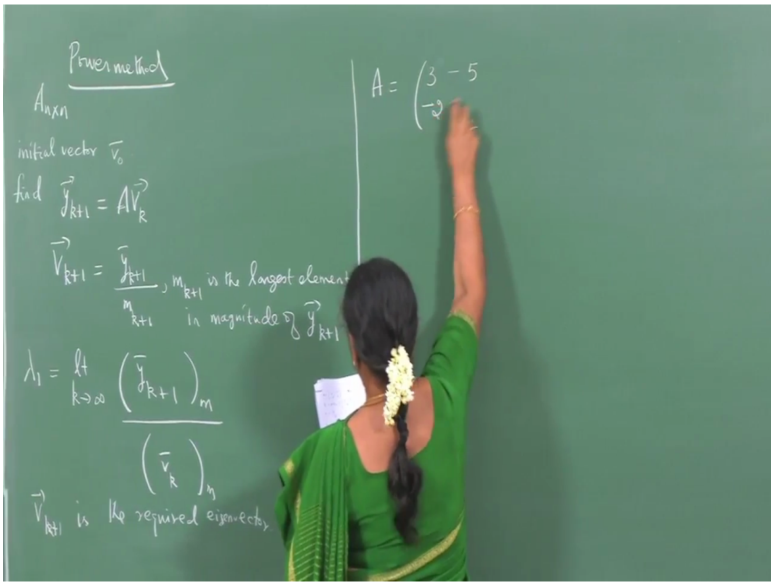
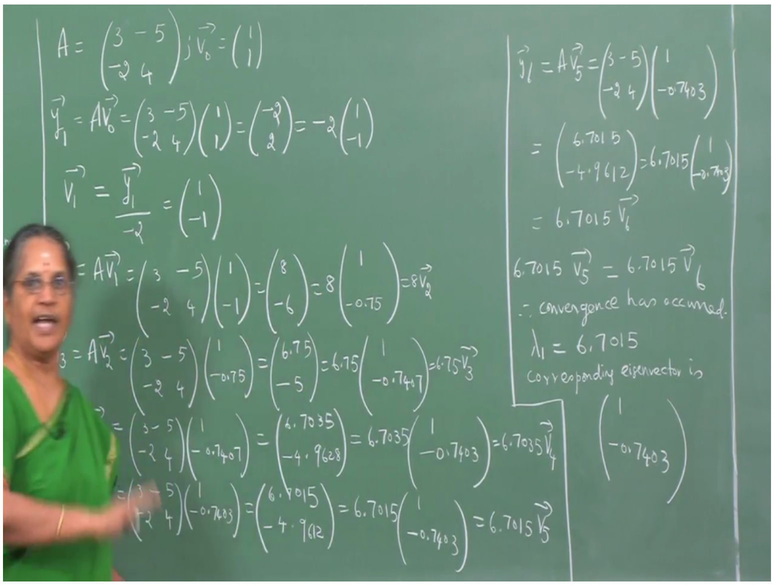
Numerical Analysis.
Professor R. Usha.
Department of Mathematics.
Indian Institute of Technology, Madras.
Lecture-47.
Matrix Eigenvalue problems-2.
Power method-2.
Gerschgorin's Theorem, Brauer's Theorem.

(Refer Slide Time: 0:38)



Good morning everyone, in the last class we discussed power method for computing numerically the largest eigenvalue in magnitude of a given matrix A. So let us illustrate this method by taking some examples. So we said that, if we are given a square matrix, say N cross N matrix, we start with an initial vector which we call as V0 and we find the vectors say YK +1 which are AVK. And the next vector VK +1 will be YK +1 by MK +1 where MK +1 is the largest element in magnitude in YK +1 vector. Then lambda 1 will be limit as T tending to infinity of the mth components of vector YK +1 by the mth component of the vector VK.

(Refer Slide Time: 2:27)



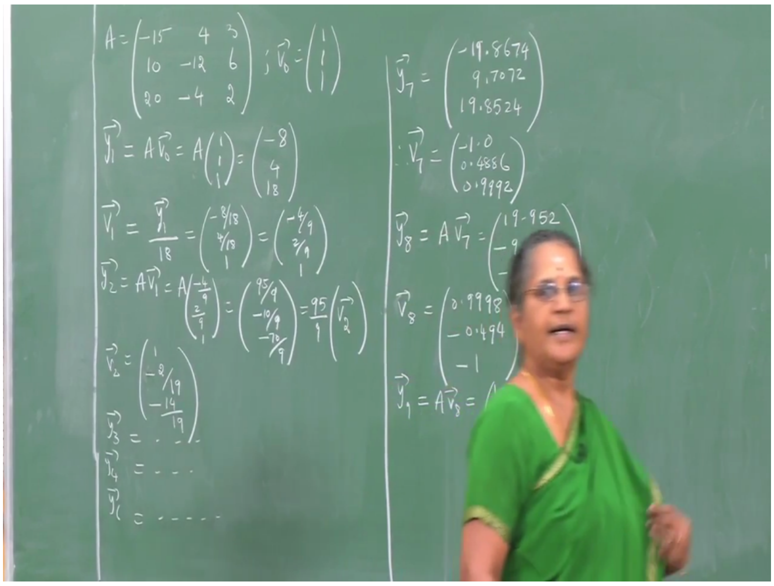
And the vector VK +1 that we have obtained will give us the required eigenvector. So let us now illustrate this procedure which is power method for computing numerically the largest eigenvalue of a given matrix A. So let us consider the given matrix A to be a 2 cross 2 matrix having entries 3, -5 and -2, 4. We start with an initial vector V0, say 1, 1. So we should find 1st vector Y1, which is A times V0, so it is 3, -5, -2, 4 into 1, 1 and that is the column vector having components -2, 2. Now I remove the largest element in magnitude and write down the vector as -2 times 1, -1. So now I should define V1, what is V1, that is Y1 - the largest value of the element that I have factored out.

