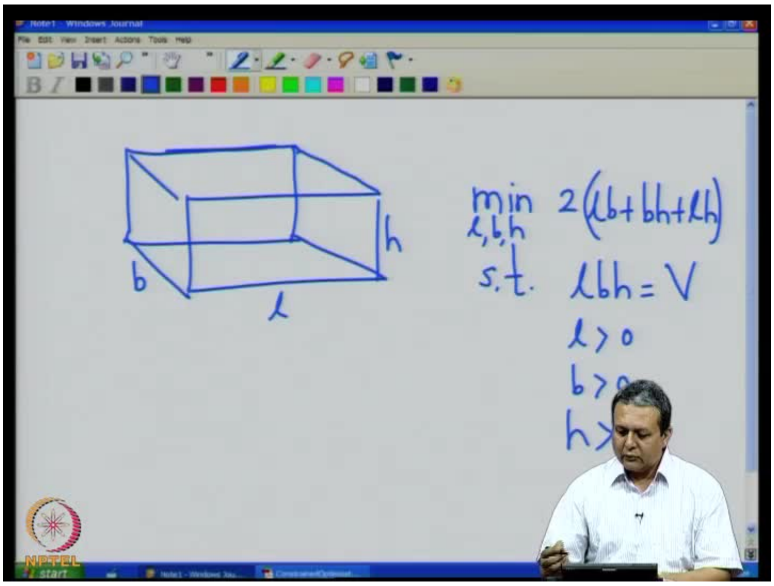
Numerical Optimization

Prof. Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 20
Constrained Optimization - Local and Global Solutions, Conceptual Algorithm

Hello, welcome back to this series of lectures on numerical optimization. In the last few classes, we discussed about unconstrained optimization problems and some algorithms to solve those unconstrained optimization problem. Now, in practice many problems are constrained optimization problems; and we need to solve those constrained optimization problems using suitable methods. For example, we have looked at this problem, where we want to find out a box of minimum surface area which can accommodates certain volume.

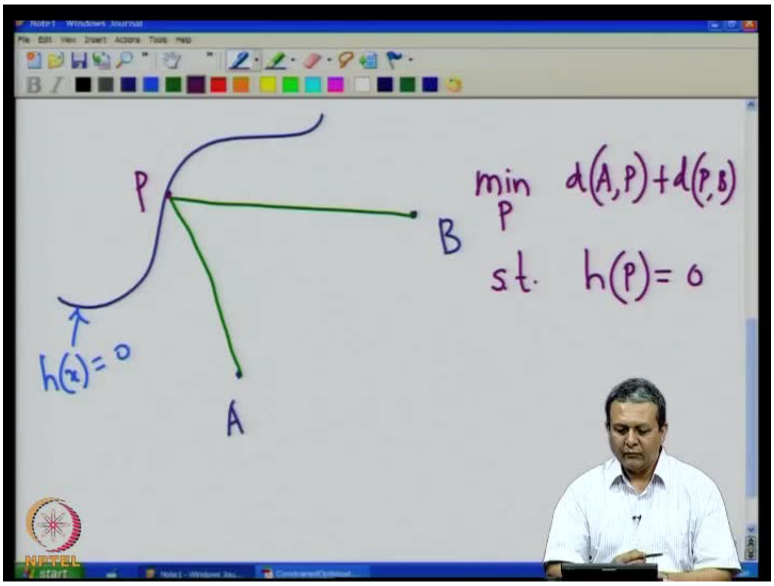
(Refer Slide Time: 01:06)



So, if we take a box, let us assume that its length is l, breadth is b and the height is h. So, we are interested in finding the box of a minimum surface area which can accommodate a certain volume V. So, we can write this problem as minimize the surface area, surface area is 2 into l b plus b h plus l h. Now, this is the surface area and the constrained that we have is that it can accommodate, it should accommodate a certain volume V. So, l b h should equal to some volume V, and the variables here they are strictly positive quantities. So, we can say that l is greater than 0, b is greater than 0, h is greater than 0 and the optimization variables are l b h.

So, you will see that finding a box of minimum surface area which can accommodate a certain volume V of us a particular matter can be written as our constrained optimization problem. So, this is the, our objective function and there are various constrains, the first constrains is that the, it should accommodate the volume V. And then the next three constraints say that all the quantities or all the variables they are all strictly positive none of them can become 0 or negative. Let us look at some other problem.

(Refer Slide Time: 03:26)

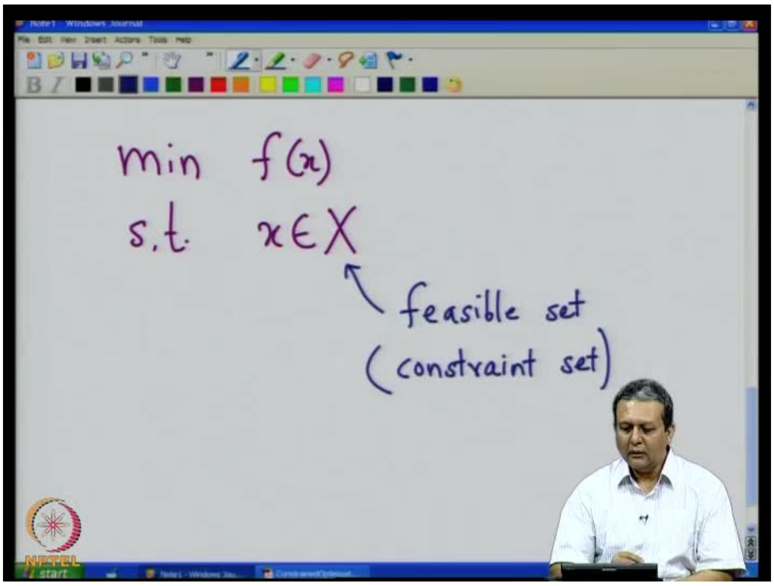


So, suppose we want to move from one place to another place. So, let us call these points as A and B, so under the Euclidean norm one can find a straight line path which minimizes the distance between the two points A and B. Now, suppose there is one constraint that we have to follow. So, we want to visit this road which is given here, a non linear road and then go to the place B.

Now, we want to find out a point say P let us call this point as point P. So, our aim is to move from A to P and then P to B. So, what we want to do is that we want to find out the point P on this curve such that the total distance travel which is the distance A P plus distance P B is minimized. So, let us assume that we are using the Euclidean distance. So, let us denote this curve by say H of x equal to 0. So, what we are interested in is to solve the following problem where we want to minimize the distance between A P plus distance between P B. Now, what are the constraint? The constraint is that the point P should belong to the curve H x equal to 0. So, the constraint is that the point P should belong to this curve H P. So, we are interested in finding out the point P, I have not written this problem formally, because we need to look at the coordinates of the point P. And then write the objective function and the constrained appropriately, but what this problem formulation essentially means is that we are interested in that point P, such that the distance between A P plus the distance between P B that is minimized where the point P lies on the curve H x equal to 0.

