Numerical Analysis

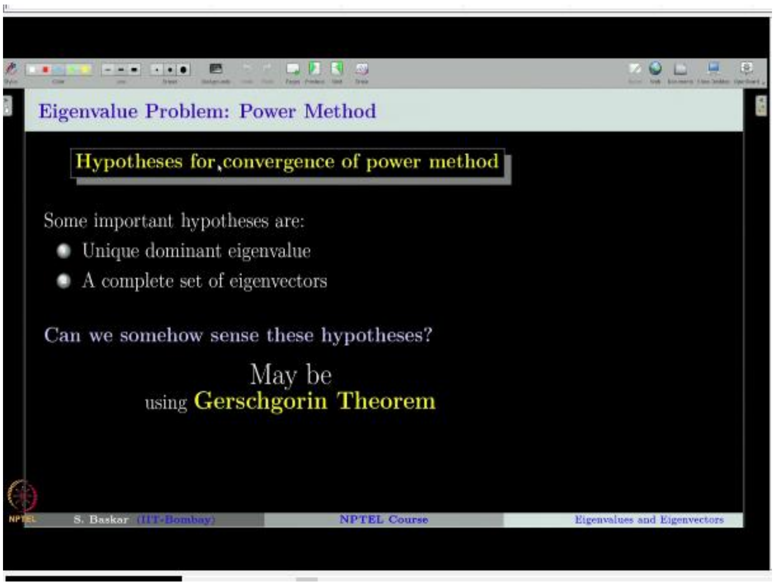
Prof. Dr. S. Baskar
Department of Mathematics
Indian Institute of Technology-Bombay

Lecture-25

Eigenvalues and Eigenvectors: Gerschgorin's Theorem and Applications

Hi, in this lecture let us learn an interesting and important theorem called Gerschgorin's theorem. Gerschgorin's theorem can be used to localize eigenvalues of a given matrix. You may be asking why are we doing this theorem in between our discussion on power method. Gerschgorin's theorem gives us how the eigenvalues are located in the complex plane that can be used to verify some of the hypothesis of power method.

(Refer Slide Time: 00:53)



If you recall, power method can be used to compute the dominant eigenvalue and the corresponding eigenvector of a given matrix. This may be viewed as a disadvantage of power method because it cannot compute other eigenvalues of a matrix when applied once; it only can capture dominant eigenvalues. In fact in the next lecture, we will see how to modify the power method to get other eigenvalues also.

But the fact is that, when applied power method once we get only one eigenvalue and the corresponding eigenvector. If we are directly applying the power method then we will get the dominant eigenvalue and the corresponding eigenvector. However, it looks like a disadvantage. In many applications we want only the dominant eigenvalue of a given matrix. One good example is the well known Google search engine.

Google search engine uses page rank algorithm to find the relevance of a site with the keyword given for searching. What Google search engine does is that for each site it assigns a number called page rank score then it arranges all these numbers in the form of a matrix called google matrix and Google will then compute a vector called page rank vector. The page rank vector is the eigenvector with certain properties of the corresponding dominant eigenvalue of the google matrix.

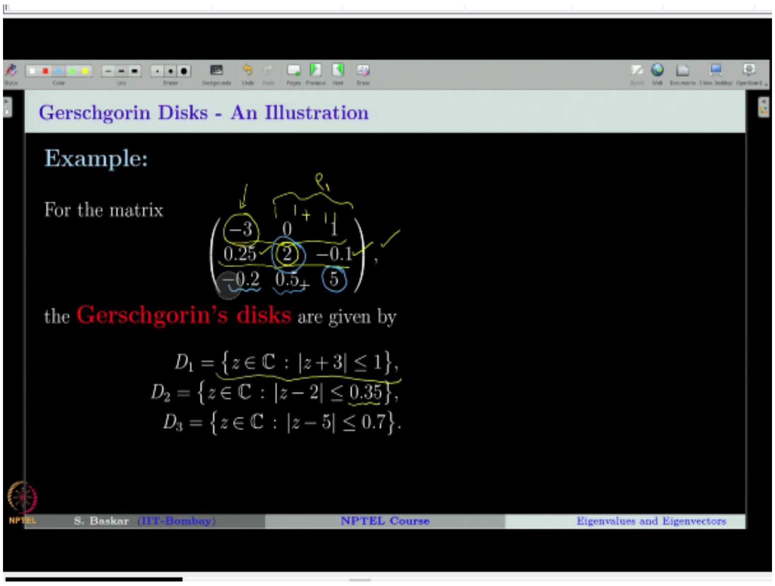
Therefore, Google does not need all the eigenvalues of the google matrix; it only wants the dominant eigenvalue of the google matrix. In that way power method is highly preferred because the google matrix in general can be with few billion dimensions therefore if you go for computing all the eigenvalues of the matrix. Say for instance if Google uses QR method which can give all the eigenvalues of the matrix then it will be very costly computationally, whereas power method will specifically capture only the dominant eigenvalue.

Therefore, power method is computationally very efficient in this case. Therefore, in fact Google search engine uses power method. Now we understood the importance of power method. The only disadvantage of the power method is when we want to understand the convergence of the power method then we have to check certain hypotheses which are not practically possible to check.

However, note that this hypothesis are not required if you want to just implement power method because to implement power method we only needs to give an initial guess vector which can be chosen arbitrarily, it is only that when we want to understand the convergence of the power method to the dominant eigenvalue then we need certain hypothesis. For instance, we need to know that the given matrix has a unique dominant eigenvalue and also we need to know that the matrix has a complete set of eigenvectors.

That is, a set of eigenvectors should form a basis for $\mathbb{R}^n$ when the matrix is $n \times n$ matrix. Now the question is can we somehow sense this hypotheses without explicitly knowing the eigenvalues and the eigenvectors? The answer is may be possible. In certain cases we can use the Gerschgorin's disk to conclude this. Now I will introduce you to Gerschgorin's theorem through a couple of examples and then I will state the theorem.

