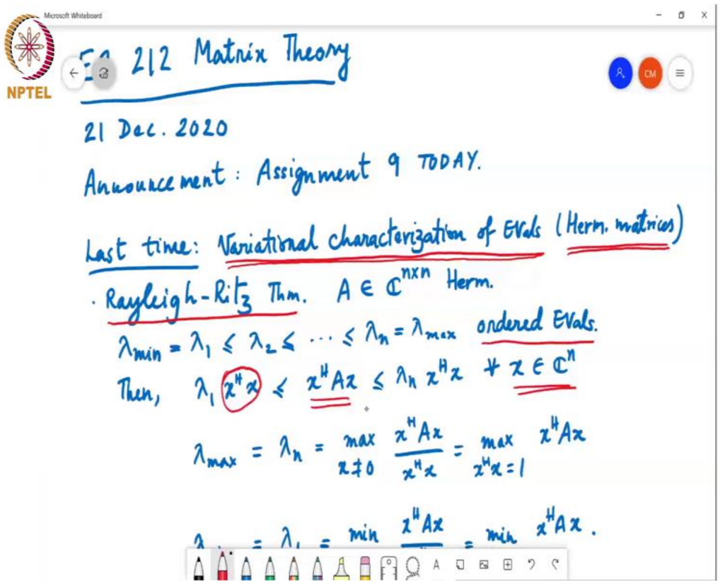
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Variational characterization of eigenvalues (continued)

(Refer Slide Time: 00:14)



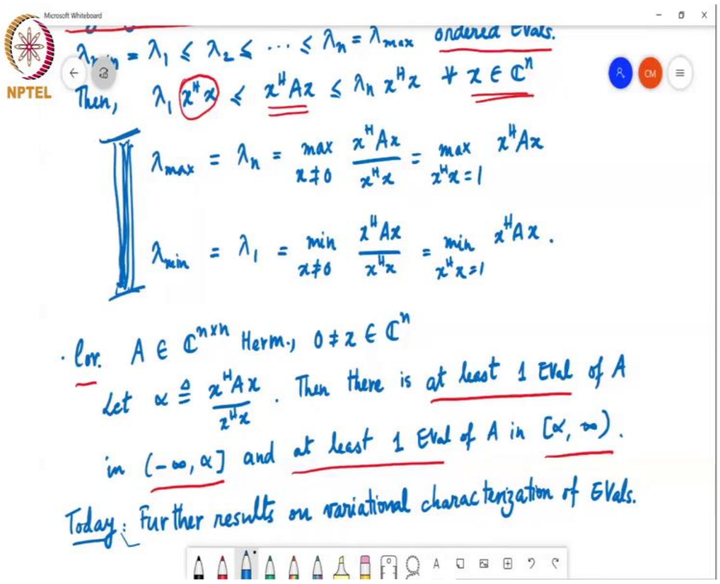
Professor Chandra R. Murthy: The last time we were looking at variational characterization of eigenvalues by which we mean that we are looking at characterizing eigenvalues as solutions to an optimization problem and this is specific to Hermitian matrices and which have the property that the eigenvalues are all real and so, you can consider ordered eigenvalues.

So, you can order them in increasing order and that is the set of eigenvalues, we saw this Rayleigh-Ritz theorem, it said that, if you have a Hermitian matrix with ordered eigenvalues lambda 1 to lambda n, then x Hermitian, Ax is lower bounded by lambda 1 times x Hermitian x and upper bounded by lambda n times x Hermitian x for any x belonging to C to the n.

So, the length x Hermitian x is the length Euclidean length squared of x, it gets scaled when you multiply. In fact, when you do x Hermitian Ax and the smallest possible scaling is lambda 1 and the largest possible scaling is lambda n, so that gives you bounds on how large or small x Hermitian Ax can become compared to x Hermitian x.

And further lambda max or lambda n is equal to the largest value that x Hermitian Ax over x Hermitian x can take over all x not equal to 0 which is the same as maximizing over vectors lying on the unit and dimensional complex ball given by x Hermitian x equals 1 of x Hermitian Ax.