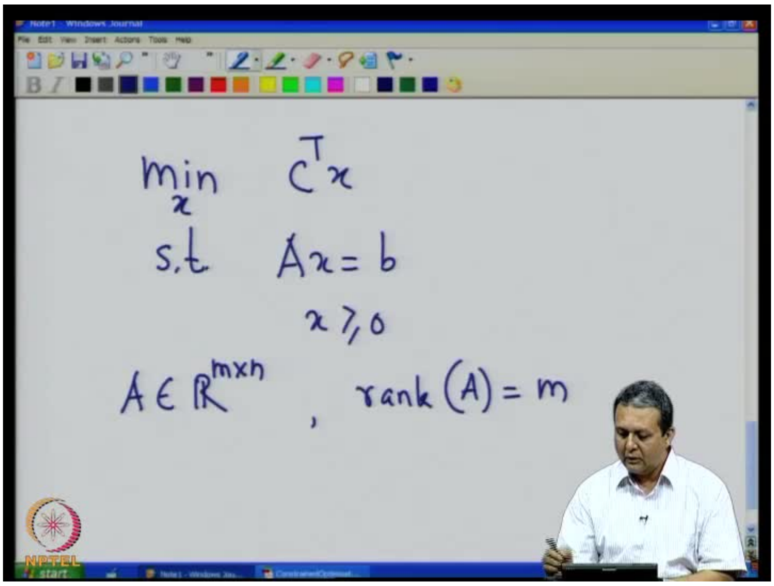
So, like that in our daily life become across lot of problems which are constrained optimization problems. And it is important to study the behavior of this problems and develop some efficient algorithms to solve such problems. Now, many times this constrained optimization problems can be written as unconstrained optimization problems and then solved. And therefore, it was important to study unconstrained optimization problem, unconstrained optimization theory earlier, so that we are in a position to study the constrained optimization theory.

(Refer Slide Time: 07:11)



Now, general constrained optimization problem would look like this, so minimize f of x subject to x belongs to X. Now, as usual this f is objective function and this x; this is called feasible set or all this is also called constraint set. So, this is the typical constrained optimization problem. So, when this x is the n dimensional space of real numbers then when the capital X is n dimensional space of real numbers then this problem becomes unconstrained problem. Now, some of the typical constrained optimization problems that one comes across are the following.

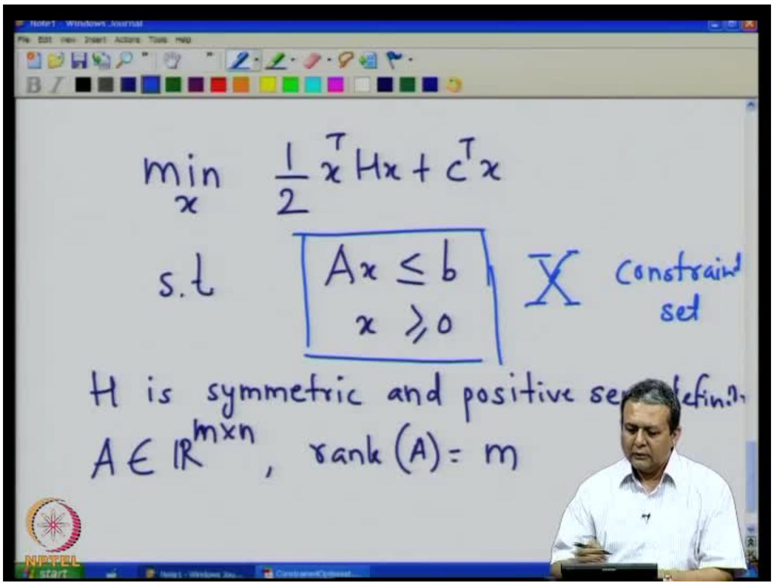
(Refer Slide Time: 08:22)



So, one problem is that if you want minimize C transpose x subject to the constraint A x equal to b, x greater than or equal to 0 where A is a m by n matrix and without loss of generate you assume that rank of A is equal to m. So C is a n dimensional vector, b is a m dimensional vector and x is non negative. So, the objective function is to minimize C transpose x subject to the constraint that x equal to b, x greater or equal to 0. This is commonly known, known as linear programming problem, because the objective function is linear in the variable, the constraints are linear in the variables. So, this is the linear programming problem.

So, later on we will see how to solve this kinds of problems, because they are the very important part of optimization problems and they are exists very good methods to solve them. So, in this course some time later, we will study how to solve this kinds of problems, so this is called a linear programming problem. Again you will see that this is the constrained optimization problem.

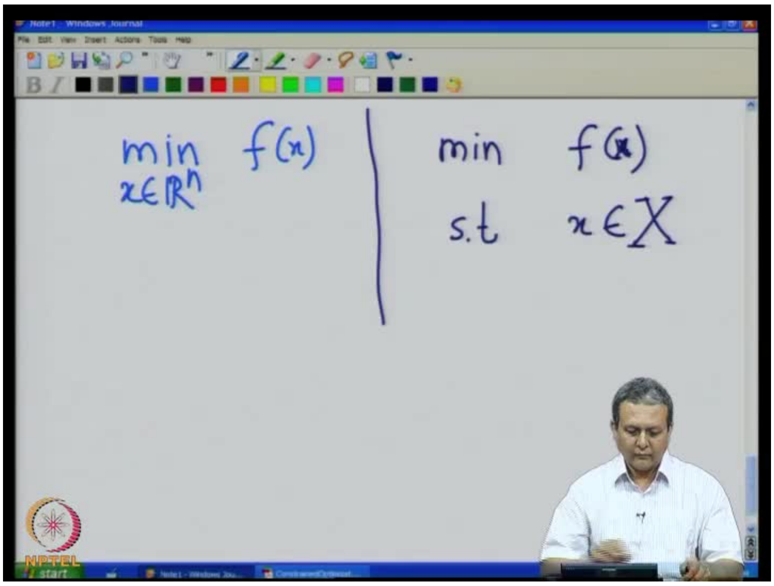
(Refer Slide Time: 09:51)



Let us look at another example, minimize half of x transpose H x plus C transpose x subject to A x less than or equal to b x greater or equal to 0 where H is symmetric and positive definite matrix. And let us assume that A is m by n matrix and rank of A is equal to m. So, the objective function here is a quadratic function the constraints of linear, so many problems in practice can be pause as a problems where the objective function is quadratic and the constraints are linear in the variables. So, here the variable is x and we want to solve this kind of problem. So, this set of constraints say x less than or equal to B and x greater than or equal to 0 in this case they from the constraint set so which we have called it as capital X, so this is our constraint set.

So, you will see that lots of problem that we come across in practice are constrained optimization problems. And therefore, it is important to study the nature of solution of this constrained optimization problems as well as some algorithms to solve this constrained optimization problems. So, in the next few lectures, we are going to concentrate more on the constrained optimization problems, the nature of the solution how does one get the solution.

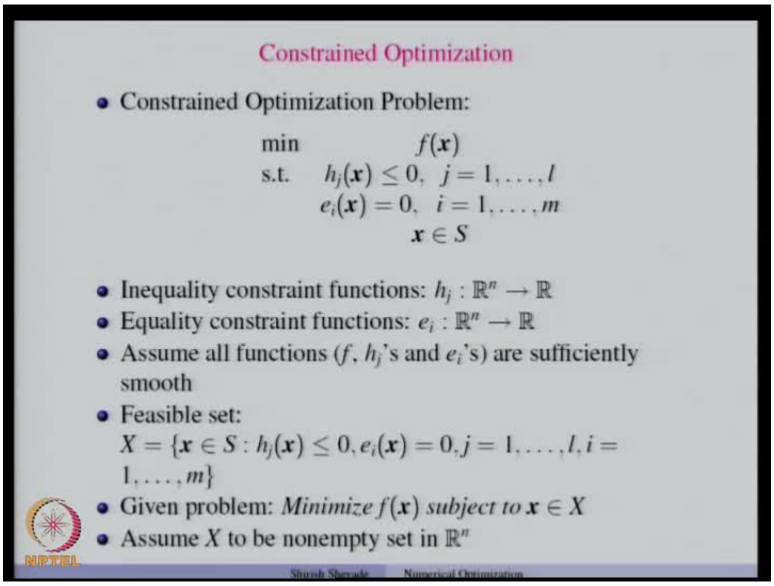
(Refer Slide Time: 12:24)



Now, if we recall in the unconstraint problem case where we solve this problem, we saw that the necessary condition for extra to be a local mean is that the gradient of f at x star 0. We assume that f is smooth so what I mean by smooth is that second derivative exists and it is continuous at every x in the domain. So, the necessary conditions for the local minimum of this problem are that the gradient of this function should be 0. And the sufficient condition is that the hessian matrix should be positive definite, positive semi definite.

