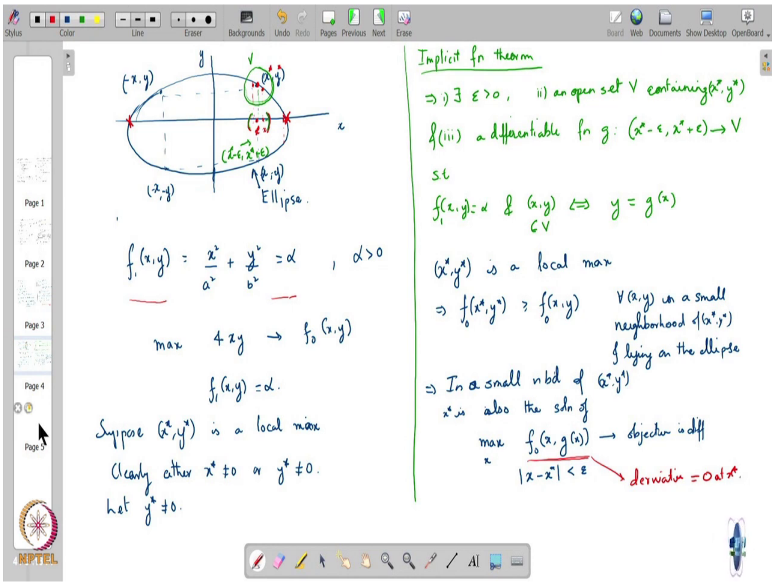
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

Lecture – 5C
Implicit function theorem

(Refer Slide Time: 00:19)



So, suppose x star comma y star is a local minimum of is a local maximum in this case of this particular optimization problem. Now, how did we argue in the case when the constraint was in open set? We said well x star and y star is a local solution. So, therefore, there is a ball around it and the constraint is an open set.

So, what we can do is we can find a ball around it such that the ball lies completely in the feasible region. If the ball lies completely in the feasible region we said we can go in any we can move in any particular any direction h and so, therefore, and then from there using

Taylor's theorem we concluded that the gradient transpose h should be greater than equal to 0, for every direction is or less than equal to 0 as the case may be ok.

So, in that case then we and then we said that well we one can take h suitably and show that this would imply that the gradient itself must be equal to 0 or the derivative of the function must be equal to 0. So, where does that argument fail here? The argument fails because, this not this is not an open set. So, it is not true that if it is just because it is a local minimum we can we cannot compare the value of the objective with we cannot compare that with every point in the neighborhood of x star y star.

We have to compare it only with those points that are in the neighborhood and in the feasible region right. So, you have to be sensitive to the shape of the feasible region near this point and that has to get factored into your analysis alright. So, it is not possible to therefore, simply work with only neighborhoods because neighborhoods are not necessarily all included in this alright.

So, now, let us look at a point like this. Suppose x star, y star is a local maximum of this optimization problem ok. Now, so the, now if you look at the so can someone tell me without loss of generality can I take either x star or y star as non zero? So, I did not mention it, but this goes without saying that alpha is positive here.

So, can I take x star and either x star or y star as nonzero? Yeah. If both of them are 0 then this equation will not be satisfied right. So, if the origin is not on the ellipse, ok. So, both of them 0 will not satisfy this equation. So, we will take suppose let us take for simplicity. So, clearly either x star is not 0 or y star is not 0 here ok. So, let us suppose let y star be non zero ok.

So, what that means is, in terms of this diagram here I am not looking at points where y star can be 0. So, I am looking at points I am not looking at this sort of point these 2 points here. These are these 2 are points here where which could where y star could put could be 0. I am not looking at those points.

So, I am looking at only points like these here either here or here where y star is nonzero. Now, when y star is not equal to 0, that means, we are in somewhere at a point like this ok. So, this let us take suppose this point itself. So, we are at a point like this.

