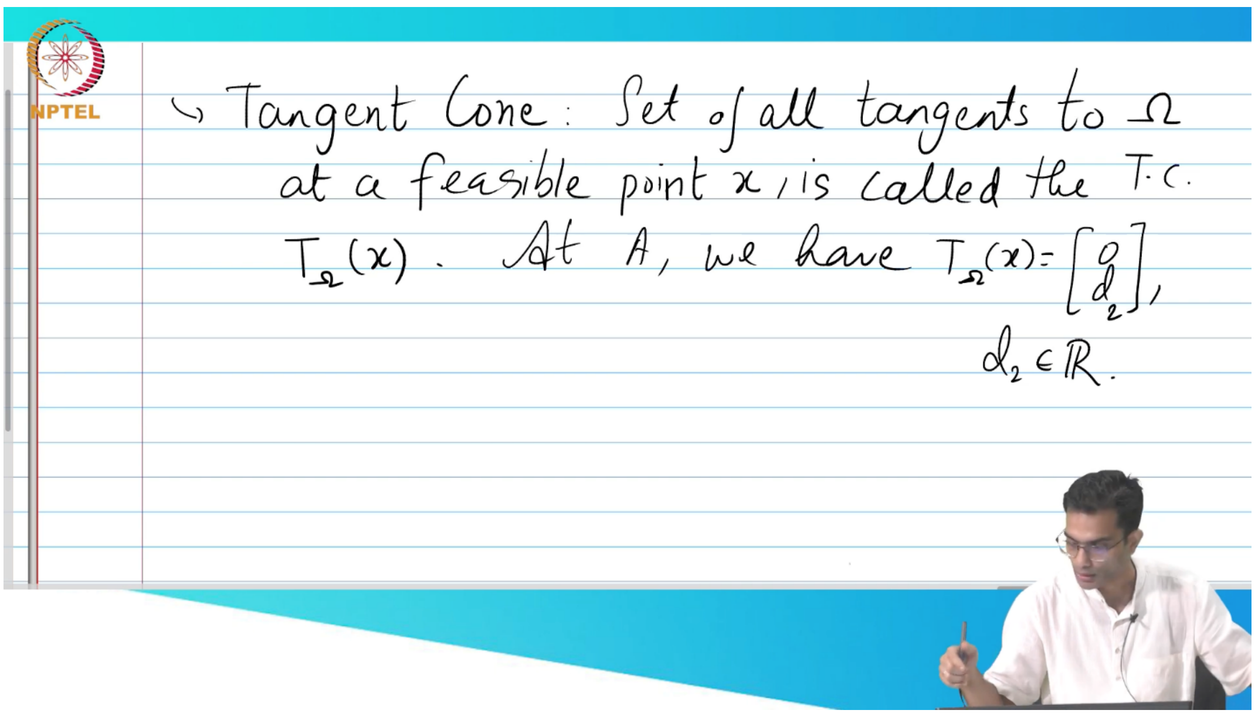
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 10
Lecture - 67

LICQ conditions

That finally, brings us once I manage to scroll this up over here, ok. Finally, we can define a tangent cone. A tangent cone, what do you think would how how would you define it? You have got all of these tangents now what should we do? The set of all tangents is going to define a tangent cone, ok. Set of all tangents to σ at a feasible point x is called the tangent cone and the notation for it is this, ok. So, it has everything that we need in it. It has the feasible set in it, the point x is in it and it is a collection of all of these tangents that are being derived from the feasible sequences.