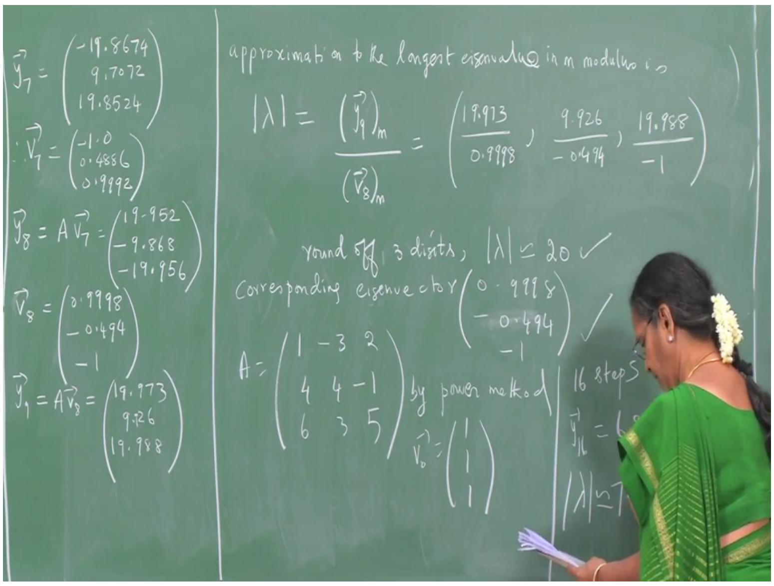
So that will give you 1, -1, so with this V1 I compute Y2 vector which is A V1. 3, -5, -2, 4 into V1 which is 1, -1 and that is 8, -6, I remove the element which has largest magnitude and write down the vector as 1, -0.75. So it is 8 times vector V2. So we compute Y3 which is A V2, 3, -5, -2, 4 into V2 which is 1, -0.75. That turns out to be 6.75, -5, so I remove 6.75 and the vector that I get is 1, -0.7407, so it is 6.75 times vector V3. We compute Y4 which is A V3, so 3, -5, -2, 4 into V3 which is 1, -0.7407 and that turns out to be 6.7035, -4.9628, so we remove 6.7035 and the vector is 1, -0.7403. So we have Y4 to be 6.7035 times vector V4. We continue this process and compute what is Y5 and that is A V4.

So 3, -5, -2, 4 into vector V 4 is 1, -0.7403 and that turns out to be 6.7015 - 4.9612, so I remove 6.7015 and we get 1, -0.7403. So we have 6.7015 times vector V5 and we see that every time we perform the computations, the vector is such that the components in the vector are coming to be closer to each other. So let us perform another iteration and see. So compute Y6, so vector Y6 will be A into vector V5, so 3, -5, -2, 4 into V5 which is 1, -0.7403 and that turns out to be 6.7015, -4.9612, so it is 6.7015 into 1, -0.7403. And so this will be 6.7015 into vector V6.

So at this stage we observe that 6.7015 into vector V5 is the same as 6.7015 into vector V6 and therefore convergence has occurred and therefore we can stop our computations and write down what the eigenvalue is and the corresponding eigenvectors. So we observe that the eigenvalue lambda 1 which is numerically the largest eigenvalue of this matrix is 6.7015. And what is the eigenvector, the corresponding eigenvector is given by 1, -0.7403. So power method helps us to compute numerically the largest eigenvalue or the most dominant eigenvalue of a given matrix A numerically. And the procedure is illustrated by means of this example.

(Refer Slide Time: 9:39)





So let us now take a 3 by 3 matrix A and then work out the details, so the problem is find the largest eigenvalue in modulus and the corresponding eigenvector of the matrix A, which has elements -15, 4, 3, 10, -12, 6, 20, -4, 2 and we are asked to obtain the dominant eigenvalue of this matrix by power method numerically. So we start with the initial vector V bar having components 1, 1, 1, I compute the vector Y1 1st. What is it, it is equal to A into V dot, so A into vector 1, 1, 1 where A is this matrix. The entries in the product of these 2 matrices are given by -8, 4, 18. So I have to define now my vector V1 which is vector Y1 by the largest element in the vector Y1.

So that will give me -8 by 18, 4 by 18, 1 and that is -4 by 9, 2 by 9, 1. Now that I know the vector V1, I compute Y2, which is AV1. So I need to -4 by 9, 2 by 9, 1 and that turns out to be the vector 95 by 9, -10 by 9, -70 by 9. And I divide by the largest element, namely 95 by 9, so if I remove that factor, I get a new vector which is V2. So V2 turns out to be the vector 1, -2 by 19, -14 by 19. So we compute Y3, Y4 and so on, so I leave these computations for you. And when it comes to obtaining the vector Y7, it turns out to be the vector -19.8674, 9.7072 and 19.8524 and therefore I compute vector V7 by dividing vector Y7 by the numerically largest element in that vector.

That gives me V7 to be -1.0, 0.4886, 0.9992. So we compute Y8 which is A into V7 and that gives 19.952, -9.868, -19.956. And so I compute vector V8 which is Y8 vector by numerically the largest element in Y8. And that turns out to be 0.9998, -0.494 and -1. And so we work out Y9 which is A times vector V8 and this turns out to be 19.973, 9.926 and 19.988. Suppose say the problem also says perform 9 steps using power method and write down the numerically largest eigenvalue of matrix A.

So we have performed 9 steps and at this stage we want to stop our computations because we are asked to do only 9 steps of computations of power method. And so we determine the eigenvalue and the Eigen vector. So what is going to be an approximation to the largest eigenvalue in modulus? So it is modulus of lambda and that will be the Mth component of vector Y9 by the Mth component of vector V8. So I have to compute the ratio of the 1st component of Y9 and 1st component of V8. Then take the 2nd component of Y9 by the 2nd component of V8 and then get the ratio, the 3rd component of Y9 and the 3rd component of V8 and obtain what the value is.

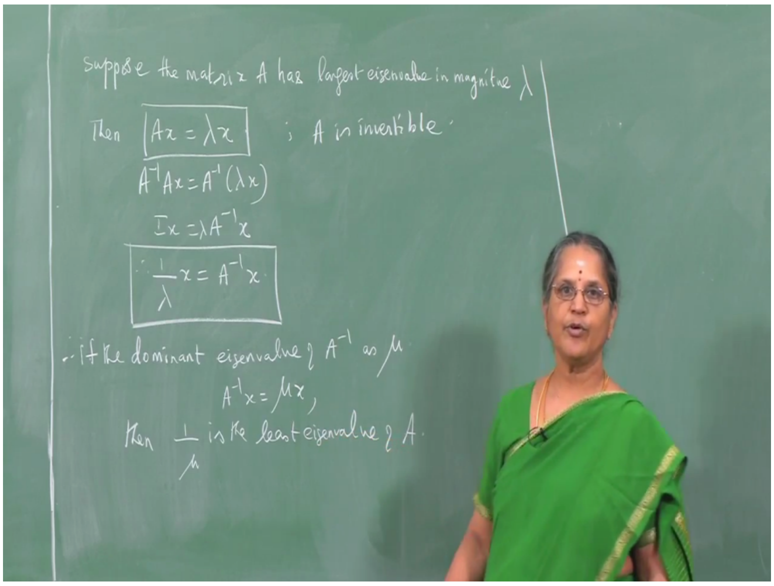
So we get, for the 1st component of Y9 is 19.973 divided by 1st component of V8 is 0.9998, the 2nd component of Y9 is 9.926 and that of V8 is -0.494, the 3rd component is 19.988

divided by -1. I observe that if I compute these ratios and take the absolute value, then round it off to 3 digits, then I end up with mod lambda to be 20. So if I compute this ratio, it is going to be very close to 19.966 etc, similarly this one and here it is going to be -19.988, I observe that that is the largest, so in absolute value correct to 3 digits and I round it off to 3 digits, mod lambda is close to 20 and therefore that is the largest eigenvalue in magnitude of this matrix A.