Now, if the hessian matrix is positive definite then we say that at that point x star where the hessian matrix is positive definite and where the gradient of the function vanishes we say that that point is the strict local minimum. Now, we are looking at a different problem where we want to minimize f of x subject to x belongs to X. Now, what kinds of optimality conditions exists for this problem? So, remember that here we said that the gradient of the function should vanish. Now, that condition alone may not be enough for this, because now we are working with some constraint. And can we arrive at a condition, which is somewhat similar to that first order necessary condition for a unconstrained optimization problem. And how do we make use of this set x in order to write that our optimization condition for a constrained problem like this and that is the part of the next few lectures.

(Refer Slide Time: 14:54)



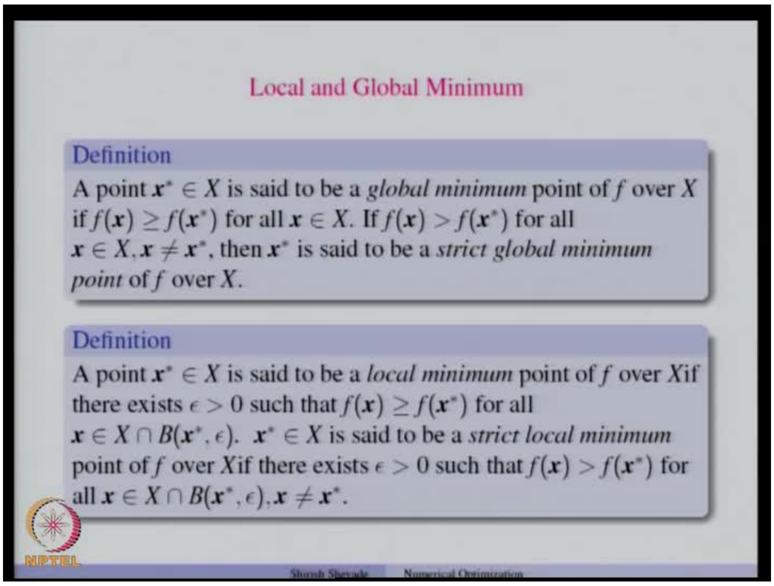
So, as I said that constrained optimization are very important and let us start looking at a general constrained optimization problem. Now, a typical constrained optimization problem looks like this, where we want to minimize the objective function f of x subject to some constraints of the type h j (x) less than or equal to 0. So, these are the inequality constraints then there could be some equality constraints e i (x) equal to 0. So, there are l inequality constraints m equality constraints and further x belongs to some set x. So, all these together they form the constraint set or the feasible set of this problem, sometimes it is also called the feasible region.

So, the inequality constraints, in this course we are going to denote it by h j and all h j's are functions from R n to R and all equality constraints we are going to denote by e i where e i is a function from R n to R. Now, we will a show assume that this f h j and e i's are all smooth. So, what I mean by smoothness is that they are the second derivatives of continuous at every point in the domain, they exist the second derivatives exists and they are continuous at every point in the domain. So, we will make this essential throughout, so that it is easy to write down the optimality condition in terms of the derivatives at a later point of time.

Now, the feasible set x is basically the set of points in the set S which satisfy the constraints h j (x) less than or equal to 0 and e i (x) equal to 0 for all j is going from 1 to l and i is going from 1 to m. So, this is going to be the feasible set of our constrained optimization problem. In many problems this S is nothing but R n and therefore, we can replace S by R n here and simply write it as x is equal to all x belong to R n such h j (x) less than or equal to 0 and e i (x) equal to 0.

So, our problem is to minimize f of x subject to x belongs to X. So, this is a compact way of describing the constrained optimization problem which we are looking at. Now, let us assume that x is nonempty set in R n, because if x is empty then we really cannot solve this problem. And suppose x is singleton then that point the singleton point in the set x will be an optimal point. So, let us assume that not only that x is nonempty, but x is also not singleton.

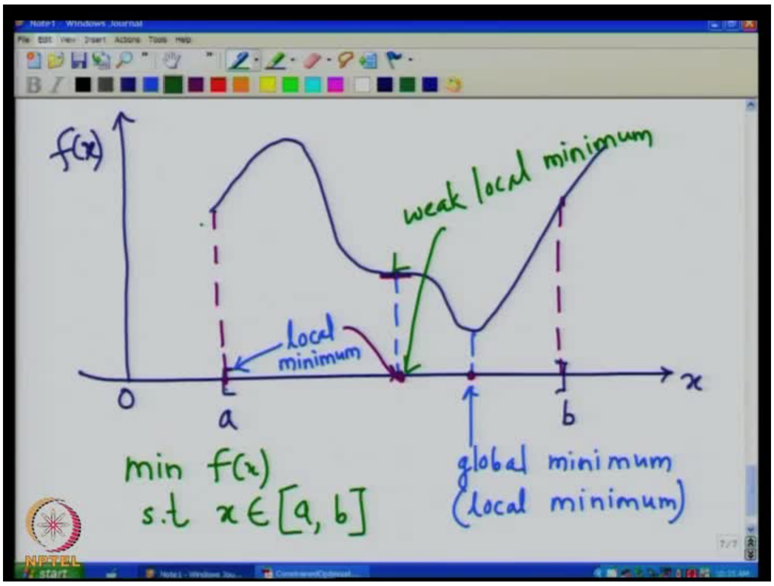
(Refer Slide Time: 18:04)



Now, similar to the constraint the similar to the unconstrained optimization problems, we have to see the notion local and global minima in the context of constrained optimization problems. So, remember that we are trying solve the problem minimize f (x) subject to x belongs to the capital set x or x belongs to the feasible region. So, a point x star in the feasible region is said to be a global minimum point of f over x if the value of f (x) is greater than or equal to f of x star of all x in the feasible region and if this inequality holds strictly then the point x star is called strict global minimum. So that is if f of x is greater than f of x star for all x in x where x is not equal to x star then x star is said to be a strict global minimum point of f over x. Now as is the case of unconstrained optimization problems is a global minimum or are to obtain.

So, one looks at the local minima, so one can define the local minima in the following way, so a point x star which is feasible said to be a local minimum point if there exists a neighborhood of x star or there are exists a epsilon neighborhood of x star such that f of x is greater or equal to x star for all x in the epsilon neighborhood. And again when this inequality is strict we called it as a strict local minimum. Now, let us see some examples of the local minima and the global minima.

(Refer Slide Time: 20:16)



So, let us consider a problem where this is x suppose this is the origin and from the y axis we will plot x. So, this is f of x; this is our origin and suppose that the function that we have is in the interval say a to b and the function is like this. So, although the function may be defined over the entire real line we are only interested in the closed interval a, b. So, what we are interested in is solving this problem minimize f of x subject to x belongs to a b. Now, now let us look at the global minimum first, now if you just concentrate on this part of the function, now this part of the domain a, b of the function. So, we are not interested really in this part what happens beyond the other points beyond the closed interval a, b. So, if, if you look at the definition of the global minimum which says that f of x star should be less than or equal to f of x for all x in the constraint set of the feasible set and remember that x star also should belong to the constraint set.

