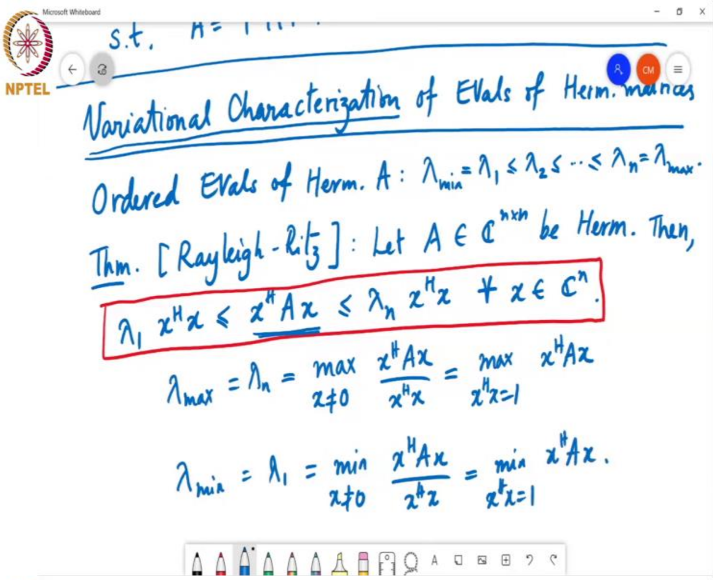
Matrix Theory

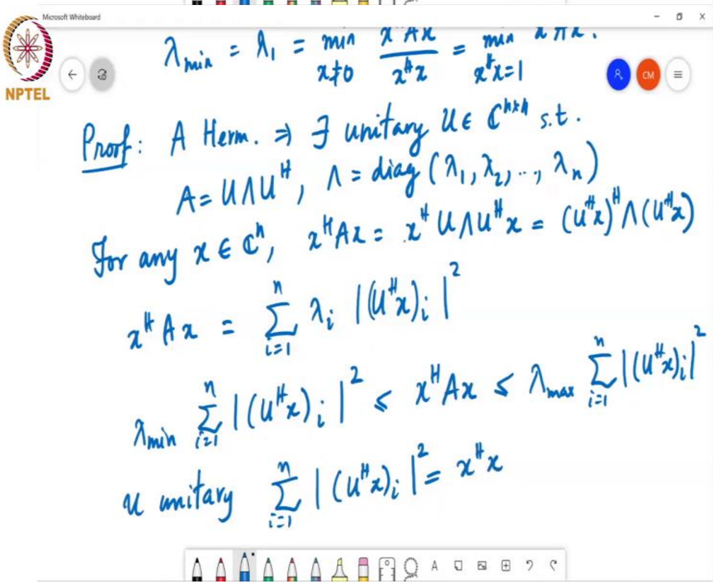
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore

Lecture 59

Variational characterization of Eigenvalues: Rayleigh-Ritz theorem

(Refer Slide Time: 00:14)





Now, the next the next thing I want to say is something also very, very useful which is that in general, the eigenvalues of A matrix are roots of the of its characteristic polynomial. But for Hermitian symmetric matrices, we can write the eigenvalues of A matrix as solutions to optimization problems. And that different way of looking at eigenvalues of Hermitian symmetric matrices is called the Variational Characterization of eigenvalues of Hermitian matrices.

Variational Characterization just means that it is the solution to an optimization problem by varying some cost function and looking for local minima, local maxima, saddle points of this cost function, you can identify all the eigenvalues of the matrix. So, recall that the eigenvalues of Hermitian symmetric matrix are all real.

So, we can consider ordered eigenvalues of Hermitian matrix A, we will order them in this order, so, we will call lambda min or lambda 1 to be the smallest eigenvalue and that is less than or equal to lambda 2 less than or equal to, etcetera, lambda n is the largest eigenvalue. So, we will consider ordered, considered ordered eigenvalues like this. So, we have the following theorem.

And this is called the Rayleigh-Ritz theorem. So, let A in C to the n cross n be Hermitian. Then, lambda 1 times x Hermitian x is less than or equal to x Hermitian Ax is less than or equal to lambda n x Hermitian x for every x in C to the n. What this means is that, if I consider for any x in C to the n, if I consider the quantity x Hermitian Ax that is lower bounded by lambda 1 times x Hermitian x and upper bounded by lambda n times x Hermitian x.

And in fact, the both these lower bound and the upper bound are achievable and you can achieve them by setting x to be the eigenvector corresponding to the smallest and largest eigenvalues respectively. So, basically we have that lambda max is equal to lambda n is equal to the max over all non-zero x of x Hermitian Ax over x Hermitian x, but then I can always say if I scale x by some constant, the numerator scales by that constant modular squared and the denominator also scales by that constant modular square.

