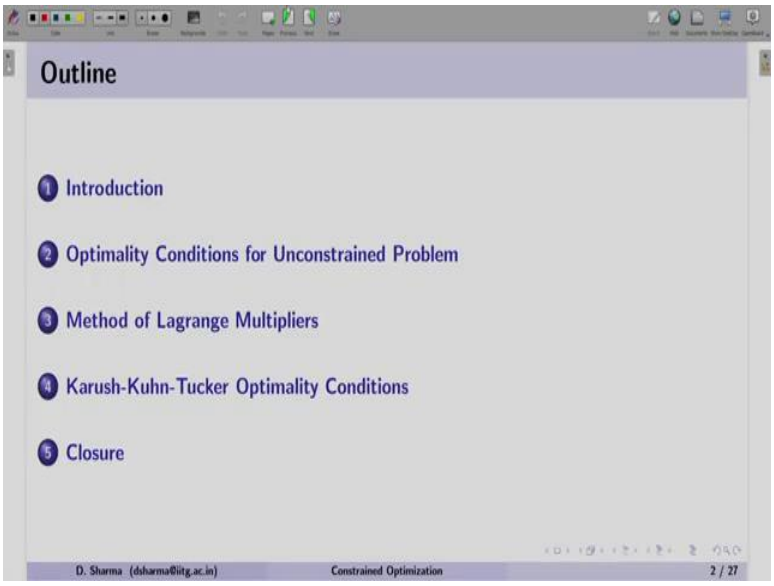
**Evolutionary Computation for Single and Multi-Objective Optimization**
**Dr. Deepak Sharma**
**Department of Mechanical Engineering**
**Indian Institute of Technology, Guwahati**

**Lecture - 12**
**Constrained Optimization: Introduction and Optimality**

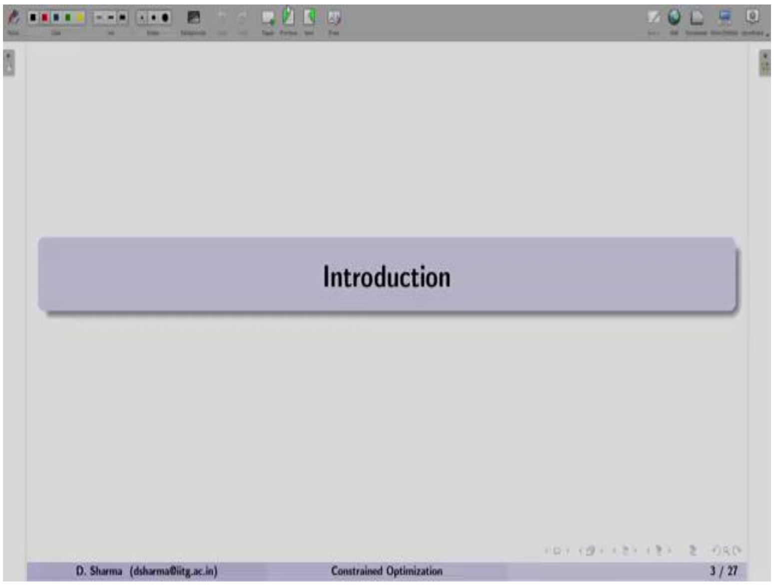
Welcome to the session. In this session, we will be focusing on Constrained Optimization. In this particular session, we will be focusing on introducing the problem and the optimality conditions.

(Refer Slide Time: 00:48)



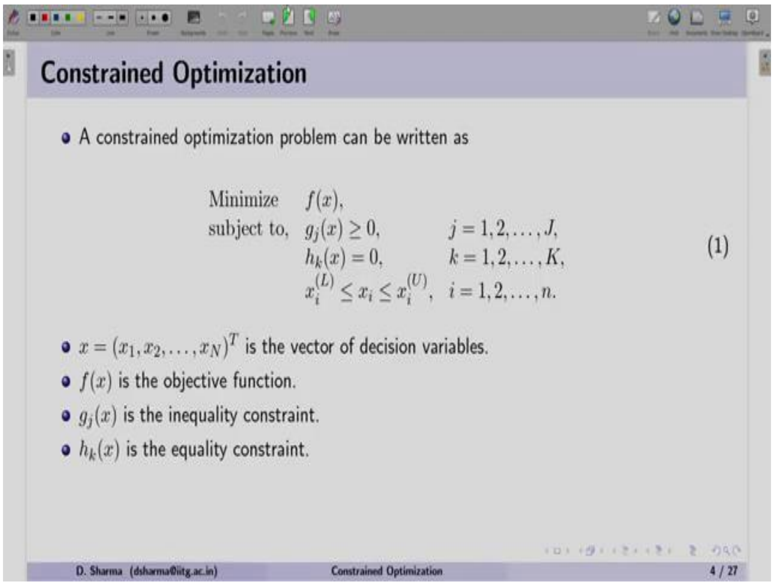

The content of this particular session is we will start with the introduction, thereafter we will discuss the optimality condition for unconstrained optimization followed that we will be discussing the method of multiplier for constrained optimization. This method of multiplier we will take it further, we will understand the KKT condition for constrained optimization and finally, we will conclude this session.

(Refer Slide Time: 01:22)




So, let us begin with the introduction on constrained optimization.

(Refer Slide Time: 01:28)




Now, as we can see here, a constrained optimization problem can be written as minimization of the function f x and this particular problem is subjected to inequality constraint which is represented as g j of x and the total number of such inequality constraint, we can have capital J. Similarly, we can have equality constraint and we can have multiple equality constraint which is represented by capital K.

$$Minimize \ \ f(x),$$

$$subject\ to,\ \ g_j(x) \geq \ 0, \quad j \ = \ 1, 2, \ldots, J,$$

$$h_k(x) \ = \ 0, k \ = \ 1, 2, \ldots, K,$$

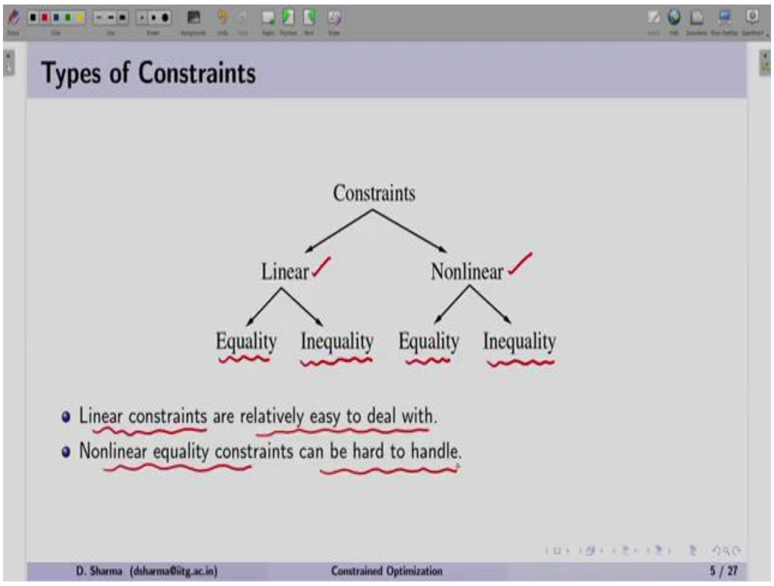
$$x_i^{(L)} \leq \ x_i \leq \ x_i^{(U)}, \quad i = 1, 2, \ldots, n.$$

Apart from that, the problem can have the variable bounds. So, as we can see in equation number 1, we are minimizing a function which is subjected to inequality constraints, equality constraints and the variable bound. Now, here the x which we are writing that is a column vector for us. So, this is a vector for a decision variables and as we have understood that since we are minimizing the function, so, f x is our objective function and g j of x is our inequality constraint and h k of x is the equality constraint.

Here, when we are looking for a solution, as and when a solution satisfying all the constraints including the variable bound, this particular solution is referred as feasible solution. And as an when the any solution which is not satisfying a single constraint or more than one constraint or the variable bound, that solution is referred as infeasible solution.

So, from this discussion, we can find that a one particular constraint or a group of constraints can make some solutions feasible and some solutions infeasible. So, as we have mentioned here, it means that so, basically because of the constraints, these constraints can restrict solutions to be feasible. So, we can divide a set of solutions into feasible and infeasible solutions.