(Refer Slide Time: 05:16)



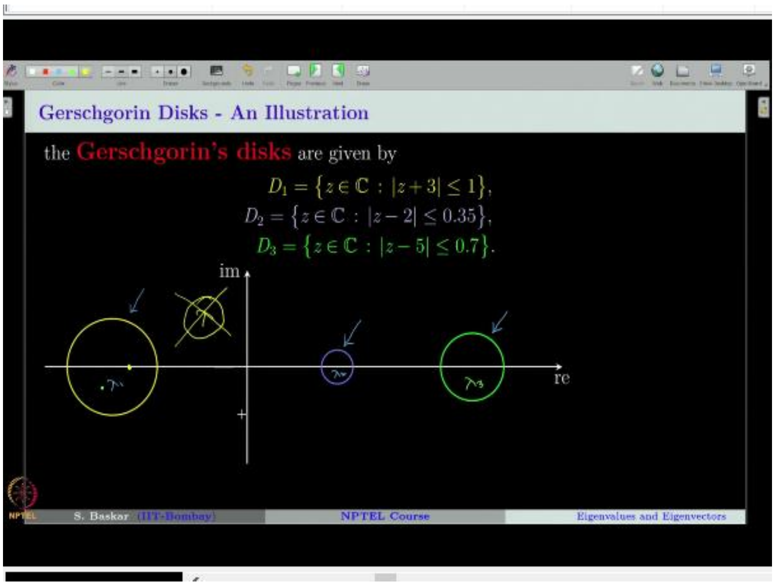
Let us first consider this 3×3 matrix. First, we have to construct Gerschgorin's disks. Each Gerschgorin's disk is associated with a row of the matrix. Therefore, if we are working with a $n \times n$ matrix then we can construct n Gerschgorin's disk. How to construct that? Let us take the first row of this matrix. You have to take the diagonal element as the center of the disk and then take the modulus of all the non diagonal elements of that row and then sum them up.

And that will be the radius which will call as row 1 and the center is the diagonal element. In that way the disk 1 which is coming from the first row is given by the set of all z in the complex plane such that $|z + 3| \le 1$. Remember these disks are constructed in the complex plane, because we know that a matrix can have complex eigenvalues.

In fact, Gerschgorin's disk theorem can localize the complex eigenvalues also. That is why we construct this disk in the complex plane. For the second disk we have to take the second row of the matrix. Again take the diagonal element as the center of the disk and take the modulus of the non diagonal elements of that row and sum them up. In this case it sums to 0.35. So, you have to take these two numbers.

Take their modulus and then sum them. That gives you the radius as 0.35 and the center is 2. Similarly, the disk 3 is given by the set of all $z \in \mathbb{C}$, such that $|z - 5|$, 5 is the center is less than or equal to mod 0.2 + mod 0.5. That sums to 0.7 which is the radius of the disk 3. So, this is the set of Gerschgorin's disk for the given matrix.

(Refer Slide Time: 07:49)

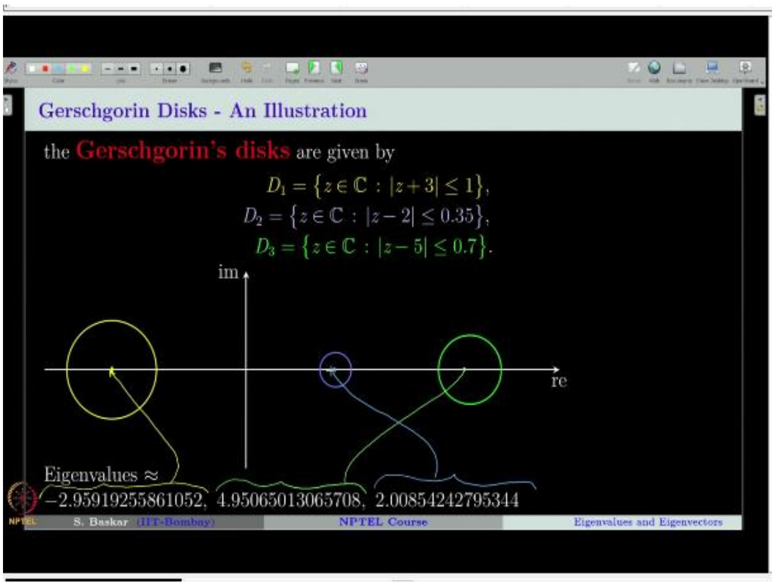


Let us visualize this disk. The disk D_1 is shown in the yellow colour, the disk D_2 is shown in blue colour and the disk 3 is shown in green colour. Now what the Gerschgorin's disk theorem says that all the eigenvalues of the matrix A will lie in this disk? In this case in fact the Gerschgorin's disk theorem says that you will have exactly one eigenvalue. Let us call it as μ_1 that belongs to this set.

And another eigenvalue say λ_2 lies in the disk D_2 and the third eigenvalue say λ_3 lies in the disk D_3. They can be anywhere in the disk. Suppose if it is here, it means it is a complex eigenvalue. If it lies on the real line it means it is a real eigenvalue, but what Gerschgorin's disk theorem surely says that you cannot find an eigenvalue which is outside the disk, that is not possible. That is what the Gerschgorin's disk says.

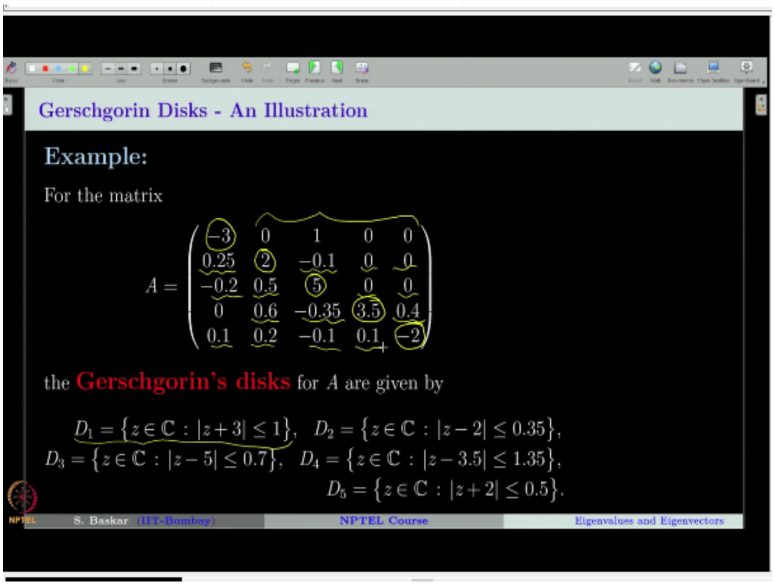
This is what we mean by saying that the Gerschgorin's theorem localizes the eigenvalues, it means it just gives you a particular set in which all the eigenvalues lie. In this case it is the union of these disks.

