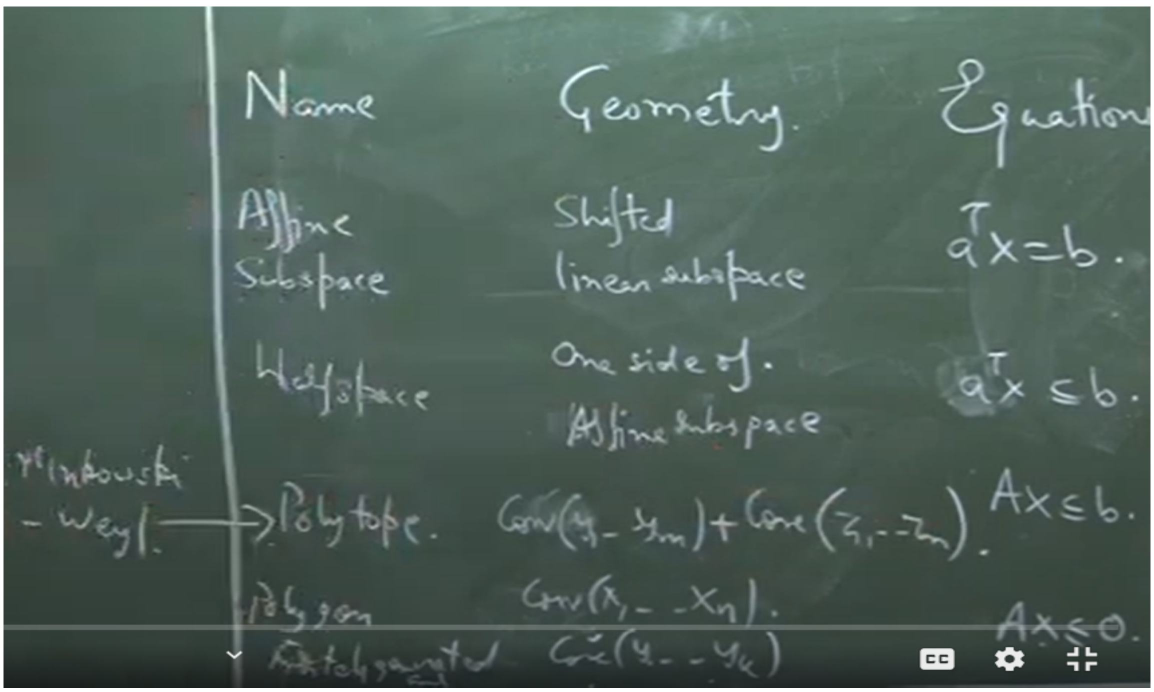
Linear Programming and its Applications to Computer Science

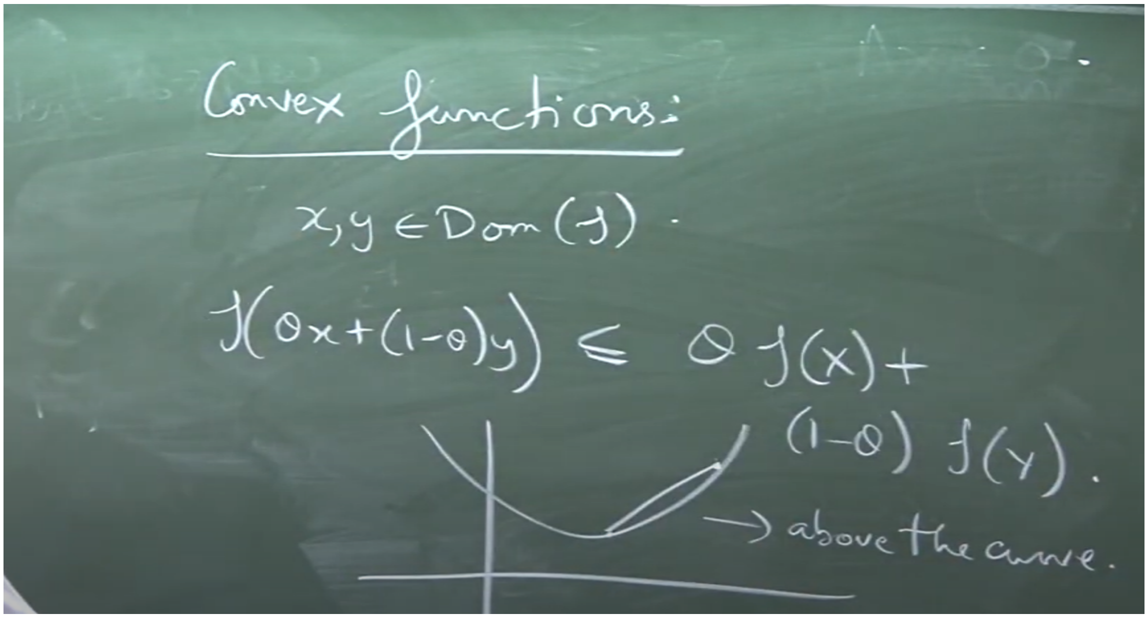
Prof. Rajat Mittal
Department of Computer Science and Engineering
Indian Institute Of Technology, Kanpur

Lecture – 15
Definition of Convex Functions

So, the punch line for all this is that for all the convex shapes which we saw there are these two ways in which we can look at these thing.



