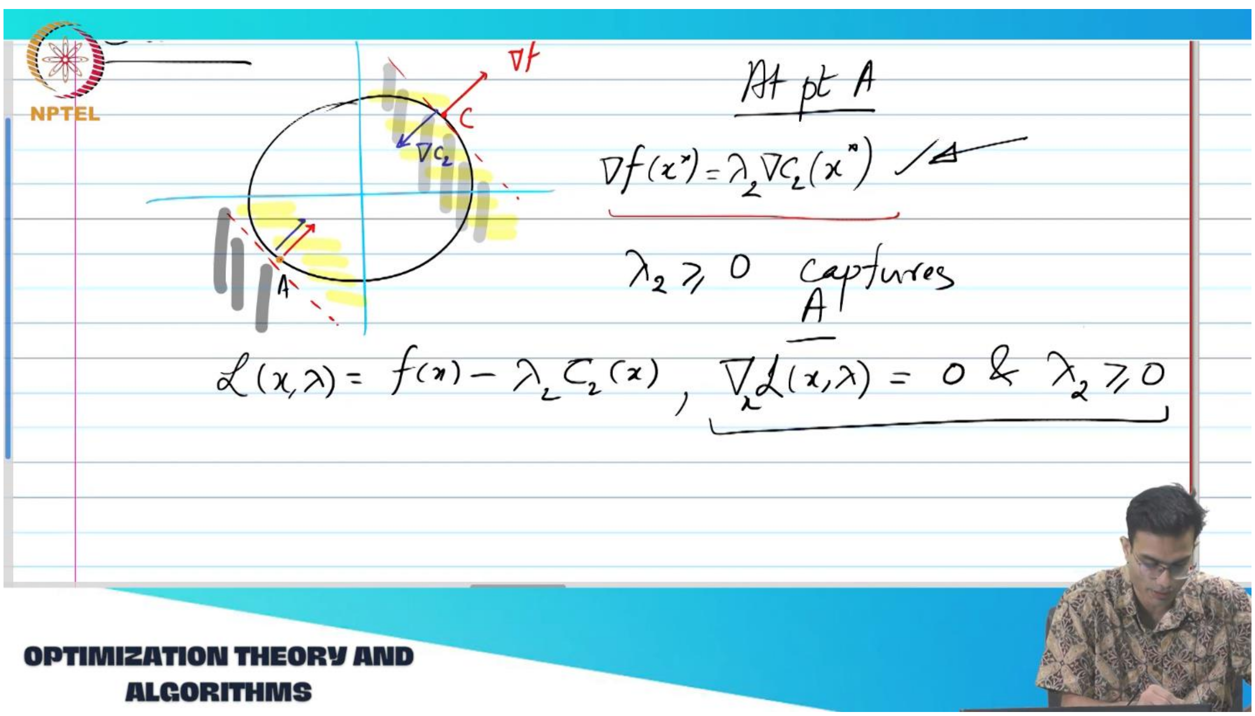
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 09
Lecture - 63

Single inequality constraint - part 2

So we are talking about constraint optimization. We have looked at a single equality constraint and now we are looking at a single inequality constraint. So this is just from last time. So in the single inequality constraint the problem that we have is simple objective function $x_1 + x_2$ and the constraint is the interior of the circle and immediately I know the, I mean not immediately but graphically I can see the solution is at point A and I am asking how do I improve. I am at some point x I want to improve. And to discuss this we came up with two possible cases, one is the point that I am currently at is either inside the feasible set, second case is I am on the feasible set boundary, so that was case one.

When x lies inside we said that this is called an inactive constraint right because the inequality becomes strict and you use first order Taylor's theorem. So you get an expansion for $C_2(x) + s$ and it turns out that any small s will work because you can as long as s is small I can maintain this inequality. And in particular, if I choose s like this in the negative gradient direction, I am bound to decrease my function value. So this was a good candidate, okay? And as we said, this is a candidate, not a unique thing.



Any feasible descent direction will work as long as the magnitude is small enough, okay? Then we said, let's look at the second case, right? So second case is where the constraint is active. That means $C_2(x) = 0$ which means I am at the boundary. Now when I am at the boundary we have both things to look at. One is feasibility which came from this equation, second is decrease. So here now feasibility said that $\nabla C_2^T s \geq 0$.

