Constrained and Unconstrained Optimization

Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 44
NLP with Equality Constrained-I

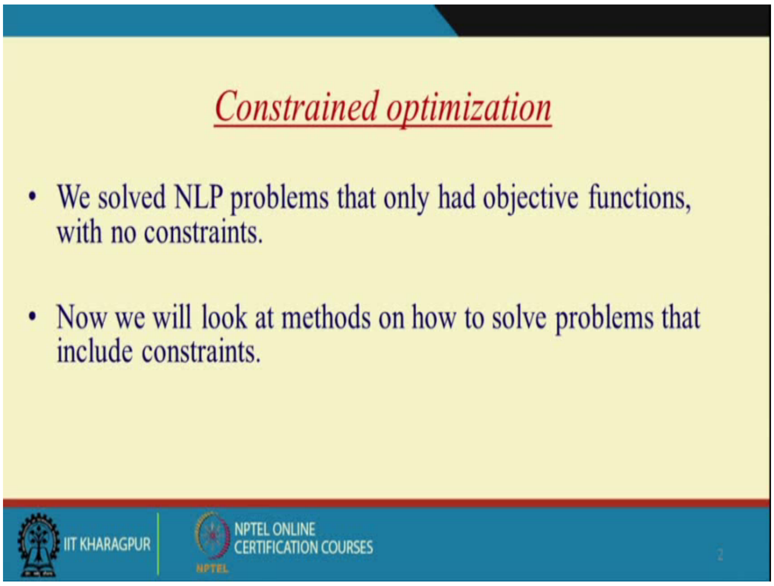
Now, today's topic is the again the non-linear programming problem, but we have done enough with the unconstrained optimization problem. Now that was the simplest form of a non-linear programming problem where there was no constraint at all, but now we are moving to the situation, where we need to optimize the objective function under certain restrictions. That is why we would say the name of this topic as a non-linear programming problem with constraints, but you know the constraints can be of 2 types the constraint can be of it, it can be linear it can be non-linear this way it could be 2 types. Otherwise we can have 2 types of non-constraints with respect to the sign of the constraint that is a either the constraint is of equality type or the constraint is of inequality type.

Inequality can be less than equal to and or greater than equal to, but equality means equality. That is why if it is draw the feasible space for the equality constraint you might be understanding that the constraint would be the feasible space would be only can will be considered the boundary of the space, not the inner space will be considered as a feasible space, but if we consider the less than equal to or greater than equal to then either of the side of the feasible the boundary of the region will be considered all right.

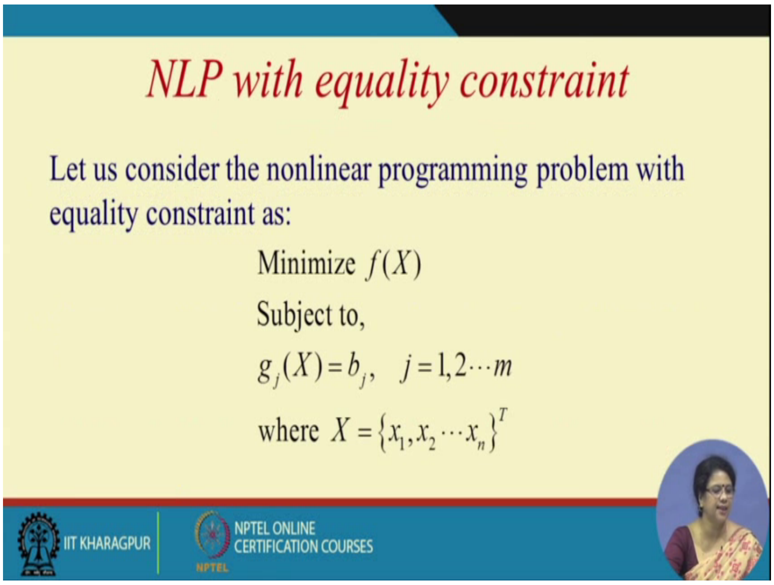
Let us today my topic is non-linear programming with equality constraint, I will show you what does it mean geometrically. Then I will show you how to handle rather how to get the optimal solution for that kind of non-linear programming problem. One thing is that in non-linear programming there is no you must have been realized that there is no unique methodology to handle the non-linear programing. That you have experienced for linear programming, there was simplex algorithm in one go you will get the optimal solution in every type of problem you can apply the simplex algorithm maybe of the normal simplex revised simplex dual simplex. Whatever it is the advanced form of simplex algorithm, but these are the simplex methods, but you must have realized in the unconstraint non-linear programming problem, there are several kind of methodologies available. Depending on the situation rather whether the function is quadratic function then is non quadratic function is discrete function is continuous depending on different situations, whether the function involves single variable multi variable all depends on that several methodologies are available. These are the tools available for us and according to a requirement we use it.

Now, here again another kind of non-linear programming problem now you are learning the process to handle non-linear programming problem in different situations. That course is offering you that kind of knowledge, so that you can apply in your practical area all right.

(Refer Slide Time: 03:37)

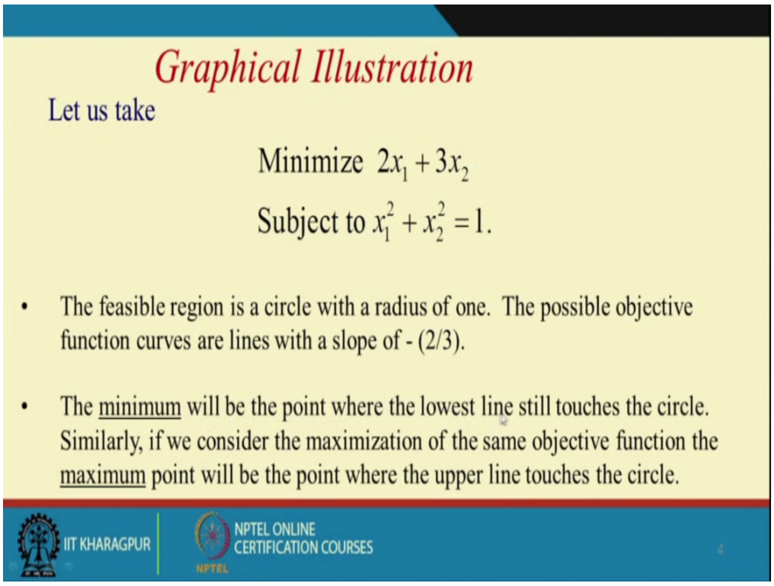


(Refer Slide Time: 03:43)



