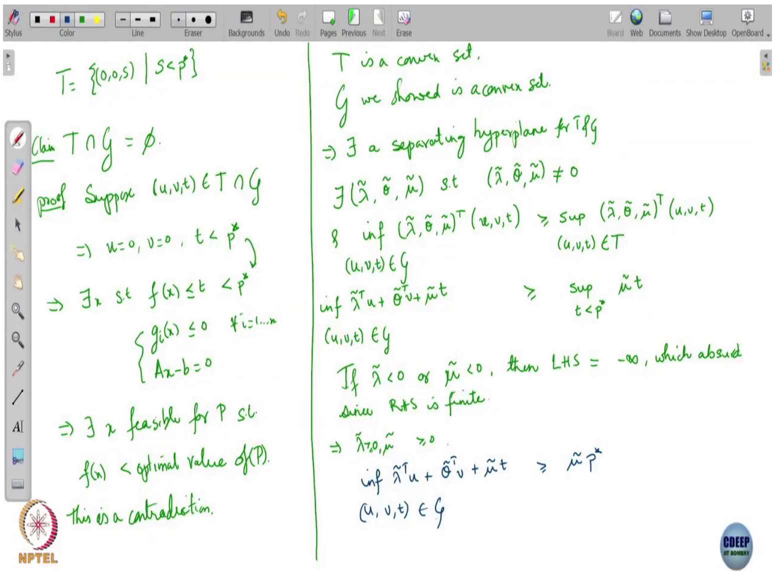
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

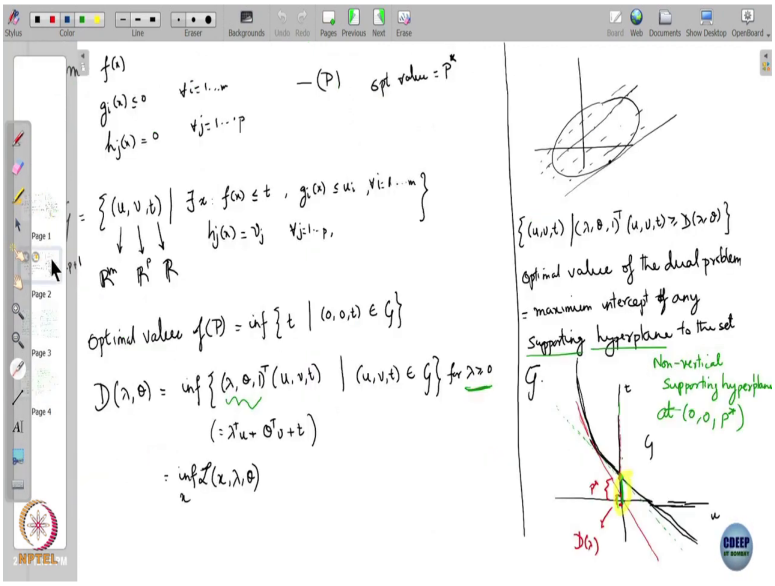
Lecture - 18C
Strong duality in convex optimization – III

(Refer Slide Time: 00:16)



Now let us look at; let us look at this even further let us go a little bit more into depth here. Now, suppose now what is this particular expression here on the left? What is this particular expression? The infimum of this linear function evaluated over T.

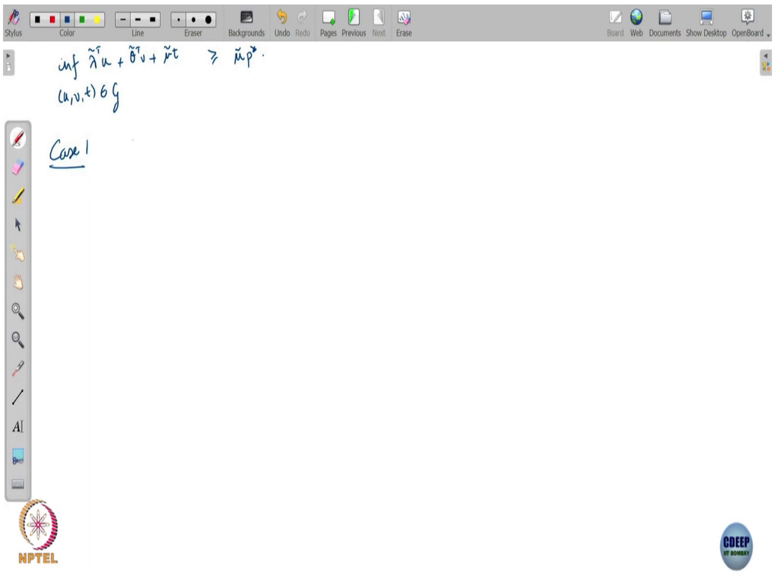
(Refer Slide Time: 00:45)



