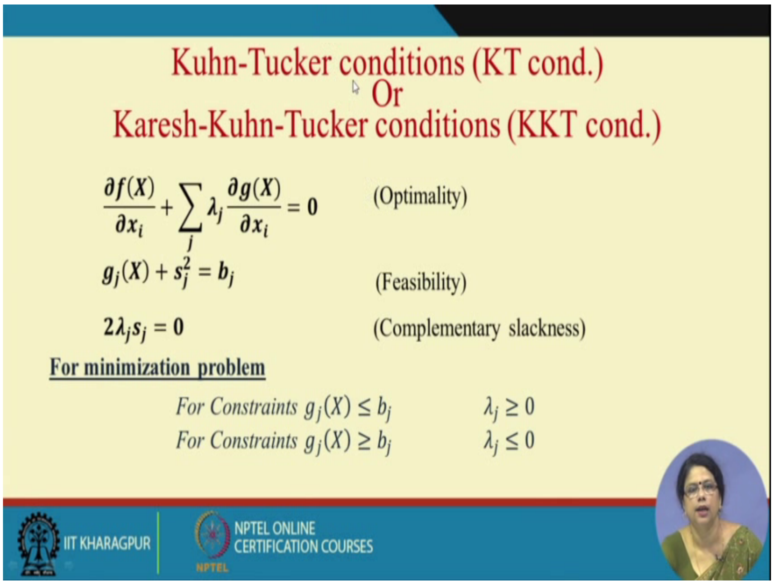
Constrained and Unconstrained Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 50
KKT

Today, I will solve certain non-linear programming problem. Problems are having constraints of different kind. And we will use the Kuhn tucker conditions as I explained in the previous classes, just let me brief.

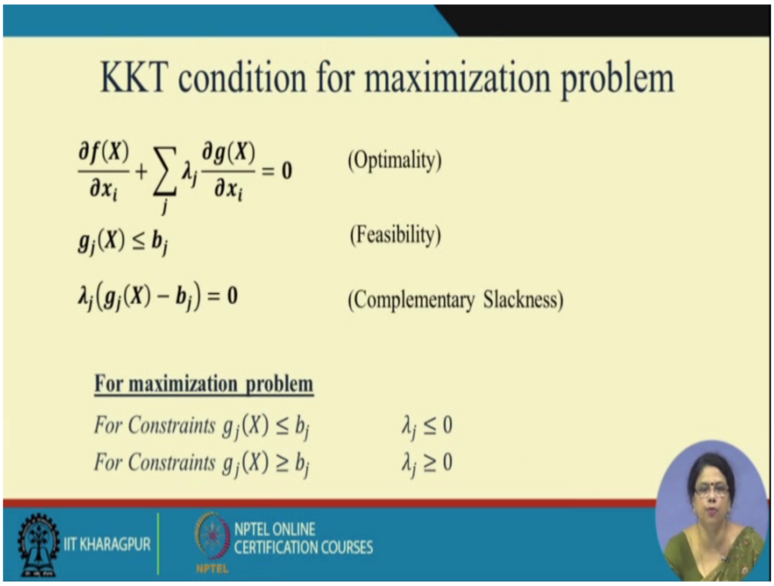
(Refer Slide Time: 00:37)



If we consider a minimization problem, where minimization of f X subject to g X less than is equal to 0. Then we will have a set of KareshKuhn Tucker conditions first set is the first order derivative of the Lagrange function with respect to decision variables we are getting the optimality conditions. Now from the second we are differentiating with respect to KKT KKT multiplier. Then we will get the feasibility condition. And if we just differentiate l with respect to the slack variables si sjs then we will get the complementary slackness condition.

Now I have already explained that depending on the sign of theKKT multiplier, we can have, we will get theKKT conditions, but these are all the necessary conditions these are not sufficient conditions. Now this conditions will be the sufficient condition under certain conditions, I will go through that. Now once at I am repeating here that for the minimization problem if the constraints are of the type less than equal to then, the corresponding multiplier would be greater than equal to . I need not to explain again that if this is equal to 0, lambda j is equal to 0; that means, the corresponding constraint is the active constraint otherwise if it is greater than 0, then it is they the corresponding constraint would be the in active constraint.

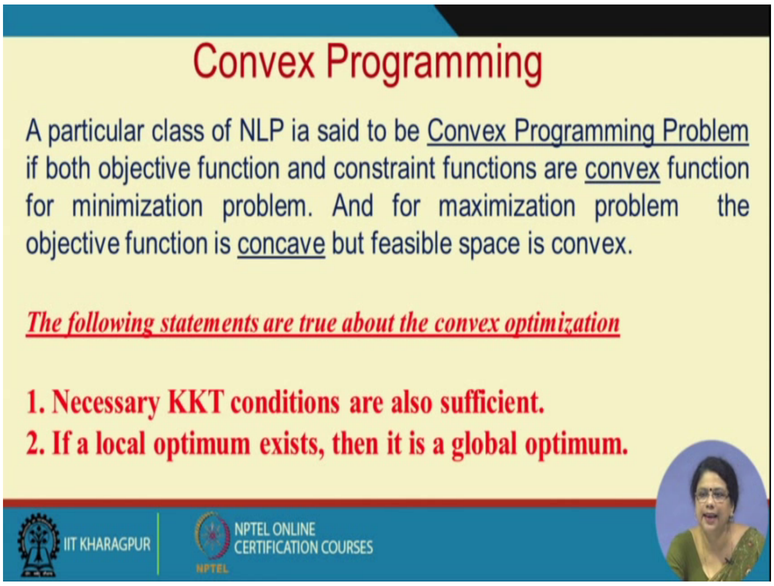
(Refer Slide Time: 02:33)



And for the constraint of type greater than equal to we will have the multiplier less than is equal to 0. Same set of conditions can be repeated for the maximization problem as well, when we are having the non-linear programming problem maximization of f X subject to g X either less than is equal to bj or gj greater than equal to bj.

Now, here now g functions and the f functions these are all the non-linear functions. Here also after taking the first order derivative with respect to different variables. We will get the KKT KKT conditions like optimality condition, feasibility condition and the complementary slackness conditions. These are the this one and if the constraint is of the less than equal to then the corresponding multiplier would be neg less than is equal to 0 non positive and if the constraint is of type greater than equal to, then the corresponding multiplier would be non negative. I have explained why these are so in my previous classes.

