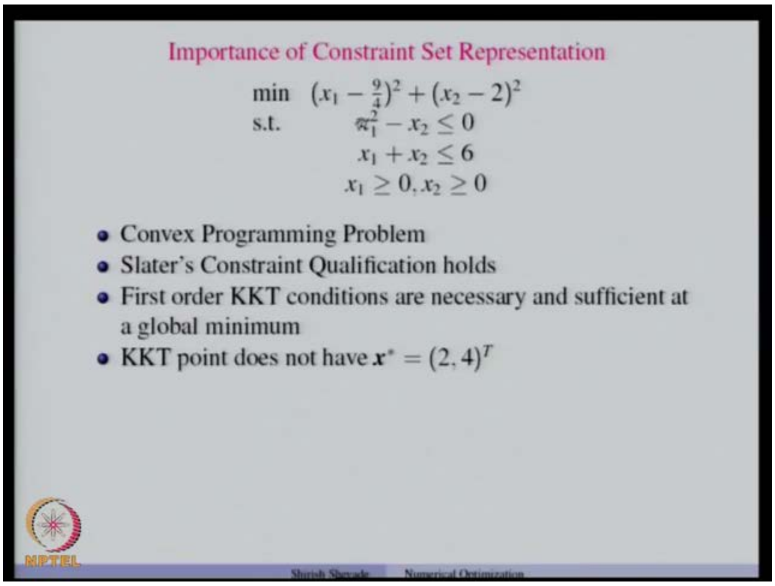
Numerical Optimization
Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 26
Second Order KKT Conditions (Contd)

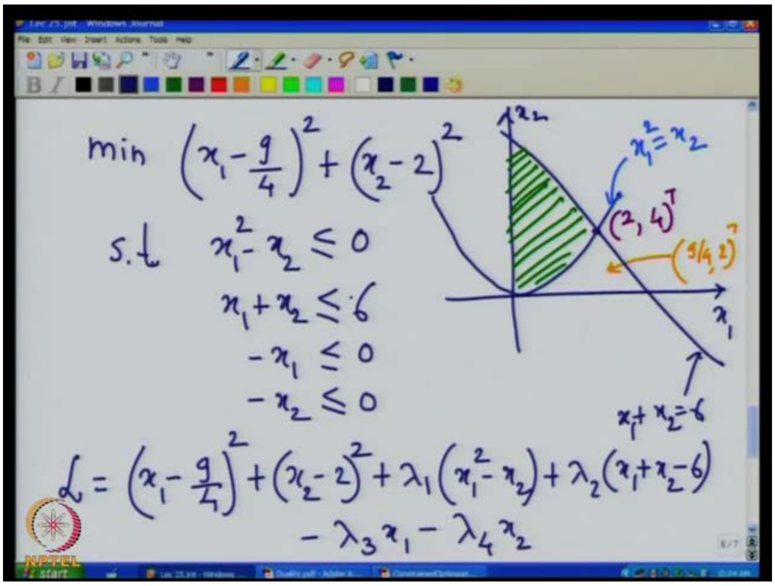
Hello, welcome back. In the last class we were looking at some examples related constraint optimization problems. And we were checking the conditions about the optimality in particular we are looking at first order KKT conditions.

(Refer Slide Time: 00:43)