(Refer Slide Time: 01:56)



And similarly, lambda 1 which is the smallest eigenvalue is the minimum of x Hermitian Ax over x Hermitian x for all x not equal to 0 or over all x not equal to 0 and is the same as minimizing x Hermitian Ax over the unit n dimensional complex over the n dimensional complex unit sphere and a corollary to this was that if A is a Hermitian matrix.

Then, if we define alpha to be x Hermitian Ax over x Hermitian x for any non zero, x and C to the n then there is at least one eigenvalue of A in the interval minus infinity and alpha and at least one eigenvalue in the interval alpha to infinity. Now, today we will continue this discussion and talk about further results on such variational characterizations of eigenvalues. So, this Rayleigh-Ritz theorem.

Student: Sir

Professor Chandra R. Murthy: Is there a question?

Student: Yes sir. Sir, what is the use of this corollary?

Professor Chandra R. Murthy: So that, it allows you to identify intervals in which eigenvalues of A must lie. So, we will see some, some examples of further results we can derive based on these results. In fact, this corollary is an easy consequence of the Rayleigh-Ritz theorem. So, sometimes we may not explicitly refer to the Rayleigh-Ritz theorem, refer to this corollary and actually go back to the Rayleigh-Ritz theorem to show it but sometimes, but in fact, it is a consequence of this corollary as well. But we will see some examples of where this will be useful.

But for now, just note that if you have, if you know any x, so for example, I could take x equal to E1, if I take x equal to E1, x Hermitian Ax will be A1 1 the 1 comma 1 element of the matrix A. Of course, x Hermitian x is equal to 1 for that vector. So, what I know then is that there is at least one eigenvalue of A, which is between minus infinity and A1 1 and at least one eigenvalue in the interval A1 1 to infinity and this applies to any diagonal entry, if I take, take x equal to Ek, I will take different diagonal entries of the matrix A.

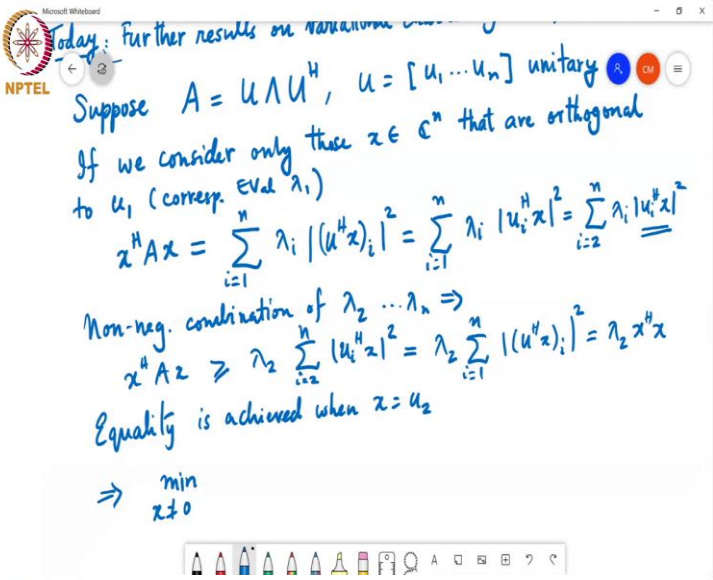
So, what this is saying is that there is at least one eigenvalue that is less than or equal to any one of the diagonal entries of A and at least one eigenvalue which is greater than or equal to any of the diagonal entries of A, and so on.