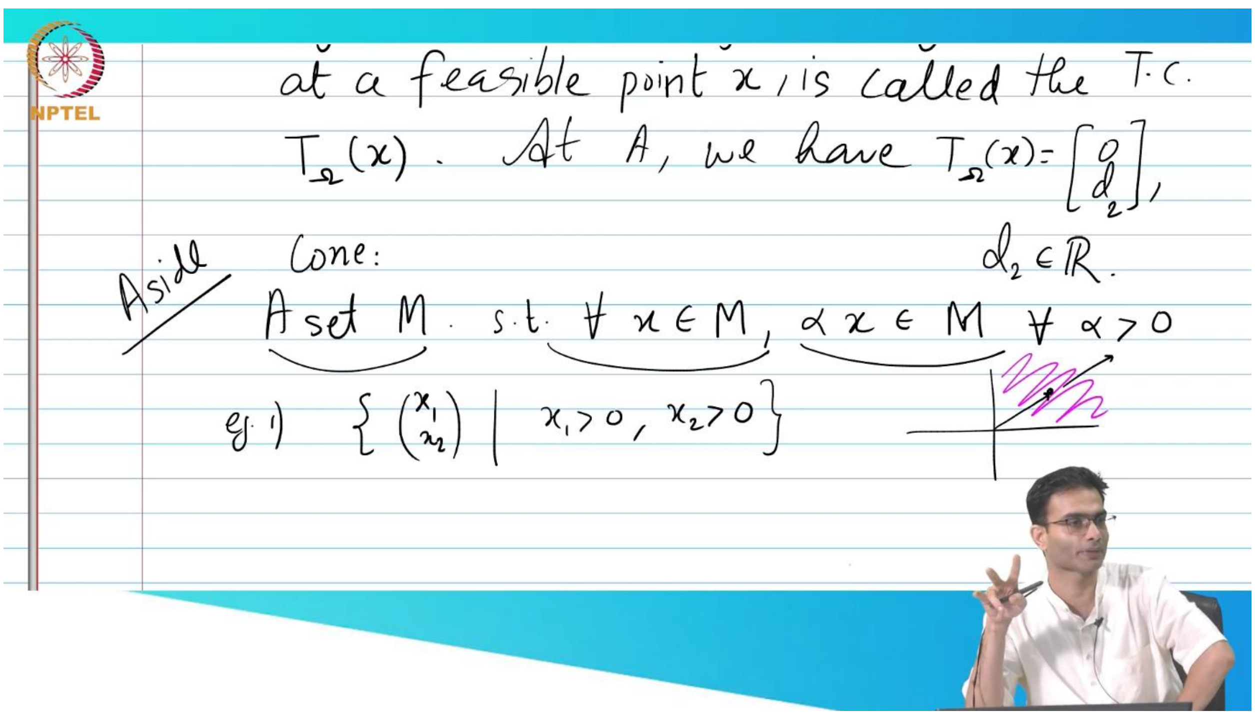
This is my tangent cone, ok. So, for the problem that we have been talking about how will you define the tangent cone? We have two sequences, right? We had the blue arrow and we had the pink arrow. So, if you were to define give me a precise definition of this tangent cone. So, at point A we have what would it be? It could either be the up arrow or it could be the down arrow.

Any sequence I take is going to give me either of these two arrows. So, how do I capture this? If I were to parameterize this I could simply write this as this, this captures plus minus, right? Now you are thinking of unit vectors, some of you are thinking of unit vectors. Why is this $0\, d_2$ and not 0 ± 1? If I had defined, so it depends on how you define your sequence over here, right. This scalar sequence t_k, the scalar sequence t_k if it was the L_2 norm of $z_k - x$, then the tangent vector that I get, either the blue or the pink arrow is going to have magnitude 1. Now, supposing I decide to choose t_k as equal to 100 times $z_k - x$, 100 times the norm of this.

It satisfies the definition because it tends to 0 or instead of 100, $1/100$ whatever it is still a tangent, right? And what is the definition of the tangent cone? The set of all possible tangents. So, that is why I could have magnitude 1, I could have magnitude 100, $1/100$ whatever I want rational irrational whatever. So, that is why when I put all of these guys together I get a tangent cone which is like the way I have defined it over here. It is not a vector space it is a cone, a cone is a cone or vector space. Is a cone a vector space? Because.

So, if I am exactly. So, we will come to this a little bit later the way I define a cone, right. If an element lives in the cone, then a positive scalar times that vector should also be in that set, but if I put negative I am out of it, right. So, it is not a vector space in that but idea is similar. So, right, I think one thing I forgot to mention over here this feasible obviously, this is tied to the definition of a cone, this scalar should be positive.

So that the flipping of sign does not happen, yeah. Well ok. So, let me come to the definition of the cone, I will revise the definition of the cone and that should make it clear, ok. So, so far the definition of tangent cone is at least clear start. So, how do we do it? Start with feasible sequence, feasible sequence is fine then I define another scalar to help me to get a limiting direction of a feasible sequence that gives me a tangent. The set of all tangents gives me a tangent cone.

