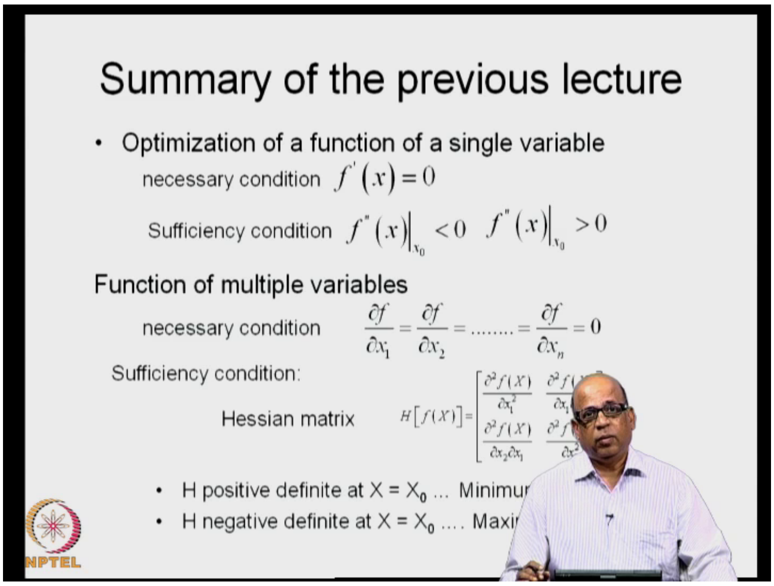
Water Resources Systems
Prof P.P.Mujumdar
Department Of Civil Engineering
Indian Institute Of Science, Bangalore

Lecture No # 05
Constrained optimization (1)

Good morning, and welcome to this lecture number five of the course water resource systems, modeling techniques and analysis; now we are discussing the unconstrained optimization problems, and in the previous lecture, we discussed problems of single variables as well as problems functions of multiple variables. So, if you recall, for the single variables, the necessary condition for a function f of x to have an optimum at a point x is equal to x 0 is that the slope of the function at that point x is equal to x 0 must be 0, that is f dash of x - the first derivative of the function with respect to the variable must be 0 at that particular point. So, to examine whether a function f of x has optimum at a given point, first you have to check whether the slope is 0 at that point; but typically what the… What we do is that given a function f of x, we get the first derivative f dash of x equated to 0, and then obtain all the solutions, corresponding to f dash of x is equal to 0; these solutions are called as stationary points. At each of the stationary points, then we will examine whether the function has a maximum or a minimum or either of them.

So, we use whatever, then we use the sufficiency conditions, we go to the higher order derivatives, second order derivative, third order derivative, fourth order and so on; and capture the first non zero derivative, and look at the sign of that non zero derivative; if the order at which the first non zero derivative is odd, then the function has neither a minimum nor a maximum at the particular stationary point, at which the derivative has been determined; if the order of the first non derivative non zero derivative is even, and the sign of this derivative at that particular stationary point is negative, then the stationary point corresponds to a maximum value of the function; if the sign is negative is positive, then the stationary point corresponds to a minimum value of the function. So, this is what we saw in the previous lecture.

(Refer Slide Time: 02:55)

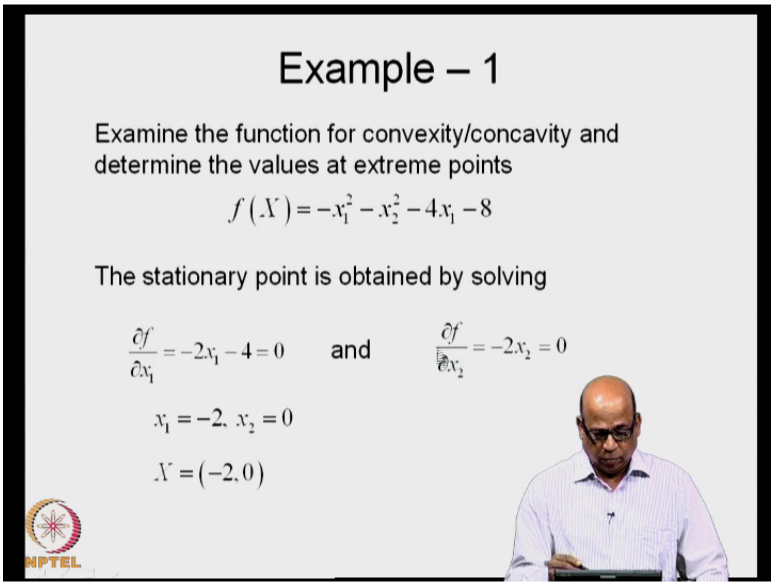


We have let me here for the second derivative that is f dash of x is equal to 0, and the sufficiency condition is f double dash of x evaluated at x 0, x 0 is the stationary point that comes from here, if it is less than 0, then it corresponds to maximum, if it is greater than 0, it corresponds to minimum; and recall that if the second order derivative is 0, then you go to the higher order derivatives and capture the first non zero derivative and so on; then look at the order of the derivative, if it is odd, then the particular stationary point x 0 corresponds neither to a minimum nor to a maximum; if it is even and the sign of the that particular derivative is negative, then it corresponds to a maximum, if the sign is positive, then it corresponds to a minimum; then we went on to extend this for functions of multiple variables x 1, x 2, x 3 etcetera; you have n number of variables, so, f is the function of n number of variables, then the necessary condition is that the first derivative of the function with respect to each of these variables must be equal to 0.

And then for the sufficiency condition, we looked at the hessian matrix formulated the hessian matrix which is in fact, a matrix consisting of the second order derivatives of the function d square f by d x 1 square d square f by d x 1 d x 2 and so on, this is for two variables is, what I have seen, what I have shown here, but typically it will be a n by n matrix; we look at the hessian matrix, and then if the hessian matrix is positive definite, then the particular point this hessian matrix is again evaluated at the stationary point, the stationary point is obtained by solving these equations - n number of equations, you obtain solutions for the n number of variables, they constitute the stationary points.

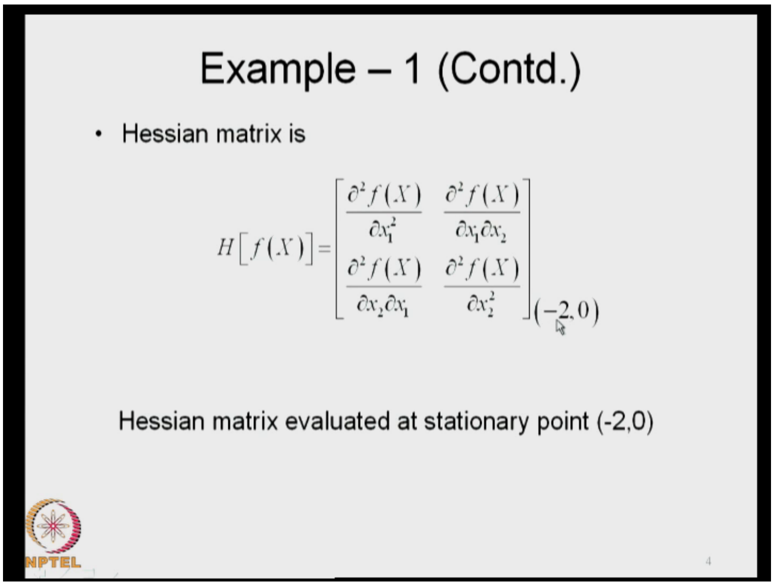
At the stationary point, you evaluate the hessian matrix; at the stationary point, if the hessian matrix is positive definite, the particular stationary point corresponds to a minimum, if at that particular stationary point, the hessian matrix is negative definite, then the stationary point corresponds to a maximum value; we have also seen, what is the positive definite matrix, what is the negative definite matrix, if the… If all the eigen values of the hessian matrix are positive, then the hessian matrix is positive definite; if all the eigen values are negative, then it is a negative definite matrix. So, we will continue with the examples of multiple variables, we are still in the domain of unconstrained optimization; remember we are talking about unconstrained optimization where we are looking at the optimum values of the functions without any… Without attaching any constraints to that, so we will now see some more examples of functions of multiple variables to examine different cases that are possible for stationary points for the optimal solutions for global minimum global maximum and so on.