So that is which shaded region over here? The blue shaded region, right. ∇C_2 and s should be in the same half space so that their angle is acute, therefore the inner product is positive or greater than equal to 0, that is my blue region. $\nabla f^T s < 0$ that means ∇f and s should be in, should have a greater than 90 relation so that gives me the red shaded region. Finally, I am left with a cone which is this yellow shaded region, right. So this has given me a possible region in which I can choose my s to move forward.

Remember, we are not currently talking about an algorithm. We'll talk about an algorithm once we finish all of this, but any s in this direction is roughly okay. So, I think I moved it to the right. Okay. So this, the question that we asked here is, how do I move to a better position? And finally, the second question that we want to ask is, when should I stop? Or when is there going to be no improvement possible? So that's the question that we'll start by asking today.

So, when do we stop improving? Okay. So let us go case by case. Case 1, okay. Remember case 1 was when I am inside the boundary. So what do you think? When will I stop moving, when do you think this thing whole process will not, I cannot move to $x + s$? I am in the interior.

So it is very difficult for me to jump into an infeasible point. So I am in the interior already. So the hint is look at the potential expression I have for s. When will I stop moving? Exactly. If $\nabla f = 0$, then I am, I do not really know which direction to move in.

If $\nabla f = 0$, that kind of kills it there, right? So in the case of case 1, one just to remind you was $C_2(x)$ was strictly greater than 0, ok. So, we have no feasible direction if $\nabla f(x) = 0$. Remember, when we made this, when we discussed the equality constraint, we had introduced a Lagrangian, right. And the idea behind introducing the Lagrangian was to make a kind of a one to one comparison between constrained and unconstrained optimization, right.

So, in the language of the Lagrangian, how can I quantify case 1? So, remember the Lagrangian $L(x, \lambda)$ is what? It was $f(x) - \lambda C_2(x)$, okay. And what is the Lagrangian with respect to x over here? It will be $\nabla f(x) - \nabla C_2(x)$, okay. And in the equality constraint problem, we had found out that the signature for being at the optimum point was what? The gradient of the Lagrangian was zero. If I want to keep the same kind of analogy over here, what do I do? Can I still set $\nabla L = 0$ to capture case 1? It doesn't seem like it'll work, right? Because it's only setting one term to zero, which is the first term. In case 1, $\nabla f = 0$ means I cannot improve anymore.

So that's a signature for a stationary point. What else can I do? Can I impose one additional condition so that $\nabla L = 0$ still captures the case? ∇C is not in my hand. λ is in my hand. So if I set $\lambda_2 = 0$ in right, then this condition seems to work.

I can say that if I am in case 1, then two things have to happen for it to be a signature of a stationary point: $\nabla L = 0$ and $\lambda_2 = 0$, okay. This is, it seems very silly like artificially I am adding in this thing and you know, but we will see what is the power of doing that in a few moments. Now let us move to case two. Case two, when is case two not going to work out? If there is no intersection between these two half-spaces, right? That is what we need to capture, okay. Now let us look at this a little bit geometrically once again.