Now, equality constraint, and equality constraint means simple equality constraints that is why we can say the general multivariate non-linear programming problem where the variable is defined in rn in n dimension. And this there is a set of equality constraints there are n equality constraints n number of variables all right. What does it mean geometrically rather graphically I will show you?

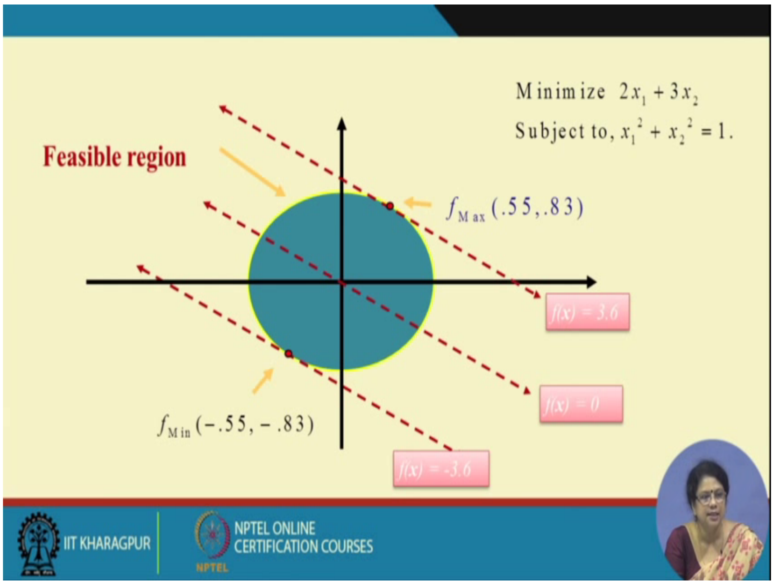
(Refer Slide Time: 04:08)



Now, there is a problem for you minimization of 2 x 1 plus 3 x 2 subject to x 1 square plus x 2 square is equal to 1. What did you see there is a constraint through this optimization problem, there is one objective function fortunately the objective function is of linear type? So, that you can visualize that space in a better way it can be non-linear as well, but the constraint is non-linear. And it is circle simple circle it is a circle centered at 0 zero radius is one all right.

Now, if I can draw the feasible space as a simple boundary of this circle all right and the objective function is moving in the x 1 x 2 region freely. There is no restriction for x 1 and x 2 all right. That is why let me see it has been said the objective function is a straight line which is having a slope of minus 2 by 3. Now where is the minimum will be the point, where the lowest line still touches the circle because the line is moving freely I will show you the figure. And what is the maximum the line is moving through the space and it will touch to that point where the if I just go beyond to that I will be out of the feasible space what does it mean, I will show you this is my problem.

(Refer Slide Time: 05:47)



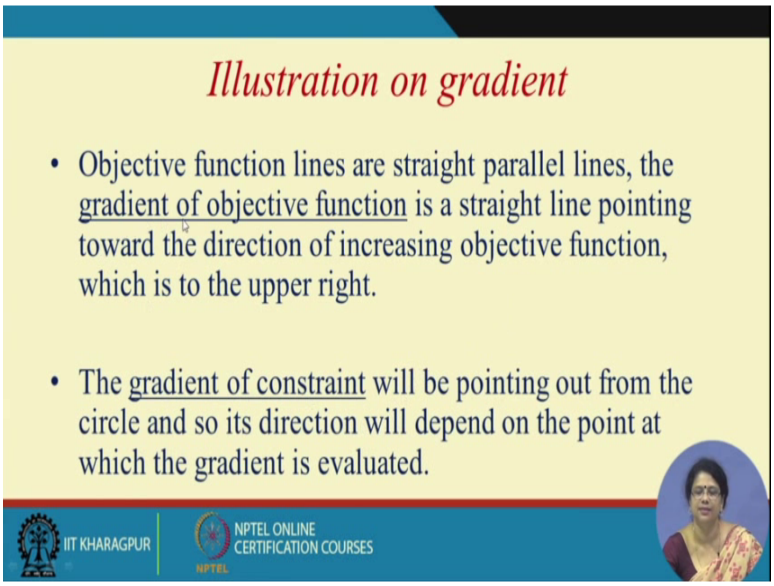
That is why x 1 direction x 2 direction this is the feasible region that is the region is only the yellow line. In a line is not included because the sign is equality type there is no less than equality all right.

Now there is a line where f x is equal to 0, this is the value of f x at 0.00 this is moving freely here. If this line is moving this way functional value of that objective function will maximize, and if it moves this way then the functional value will be the value will reduce now for minimization always it will be it will go this way for maximization, it will go that way that is for sure. That is why in other way we can say that the gradient of the line is positive increasing this way and decreasing this way all right. That is why you see this is the minimum for us, because if I go beyond to this point then I will be out of the feasible space and that would be the minimum point I need not to explain this part because you must have done you the graphical solution of linear programming here is the same way only difference is that feasible space is non-linear in nature. And where we will get the maximum value maximum value we will get here because of this symmetricity of the space we are getting the functional value is 3.6 here and here minus 3.6 all right.