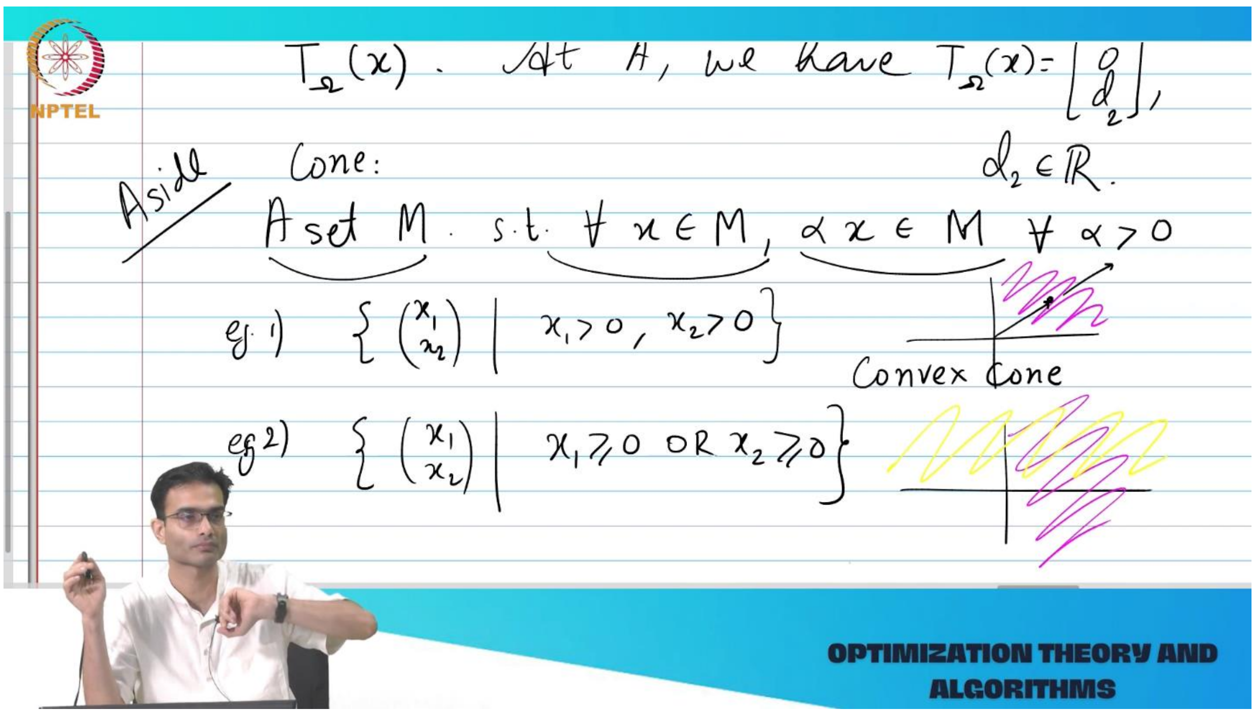


If you if you look back at your notes when we define the set of linearized feasible directions, did I get the same set $0, d_2$? Yes, when I defined the constraint in the normal way as $x_1^2 + x_2^2 - 2$ and I did the linearized feasible directions that is exactly the set that I got. So, there is an agreement there between algebra and geometry over here, but when I squared the constraint I got, algebra said any direction is possible and geometry said geometry is going to be unchanged.

Why? Because I can square the constraint, my feasible sequence will still remain the same. It has to be on the circle whether I square it, cube it, whatever, the feasible sequence I mean the feasible set will remain the same. So, that way this definition of tangent cone is immune to all of these funny tricks that I do with algebra, right. So, that is this is what we needed to arrive at, ok. Now let us there are some questions about cone.

So, let us talk about cone as an aside. That is the definition of a cone. So, it is a set M such that for every x that belongs to that set $\alpha \times M$ also belongs to that same set. Sorry, this should be $\alpha \times x$ also belongs to M and α should be positive, right. So, let us you can take two examples very quickly for example, x_1, x_2. So, I am talking two-dimensional plane, ok.

Supposing I say $x_1 > 0$ and $x_2 > 0$. So, if I want to sketch this what does this look like? First quadrant. First quadrant, right. So, I am basically saying this region. So, what are we saying pick some point over here, if I multiply this by 2.75 will it still be in the pink region? It would be it will just be like this, right. Can I multiply it by minus 1? That is not allowed as per the definition of the cone, right. So, this is a cone for sure because it obviously satisfies the definition over here. We can further qualify this. Is there some special property you can infer in this cone? Can we talk about convexity? Is this a convex set? The linear combination if I take any two points does it live in the same set? Yes.

