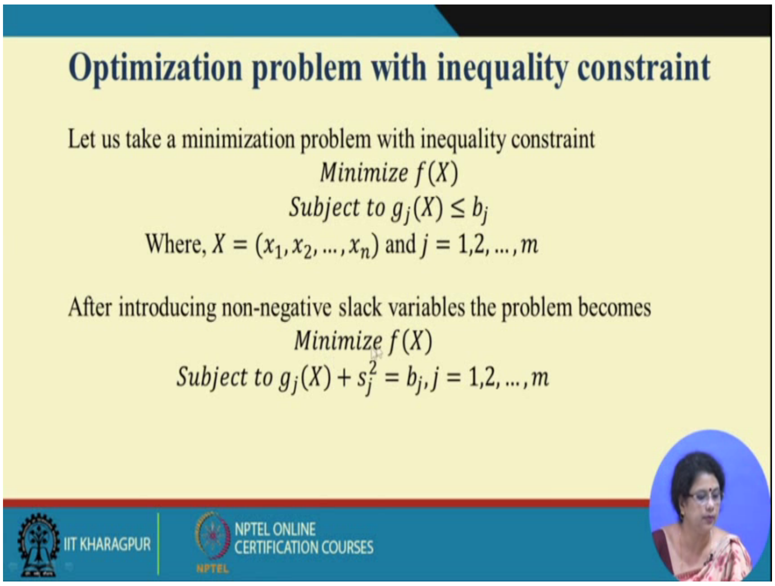
Constrained and Unconstrained Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

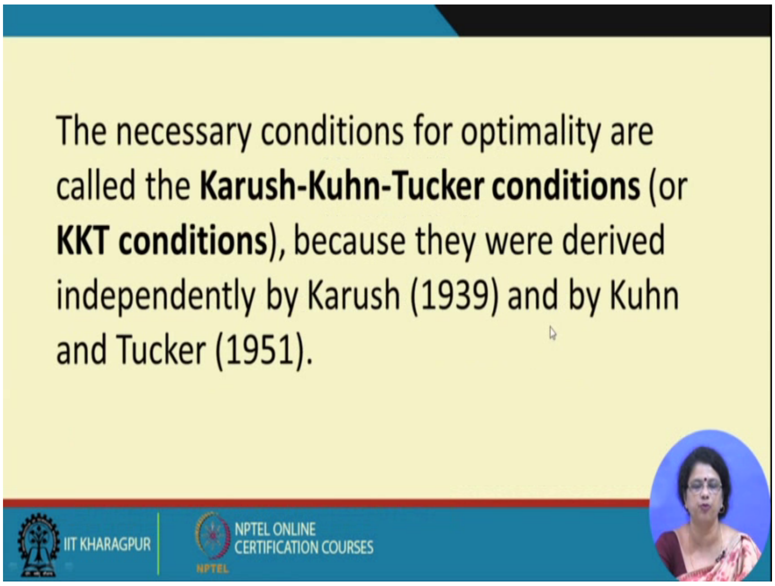
Lecture - 47
Constrained NLP – II

(Refer Slide Time: 00:39)



Now, in continuation to my previous class, I was dealing with the non-linear programming problem with inequality constraints. I just introduced the how we are getting the KKT first order necessary conditions FUS, but I did not say about the equations etcetera today I will talk on that. Now if I just start my class with again with the same general minimization problems non-linear programming problem, where the con there are m number of constraints which are of less than inequality type and we are introducing the slack variables slack variables are nonnegative slack variables as s j.

