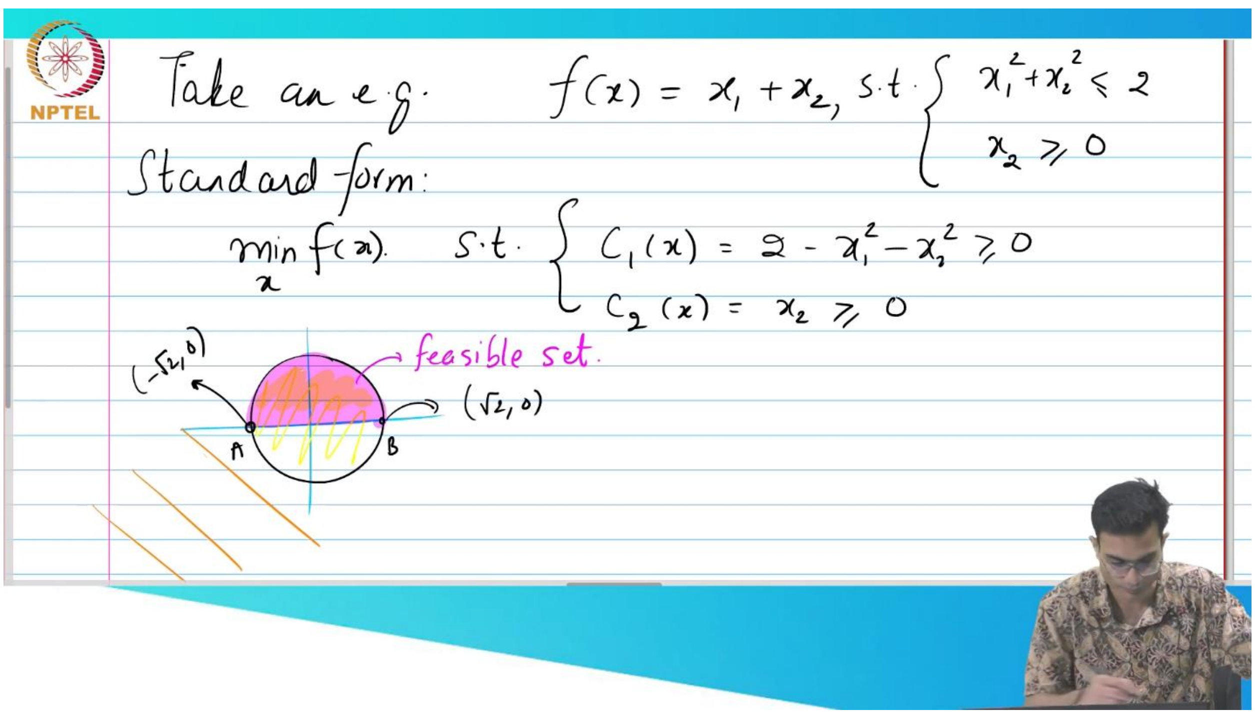
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 09
Lecture - 64

Two inequality constraints example

Okay. So let us further drill this in. Let us take one more example now. Okay. Because in constraint optimization the more examples we study the more rooted we will get in the understanding. So let us take another example.

This time let us make the example a little bit more challenging in the sense let us include multiple inequalities. Okay. So let us take an example. I am going to take my good old objective function $x_1 + x_2$ and I am going to give your constraints, two constraints in the following way, $x_1^2 + x_2^2 \leq 2, x_2 \geq 0$.

So this is like a general problem that has, that comes to you, right. What should be our first step? First step is actually very, very simple. It is called, I would say standardize the problem. So in standardizing the problem, I need to make sure that all the, remember when we wrote down the form, the expression for a constraint optimization, all the inequalities were what? Greater than or equal to zero. So let's convert it into that because if we don't do it, all our answers will be off by a minus sign.



Standard form. So, let us denote this $C_1(x)$ as the first constraint. So, how should I write it? $2 - (x_1^2 + x_2^2) \geq 0$, $C_2(x) = x_2$, this is already in the standard. So, this is my feasible set. So, we can sketch the feasible region quite easily.

