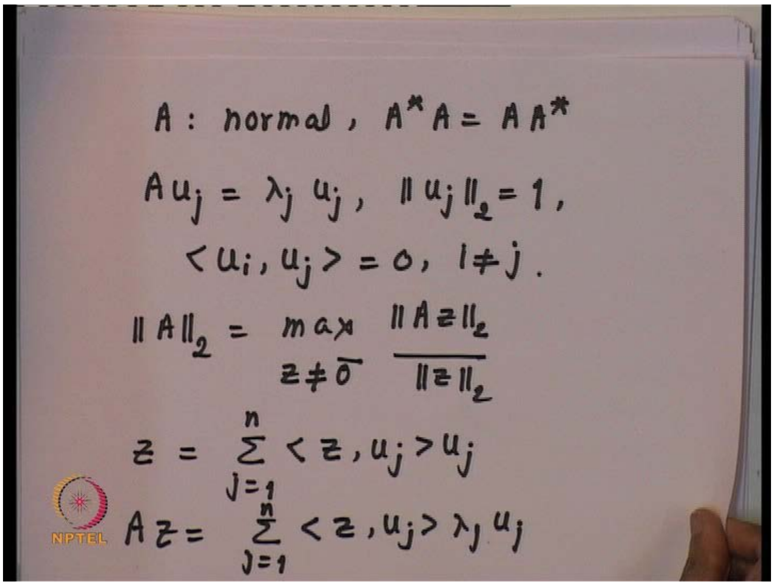
Elementary Numerical Analysis
Prof. Rekha P. Kulkarni
Department of Mathematics
Indian Institute of Technology, Bombay

Lecture No. # 37
Power Method

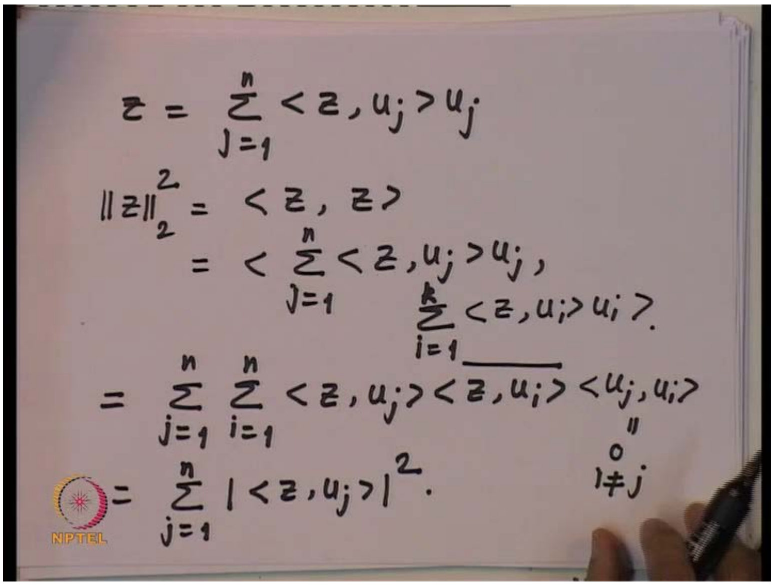
.In this lecture, we are first going to show that if A is a normal matrix; that means, A star A is equal to A A star, then 2 norm of the matrix is modulus of the biggest Eigen value. Then we will look at some of the localization results; that means, we will look at some region in the complex plane which are going to contain Eigen values of given matrix, after that we are going to look at power method, which is a method for finding approximation to the dominant Eigen value and then we will look at some of the extensions of this power method. So, let me first show that 2 norm of A is maximum of the biggest or it is the modulus of the biggest Eigen value provided A is normal.

(Refer Slide Time: 01:26)



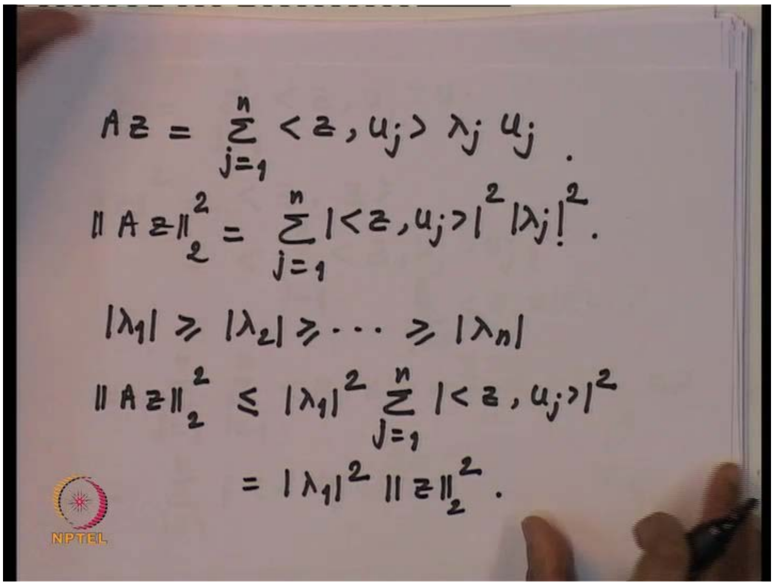
So, we have got A to be normal matrix A star A is equal to A A star, by spectral theorem we have got A u j is equal to lambda j u j, where norm u j, its 2 norm is equal to 1 and inner product of u i with u j is 0, if i not equal to j. Our definition of 2 norm is maximum of norm A z its 2 norm divided by norm z 2 provided, z is not equal to 0 vector. A z in c n can be written as summation j goes from 1 to n, inner product of z with u j u j. So, this is using orthonormality of u j s. So, let us calculate norm z, norm A z, A z is going to be summation j goes from 1 to n, inner product of z with u j and then A u j. So, that will be lambda j u j.

(Refer Slide Time: 03:05)



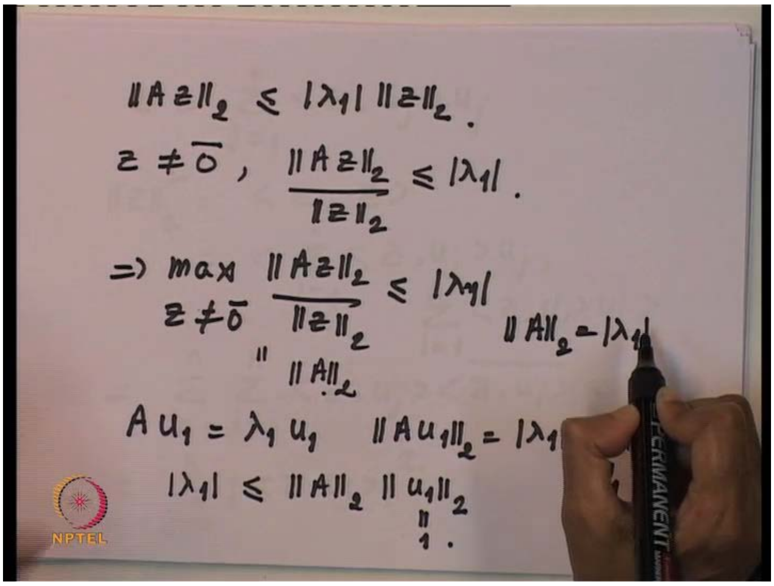
So, we have got z is equal to summation j goes from 1 to n inner product of z with u j u j. So, norm z its 2 norm square, it is going to be inner product of z with itself. So, this will be summation j goes from 1 to n, inner product of z with u j u j and summation, say, i goes from 1 to n, inner product of z with u i u i. Now, use the linearity in the first variable conjugate linearity in the second variable to obtain summation j goes from 1 to n, summation i goes from 1 to n, inner product of z with u j this is coming out as it is, complex conjugate of z u I, because it is coming out from the second variable, u j comma u i, now this is 0 if i not equal to j. So, you will get this to be equal to summation j goes from 1 to n modulus of z comma u j square. So, this is norm z its 2 norm square.