(Refer Slide Time: 06:13)



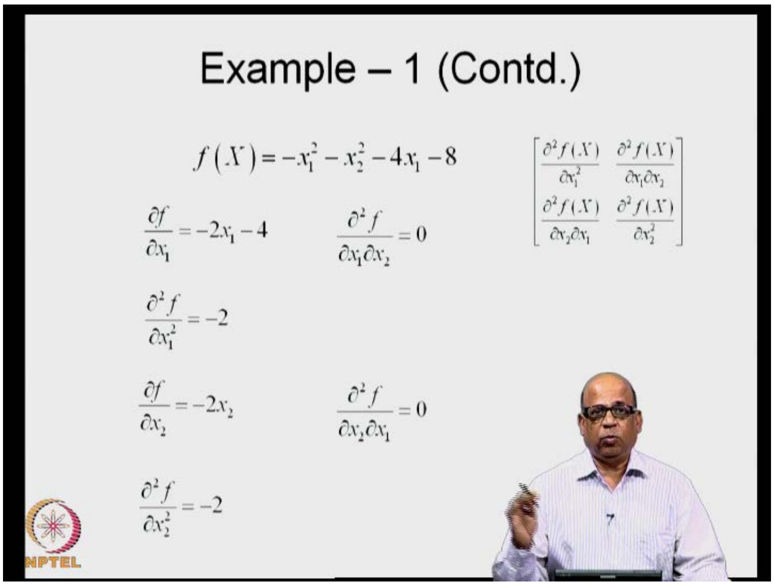
So, we look at function minus x 1 square minus x 2 square minus 4 x 1 minus 8; so, this is the function. We will look at first the stationary point, so we take the first derivative d f by d x 1 and d f by d x 2 and get solutions for x 1 and x 2. So, x 1 and x 2 turn out to be minus 2 and 0. So, x which is the stationary point is minus 2, 0; there is only one solution here. So, this is how we obtain the stationary point, by equating the first derivatives to 0 with respect to each…First derivatives with respect to each of the variables.

(Refer Slide Time: 07:08)



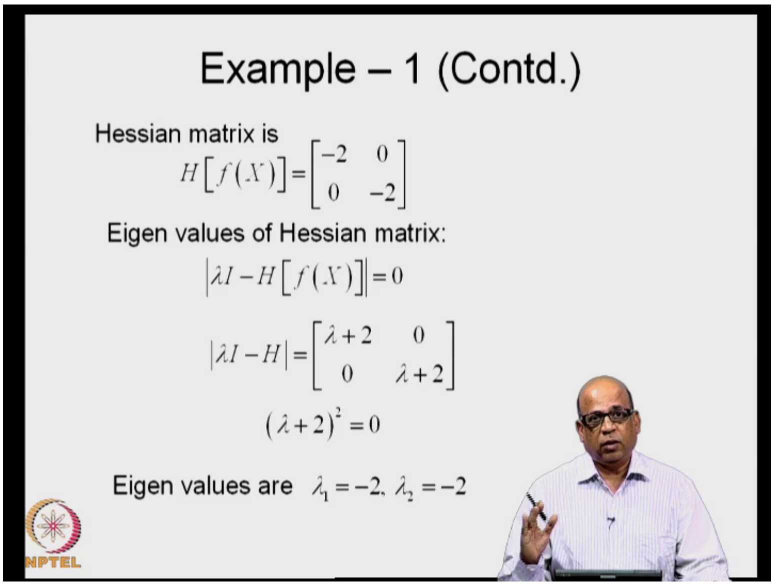
Then we formulate the hessian matrix; so, the hessian matrix of f of x is the matrix of the second derivatives, so d square f by d x 1 square d square f by d x 1 d x 2 d square f by d x 2 d x 1 and d x d square f by d x 2 square; this has to be evaluated at the stationary point minus 2, 0. So, we evaluate the stationary we evaluate the hessian matrix at minus 2, 0.

(Refer Slide Time: 07:41)



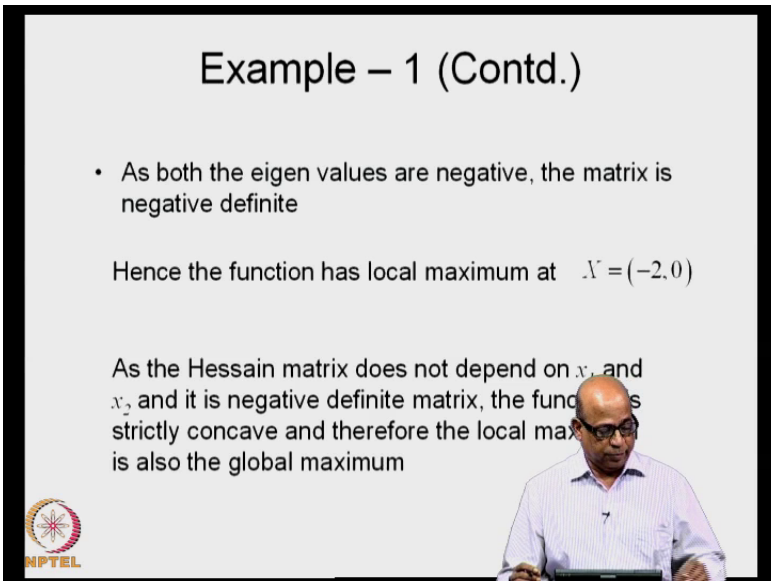
So, we will get the second order derivatives now. So, d f by d x 1 was minus 2 x 1 minus 4. So, we differentiate this again with respect to x 1, you get d square f by d x 1 square minus 2; we differentiate this with respect to x 2, you get a 0; then d f by d x 2 is minus 2 x 2, you differentiate that again with respect to x 2, you get minus 2; and this will be same as d square f by d x 1 d x 2. So, you got all these values, we now formulate the hessian matrix; look at these values now, the second order derivatives are all constants here; that means, they are not functions of either x 1 or x 2; in general, however there will be functions of x 1 and x 2 and therefore, you have to evaluate the hessian matrix at the particular stationary value - stationary point x 1 and x 2 by putting substituting the values of x 1 and x 2; the fact that these are independent of the values of x 1 and x 2, mainly to the conclusion that the hessian matrix is positive definite or negative definite in the entire range of a function, entire range where the function is defined; and therefore, it may lead to… If it is positive definite, it may lead to a positive definite means it is minimum value and therefore, it may lead to a convex function; if it is negative definite, it may lead to a concave function.

(Refer Slide Time: 09:25)



So, we have the hessian matrix minus 2 0 0 minus 2, starting with these derivatives here, we are substituting these values and you get h is equal to minus 2 0 0 minus 2. To examine whether this is the positive definite matrix or a negative definite matrix or neither we evaluate the eigen values. So, to get the eigen values you get determinant lambda i minus h is equal to 0. So, we write lambda i minus h determinant of that as lambda plus 2 here 0 and 0 and lambda plus 2 again here; which gives the lambda values, which is the which are the eigen values as minus 2 and minus 2. So, lambda 1 is minus 2 lambda 2 is minus 2 that is there are two roots of this both the roots are minus 2, because both the lambda values are negative, this hessian matrix is a negative definite matrix; and further, because the lambda values are not dependent on the x 1 and x 2, it means that the lambda value remain negative for all the values of x 1 and x 2 over which the function has been defined; and therefore, the hessian matrix is the negative definite matrix over the entire range of the function; this means that the local maximum that we get is also in fact, the global maximum.

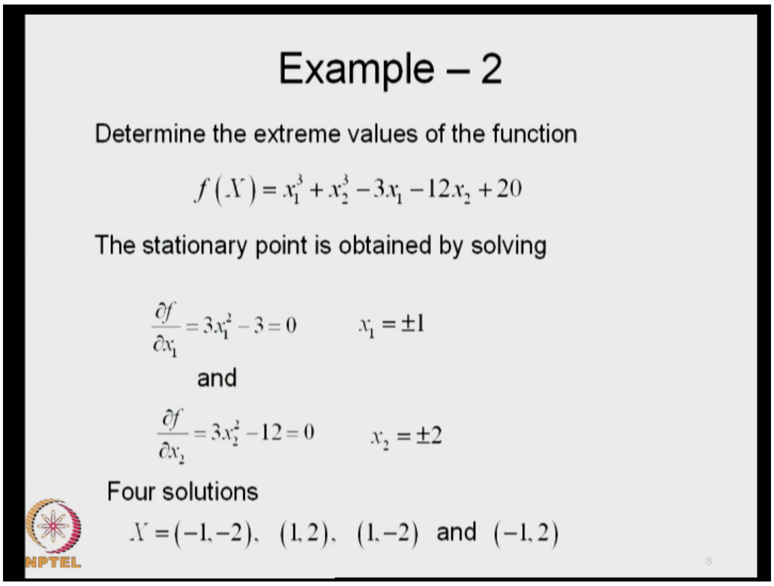
(Refer Slide Time: 11:05)



And therefore, it it is a concave function, the function is the concave function, I will just read out the whole thing now, for completeness; because both the eigen values are negative, first of all the matrix is negative definite; and therefore, the function has a local maximum, because of negative definiteness, it has it corresponds to a maximum value; always remember, when we are making decisions based on the derivatives - higher order derivatives, negative magnitude corresponds to a maximum value, positive magnitude of that derivatives are the matrix hessian matrix corresponds to a minimum value. So, therefore, the local maximum occurs at minus 2, 0; further as the hessian matrix does not depend on x 1 and x 2, and it is negative definite matrix, the function is strictly concave, and therefore, the local maximum is also the global maximum.