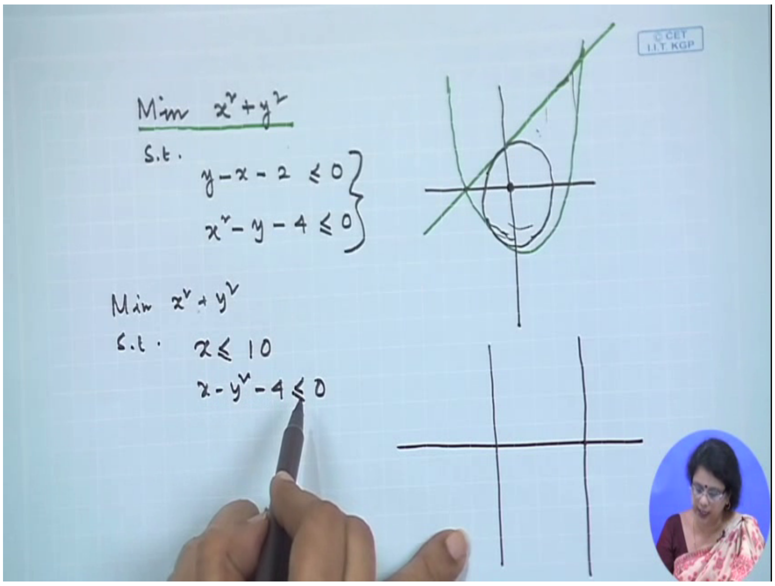
(Refer Slide Time: 01:09)



Now we will deduce the Kuhn Tucker condition necessary condition for this type of problem. Now this why we say this is the KKT condition there is a big history to it. In 1930, in 1951 Kuhn and Tucker; these are 2 scientists, they only developed the first order necessary condition and they declared that this conditions are when this conditions are sufficient as well for non-linear programming problem that was invented in 1951.

Since then it was being named as the Kuhn Tucker conditions, but after few years it has been seen that there was another person by the name of Karush; he already did in his Master's thesis the work he already in 1939; that is why since then people are calling it as a Karush Kuhn Tucker conditions, these are the names of 3 persons Karush, Kuhn and Tucker and few people are still say it Kuhn Tucker condition both are same. You can name it as either KKT condition or KKT condition, but we I we prefer the name KKT condition and this is the necessary condition, but there is certain properties are there for this necessary for the non-linear programming problem when this necessary condition will become the sufficient condition to get the global optimality.

(Refer Slide Time: 02:57)



Now, there are before going into detail about the KKT condition, I will explain few problems to you. Let me start with the first problem this is one non-linear programming problem minimization of x square plus y square subject to y minus x minus 2 less than or is equal to 0 and another one x square minus y minus 4 less than is equal to 0.

Now, as I showed you in the last class, we can get the Lagrange function from here by introducing the slack variable after introducing the slack variable both the in equations will become the equation, then after attaching the KKT multipliers instead of Lagrange multiplier because as I said KKT condition the process is the extension of the Lagrange multiplier process Lagrange theorem rather we can do it, but let us try to understand the that what is the geographical picture of this problem. Now the first one is the line, there is no doubt about it let me draw the line with the green y minus x minus 2 that is why I can say this is the line for us.

What is the next; next constraint is x square minus y minus 4 is equal to 0; that means, it is a parabola the parabola if I just draw again the parabola; parabola will be like this, this is the line this is the parabola. Now what is the objective function? Both are less than type that is why we can say that this is my feasible space. Now what is the objective function objective function is minimization of x square plus y square. Now within the feasible space, if we just look at the feasible space you can very easily find out the optimal solution is only one that is 0 0, because there only the objective functional value is minimum that is why your optimal solution will come at this point.

Now this is different from the example I explained you in the last class, where we have considered the constraint as the ellipse and the objective function as a circle, but in the ellipse we considered the sign as equality type that is why the feasible space was only the boundary of the ellipse, but here we are considering the space where we are considering the inequality constraints less than is equal to 0; both the cases. That is why we are having the feasible space is this whole feasible space, this one the feasible space together and here if you just find out the objective functional value, very easily we can say that it is the minimum value of the objective function, but if you ask what is the maximum value of objective function then we can say something else.