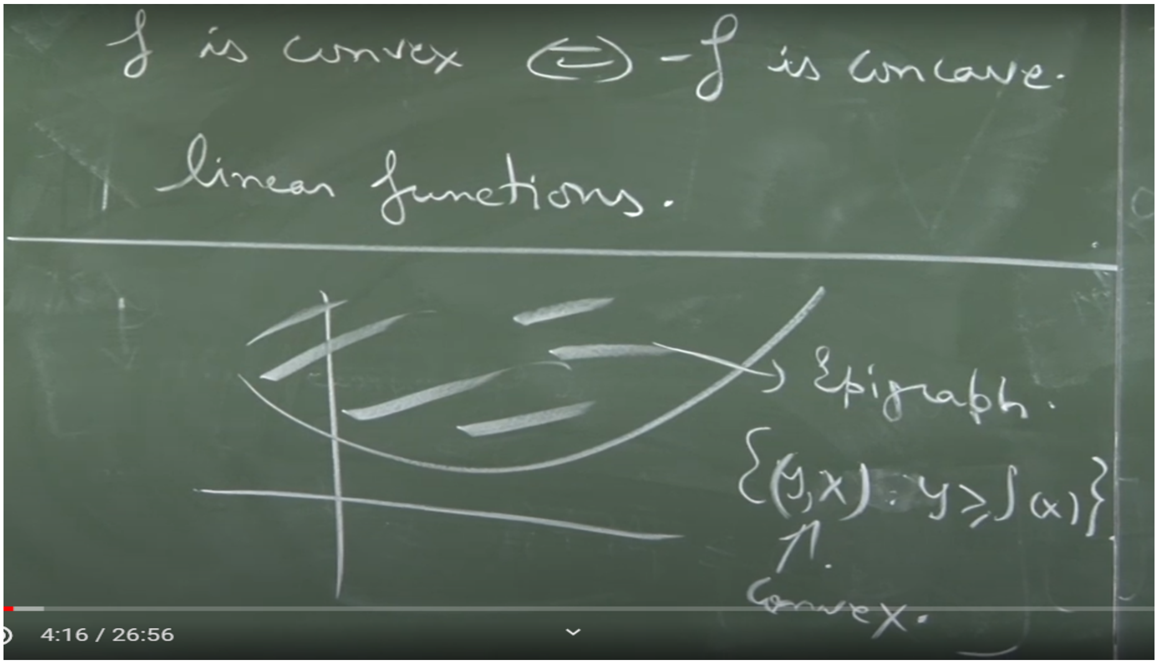
So, this equivalence is kind of nice the proofs are not really important for us, but it is good to see at least one proof.



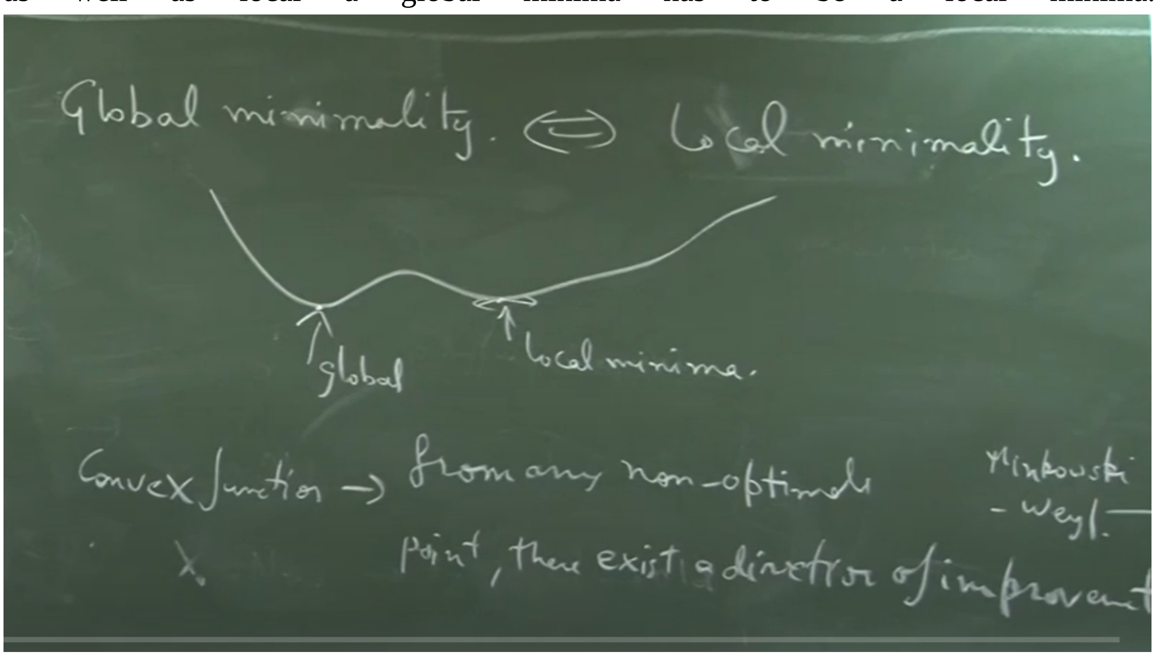
So, what will definitely remember is that these spaces have multiple ways to understand right. We can think of them as convex or conic combinations or we can think of them as in terms of inequalities. Now, we have this about the feasible set another thing we want to worry about is our optimization function.