What is it? The first constraint is saying I should be inside the circle. And the second condition is saying in the positive y-plane, right. So, that in fact is. So, this is our, quite easy and the contours of the cost function are anyway like this. So, where do you think is the solution? $(-\sqrt{2}, 0)$, right. So, I have my point A, there is another point over here which is also interesting.

So, let us call it. So, this is $(-\sqrt{2}, 0)$, this. So what we will do is we will look at these conditions which we have just formulated now and we will analyze it for both point A and point B and see what is the difference between these two guys. So now let's, since we're doing this in the beginning, let's write it down everything explicitly, right? So at points A and B, can we say that the constraints are active? Right? Are both constraints active at point A and point B? Yes, right? So both constraints are active, okay? So, anyway that is bookkeeping okay. Now for there to be a feasible direction to move in right, for there to be a, what do I need to happen? I need to have feasibility and a decrease in the function, those are the two kind of poles right. So, I had $\nabla f^T d$.

This is less than 0. This gives me what? Decrease of feasibility. Decrease, right? This gives me decrease. And feasibility is going to be given by this. This is going to be. Quick question, is this second condition like case one or case two of last time? Case two, right? Because the constraint is active, right? So let's just write it like case two.

Okay, I'll move to the next page then. Now I want to define the Lagrangian of this problem. So, how do I do it? $L(x, \lambda) = f(x) - \lambda_1 C_1(x) - \lambda_2 C_2(x)$. So if I had n constraints, whether they were equality or inequality, what do I do? I just keep appending it, right.

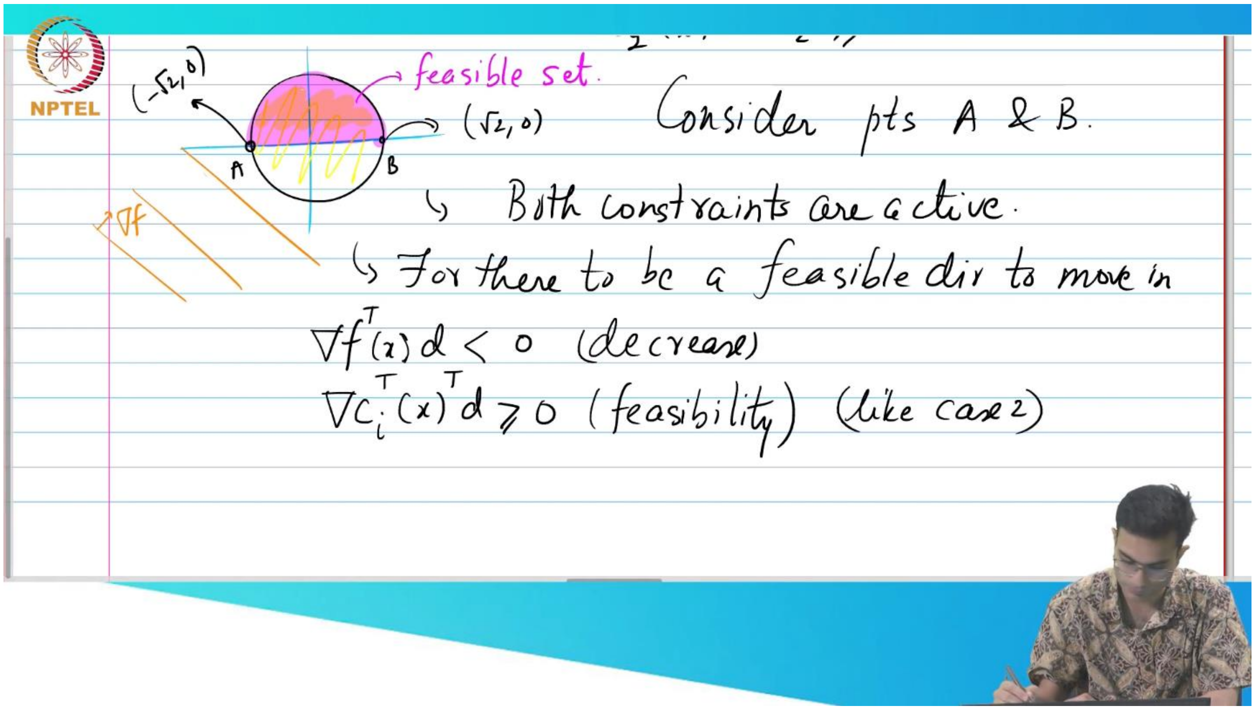
So that is, you will also find a more compact way of writing this. You will find in many places this is written either like so, $\sum \lambda_i C_i(x)$, this is one way and another way is to simply write it as $\lambda^T C$, where C has all the rows of C as the constraints and λ is a Lagrangian multiplier vector. So, this is a more compact way of writing it. Now, in order to find, you know what we have just discussed previously, what do I need? If I am at, so I want to investigate whether point A or whether point B, how does it look with respect to the condition that we have previously derived? The previously derived condition was $\nabla L = 0$ and $\lambda_2 C_2 = 0$, complementarity conditions, right? If those were satisfied, we said that what? It's a candidate, a first order necessary condition for a stationary point.