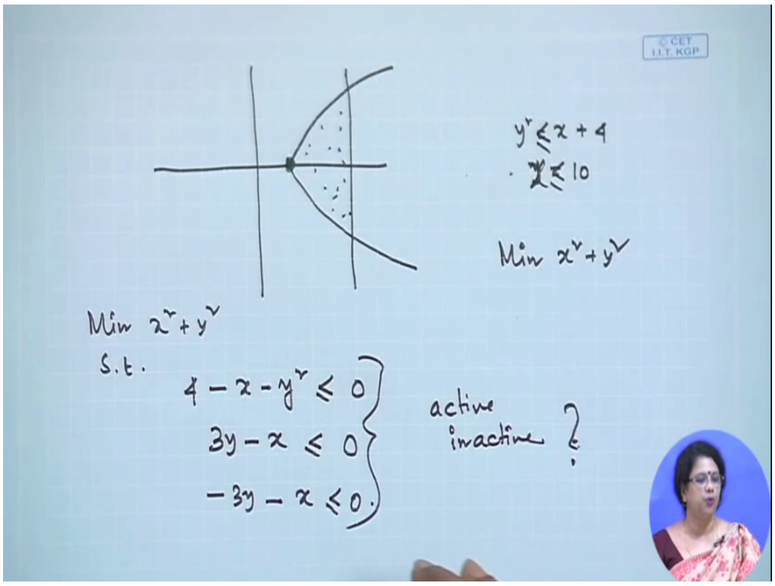
But if I considered the problem as a minimization problem what we could see that we could see that the both the constraints are inactive constraints; that means, both the constraints are not contributing anything to the optimal solution, in other way we can say that if we just change the constraint if we just tighten the constraint or if we relax the constraint, it will not affect the optimal solution as well at all.

But the question comes if this constraints are not contributing anything to the optimal solution for this minimization problem, but if we considered the maximization problem the situation is different because the circle will go further and further probably at the space somewhere, we will get the solution my figure is not symmetric that is why it is not giving you the exact solution; if you want, you can find out the maximum value for this objective function. In that case you can say that both the constraints or other of these 2 constraints are contributing on the optimal solution, but once we are considering the minimum of minimization problem nothing no constraint is contributing then the question comes why we will carry both the constraints further and further.

For our calculation can we not drop this is the question comes to your mind because unnecessary if you do the unnecessary conditions these and that it will take unnecessary longer process that is a bigger question comes in our mind. Now let me consider another problem for you minimization of x square plus y square subject to x less than or is equal to 10 and x minus y square minus 4 less than is equal to 0.

If this is if we just draw the graph of it what we could see that x less than is equal to 10 if it is 10 for us x less than is equal to 10 all right this side and y square is equal to x plus four; that means, it is a parabola again the parabola will be oh sorry, if we just draw it x is equal to 10 this is 0 0 then the parabola will be this one all right because we are having y square is equal to x plus 4 and y equal to 10 and both are of less than type less than equal to type then we can say this is my feasible space, all right.

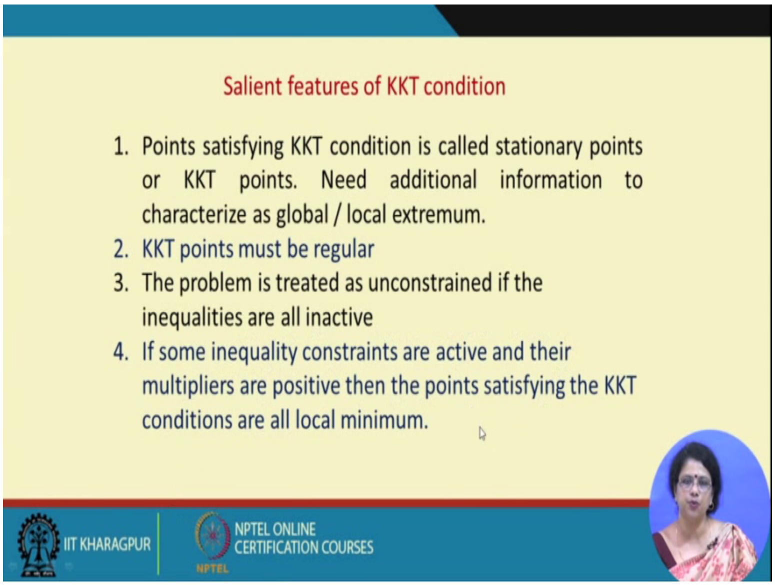
(Refer Slide Time: 08:47)