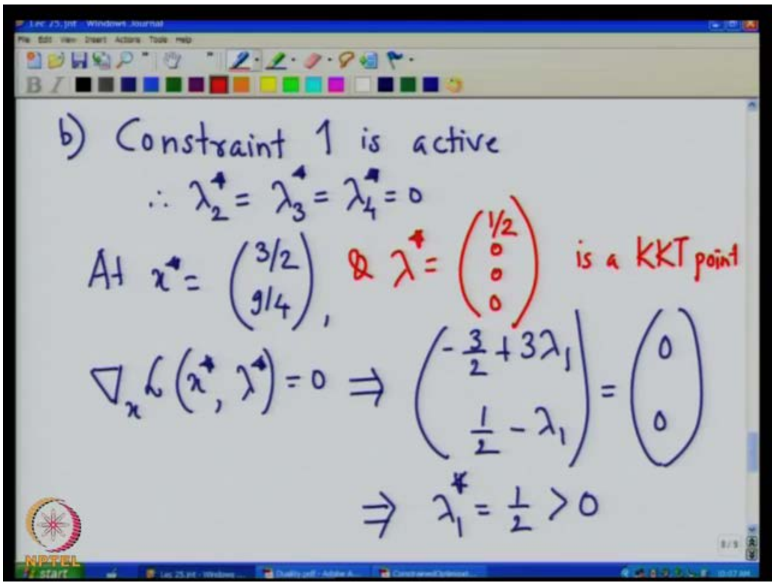
So, this was the example that we are looking at so minimize x 1 minus 9 by 4 square plus x 2 minus 2 square and subject to these constraints, and you have a nonnegative constraint on x 1 and x 2. And, we saw that the constraint set is convection, the objective function is a convex function. So, this is a convex programming problem. Further there exists at least one point in the inter here. So, let us constraint qualification holds. And it is enough for us to find out KKT point which satisfies first order conditions, because under the convexity; this first order KKT conditions are necessary and sufficient.

(Refer Slide Time: 01:31)



So, the constraint set it is shown her by the shaded region, and this was the problem that we want to solve. So, we checked that this 0.24 does not satisfy the KKT conditions. And, that is mainly because the Lagrangian multipliers corresponding to the second constraint, second unit quality constraint become negative. And, therefore this point this point cannot be a KKT point. So, now let us assume that the first constraint x square minus x 2 less than or equal to 0 is active. So, this is the first constraint; and suppose that that is active. Now, let us consider that case.

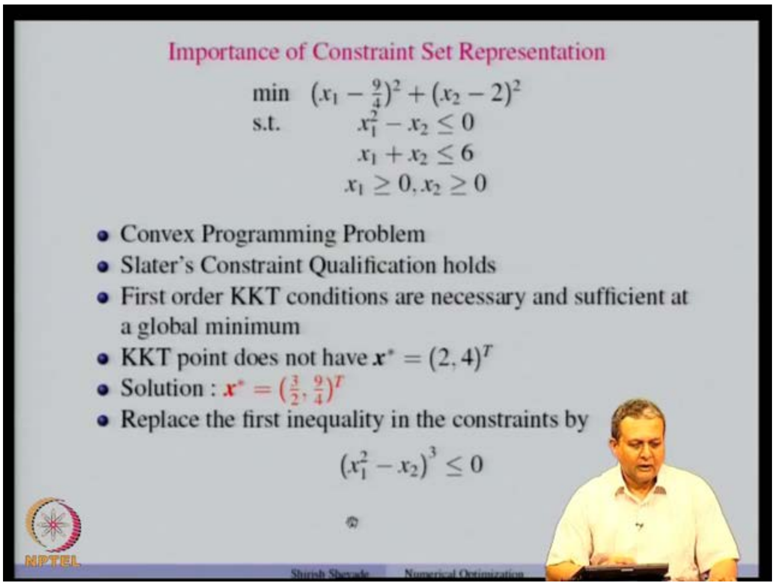
(Refer Slide Time: 02:35)



So, let us assume that so let us consider the part b where constraint 1 is active, and since constraint 1 is active and the rest of the constraints are inactive. So therefore at optimality what we expect is that the Lagrangian multipliers corresponding to the remaining constraints are all 0. So, lambda 2 star equal to lambda 3 star equal to lambda 4 star equal to 0. Now, let us consider the point at x star to be 3 by 2 and 9 by 4; at this point let us see what happens? So, gradient of L x with respect to x star lambda star equal to 0. So, if you look at the gradient, so if you look at the Lagrangian which is given here; and if you take the gradient of that with respect to x and evaluate it at to given point x star. And, you know that lambda 2 star, lambda 3 star, lambda 4 star equal to 0. Our aim is to get lambda 1 star and that is what we will do now.

