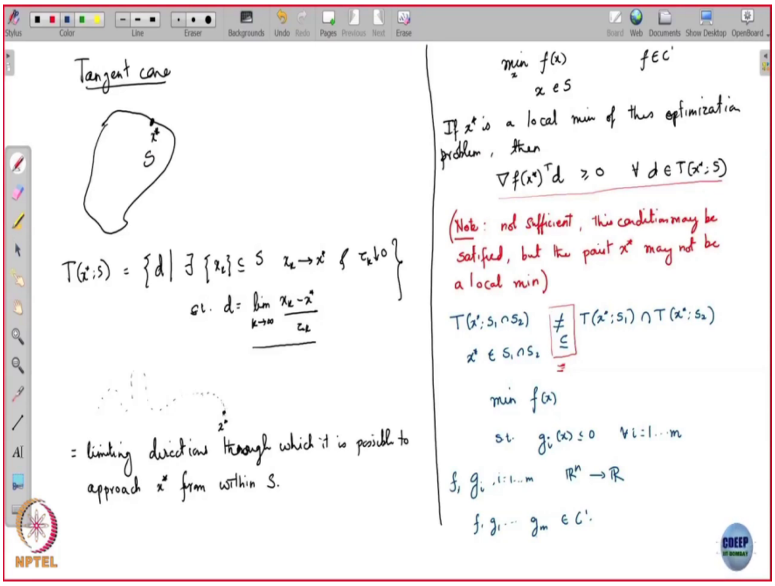
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

Lecture - 15A
Tangent cones (continued)

(Refer Slide Time: 00:23)



Ok. In the previous class, we defines this quantity called the Tangent cone, ok. So, the tangent cone remember how the tangent cone was defined for two things; one is a set, the set S and say a point, point x star that lies in the set. So, for example, if this is a point x star and this is your set S; then the tangent cone T of x star with respect to x was defined as those directions d, by which it is possible to approach the point x star from within the set S, ok.

So, d is. So, tangent cone is those d's, such that there exists a sequence x k that lives in the set S, x k converge into x star and a tau k that decreases to 0, such that d can be written as this limit x k minus x star divided by tau k, ok. So, what we learnt was that, if you look at this sort of sequence that is eventually converging to x star. So, here is my point x star suppose and my, this is my sequence that is it dances around all over and then eventually gets converges to x star.

Then this ratio limit as k tends to infinity x k minus x star divided by tau k; this ratio is actually capturing the tangent to the trajectory of the x k's. With as x k equals to x star, it is this is the this is the tangent to that trajectory and my intuition for that I explained was that, you should think of x k minus x star as the distance between x k and x star and tau k as a unit of time.

So, this is like a velocity right; it is the rate at which x k is approaching x star. So, it will eventually become tangent to the to be precise curve that x star trajectory that x star x k traces, alright. So, these were the. So, this as you can think of the tangent cone as the limiting directions, limiting directions through which it is possible to approach x star from within S, ok.

And what was the significance of this tangent cone? The significance of this tangent cone was that, it captured a very general necessary condition for optimization. So, if you have a general function f, just that f has to be differentiable. So, let us take continuously differentiable and x and you are minimizing this function f over all x in a set S; S can be any kind of set can be a polyhedron, can be not a polyhedron, can be convex, not convex it does not matter.

Any set S and you are optimizing of an arbitrary differentiable function over it, then we had this result that x star; if x star is a local minimum of this optimization problem, then it must be the case that the gradient of f at x star makes an acute angle with all vectors d that lie in the tangent cone at x star with respect to S.

So, this is an extremely sweeping and general condition and it is probably the most general condition known to us, about that captures a necessary condition for a point to be a local minimum.

Unfortunately there are just for you to know, this is not actually sufficient. So, note that even though this is the most general condition, it is not in general sufficient, ok. There are counter examples that of where you have a function that is, where you have of a function and a set where this condition is satisfied.

That means this condition here, the underlying condition is satisfied; but the point x star is still not a local minimum, ok. So, it is not sufficient, there are. So, this condition may be satisfied, may be satisfied point x; but the point x star may not be a local minimum, alright. The other thing we saw about the tangent cone was that, the tangent cone is actually a treacherous object; means that you might think that, if you have suppose two sets S 1,S 2 and a point x star that lies in both S 1 and S 2.