Now, there are certain properties of these whole fact what is the property of these just to think the gradient direction of the constraints it in the gradient direction of the objective function. This is the gradient direction of the objective function this way. And this is the reverse direction, reverse gradient direction. For the constraint what is the gradient direction gradient direction, will be if I consider this point that will be a point at that point if I just draw a tangent normal to that tangent outside would be the positive gradient direction at that point.

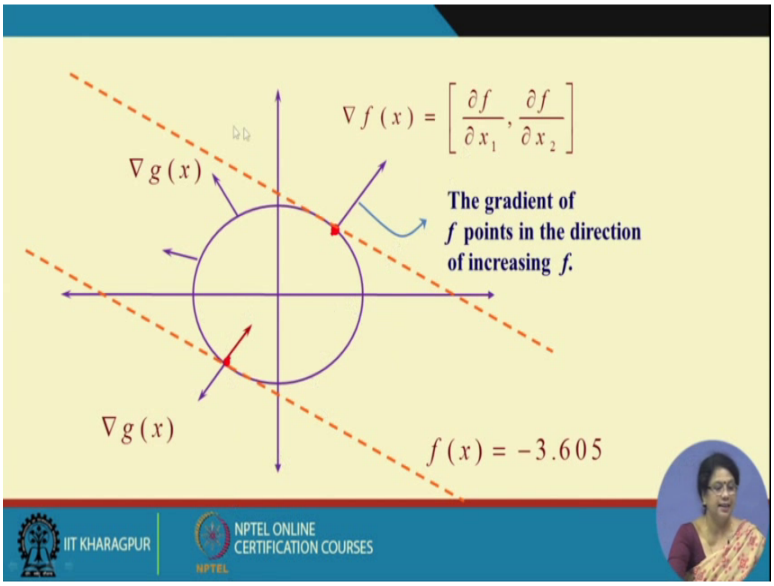
Similarly, if I just move through the whole boundary of the circle at each point we will get a gradient direction that will be nothing, but the rays these are emanating from the circle outside all.

(Refer Slide Time: 08:36)



And these are all the gradient direction with this idea this stage just sees a nice fact in the next that if I say very variably it is being said that gradient of objective function is a straight line pointing toward the direction of increasing objective function. And the constraints would be pointing out from the circle, it is direction will depend on the point at which the gradient is evaluated that way I can see.

(Refer Slide Time: 09:00)

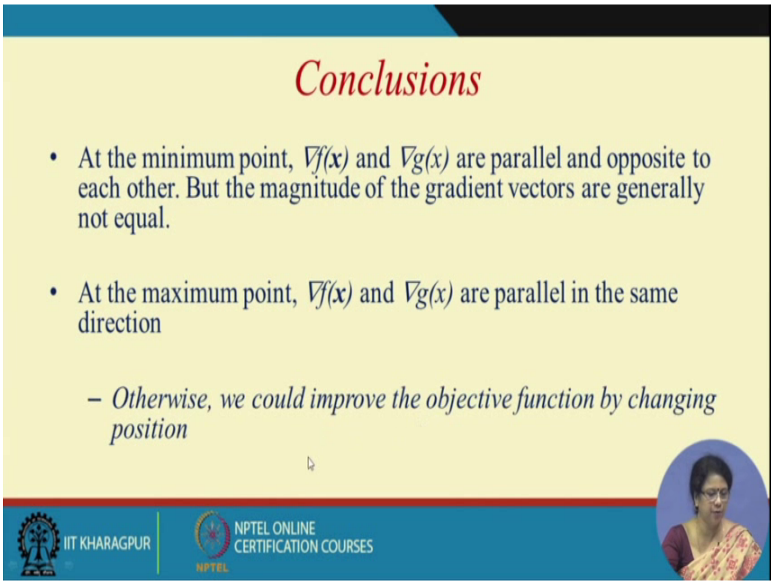


If this is the circle at this point, this one this is the gradient this is the gradient it that should be the normal this is the gradient this is the gradient this way understood.

Now, for the same figure the fact you see this is the circle this is the maximum value gradient direction is this one grad f this is the minimum value where is the gradient direction gradient direction of the objective function this way, but at this point the constraint is this way what did you see at, but the fact you got from here that at the minimum point you must be seeing that the gradient of constraint and the gradient of objective function just opposite to each other. Did you see and what about at the maximum value? At the maximum value gradient of the objective function and the gradient of the constraint is in the same direction. It will happen this is the no this is the very nice fact for non-linear programming. Any non-linear programming problem always is in the minimization problem, this is the gradient direction of constraint this is the gradient direction (Refer Time: 10:31) always it will make 80 degrees, but here it will make the 0 degree.

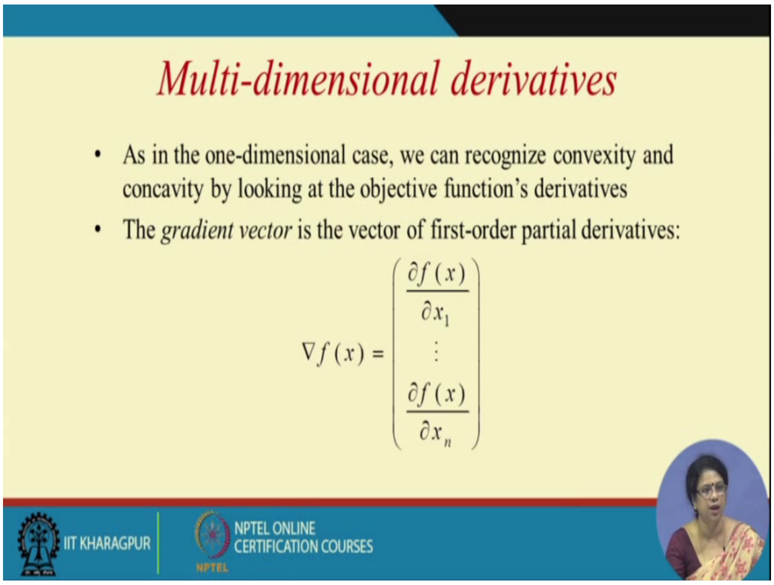
We will use this fact we will revisit this fact, in the next when we will study the next level of Kuhn tucker condition. There also we will just reconfirm this fact.

