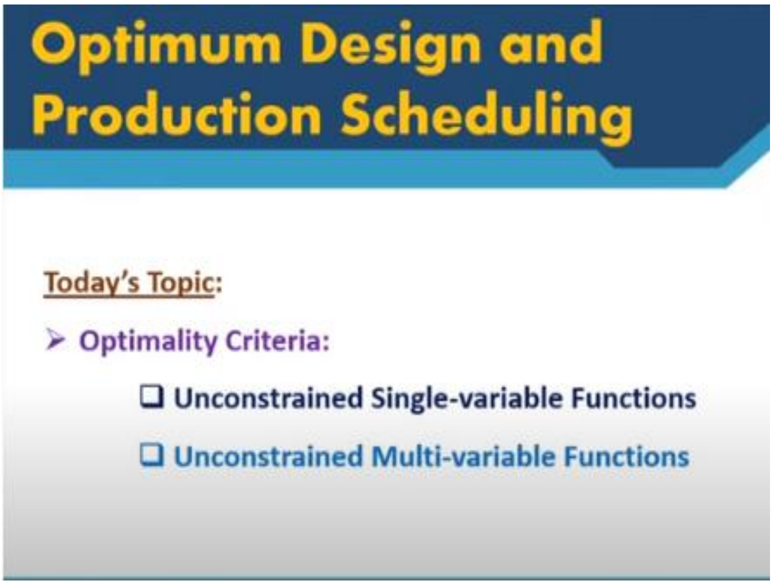
Plant Design and Economics

Prof. Debasis Sarkar
Department of Chemical Engineering
Indian Institute of Technology-Kharagpur

Lecture - 56
Optimality Criteria for Unconstrained Functions

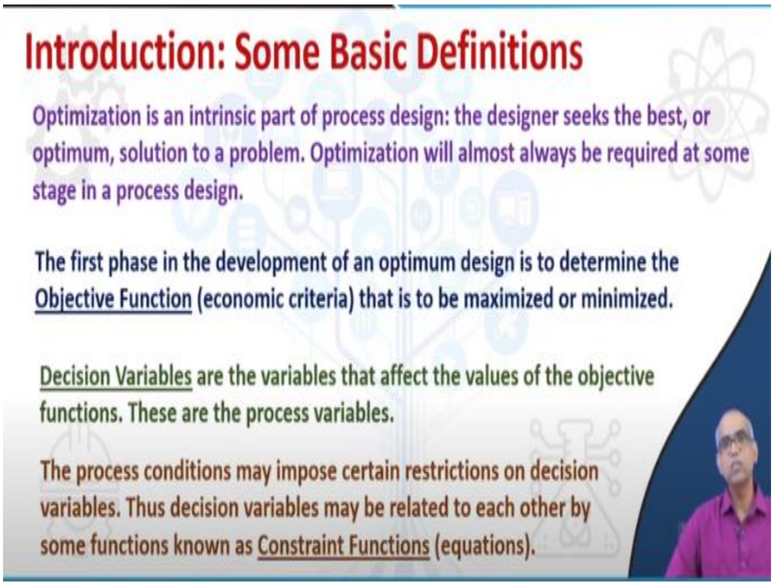
Welcome to module 12 of Plant Design and Economics. In this last module of the course, we will talk about optimum design and production schedule. Essentially, we will look at briefly the scope of optimization in process design. So in this first lecture of this module, we will talk about optimality criteria for unconstrained functions.

(Refer Slide Time: 00:36)



We will talk about optimality criteria for both unconstrained single variable function as well as unconstrained multivariable functions.

(Refer Slide Time: 00:49)



Optimization is an intrinsic part of process design. While designing, the designer will try to obtain the best or the optimal solution to a problem. Now when we try to obtain the optimal solution to a problem, we must have a criterion to judge whether the solution at hand is best or not. So while the designer seeks to obtain the best or optimum solution, the designer evaluates the solution in hand using certain criterion.

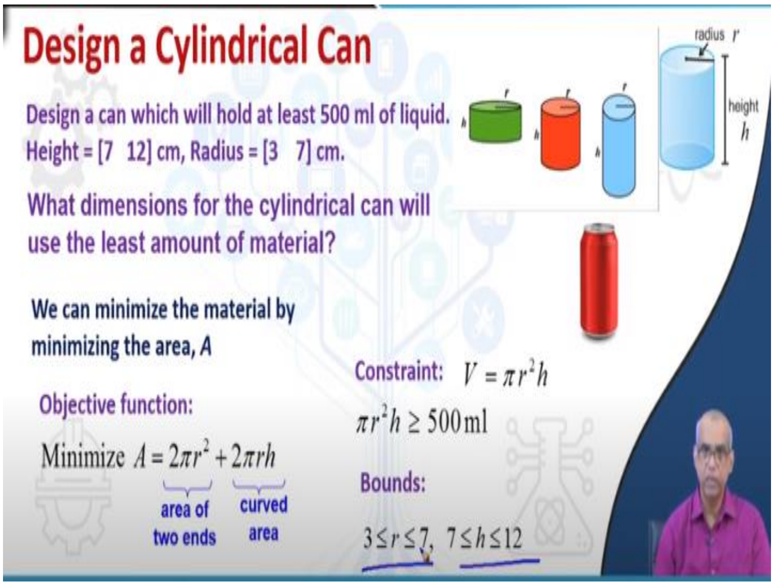
So this criterion is known as objective function, and oftentimes is an economic criterion. So the designer will try to maximize or minimize an economic criterion, which we call objective function. This objective function is basically a function of certain variables known as decision variables. So these are process variables. These are the variables that you would like to have the optimal values.

And these optimal values for the decision variables will maximize or minimize the objective function. The process conditions may impose certain restrictions on these decision variables, and thus, the decision variables may be related to each other by some functions, which we call constraint functions or constraint equations. For example, I am designing a cylindrical can and I say that volume of the can will be exactly equal to say 100 ml.