And if you wanted to find the tangent cone with respect to the set S 1 intersection S 2 at x star; then this is not in general this set. So, it is not in general, you cannot construct it by simply taking the tangent cone of the point S star with respect to S 1 and with respect to S 2 and then taking the common region of the other two.

If you do this, you would in general get a much larger set. So, this is. So, the intersection is actually just a subset of this larger set, ok. So, but still having said that, this turns out to be our; it turns out that in optimization this is still our best bet what we. So, the way optimization theory proceeds is that, it tries to see put conditions to make this equal, ok.

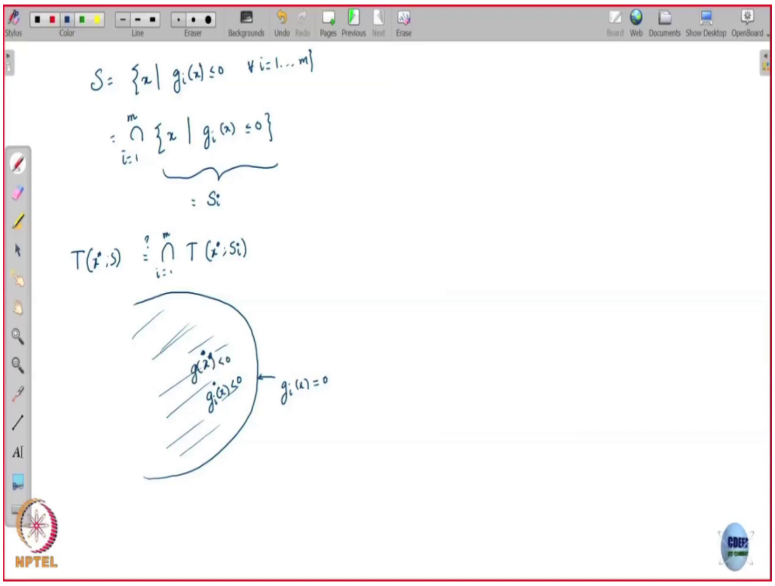
So, we have conditions to make, somehow make this equal and that is what I will try to discuss today, ok. So, what, so for this let us go back to a let us go back to a optimization, the form of an optimization problem. So, now, instead of taking an arbitrary set S and an arbitrary function like this; we will describe the set in a much more specific way. So, we will now write the set using its constraints.

So, consider now an optimization, optimization problem of the following kind; you are minimizing a function f, subject to constraints that look like this g i of x less than equal to 0, where i goes from 1 to m. So, you will notice that I have not used equality constraints here and that is for a reason I will bring those equality constraints in subsequently; but for the moment let us just focus on this particular type of optimization problem, right.

So, here you have f of x as the objective and g i of x is from i equal to 1 to m, these are your m constraints; each g i is and f itself they are all functions from R n to R, ok. So, each g i of x is a scalar function and we will assume that they are all continuously differentiable.

Let us assume that they are all continuously differentiable, ok. So, now what we will what we will do is, we will see if we can somehow get to an understanding of the tangent cone of this particular set, ok.

(Refer Slide Time: 09:15)



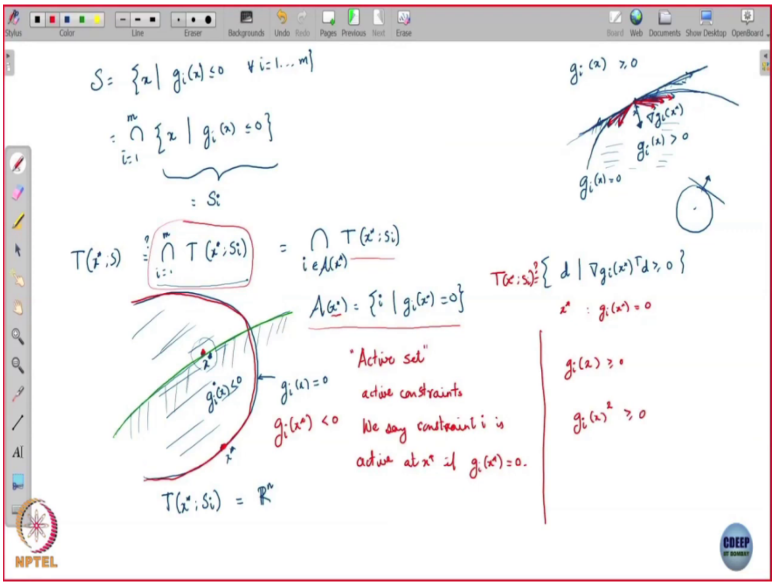
So, we know that if you. So, this here the feasible region now S, which is x such that g i of x less than equal to 0 for all i from 1 to m; this is nothing but the intersection of these regions, it is the intersection of these individual sets. Let us call these S i, right.

