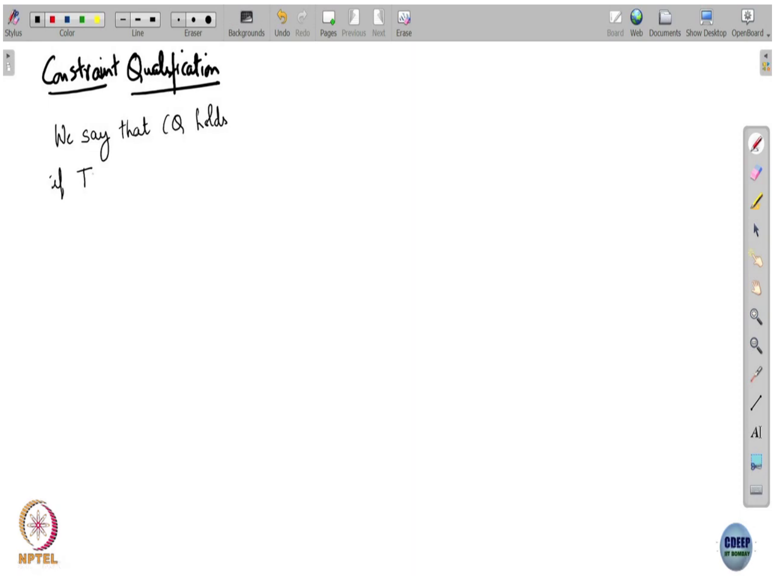
Optimization from Fundamentals

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

Lecture - 17A
Slater condition and Lagrangian Dual

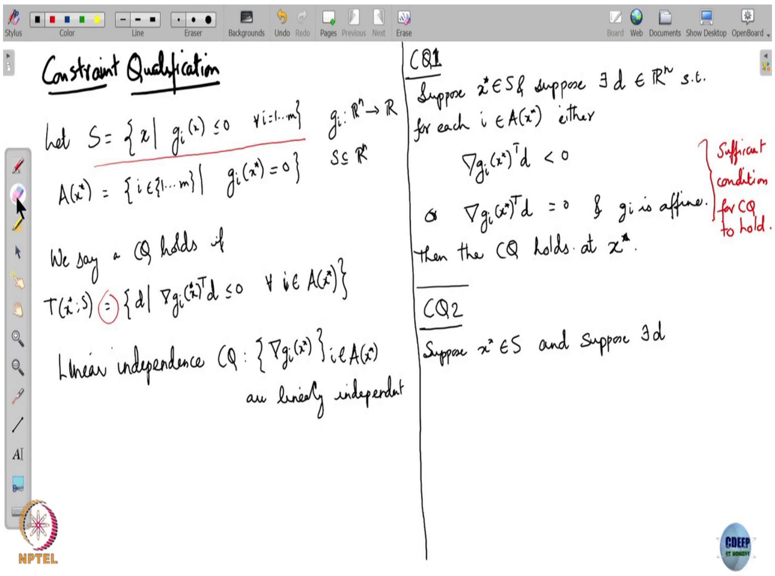
Ok. Welcome, everyone. So, today, what I will talk about is a little bit more about constrained qualifications and then a little bit and then we will move on to the theory of duality for convex optimization. Now, you may you will recall that we had I told you about what is in an abstract sense what is called a constraint qualification.

(Refer Slide Time: 00:41)



So, a constraint qualification if you recall I had mentioned that we say that a constraint qualification holds. If the tangent if the tangent cone can be described in terms of the constraint in terms of the gradients of the constraint. So, let us say suppose so, to describe this properly let me let us first let us take.

(Refer Slide Time: 01:37)



Let S be the set g i of x less than equal to 0 for all i from 1 to m. If you remember A of x star was what we call the active set it is those i in 1 to m such that g i of at x star is equal to 0. And, we say that we said that we say that a constraint qualification holds. If the tangent cone with respect to the set S evaluated at x star is equal to the set of d such that. So, these g i's remember are functions from R n to R. So, S itself is a subset of R n.

