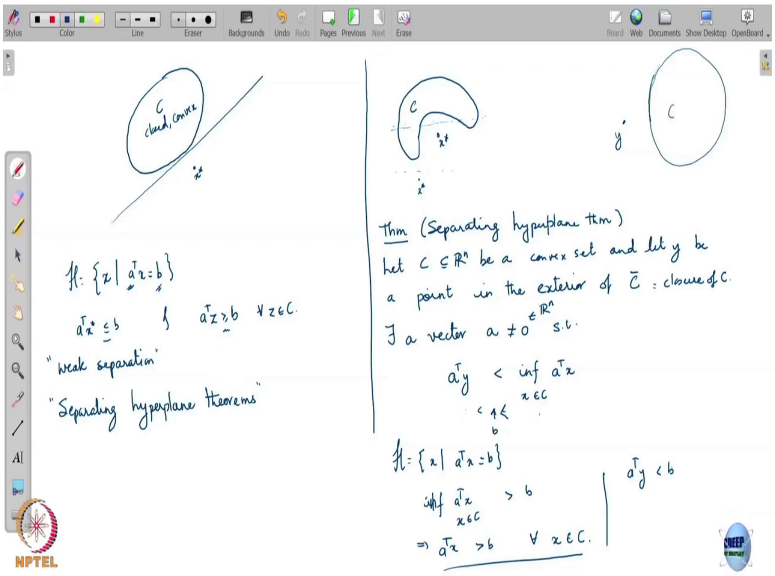
Optimization from Fundamentals

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

Lecture - 12B
Proof of Strong Duality

Ok. So I will state the separating hyperplane theorem that we need.

(Refer Slide Time: 00:21)



So, the theorem is the separating hyperplane theorem. Let C subset of R n be a convex set and let y be a point in the exterior of C bar; C bar is simply the closure of C. So, I am allowing the set to be not closed; but let us take the closure of the set and then consider a point y that lies outside. So, here you have a set C, which is probably not closed; but then you take it is closure, that closure will the closure is also a convex set.

And then you take a point y ok take a point y that lies outside the closure. Then the claim is there exists a vector a not equal to 0, such that a transpose y is less than; if I take the infimum of a transpose x as x ranges over C.

So, there exists a vector a not equal to 0, such that a transpose y. So, this is a vector in R n, such that a transpose y is less than the infimum of a transpose x as x ranges over C. Now, what, how does this define a separating hyperplane?

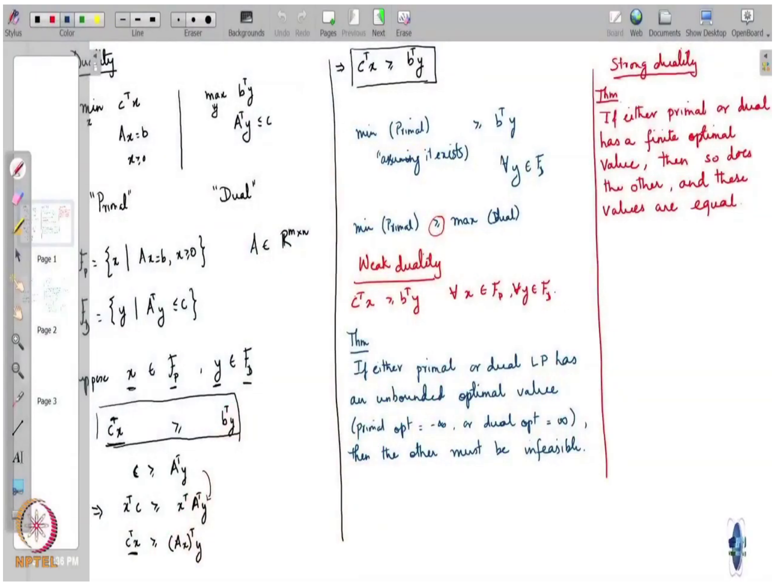
So, for that the normal to the hyperplane is the vector a, a is the normal of the hyperplane. Where is my b the b? Where is, for a hyperplane I need to define, I need to talk of a as a b right, a scalar b yes. So, I can I have that these two values are one, the left one is strictly less than the right one, right. So, I can always fit in a, take a value b that lies in between them. So, this is right a value b that is greater than a transpose y; but at the same time less than the infimum of a transpose x over x in c.

So, look at a consider a value b that lies in between; then clearly and then look at the hyperplane H defined by such that a transpose x is equal to b. Then in that case, y lies on one side of the hyperplane and the entire set C lies on the other side. Why does the entire set C lie on the other side? Because we know that infimum of a transpose x, x in C is greater than b; which means that a transpose x is greater than b for all x in C.