So, if I want to know what the tangent cone at a point x star in S is; then that, I can try to see, if I can get to this by looking at the tangent cone of x star with respect to each of these S i's, right. And the hope is that this would turn out to be the same as the intersection of all of these.

Now, we know that this is in general not true ok; but we will try to see when we can make this work, ok alright. So, now let us, now let us do a few simple things; now if you have a, what is a constraint like this g i of x less than equal to 0? If I have a constraint like this, where g is a continuously differentiable function; what sort of, what does that constraint look like?

So, that constraint if you see it is. So, here is the region that constraint describe. So, the boundary here of this region is say g i of x equals 0; the inner the interior here is g i of x strictly less than 0. So, if you have a point x star that lies here in the interior of this, for such a point g i of x star must be strictly less than 0. So, here that is that must be the case. So, it is a point. So, it is a it is a point for which g i of x star is strictly less than 0.

(Refer Slide Time: 11:31)



So, for this sort of point, what would be the tangent cone with respect to S i? Now, this sort of point lies in the interior of the set and we had seen this last time that, if you have a point that lies in the interior of the set; then the tangent cone is the entire space, right. So, you can approach the point from every possible direction in the ambient space, right. So, consequently for such a point, you must have that the tangent cone is actually R n.

So, if you if. So, what this means is? So, if out of these sets or out of these constraints g 1 to gm, if any of them are satisfied with strict equality at x star; then they do not, the tangent cone with respect to those sets do not appear, will not appear in this intersection, because the tangent cone there is just R n right.

So, this here is therefore, can be written in the following way; you can write this as the intersection over i in a set A of x star and I will define what A of x star is, where A of x star is nothing but those i’s for which g i of x star is exactly equal to 0. Because once g i of x star is strictly less than 0, the tangent cone is R n and it has it plays no role in the intersection, right.

So, in other words what, when I have to try to describe my the tangent cone with respect to my full set; what I, I have now brought the problem down to just this one part, which is that I need to now describe the tangent cone with respect to the set S i, right. So, I need to with respect to the set S i for those i’s that lie in the set A, A of x star; A of x star is what is called the active set.

Student: (Refer Time: 13:46).

Yes.

Student: (Refer Time: 13:49).

Yeah, so, T of x star with respect to S is ok; we are hoping that this can be somehow made equal to this intersection here, right. I am hoping that this can be made equal to this intersection; but then in this intersection if I look at any those i‘s for which gi of x star is. So, if I look at an i for which say g i of x star is strictly less than 0; then such an x star will lie in the interior right and then around that x star, I can fit a ball and also in the that will lie completely in the set.

So, the tangent cone with respect to that particular set S i would be R n right, at the point x star. So, that is why in when I look at this intersection here, those i’s the i’s for which gi of x star is strictly less than 0 will not count in this intersection; because they are all R n, right. So, what will remain is only the i’s for which gi of x star is exactly equal to 0, ok.

And those are what we call the, that is the set of such i’s they are called the active set or the i’s are themselves called the active constraints. So, we say that a constraint is active. So, a constraint is said to be active, if the at a particular point ok; we say the constraint i is active at x star, if gi of x star equals 0.

Now, to be this the active set ok, remember is a function of will change with x star; at a point x star, at this point the active set will be, the constraint that I have drawn here it is not in the active set, ok. And maybe another constraint like this, this may be in the active set here, this sort of constraint; this is in the active set, because it you know it x star lies on the boundary.

So, if I take this sort of constraint here, this is in the active set ok; but if I. So, and if, but if I change my point x star to this point here; suppose if I take this as my point x star, then this constraint will now become active. My earlier one which was not active earlier will now become active; whereas the and whereas this one will not be active anymore, the green one will not be active anymore, right.