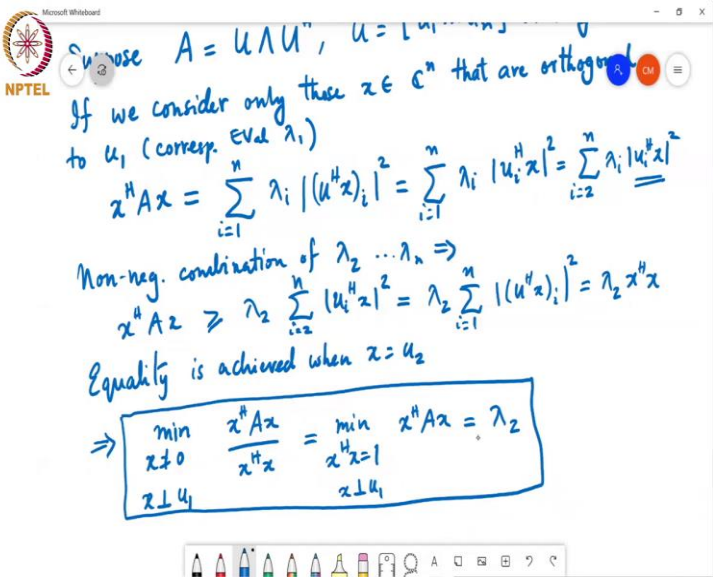
So, in fact, it is often useful to approximately locate these eigenvalues, you may not want to get the exact eigenvalues simply because computing the exact eigenvalues is a computationally expensive task especially for very large dimensional matrices. And so, finding bounds or intervals in which these eigenvalues may lie is actually very useful.

Student: Thank you sir.

Professor Chandra R. Murthy: Yeah. So, so we will continue now this, this result tells us something about lambda max and lambda min. And, natural question is; what about the other eigenvalues? Can we have a variational characterization of the other eigenvalues?

(Refer Slide Time: 05:31)





So, now, suppose any Hermitian symmetric matrix is diagonal, unitarily diagonalizable. So, suppose A can be written as u lambda u Hermitian, where u is unitary and we will denote its columns as u1 through un unitary and lambda is a diagonal matrix containing the eigenvalues of the matrix A.

Now, suppose, we consider only the vectors x that are orthogonal to u1. So, if we consider only those x n C to the n that are to u1, the first column of you, which has the corresponding eigenvalue lambda 1, this is the smallest eigenvalue, then we have the following. So, if I consider x Hermitian Ax, this is equal to, I will expand it out. So, A is u lambda u Hermitian.

So, I can write this as summation i equal to 1 to n lambda i times the entry of u Hermitian x the ith, ith entry square, which in turn is equal to the ith entry of u Hermitian x is simply ui

Hermitian times x, because u has columns u1 to un. And so, I can write that as sigma i equal to 1 to n, lambda i times ui Hermitian x square. Now, u1 Hermitian x is equal to 0 because I am assuming that I am considering only an x which is orthogonal to u 1.

And so, I can further drop the i equal to 1 term and write this as i equal to 2 to n, lambda i, ui Hermitian x square. Now, this is a non-negative number. So, this is a non-negative combination of lambda 1 to lambda n and lambda 2 is the smallest number. So, if I replace all these eigenvalues by lambda 2, I am only making this this summation here smaller. So, then I get, so it is a non-negative combination of lambda 2 to lambda n.

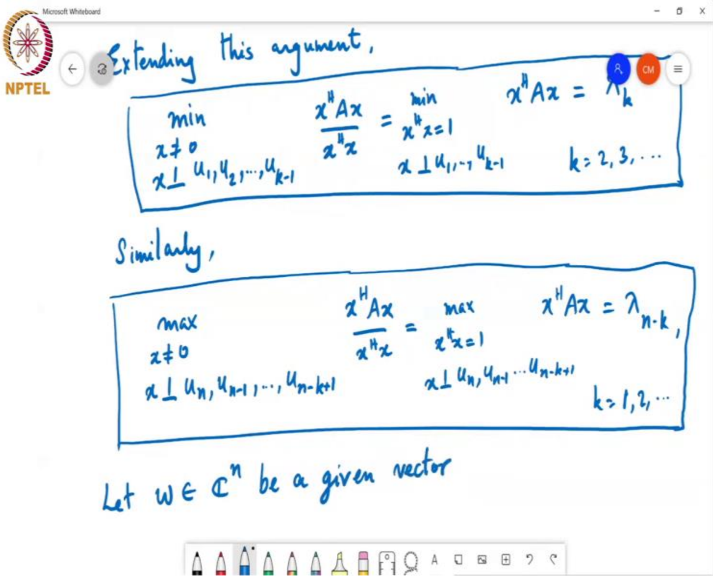
So, I have x Hermitian Ax is greater than or equal to lambda 2 times sigma i equal to 2 to n, ui Hermitian x square and this again, see u is a unitary matrix. And so, this is actually equal to lambda 2 times sigma i equal to 1 to n. I mean, I am reinserting that 0 which was u1 Hermitian x and I will write it as u Hermitian x ith component square and this is just nothing but x Hermitian u, u Hermitian x and u is a unitary matrix.

