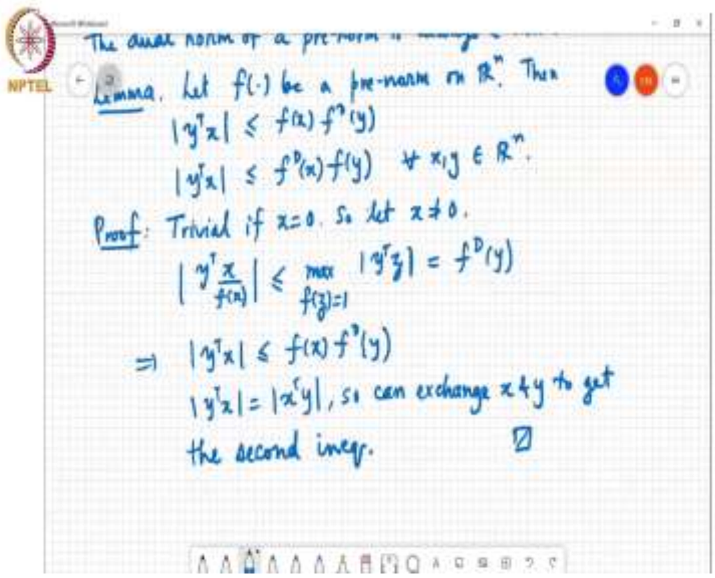
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Properties and Examples of Dual Norms

(Refer Slide Time: 00:25)



So, we have a couple of Properties of these Dual Norms. The first is that let f be a pre-norm on R to the n, then mod of y transpose x is less than or equal to f of x, f of D of y and it is also less than or equal to f D of x times f of y. So, this is a little reminiscent of the Cauchy–Schwarz inequality which says that model y transpose x is less than or equal to the l 2 norm of x times the l 2 norm of y.

It turns out that the dual norm of the l 2 norm is the l 2 norm itself. So, when I, when I consider F to be the l 2 norm, which is , which is of course a vector norm, but it is also a pre-norm, then this these inequalities reduced to the Cauchy–Schwarz inequality and this is true for every x y belonging to R to the n. So, but this is a more general inequality.

Now, how do you show this. So, of course, if for example, x was equal to 0, then the left hand side is 0, the right hand side is also equal to 0 in either case and so the inequalities certainly holds when x equals 0. So, if x equals 0, so let x be non-zero, then what we have is, let me consider the quantity y transpose x divided by f of x.

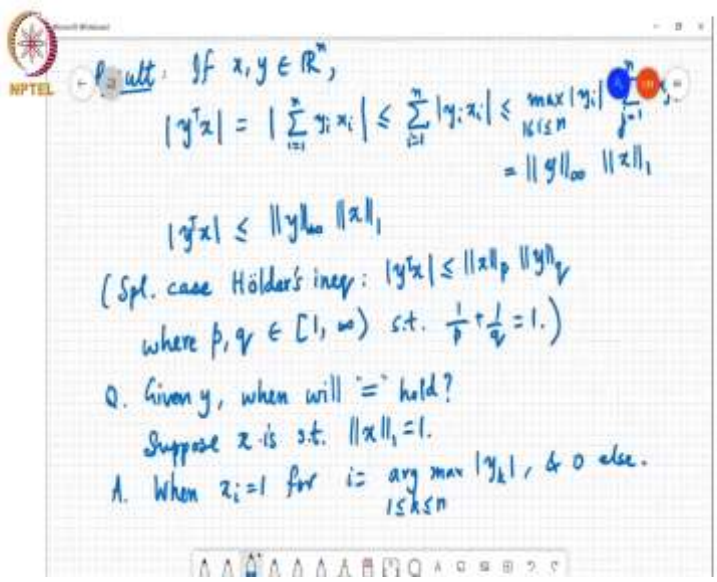
Since x is not equal to 0, f of x is nonzero. So, I can consider the vector x over f of x and this is less than or equal to the max over all these z such that f of z equals 1 of mod of Y transpose z, y is this true, of course, this is equal to f d of y by definition and the inequality immediately follows. So, this implies I am just going to take this f of x to the other side, it is not, it is strictly positive.

So, mod y transpose x is less than or equal to f of x f D of y. But why is this inequality true? This is trivial, I just made the same argument, x over f of x is one such vector such that f of z equals 1, if I took f of x over f of x, I get, I will get the value equal to 1. So, it satisfies this and so here what I am doing on the right hand side is, I am not restricting myself to x over f of x, instead, I am considering all possible z such that f of z equals 1 and I am maximizing y transpose z.