So, by that definition this point; this point becomes a global minima. So, in the constraint set there is no other x which has a value functional value which is lesser than this so this is a global minimum. Now if you look at the definition of a local global minima local minimum, we will see that those are the points where in the epsilon neighborhood of those points, if you look at all the feasible points in the neighborhood the functional value is at least the functional value at the given point.

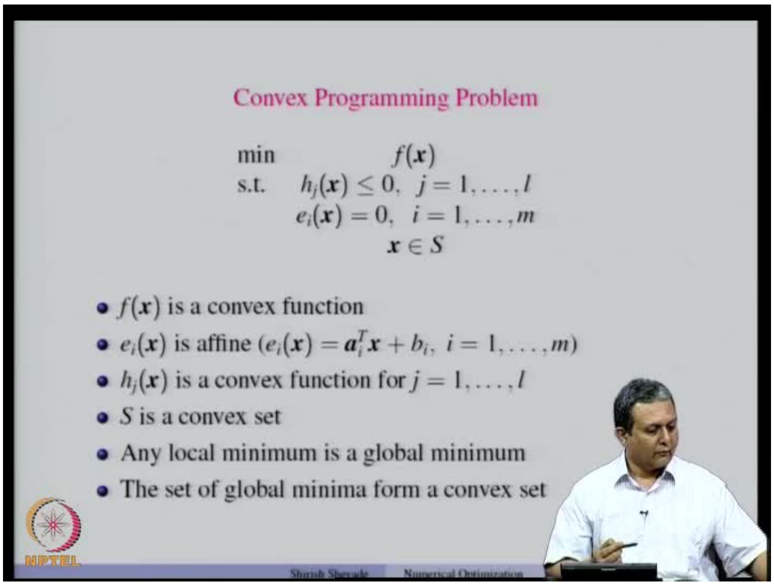
So, for example, so this point is also a also a local minimum. Now, if you look at this point x, x equal to a, you will see that in the neighborhood the functional value is increasing, so this is also a local minimum and if you look at this point. So, in the neighborhood, you will see that the functional value does not increase or in the neighborhood the functional value does not decrease, so this is also is a candidate for a local minimum. So, the functional value is at least the value at this point, so, so this point is the local minimum. So, there are 3 local minima for this one is x equal to a then this point and this point these are the 3 local minima, because in the neighborhood the function does not decrease there exists at least some epsilon neighborhood around these points where the functional value does not decrease. So, these are the local minima, now this point is also global minimum with respect to the constraint set a, b.

Now, if you look at these 2 points here the functional value locally does not decrease but does not increase either it remains constant. So, this is a weak local minimum, so the corresponding point, so this point is also a weak local minimum, so where the functional value does not increase or decrease but remains constant in the epsilon neighborhood. So, these are different characteristics of the local or the global minima. Now, as is the case in case any optimization problem we are interested in finding the global minimum of our, the optimization problem but that is difficult to find, because of the same reasons that we discuss for an unconstrained problem. So, we will be interested in looking at the local minima their characterizations and how do we design efficient algorithms to find out the global minima.

Now, you will also notice that the first order necessary conditions that we discuss for a unconstrained optimization problem may not always hold at local minima. So, at this point the gradient of the function vanishes. So, this is fine; this also this point is also fine, but if you look at this local minimum if you take the one directional derivative of the function at this point it is not 0. So, you will see that the first order conditions that we studied for unconstrained optimization problem cannot be directly used to characterize the local minima of a constrained optimization problem.

So, we need some extra conditions to characterize the local minima of constrained optimization problems. Now, under suitable conditions the first order necessary conditions for constrained optimization problems are also sufficient and those problems are called convex programming problems.

(Refer Slide Time: 29:21)



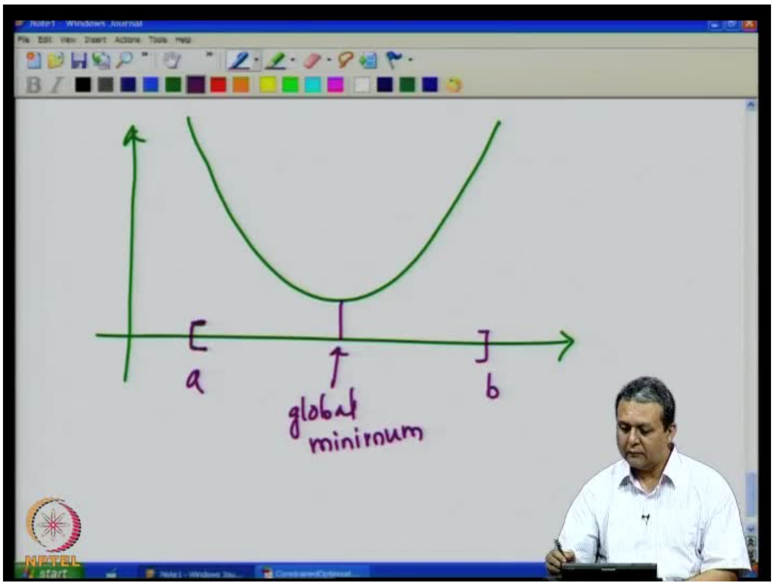
Now, in the unconstraint case we saw that if we want to minimize a convex function over the set R n then the first order necessary conditions are also sufficient. Now, similar result holds in the case of convex programming problem where we are writing a general constrained optimization problem, but then the functions which are mentioned here the objective function as a there is a constraints they have certain characteristics and because of which these problem are called convex programming problem.

So, the objective function f (x) is a convex function. So, the objective function that you want to minimize is a convex function the e i (x) equal to 0 is affine constraint that means e i (x) is of the form a i transpose x plus b i for all i is going from 1 to m. So, all the equality constraints are of, of the type a i transpose x plus b i equal to 0 or they are affine constraints then the h j (x) less than or equal to 0 constraint or the inequality constraint is such that the function h j (x) is a convex function and further S is a convex set. So, now if you look at this constraints h j (x) less than or equal to 0 where h j (x) is the convex is a convex function. So, we have one such constraints so in intersection of these, so h j (x) the set of all x's that h 1 x less than or equal to 0 where h x h 1 x is a convex function is a convex set.

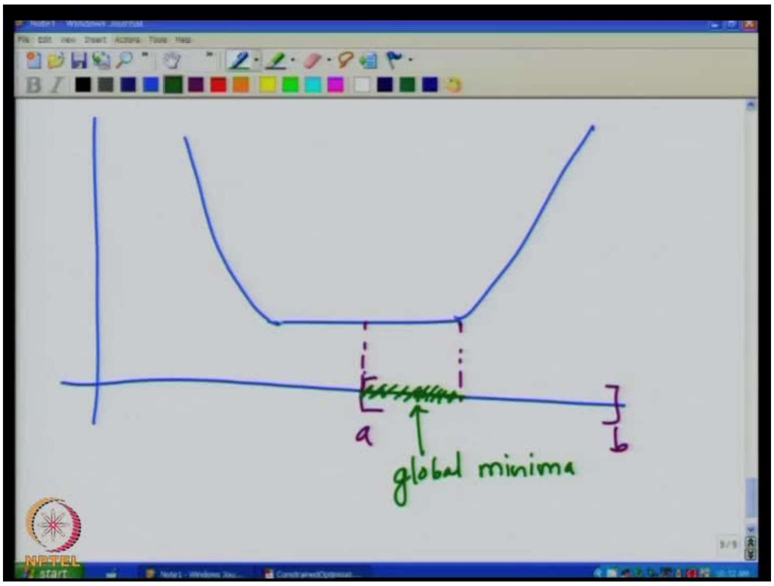
Now, we have one such constraints. So, the intersection of all convex sets is a convex set then the intersection of that with affine set is also a convex set and the intersection of this convex set with another convex set is a con is a convex set. So, we have studied this property of convex sets earlier that the intersection of any collection of convex sets is a convex set. So, the constraint set is a convex set the objective function to be minimized is a convex function and such problems are called convex programming problems.