So these are conditions to be evaluated at some given point in the feasible set. These conditions are not going to help you to arrive at a feasible point, that the algorithm will do. But that is why we have taken two points for investigation, point A and point B. So at both of these points, let us see what do these conditions say. It should work for point A, it should fail for point B.

So let us see is that actually happening. Okay. So what is our condition once again? So ∇L with respect to x, λ, this should be 0 and what was the statement 1? and λ_1, now it will be λ_1 and λ_2, both should be greater than or equal to 0 because both are inequality constraints, right. So, I have $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, this was statement 1. Statement 2 was complementarity.

Complementarity would say that $\lambda_1 C_1 = 0$, $\lambda_2 C_2 = 0$. This is complement, right. So, if you had more constraints you would just keep expanding so many, if there are n inequality constraints then you will have n complementarity constraints, ok. So, now let us write down a general expression for ∇L. What is $f(x)$? What is ∇f? We have done this many times, $(1,1)$.

So, this guy is $(1{,}1)$, minus λ_1. What is ∇C_1? Let us look back at how we defined it. It is $(-2x_1, -2x_2)$. So that minus can, that minus will make this guy plus and I have $(2x_1, 2x_2)$ and then I have a minus λ_2.



What is C_2? x_2. So its gradient is $(0{,}1)$. Ok. The minus sign stays there.

So $(0{,}1)$. Ok. And I want this to be equal to 0. So, if I want this to be equal to 0, does this look like a system of linear equations, right, in either x_1, x_2 or in λ_1, λ_2 right? So, let me just write it here. So, I am going to get $1 + \lambda_2\lambda_1 x_1$ and 0, and then I have $1 + 2\lambda_1 x_2 - \lambda_2$. This should be equal to 0, 0.