And our optimization function is also special it is a linear function. In other words it is all it is a convex function as well as concave function. So, let me define what a convex function is. Given two points x y in the domain of f. a function is convex if in some sense. In some sense.

It is always good to view them in one dimension and if you want to view it what it says is that if you look at two points and draw the line between these two points it is above the curve. This is the intuitive definition of function being convex. And then and what are the functions ok and what are both convex as well as concave? Linear functions good that you know the geometry of it, but now we call them linear functions right. And that means we can use this because our optimization function linear program is a linear function that means we can use both that our function is convex as well as concave. But why is convex function nice or why is optimizing convex function nice? Just before that why is it called convex function? Exactly.

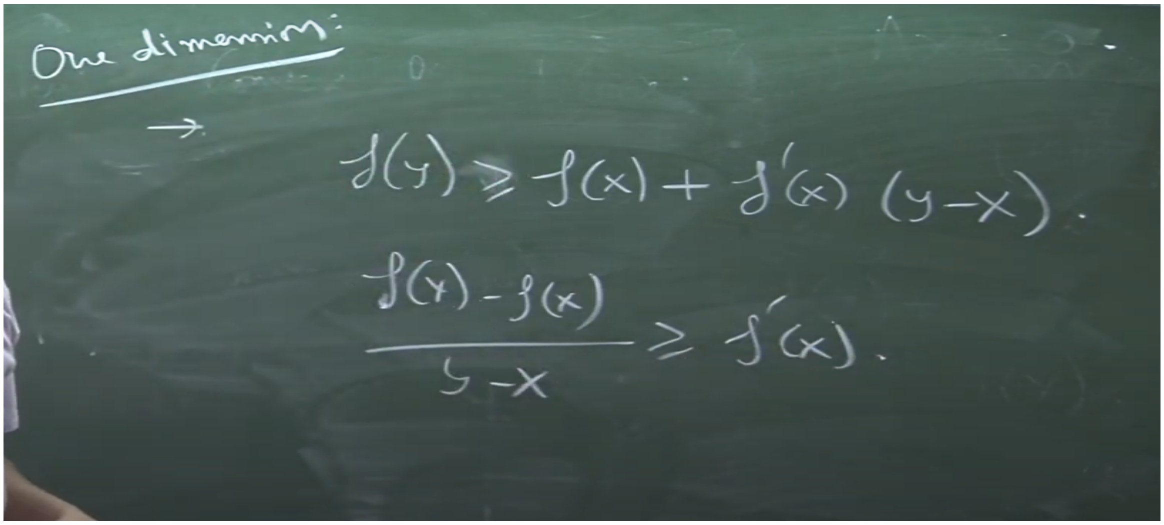


So, if you look at the shape of this. This is called the epigraph right. This is the set of all y such that all greater than equal to f x right. And for a convex function it is epigraph is convex. Now, we know what convex functions are and the main reason why convex functions are nice is because global minimality is same as local minimality. What does it mean is that in general let us say I have this function right. This function is a local minima. What does it mean local minima? If I move in any direction my function value increases. So, this is a local minima, but this is not a global minima. This is global as well as local a global minima has to be a local minima.