(Refer Slide Time: 09:26)



In fact, let us see what are all the eigenvalues of the matrix A that we have taken. You can see that 1 eigenvalue is given by -2.9591 and so on. You can clearly see that it lies in this disk and the second eigenvalue that we have computed is 4.95 and so on. And that you can clearly see that it lies in this disk, whereas the third eigenvalue that we have computed is 2.0085 and so on and that lies in this disk. So, this particular matrix respects the Gerschgorin's disk theorem very well.

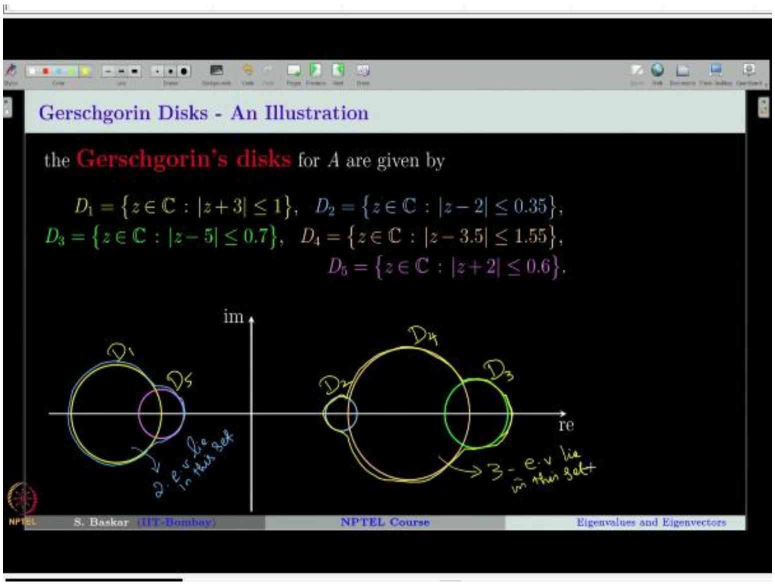
(Refer Slide Time: 10:13)



Let us take another example. Here we have a 5×5 matrix. Again, you have to construct the Gerschgorin's disk; since there are 5 rows we will have 5 Gerschgorin's disk; disk D_1 is given like this where the center is -3 and the radius is obtained by summing the absolute values of these elements. Similarly, the second this D_2 has centered 2 and the radius is obtained by summing 0.25 + 0.1.

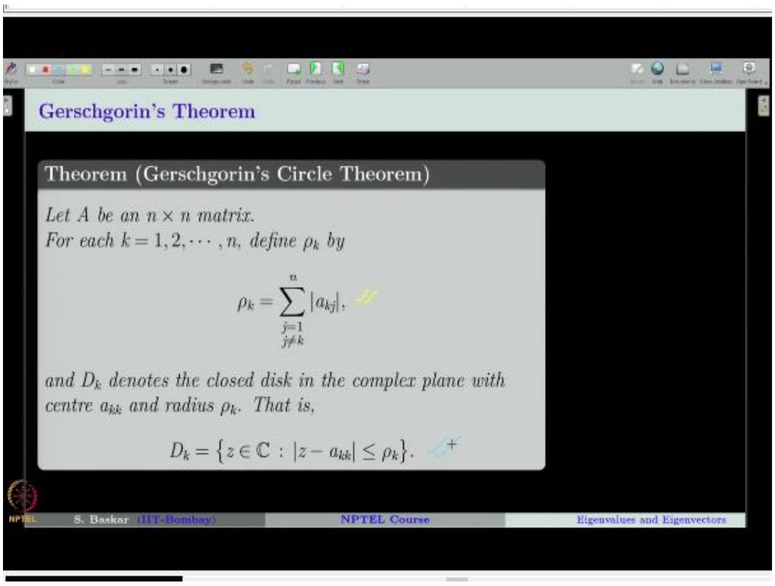
The third disk is again the center is at 5 and the radius is 0.7 because you are summing the absolute values of these elements. So, the disk 4 has center at 3.5 and the radius as the sum of the modulus of these elements that adds to 1.35 and finally the disk D_5 has center at -2 and its radius is 0.5.

(Refer Slide Time: 11:33)



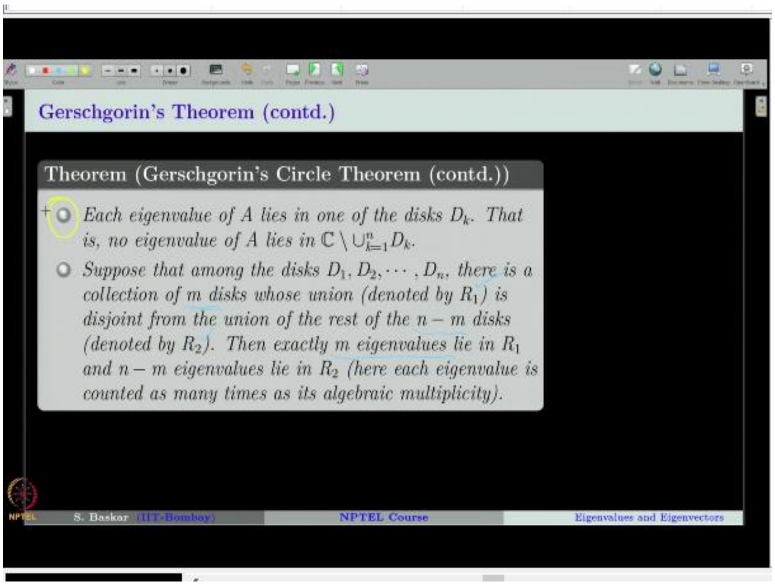
Again, let us visualize this disk D_1 is given like this, D_2 is this and D_3 is this disk, D_4 is this and D_5 is this disk. Now Gerschgorin's disk theorem also says that you have precisely 2 eigenvalues in this region. Why? because this is the union of 2 disks. Therefore 2 eigenvalues lie in this set. Similarly, 3 eigenvalues will lie in this set. That is what the Gerschgorin's disk theorem says. That is you can find 3 eigenvalues lying in this set.