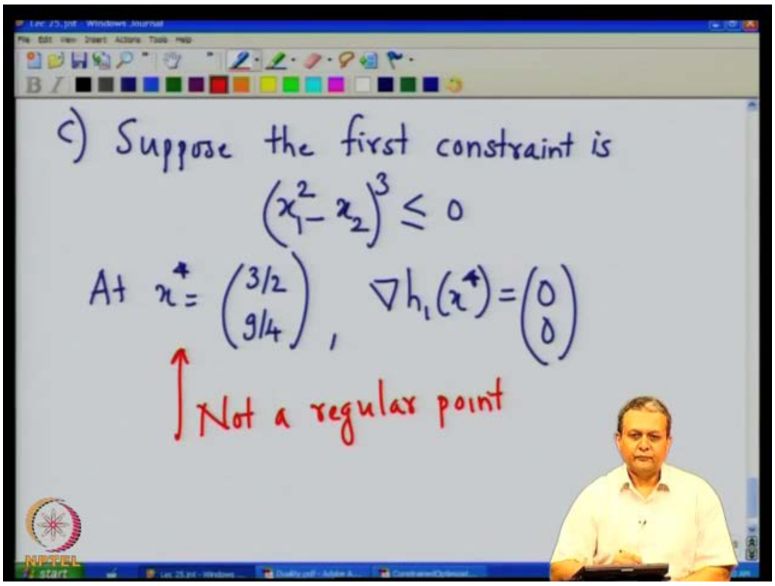
So this implies minus 3 by 2 plus 3 lambda 1 and half minus lambda 1 is equal to 0. So, we have substituted this value of x star and lambda star. So, we just have equations in terms of lambda 1; and this implies lambda 1 star to be half and that is greater than 0. So, this will satisfy our requirements. So, therefore this point x star equal to 3 by 2 and 9 by 4 and lambda star equal to half 0 0 0. So, x star lambda star is a KKT point because it satisfies all our requirements. And, therefore the 1st order necessary condition are satisfied and this being a convex programming problem. These conditions are also sufficient as the latest constraint qualification holds. Now, let us consider the problem.

(Refer Slide Time: 05:34)



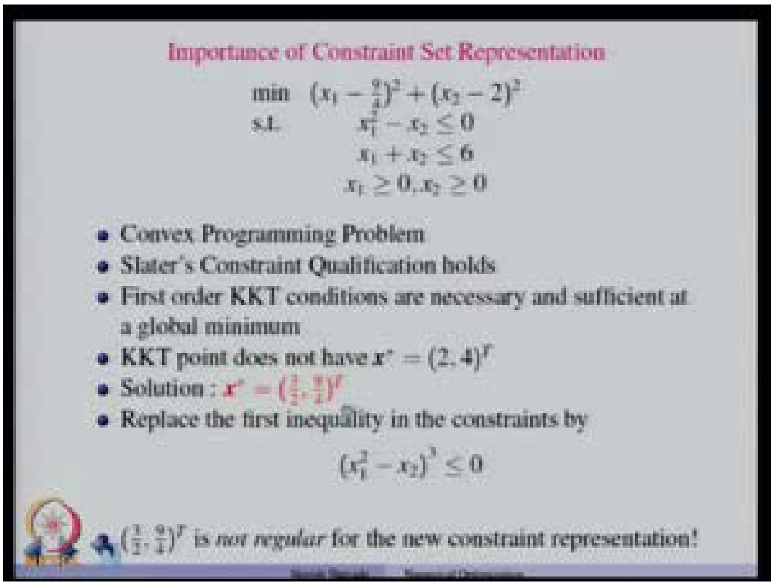
So, as we saw that x star equal to 3 by 2 9 by 4 is the strict local minimum of this problem. Now, suppose we replace this the first inequality constraint by the following inequality constraint; that x 1 square minus x 2 square the whole cube is less than or equal to 0. So the other constraints remain the same only the representation of the 1st constraint changes. Now, if you look at the so the only change that is made is in the representation of this constraint. And so this constraint is changed to x 1 square minus x 2 square cube less than or equal to 0. And now if you look at the 0.3 by 9 by 4. So, let us see what happens?