Sometimes, when one wants to maximize maximize the quantity of function that problem also subject to convex set that problem also can be written as minimization of a convex function, because maximization of a convex function is same as the minimization of the corresponding convex function so such problems are called convex programming problems. And for such problems every local minimum is a global minimum and there are no, the, the set of global minima form a convex set.

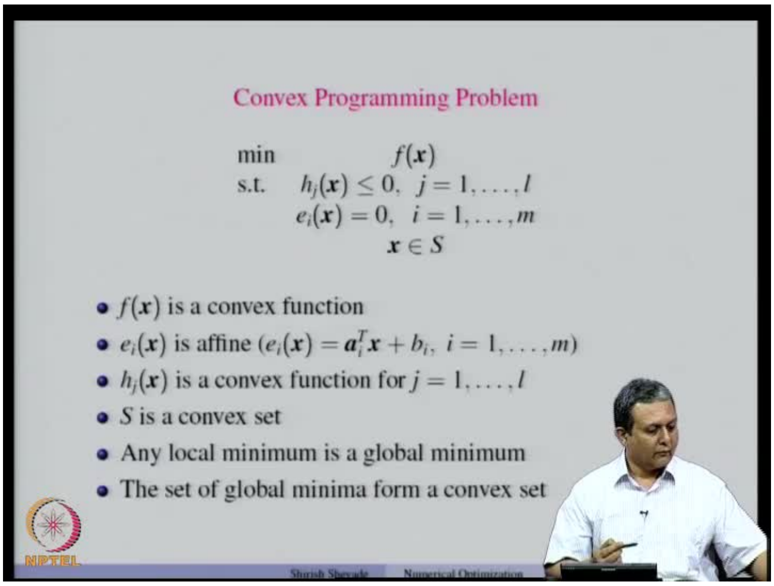
(Refer Slide Time: 32:40)



(Refer Slide Time: 33:31)



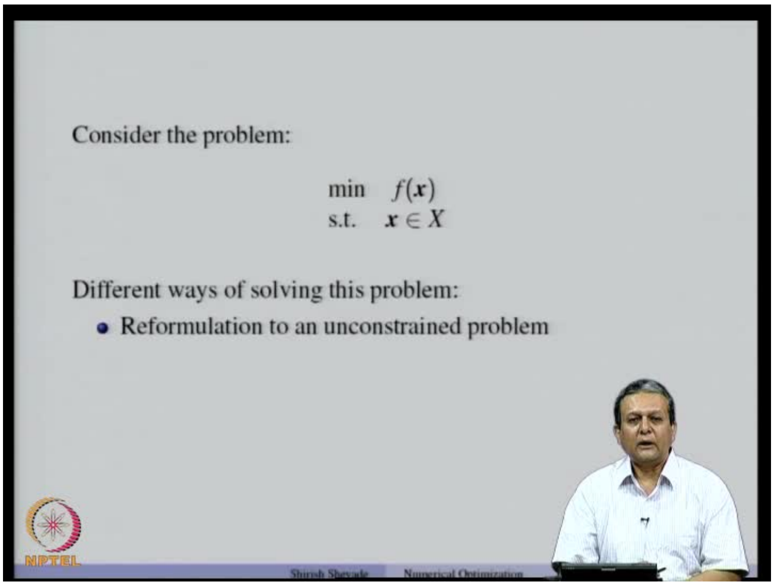
(Refer Slide Time: 29:19)



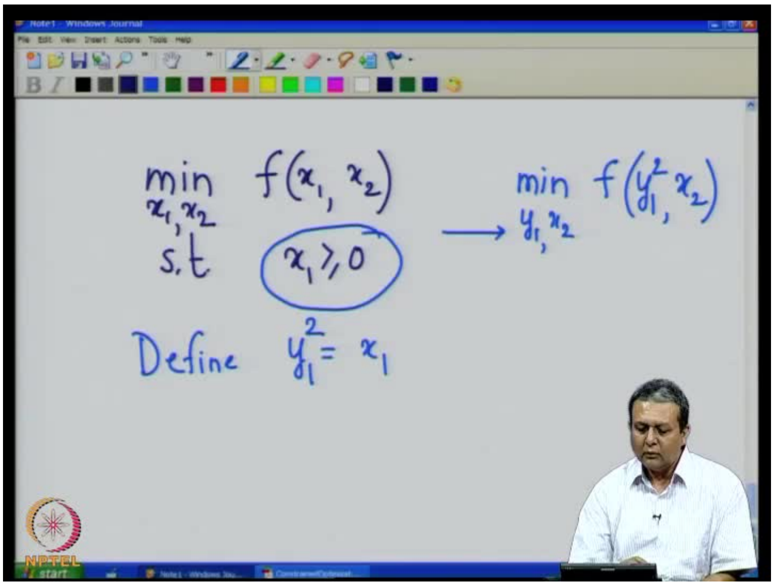
So, if you want to minimize a convex function, now you will see that suppose we want to minimize it with respect to the set a, b. So, you will see that this is a so this point is a global minimum and if you look at another convex function. So, let us look at another convex function and we want to minimize it with respect to the set a, b. Now, you will see that all these are the. So, this entire set of points, so this entire set of points they are global minima. So, you will see that they form a convex set, so not only that the glob there no question of local minima in the convex programming problem, in fact the all the global minima which are possible they form a convex set.

So, this is an important property of a convex programming problems. And further, as later on we will see that the first order necessary conditions for of a next programming problems are sufficient under certain conditions. So, we will look at those things sometime later, so these are some important properties of convex programming problems and later on we will see this problems in detail,

(Refer Slide Time: 35:32)

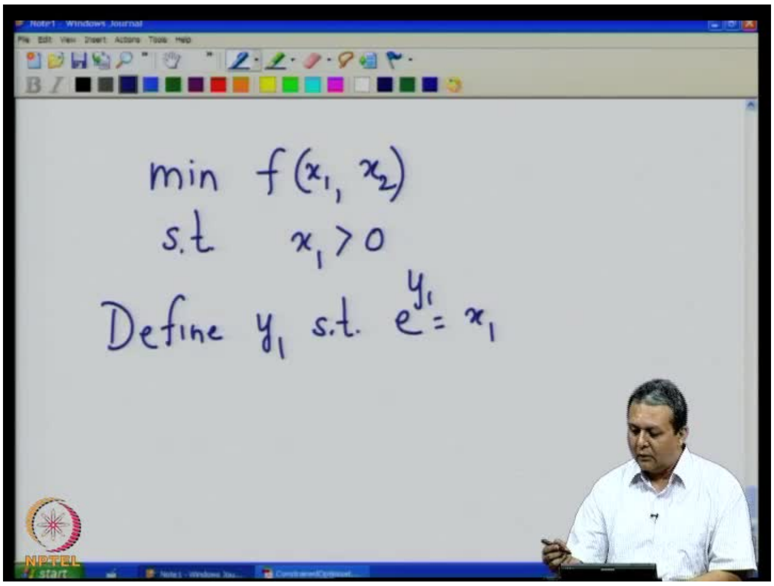


(Refer Slide Time: 36:19)



