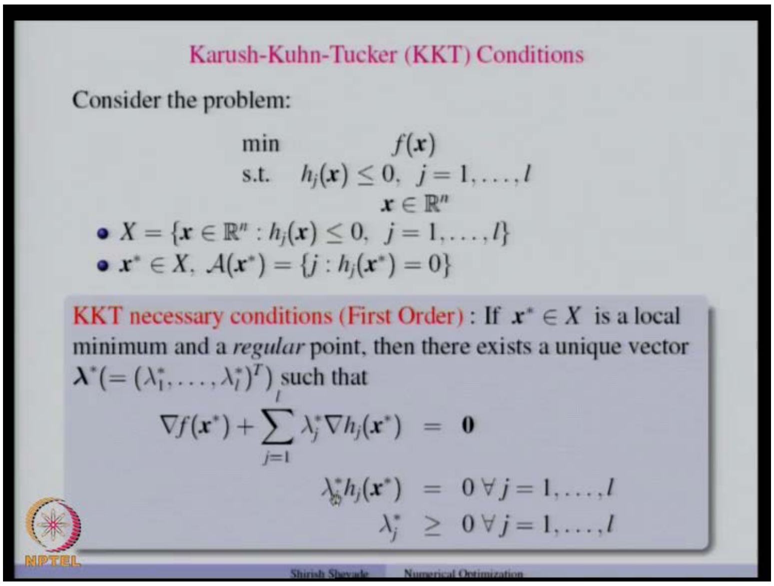
Numerical Optimization
Prof. Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 23
Constraint Qualifications

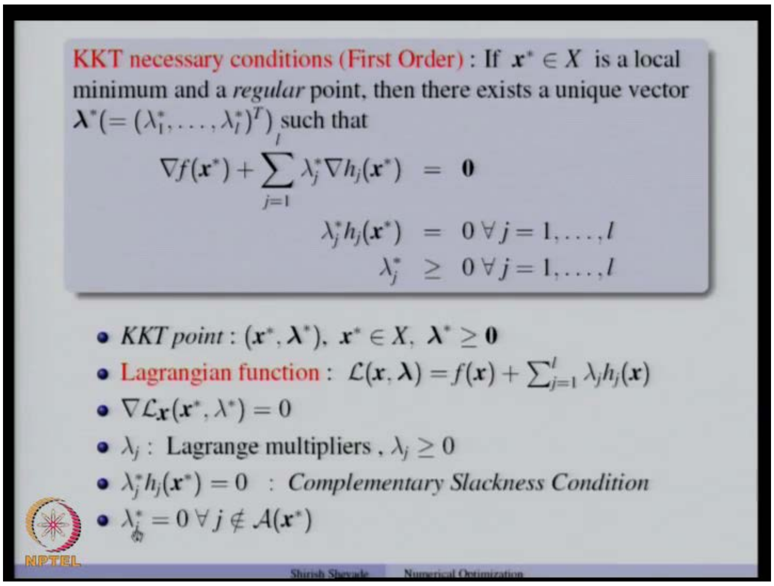
(Refer Slide Time: 00:31)



Hello, welcome back. So, in the last class, we started looking at KKT conditions for a constrained optimization problem. In particular, we considered this problem where we want to minimize f of x subject to the inequality constraints of the type h j (x) less than or equal to 0. There are one such inequality constraints. So, we define the constraints said to be capital X like this and for a given feasible point x star, we defined in active set A x star to be a set of all constraints which are satisfied with equality. So, set of all j's such that h j (x star) is equal to 0 and then, we have the necessary conditions for a local minimum. These are the first order condition because they use the first order derivative information.

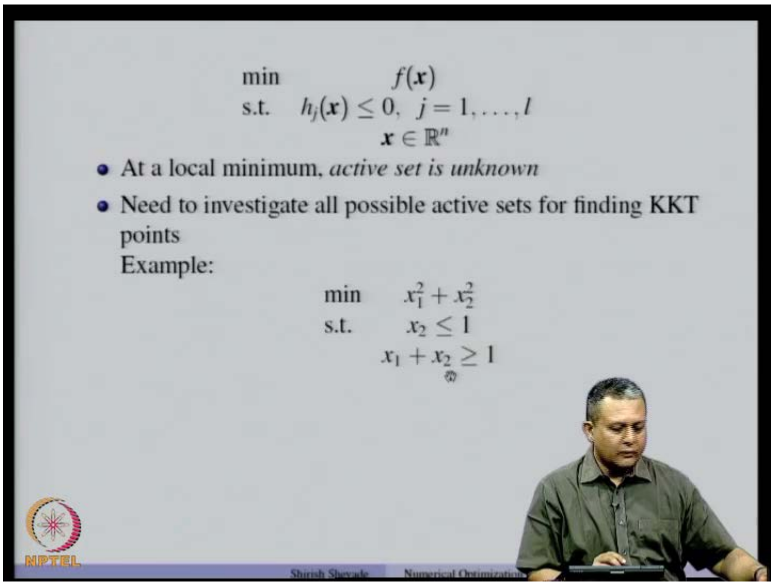
So, if we take a feasible x star, then that feasible x star is a regular point and is also a local minimum. Then, there exists a unique vector lambda star, such that these conditions hold. So, gradient effect star is written as a non-negative linear combinations of the gradients of the active in equality constraints and for the inactive inequality constraints, we can say lambda j star to be 0. So, in all we have all lambda j star non negative and on this condition holds.

(Refer Slide Time: 02:12)



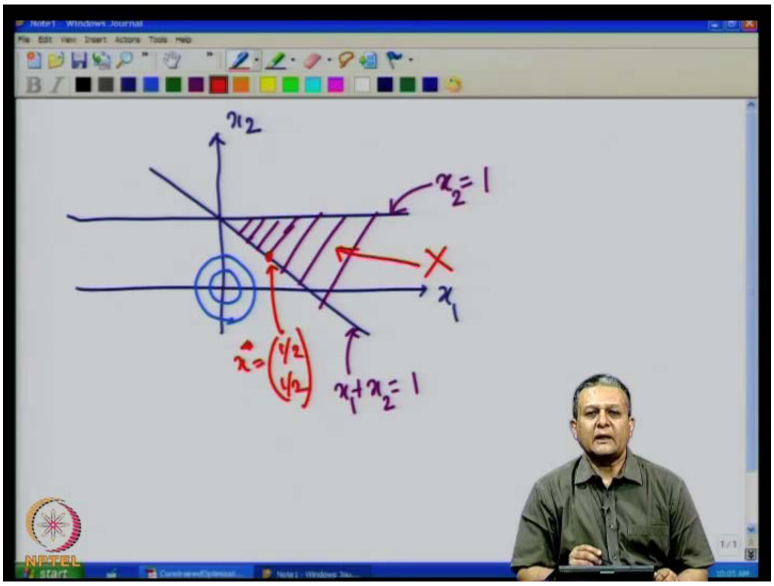
So, for the inactive inequality constraints, x j x star is less than 0. So, for those constraints lambda j star will be 0 and for the active inequality constraints, lambda j star will be non-negative. So, we saw that x star lambda star is KKT point, where x star is feasible point and lambda star is non-negative. So, such point we are going to call them as KKT points and if you define the lagrangian function to be f of x plus sigma lambda j h j x, where j is from 1 to l, then this condition can be written as the gradient of the lagrangian with respect to x evaluated at x star and lambda star 0 or evaluated at a KKT point is 0 and multiplies lambda j is or call the Lagrange multipliers and they are non-negative.

