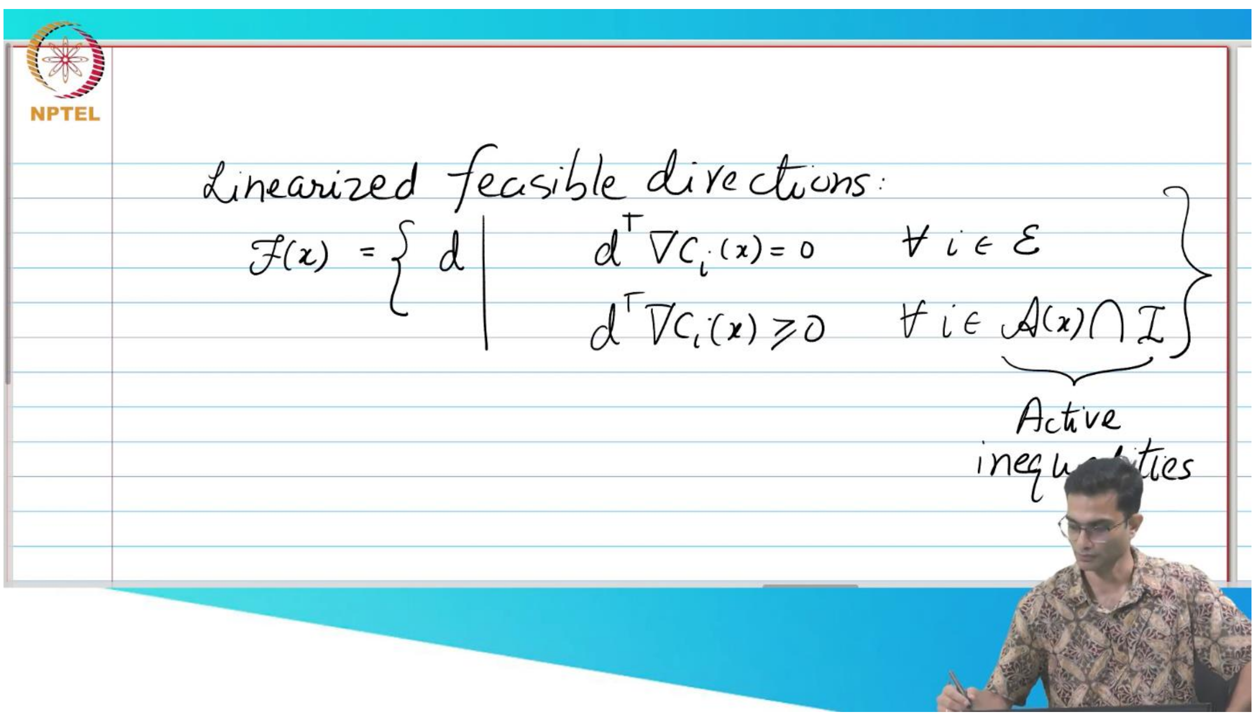
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 09
Lecture - 65

Linearized feasible directions

Which order of Taylor's theorem did we use to derive them? First order. So, there is a formal term for these directions, they are called linearized feasible directions. So, linearized linearized feasible directions, how are they formally defined? This is a script F as per Nocedal's notation, is the set of all d, ok. Just go back to how we have defined it in the previous case for the inequality constraints. When I had an equality, which was the first example, very first example, what was the relation between d and ∇C? Look back at your notes. Greater than 0, less than 0, equal to 0 for an equality constraint.

Actually equal to 0. So, $d^T \nabla C_i(x) = 0$ for all i belonging to... So, if I were in the equal set, then how do I move? I had better move, I mean this makes sense in plain English also, right? If I am at an equality constraint, then by moving orthogonal to the contours of ∇C, I will maintain the same value of C, like when I am climbing up a hill. So, this is equal to 0, that makes sense.