So, the entire set C lies on one side of the hyperplane and on the other on and as far as y is concerned, we know a transpose y is less than b. So, it lies on the other side, right. So, this theorem basically guarantees for you implicitly a separating hyperplane; it is giving you that, it is giving you the slope as well as as well as the implicitly through this is also letting you pick an intercept.

You can pick any intercept is ranging from anywhere in between here the, that is that is perfectly ok, alright ok. So, now using this we will be proving the strong duality theorem, ok.

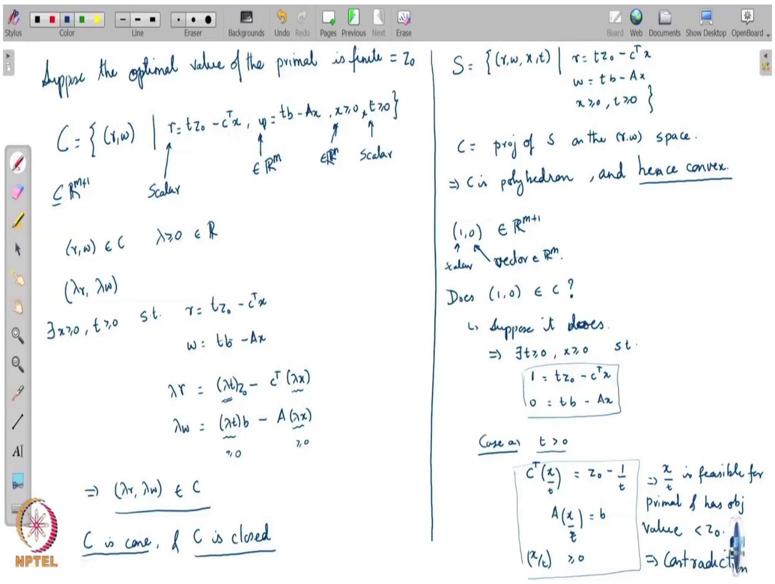
(Refer Slide Time: 05:13)



So, now, what are we trying to, what are we going to prove; let us go back to the, let us go back to the statement of the theorem, we want to prove this red statement here. The statement says that, if either primal or dual has a finite optimal value; then so does the other and these values are equal, ok.

The other theorem, other statement which is that; if either primal or dual is unbounded, then the other must be infeasible, that is trivial we just, we argued that directly by looking at weak duality, ok. So, now we are going to look at the strong duality theorem.

(Refer Slide Time: 05:42)



So, suppose the optimal value of the primal is finite and equal to say some Z 0, ok. Let me just denote that by Z 0, that is the optimal value of the primal, ok. Now, I am going to define the set, the set C as follows; it is the set of r comma w such that r is equal to t Z 0 minus c transpose x. And w is equal to t b minus A x, as and I allow x to the to be greater than equal to 0 and t to be greater than equal to 0. So, C is this set of r comma w that satisfy this, ok.

So, it is the set of r comma w that can be expressed in the following way, it can r can be expressed as t Z 0 minus c transpose x for x greater than equal to 0 and t greater than equal to 0. And w can be expressed as t b minus A x, for x greater than equal to 0 and t greater than equal to 0, ok. For the same x and t greater than equal to 0, your r can be expressed in this way and w can be expressed in this way; this is the set c, alright.

Now, what is the, what kind of a set is this set C? If you look at this set C, how can you tell me what is its dimension; how many in what space does it lie? What is the dimension of r and what is the dimension of w? So, r here is; is r a scalar or a vector? Yes, r is a scalar; r is t Z 0 minus c transpose x, ok, t here is has to be a scalar, otherwise I will not be able to add it to c transpose x right; c transpose x is a scalar, so and Z 0 is another scalar, so t is also a scalar, ok.

So, t is a scalar. So, here t is a scalar; x is a vector, x is a vector or your original vector this is R in the original space. So, x is a vector in R n. So, r therefore, is a scalar. What about w, what is the dimension of w? Yeah w has the length of w is as many as the rows of A, right.

So, the w is in R m, ok. So, C itself is a subset of R m plus 1, alright ok. Now, so now, let us ask what type of, what kind of a set is C? So, suppose r comma w belongs to c ok, and I give you a lambda greater than equal to 0.

Then what can you say about lambda w comma lambda, lambda r comma lambda w? So, if r comma w belong to C; that means r it just means that, there exists x greater than equal to 0, t greater than equal to 0, such that r is equal to t Z 0 minus c transpose x, and w is equal to t b minus A x.

