue of this matrix are if us calculate with characteristic polynomial i and minus i. Now this row gives me the gershgorin disc as lambda minus 1 less than equals to 1 and the another one is giving lambda plus 1 less than equals to 2. So disc are 1,2,3,4,5 1,2,3,4,5. So real axis and imaginary axis so if I plot first disc so center at 1 and radius is 1 so this is the center and disc will be like this. Now if I see the second disc, center is minus 1 that is the this point and radius is 2. So it will pass through here and then here. So this will be the another Gershgorin disc.

Now check the eigenvalue, eigenvalues are i and –i and here, you can see both the eigenvalues are in union of gershgorin disc. Basically here both the eigenvalues are coming in the second disc. They are not coming in this disc. So the claim that the eigenvalues will lie in the union of gershgorin disc is correct. If someone say that each eigenvalue will lie in its respective gershgorin disc that is not true and how to prove this.

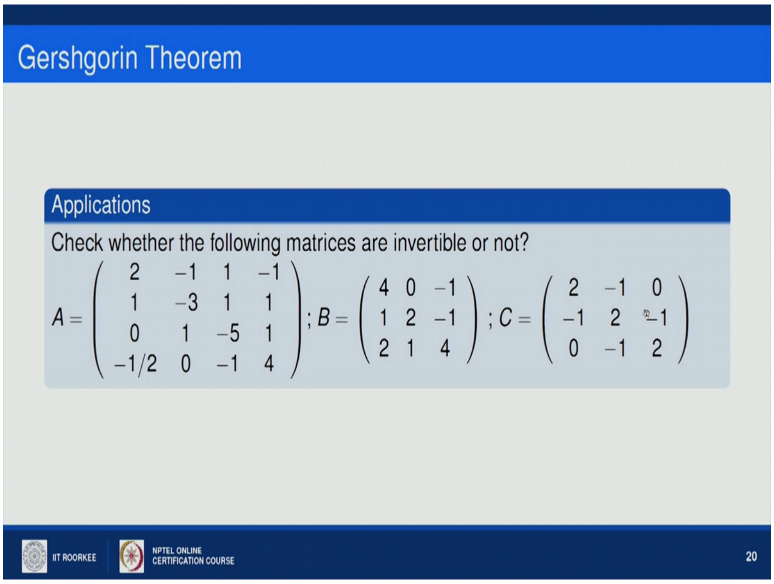
(Refer Slide Time: 31:30)



So proof is quite easy, let us say I am having a square matrix A, which is having eigenvector x corresponding to eigenvalue lambda. Let us say in this eigenvector xi is the largest largest component. So if A is n by n matrix, the vector x will be having n components and out of n ith component is the largest. Now what this particular relation is telling to me, suppose I am having see the ith row of tis matrix, this particular relation. So from this ith row will multiply each component of x and this will be calls lambda times xj xi ith component. So it like this j equals to 1 to n aij xj.

So if I see the ith row of this product and this equals to, because this will be i equations, so I am taking n number of equations out of n I am taking the ith equation. This is equal to lambda times xi or I can write this, j not equals to i, so one element it is j is 1 to n. So I am taking 1, when j is i into other side, because there it will be xi. So it will become aij xi j not equals to i from 1 to n and here, it will be lambda minus aii xi or let us take this into this side take a mod on it will be summation j not equals to i, aij xj upon xi, okay and it is up to n. so what I have in the beginning, I told you that ith entry is the largest one. So xj upon xi will be always less than 1. So if I replace it by 1 what will happen, it become less than and this is our gershgorin disc and this is the poof, very simple proof, just coming from the definition of eigenvectors and eigenvalue.

(Refer Slide Time: 35:17)



Let us see some of the applications of this theorem, just look at this 4 by 4 matrix, can you tell me just looking at this matrix whether it is invertible or not, you can tell this to me if you know the determinant of tis matrix or you know the eigenvalues of this matrix; however just look here, if I apply the gershgorin theorem on this particular problem what I am getting, the first is disc is lambda minus 2 less than equals to 3. The second is lambda plus 3 less than equals to 3, the third one is lambda plus 5 less than equals to 2 and the last one lambda minus 4 less than equals to 3 by 3 by 2 ya.

So from here we cannot say anything, but if you apply the gershgorin circle theorem on the columns of A, because A eigenvalues of A and A transpose will be equal and hence, the theorem holds for columns also. So from columns it is quite clear that in each column I am having this inequality as strict inequality, it means none of the disc will contain the origin and hence, union of all the disc will not contain the origin and hence, 0 cannot be an eigenvalue of this matrix and hence, it is a invertible matrix. So same kind of analysis you can make for these two examples also like, again it is a from the columns we can see, it is invertible.

Here, we cannot see from the columns as well as from the row due to this second thing, second row or second column, because second column or the disc corresponding to second row will touch the origin and since, it will touch the origin, it may happen that 0 may be an eigenvalue of it. So what to do in this case? Here, we are not having any result from the gershgorin theorem, we cannot talk about the inevitability, but from these 2 examples, I have told you how to apply gershgorin theorem just to check whether the matrix is invertible or not. Thank you for this lecture, so in this lecture we learn about Gershgorin theorem as well as we have seen some examples of similarity transformation, thank you.