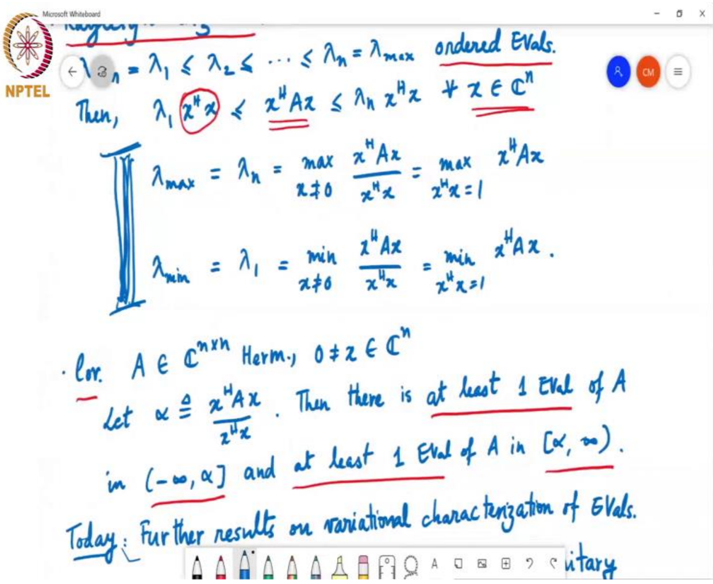
So, this is equal to lambda two times x Hermitian x. So, we have now that x Hermitian Ax is at least or x is greater than or equal to lambda 2 times x Hermitian x for any x that is orthogonal to u1. Now, we can achieve equality in this by choosing x equal to u2. So that means that, I mean, you can see that from here itself, if x equal to u2, then only the u2 term will survive, and this will become E equal to lambda 2 times u2 Hermitian x square, all the other terms will be equal to 0 because these are orthonormal eigenvectors.

And so, then this will become equal to lambda 2 times x Hermitian x, so, or u2 Hermitian u2 which is equal to 1, so u2 Hermitian A u2 is equal to lambda 2, so that means that the minimum over all non zero x, that are perpendicular to u1 of x Hermitian Ax over x Hermitian x, which is actually equal to instead of considering all x here, I can as well minimize over all x such that x Hermitian x equals 1 and retaining this in this constraint x is perpendicular to u1 x Hermitian Ax, x Hermitian x equals 1, so I do not have to divide by that, and this is equal to lambda 2.

So, this shows how I can characterize other eigenvalues in terms of, as a solution to an optimization problem. So, if I want lambda 2, I need to insert a constraint, x should be perpendicular to u1.

(Refer Slide Time: 12: 25)





By making the same exact argument and extending it. We have that, the min over x not equal to 0, x perpendicular to u1, u2 up to uk minus 1, x Hermitian Ax over x Hermitian x, which is equal to the min over x Hermitian x equals 1, x perpendicular to u1 up to uk minus 1, x Hermitian Ax is equal to lambda k.

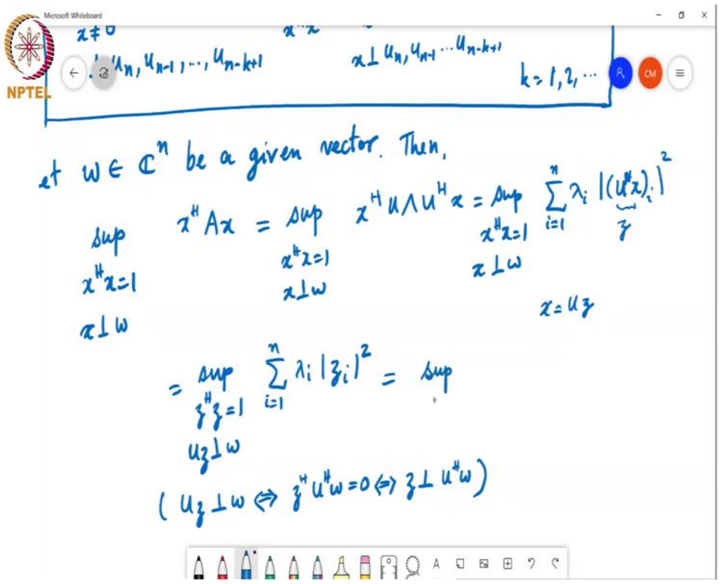
And this is true for k equal to 2, 3, etc. It is also true for k equals 1 except that this inequality and this constraint here, x perpendicular to u1 drops off when I consider k equal to 1. So, we will follow that convention going forward. And so, we may even write k equal to 1, 2, 3, etc. But when I say x is perpendicular to u1 through u k minus 1, and if I say k equals 1, it is kind of saying x is perpendicular to u 0, but there is no such factor like u 0. So, what that means is that this constraint drops off. So, this is 1 way to write all the eigenvalues of the matrix A in terms, in terms of an optimization problem.