So we have to get the corresponding eigenvector, so the corresponding eigenvector is going to be the vector having the vector having components namely V8, that will give you 0.9998, -0.494 and -1. So the numerically largest eigenvalue and the corresponding eigenvector of the given matrix A have been computed using power method. I will give you some problems, you can try to work about that home and I will include some more problems in the assignment sheet.

So find numerically the largest eigenvalue of the matrix 1, -3, 2, 4, 4, -1, 6, 3, 5 by power method. You start with the vector V0 having components 1, 1, 1, I will give you the final answer, you can try to see, compute 16 steps of computations, I am sure that at the 16th step you will get Y 16 to be 6.9998 into 0.3000, 0.06661 and therefore eigenvalue, the numerically largest eigenvalue is close to 7. So using power method we have been able to compute the most dominant eigenvalue. The question now is, is it possible to obtain the smallest eigenvalue of this matrix A using power method.

(Refer Slide Time: 19:22)

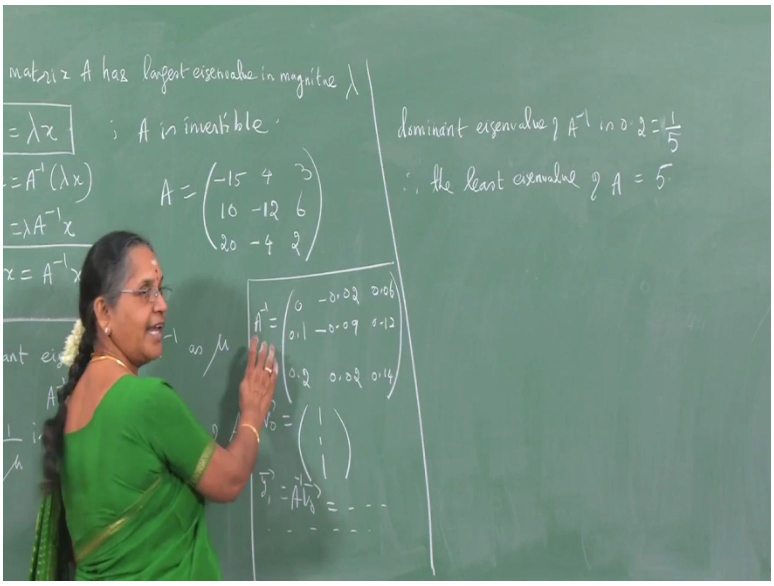


The answer is yes and let us see how we can obtain the smallest eigenvalue. So suppose that we assume that the matrix A has largest eigenvalue say lambda. Then what does that mean, it means AX is equal to lambda X and let us assume that A is invertible. Then when I pre multiply this by A inverse, then I get IX to be equal to A inverse into X multiplied by lambda. Therefore 1 by lambda X will be equal to A inverse X. So when we say that AX is equal to lambda X, we say that lambda is an eigenvalue of the matrix A and the corresponding eigenvector is X. Now this statement tells us that 1 lambda is an eigenvalue of A inverse and the corresponding eigenvector is the same vector X for the eigenvalue lambda of the matrix A.

So this tells you that 1 by lambda is an eigenvalue of matrix A inverse with the same eigenvector X. If I call the dominant eigenvalue of A inverse as mu, so that A inverse X is equal to mu X, then it is clear that 1 by mu is the least eigenvalue of the matrix A. Why, what have we shown? If lambda is the dominant eigenvalue of matrix A, then A inverse X is 1 by lambda into X, so that if mu is the dominant eigenvalue of A inverse, such that A inverse X is equal to mu, then 1 by mu will be the least eigenvalue of the matrix A. So if I ask you to compute the smallest eigenvalue of a given matrix A, what is it that you should do?

Given the matrix A, compute its inverse, say by Gauss Jordan method and then compute its dominant eigenvalue using power method, take the reciprocal, that gives you the least eigenvalue of the given matrix A. And therefore it is possible for you to compute the smallest eigenvalue in magnitude for the given matrix A. So let us work out an example illustrating this.

(Refer Slide Time: 23:19)

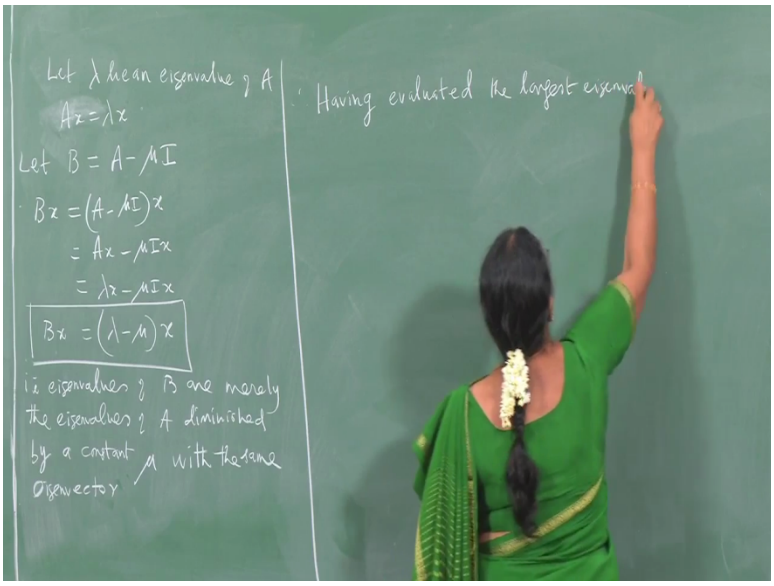


So if I consider matrix A to be given by -15, 4, 3, 10, -12, 6, 20, -4, 2 and I am asked to compute the smallest eigenvalue of this matrix A. So I immediately compute the inverse of the matrix A and it turns out to be 0, -0.02, 0.06, 0.1, -0.09, 0.12, 0.2, 0.02, 0.14. So I have to now compute the most dominant eigenvalue of this matrix A inverse. So use power method to compute the most dominant eigenvalue. Start with the vector say V0 which is 1, 1, 1 and compute Y1 which is A inverse into V0 because you are computing the dominant eigenvalue of the matrix A inverse.