(Refer Slide Time: 03:16)



Moreover, we have this complimentary slackness condition which is mentioned here that lambda j star h j x star is equal to 0 and lambda j star is 0 for all constraints which are inactive at x star. So, all this analysis holds, provided the active set is known and typically in practice, the active set is not known.

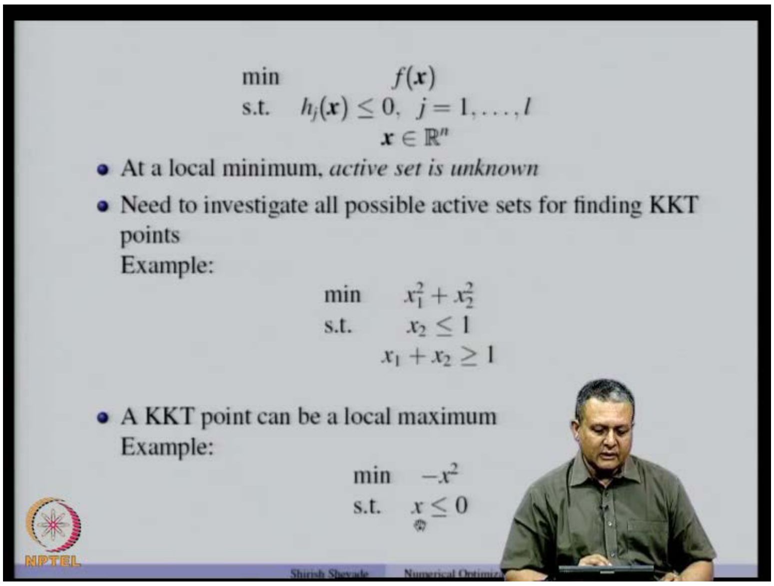
(Refer Slide Time: 03:56)



So, one has to investigate all possible active sets for finding KKT points and we saw one example in the last class that if you want to solve this problem, minimize x 1 square plus x 2 square subject to the x 2 less than or equal to 1 and x 1 plus x 2 greater than or equal to 1. So, we have x 1 x 2 and here, we have one constraint which is x 1 plus x 2 greater than or equal to 1 and other constraints which is x 2 less than or equal to 1. So, if you look at the feasible region, so this is the constraints which satisfied x 2 equal to 1 and this is the constraint which satisfies x 1 plus x 2 equal to our set x. If there were a feasible region where x 2 is less than or equal to 1 and x 1 plus x 2 greater than or equal to 1 and our aim is to minimize the objective functions which contours are like this.

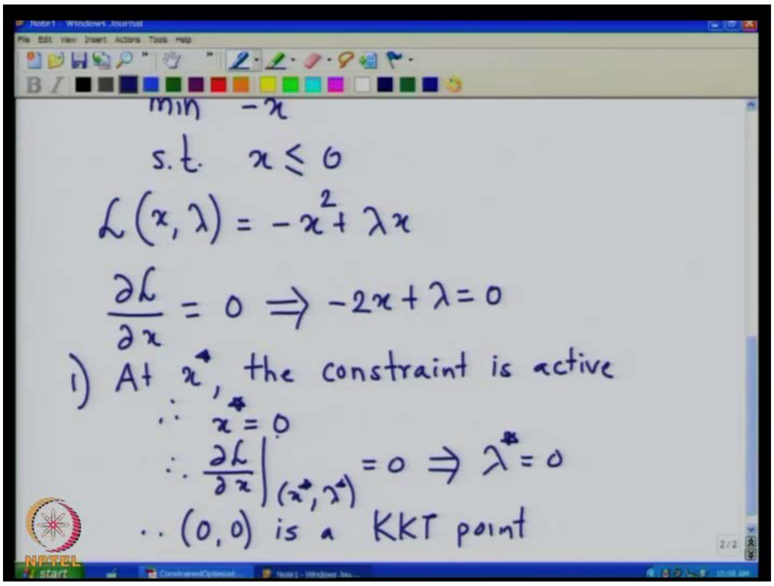
So, you will see that at this point, the minimum occurs and that point is, but then we do not know the active constraints beforehand. So, we have to check all possibilities. So, if we assume that at this point the minimum occurs, then we satisfy the KKT condition or only this constraint x 2 equal to 1 is active and this constraint is inactive. That means a solutions lies somewhere on this line. Then, what happens is we saw that the solutions cannot lie here and we saw in the last class that the solution indeed lies here, we were able to take a x star to be half-half and the corresponding lambda which is positive, such that this condition, the KKT conditions are satisfied. Now, I will leave it is an exercise to check whether the solution does lie on this line. So, in other words, we have to check all possible combinations of constraints and treat them as active sets and see, which are the possible KKT points.

(Refer Slide Time: 06:53)



So, we saw this example in the last class. Now, another important point that needs to be noted is that KKT point; it can also be a local maximum. So, satisfying KKT conditions does not mean that we will always get a local minimum. So, let us consider one example where we want to minimize minus x 1 minus x square subject to the constraints x less than or equal to 0. So, this is just a one-dimensional optimization problem.

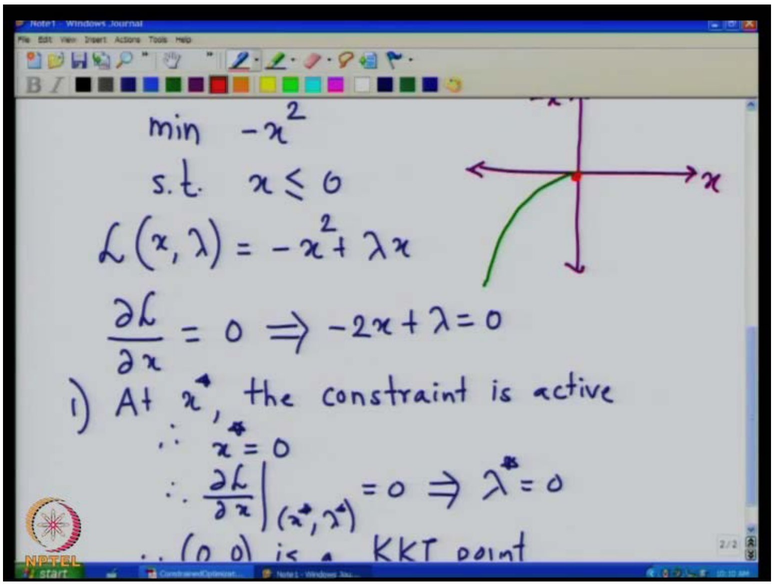
(Refer Slide Time: 07:32)



So, we want to minimize minus x square subject to the constraints x less than or equal to 0. Now, if you write the lagrangian, so lagrangian of x lambda will be minus x square plus lambda x. So, partial derivative of l with respect to s and that is 0 at the first order KKT condition. So, we will equate that to 0 to get a KKT point and this implies that minus 2 x plus lambda is equal to 0. Now, if you consider the case one, where we divide the constraints into two parts, so at the solution suppose this is active constraints. So, in other words, at x star, the constraints is there, only one constraints is there. So, suppose that that constraint is active.