(Refer Slide Time: 06:42)



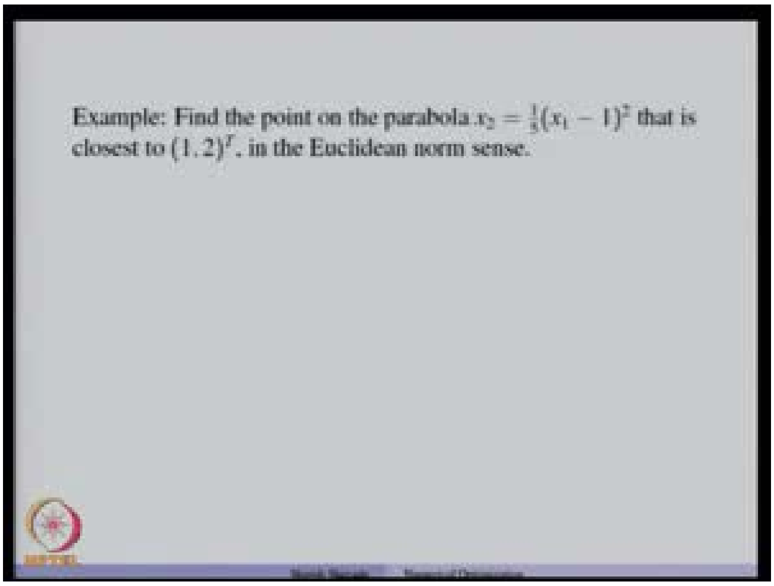
So, let us look at the third case. So, suppose the 1st constraint is x 1 square minus x 2 whole cube less than or equal to 0. Now, let us look at x star which is a local minimum of this problem which is 3 by 2 and 9 by 4 at x star; we will see that gradient x 1 star is equal to 0. And, therefore if you represent this constraint like this; then this point is not a regular point because this is the only constraint which is active at this and it is not a linearly independent set. So, this is not a regular point. So, although the constraint set remains same the representation of the constraint is very important. And, this example is juts to illustrate that fact that if you change the representation a local minimum may not be regular point. So, this is a very important point that 1 needs to understand.

(Refer Slide Time 08:22)



Now, let us look at another example. So, you will see that this point which was a point which was a local minimum strict local minimum is not regular. So, we really cannot apply the KKT conditions. Now, let us look at another example.

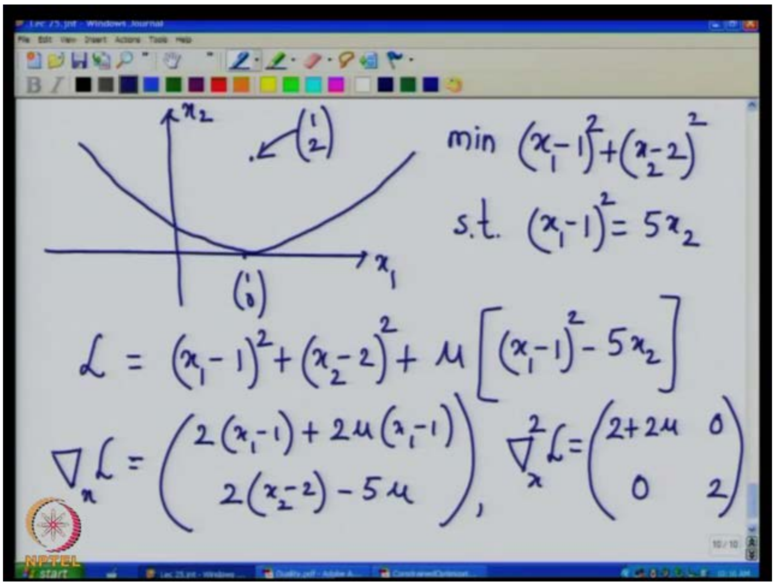
(Refer Slide Time 08:42)