Now, let us look at the problem minimize a f (x) of that 2 x belongs to x which is a compact form of a constrained optimization problem where x is a feasible set. Now, there are different ways of solving this problem and one of the ways is to converted to an unconstrained problem or reformulation to an unconstrained problem. So, let us look at a some ways of doing that. So, suppose that we have to solve the following problem minimize f of x 1 x 2 subject to the constraint that x 1 greater than or equal to 0. Now, here there are a constraints on the variable x 1. So, what one can do is that since x 1 is going to be non negative, we can replace this constraint using some new variable of i defining some new variable. So, let us define y 1 square is equal to x 1, so since y 1 square is always a non negative quantity that means that x 1 will also be a non negative quantity. And then we can write this problem as minimize f of y 1 square x 2 and now the variable is y 1 and x 2. So, you will see that by replacing a non negative variable by another variable of the square of another variable y 1, we have converted the given problem into an unconstrained problem where the variable is now y 1 and x 2. So, in the earlier case the variable was x 1 and x 2, so we written x 2 as it is, but then replaced x 1 by y 1 square in the objective function and that become the unconstrained optimization problem.

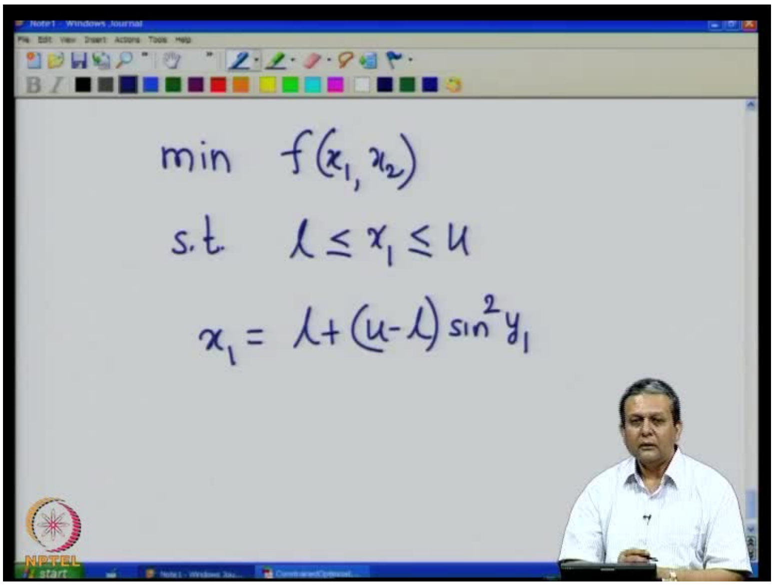
(Refer Slide Time: 38:35)



Now, sometimes we may want to solve the following problem minimize f of x 1 x 2 subject to constraint that x 1 is greater than 0. So, now we are interested in those x 1's which are strictly positive, so here one can use a new variable. So, suppose if you define y, y 1 such that e to the power y 1 is equal to x 1. So, we have defined new variable y and we know that e to the power y 1 is always a positive quantity. And therefore, in this objective function if we replaced x 1 by e to the power y 1. And then treat y 1 as a variable along with x 2 then we have converted this problem to of conse to an unconstrained optimization problem and we know different ways to solve unconstrained optimization problems.

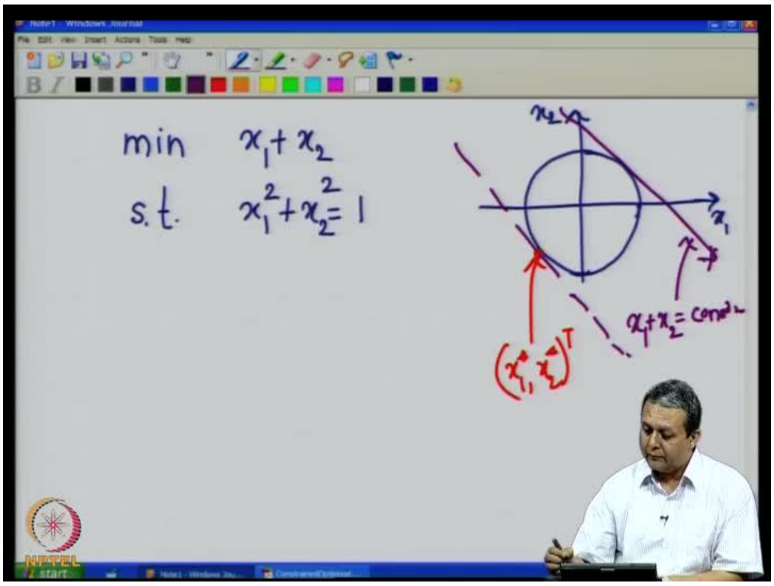
So, many a times it may be a good idea to see whether a problem can be converted to and unconstrained problem although there is no rule that we always have to convert an unconstrained optimization problem to a constrained optimization problem. Sometimes the constrained pro optimization problems are easier to solve compare to unconstrained optimization problems.

(Refer Slide Time: 40:27)



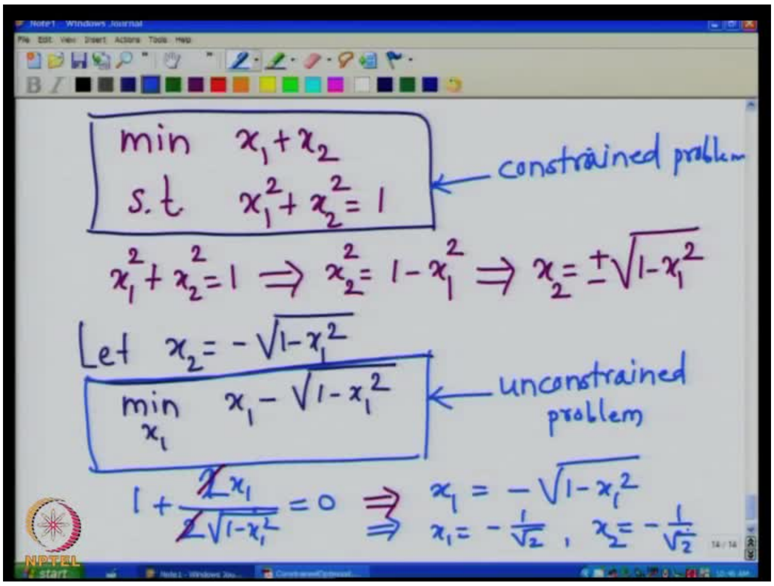
Now, one can also have a problem which is of the type say minimize f of x 1 x 2 subject to the constraint that l less than or equal to x 1 less than or equal to u. Now, such problems also can be converted by defining a new variable. See for example, one can use some functions which very over the certain range for, for example, one can write x 1 as say l plus u minus l sin square y 1. Now, sin square y 1 takes the values in the range 0 to 1, so when sin square y 1 is 0 we have x 1 equal to l and when sin square y 1 equal 1, we have x 1 equal to u so and since sin square y 1 varies in the range 0 to 1, we have x 1 varying in the interval l to u and y 1 can be any real number. So, you will see that by using such transformations the original problem can be converted to an unconstrained optimization problem. Note that this transformations are not unique one can always come up with some new transformations depending upon the problem. Now, as I mentioned earlier that one has to be very careful when one solves this constrained optimization problems by reformulating by them is as unconstrained optimization problems.