(Refer Slide Time: 10:49)



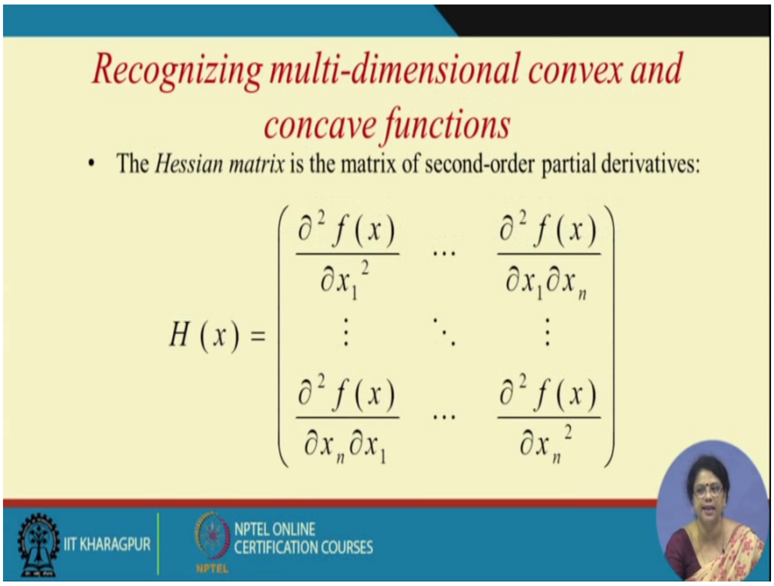
That is why if we just conclude from here that at the minimum 4 point del f and grad f and grad g are parallel, but opposite to each other and the magnitude of gradient vectors are generally not equal certainly, that depends on the point on which we are calculating the gradient and at the maximum point both her parallel in the same direction.

(Refer Slide Time: 11:17)



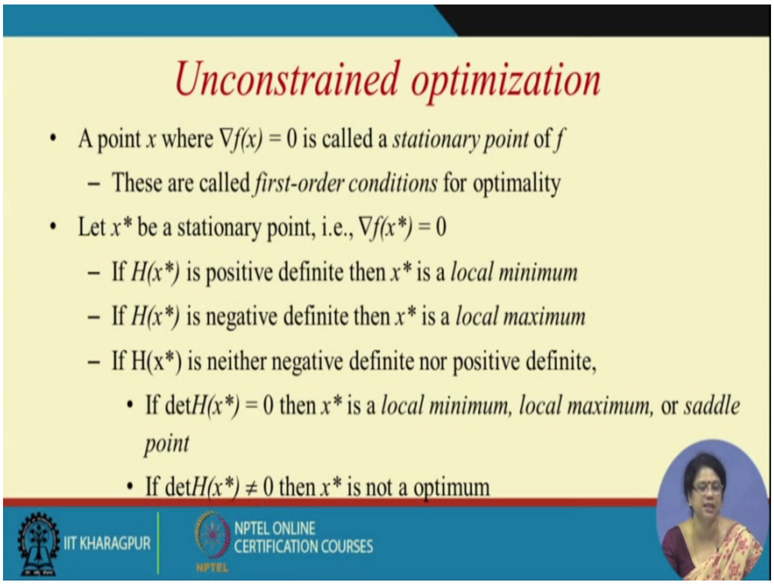
Now, if this is the fact how really we can consider a multi-dimensional non-linear programming problem by using this gradient vector and we want to solve a constrained non-linear programming problem, we will use this fact. How we can use it we know the general classical optimization technique that if we consider the necessary condition first order derivative p z is equal to 0. And in case of multi-dimensional case we say the grad f is equal to 0. That is a vector which consists of first order partial derivatives of the function with respect to different variable that is equal to 0. And what about the sufficient condition for the sufficient condition we are going to the second order derivative, and that is why there is a concept that is called the hessian matrix we form the hessian matrix in this way, del 2 f by del x 1 square del 2 f by del x 1 x n.

(Refer Slide Time: 12:09)



This way that is the n by n matrix all right square matrix and depending on the property of this hessian matrix we declare whether the sufficient condition whatever extreme points we rather the stationary points, we got from the necessary conditions at this point whether it has the function has minimum or maximum or the saddle point this fact you already know, all right.

(Refer Slide Time: 12:47)

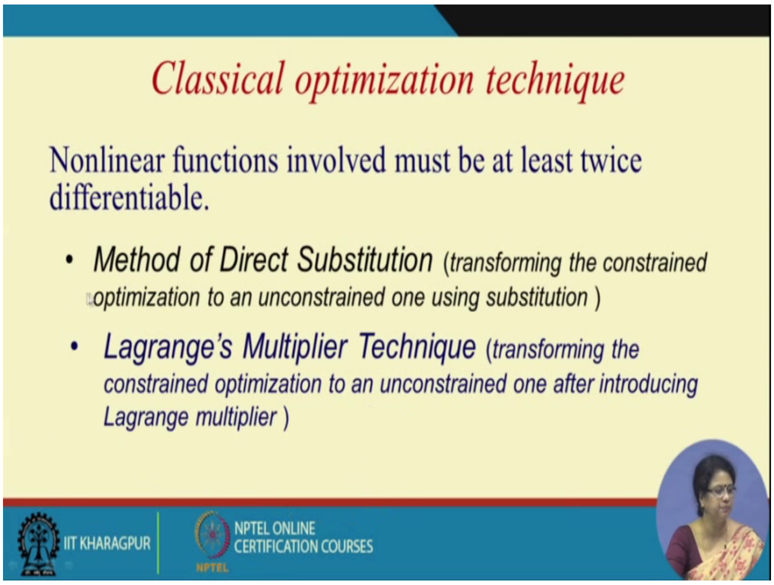


That is why what are the facts, it says you that if we see the hessian of the matrix is positive definite then it has the local minimum.