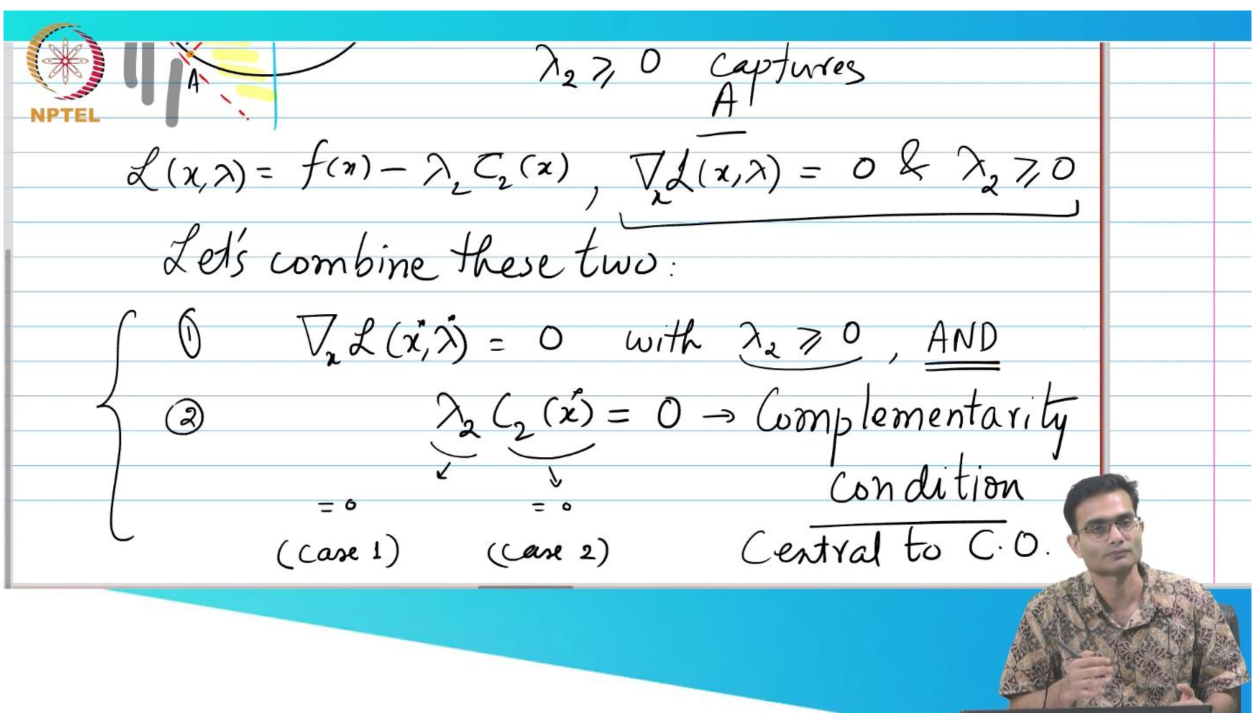
Let us look at this point over here, sorry I had used another color over here. Let us look at this point over here A and let us, this point was marked C, this point was marked B and this point was marked D. At point at this point which I have chosen obviously there is a yellow shaded cone which is okay. Let us look at point A which is the solution, okay. Now at point A in the solution, maybe I should draw this once again over here for clarity.

Which way is ∇f pointing? At all points ∇f is always the same which is $(1,1)$, right. So my ∇f is here, my ∇f is here, this is my ∇f. What way was my ∇C_2 pointing? Outward or inward? Inward, right because I had defined C_2 with a $2 - x_1^2 - x_2^2$. So ∇ gave me $-2x_1, -2x_2$. So that makes my ∇C_2 here.

Now, let us look at this point over here. So, $\nabla f^T s$ should be less than 0, that means where am I? Here, oops, ok, great it is. And $\nabla C_2^T s$ should be greater than 0. That means where am I? Here.

Is there any intersection? There is no intersection. The inequality is strict. $\nabla f^T s$ is strictly less than 0, there is no intersection, right. So, these together seem to be capturing it. So, is there a relation that I can catch between ∇f and ∇C_2? What can I say? In this situation at point A over here, is there a relation that you can observe between ∇f and ∇C_2? It is staring at us in the face, what is it? They are parallel, okay.

So at this point I can say that $\nabla f(x^*) = \nabla C_2$. This seems to work at point A and what, ok, so we will come to that. Let us look at the same thing for point C over here. So, this is point C. At point C I have once again let us look at this gray shaded region is going to give me $\nabla f^T s$ less than 0 I am here.



$\nabla C_2^T s$ where is it? $\nabla C_2^T s$ should be right. So, in this case I am going to have something like this. Right. That means is there an intersection possible? The whole half space is possible and we know that C is not a legitimate, is not the solution. In fact, this is the maxima of the problem, right.

So now looking at this, how can you make this expression over here? How can we fine-tune it just a little bit more? λ_2 should be greater than or equal to 0. If I say $\lambda_2 \geq 0$, it captures point A and not point C. So, $\lambda_2 \geq 0$ captures A correctly, ok. Now, in terms of the Lagrangian how do I write this? Remember Lagrangian was $f(x) - \lambda_2 C_2(x)$. Now, I want to capture this case using the Lagrangian.

What is the first thing I can say? I want to capture this condition. How do I capture it? ∇L with respect to x, this if I set it equal to 0, it captures this and λ_2. So in case 2, this is what I have. In case 1, you have seen what we had. We had $\nabla L = 0$ and $\lambda_2 = 0$.

So this looks a little tedious to have these separate cases each time. So it turns out if you just stare at both of these cases a little bit more, you can come up with a very elegant way to combine them. So let's do that. So let us say, what is the common thing between both of them? ∇L with respect to x is 0. So, let us write that as a first condition because that is clearly common over here.

