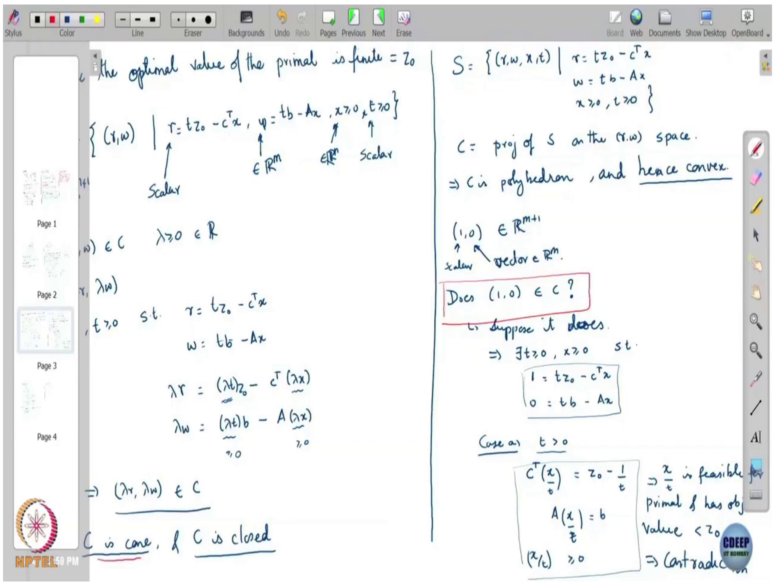
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

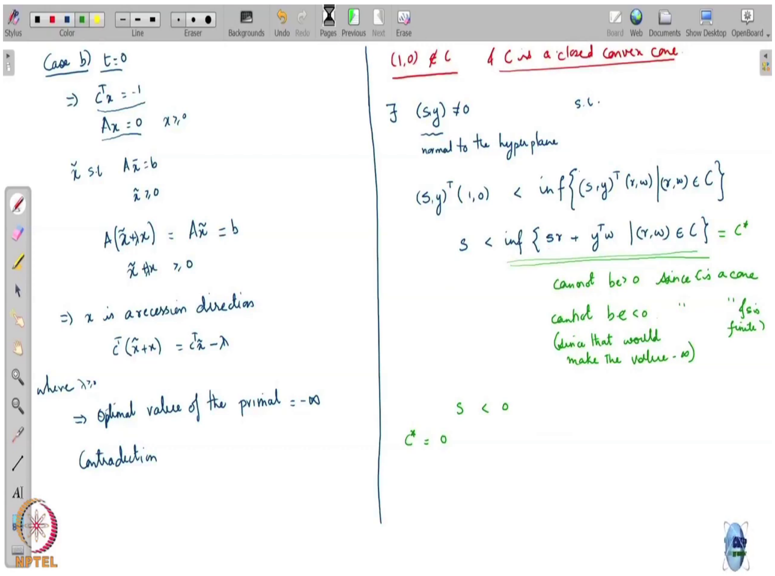
Lecture - 12C
Proof of strong duality (continued)

(Refer Slide Time: 00:25)



Now, go back and let us take a step back and see what where we are. We have a set c which is a closed convex cone and now we have a point 1 comma 0 that lies outside C, ok. And C is a closed convex cone.

(Refer Slide Time: 00:37)



So, C is a closed convex cone 1 comma 0 lies outside C; what does this mean? It means there exists a separating hyperplane; that separates 1 comma 0 from C, right. So, this point can be separated from C, ok. So, which means there exists, there exists a separating hyperplane.

Now, the separating hyperplane remember C is a set in r m plus 1. So, let the hyperplane should also have, its normal should also have m plus 1 coordinate, ok. So, let me write it like this. So, there exists a s comma y not equal to 0, ok. So, that is your that is this is the hyperplane normal, such that s comma y transpose 1 comma 0 is less than the infimum of s comma y transpose r comma w, where r comma w belong to the set C. We write this more neat, a little neatly.

So, s comma y transpose 1 comma 0 right, this is less than the infimum of s comma y transpose r comma w for as r comma w ranges in C.

Now, let us just evaluate this s comma y transpose 1 comma 0 is simply s, that is my left hand side; right hand side is infimum of s into r, because they are both scalars plus y transpose w, these are both vectors of length y right as r comma w belongs to C.

So, now s must be less than this and we know that s comma y as a vector as a whole is not equal to 0; that means it cannot be that it all its components are 0, alright ok alright.

So, now tell me, what can you say about this thing, the green thing that I have just underlined; the infimum of s r plus y transpose w as r comma w ranges over C? Remember C is a closed convex cone, right. So, if r comma w is in C; then a scaled version of it is, positively scaled version of it is also in C.