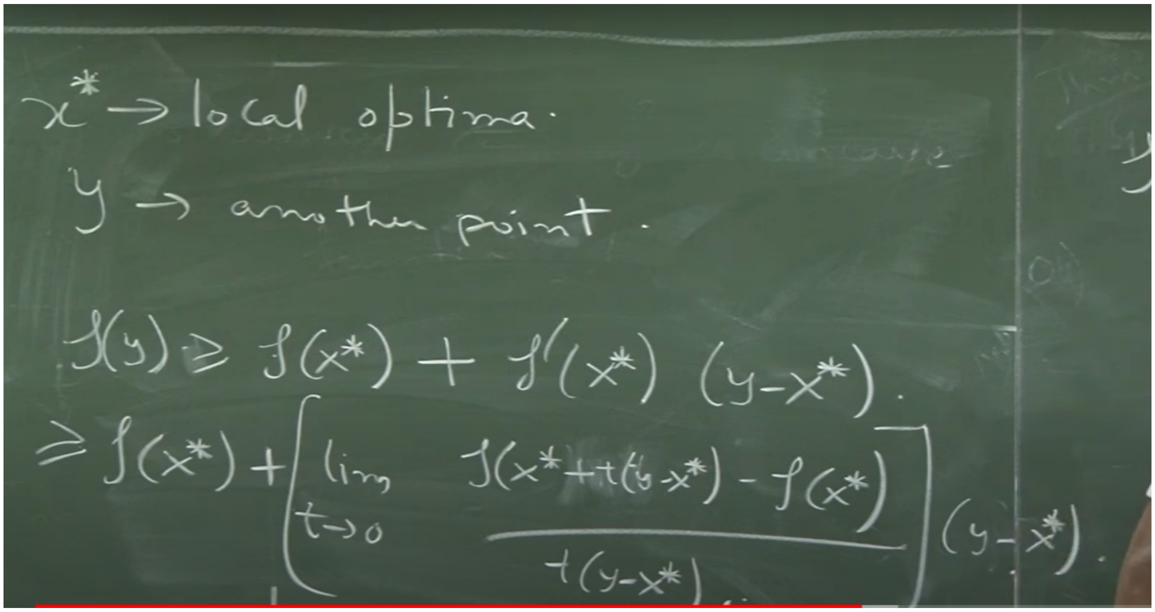


This is only a local minima right. And why is it bad for optimization? Exactly, while we are optimizing while we are changing directions going towards the right you know the best possible solution we might get stuck in local minima. In other words if you are optimizing or if you are minimizing a convex function from any non optimal point there exist a direction of improvement right. So, wherever we are in if we are optimizing a convex function we are minimizing a convex function we do not have to worry. There is a direction where we can move and move towards the right solution.

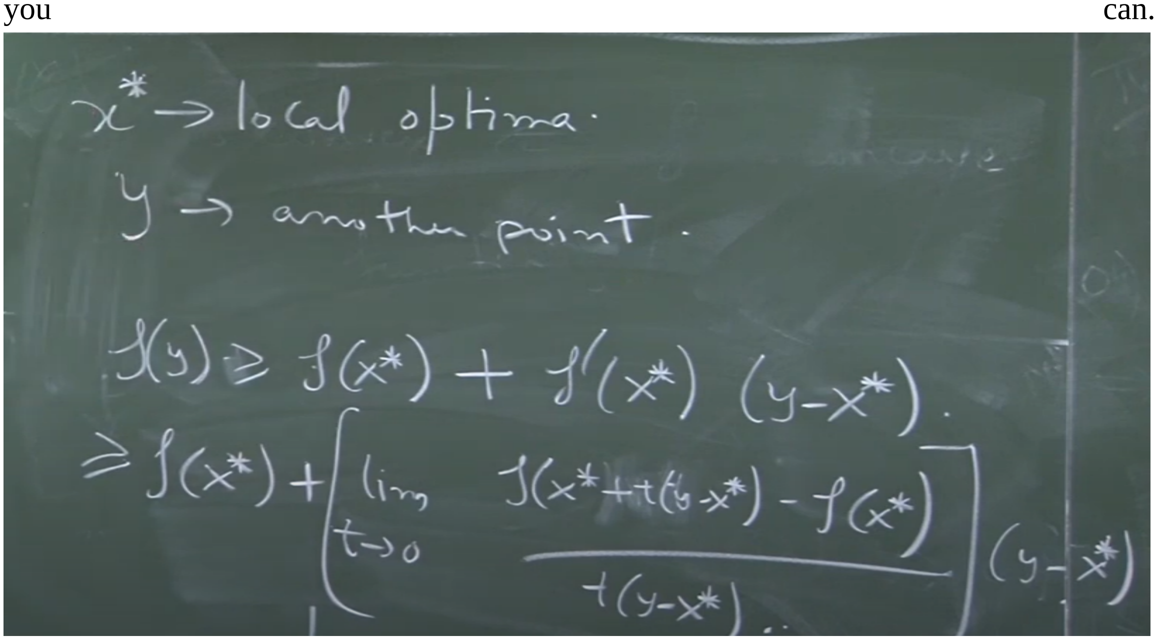
This is why there is so important convex function is so important. You are saying you are confusing between greater than equal to and strict inequality and. So, even the same direction you can move right without like you want your function value direction where direction of not bad performance is that what you want me to write. So, he is worried about the case simply put if this is your function then everything is a minima right and he is saying where do I move from here. Sorry So I think when we do this simplex method we would want to go on to vertices that are probably that is why I do not know why, but this is not a problematic technicality.