So your Y1 is A inverse into V0. So continue your computations and show that the dominant eigenvalue of A inverse turns out to be say 0.2, that is 1 by 5. Show that the dominant eigenvalue of A inverse is 1 by 5 and therefore the least eigenvalue or the smallest eigenvalue in magnitude for the given matrix A is going to be 5. So given a matrix A, you know how to get the largest eigenvalue by power method and how to get the least eigenvalue by computing the dominant eigenvalue of A inverse and taking its reciprocal and that will give you the smallest eigenvalue of A.

So you now know using power method how to compute both the largest eigenvalue and the least eigenvalue. There is also another of computing the smallest eigenvalue of a given matrix A even without computing what A inverse is. So let us see how this can be done.

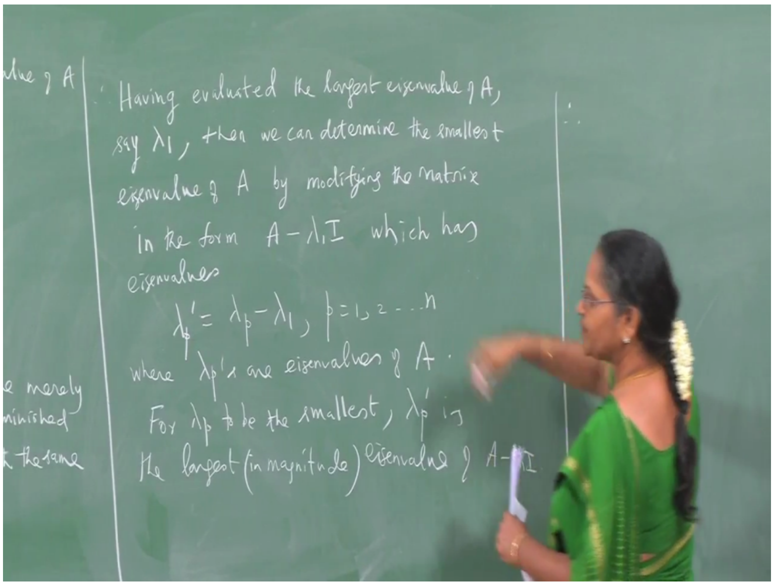
(Refer Slide Time: 26:08)



So let lambda be an eigenvalue of A, so A is a given matrix, suppose lambda is an eigenvalue of A, that AX is equal to lambda X. Let us find a matrix B where B is given by A - mu times I, mu is some constant. Then let us compute what is BX, so BX is A - mu I into X. So it is AX - mu IX, but what is AX, AX is lambda X. So this tells you it is lambda - mu into X. So what have we shown, matrix B which is A - mu I is such that it satisfies this equation namely BX equal to lambda - mu times X. What are lambdas, lambdas are eigenvalues of given matrix A.

And therefore A - mu I has eigenvalues lambda - mu. And therefore the eigenvalues of B are nearly the eigenvalues of A diminished by a constant mu. And what about the eigenvector, the eigenvector is the same, namely if lambda is an eigenvalue of A and the corresponding eigenvector is X, then lambda - mu is an eigenvector of B having the same eigenvector X. So that is what we want to write, the eigenvalues of B are nearly eigenvalues of A diminished by a constant mu with the same eigenvector.

(Refer Slide Time: 29:06)

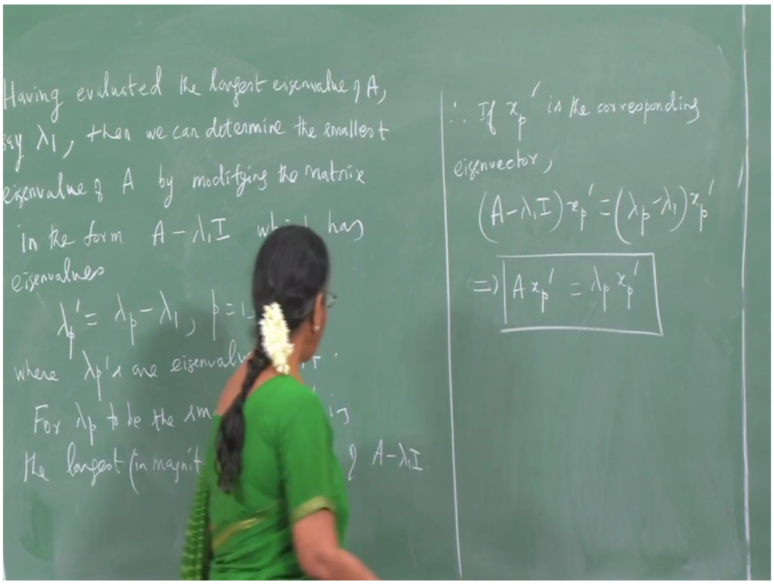


So therefore if you evaluate the largest eigenvalue of A and call it as lambda 1, then we can determine the smallest eigenvalue of A by modifying the matrix. We are given a matrix A, we would now like to consider another matrix which is different from A, so we want to do some modification in the matrix A, what is it? Take it in the form A - lambda 1 times I, take the matrix A and subtract the matrix which is lambda 1 times the identity matrix. By what we have discussed, what can you say about the Eigen values of this matrix, this matrix has eigenvalues given by say lambda P dashed which is equal to lambda P - lambda 1 for P is equal to 1, 2, 3 up to N where what are lambda Ps? Lambda Ps are eigenvalues of A.

Now I want lambda P to be the smallest. So for lambda P to be the smallest, lambda P dashed must be the largest in magnitude eigenvalue of which matrix, A - lambda 1 into I. What are lambda Ps, they are the eigenvalues of the given matrix A. So lambda P - lambda 1 for P taking values 1, 2, 3 up to N and if you call that as lambda P dashed, then the lambda P dash are such that they are the eigenvalues of A - lambda 1 into I. So in order that you are lambda P which is an eigenvalue of A to be the smallest, you must have lambda P dashed to be the largest eigenvalue in magnitude of A - lambda 1I.