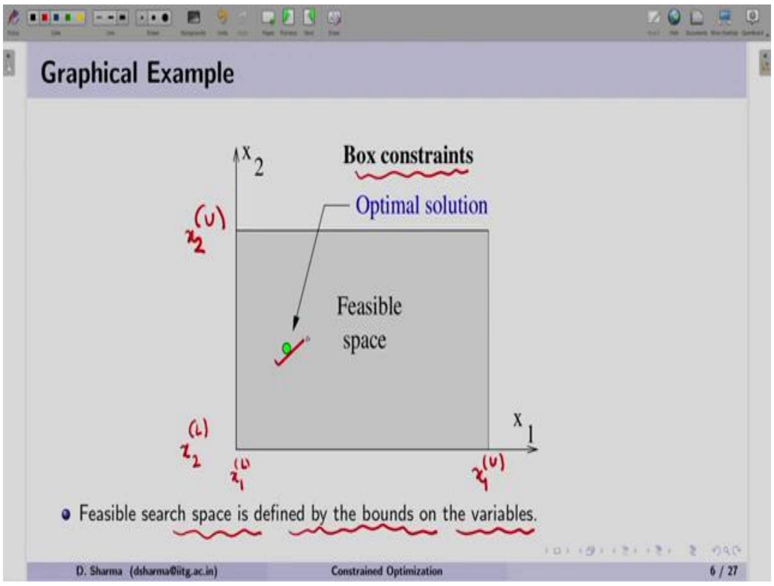
(Refer Slide Time: 03:54)




There are various kind of constraints, now, let us see that. We can have constraint like linear equations, or we can have non-linear equations. Inside these two category, we can have equality constraint as we have shown in the previous slide and we can have inequality constraints. Under non-linear constraints as well, we can have equality and inequality constraints.

So, these are the different types of constraints we can have in an optimization problem. Generally, it is found that the linear constraints are relatively easy to handle with. However, the non-linear constraints can be hard to handle. So, it is a relative term. So, we found that from the various simulations, linear constraints are easy to handle.

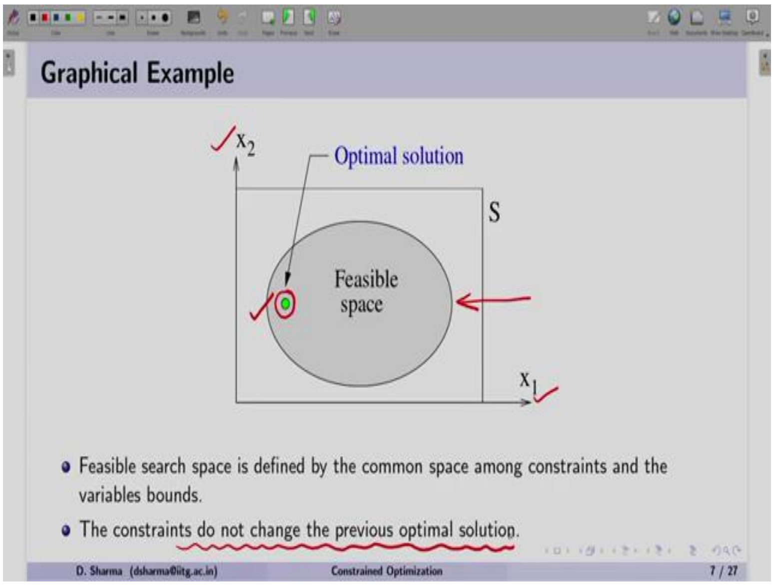
(Refer Slide Time: 04:51)




Now, let us see what kind of examples or constrained optimization methods we can have. So, instead of a methods, we should say constrained optimization problem. So, the first graphical example is shown for box constraint meaning that the feasible search space is defined by the bounds on the variables.

So, as you can see in this picture, we can have this lower bound on x 1, similarly, I can have upper bound on x 1. Similarly, we can have lower bound on x 2 and the upper bound on x 2. So, these bounds are making this particular region feasible and when we run any optimization algorithm, we will get this optimal solution which is represented by the green color.

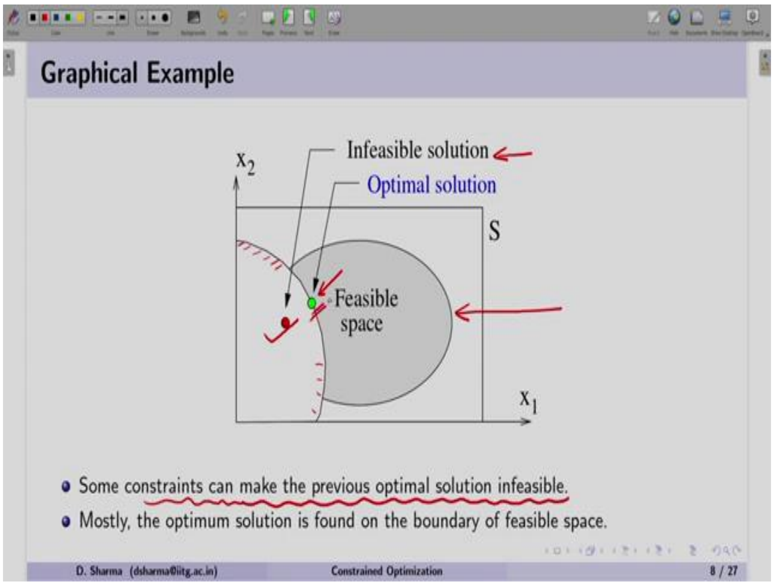
(Refer Slide Time: 05:59)




The another problem is that suppose we have along with the variable bounds which is you can see on x 1 and on x 2, we can have other constraints, that constraints make this particular feasibly space. So, this oval shape is now our feasible space. Now, here what we can see that the feasible, the optimum solution which we had in the previous slide that is the same as the optimum for this scenario meaning that the constraints do not change the previous optimal solution.

So, in that case, we can understand that even if there are constrained into the problem, but the optimum solution remains the same as we have found without those constraints. So, with the help of variable bounds only, we can locate where is the optimum.

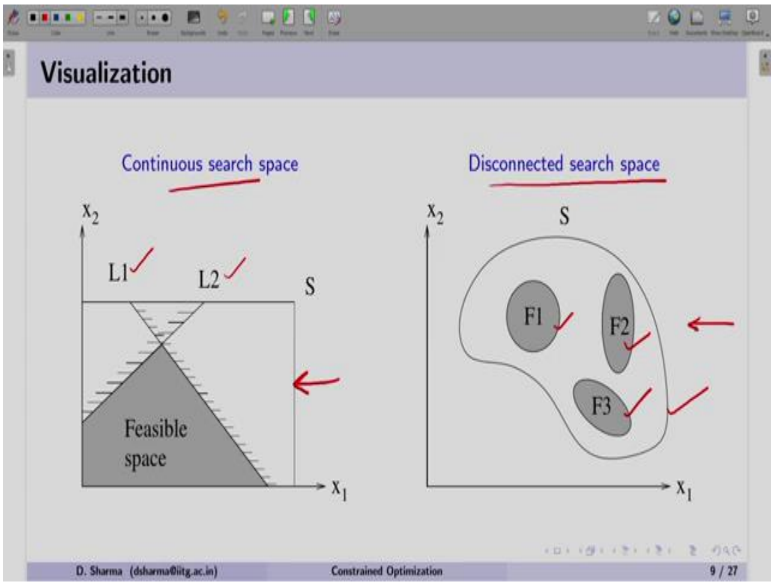
(Refer Slide Time: 07:02)




Look at the third situation which is generally the case that we have one constraint as you can see here. This particular constraint is making our previous optimum solution so, that is the red dot here and this red dot is now infeasible. Meaning that, the constraint that is introduced right now that makes our previous solution infeasible and you can also see that our feasible space is also changed because of this constraint.

So, that is why, some constraint in the problem can make our previous optimum solution infeasible. And in such kind of scenario, most of the time, we found that the optima solution is generally lying on the boundary of this feasible space.

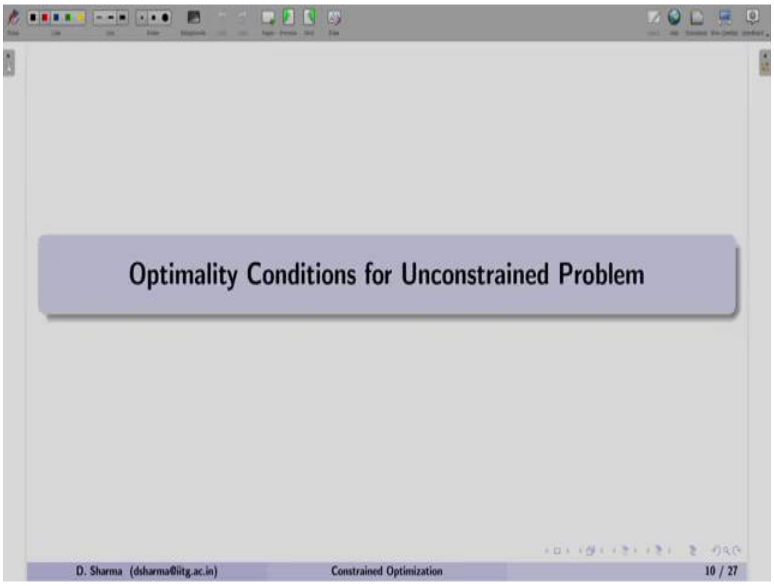
(Refer Slide Time: 08:00)




Let us see what could be the search space we can have it. As we can see here, we can either have a continuous search space or we can have a discontinuous search space. Looking at the figure on the left-hand side, this particular space is made by the constraints as well as the variable bound. So, we have just one feasible space and that is continuous as you can see from this left-hand side picture.