(Refer Slide Time: 03:36)

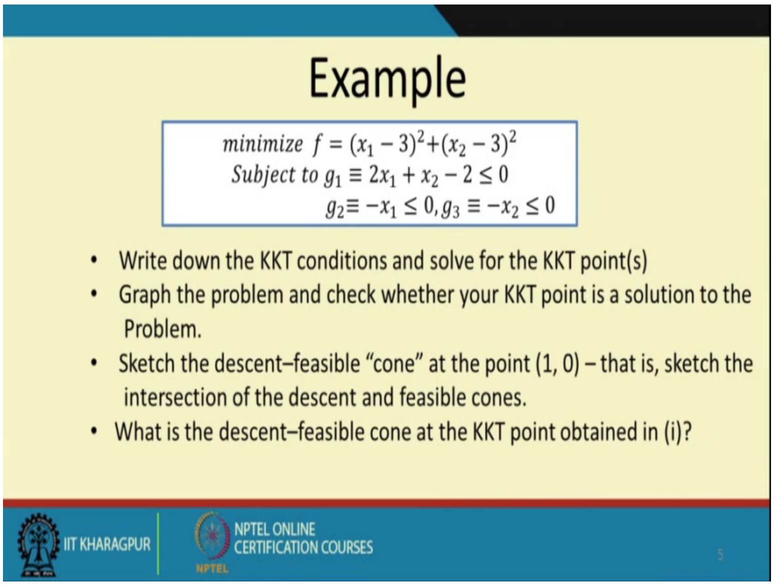


Now, this conditions KKT KKT conditions need not be the sufficient conditions. That is why in optimizations problems there is a specific branch of problems non-linear programming problems which are of type convex optimization problems. This is very important in optimization problem because if we consider the convex optimization problem then we will see whatever local optimality we will get through the KKT conditions this will be the global optimal solution as well.

Now how to define convex optimization problem rather how to define convex programming problem. Now if we consider a non-linear programming problem, where the feasible space is of the problem, if it is of minimization type if the feasible space is convex and the objective function as well as in convex then corresponding optimization problem will be the convex optimization problem, but for the maximization problem the pattern of the objective function will be different for convex optimization problem. Here the objective function must be concave, but the feasible space must be convex, then whatever local optimal whatever optimal solutions we will get all the local optimal solutions will be the global optimal solution that is the specific sufficient condition for through the KKT conditions.

Now, I if this is not so if the problem is not convex optimization problem; that means, the minimization problem both are not of convex type for maximization objective function is not concave or a feasible. Space is not convex if either of these conditions are violated then we have to go through the process, as we did we have to do the second order of the Lagrange fun function to check the global optimality or the optimality conditions as a sufficient condition. That is why the KKT points whatever we are getting through the KKT conditions. We can say this KKT points are the global optimal solution for the convex optimization problem.

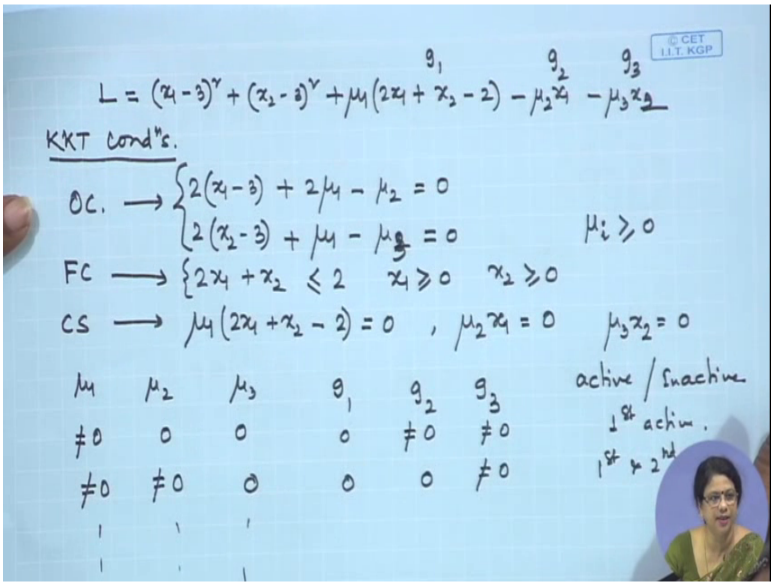
(Refer Slide Time: 06:02).



Now, I have given this problem in the last class, that if there is a minimization type of problem minimization of f, the function objective function is given as a that is the function is the circle centered at 3 3. Now function is we have to minimize; that means, only the variable is the radius of the circle. It will vary now this is of type convex there is no doubt about it. Now the constraint set the first constraint g 1 constraint is of linear type, linear functions are always the convex functions. That is why whatever half space we will get through g 1 this is linear and convex as well and other 2 constraints a as well, g 2 and g 3 that is the non negativity constraints of the decision variables we have considered the reverse because reverse notation because we wanted to have all the constraints of less than equality type. Now we have to apply the KKT conditions here. One thing is sure that through this whatever KKT points we will get through the KKT conditions, those are the optimal solution no doubt about it through KKT we are getting the necessary, but for this specific problem that would be sufficient as well because the problem is of convex optimization type problem.

Now, let us try to answer one by one. I have given to you in the last class hope you have tried in between let me go through all the answers like how to formulate the KKT conditions, how to get the KKT points one by one.