But again the positive definite matrix what is the meaning of it, yeah there is a matrix now we will say that is a positive definite matrix. If all the principal minors of the matrix these are all positive or in other way, we can say that if the eigenvalues of that matrix are all positive then the corresponding matrix is positive definite matrix. That is why at the local minimum point if x star we got it from the necessary condition that grad f x star is equal to 0. And if we see the hessian matrix value at x star is positive definite then we will say local man minimum if the hessian matrix is negative definite, we will say the local maximum.

And if we see the principal minors are fewer positive fewer negative and there is no pattern to it then we will say for at the point at that point function has neither maximum nor minimum, but still there is certain condition we say if we see the hessian matrix hessian matrix value at the determinant value is 0, then the for the function may be has a local minima or maxima or saddle point because we need to go for the third order derivative of the partial derivative of the function. But if we see the this value is not equal to 0 and, but there is no pattern to it will declare there is no optimum value for it this the in the detail of it I will explain you later on, but with this fact with this gradient of the function how it is being used for the non-linear optimization problem I will show you in the next.

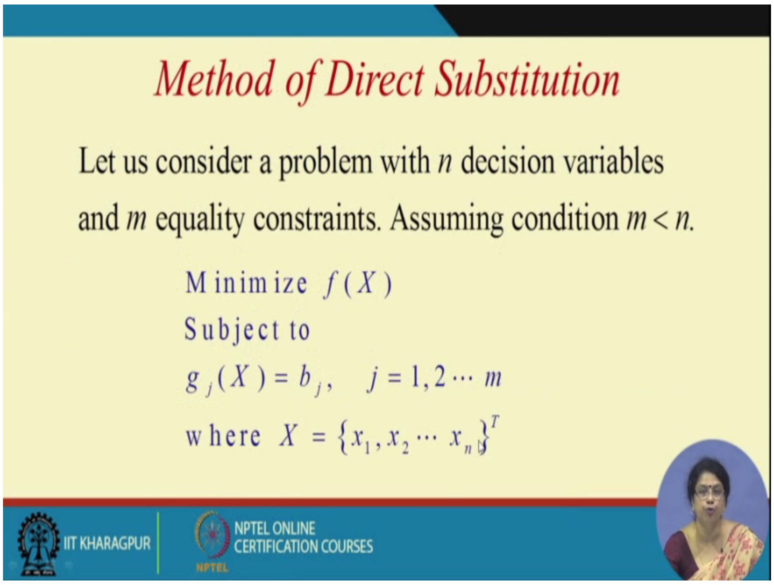
(Refer Slide Time: 14:41)



Now, let us consider a non-linear optimization problem where the non-linear functions are twice differentiable. If these are not one is differentiable we cannot check the sufficient and the necessary and sufficient conditions. Because hessian matrix needs that it must be differentiated twice all right. That is necessary condition for us and there are 2 methods available for solving non-linear programming of equality type one is the method of direct substitution. And the second method is the method of Lagrange's multiplier today I have a plan to explain you the basics of Lagrange's multiplier method.

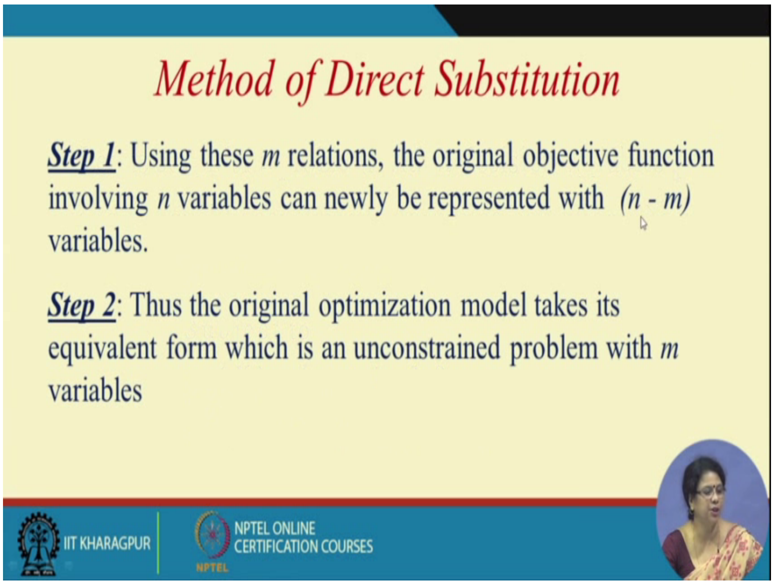
Now, and the method of direct substitution is a simple one. Let me just go through quickly of this method one by one.

(Refer Slide Time: 15:40)



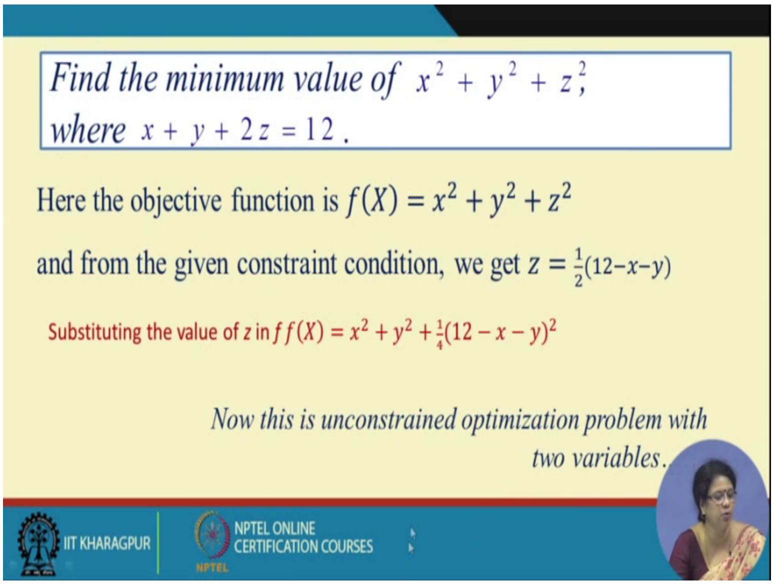
Let me consider the direct substitution method for a non-linear of this type all right always you need to consider a must be lesser than in then, only the problem is complicated if m is equal to n problem is very simple for us. We will get exact number of points and we will get the optimal solution very easily from there.