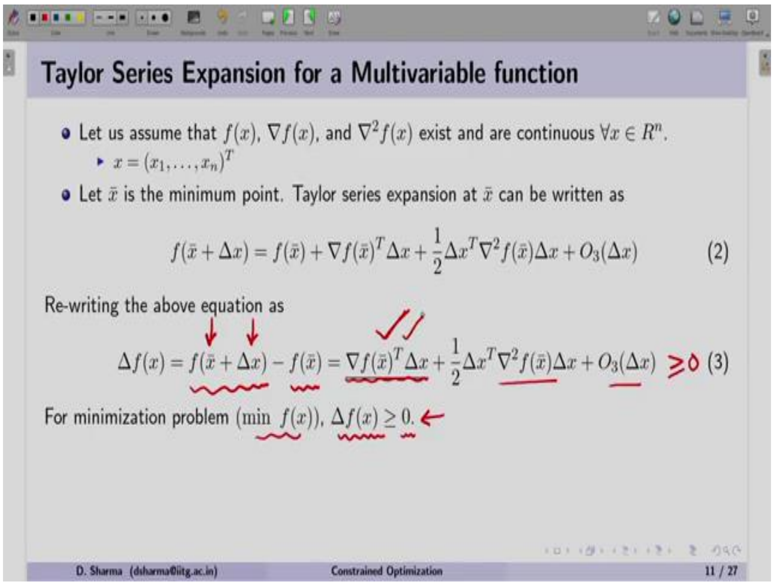
Now, looking on the right-hand side, we have a search space S, inside this we can have disconnected feasible spaces. Now, from these two figures, we can see that when we are solving a problem which is making a continuous search space, those are relatively easy to solve, but if the problem is making our disconnected search space, those problems are difficult to handle.

(Refer Slide Time: 09:03)




Now, since we have understood that what could be the constraints, the various types and those constrains are making some solution feasible, some solution infeasible as well as those constraints can make the search space continuous or discontinuous. Now, let us move towards the optimality condition. We will begin our discussion on the optimality of unconstraint optimization problem.

(Refer Slide Time: 09:35)




Now, if I am saying unconstraint, this means that there is no constraint on the problem, we are either minimizing the problem or maximizing it. For our reference, let us take minimization of a function f x. So, let us look here. So, we will be using the Taylor series expansion for multivariable function. So, the function f of x is a multivariable as you can see the x is a column vector here.

$$f(\bar{x} \;\; + \Delta\, x) \; = \; f(\bar{x}) \, + \nabla\, f(\bar{x}^T) \,\Delta\, x \, + \, \frac{1}{2}\,\Delta\, x^T\, \nabla^2\, f(\bar{x})\,\Delta\, x \;\; + \;\, O_3(\Delta\, x)$$

$$\Delta\, f(x) = \; f(\bar{x} \;\; + \Delta\, x\,) \; - \; f(\bar{x}) = \; \nabla\, f(\bar{x})^T\,\Delta\, x \; + \, \frac{1}{2}\,\Delta\, x^T\, \nabla^2\, f(\bar{x})\Delta\, x \; + \; O_3\,(\Delta x)$$

So, here let us assume that the function, the gradient of the function and this is gradient of the gradient of the function, they exist and as well as continuous for all value of x that is belonging to the real space of n. Let us assume that we have x bar which is the minimum point.

We are going to use the Taylor series expansion at a current point say x bar and we can write this particular equation as f of x bar plus Δx, we are going to expand using the Taylor series expansion so, we get f of x bar plus gradient of the function at x bar transpose Δx. So, this transpose means that we are performing a dot product here. The third term is half of Δx transpose gradient of the gradient of the function at x bar and Δx plus the other terms. Now, here you know that this Δx is also a vector.

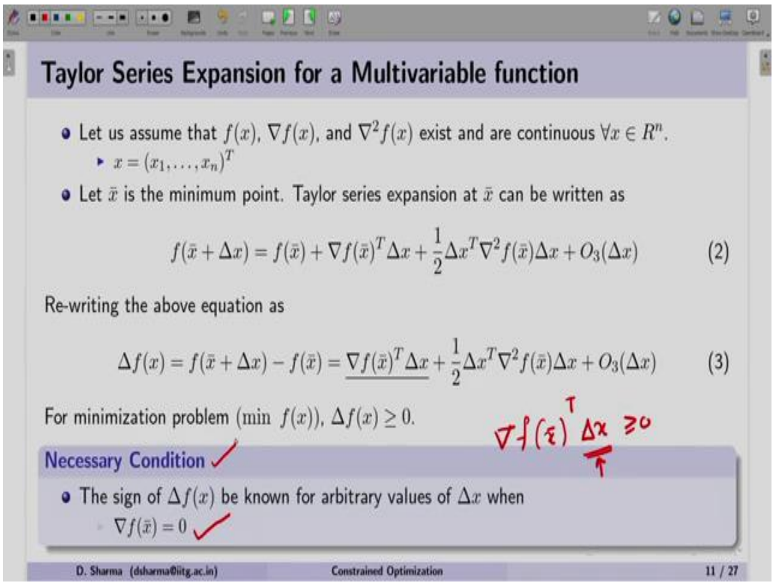
Now, we if we are going to rewrite this equation meaning that the equation number 2, the this particular term on the right-hand side we are taking into the left-hand side. Looking on looking at the equation number 3, we have Δ of f x. This is f of x plus Δx minus f of x bar and then, we have the rest of the terms which is as per the equation number 2. So, we have the first term and then the second term and we have the rest of the terms here.

Now, since we have assume, we are using a minimization problem here. So, we know that for a minimization problem, this Δf of x should be greater than 0. Why? Because, we know that this x bar is the minimum point so, whether the Δx value is positive or a negative, the function value of x bar plus Δx will be greater than the function value at x bar. So, this says that the Δof f of x should be greater than 0.

If we are including this particular condition here so, in equation number 3 so, this right-hand side should be greater than and equals to 0. While doing so, we have this particular first term which is underlined here, the second term and the third term. Now, since we want the right-hand side of the equation should be greater than equals to 0, so, we know that the dominating term in this particular side is gradient of the function at x bar transpose Δx should be greater than 0.

If this dominating term will be greater than or equals to 0, then, rest of the terms will be quite smaller as compare to the first term.

(Refer Slide Time: 13:28)

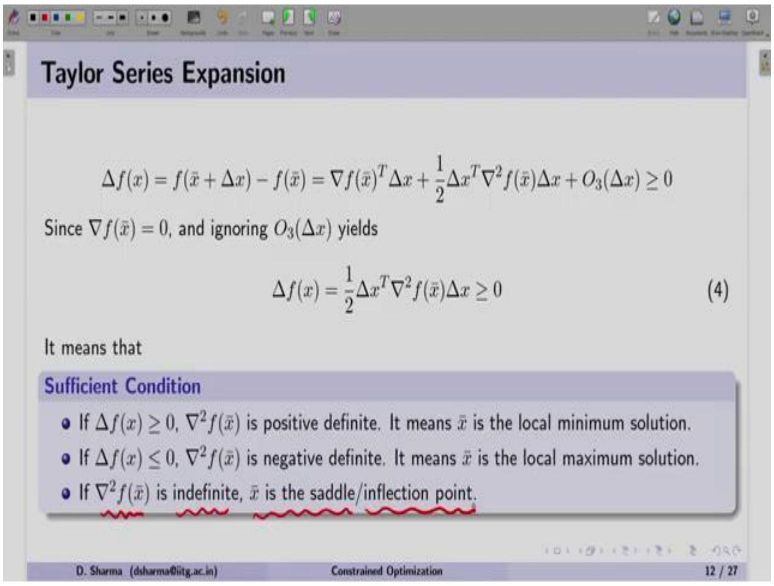



So, in this case, what we will see that if we want the gradient of the function at x bar transpose Δx, if we want this should be greater than the 0 for any arbitrary value of x, this means if Δx can take positive value as well as negative value, this condition can only true when we say the gradient of the function at x bar is equals to 0.

$$\Delta f(x) = f(\bar{x} + \Delta x) - f(\bar{x}) = \nabla f(\bar{x})^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(\bar{x}) \Delta x + O_3(\Delta x)$$

So, this particular condition is known as the necessary condition for the unconstrained optimization problem. So, let us take move a take further and we have to find the other conditions.