(Refer Slide Time: 07:59)



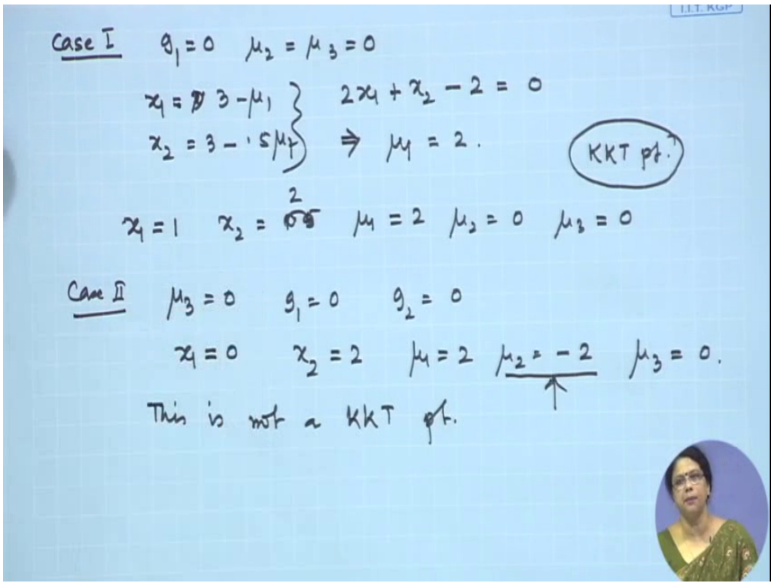
Let me first formulate the Lagrange functionL, L is equal to f x 1 minus 3 whole square , plus x 2 minus 3 whole square, plus the first KKT multiplier mu 1 into 2 x 1 plus x 2 minus 2 plus mu 2 minus x 1 mu 2 is the Lagrange multiply KKT multiplier minus mu 2 x 1 minus mu 3 x 3. Now let us get the KKT conditions first is the optimality conditions that is why we will differentiate with respect to x 1 first then what we will get? We will get 2 into x 1, minus 3 plus 2 mu 1 minus mu 2 is equal to 0. If we differentiate with respect to x 2, then we will get 2 into x 2 minus 3 plus mu 1 minus this is x 2 minus mu 3 is equal to 0. This is these are the feasibility I am sorry the optimality conditions.

If we want to have feasibility condition what will be the feasibility condition? Feasibility condition will be 2 x 1 plus x 2 must be less than is equal to 2 what else minus x 1 less than equal to 0, rather x 1 must be positive x 2 must be positive. These are the feasibility conditions and let us go for the optima the complementary slackness condition. The first condition is mu 1 2 x 1 plus x 2 minus 2 is equal to 0 and the second condition is mu 2 into x 1 is equal to 0 and mu 3 x 2 is equal to 0. This is the complementary slackness condition, this is the feasibility condition and this is the optimality condition. Now you see if you have so many conditions together we have unknowns these are 2 we have a unknowns like x 1 x 2 mu 1 mu 2. Now for different value of mu 1 mu 2, mu 1 mu 2 mu 3 as well this is 3, here this is 3 x 2. Now we can have mu 1, mu 2, mu 3. We are having g 1 if we consider this is as g 1, this is as g 2, this is at g 3. Then with different conditions of g 1, g 2, g 3 mu 1 mu 2 mu 3.

We will have either active or in active conditions, is it not? That is why if we consider first with mu 1 not is equal to 0, mu 2 is equal to 0, mu 3 is equal to 0. Then if mu 1 not equal to 0 then certainly g 1 must be is equal to 0, because we are having conditions condition like mu 1 into g 1 is equal to 0. And if mu 2 is equal to 0 then this is not is equal to 0 this is not is equal to 0. Then what is the active constraint here only first constraint is the active constraint. Let us take the other conditions that let us consider mu 3 is equal to 0 mu 1 not is equal to 0 mu 2 not is equal to 0 then what we will have we will have this is equal to 0. This is equal to 0 and this is not is equal to 0, then who are the active constraints active constraint will be first and the second would be the active constraint.

In this way we can have different combinations of mu 1 mu 2 mu 3. Depending on the value of the combinations of mu 1 mu 2 mu 3, whether this is 0 or non 0 we will get a set of values for x 1 x 2 x 3 what we need to do we need to find out for the corresponding cases what would be the value of x 1 x 2 x 3 are these satisfying the constraints KKT conditions rather then we will declare this is a KKT point otherwise not. Let us do certain calculations for this.

(Refer Slide Time: 12:57).



Let us consider the first case we have considered that, is case one where we have considered that g 1 is equal to 0. Probably this is the second condition here g 1 is equal to 0 mu 2 0 mu 3 0 all right. If we get mu 2 is equal to 0 mu 3 is equal to 0 then what will get we will get x 1 mu in in terms of mu 1, mu 1 and x 2 in terms of mu 2 all right and.

If we have this if you just substitute the value in the let us see what we are getting we are getting, x 1 is equal to 2 mu 1 2 2 will cancel only we will get 3 minus mu 1. And for x 2 we will get 3 minus 0.5 mu 2 this 2 all right. Now we have the condition 2 x 1 plus x 2 minus 2 less than is equal to 0, but we have considered that mu 1 not is equal to 0, that is why g 1 is equal to 0 that is why we can say this is equal to 0. If we substitute both the values of x 1 and x 2 what we will get we will get from here, the value this is mu 1 because these are both in terms of mu 1 we will get the value for mu 1 is equal to 2 if we get mu 1 equal to 2. Then certainly x 1 would be is equal to 1 x 2 would be is equal to 2, 0.5 is that x 1 is equal to 1, mu 2 is equal to 0. Then 2 2 will cancel then it must be is equal to x 1 must be is equal to 3 minus mu 1 certainly x 1 is equal to 1 x 2 would be is equal to 1 0.5 mu 1 is equal to 2, mu 2 is equal to 0 mu 3 is equal to 0.

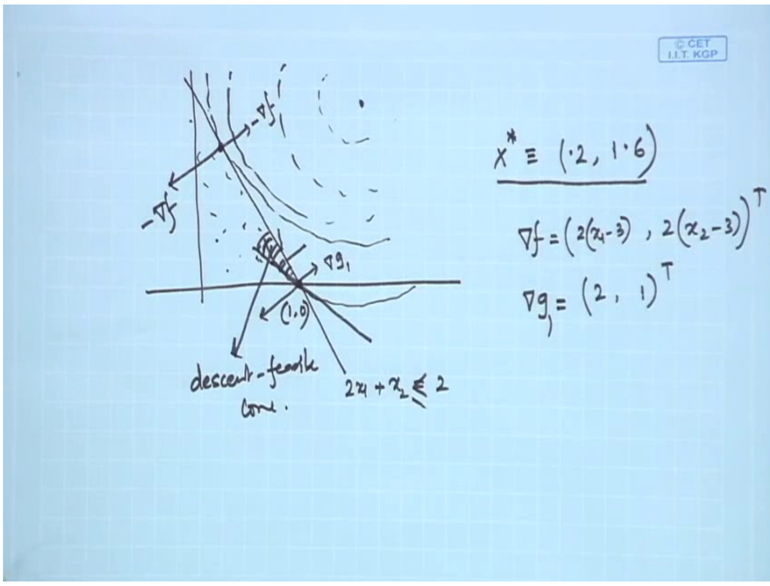
We will get this value this would be 2 0.5 not this is one, this is x 2 is equal to this would be is equal to 2 only because 2 into 0.5. That is 1 3 minus 1 2 mu 1 is equal to 2 mu 2 is equal to 0, mu 3 is equal to 0. We will get this value. Let us see we are (Refer Time:

15:50) all the KKT conditions are being satisfied or not. The first condition satisfied second satisfied both are positive there is no doubt about it this is satisfied all are being satisfied that is why we can declare that this is a KKT point all right, but this can need not to be the this is one of the optimal solution, it could be the local optimal.

We have to find out all conditions together all right now we will go to the second condition case 2 where we are considering mu 3 is equal to 0 only, g 1 0 and g 2 0 nothing else. From here after doing the calculations we will get x 1 is equal to 0 x 2 is equal to 2 and mu 1 is equal to 2 u 2 is equal to minus 2 mu 3 is equal to 0. Now one thing is that this is a minimization type of problem, is it not. That is why here we missed one set of conditions that mu I must be greater than equal to 0.

All right that is one of the KKT conditions, but you see here for the second case we are getting mu 2 negative. That is why KKT conditions are not being satisfied that that is why this is not a KKT point. In this way for all other combinations of mu 1 and mu 2 and mu 3 you have to find out a set of values of objective function and we will get after that the first question was write down the KKT condition, solve for the KKT points we already got one KKT point like that you can get other cases as well and graph the problem and check whether your KKT point is this solution to the problem or not. And sketch the descent feasible cone. Let us try to sketch what is the solution what is the how what would be the solution fir this problem.

(Refer Slide Time: 18:03)



Now, our constraint is 2 x 1 plus x 2 less than is equal to 2, that is why we can have this is the constraint 2 x 1 plus x 2 is equal to 2. And at 3 3 we have the center of the circle that is why this is the contour of the different circle. And here would be the optimal solution. This would be the optimal solution for this problem, we will get the optimal solution as 0.2, 1.6 all right for different cases if you solve it you will get it. Now we need to find out that this is a KKT point that you can get it from different cases. Now there is a point is at 1 0 this is 1 0. Now in this point 1 0 we need to find out the descent feasible cone all right.

Now this is my feasible space because my condition was this is less than is equal to 2. That is why this is the minimum of the objective function. That is why we can say that at this point, at this point if I consider grad f and grad g differently what is my grad f grad f would be 2 x 1 minus 3, 2 x 2 minus 3. And what it what is grad g grad g would be I mean to say grad g 1 g 1 would be 2 1 T. That is why if I just stand here what is the value for 2 1, 2 1 would be here somewhere here. That is why we can say this is the this is the direction for grad g, and the reverse direction would be here and what about the objective function this is the minus grad f and this is the plus grad f direction all right.

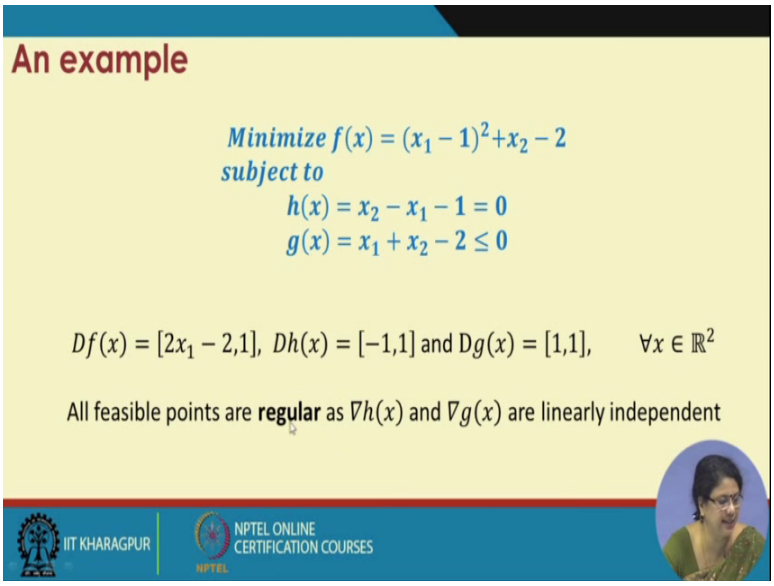
Thats why here also the plus grad f would be this way all right. And if I consider here is say this there is one circle here, then at this point if I just consider the tangent. This tangent will give us the direction of the feasible cone. What is feasible cone I explained already in the last class with the figure that feasible cone is the combination of the half spaces we are getting through the constraints. Here we are having 3 constraints.

What is the half spaces we are getting half spaces we are getting at this point ne from g 1 one from g 2 and one from g 3 all right? And that is why we can say this is my the feasible direction at this point this is the feasible direction from here. If I move any one of these direction I will be in the feasible direction, and what will be the descent direction descent direction would be at this point this is the descent direction this is minus grad f this is the ascent direction of the objective function feasible space.

Feasible cone in terms of constraints, but the descent direction in terms of objective functions since we are interested to find out the descent feasible cone that is why we have to consider the intersection of these 2 spaces all right. Intersection of 2 of 2 spaces can only be happen at this point will be this region only. That is why we can say this region is the descent feasible region that is the cone that was the question given to you all right. Now what is the descent feasible cone at the KKT point obtain in one, now if I consider this as a kk this is a KKT point though this is the global optimal. But thats a KKT point what we are getting at this point we are getting if I consider this at this point the tangent at this point here.