So, what should be this the value of this, this infimum? This infimum is minimize; what you are doing is effectively minimizing a linear function over a cone right, a linear function, the slope of the linear function is given by s and y. And the point in the cone is r comma w, r comma w ranges in this cone.

So, what can you say about the, about this the infimum here? So, let us suppose it can can it be can it be a positive value, can it be 5, 3, 1.5? If it is any positive value right, then because C is a cone, I can always scale it down and bring it down to 0, right. I can all if I if some r comma w gives me a value 5 suppose; I can take one tenth of r comma w and get a value of 0.5 right, 5 by 10, right. Essentially I can bring the value down as much as I want and eventually bring it to 0.

So, the infimum cannot be any some positive value like this. Can it be a negative value? So, this infimum cannot be positive since C is a cone. Can it be negative, can it be minus 1? Why? Again you can scale it in the opposite direction; you can magnify also, right.

So, if it is minus 1, I can multiply; if an r comma w gives me a value minus 1 right, I can take twice that and get minus 2, 10 times that and get minus 10 and so on and I will not lose feasibility at all. So, I will remain in this cone, right. So, I can scale up and then go to minus infinity.

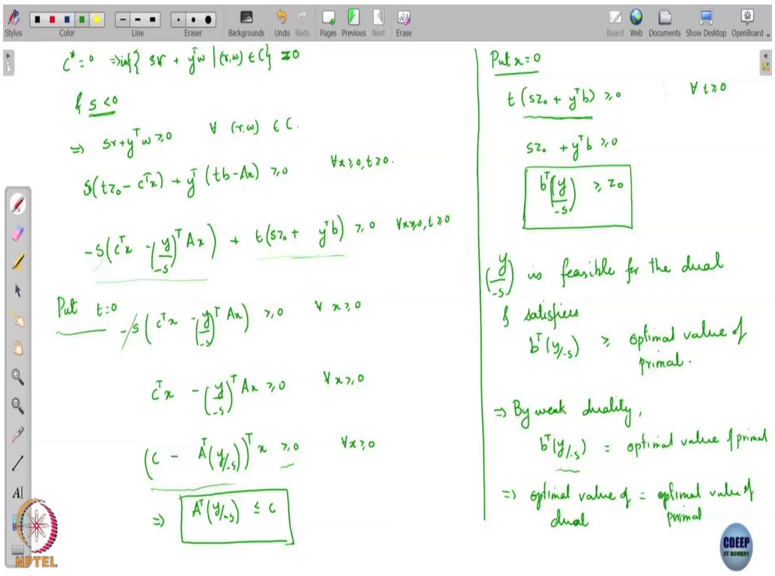
Now, what is the problem with minus infinity? I know that this infimum is definitely greater than some s, where s is some finite value; we have been guaranteed that there exists such an s, right. It is some finite value; I do not know what it is positive, negative whatever, it is definitely some finite value, right.

So, if it is if it is so, the infimum therefore, cannot be minus infinity. So, it cannot be positive. So, infimum if it is positive, then you can you have a contradiction, you can bring it down to 0; positive value cannot be an infimum. If it is negative, then it had to become minus infinity; that is also not possible because of this inequality. So, you know in other words we are left with just one choice; what is the infimum then? The infimum must be 0, right.

So, let me summarize this. So, cannot be negative, again since C is a cone and C is a cone and s is finite, make the value. So, in other words the optimal value has to be 0, right. So, consequently what we have is that s therefore, is strictly less than 0, ok. And the optimal value let us call this here, let us call this optimal value say C star; C star is equal to 0, ok.

Now, the optimal value is 0 and s is strictly less than 0; this is what we have concluded so far.

(Refer Slide Time: 07:59)



Now, since the optimal value is 0 what this means is, let me go to an new page. So, C star is equal to 0, which is means basically infimum of s r plus y transpose w as r comma w ranges in C ok. This the infimum here is equal to 0 ok and we have that s is strictly less than 0.

Now, s r plus y transpose w is it is, the infimum of that is equal to 0; which means s r plus y transpose w is greater than equal to 0 for all r comma w belonging to C. Now, r comma w belongs to C, then go back to the definition of the set C. We define the set C here; this was this is how we defined the set C, r is t z 0 minus C transpose x and w is t b minus A x for some x greater than equal to 0 and t greater than equal to 0, right.

So, let me substitute that, this is greater than equal to. So, I have s r plus y transpose w greater than equal to 0; I just substitute this, I get s times t z 0 minus C transpose x plus y transpose t b minus A x greater than equal to 0, for all x greater than equal to 0 and t greater than equal to 0. Let me rearrange this a little bit.