So you know the volume of a cylindrical tank is pi r square into h, where r is the diameter of the can and h is the height. So in this case, if you are trying to find out which r and which h will be best for the designing of the cylindrical can, this r and h are decision variables and they are related as pi r square h equal to 100 ml. So every optimization problem will have an objective function that let us call that as an economic criterion.

Oftentimes, it is an economic criterion. And then a set of decision variables whose optimal value we are trying to arrive at and then the process can impose certain restrictions on the decision variables. The decision variables cannot take any values they want. Then, these restrictions that the decision variables will satisfy will be expressed by a set of equations known as constraint functions or constraint equations.

(Refer Slide Time: 04:13)



Let us again look at that design of cylindrical can example. So we want to design a cylindrical can which will hold at least 500 ml of a liquid. We also specify that height of the can will be between 7 cm to 12 cm and the radius of the can will be between 3 cm to 7 cm. Note that there may be various values of r and h which may satisfy this restriction that the volume will be enough to hold at least 500 ml of liquid.

So you have to find out that r and that h which will definitely hold at least 500 ml of liquid but will also have minimum cost for this cylindrical can. Now what will be a measure of this minimum cost? We can minimize the material required. So that will minimize the construction cost or fabrication cost. And you can minimize the material by minimizing the area of the cylindrical can.

So the area of the cylindrical can will be expressed as a function of radius and height of the can and that will give me the objective function. Note that the area of the can,

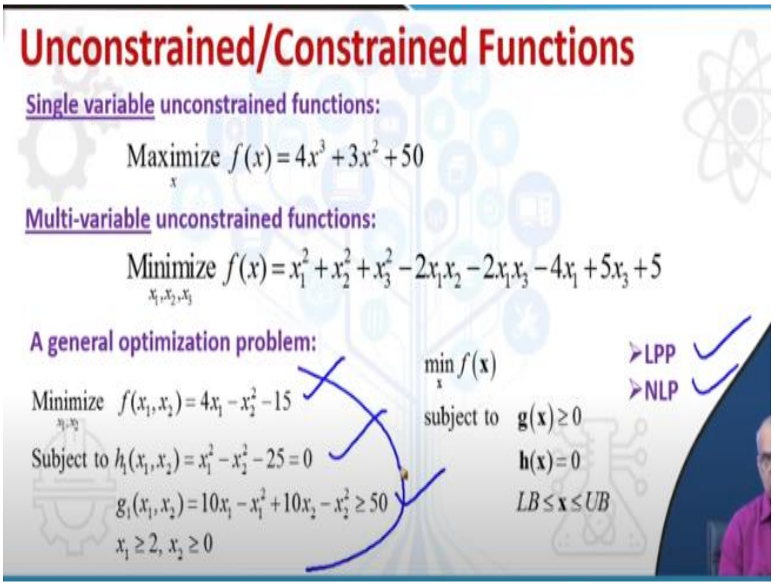
can be expressed as a function of r and h where the curved area is 2 pi r h and also the area of two ends at the bottom and at the bottom are pi r square plus pi r square is 2 pi r square. So this is the objective function.

Here we would like to minimize the objective function, so that we are minimizing the area. And by minimizing the area, we are minimizing the amount of material required. We are assuming that we have we are we have been given a metal sheet of constant thickness. So in that case, minimizing area and minimizing the volume or amount of material are all equivalent. Now is there a constraint?

Yes there is a constraint because we are saying that the can must hold at least 500 ml of liquid. So V equal to pi r square h represents the volume of the can. So pi r square h must be greater or equal to 500. So that is the constraint. Now the decision variables r and h are not free to take any values. In certain forms they can take, they can take any values, we will come to that later.

But in this particular case, we have seen that the height must be bounded between 7 and 12 and the radius must be bounded between 3 and 7. So these are the bounds on the decision variables. So this represents an optimization problem for the design of a cylindrical can. So we see the components as objective function, decision variables and constraints. There are simple bounds which is also one type of constraint.

(Refer Slide Time: 07:57)



So if the objective function is such or the optimization problem is such that the decision variables are free to take any values we call those problems as unconstrained problems or unconstrained functions. Now these functions can be single variable functions. It can also be multivariable functions. So here you see an example of single variable unconstrained function where we just maximize 4 x cube + 3 x square + 50.

But do not restrict the value of x. So x can take any values in the real space. The next one is a multivariable unconstrained function. It is the function of three variables. We minimize this function that means, we have to find out the value of x 1, x 2, x 3 such that this function takes on minimum value. Again the values of x 1 and x 2, and x 3 are not restricted.

So such problems are known as unconstrained functions or unconstrained problems. But a general optimization problem will be constrained problems. And these constraints as we say that they are functions of the decision variables. They can be of equality type, they can also be of inequality type. So you have a general formulation of an optimization problem, where you have this objective function.

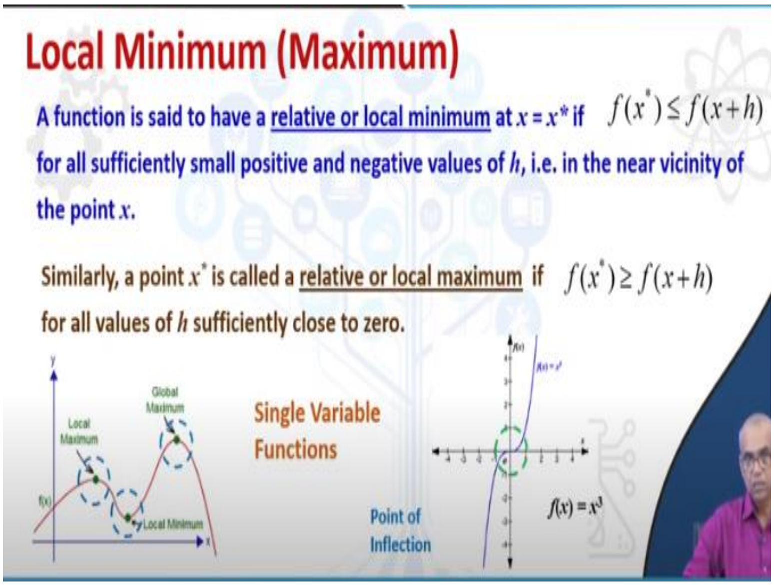
Here it is shown for two variables, it can be of n variables. You have a constraint which is of equality type. Note that x 1 square - x 2 square - 25 = 0. And then you have an inequality constraint. Here it is greater or less than 50, the right hand side. And also x 1 greater or equal to 2, x 2 greater or equal to 0. So this is a general problem. So this is written where there is only one equality type constraints and there is only one inequality type constraints.