Now, if I multiply both sides by lambda in both equations; then what would I get? I would get lambda r equals lambda t Z 0 minus C transpose lambda x, and lambda w as lambda t into b minus A into lambda x, lambda is a scalar, right. So, what this means is, if I have if lambda, if r comma w belongs to c and lambda is some scalar greater than equal to 0.

Then lambda r comma lambda w can be are also in c and the corresponding value of t and x is just lambda times the earlier erstwhile t and the lambda times the erstwhile A, ok. So, by, so these are also greater than equal to 0. So, which means lambda r comma lambda w belong to C.

Now, if a set C is of this kind, where if a, if it is if r comma w belongs to C; then lambda r lambda w belongs to C for lambda greater than equal to 0, then what kind of a set is this? It is a cone ok, so C is a cone.

It is also very easy to show that, C is actually a closed cone. So, in fact C is a cone and C is closed, C is a closed cone. What about the convexity of C? Is C convex? It is convex; you see C is defined using a bunch of linear equations and inequalities, right.

So, it is just r comma w that satisfies some linear equations; you can think of it this way, you can think of a set comprising where the variables are its, you can think of a set s like this, which is r, w, x, and t such that r is equal to t Z 0 minus C transpose x, w is equal to t b minus A x and x is greater than equal to 0 and t greater than equal to 0.

If you look at a set s like this, the projection of this set on to the r, w, on the r, w axis is my set c, right. So, if I take the shadow of this set on the r, w space that set is actually C; C is projection of s on the r comma w, so on the r, w space.

And what kind of a set is S? Well S is actually a polyhedron, it is just a bunch of linear; it is a vector that satisfies some linear inequalities right, it is a set of vectors that satisfy linear inequalities. So, S is a, S is itself a polyhedron. So, the shadow of a polyhedron has to be a polynomial, right.

So, this is a projection. So, therefore, C is a polyhedron and hence convex. So, many proofs in optimization begin with this one question first; what kind of a set is this, right? You encounter a set, you ask what kind of a set is this right; more you can say about the set, more you will be able to say about the problem, alright. So, now, what have we concluded? Well that C is a cone, C is convex closed and C is also convex, alright. So, what we have is a closed convex cone, alright.

Now, let us look at this point, let us say a point like this; say a point 1 comma 0. Is this 1 comma 0, what I mean by that is; so one here is a scalar is a scalar and 0 is the vector. So, this is a vector ok; this is a point in R m plus 1, right. So, this vector is a 0 vector in R m. Now, this point can you, does this question; does this, does 1 comma 0 lie in C, is this 1 comma 0 out in C, right?

So, let us check. So, suppose it does ok, suppose it does, then means that, there exists t greater than equal to 0 and x greater than equal to 0, such that your 1 which is in the plane. So, your r equals 1 and w equals 0 now. So, r, that means 1 is t Z 0 minus c transpose x and w equal to 0 means that 0 is equal to t b minus A x, right. Now, here again; so if 1 comma 0 belongs to c, then it means that you can find t greater than equal to 0 and x greater than equal to 0, that means that satisfy these equations.

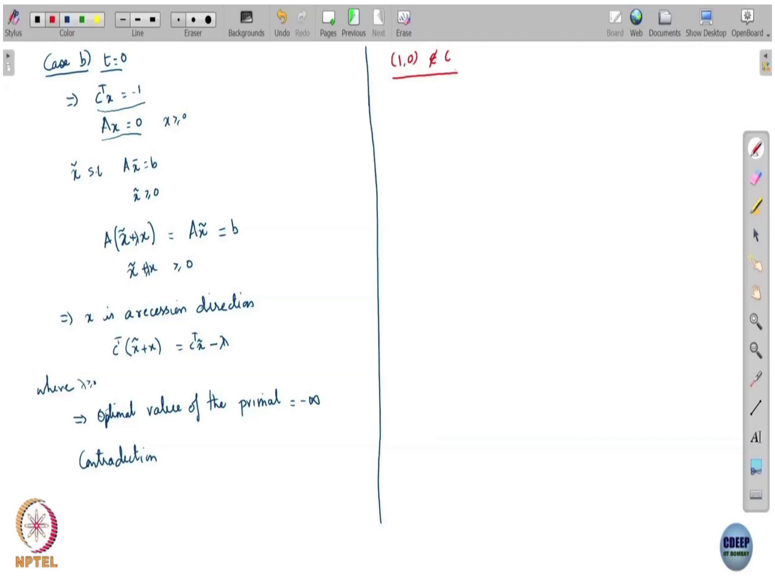
Now, let us take two cases. So, case one. So, case a suppose is that, here the t that you got is positive, ok. If the t that you got is positive, if you, if the t that you got is positive; then I can divide throughout by t and I get that, I get that 1 by t minus z. So, I get, let me write this again; I get that C transpose x by t is equal to z 0 minus 1 by t. I can I and the second equation if I divide throughout by t, I get, also that x A, A into x by t is equal to b.