So, we say that the we say that the constraint qualification holds if the tangent cone can be written in this, it can be expressed in this sort of form. It is all these such that the gradient of g i transpose d is less than equal to 0, for all i belonging to the active set, right. So, if so, we had we if this holds then we say that the that a constraint qualification holds.

Now, we do not we I also alluded to you that one way by which one condition that ensures that these this holds right; the one condition that ensures that this holds was what was called the was the linear independence condition. So, it said that if these gradients so, the linear independence constraint qualification was one such constraint qualification.

Linear independence constraint qualification was it simply said that all these gradients the gradients at star these are linearly independent. Now, if these are linearly independent, then some then that guaranteed as equality in this in this relation. So, then essentially we could characterize the tangent cone as given by this particular set of inequalities.

Now, what we will what I will mention to you now are some more constraint qualifications. So, some of these have names some of these are have been just referred to in the literature by kind of directly. So, here is one constraint qualification ok. So, suppose we are referring to some x star in s and suppose.

So, we are I am here in this in the in these constraint qualifications I am going to be referring to S in this sort of form ok. In the form that it is all those x such that g i of x is less than equal to 0. So, this gives us constraint qualifications for sets that are described completely by inequality constraints. But then you can always extend it directly to those which have equality constraints by posing the equality as two opposing inequalities ok.

So, suppose x star in suppose x star is easiness and suppose there exists and suppose there exists an h star in R n such that for each i in the active set we either have either gradient of i of gradient of g i transpose d is less than 0 strictly less than 0 or gradient of g value g i at evaluated at x star transpose d is equal to 0 and g i, I do not need a bracket here and g i is affine, ok.

Then we can then the constraint qualification is satisfied constraint qualification holds. So, in short this here what I have written here is a you can say is a sufficient condition for the for a constraint qualification to hold. So, this is a sufficient condition for C Q to hold.

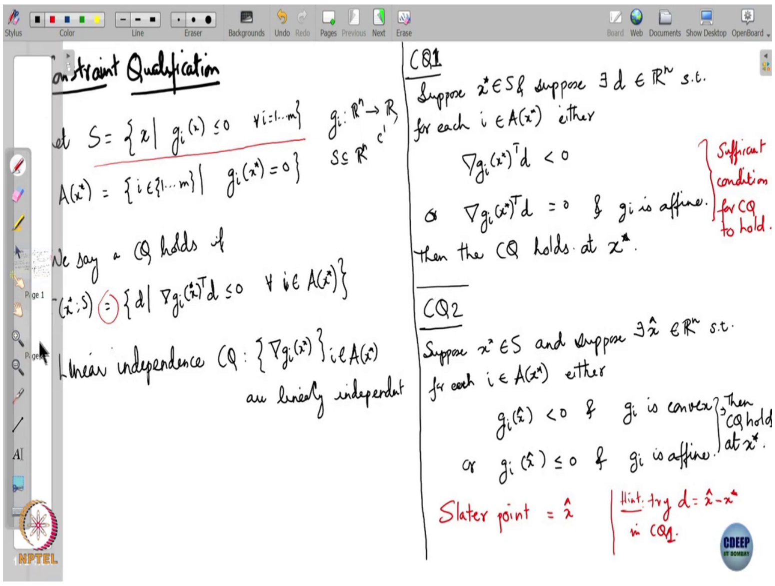
Now, you can be you can make this. So, what does this say? This says that my mistake here I should not say let me erase this mistake here in the notation. I instead of h star let me write d let me write d. So, if you can find the d direction d such that for every constraints you the in that direction you are if you move in that direction for every constraint that is active.

And, from that if you move in that direction for each constraint you become strictly feasible or you remain on the constraint, but then the constraints is are fine, ok. In that case in that case we say that the then from there it follow it would follow that there is equality there is equality here that the constraint qualification holds alright, ok.

Now, there is a somewhat longest proof for this is somewhat longest proof for this I am going to skip the proof. If you want to look up the proof you can look up the notes of in and around on the internet. So, here is another here is another constraint qualification actually this one is very popular it is a very popular constraint qualification and it is it comes out as a weakening of CQ 1, right.