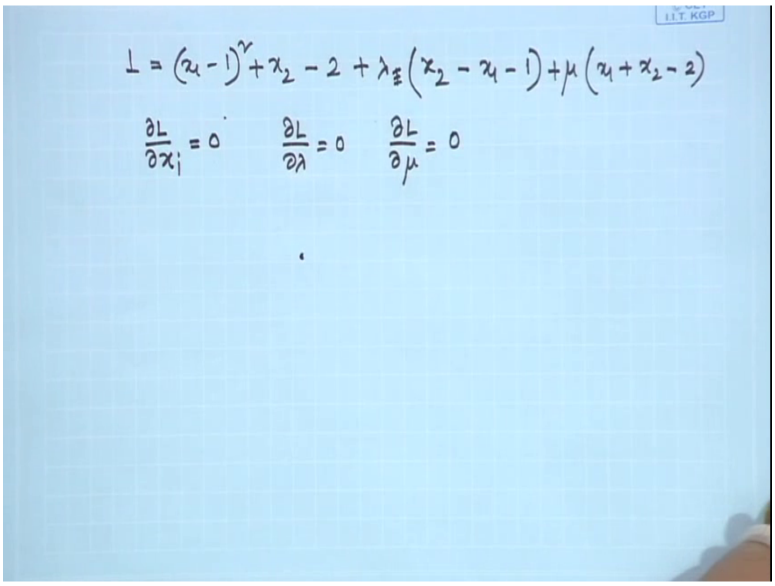
Now; that means, this tangent this objective functional tangent will be just tangent here. Because it will this will just move this way at this point. And this direction would be the descent direction. What is the feasible direction? If I ask you what would be the combination of the feasible and descent direction that is no that is empty, because there is no point which false under feasible as well as the descent direction at the optimal solution. The because there is no space there is no intersection region of these 2 spaces one this side another this side all right. That is why at the KKT point we at the KKT point in the sense that at the optimal point all at the KKT points we may take different KKT points, but all are not the global optimal solution at least at the global optimal point we can say that feasible descent cone is empty. That is the conclusion for us. Now this is all about the KKT conditions KKT theory, now I will go through few examples.

(Refer Slide Time: 23:58)



For understanding more about it you see we have considered 2 types of constraints one is of equality type another is of less than equal to type. Once we are having the equality type we know the Lagrange multiplier process and it if it is of inequality type we will use the KKT condition that is why both together we will introduce the Lagrange multiplier and the KKT multiplier together.

(Refer Slide Time: 24:35)

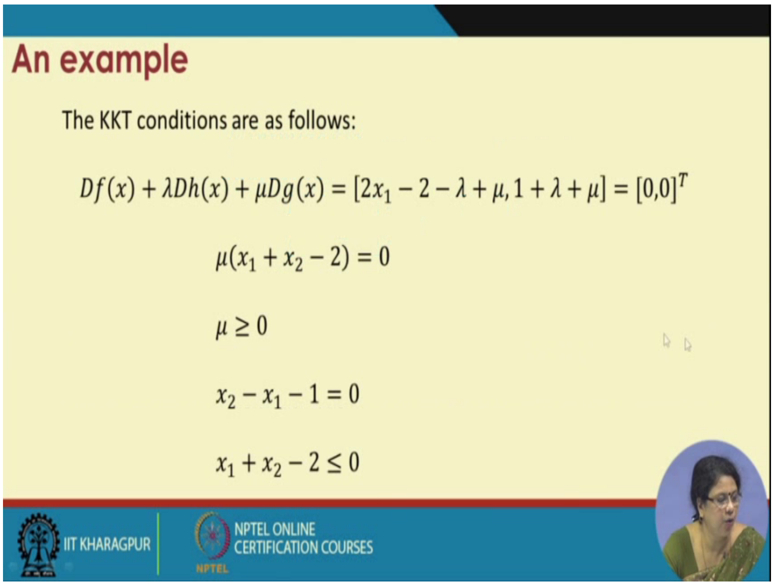


And we will formulate the Lagrange function for this. And we can solve this function what would be the Lagrange function for this Lagrange function would be l is equal to x 1 minus 1 square plus x 2 minus 2 plus lambda one this is the Lagrange condition.

Since we are having only lag one Lagrange multiplier that is why I let me just put it as lambda plus mu into x 1 plus x 2 minus 2 where mu is KKT multiplier. Now from here again we will get the first order derivative of Lagrange functions second order derivative of Lagrange function, and here another thing we can check that we know all the feasible points are regular if grad g and grad g are linearly independent.

Now, this thing also we can check it very easily, we have grad of h this one grad of g is this one we can check both are linearly independent. That is why these are the regular points.

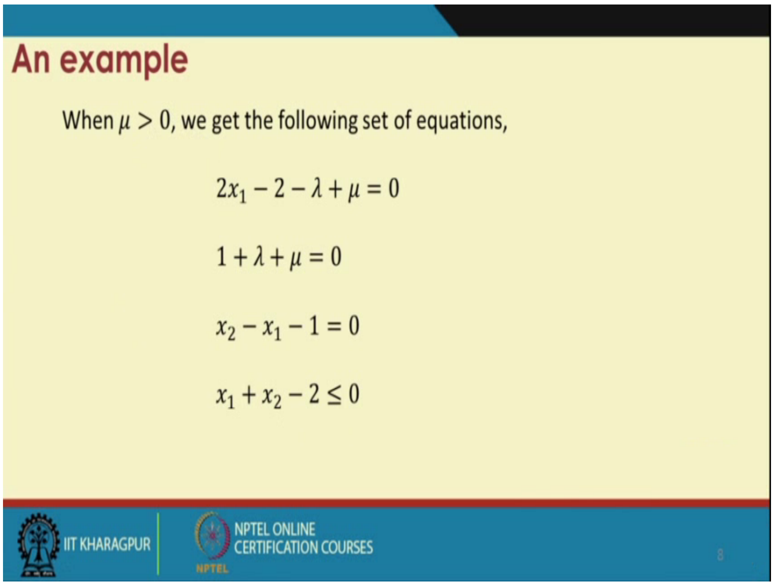
(Refer Slide Time: 25:41)



Now, if we consider the KKT conditions, then we will get set of KKT condition this is the optimality condition. Optimality condition we will get through del l by del x a x 1 equal to 0 x 2 is equal to 0 del l by del lambda is equal to 0 del l by del mu equal to 0. If we just consider all together then sorry we will get this set of conditions all right, 2 x 1 etcetera. Now this is the complementary slackness condition. Since we are having the minimization type of problem where the in equality type is less than equality type that is why KKT condition would be mu less than is equal to 0, but you see we can only say lambda not is equal to 0. Nothing we can say about the greater than or less than and this movement, we will see later though you know what would be the value for lambda.