Now, we are considering minimization of x square plus y square, this is my objective function that is why again the circle center is 0, 0. It is going further, then where will be the optimal solution optimal solution will come here again you can ask that if the constraint y call to me; this is x equal to this is x equal to 10 if x equal to 10 is not contributing anything to the optimal solution why we really will carry this constant at all this is the inactive constraint.

Is it not now the last problem let me considered for you that is minimization of x square plus y square I am leaving to you; you find out which one is the inactive constraints and which are the active constraints or all are at active for this problem minus 3 y minus x less than equal to 0 you need to find out which are the active constraint which are the in active constraint now that is the question given to you. Now this is one part, but the then there is a big question in front of us in non-linear programming problem.

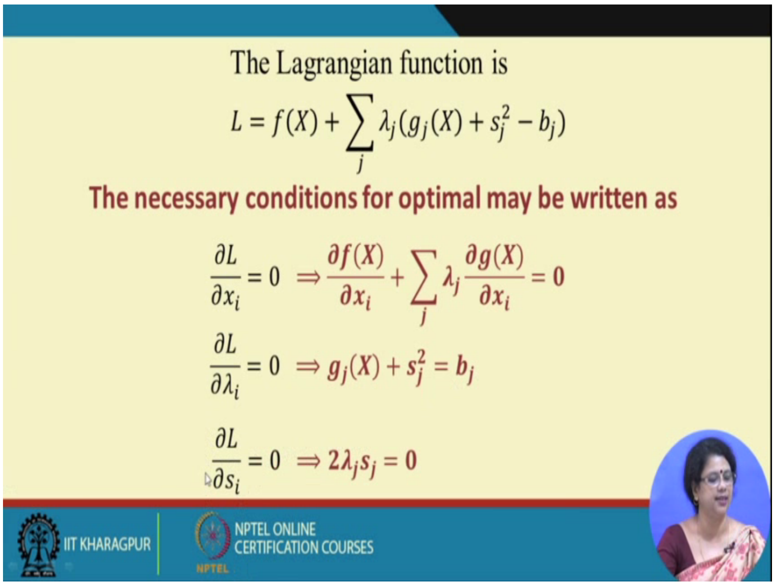
(Refer Slide Time: 11:08)



Now, KKT condition is such a condition through the necessary condition we very easily we can find out which one is the active constraint and which are the inactive constraint very easily and there are few feature for the KKT condition that since this is the necessary condition that is why the whatever points we are getting through the KKT conditions these points are being named as the KKT point and there are few more properties f the KKT point all the KKT point need will be either local minimum or local maximum and KKT points must be regular; that means, regular means this must be this must satisfy all the constraint not only that the gradient of the constraints at the KKT points must be linearly independent.

Then only this is the property of the KKT condition that the KKT points are regular and if we consider the unconstrained problem the just I showed you the first problem first problem can be treated as a unconstrained problem where there is no constraint at all because constraints or not contributing anything that is why there is no we can see that all the inequalities are inactive and if we have the inequality constraints few inequality constraints which are active when few in inequality constraints which are inactive through the KKT condition. We can very easily can find out for the inactive constraints then the KKT multiplier is 0 and for the active constraints KKT multiplayer will be non 0 that is the symptom will be given through the KKT condition that is why let me just go through the KKT condition in detail this is the Lagrange function when we have m number of constraints with us all are of less than type and we are introducing m number of slack variables there, then L can be of type f X plus this way; now let us deduce the necessary condition.