And so, I can also write this as max over all x such that x Hermitian x equals 1 of x Hermitian Ax. So, this is what you would call an unconstrained optimization problem except there is a small constraint that x cannot be equal to 0 and this is a constraint optimization problem. So, I can find out lambda n by solving this optimization problem, which is to maximize x Hermitian Ax subject to x Hermitian x being equal to 1.

And similarly, I can write lambda min equal to lambda 1 equal to the min over all x not equal to 0 x Hermitian Ax divided by x Hermitian x which is equal to the min over x Hermitian x equals 1 x Hermitian Ax. So, let us see this, this is a very important result, which we will use many times in the coming classes. So, if A is Hermitian that means that there exists a unitary u, such that A equals u lambda u Hermitian where lambda is a diagonal matrix containing the eigenvalues along the diagonal.

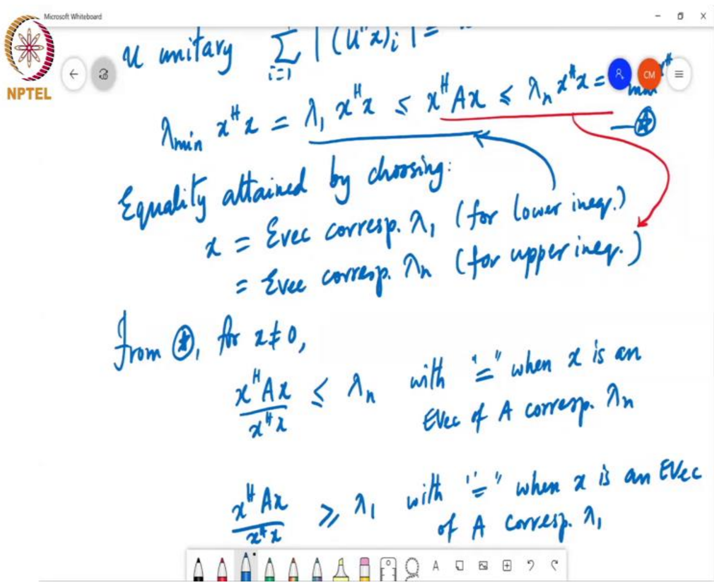
Now, consider for any x and C to the n, x Hermitian Ax is the same as x Hermitian u lambda u Hermitian x, which is of course, equal to u Hermitian x Hermitian times lambda times u Hermitian x. Now, lambda here is a diagonal matrix, so I can expand this out and write this as sigma i equal to 1 to n lambda i times u Hermitian x and then I take the i'th component of it and then mod squared.

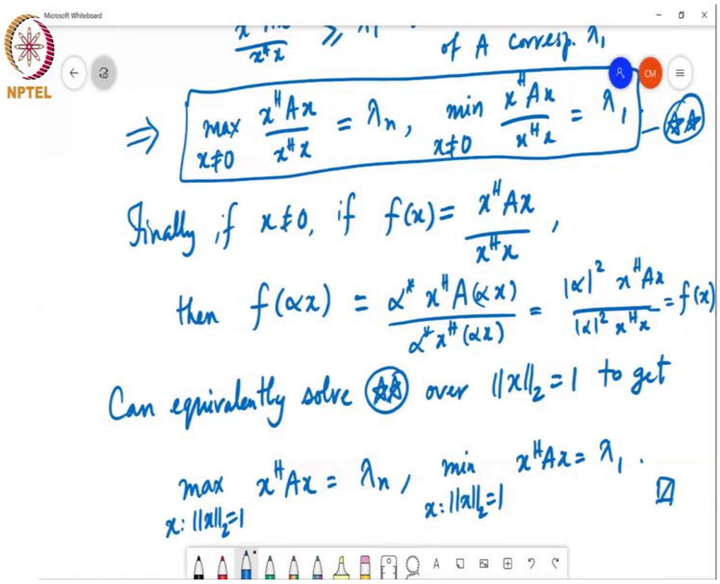
And each of these terms is non-negative. So, that means that if I am taking a linear combination of these terms scaled by lambda i's if I replace all these lambda i's by lambda min that will be a lower bound on whatever value this can achieve. And if I replace all these values by lambda max that will be an upper bound on whatever this can achieve.

So, that means that lambda min times summation i equal to 1 to n mod u Hermitian x i'th components squared is less than or equal to x Hermitian Ax is less than or equal to sigma i equal to 1 to n lambda i, this is equal to this. So, I wanted to write, lambda max sigma is equal to 1 to n mod u Hermitian x i'th component square. And because u is unitary, if I take the summation, this summation is nothing but the summation of mod xi squared which is x Hermitian x.