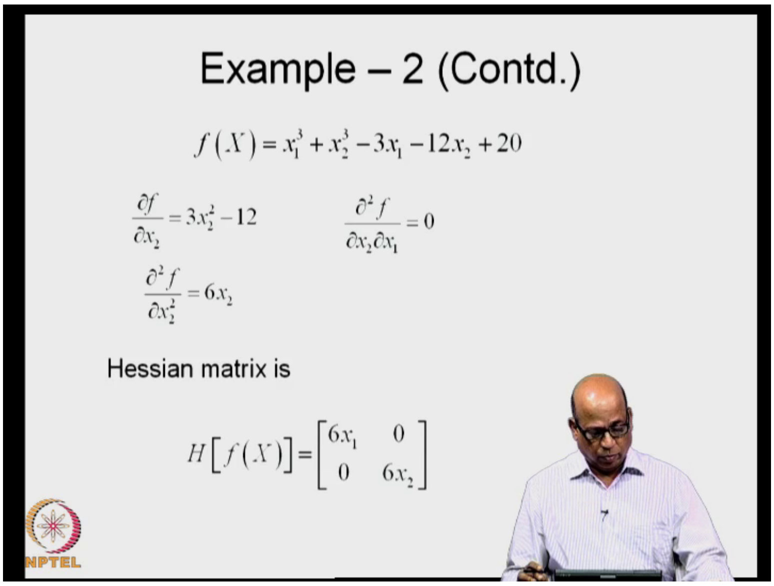
If you had the hessian matrix and therefore, the eigen values as functions of x 1 and x 2, the hessian matrix would have been negative definite only at that particular stationary point; and therefore, you would have had you would have obtained only the local minima, because the eigen values are independent of the values x 1 and x 2, which means that they were the entire range of function the eigen values will remain the same, in this particular case lambda one is equal to minus 2 lambda 2 is equal to minus 2 will remain the same, and therefore, the function is a concave function, and therefore, the local maximum that you obtain is in fact, the global maximum.

(Refer Slide Time: 12:57)



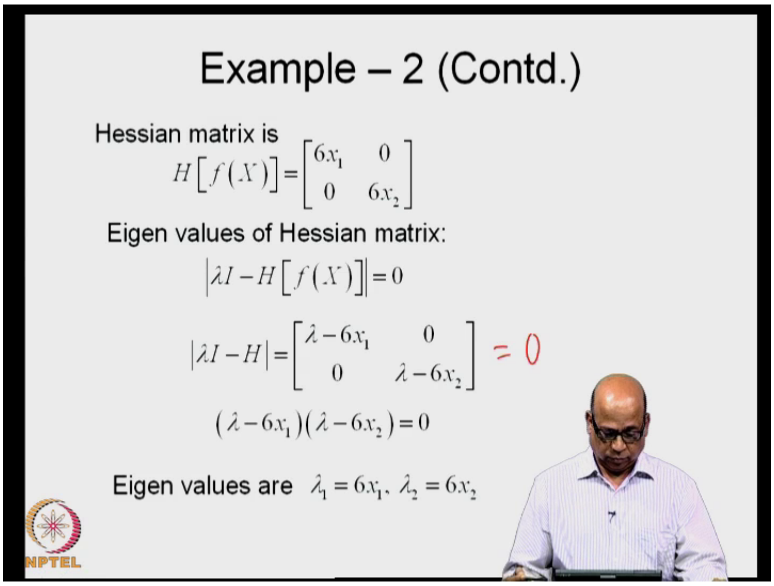
Let us look at for variety one more example; f of x is equal to x 1 q plus x 2 q minus 3x 1 minus 12 x 2 plus 20. So, this is the function of two variables, we will get the stationary points first, by taking the first differential first derivative with respect to each of the variables. So, d f by d x 1 is 3 x 1 square minus 3 is equal to 0, set it to 0; and you get x 1 is equal to plus minus one; and d f by d x 2 is equal to 3x 2 square minus 12 is equal to 0. So, you get x t is equal to plus or minus 2, and therefore, you get four solutions: x is minus 1 minus 2, plus 1 plus 2, plus 1 minus 2, and minus 1 plus 2. So, these are the four solutions that you get. So, these are the stationary points, and at these… Each of these stationary points, you have to examine, whether the point… Whether these point corresponds to a maximum or a minimum or neither of them; to do that, we formulate the hessian matrix, evaluate the hessian matrix at each of these stationary points, and then examine, for each of these stationary points whether the hessian matrix is positive definite or negative definite or neither.

(Refer Slide Time: 14:37)



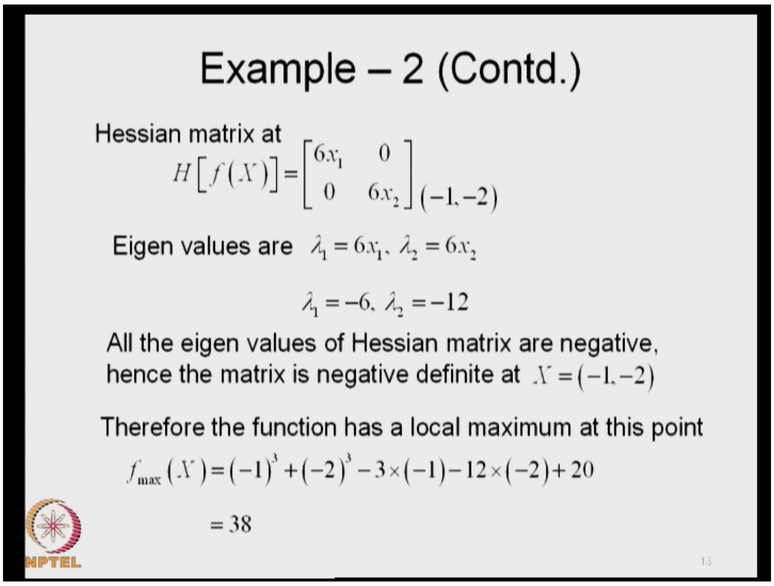
So, let us go to the hessian matrix now. So, the hessian matrix where there are two variables here, therefore d square f by dx 1 square d square f dx 1 dx 2 d square f dx 2 dx 1 and so on; this has to be evaluated at their stationary point; so, we take one by one stationary point, and then evaluate that. So, d f by d x 1 is 3x 1 square minus three and therefore, d square f d x 1 square will be 6 x 1 d square f by dx 1 dx 2 is 0, and similarly, you have d f by d x 2 is 3 x square minus 12 and d square f by dx 2 square is 6x 2, and the second order derivative with respect to x 2 and x 1 is 0. So, the hessian matrix is 6x 1 here d square fd x square 6x 1 and 0 0 and 6x 2. So, this determines the hessian matrix; unlike in the previous example, the hessian matrix is now a function of x 1 and x 2; and therefore, you have to evaluate the hessian matrix at each of the stationary points that you have obtained, you have obtained four stationary points now, for each of the stationary point, we will have to examine whether the hessian matrix is a positive definite matrix or a negative definite matrix.

(Refer Slide Time: 16:08)



So, we will evaluate this for different stationary points; either you evaluate the stationary… Hessian matrix at each of the stationary points, and then obtain the Lagrange multipliers I am sorry obtain the eigen values I am sorry we will introduce the Lagrange multipliers presently in this course, which are also denoted by lambda, but the eigen values we determine based on the … Based on the hessian matrix, and then substitute the values of x 1 and x 2. So, that is what we do here; eigen values we determine determinant of lambda i minus h, which is lambda minus 6 1 0 0 lambda minus 6x 2 and this is set to 0. So, he set this to equal to 0; and we get the solution lambda minus 6x 1 lambda minus 6x 2 is equal to 0; which gives values of lambda 1 is equal to 6x 1 and lambda 2 is equal to 6x 2.