(Refer Slide Time: 04:48)



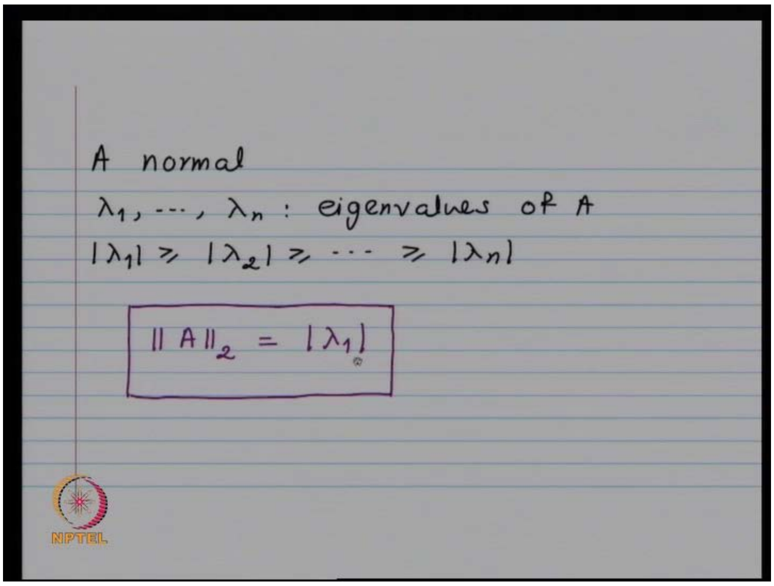
In a similar manner, we are going to have A z to be equal to summation j goes from 1 to n, inner product of z with u j lambda j u j. So, norm of A z, its 2 norms square will be summation j goes from 1 to n modulus of z comma u j square mod lambda j square. Suppose the Eigen values are ordered in this manner, mod lambda 1 are bigger than or equal to mod lambda 2, bigger than or equal to mod lambda n, then norm A z its 2 norms square will be less than or equal to each mod lambda j will be less than or equal to mod lambda 1, take out of the summation sign. So, we will have mod lambda 1 square, summation j goes from 1 to n modulus of z comma u j square. Now this is nothing but 2 norm of z whole square. So, we have got mod lambda 1 square and then norm z 2 square and thus we will get norm A z 2 norm to be less than or equal to mod lambda 1 times norm z 2 norm.

(Refer Slide Time: 06:33)



And then for z not equal to 0 vector, norm a z divided by norm z will be less than or equal to mod lambda 1, which will imply that maximum of norm a z by norm z, z not equal to 0 vector is less than or equal to mod lambda 1. This is our norm a 2. We have got a u 1 is equal to lambda 1 u 1. So, norm a u 1 is going to be mod lambda 1 times norm u 1 2 norm. So, this is equal to 1 and thus mod lambda 1 will be less than or equal to norm a 2 norm u 1 2 which is 1. So, here you have got norm a 2 to be less than or equal to mod lambda 1, here you have got mod lambda 1 to be less than or equal to norm a 2, so combining you get norm a 2 to be equal to mod lambda 1.

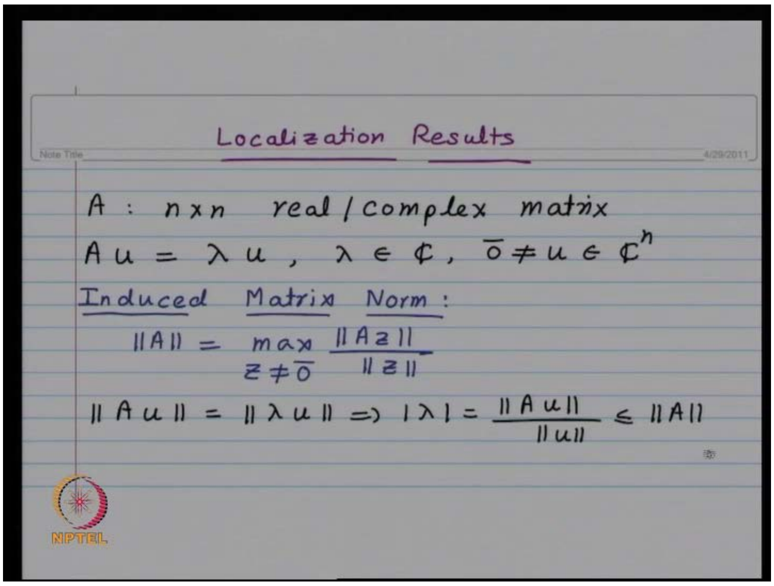
(Refer Slide Time: 07:54)



There is the result that, if A is normal, lambda 1 lambda 2 lambda n are Eigen values of A, which are arranged in such a manner that mod lambda 1 bigger than or equal to mod lambda 2, bigger than or equal to mod lambda n, then 2 norm of A is nothing but modulus of lambda 1. Now, here we have just obtain a formula, but norm A 2, it is still we cannot compute, because we cannot compute the Eigen values for all matrices. For special matrices, if it is a upper triangular matrix, I know how to calculate its Eigen values.

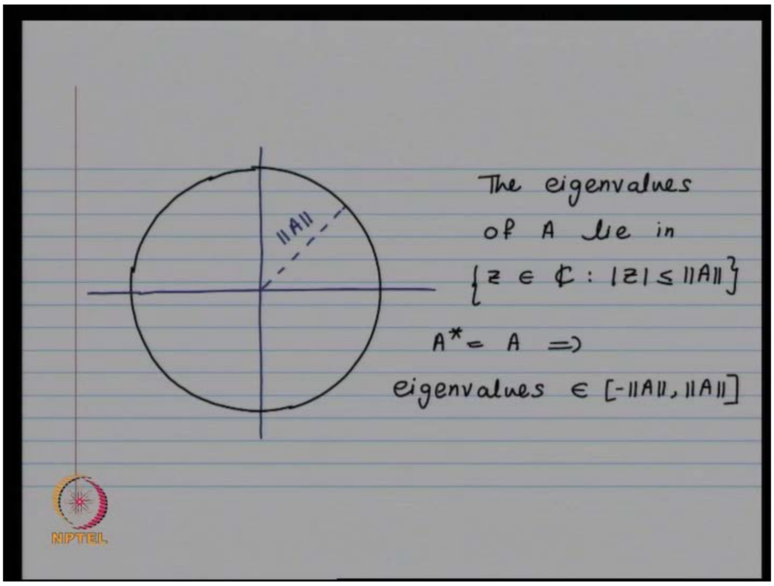
So, for norm A 2, we did not have a formula, in terms of the elements of the matrix as in case of 1 norm or infinity norm, and now we have got norm a 2 is equal to mod lambda 1, where both left hand side and right hand side are not computable, but still we have got a relation. So, now we are going to consider localization of Eigen values.

(Refer Slide Time: 09:16)



So, we have the simplest one which we get is, if A is a, this is going to be always the case. A will be either real or a complex matrix, if lambda is a eigen value, then you have A u is equal to lambda u, where u is a not non-0 vector. If you consider induced matrix norm, then we have got norm a to be maximum of norm a z by norm z. So, taking norm here, you get norm A u is equal to norm lambda u. So, that gives you mod lambda to be equal to norm A u by norm u, I can divide by norm u, because u is a non-0 vector and this will be less than or equal to norm a.