Can you compute this, yes, we know we can use power method to obtain the largest or the dominant eigenvalue in magnitude of a given matrix. What is the matrix that is given to you now, the given matrix is say B which is A - lambda 1 into I. Do you know lambda 1, yes, it is the dominant eigenvalue of the given matrix A. So compute a new matrix, call it is B which is A - lambda 1 into I and compute its dominant eigenvalue, so you are using again power method to compute its dominant eigenvalue. Then once you know it is the dominant eigenvalue of A - lambda 1 into I, then your that eigenvalue + lambda 1 will give you lambda P which is the least eigenvalue.

(Refer Slide Time: 33:05)



So let us discuss about the corresponding eigenvector. So if suppose I call XP dashed as the corresponding eigenvector, then A - lambda 1 into I into XP dashed must be lambda P - lambda 1 into XP dashed because lambda P - lambda 1 are eigenvalues of A - lambda 1 into I and the corresponding eigenvector is XP dashed. So what does this give, this gives you A into XP dashed is equal to lambda P into XP dashed. While the next, the next term will be - lambda 1I XP dashed that will cancel with - lambda 1 XP dashed. So we get A XP dashed if lambda P XP dashed.

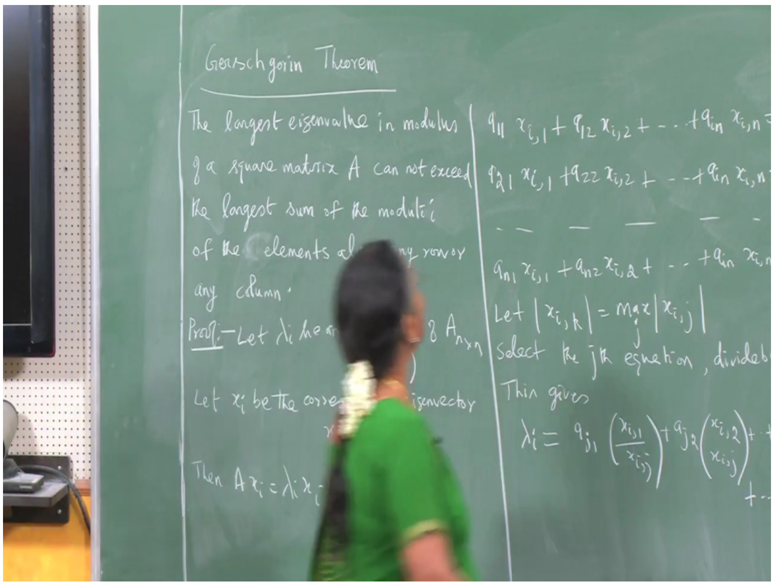
What does that give, this gives you that the Eigen vector corresponding to the eigenvalue lambda P which is the smallest eigenvalue of A is the same as the corresponding eigenvector associated with the eigenvalue lambda P - lambda 1 of the matrix A - lambda 1 into I. So you not only get the smallest eigenvalue in magnitude of the given matrix A but you also get the corresponding eigenvector, what is it, it is a XP dashed, what is this eigenvector, this is the eigenvector which is associated with the eigenvalue lambda P - lambda 1 of the matrix A - lambda 1 into I.

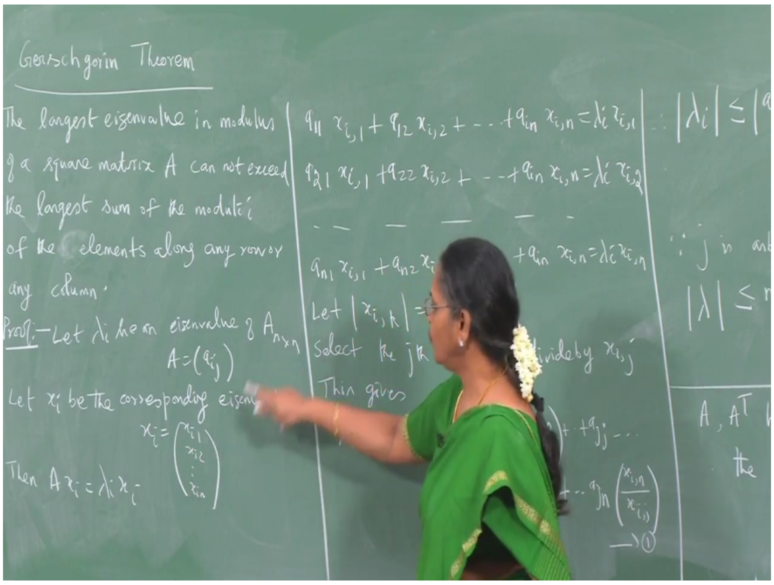
So you are able to get the smallest eigenvalue as well as the largest eigenvalue by using the power method. So let us just recall the steps for computing the smallest eigenvalue of a given matrix A. So you start with the matrix A, get its largest eigenvalue, call that as lambda 1, so it is the largest eigenvalue in magnitude for the matrix A, call that as lambda 1. Now form a matrix B which is A - lambda 1 I, for this matrix, you compute the largest eigenvalue, then the smallest eigenvalue of the given matrix A will be lambda P such that if lambda P dashed is the dominant eigenvalue of A - lambda 1I, then the 2 are related by lambda P dashed is lambda P - lambda 1.

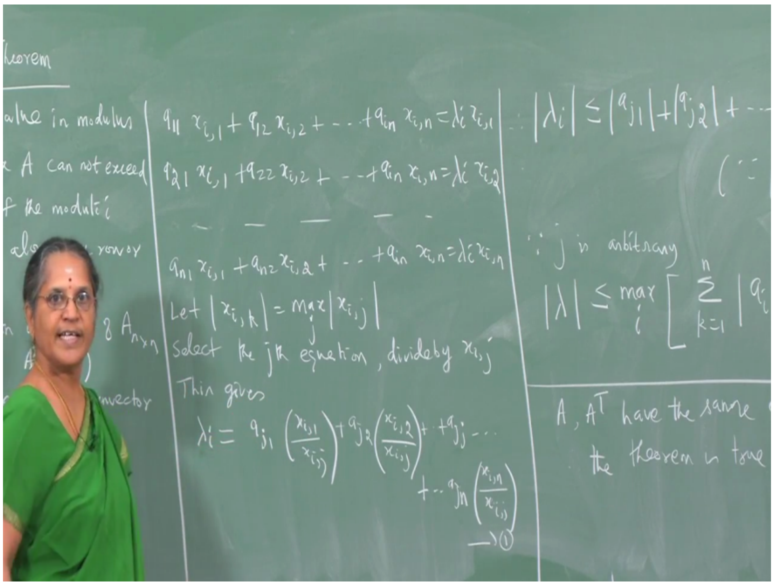
So you immediately know what is the smallest eigenvalue of the given matrix, what is the associated eigenvector. You compute the corresponding eigenvector associated with lambda P dashed for the matrix A - lambda 1 I, that Eigen vector is the Eigen vector associated with the smallest eigenvalue lambda P of the matrix A. So that completes our discussion on how we can get the smallest eigenvalue. So the method does not involve computation of inverse of the matrix A, so it is much easier to work out because you again apply power method to a new matrix and therefore the computations are very simple.

