Water Resources Systems

Modeling Techniques and Analysis

Prof. P. P. Mujumdar

Department of Civil Engineering

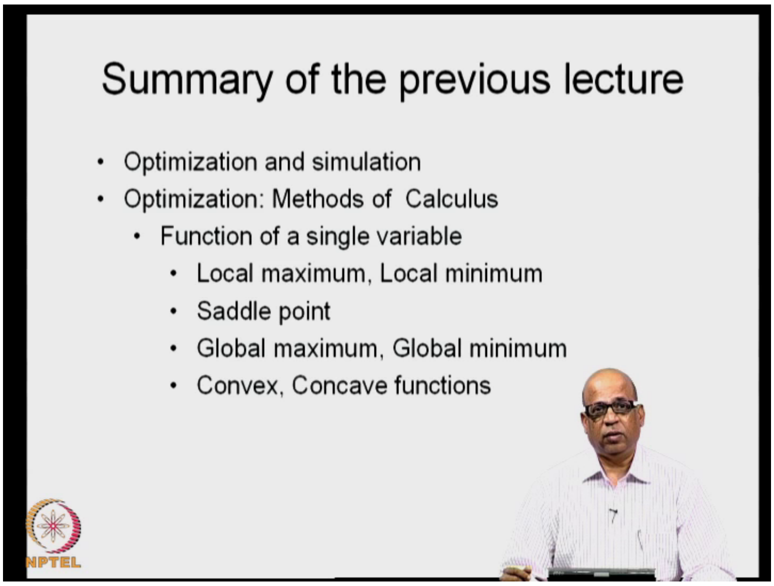
Indian Institute of Science, Bangalore

Lecture No. # 04

Optimization: Functions of Multiple Variables

Good morning, and welcome to this the lecture number 4, of the course, Water Resources System - Modeling Techniques and Analysis. In the previous lecture, we have seen for functions of single variables.

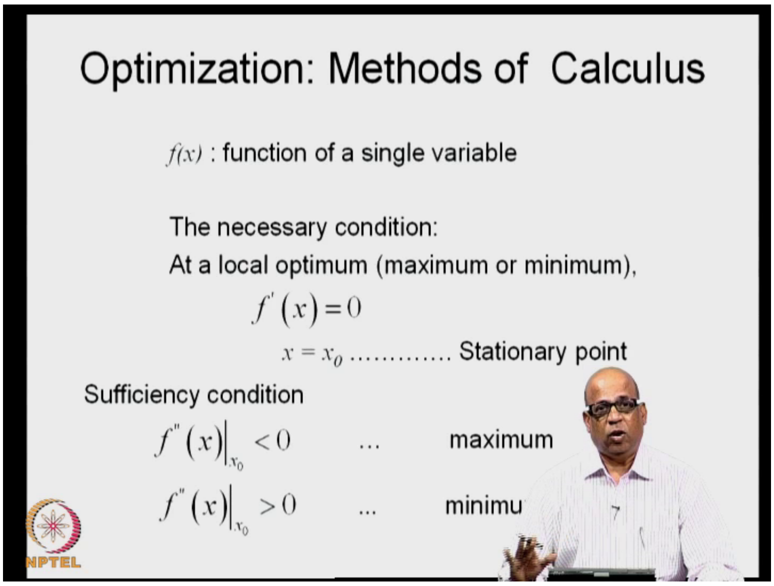
(Refer Slide Time: 00:35)



How we pose the necessary conditions for the optimum values, and we were just about to start with the sufficiency conditions in the previous lecture. So, just look at the summary of what we covered in the previous lecture; we started with the distinction between optimization and simulation, in optimization we are looking for maximum or minimum value of a function. Whereas, in the simulation we are essentially trying to mimic the behavior of a particular system, and simulation is also a very powerful technique especially when we are looking at screening of alternatives in large river basins and so on. Whereas, optimization why you used to get the optimal values of a system performance. Let us say we are talking about optimum hydro power development or optimum flood control and so on. So, in situations where you are really interested in the best values that you can derive out of the system you use optimization.

However in many situations, as I mentioned we may not be interested in single optimum values, but we may be interested in answering questions such as what if types of questions, in which simulation will be a powerful technique. Then we went on to examine functions of single variables, and specifically we talked about local maximum, local minimum, and then the saddle point where the first slope of the function is 0, yet it may not correspond to either a maximum or a minimum value, and then we have a seen the definitions of a convex function and concave function; the in the case of convex function the local minimum also corresponds to the global minimum, in the in the case of concave function - the global the local maximum also corresponds to the global maximum. So, this is what we covered in the previous lecture. Now, we will start with the optimization of functions of a single variable and post on necessary and sufficiency conditions more formally.

(Refer Slide Time: 03:00)



So, you have a function f of x as function of a single variable. Now for this function to have an optimum at a particular point x is equal to x naught; the necessary condition as I have been mentioning is that the slope of that function at that particular point x is equal to x naught must be 0; so, the slope must be 0 that is an necessary condition. So, at a local optimum which is either a maximum or a minimum f dash of x is equal to 0. So, typically what we do is given the function you obtain f dash of x it is the first derivative set it to 0, and solve for x; you may get many several solutions are x 1, x 2, x 3 etcetera, all of its satisfy this f dash of x is equal to 0. And that we define you also stationary prime; so, x is equal to x naught is the stationary point which satisfies f dash of x is equal to 0.

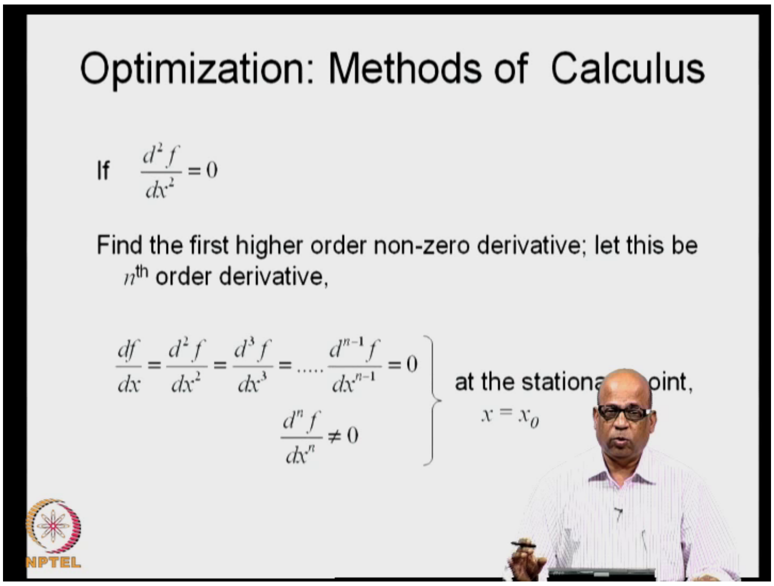
Now, this is a necessary condition; for f of x to have a minimum or a maximum at the stationary point x is equal to x naught, the sufficiency conditions are f double dash of x evaluated at x naught less than 0, if f double dash of x which is the second derivative of the function evaluated at x is equal to x naught, if it is negative then the point x is equal to x naught - x naught corresponds to a maximum value. If f double dash of x evaluated at x is equal to naught is greater than 0 which is positive, then the point x is equal to x naught corresponds to a minimum value which means the function will have a minimum value at x is equal to x naught. So, these are the sufficiency conditions. So, the necessary condition is that the slope must be 0 at that point.

So, what you do is you set f dash of x is equal to 0 solve for x is equal to x naught, you get the stationary point, you may not get just one solution you may (( )) you may get 2, 3, 4, etcetera depending on the nature of f dash of x. At a given x is equal to x naught, you go to the second order derivative, evaluate the second order derivative at x is equal to x naught, and then examine whether f double dash of x is less than 0 at x is equal to x naught, if it is less than 0 at x is equal to x naught the point x is equal to x naught corresponds to a maximum value, if f double dash of x evaluated at x is equal to x naught is greater than 0 or it is positive, then the point x is equal to x naught corresponds to a minimum value.

Now, what if f double dash of x evaluated at x is equal to x naught is also 0; that means, we started with f dash of x is equal to 0, solve for x is equal to x naught, solve for x naught, and then evaluated the second order derivative at x is equal to x naught; now, what if f double dash of x evaluated at x is equal to x naught is also 0. In such case you go to the higher order derivative, get f triple dash x naught, x f triple dash x at x is equal to x naught; that means, the third derivative third order derivative evaluated at x is equal to x naught; what if this is also 0, go to fourth order derivative, evaluate at x is equal to x naught; what if that is also 0 go to fifth order derivative, evaluate at x is equal to x naught; what if that is also 0 and so on. You keep continuing until you get the first non-zero derivative, evaluated at x is equal to x naught.

And then, you look at the order of the derivative. I will state this more formally. So, we use the second order derivative our sufficiency condition, if the second order derivative is also 0, go to the third order derivative, fourth order derivative, fifth order derivative and so on, until you get the first derivative evaluated at x is equal to x naught which is not not zero. And then look at the order of the derivative at which you are getting a non-zero derivative value.