Now, if you are at this sort of point you are on the surface of the ellipse and the thing that we want to exploit is because you are; because you are at this at a point like this, I should be able to if I tell you what the if I tell you what the x is, I should be able to solve for y in terms of the x knowing that I am on the ellipse ok.

Being on the ellipse gives me this particular thing that if I tell you what the if I tell you what the x is I know precisely what the y is ok. So, for example, if this is my point x I know that this is what the y is. If I if this is my point x then I know this is what the y is. This is because you are on the ellipse.

On the ellipse I can solve for y in terms of x. I cannot do that actually at this sort of point because here if I tell you an x here there are two possible y's. There a y above and the y below, but here near my point near the point x star y star that I am considering, I can solve for y uniquely in term of x on the ellipse ok. So, let explain this little more slowly.

So, suppose this is my point x star y star and I look in the very in the neighborhood of my of x star. In the neighborhood of x star ok; in the neighborhood of x star every x maps to a different y and to a unique y right. And that is so, which means that if I give you an x, I can solve through the equation of the ellipse for what the value of y should be right.

What that is effectively saying is I can look at this equation f of x comma y equal to f 1 of x comma y equal to alpha and then fix the x here and solve for the y using this equation and so long as I am not at this sort of point, I should get back a unique answer right, I should be able to get y in terms of x.

Now, this thing which I am just I was I am just telling you geometrically actually is the theorem it is what is called the implicit function theorem. So, this is a result of the implicit function theorem. Again I will tell you what the result is in this particular context. So, the implicit function theorem in our in this particular context guarantees that, there exist an epsilon greater than 0 an opens so, let us number this. It is 1, an epsilon greater than 0. 2, an open set V containing my point x star y star.

And this function g and a differentiable function g, g will this function is what will give me back my y in terms of x. That means, the argument to this function is going to be an x and its output will be y right. So, this function will take argument x. So, it but this function will take an x and generate for me y, but it will generate for me this y in a limited scope.

Means that only in the neighborhood of x star can I talk of such a existence of such a function, only in this neighborhood of x star will I be able to get my y back in term of x. If I make this neighborhood too large then you know all hell breaks loose I get multiple y's there is no more a correspondence between x and y.

So, if but so long as I am close enough I should be able to solve for x in terms of y ok, solve for y in terms of x. So, there is a differentiable function g whose domain is like this. It is just x star minus epsilon to x star plus epsilon mapping back to mapping back to V.

So, in this figure here V is this open set ok, x plus minus epsilon is this, sorry x star plus minus epsilon is this set, x star minus epsilon to x star plus epsilon and g is the function that you know taking a value of x returns the value of y from it ok. So, now, I will tell you what the theorem says.

There exist an epsilon greater than 0, an open set V containing x star y star and a differentiable function g; g whose domain is x star minus epsilon to x star plus epsilon mapping back to V, such that if you are on the ellipse that means, f 1 of x comma y equals alpha and you are in this neighborhood V that is equivalent to; that is equivalent to evaluating g on that neighborhood ok.

So, y equals g of x if and only it lies in the on the ellipse sorry and x comma y belongs to V and is in the neighborhood V that is being considered here right. So, one direction means that if you look at any point that is in the ellipse that is on the ellipse and in the small neighborhood V of x star y star, then that point can be expressed in this sort of way. y is a function of x, y equal g of x and x lies x of course, lies in this sort in this domain.

So, take any point on the ellipse that can be expressed this way. The reverse is also true. Take any x that lie that is in this neighborhood here evaluate the function lo and behold you get a point that is bang on the ellipse. This is what formerly it means to solve for y in terms of x.

So, you have got y in terms of x ok, alright. So, you can actually work this out you know solve this as a quadratic equation and get what the what this function actually is. We do not need the form of it. I just want you to know that this is we will see how this generalizes there is such a function that is the point alright ok.

So, now, because there is such a function so, now, that it means the following. That since x star y star is a local minimum right x star y star is the best local maximum. x star y star is the best in amongst all points on the ellipse and in a small neighborhood around x star y star. It has the maximum value of maximum value for this area.

