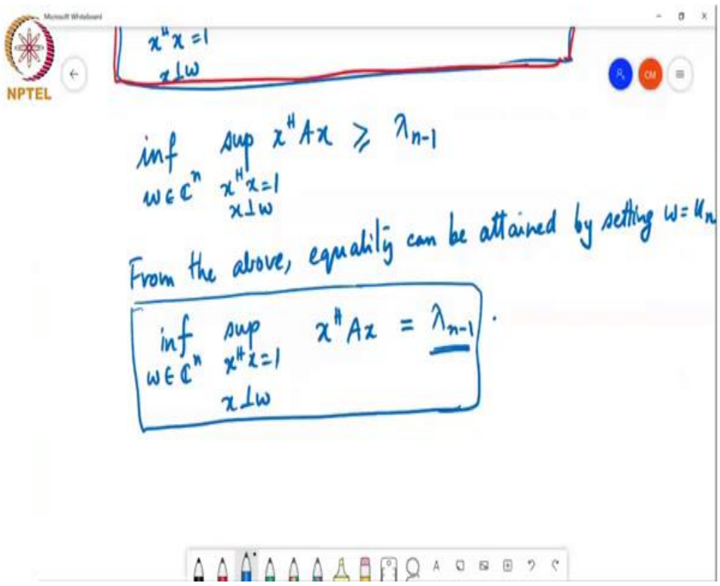
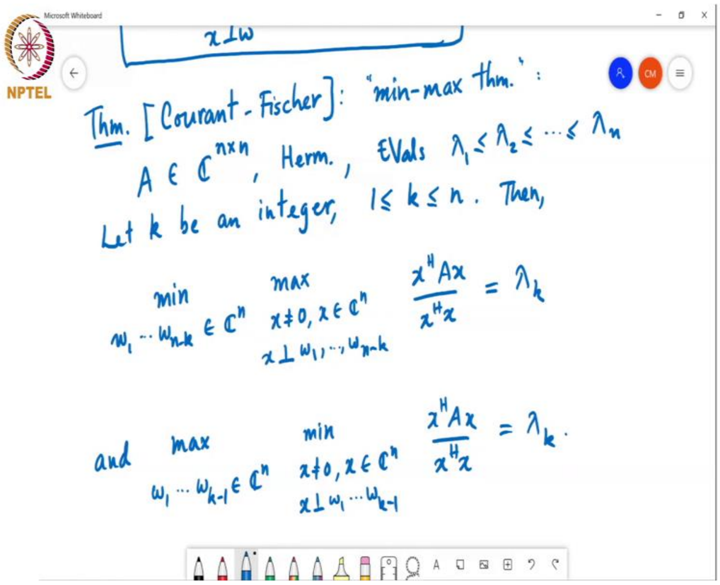
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Courant - Fischer Theorem

(Refer Slide Time: 00:13)





Professor Chandra R. Murthy: So, the generalization of this for all eigenvalues is the following theorem, which is a very central theorem in this variational characterization of eigenvalues. This is the theorem by two people Courant and Fischer, this is called the min max theorem.

So, as usual our setup is A is an n cross n Hermitian symmetric matrix and it has eigenvalues lambda 1 less than or equal to lambda 2 less than or equal to etcetera lambda n and suppose or let k be an integer 1 less than or equal to k less than or equal to n, then we have two, two results, the minimum over w1 through wn minus k in C to the n.

So, I am allowed to choose n minus k vectors in C to the n over which I am doing this minimization the maximum over x not equal to 0, x in C to the n, x perpendicular to w1 through wn minus k, x Hermitian Ax over x Hermitian x is equal to lambda k and the maximum over w1 to wk minus 1, the minimum x not equal to 0, x in C to the n, x perpendicular to w1 through wk minus 1, x Hermitian Ax over x Hermitian x is also equal to lambda k.