(Refer Slide Time: 14:21)




Now, here, we have written this expression as Δof f of x which is f of x bar plus Δx and the function value at x bar so, difference of that is the right-hand side which we have found in the previous slide and everything should be greater than 0. Since for an arbitrary value of a Δx, gradient of the function at x bar should be 0.

$$\Delta f(x) = f(\bar{x} + \Delta x) - f(\bar{x}) = \nabla f(\bar{x})^T \Delta x + \frac{1}{2}\Delta x^T \nabla^2 f(\bar{x})\Delta x + O_3(\Delta x) \geq 0$$

$$\Delta f(x) = \frac{1}{2} \Delta x^T \nabla^2 f(\bar{x}) \Delta x \geq 0$$

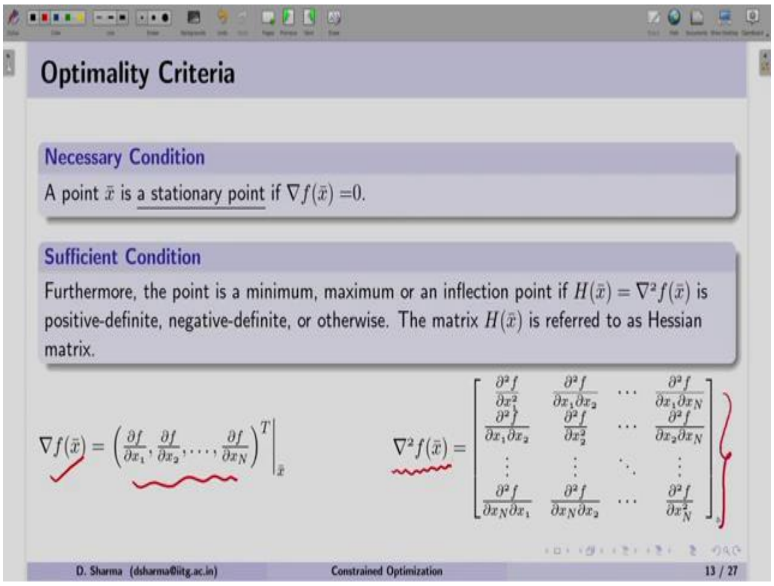
And let us ignore the other term, we will get the equation number 4 which says that the Δof f of x is equals to the second term which is half Δx transpose gradient of the; gradient of the function at x bar and Δx and that should be 0. Now, this particular condition is important for us. Why?

Because, this set of a condition will give us another condition meaning that we are going to get the sufficient condition for us if our Δ of f of x is greater than equals to 0 that is what we are doing right now in an equation number 4. In this case, this gradient of the gradient of the function at x bar is a positive definite. If this is positive definite meaning that x bar is the local minimum solution.

Similarly, if we use the Taylor series expansion for the maximization problem, what we will see, here is that if Δ of f of x is smaller than 0 and this gradient of the gradient of the function at x bar is negative definite, then we can say that x bar is the local maximum solution. In case, this gradient of the; gradient of the function at x bar is indefinite meaning it is nor positive definite and not the negative definite in that case, we can say that x is the saddle point or the inflection point.

Now, the important question is that when we get a point say x bar, this x bar is going to be a stationary point. Since, we do not know this is stationary point is a minimum, maximum or inflection point, at that particular case, we need sufficient condition as described here. So, let us see all these conditions in a together.

(Refer Slide Time: 17:13)

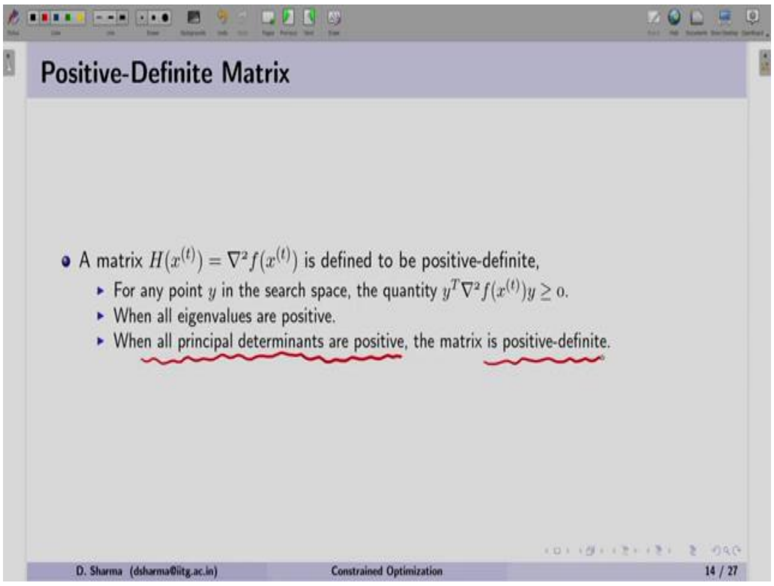



So, here necessary condition says that a point x bar is a stationary point when the gradient of function at that point is 0. So, that is the necessary condition we have to check it. Thereafter, the sufficient condition says that the point which is x bar right now is minimum or maximum or an inflection point if now, we are introducing this letter H which is being referred as Hessian of the matrix.

So, this Hessian matrix is positive definite, negative definite or it can be otherwise. So, based on the Hessian matrix, we can find out whether the stationary point is minimum, maximum or an inflection point. As we know, if we have the function which is the scalar function, when we find the gradient of the function, it is going to be a vector for us.

For the same scalar function, when we are going to find gradient of the gradient of the function meaning the matrix so, as you can see a symmetric matrix will come into the picture.

(Refer Slide Time: 18:31)




$$H\left(x^{(t)}\right) = \nabla^2 f\left(x^{(t)}\right)$$

$$y^T \nabla^2 f\left(x^{(t)}\right) y \geq 0$$

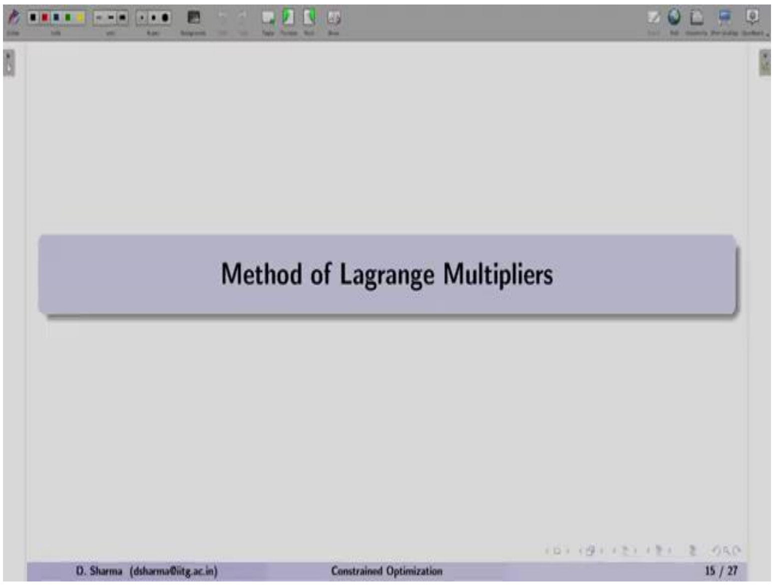
Now, the question is that when we are going to get a point says stationary point x bar, then how would I know that the given matrix is positive definite or negative definite or it is indefinite? So, in order to know this, we have certain conditions. So, the matrice matrix H at a point say x t is defined to be a positive definite so, we can check a condition. So, we have three conditions.

So, first condition is let us take any point y in the search a space and if you find this particular quantity as y transpose Hessian y greater than 0, then we can say that the matrix is a positive definite matrix. Similarly, the second condition is since we already have a Hessian matrix, we can find the eigenvalues. If all the eigenvalues are positive, we can say the matrix is a positive definite matrix.

Now, what could be the third condition? This is when we are going to find the principal determinants and all of them are positive, then we can say the matrix is a positive definite matrix. So, we have three ways to identify whether the matrix is positive definite or not. Now, how I can find the negative definite?

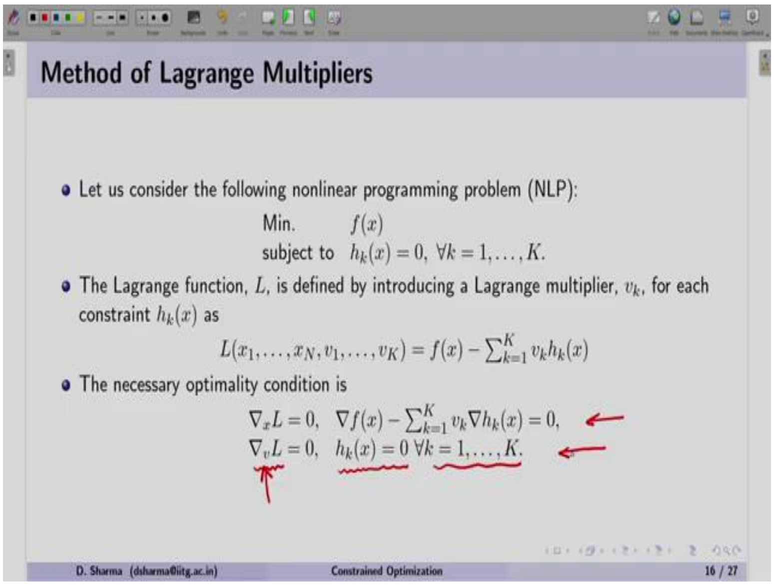
As we know if we multiply minus 1 to the Hessian and if this minus times of the Hessian is coming out to be positive definite, then we can say that it is a negative definite. If it is not positive, not negative, then we can say it is an indefinite matrix.