So, here the problem is the following; we want to find a point on the parabola? The equation of the parabola is x 2 equal to 1 by 5 x 1 minus 1 square; and that point should be closest to the point 1 comma 2 in the Euclidean norm sets.

So we are interested in finding a point in the parabola which is closest to this is. Since, this is a problem in two-dimensional space; one can think of the interpretation of this problem as we want to find a circle of minimum radius centered at the point 1 comma 2. So, we want to find a circle centered at 1 comma 2 which has a minimum radius and which touches the parabola. So, let us write down the problem formulation.

(Refer Slide Time: 10:00)

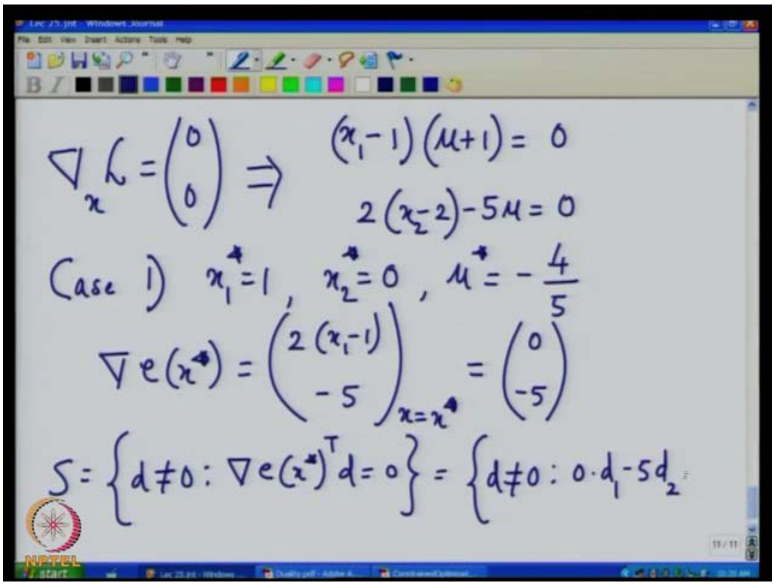


So, we have the x 1 and x 2 as the 2 axis. And this parabola the given point is 1 comma 2. So, what we are interested in finding out the circle of minimum radius which touches the parabola. So, the objective function is to minimize x 1 minus 1 square plus x 2 minus 2 square; and then that point should touch the parabola. So, the equation of the parabola is x 1 minus 1 square is equal to 5 x 2. So, the parabola would look so it will touch the x 1 axis at 1 0; and the parabola would be something like this and the point 1 comma 2 is a point somewhere here.

So, this is the point 1 coma 2. So, we want find out the circle of minimum radius which touches this parabola. So, note that there is only 1 equality constraint. So, let us write down the Lagrangian 1st. And, the Lagrangian is the objective function which is x 1 minus 1 square plus x 2 minus x 2 square plus mu into the constraint e of x equal to 0. So, mu into x 1 minus 1 square minus 5 x 2. So, let us write of the gradient of Lagrangian with respect to x; and that gradient is nothing but 2 into x 1 minus 1 plus 2 mu into x 1 minus 1 and 2 into x 2 minus 2 minus 5 mu.

So this is the gradient of the Lagrangian. And, we can write down the k c l of the Lagrangian with respect to x. Note that this Lagrangian is a function of x as well as mu but just to avoid notational clutter we are not writing the dependence of L on x 1 x 2 and mu but we assume that L is dependent on all this parameters. So, the a c m of this Lagrangian will be 2 plus 2 mu in the derivative of the this term with respect to x 2 is 0; and the derivative of this term with respect to x 2 is 2. So, this is going to a c m of the Lagrangian.