So, x star y star is a local maximum. So, let us denote this objective here by f 0 of x comma y. So, utilize; f of x star y star is greater than equal to f of x comma y, for all x comma y in a small neighborhood of x star y star and lying on the ellipse f 0 sorry right. Now, so, in a small neighborhood around x star y star x star is the best value gives you the best value of f 0. In a small neighborhood of x star y star, y can be solved in terms of x right. So, if I take the smaller of the two neighborhoods what must be the case?

Student: (Refer Time: 14:51).

Both should hold right, both should hold, that means in a there is a small enough neighborhood in which I can solve for y in terms of x as given by the implicit function theorem. And moreover, x star y star is the local maximum is the maximum over that neighborhood, ok.

So, therefore, guys in a small neighborhood of x star y star x star, so, also the solution of this optimization right. So, you can; so you can look at this particular optimization in which y has now been substituted for as g of x and you are optimizing only over x right. So, in a so small enough epsilon this, the original your x star will also be an will also be the optimal solution of this problem ok.

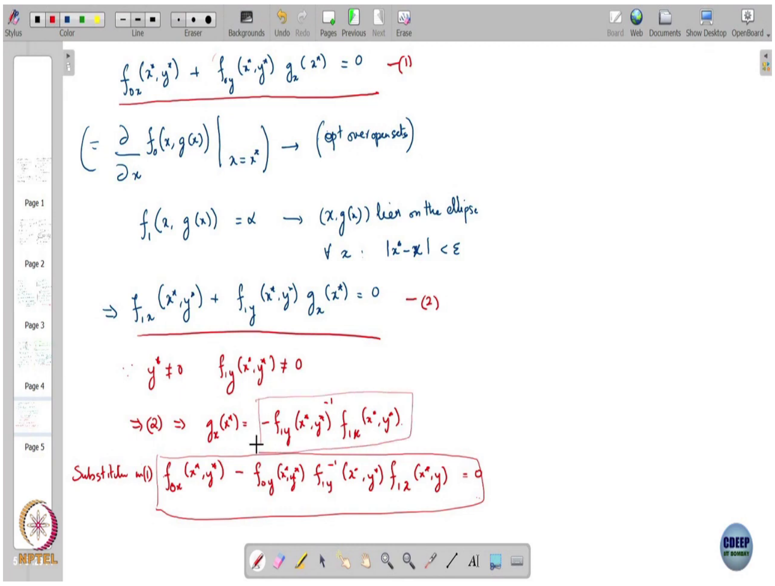
So, what does now what kind of problem has this become? You have f 0 which was a differentiable function, I told you g is also differentiable. So, this is all nicer. Is not it all nice? Your objective is differentiable, objective is differentiable, but importantly what is the shape of the constraints? What or what is the nature of the constraints? It is now an open set right.

So, as I this is what I wanted to tell you. So, you take a problem which where you are optimizing a function over a surface ok. The implicit function theorem if it kicks in let us you convert that to an optimization over an open set, over only if some of the variables. If this was an optimization over x and y to begin with, we now have an optimization over only x ok, but it is an, the earlier problem in x and y was an optimization over a close set ok.

We what we have called, been able to conclude is that its x star y star which was the solution of that earlier problem is also a local minimum of this particular new optimization problem which is an optimization of a differentiable function, but over an open set, but open set only in x not in x and y both ok.

So, now, we can simply apply what we know from our previous lecture which is that the this means that the derivative of this new function the derivative of this thing must derivative of this should be equal to 0 at x star right and that is just an application of chain rule of differentiation.

(Refer Slide Time: 18:45)



So, we therefore, get that. So, I can write this in the following way that f 0 of f 0 x evaluated at x star y star plus f 0 y evaluated at x star y star times g x evaluated at x star this must be equal to 0. What is this quantity? This the left hand side here is this is simply partial derivative with respect to x of the function f 0 of x comma g of x evaluated at x equal to x star. I have just applied chain rule to get this.