And this is about equality. What am I left with? Inequality constraints. Inequality constraints, I think if I wake you up at 2 am you will tell me this is the condition now. $d^T \nabla C_i(x) \geq 0$. Now for all i belonging to what? Should I write inequality or should I refine it a little bit? Is it all inequalities for which this should be true? Think of the distinction between case one and case two.



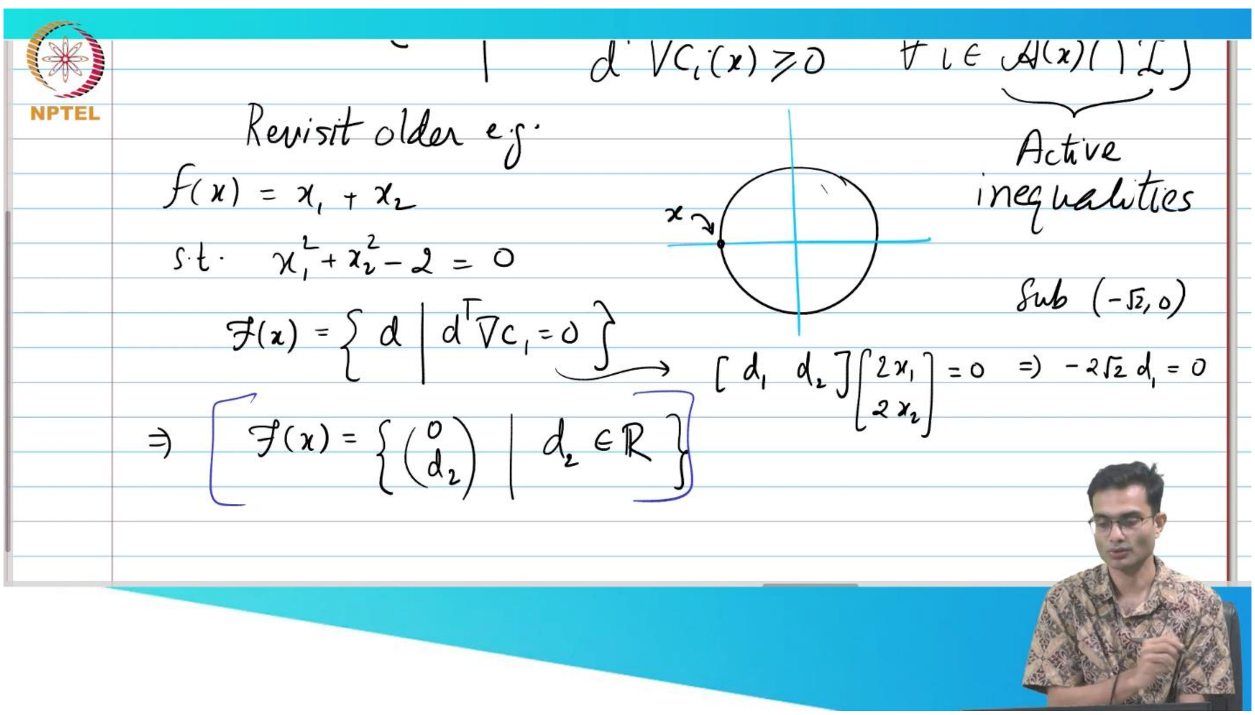
What are the differences between case one and case two? Active versus inactive. If I am inactive, think of it, I'm way inside the feasible region. Any direction works. But if I'm on the boundary, I had better be careful about this, right? So, this is the intersection of the active set. Active set contained equalities and inequalities that are active.

So, I have to intersect this with the set of inequalities. Looks tedious, but what does it mean? The active inequalities. So, what we already know has been sort of formalized in some fancy language, this captures that if you are at some point x and you want to remember this is the linearized feasible directions at a point x, right? So, x is over here. So, you calculate this, the d which satisfies this together, both of these sets is the d in which I can potentially move to.

Then I have to see whether that d gives me an increase or a decrease and so on. That is the rest of the story. But this is the set of feasible directions as given by the constraints. So, now let us revisit an older example which was simply $x_1 + x_2$ was my condition and the simplest case which was $x_1^2 + x_2^2 - 2 = 0$. This was our equality constraint.