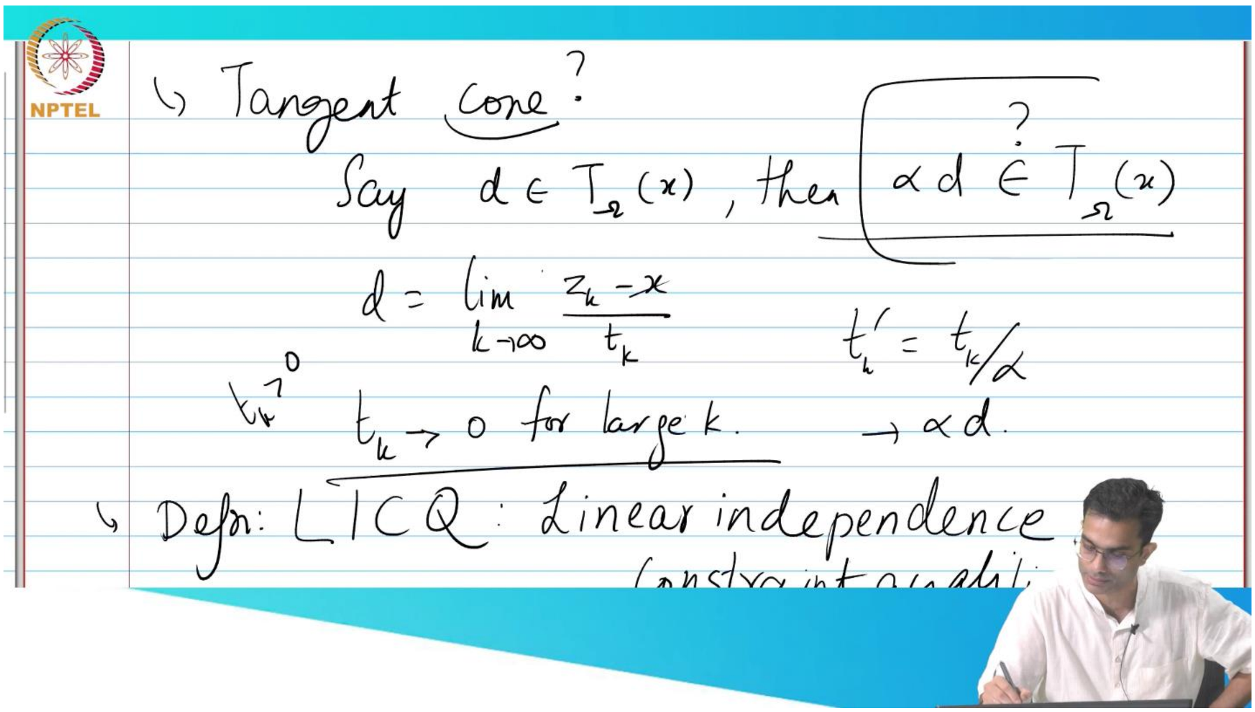


So, this is actually an example of a convex cone, the pink region, ok. Let us just to make sure that you do not get the idea that all cones are always convex let us take another example. I am going to define a new set like this. So, $x_1 \geq 0$. So, $x_1 \geq 0$ is going to be which region? The right half-plane, right? It is going to be this and $x_2 \geq 0$ is going to be what? It is going to be the upper half-plane and do I have an and or a or condition? Or condition.

Or condition. Therefore, what is the set? This quadrant number 1, 2 and 4. The third quadrant is gone, right? So, now, you can see that if I take some vector, let us say I take a vector over here, this belongs to the set. Does 2 times this belong to the set? Right, you can see take any positive scalar it will live in this. In this particular example of the blue arrow, if I do minus 2, will it work? It works, but let us not push our luck because if I take a vector like this, then that minus 2 will not work, ok. So, this is an example of a non-convex cone, ok.

So, now having revised what we mean by a cone, let us answer the question that why did we call this thing or why is the tangent cone a cone? Is it a cone?

So, we just have to use the definition of a cone. So, say that some vector d belongs to the set T is the cone. Then we have to ask does $\alpha \times d$ also live in the same set, that is the question. So, remember d was given as the limiting direction as k tends to infinity, $\frac{z_k - x}{t_k}$. This is the definition of your tangent.



How do you think we can prove this statement? He has the right answer. Remember the only property of t_k was what? t_k should tend to 0 for large k, that is the only property of t_k that we cared about. So, if I instead of using this t_k, supposing I define a $t_k{}'$ as equal to $\frac{t_k}{\alpha}$, still fine, right? And if I define it as $\frac{t_k}{\alpha}$, the vector that I get will be α times d and therefore, this also is this is the same vector, but I just change my definition of t_k. So, along with the fact that t_k is greater than 0, you can see that this object that we have defined is a cone. In a few moments we will start doing some further geometry with this which will convince you that this is a cone. I will draw some diagrams.

But at least if you go from the dry algebraic definition, you can see that this is a cone. A cone because take a vector, its multiple is also in it. That multiple, the freedom to choose that multiple, comes from the fact of the freedom in defining my t_k, that is all. Yeah, at a point, no, we just did that, right? So, for example, if you look back here, where did my famous diagram?