So, any one this is like one candidate solution to this optimization problem and since this is maximizing this quantity, this has to be at least equal to its value at one of the feasible points and so then that is equal to f D of y. Now, the rest of the proof follows immediately because mode of y transpose x is equal to mod of x transpose y and so I can just exchange x and y and then I will have y transpose x magnitude is less than or equal to f D of x times f of y.

So, now that we have defined dual norms, we can ask there is some, we looked at some examples of norms and we can ask what are the duals of those norms? So, here is one, one result, which will help to answer that.

(Refer Slide Time: 05:40)



If x and y are vectors in R to the n then mod of Y transpose x is equal to mod of sigma i equal to 1 to n, y i x i, which is less than or equal to, I will take the mod inside, that will only increase the value or it cannot decrease the value i equal to 1 to n mod y i, x i, which is less than or equal to.

Now, what I can do is in this summation, I can pull out the largest value of y and a magnitude and that will only in other words, all these y is I will replace with the biggest of these, biggest of y 1 through y n and then that is just some single number that is multiplying all these x's. So, I will write it like this. So, max 1 less than or equal to i less than or equal to n mod y i times sigma i equal to, let me just write it with j, so that you do not get confused.

j equal to 1 to n mod x j and this is equal to by definition the max entry of, max modulus entry of y is what we call the infinity norm and the sum of the magnitudes of x is what we call the l1 norm. So, what we have is that y transpose x is less than or equal to y infinity times x and y, this is like similar to the Cauchy-Schwarz inequality, there the two norms being operated were the l 2 norm and the l 2 norm.

Here I have the infinity norm and I have the l 1 norm. This is in fact, special case of an inequality known as Holder's inequality which says that mod of y transpose x is less than or equal to norm x l p norm times y l q norm where p and q are such that 1 over p plus 1 over q equals 1. So, for example, if I choose p equals 1, then I must choose q equals infinity.

So, that 1 plus 1 over infinity which is 0 equals 1. So, it reduces to this inequality when I said p equals 1. So, the question now I can ask is, if I am given y when will equality hold here, or for what x will equality hold, then for that you have to examine, where we did this inequalities here and ask when will this, these inequalities hold with equality.

Now, the when you take the modulus inside, this will be, this will hold with equality if each of these terms were already non-negative and so then there is no cancellations that are happening across the terms here and so taking the mod inside does not really change this value. So, the two will be equal.

So, this should be non-negative and then when will this not affected, it would not affect it, if the x is such that pulling out the maximum value of y i does not affect this overall summation. In other words if x was chosen such that it has a nonzero entry only for the entry of y which solves this optimization problem that is, suppose the first entry of y was the maximum magnitude entry.

Then if only the first entry of X was nonzero and all other entries of x were equal to 0, then the other terms in the summation are not contributing to the sum anyway and so then pulling out this maximum value, which is the first entry of y is not going to change the value of, value of this quantity, I mean, these two will become equal. So, basically, so suppose x was such that.

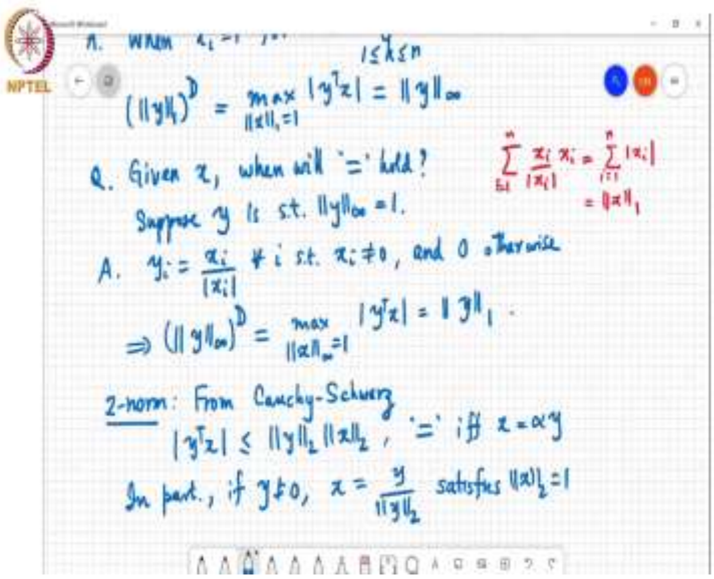
Student: Sir? Alternatively, we can say that all the values of y are equal, will that make sense?