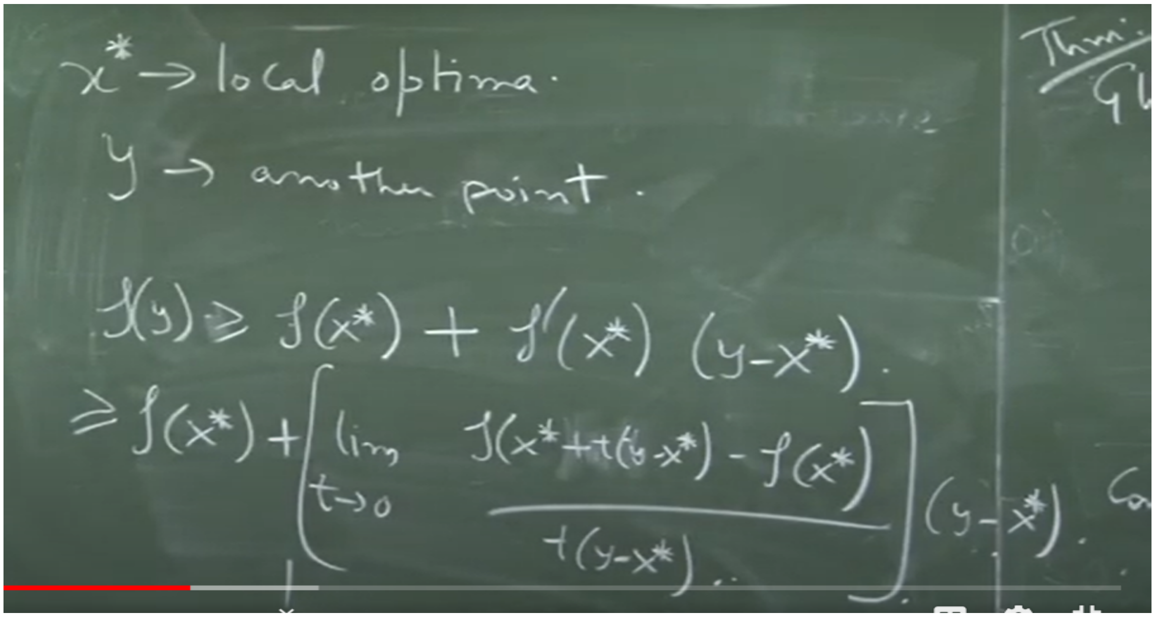
So, this is the important theorem for us and let us try to see a proof of it. So, I will show a proof for 1 dimension and a similar proof can be extended for multiple dimensions I want to keep my life easy and just give you an idea of how this thing works right. So, the first claim I want to make is that for a convex function I can have another definition this is the derivative at the x point x and if it is multiple dimension then this will be a vector this will be a vector of derivatives and then you take the dot product, but does this seem believable. So for us we will assume that they are defined by these nice inequalities and equalities the space is nice the function is for us the function is linear. So, it has good derivative, but in general you can assume that your function has nice derivative.