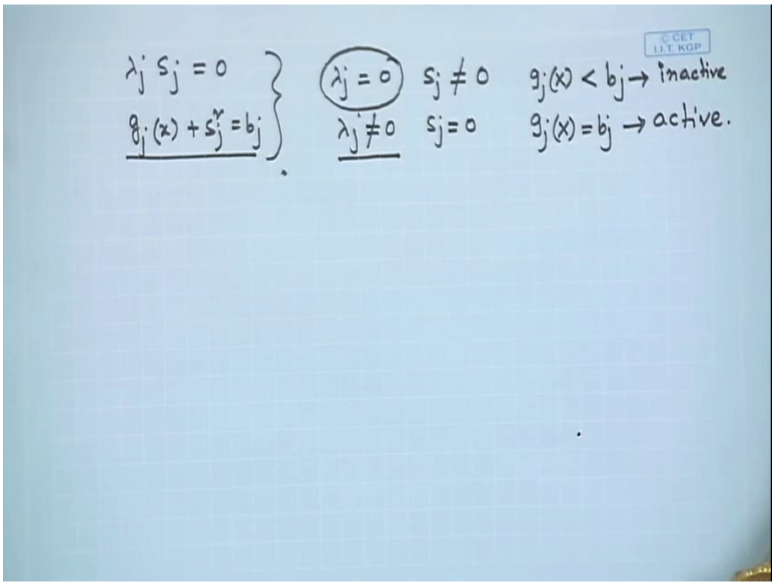
(Refer Slide Time: 12:49)



How many variables you are having you are having n number of decision variables you are having m number of KKT multipliers these are all lambda js we are having n number of slack variables which are nonnegative and the constants, these are the constant that is why we are having 2 m plus n number of variables . Now, therefore, we can in the first order condition we will have 2 m plus n number of equations; the first equation would be by considering the first order derivative with respect to x i there are n number of x i because we are considering the in the n dimensional space that is why very easily from here we can write it there is no confusion about it what about the next that if we consider with respect to lambda j because lambda j is a varying we do not know whether lambda j is positive or negative or 0 that could be anything we need to find out the values for this that is why these are the variables with respect to this variables if you just equate to 0.

Then we will get the original constraint from here we can write the gj x must be less than is equal to b j there is no; no problem at all and then is another constraint that is del I by del S i is equal to 0 because we are having the having the slack variables here and from here, if we just do we are getting 2 into lambda j into S j is equal to 0 this is being named as the feasibility condition the set n number of equations these m number of equations are being named as the these are the optimality conditions this is the feasibility condition and this is about the complementary slackness condition.

(Refer Slide Time: 15:24)

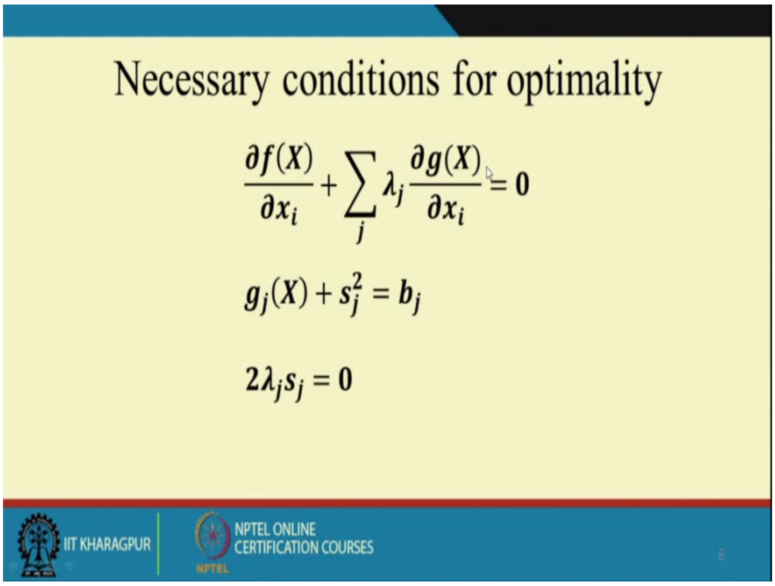


I will tell detail about it, but before to that I want to say few things to you just you see we are getting the condition that lambda j S j is equal to 0, I can just remove 2 from here there is no problem at all and we are getting g j X plus S square S j square is equal to b j; we are getting both the equations together, all right; what exactly we can say from the first equation that lambda that j can be j is equal to 0 it can be other option is that lambda j not is equal to 0 if lambda j is equal to 0 what exactly we are getting that S j is equal to 0 1 option what is the meaning of it.