(Refer Slide Time: 03:00)



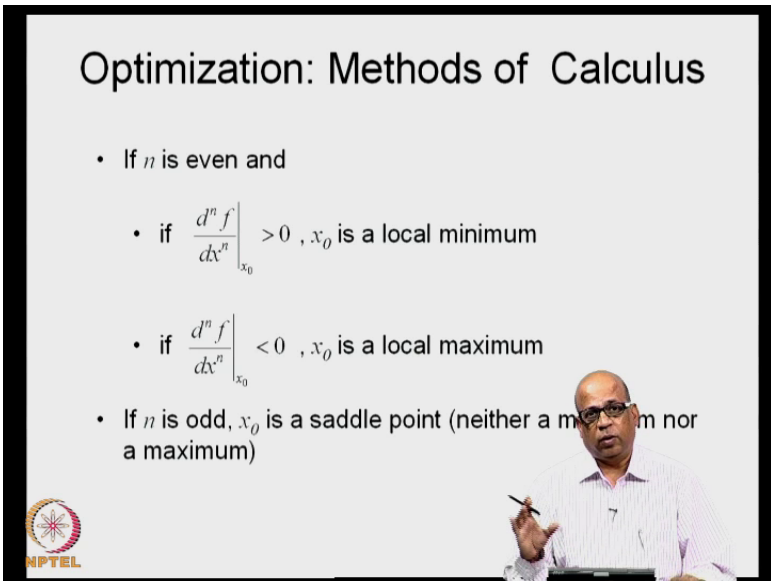
So, if the second order derivative is 0, find the first higher order non-zero derivative. Let this be the n th order derivative, what I mean by that is d f by d x is equal to d square of f by d x square d q f by d x cube etcetera n minus 1 n th order derivative they are all 0, and the n th order derivative is the first non-zero derivative that you are getting when you are evaluating the derivative at x is equal to x naught. Then you look at the order of the derivative; if the order of the derivative is even, and this value of this derivative that you get is negative, then it the point corresponds to a maximum value. If the order of the derivative n is the; so, (( ))… So, obtained is even and the derivative value is positive, then the value the point x is equal to x naught corresponds to a minimum.

Remember always here, the higher order derivative negative always corresponds to maximum; higher derivative positive always corresponds to minimum. This you must keep in mind. So, the same thing we apply you keep going higher and higher orders of the derivative, until you get the first derivative which is nonzero. Then, you look at the order of the derivative; if that order is even and the magnitude of the derivative is negative, then it corresponds to maximum value. If the order is even and the magnitude is odd, I am sorry and the magnitude is positive then the value corresponds to a minimum value.

If the first derivative which is non-zero, and the order of that non-zero derivative is odd; then the point x is equal to x naught neither corresponds to a minimum value nor corresponds to a maximum value. I repeat this again before; so, formally stating it; we use the second order derivative as a sufficiency condition, if the second order derivative is negative then the point x is equal to x naught corresponds to a maximum value. If the second order derivative is positive then the point x is equal to x naught corresponds to a minimum value.

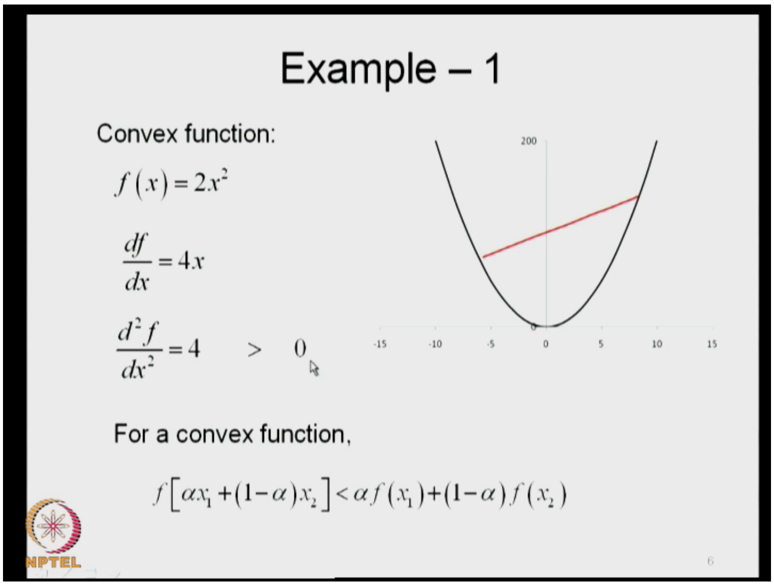
If the second order derivative is 0 then we go to higher order derivative, third order derivative; if it is also 0 go to fourth order derivative that is also 0 go to fifth order derivative and so on. Get the first derivative which is non-zero, then you look at the order of the derivative, let us say this was sixth order derivative which was non-zer0. The order of derivative is 6 which is a even number. So, if the first non-zero derivative is of the order which is a even order, then if the magnitude of the derivative is negative then the point x is equal to x naught which is stationary point corresponds to a maximum value. If the magnitude is positive then the point x is equal to x naught corresponds to a minimum value; if however, the order of the derivative is odd then the point x is equal to x naught corresponds neither to a minimum value nor to a maximum value. So, this what we state formally (( )).

(Refer Slide Time: 03:00)



So, if n is odd, and d n f by d x n which is the n th order of derivative of the function evaluated at x is equal to x naught is positive, then x naught is a local minimum. This is a more informal way of putting it what we mean by that is x naught corresponds to the function having a local minimum at that point or the function has a local minimum at x is equal to x naught. If n is even and d n f by d x n evaluated at x is equal to x naught is negative, then x naught corresponds to a local maximum. If n is odd then x naught is a saddle point which means it is neither a minimum nor a maximum who is call it as a saddle point, because your first derivative is 0 which means the slope is still 0 there, and this is how we make a decision on whether x is equal to x naught corresponds to a minimum or a maximum or neither of that.

(Refer Slide Time: 13:02)

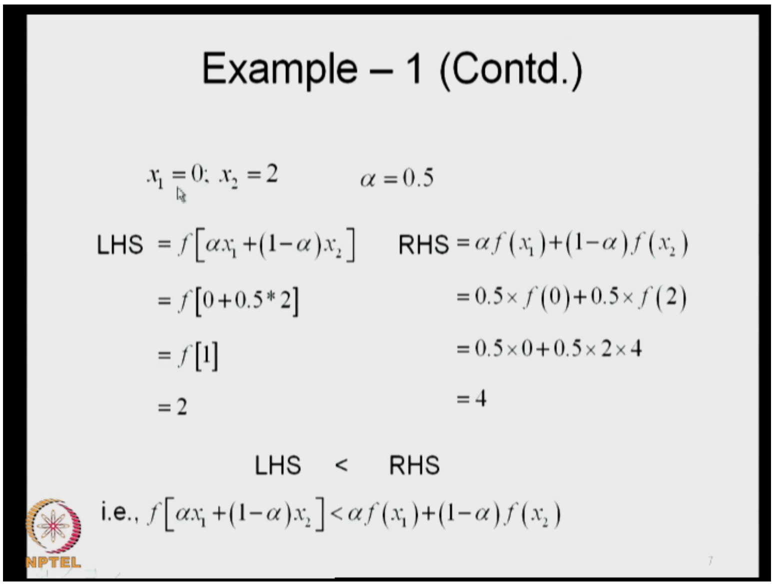


Let us look at some examples now. So, we started with the functions with the definitions of a concave function, convex function and then we enumerated the necessary, and sufficiency conditions for local minima of a local minima, local maxima of a given function and we are talking about functions of single variables. Remember we have also not added any conditions or constraints. So, we are talking about unconstraint optimization of a single variable or functions of a single variable.

So, let us look at a function f of x is equal to 2 x square, if you plot this the function will look something like this, the function will look exactly like this not something, it looks exactly like this. Now, you take the first derivative d f by d x is equal to 4 x, because we are talking about of a single variable, I was the full derivative notation d f by d x this is four x and take the second derivative d square f by d x square which is equal to 4 which is positive which means irrespective of the value of x, the second derivative is always positive.