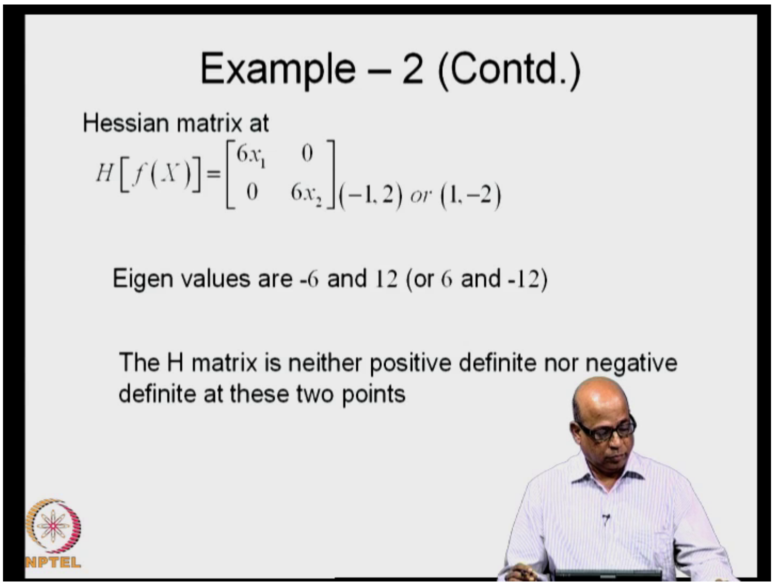
(Refer Slide Time: 17:25)



So, the hessian matrix at 1 comma 2 will give us lambda 1 of 6 and lambda 2 of 12, both of which are positive. So at 1 comma 2, you have 2 lambda values both of which are positive, and therefore, the hessian matrix is a positive definite matrix at the stationary point 1 comma 2. So, it is positive definite at x is equal to 1 comma 2, because hessian matrix is positive definite, the point corresponds to a minimum value of the function; and the minimum value of the function, we can determine by substituting x 1 is equal to 1 and x 2 is equal to 2 in the original function definition. So, in this you put x 1 is equal to 1, and x 2 is equal to 2, and that is what you get as the minimum value - local minimum; hessian matrix being positive means that the stationary point is a minimum value and the minimum value is obtained by substituting the values x 1 is equal to 1 and x 2 is equal to 2. So, this is what you get; then we could move to the next stationary point.

We remember we have four stationary points now, so we go to the next stationary point we evaluate this at minus 1 minus 2 that is a hessian matrix at minus 1 minus 2. So, the lambda values will be minus 6 and minus 12. So, at minus 1 minus 2 at the stationary point minus 1 minus 2, you get the lambda values as minus 6 minus 12 both of which are negative, and therefore the hessian matrix is the negative definite at the stationary point minus 1 minus 2, because the hessian matrix is negative definite, the point minus 1 minus 2 corresponds to a maximum value. So, all the eigen values of the hessian matrix are negative, in this case there are two eigen values both of which are negative; and therefore, the hessian matrix is negative definite and negative definite means, it is a maximum value and we calculate the maximum value as 38, by substituting x 1 is equal to minus 1 and x 2 is equal to minus 2; like this we go to the next available values.

(Refer Slide Time: 19:58)



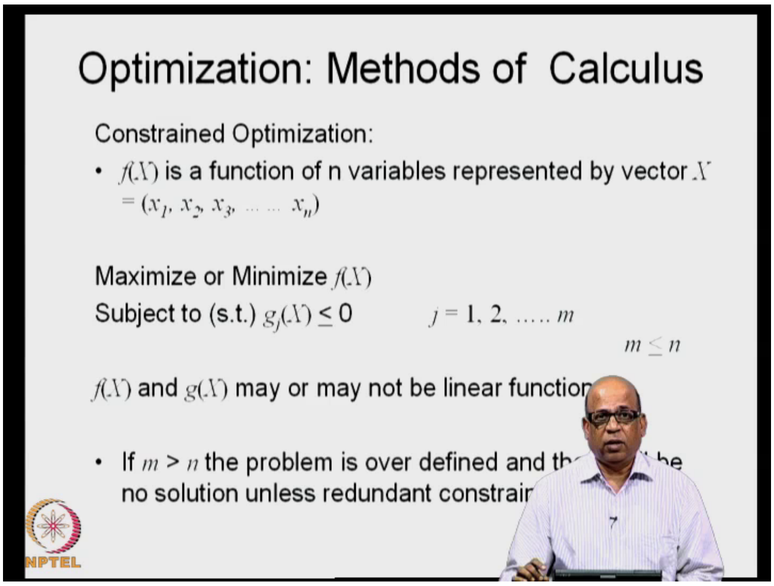
So, we will exhaust minus 1 comma 2 and 1 comma minus 2; that means, x 1 is equal to minus 1 x 2 is equal to 2; reevaluate your eigen values, x 1 is equal to 1 x 2 is equal to minus 2; reevaluate the eigen values and examine whether the whether all the eigen values are negative, all the eigen values are positive or some of them are positive and some of them are negative. So, in this case, you get the eigen values as minus 6 and 12 for this stationary point; and 6 and minus 12 for this stationary point; which means out of the two eigen values, one is positive and one is negative in both the cases; and therefore, we will not be able to… Therefore, the points - stationary points of minus 1 comma 2 correspond neither to a local minimum nor to a local maximum; similarly, the point 1 comma minus 2 corresponds neither to a local minimum nor to a local maximum.

So, out of the four stationary points that we obtained in this example we could identify one as a local minimum, one as a local maximum and the other two as neither local minimum nor local maximum; and in fact, the from the hessian matrix, we also saw that the hessian matrix and therefore, the corresponding eigen values are in fact, functions of x 1 and x 2 and therefore, the function is neither a convex function nor a concave function and therefore, it behaves differently at different stationary points.

So, this is what we did in unconstrained optimization we started with the single functions of single variables then extended it on to functions of multiple variables, but we did not attach any constraints; these are by for the simplest of problems, and they are introduced typically as just a background necessary for further optimization; all most (Audio not clear from 22:15 to 22:18) we will have a set of constraints, set of conditions that we want to impose on the problem, for example, let us say you are trying…You are minimizing the cost of a reservoir, cost of a dam; then you need to the conditions on what are the functionalities that would like to expect, because obviously, if you want to have a minimum cost, it can be a 0 cost, if you do nothing is a solution, but you would like to achieve some functional functional performance of the particular structure, and therefore you would like to put conditions that minimum we need to produce certain amount of hydro power or you want to maintain a minimum flood control or you want to maintain the mass balance the from one time period to another time period and so on.

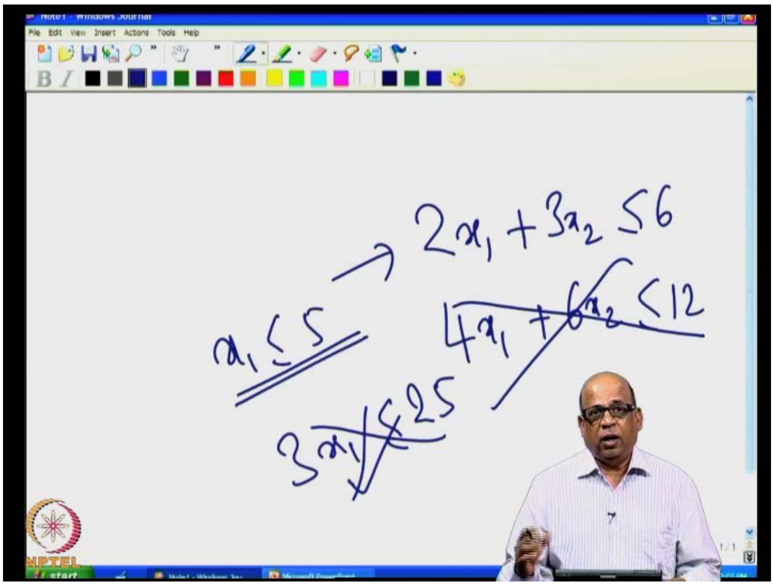
So, all the physical conditions that the problem has to satisfy often come as constraints and these constraints will have to be included in the mathematical formulation of the optimization and we will be we will be optimizing the function subject to several of these constraints. So, any engineering optimization problem is almost always unless we have some means of removing the constraints, a generally stated optimization problem almost always constraint contains constraints, and therefore, we should be talking about constrained optimization. So, we will now start talking about constrained optimization, where we are we are looking at optimal values of the function, subject to a series of, a set of constraints. So, let us see, we will start with the simplest of the constraints namely, the equality constrain, and then we will go to more generalized non equality in equality constraints.

(Refer Slide Time: 24:27)



So, in the constrained optimization, we have a function f of x of n variables that is multiple variables x 1, x 2, x 3 etcetera x n, and then we are subjected to a set of constraints of the type g j of x less than or equal to 0, j is equal to 1, 2 etcetera n. So, there are n variables, and there are m number of constraints; let us say, you have four variables and three constraints, four variables four constraints, four variables ten constraints and so on; if the number of constraints m is larger than the number of variables n, then in general the problem is over defined, and therefore, there will be there will be no solution, unless there are what are called as redundant constraints; what do we mean by redundant constraints? If one of the… If you state one of the constraints and the other constraint is automatically satisfied by stating this particular constraint, then the other constraint becomes redundant.