But in general there can be any number of equality constraints and there can be any number of inequality constraints. Now here a note on the classification of such general optimization problems. In case the objective function is linear in decision variables and also all constraints are linear, we call that kind of optimization problem as linear programming problem.

So in case of linear programming problem, you have the objective function linear and all the constraints are also linear functions of decision variables. However, in case we have objective function nonlinear or the constraints are nonlinear or both are nonlinear, we have a nonlinear programming problem. For example, this is a nonlinear programming problem.

But, had this been linear, had this been linear and had this been linear would have got a linear programming problem. There is a special class of problem known as quadratic programming problem, where the objective function is quadratic, but all the constraints are linear.

(Refer Slide Time: 12:01)



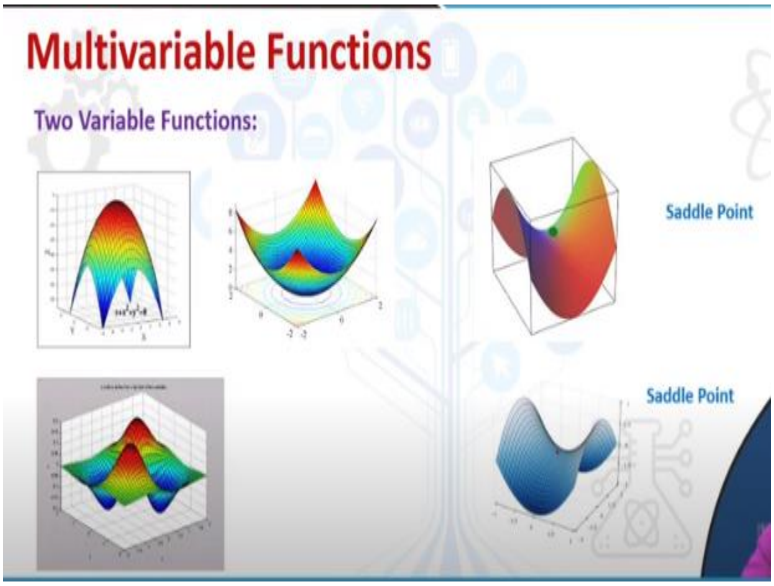
Now let us define what we mean by minimum or maximum or local minimum or local maximum. A function is said to have a relative or local minimum at x equal to x star if the value of the function at x star is small is minimum compared to value of the function at any other point in the neighborhood of x star or in the vicinity of the x star.

Similarly, a point x star will be called a relative maximum or a local maximum if the value of the function at x star is greater or equal to the value of the function at any other point in the neighborhood of x star. So these neighborhoods are shown here. The neighborhoods can be small. It can be sufficiently small. So for single variable functions I have plotted safe x versus x and you see these two represents local maximum.

And it happens that there are two maxima here and these local maxima also happens to be global maxima. And this is a minimum. So in the neighborhood of the point shown these values, the function values are minimum or maximum. Now consider a function f x equal to x cube and consider its value at x equal to zero. So the neighborhood is shown.

At that point if x equal to zero, the function x cube neither attains maximum nor attains minimum. Such points are known as point of inflection.

(Refer Slide Time: 14:14)



So we talked about single variable function. Here there are images of two variable functions or multivariable functions. The simplest multivariable function is two variable function. Note this is represents maximum, this represents the function minimum whereas this and this if you look at these shown points, that dots, at those points the functions neither attains minimum nor at its minimum. These are saddle points.

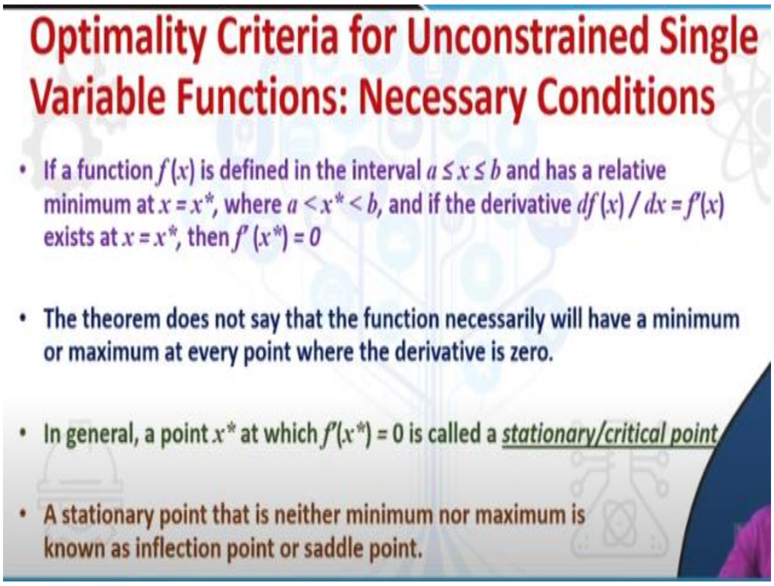
So these are equivalent to point of inflection for single variable function. So we have maximum, we have minimum, we also have point of inflection or saddle point.

(Refer Slide Time: 15:21)

Now let us define necessary conditions for optimality and sufficient conditions for optimality. The conditions that must be satisfied at the optimum point are called necessary. If a point does not satisfy the necessary conditions it cannot be optimum. But remember satisfaction of necessary conditions does not guarantee optimality of the point. So necessary conditions are those that must be satisfied at the optimum point.