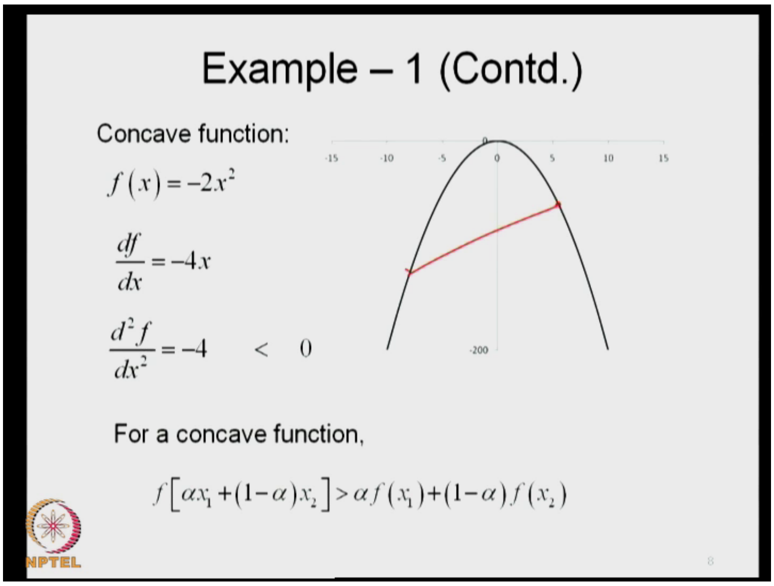
So, this is a strictly convex function. So, this function f of x is equal to 2 x square is a strictly convex function. One of the features of the convex function as we saw in the are previous lecture is that you join any 2 points on that function, that line will be always above the function itself. Now, this we can verify. So, f of alpha x one you take any 2 points. So, this is x. So, I will take x 1 here, and x 2 here. You take any 2 points and choose a value of alpha between 0 and 1 and this condition must be satisfied; that means, f of alpha x 1 plus 1 minus alpha x 2 which means a function value corresponding to that particular point must be less than the point alpha f x 1 plus one minus alpha f x 2 itself; which means that the straight line joining 2 points will be above the curve will be enclosed in the curve in this particular case.

(Refer Slide Time: 13:02)



So, let us seen will choose some x 1 value x 1 is 0 and x 2 is 2 which means I am taking x 1 is 0 and x 2 is 2 somewhere here. So, I am joining those 2 points and I choose alpha is equal to point 5, just to verified. So, the LHS which is this part f of alpha x 1 plus 1 minus alpha x 2, f of x is 2 x square. So, I will use this function at the value alpha x 1 plus 1 minus alpha x 2. So, I get f of 1; so, f of 1 is 2 2 into x square that will be 2, alpha x 1 alpha I am choosing it point 5 and x 1 is 0 1 minus alpha x 2 x 2 is 2; so, this would be f of 1 and that is equal to 2. Similarly, value it take alpha f of x 1 plus 1 minus alpha f of x 2. So, alpha which is point 5 f of x one which is 0 then 1 minus alpha again point 5 f of x 2 is 2; so, I get four. So, LHS less than r h s; so, that will be verified that. This is just a feature of convex functions, remember this not the definition of convex functions - the definition of convex functions is here that is d square f by d x square must be positive for all values of x in that particular range, that is a definition.

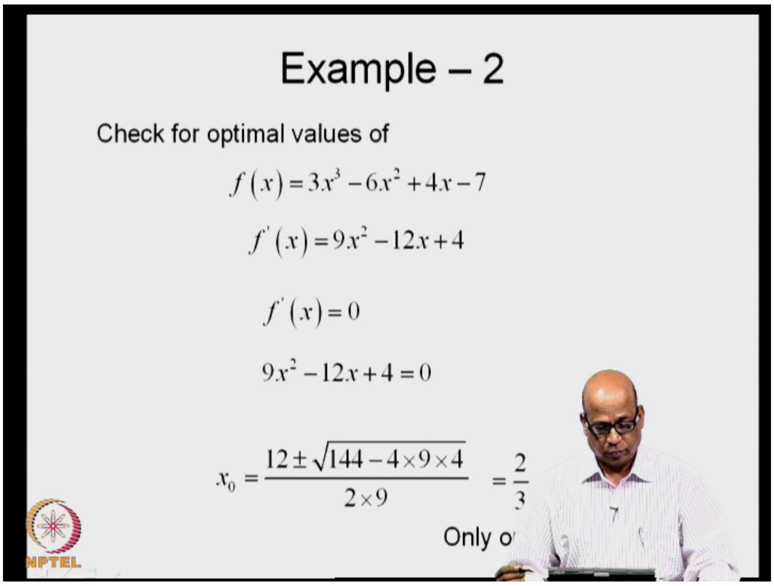
(Refer Slide Time: 13:02)



Let us look at another function. We just take the mirror image of that will say f of x is equal to minus 2 x square. So, the function plots like this. So, if you join any 2 points here in this curve, this would be below that curve, this line will be below that curve and this is what we have seen in the previous class. So, f of x is equal to minus 2 x square and therefore, I take the first derivative that will be minus 4 x and d square f by d x square would be minus 4 which is always negative irrespective of the values of x, the second derivative is always negative. And therefore, you get the global maximum; the local maximum also corresponds to the global maximum. In this case it occurs at point 0.

So, this is how you determine whether a given function is a concave function or a convex function. Now, you can also verify this for the concave function you have f of alpha x 1 plus 1 minus alpha x 2 must be greater than alpha f of x 1 plus 1 minus alpha f of x 2, choose any values of x 1 and x 2 convenient values of x 1 and x 2 choose a value of alpha between 0, and 1 and then verify this. I I want do that, because I just demonstrate that for the convex functions, similar way you can do it for concave function. Now, we will see for given functions how we identify the stationary points and then see whether the stationary point corresponds to a minimum value or a maximum value and so on.

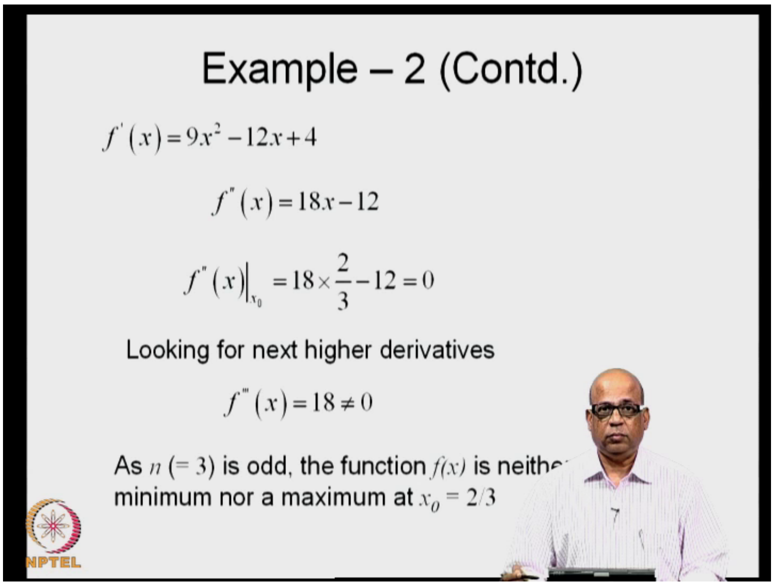
(Refer Slide Time: 19:10)



So, we take of function f of x is equal to 3 x cube minus 6 x square plus four x minus 7. The first condition for this to have optimum values or either minimum value or maximum value at a given point is that the first derivative of the function must be equal to 0. So, we take the first derivative f dash of x and that will be 9 x square minus 12 x plus 4 and this must be equal to 0. So, f dash x equal to 0. So, 9 x square minus 12 x plus 4 is equal to 0. You solve for x using this and that we call it as x naught. So, the solution for f dash of x is equal to 0 we call it as x naught, and this we obtain it as 2 by 3. Now, in this particular case all though we had quadratic quadratic equation, we get only one solution, but in general you may get 2 solutions in that situations; if it is the polynomial of order 3 you may get 3 solutions or less than 2 of them may be common and so on.

So, you solve f dash of x is equal to 0 obtain as many solutions as it yields, and then corresponding to each of these solutions of a x naught; you determine f double dash x that is the second order derivative. So, what it did we do, we took the first order derivative equated (( )) 0, solve for that equation get the stationary points. In this particular case you obtained only one stationary point, but there are in more generally you will get several such stationary points; that means, x naught is equal to 2 by 3 may be 1 of this solution, x naught is equal to 0 may be another solution x naught is equal to minus 1 may be another solution and so on. So, you may get several stationary points at each of the stationary points you have to examine whether the function corresponds to a minimum or a maximum or neither of them. So, in this case we will examine at x naught is equal to 2 by 3 whether the function corresponds to a minimum or a maximum or neither of them; how do we do this we go to the second order derivative and then evaluate the second order derivative at x naught is equal to 2 by 3.