So, sigma i equal to 1 to n u Hermitian xi squared is equal to I can write this the other way to think about this is that it is ux Hermitian ux which is x Hermitian uu Hermitian x and uu Hermitian is equal to the identity matrix for u being unitary and so this is nothing but x Hermitian x.

(Refer Slide Time: 09:19)





And so, substituting that in here I get lambda min times x Hermitian x is, which is equal to lambda 1 lambda min is the same as lambda 1 in my notation, x Hermitian x is less than or equal to x Hermitian Ax is less than or equal to lambda n times x Hermitian x which is equal to lambda max x Hermitian x, lambda max is the same as lambda min, as lambda n by my notation.

So, we will call this start for later use. Now, so, we found these bounds. Now, when can this equality be attained here and here. So, equality can be attained by choosing x equal to the eigenvector corresponding to lambda 1 for the lower inequality the first one and equal to the eigenvector corresponding to lambda n for the upper inequality, this part here.

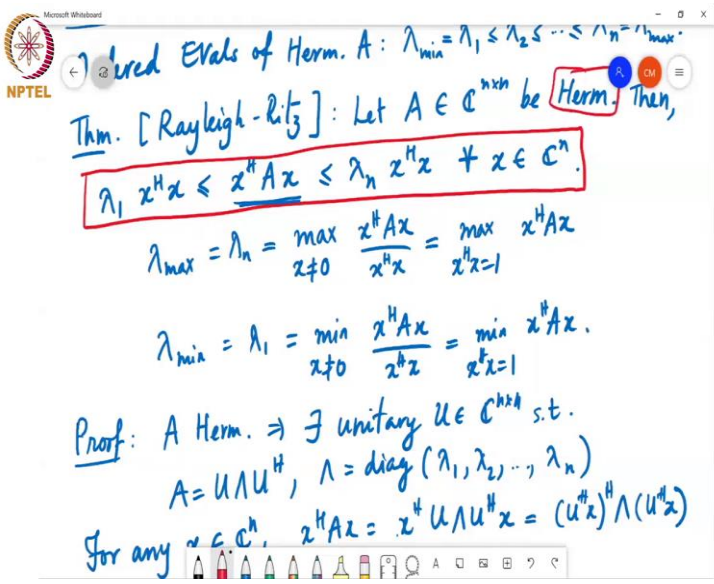
Further from this equation for x not equal to 0, x Hermitian Ax over x Hermitian x, I am just taking x Hermitian x to the other side, and that is less than or equal to lambda max or lambda n with equality when x is an eigenvector of A corresponding to lambda n. And similarly, x Hermitian Ax over x Hermitian x is greater than or equal to lambda 1 with equality when x is an eigenvector of A corresponding to lambda 1.

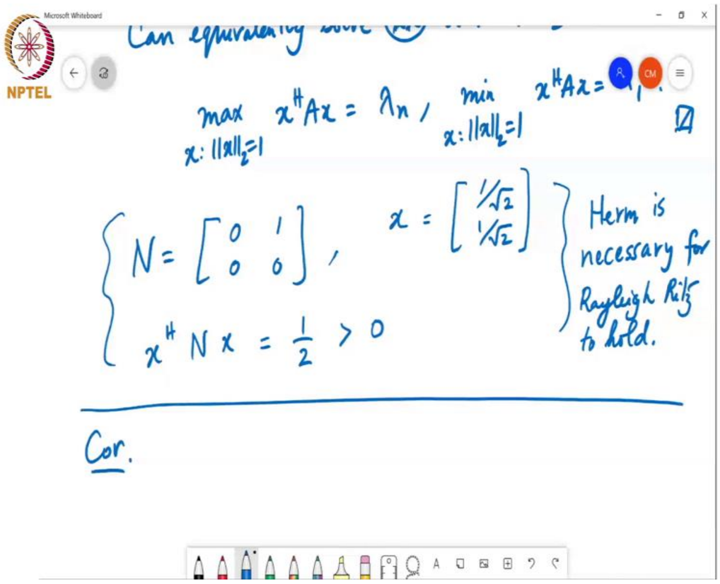
So, this two in turn imply that the max, so this is always less than or equal to this, and it retains equality when x is an eigenvector. So, then what this means is that max over x not equal to 0, so, it is an upper bound that is achievable. So, x Hermitian Ax over x Hermitian x is equal to lambda n, and min over x not equal to 0 of x Hermitian Ax over x Hermitian x is equal to lambda 1.