Now, if you recall I had discussed this a little bit earlier right. If you evaluate a linear function over G right. And when lambda is greater than equal to 0 ok alright and when this quantity here is 1 right, then that linear function was actually nothing but the minimizing that linear function was actually the same the minimizing that linear function was actually the same as minimizing the Lagrangian over the entire space.

So, its value was actually the value of your dual function. So, now, I want to be a I want to try and relate this relate this in that way. So, so what I have on my if I if you if we just take this here. So, I have infimum of let us write this out here once again.

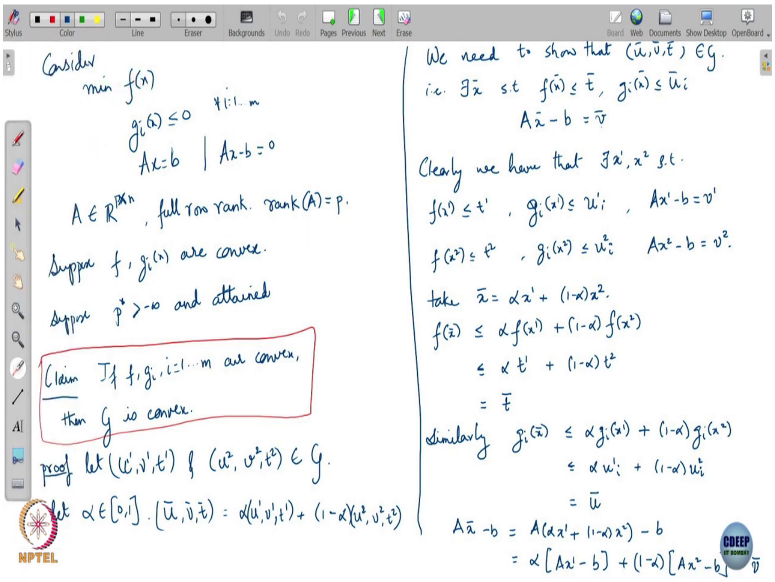
(Refer Slide Time: 01:37)



So, infimum of lambda tilde transpose u plus theta tilde transpose v plus mu tilde into t over u, v, t in G. This is greater than equal to for us mu tilde into p star this is where we are doing ok. Now suppose just suppose or you can say let us take two cases ok. So, case 1 case 1 is when ok. So, actually. So, in fact, before I take before I take these two cases let us come let us look at this once more.

So, what we have done is, we have said that well T and G ok they are convex sets and we applied the separating hyper plane theorem and that has got us to this stage that has got us to this particular inequality here ok.

(Refer Slide Time: 02:52)

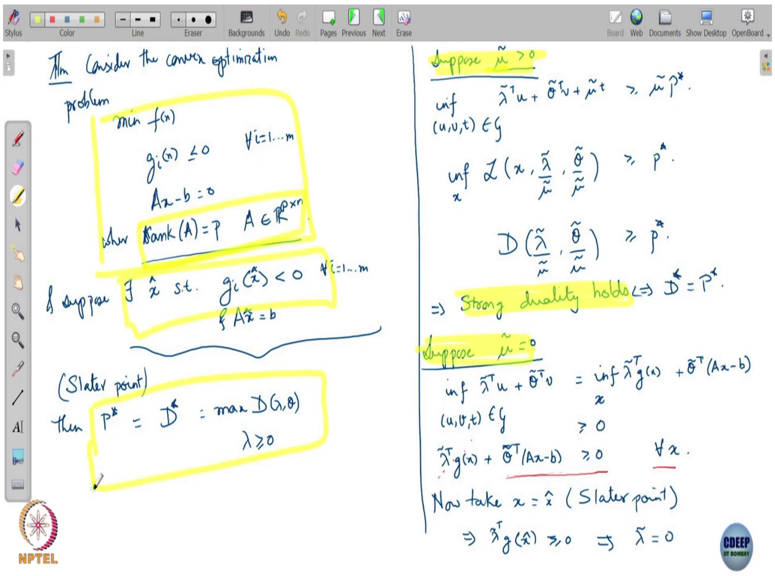


So, now, we have so, far what have we; what have we assumed about the optimization problem? We have assumed that well you have this optimization problem where A is full row rank g has a convex and this is attained.

