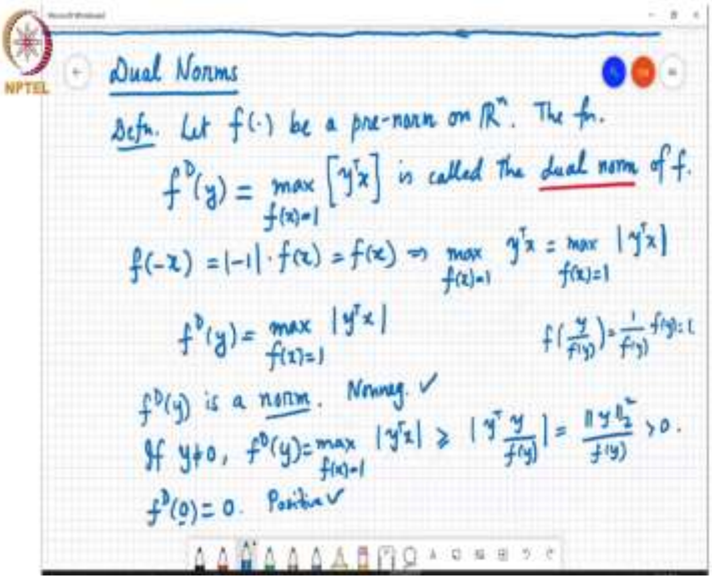
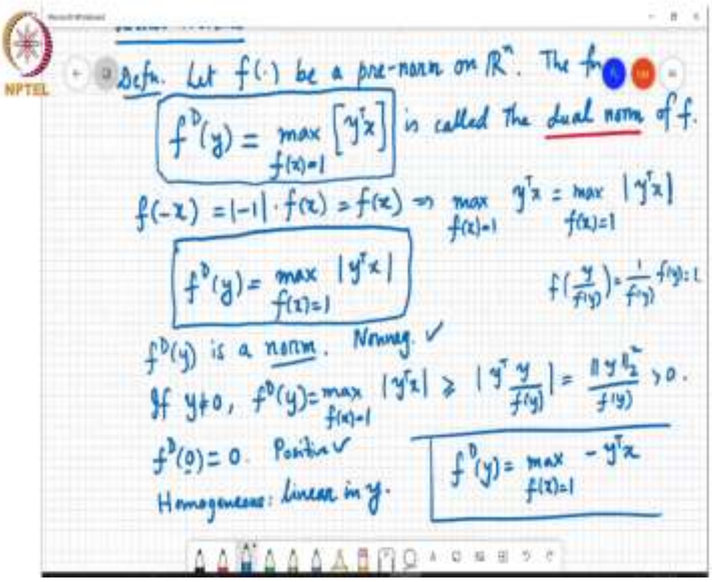
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Technology, Bangalore
Dual norms

(Refer Slide Time: 00:16)





And in particular, I want to talk about dual norms. This is also another way to generate norms from other norms. So, the definition is like this: If f denotes a pre-norm, what is a pre-norm, it is satisfies all the properties of a norm except for the triangle inequality, may not satisfy the triangle inequality, then the function fD of y defined as the max over all x such that f of x equals 1 of y transpose x is called a dual norm.

So, notice that the dual norm is defined via an optimization problem, in order to find the dual norm at a point y you need to solve this optimal problem where you are asked to maximize y transpose x over all x satisfying f of x equals 1. So, it is a constrained optimization problem, the cost function, namely y transpose x is linear in y. And it is also linear in x. But the space of optimization could be a somewhat complicated space because set of x is such that f of x equals 1.

Also notice that because f is a pre-norm. And by homogeneity of f a pre-norm, satisfies this homogeneity property. So, if you do like f of minus x, that is equal to minus 1 magnitude of minus 1 times f of x, which is equal to f of x. And so, to write this optimization problem, as so this basically means that max over f of x equals 1, y transpose x is that is actually the same as the max over f of x equals 1 of mod of y transpose x.

So basically, because of this, we sometimes also write fD of y is equal to the max over f of x equals 1 mod of y transpose x. So, both are both are equivalent optimization problems, you can write it either way. So, this is called a dual norm, not without reason, it is because the dual norm of f is a norm.

So, in order to show that you need to show that this fD of y satisfies the four properties we need, that is non negativity, positivity, homogeneity and triangle inequality. And that is why we are writing it in this alternative form that it is the maximum of the mod y transpose x helps because when you take a quantity and you are maximizing it over a set of points, that is always going to be non negative.

And so clearly, fD of y is non negative. If so, the other point is that it is positive unless y is equal to 0 in order to see that if you take so I will just write this here. fD of y is a norm, and it is non negative obviously. And for positive if y is not equal to 0, then fD of y which is the max f of x equals 1 mod y transpose x, this I can lower bound by choosing a specific value of y. And I will particularly choose mod of y transpose y over f of y.

And this y over f of y satisfies f of x equals 1 and write it up here, f of y over f of y, f of y is just a scaling it is non negative. So, I can write that as 1 over f of y times f of y, which is equal to 1. And so, this satisfies this constraint. And so, the max over all possible x such that f of x equals 1 is at least equal to the value y… mod of y transpose x takes for a particular value of x satisfying the constraint and that is this one.

And this is equal to the norm y l2 square divided by f of y. And this is strictly positive, because y is not equal to 0, f of y is strictly positive, because y is not equal to 0, and therefore this is strictly greater than 0. And we also have that fD of 0. Because if I take the 0 vector, no matter what I multiply which x I chose here, y transpose x is always equal to 0.