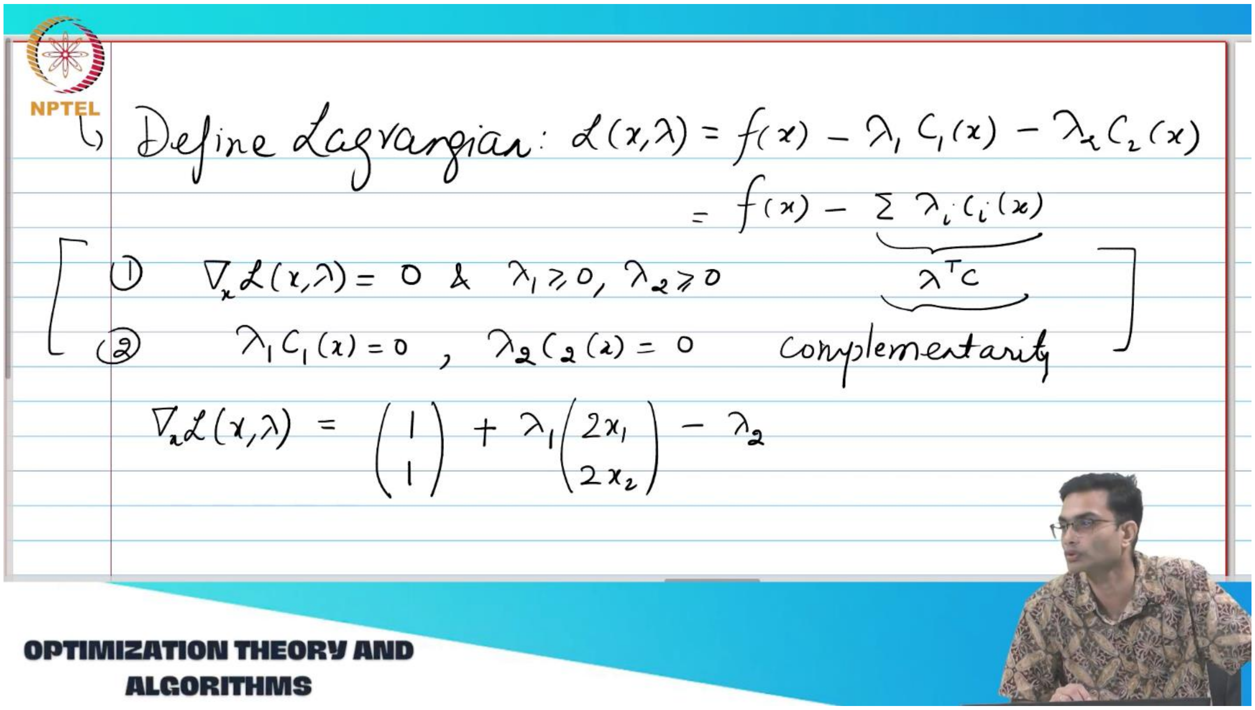
So, now let us do this analysis for point A and point B. So, at A and... So, this is my point A. At point A, where is my ∇f pointing? ∇f is always pointing in the same direction, $(1{,}1)$. So, this is my ∇f. Which way is ∇C_1 pointing? ∇C_1 is the interior of the circle.

So, it is radially inwards, right. So, this is my ∇C_1. Which way is ∇C_2 pointing? Upwards, $(0{,}1)$, right. So, this is my ∇C_2, ok. Now let us, let us start peeling these apart, okay.

So at point A, so first we will just do the intuitive thing. Let us see is there any intersection giving me a feasible direction to move in. $\nabla f^T d$ should be less than 0. That means I should be to the bottom part away from ∇f, right, this region over here, okay. $\nabla C_1^T d$ should be what? Greater than or equal to 0.

So, that is going to give me which region? Right side, right? It is going to give me this region. So, is there any intersection of ∇f and ∇C_1? Yes, currently which is the fourth quadrant. The fourth quadrant is common. And then comes our second friend ∇C_2. $\nabla C_2^T d$ should be greater than or equal to 0.

That gives me which? Top half, right. So it gives me this region. Now is there any intersection? There is no intersection. So looking at our analysis that we did of the set of feasible directions, we can conclude there is no feasible direction. Now let us see if the algebra that we wrote over there captures the same thing, ok.



So here. What was the coordinate? $(-\sqrt{2}, 0)$. I can substitute this coordinate into the equation that I have over there. So what does it say? I am going to get $1 - 2\sqrt{2}\lambda_1 = 0$. And the second condition is going to give me what? $x_2 = 0$. So I am left with $1 - \lambda_2 = 0$.

Right? I am just using this condition over here. I substituted the value of $x_1, x_2 = -\sqrt{2}, 0$. Put it into this simple equation and I solve for λ_1, λ_2. So this gives me $\lambda_1 = \frac{1}{2\sqrt{2}}$ and $\lambda_2 = 1$, ok.