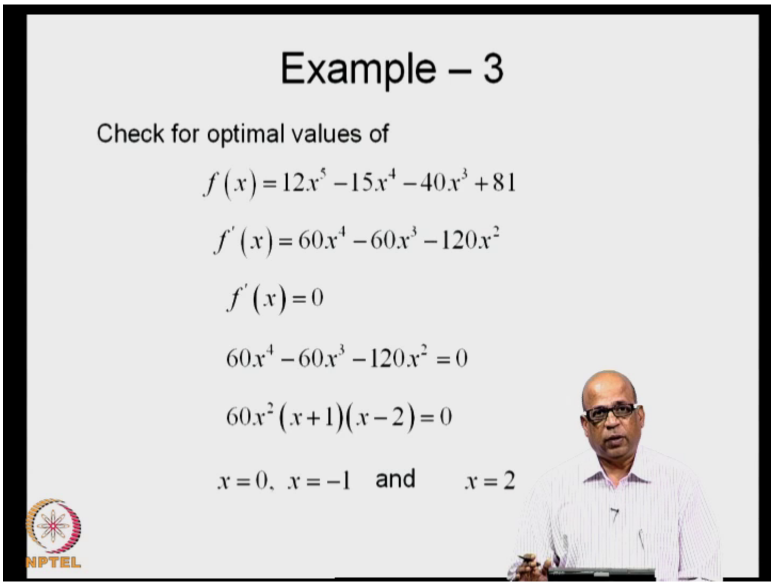
(Refer Slide Time: 19:10)



So, f dash of x is 9 x square minus 12 x plus 4, we take the second order derivative which means we differentiate this f dash of x again with respect to x therefore, f double dash of x will be 18 x minus 12. And this second order derivative we evaluate at x is equal to x naught and x naught is 2 by 3. So, f double dash of x we evaluate at x is equal to 2 by 3 and these terms have to be equal to 0. As I said you take the first order derivative equated to 0 obtain the stationary point go to the next derivative next order derivative which is second order derivative and evaluate the second order derivative at the stationary point which is 2 by 3 a which is x is equal to 2 by 3 in our case, and that terms out to be 0.

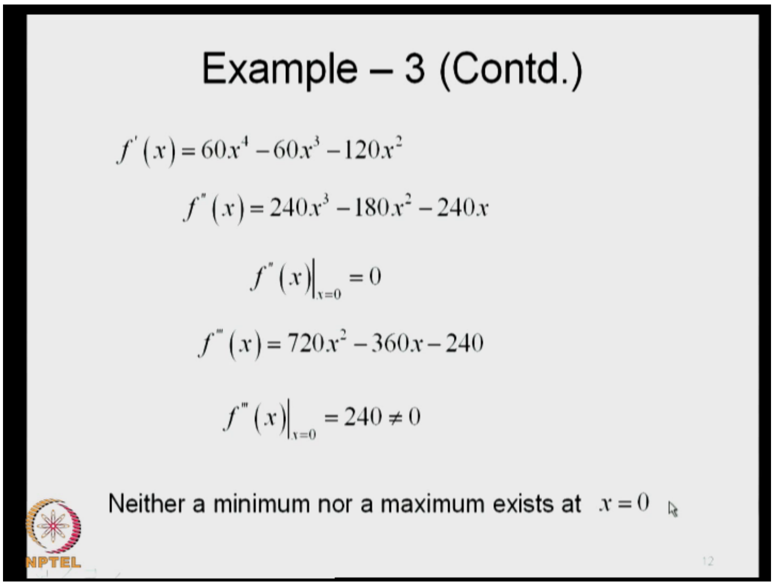
If the second order derivative terms out to be 0, go to the next order derivative which is the third order derivative and in that case the third order derivative terms out to be 18 which is not equal to 0. Once you obtain the first non-zero derivative, you look at the order of the derivative, in this particular case the order of the derivative is 3 and because the order of the derivative is odd, the point x is equal to 2 by 3 which is a stationary point neither corresponds to a maximum nor a minimum. So, that is the conclusion here; as n is equal to 3 is odd n being that order of derivative for which you are getting a non-zero derivative value, the function f of x is neither a minimum nor a maximum at x naught is equal to 2 by 3. So, this is the conclusion that you derived.

(Refer Slide Time: 23:45)



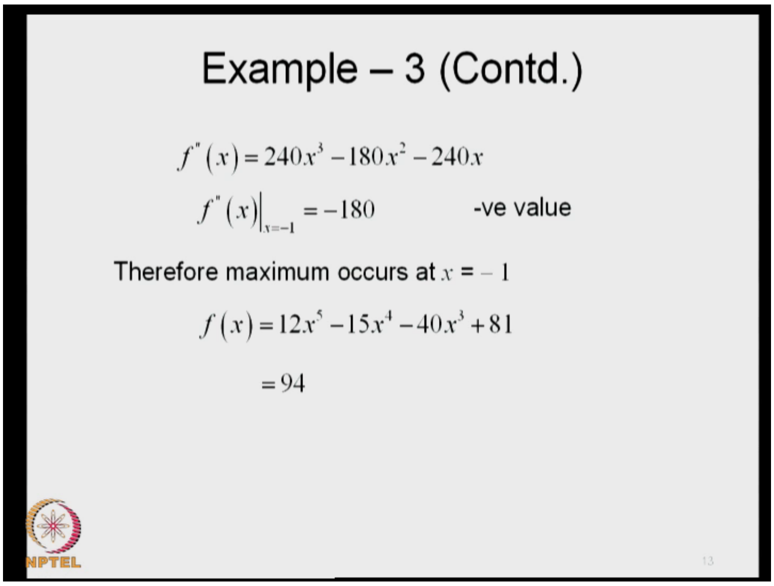
Let us look at another example, where we look at a function 12 x to the power 5 minus 15 x to the power 4 minus 40 x to the power 3 plus 18 81. We take the first order derivative f dash of x and equated to 0. So, 16 x to the power 4 minus 60 x cube minus 120 x square equated to 0. And you get such a equation; so, 60 x square you take out x plus 1 into x minus 2 is equal to 0. So, you get solutions x is equal to 0, x is equal to minus 1 and x is equal to 2. So, you got 3 stationary points here, all of which satisfy the equation f dash of x is equal to 0. We examine for local minimum or local maximum or existence of neither of them at each of these stationary points. So, you got 3 stationary points corresponding to each of the stationary 3 stationary points we examine whether the function has a local minimum or local maximum or neither of them. So, we start with x is equal to 0.

(Refer Slide Time: 23:45)



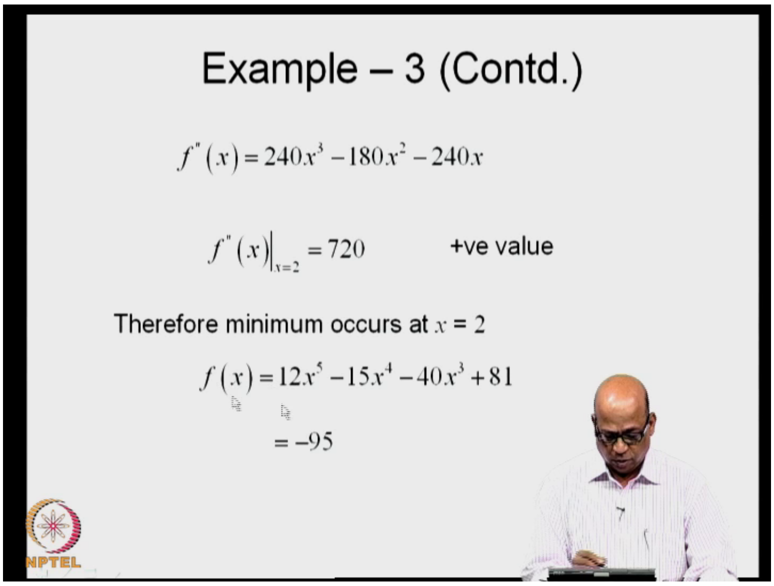
So, you got f dash of x from that you get the second order derivative f double dash of x. So, f double dash of x in this case will be 240 x cube minus 180 x square minus 240 x. This second order derivative you evaluate at x is equal to x naught; let us say x is equal to 0. So, at x is equal to 0, because of this nature it terms out to be 0 itself. Because the second order derivative terms out to be 0; you go to the third order derivative still evaluated at x is equal to x naught. So, third order derivative at x is equal to x naught terms out to be 240 which is non-zero, and order is 3 which is odd and therefore, the point x is equal to 0 corresponds neither to a minimum nor to a maximum, because the order is odd here. So, that we exhausted for x is equal to 0.

(Refer Slide Time: 23:45)



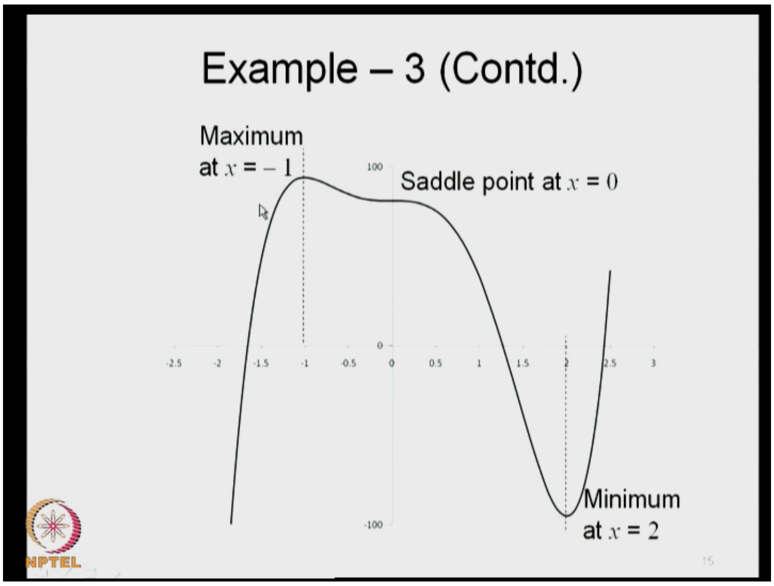
Then we take x is equal to minus one, remember we got 3 solutions here x is equal to 0, x is equal to minus 1, and x is equal to 2. We exhausted x is equal to 0 then we go to x is equal to minus 1. So, at x is equal to minus 1 we evaluate f double dash of x and this terms out to be minus 180. So, f double dash of x evaluated at x is equal to minus 1 is negative and therefore, it corresponds to a maximum value. So, maximum value occurs at x is equal to minus 1; once we determine this then you also evaluate the maximum value of the function by putting x is equal to minus 1. So, in the f of x which is the original expression for the function, you put minus 1 and you get f of x is equal to 94. So, the maximum value of f of x which is a local maximum of f of x occurring at x is equal to minus 1 is 94. So, we obtained x is equal to 0, x is equal to minus 1, we we exhausted x is equal to 0 and x is equal to minus 1.