Draw this. Feasible set is the border of the circle. Okay. Now let us try to work out, for example, let us take this point. x is over here.

I have taken this one point over here. I am purposely not taking the solution just to make it a little bit more interesting. So, having this direction, this definition in mind, at this point x, remember this is at some point x, if I ask you what is the d? How do I define it? What will you, what will we write? $d^T \nabla C_1 = 0$, right? So $d^T \nabla C_1 = 0$. This is a set of feasible directions, right? So, let's open this up. What is d^T? d is a two-dimensional vector because it's a direction in x_1, x_2 space.



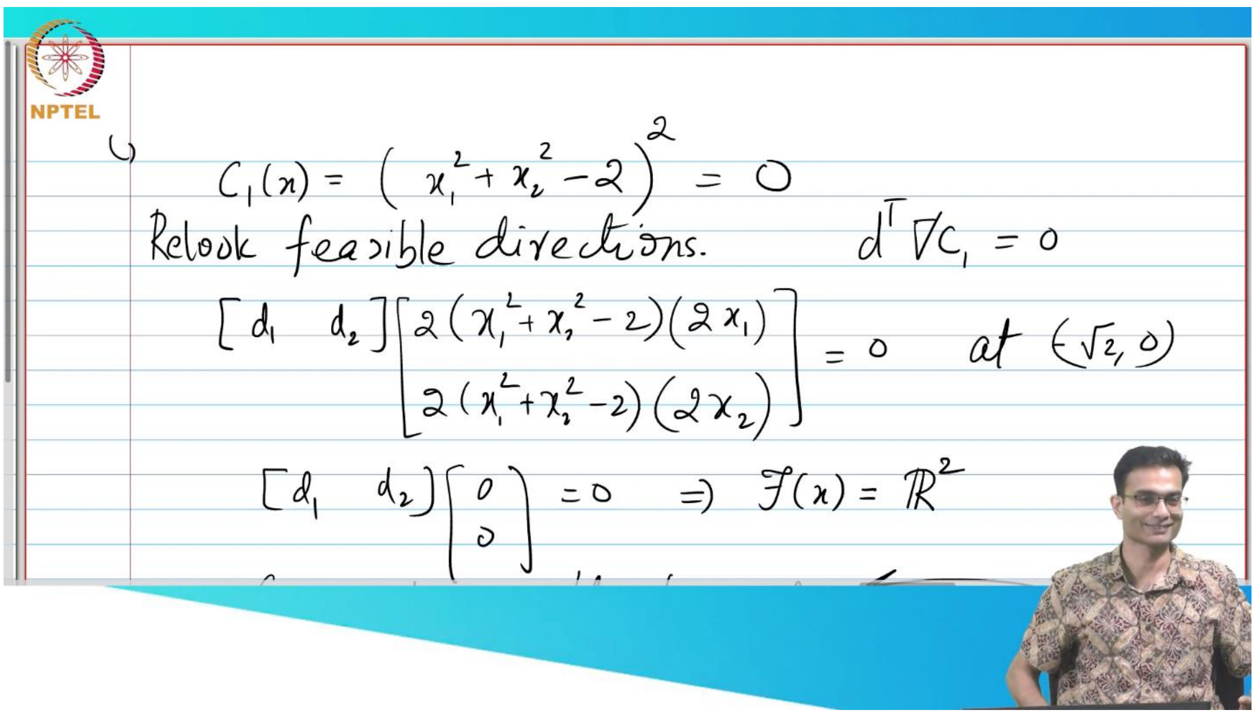
So, I can simply write it as d_1, d_2 like this. What is ∇C_1? $2x_1, 2x_2$. And I need this to be equal to 0. Right. So, now what do I do? Substitute.

What is this point? $(-\sqrt{2}, 0)$, right. So, this is going to give me what? I mean rather this will lead to, so if I substitute this over here, $x_1 = -\sqrt{2}$. So, $d_1 \cdot (-2\sqrt{2}) + d_2 \cdot 0 = 0$. So, this is simply going to give me $-2\sqrt{2}d_1 = 0$. I just substituted the coordinates x_1, x_2 into this. So I got $d_1 = 0$.