(Refer Slide Time: 42:28)



So, let us suppose consider case where we want to minimize x 1 plus x 2 subject to the constraint x 1 square plus x 2 square is equal to 1. Now, let us see the problem graphically, so we have x 1 and x 2 as the 2 axis and the constraint set is a circle with centre 0 and radius 1 and the objective function is x 1 plus x 2 is equal to constant. So, this is the, so this is the objective function x 1 plus x 2 is equal to constant. Now, we want to minimize this objective function so you will see that the minimum is achieved at this point so, this point is the point. So, this is x 1 star x 2 star, so this is the point at which the minimum is achieve. So, you will see that this line the minimum occurs at this point and this is our solution of this problem.

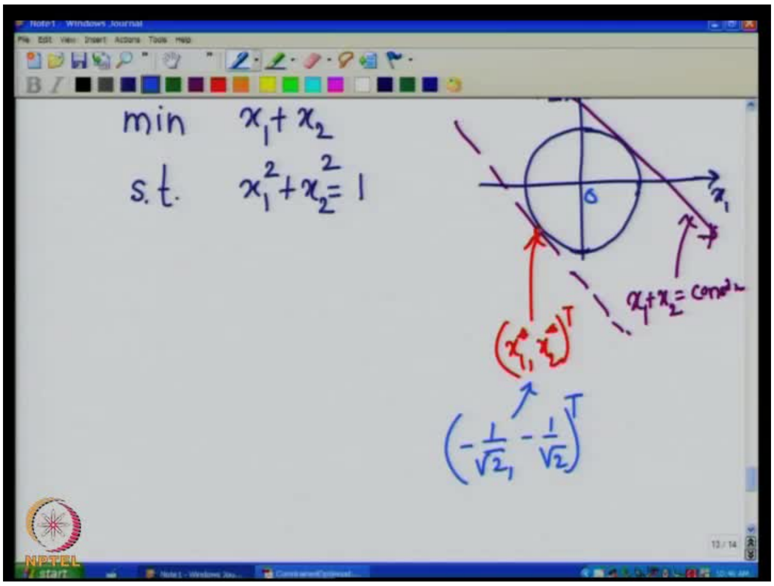
(Refer Slide Time: 44:42)



Now, now let us look at this problem and then converted to an unconstrained optimization problem. So, we have minimize x 1 plus x 2 subject to x 1 square plus x 2 square is equal to 1. Now, if you look at the constraint x 1 square plus x 2 square is equal to 1, what this gives us is that x 2 square is equal to 1 minus x 1 square and this implies that x 2 is equal to plus or minus square root of 1 minus x 1 square. So, there are 2 possibilities; one is that x 2 is square root of 1 minus x 1 square and other possibility is the negative of the square root of 1 minus x 1 square.

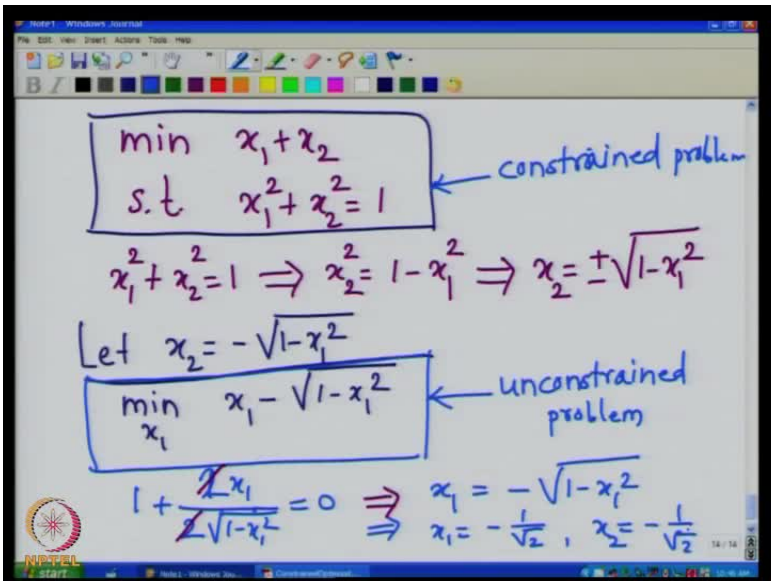
So, suppose you chose one, one of this possibilities so let us take, so let x 2 to be minus root of 1 minus x 1 square, so when we minimize x 1 x 2. So, this was our earlier problem and we use this value of x 2 in this, so we this value of x 2 was there are using this constraint, so really now do not have to worry about the constraints. So, let us plug in this value of x 2 in this and what we get is minimize x 1 minus square root 1 minus x 1 square with respect to x 1. So, you will see that this is now an unconstrained optimization problem. So, this was constrained problem and this is now un constrained problem.

(Refer Slide Time: 48:51)

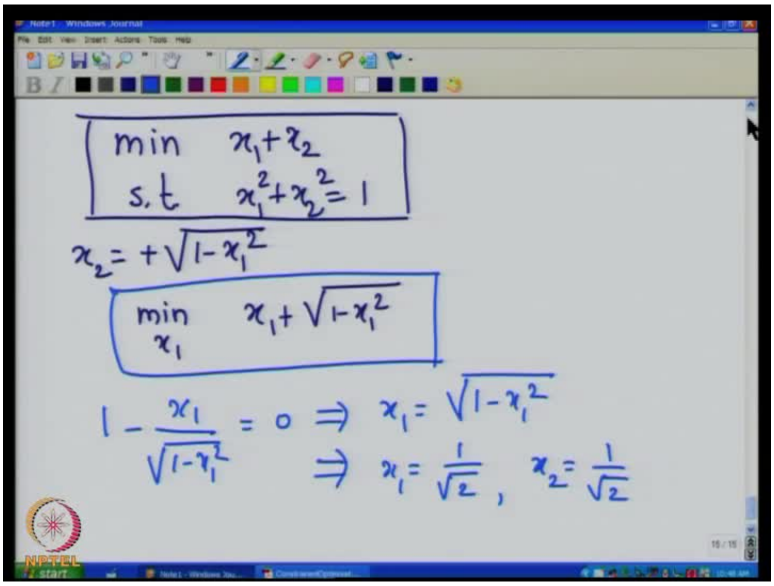


Now, we are lead in a how to solve this unconstrained optimization problems. So, we take the derivative of this objective function equated to 0 and see the candidates for the local minima. Now, if we take the derivative of this objective function what we get is that 1. So, the derivative of root of 1 minus x square is 1 over 2 root 1 minus x 1 square into minus 2 x 1. So, what we have is plus 2 x 1 that is equal to 0 and this quantities get canceled and what we get is x 1 to be minus square root of 1 minus x 1 square and this implies that x 1 to be minus 1 over root 2. And if we substitute this value of x 1 in this formula we get x 2 is equal to minus 1 over root 2 and if you look at the minimum of this objective function. So, the x 1 star x 2 star that we get here is minus 1 over root 2 minus 1 over root 2, remember that this is our origin and so the minimum lies in the third quadrant and this is our minimum.