So, therefore, x star is equal to 0 because at x star h j x star equal to is an active constraint. So, if we put x star equal to 0, what do we get? Therefore, gradient of l with respect to x evaluated at x star lambda star equal to 0 implies lambda star equal to 0 because x star is 0. Therefore, 0, of 0 is a KKT point because our KKT necessary condition just say that lambda star has to be non-negative and so in this case, it is 0 and a corresponding point x star is 0.

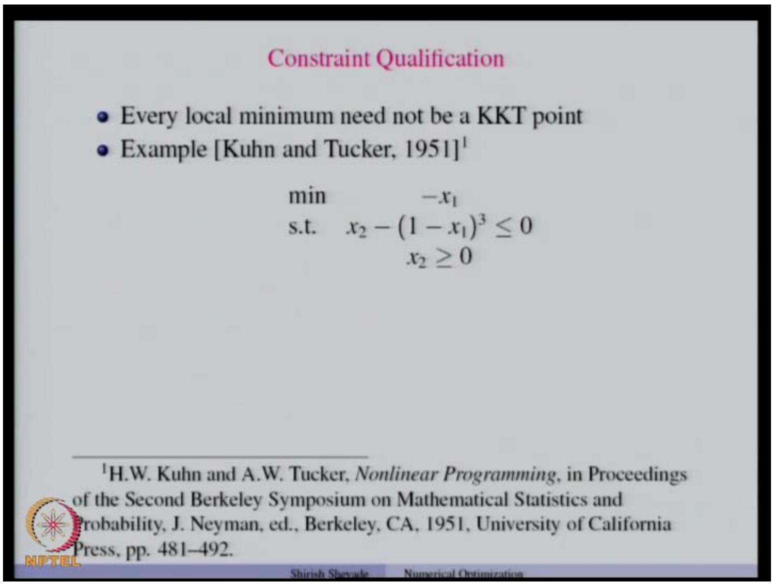
(Refer Slide Time: 10:15)



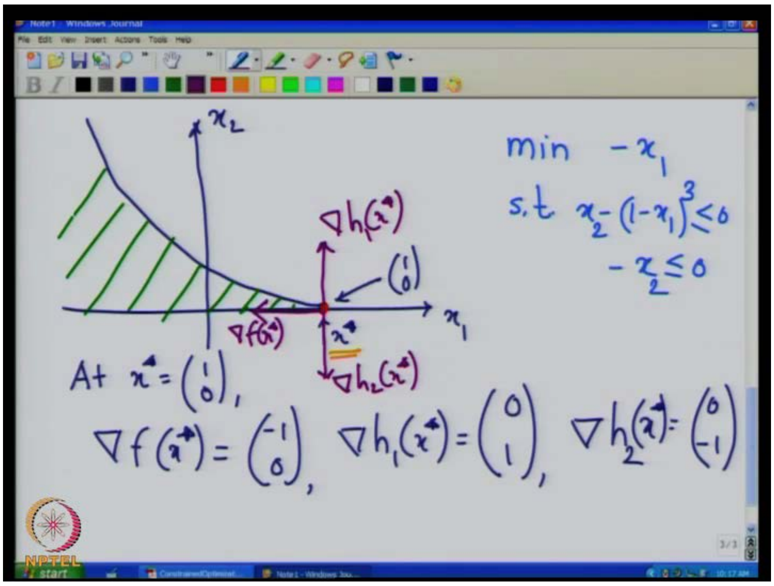
Now, let us see if we look at this function, so if you take the function minus x square and the constraints x less than or equal to 0. So, x that is less than or equal to 0 means we are interested in non-positive values of x and the function would be like this. So, you will see that the function is unbounded. So, the minimum of minus x square will not be active because the function is unbounded.

Now, let us look at the point. So, let us look at the point x star is equal to 0 which is this point which we were considering and we were able to get a non-negative lagrangian multiplier corresponding to x star and we got 0 as KKT point and we will see from this figure that in fact, this point is a local maximum. So, although this point satisfies the KKT conditions, it is the local maximum. Now, therefore, satisfying KKT conditions does not guarantee local minimum. Now similarly, one can work out the other condition that at x star, the constraints is inactive. That means, at x star lambda is less than 0 and then, one can find out what happens to this condition as well as the condition that lambda star has to be non-negative. So, I leave it as an exercise, ok.

(Refer Slide Time: 12:27)



(Refer Slide Time: 13:15)



Now, let us look at another important point that needs to be understood that is called a constraint qualification. Now, every local minimum that we get need not be your KKT point and one example which was given by Kuhn and Tucker sometime back in 1951 was problem where we want to minimize minus x 1 subject to re-constraint at x 2 minus 1 minus x 1 cube less than or equal to 0 and x 2 greater than or equal to 0.

So, let us look at this example. So, suppose this point is 1, 0 and so this is one constraint and the other constraint is x 2 greater than or equal to 0 and we want to minimize minus x 1 subject to the constraint x 2 minus 1 minus x 1 cube less than or equal to 0 and x 2 greater than or equal to 0. So, we will write it as minus x 2 less than or equal to 0. So, this is our constraint set and we want to minimize minus x 1.

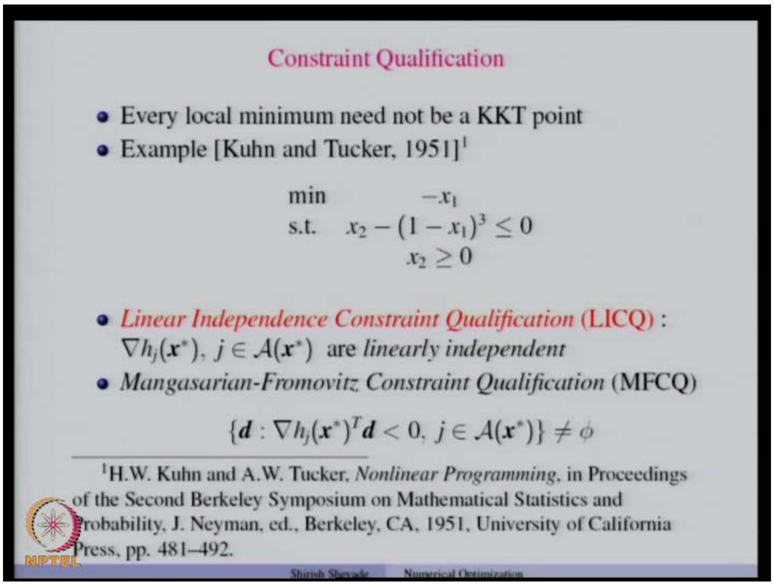
Now, let us look at this point. Now, at this point, we will see that this is point and this is the solution to this problem because we want to find out the maximum value of x 1, where these constraints are satisfied and that turns out to be this value, but let us look at the gradients of the objective functions and the constraints. So, this is x star. So, x star is given here. Now, at x star which is equal to this point, what is the gradient of f of x star? So, gradient of f of x star will be minus 1, 0.

Let us take the first constraints. So, gradient h 1 x star will be 0, 1 and where evaluating it at x 1, 0. So, the x 1 component of this first gradient will be 0 and the gradient of the second component, the gradient of the inequality constraints will be 0, minus 1. Now, let us plot these constraints on this figure. So, gradient f x star is pointing in this direction. So, this is gradient f x star and gradient h 1 x star is pointing in this direction. So, gradient h 1 x star and gradient h 2 x star x star is pointing in this direction.