So, whether which constraints are active really depends on the point x star ok; some certain constraints may be active, certain constraints may be inactive etcetera. But the point is once we, once I give you the x star; I can I know immediately which constraints are active and to. And if I have to look at the tangent cones, then all I have to do is focus on the tangent cones of those individual constraints, alright ok.

Now, let us look at one of these individual constraints now; let us see if we can make sense of the tangent cone of any of these individual constraints when they are active, ok. So, let us see suppose I have a constraint like this; let us for simplicity let me just change the direction for the moment, I will just write this as a greater than equal to 0, because that is easier to explain, ok.

So, suppose we have a g i of x greater than equal to 0 is my constraint, ok. Now, if I look, you what you are tempted to think is that, well the tangent cone and because g is continuously differentiable; what we are tempted to think is that, well I can capture the tangent by looking at the gradient of the function g i, right.

So, I am looking at a point x star which is on the boundary, ok. So, what I need to do, what I really geometrically what I need to do is, look at all these directions by which I can approach the point x star, all of these directions and eventually all the way till I get to the tangent surface, the surface that is tangent to this constraint at x star, alright. So, this is. So, the because I have written it as a greater than equal to, I have this g i of; this region is g i of x strictly positive, this boundary is g i of x equals 0 and so on.

So, what I want to do is to get to this tangent. So, geometrically this is actually correct. The trouble is, how do I get the formula for this? Means, I have to I need to, I have the right geometric picture here, yes I need a tangent and so on; but what is the formula for the tangent and what should how do I express that tangent using the function g?

Now, if you if I drew a circle for example here; you would know how to draw the tangent, well the you would know how what the tangent is and you would know what the formula for that tangent also is.

Well, what do I need to do; I need to look at the gradient here, the gradient will be pointing is going to be an is going to be an normal, look at the normal the direction, the direction in should be in the direction of increase of the function ok, that is where the gradient will point. And then using that I perpendicular to that would be my tangent, right.

The space that is that whose normal is this particular gradient is would be my, that would be my that would be my tangent, ok. So, let us see if we follow through on this, what you would think would be; how would you capture this particular space? Well, it would be the space that is perpendicular to the gradient at this point, ok. And now where would the gradient at x star point? See gradient of the gradient.

So, the this region here, this region here has g i of x strictly positive and as you get to the boundary, g i of x has become 0; the gradient should be pointing in the direction of increase of the function, right. So, the gradient of g i at x star will be normal to it will be normal to this boundary and pointing inwards, right.

So, this is where you would think, this is where the gradient would be. So, this here would be gradient of g i at x star. So, if and then you would think that, well what would be this space; then what would be the tangent cone in that case? Well, you would think that the tangent cone would be those d's that make, such that gradient of g i of x star transpose d is greater than equal to 0, right.

So, it is all these directions d, which make a acute angle with this gradient. So, all of these directions, up till the point where you become tangent; all of these directions are you would think are in the tangent cone.

Student: If you (Refer Time: 22:24) other way around them; we are going from this point to this point. So, (Refer Time: 22:28).

No, no. So, the unfortunately the convention over the tangent cone is that, it centers it at x star. So, it is x k minus x star. So, the origin has been shifted to x star. So, it is always origin form x star into towards x k, ok.

So, that is how you would define the tangent cone; that is unfortunately that is the been the convention, although I mean it is easier to explain it as the directions by which you can approach, but it seems like you are point heading towards x star, but actually if you see it is the other way around, ok alright.