If the point does not satisfy necessary conditions, it cannot be optimum. But if it satisfies, it does not guarantee that that the point will be optimum. So that guarantee will come through sufficient conditions. If a candidate optimum point satisfies the sufficient condition, then it is indeed an optimum point. If the sufficient condition is not satisfied or cannot be used, we will not be in a position to draw any conclusion about the optimality of the point.

(Refer Slide Time: 16:28)



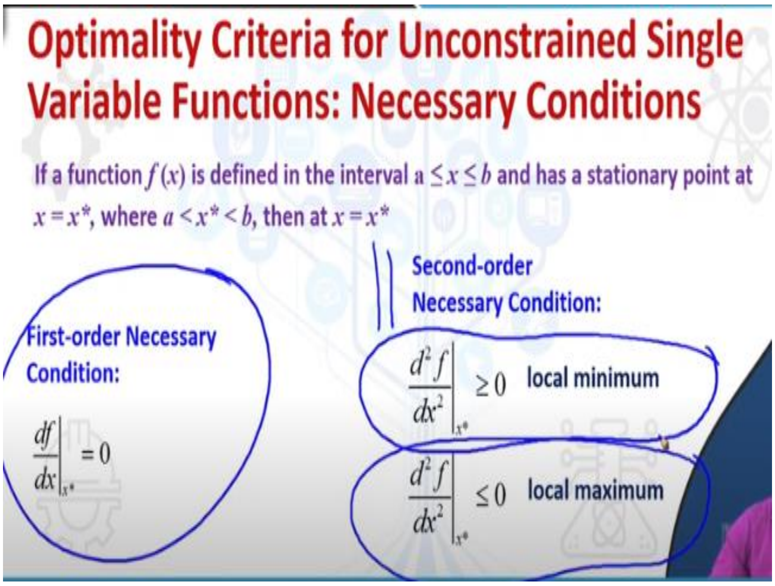
Now let us write down the optimality criteria for unconstrained single variable functions. First, the necessary conditions. Consider a function f x in the interval a to b. We say that it has a relative minimum at x = x star. So x star lies between a and b. Now if the derivative of this function at x star exists and if x star is a relative minimum, we will have the derivative of the function evaluated x star equal to zero.

So that is the necessary condition. So at x star the function value will be zero. The problem is, this theorem does not say that the function will necessarily have a minimum or maximum at every point where the derivative is zero. If the function attains a minimum or maximum at a point x = x star, the derivative of the function at evaluated x star will be zero.

But, the theorem does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero. In general, a point x star at which the derivative of the function is zero we call the stationary point or the critical point. The stationary point can be minimum or can be maximum, but it can also be point of inflection or saddle point.

So if the function attains, if the derivative of the function at x star is zero, we cannot say this is minimum or maximum. It can be minimum, it can be maximum, it can be point of inflection or saddle point.

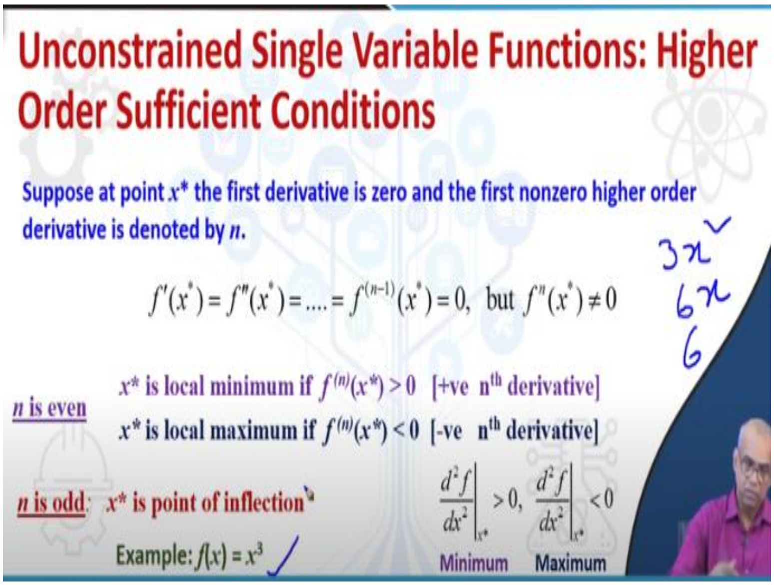
(Refer Slide Time: 18:43)



So the necessary condition is this that the derivative of the function at x star will be equal to zero. So that is the first order necessary condition. The second order necessary conditions for local minimum is that the second order derivative evaluated x star will be greater or equal to zero. Second order necessary conditions for local maximum will be the second order derivative is less or equal to zero.

So first order necessary condition the derivative of the function evaluated x star will be equal to zero. Second order necessary condition if x star is local minimum the second order derivative is greater or equal to zero. If x star is local maximum the second order derivative is less or equal to zero when evaluated at x equal to x star.

(Refer Slide Time: 19:49)