(Refer Slide Time: 12:46)



Let us now precisely state the Gerschgorin's disk theorem, you are given a matrix A which is a $n \times n$ matrix. First you have to find the Gerschgorin's disk, each corresponding to a row of the matrix. Therefore, if it is a $n \times n$ matrix you have n Gerschgorin's disk, the radius is given like this and the disks are given like this.

(Refer Slide Time: 13:14)

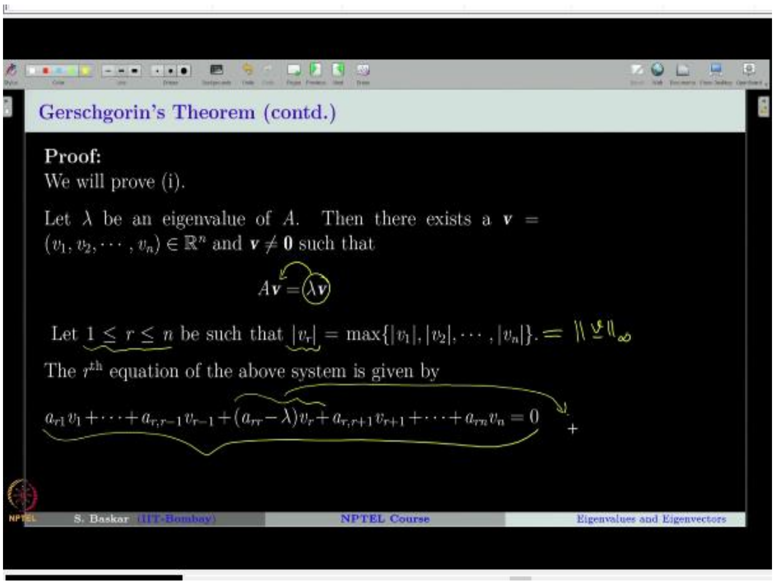


Now the conclusion of the theorem is that, each eigenvalue of the matrix A lies in one of the disk D_k, it means what, you cannot find an eigenvalue lying outside the union of this disk. That is what it means. The theorem also concludes that if you have some m number of disks whose union is disjoint from some n - m number of disk then precisely you have m eigenvalues in the first set and the n - m eigenvalues in the second set.

That is precisely what we have visualized here, you have m number of disk whose union is disjoint from n - m number of disks, here the m = 2. So, you have 2 disks which are intersecting each other and their union is disjoint from another set of disk which consists of 3 disks; here in this particular example. Therefore, you can precisely have m number of eigenvalues in this set.

Let us call this as R_1. So, R_1 will have 2 eigenvalues in this example and this set can have 3 eigenvalues. Let us call this as R_2. That is what the theorem says. Suppose among the disks there is a collection of m disk whose union is disjoint from the union of the rest of all n - m disk. We are calling the first set as R_1 and the second as R_2. Then the theorem says that you have exactly m eigenvalues lying in R_1 and n - m eigenvalues lying in R_2. Let us prove only the first part of the theorem.

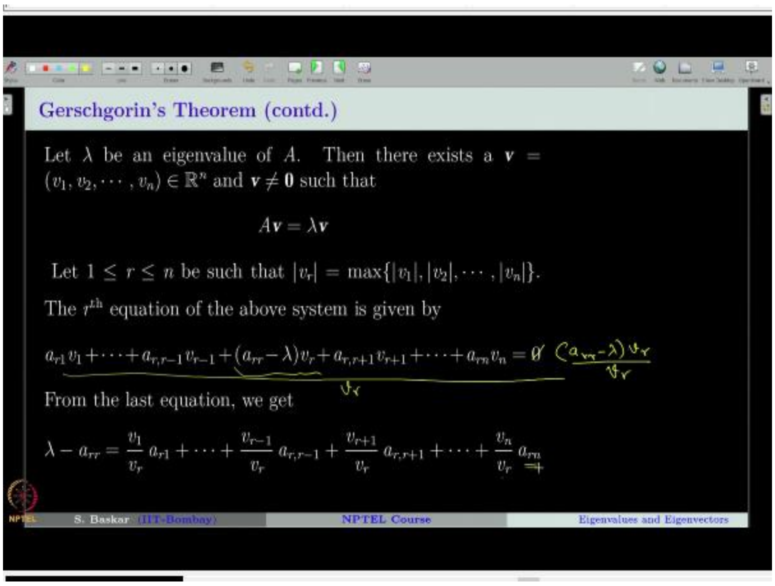
(Refer Slide Time: 15:26)



Let λ be an eigenvalue of the matrix *A* and let $\boldsymbol{v}$ be the corresponding eigenvector that we have chosen. Then you can write $A\boldsymbol{v} = \lambda\boldsymbol{v}$. That is the definition of the eigenvector. Also note that the eigenvector should be non zero. Now let us take the maximum norm of the vector $\boldsymbol{v}$ and that is given by maximum of modulus of $v_1, v_2, \cdots, v_n$. Let this maximum norm be achieved at the coordinate say *r*, where *r* is an integer between 1 to *n*.