(Refer Slide Time: 23:45)



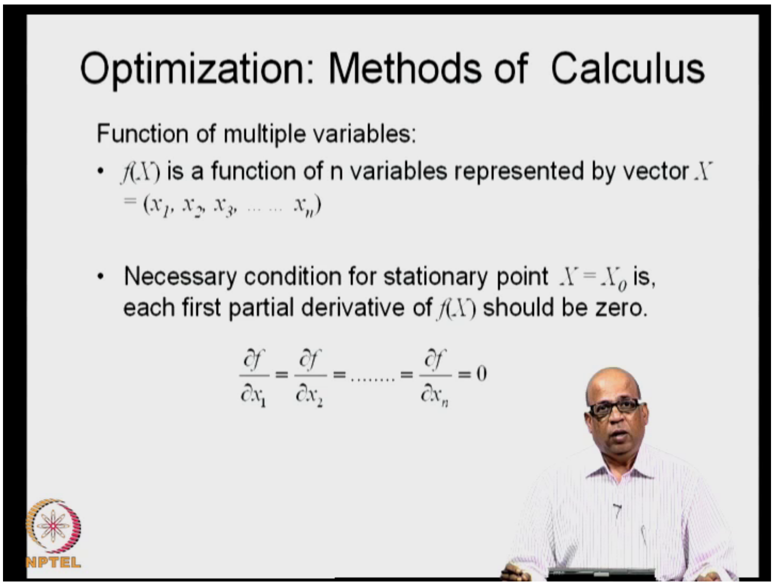
Next we go to x is equal to 2; determine the second order derivative evaluated at x is equal to 2 and see what happens at x is equal to 2. So, f double dash of x we are determined as this expression, we substitute x is equal to 2 which is the third stationary point and then obtained f double dash of x is equal to 720 which is the positive value. Because the order is even which is the second order derivative, and it terms out that the magnitude is positive and therefore, the value x is equal to x naught which is x is equal to 2 in this particular case corresponds to a minimum value. So, therefore, minimum occurs at x is equal to 2, and we determine the actual minimum value as f of x is equal to 12 x to the power 5 minus 15, we substitute x is equal to 2 in the original value of (( )) function and that is the minimum value that you get which is minus 95.

(Refer Slide Time: 23:45)



And if you plot the function, the function plot looks like this. So, at x is equal to 0 you obtained the saddle point. So, the slope is 0, but it neither corresponds to a local minimum nor a local maximum. So, this is a saddle point at x is equal to 0; at x is equal to minus 1 there is the slope of 0 and it corresponds to a maximum value, because your second derivative was negative. At x is equal to 2, you get the slope as 0 and you get a minimum value, because the second derivative was positive. Remember these are all local these are here you get a local maximum, local minimum and this is the saddle point. So, all are at all of these places the slope is 0 and therefore, we are able to obtain the local minimum and local maximum in this locations.

(Refer Slide Time: 29:36)



So, we just completed functions of single variables first quickly recapsulate what you did power functions of single variables, we know how to identify whether the function is a convex function or concave function; convex function has a global minimum and therefore, f double dash of x is positive there; concave function has a global maximum and therefore, f double dash of x negative term. So, for a function of a single variable if the second order derivative is positive for all values of x in the particular range, then it corresponds to a convex function, because you are talking about a minimum value.

Then we went on to state the necessary conditions for a function to have a local minimum at point x is equal to x naught, the necessary condition is that at that particular point the slope must be 0 which means f f dash of x must be equal to 0. So, typically what we do is given a function you take the first order derivative equated to 0 solve for a x is equal to x naught, you may get several solutions at each of the is each of this solution which are called as stationary point, you examine whether the point corresponds to a local minimum or a local maximum by looking at the second order derivative, third order derivative and so on. And we we are now in position to say whether the stationary point corresponds to either local minimum or a local maximum or neither of them in which case it will be a saddle point.

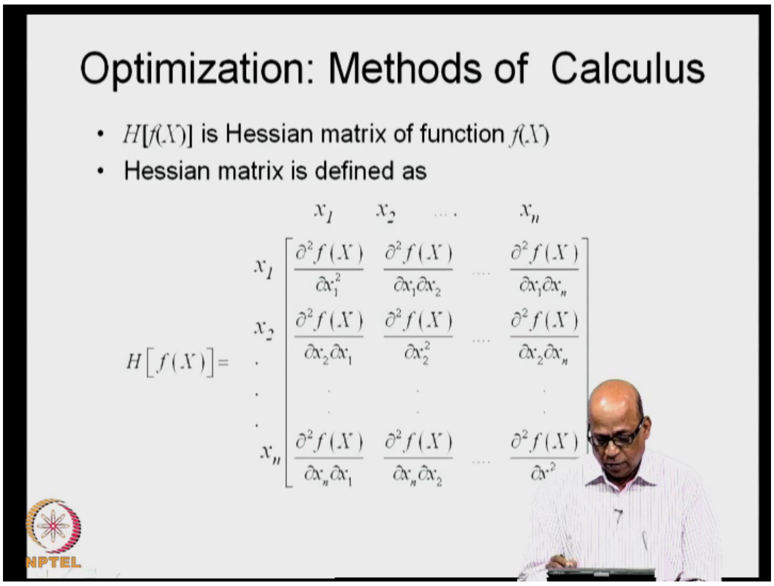
Now, we will generalize this procedure for functions of multiple variables. What we did in the case of single variable, the same principal we apply to functions of multiple variables also. Let us say that you have a function of n variables f of x where x is the vector of x 1, x 2, x 3, etcetera n variable x n. The necessary condition is just for excitation of what we did for the single variable case that is the first order derivative of the function with respect to each of the variables x 1, x 2, x 3, etcetera, x n must be equal to 0. So, we say d f by d x 1 why use the partial derivative, because there are n number of variables d f by d x 1 is equal to d f by d x 2 is equal to etcetera d f by d x n with respect to each of the variables the first order derivative with respect to each of the variables must be equal to 0, there is an necessary condition.

So, let us write it more formally round. So, you have functions of multiple variables and we denote it has f of x we use capital x to denote that, this is the vector f of x is a function of n variables represented by vector x; x 1 is equal to x 1, x 2, x 3, etcetera, x n. So, the necessary condition for stationary point x is equal to x naught is each first order partial derivative of f of x should be 0. So, d f by d x 1 is equal to d f by d x 2 etcetera, d f by d x n is equal to 0. So, given f of x you obtain the first order partial differentials with respect to each of the variables. So, you get n number of equations, solve for the n number of equations, get the stationary point x is equal to x star or x is equal to x naught capital x is equal to x naught. This is the stationary point between what you get x 1 naught, x 2 x 2 naught, etcetera x n naught, as the stationary points.

Then we go to the sufficiency conditions; what we do in the case of single variable, we went to the second order derivative. Let us say you have function of 2 variables; x 1 and x 2 - just look at 2 variables, x 1 and x 2. How many second order derivatives would be there, you will have 3 second order derivative d square f by d x 1 square, d square f by d x 2 square and d square f by d x 1 d x 2. So, you will have 3 second order derivatives.

So, unlike in the first in the case of single variable, you will have many second order derivatives in the case of multiple variables; in case of single variable you had exactly one second order derivative and therefore, you variable to make a decision based on the second order and higher order derivatives whereas, the number of variables increases the higher order differential derivatives will be more than 2 and therefore, the therefore, you need to construct matrices of second order derivatives to decide whether the stationary point that you so, obtain x is equal to x naught corresponds to a minimum or a maximum or a neither of them. So, we formulate what is called as the hessian matrix. The hessian matrix is a matrix of the second order derivatives, and we make our decisions based on the hessian matrix.