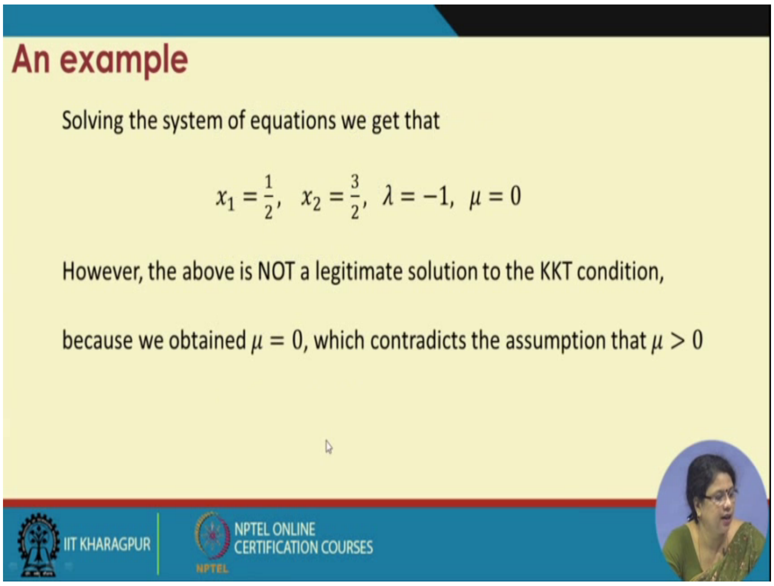
Because you know the significance of the lambda in the Lagrange multiplier what it means and for minimization what would be the sign of lambda you know very well, but that will not come under the condition of the KKT condition.

(Refer Slide Time: 27:16)



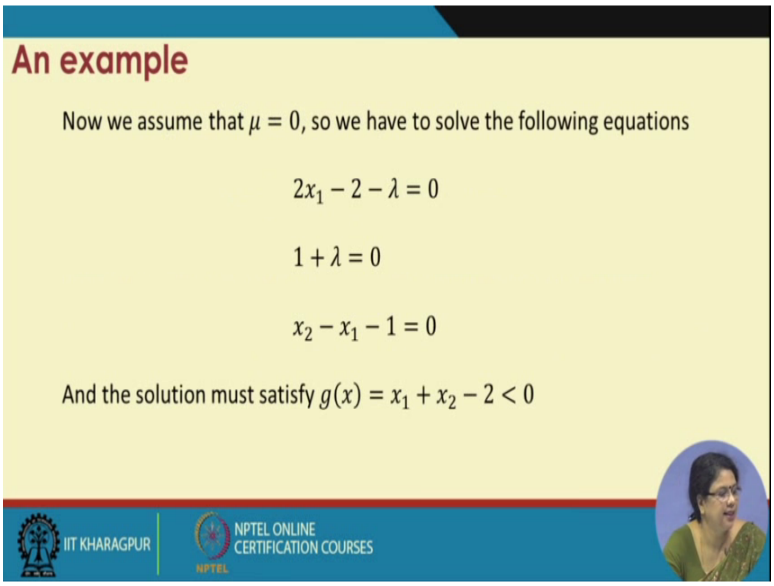
Now, what else we have we have these 2 feasibility conditions. These are all the KKT conditions together from here we can have since mu can be greater than 0, mu can be equal to 0.

(Refer Slide Time: 27:29)



If we consider mu greater than 0 then these are the conditions we are getting from that. Now if this is so we are getting mu is equal to 0 x 1 is equal to half x 2 is equal to 3 by 2. And lambda is equal to minus 1, is it a KKT point it is not a KKT point because we are getting certain condition which is contradicted in nature. Because we have assumed mu greater than 0, but after doing the calculation we are getting mu is equal to 0. That is why we cannot say this is a KKT point at all, that is why mu greater than 0 cannot happen.

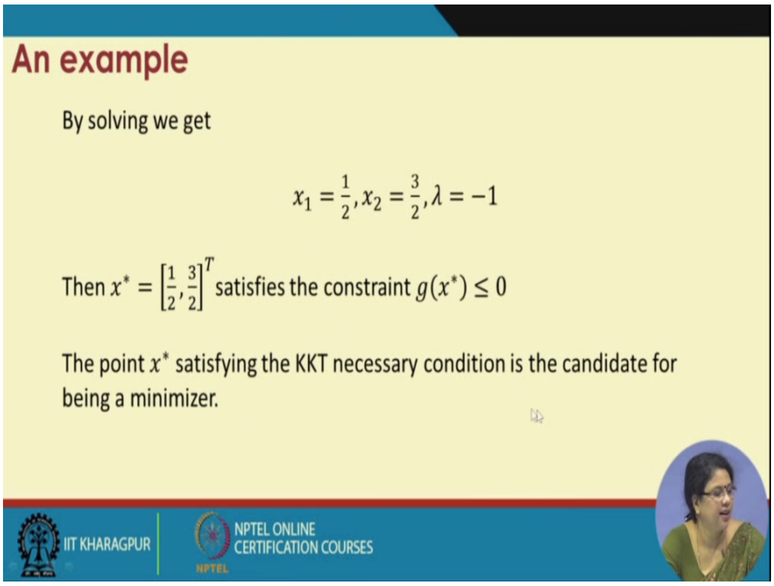
(Refer Slide Time: 28:07)



Let us go to the next that is mu is equal to 0 if we consider mu is equal to 0 what does it mean the corresponding constraint that is the x 1 plus x 2 less than minus 2 that is of type less than, that is why that constraint is inactive constraint that is not the active constraint, since mu is equal to 0 all right.