Now, the first order KKT condition says that gradient f x star is written as a non-negative linear combinations of the gradients of the active inequality constraints. Now, at this point x star, both the constraints are active. So, both the inequality constraints are active at this point x star and you will see that gradient h 1 x star and gradient h 2 x star, all are not linearly independent in this two-dimensional space. Therefore, since they are not truly linearly independent vector seen two-dimensional space, therefore, they cannot form basis for a two-dimensional space. Therefore, gradient f x star cannot be written as a non-negative linear combinations of gradient h 1 star h 1 x star or gradient h 2 star.

In fact, you will see that now gradient f x star is an orthogonal 2 gradient h 1 x star as well as gradient h 2 x star. So, even though x star is an optimal point of this problem. It is not a KKT point and that is because x star is not a regular point. The gradients of the active inequality constraints are not linearly independent in this case and therefore, KKT conditions are not satisfied at this point.

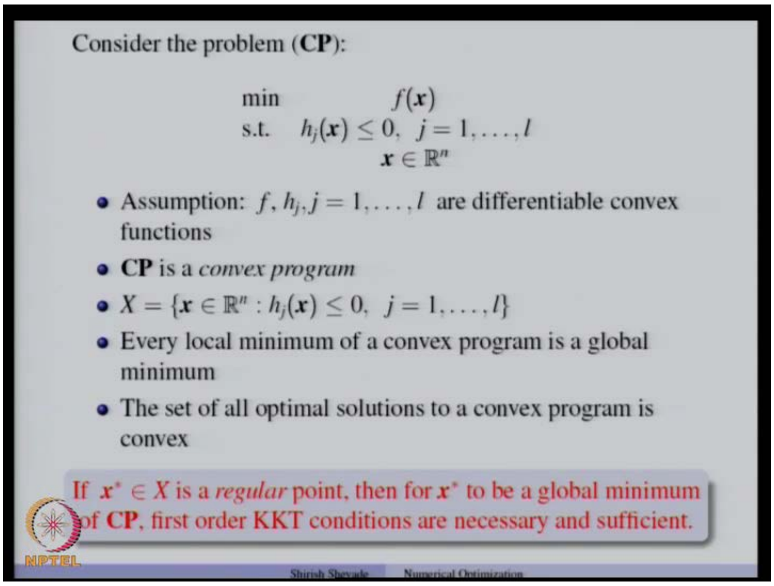
(Refer Slide Time: 18:40)



So, it is very important to have the regularity assumption for getting the KKT points and therefore, this linear independence constraint qualification is very important constraint qualification condition and that says that the gradients of the active constraints active at x star how to be linearly independent. So, that implies that x star becomes a regular point and once it is a regular point, once we have this linearly independent vector, we can write the gradient of f of x star as non-negative linear combinations of the variance of these points. If they are not linearly independent, we can write gradient f x star as the non-negative linear combinations of x.

So, this is the very important condition that needs to be satisfied. Then, we can say that under the linear independence constraint qualification if x star is the local minimum, then first order KKT conditions are satisfied or they are necessary. Now, Mangasarian and Fromovitz also gave another constraints qualification condition and that says that if we consider the problem of minimizing f x subject to h j x star less than or equal to 0, then there exists at least one direction which makes an off choose angle with all the gradients of the active inequality constraints. So, gradient h j x star transpose d is less than 0 for all j belonging to the set of active constraints at x star.

(Refer Slide Time: 20:37)



So, there exists at least one direction which satisfies this. So, this is another constraint qualification condition that was proposed by Mangasarian and Fromovitz. Now, let us consider a problem to minimize f x subject to the constraint h, j, x less than or equal to 0, but this time we make an assumption that f and h, j, x are differentiable convex functions. So, the important point here is that both f as well as h, j, x are convex functions and also they are differentiate.

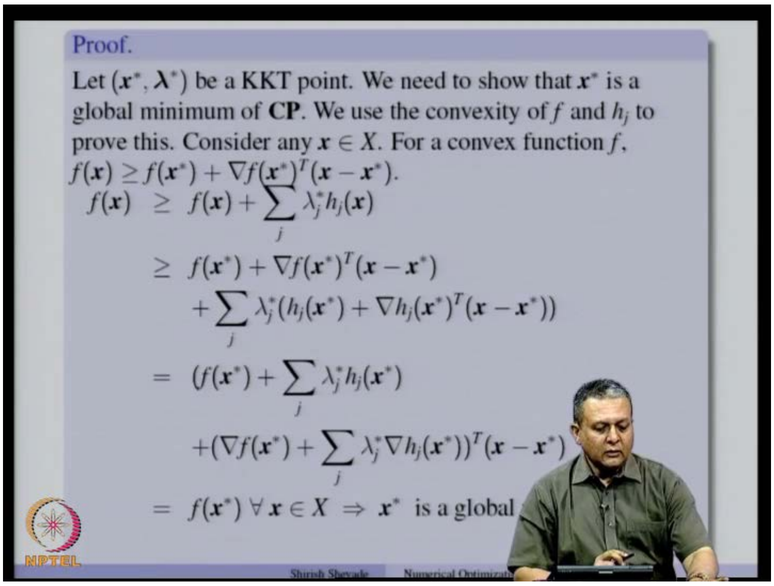
Now, under this situation turns out that the first order KKT conditions are sufficient if f star is the regular point. Now, we are optimizing or minimizing a convex function subject to the constraint that h j x is less than or equal to 0. Now, as we have seen earlier that if h j x is less than or equal to 0, where x h is a convex function and then, a set x such that h j x less than or equal to 0 is a convex set. Now, if you combine different convex sets or intersect different convex set, then we know that the intersection of convex set is also the convex set.

So, the constraint here is the convex set. The objective function to minimize is the convex function. So, such a program is called a convex program and clearly the set x is a convex set. Now, there are two important results that I would like to mention here to a convex program and the first result is that every local minimum of a convex program is a global minimum. We saw a variant of this result when we studies convex functions that under the consent case, we saw this result and those ideas can be extended to a constraint problem like this which is in this case is a convex program.

So, every local minimum of a convex program is a global minimum. So, there is no question of local minima as for as convex programs are concerned. Every local minimum is a global minimum and not only that, all these global minima when they are found, they form a convex set. So, the set of all optimal solutions to a convex program is convex. So, this is a very important result again that not only do we have the problem of local minima for convex programs, but the set of all solutions to a convex program is a convex set.

So, this result again we have seen earlier that even if you take a unconstraint of optimization problem where we want to minimize a convex function, then the set of all optimal solutions is a convex set and that result can easily be extended to a constraint problem constraint convex program which is given here that a set of all optimal solution form a convex set. So, here we have important result that says that x star is a regular point and then, for x star to be a global minimum of a convex program c p, the first order condition KKT conditions are necessary and sufficient.

(Refer Slide Time: 24:56)



So, we have already seen that if x star is local minimum, then first order condition are necessary. Now, we will show that for convex program, these conditions are sufficient and to show that we will need the convexity of f and the function h j x. So, let us see how to prove that. So, let us take a KKT point x star lambda star and since, it is a KKT point x star belongs to the feasible region and lambda star is a non-negative vector. Now, what we want to show is that x star is a global minimum of a convex program, where f is a convex function and h j are convex functions and we also assume that they are sufficiently smooth. Now, in this case, we just need a first order derivative. So, it is enough to assume that they belong to class c 1 of function. Now, to prove that the KKT, first order KKT conditions are sufficient, we will use the convexity of the functions f and h j and for that purpose, let us consider any point x in the feasible region.