And from there we said we just we only showed that well G is G must be a convex set t must be a convex set G and t are and we said that there must be a separating hyper plane that separate G and t. I have not yet made any assumption about constraint qualifications ok. So, now, what I will do is I will actually introduce a constraint qualification ok.

So, what I will; what I will do is ok. So, we have reached this stage ok. So, now, before we move any further.

(Refer Slide Time: 03:42)



So, let me erase this, before I move any further now suppose a ok. So, this is going to be my theorem. Consider the convex optimization above convex optimization problem.

So, A is P cross n rank of A is P and suppose there exists an x hat such that. So, we want a slater point right. So, there exists an x hat such that all these g i of x hat is strictly less than 0 and Ax hat is equal to b ok. So, this sort of point is what is called a slater point ok. So, consider the convex optimization this and suppose there exists an x hat such that g i of x hat is strictly less than 0 for all i from one to m and Ax hat is equal to b ok.

So, this is what is called suppose there exists a slater point and this condition is what is called as slater condition ok then P star is equal to D star and what is D star? D star is simply the max of D of lambda comma theta over lambda greater than equal to 0. So, it is the. So, at P star. So, the optimal value of the primal is equal to the optimal value of the dual ok.

So, that is what we are now we are setting up to show. So, we came up till this stage we wrote this set we wrote this set G, we wrote this set T and we found that these two sets are disjoint and then therefore, there is a separating hyper plane and the separating hyper plane gave us gave us that there exists this sort of quantities lambda tilde mu tilde etcetera such that lambda tilde is greater than equal to 0 mu tilde is greater than equal to 0 and this inequality holds.

Now from here onwards I will need I will have to invoke my slater I will have to invoke that we have a slater point ok. So, now, suppose here ok suppose out here mu tilde is positive ok. Now, suppose mu tilde is positive or let us say we are taking the case where mu tilde is positive. So, if mu tilde is positive, then what can I; what can I look do with this expression? This expression let me just write this expression once again there on the next slide.

So, I have an I have this expression infimum over u, v t in G, lambda tilde transpose u plus theta tilde transpose v plus mu tilde times t this is greater than equal to mu tilde times P star that is what we have so, far. Now, suppose mu tilde is positive and I have been I have that lambda tilde is greater than equal to 0 already. So, this is actually nothing, but the infimum this is nothing but the infimum of the Lagrangian evaluated at. So, what I can do is the following.

So, if mu tilde is greater than 0, I can just divide both all sides by mu tilde and then I would get that this is nothing but the infimum of the Lagrangian evaluated at lambda tilde divided by mu tilde and theta tilde divided by mu tilde and that would be greater than equal to P star and what is this infimum? How did I get this?

This is because of what we just wrote here that we just we wrote that D the dual function is actually equal to the infimum of a linear when lambda is greater than equal to 0 and this coefficient here is 1, the dual function is actually equal to the minimum of this linear function over this entire over this set G and if you see that is exactly what this was a minimum of this linear function over the set G.

Lambda tilde was greater than equal to 0 we did not have we did not have a coefficient 1 here, but that coefficient can be made 1 by just dividing throughout by mu tilde and that you can do because mu tilde is actually mu we assume mu tilde is greater than 0.

So, when mu tilde is greater than 0 this becomes the infimum of the Lagrangian which is simply D of lambda tilde divided by mu tilde theta tilde divided by mu tilde and that is greater than equal to P star and what is that mean? What is this means that? The this means that we have got the inequality which is opposite of that of the equality.

So, which means this implies strong duality holds equivalently D star equals P star. So, when mu tilde is positive effectively what we have done is we have automatically found a non vertical supporting hyper plane and that is given us this particular direction. Now, for the when mu tilde is now suppose mu tilde is equal to 0 see mu tilde is greater than equal to 0 we take care of one case where mu tilde was strictly positive and there we got that strong duality holds.