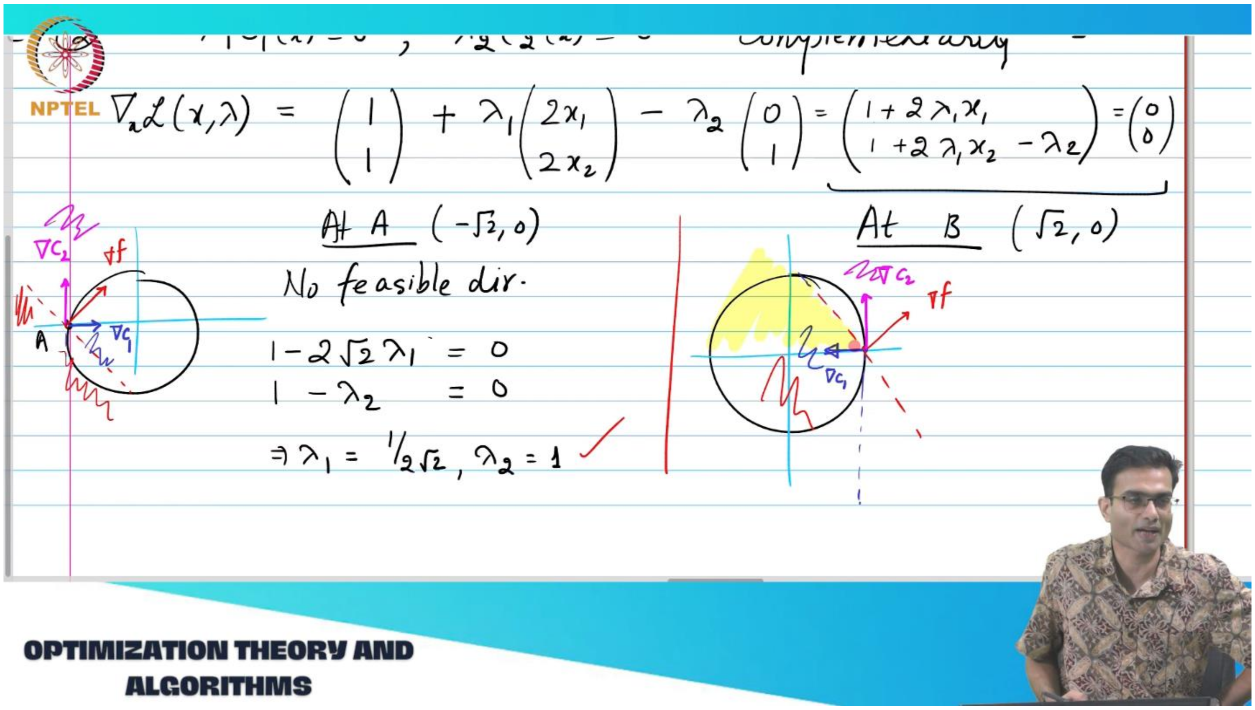
That seems ok. Why? Because. Because the Lagrange multiplier should be positive because it is an inequality. So that seems to be satisfying my statement 1, right? So over here. If I choose these values of λ_1 and λ_2, statement 1, right? Statement 1 was this guy over here. Obviously, the way I derived it was by setting $\nabla L = 0$. That's how I got these values of λ_1 and λ_2 in the first place.

So I've satisfied that. $\lambda_1 \geq 0$, yes. $\lambda_2 \geq 0$, yes. What about the complementarity conditions? It's an active constraint.

So C_1, C_2 are anyway zero. So I don't care. This is automatically satisfied. So you can see that what I very tediously arrived at by drawing half-spaces and shading their intersection and all, I can see I'm just, by algebra I'm getting it, right? So this is, I can put a tick mark. This point at $(-\sqrt{2}, 0)$ is giving me a set of legitimate Lagrange multipliers. First order conditions are satisfied. Therefore, as a necessary thing, I can say point A satisfies the necessary conditions for being a stationary point.