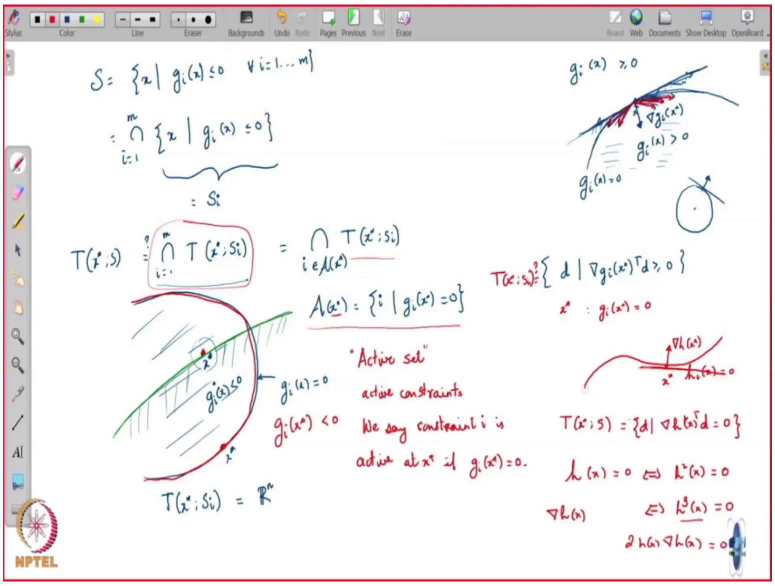
So, this is what you would guess would be the tangent cone; at a point x star, so what is this. So, just to summarize this; you have a point x star, such that g i of x star is equal to 0 and you would think that such for such a point, the tangent cone would be equal to this, this is what you would think. But the trouble is, this also is not true.

That means, so if you a simple case like this, where you have a point x star which is on the boundary; you have a differentiable function, you will know the you know what the normal is, you can calculate the normal and so on. And you would, yes you would think that well the tangent cone should be equal to this; the problem is that also is not true.

And the reason for that is this simple problem the which is the following. So, this constraint, if I look at this constraint g i of x greater than equal to 0 right; this constraint is actually the same, ok. So, this constraint is actually the same as the constraint g i of x the whole squared is greater than equal to 0, ok.

So, or more specifically we can even, I will it is probably easier to see this with the in the case of an equality constraint. So, maybe let us do this for an equality constraint and I will be able to convince you more easily.

(Refer Slide Time: 24:49)



So, suppose. So, this was for the case of a greater than equal to constraint; suppose if I had if I had a surface like this, if I had a equality constraint, suppose I had a surface h of x equal to 0 and I had I took a point x star on this surface. You would think that the tangent to this would be something like this; this would be the direction of the gradient of h at x star. And you would think that the tangent cone with respect to that sort of set would be d, such that gradient of h at x star transpose d would be equal to 0 right, almost following the same sort of reasoning as before.

The trouble is this is where the it starts begin begins to break down; because h of x equal to 0 is equivalent to h square of x equal to 0 and h square of x which is equivalent further to that any you know say h cube of x equal to 0 and so on. I can keep taking powers of h and I will continue to get the same set.

And so, what is the reason for this? The reason for this is you; this h of x equals 0 is the same as you know any power of h of x also equal to 0. But then what is the problem that this creates? The problem is that yes I am looking, if I look at the gradient of h of x ok, for the same set; for the same set I apply this particular formula; on the one hand I will get gradient of h, for on the other hand I will get, I will get something like this, I will get 2 h of x gradient of it. And what is h of x at that point?

Student: Zero.

Zero, which is equal to 0. If I apply it for this one, I will get 3 h square again gradient of it. You see the problem that has occurred. The problem is that, the same set of points, geometrically the same set of points can be represented in multiple ways using algebraic constraints. And when we say we have we know the formula for a normal and so on; all of that presumes something about the about how we have represented the constraints algebraically, right.

So, surely it can, this is not fair right; this is not acceptable that, the same set give will end up having multiple, will have very different formulae based on how you have chosen to represent it, right. So, this is the trouble. So, the trouble where is the challenge; what is the reason for the problem?

The reason for the problem is that basically the, there is a gap between the way we, the optimization is proposed as a geometric problem as a set of points over which you want to optimize versus how you want to represent those geometric points using constraints, right. So, the representation is which was coming in the way here.

Student: Sir, (Refer Time: 28:26)

But how do we know, what is the original h of x?

Student: Because that is what is taking place in (Refer Time: 28:32).

But how do you know that it cannot be reduced further? See the point is that it is the normal, the geometric normal is not, the issue is exactly this; the geometric normal is not necessarily the gradient, the gradient it could be zero, even though there is actually a normal to that side, geometrically there is actually a normal to that side and there is how will we know what the actual normal is from the formula right.

So, what I wanted to illustrate in this was that, there are basically multiple ways by which you can represent the same geometry.