(Refer Slide Time: 20:23)




As of now, we have gone through the optimality condition for unconstraint optimize problem. Now, let us move towards the constrained optimization problem.

(Refer Slide Time: 20:48)



$$Min.\ f(x)$$

$$subject\ to\ \ h_k(x) = \ 0, \sim\forall\ k = 1, \dots, K.$$

Now, to start with let us understand the method which is known as method of multiplier. So, we will take first example. Here, let us consider the following non-linear programming problem so, NLP. In this case, we want to minimize the function effects so, it is again a multivariable function and this particular objective function is subjected to equality constraint which is defined as h k of x equals to 0. So, we are taking a simple case right now and we can have multiple search equality constraints.

$$L(x_1, \dots, x_N, v_1, \dots, v_K) \ = \ f(x) \ - \sum_{k=1}^{K} v_k\, h_k(x)$$

$$\nabla_x\, L \ = \ 0, \qquad \nabla\, f(x) \ - \sum_{k=1}^{K} v_k\, \nabla\, h_k(x) \ \ = \ 0,$$

$$\nabla_v\, L \ = \ 0, \qquad h_k(x) = \ 0 \sim\forall\ k \ = \ 1, \dots, K.$$

So, here, we are introducing a function called Lagrangian function. In this Lagrange function, we introduce some Lagrange multiplier say v k for each constraint h k. So, meaning that if we have say five constraints, then we are going to have five Lagrangian multiplier. How we can write this Lagrangian function? So, let us see here.

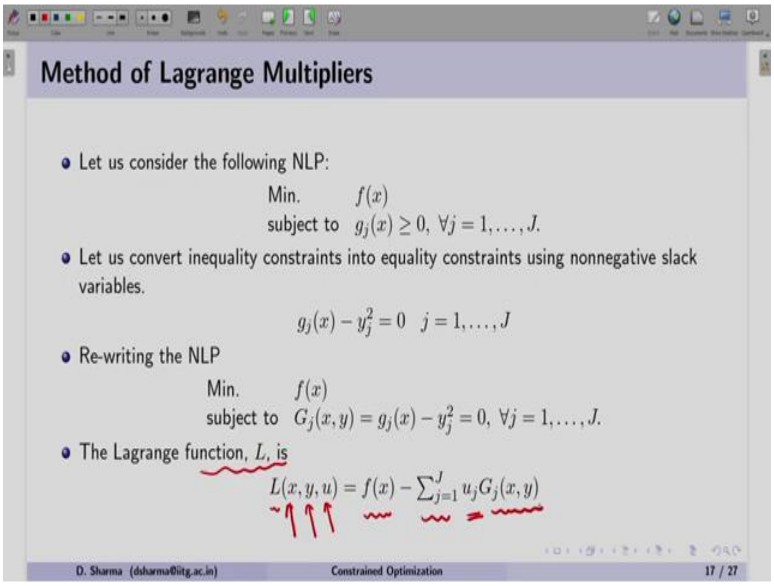
So, the Lagrangian function is made of the variable x 1 to x N and the Lagrangian multiplier from v 1 to v K. So, we have K number of constraint as of now. So, we have introduce that many Lagrangian multiplier. On the right-hand side of the equation, we have the function minus times of now we are making a summation. Now, this summation is that we are multiplying the Lagrangian multiplier with equality constraint.

What we can see from this particular equation? That we started with the constraint problem where we want to minimize a function subjected to various equality constraints. With the help of Lagrangian method or the Lagrangian multiplier, we have converted this constraint problem into the unconstraint problem. So, the question is can we use the optimality conditions which we have understood for unconstrained optimization problem? So, the answer is yes.

So, in this case, the necessary optimality condition for the Lagrangian function is so, let us find the gradient of this Lagrangian function with respect to x and we can make it 0. So, this is our; this is our necessary optimality condition. So, in this case, when we are going to find the gradient so, we can get gradient of the function minus times of the summation v k and the gradient of the constraint equal to the 0. So, this is the condition which we know as a necessary optimality condition.

Apart from x, we have other variables, which is v k. So, here, when we are finding the gradient of the Lagrangian function with respect to the Lagrangian multipliers v, we will get h k of x equals to 0 for all K's. This means that the necessary optimality conditions for a Lagrangian function will be that is based on the first equation and the second set of equations. So, both of them will be giving us the optimality condition for the given non-linear programming problem.

(Refer Slide Time: 24:14)




As of now, we have just taken the equality constraint. Now, let us take the inequality constraint into our problem and we will see how this Lagrangian or the method of multiplier can be used. Given here is again we will consider an NLP problem here. We are going to minimize a function and this function is subjected to the inequality constraints.

$$Min. \quad f(x)$$

$$subject\ to \quad g_j(x) \geq \ 0, \sim\forall\ j = 1, \dots, J.$$

$$g_j(x) - \ y_j^2 \ = \ 0 \ \ j = 1, \dots, J$$

So, this particular sign greater than equals to 0 that makes the inequality constraint and we can have various kinds of constraints as well.

So, here let us convert this inequality constraint into the equality constraint. So, for this conversion, we need slack variables. So, we are going to use nonnegative slack variables to convert them. The concept of slack variable, we can also understand from the linear programming problem in that case, the inequality constraints are generally changed into equality constraint with the help of slack variable.

$$Min. \quad f(x)$$

$$subject\ to \quad G_j(x, y) \ = \ g_j(x) - \ y_j^2 \ = \ 0, \sim\forall\ j = 1, \dots, J$$

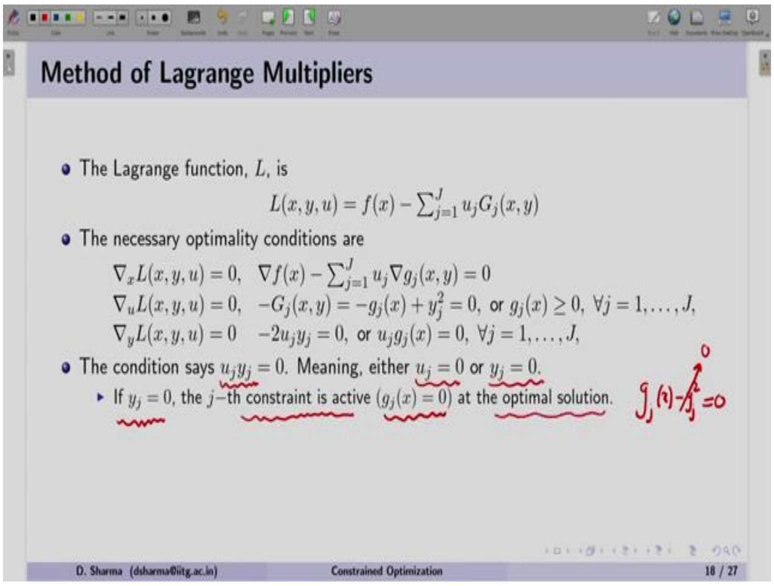
Now, looking at the equation now, g j of x minus y j square greater than 0 and all these j so, for every constraint we are putting this slack variable. What is important point here being this y, y j can take any value, it can be positive or a negative, but y j square will always take a nonnegative value why because the square makes our negative term positive and positive remains positive.

So, in this case, we are subtracting this slack variable y j square in this particular equation so that the inequality constraint given here can we converted into the equality constraint. So, let us rewrite this NLP now. In this NLP, we are going to minimize a function say f x and now, we are writing as capital G j and this is the function of x and y and we know x is the set of variables for the given problem and y, a y is the set of slack variable.

So, the capital G j of x, y is equal to g j of x minus y j square equals to 0. So, the same equation which we have written on the top, we are writing here. Since, the constraint is converted into the equality constraint, we can use the Lagrangian function. Here, the Lagrangian function is as you can see, it is made of variable set of x, set of y and set of u.

On the right-hand side, we have function f of x minus we have a summation and then, we have a Lagrangian multiplier so, u j is now Lagrangian multiplier for the inequality constraint and we are writing this G j which is the function of x and a y. So, we are going to have x, y and u.

(Refer Slide Time: 27:43)




$$L(x,y,u) \;=\; f(x) \;-\sum_{j=1}^{J} u_j\, G_j(x,y)$$

$$\nabla_x\, L(x,y,u) \;=\; 0, \qquad \nabla f(x) \;-\sum_{j=1}^{J} u_j\, \nabla\, g_j(x,y) \;=\; 0$$

$$\nabla_u\, L(x,y,u) = \;0, \qquad -G_j(x,y) = \;-g_j(x) + \;y_j^2 \;=\; 0, g_j(x) \geq\; 0, \forall\, j \;=\; 1,\ldots,J,$$

$$\nabla_y\, L(x,y,u) = \;0, -2u_j\, y_j \;=\; 0, \qquad u_j g_j(x) = \;0, \forall\, j \;=\; 1,\ldots,J,$$

For this given Lagrangian function, we can find the optimality condition. So, we are going to find the gradient of the function first with respect to x, thereafter with respect to y, thereafter with respect to u and then, we will get set of optimality condition. Now, from this particular equation, we can see that a constraint problem is converted into an unconstraint problem.