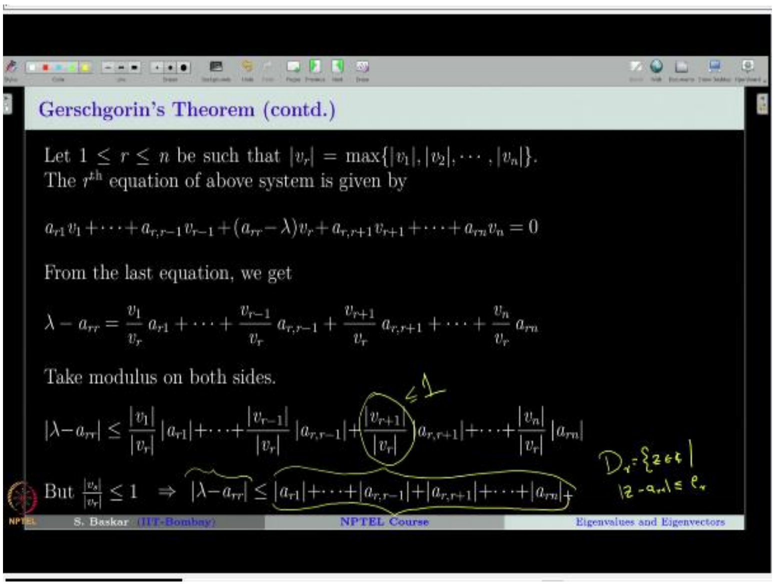
Then the *r*th equation can be written like this, why because I am just taking the *r*th coordinate of this vector to the left hand side and writing this equation. Now what we will do is we will take this term alone on the right hand side and keep all the other terms on the left hand side and then divide both sides by v_r.

(Refer Slide Time: 16:41)



That is, we are taking this to other side that gives us $\frac{(a_{rr}-\lambda)v_r}{v_r}$ and that gives us $\lambda - a_{rr} = \frac{v_1}{v_r}a_{r1} + \cdots + \frac{v_n}{v_r}a_{rn}$. Now take the modulus on both sides.

(Refer Slide Time: 17:08)

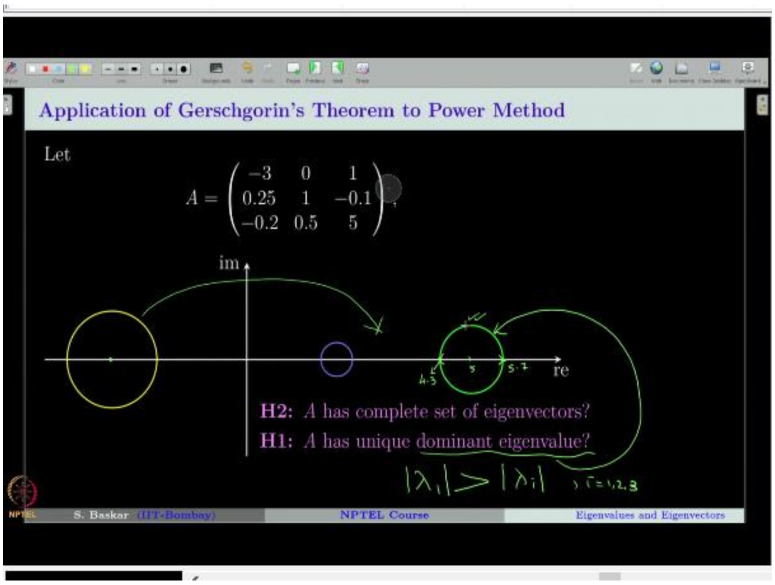


And use the triangle inequality, we get $|\lambda - a_{rr}|$ is less than or equal to this expression. Now you see $v_r \geq |v_i|, i = 1,2,\cdots,n$. That implies $\frac{|v_i|}{|v_r|}$, this less than or equal to 1. Therefore, each of this quantity is less than or equal to 1. So, you use this property and write this right hand side expression like this.

You can see that $|\lambda - a_{rr}|$ is less than or equal to this where each of this term is dominated by 1 here. Now if you recall the disk $D_r = \{z \in \mathbb{C} : |z - a_{rr}| \leq e_r\}$. The radius of the *r*th disk is precisely given by this expression. So, that is what we want to prove for the first conclusion of the theorem. The theorem says that each eigenvalue of *A* lies in one of the disks.

How we picked up that disk, is by first picking up the maximum coordinate of the vector ***v*** and then we are trying to pack λ into that disk D_r, where the maximum norm is achieved. So, this completes the first part of the proof of Gerschgorin's theorem.

(Refer Slide Time: 19:01)



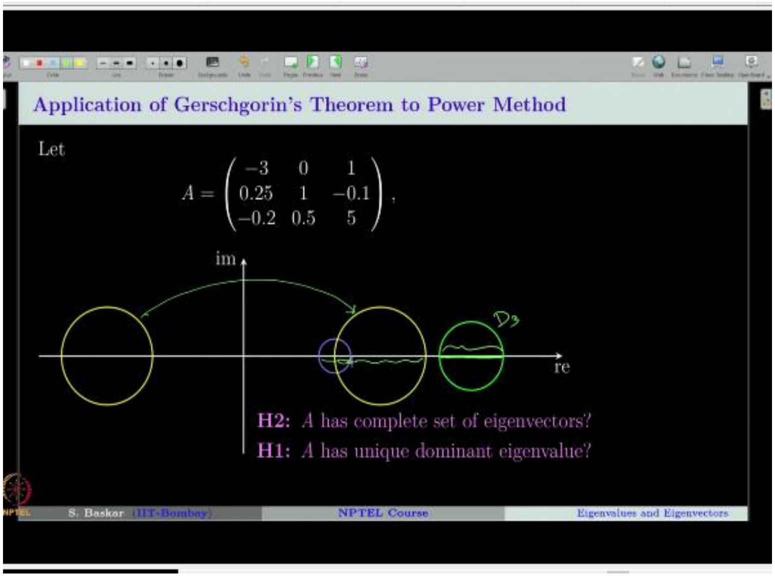
We will not prove the second part. Now let us see how to use Gerschgorin's theorem to check the hypothesis of power method. Let us again take our first example where we have taken a 3×3 matrix. If you recall we have also constructed the Gerschgorin's disks for this matrix. Now let us see whether we can get some idea about the hypothesis of the power method. Recall that the second hypothesis of the power method demands that A should have a complete set of eigenvectors.