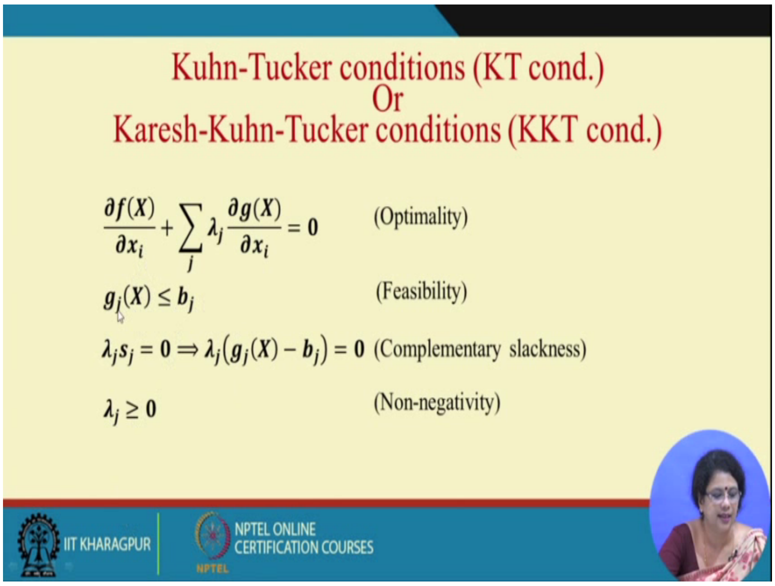
It means that from this equation if S j is not equal to 0, then we can get that g j X is strictly less than b j, all right; what does it mean? It means that this constraint is in active constraint at the at this stage because it is not being satisfied with the equality type just see the other way if lambda j is not equal to 0, then S j is equal to 0 then what exactly we are getting if S j is equal to 0; we are getting from here g j X is equal to b j hats why in this condition; we are getting these are all the active constraints all right that is why from here you can see you can conclude very easily for lambda j is equal to 0 corresponding constraint must be inactive and when lambda j not is equal to 0 corresponding constraint must be active that, but whether lambda j is positive or lambda j is negative that depends whether we are considering the minimization problem or the maximization problem.

The physical significance of lambda j; already I explained to you in the last class, but still I will say once more about when once more about the sign of lambda j for the minimization and for maximization both all right now again this is a necessary condition.

(Refer Slide Time: 17:47)

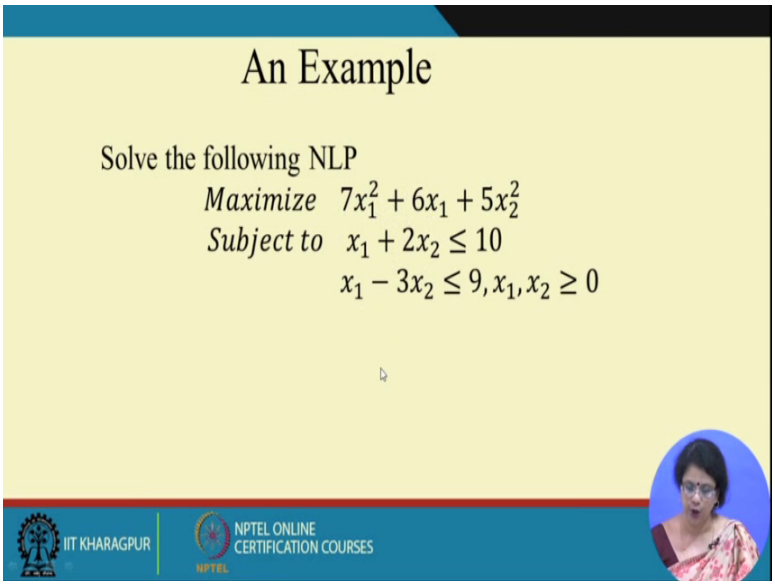


(Refer Slide Time: 17:59)