(Refer Slide Time: 25:51)

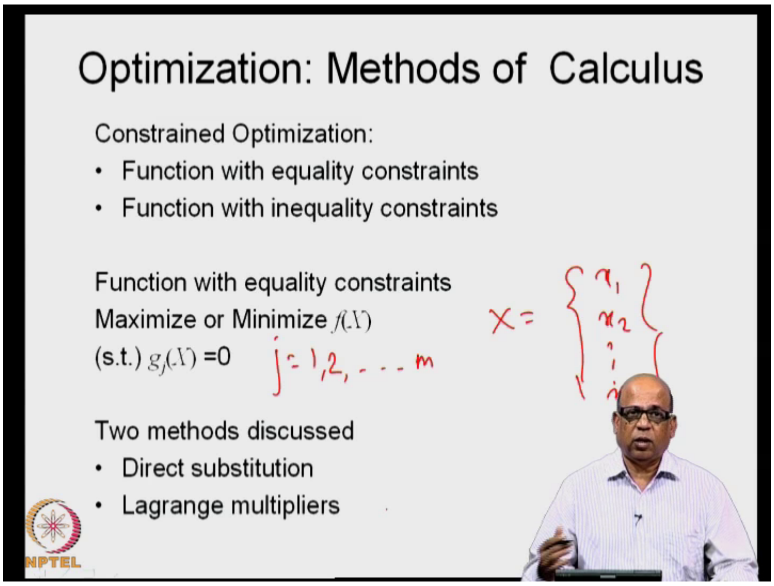


For example, let us say that you have a constraint of the type 2x 1 plus 3x 2 is less than or equal to 6; this is one of the constraint into variables; let us say, I have an another constraint 4x 1 plus 6x 2 is less than or equal to 12; that means, simply multiply this by 2; the naturally, whenever this is satisfied, this is always satisfied, so one of the constraints becomes redundant constraint; similarly let say you have x 1 is less than or equal to 5 as one of the constraints; then elsewhere in the model or in the formulation, you may have some constraint of the type 3x 1 is less than or equal to let us say 25; the moment it satisfy this constraint x 1 is less than equal to 5, automatically this constraint is satisfied 3x 1 is less than or equal to 25. So, in general when you have redundant constraint, and therefore this becomes redundant constraint, in general whenever you have redundant constraints it means that not enough thought has gone in to the formulations of a constraints; the constraints are after all formed based on the physical understanding of the system, and therefore, you may tend to add constraints which are in fact, functions of other constraints have to, you have already included in the problem.

So, whenever you have m greater than n that is the number of constraints higher than the number of variables, the problem is in general over defined and therefore, you should look for, you should be alert to the situation where you have redundant constraint in this. So, you should be able to identify the redundant constraints and remove the redundant constraints; we will see this more when we look at large applications, large scale applications of multiple reservoir operations and so; on where generally we tend to introduce redundant constraints.

So, this is the concept of redundant constraints; now, so typically, we will have m less than or equal to n. So, when you have m less than or equal to n that means, the number of variables is larger than the number of constraints, and that is the type of problems that we will be addressing now; this is the general generally stated problem where we are looking at maximization or minimization of a function, subject to g j of x; we are not at this stage putting any conditions on the nature of the function itself. So, f of x and g j of x can assume any type of functions, there may be linear, non-linear and and so on. So, we are not saying that either f of x or g j of x must be linear functions from this general optimization problem, we will start the discussion with the simpler version of this general statement where we will consider all the constraints to be equality constraint. So, from g j of x less than or equal to 0, I will start with g j of x equal to 0 that is some functions of the decision variables x 1, x 2, x 3 etcetera, all these functions are equal to 0. So, we start with problems with equality constraints.

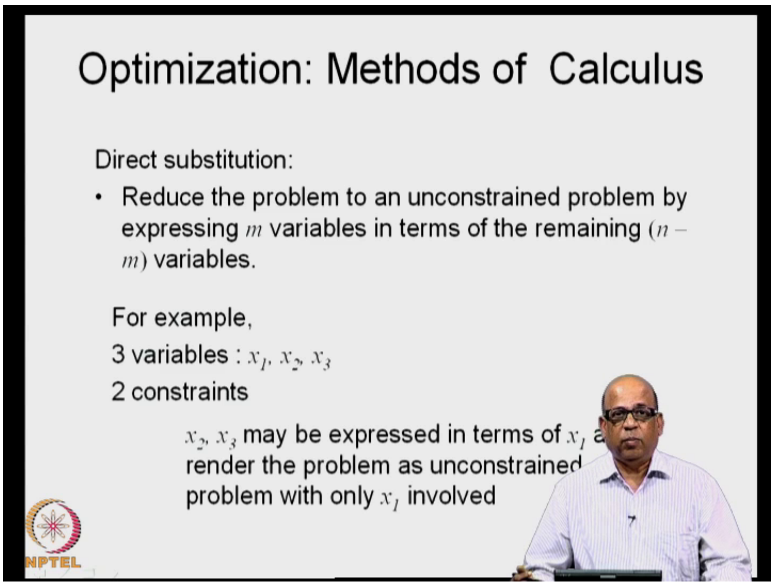
(Refer Slide Time: 29:37)



So, we have starting with this function with equality right; how do we state that? Maximize or minimize f of x subject to g j of x is equal to 0, and this is for j is equal to 1 to m, and x is a vector x 1 x 2 x n, it has n variables and we have m number of constraints; now, in this general problem, we will discuss two simple methods: One is direct substitution, another is the method of Lagrange multipliers; in the direct substitution what we do is, let us say have m number of constraints, because they are all equality constraints, it is possible for you to express m number of variables in terms of the remaining n minus m number of variables, there are n number of variables there are m constraints. So, you express using these m constraints which are all equality constraints, using these m constraints, you express m number of variables in terms of the remaining n minus n number of variables, and substitute them and therefore, you are you have taken care of the m conditions, m constraints and then converted the original constraint problem in to an unconstraint problem with n minus m number of variables. So, this is what we do in direct substitution.

A direct substitution is almost always handy when you have less number of constraints in class room type of examples, where you may have two variables one constraint, two variables two constrain three variables two constraints and so on, but in most practical situations where you have to deal with large number of variables, large number of constraints and so on; the direct substitution… And also the nature of the construct themselves; if there may be non-linear constraints and so on. So, the substitution may not be always very handy, in which case you go with Lagrange multipliers I will introduce in the today's class.

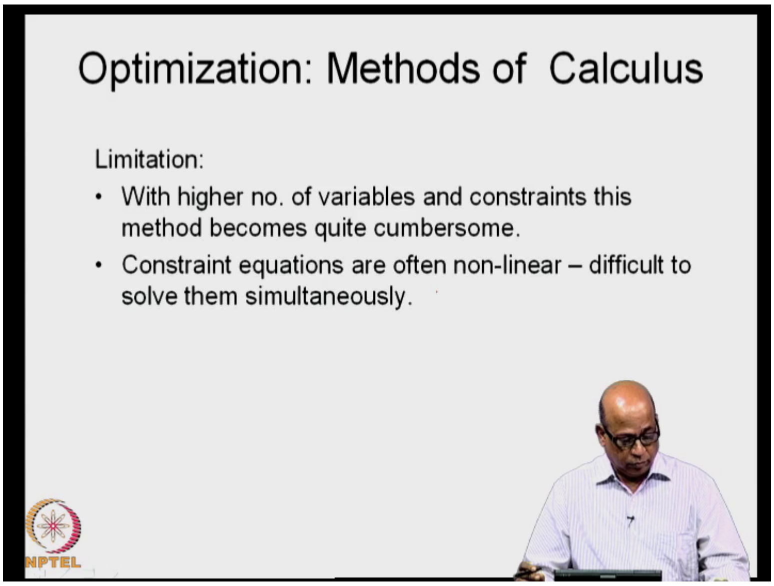
(Refer Slide Time: 32:03)



So, in the direct substitution method, we reduce the problem to unconstrained problem by expressing n variables in terms of the remaining n minus m variables, you have m constraints and you use this m constraints to express the m variables in terms of the remaining n minus m variables; remember all these m constraints are in terms of all the n variables, and we use this m constraints to convert the problem in to a unconstrained problem of n minus m variables; for example, you have three variables x 1, x 2, x 3; we have two constraints both of the constraints contain all the three variables in general, then we may pick two of the variables let say x 2 and x 3, and solve these two constraints and express both x 2 and x 3 in terms of x 1; and that is what we do.