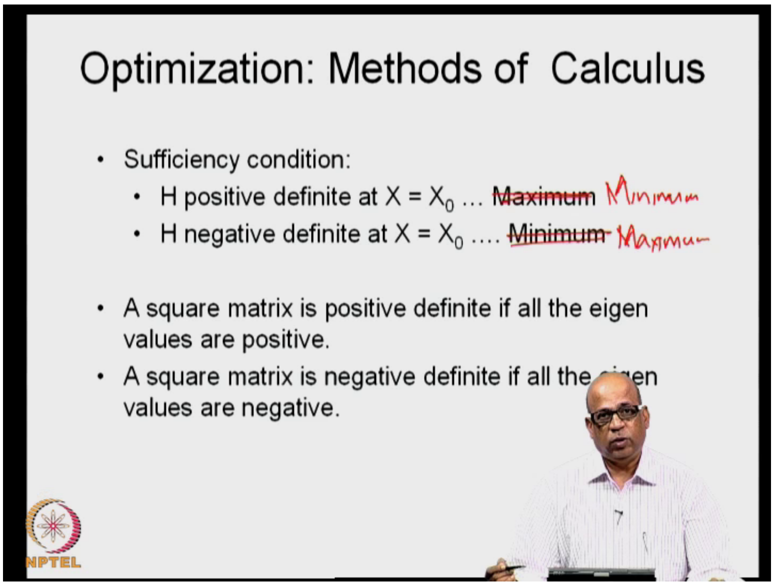
(Refer Slide Time: 29:36)



How do we formulate the hessian matrix, the h of a given function h of f of x in the particular case is a hessian matrix of function f of x; we define the hessian matrix as the matrix which is the n by n size, n being the number of variables in this case of second order derivatives. So, we are write x 1, x 2, etcetera, x n here; x 1, x 2, etcetera x n, here. Take the second order derivative d square f of x by d x 1 square, there is x one with respect to x 1, x 1 with respect to x 2 which is d square f by d x 1, d x 2 then x 1 and x n d square f by d x 1 d x n like this for example, here you get d square f by d x 2 x 1; the d square f by x 2 x 2 that is the x 2 square x 2 x n like this you formulate the second order derivatives; for x n it will be d square f by d x 1 d x 2 and so on. So, you get a n by n matrix of the second order derivatives.

Now, this is the hessian matrix. The hessian matrix has to be evaluated at x is equal to x naught. So, this hessian matrix which so formulate, we evaluate this hessian matrix at x is equal to x naught; what is the x is equal to x naught? That is the stationary point that we obtain here. So, you obtain the stationary point at x is equal to x naught, formulate the hessian matrix evaluate the hessian matrix at x is equal to x naught, and then make your decisions of whether the point x is equal to x naught, corresponds to a global a local minimum or a local maximum; why we will do that now.

(Refer Slide Time: 29:36)

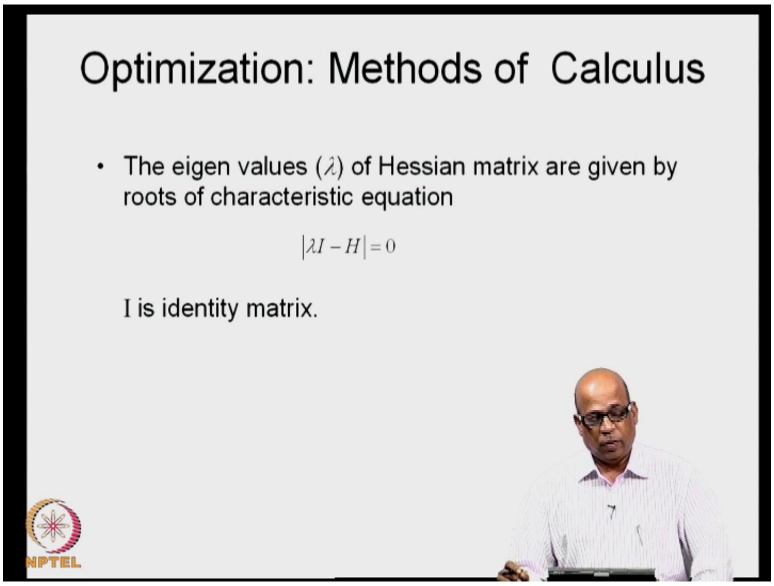


So, the sufficiency condition is if the hessian matrix that you show determine just now and evaluated at x is equal to x naught, if the hessian matrix is positive definite; there is a mistake here. I will just convert this, if the hessian matrix is negative definite at x is equal to x naught then it corresponds to a minimum value. And if it is negative definite at x is equal x naught then it corresponds to a maximum point. So, I repeat you formulate the hessian matrix which is the n by n matrix of the second order derivatives. Evaluate the hessian matrix at x is equal to x naught. If the hessian matrix is positive definite then the point corresponds to a minimum and if it is the negative definite it corresponds to a maximum.

Now, I use the terms positive definite and negative definite; how do we determine whether a given matrix is positive definite and now or or it is the negative definite; of we define the matrix to be positive or positive definite or negative definite only for a square matrices. So, if a square matrix has its eigen values all of which are positive, then it is called as the positive definite matrix; that means, if all the eigen values are positive then the square matrix is a positive definite matrix. If all the eigen values are negative then it is the negative definite matrix. If some eigen values are positive and some are negative then it is the neither of positive definite nor a negative definite.

So, we make decisions at x is equal to x naught based on whether the h matrix or the hessian matrix evaluated at x is equal to x naught is in fact, of positive definite matrix or a negative definite matrix. If it is neither of them then, the point x is equal to x naught which is the stationary point corresponds to neither a minimum nor a maximum.

(Refer Slide Time: 29:36)

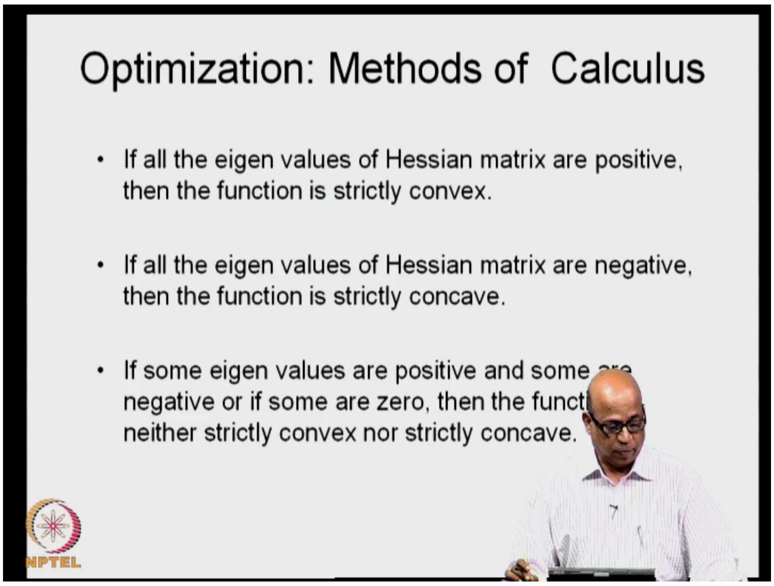


And then I I said that the hessian matrix is positive definite, if all the eigen values are positive. So, we must know how to determine the eigen values; I hope all of you know how to determine, this is the simple characteristic equation lambda I minus H, H is the a symmetric determinant of that you set it to the 0 and then solve for lambdas. So, the lambdas are though eigen values when we solve the numerical examples it will become (( )).

So, in the case of multiple variable, in the case of functions of multiple variables we get the first order derivatives equated to 0, handle if you have a n variables you have n first order derivatives and therefore, you have n equations solve for the n equations you get the stationary point, capital x is equal to capital x naught at the stationary point you evaluate the hessian matrix - hessian matrix is the n by n matrix of the second order derivatives.

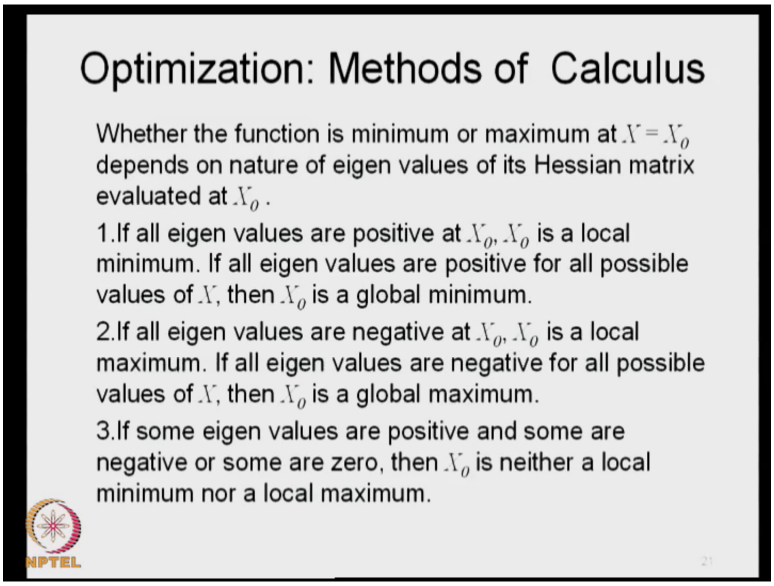
You evaluate the hessian matrix at the stationary point x is equal to x naught examine whether at x is equal to x naught, the hessian matrix is positive definite or negative definite. If the hessian matrix is positive definite the point x is equal to x naught corresponds to a minimum; if the hessian matrix is negative definite the x is equal to x naught corresponds to a maximum. To determine whether the hessian matrix is positive definite or negative definite you use the eigen values; if all the eigen values are negative, then the hessian matrix is negative definite matrix, if some of them some of the eigen values are positive, some of them are negative then the hessian matrix is neither positive definite nor negative definite; in which case the stationary point x is equal to x naught neither corresponds to a maximum value nor corresponds to a minimum value. So, this is what we do in the case of multiple functions of multiple variables.