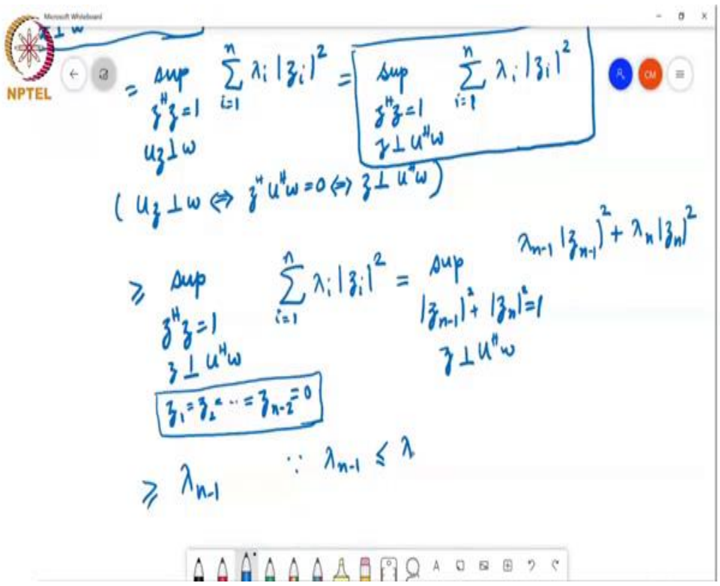
And similarly, so, if you remember we started by looking at, I mean in this we have lambda max also, which is the max solution to a maximization problem. So, starting from lambda n if you had considered all x that are perpendicular to un and then proceeded with exactly these arguments, what you can show is that this is also equivalent to saying that the max over x not equal to 0, x perpendicular to un, un minus 1 all the way down to un minus k plus 1, x Hermitian Ax over x Hermitian x is equal to the max over x Hermitian x equals 1, x perpendicular to un, un minus 1 all the way up to un minus k plus 1, x Hermitian Ax is equal to lambda k.

So, sorry, this is lambda, I have gone up to n minus k plus 1. So, this is lambda n minus k and again k equal to 1, 2, etc, because when I put k equal to 1, I get lambda n minus 1. So, this is another way to characterize the eigenvalues of A as a solution to a maximization optimization problem.

So, we have seen these variational characterizations of all the eigenvalues of matrix A, now this is nice, but it has a small drawback which is that in order to set up the optimization problem, in this case, for example, you need to know what u1, u2 up to uk minus 1 are, or in this case you need to know what un, un minus 1 up to un minus k plus 1 are, we can overcome this dependence on the knowledge of these eigenvalues as follows.

(Refer Slide Time: 16:29)





So, let w be an arbitrary vector in C to the n, then the maximum, this will write a sup for, for no particular reason. So, for the purpose of this course, sup and max are the same the textbook write sup. So, I am also writing sup here x Hermitian x is equal to 1 and x perpendicular to this vector w that is given to us of x Hermitian Ax is equal to the sup of x Hermitian x equal to 1, x perpendicular to w, x Hermitian, I will substitute x A is equal to u lambda u Hermitian. So, this is what I get and this is again equal to just expanding this in terms of a summation, summation i equal to 1 to n lambda i times u Hermitian x ith component square.