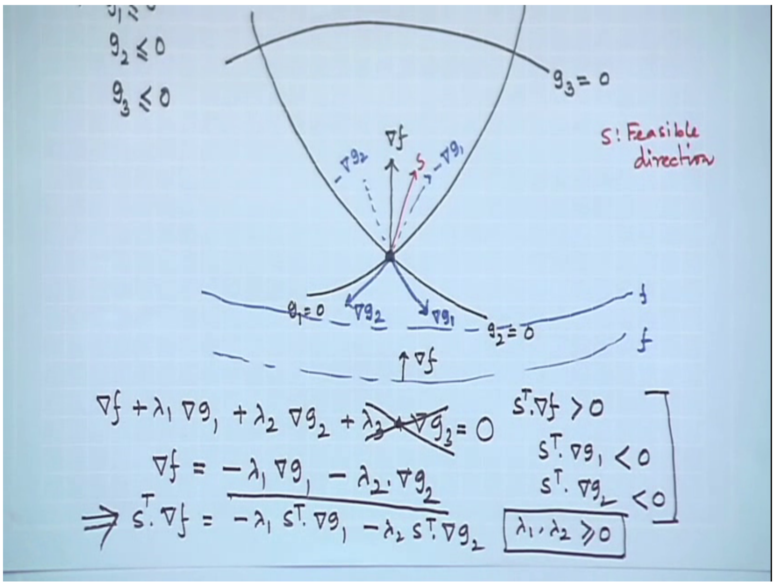
We can say and we are having how many number of unknowns we are having 2 m plus n number of unknowns and 2 m plus n number of equations, but we can get the solution for it this is the optimality condition this is the feasibility condition and the complementary slackness condition gives us says us many thing now there are few properties of lambda j, I will discuss this part later on this is for them that was for the minimization and at least we can say that lambda j is non 0 value; if we consider the less than type we will see if you remember I told you if lambda j is positive then for minimization problem this must be less than is equal to that is why we say it all. These thing I will detail and this is the reverse case for maximization problem.

(Refer Slide Time: 18:44)



Now, let us consider one example for here from here before we are saying before explaining the example I will choke few properties of these geometrical interpretation of this Lagrange multiplier we are considering for explaining these things we will consider 2-3 constraints.

(Refer Slide Time: 19:16)



One is g 1 less than is equal to 0, g 2 less than is equal to 0, g 3 less than is equal to 0, 3 constraints we will consider and we will explain for some meaning of and we have the objective function minimization of f let this is my g 1 equal to 0 this is my g 2 is equal to 0 and this is my g 3 is equal to 0; this is my g 3 is equal to 0. Now say this is the objective function for us; this objective function is moving. Now g 1 less than is equal to 0 g 2 less than is equal to 0 g 3 less than equal to 0 gives you the space this is a convex space.

Now, if the objective function if this is the level curves of the objective function could understand that here we will get the minimum value of f, but we can see here see this is the gradient direction of f is it not because objective functional value increasing this way that is why we can see that this is the direction of grad f at this optimal point this is my optimal point x star.