Now, let us hope that this whole thing fails for point B. So, at point B my directions are like so, draw this over here, ok. So, we are sitting, this is a terrible, draw the circle first, okay, and now we are sitting over here at this point.

∇f is $(1,1)$. This is ∇f. ∇C_1 is radially inwards pointing towards the circle. This is ∇C_1. And my ∇C_2, which was ∇C_2? (1,0,1), right. So this guy is pointing over here.



This is my ∇C_2. So, $\nabla f^T d$ should be less than or equal to 0 for improvement, that means I should be somewhere in this region, ok. $\nabla C_1^T d$ should be greater than or equal to 0, that means where should I be? In the left half of this point, right. So, I should be here. So, there is so far I am seeing an intersection over here, right, that which is there. In fact, it is an interesting space because it is going to be bounded over here and then like this, right.

So, it is a different, like a paper folded like that. And $\nabla C_2^T d \geq 0$, where am I? Top half, right? So that is this region. Is there an intersection? Yes, that intersection actually, if I shade it with yellow, is going to give me, oops, something like this, right? And that makes sense, right. You can imagine that if your algorithm started from point B, for example, I know the solution is A, I need to move towards A, and this cone is giving me a legitimate direction to move towards A, right. So, that is what I expect from all this tedious drawing of planes and so on.

Now, let us see what the algebra does, ok. So, what will I get over here when I substitute $(\sqrt{2},0)$ into this equation? So, I am going to get 1 here, just the sign will flip over here. This is going to be the first one and this is going to be the second one. So, implies that $\lambda_1 = -\frac{1}{2\sqrt{2}}$ and $\lambda_2 = 1$. So, what has failed us? This guy, the first guy has failed us. So, now this is a way by which we are translating our geometric intuition into algebra.

And the moment I do convert it to algebra, it is easy for me to then code it, right. So this is my sort of proof that B is not a stationary point, right. So you notice that this example was really very, very straightforward, right? It had a very simple objective function. It had two inequalities. And what we did was, and what you should keep in mind as you try to solve problems in the future, first get it into standard form so that the inequalities are all in the way that you are familiar with, i.e., greater than or equal to zero, right? And your algorithm has given you some points to consider.

And at those points, we are trying to evaluate whether or not the conditions which we derive hold. The conditions being gradient of the Lagrangian equal to 0 and the complementarity condition, right. So, then we basically looked at, we started with the definition of the Lagrangian. As many inequalities and equalities are there, those many Lagrange multipliers get added.

That is my big hefty Lagrangian. It looks big, but it is actually very simple. I am just mechanically adding terms after each other. And I look at these two conditions and I can see that the conformation between geometry, the intersection of half planes, and algebra has given me a proof, right. So A, should I, what should I say? A, should I say A is a stationary point? Strictly speaking, can I say A is a stationary point? If I were being very, very pedantic.

I should actually say is a candidate. This problem is very simple. I know A is the stationary point. That's because I drew it and calculated it explicitly. But I should say A is a candidate for a. This is nothing fancy. In the world of unconstrained optimization, if I found a point where $\nabla f = 0$, I didn't say that this is the minima of the problem.

Okay, if I define a stationary point as $\nabla f = 0$, it's a stationary point, but it may be a maxima also. So, actually, wait, this is a candidate. It is a stationary point. We should say it's a candidate for a minimum point. We will not be doing it in this course, but like we had in the case of unconstrained optimization, first order gave me necessary conditions. Where did I get my sufficient conditions from? Second order. So in constrained optimization also there is a Hessian-based sufficient condition.

If we get time, we will go into it. Otherwise, we will work mostly with first order conditions. Any questions on this? So, the next thing that we want to talk about is kind of an opposite step from what I just discussed. What I just discussed was geometry got confirmed by algebra. I had some intuition of geometry of intersecting half-planes and I confirmed that with algebra. Now, in what we will talk about next, I will show you that there is a problem with this sort of dance between algebra and geometry.

So, let us, to get this into more formal language, let us define a few things, okay. These directions that we just computed previously on the example, we call them feasible directions, right.