Yeah, at that same point over here, I was able to define at least two sequences and therefore, at least two tangents. I have a sequence z_k, I have a sequence w_k, correct?

So, in this case, but you agree that I have multiple tangents. I have multiple tangents and the tangent cone in this case is an entire line. I mean, in 2D, there is not much more I can visualize. If this were for example, a sphere, then the tangent cone would be like an entire plane, can think of it like this and then it will get more and more complicated, but in 2D this is about the maximum I could do.

Yeah, yeah, do not forget cornetto for now, I mean. Yeah. So, just keep this the definition of cone that we said over here, right? This is what you have to keep in mind. If an element belongs to that set and it is a positive multiple, then you call that set a cone. So, now it is not a vector space, right? What is the how would I define a vector space? That any linear combination of the basis elements will also live in the same space, including the origin, but here every linear combination is not allowed. I am only allowing you to take positive, positive combinations.

So, you can come up with pathological examples like that, which in which case it turns out to be it will be like a one-dimensional vector space, right? So, what I defined over here $0, d_2$, d_2 lives in all $\mathbb{R}$, this is, you can think of this as a one-dimensional vector space, but this is by chance, we cannot make a claim that this is a vector space. No, the tangent cone idea is valid for any constraint. Here we have taken this example. That would be the exterior of the thing.

Interior of the circle, right. There will be no multiple. There will be multiple feasible. Different tangent. Correct, yeah. So, you will get a whole bunch of different tangents, right?

But the idea is. Idea will still remain. But then you are violating the basic definition of a feasible sequence, property b. If it oscillates. It tends to the point you are saying, but it oscillates about the point, right?

So, then in that case. So, in that case the definition of the tangent is a little problematic, right? Because. Yeah. What I would do then is I would split that sequence into two sequences, keep all those points that are staying on one side of the point to be sequence 1, sequence 2. Then I can define two tangents and then putting it together, I can define a tangent code. If you are talking about points living in the interior, then there is no oscillations all on one side. It oscillates like this, it reaches x, but then whether it reaches x from down or from up.

Correct. So, that is what I am saying. In that case, I would split the sequence definition such that I get one part that has a kind of monotonic behavior. Then from each of those, I can define a tangent direction, and then put it together. Finally, I am putting it all together in the tangent cone. That is what I care about. So, may as well make my life simpler, right?

So, we have got, we have reached all the way to the tangent cone. Now comes the part where I connect the whole idea of tangent cone and algebra, ok. So, let us note that down over here. So, we have a new definition. So, this is called LICQ. Remember in the previous section I had mentioned that there is something called constraint qualification that we have to do.

So, this is that constraint qualification, which is called LICQ, which is Linear Independence Constraint Qualification. What does this say? So, it says it is a little bit of a long definition. Let us write it down. So, given a feasible point x and the active constraint set A_x, remember what does the active constraint set consist of? Equalities and inequalities, which are active, yeah.

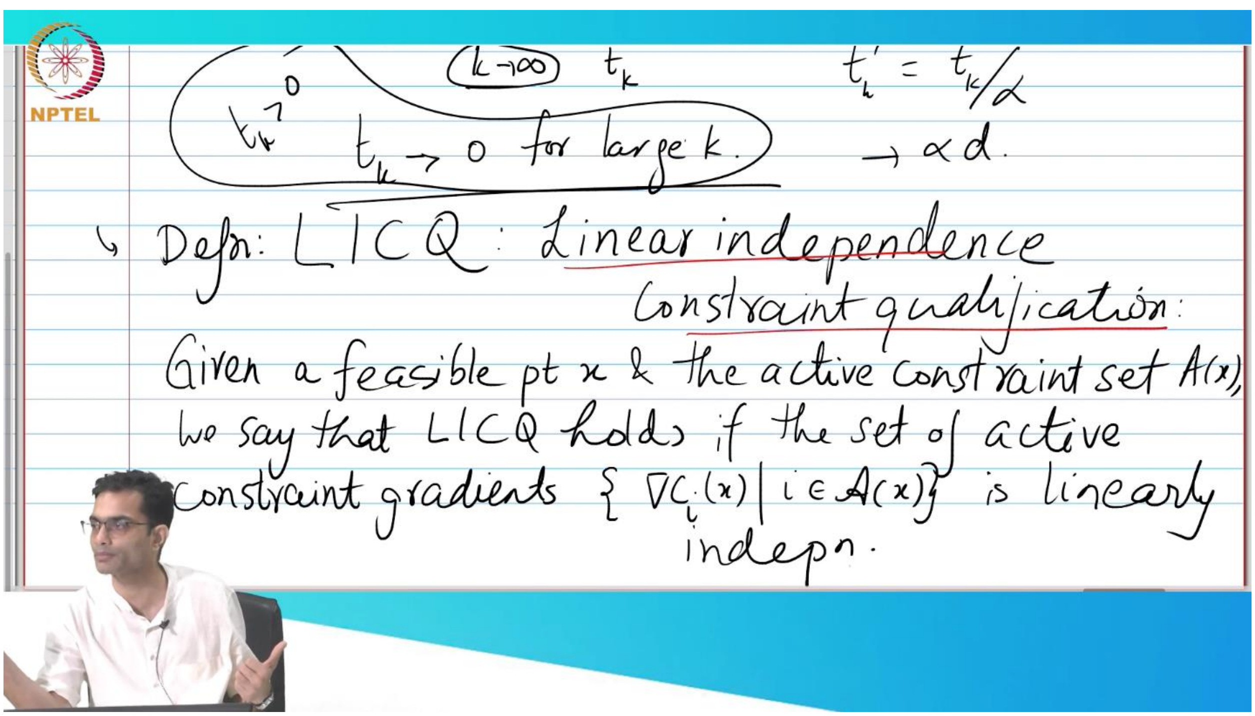
No, for testing the tangent cone, I will always have this question mark when I ask. I am going to choose α to be positive only; then it is in the definition of a cone. Which limit? The limit is only coming from here, from $k \to \infty$.