(Refer Slide Time: 13:32)



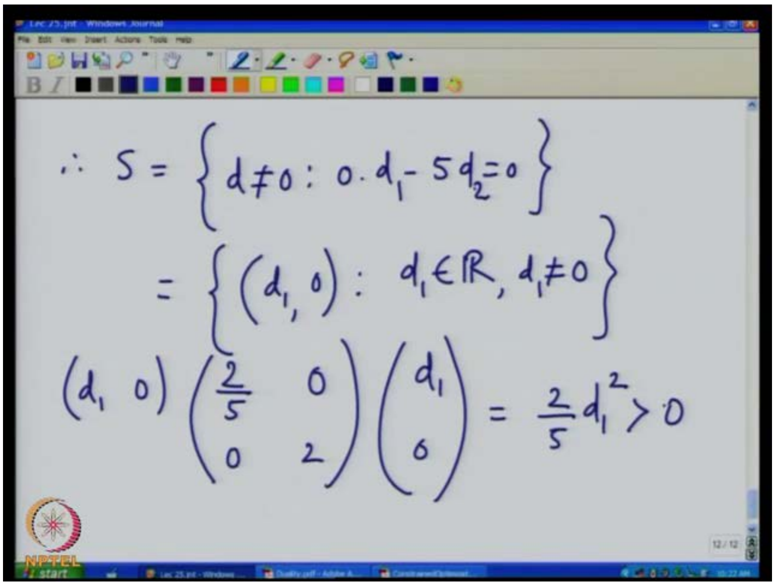
Now, now let us check when does the Lagrangian becomes 0? So, this implies that so if you look so you have to equate this 2 quantities with 0. So, this quantity is nothing but x 1 minus 1 into 2 plus 2 mu that should be equal to 0; and 2 into x 2 minus 2 minus 5 mu should be 0. So, let us write down those quantities. So, x 1 minus 1 into mu plus 1 equal to 0; and 2 x 2 minus 2 minus 5 mu equal to 0. Now, if you look look at the 1st equation, so we will have either x 1 equal to 1 or mu equal to minus 1. And, note that this mu is the Lagrangian multiplier corresponding to the equality constraint.

So, it can be negative. So, let us consider the 1st case; where we say we x 1 star equal to 1 because then this 1st equation is satisfied. Now, once we know x 1 star you have to get x 2 star. And, how do you get x 2 star? So, we look at the constraint. When x 1 star equal to 1 we need to satisfy the equality constraint; and therefore, x 2 star will be 0. So, x 2 star will be equal to 0. And, if you plug in that value here what we get is mu star is minus

4 by 5. So, we get a KKT point where x 1 star equal to 1, x 2 star equal to 0 and mu star equal to minus 4 by 5. Now, now let us look at the second order sufficiency conditions.

So for that purpose we will need gradient e x star. Now, this e x star is basically this constraint; so we can write this constraint as e x star e x equal to 0. So, x 1 minus 1 square minus 5 x 2 equal to 0. So, this is our equality constraint. So, gradient e x star will be 2 into x 1 minus 1 and minus 5 evaluated at x equal to x star. So, let us evaluate it at this point 1 comma 0 and it will be 0 comma minus 5. So, let us define the set s to be the set of all d is not equal to 0 such that gradient e x star transpose d equal to 0. And this set is nothing but set of all d is not equal to 0 such that gradient e x star is 0 comma minus 5. So, 0 into d 1 minus 5 d 2 is equal to 0.