Now, I already have x greater than equal to 0, or t greater than equal to 0; this means that, x by t is greater than equal to 0. Now, what has, what have I got now as a result in this case? In case a, if t is positive; then I have concluded that I these equations must hold. But what are these equations? This is saying x by t satisfies A into x by t equals b and x by t greater than equal to 0 and the.

So, x by t is therefore feasible, x by t is feasible for the primal; but then what is the, what is the value of x by t? It is Z's naught minus 1 by t; that means it strictly, its value is strictly less than Z naught, right. So, it is a feasible value, the feasible point for the primal whose value is less than Z naught. And what was Z naught here? Z naught we had said that, Z naught is the optimal value of the prime, right.

So, what this means is, if t is greater than 0; we have effectively constructed a point, whose value is even better than the optimal value that is a contradiction, right. So, x by t is feasible for the primal and has objective value strictly less than Z naught, which is a contradiction, alright. So, case a therefore is impossible; means you cannot have t to be strictly positive, ok.

(Refer Slide Time: 19:08)



Now, let us go to the other case, case b. Case b would be, since t is greater than equal to 0; the only case remaining is t equal to 0. Now, if t has, if t is equal to 0; which means that these boxed equations here, these boxed equations must are hold with t equal to 0, right. So, if I what do I get if I put t equal to 0? I get C transpose x equal to minus 1 right and I get a x equals 0, right. So, this implies C transpose x equals minus 1 and A x equals 0.

Now, what does this mean ok? So, remember x and x and t ok, x and t are simply, you simply have x greater than equal to 0 and t greater than equal to 0. The x is here is not necessarily feasible for the primal; it is just some vector of length n, we are not saying A x equals b as well.

Here we constructed an x by t which became feasible, right. Now, if A x equals 0 and C transpose x equals minus 1. What do we get? What is this saying? If A x equals 0, it means that x is in the null space of A, right.

So, we have found a direction ok, such that it if I go in that null in that direction my A, A times that direction will not change, right. So, what this means is. So, if I have say any feasible value. So, if I have say an x tilde, such that A x tilde equals b and x tilde is greater than equal to 0; then x plus x sorry, x tilde plus x would also satisfy this. Because x tilde plus x would give me still, would still be A x tilde and x tilde plus x would also be greater than equal to 0.

Now, what does this; remember what we used to, we had a name for this, a direction along which you can keep going by without leaving the set, this is a recession direction, right.

So, what this means is, your x here is a recession direction right; x is also remember greater than equal to 0. So, it is greater than equal to 0, it satisfies A x equal to 0. So, x is means x is a recession direction or a ray of the set right; it is a direction along which you can keep travelling and not good.

But then every time I travel one unit of distance along x, look at what is happening to my objective. I keep get incurring my C transpose x, goes down by minus 1, goes down by 1, right. So, if I take this point x tilde ok, and look at C and add an x to it, where x plus x tilde is still in the set.

But what about C transpose x plus x tilde? C transpose x plus x tilde is now C transpose x tilde minus 1, right. So, I can keep go. So, what and now if I. So, this is because I took a unit step in the direction x; if I took a step of length lambda in the direction x, I would still be within the set and my objective value would have dipped by lambda, lambda greater than equal to 0, right.

Eventually I can go along, I can make my lambda as large as I want and that will drive my optimal value as low as I want, right. So, this is again a contradiction; contradicts what? Contradicts the assumption that the primal has a finite optimal value, right. So, this so optimal value of the primal is minus infinity, which is a contradiction.

Contradiction since we assume that the primal must have a as a finite optimal value; which means again case b is also not possible, so right. So, what have we, what so far we concluded case a is not possible; because in case a, you get something better than the optimal, which means you are contradicting optimality.

In case b, we are getting something that the optimal value is minus infinity, which hence we are, where we are contradicting that the optimal solution exists right, it is that it is finite, alright.

So, in summary what we have shown is after all this, what we have basically shown is this, which I am; let me put this here. We ask we began through this analysis asking, does the point 1 comma 0 belong to C. And what have we shown; we have considered both cases and shown that, no, it cannot belong to C, right. So, the point what we have effectively shown is the 1 comma 0 does not lie in C, this is what we have shown.