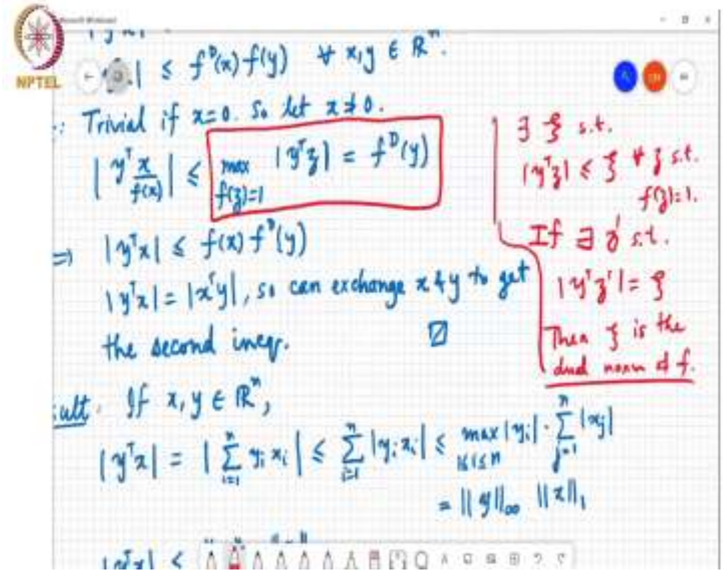
Professor: No, I am saying y is given. I will come back to that. That is a good point. I will come back to this point in a moment, when I asked me with the alternative question, which is when given an x, when will the equality hold? Then your answer is absolutely right. You want to choose all the values of y to be equal. But now I am looking over all xs such that the l 1 norm of x is equal to 1 and I am asking when will equality hold in here.

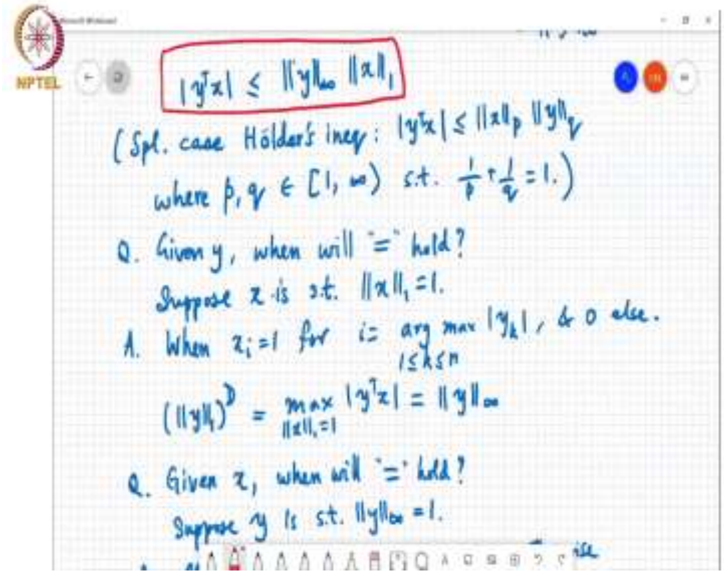
Student: Okay, sir, yeah.

Professor: In this inequality and so if, so the equality holds when x i equals 1 for i equal to the argument, 1 less than or equal to k less than or equal to n, that maximizes, more y k and 0 otherwise. Then, so that is the x.

(Refer Slide Time: 12:51)







In other words, what I have just done is I have actually solved the problem of, I have actually solved what is the maximum over all norm x l 1 equals 1 of mod y transpose x. So, I am given a y I am asking what is this value and this is equal to the max magnitude entry of y, which is equal to norm y infinity. But by definition, this is equal to the dual norm of the l 1 norm, so I will write that as norm y l 1 dual.

Basically, the dual norm of the l 1 norm is the l infinity norm. And similarly, I can ask the alternative question given. So in other words, what I am trying to show you here is how to find the dual norm. Given a norm, we ask, what is the dual norm of a given norm this is how you solve it. So, for given x when will equality hold, and it holds when again suppose y is such that, then basically, what I have to do is to choose x, choose y i is equal to x i over mod x i for every i such that x i is not equal to 0.

And if x equals 0 it is eventually going to multiply with 0 anyway, but I can choose it to be 0 otherwise. I can choose it to be any value less than 1 in magnitude. But because the y infinity is bounded by 1, so I should not the entry of y to be greater than 1, but I can choose it to be 0 otherwise, then this implies y infinity the dual norm which is equal to the max over all x such that x infinity equals 1 of mod y transpose x.