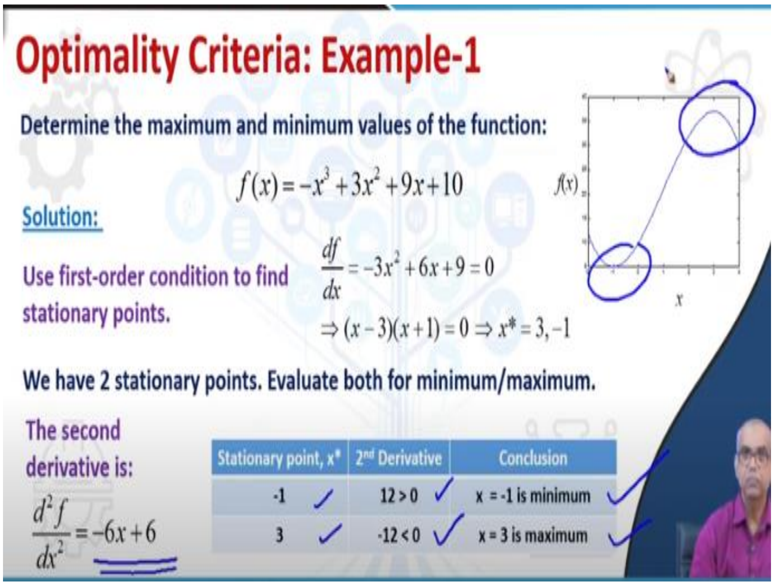
Now let us consider that at point x star the first derivative is zero df dx evaluated x star equal to zero. Also the second order derivative is zero, third order derivative is zero, so on and so forth. And the first nonzero higher order derivative is denoted by n. So first nonzero higher order derivative is denoted by n. Now this n can be either even or it can be odd.

Now if n is even, then the x star is local minimum if that derivative is positive. And x star is local maximum, if that derivative is negative. So in terms of second order derivative, second order derivative so it is even, n equal to even; n equal to 2, equal to even. So if the second order derivative is greater than zero, then the point x star is minimum. If the second order derivative is negative at x equal to x star, the point x star is maximum.

Now let us consider the other case where the first nonzero higher order derivative n is odd. In that case, the point x equal to x star represents a point of inflection. For example, let us say f x equal to x cube. So the first derivative is 3 x square. So 3 x square equal to zero will give you x equal to zero. So x equal to zero let us say, we are evaluating x equal to zero, which is a stationary point, is minimum or maximum. So evaluate.

So this df dx is equal to zero. Then second order derivative, which is 6x. So that is also zero at x equal to zero. Third order derivative is 6 which is greater than zero. But, this is third order derivative, so n equal to 3 equal to odd. So x equal to zero for function f x equal to x cube is neither minimum nor maximum, it is a point of inflection.

(Refer Slide Time: 22:18)

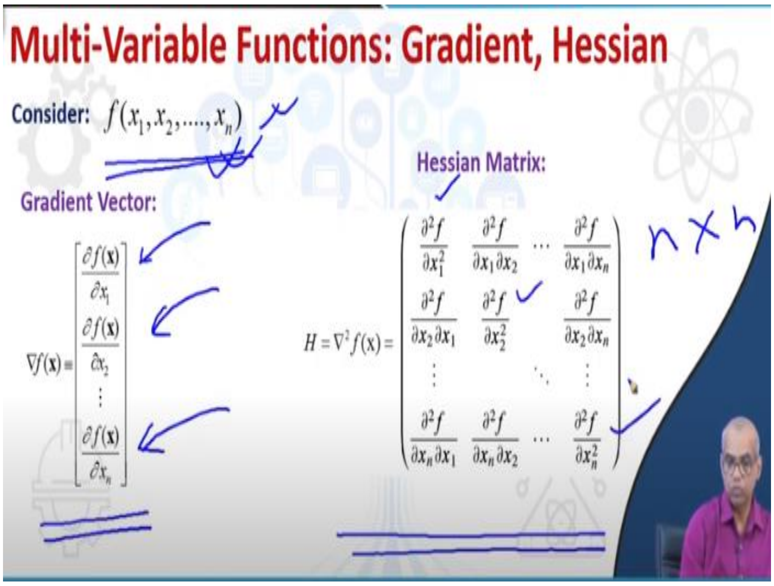


So let us take a quick example. Determine the minimum and the maximum values of the function f x = - x cube + 3x square + 9x + 10. We have plotted the function is a single variable function. So you can plot. You can clearly see that x equal to -1 and x = 3 minimum and maximum of the function. So let us use so we must get these values when we apply the conditions of optimality for this function.

So let us use first order condition to find stationary points. So take derivative of this function set that equal to zero and we obtain x star equal to 3 and -1. Now we have to evaluate whether x star = 3 is minimum or maximum. Same for x star = -1. So for that we must take help of second order or higher order conditions. Let us take second derivative which is d square f d x square = -6x + 6.

Now put x = -3, put x = 3 and x equal to -1 in these expression for second order derivative. When x = -1, the second derivative takes 6 + 6 12 which is greater than zero. So it is minimum, x = -1 is minimum. When x star = 3, we have -18 + 6 which is -12, the value of the second derivative is -12 which is less than zero. So x = 3 is maximum.

(Refer Slide Time: 24:15)

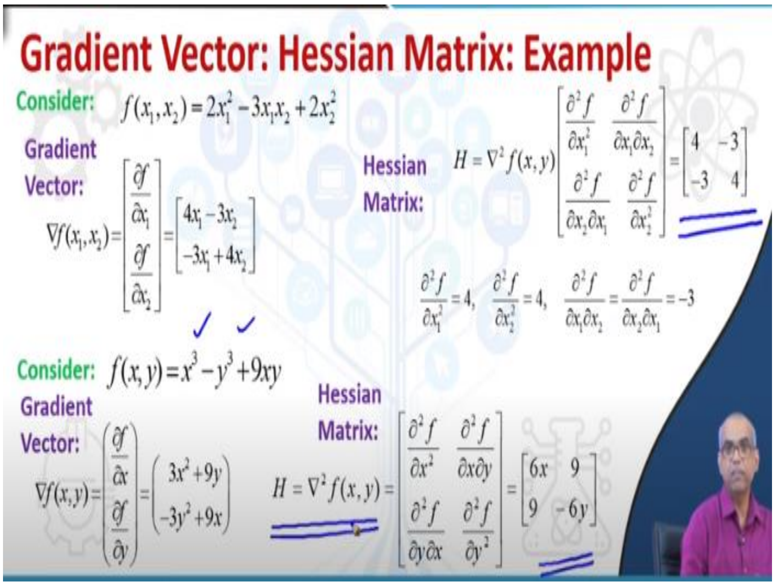


Now we will look at the optimality criteria for multivariable functions. Now in case of single variable functions, we talked about the first order derivative equal to zero as the first order necessary conditions for optimality. Here we will use the gradient vector because it is a multivariable function. So we will consider the gradient vector. And the gradient vector will be equal to zero will be the first order necessary conditions.