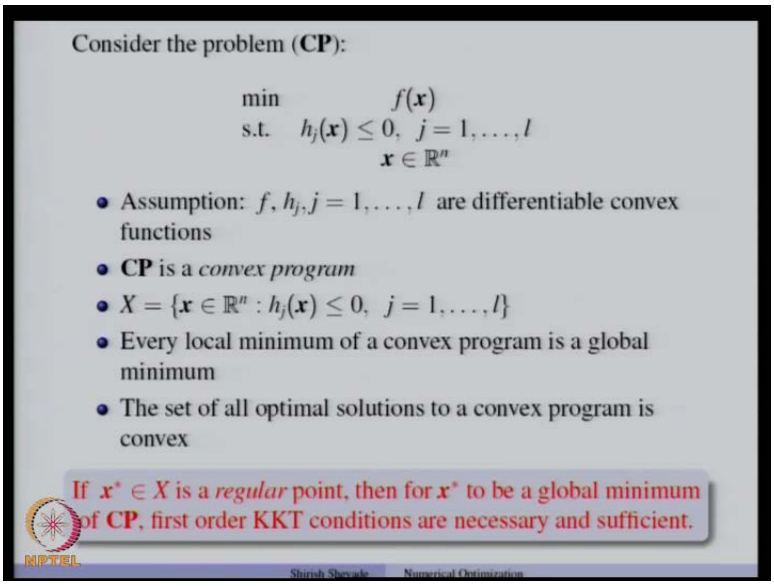
Now, we have seen this result earlier when we studied convex functions that the (( )) approximation of a convex function at x star does not over estimate the function. So, f of x is always greater than or equal to f of x star plus gradient f x star transpose x minus x star. So, these results we have seen earlier. So, let us make use of these results to prove that a KKT point under the regularity assumption is a local minimum is a sufficient condition for a local minimum of a convex programming problem. Now, since x is a feasible point, h j x is less than 0 and x star lambda star is a KKT point. So, lambda star has to be non-negative. So, this quantity here is a non-positive quantity and therefore, we can write f of x to be greater than or equal to f of x plus sigma j lambda j star h j x because f j x is less than or equal to 0 and lambda j star greater than or equal to 0.

Now, let us make use of these results. So, expand f x around f x star and use this inequality. So, f of x is greater than or equal to f x star plus gradient f x star transpose x minus x star and since, each h j x is also a convex function, h j x is greater than or equal to h j x star plus gradient h j x star transpose x minus x star. So, in this we have used the convexity of both f and h j x. Now, let us rearrange these terms together. So, let us combine this term with this term and then, all the terms in all the gradients of f and gradient of h j x will combine them together. So, this gives us, so this quantity is nothing, but f x star plus h j x star sigma lambda j star h j x star plus. Let us combine terms in one gradients, so gradient f x star plus sigma lambda j star gradient h j x star transpose x minus x star.

Now, we have seen that x star lambda star is a, we have assumed that x star lambda star is a KKT point and satisfies necessary KKT conditions. So, since the complimentary slackness condition is satisfied, we have lambda j star h j x star to be 0. So, every term in this summation is 0.

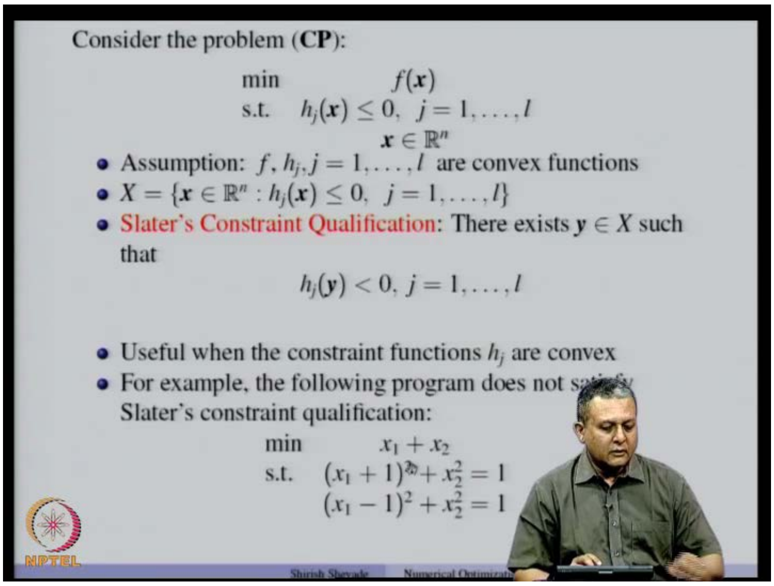
Now, let us look at this point. Since, the first order KKT conditions are satisfied gradient f x star plus sigma j lambda j star gradient h j star is 0. So, this is 0, this is 0 and what we are left with? It is left with f of x star and therefore, f of x is greater than or equal to f of x star for all feasible x and that means that x star is a global minimum of the convex program. So, x star is a global minimum of the convex program that we have considered.

(Refer Slide Time: 20:37)



So, if we find the KKT points of a convex programming problem, then under the regularity assumptions, those KKT points are indeed or those KKT points give us the optimal solution x star to a given convex program. So, this is a very important result as far as convex programming problems are concerned. Now, interestingly this result can be extended when we add the equality constraints to this program. The only thing that one has to keep in mind is that when we add the equality constraints, those equality constraints have to be of the type a i transpose x minus b i because only then on the constraint set will remain a convex set. So, we will see those results when we introduce the equality constraints to our optimization problem. So far, we have been dealing only with the problems of this type where we want to minimize f x subject to any quality construction, but those results can easily be extended to general convex programming problem. We will see those things sometime later.

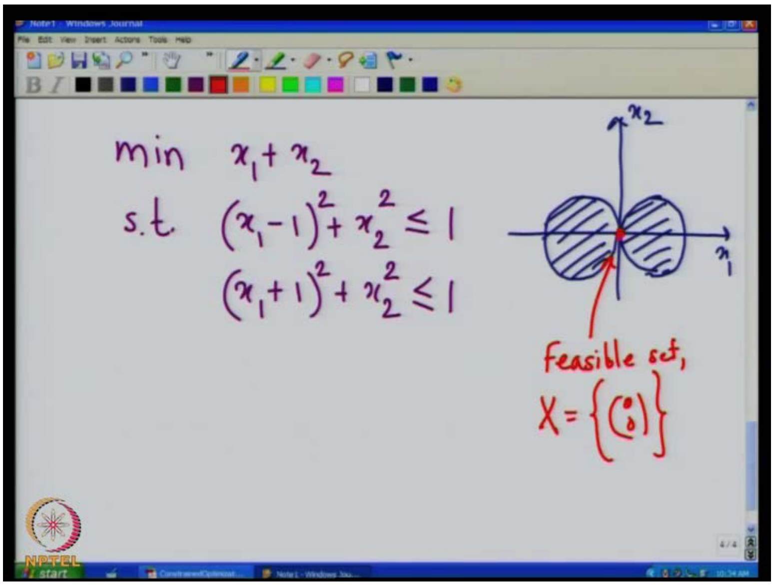
(Refer Slide Time: 30:57)



Now, let us again go back to our convex program where the function f is convex function and h j x are again convex function and the constraint set is a convex set. So, the feasible set x becomes a convection because it is an intersection of all convex set. Now, for convex programs, there is a constraint qualification condition propose by later and that condition says that there exists of feasible point y, such that h j y is less than 0 for all for all the inequality constraints.