(Refer Slide Time: 44:42)

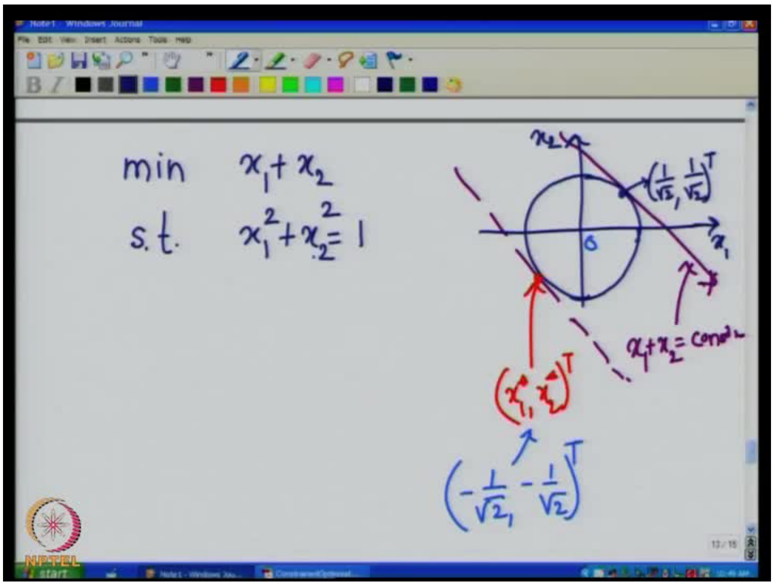


(Refer Slide Time: 49:35)



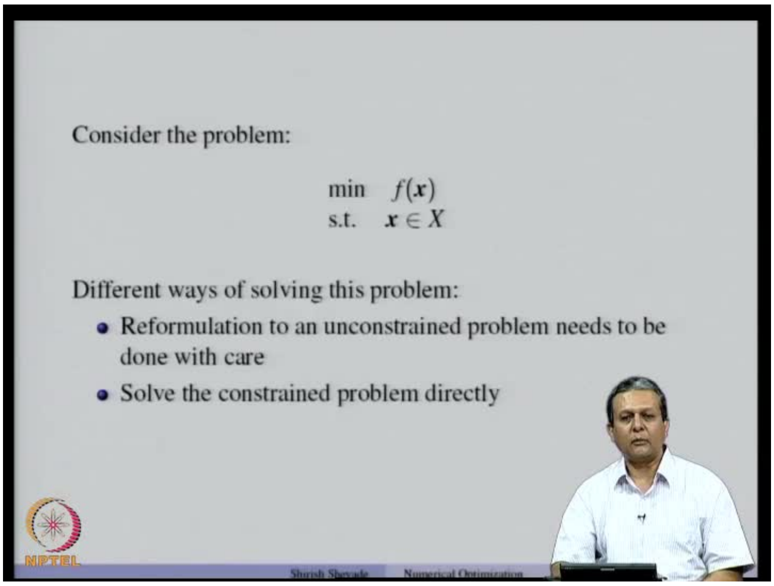
Now, so we use this value of x 2 to be minus root of 1 minus x 1 square to write it as a unconstrained optimization problem and then solve it. Now, the same problem, minimize x 1 plus x 2 subject to x 1 square plus x 2 square equal to 1. Now, we use x 2 to be plus square root of 1 minus x 1 square. And therefore, the given problem becomes, so this was the constraint problem and we write it as an unconstrained problem as minimize x 1 plus square root of 1 minus x 1 square with respect to x 1, so this becomes our unconstrained optimization problem.

(Refer Slide Time: 51:08)



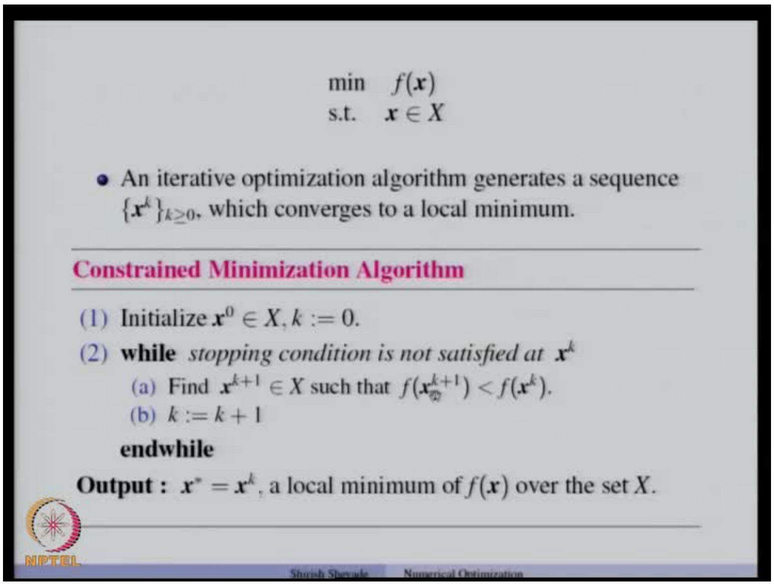
So, again taking the derivatives, so what we get is 1 minus x 1 by root of 1 minus x 1 square equal to 0 which implies x 1 is equal to root of 1 minus x 1 square, and this implies that x 1 is equal to 1 over root 2 and if you plug in this value of x 1 here what we get is x 2 to be 1 over root 2. So, now let us look at this point 1 by root 2 1 by root 2 in the earlier figure. So, you will see that this point; this point is 1 over root 2, 1 over root 2 transpose and this in fact is a. So, you will see that this in fact is a local maximum. So, when we transform the given unconstrained problem given constrained problem to an unconstrained problem, we have to be very careful, because you saw that in this case when we 2 case 2 be the minus root of 1 minus x 1 square we got the actual local minimum. And if we used x 2 to be plus root of 1 minus x 1 square and plugged in that value here we got the local maximum. So, whenever we do any transformation of our constrained optimization problem to an unconstrained problem optimization problem we have to be very careful.

(Refer Slide Time: 52:31)



So, the reformulation of to an unconstrained problem is always possible, but that needs to be done with some care the other possibilities to solve the constrained problem directly. And it is this way of solving the constrained optimization problem that we are going to concentrate of all.

(Refer Slide Time: 52:56)



So, let us consider this compact way of writing the constrained optimization problem and as in the case of unconstrained optimization problem, we are interested in deriving an iterative optimization algorithm to solve this problem. And that optimization algorithm it generates a sequence x k which converges to a local minimum. Now, one has to be very careful when one writes this that this x k its always assume that this x k always belongs to the set x. So, it is not a any arbitrary sequence x k that we want to generate, but we want to generate the sequence which lies in the feasible set x.

So, typical constrained minimization algorithm is given here the reason why I said that this is typically is that there could be some algorithms which not necessarily insure that x k lies inside the set x. But when the algorithm converges to a local minimum the point convert the point of converges will be in the set x. So, those are the different types of algorithms, but the typical algorithm would look like this where we initialize x 0 in the set x set the iteration counter k to 0. And then while some stopping condition is not satisfied at x k, we find out x k plus 1 in x k a in the set feasible set x such that the function value decreases then increase the iteration counter. And the procedure is repeated till one gets a point x star at which at which point the stopping condition is satisfied and that x star is a local minimum of f (x) over the set x.

So, you will see that in this typical algorithm of the conceptual algorithm for a constrained minimization problem every, at every point or at every iteration the point x lies in the feasible set x. So, the initial point itself is in the set x and at every iteration you get x k plus 1 in the feasible set x and such that the value of the function decreases. Now, there are 2 important points that one needs to consider here is that what is the stopping condition that needs to be satisfied at x k? In the case unconstrained optimization problem, we said that the norm of the gradient should be less than or equal to some epsilon can be one of the stopping conditions. Now, what are the analogous conditions in the case of constrained optimization problem?

Now, secondly how do we get x k plus 1 in x, the feasible set x, such that the function value decreases. So, these are two important points that one needs to consider when one writes a constrained optimization algorithm. And in the next few lectures, we will look at what are the stopping conditions for a constrained optimization problem and how do we find the new points x k plus 1? Now, there exists different types of algorithms to solve constrained optimization problems and depending upon the application one has to choose one of this algorithms. So, we will study those things in the next few lectures.

Thank you.