What is gradient vector? So consider an n variable function, so the gradient vector will be the vector of all first order derivatives of this function with respect to each decision variables. So the gradient vectors elements will be del f del x 1, del f del x 2, del f del x 3 up to del f del x n. Now the higher order conditions, second order condition which was represented by second order derivative for single variable function.

For a multivariable function, we will use Hessian matrix. This is the Hessian matrix of n variable function. Note that this is a symmetric matrix. So for a function with n variables, this will be an n cross n symmetric matrix. The elements are all second order partial derivatives of this function with respect to its decision variables. So the diagonal elements will be like del square f del x 1 square, del square f del x 2 square so on and so forth.

(Refer Slide Time: 26:22)

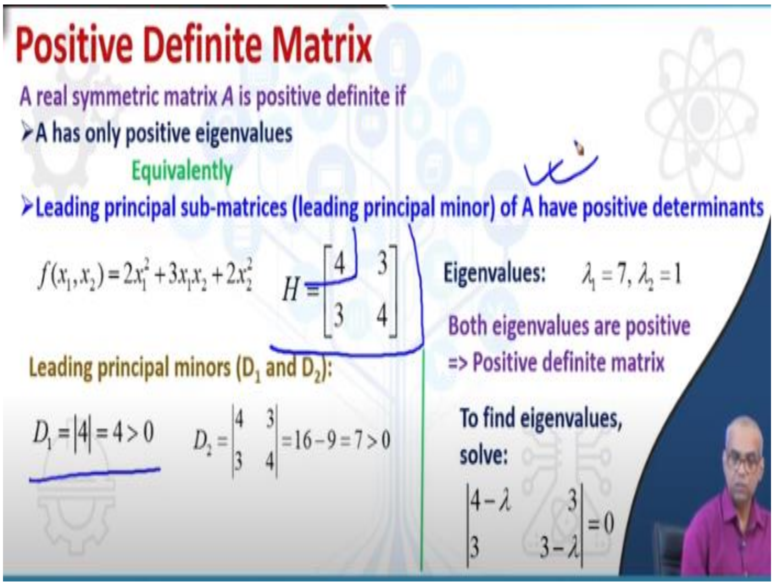


For example, let us consider this function f = 2x 1 square – 3x 1 x 2 + 2x 2 square, a two variable function. So what is the gradient vector? Del f del x 1, del f del x 2. So this represents the gradient vector. Del f del x 1 equal to we evaluate 4x 1 – 3x 2 and del f del x 2 will be -3x 1 + 4x 2. So this represents my gradient vector. And what will be Hessian? So this is a two variable function. So Hessian matrix will be a two cross two symmetric matrix.

So the elements will be del square f del x 1 square, del square f del x 1 del x 2, del square f del x 2 del x 1 and del square f del x 2 square. Now their symmetric matrix this and this will be same. So you just evaluate these quantities and put it there and then you will have this as Hessian matrix. So this is how you can evaluate the elements. Let us consider another function x cube – y cube + 9xy.

Evaluate the gradient and evaluate the Hessian. Now you see, in case of the previous function which was a quadratic we got the Hessian as a matrix with constrained numbers. But here the Hessian is a function of x and y, this x and y. So the Hessian is a function of x, y. For quadratic function it will be a constraint matrix.

(Refer Slide Time: 28:15)

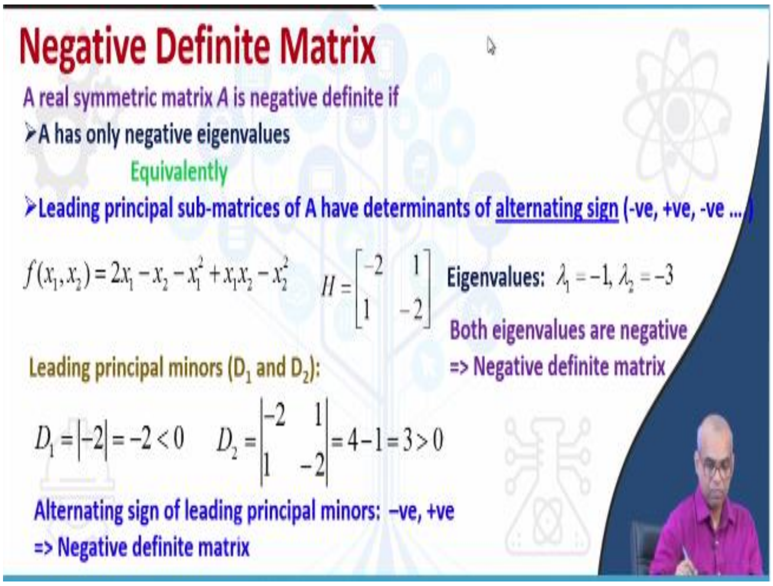


Now we have to define two, three terms here. Positive definite matrix, negative definite matrix and indefinite matrix. A real symmetric matrix which is a Hessian matrix actually, a real symmetric matrix A is positive definite if it has only positive eigenvalues or equivalently the leading principal sub-matrices of the matrix A have positive determinants. So the leading principal minors all have positive values.

So let us consider this Hessian matrix so it has eigenvalues 7 and 1. Both the eigenvalues are positive. So this matrix is a positive definite matrix. You know that to find out the eigenvalues you can solve this determinant equation. So solve this for the values of lambda, you will get these. So those are the eigenvalues. Both the eigenvalues are positive. So it is a positive definite matrix.

We can also say in terms of leading principal sub-matrices, find out those determinants. So the first one, the first leading principal sub-matrices is 4, which is greater than zero. The second one is the determinant of the matrix itself. So both are positive, both are positive. So we get the same conclusion what we obtained using a eigenvalues that the given matrix is positive definite.