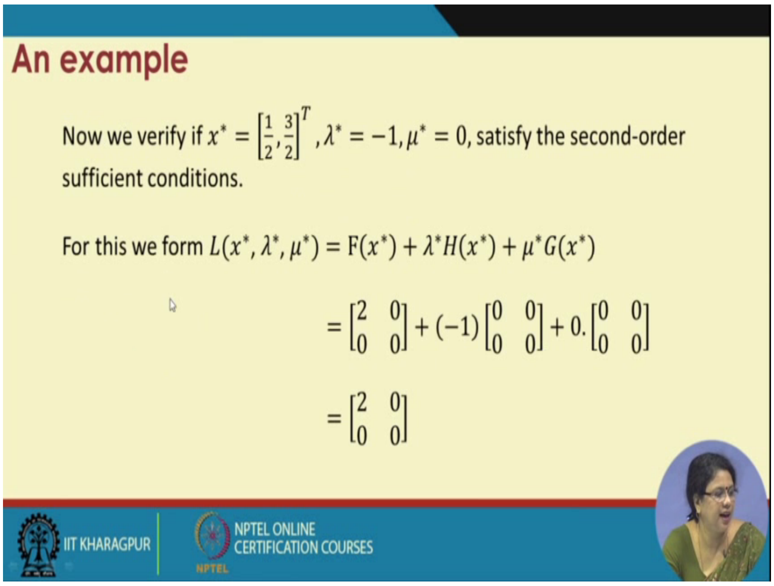
Now, in that case we are getting this one along with this condition. And whatever value we are getting we could see that the that is satisfying the kk all the KKT conditions together.

(Refer Slide Time: 28:34)



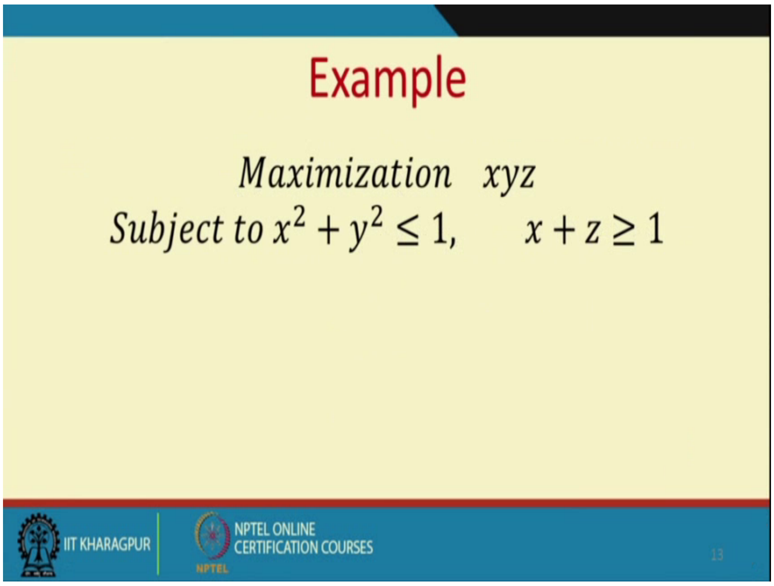
And you must have been seen that the lambda value is negative. Because in the Lagrange multiplier method you know for the minimization type of problem lambda is negative. We are getting that as well that is why we can say that the corresponding x will be the optimal solution for the problem.

(Refer Slide Time: 29:02)



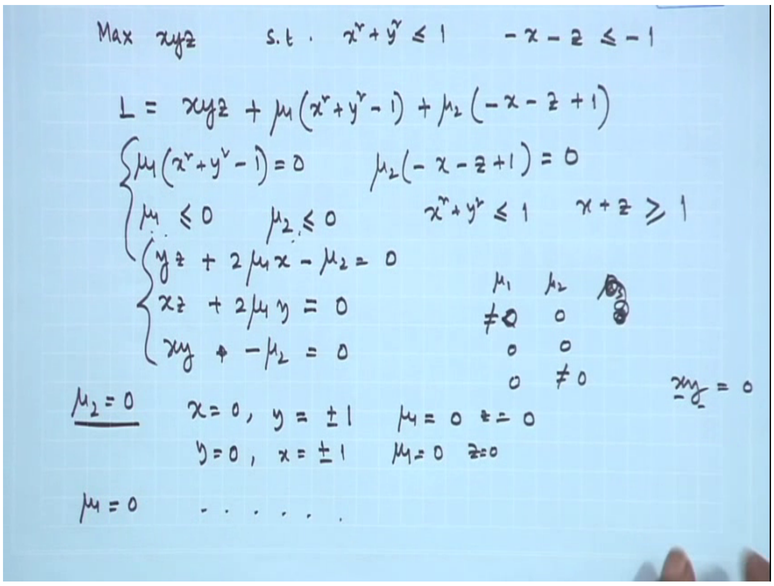
And if I want to check whether this is the sufficient condition or not, first we have to check about the property of the objective function property of the constraints and the constraint space as well. Since the constraints of all are all of linear type and this may that is why constraint is convex we can say, but about the objective function we could not say anything that is why we are going to the second order derivative of the function of the Lagrange function. We could see that this is that is the positive definite, that is why we can conclude that whatever optimal solution we got that was the global optimal for this point.

(Refer Slide Time: 29:54)



I am considering the second problem you see one is of less than type one is of greater than type and to handle this problem.

(Refer Slide Time: 30:05)