So, the necessary optimality conditions are we are finding the gradient of the function L with respect to x and making it equals to 0 that gives me the first set of equation as it is mentioned here which is gradient of the function minus the summation over all the constraint u j gradient of g j x, y. Thereafter, we can have the gradient of the L with respect to u and we are making it equals to 0.

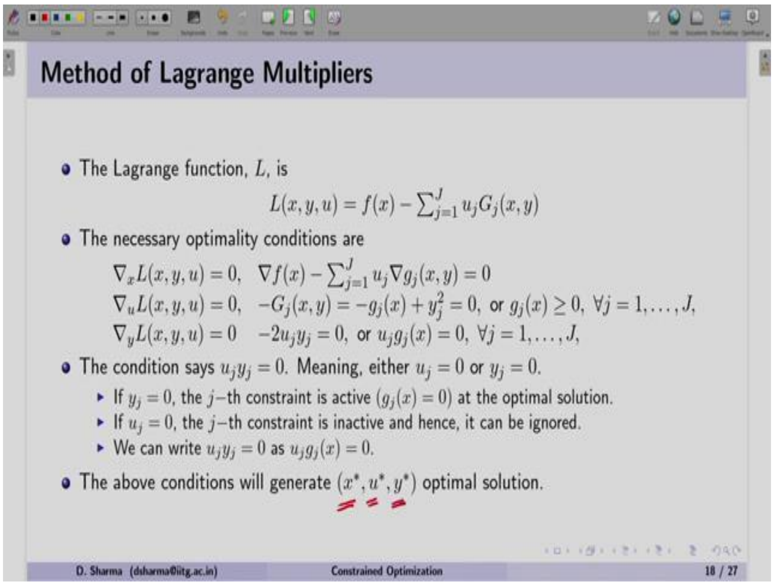
Now, since we know the u as you can see on the Lagrangian function, u is available with the constraint only so that is why we are going to get minus G j of x, y or minus g j of x plus y j square equals to 0. Now, looking at this condition, we know that this y j square will either take positive value or it will be non-negative. Either it can take 0 value or it will be a positive value meaning that if you take the whole term on the right hand side, what we can see that we can write this particular condition as g j of x greater than 0 for all the constraint.

So, this is the constraint for our problem. The third optimality condition we can get when we will be finding the gradient of the function L with respect to this slack variable y and then, we are making it equals to 0. Now, this case, the condition which we will get is minus 2 u j y j equals to 0 or we can write u j g j is equals to 0 and this is for all constraints.

Now, here, the important point is that we have a condition called y j; u j y j equals to 0 meaning that either u j equals to 0 or y j equals to equals to 0. So, let us take the case one by one. If y j equals to 0, meaning that, the slack variable is 0. So, if we rewrite our constraint, this says that g j of x minus y j square is equals to 0.

So, this current condition says that y j is 0 meaning that g j of x is going to be 0 as and when at any point, any of the constraint is equals to 0, those constraints are referred to as active constraints. So, if y j equals to 0, the j-th constraint is active at the optimal point.

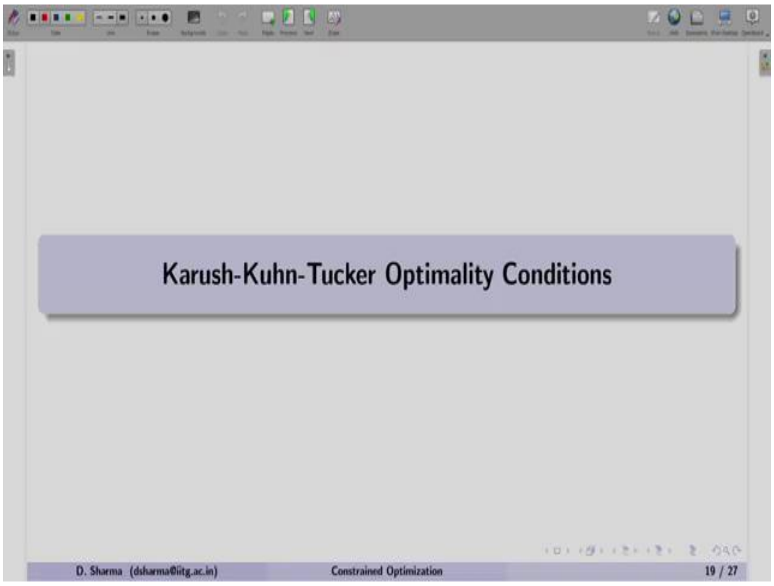
(Refer Slide Time: 31:16)




If u j equals to 0 meaning that the j-th constraint is inactive and since it is inactive, it can be ignored because this inactive constraint is not going to change or affect our optimal solution. So, the overall condition y j u j equals to 0 can be written as u j g j of x equals to 0. So, as you can see y j equals to 0 means g j of x equals to 0 so that is why we are writing; rewriting the equation as we can get it in the terms of u j g j equals to 0.

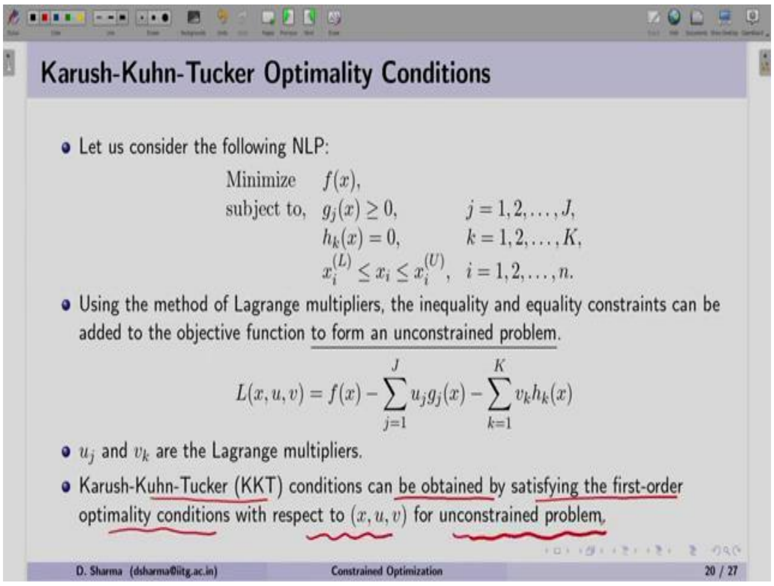
Now, the above condition will generate the x star which we are generally refer this x star for the optimal point or the optimal values. So, we have x star, u star, y star and that is the optimal solution for the given problem.

(Refer Slide Time: 32:19)




Now, these are the method of multiplier that is actually helpful to find out the KKT condition. The KKT conditions are referred as Karush-Kuhn-Tucker optimality condition for constrained optimization. So, let us look the KKT condition now.

(Refer Slide Time: 32:40)




$$Minimize \quad f(x),$$

$$subject\ to, \qquad g_j(x) \geq 0, \qquad j = 1, 2, \ldots, J,$$

$$h_k(x) = 0, \qquad k = 1, 2, \ldots, K,$$

$$x_i^{(L)} \leq x_i \leq x_i^{(U)}, \qquad i = 1, 2, \ldots, n.$$

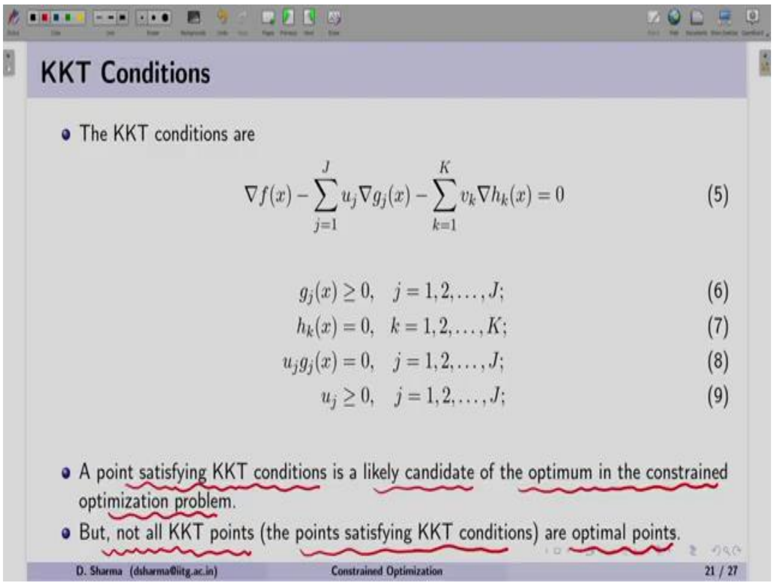
$$L(x, u, v) = f(x) - \sum_{j=1}^{J} u_j g_j(x) - \sum_{k=1}^{K} v_k\, h_k(x)$$

Now, here, let us consider the NLP as given on the top. Here, we are minimizing the function f of x, we have inequality constraints, we have equality constraints as well as we have variable bounds. So, this is the original problem which we had it earlier. Now, since we have to find the optimality condition so, we know we are going to write the Lagrangian function so that we can convert this constrained optimization problem into unconstrained optimization problem.