Now, what I have to do is, I will call this vector z, u Hermitian x equal z or x equal to uz, so multiply by u Hermitian, u Hermitian x equals, so then I get that this quantity here is equal to the sup. So, x Hermitian x equals 1 is the same as saying z Hermitian u Hermitian uz equals 1 but u Hermitian u is the identity matrix.

So, I can write the constraint as z Hermitian z equals 1 and x is uz, so uz is perpendicular to W of the summation i equal to 1 to n lambda i times mod zi square which is equal to, now if uz is perpendicular to w that means that mathematically this is the same as saying z Hermitian u Hermitian w equals 0 which is the same as saying z is perpendicular to u Hermitian w. Actually, these are both all equivalent statements. So instead of constraining uz to be perpendicular to w, I can say that z should be perpendicular to u Hermitian w, so I will write this as the sup.

And then that is greater than or equal to the supremum over the summation i equal to 1 to n, lambda i mod zi square, subject to z Hermitian z equals 1, z perpendicular to u Hermitian w, and z1 equals z2 equals, etcetera up to zn minus 2 equals 0. Are you able to hear me?

Student: Yes sir, only just now sir.

Professor Chandra R. Murthy: Okay, so I am not gone much further ahead. All I did was I said that this quantity is greater than or equal to the same quantity, but with the extra constraint z1 equals z2 equals zn minus 2 equals 0. Are you able to hear me now?

Student: Sir, your voice starts breaking up every now and then.

Professor Chandra R. Murthy: Hello, can you hear me? Yes, I understand. But unfortunately, I do not have a very good internet connection right now. So, you have to tell me if you are able to follow up the argument I am making, I am making one small argument (())(22:00) that this quantity that we came up to is greater than or equal to this quantity here, which is the same as this except that there is this additional constraint that z1 through zn minus 2 equals 0. Are you able to hear me?

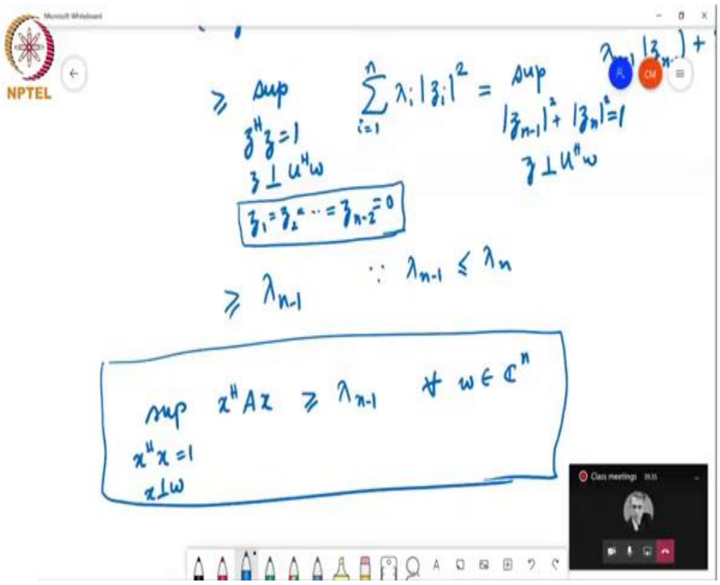
Student: Yes, sir.

Professor Chandra R. Murthy: All I have done is to add a few extra constraints making or forcing some of the zi's to 0 can only decrease the value of this cost function, whatever supremum you could achieve here, you may or may not be able to achieve it here because you have this additional constraint that z1, z2, etcetera up to zn minus two must be equal to 0. So, the cost function value will decrease. And so, this is the same as supremum.

So, since z1 to zn minus 2 equals 0 and z Hermitian z equals 1, I can write that as zn minus 1 square plus zn square equals 1 and the vector z is perpendicular to u Hermitian w of since the 1st n minus 2 zi's are equal to 0, I can drop those terms here and write the cost function as lambda n minus 1 times n minus 1 square plus lambda n times zn square.