(Refer Slide Time: 29:36)



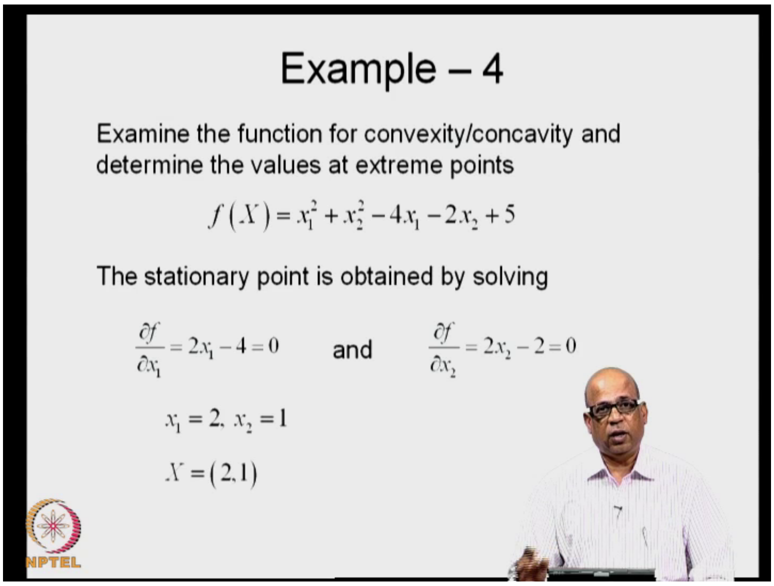
Now, similar to what we did in the case of functions of single variables; let us say that the hessian matrix which is the matrix of the second order derivatives is positive definite irrespective of a values of x at are using which means the entire range of the entire range of capital x over which the functions has been defined, the hessian matrix is positive always positive. Then the function is strictly convex. So, the multiple function of multiple variables is strictly convex which means the hessian matrix being positive, irrespective value of a values of the variables x, then the function is strictly convex in which case the local minimum also corresponds to the global minimum. If all the eigen values are negative, then the functions corresponds to function is strictly concave which means the local maximum is also is local global maximum; if some over eigen values are positive and some are negative as I said it neither corresponds to it is neither convex nor concave in that particular range.

(Refer Slide Time: 29:36)



We will just summarized this whatever I just said whether the function is minimum or maximum at x is equal to x naught depends on nature of eigen values of its hessian matrix evaluated at x is equal x naught; that is the hessian matrix is evaluated at x is equal to x naught. If all eigen values are positive at x naught, x naught is a local minimum this is what we solve. If all eigen values are positive for all possible values of x then x naught is a global minimum, because that is corresponds to a convex function. So, if if the function is convex then the local minimum, there is a local minimum also corresponds to the global minimum. So, that is what you mean here. If all eigen eigen values are negative at x naught, x naught is the local maximum; then if all eigen values are negative for all possible values of x, then x naught is a global maximum. Similarly, if some of them are positive, some of them are negative where you are able to make decision on whether it is a local minimum or I am sorry you are able to stay that the point x is equal to x naught corresponds neither to a local minimum nor to a local maximum.

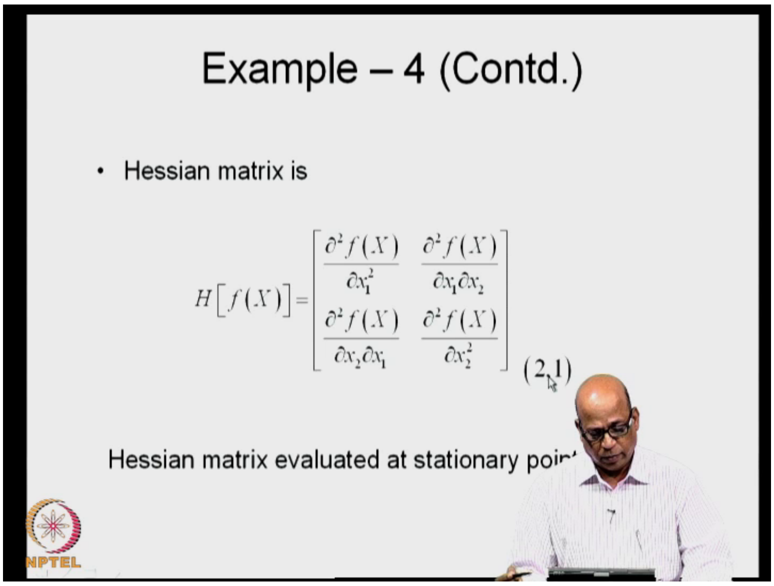
(Refer Slide Time: 44:47)



Let us look an example of multiple variables. We take 2 variables now. So, f of x is equal to x 1 square plus x 2 square minus 4 x 1 minus 2 x square 2 x 2 plus 5. So, the first step in optimizing in obtaining the optimal values of functions of multiple variables is to take the first derivative with respect to each of the variables; you have 2 variables x 1 and x 2 corresponding to each of the 2 variables, we take the first derivatives. So, d f by d x 1, I use the notation for partial derivatives d f by d x 1 is equal to 2 x 1 2 x 1 minus 4, I am differentiating this function with respect to x 1, and set to 0 2 x 1 minus 4 is equal to 0.

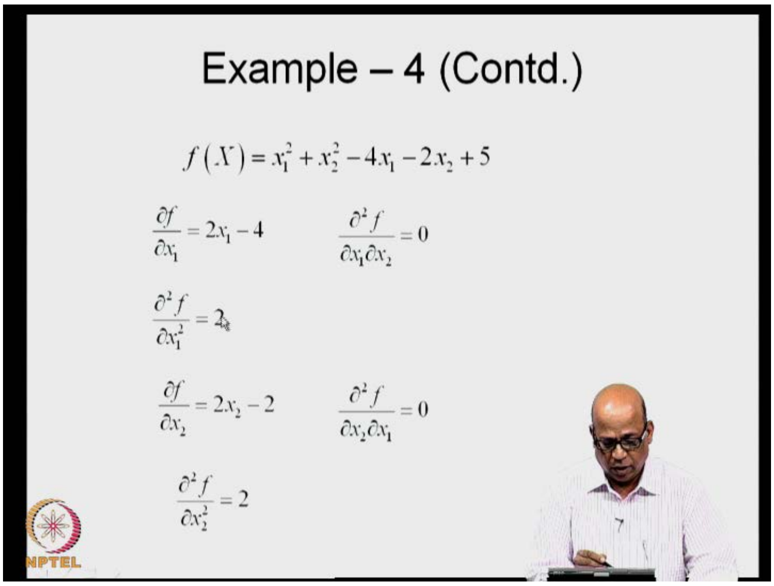
Similarly, I differentiate this function with respect to x 2 that will be 2 x 2 minus 2 is equal to 0, and by solving this I get x 1 is equal to 2 and x 2 is equal to 1. So, the stationary point I obtained it as (2,1) that is the x 1 is equal to 2 and x 2 is equal to 1, this is the stationary point. Now, corresponds to this the stationary point I will now evaluate the hessian matrix for the function. So, formulate the hessian matrix which will be a 2 by 2 matrix of the second order derivatives.

(Refer Slide Time: 44:47)



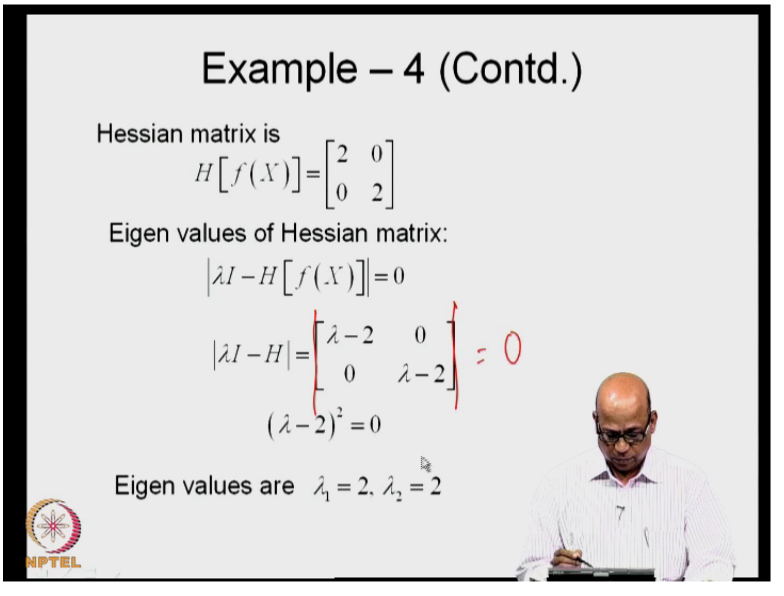
Let us look at the hessian matrix. There is a x 1 here, x 2 here, x 1 here, x 2 here. So, we have d square f by d x 1 square, d square f by d x 1 d x 2, d square f by d x 2 d x 1 and d square f by d x 2 square; these 2 will be the same derivatives. This has to be evaluated at (2,1) which is the stationary point that you have obtained. So, we formulate the hessian matrix and evaluate it at (2,1).