I would not do that, ok. Right, but keep in mind one thing: when I choose t_k, I have these two properties that t_k should satisfy. t_k should tend to 0 for large values of k, and t_k should be positive. So, I cannot choose a negative t_k. So, at that point, I would stop and flip the side, ok. So, then I do not run into that issue. Fine, that clear to everyone? Ok.

So, the active set, everyone knows equalities and inequalities which are active. So, given a feasible point x and the active constraint set, okay. We say that LICQ holds if the set of active constraint gradients, so what are the active set? These guys. ∇C_i, these are the gradients and, of course, i belongs to the active set. If this set is linearly independent, we say LICQ holds.

So, I am just making a statement over here to define what I mean by LICQ. So, we will just go step by step. So, LICQ simply is Linear Independence Constraint Qualification. What does it say? When do I say that LICQ holds? It is a yes-no question. How do I answer it as yes-no?

At a point x where I am, I look at ∇C_i for all the constraints which are active. What are the constraints that are active? Equalities are always active. Some inequalities may be active. I look at the gradients. Now gradients are going to give me, for example, in the example that I have, a gradient is a two-dimensional vector. For all the gradients, I will have a bunch of two-dimensional vectors. Can I investigate whether or not this set is linearly independent? Yes, I will get a yes or a no answer.



If the answer is yes, then we say LICQ holds. If I get no, we say LICQ does not hold, ok. So, this is just our definition, we are going to do something useful with this definition in a moment. Ok, I ran out of pages, ok. Yeah, yeah, I mean if there are more gradients, more constraints than the dimension, then you can immediately tell.

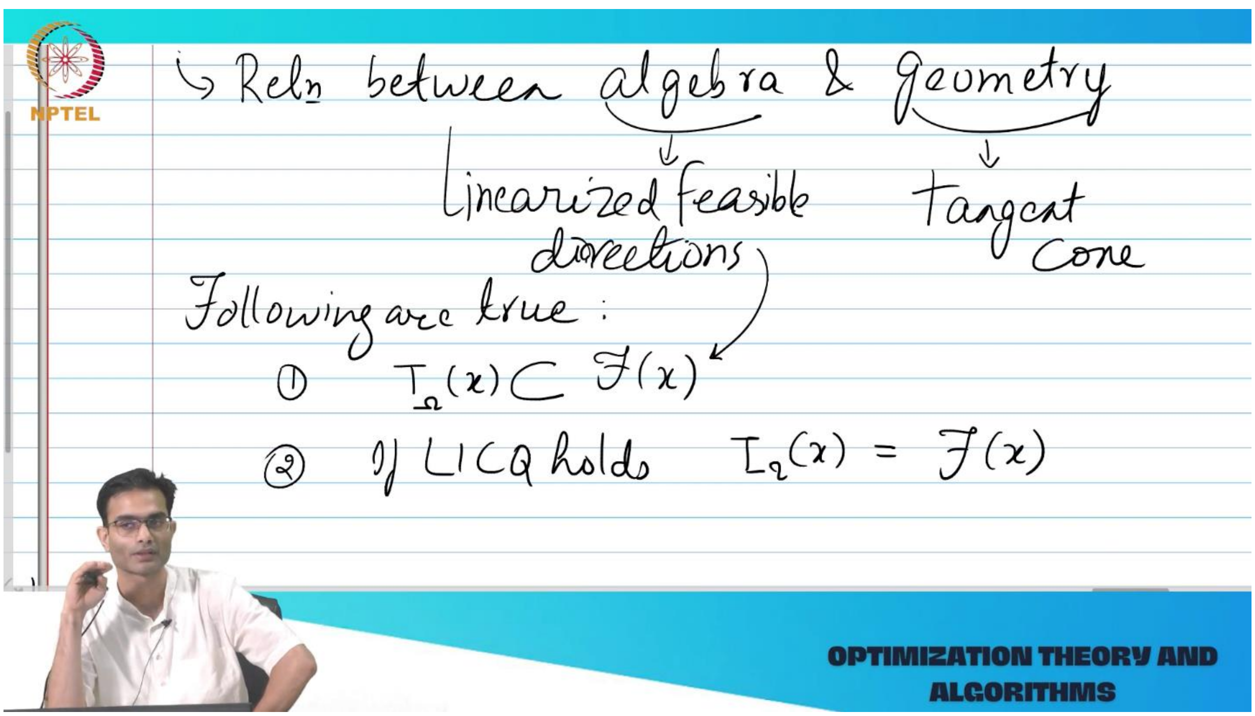
Okay, so what do I do with this LICQ business? So, this is sort of the main thing that we are heading towards, the relation between algebra and geometry, right?