(Refer Slide Time: 30:11)

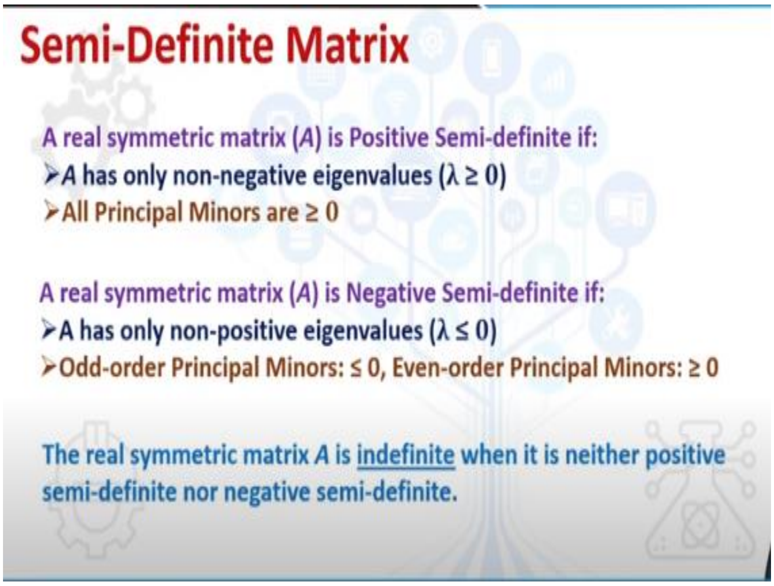


Similarly, a real symmetric matrix is negative definite if all the eigenvalues of the matrix are negative or equivalently. The leading principal sub-matrices have determinants with alternating signs, this is important, alternating signs. So it starts with negative then positive then negative then positive so on and so forth.

The first determinant of the leading principal sub-matrices will have negative determinant, then it will have positive determinant and so on and so forth. So alternating signs starting with negative. Let us take this example. Consider this Hessian matrix or any matrix, real symmetric matrix. Find out the eigenvalues. Its eigenvalues are -1 and -3. So both are negative. So it is negative definite matrix.

Find out the determinants of the leading principal sub-matrices. First is -2, which is less than zero. Next is the determinant of the matrix itself which is 3, which is greater than zero. So first negative next positive. So this pleads the definition of negative definite matrix.

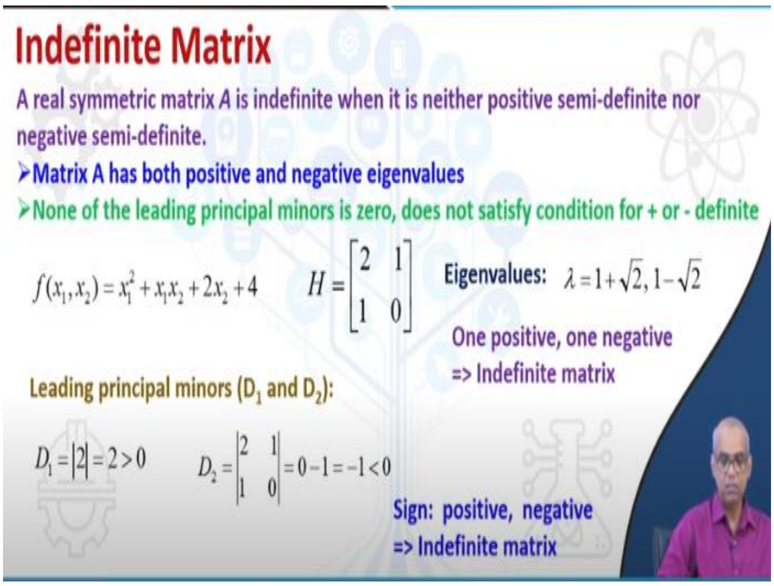
(Refer Slide Time: 31:33)



If a real symmetric matrix, a real symmetric matrix is positive semi-definite if some eigenvalues are positive, some eigenvalues are zero. So it has non-negative eigenvalues. Equivalently all principal minors are greater equal to zero, all principal minors are greater equal to zero. Similarly, a real symmetric matrix A is negative semi-definite if the matrix has some eigenvalues zero some eigenvalues negative.

So that that basically means that eigenvalues are non-positive. And the equivalent condition in terms of principal minors are odd order principal minors will be less or equal to zero and even order principal minors will be greater or equal to zero. The real symmetric matrix A is indefinite, when it is neither positive semi-definite nor negative semi-definite.

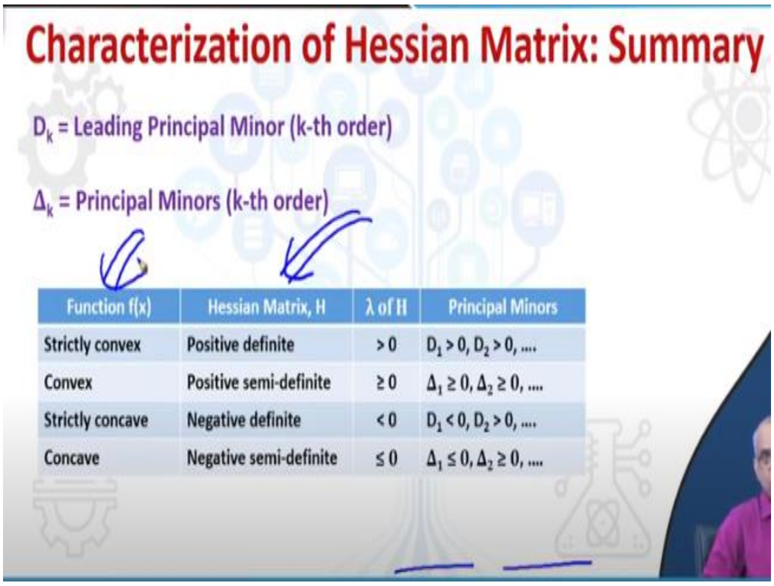
(Refer Slide Time: 32:33)



So let us consider this matrix. Let us evaluate the eigenvalues. One eigenvalue is positive another Eigenvalue is negative. Note that square root of 2 is greater than 1. So one eigenvalue is positive one eigenvalue is negative, so it is an indefinite matrix. In terms of principal minors first one is greater than zero, second one is negative.