So, let me mention this to you. So, let me once again suppose I should be more clearer here CQ then the CQ holds at x star, ok. Suppose, x star is in S and suppose, there exists a direction d sorry, I do not need the direction d here my mistake.

(Refer Slide Time: 10:07)



Suppose, x star is in S, ok and suppose there exists an x hat ok in R n such that the following is following holds. Such that for each i in the active set A of x star e either g i of x star is strictly less than 0 and g i of g i is and g i is convex, ok.

So, this for us are all and also in C 1 ok I did not mention it, but these functions are all differentiable functions, otherwise I would not be able to take derivatives here yeah. So, these are. So, this is strictly less than 0 and g i is a convex function or g i of x star, sorry. This is not x star sorry, this is x hat my mistake today.

So, and g i of x hat is less than equal to 0 and g i is affine ok. So, suppose you can find a point x hat here a point x hat like this an x hat in R n. Such that for each constraint in the active set either the constraint holds with constraint holds strictly at x hat ok either the constraint holds strictly at x hat and the constraint the function the constraint itself is a convex constraint or the constraint just simply holds its g i of x hat less than equal to 0 and the constraint is affine, ok.

In other words, for affine constraints all you are asking for is feasibility and for convex constraints you are asking for strict feasibility ok, for convex constraints you are asking for strict feasibility, ok. So, naturally an affine constraint is also convex constraint. So, those that are so, this one is to be applied it can be applied only to those constraints that are non affine and get convex ok.

So, for this you have you what we are asking for is strictly feasible. So, this sort of point this point x hat is what is called a slater point and I will make a mention of it again. It comes up in a very important way, later also it is what is called a slater point. So, x hat is what is called as slater point ok and this condition itself is sometimes referred to as the slater condition ok that there exists such a point.

Now, there is a weaker version of this where you do not where you do not ask for the trouble with the way this way of writing constraint qualifications is that it asks us to check this for every i in the active set. A much easier thing to do is not bother about the active set and at all and to check that this holds for all i, ok.

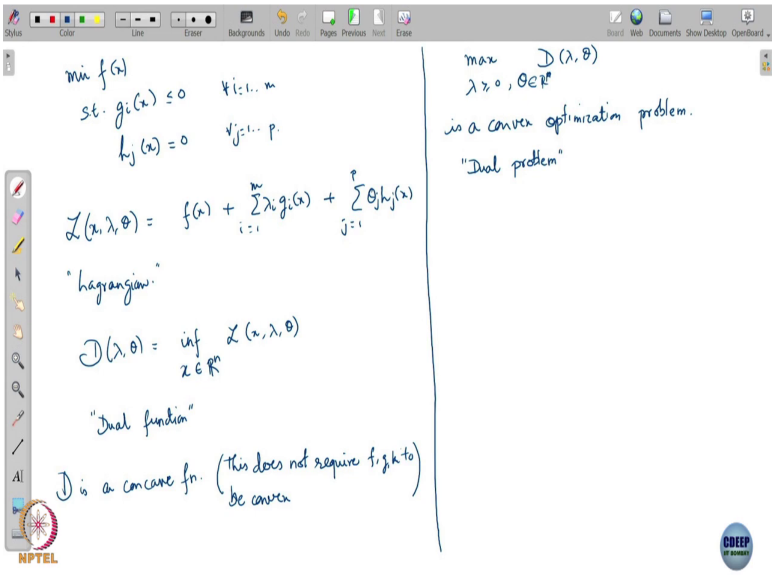
And so, sometimes when we refer to a slater point we refer to that to this to a point where this holds. These hold not just for i in the in a certain active set, but rather for all i. So, that so, essentially what it refers to is that your the existence of in that case what this is referring to is that your the feasible region S is such that there is a there is a point that is in the interior of all the convex constraints and is feasible for all the affine constraints.