What else we can say at this point the gradient of g 1 would be where the gradient of g 1 would be this one all right and gradient of g to would be these 2 because these are all the at the optimal point if we just consider the tangent then gradient would be the normal to that tangent that is why grad g 1 and grad g 2 would be this way if I just extend it further, this is minus grad g 1 and this is minus grad g 2; what you can see one thing we can say that if we consider a any direct from this point through the feasible space I can move in different direction is it not let me consider one of the direction as this one which can be named as the feasible direction there is nice property of this feasible direction I will discuss in later classes, but I can say from here s is the feasible direction because if I move through this direction I will be still in the feasible space I can move further and further once I am coming here I this is not the feasible direction that could be the feasible direction that could be the feasible direction this could be the feasible direction, but from here at least if we see that this is the feasible direction s is the feasible direction all right now here up to this graphical.

Now, I am coming to the which part I am coming to the KKT condition what is my KKT condition grad f plus lambda one grad g 1 plus lambda 2 grad g 2 plus lambda 3 plus grad into grad g 3 is equal to 0. This is the optimality condition we got from the first equation now one what if this is for the minimization problem if this is the optimal solution then lambda 3 must be 0 because this is in active constraint g 3 is in active here that is why; this term will go out what exactly we are getting from here we are getting grad f is equal to minus lambda 1 grad g 1 minus lambda 2 grad g 2 into grad g 2 there is no problem.

But just if I just see here what exactly you could see that grad f can be represented in terms of combination of 2 grad g 1 and grad g 2 and with the minus sign that is quite acceptable to you what else you can see you can say that if s is the feasible direction here if grad g 1 is this side and grad 2 is this side, then we can say s is making acute angle with grad f, since it is making acute angle with grad f these are all vectors if we just take the dot product of it, then we can say grad f rather S T dot grad f S T means that is the transpose of S and dot grad f must be greater than 0.

Because its making the acute angle what else we are getting we are getting from here that S T dot grad g 1 and what you can say about it grad g 2 s is making acute angle with grad f s is making acute angle with minus grad g 1 S is making acute angle with minus grad g two, but on the other hand s is making obtuse angle with grad g 1. Similarly with grad g 2 that is why both are coming negative that is very much reflected from this equation because if we just pre multiply with s this one s t this is equal to minus lambda one S T grad g 1 minus lambda 2 S T grad g 2 clear.

That is a very much reflected here and what is the necessity for us the necessity is that lambda one and lambda 2 both must be greater than equal to 0 because this conditions are holding because of this only, all right that is why from here what we can conclude that if we consider s as a feasible direction in the feasible direction S if it is. So, then S into the gradient of the objective function is greater than 0 s into mean say dot product inner product I am considering grad of g 1 is less than 0 s inner product with grad of g 2 is less than 0 since we are considering a minimization problem, then can we conclude that lambda one lambda 2 greater than or equal to 0.

From this condition think about it that is why as I said before that for the minimization problem, there was a cons condition that if g 1 is less than is equal to 0 we have considered that that KKT multipliers must be if I just go back to that page just you see a sorry look at this if we consider the constraints of the type g j less than is equal to b j, then we can conclude that all the KKT multipliers must be positive and if we just extend this geographical interpretation with the constraint g j greater than or equal to b j, then we can prove that the KKT multipliers are always less than is equal to 0 that is the task I am giving to you for the set of constraints try graphically and geometrically try to prove that lambda j always will be negative for the type g j greater than equal to b j.

Similarly, for the maximization problem the cases is reverse if we consider here also we can explain because is if the function is moving further and further we can find out the optimal solution here is it not we can see the optimal solution will come here because the objective function is moving this way these are the level curves of the objective functions this is the optimal solution just try to prove for less than equal to type always lambda g would be lesser than equal to or lesser than 0 equal to 0; we are including because few constraints are the in active constraints for the in active constraints always the lambda j equal to 0 and for the active constraints only lambda j is non 0 that is why try to prove that for maximization problem we will get once we are considering the set of constraints as g 1 less than is equal to 0; g 2 less than is equal to 0 etcetera then you have to prove that lambda the KKT multipliers must be non positive.

Similarly, for the constraints of the type g j greater than equal to b j the KKT multipliers must be nonnegative with that; today I am concluding my class. In the next, we will solve few problems on KKT conditions.

Thank you very much.