So it does not fit the alternating sign criterion, because there first one has to be negative next it will be zero. So this is indefinite matrix.

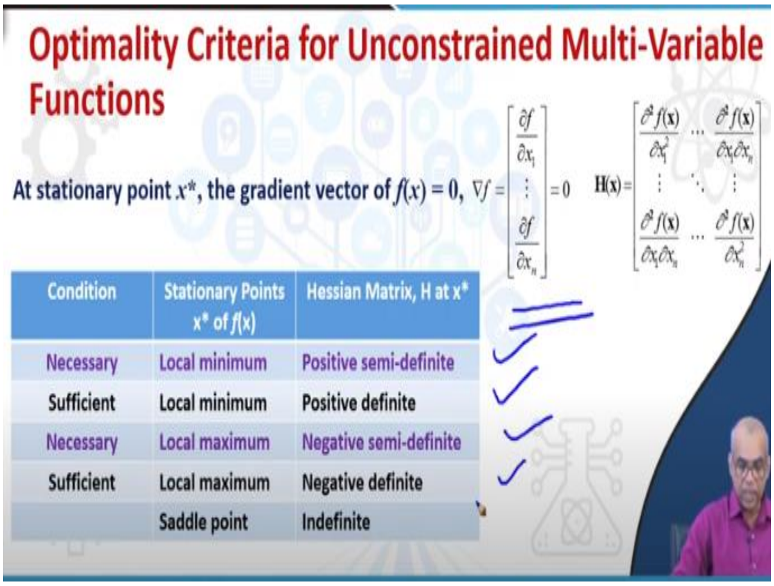
(Refer Slide Time: 33:14)



So in summary the Hessian matrix will be positive definite when all eigenvalues are greater than zero, positive semi definite when all the eigenvalues are greater or equal to zero. Hessian matrix will be negative definite when all the eigenvalues are less than zero and Hessian matrix will be negative semi-definite when all the eigenvalues are less or equal to zero.

Note that these characterization of Hessian matrix also tells us whether the function is strictly convex, convex, or strictly concave or concave.

(Refer Slide Time: 33:56)

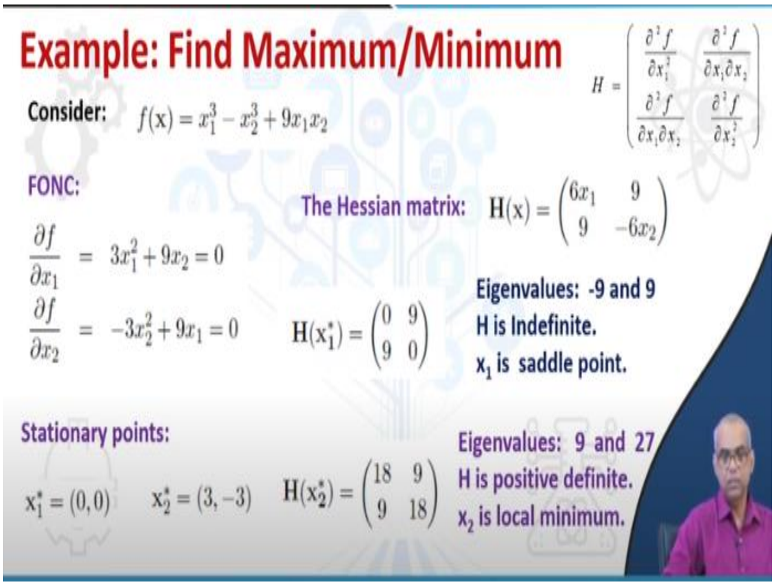


Now let us tell the optimality criteria for unconstrained multivariable function. So the necessary condition is this that the gradient vector evaluated at x star will be equal to zero. So gradient vector is equal to zero at the stationary point. Now evaluate the Hessian matrix at the stationary point. If the Hessian matrix is positive semi-definite, then this is a necessary condition for local minimum.

If the Hessian matrix is positive definite at stationary point x star, then this is a sufficient condition for local minimum. If at the stationary point the Hessian matrix is negative semi-definite then this is a necessary condition for local maximum. And if at x equal to x star the Hessian matrix is negative definite, then this becomes a sufficient condition for the point to be local maximum.

If at the stationary point the Hessian matrix is indefinite, then the point is saddle point.

(Refer Slide Time: 35:22)



So now let us apply these conditions to this function of two variable. f x = x 1 cube – x 2 cube + 9x 1 x 2. So evaluate the first order necessary condition. I take the gradient, set that equal to zero. That means take del f del x 1 equal to zero. Take del f del x 2, set that equal to zero. Solve these two equations. Simultaneously two variable two equations take and solve. I obtain two solutions 0, 0. That means x 1 = 0, x 2 = 0.

And also another solution x 1 = 3, x 2 = - 3. So you have to evaluate both whether they are minimum, maximum or saddle point. So evaluate the Hessian. Evaluate the Hessian at first, say point 0, 0. We see that the eigenvalues are -9 and 9. That means one eigenvalue is positive another eigenvalue is negative. So Hessian matrix is indefinite. So x 1 = 0, x 2 = 0 this point is saddle point.

Now let us consider the other one 3, -3. Evaluate the Hessian at x = 3, x 1 = 3, x 2 = -3. Eigenvalues are 9 and 27, both the eigenvalues are positive. So Hessian matrix is positive definite. So this point is a local minimum. So these are the criteria for optimality criteria for single variable unconstrained function and multivariable unconstrained function. With this we conclude our discussion here.