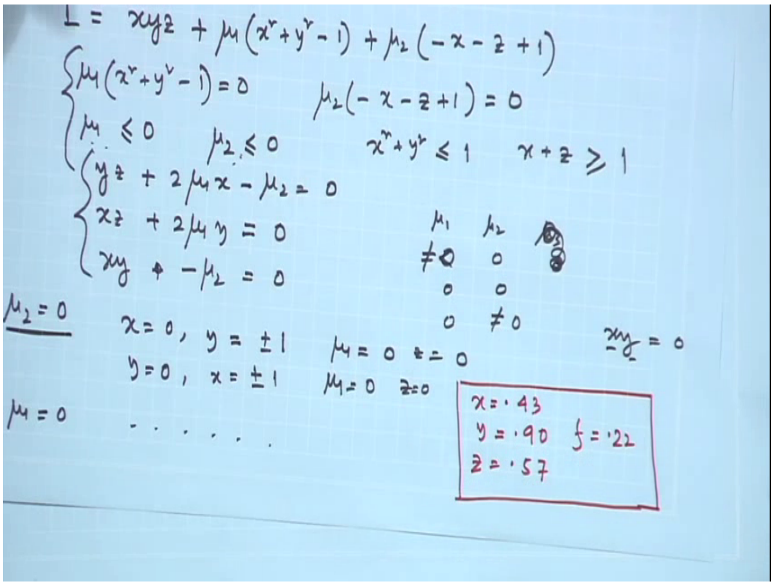


We will have the maximization of xyz subject to x squared plus y squared less than is equal to 1 let me take both as less than is equal to 1. Then less than equal to minus 1 then we can do it very nicely let us formulate the Lagrange function, xyz plus mu 1 x squared plus y squared minus 1 plus mu 2 minus x minus z plus 1 all right. What else we can say? We can say mu 1 x squared plus y squared minus 1 is equal to 0 mu 2 minus x minus z plus 1 is equal to 0. We can say mu 1 is maximization type less than equal to that is why mu 1 must be less than is equal to 0, mu 2 must be less than is equal to 0. We know what else we can have we can have x square plus y square less than is equal to 1. And x plus z greater than equal to 1. These are all the KKT conditions for us anything else I left. No it is oh I did not do the that is the optimality conditions.

We can have a set of optimality conditions yz plus 2 mu 1 x minus mu 2 is equal to 0, xz plus 2 mu 1 y is equal to 0, xy with respect to z, I am differentiating then it would be xy minus mu 2 is equal to 0 these are all the optimality conditions for us. Optimality feasibility that is a check for the complementary slackness can stations, check for the KKT multipliers the sign of KKT multipliers all these things. We have done what are unknown here again we can have the unknowns like mu 1 mu 2 mu 3, mu 1 mu 2 mu 3 can be less than equal to 0. Let us consider different conditions together less than 0 0, mu 3 is not there only mu 1 and mu 2. That is why both can be 0 both can be one less than type one can be 0 one can be non 0, rather not less than type non 0 0 0 0 0 non 0. These are all the conditions together. If we just solve mu 2 is equal to 0 rather mu 1 is not 0 we are getting solution as x is equal to 0 y is equal to plus minus 1, and mu 1 equal to 0 just to check whether we are getting this is a KKT condition or not. If mu 2 is equal to 0, this we are getting for x equal to 0 you see one thing if mu 2 is equal to 0 mu 2 is equal to 0.

Then we are getting xy is equal to 0, if mu equal to 0 then either x equal to 0 or y equal to 0. If x equal to 0 we are getting this if y equal to 0. We are getting another set of values x equal to plus minus 1, and mu 1 is equal to 0 and z is equal to 0 here also here also z is equal to 0 if you substitute the values here x plus z greater than one. It is not being satisfied in both the cases second case it is satisfying now that is why we will get a set of KKT point. Here at least I am just wanted to say that thing only if mu 1 equal to 0 we will get another set of KKT points. And if mu 1 and mu 2 not is equal to 0, we will get another set of KKT points ultimately we will get an optimal solution from here.

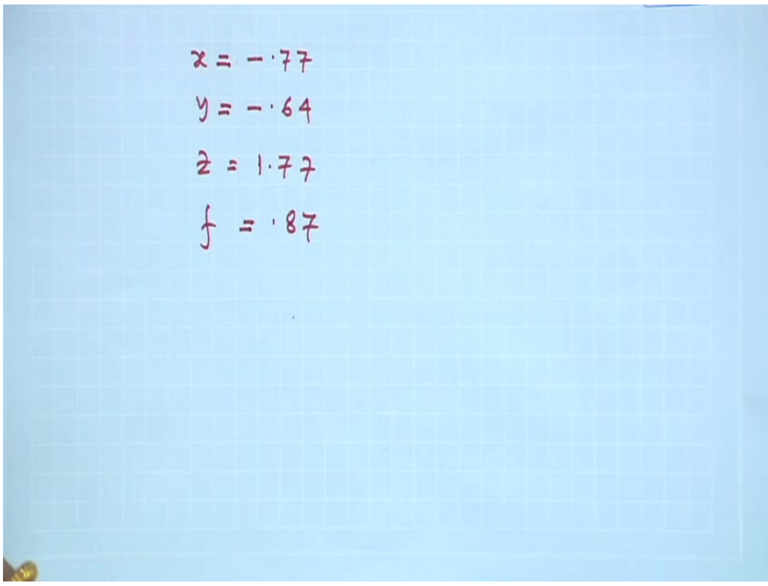
(Refer Slide Time: 34:36)



This is the you have to check it with different conditions of mu 1 mu 2 considering 0 non 0 etcetera.

We will get x is equal to 0.43 y is equal to0.90 z is equal to 0.57 and f is equal to 0.22. But if we relax the values see for the original problem there is no condition like x is the non negativity of the decision variables. That condition has not been imposed here that is why in other way we can say that xy and z these are all free variables you may consider that way as well, x can take negative value y can take negative value z can take a negative value.

(Refer Slide Time: 35:34)



If we just relax that condition, then we will get a better optimal solution for the same problem as x is equal to minus 0.77 y is equal to minus 0.64 z is equal to 1 0.77 and corresponding f we will get a better f 0.87, that is why you have to check this problem with a relax decision variables as well as you have to consider the case for the decision variable these are all non negative. And you have to find out the KKT points for this problem you have to find out the KKT point.

Which is optimal solution you have to check for the global optimality, this is of maximization time global optimality as well and you have to find out the decent feasible cone at different point. In that way just play with the problem and try to understand the whole geometry of the problem. And since this is a 3 dimensional problem geometry have to visualize in different way, but for understanding the geometry of the problem better you take a problem which is having 2 variables. That you can visualize nicely and what I suggest that we are giving the assignments to you and you just try all the assignments though we are asking for the KKT points and the optimal solution, but do not depend on that for each and every problem try to find out all the characteristic of the KKT theory and try to understand the non-linear programming problem geometrically.

Now, this is the part I have covered. I completed this week regarding the non-linear programming when the non-linear programming is of constraint type. Now in the next session I will take classes and the classes will be on the handling non-linear programming problem with different methodologies. There are different methods direct and indirect processes are there to handle and how really we are using the gradient etcetera, there let us see in the next.

Thank you for today.