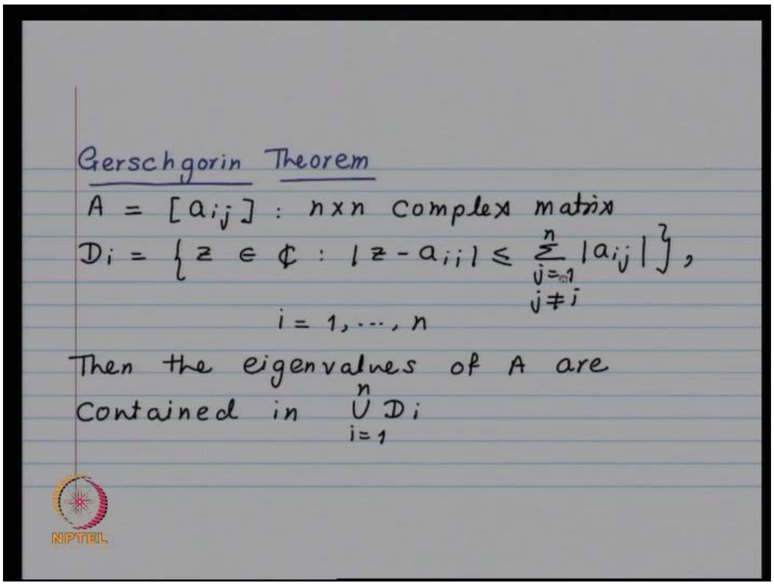
(Refer Slide Time: 10:03)



So, if you consider a disc with center 0 and radius norm A, then the Eigen values they are going to lie in this close disc, set of all z belonging to c such that mod z is less than or equal to norm A. This norm can be any induced matrix norm. It can be 1 norm, it can be infinity norm, it can be 2 norm or whichever norm you take. You fix a vector norm consider the induced matrix norm, then your mod lambda is going to be less than or equal to norm of a. This result tells us that if i take any z outside this disc, then my a minus z i is going to be invertible. If your matrix is a special matrix like, if it is a self-adjoint matrix, then what we know is the Eigen values they are going to be real. So, we can say that for self-adjoint matrix, the Eigen values they will lie in the interval, closed interval, minus norm a to plus norm A. We have to take intersection of the close disc with the real line.

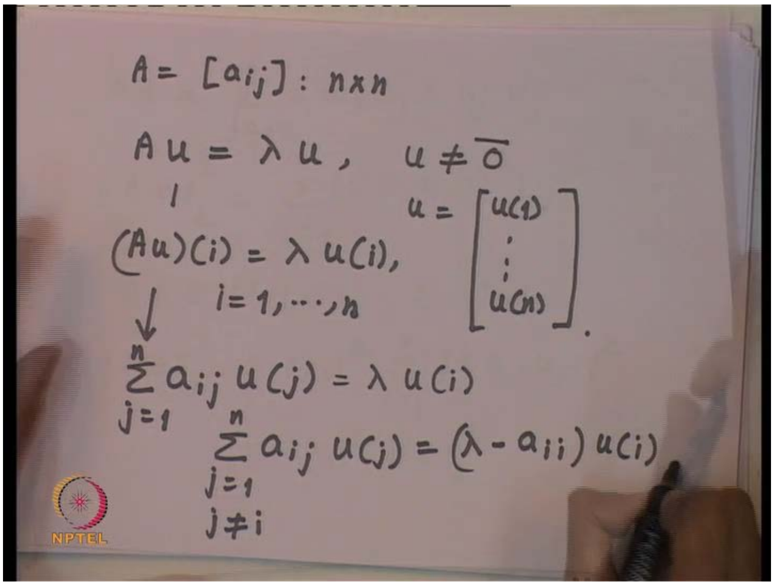
If you are considering or if your matrix is skew self-adjoint, then you have to take intersection of this close disc with the imaginary axis. Now, we are going to consider what is known as Gerschgorin theorem. So, that again gives us regions in the complex plane, which are going to contain all our Eigen values and as a consequence of Gerschgorin theorem, we will show that if the matrix is diagonally dominant, then such a matrix is invertible.

(Refer Slide Time: 12:19)



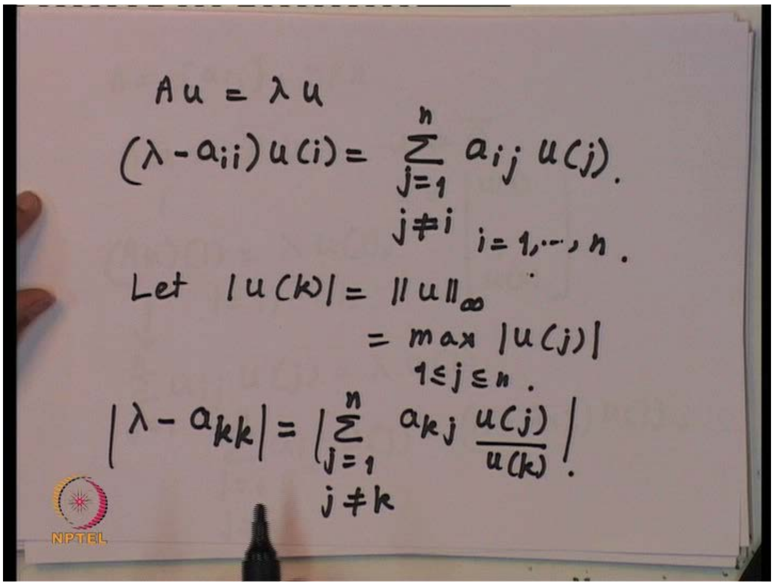
So, here is the Gerschgorin theorem A is n by n complex matrix, you look at D i to be the disc with center a i i and radius to be sum of half diagonal entries in that particular row. So, you are going to have n such discs, then the eigen values of a, they are going to be contained in the union of this disc. So, the matrix a is given to us. We look at the disc with center a i i. So, look at the ith row. So, take the center to be a i i and radius to be the entries in the same row except the diagonal entry, take their modulus and then sum. So, that is going to be the radius. So, you have such n disc, so all our eigen values they will lie in the union of this n disc and the proof is not difficult. So, let us prove this.

(Refer Slide Time: 13:47)



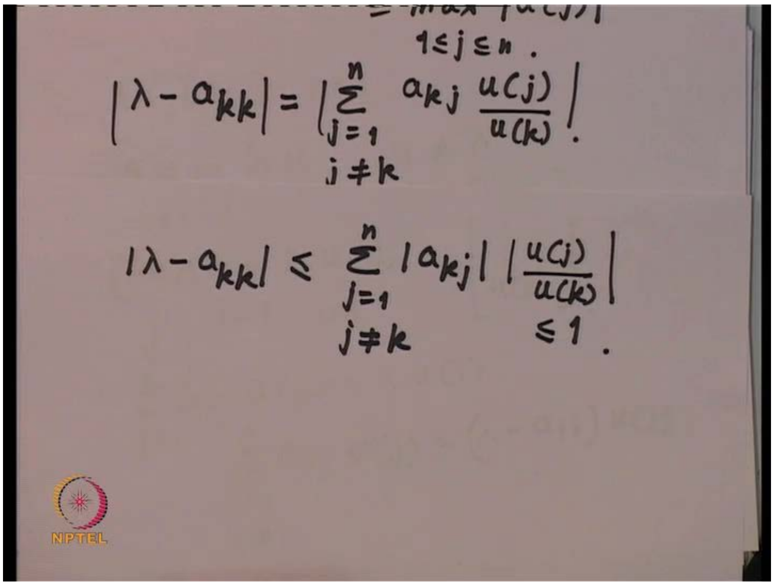
So, we have got A to be a n by n matrix whose i jth entry is given by a i j. Then suppose you have got a u is equal to lambda u, where u is a non-0 vector. Let me write u as u 1 up to u n, a column vector. Now, here a u is equal to lambda u, this is equality of 2 vectors. So, this means you have got ith component of a u is equal to lambda times u i, i goes from 1 to up to n. Now this is nothing but, summation a i j u j, j going from 1 to n lambda times u i . So, what I will do is, the term which contains u i, i will take on the other side and the remaining terms, i will keep here. So, we have got summation j goes from 1 to n a i j u j, j not equal to i is equal to lambda minus a i i u i.