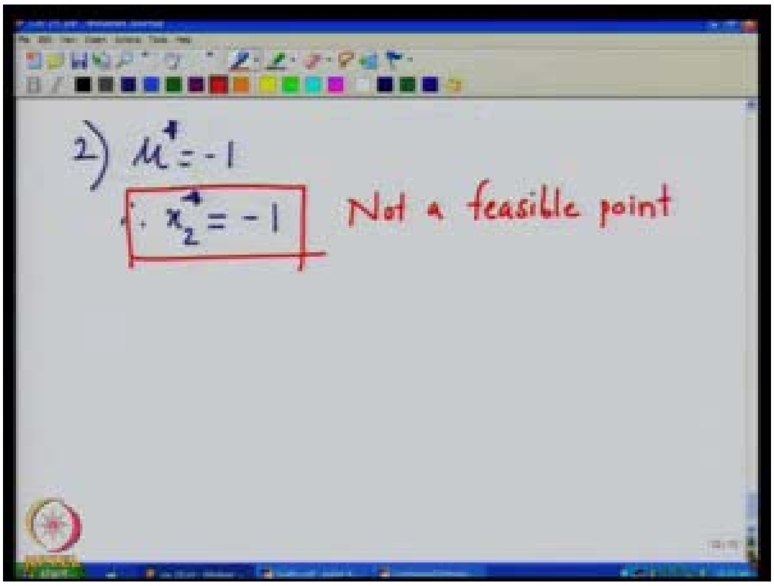
(Refer Slide Time: 17:42)



And, and this therefore s becomes the set of all d is not equal to 0; 0 into d 1 minus 5 d 2 equal to 0 which is what we saw earlier. And, this is nothing but so you will see that in order that this equation is satisfied d 2 has to be has to be 0; while d 1 can take any value. So, d 1 belongs to R and d 2 can should be 0. So, what we have is the set of all points d 1 comma 0 such that d 1 belongs to R and we do not want d to be at 0 direction. So, d 1 not equal to 0. So, this is our set s; so our aim is to get a vector from this space, mean vector from the space. And, see whether with respect to this the matrix is positive definite. So, let us take a vector d 1 0 from the space. Now, the matrix evaluated at mu star equal to minus 4 by 5 is 2 by 5, 0, 0, 2 into d 1 0.

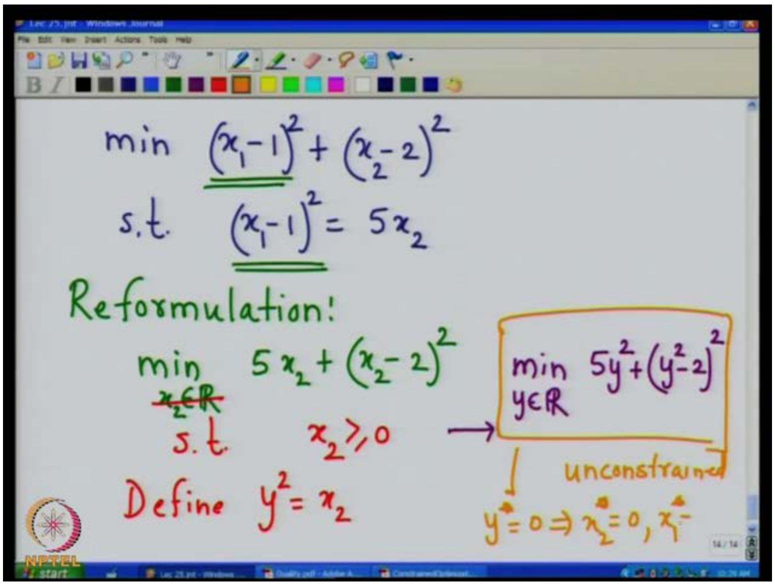
So, this is d transpose (( )) of L into d; where d belongs to the space of vectors which satisfy that gradient e x star transpose d equal to 0. So, this quantity is nothing but 2 by 5 d 1 square. Now, d 1 is a real number nonzero real number; and therefore, 2 by d 1 2 by 5 d 1 square is greater than 0. And therefore the point that we saw here is a KKT point and is a strict local minimum. So, this point is a strict local minimum. Now, let us look at a another case; if recall that if you consider the case where x 1 was set to 1 and this equation was satisfied. Now, let us consider the equation where mu is set to minus 1.