Now, in addition to this I also have that the I also have that g of x returns me a point that on the ellipse right in the neighborhood of in the neighborhood of x star g of x is a point on the ellipse x comma g of x is a point on the ellipse. So, which means that x f 1 of x comma g of x must be equal to alpha. So, this first equation comes by optimization result on optimization over open sets; over open sets. This here is that x comma g of x lies on the ellipse.

Now, for what x does this lie on the ellipse? The implicit function theorem tells you right. The implicit function theorem tells you that this is true only if you are in a small neighborhood of x star. So, x should be within plus minus epsilon of x star right. So, this is lies on the ellipse for all x such that x star minus x in absolute value is less than epsilon right.

But then this since this is true for all x like this that are in this interval from x minus epsilon to x plus; sorry x star plus minus x star minus epsilon to x star plus epsilon. Since this is true for all of them. What does it mean? The left hand side the left hand side is actually independent of x as I vary x from x star minus epsilon till x star plus epsilon in that open interval because it is all it is for all these x's it is equal to alpha.

What does that mean? What does that means is I can put the derivative of the left hand side should be equal to 0 because it is independent of derivative with respect to at x right. So, derivative at x so derivative of this at x star must be equal to 0. So, which means that so that just because f 1 x evaluated at x star y star plus f 1 y evaluated at x star y star times again g x x star this should be equal to 0.

So, I want to highlight this. We have got 2 equations like this. Yeah so, this is so this is true for this is identically equal to alpha, for all the x in that interval right. So, at x star; x star is the is in that interval right. The derivative should it I mean it is a constant. So, it is derivative should be equal to 0, if I view this as a function of x right. So, that is the; yeah so that is the reason.

So, this derivative must be equal to 0 ok. Now, what is this f 1 of; f 1 of y evaluated at x star y star? f 1 y evaluated at x star y star. This is the partial derivative of your ellipse equation evaluated at x star with respect to y evaluated at x star y star ok. Now, so long as my so if you look at that equation here. So, if I take the partial derivative with respect to y, I am going to get a 2 y by b square right and I had assumed that y star is not equal to 0.

So, when y star is not equal to 0 that f 1 y evaluated at x star y star is not 0 right. So, since y star is not equal to 0 you get f 1 y evaluated at x star y star is not equal to 0 ok. So, what that means is I can just use this equation 2nd equation here. So, let us call number these. Let us call this equation 1, let us call this equation 2.

So, I can use 2; from 2 I can just solve for gx of x star. So, gx of x star is equal to negative of f 1 y evaluated at x star, y star times f 1 x evaluated at x star, y star. And this then putting this into equation 1 gives me the condition, that substitute in 1, sorry I think I missed an inverse here. So, this is there should be an inverse here. Now notice what has happened.

So, one of you asked me we do not we need not know g enclosed form, yes it is true we did not know g enclosed form, but what I have done is I in using equation 2 and using that f 1 y evaluated at x star y star this is not 0. I was able to eliminate g from the equation or completely and I have now got here an equation that is only in terms of x star and y star and the known functions f 0 and f 1 right.

So, this therefore, is a necessary this gives me a new necessary condition for my for x star y star to be a local maximum. So, go back to the optimization problem that we had this was the optimization problem. So, x star y star is a local maximum then it is necessary that x star y star satisfies this boxed equation ok.

So, what this is giving us is that is we as a result we by what we have done is we took an optimization of a differentiable function over a surface use the fact that we are on a surface, use the equations of the surface to solve for one some variables in terms of the others turned that around and substituted and then use chain rule some more calculus and so on.

And then got back an bunch of equations that must be satisfied you can see that if you did not have constraints some of these terms would vanish and all you will be left with is just you know derivative equal to 0 that is all you will be left with, derivative of the objective equal to 0. So, this is how that earlier thing generalized, what we have done earlier generalized ok.

So, there is a this the implicit function and the and this particular term in particular this term in particular has an important meaning and I will discuss all these things in the next class ok, how this generalizes to a to the to more general settings.