(Refer Slide Time: 15:32)



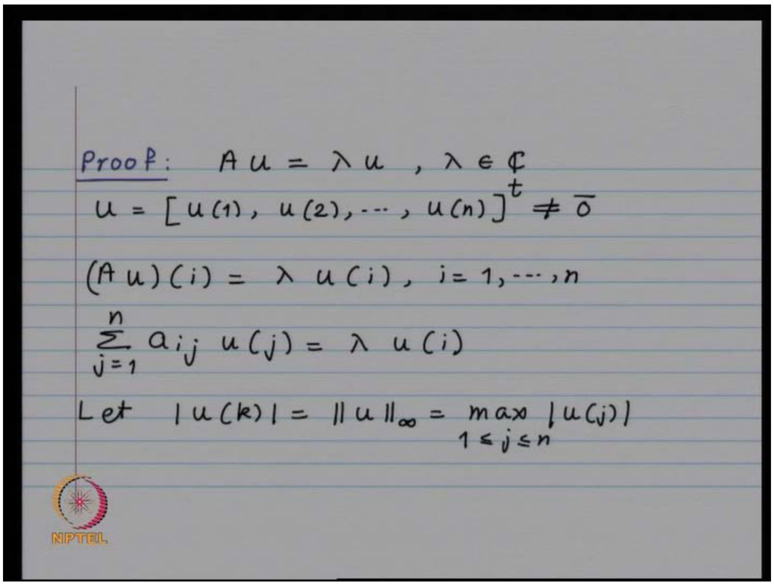
So, we have got a u is equal to lambda u and from here we deduced that lambda minus a i i u i is equal to summation j goes from 1 to n, j not equal to i a i j u j, i going from 1 to up to n. Let modulus of u k be norm u infinity norm; that means, maximum of modulus of u j 1 less than or equal to j, less than or equal to n. This is true for i is equal to 1 to up to n. So, in particular it will be true for i is equal to k. So, i will have lambda minus a k k is equal to summation j goes from 1 to n, j not equal to k, a k j u j divided by u k. u is a non-0 vector. So, u k will not be equal to 0. Take modulus and use triangle in equality. So, we will have modulus of lambda minus a k k to be less than or equal to summation j goes from 1 to n, modulus of a k j j not equal to k and then mod u j by u k, and this is going to be less than or equal to 1. So, we have got this.

(Refer Slide Time: 17:16)



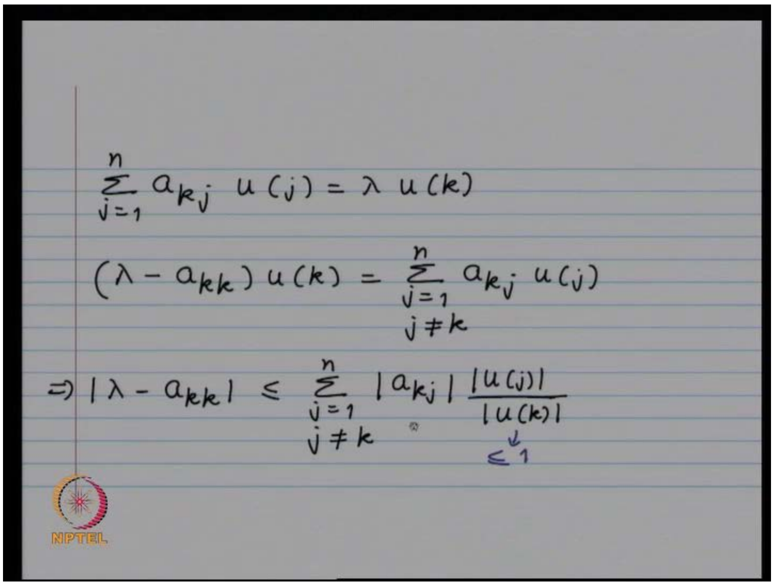
So, we started with a eigen value lambda. Then we looked at corresponding eigen vector. For this eigen vector, we looked at component u k, where mod u k is equal to norm u infinity. There can be more than 1 such k, if your vector is a constant vector then k will be any component. Any way it does not matter. So, mod u k is equal to norm u infinity and then we showed that modulus of lambda minus a k k is less than or equal to summation j goes from 1 to n, j not equal to k and then modulus of a k j. So, here the catch is, we do not know lambda, we do not know eigen vector u. Then if I do not know eigen vector u, I cannot know what is the k, where mod u k is equal to norm u infinity. So, this whatever estimate I have got it is not of much use. It says that your eigen value lambda is going to lie in the disc with center a k k and radius to be some of the moduli of half diagonal entries in that kth row, but we do not know what is the that row. So, that is why what we do is I do not the row. So, I will do it for each such row and then take their union. So, then I will know that my lambda has to be in the 1 of the disc.

(Refer Slide Time: 19:54)

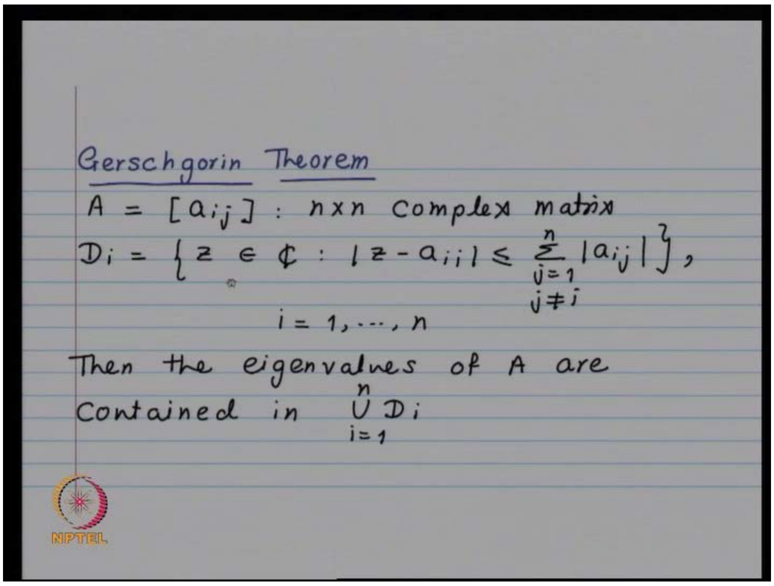


Is your a u is equal to lambda u, u is vector non-0 vector, then we have got we looked at a u i is equal to lambda times u I, i is equal to 1 to up to n, wrote down what it means mod u k was norm u infinity. And then we got modulus of lambda minus a k k to be less than or equal to summation j goes from 1 to n j not equal to k modulus of a k j. Since we do not know k, we are going to look at all such discs. So, d i is set of all z belonging to c such that modulus of z minus a i i to be less than or equal to this and then the eigen values of the A they are going to be contained in union of d i i goes from 1 to up to n.

(Refer Slide Time: 20:15)

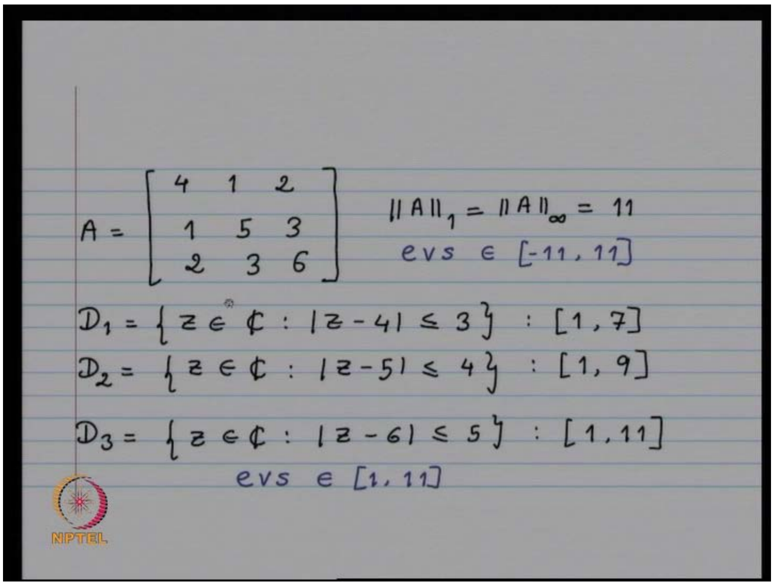


(Refer Slide Time: 20:36)



So, let us look at a example and try to find that the region in which your eigen values are going to lie by using our first estimate, that modulus of lambda is less than or equal to norm A. So, that norm we can take either 1 norm or infinity norm and another result which we have got is this Gerschgorin theorem.