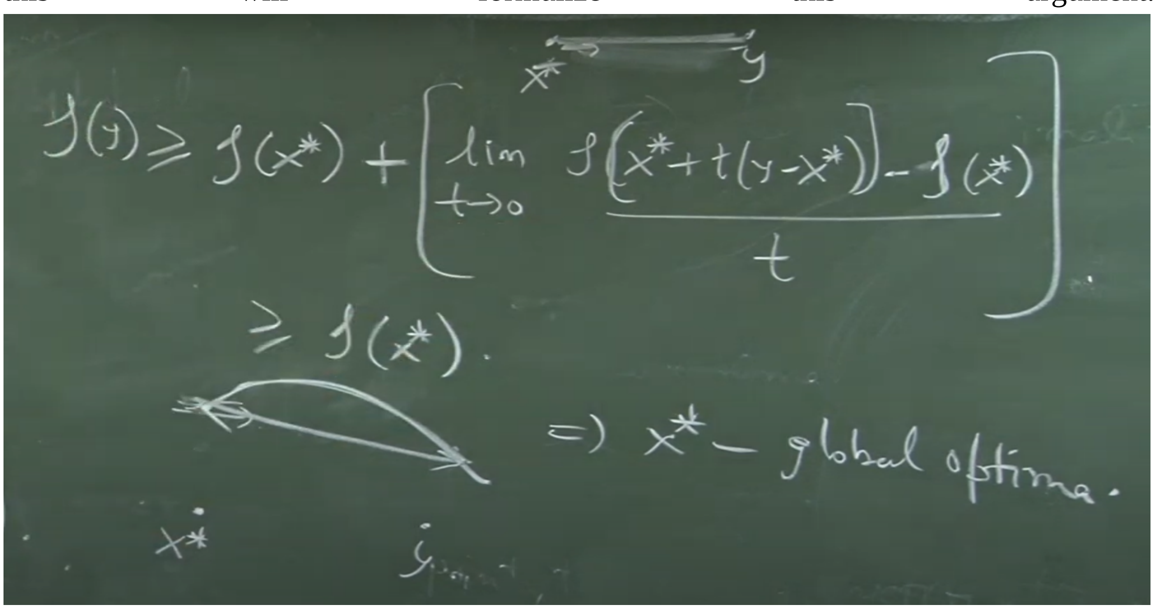
Again this is not like a 100 percent solid proof it is 1 dimension it is assuming things just to give you an idea of why such a thing is true for a convex function. So, I am not saying that takes care of all the cases and everything it is just giving you an idea of why this is working why local optimality local minimality without global minimality cannot exist for a convex function. Assuming the derivatives exist does this sound believable right other way to see it. This is saying that the line y to x is above the tangent. So, this you can.



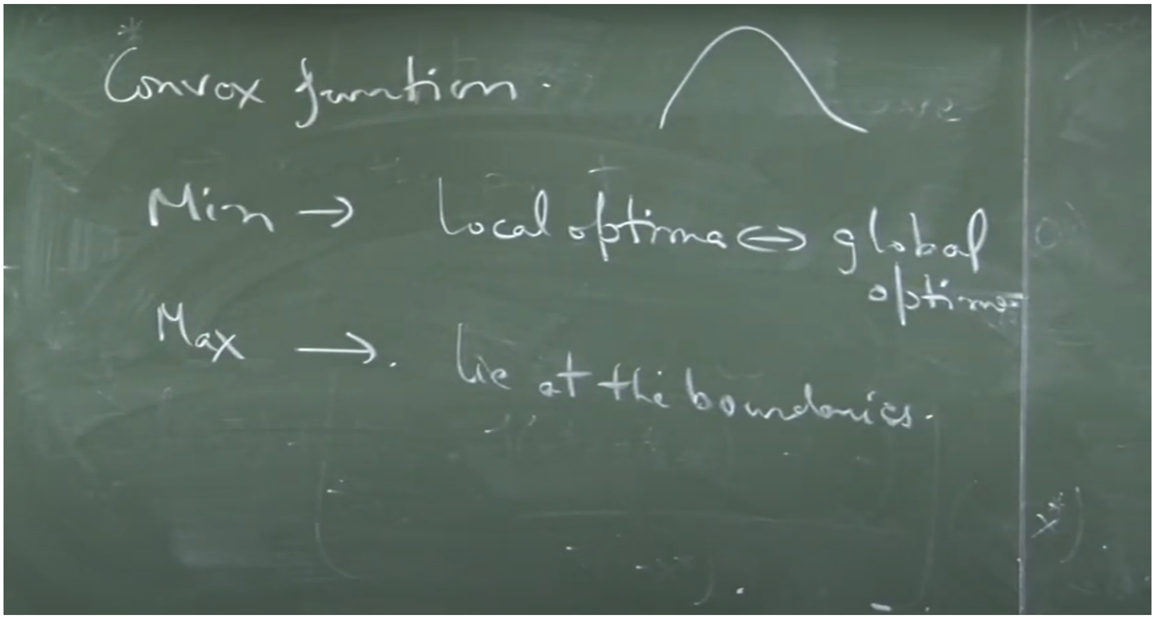
So, now let us say x star is your local optima and for the sake of contradiction or let us say y is another point right. What do I know from x star if I move in any direction my value increases and now I want to show that f y is more than f x star right. So, let us say f x star is your local optima. So, this is where I am always I always have less confidence because this something I did 20 years back right is that. Let us say x star is a local minima if I want to look at the point which is.