That is all the eigenvectors should be linearly independent, but by seeing the Gerschgorin's disk you can see that one eigenvalue is sitting here, one eigenvalue is sitting here and one eigenvalue is sitting here. Therefore, this matrix has distinct eigenvalues. Therefore, all the eigenvectors are going to be linearly independent. So, Gerschgorin's theorem gave us the information that all the eigenvalues of the matrix A are in fact distinct.

Now let us check the first hypothesis, where we need to see whether a has a unique dominant eigenvalue. First observe that the dominant eigenvalue is coming from this disk because this point is 4.3 and this is 5.7. The center is 5 therefore assuming that A has real eigenvalues because we always work with real eigenvalues. In that way the eigenvalue of A which is coming from this disk is going to be something between 4.3 to 5.7.

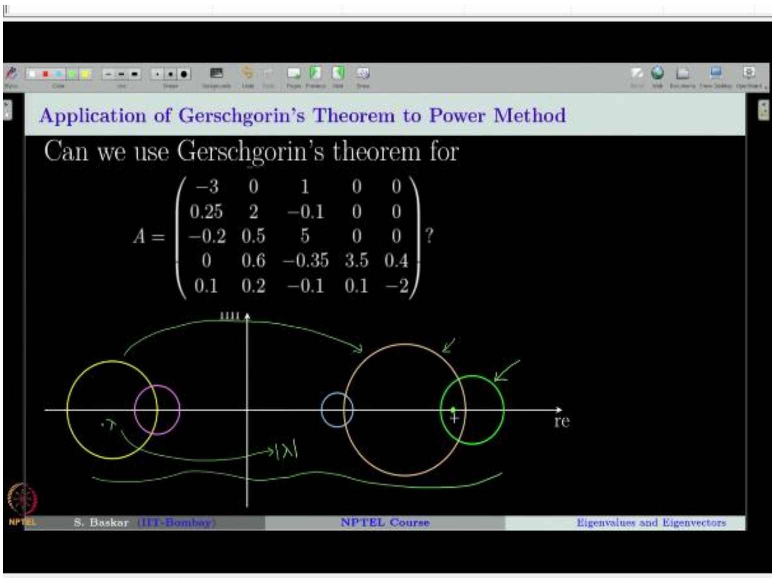
And that is surely going to be the dominant eigenvalue why because we have to check this condition $\lambda_1 > \lambda_i, i = 1, 2, \cdots$, here it is only 3. So, to check that what you have to do is you see whether the disk which is lying on the negative side of the plane when it shifted to the positive side it should not intersect the circle which has the dominant eigenvalue. That is the idea. Let us shift this disk to the positive side and see.

(Refer Slide Time: 21:48)



That this when you shift to the positive side it is not intersecting the disk D_3. That shows that whatever may be the value that is coming from this part of the interval is surely going to be different from this path as well as this part. Therefore, from here you can in fact conclude that A has a unique dominant eigenvalue.

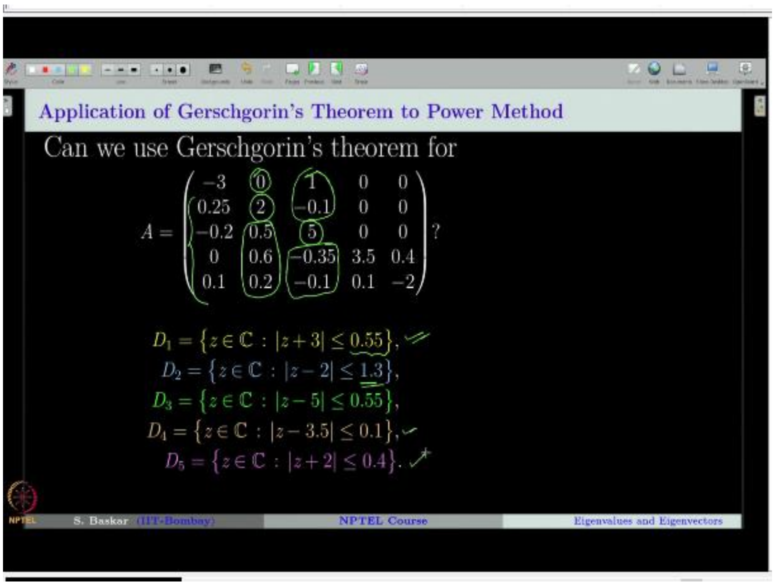
(Refer Slide Time: 22:18)



Let us take the next example where we had a 5×5 matrix and here, we have seen that the Gerschgorin's disk are given like this, you can clearly see that we cannot use Gerschgorin's disk theorem to check the hypothesis of the power method, why? because some of the disks are intersecting. For instance, the dominant eigenvalue is coming from this disk and it is intersecting this disk also.

Therefore, it may happen that there is a eigenvalue with algebraic multiplicity as 2 or even 3 and also when you shift this disk to the other side that is when you take modulus of the eigenvalue which is sitting here it may also coincide with the dominant eigenvalue. So, therefore Gerschgorin's disk theorem cannot be used to verify the hypothesis of the power method. It does not mean that A is going to violate the hypothesis of the power method, it is only that the Gerschgorin's theorem cannot be used.