Now, when mu tilde is equal to 0 there should be we should be able to rule out that we should be able to still show that either still show that strong duality holds or rule out that this case is not possible ok. Still until this point we have not yet made use of the slater condition we have only use convexity. So, now, we are going to actually going to need this we are actually going to need slater condition ok. So, let us look at that this way.

So, now, suppose mu tilde is equal to 0, then what happens to my right hand side? My right hand side here is becomes equal to 0 when mu tilde is equal to 0 ok. So, if my right hand side is equal to 0, then I am left with and on the left hand side my mu tilde times T that term is also equal to 0.

So, when mu tilde is equal to 0 what I have left with is something like this lambda tilde transpose u plus theta tilde transpose v infimum of this where u v t belonging to G. Now, this here I can again since lambda tilde is greater than equal to 0 and v should be equal to Ax minus b because of u, v, t belonging to G this in fact, can be written in this way infimum of lambda tilde times this itself.

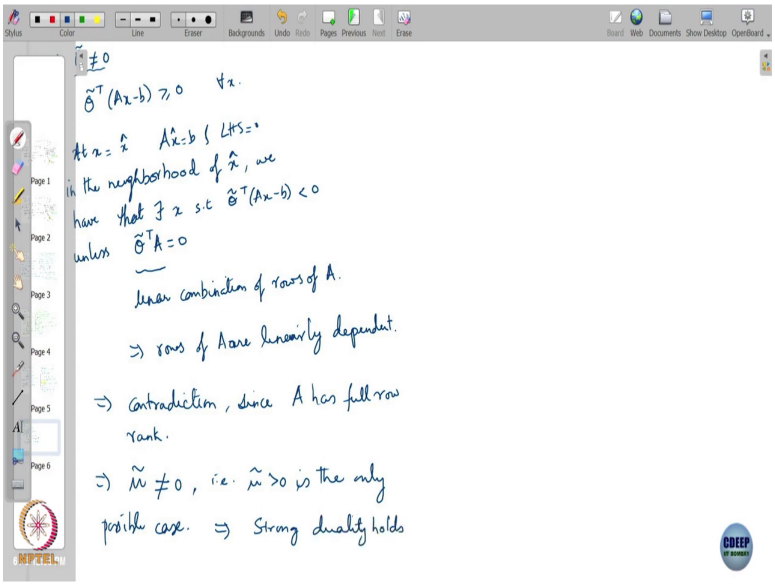
So, the least value of u will be get attained when u that u is actually equal to g of x. So, its infimum of this into lambda tilde transpose g of x plus theta tilde transpose Ax minus b ok and over all the infimum is over all x ok. Now, let us take x to be a slater point ok. So, suppose this yeah suppose when ok sorry this is equal to this and my right hand side on the other hand is my right hand side here remember this quantity evaluates 0.

So, this whole quantity is what I get is that this whole quantity is greater than equal to 0. In other words, I am getting that lambda tilde transpose g of x plus theta tilde transpose Ax minus b is greater than equal to 0 for all x. Now, take x equal to x hat which was my slater point ok. So, I had this x hat here which was my slater point. So, let us take x equal to x hat.

So, what would we get from there? We would get that g i of g of x hat remember would be strictly negative and Ax hat would be equal to b right. So, if g of g i of x hat is strictly negative and Ax hat would be equal to b and Ax hat is equal to b. So, if I put x as the slater point that would imply that the only way possible is so, that would give me. So, we have a tilde.

So, lambda tilde g of x hat greater than equal to 0 and the only way that is possible is that lambda tilde is equal to 0. So, not only is mu tilde equal to 0, I have now also got lambda tilde is equal to 0 alright. You know we are now almost at the end. So, if mu tilde is equal to 0 lambda tilde is equal to 0 the only the only thing the only way remember the separating hyper plane theorem told us that there is a non-zeros what did the separating hyper plane theorem tell us that we told us this should be non-zero right.

(Refer Slide Time: 15:17)



So, the only way this slope should be non-zero is that now is that theta tilde is not equal to 0 write this more neatly. So, only the only way that is possible is that theta tilde is not 0 now theta tilde is not 0. So, let us what does this mean? But anyway we have still we have got that lambda tilde is equal to 0.