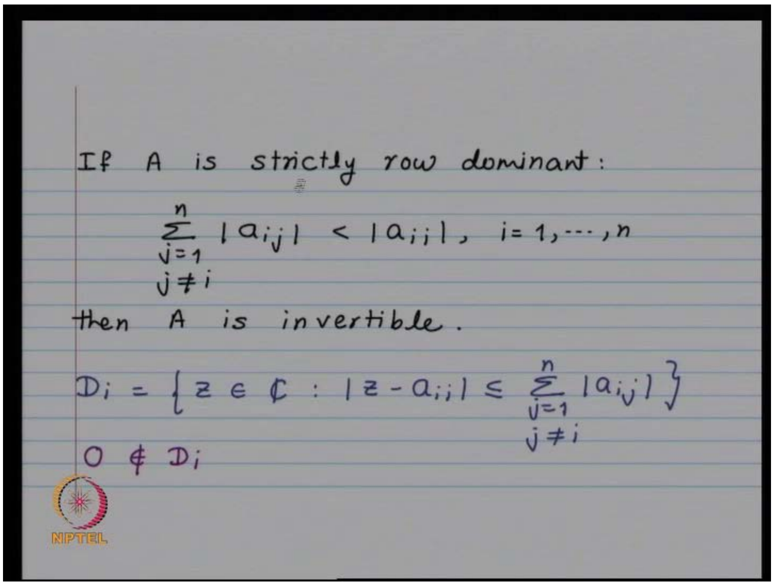
(Refer Slide Time: 21:27)



So, here is a 3 by 3 matrix, the matrix is a symmetric matrix, you can see that 1 norm is equal to infinity norm is equal to 11, because for infinity norm, we are going to look at the rows. So, it will be 4 plus 1 plus 2. So, that is going to be 8, then second row it will be 9 and third row it is 11. So, norm a 1 is equal to norm a infinity is equal to 11. As the matrix is a real matrix, eigen values they are going to be real. So, eigen values they will lie between minus 11 to plus 11.

Now, let us look at Gerschgorin disks. So, D 1 will be modulus of z minus 4 less than or equal to 1 plus 2 that is 3. D 2 is going to be set of all z belonging to c, such that modulus of z minus 5 is less than or equal to sum of half diagonal entries. So, 1 plus 3 is equal to 4. Then d 3 will be set of all z belonging to c, such that mod z minus 6 less than or equal to 5. As I know that the eigen values are real, it is suffices to look at the intervals. So, D 1 will be interval 1 to 7, d 2 will be interval 1 to 9 and d 3 will be interval 1 to 11. When you take the union of these 3 intervals, it is going to be 1 to 11. So, here Gerschgorin theorem tells us that eigen values, they will be in the interval 1 to 11 whereas the norm thing, it gave us interval to be minus 11 to plus 11. Also look at the interval. It does not contain 0, so that means, 0 cannot be an eigen value of a, because all eigen values they have to be in the interval 1 to 11. So, 0 not an eigen value; that means, A is going to be invertible matrix. So, now is the result which I said that if the matrix is strictly row dominant; that means, the diagonal entry modulus of a i i is bigger than sum of the moduli of remaining entries. Then such a matrix is going to be invertible.

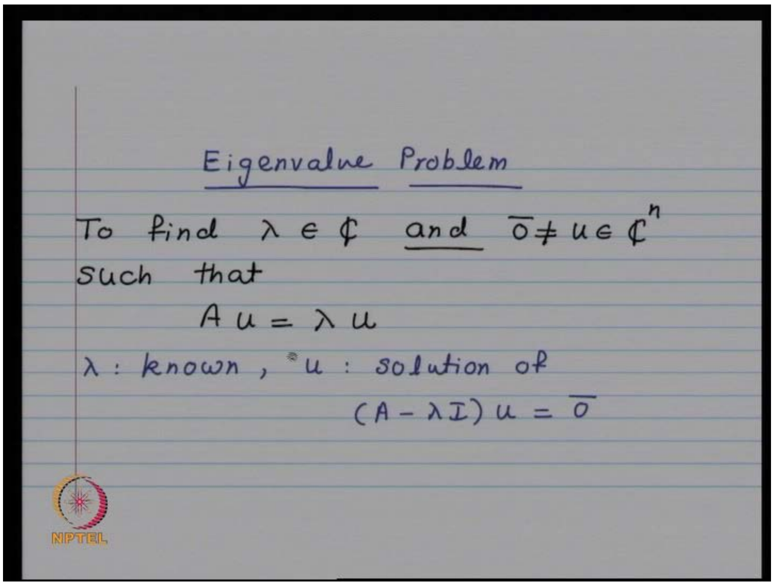
(Refer Slide Time: 24:30)



So, you have summation j goes from 1 to n modulus of a i j, j not equal to i to be strictly less than modulus of a i i, then A is invertible. Our D i is set of all z belonging to c, such that modulus of z minus a i i is less than or equal to this number. Now, 0 cannot be in D I, because if 0 belongs to D I, it will mean that modulus of a i i is less than or equal to this. So, it will contradict this. So, 0 does not belong to any of the discs and all eigen values they are in the union of these discs. So, a strictly row dominant matrix is going to be invertible.

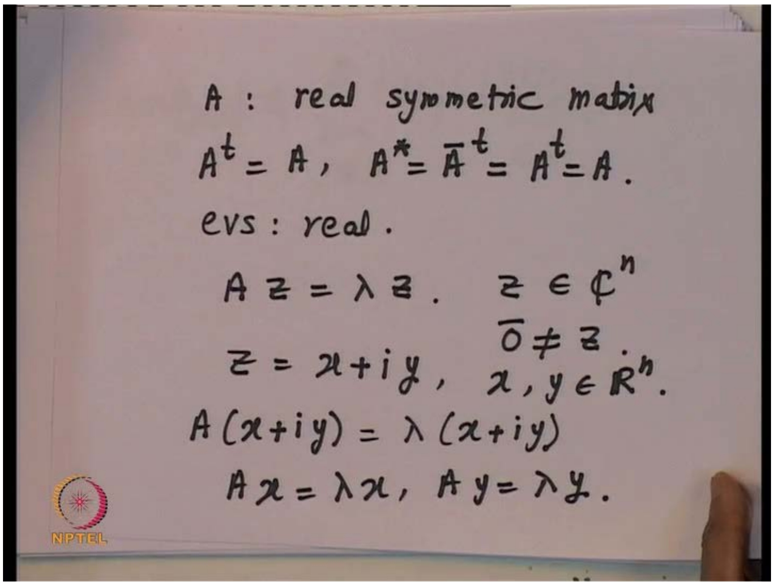
Now, what about strictly column dominant? That means, in a column, suppose the diagonal entry modulus of a i i is going to be bigger than sum of the half diagonal entries in the column. Earlier we have looked at rows, now we are looking at columns. Will such a matrix be invertible? So, the answer is yes, because what we can do is, if you have such a matrix look at its transpose. So, if a matrix is diagonally column dominant, then a transpose will be diagonally row dominant and diagonally row dominant means invertible we have seen just now. So, if A is diagonally column dominant, then a transpose will be invertible, but if a transpose is invertible, then a also is invertible, because a transpose inverse is same as a inverse transpose.

(Refer Slide Time: 26:40)



So, when we started eigen values, I said that till now we were restricting ourselves to real numbers. Now we have to go to complex numbers. If your matrix is real symmetric matrix, then you can restrict yourselves to real numbers. So, the matrix is real symmetric; that means, it is going to be self-adjoint a star is equal to a. So, the eigen values they are going to be all real. Now the question is what about eigen vector? If i can show that it has a real eigen vector, then i need not go to complex numbers. Now that is so, let me show you, so, we have got a to be real symmetric matrix. So, we have got a transpose is equal to a and a star which is a bar transpose, it will be equal to a transpose, it will be equal to a. So, our eigen values, they are going to be real.