(Refer Slide Time: 16:04)



And method of direct substitution what we do if I just say, where while it says that aim relations are there in variables are there you replace aim relation with n minus m variables what does it mean just.

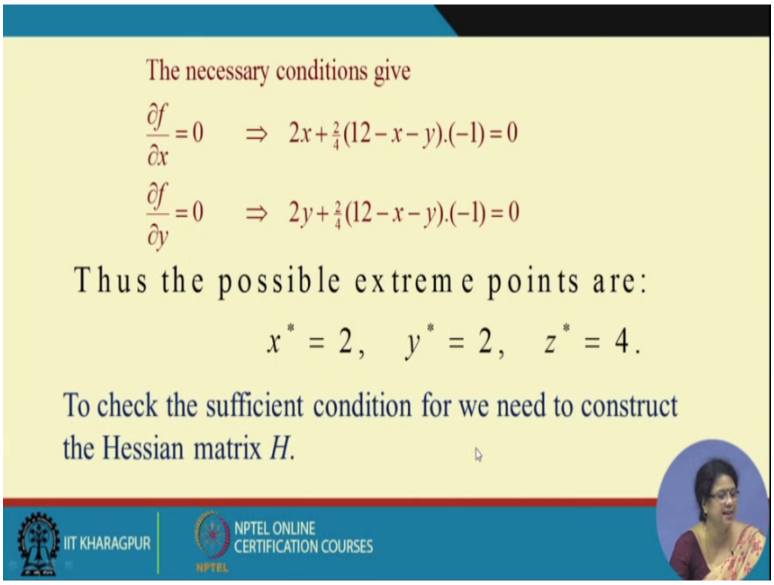
(Refer Slide Time: 16:19)



I am taking one simple example here. This example is having 2 3 variables one objective function, one constraint from the constraint. What we will do we will make one variable in terms of n minus m number of variables, n is here 3 m is here one. That is why z can be represented in terms of x and y you can represent x in terms of y and z you can represent y in terms of x and as you wish.

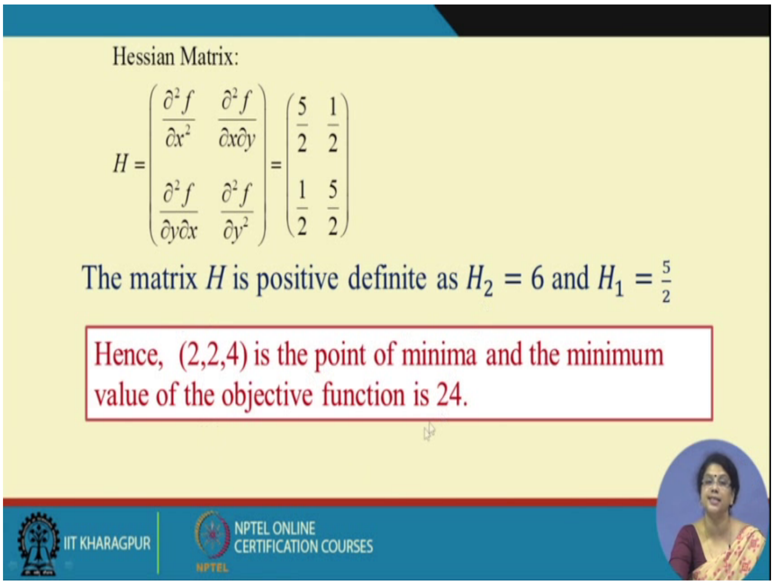
Once we could represent it what we will we will redefine the objective function this way. Once we are redefining the objective function by substituting the value of z here and what is the problem is coming ultimately the problem comes as a unconstraint optimization problem for me. That is a very simple that is why the method is a simplest method of direct substitution, but this method fails when this constraint is of complicated type you cannot make one variable in terms of other 2 variables. After substitution the function may become a very difficult function. There this method fails for the simple cases it made method works very nicely all right.

(Refer Slide Time: 17:43)



Now, if this is the case just consider as an unconstraint optimization problem go for the necessary condition go for the sufficient condition. And you get the solution for it these are the extreme points that we could get from the necessary conditions because unconstrained optimization problem was having 2 variables that is why this is the grade f values if we just equate to 0. Then we are getting the extreme points, but the ness sufficient condition has to be checked through the hessian matrix.

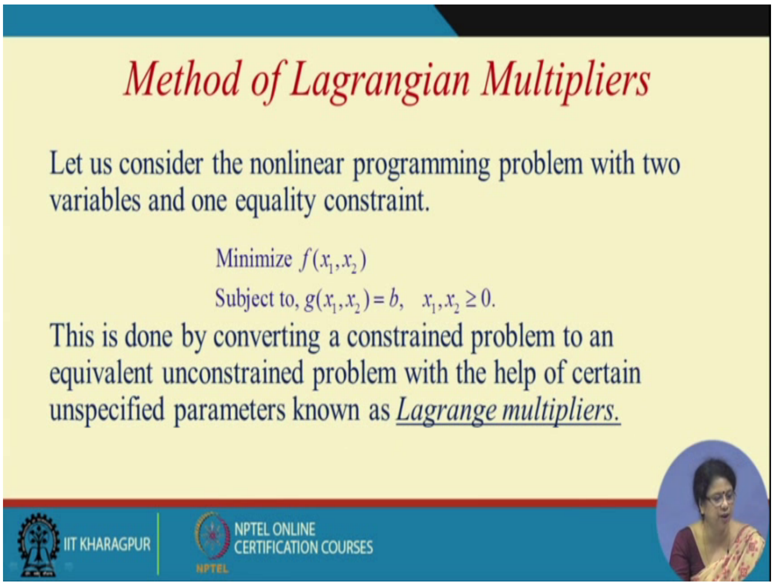
(Refer Slide Time: 18:10)