(Refer Slide Time: 23:39)



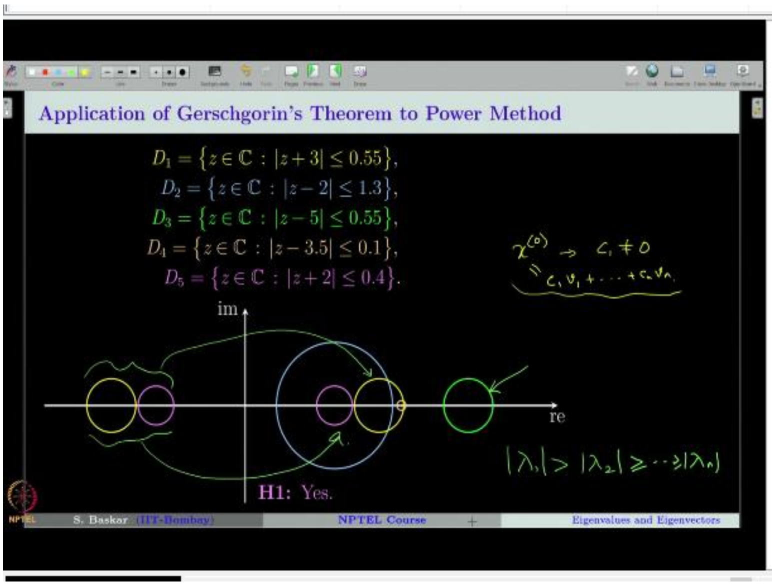
Now what is the idea? Should we give up or is there any other way to check the hypothesis of power method through Gerschgorin's theorem? There is a nice idea that why not we apply the Gerschgorin's theorem to A^T; why because the eigenvalues of A and the eigenvalues of A^T are one and the same. Therefore, we can also apply the Gerschgorin's theorem to A^T and see if that gives us a good information.

If so, then that can be borrowed to conclude how the eigenvalues are going to be located for the matrix A itself. So, in this case what we have to do instead of taking the radius along the rows you have to now take the radius along the columns. That is take the absolute values of the non diagonal elements of the columns. Let us see in that way how the Gerschgorin's disk for A^T are looking like.

You can see that the disk D_1 is given like this. Again, you can observe that the centers of these disks are not going to change; only the radii will change. For the first disk, now you have to sum along the columns. In the previous case we took the sum along the row and therefore we had the sum as 1 for the radius of the first disk. Now the first disk will have radius 0.55 only.

So, that is the advantage. In this particular example it may be that A^T may give a bad information than A sometimes, but in this case A^T seems to be giving a better information than A. Similarly, for D_2 again you have to take the diagonal element 2 as the center of D_2 and then now the radius is computed along the column elements. In that way we got 1.3 for the disk 2 and similarly disk 3 has radius 0.55 again and the center is at 5 and similarly the other 2 disks.

(Refer Slide Time: 26:17)



Let us see how they look like in the complex plane, you can see now that all the disks for A^T are disjoint. Therefore you can have the hypothesis of power method verified from the Gerschgorin's theorem applied to A^T. You can in fact shift these 2 disks to the positive side and see whether they are intersecting the disk from where you get the dominant eigenvalue. Note that the dominant eigenvalue is coming from this disk.

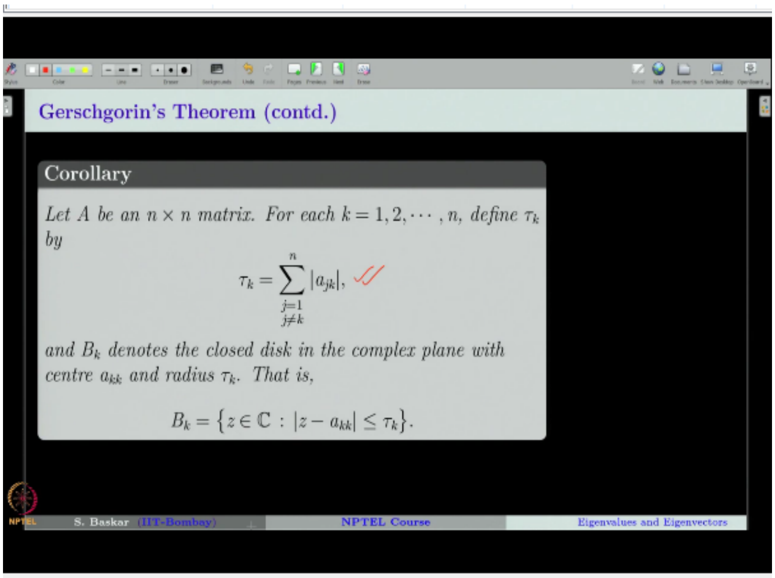
Let us see how they look like. When I shifted these 2 disk from the negative side to positive side, why we are doing because the dominant eigenvalue condition is checked with modulus of the eigenvalues, it goes like this. So, therefore we have to check whether the absolute value of the eigenvalues are going to coincide with the dominant eigenvalue. That is why we are shifting this disk to the positive side.

And now you can see that even when you shift them to positive side they do not intersect the disk from where you get the dominant eigenvalue. Therefore, the Gerschgorin's theorem applied to A^T gives us a good information and it tells us that the power method is going to converge. If at all you choose your $x^{(0)}$ correctly. Again, that is not a problem because what you need is that you have to choose your $x^{(0)}$ such that c_1 should not be equal to 0.

If you recall c_1 should not be equal to 0. What is c_1? c_1 is nothing but the coefficient of the first term in the representation $c_1 v_1 + \cdots + c_n v_n$. This is not something very serious because we may choose 2 or 3 vectors and run the power method. If all these are going to converge to the same eigenvalue, then it is very likely that that is going to be dominant eigenvalue.