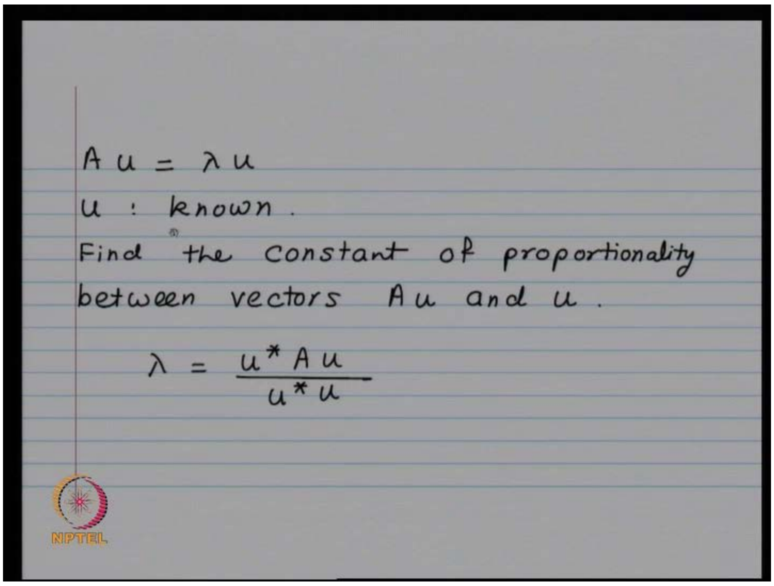
(Refer Slide Time: 27:40)



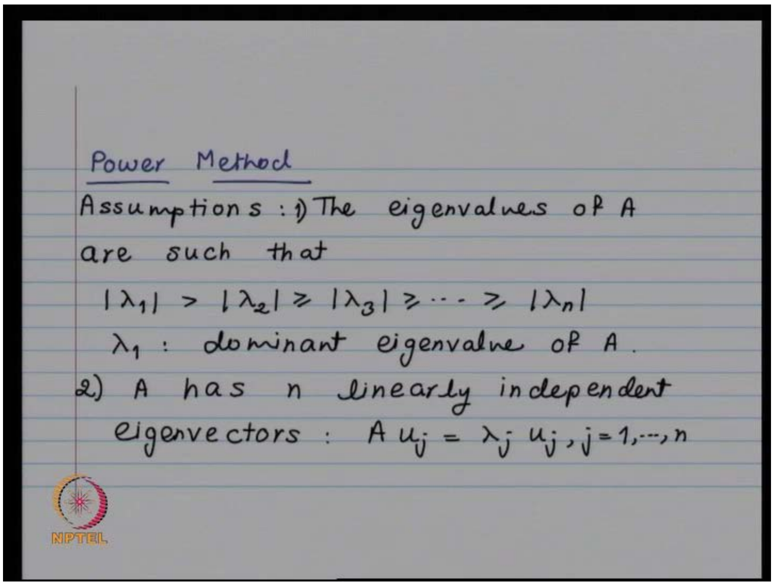
Now, look at A z to be equal to lambda z, where z belongs to c n and z is not equal to 0; that means, z is eigen vector. So, this z, I can write as x plus i y where x and y, these are in r n. Now let me substitute, so i will get a of x plus i y is equal to lambda times x plus i y. Elements of a are real lambda is real. So, this gives you a x is equal to lambda x a y is equal to lambda y. Since z is not a 0 vector, we have got either x is not 0 or y is not 0 or both are not 0. So, we have got, we started with z to be a eigen vector, then we looked at its real and imaginary parts. So, you have got vector x, vector y, both x and y are going to be satisfying a x is equal to lambda x a y is equal to lambda y. So, they will be eigen vectors provided they are non-0

Since z is not equal to 0 vector, either x is not 0 or y is not 0 or both are not 0 and x and y, these are real vectors. So, thus for a real symmetric matrix, eigen values are real, the entries of the matrix are real and you can choose your eigen vector to be a real vector. So, there we can restrict ourselves to the real numbers. Now when we look at eigen value problem, it is that simultaneously we have to find a complex number lambda and a complex vector u, such that a u is equal to lambda u. So, this is what it makes it difficult, that you have to simultaneously calculate or find lambda and u. Suppose lambda is given to you and you want to find u, such that a u is equal to lambda u, then it is easy. You have got lambda is known, you want a u is equal to lambda u, so u is going to be solution of a minus lambda I, u is equal to 0 vector, lambda is eigen value. So, a minus lambda i will be a singular matrix. So, you get a homogeneous system with coefficient matrix to be singular. So, it will have always have a non-trivial solution and we know how to calculate the solution of system of linear equations. So, if lambda is known, you can calculate eigen vector. On the other hand if eigen vector is known then lambda is nothing but constant of proportionality. So, u is given to you, matrix a is known. So, calculate a u, calculate u and then they are going to be multiple of each other. So, whatever is that multiple, that is lambda. Or you can look at lambda to be equal to u star a u divided by u star u. So, thus if eigen value is known you can calculate eigen vector, if eigen vector is known you can calculate eigen value.

(Refer Slide Time: 31:53)

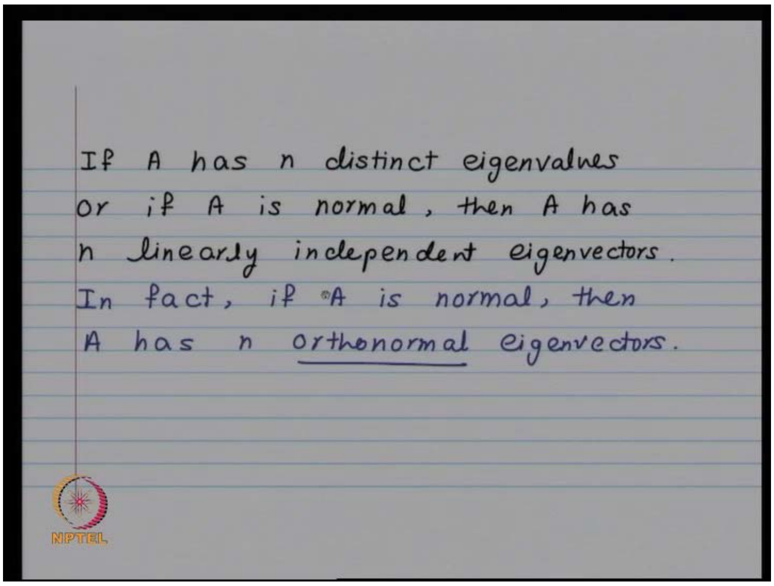


(Refer Slide Time: 32:40)

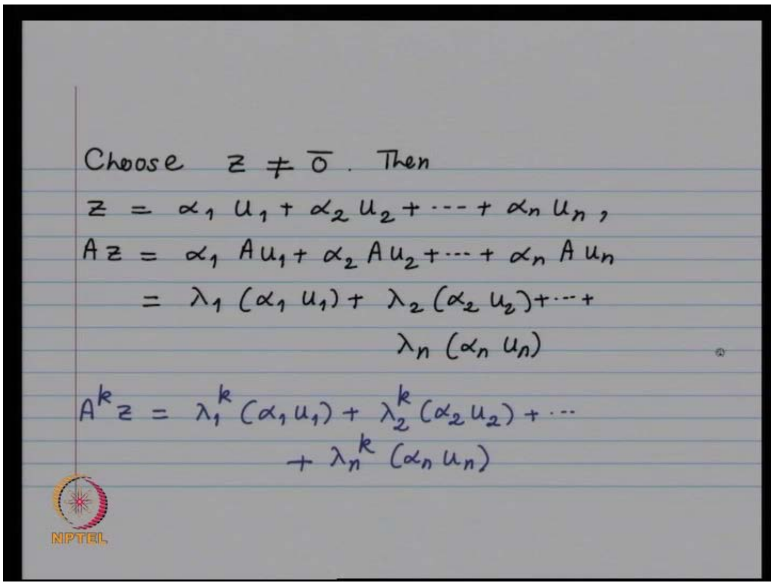


So, now here is the method, which is known as power method for calculating approximation to a dominant eigen value. So, these are our assumptions, that the eigen values of a are such that mod lambda 1 is strictly bigger than mod lambda 2, bigger than or equal to mod lambda 3, bigger than or equal to mod lambda n. So, such a matrix, this is known as dominant eigen value. So, here you have to notice this strictly bigger than, this is essential. The second assumption is a has n linearly independent eigen vectors. So, you have a u j is equal to lambda j u j, j is equal to 1 to up to n, lambda j's can be repeated except for the first 1. This biggest eigen value, this should be simple, it should not be repeated. So, now these are our assumptions and the second assumption will be satisfied for class of normal matrices or if a has n distinct eigen values, then also the second assumption is satisfied and if the matrix a is a normal matrix then. In fact, a has n orthonormal eigen vectors.