Now, if I substitute this y i into that summation above, you can see that this is equal to the mod x i will cancel with the mod x i and so, you will be left with summation of mod x i and so this is equal to norm y l 1. So, the dual norm of the l infinity norm is the l 1 norm. So, specifically what did I do here, I will just repeat this for the sake of clarity, what I did was, recall that the by definition, the dual norm is like this.

If I want to find the dual norm, I want to find, I want to solve this optimization problem maximize y transpose x magnitude subject to f of z equals 1. So, I can choose this f to be some particular norm or a pre norm and I can ask what is the dual norm. In order to find this again I have to solve this optimization problems.

Now, one typical trick in solving optimization problems is to find an upper bound on the cost function and try to see if there is a z satisfying the constraint where you actually achieve this upper bound, if you can find an upper bound. So, suppose I can find some, if there exists some say zeta such that mod y transpose z is less than or equal to zeta for all z, such that f of z equals 1, then if and if there exists some z prime such that mod y transpose z prime is equal to this zeta, then this quantity, then zeta is the dual norm.

So, that is the process I am trying to follow here. What I did first is I found an upper bound on mod y transpose x in terms of the l 1 norm of x and the infinity norm of y and so, then if I restrict the l 1 norm of x to be equal to 1, then I know that model y transpose x is less than or equal to norm at the l infinity norm of y for all y, then I asked can I ever, can achieve this upper bound?

Yes, answer is yes, I can achieve this upper bound of non y infinity, if I choose x such that x i equals 1 for the argument i which maximizes mod y k and 0 otherwise and this in turn shows that the dual norm of the l 1 norm is equal to the l infinity norm of y. And similarly, if I restrict the lawn y to be y infinity equal to 1, then I asked whether this mod y transpose x can never be equal to the l 1 norm of x.

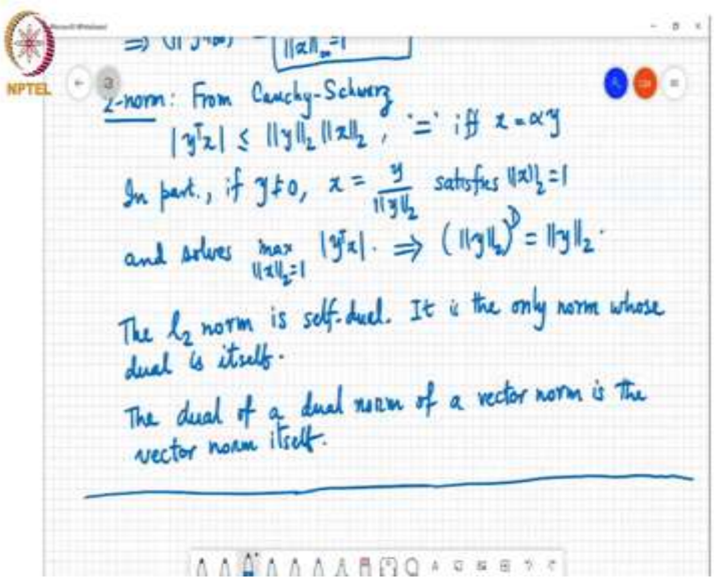
This is the upper bound for norm y infinity equal to 1 and I find that the answer is yes, if I just choose y i equal to x i over mod x i, then if you look at what happens to y transpose x, that becomes summation y i, summation of y i x i, but y i itself is x i over mod x i times x i i equal to 1 to n and this is x i times x i over mod x i is.

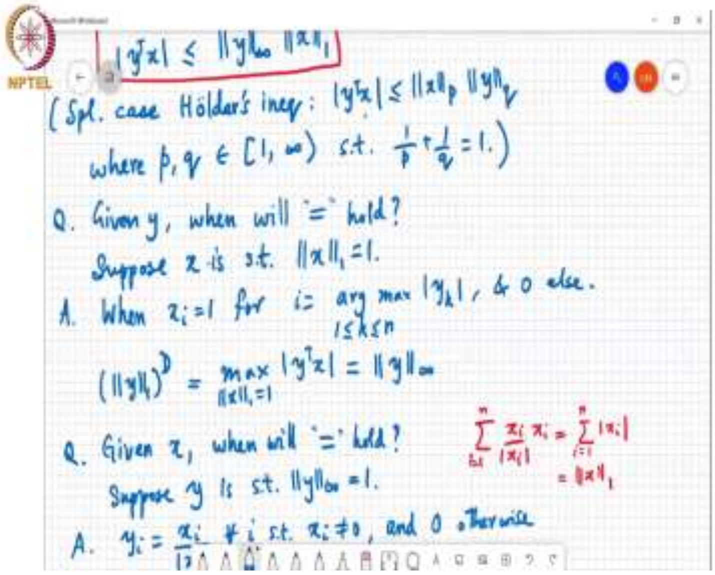
So, this is equal to summation i equal to 1 to n mod x i. So, because this is x i square and I can write that as mod x i square and that divided by x i is just mod x i and which is in turn equal to the l1 norm of x and as a consequence the dual norm of the l infinity norm is the l 1 norm.