Can we give locations of the eigenvalues for a given matrix A? The answer is yes, there are results which clearly specify the regions within which circular disks within which the eigenvalues are located. So we look into these results and therefore we will be able to know where the eigenvalues of a given matrix are located. The following result gives us a way to find the location of eigenvalues of a given matrix A. What does the result say, it says that the largest eigenvalue in modulus of a square matrix A cannot exceed the largest sum of the moduli of the elements along any row or any column.

(Refer Slide Time: 38:23)



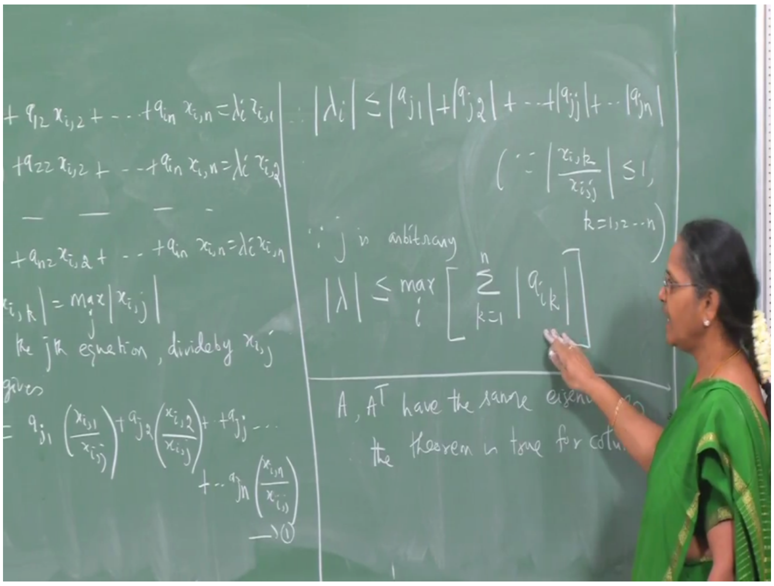




Suppose say lambda I is an eigenvalue of an N cross N matrix A, let us take A has entries denoted by AIJ where I and J run from 1 to N. Suppose X I the corresponding eigenvector and having components X I1, X I2, etc. X IN, then we know that A X I is lambda I XI. So if I write out these N equation is, then they are of this form. A1 1 into X1 + A1 2 XI2 + etc. + AIN into XIN will be equal to, on the right-hand side I have lambda 1, lambda I into XI 1. Similarly I take the 2^{nd} row of A and multiply this column vector and equate that to the 2^{nd} entry in lambda I XI. So I continue and write down the Nth equation. Now I take modulus of XIK to be the maximum of modulus of X I J over J.

So what does , once I get that, I take the Jth equation and divide the Jth equation by X IJ. So what will I get, I will get lambda I, so if I take the 2^{nd} equation and divide by X I2, then on the right-hand side I have just lambda I. So here I select the Jth equation and divide the Jth equation by X IJ. So that will give you lambda I is equal to A J1, that will be the 1st term here, A J1 X I1, that divided by X IJ. Then the next term, A J2, X I2 divided by X IJ and so on. Then I will have a term A JJ, X IJ, I divided by X IJ, so that simply gives me A JJ + etc. + A JN into X I N by X IJ.

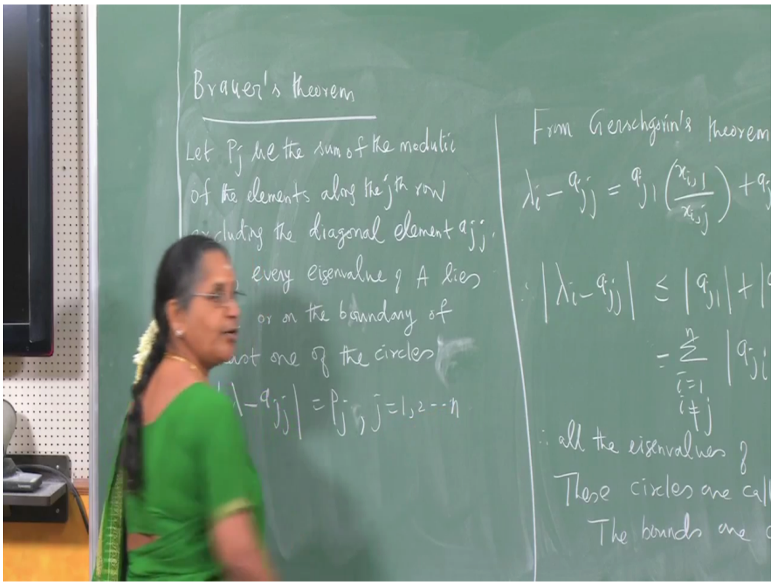
(Refer Slide Time: 41:11)



So if I take the absolute value on both sides, then mod lambda I will be less than or equal to mod A J1 + mod A J 2+ etc. + mod A JJ + etc. + mod AJN. Why modulus of X IK by X IJ is less than or equal to1, why X IK in absolute value is the maximum of modulus of X IJ? And therefore modulus of X IK by X IJ in absolute value is less than or equal to1 and therefore mod lambda I is less than or equal to mod A J1 + mod A J2, etc. + mod A JN. What are these, these are the absolute values of the entries that appear in the Jth equation. And so J is arbitrary and therefore mod lambda is less than or equal to maximum over I Sigma over K equal to1 to N modulus of A IK.