(Refer Slide Time: 44:47)



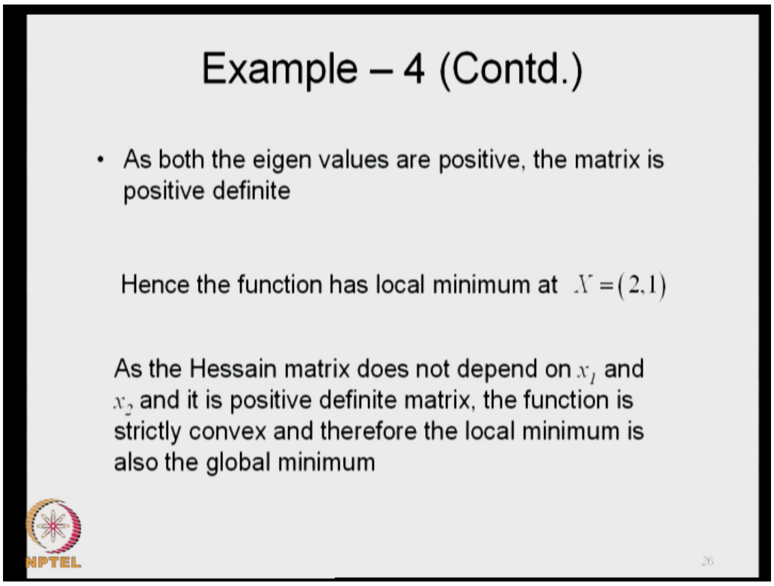
So, f of x is this d f by d x 1 is 2 x 1 minus 4 and therefore, d square f by d x 1 square is we are differentiating this with respect to x 1 and therefore, we get d square f by d x 1 square is equal to 2, d square f by d x 1 d x 2 you differentiate this with respect to x 2 that till 0, and you have got d f by d x 2 as 2 x 2 minus 2 and d square f by d x 2 d x 1 will be 0 similar to this d square f by d x 2 square will be equal to 2. What happens in this case is the irrespective of the values of x 1 you are second order derivative is 2 here; irrespective of the values of x 2 your second derivative is 2 here; and irrespective of the values of x one and x 2 your second order derivatives with respect to x 1 and x 2 or 0 here. So, we will examine whether this is the positive definite matrix or an negative definite matrix; first and see whether this remains irrespective of the values of x 1 and x 2. So, the hessian matrix is 2 0 0 2 these are the values, d square f by d x one square and d square f by d x 2 square and d square f by d x 1 d x 2.

(Refer Slide Time: 44:47)



So, the hessian matrix is 2 0 0 2, let us determine the eigen values of this which means I take the determinant lambda I minus H is equal to 0. So, H of f of X, but we can typically write it as H itself. So, we will write lambda I minus H which will be lambda minus 2 0 0 lambda minus 2; that is what we get here. So, you get lambda minus 2 the whole square is equal to 0 this is the determinant, here this is determinant and this we are setting it as 0. So, lambda minus two the whole square is equal to 0. So, we get the solutions for lambda as lambda 1 is equal to 0 and lambda 2 is equal to 2.

(Refer Slide Time: 44:47)



Which means the both the eigen values that you obtained or in facts same lambda is equal to 2 if what you got both the solutions are positive and therefore, the H matrix is positive definite. And therefore, the solution that we obtained that is the stationary point that we obtained namely x is equal to (2,1). In fact, corresponds to a local minimum further why we set local minimum, it is because the H matrix is positive definite, because it is positive definite it corresponds to a minimum value and therefore, the stationary point that we just obtained namely (2,1) is in fact, a local minimum. Further, because the h matrix is positive definite irrespective of the values of x, remember here we got this matrix (2,0), (0,2) irrespective of the values of x that we substitute.

If this where a function of x 1 and x 2 then you could have substituted at x 1 is equal to 2, x 1 x 2 is equal to 1 and then obtained the h matrix evaluated at that point, but because this is independent of x 1 and x 2 and we have a obtain this matrix to be positive definite, it means that H remains to be positive definite irrespective of the values of x and therefore, the function is a convex function. Because the function is concave function the local minimum that you obtained at this particular point x is equal to (2,1) is also a global minimum.

So, as the hessian matrix does not depend on x 1 and x 2, and it is positive definite matrix the function is strictly convex and therefore, the local minimum is also the global minimum. Let us look at I hope this problem is understood correctly what was essentially did was that we obtain the stationary point by taking the first order derivatives, there were 2 variables in the case and therefore, you get two equations when we take the first order derivative, solve for the the variables x 1 and x 2; this consecutive stationary point - at the stationary point you evaluate the hessian matrix - the hessian matrix in this particular case will be a square matrix of order two; this hessian matrix has to be evaluated at the stationary point, and then examine if this is the hessian matrix if the hessian matrix is positive definite or negative definite.

If a hessian matrix is positive definite, then the stationary point corresponds to a minimum value; if the hessian matrix is negative definite, then the stationary point corresponds to a maximum value. In the in the particular example that we examine the hessian matrix is remained positive definite irrespective of the values of x 1 and x 2; which means all the values of x 1 and x 2 hessian matrix remains positive definite. And therefore, the function corresponds to a a function is a convex function and therefore, the local minimum that you obtain at the stationary point is in fact, the global minimum.

The same thing is valid for concave function also where the stationary point corresponds to a local maximum and therefore, the local maximum also corresponds to a global maximum, if the function is a concave function. We will see some are examples where we may get a concave function or we may get neither a concave nor a convex function in the entire range and therefore, we may get some stationary points corresponding to a local minimum, some corresponds to corresponding to a local maximum and so on.

That we will examine throws some other examples in the next lecture. So, to summarize in today lecture, essentially we have started with functions of single variables and seen the necessary and the sufficiency conditions for functions of single variables to have a local minimum or a local maximum at a given point. So, what we do is we take the first order differential, first derivative we take and equated to 0, solve for that and you get the stationary point.

Now, at the stationary point the examine the higher order derivatives, second order derivative, if the second order derivative is negative then the stationary point x is equal to x naught corresponds to a maximum; if the second order derivative is positive then it corresponds to a minimum; or I again repeat always associate positive values with minimum, that is positive values of the second order derivative or the higher order derivative where we are able to make a decision positive always to corresponds to a minimum, negative always corresponds to a maximum.

And if it is if the second order derivative was also 0 then you go to higher order derivative, third order, fourth order, fifth order, etcetera, and get the first derivative which is non-zero, then look at the order. If the order is odd then this stationary point is neither a minimum nor a maximum. If the order is even then you then you check the magnitude of that particular derivative, if the magnitude is positive it corresponds to a minimum, if the magnitude is negative it corresponds to a maximum. We saw some examples, where we got a combination of these cases where you can you can identify a local minimum or local maximum and we are also able to stay that the point corresponds to neither a minimum nor a maximum.

Then we went on to examine functions of multiple variables where essentially the same principle wholes you first take the first order derivative equal to 0, if there are n number of variables you get n number of equations, solve for this n variables using these n number of variables that defines the stationary point, capital x is equal to x naught where x is the vector of n variables then you formulate the hessian matrix hessian, matrix is the n by n matrix n by n square matrix of second order derivatives, evaluate the hessian matrix at x equal to x naught which is the stationary point; if at x equal to x naught your hessian matrix is a positive definite matrix, a point of x equal to x naught corresponds to a minimum; if the hessian matrix evaluated at x is equal to x naught is the negative definite matrix then the point x is equal to x naught corresponds to a maximum.

Always associate positive with minimum values negative with maximum values. When you are making their decisions based on the derivatives. Then, we also saw that if the hessian matrix remains positive definite illustrative of the values of the x, that will use which means in the entire range of the definition of the function, if the hessian matrix remains positive definite then the function corresponds to a convex function and therefore, the local minimum also corresponds to the global minimum; local minimum is also the global minimum.

Similarly, if the hessian matrix remains the negative in the entire range of a function definition then the local maximum corresponds to the global maximum. So, we will continue this discussion on multiple variables in the next class, I will cover another two examples and that completes the unconstrained optimization. Remember the class of optimization technique that we are dealing with or unconstrained optimization, because we are not putting any conditions, we are just stating a function and for that function to have a local minimum local maximum, global minimum global maximum, we are setting out the conditions for that; when we start put constrains on to this, then the problems becomes much more complicated. We will see those in the subsequent classes, thank you for your attention.