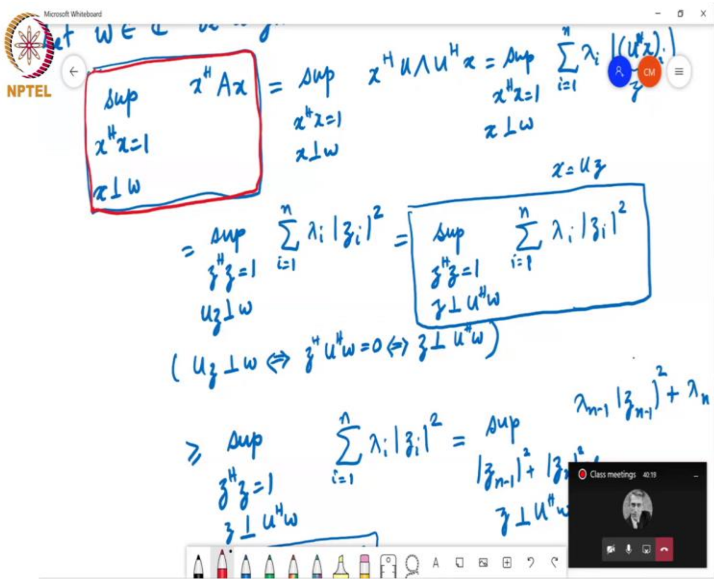
And of course, between these two this quantity is the smaller quantity and we are taking essentially a convex combination of these two terms, because zn minus 1 square plus zn square equals 1 and so, when you take a convex combination of two numbers lambda n minus 1 and lambda n, then whatever (con) this is is going to be some number between lambda n minus 1 and lambda n and so, this is greater than or equal to lambda n minus 1.

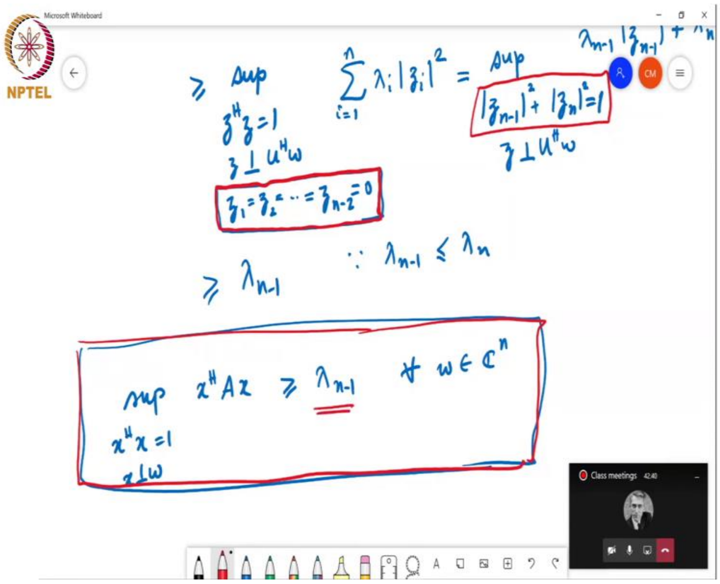
(Refer Slide Time: 25:02)



So, what this says is that what we have just shown is that sup x Hermitian x equals 1, x perpendicular to w, x is greater than equal to lambda n minus 1. And this is true for every w. Now it should be okay. So, I will just very quickly review, what I was saying.

(Refer Slide Time: 25:35)





So, our starting point was, we were looking at the largest value, or the supremum of x Hermitian Ax over all x such that x Hermitian x equals 1, and x is perpendicular to w, we went through a few simplifying steps, and we came up to a point where we showed that this is exactly equal to the supremum of summation i equal to 1 to n lambda i times mod zi square, subject to z Hermitian z equals 1, and z is perpendicular to u Hermitian w.