Hessian matrix construction it is not a difficult for us we will simply calculate the hessian matrix. This we just look at the matrix hessian matrix the principal miners are the first principle minor, if first 5 by 2 positive second principle minor would be the value would be 5 by 2 into 5 by 2 minus half square. And the values are coming as one is 6 and the other one is 5 by 2. Both are positive. That is how we can declare that we are getting a positive definite hessian matrix that is why whatever extreme points we got it from the from the necessary condition, that is that gives us the minimum value of the objective function.

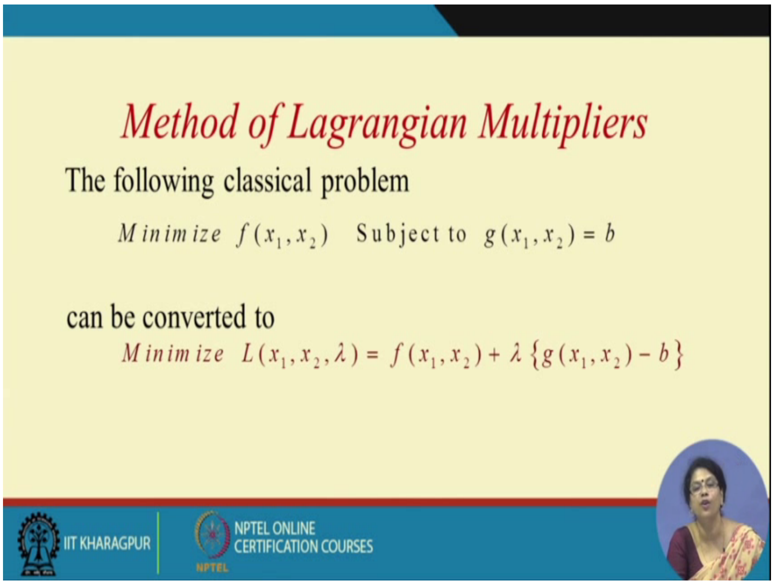
Now similarly we can go for we can search for the, if at all exists for the maximum value. This is the simple method of direct substitution, but there was there is another method that is a very popular method that is called the method of Lagrange multiplier. Then new thing of this method is that this is the name has come from the mathematician by the name of leg professor Lagrange from here, it is it has come we are formulating one Lagrangian function for the given optimization problem by using the Lagrange multiplier, and after that let us see what is happening. And for this problem again we are considering same kind of problem.

(Refer Slide Time: 19:40)



Let us consider the simplest case 2 variables are there one simple equality constraint is there. Because if there are more number of equality constraint some complications are arising out of feet, that I will discuss in the next class. Today I am considering the simplest case of Lagrangian multiplier method where only one constraint is there. From here we can construct a Lagrangian function by using a Lagrange multiplier, Lagrange multiplier is we are considering as a parameter the value of the parameter is unknown to us beforehand, but this parameter will help us to detect whether the optimal solution we got there is a maximum or minimum. That is the beauty of that Lagrange multiplier. There is a physical significance of the Lagrange multiplier, I will come to that point if after explaining the methodology in detail.

(Refer Slide Time: 20:44)

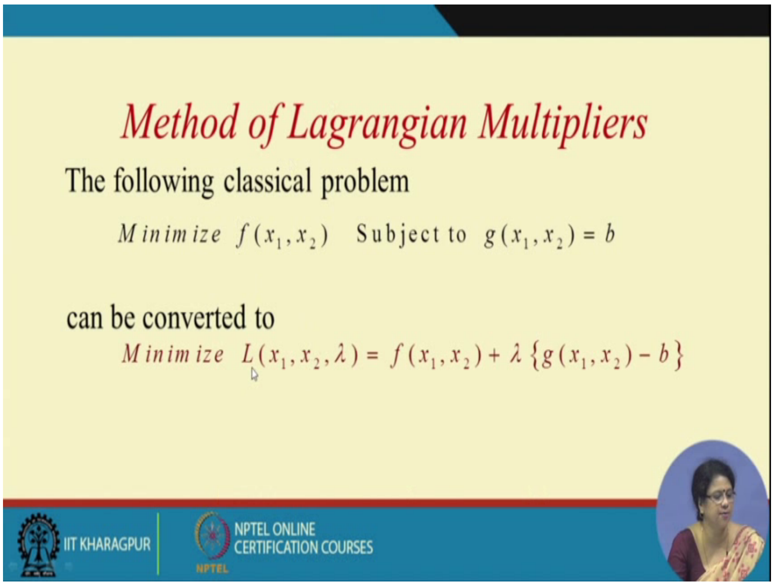


Now, this is a problem how the Lagrange function is being formed this way Lagrange function, if we just introduce one variable lambda then the Lagrange function will become a Lagrange a function of 3 variables x 1 x 2 and lambda. What is that? That would be is equal to f plus lambda into g minus b. Because g minus b is being considered as a from the constraint only all right.

Now, what we have to do we have to minimize this Lagrange function. Now whatever tool you know till now, again I would say just try to apply it here. We have you see we have incorporated the constraint part as well within the objective function. That is why without any substitution that is. So, even the constraint is a is a complicated in nature we need not to bother about it because we need not to do any simplification, just we will introduce one lambda there and we will just put it at the end of the objective function we will append it there. Now it will become again an unconstrained optimization problem.