So, let us say, let us write this down. Algebra and geometry. This is our casual way of putting algebra. I am going to call it as what do I mean by algebra? Linearized feasible directions. What do I mean by geometry now? Tangent. Tangent cone, exactly, ok.

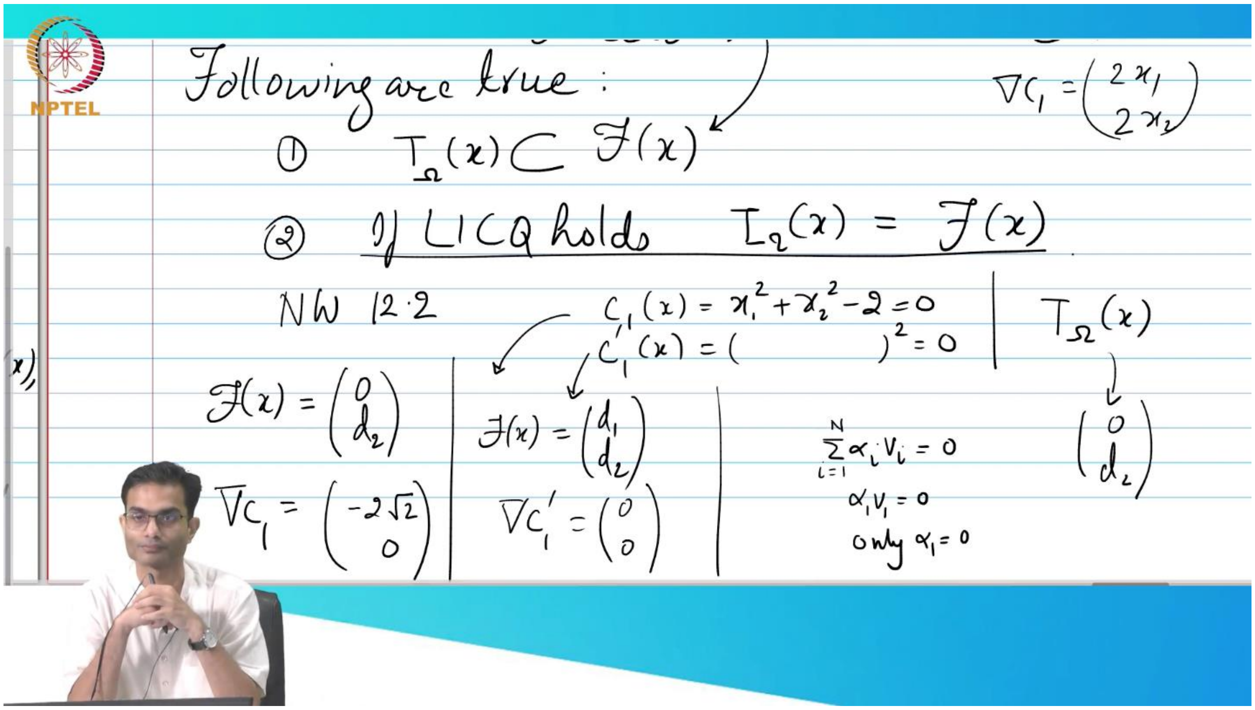
So, the following are true statements: The tangent cone is a proper subset of the linearized feasible direction. Remember, we had given it a fancy F symbol, right?