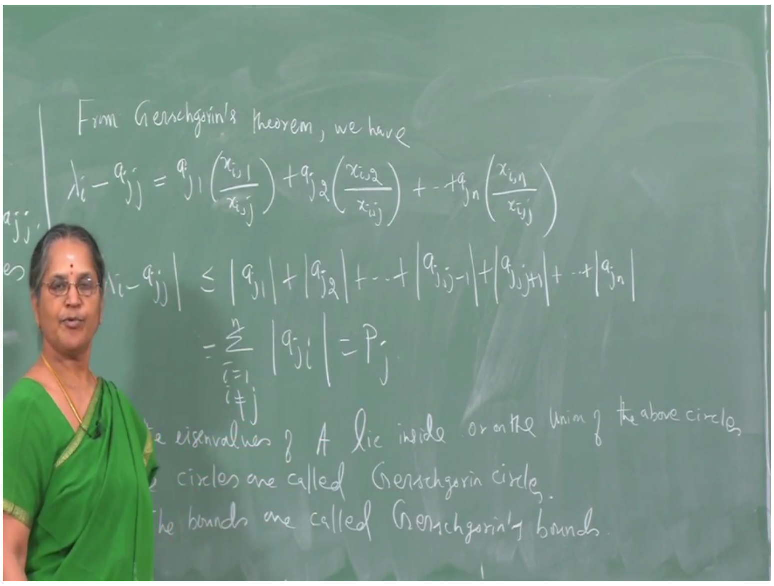
And that completes the result that the largest eigenvalue in modulus of a square matrix cannot exceed the largest some of the moduli of the elements along any row or any column. What about the proof for any column? The matrix A and A transpose have the same eigenvalues and therefore the proof of the theorem holds good also for columns and hence Gerschgorin's theorem is completed.

(Refer Slide Time: 41:54)



So we now have another result which gives the location of the eigenvalues of the given matrix. It is given by Brauer's theorem and it states that if PJ is that some of the moduli of the elements along the Jth row, excluding the diagonal elements that are A JJ, then it says every eigenvalue of A lies inside or on the boundary of at least one of these circles, which are described as mod lambda - A JJ equal to PJ for J is equal to 1, 2, 3 up to N. So these are circles which have A JJ as Centre and radius as PJ. So there are N such circles and the result says that every eigenvalue of the given matrix A live inside or on the boundary of at least one of these circles, that is what Brauer's theorem states.

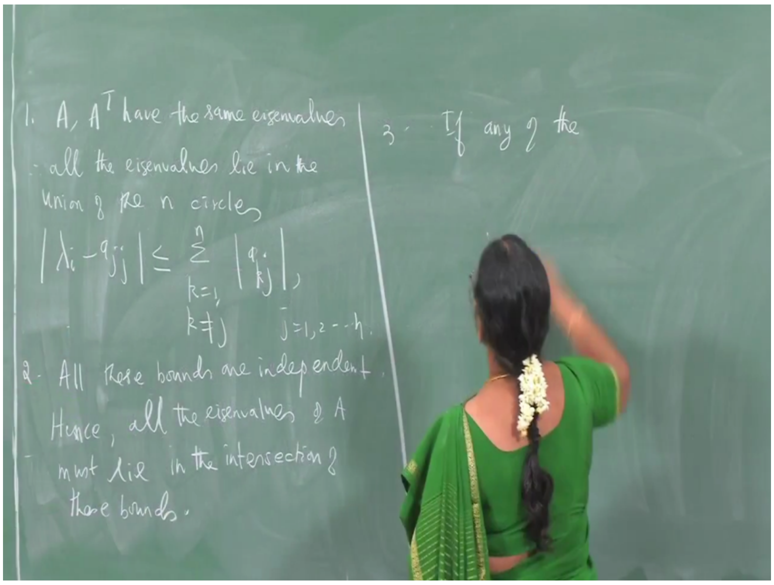
(Refer Slide Time: 43:17)

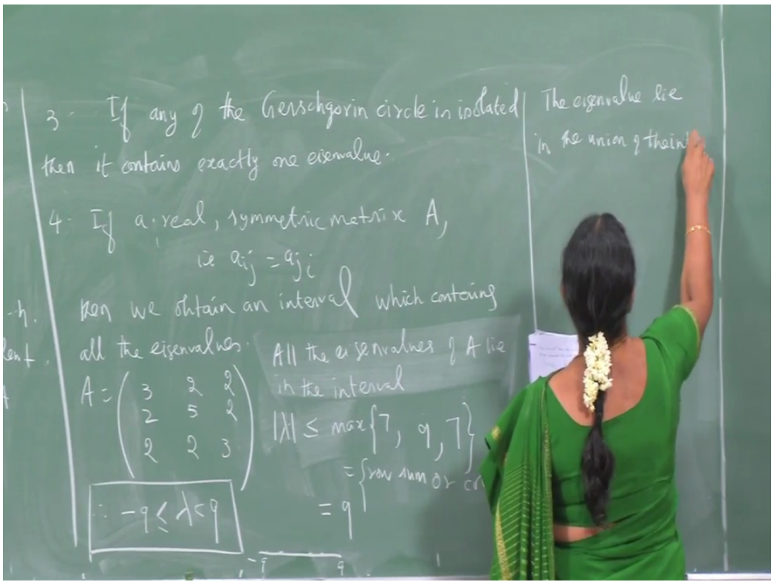


The proof is very simple, so let us work out the details. We already have seen in Gerschgorin's theorem this step. Namely lambda I was written as all these terms + A JJ after dividing by X IJ. So I can bring A JJ here and write lambda I - a JJ is the sum of the rest of the terms. And therefore in absolute value, mod lambda I - A JJ will be less than or equal to mod A J1 + mod A J2 + etc. + mod AJJ -1 + modulus of AJJ +1 + etc. + mod A JN. Why, the reason is modulus of X I K by X I J is less than or equal to1 for K running from 1, 2, 3 up to N. And what is this, this is summation I is equal to 1 to N, you see that that is no term mod A JJ, so I0 equal to J of modulus of A JI but that is what is denoted by PJ in the theorem.

So this is equal to PJ, so what are your circles, your circles are such that they have their centres at A JJ for J is equal to1, 2, 3 up to N and the radii are given by PJ for J is equal to 1, 2, 3 up to N. So you now know the circles within which the eigenvalues can be located and the result says every eigenvalue of A lies inside or on the boundary of at least one of these circles which have their Centre at A JJ and radii at PJ. So let us now use Gerschgorin's theorem and Brauer's theorem to find out the location of eigenvalues of a given matrix A.