So, this is x star this is y what do I do I move in this direction right and I know that the value is increasing right and if at y it comes down the function does not remain convex right. Just this picture thing I have done like this from here to here since this is a local minima this quantity is positive right. So, again you look at x star x star is a local minima sorry this is what you are saying yes my handwriting yes. So, you start from x star you will go move in the direction of y even in the case of multiple dimensions and you know that the value should increase. So, if you just concentrate on that plane if the value goes down then your function is not convex that is a problem ok and any proof of this will formulize this argument.

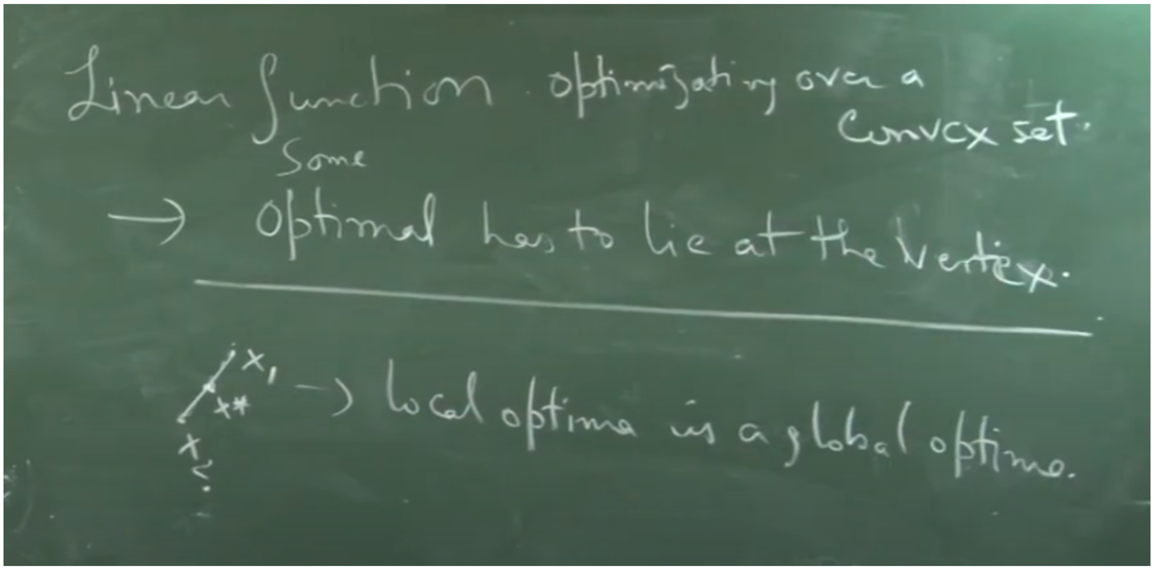


If you assume derivative is nice and everything then this proof works. So, what I am saying is that when the derivative is defined or something this is my definition of convex function those 2 you can switch between. I have not proved it, but I had started with right. So, what I am saying is that if the derivative is defined then you can convert and go between these 2 definitions. When your function is super nice these 2 definitions are equivalent and I only gave a picture kind of just with this I just give you an idea of why this equivalence makes sense.



I started with the local optima. It is a local optima and I have shown that for any y in the domain it is a local optima then for any point y f y is greater than f x star. That means, it is also a global optima that is what I wanted to show right. So, this implies you are saying that it is coming from the other side. You have to just take care of some direction.

You have to go from y to x star right like you are going in this direction this convexity right. So, t less than 0 that will not come as a convex part right. So, we do not have to worry about that. So, I just have to show it from one direction. Right, but then this should not be y minus x star this should be x star minus y or something because no no.