$\nabla L(x, \lambda^*) = 0$. This is common to both cases. But we also need over here to catch case 2, let us say $\lambda_2 \geq 0$. This is needed. The second condition is, the second way to combine it is quite interesting. It says, and have a look at this, that $\lambda_2 C_2 = 0$.

So, does this work for us? In case 1, what did I want? $\lambda_2 = 0$ right. So this guy equal to 0 captured my case 1, right. Did this constraint cause any problem for us? It does not cause a problem because $\lambda_2 \geq 0$ the equal to part is getting set over here. What about case 2? $C_2 = 0$ in case 2.

So this is equal to 0 in case 2. That means in case 2, I do not have to force λ_2 to be equal to 0. λ_2 can be some number greater than or equal to 0. It is consistent because $C_2 = 0$ is capturing it. So you notice this is quite clever, right? I have combined both cases. Now I do not have to talk about case 1, case 2.

I say that these are the two conditions necessary for sufficing as a check for a stationary point. $\nabla L = 0$, the Lagrange multiplier is greater than or equal to 0 and $\lambda_2 C_2 = 0$. Ok. So this second condition turns out it is used so often in optimization there is a name given to it, it is called the complementary condition. Complementarity, ok, it is very very central to constraint optimization.

So let us just quickly summarize. Case 1 was setting the gradient of the Lagrangian equal to 0. That seemed a little superfluous because what do I do with λ_2? So I said $\lambda_2 = 0$, I am done with that. Case 2, what did I have to do? I basically had to set, along with the Lagrangian equal to 0, I had to set the parallel condition.

Anti-parallel did not work, right? So this was the parallel and this was anti-parallel. This was for a negative λ, this did not work, right? So that is why we had to force $\lambda \geq 0$ and then the nice way to combine these two is over here. It looks similar to unconstrained optimization, it is just that way. Earlier in unconstrained I had $\nabla f = 0$, now I have $\nabla L = 0$ with now a few extra things: Lagrange multiplier ≥ 0 and complementarity conditions.

Notice one very big hole or gap right now. I have not proven any of this. I am just going from geometric intuition to formulating this. So we will come to proofs later time permitting, okay? But the intuition from geometry I hope is very clear how we have arrived at this. There is a proper rigorous proof behind all of this which we can come to.

And you can see they don't contradict each other. So let's look at it again. In case 1, case 1 I needed $\lambda_2 = 0$. It was an inactive constraint. So $C_2 > 0$.

But I needed $\lambda_2 = 0$. So case 1 is covered. And case 2, the constraint is active.

$C_2 = 0$. λ_2 can be in it. So it's captured over. But that is why the $\lambda_2 \geq 0$ is included in the first case, in the first statement. Because the first statement is saying $\nabla L = 0$ that means the gradient of the function and the gradient of the constraint, they are parallel or anti-parallel, both are possible. So, I additionally specify $\lambda_2 \geq 0$ to make it such that they are parallel and not anti-parallel. So, this very simple analysis with a single inequality constraint has given us this complementarity condition. In the case of equality constraint, we never needed this complementarity condition, right?

So, later on as we go, we will talk about a general situation where I have equalities and inequalities. Together we will formulate one big hefty theorem which tells us what are the first-order necessary conditions for optimization, but you have all the nuts and bolts are here itself. As you can guess, in the case of equality constraint, did I have a condition that the Lagrange multiplier should be greater than or equal to 0? No, right? So that is what will turn out in the final theorem statement where I have a combination of equalities and inequalities. The equality-related Lagrange multipliers can take any sign.

The inequality-related Lagrange multipliers must be greater than or equal to 0, and I have this complementarity condition. This is together going to form our sort of canvas for doing constraint optimization. The first, you mean this guy? So this captures both case 1 and case 2.

So let me put them side-by-side. So look at this. This is over here, case 1 is happening over here. So case 1, I can $\nabla L = 0$ is along with $\lambda_2 = 0$. So both, actually both of these, let us call these statements, both of these statements together work for case 1 and case 2. The only difference is when I say $\lambda_2 C_2 = 0$, there are either λ_2 could be 0, that is my case 1, along with statement 1. $C_2 = 0$ along with statement 1 gives me case 2. That's how we do it. So I've combined it in a slightly neat way.