So, d_1 has to be 0. So it tells me that at this point, my set of feasible directions is $(0, d_2)$. $d_1 = 0$. I'm left with d_2. And I didn't get any constraint on d_2.

So, d_2 could be any number. So I will just simply write that as d_2 belongs to the set of all real numbers. Everyone with me so far? So, what is this saying? If I were to draw this, if I ask you rather to draw this, draw this feasible set at this point x, what will you draw? A vertical line, will that vertical line be up or down or can't say? It should be up, why? Is d_2 belonging to $+\mathbb{R}$? No, we are not worrying about decrease right now. Just the set of feasible directions, don't take the function decrease or increase into account, it is just feasibility. So, it is saying you can either, you are basically on this line, you could go up, you could go down, right.

And that kind of makes sense because we know what is the solution to this guy, right? The solution to this problem is in fact this point over here. Now for me to stay feasible, I have to go tangential to the circle. I will go tangential. Between these two directions, which direction is the direction that makes sense? Minus x_2. And that minus x_2 is included in the two possibilities.



So it kind of makes sense. If I had an algorithm, out of these two directions, it will pick the correct one going downwards, okay? So, great. So, so far everything is fine. So, we are yet to decrease condition, correct. Now, I am going to do a little bit of a trick, cheap trick, very very cheap trick.

To break this connection that we have going on between algebra and geometry, I am going to leave geometry as it is, but I am going to change algebra a little bit in a very, very tricky way. $C_1(x)$ I have defined as, how have I defined $C_1(x)$? $x_1^2 + x_2^2 - 2 = 0$, right? That everyone agrees is the border of the circle. Now what if I define this instead? This is what I had but now supposing I do this. Looks like a very cheap trick. Has the feasible set changed? Not at all.

Geometry therefore is unchanged. But this is going to give us some strange problem with algebra. Okay. So let us relook feasible directions. What did I need for feasible directions? $d^T \nabla C_1 = 0$.

This is what I am looking for. So, d_1, d_2. This is my d vector. Now, let us put down ∇C_1.

So, $\frac{dC_1}{dx_1}$. What am I going to get? $2(x_1^2 + x_2^2 - 2) \cdot 2x_1 + 2(x_1^2 + x_2^2 - 2) \cdot 2x_2$. This whole thing should be equal to 0. At which point am I investigating this? The point is the same, right? $(-\sqrt{2}, 0)$. What does this translate to? When I substitute this, I am in deep trouble. $0 \cdot 0 = 0$, right.

So, does this mean $d_1 = 0$ or $d_2 = 0$ or no? It means d_1, d_2 could be anything, right. So, this tells me that my feasible set is equal to, pick your favorite d_1, d_2 it satisfies these conditions, right. So this, we get a very important conclusion. Geometry is unchanged. I am at the same point, satisfying the same constraint.

Okay. What happened to algebra? Changed in a very dramatic way. Okay. In fact, it changed the set of feasible directions. From being constrained to just be that vertical blue line, now I am saying go anywhere in the plane, you are okay. So algebra has changed, geometry is not changed, ok.

So there is something more that I have to do to algebra to make it be in sync with geometry. There is a missing link over there, right. Now that missing link is what is called constrained qualification, ok. So which we will talk about. So you have to pass this through a filter, think of it as a filter of constrained to kind of restore the balance with geometry.

Correct. To further terms? Possible. Then I have to, yeah, I mean this entire analysis is assuming first order conditions. Yeah, I did. So if I were living with first order, it's like I have a toolbox that works only on first order conditions. And by doing this cheap trick, your tools stop working. So before we go to second order, third order, remember going to higher orders is extremely expensive.