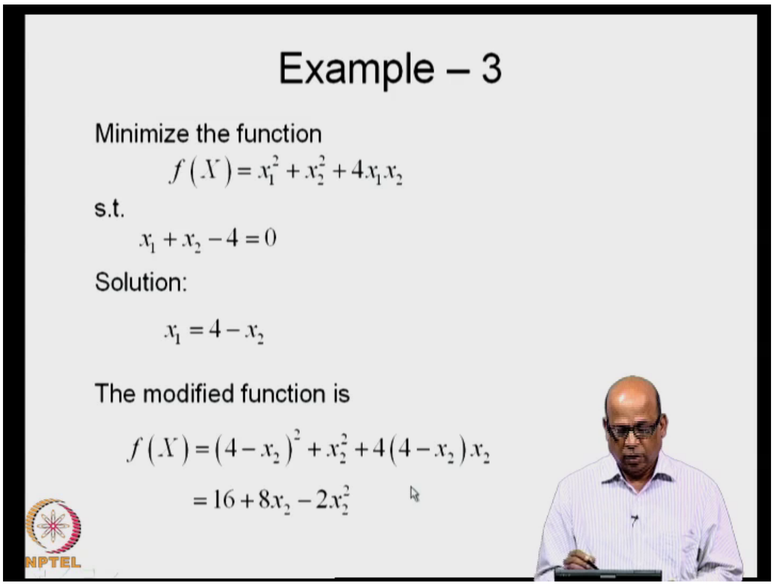
So, two variables we can we express in terms of the n minus m variable, it is 3 minus 2 variables, which is one variable, and that is x 1; and then render the problem as unconstrained problem with only x 1 involved; and once it is rendered in un constrained problem, we know how to solve this; we have just seen the unconstrained optimization method, so, we use the unconstrained optimization method to solve such a problem. So, this is what we do in direct substitution, I again repeat direct substitution is good only as the class room type of examples where you you have less number of variables, less number of constraints and so on where you can easily solve and substitute for some variables in terms of the other variables.

(Refer Slide Time: 33:47)



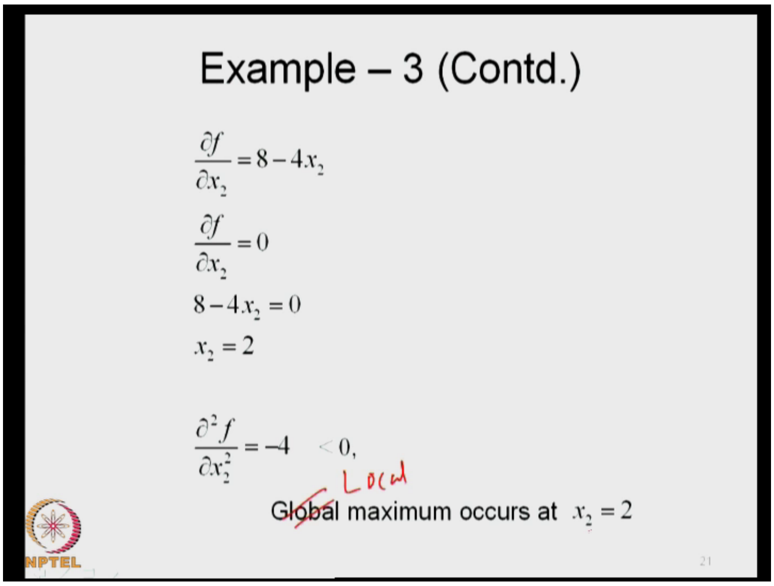
Constrained equations that we come across in practical problems are often non-linear and therefore, direct substitution also becomes quite involved and difficult.

(Refer Slide Time: 34:07)



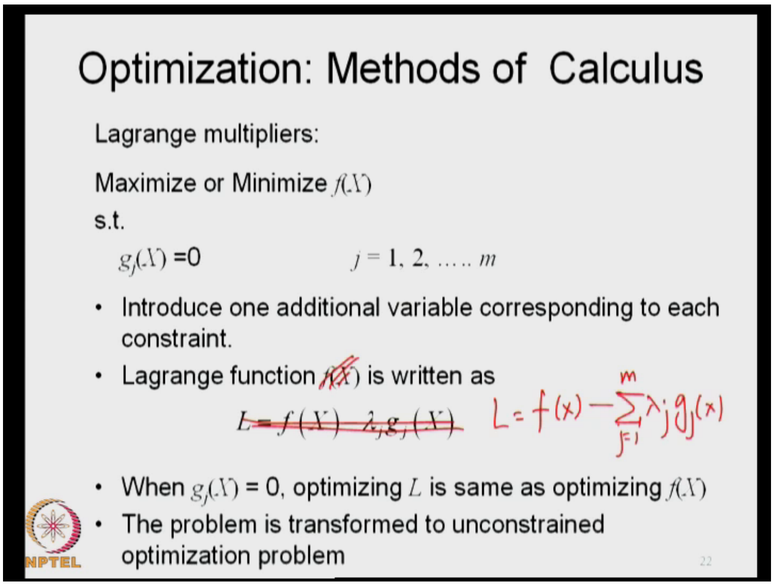
So, let us see one example, let us say that you have a function f of x is equal to x 1 square plus x 2 square plus 4 x 1 x 2 subject to 1 a single constraint x 1 plus x 2 minus 4 is equal to 0; now this is of the form g j of x is equal to 0. So, this is the equality constraint; we have an equality constraint. So, let us say that you want to substitute for one of the variable; there are two variables one constraint. So, n is equal to 2 m is equal to 1. So, we express m number of … m number of variables which is one number of variable in terms of n minus m variables, which is 2 minus 1, which is 1. So, express x 1 in terms of x 2 x 1 is equal to 4 minus x 2, and then substitute it in the original function. So, x 1 square which is 4 minus x 2 square plus x 2 square plus 4 x 1 4 minus x 2 in to x 2. So, we have converted the original problem of two variables in to a problem of one variable using the constraint. So, this becomes 16 plus 8x 2 minus 2x 2 square. So, this is the function of a single variable; once we know, once we have converted the function into function of single variable, and then got read of all the constraints we can solve this problem as an unconstrained optimization problem, and in this particular case of a single variable.

(Refer Slide Time: 35:50)



So, we take the first derivative d f by d x 2, the x 2 is the only variable here; and then equated to 0, you get x 2 is equal to 0, you get x 2 is equal to 2 as a stationary point; go to the second derivative that becomes minus 4 from here, it becomes minus 4 which is negative and therefore, the global maximum occurs. Actually from here, we conclude that it is a local maximum occurs at this location, at x 2 is equal to 2. So, this is what we do in direct substitution that is we substituted for one of the variables in terms of the other variable, m number of …We substitute for n number of variables in terms of the other n minus m variables, and then convert the problem in to an unconstrained optimization problem in terms of n minus m variable and we use the methods of unconstrained optimization and then solve the problem.

(Refer Slide Time: 36:59)

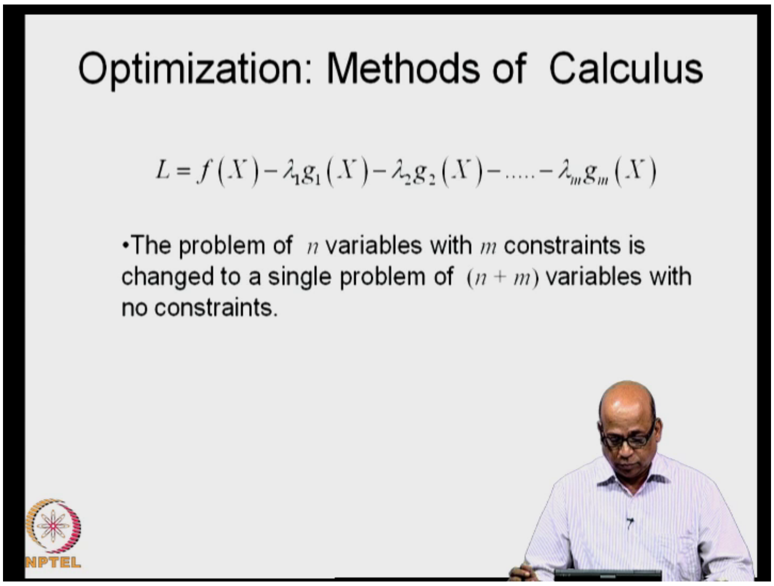


However, for large problems, this method is not a elegant method it will be quite cumbersome and therefore, we evolve methods of Lagrange multipliers, which are much more regress, which is much more regress method of calculus. We have n number of variables, we have m number of constraints; in the direct substitution method, what we did is, we eliminated a certain number of variables, in fact, we eliminated m number of variables; in the Lagrange multiplier method, we add n number of variables which means, we introduce an additional m number of variables; let us see what we do in this, corresponding to each as a the constraints g j of x we introduce a Lagrange multiplier.