So, let us see that. So, using the method of Lagrangian multiplier, the inequality constraints can be added to the objective function to make the problem unconstrained. So, here, the Lagrangian multiplier now, it is a function of x, u and v and we can have a right-hand side f of x minus times so, this summation for the inequality constraint minus the summation for the equality constraint and u j's, this is the Lagrangian multiplier for g j and v k that is the Lagrangian multiplier for equality constraints.

So, KKT condition or Karush-Kuhn-Tucker condition can be obtained by satisfying the first order optimality condition with respect to x, u and v for unconstrained problem. So, this practice, we have already done when we have shown the optimality condition for the equality constraint and then, thereafter for the inequality constraint. So, if you follow the same procedure, we are going to get the KKT conditions. So, what are those conditions? Let us see that.

(Refer Slide Time: 34:41)




$$\nabla f(x) - \sum_{j=1}^{J} u_j \nabla g_j(x) - \sum_{k=1}^{K} v_k \nabla h_k(x) = 0$$

$$g_j(x) \geq 0, \quad j = 1, 2, \ldots, J;$$

$$h_k(x) = 0, \quad k = 1, 2, \ldots, K;$$

$$u_j g_j(x) = 0, \quad j = 1, 2, \ldots, J;$$

$$u_j \geq 0, \quad j = 1, 2, \ldots, J$$

So, the KKT conditions are we have the gradient of the function minus this is the summation u j gradient of the g j minus then, we have a summation v k h k of x equals to; h k of x equal to 0. So, this is you know we get it corresponding to the optimality condition when we are going to find the gradient with respect to x.

Similarly, if we do the same thing, then the equation number 6 is our the original constraint, inequality constraint. Equation 7 is the original equality constraint. Equation 8, we get as u j g j equals to 0 so, this condition also we got it and finally, that all the Lagrangian multiplier corresponding to the inequality constraints all of them either should be greater than 0 or it can be equal to 0.

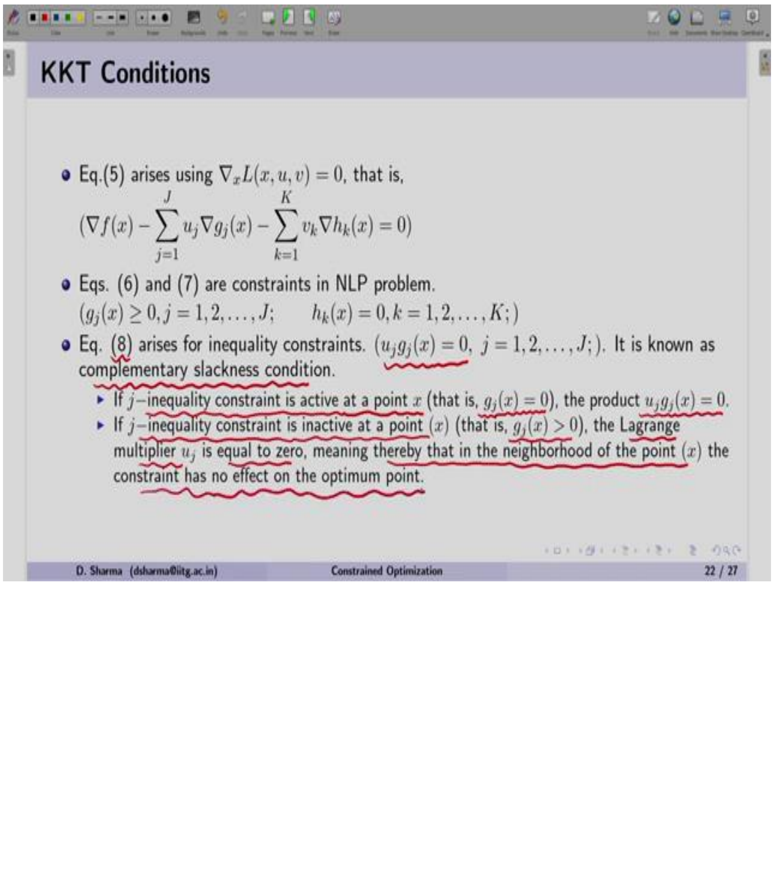
For the KKT a condition, there are certain situations. First is if a point satisfying the KKT conditions is a likely candidate of the optimum in the constrained optimization problem. So, this statement says that if I get stationary points x bar and if I find the KKT conditions and suppose all the conditions are satisfy, then this x bar is likely candidate.

A still we cannot say that it is optimum or not, but it is going to be a likely candidate for the optimum solution. However, we can use this KKT condition to say that that if the point is not a KKT point, then it can never be an optimum solution for the given constrained optimization problem.

In the bottom, we can see that not all KKT points, the points satisfying the KKT conditions are optimal. So, let us take a case that as since EC techniques they work on multiple solutions so, in that case, you will get many feasible solutions. When you get those solution satisfy the KKT condition, we may find that there will be many solutions which are satisfying the KKT conditions.

So, here the statement says that all the solutions are the likely candidate of optima. However, it is not necessary that all of the solutions are optimum. So, the optimum point will definitely be a KKT point, but all point, all KKT point need not be an optimal solution for our given problem. So, let us understand those equations one by one.

(Refer Slide Time: 37:47)



$$\nabla_x L(x, u, v) = 0, \text{that is,}$$

$$\left(\nabla f(x) - \sum_{j=1}^{J} u_j \nabla g_j(x) - \sum_{k=1}^{K} v_k \nabla h_k(x) = 0\right)$$

$$\left(g_j(x) \geq 0, \quad j = 1, 2, \dots, J; \;\; h_k(x) = 0, k = 1, 2, \dots, K;\right)$$

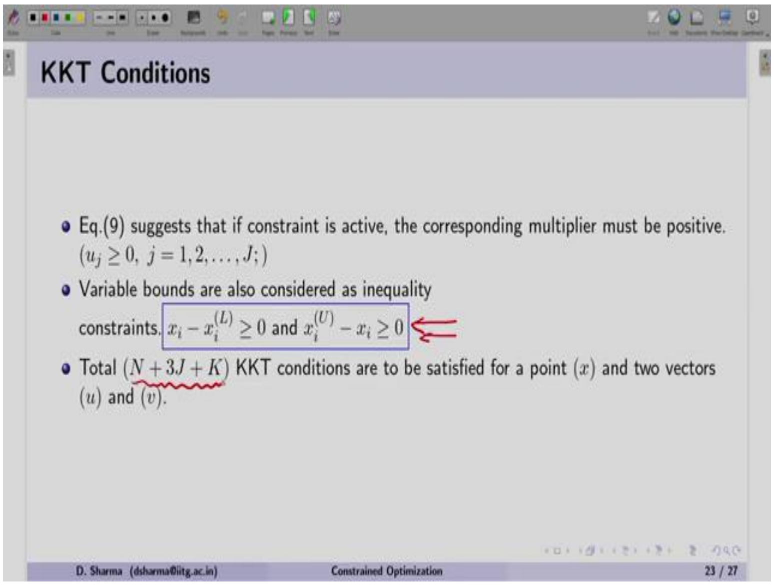
$$(u_j \, g_j(x) = 0, \quad j = 1, 2, \dots, J;)$$

Equation 5 that arises when we are finding the gradient of the Lagrange function with respect to x make it equals to 0 so, this will give me the optimality condition. Thereafter, as we have understood the equation 6 and 7, these are our original constraint of the non-linear programming. So, that is why, g j of x is greater ah; g j of x are greater than equals to 0 and h k of x equal to 0.

Thereafter, we get a another condition in equation number 8 that is u j g j equals to 0. Now, these conditions generally are referred to as complementary slackness condition. As we have discussed earlier, the if the j-th inequality constraint is active at a point meaning that g j of x is equals to 0, the product u j g j is 0.

Second is if inequality constraint is inactive at a point, then g j of x is greater than 0. So, at that particular point, the Lagrangian multiplier u j is equal to 0 meaning there by that in the neighborhood of the point x, the constraint has no effect on the optimum point, and this is what we discussed earlier. If the constraint at a given point is not active meaning that this constraint is not going to affect our optimum solution. So, in this case, we should make our Lagrangian multiplier u j should be equals to 0.