(Refer Slide Time: 34:13)



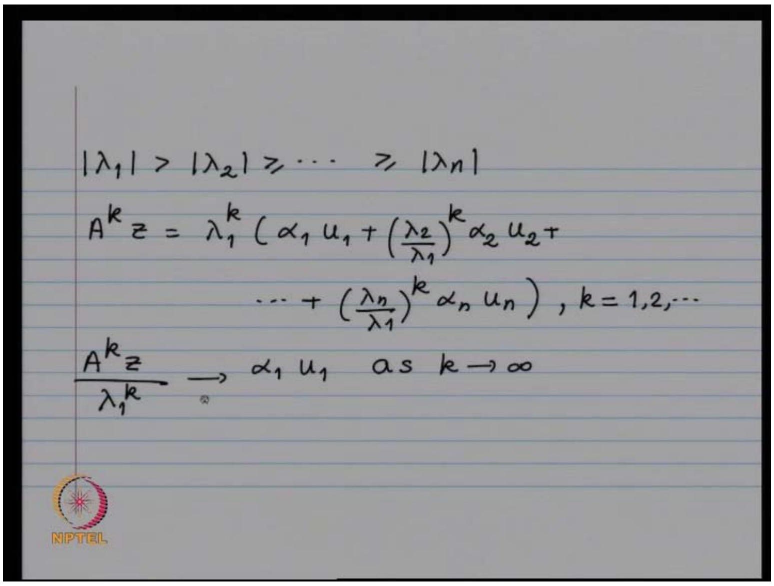
(Refer Slide Time: 34:24)



So, what is the method? The method is you choose z to be a non-0 vector. This z you can write as a linear combination of u 1 u 2 u n. So, z is equal to alpha 1 u 1 plus alpha 2 u 2 plus alpha n u n. Apply a to this. So, a z will be alpha 1 a u 1 plus alpha 2 a u 2 plus alpha n a u n. a u 1 is equal to lambda 1 u 1, so this will be lambda 1 alpha 1 u 1. a u 2 is lambda 2 u 2 so it is lambda 2 alpha 2 u 2 plus lambda n alpha n u n. a raise to k z will be lambda 1 raise to k alpha 1 u 1 plus lambda 2 raise to k alpha 2 u 2 plus lambda n raise to k alpha n u n.

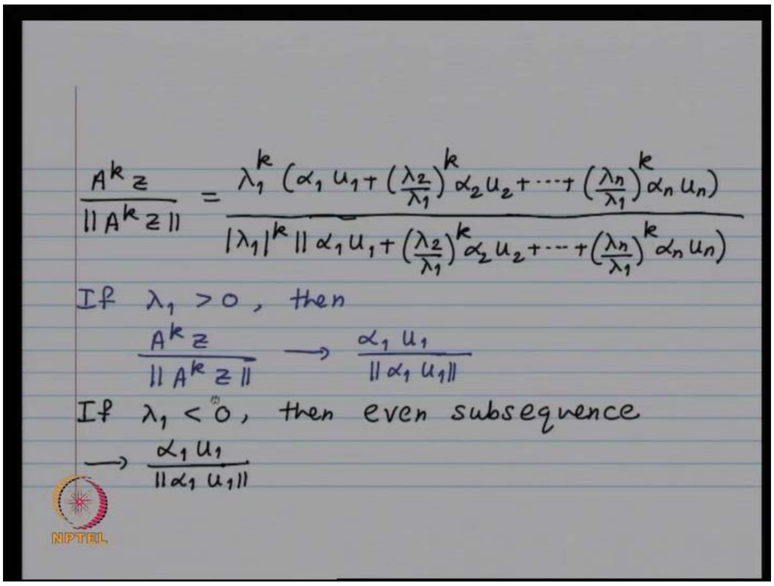
So, see what is happening? You start with a non-0 vector, any vector z. This z will have component in the direction of u 1. So, you write z as alpha 1 u 1 plus alpha 2 u 2 plus alpha n u n, then you keep applying a. So, the component in the direction of u 1, which was that alpha 1, that is getting multiplied by lambda 1 raise to k. The component in the direction of u 2, that is getting multiplied by lambda 2 raise to k. We are assuming mod lambda 1 to be bigger than mod lambda 2. So, this component in the direction of u 1 will become more significant. So, that is the idea of the power method that if you consider a raise to k z divided by lambda 1 raise to k. So, we have got a raise to k z to be alpha 1 lambda 1 raise to k u 1 plus alpha 2 lambda 2 raise to k u 2 and so on.

(Refer Slide Time: 36:34)



So, take lambda 1 raise to k common and when you look at a raise to k z by lambda 1 raise to k, this is going to converge to alpha 1 u 1 as k tends to infinity. Now, we do not know lambda 1, we do not know lambda 1, we do not know u 1. So, this a raise to k z by lambda 1 raise to k, even if it is converging to a multiple of eigen vector u 1, this is not something which you can calculate.

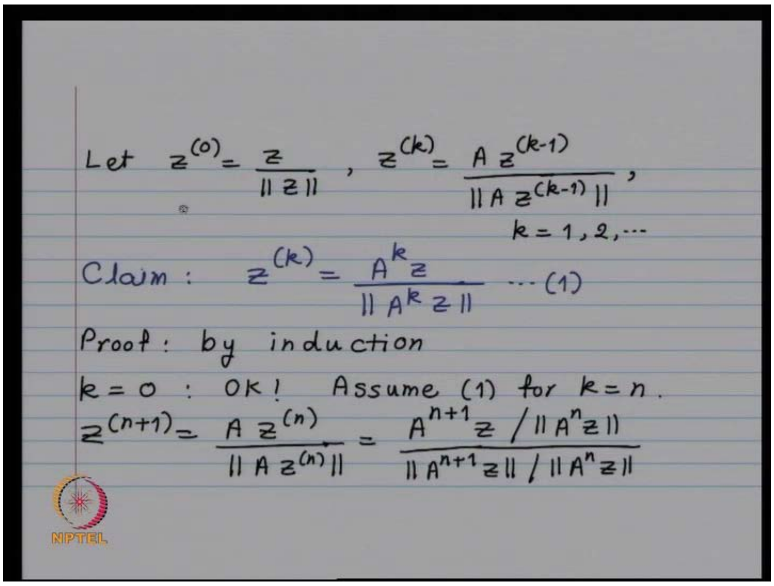
(Refer Slide Time: 37:35)



But this is just a normalization. So, you look at a raise to k z divided by norm of a raise to k z. So, here you have lambda 1 raise to k, alpha 1 u 1 plus lambda 2 by lambda 1 raise to k alpha 2 u 2 and so on, when k tends to 0, this is going to tend to 0, this is going to tend to 0, then if lambda 1 is bigger than 0, then lambda 1 will be equal to mod lambda 1, this will get cancelled and a raise to k z divided by norm of a raise to k z will tend to alpha 1 u 1 divided by norm of alpha 1 u 1.