So, g 1 of x has lambda 1, g 2 of x has lambda 2, g 3 of x has lambda 3 etcetera, associated with each of the equality constraint we have a Lagrange multiplier. So, that is what we do introduce 1 additional variable lambda j corresponding to each constraint j, which is g j of x is equal to 0 as our constraint, we associate with each constraint a Lagrange multiplier lambda j; and then we formulate the Lagrange function you have the original function f of x which is the function of n number of variables we have the constraints g j of x is equal to 0, associated with each of these constraints g j of x, we have now a Lagrange multiplier lambda j; we formulate what is called as the a Lagrange function using the function f of x and the constraints g j of x along with the Lagrange multipliers lambda j, and the Lagrange function l is written as… Lagrange function is written as l is equal to f of x, which is the original function minus lambda j g j of x. So, you have g j of x here there will be summation here; sorry let me write; I will rewrite this here l is equal to f of x minus summation j is equal to 1 to m, there are m constraints lambda j g j of x now this is how we formulate the Lagrange function that is Lagrange associated with the original function f of x and the set of constraints g j of x with Lagrange multipliers; these lambda js are called as the Lagrange multipliers, you have one Lagrange multiplier associated with one constraint. So g g 1 of x will have lambda 1, g 2 of x will have lambda 2 and so on. So, there are m number of Lagrange multipliers and this is the function l Lagrange function.

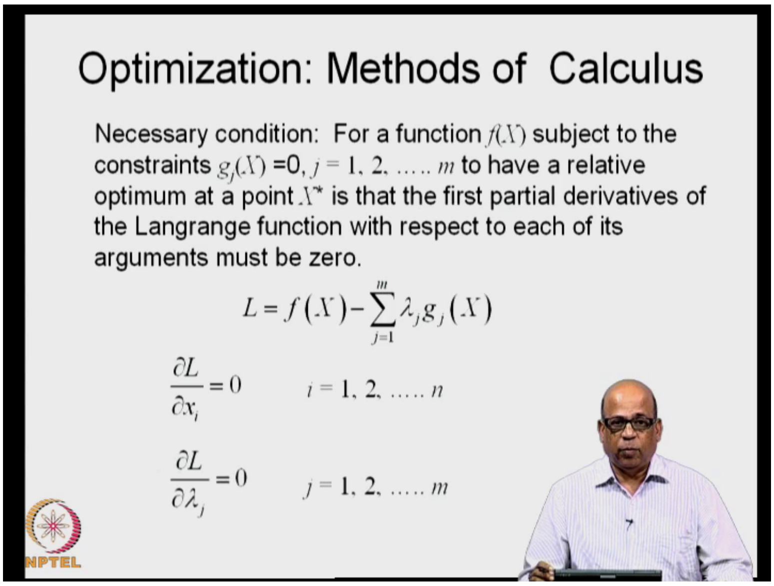
Now, because g j of x is equal to 0, what does that mean? Optimizing l also means that you are optimizing f of x. So, the problem is now transformed to transferred to unconstrained optimization problems. So, you have, you are looking at only l is equal to f of x minus lambda j summation lambda j g j of x as the Lagrange function. So, we will work with this function now, by converting the constraints using lambda j g j of x. So, we have incorporated the all the constraints in to the Lagrange function. So, we will work with Lagrange function; how many arguments are there in Lagrange function? You had n number of variables here x 1, x 2, x 3 etcetera x n; then you have m number of multipliers lambda 1, lambda 2, lambda 3 and so on. So, when you are optimizing l, you have to the necessary condition is that you differentiate l with respect to each of its augments. So, differentiate l with respect to x 1 with respect to x 2 etcetera, first order derivatives with respect to x 1, x 2, x 3 etcetera x n, also with respect to lambda 1, lambda 2 etcetera lambda n; and equated to 0.

(Refer Slide Time: 42:05)



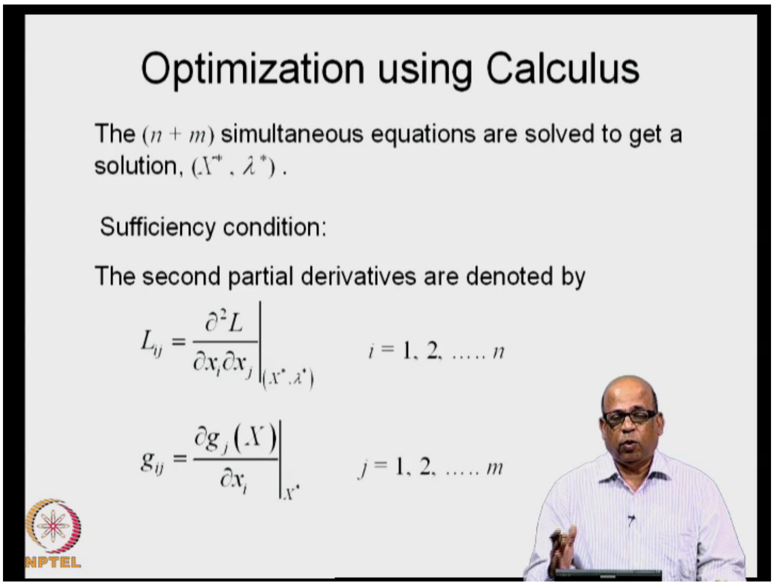
So, this is what we start with as necessary condition. So, L is written in the longer form as f of x minus lambda 1 g 1 of x minus lambda 2 g 2 of x etcetera minus lambda m g m of x. So, the problem of n variables with m constraints is now change to a single problem of n plus m variables with no constraints; just distinguish this with respect to what we did in the direct substitution method; in direct substitution method, you had m number of constraints, we express the m number of variables, which actually means we eliminated m number of variables, and then express the original problem with n minus m number of variables; still making it unconstrained optimization; in the Lagrange multiplier method on the other hand, we introduced m number of additional variables which are called as the Lagrange multipliers and therefore, the original problem of n variables with m constraints have now been converted in to an unconstrained optimization with n plus m number of variables, we have added m number of variables in this case.

(Refer Slide Time: 43:10)



So, the necessary condition, we formulated the Lagrange now; the necessary condition is, we state it formally, for a function f of x subject to the constraints g j of x is equal to 0 j is equal to 1 to n, you have a relative optimum, which means relative minimum or relative maximum at a point x star, this is the stationary point x star, is that the first partial derivatives of the Lagrange function with respect to each of its arguments must be 0. So, we have formulated the Lagrange here l is equal to f of x minus summation j is equal to 1 to m lambda j g j of x, with respect to each of its argument namely, x its arguments are x 1, x 2, etcetera x n; lambda 1, lambda 2, etcetera lambda m; first derivatives must be 0. So, d l by d x i is equal to 0 for i is equal to 1 to n d l by d lambda j must be equal to 0 for j is equal to 1 to m. So, you formulate n plus m equations, there are n variables here, n equations here; there are m equations here. So, n plus m equations solve them simultaneously, you will get solutions for these m plus n variables, what are the n plus m variables x 1, x 2, x 3 etcetera x n; lambda 1, lambda 2, lambda 3 etcetera lambda m. So, you solve for x i as well as lambda js and that constitutes the stationary point.

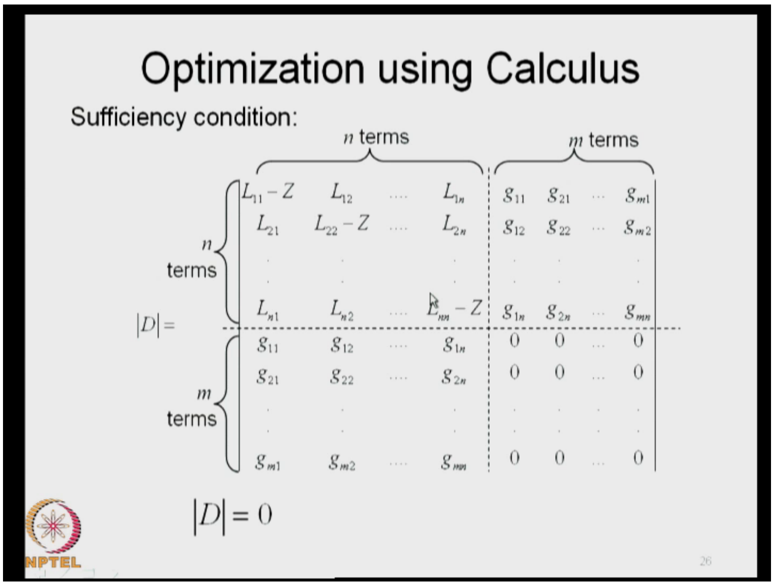
(Refer Slide Time: 44:43)



So, you get after solution of this n plus m equations, you get x star and lambda star as the solution, this is in fact, called as the stationary point; then we go to sufficiency condition. So, this is the necessary condition; that means, the first order derivatives of the Lagrange which respect to each of the arguments must be equal to 0, from that you solve n plus m equations, you get x star which is the x 1 star, x 2 star, etcetera x n star; lambda 1 star, lambda 2 star etcetera up to lambda j star lambda m star. So, you get the solutions of that, then we move to the sufficiency conditions; sufficiency conditions are always in general in terms of the second order derivatives; now, you have Lagrange function l, which is a function of x 1, x 2 etcetera x n; and lambda 1, lambda 2, etcetera lambda m; and you also have the constraints, you have m number of constraints, each of which is the function of n number of variables.