Because then I have to calculate higher order derivatives. So the very natural question people asked is, can we still make our tools work while staying in first order? So these constraint qualification ideas, which we will talk about in the next session, they help us capture, they help us to retain this kind of a balance between algebra and geometry. Yeah, I think that's basically what I wanted to mention about this. Lack of sync between the two. So hopefully you all got an idea of when I say geometry what I mean and when I say algebra what I mean.

Algebra is sort of just manipulating symbolically various symbols and expressions and getting a gradient or whatever out of it and it conveys something. And geometry is what I am getting by looking at this drawing. The drawing that I have done, it is okay for simple two-dimensional case. It is not going to work for something, you know, in n dimensions. So, we need to figure out a way to deal with that.

So, we will do that in the next session. So, any questions? Okay. Yeah, so this definition of linearized feasible directions is simply formalizing what we already knew, which is the what are the directions in which I can go so that I maintain or rather so that I remain in the feasible set, that is all. For equality constraints. For equality constraints. Second one, okay, so forget this math that I have written there, you can just say active inequalities, that you understand? because let us go back to case 1 and case 2 earlier, which was here, right.

So this is the same problem. Case 1 was when x was inside this, right. At that point if I ask you what is a feasible direction? Feasible direction, I am not even interested in degrees. So I am sitting here in the middle or somewhere inside. I ask you feasible direction, what would you say? At least 0.

$C_2(x)$ is strictly positive. So the second term can be actually anything as long as overall I am greater than or equal to 0. So basically what is it saying? There is no real constraint. You can be in any direction. Right? So this is a constraint which is inactive. So I need not bother about it because anything is possible.

Well, a greater than or equal to zero needs to be satisfied. So here I'm talking about the direction, not the length of the vector. So any direction will work. The total, as long as the total term is greater than or equal to zero.

It can be negative. According to the definition of feasible direction, this guy didn't appear only because this is an inactive constraint. This guy drops out of the discussion altogether. Exactly, that is why I need to be little careful about active inequalities because it is very, if I take any direction if I am in case 2, if I take any direction when I am in case 2, I am in trouble, I can fall out, right. So that is why the final definition that we have of this, of the linearized feasible directions takes care of the active inequalities. because active set includes equalities and inequalities.

If you look at the original definition, it has both. I've already caught the equalities above. So I just want to pull out the inequality constraints, but I want them to be active. It looks complicated to write, but we're just saying equalities and active inequalities, that's what. Constraint qualification, well I have not told you very much yet.

We will discuss it in the next module. What we are saying is that by changing the constraint by simply in this case we just squared it. We found out that $d^T\nabla C_1$, it gave us nothing meaningful. It said take any direction. But we know that taking any direction here is not good because if for example I go along the plus minus x axis I will become infeasible.

So I don't want that to happen. So there is something shaky about this way of coming up with feasible direction. So I need one more kind of a filter to kind of tame it down. Those filters are called constraint qualification, which we will come to. I'm just giving you a trailer of what is to come. But we identified the problem here that I could, I squared it, I could take a square root of it, I could cube it, whatever.

Other ways of writing the same constraint will get me into trouble. And when someone gives you a problem they may not you know filter and boil down the constraint in the most simple logical way. It may be written with a whole square in it you would not really know.

Yeah. Whether that point is a stationary point. Yeah, yeah, yeah. So it is a test to be, to subject a point x to. We have not discussed any algorithm. Okay. So this is, it's kind of a, you know backwards in the sense that I am not giving you an algorithm but I am giving you the convergence like you know when you decide whether you want to exit the while loop of optimization.

You check is grad, is norm of grad f small, is number of iterations exhausted and then you quit. Similarly here you have one instead of grad f equal norm grad f being small here you have these set of conditions which have to be checked. But I have not told you what is in the body above it. right. So, once we finish this discussion on first getting geometry and algebra into sync ok, some techniques are there for that.

After that we will talk about one algorithm and the only algorithm that we will discuss in this course for constraint optimization which is called the projected gradient descent method. This projected gradient descent method is a algorithm that tells you how to move from x to $x + s$. and how to exit that algorithm is what we first spoke about. It's possibly the simplest constraint optimization algorithm.