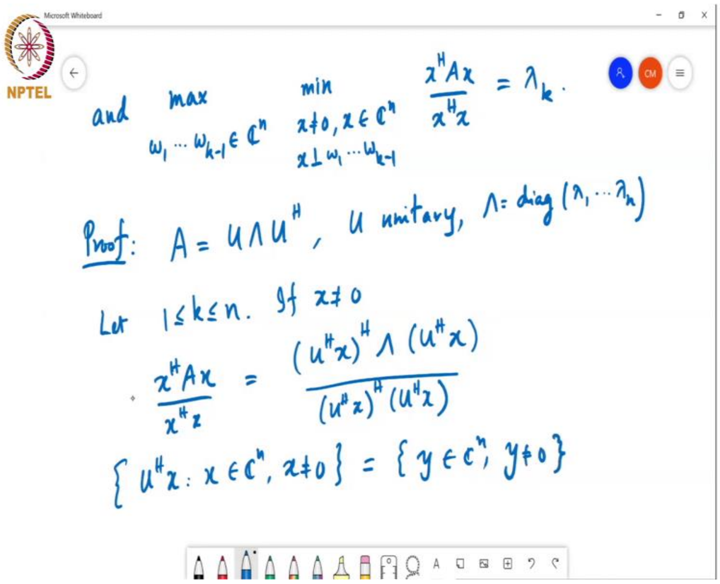
So, there are two ways to write lambda k as a solution to an optimization problem and in both cases, it is a double optimization there is a min max in one case and a max min in the other case.

And, of course, like I mentioned before if k equal to 1 or n in, in this case, when k equals 1, this goes over w1 through wn minus 1 and x perpendicular to w1 to wn minus 1 whereas, in this case this maximization step you can drop it because there is no such thing as w0, so this constraint does not arise. So, you just have to do the minimum over all x not equal to 0, x in C to the n x Hermitian Ax over x Hermitian x that will be equal to lambda 1. So, this constraint also drops off.

Similarly, when k equals n, there is no such thing as w0 here, so, this minimization drops off and similarly, x perpendicular to w1 through w0 there is no such thing. So, this constraint also drops off. And so, the maximum over all non-zero x of x Hermitian Ax over x Hermitian x is equal to lambda n.

So, when k equals 1 or n in one of the two cases, we will be omitting one of the outer optimizations, so we will prove this result, we will only prove the first, this first part, the other part is almost, exactly the same but you just have to modify the steps a bit.

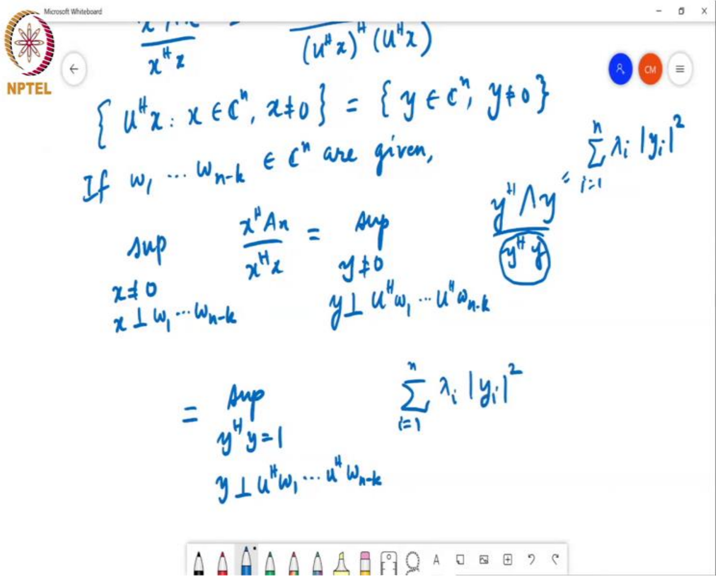
(Refer Slide Time: 4:53)



So, as usual, we will write A as u lambda u Hermitian, u is unitary and lambda is diagonal of lambda 1 through lambda n, then let k be some number which is between 1 and n. Then if x is some vector which is not equal to 0, then x Hermitian Ax over x Hermitian x is as usual, u Hermitian x Hermitian lambda u Hermitian x divided by u Hermitian x Hermitian times u Hermitian x.

And further, if I look at all vectors such that x is over all non-zero x, is the same as, so I will write it this way, u Hermitian x, x in C to the n, x not equal to 0 is the same as the set of vectors y in C to the n, y not equal to 0. In other words, I am thinking of u Hermitian x as y and if I want to optimize this over all x not equal to 0, I can as well optimize it over all y not equal to 0.