So, I have just rearranged a few terms here; I picked up the terms that a couple of terms that involve x and put them together here. So, this is C transpose x; there is a there is a minus s which I pulled out outside, I took up, I took y transpose A x ok, there is a minus out minus sign here, that minus sign has come here. And since I pulled a minus s outside, that I have divided by a minus s, right. Just I am writing this specifically in this kind of little non intuitive form for a particular reason.

And then everything else I have put here. So, I have a y transpose b, that was multiplied by T. So, T has come out and I have an s t z 0, which is left here, right. So, all of this is greater than equal to 0 and remember this has to be true for all x greater than equal to 0 and t greater than equal to 0.

Now, this claim here, this inequality has to be greater than equal to 0 for all x greater than equal to 0 and t greater than equals to 0 right; which means I can put my favorite values of x and t and check, right say first. So, I can put say first let us, suppose I put, suppose we put say t equal to 0; what if I put t equal to 0, what do I get?

If I put t, so put t equal to 0 here; then this term disappears, the second term here disappears, because it is been multiplied by t. So, all I am left with is then minus s C transpose x minus y by minus s transpose A x greater than equal to 0, for all x greater than equal to 0.

Now, remember s was negative; s was strictly negative, we just concluded that. So, I can just divide throughout by minus s; minus s is a positive quantity. So, I can just cancel this minus s out and all what I am left with then is, C transpose x minus y by minus s transpose A x greater than equal to 0 and this is true for all x greater than equal to 0, ok.

So, let me again gather some the x terms together. So, this this is saying. So, this has to be greater than equal to 0 for. So, what I have got is C transpose, C minus sorry C minus A transpose y by minus s; the whole thing transpose x greater than equal to 0, for all x greater than equal to 0, right.

Now, this has to be greater than equal to 0 for all x greater than equal to 0. What does this mean? What? So, it is a these I have taking some an inner product of some vector with x and I want that inner product to be greater than equal to 0, whenever x is greater than equal to 0.

So, what sort of component should that inner product, that vector have? Non negative components right; because if any of the components is negative, then what I can do is, make my x very large for that component and 0 for everything else and that will give me a negative number, that would not give me greater than equal to 0, right.

So, the only way this is possible is that, the term in the bracket is less than equal to 0. So, A transpose this is less than equal to C, ok. So, let me just box this, this is my first conclusion; A transpose y by minus s is less than equal to c, alright. Look let us, we got to this by putting t equal to 0; let us now put x equal to 0.

Suppose I put x equal to 0, what do I get? So, now, this term, the first term here is going to disappear; if I put x equal to 0; if the if I put x equals 0, this all these terms are going to all be 0. So, all I am left with is the second term; which means I have t times s z 0 plus y transpose b greater than equal to 0, for all x greater than equal to 0 and t greater than equal to 0, alright ok.

So, now let me once again do the same thing that I did earlier. I know that I know that minus s is positive. So, s is negative, so minus s is positive. So, let me look, let me write it like this; let me write this whole thing in the following way. So, I so, since I put x equal to 0, I do not need the x greater than equal to 0 here; I just need this to be true for all t greater than equal to 0.

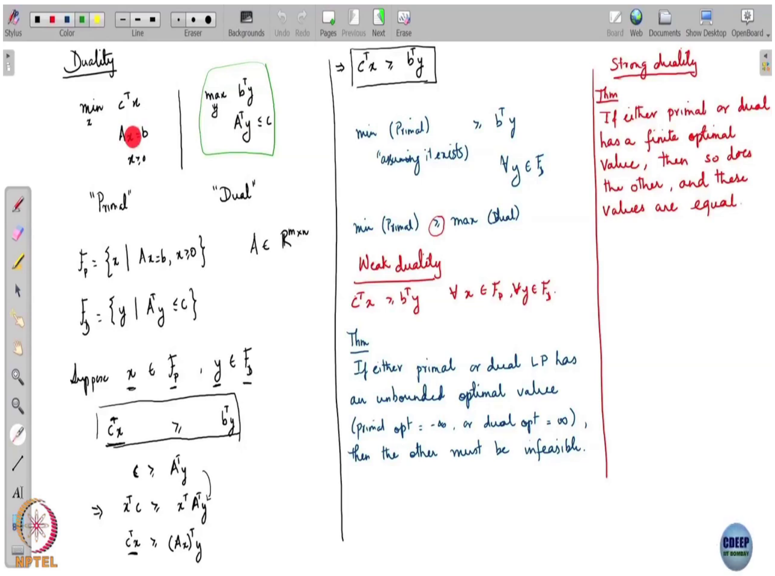
Now, ok in fact we can do a quick analysis here also; since this has to be true for every t greater than equal to 0, it just means. What does this mean? t is greater than equal to 0, this has to be true for every t greater than equal to; say t equal to 100, this must be true for t equal to 100. So, what does this mean?