Similarly, if I take the l 2 norm then I have from Cauchy-Schwarz mod y transpose x is less than or equal to norm y l 2 times norm x l 2 and equality if and only if x and y are dependent or x is equal to some alpha times y. So, in particular if y is nonzero, then x, choosing x to be equal to y over the l 2 norm of y satisfies x l 2 norm equals 1 and so this quantity it will end and it solves max norm x l 2 equals 1 mod y transpose x.

So, that implies the dual norm is equal to the l 2 norm itself. So, we say that the l 2 norm is its self-dual. In fact, it is the only norm that has this property.

(Refer Slide Time: 21:26)





Student: Sir in the previous one the dual norm of phi infinity should be x norm of y x 1? Or should be y 1? The previous equation dual norm of y infinity should we 1 norm of x not y?

Professor: No, no, no. See, when an optimization problem like this, the solution to the optimization cannot contain x, I have already searched over all possible access and I am asking what is the maximum value of this, it will only depend on y there is no meaning to writing the dual norm of y as something that depends on x axis, x is like a local variable to this optimization problem.

It is the only norm which has self-dual. Another, so this is also something that can be shown, I would not show it here, but it can be shown. So, if you take, if you start from the property that the dual norm of a given norm is equal to the norm itself, you can then derive and show that norm must be the Euclidean norm and also another property is that the dual norm of the dual norm of a vector norm.

So, we start with a norm, we find its dual norm and then we ask what is the dual of the dual norm and you see from these two examples, that if I started with the l 1 norm the dual of that is the l infinity norm, then if I asked what is the dual of the l infinity norm I get back the l 1 norm. Here also in the second example, anyway if I take the dual norm of the Euclidean norm, I get the Euclidean norm. Then if I ask what is the dual of that, it is again the Euclidean norm and this property is true for all norms.

So, the dual of a dual norm of a vector norm is the vector norm itself. So, you can go on producing new norms by finding the dual of the dual of the dual and so, you can do two of them and that is it, you stop there. Yeah. What is the question?

Student: Yeah. My question is, should you taken a norm, should the dual norm always exist? I mean, for any norm, do we have a dual norm for it?

Professor: Yes. So essentially, the point is you are maximizing a linear, which is a convex function over a compact set. So, it will always have a maximizer within that compact set and so the dual norm always exists, it may not always be easily computable; you may need to solve an optimization problem like this. In some special cases, like the ones we considered, it is possible to work out what is the solution to this optimization problem, but that need not always be the case.

Student: And, sir one more doubt in the l p norm that we have defined, can p be any rational number or should it be only a positive integer?

Professor: It can be any number, any rational number, it can even be an irrational number.

Student: Okay, one of the follow up questions sir. So can I, so given an norm say l p norm, can I say that its dual norm will be an l q norm, where p and q satisfy the holder inequality? Is it necessary?

Professor: I think so. But I need to double check that. Yeah, so this inequality itself suggests that, that this inequality is holder's inequality, it is a generalization of the Cauchy-Schwarz inequality and for a given x such that x l p norm equals 1, if you ask what is the, for a given y, if you fix x such that x l p norm equals 1.

And then you ask among all such vectors whose l p norm equals 1, when can the maximum of y transpose x be attained and if you solve that optimization problem, you will find that the maximum will be attained at an x such that the value of y transpose x equals this upper bound which is the l q norm of y, where q is a number such that 1 over p plus 1 over q equals 1.

So, those are not that easy to show algebraically, but it is possible to show that and therefore, the dual norm of the l p norm, where p is any number between 1 and infinity is the l q norm, where q is the number satisfying this equality here, 1 over p plus 1 over 2 equals 1.

Student: Yes, okay. That was my doubt. Okay. Thank you.