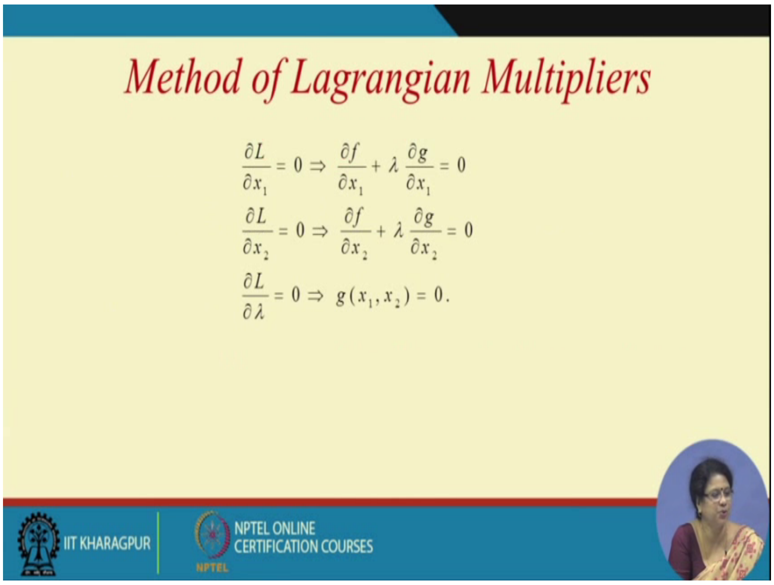
Again I would say for an unconstrained optimization problem, if the function is twice differentiable then the best way to handle it go for the first order derivative grad of the function Lagrange function equate to 0 get the stationary points then after that, find out the hessian matrix and see the nature of the hessian matrix. And from the nature of the hessian matrix try to calculate the optimal solution. That is the simplest way again that is why in the Lagrange function you see there are a few things.

(Refer Slide Time: 22:33)



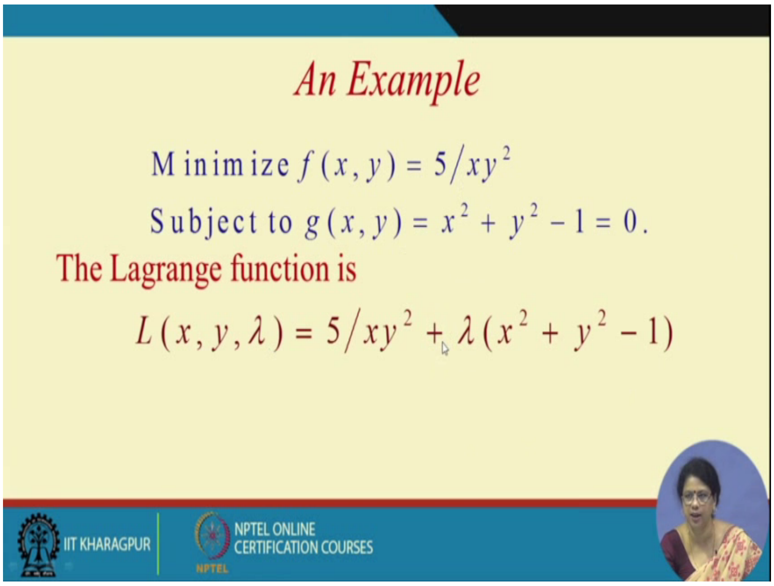
One is the function one is the parameter. Just remember it because you may ask me why I have deliberately we are introducing one parameter in it has no meaning of it is not like that this is the necessary. Condition through the Lagrange because since it has now 3 variables.

(Refer Slide Time: 22:53)



Now, we can say that del l by del x 1 must be is equal to del f by del x 1 plus lambda into del g by del x 1 b is a constant for us that is why that part will not come all right. Now del l by del x 2 must be is equal to del f by del x 2 plus lambda into del g by del x 2. But from the third constraint del l by del lambda equal to 0 you see you see we are getting the original constrain, but that should not be 0 that must be is equal to be all right, g x 1 x 2 must be is equal to b from the Lagrange function, if we consider b as a 0 then only it will become all right that is why let us see one example.

(Refer Slide Time: 23:49)



For applying this method, you see the we are getting in just now we are we are landing into a situation where the situation is simple, but you see objective function is having a non-linear function constraint is also having a non-linear function. That is why we will see how we can handle through the Lagrange multiplier.

Now, one thing is you see if we just do this direct substitution here, we may face difficulty all right. May face may not face because y square can be substituted with x square nicely, we can do for hip for this problem otherwise. Now for this problem let us apply the Lagrange multiplier method, let us formulate the Lagrange function. After introducing the Lagrange multiplier. Since there is only one constraint we are introducing one Lagrange multiplier, if we have 2 constraints we will introduce 2 Lagrange multiplier.

Now, you see this is the thing now we can go for the necessary conditions, just do the calculations to get the necessary conditions. We will get this thing what you have you have 3 equations with 3 unknowns. Is it not? From here can you find out the values of x y lambda, we can very easily we can find it out. That just task of yours I need not to do it, but you will get you we can do certain simplification from here. And from here we will get one relation between x and y, we can say at optimal point this relation will hold all right. That is why we can declare one extreme point could be this one from the relation. We can have another extreme point as well.

Now, if you want to check in the individual extreme points whether it has maximum or it has minimum you need to do the you need to calculate the you need to see the property of the hessian matrix all right. That part you should do to get the sufficient condition this I am keeping as a as an assignment to you do it on your own. Find out the hessian matrix find out the principal minors or you can go for the Eigen values calculate the thing and just check whether this extreme points is a minimum value or not all right.

And with that, I am concluding today and in the next class we will deal with the Lagrange method for multi-dimensional cases where we will have several constraints. And together all are of equality type and we will see the Lagrange multiplier method there.

Thank you very much for today.