So, what this means? This means that there exists some feasible point which lies in the interior of the feasible set x. So, that means that the interior of the feasible set is not empty. So, this Slater's constraints qualification is useful especially for convex program problems and it is easy to check that this condition to check whether this condition is satisfied or not.

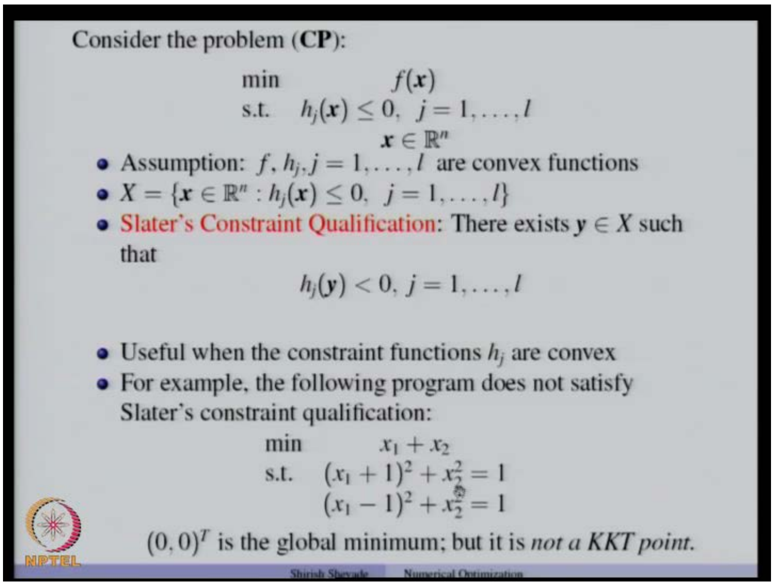
(Refer Slide Time: 32:57)



So, typically for convex programs, we will use Slater's constraints qualification and to show the importance of this Slater's constraints qualification, let us consider an example where we want to minimize this problem. So, suppose if we consider this problem minimize x 1 plus x 2 subject to the constraints that x 1 minus 1 square plus x 2 square less than or equal to 1 and x 1 plus 1 square plus x 2 square is less than or equal to 1.

So, the first constraint is a circle of radius one centered around 1, 0 in the second constraint is again, so this is the point inside the second circle and these are the points inside the first. So, the feasible region will be only this point feasible set x is nothing, but the origin, the dashed portion here shows the points inside the circle. So, from one set, so this is the feasible region corresponding to the second constraint. This is the feasible region corresponding to a first constraint and when we intersect, we get only one point which is a single term point.

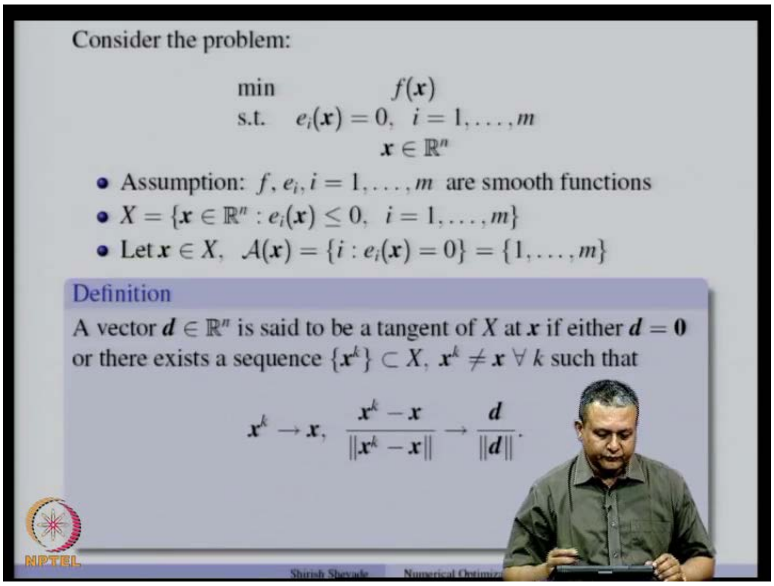
(Refer Slide Time: 36:00)



Now, this point, this feasible set being a single term set has no interior. So, one can check that this point is a local minimum, but it does not satisfy the Slater's constraints qualification condition. So, this point which we have got which is the origin, which is the feasible set, is the global minimum, but it is not a KKT point. So, one can check that this point is not a KKT point, although it is a global minimum and that to happen because we had a constraint set which is also single term set in this case and that did not satisfies Slater's constraints qualification.

So, many times for convex programs which is possible to have a minimum, but that minimum may not satisfy KKT conditions and that is mainly because the Slater's constraint qualification is not satisfied, that is that did not exist a point which slice in the interior of the feasible set or in other words, the feasible set does not have interior or has empty interior, ok.

(Refer Slide Time: 37:22)



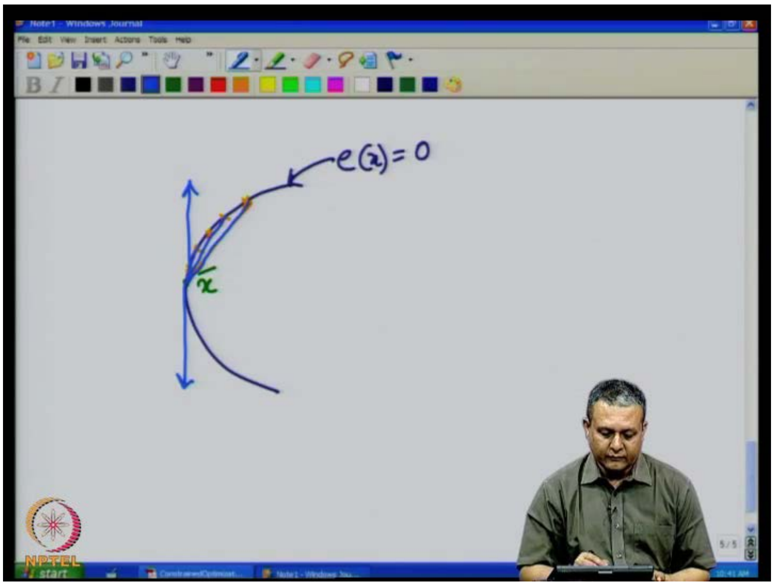
So far, we have studied problems of type where belongs to minimize effects subject to the constraints h j x is less than or equal to 0. Now, the idea that we have studied so far cannot be directly used for equality constraint problem. Now, let us start looking at the equality constraint problems. Now, you will see that the constraints are of the type e i x equal to 0. I am going from 1 to m. Now, there could be e i x could be as simple as a i transpose x minus b i equal to 0 as it could be different function non-linear function of x.

Now, we can arrive at some algebraic condition for the local minima of this problem. So, that is what we are going to study now. So, let us again assume that the objective function f and e i’s are smooth functions and for the first order conditions will require them to continuously differentiable, and for the second order condition will require them to be quite continuously differentiable. So, here I just mention them as smooth functions. So, let us collect all possible x. We satisfy e i x equals to 0, it should be e i x equals to 0.