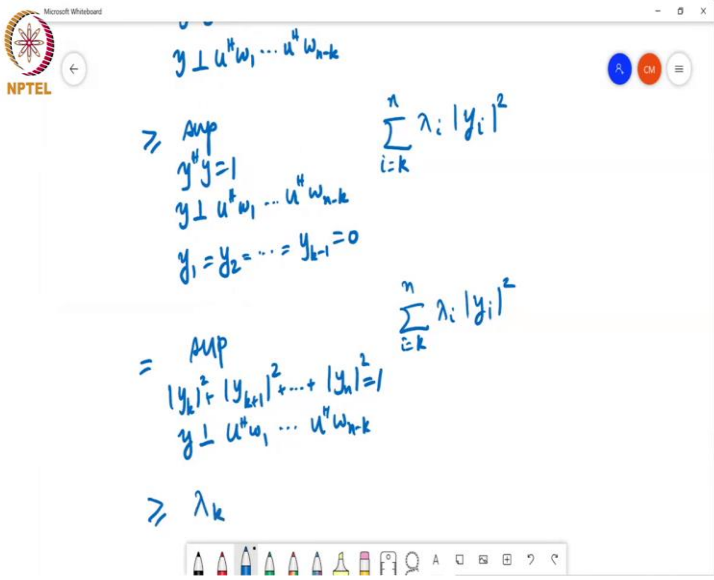
(Refer Slide Time: 7:00)



So, if w1 through w n minus k in C to the n are given, then the sup of x not equal to 0 x perpendicular to w1 through wn minus k of x Hermitian Ax over x Hermitian x is equal to the supremum over y not equal to 0, y perpendicular to instead of since x is, since y is u Hermitian x, I can write this as y is perpendicular to u Hermitian, w1, etcetera up to u Hermitian w n minus k of y Hermitian lambda y, over y Hermitian y.

And as before, I will expand this out. And this numerator is equal to sigma i equal to 1 to n lambda i times mod y square. And further, I can impose a constraint that y Hermitian y equal to 1 and optimize this over all y Hermitian y equal to 1. So, just for the sake of completeness, let me write this step. This is equal to supremum over y Hermitian y equals 1, y perpendicular to u Hermitian w 1 up to u Hermitian w n minus k of sigma i equal to 1 to n lambda i mod y square.

(Refer Slide Time: 8:52)



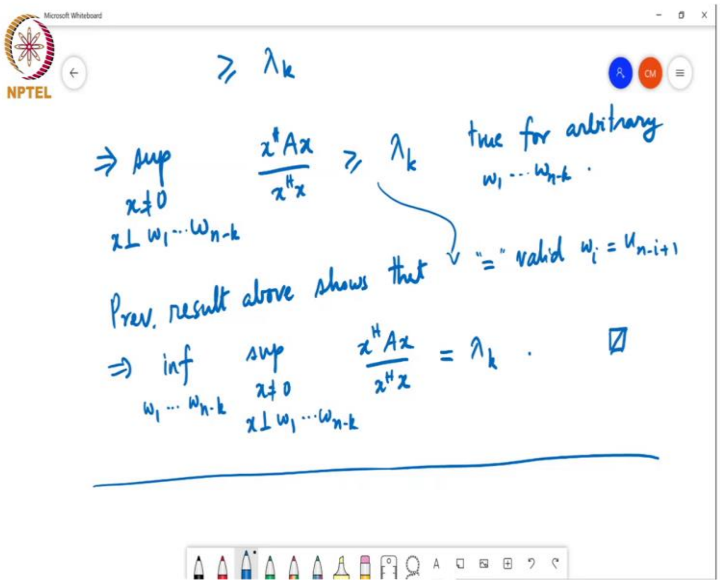
Now, I will do my brilliant thing from the previous discussion. And I will say that this is greater than or equal to the supremum over y Hermitian y equals 1, y perpendicular to u Hermitian w1 up to u Hermitian wn minus k. And I will further set y1 to yk minus 1, yk minus 1 equal to 0 of the summation i equal to 1 to n, y i mod lambda i, lambda i mod y i square.