So, these points here which satisfy the convex constraints with strictly this point here where all the const convex constraints are satisfied strictly. That sort of point is in the interior of all the convex constraints and this ensures that it satisfies all the affine constraints. So, this put these two put together is effectively saying that you have a what is a point that is in a strict strictly in the interior of the convex constraints and satisfying the affine constraint.

Now, why does this have what does this have to do with the constraint qualification? Well, you can actually check once I will just mention a hint here, the hint is to check try hint is try d equal to x hat minus x star in constraint qualification 1, ok. In constraint qualification 1 you try out d as x hat minus x star, you use the fact that g is convex and g is differentiable etcetera and that should bring you back to something like this right, ok.

So, the I would not say more than that, ok. I did not complete my sentence here. So, let me complete do that ok. So, with these constraint qualifications done, we will now move on to the study of duality in convex optimization ok.

(Refer Slide Time: 16:52)



So, here is my optimization problem. So, minimize f of minimize f of x subject to g i of x less than equal to 0 and for all i equal to 1 to m and although we are talking of convex optimization. Let me just write this out first for the moment in a little bit in a you know general sort of way so, with h j of x equal to 0. So, in a convex optimization problem obviously, these would be linear, but let this issue come up when we actually need when we actually need it, ok.

Now, if you recall we I had defined this function called the Lagrangian function. So, I had defined it as this summation lambda i g i of x i going from 1 to m plus summation theta j h j of x, j going from 1 to p, alright, ok. So, this is what. So, this is this was what was called the Lagrangian function.

And, then now let me define the following other function. This function is this function is D which is a function of just lambda and theta. This D is simply the infimum of the Lagrangian over the entire space; infimum over x for a fixed lambda in theta over the entire space. So, this function is what is called there is a name for this is called the dual function.

Now, there is a reason why it is called the dual function because you will soon see that it is actually related in very in a very close way to duality itself. So, the problem of the dual problem and so on comes up comes up exactly from there. Now, before we before I before we actually go a little deeper into this so, do you is there something that can be noticed directly first?

So, here are a couple of things that that that need to that you should note ok first is that D b. So, D is a function which is a point wise minimum of linear functions, what does that mean? So, if you look at the if you look at the Lagrangian as a function of if you look at the Lagrangian function as a function of just lambda and theta then it is actually linear in lambda and theta.

And, what you are doing in this definition here is taking the minimum of this sort of linear function over a third variable x. So, you are take you have if you look at this as functions of lambda and theta you have in fact, a family of linear functions of lambda and theta. What you are doing is taking for each lambda and theta the least of those.

Now, what happen what sort of function would result from this? A point wise infimum of linear functions would end up becoming actually a. So, that actually becomes necessarily a concave function ok. So, this is something you can prove for yourself that D is a concave function. It is a concave function of lambda and theta.

Now, this this fact does not require this does not require f, g etcetera f, g h to be convex. So, the problem does not need to be a convex optimization, this the fact that this is that d is always a concave function is holds for any optimization problem like this ok.

So, as a consequence what we can say is if I suppose I pose this following. Other type of problem which is suppose I say look at this problem where I am looking to maximize with subject to lambda greater than equal to 0 and all theta, the function D of lambda comma theta. This sort of a problem is therefore, is a convex optimization problem.

This sort of problem is always is a convex optimization problem and why is that the case? The reason for that is because the it is objective is concave and you are maximizing the objective. So, effectively is equivalent to minimizing the negative of D. So, if you are minimizing the negative of D, that would be a convex minimizing a convex function over a convex feasible region which is just lambda greater than equal to 0 and theta unconstrained or in short theta in R p, right.

So, this is always a convex optimization problem. This problem is what we is let me give it a name it is what is called the dual problem and you will soon see that this is in fact, nothing, but the dual problem that you have encountered the same as the dual problem that you have encountered as part of your study in linear programming alright, ok.

So, well if this is supposed to be the dual problem then what where are the relations of duality which is strong duality and which is strong duality and weak duality? So, let us first look at weak duality strong duality is where the is where we will spend most of our time subsequently so, then.