Yeah, so, it means that, the term that is in the bracket, must be must also be positive right, must also be non-negative; which means s z 0 plus y transpose b is also greater than equal to 0, ok.

So, let me rewrite the whole thing, rewriting it basically gives me b transpose y by minus s is now greater than equal to z 0. I can divide throughout by minus s and thus the inequality does not flip sign; because minus s is positive, right. So, I got b transpose y by minus s is greater than equal to z 0, right.

So, here is my second conclusion. So, now what are, what do these two box conclusions mean? The first conclusion here says that, y by minus s, y by minus s satisfies A transpose y by minus s less than equal to C.

(Refer Slide Time: 17:15)



Which means y by minus s is feasible for my dual; y by minus s is a feasible point for the dual, right. And the second one says that, the optimal value the sorry, the value of y by minus s; the objective value is greater than equal to z 0. And what was z 0? z 0 was the optimal value of the primal. So, which means I have found a feasible point for the dual, ok.

So, y by minus s is feasible for the dual and satisfies b transpose y by minus s greater than equal to optimal value of primal. But then what did we know from weak duality? Weak duality told us that, this inequality is actually the opposite direction, right.

So, weak duality told us that b transpose y has to be less than equal to the optimal value of the primal for every feasible y. So, now, we have found a new feasible, a feasible point y by minus s, such that its optimal its value is greater than equal to the optimal value of the prime.

So, which means we have got, which means the only way this is possible is that there is equality throughout right; which means that now by weak duality, it is the optimal value of the primer. And in fact, this is this or this automatically shows that y by minus s must be the optimal, must be the optimal solution of the dual also right; so which means in fact that, the optimal value of the dual is equal to the optimal value of the prime.

So, it has shown that, the optimal value is attained by y by minus s that is the point; that attains it, we know that it exists, we know that, we also know that the optimal value of it is must be equal to that of the prime, right.

So, weak duality gave us the inequality in one direction, the other inequality came by the in through this theorem basically; and what was the key idea we used in it? Separating hyperplanes right, we have the existence of a separating hyperplane.

We basically constructed a set C and show and strategically chose a point outside it and showed that you can separate the two; using the fact that you can they can be separated, we got our result.

Now, you might wonder obviously, how did someone think of this particular set, how did you think of that particular point etcetera; that we will build the intuition for that later in this course, in the following lectures also. The basic intuition for it is to remember that, this set C is actually nothing but the projection of this set s; the set s that I have shown here.

And what does this set s have? The set s has a representation here of the feasible points for the primal as well as the feasible points for the dual. You can say for, effectively it is it has a x which captures the feasible point of the primal and it has a w which captures a variable for the constraints, one for each constraints.

And how many variables are how many variables are there in the dual? You go back and see the variables of there are as many variables in the dual as there are constraints in the primal, every constraint in the primal gets one dual variable and every constraint you can say in the dual gets one primal variable, right.

So, this is obviously reminding you of something that you have already seen, which is that the dual must be have some relation with Lagrange multipliers. You have seen Lagrange multipliers as variables that you have one per constraint right, in effectively that is what, that is what we are seeing here. So, in fact the optimal value of y is actually the optimal value of Lagrange multipliers for this particular problem.

Now, I have not taught you Lagrange multipliers as yet for inequality constraints and so on.

So, we will come to that, that will become evident very quickly; but essentially the main reason why this, the thinking behind this is that is to capture the problem in both the space of the primal as well as the dual, means both in the space of the decision variables as well as in the space of the constraints. And that is the, so you can there are many different ways in which such sets can be constructed and they all lead you to duality type theorems like this, ok.

So, the correct tension that is that, that is present in an optimization problem is always the tension between not multiple decisions that you have to make; but rather the decision between, the tension between the decision variables and the constraints. And are essentially the Lagrange multipliers, and that is being captured by this particular set. So, you will see more of this as we go later in this course, ok.

So, to summarize essentially what we have shown is that, you know the first thing which is weak duality that is a claim that holds for free; I mean almost for free, it just comes by the way this is constructed, you can immediately argue that. And, but moreover there is also this claim that, if a primal has an optimal solution; then so does the dual and the optimal values are equal.

Now, this also means there are the, this also has computational implications; for example, it may be that primal looks much harder to solve than the dual, does not matter you can solve either and you will get the same answer, right. So, it has it is it that is one of the motivations behind algorithms also, ok.

I will in the next class, I will give you another application of strong duality ok, that arises in network flow problems, ok. So, we will end.