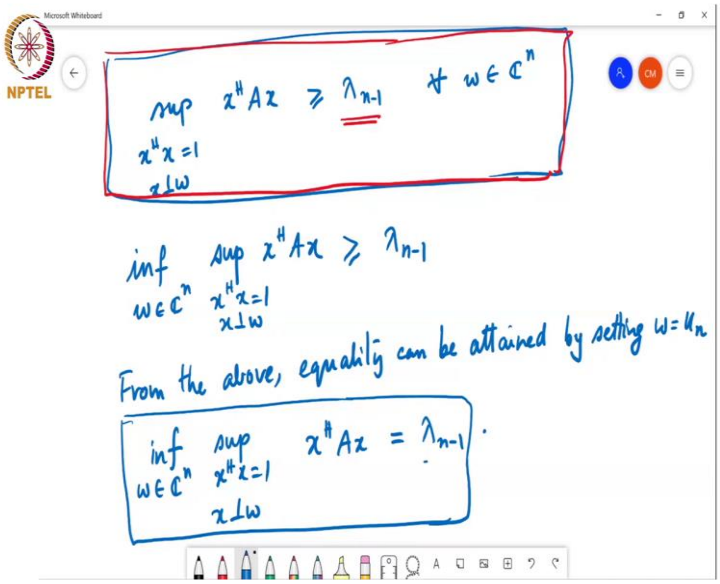
And then we did something which I consider quite brilliant, which is to say that this is greater than or equal to the supremum of the same quantity summation i equal to 1 to n lambda i times mod zi square, subject to z Hermitian z equals 1, z perpendicular to u Hermitian w. But we threw in one extra constraint that z1, z2 up to zn minus 2 are all equal to 0. That is because throwing in an extra constraint can only reduce the value of the cost function, because not all points that are feasible here are going to be feasible here.

Here, you are only allowed to search, you not only have to respect these two constraints, that z Hermitian z equals 1, and z is perpendicular to u Hermitian w, you also have to respect another additional constraint that z1, z2 up to zn minus 2 equals 0. So, this cost function cannot be as, may not be as large, can never be larger than this cost function. And so, this is greater than or equal to this. And, this now that I have set z1 to zn minus 2 equals 0, I can drop the first n minus 2 terms in the summation, and write this as the supremum over.

And similarly this constraint, this is nothing but z1 square plus z2 square plus etcetera up to zn square equals 1, and the 1st n minus 2 terms are equal to 0. So, I can replace the constraint with this constraint here, zn minus 1 square plus zn square equals 1, and the cost function becomes lambda n minus 1, zn minus 1 square plus lambda n zn square. Now, this is these two things in these two quantities add up to 1, so they are numbers between 0 and 1, they are non-negative numbers.

And so, this is just a convex combination of lambda n minus 1 and lambda n. And lambda n minus 1 is smaller than lambda n. So, the smallest this can ever be is just lambda n minus 1. So, in fact, what we ended up showing is that the supremum of x Hermitian Ax subject to x being perpendicular to w, and x Hermitian x equals 1 is at least equal to lambda n minus 1. And this is true for any arbitrary w, which is in C to the n. So, since it is true for any w, even if we throw in an infimum, even if we take the minimum of the left-hand side, that will still satisfy this inequality.

(Refer Slide Time: 28:52)



In other words, I can fix my w to be anything, I will fix it to be the one that achieves the minimum overall w in C to the n of the supremum x Hermitian Ax subject to x Hermitian x equals 1, x perpendicular to w, this is equal to lambda n minus 1, sorry is greater than or equal to lambda n minus 1. So, but then, from what we saw above this quantity will achieve equality if I said w equals un,

Student: Sir.

Professor Chandra R. Murthy: Yes.

Student: Sir, in the infimum statement, why is there are not equality?

Professor Chandra R. Murthy: So, that is what I am coming to next. So that is the, that is the point I can achieve equality here. So, what showed is that the infimum is at least equal to lambda n minus 1. But then when I said w equals u n, I will get lambda n minus 1, that is what we showed earlier. And so, the conclusion is that the infimum over all w in C to the n of the supremum over x Hermitian x equals 1 x perpendicular to w, x Hermitian x is equal to lambda n minus 1.

So, in other words, in this particular optimization problem, this is a different optimization problem that characterizes lambda n minus 1. And in this optimization problem, instead of saying, I will take the supremum over x perpendicular to un, I am doing a supremum over an arbitrary w and then taking an infimum over all such possible w's.

And so, I do not need, I mean, at least technically, the way this optimization problem is set up. I do not need to know what un is in order to solve the problem. It is another matter that the solution to this optimization problem occurs at w equal to un. But, in the problem set up itself, I do not have a requirement that I need to know what u, what un is.