So, this term is. So, this term is now gone it means now that with theta tilde not equal to 0 we have we must have that this expression is greater than equal to 0 for all x ok. So, let us go back there. So, which means that theta tilde transpose Ax minus b is greater than equal to 0 for all x ok.

Now, if you take. So, let us think about it this way. So, here is Ax minus b ok a at x at x equal to x hat ok at x equal to x hat which you have Ax equals b. Now; that means, that is a point that is exactly on the intersection of all of these; all of these hyper planes right. So, in this in the neighborhood of this point you should be able to find other there exist other points. Where the you should be able to find other points for which for you know whatever slope you take you should be able to find points for which the inequality actually gets reversed.

See its this quantity is great in a neighborhood of the slater point there would exist points like this points where this inequality gets reversed unless see because this is after all a linear function this is a linear function. So, at and at x equal to x hat this the left hand side is becoming exactly 0. So, you should be able to find other points where in the in the neighborhood of x hat where it also becomes negative right of because this is the linear scalar function.

So, since it is the only way it is not that it that a linear that this function will not become negative is that its slope is actually 0 ok. At x equal to x hat you have Ax hat equals b and LHS equals 0 in the neighborhood in the neighborhood of x hat we have we have that there exist x such that theta tilde transpose Ax minus b is less than 0 unless it's the slope itself is 0 unless theta tilde transpose A is equal to 0. Now, so, in that case then you will not be able to change x and get to a you know you will not be able to vary x and get to a negative value.

So, unless it is a constant function the linear function around that point would take values that are both positive and negative. So, now, so and it will be a constant function only if the coefficient itself is 0, but then can this coefficient actually be 0? What is this; what is this actually saying? Well for this coefficient to become equal to 0.

What we are saying is that theta tilde transpose A is equal to 0 which means that if you take the rows the rows of A and you sum them up using some linear combination which is with these non-zero weights with a way with a bunch of weights that are not all 0 ok. So, you take a linear. So, this is actually the left hand side is a linear combination of rows of A.

And that linear combination we are saying is equal to 0 well; that means, that what does this mean? That means that the rows of A are linearly dependent. Now, when so, but then we if you recall here what did we just assume? We assume that the rank of A is equal to P; that means, it has full row rank which means which is a contradiction to that.

Contradiction since A has full row rank. So, the only way this is possible is that A has linearly dependent rows, but we assume that A has its full row rank. So, the problem we are starting with is one where A has full row rank. So, which means that this is also not possible ok.

So, what does this mean? What this has got us to is that this case right. This case here where we said suppose mu is equal to 0 this case is not possible. So, which means mu tilde is not equal to 0 which means the only case that is the that is mu tilde has to be greater than 0 right and is the only possible case.

And then which means that strong duality holds ok. So, the only possible case is this case where mu tilde is where mu tilde is positive and in that case we concluded that strong duality holds and that is the proof. So, to summarize we have a convex optimization problem as written here with the following requirements that the constraints the linear constraints have are full row rank and there exists a slater point for the constraints right.

You put these together then it has to be that the optimal value of the primal is equal to the optimal value of the dual that is what that is what this theorem as shown ok and improving this basically up to the point we got to this got to the point where we wrote this set G we wrote this set T.

We showed that the sets T and G are both convex and then we showed that there is there say must be a separating hyper plane. Now, the problem from there onwards was that we wanted to show that the existence of a non vertical separating hyper plane which means that we wanted that this mu tilde is the mu tilde here has to be basically has to be positive ok.

And that this slope here is greater than equal to 0. The greater than equal to 0 came quite easily, but in showing that the separating hyper plane is non vertical; that means, this mu tilde has to be positive that is where we needed that is where we needed a constraint qualification. So, until that point everything works without constraint qualifications.

The constraint qualification ensures that your slope is that your separating hyper plane is actually a non-vertical one and that is what actually get makes your makes your strong duality work ok. So, with this then we have now shown strong duality also for convex optimization.

If you will recall that the slater condition which is the constraint qualification for ensuring strong duality that condition is was also what gave us was also used in KKT conditions. So, all these things come together when you actually solve an optimal when we when strong duality holds.

The optimal values of these Lagrange multipliers are also optimal values of your also what will the optimal values of the dual variables are also what will solve your KKT conditions as Lagrange multipliers ok. So, this. So, with this I will end here we will next move on to a new topic.