Now, let us consider a feasible point and then, see what constraints are active at the feasible points. So, as per our definition, the set of constraints which are satisfied with equally, so a set of all i’s such that e i is equal to 0 and since, x is a feasible point. So, that means that e i x equal to 0 for all i which means that all the equality constraints are active at any given feasible x. So, this set is nothing, but all m equality constraints. Now, we look at the definition of a tangent of the feasible set at a given point x. That definition helps us to characterize the feasible directions using algebraic conditions. So, a d dimensional vector d is said to be a tangent of x at x tangent of the feasible set at a given point x.

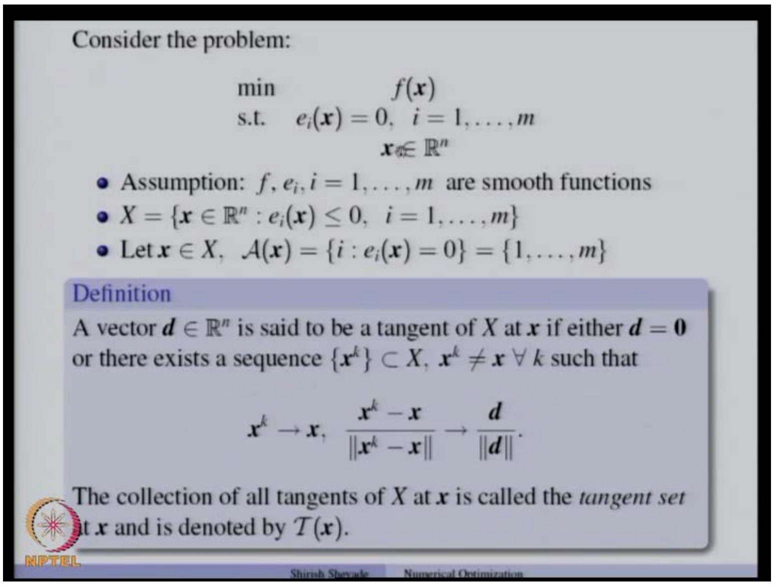
Remember that x also is part of the feasible set here. So, if either d is equal to 0 or if we consider a sequence x k in the feasible set, where x none of the x k is equal to x for all k, such that the sequence converges to x and if we take the quad joining the x and x k and then, like though vector in the normalize vector x k minus x by norm of x k minus x that converges to d by norm d.

(Refer Slide Time: 40:59)



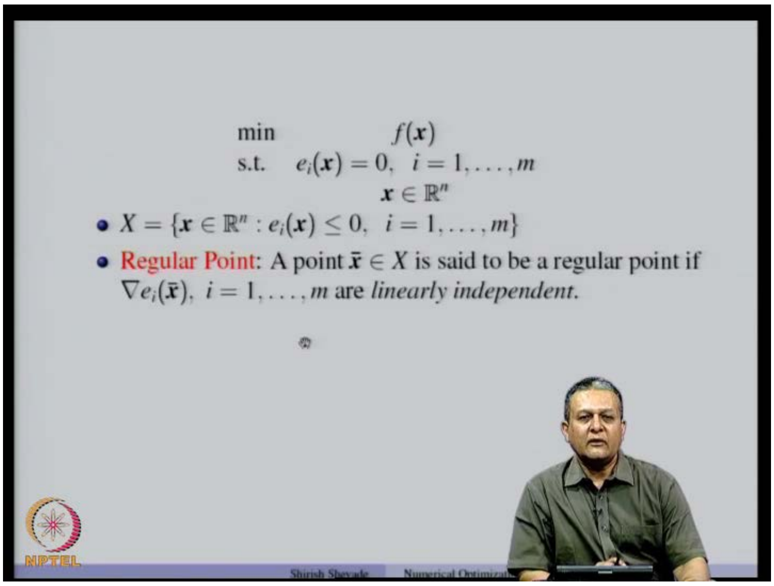
So, such a vector d is called a tangent vector. So, suppose we have aa constraints. So, this is the constraints e x equal to 0. Now, if we take a point x bar. Now, we take a sequence x k which is part of this feasible set. So, let us take the sequence, point which converges to this point. Now, if we take a quad joining each of this, so in the limit this quad will have this direction and this is the tangent to this surface at x star can also have a tangent which is in this other direction by considering another sequence. So, this becomes a tangent set to this feasible region at x bar. Now, in two-dimensional space, this tangent, this one line, now three-dimensional space, it will be a tangent length and so on and so forth.

(Refer Slide Time: 43:26)



So, we collect all feasible sequences in the region which converse to x bar and then, if we take a quad joining those x k and x bar and normalize them to a unit vector, then we get a and then, a find limiting direction of those quads that will give us the tangent to the feasible set at x bar. So, if we collect all the tangents of the feasible set x at x at a feasible point x and that is called the tangent set, and we are going to denote it by T x. So, T x denotes the set of all tangents of the feasible region at a given point x.

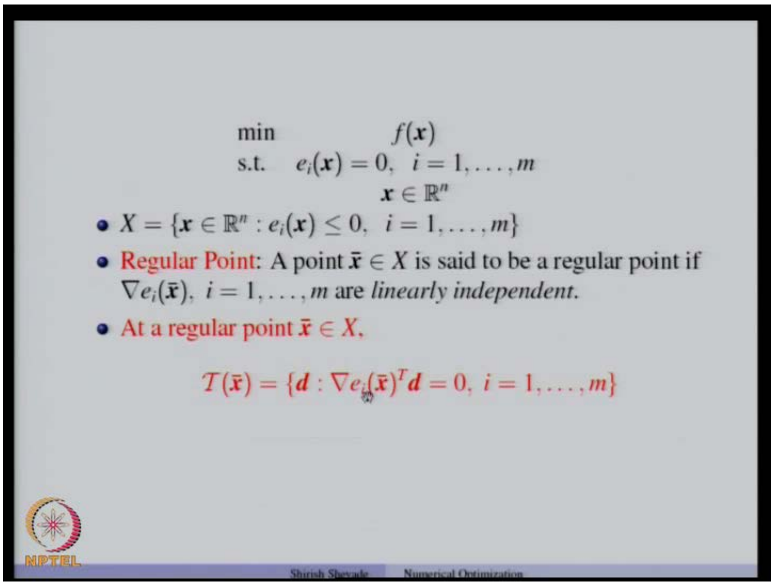
(Refer Slide Time: 44:18)



Now, let us see how to characterize this tangent set? If you recall for the inequality constraints, we use the feasible set f of x script f x and then, that script f x feasible set was characterized using f tilde. So, similarly, how do we characterize T x? So, let us look at that part. Now, so far that purpose will need the definition of a regular point. So, a point x bar in the feasible set, a set to be regular point if gradient e i x bar are linearly independent.