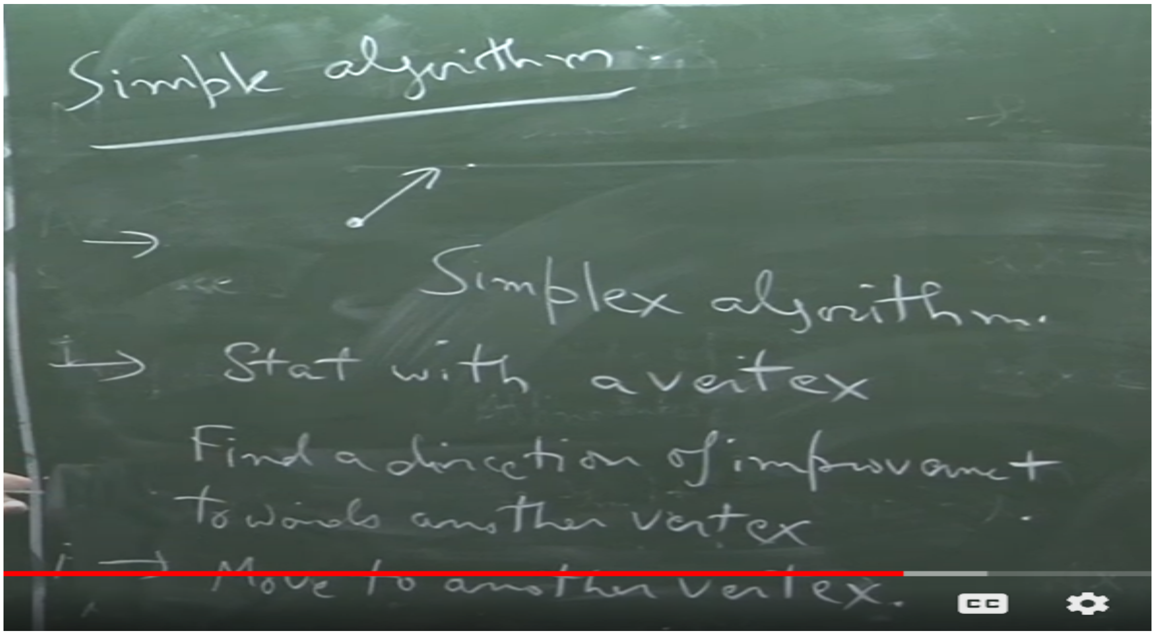


But in this equation limit will evaluate to 0. No, why would limit evaluate to 0 that is not correct. You are saying that this quantity is 0. I guess going from here to here. No, but his problem is after assuming this right.

This is 0 you are saying. So, at least one direction of this should be easy. So, we will worry about the proof of this later, but you believe that this is true right that local optima is global optima right. So, what we now know for convex function minimization local optima is global optima and what else what about the maximum this will lie at the boundaries. Right, because you cannot have this kind of a shape right. And now if you have a linear function this will imply that your optimal or some optimal has to lie at a vertex right.

Because if you take interior point you have 2 points here this point is a convex combination this x star is a convex combination of x 1 and x 2 then at least one of them has the value same as x star right. So, I can keep going outside and my value keeps decreasing or keeps increasing whatever way I want to do it right. So, this shows that for a linear function the optimal will lie at a vertex. Sorry. What is the vertex? No, now I am thinking of linear function it is optimizing over a feasible set.

Convex set ok. So, if I am optimizing a linear function over a convex set then my optimal has to lie at a or I there might be multiple optimal at least one of the optimal has to lie at the vertex. Sounds good? And is a global optimal right. Using these two what is a simple dumb algorithm? Right how to do it is a different question, but what I can do is I can go over all possible vertices look at all possible values and find the minimum or maximum. And that is obviously, yes because the number of vertices could be exponential.



But then we also realize that since local optimizer global optimer that means, if I start from any vertex I always have a direction to move right. I can always move in a direction where my I get to an improvement right. So, then this is just an algorithm. So, let us move to another vertex. So, this is the idea and this is called the simplex algorithm.

There are many things which are not defined here how do I start with the vertex right. How do I know that there is a direction of improvement, but why should it be an edge and which edge should I take? The other thing is why does this have to stop after small number of steps? What we know is that we still do not have any proof of efficiency of this algorithm. For whatever natural simplex algorithm people have designed. So, again these three steps can be defined in multiple ways. For all those ways there are always counter examples which are in exponential time.

Those are kind of boundary cases for most of the natural cases simplex algorithm works in polynomial time. So, people are pretty happy many of the industries still use simplex algorithm to solve any linear program because it works well in practice not the dumb one of looking at all the vertices, but this intelligent moving around. So, this is from Dantzig this is 1948 or something. So, it is some 60 years old 70 years old, but still works well in practice. So, now what we want to see is how to implement this thing.

We will see an implementation, but still I can tell you this is not we do not know whether it is polynomial time. If you show simplex algorithm is there is some version of simplex algorithm which is polynomial time it will be a huge result. This still open, but what you want to focus on is one way to implement this and that will tell us a lot about the structure of a linear program. This is the plan for next classes. Now, realize that so take a minimum let us say.

So, in the minimum case it is a I will consider that my linear function is convex. So, if I take 2 points I can write them as a convex combination of 2 points and at least at one of the points the value should be lower. If at this value it is higher if at this value it is higher than by convexity property the value at this will be higher. Like you have to take the correct direction convex or concave, but now since a linear function it is convex as well as concave. So, with this either x 1 will be direction of improvement or x 2 will be direction of improvement.

And I keep doing this I reach the boundary, I reach the edge, I reach the vertex. Sorry. It is the same thing we are doing is graph like in the starting classes we are finding the point 2 intersection. Yeah, but now we have to do it in multiple directions and we have to realize what does it mean to move in a direction. For a for a you mean the graphical approach we can do it like easily because we know where direction to.