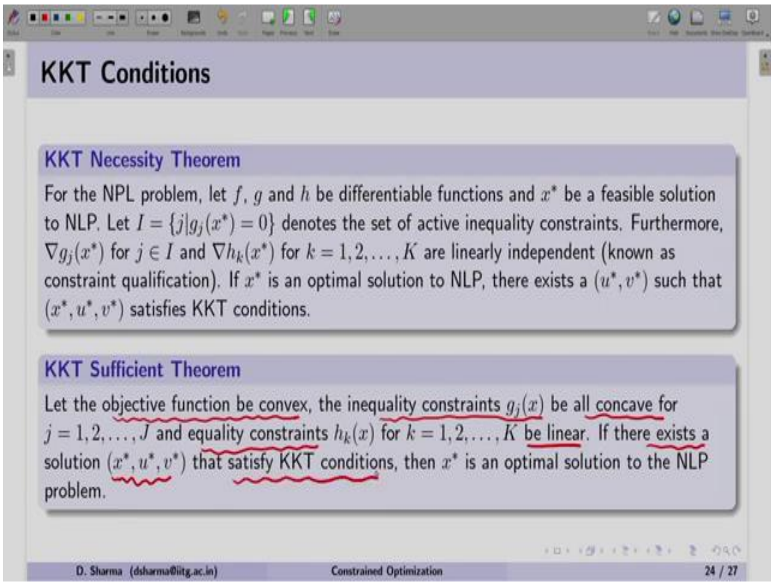
(Refer Slide Time: 39:48)




The another important condition that is the last condition. We have that in equation 9, all these u j should be equals to 0. So, for every Lagrangian multiplier corresponding to the inequality constraints should be greater than or equals to 0. Now, here, the important point is that when we are going to solve this, the bounds on the variable should also be taken as inequality constraint as given here.

So, if you are going to consider the so, we are going to have N plus 3J plus K so, total number of KKT condition will be that many. So, if you have to say that a point is satisfying the KKT condition, it has to satisfy many equations as given here.

(Refer Slide Time: 40:46)




Now, using this KKT, condition we have two theorems. First is called KKT necessary theorem. So, let us look at this particular theorem first. For the NLP, the function f so, this is our objective function, g means our constraint, inequality constraint and h is our equality constraint all of them are differentiable function and x bar is a feasible solution to the given NLP.

So, let us find out what are the solutions which are active. So, here let I is equals to all those constraints which are active at a point say x star. So, that denotes that the set of active constraints are copied to I. Furthermore, when we are finding the gradient of the inequality constraint which are active and the gradient of the equality constraints and they are linearly independent so, this condition has to be satisfied.

In that case, if x star is an optimal solution to NLP, then there exists u star and v star such that x star, u star, v star satisfies KKT condition. So, meaning that if we are going to have active constraints, active inequality constraints and then, we have equality constraints, the functions are differentiable, these gradient of the constraints are linearly dependent.

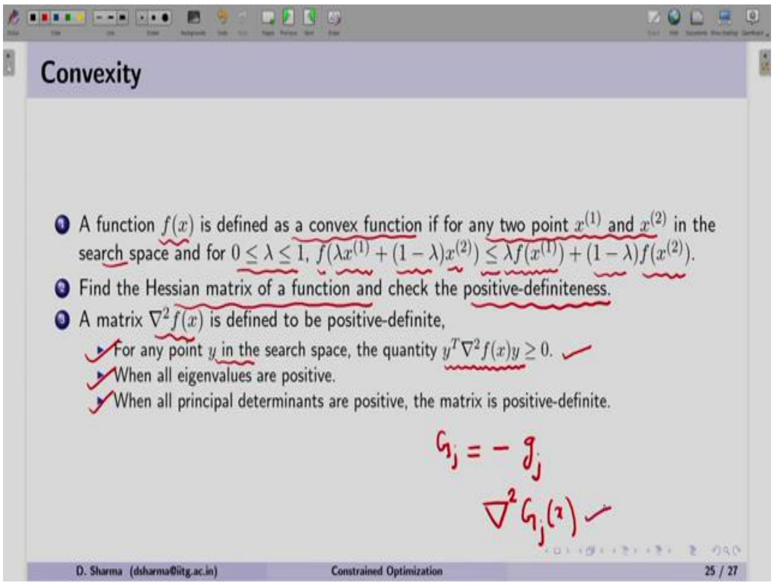
So, for a given points x star, there will be other set of point u star and v star so, we know u and v are our Lagrangian multiplier for equality and inequality constraints so, those point will be satisfying all the KKT condition. Based on that, here as we know, the point which is satisfying KKT condition is a likely candidate for the optimum solution.

Then, in which solution this KKT point can become the optimal point? So, in that case, we have another condition called KKT sufficient theorem. Let us look into that. The KKT sufficient theorem says that let the objective function be convex. How we can find convex? We can find the Hessian of this function and if it is positive definite, we can say it is a convex we will discuss more on a on it.

So, let us assume that the objective function be convex, inequality constraints g j of x be all concave and equality constraints are linear. If there exists a solution say x star, u star, v star satisfying the KKT condition, we can say that x star is an optimum solution to the NLP problem. With the help of necessity theorem, we can find a set of x star, u star and v star that are satisfying all KKT conditions.

But when we are solving a problem in which the objective function is convex, the inequality constraints are concave and equality constraints are linear, then we can say that x star, u star, v star are the optimum solution for the given problem.

(Refer Slide Time: 44:51)




Now, the question is how we can find a function is convex or a concave? So, here we will just have an introduction of convexity. So, the function f of x is defined as a convex function if for any two point say x 1 and a x 2 in the search space and the lambda is lying between say 0 and a 1 and this condition satisfies which says that function value at lambda x 1 plus 1 minus lambda times of x 2 is a smaller than equal to lambda times of f of x 1 plus 1 minus lambda times of f of x 2.

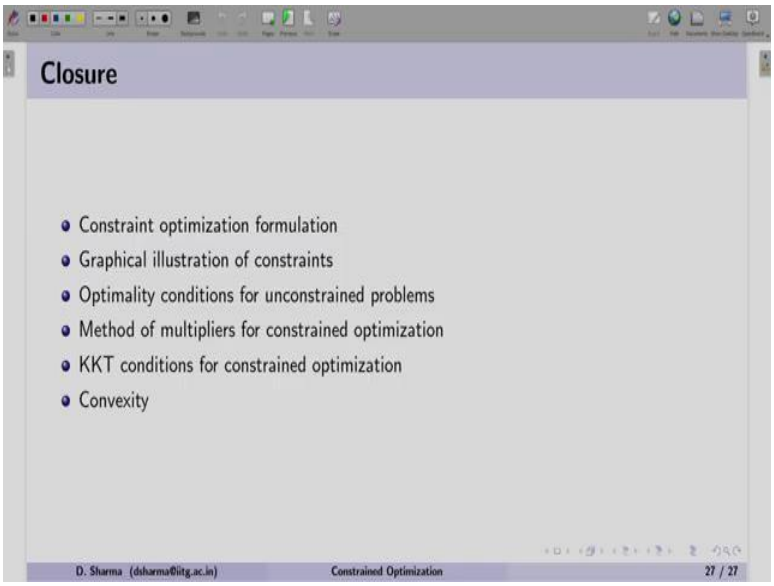
So, this particular relation is valid for single variable function as well as for multi-variable function. Similarly, we can find the Hessian of the function and we can check whether this Hessian matrix is positive definite or a not. So, as we have understood earlier that if we find the Hessian, in that case, a the first condition says that we are going to, we have any points say y in the search space

If the condition given here is satisfied that says that y transpose Hessian y greater than 0, then definitely the matrix is positive definite. Second condition is all eigenvalues are positive, then the matrix is positive definite. Similarly, when all principal determinants are positive, then the matrix is positive definite. So, in while satisfying any of these condition, we can say the matrix is positive definite eventually, we can say the function is convex.

Then, what will happen to the concave function? So, here, we can do simple simplification that we will take say G j of x as minus times of g j of x, we can find the gradient of the gradient of the function of G j of x and if this is coming out to the positive definite, then we can say the constraint is concave. So, that will help us to identify the function is concave convex or a concave.

So, with this introduction that we have a understood the optimality condition for unconstrained optimization as well as the optimality condition for the constrained optimization mainly those condition are referred to as KKT condition, we have come to the closer of the session.

(Refer Slide Time: 47:41)



So, in this particular session, we have gone through the first understand the constrained optimization formulation in which the objective function was subjected to inequality constraint and the equality constraint. Thereafter, we have understood the constrained optimization problem with the help of graphical example where we have continuous search space, we can have discontinuous search space.

Thereafter, we have gone through the optimality condition for unconstrained problem. Now, the unconstrained problem we have going to find a stationary point so, we have necessary optimality condition, we have sufficient optimality condition. Thereafter, for constrained optimization, we have gone through the method of multipliers.

So, these method of multiplier will help us to find out various conditions. So, finding the gradient of the Lagrangian function. So, what we understood from this method that a constrained problem can be converted into the unconstrained problem and the optimality condition of unconstrained problem, we can use it here to find whether the point is optimal or not.

Thereafter, we are gone through the KKT condition for constrained optimization and at the last, we discuss the convexity. Since, it will be helpful to find out whether the function is convex or a not. With this background of the optimality condition on constrained and unconstrained problem, I conclude this session.

Thank you.