So, we will see here, we will define L i j this part will just be clear on the notation we define L i j as the second derivative of the Lagrange L, with respect to the two variables x i and x j. So, L i j is second derivative of the Lagrange L with respect to the variable x i and x j, evaluated at the stationary point x star, lambda star; we denote by g i j the first order derivative of the j th constrained g j of x with respect to the variable i. So, d g j by d x i is what we denote as g i j. So, i is the i th variable, j is the j th constraint, and we are talking about the first order derivatives. So, L i j corresponds to second order derivative, g i j corresponds to the first order derivative, and this is done for j is equal to 1 to m for each of the values of i. So, you have L i j and you have g i j i corresponds to the variable j corresponds to the constraint in this particular case g i j, and in this case it corresponds to the two variables x i and x j. So, just be clear about the notation that we are using; then we formulate a determinant, so you got the second order derivative with respect to each of the variables and you got the first order derivative of the constrain with respect to each of the variables; using these derivatives, now we formulate a determinant.

(Refer Slide Time: 47:59)

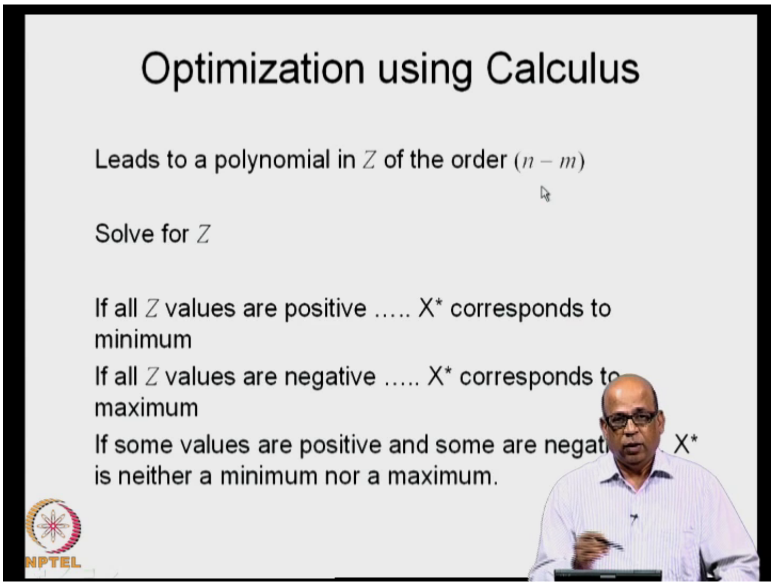


This determinant is formulated as follows; we will not going to the theory of how the sufficiency condition is arrived at and so on; those who are interested are referred to the classical book by S. S. Rao that is optimization theory and practice; but we will not worry about the theory, but we will see how the sufficiency conditions are formed for a given particular problem, and then how we arrive at the conclusion of whether the stationary point corresponds to a maximum or a minimum; we formulate this determinant thus, you have L i j defined here. So, L 11 minus Z L 12 etcetera L 1n, which means with respect to first variable 1, 2, 3, 4 etcetera up to n. So, that is what we are taking L 11 L 12 etcetera L 1n, and for the first term you are taking minus Z similarly, L 21 L 22 minus Z etcetera L 2n, L n1 n 2 nn minus Z. So, these are n terms by n terms, this is what you formulate here, followed by g 11 g 21 etcetera g m1, which means the first constraint with respect to first variable, second constraint with respect to first variable, m eth constraint with respect to first variable, then g 12 g 22 g m2 first constraint with respect to second variable, second constraint with respect to second variable, m eth constraint with respect to second variable, like this, you come to first constraint with respect to n eth variable, second constraint with respect to n eth variable, n eth constraint with respect to m eth variable. So, these are defined from here; so, these are the first order derivatives of these constraints with respect to the variables here.

So, these are m terms, and this is n terms here on along this, then we come to the lower left part of the determinant; we take the transform of this and then sorry we take the transpose of this part, and then put the transpose of this part here. So, g 11 g 12 etcetera g 1n, we arrange them horizontally, g 21 g 22 etcetera g 2n we arrange horizontally, and g m1 g m2 etcetera you arrange horizontally. So, this is how we formulated the three parts of this determinant, there are remaining m by m terms here, this will be m and this is m. So, n by m terms, all of them you put it as 0. So, this is how we formulate the determinant D and set the determinant D to 0, equal to 0; there are only the unknowns are Z here; just for clarity, I will again explain, how this determinant is formed, we have L ij d square L by d x i d x j, we know the function L difference is take the second order derivative with respect to variables x i and x j that defines L i j; for example, L 21 will be second order derivative of L with respect to d x 2 d x 1. So, that is how you formulate L ij, L 22 for example, will be d square L by d x square d x 2 square, then g ij you define as the differential of the j eth constraint with respect to i eth variable. So, you know g i j j eth constraint with respect to i eth variable; now you come here, you write m terms g 11 g 21 g m1 g 12 g 22 g m2 similarly, L 11 minus

Z L 12 minus (Audio not clear. Refer Time: 52:28) etcetera, and transpose this and put it in terms of g 11 g 12 etcetera g 1m, and the remaining m by n terms you put it as 0. So, this determinant d, you set it to 0, you get a polynomial in Z; when you formulate this determinant; I am set it to 0, you get a polynomial in Z of the order n minus m, these are n plus m terms here, n plus m terms here and only these n terms have Z, you get a polynomial of the order n minus m in Z, solve for that solve the polynomial for Z, you obtain z values.

(Refer Slide Time: 53:13)



So, it leads to the polynomial in Z of the order n minus m; we solve for Z, if all the Z values are positive, then the x star which is the stationary point corresponds to a local minimum; if all Z values are negative, x star is correspond to a local maximum; if some values are positive and some are negative x star is neither a minimum nor a maximum. So, just quickly we will recapitulate, what we did for in the Lagrange multipliers; we associate with respect to each of the constraints a Lagrange multiplier g j of x with associate lambda lambda j of x lambda j we associate with each of the g j of x and therefore, we have as many lagranger multipliers as we have number of x constraints; then formulate the Lagrange, a Lagrange function L as f of x minus summation j is equal to 1 to m, we have m number of terms, lambda j g j. So, that is the Lagrange function.

Then the necessary condition for that is that the first order derivative of the Lagrange with respect to each of the arguments must be equal to 0; what are the arguments x 1, x 2 etcetera x n, which are the original variables and lambda 1, lambda 2, lambda 3 etcetera lambda m which are the additional variables, multipliers that we have a added. So, you formulate n plus m equation like this; solve for the n plus m equations, you get the solutions for x 1, x 2 etcetera x n, which defines a x star and lambda 1, lambda 2, lambda m which will define lambda star; now this is the stationary point x star lambda star, then we go to the second order derivative to get the sufficiency conditions, we get the second order derivative L i j, which is d square L by d x i d x j that is the second order derivative of the Lagrange function with respect to the variables x i and x j; similarly, we get g i j which is the first order derivative of the j th constraint g j with respect to the i th variable x i; then we formulate the determinant and introduce this is the determinant that you formulate, and then introduce Z in the diagonal elements of L i j only, and then solve for Z this determinant when you set it to 0, it gives an a polynomial of the order of n minus m and solve for Z; if all the z values are positive, then the stationary point x star corresponds to a minimum; if all z values are negative, then it corresponds to a maximum; some of them are positive some of them are negative, then it corresponds to neither a minimum nor a maximum.

So, we will solve some numerical examples in a next class on Lagrange multipliers; remember Lagrange multipliers is the most powerful method and in fact, the physical significance of lagrange multipliers etcetera, we will be using subsequently when we deal with sensitivity analysis in the linear programming and so on. So, understand the Lagrange multiplier method correctly, and we will also see the interpretation of the Lagrange multipliers. So, in today's class quickly we just started off with the unconstrained optimization problem, two examples we solved; and then we went on to introduce the constrained optimization problems, and in the constrained optimization problems, we have started with problems with equality constraints; an easy way of handling the equality constraints in simple problems is by direct substitution, where you can replace m variables by expressing them in terms of the remaining n minus m variables, and rendering the problem to be an unconstrained optimization problem.

In the more regress Lagrange multiplier method, we add m number of variables in fact, we add m number of Lagrange multipliers, one associated with each of the constraints, you have m constraints, you add one Lagrange multiplier associated with each of the constraints; and then we formulate the necessary conditions and the sufficiency conditions; we will continue the discussion on Lagrange multipliers in the next class. In fact, we solve some numerical example, so that all points are clear on the Lagrange multipliers; thank you for your attention.s