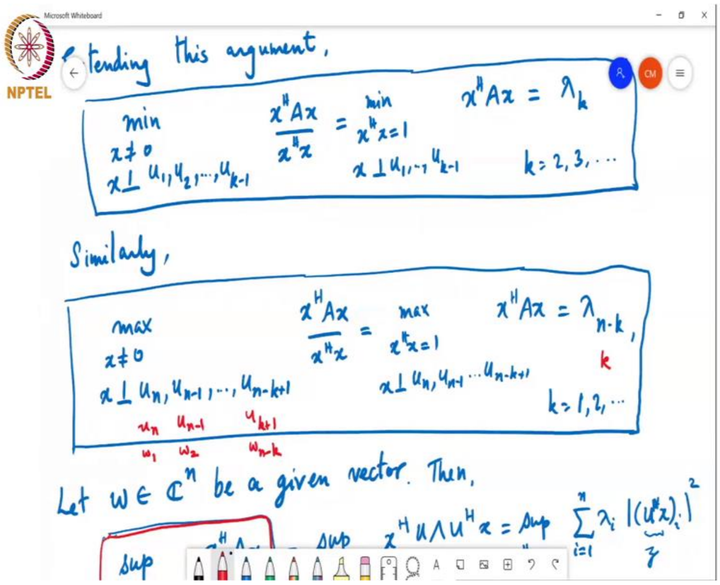
But, since I have set all these guys equal to 0, I can go i equal to k to n lambda i mod yi square and this in turn is equal to the supremum since the first k minus 1 terms are equal to 0, this constraint y Hermitian y equals 1 reduces to yk square plus yk plus 1 square plus etcetera up to yn square equals 1.

And y should still remain perpendicular to u Hermitian w1 up to u Hermitian wn minus k of sigma i equal to k to n, lambda i mod yi square. And of course, as before this is a convex combination of lambda k lambda k plus 1 up to lambda n. And this convex combination is at least equal to the smallest value here.

Another way to think about it is I will replace all these lambda i's with lambda k, then I am only decreasing the value. And summation k equal to i equal to k to n mod yi square equals 1 because of this constraint here. And so, there is nothing left to optimize. And so, I can say that this is greater than or equal to lambda k.

(Refer Slide Time: 11:08)





So, what we have shown then is that, the supremum over all x not equal to 0, x perpendicular to w1 through wn minus k, x Hermitian Ax divided by x Hermitian x is greater than or equal to lambda k. And this is true for arbitrary w1 through wn minus k. But, again the previous result above shows that which is I am referring to this one here, scrolling up. So, the minimum over, so, as the first step I need to change k to n minus k, then what will happen is this un I would be doing.

So, if I replace k by n minus k, I would be going from un, un minus 1 up to u, I have replaced k with n minus k and I would have, I am counting down, I have to count down up to uk plus 1.

Student: Sir k plus 1. Yes sir.

Professor Chandra R. Murthy: Okay, and this, if x is perpendicular to all of these vectors, then what I will get here will be let me write this with a different color so that it is not confusing later on, un, un minus 1 uk plus 1 will give me I have replaced k with n minus k, so, I will get lambda k here.

Now, what I will do is, I will call this vector w1, this vector w2, etcetera, then this will become w. So, if, so this will, so how many vectors do I have here? I have, I am going down from n to k plus 1. So, there is.

Student: n minus k.

Professor Chandra R. Murthy: And, there are exactly n minus k vectors here and so, I have w1 through wn minus k. So basically if I set w1 equals un, w2 equals un minus 2, and w, u, wn minus k equals uk plus 1 the maximum of this subject to x perpendicular to all these vectors will be equal to lambda k. So, the largest value this can take for that specific choice of w1 through wn minus 1, w1, w2 up to wn minus k is equal to lambda k.

So, therefore, just go back and think about it. So, this result shows that equality, equality here is valid when wi equals u n minus i plus 1, that implies that the infimum over all w1 through wn minus k of the supremum x not equal to 0, x perpendicular to w1 through wn minus k, x Hermitian Ax over x Hermitian x is equal to lambda k, which completes the proof. So, I do not want to go further ahead, the next theorem is a is another very, very crucial theorem, which is called Weyl's theorem. We will discuss that the next class.

But what I strongly suggest is, you know, you guys should definitely go over the proof on your own and make sure you understand every step of this proof, because the ideas in this proof will use again and again and again to prove many, many more results. And these arguments are slightly tricky to convey orally, especially in this online mode.

And since I cannot see you guys, I do not know if you are able to follow the proof as I explained it or not. But based on whatever I said, if you now go back and look at the proof on your own, you will be able to fill in the steps. And if you are not, then please stop me at the beginning of the next class and ask me which step you were not able to follow. And we can go over the argument again.

But the arguments we made today are crucial, we are going to reuse them in many proofs going forward. And at that time, if this proof is not completely clear to you, you would not follow many of the proofs that we are going to discuss in the following classes. So please spend some time on it. So, we will stop here for today and we will continue in the next class.