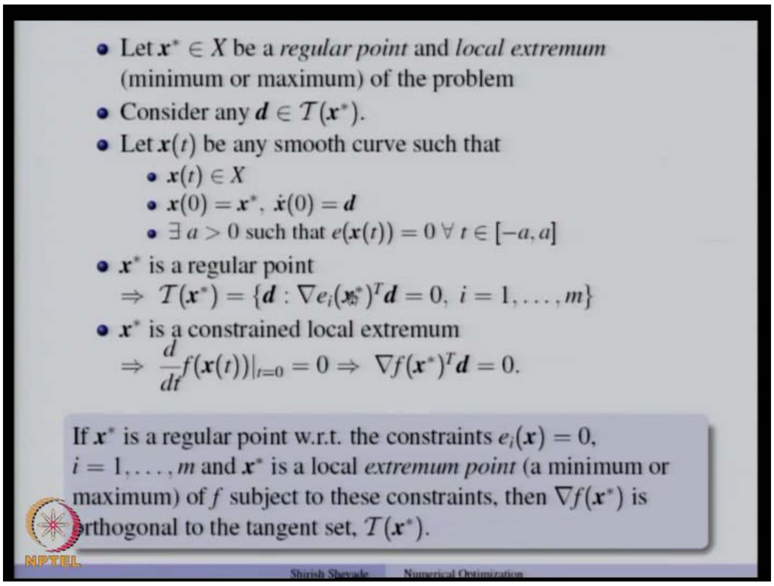
So, this definition is similar to what we saw for inequality constraints, though only difference is that for the inequality constraints, we just had to take only those inequality which we were active or for which h j x bar was 0 and then, the gradient of those points have to be linearly independent while here since we are talking about the active equality constraints. All the constraints are active at any given point x bar and therefore, a point x bar in the feasible set is a regular point, if the gradients of all active or all equality constraints are linearly independent or gradient e i x bar i going from 1 to m is a linearly independent set.

(Refer Slide Time: 46:02)



Now, this result will not prove these results, but use these results that at a regular point x bar, the tangent set is characterized by the set of all directions d which are orthonormal to gradient T i x bar which are orthonormal to the gradients of the equality constraints. So, note that we are again using all i’s going from 1 to m. That means, we are using all active or all the equality constraints, since they are active at any given feasible x bar.

(Refer Slide Time: 47:07)



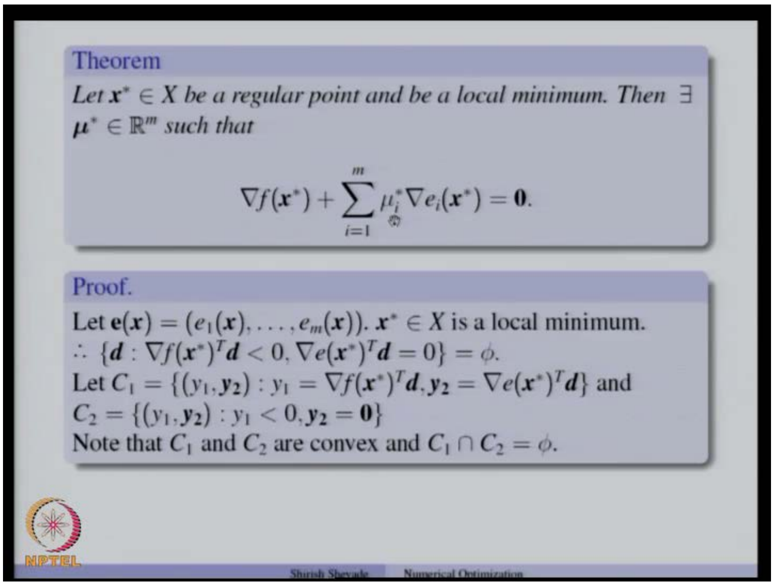
So, the set of all d’s which are orthogonal to the gradients of the equality constraints at x bar. So, under the regularity assumptions, one can show that T x bar is this set and we will use this fact to derive conditions for the local minimum of constraints problem. So, let us assume that x star which is feasible point, also a regular point and is the local extremum. Now, note that we have mentioned here local extremum which means I can be a local minimum or a local maximum.

Now, let us consider a new vector d which is in the tangent set T x. So, since it is a regular point, we have already seen that gradient e i x star transpose d equal to 0 for all i going from 1 to m. Now, let us consider us any smooth curve in the feasible set. Let us denote that curve by x t, where t is a parameter is a real number and x t denotes some point x in the feasible region. So, all x t's are feasible, moreover at t is equal to 0. We have x 0 equal to x star and then, the tangent at 0 where going to denote by d and there exists some positive a, such that e x t equal to 0. So, that means that around or in a close interval minus a to a, if you take any t, then e of x t is 0. So, this is some set of feasible points in the neighborhood of x 0 and the points are parameterized by the parameter t. Now, we have assumed that x star is a regular point. So, that means that the tangent set at x star which is characterized by the set of all directions which are orthogonal to the gradients of the equality, all the equality constraints.

Now, x star is also given to be a local extremum of a given objective function subject to the constraints. So, that means that the derivative of f of x t with respect to t evaluated at t equal to 0 because x 0 is nothing, but x star. So, these derivative is 0 and that means that gradient f x star transpose d is 0 or d is orthogonal to gradient f x star. Now, we will see that we have considered any d which was from the tangent set and the tangent set is characterized by these under the regularity assumption, and if x star is a local minimum, then this holds. So, from this we will see that gradient f x star, this orthogonal to gradient e i x star for a lies or in other words, gradient f x star is orthogonal to the tangent set T x star.

So, if x star is a regular point with the constraints e i x is equal to 0 over i going from 1 to m and x star is a local extremum point, whether it is a minimum or maximum, it is a local minimum of x this subject to these constraints, then gradient f x star is orthogonal to the set T x star. So, this is a very important observation that under these conditions, the gradient of the objective function is orthogonal to the tangent set T x star and tangent set T x star is nothing, but the set of all directions which are orthogonal to gradient e i x star. So, we will use this observation to derive the optimality conditions for the local minimum of an equality constraint problem.

(Refer Slide Time: 51:55)



Now, here is the important theorem which says that if x star which is a feasible point is also a regular point and is a local minimum, then there exists some mu star which is in m-dimensional space, such that gradient f x star is a linear combination of gradient e i x star. Now, you would notice the difference between the necessary condition that we obtain for inequality constraints problem and the necessary condition for an equality constraint problem.

So, the important deference says that the multipliers corresponding to this gradient e i x star are not constraint to be non-negative. So, mu i star, this comes from the space of m-dimensional space. So, this m corresponds to the set of or the number equality constraint. So, there is one multiplier associated with each equality constraints and since, all equality constraints are active, so there exists m multipliers corresponding to all the active equality constraints and gradient f x star is written as a linear combination of gradient e i x star. Now, so what this means is that if we collect all active equality constraints find out their gradients, then if under the regularity assumption we can always write gradient f x star to be linear combination of gradient e i x star.

So, let us look at the proof of this theorem. Now, I have combined all the constraints and write them in the form e 1 x to e m x as e x. So, let e x denote the set of all equality constraints. Now, x star is a local minimum. So, that means that the set of all d's, such that gradient f x star transpose d is less than 0 and gradient e x star transpose d equals to 0 because this is a condition corresponding to the constraints, this is the condition corresponding to the objective function. So, we do not get a descend direction d which is feasible and therefore, we can write it as this condition.

Now, if we define two sets c 1 and c 2 to be all pairs y 1 y 2, note that y 2 is a vector here, where y 1 is gradient f x star transpose d and y 2 is gradient e x star transpose d because e is matrix. So, gradient e x star transpose d is vector and then, you take another set c 2 where y 1 is less than 0 and y 2 is equal to 0. Then, you will see that this is a convex set, c 1 also is a convex set. So, c 1 and c 2 were convex set and they have an empty intersection. So, if we have two non-empty convex set which are these joint, then we saw on the result earlier, we say that there exists (( )), which separates c 1 and c 2. So, it makes use of this result to get this condition and we will see that in the next class.

Thank you.