(Refer Slide Time: 46:08)



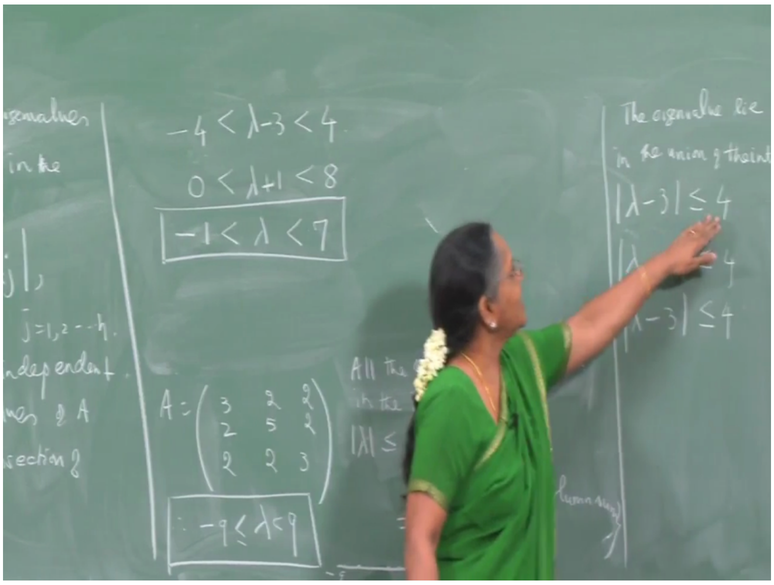


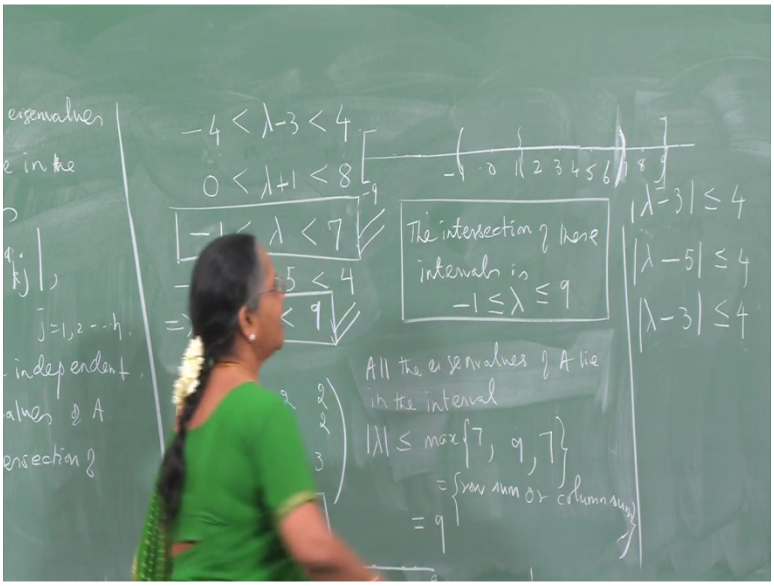
The Eigen, the circles that we have obtained are called Gerschgorin circles and the bound that we have obtained is called the Gerschgorin’s bound. So let us try to determine Gerschgorin’s bounds so that we know where all the eigenvalues of a given matrix are located. So we have some important observations to make before we proceed with demonstrating these results. So let us consider these observations. Now we know that A and A transpose have the same eigenvalues and therefore all the eigenvalues lie in the union of the n circles, namely modulus of lambda I - A JJ is less than or equal to Sigma. K equal to1 to N, K dot equal to J, modulus of A KJ or J is equal to1, 2, 3 up to n.

Then secondly all these bounds are independent of each other and therefore all the eigenvalues of A must lie in the intersection of these bounds. If any of the Gerschgorin circles is isolated, then this circle contains exactly one eigenvalue and if you are given a real symmetric matrix A such that A IJ is A JI, then in this case we obtain an interval which contains all the eigenvalues. So in case you are given a matrix A which is the real symmetric matrix, then you can see that, you can obtain an interval which contains all the eigenvalues of this real symmetric matrix.

So let us take an example, suppose say A is a matrix having entries 3, 2, 2, 2, 5, 2, 2, 2, 3. So let us apply Gerschgorin’s theorem and Brauer’s theorem and determine the bounds. So all the eigenvalues of this matrix lie in the interval mod lambda less than or equal to maximum of 7, 9 and 7, what is it, it is the row sum or it is also the column sum. So take the 1st row, take the entries in the 1st row and compute the absolute value of each of the entries and add it or take the entries in the 1st column and compute the absolute value of the entries in the 1st column and add that, say add them, then you get value to be 7. Do that for the 2nd row and the 3rd row or the 2nd column and the 3rd column.

(Refer Slide Time: 51:40)





Which theorem have I applied, I have used Gerschgorin's theorem. So I obtain mod lambda to be less than or equal to 9 and therefore lambda is such that it lies between -9 and 9, this is what my Gerschgorin's theorem says. What does that mean, I have an interval between -9 and 9 in which all the eigenvalues of the given matrix lie. Let us now use Brauer's theorem and see what the result is. The Brauer's theorem tells us the eigenvalues lie in the union of the intervals given by modulus of lambda - A1 1 that is 3 less than or equal to P1. What is P1, the sum of the absolute value of the entries in the 1st row, so it is 4.

The 2nd condition is modulus of lambda - A 2 2 which is 5 and that is less than or equal to the sum of the absolute value is of the other entries so that is again 4. And from the 3rd row we get modulus of lambda -3 is less than or equal to 4. Let us look at the 1st inequality, it gives mod lambda -3 less than or equal to 4, mod lambda -5 less than or equal to 4. So -4 less than lambda -5 is less than 4. 1 less than lambda less than 9 and the 3rd inequality is mod lambda -3 less than equal to 4 which is the same as the 1st inequality. So Brauer's theorem tells us that all the Eigen values lie between -1 and 7 and lambda, all the Eigen values lie between 1 and 9.

Therefore we make use of the results obtained using Gerschgorin's theorem and using Brauer's theorem. The 1st inequality says the eigenvalues live between -1 and 7, so it lies in the interval -1 to 7. The 2nd inequality says it lies between 1 and 9, so between 1 and 9. And results from Gerschgorin's theorem tells us that lambda lies between -9 and 9, so it lies between say -9 and 9. So now we have to make a conclusion about the interval within which all the eigenvalues lie. So since the bounds are all independent of each other, so we have to find the intersection of these intervals and write down the interval within which all the eigenvalues lie.

So we observe that the intersection of these intervals is -1 less than or equal to lambda less than or equal to 9. So the Eigen values of this given matrix lie in the interval -1 to 9. So if the given matrix A is a real symmetric matrix, then you can find an interval within which all the eigenvalues of the given matrix lie. And I will include more problems on Gerschgorin's bounds in the assignment sheets.