And so, the max that mod y transpose x can achieve for overall x such that f of x equals 1 is just equal to 0, so it is equal to 0 when y equals 0, and it is strictly greater than 0 for y not equal to 0. So, it satisfies the positivity property also.

Student: Sir?

Professor: Yeah.

Student: Sir, would you explain how f of x is equals to max of mod y transpose x, because we are not multiplying anything with like mod of minus 1 there was like nothing as multiplied (())(07:10).

Professor: It is trivial actually. So, suppose that, in fact, so if you want to minimize this quantity, y transpose, suppose you solve this problem, and you get an x where this is maximized. If you just take minus x, then obviously, this quantity will get minimized when you substitute minus x that also satisfies this constraint. And since this attains its maximum at that particular x, this will attain its minimum at minus x.

So, if there is an x for which this quantity is negative, and very large, by just substituting minus x, you can achieve the same large pause large value but in the positive direction by replacing x with minus x. And that is why the two problems, were one where you are writing fD of y equal to the max over f of x equals 1 of y transpose x is equivalent to writing it as fD of y equal to the max mod y transpose x over all x such that f of x equals 1.

Student: So the mod has been done so that we always get the maximum value and never the minimum value.

Professor: No, even if you did not take the mod, you will get the maximum value. So, let me put it this way. Suppose let us do it by contradiction. Suppose these two have two different solutions, you will get a different fD of y, if you solve this problem, instead of solving this problem, then suppose this problem gave you a solution. Let us say for example, the all ones vector is a solution to this problem, it gives you the max of mod of y transpose x.

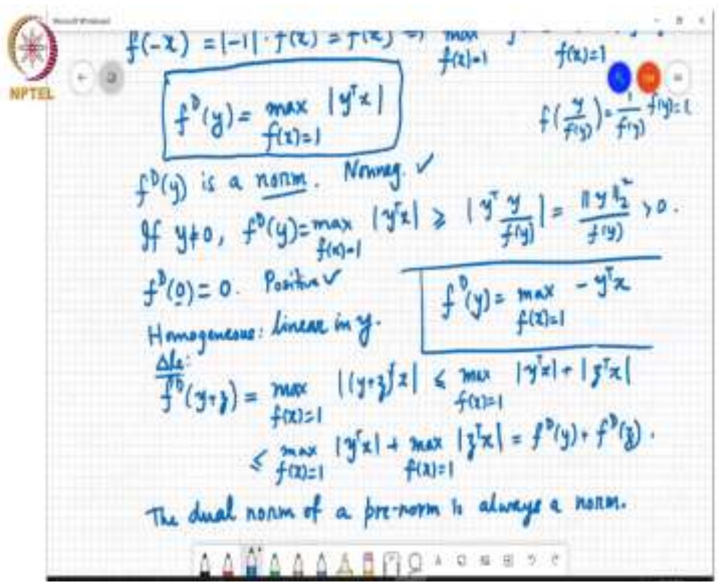
And suppose that happened, because y transpose x for the all ones vector was like, minus 100. And when you took the modulus, you got the got plus 100. And that was the biggest value that the second optimization problem here could take. Then in this problem, I can equivalently use minus all ones vector, and if this was getting a value of minus 100, this will get me a value of plus 100 and that will be the maximum value that this optimization can attain.

So, another way, if you want me to tell it to you in yet another way, this problem that I have written here, is I can also write this as and write it down here. So, I can also write it as fD of y. Like this, because for any x for which this attains the maximum value, if I take minus x, this will attain its maximum value, because minus x also satisfies f of x equals 1. So, I could even write it like this. That is why it is okay to write fD of y to be mod y transpose, maximum of mod y transpose x. So, does that help clarify?

Student: Yes sir, thank you sir.

Professor: So, this, cost function itself is linear in y. So, it is obviously homogenous. The only non-obvious property to show here is the triangle inequality. Remember that f of x itself may not satisfy the triangle inequality. But fD of y does satisfy the triangle inequality.

(Refer Slide Time: 10:55)



So, what we need to show is that if I take fD, if I take fD of y plus z, I need to show that this is less than or equal to fD of y plus fD of z, that is what the triangle inequality says, the norm of x plus y is less than or equal to norm of x plus norm of y. So, I need to show that is this fD of y plus z is less than or equal to fD of y plus fD of z. So, this fD of y plus z, by definition, is the max over all x such that f of x equals 1, mod of y plus z transpose times x.

And I can write this as. So, this is y transpose x plus z transpose x. And if I split this, the mod of the sum of two numbers is at most, the sum of the mod of the two numbers, so I just write that as max over f of x equals 1 of mod y transpose x plus mod z transpose x, all I have done is to split the mod in across the two terms, and that can only increase the value or leave it unchanged, but it cannot decrease the value of the y plus z transpose x the mod of that.

And this in turn is less than or equal to so I am taking the maximum of the sum of two terms, if I individually took the maximum of these two terms and added them up, that will only increase the value because it gives me more flexibility in optimizing this the objective function here. So, this is less than or equal to max over f of x equals 1 mod y transpose x plus the max over f of x equals 1 of mod z transpose x and this is just by definition, this is fD of y and this is fD of z.

Student: Sir.

Professor: Yeah.

Student: Initially, you told that all the properties will be satisfied for pre-norm except the triangle inequality may not satisfy…

Professor: May or may not be satisfied, yes.

Student: So, but you are proving that it will be satisfied.

Professor: No, I am sorry, showing that fD of y satisfies triangle inequality not f of y.

Student: Okay, sorry.

Professor: Yeah, so basically fD of y satisfies all the four properties needed to be called a norm and therefore, the dual norm of a pre-norm is always a norm.