So, that is this guy. So, I will not prove it, although the proof is quite easy. The tangent cone is a subset of this. This is not very interesting. The very interesting thing is this.

If LICQ holds, then these two are equal. The tangent cone becomes equal to this linearized, the set of linearized feasible directions. In other words, if this LICQ holds, algebra and geometry will give me the same picture, which is what we were after, ok. So, this is, I think, Nocedal and Wright has this in fact as a lemma 12.2, if you want to look at the proof. But I mean this is really what you want to check, that if LICQ holds, the tangent cone and the set of linearized feasible directions agree.



And having sort of written this without proof, let us go back to our example where we had that problem and see: Does this help us out? Ok, what did we have earlier? If you remember, our constraint was $C_1(x) = x_1^2 + x_2^2 - 2 = 0$ and $C_1{}'(x) = (x_1^2 + x_2^2 - 2)^2 = 0$, right? That was what we had. When I had $C_1(x)$ like this, what did my set of linearized feasible directions turn out to be? Does anyone remember? We used the definition of $(0, d_2)$, this is what I had got $(0, d_2)$. Ok, when I chose this guy with the whole squared, what did my set turn out to be equal to, right?



So, this turned out to be (d_1, d_2), right. What did my tangent cone at that point turn out to be equal to? We just computed it; this turned out to be $(0, d_2)$. Right? Now, what was the gradient of C_1, does anyone remember? So, in this case, for C_1, $\nabla C_1 = (2x_1, 2x_2)$. So, when I substitute $\left(-\sqrt{2}, 0\right)$, what does ∇C_1 turn out to be equal to? $\left(-2\sqrt{2}, 0\right)$, right? And in this case, when I calculate $\nabla C_1{}'$, what do I get? The whole thing squared, I take the gradient, and what do I get? $(0,0)$.

Now, this is a little bit of a funny thing to ask: when I give you a set of n vectors and ask for them to be linearly independent, what does it mean? So, I give you a set of vectors v_1 to v_n and ask you, what is the definition of linear independence? So, we will say what? $\alpha_i v_i = 0$ happens only when all $\alpha_i = 0$, right?

Now, this summation, I went from 1 to n, right? Now, supposing I make $n = 1$, it is legit. So, $\alpha_1 v_1 = 0$ should happen only when $\alpha_1 = 0$. That is the meaning of linear independence for a set of one vector. Now, this is my set of one vector $\left(-2\sqrt{2}, 0\right)$. Is it a linearly independent set? It is because the only case in which $\alpha_1 \nabla C_1 = 0$ is going to happen only when $\alpha_1 = 0$, right?

So, this passes the LICQ test. This vector $(0,0)$, $\alpha_1 \nabla C_1{}' = 0$ for any value of α, including 0. So, this one, this set of which consists of just one vector, this is not linearly independent. So, it is funny to talk in terms of just one vector, but it is correct, right? So, this does not satisfy LICQ. Now, when it did not satisfy LICQ, we can see that we should not expect the tangent cone and this linearized feasible directions to be the same, and that is correct.

The tangent cone is telling me, and it is only agreeing in the case where LICQ holds. When LICQ does not hold, we do not expect it, right? So, see, it is very neat: we managed to tie together something from geometry and something from algebra, right? It is a little funny that our summation boiled down to just one term, but once you get past it, you can see linear independence. Any questions on this?

Ok, so I want to mention that LICQ is one type of constraint qualification. In the literature, there happen to be several other different ways of constraint qualification, but in this course, we are going to talk about one of them, ok. And you notice that in terms of programming or coding, this is not something very difficult to check. If you wanted to check whether or not a set of vectors is linearly independent, how would you do it via code? You can put all of those vectors in a matrix and look at the rank of that. So, it is easy enough for you to implement and check whether or not LICQ holds.

If LICQ holds, what is the great advantage? I mean, you can look at all of this and say, "big deal," how does it help me? It helps you because I do not have to start constructing tangent cones at every single point. Constructing tangent cones seems to be a lot of work, right? First, get a feasible sequence, then get a limiting direction, then put all of them together. Whereas if I know that the tangent cone equals the linearized feasible set, to get the linearized feasible set, what do I need to do? Just half-spaces, right? I mean, $\nabla C_1^T d \geq 0$, whatever, right? So, it is relatively easier for me to come up with a direction, right? So, in most cases, that is the easier thing to work with rather than tangent cones, ok, but the real thing that will not fail you is the tangent cone, right.