If lambda 1 is less than 0, then even subsequence will tend to alpha 1 u 1 divided by norm of alpha 1 u 1. So, now we have got this convergence. It is a raise to k z divided by norm of a raise to k z. The matrix a is given to us its very simple to implement. If your matrix a, it is a spars matrix; that means, there are lot of 0. Then calculating a raise to k z will not involve much computations and then it is going to give us approximation to the eigen vector to an eigen vector corresponding to dominant eigen value. For the sake of stability, this a raise to k z by norm of a raise to k z, we will write it in a different form. So, we have, suppose you consider so, z is our starting vector, u divide by its norm and then call it z 0. Then you apply a to this, so, you will be a z 0 and immediately divide by its norm.

(Refer Slide Time: 39:58)



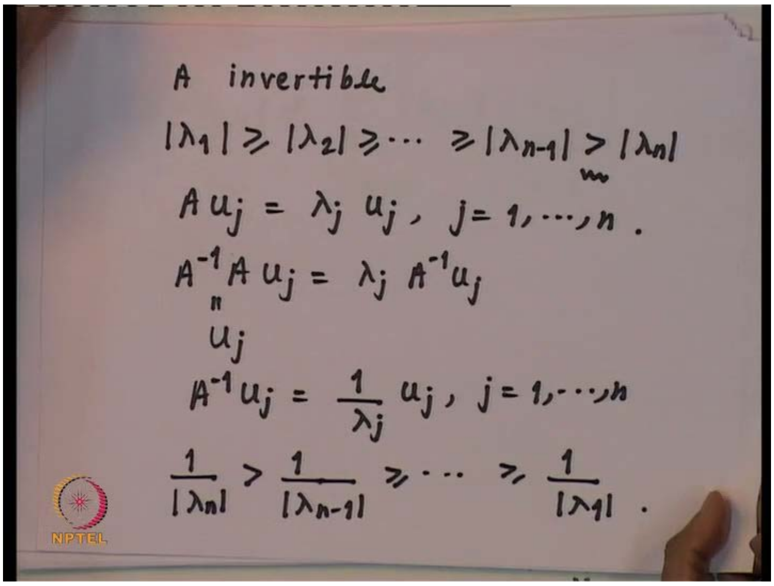
So, here is the sequence, z 0 is equal to z upon norm z and z k is a z k minus 1 divided by norm of a z k minus 1. Now, it is equivalent formulation. This was our z k is equal to a raise to k z divided by norm of a raise to k z. So, here what we are doing is, we are applying a raise to k to z and then dividing by its norm. Here at each stage, you are applying a and immediately dividing by its norm. And this proof is by induction, if k is equal to 0, then it will be a raise to 0; that means, identity. So, z upon norm z. So, that is our z 0, assume the result to be true for k is equal to n consider z n plus 1. This is our definition. z n plus 1 will be a, z n divided by norm of a raise a z n by induction hypothesis, z n will be a raise to n z divided by norm of a raise to n z. You are applying a. So, the numerator will be a raise to n plus 1 z divided by norm of a raise to n z and denominator will be norm of this. So, it will be norm of a raise n plus 1 z by norm of a raise to n z. So, this will get cancelled and we get 1 for n plus 1. So, you assume for k is equal to n and then u obtained for n plus 1. So, this is power method.

Now, there is only one slight catch, what we said was z is arbitrary vector. You write z as alpha 1 u 1 plus alpha 2 u 2 plus alpha n u n, then you define our power iterates and it is going to converge to alpha 1 u 1 divided by norm of alpha 1 u 1, z is a arbitrary vector. What if alpha 1 is equal to 0? We are saying take any vector z which is non-0. So, it can very well happen, that alpha 1 is equal to 0 and still z is a non-0 vector, because z will be alpha 1 u 1 plus alpha 2 u 2 plus alpha n u n, alpha 1 is 0. In that case, what we will get is, our vector that a raise to k z divided by norm of a raise to k z, that will tend to a 0 vector. We are trying to find an eigen vector. So, this is of no use, but what we are doing is z is chosen randomly c n is n dimensional space. So, it is unlikely that the vector which is chosen arbitrarily lies in n minus 1 dimensional subspace, like look at r 2, in r 2 1 dimensional subspace will be a straight line. So, it is really highly unlikely that the vector which you are choosing arbitrarily is going to lie along that particular line.

We have been always talking about round off errors, the problems it creates. Now here is an example, where it is rather useful, that suppose by stroke of luck, the vector which we are choosing, it is in the span of u 2 u 3 u n; that means, alpha 1 is equal to 0, but round of error is there. So, alpha 1 will never be equal to 0, it will be a small number, and now when you go on applying a, because our lambda 1 is the dominant eigen value that component will go on increasing. So, even though the starting point alpha 1 was a small number, when you perform the iterates, it is going to become big.

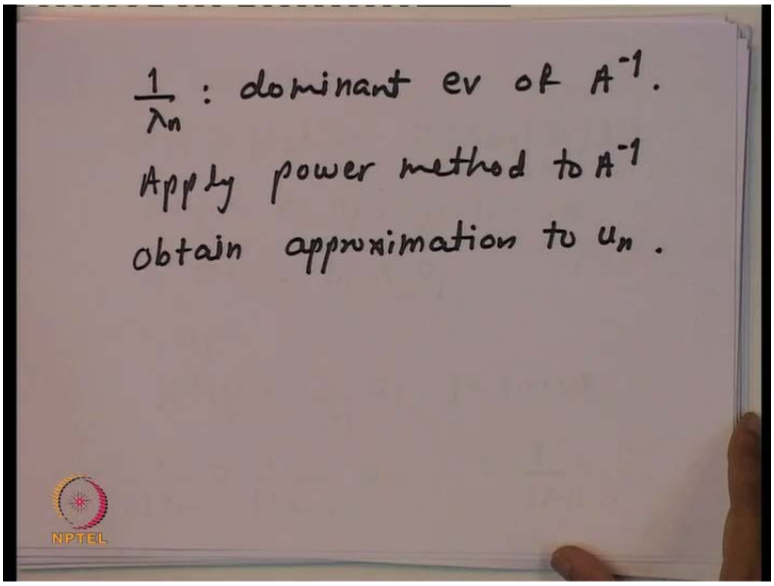
So, this power method, it is going to give us approximation, 2 eigen vector corresponding to the dominant eigen value. So, this is rather restrictive, that only the largest eigen value in modulus we can approximate. But then there are some extensions, suppose your matrix a is invertible matrix and lambda 1 lambda 2 lambda n, these are eigen values of a, then 1 upon lambda 1, 1 upon lambda 2, 1 upon lambda n, they will be eigen values of a inverse. So, apply power method to a inverse and then you can approximate 1 by lambda n.

(Refer Slide Time: 46:08)



So, we have a invertible mod lambda 1 bigger than or equal to mod lambda 2, bigger than or equal to mod lambda n minus 1 and here now, we have got strict in equality. We have a u j is equal to lambda j u j, j is equal to 1 to up to n. So, a inverse a u j will be equal to lambda j a inverse u j, this is identity. So, it will be u j, so a inverse u j will be 1 upon lambda j u j, j is equal to 1 to up to n and 1 upon mod lambda n will be strictly bigger than 1 upon mod lambda n minus 1 bigger than or equal to 1 upon lambda 1. So, thus 1 upon lambda n is dominant eigen value of a inverse. So, you apply power method to a inverse.

(Refer Slide Time: 47:41)



So, you consider. So, 1 upon lambda n is dominant eigen value of a inverse, apply power method to a inverse and obtain approximation to u n. So, I know how to calculate the eigen value of biggest modulus, how to calculate or how to rather calculate approximations to eigen value of least modulus. Now what about the eigen value in between? So, if you have some initial approximation to an intermediate eigen value, then we have got a inverse power method and using that inverse power method, we can find eigen vector corresponding to the intermediate eigen value.

So, that is going to be that we are going to do in the next lecture. So, in the next lecture, we will consider inverse power method and then we will start our discussion towards q r method, which is the most popular method for calculating eigen values of matrix a at present. So, thank you.