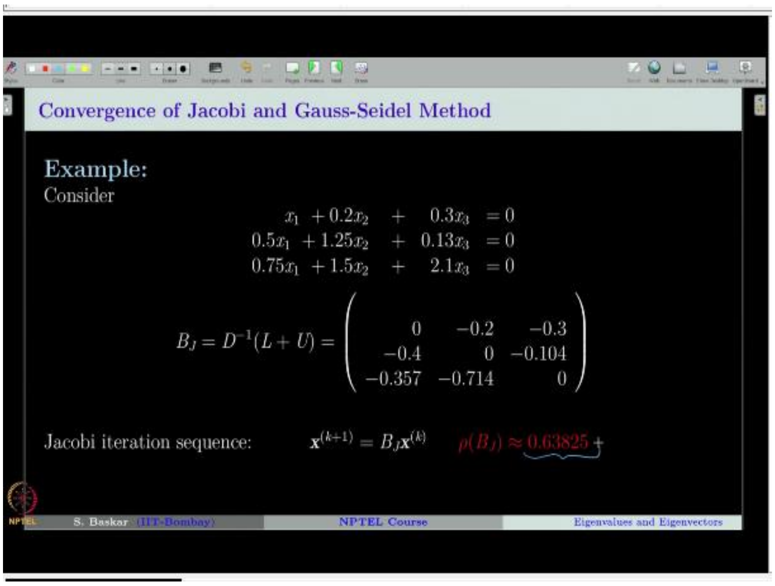
Because it is very unlikely that you will choose $x^{(0)}$ 3 times for which all the 3 vectors will have zero c_1 So, that may be very unlikely. In that way this is practically not very serious hypothesis, whereas the other 2 are really serious concern to check and Gerschgorin's Theorem gives us a possibility that we may verify those hypotheses without having explicitly the idea about how the eigenvalues are.

(Refer Slide Time: 29:16)



Let us also state the Gerschgorin's theorem on A^T. So, there is nothing new here. You just have to apply the radius column wise now that is the only difference in this theorem when compared to the theorem that we have stated previously. Otherwise, the conclusion reminds the same.

(Refer Slide Time: 29:43)



Before ending this lecture let us have another interesting application of the Gerschgorin's disk theorem. If you recall in the last class, we have studied the convergence theorem for iterative methods in particular for Jacobi and Gauss-Seidel method we had given a necessary and sufficient condition for the convergence what is that? That is using the spectral radius of the iterative matrix *B*.

That is the spectral radius of the iterative matrix should be less than 1. This is the necessary and the sufficient condition for the convergence of an iterative method. What is spectral radius? If you recall the definition of spectral radius is nothing but the maximum over all the eigenvalues λ_i. In way you can see that the spectral radius is nothing but the dominant eigenvalue of the matrix *B*.

Therefore, you can in fact compute this spectral radius using power method. If your purpose is to only see whether the Jacobian method or the Gauss-Seidel method is going to converge or any iterative method is going to converge or not then even you can go for checking the Gerschgorin's theorem. How you will do? For instance, let us take the Jacobi method. The iterative matrix for the Jacobi method is given like this.

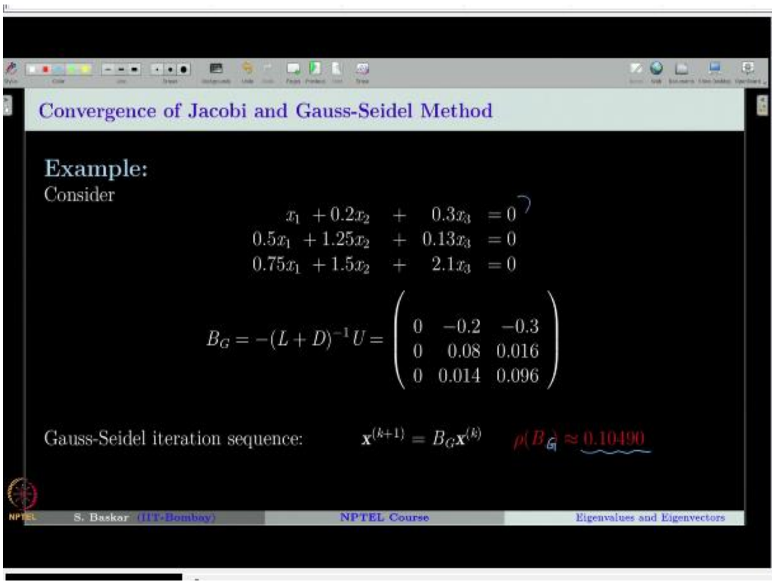
You can see that all the Gerschgorin's disks of this matrix are having their centers as 0. In fact Jacobi method for any system will have the diagonal elements as 0, you can verify that. In this case what you have to do is, you either go row wise and get the disk, also you go by column wise and see that this whether all the disks are having their radii as something less than 1. If that is something less than 1 it means what all the eigenvalues are going to lie in these disks.

And if you know that the largest disk has its radius something less than 1 it means the spectral radius which is the dominant eigenvalue in the worst case be sitting somewhere here and that will be strictly less than 1. So, that is the idea. Let us see how the Gerschgorin's disks is placed for the matrix B_J, you can see that the last disk has radius more than 1. Therefore, you cannot conclude the convergence of Jacobi method using the Gerschgorin's disk applied on B_J.

Let us see whether we can conclude the convergence by applying Gerschgorin's disk on B_J^T? You can see that when you take the disks with B_J^T the radius of the first disk is the sum of these 2 numbers which amounts to be around 0.757 and the second is as the radius something very near to 0.9 but still it is less than 1 and the third disk is pretty small, its radius is only around 0.4.

Therefore, the Gerschgorin's theorem when applied to B_J^T tells us that the Jacobi method is surely going to converge. In fact you can use the power method to compute the spectral radius also. In this case the spectral radius of B_J is 0.63825 and so on. The convergence may be little slow but still the Jacobi method converges. Without going to the explicit calculation of the Jacobi iteration we can understand either by using power method or more easily using the Gerschgorin's disks theorem we can conclude this.

(Refer Slide Time: 34:17)



Let us go to see whether the same exercise can be done for Gauss-Seidel method. In the Gauss-Seidel method the Gerschgorin's disks for the iteration matrix B_G is given like this and you can see whether you take row wise and apply the Gerschgorin's theorem or you take the this column wise you can see that all the disks have radius something very less than 1. In fact, you can see that the spectral radius of the iterative matrix of the Gauss-Seidel method is very, very small.

That in fact shows that the Gauss-Seidel method in fact is going to converge very fast to the solution of this system. With this note, let us end this lecture. Thank you for your attention.