So, and finally, if x is not equal to 0, then if f of x equals x Hermitian Ax over x Hermitian x then f of alpha x is alpha star x Hermitian A alpha x divided by alpha star x Hermitian alpha x which is equal to mod alpha square times x Hermitian Ax divided by mod alpha square x Hermitian x and alpha squares cancel each other. And so, this is equal to f of x. So, basically I can solve these optimization problems equivalently by considering, I can just scale any non-zero x, I can just scale it to have unit norm and so I can equivalently solve, by or over x l2 norm equals 1 to get max x such that x l2 equals 1, x Hermitian Ax equals lambda n, min over x such that mod x equals 1 x Hermitian Ax equals lambda 1. So, that is the proof of this theorem.

So geometrically, what is happening is that the largest eigenvalue is the largest scaling that can happen to the norm or is the largest value of x Hermitian Ax, as I vary x over the unit sphere, unit, complex, n-dimensional sphere. And lambda 1 is the smallest value of x Hermitian Ax as A vary x over t complex unit sphere in n-dimensions. So that, so that is this result. So, this Rayleigh-Ritz theorem, I will go through the statement of the theorem one second, because it is, and this result is very, very crucial, and we will be using it quite extensively.

(Refer Slide Time: 16:49)



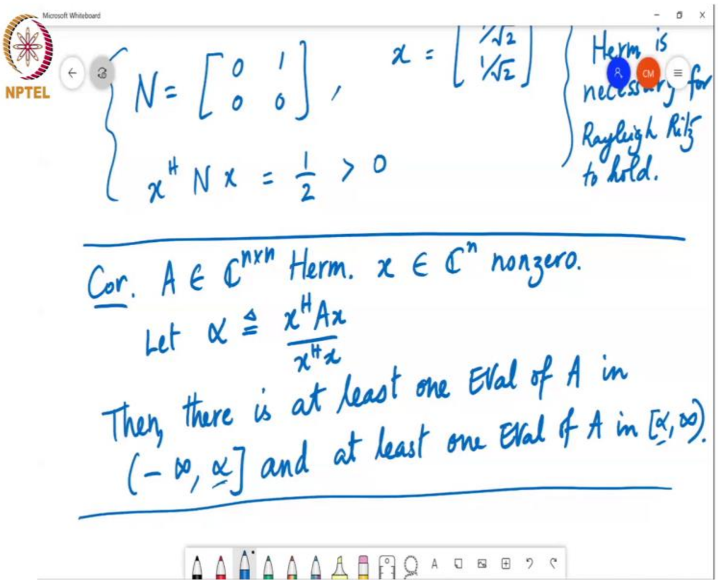


So, if A is a Hermitian matrix, then lambda 1 times x Hermitian x is a lower bound on x Hermitian Ax. and lambda n times x Hermitian x is an upper bound on x Hermitian Ax for all x in C to the n. So, for example, I mean this, the fact that the matrix is Hermitian is crucial for this result to hold. If A is not Hermitian, then this result may not hold. So, just to give you a silly example, to illustrate that, we go back to our favourite defective matrix.

So, if I take N equal to 0, 1, 0, 0, and if I take x equal to 1 over square root of 2 and 1 over square root of 2, then if I compute x Hermitian N x, that is going to be equal to 1 over square root of 2 times 1 over square root of 2. So, that is equal to half which is greater than 0, which is all the eigenvalues of A.

So, in other words, the inequality required by the Rayleigh-Ritz theorem does not hold for this example. So, also from this, the fact that lambda 1 x, x, x Hermitian x is a lower bound on x Hermitian Ax and lambda n times x Hermitian x is an upper bound on x Hermitian Ax, we have the following eigenvalue inclusion result.

(Refer Slide Time: 19:04)



So, A is again a Hermitian matrix and x is some vector in C to the n which is non-zero. Then, let alpha be defined as x Hermitian Ax over x Hermitian x. Then, there is at least one eigenvalue of A in minus infinity alpha and at least one eigenvalue in alpha infinity, because if I take an arbitrary x, this x Hermitian Ax over x Hermitian x is going to be between lambda 1 and lambda n and so there is at least one eigenvalue to the left of this thing including the point alpha and there is at least one eigenvalue to the right of this thing including the point alpha.

Now, of course, this result talked about essentially bounding x Hermitian Ax in terms of the smallest and largest eigenvalues of the matrix A. And so, one could wonder what about the other eigenvalues? So, can we also develop variational characterizations for, for example, lambda 2, lambda 3, and the other eigenvalues turns out the answer is yes. And that is another very cool theorem that we will cover in the next class, Courant-Fischer theorem.