(Refer Slide Time 20:30)



So, we have now case 2 where mu star is equal to minus 1. Now, if mu star equal to minus 1 and if we substitute that here what we get is x 2 star to be also minus 1. So, we get x 2 star equal to minus 1 and if we look at the constraint side. So, x 2 star is never minus 1; it is always nonnegative. So, x 2 star equal to minus 1 this is not a feasible point. And, therefore we cannot continue this case further; and therefore, you will see that this point 1 0 is a strict local minimum. And, if you look at the figure this point this point is a strict local minimum. Now, let us consider this problem again.

(Refer Slide Time: 20:29)



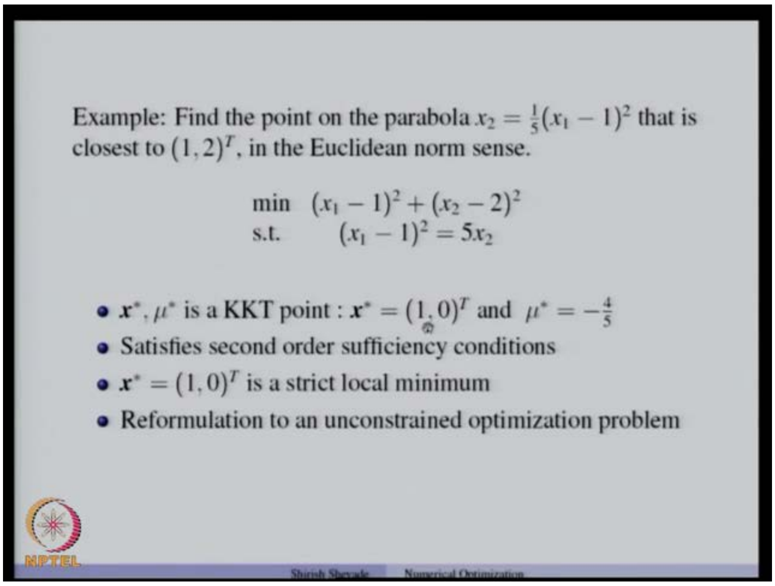
So, we have minimize x 1 minus 1 square plus x 2 minus 2 square subject to x 1 minus 1 square is equal to 5 x 2. Now, if you look at the constraint and the objective function; 1 will be tempted to make the substitution for x 1 minus 1 square because x 1 minus 1 square equals 5 x 2. And, 1 will be tempted use that equality in the objective function. So, let us see the reformulation of this problem.

So, minimize 5 x 2 plus x 2 minus 2 square and this is the problem with respect to x 2. So, this is the common mistake which is made while the reformulation is that we simply substitute x 1 minus 1 square to be 5 x 2; and solve it as a unconstraint problem. But this is not correct because if you look at this constraint; we have not fully utilize this constraint in the objective function. Note that this x 1 minus 1 whole square is a nonnegative quantity; and therefore we cannot have this constraint. But rather we should write the constraint minimize subject to the constraint that x 2 greater than or equal to 0. So, this is going to be the correct reformulation of the original problem. So, the original problem was constraint and we now have another constraint problem which is equivalent problem. Now, interestingly this problem can be converted to an unconstraint problem. So, suppose we define define y square to be x 2.

So, y square is a nonnegative quantity and therefore, it always satisfies y square greater than or equal to 0. And, then we can substitute this y square in this objective function. And, what we get is the following problem where we minimize 5y square plus y square minus 2 whole square; and then we can say that y belongs to R. Now, this problem is a unconstraint problem. So, while reformulating a given unconstraint problem 1 has to be extremely careful. And, this problem we can solve it using our our earlier optimality conditions.

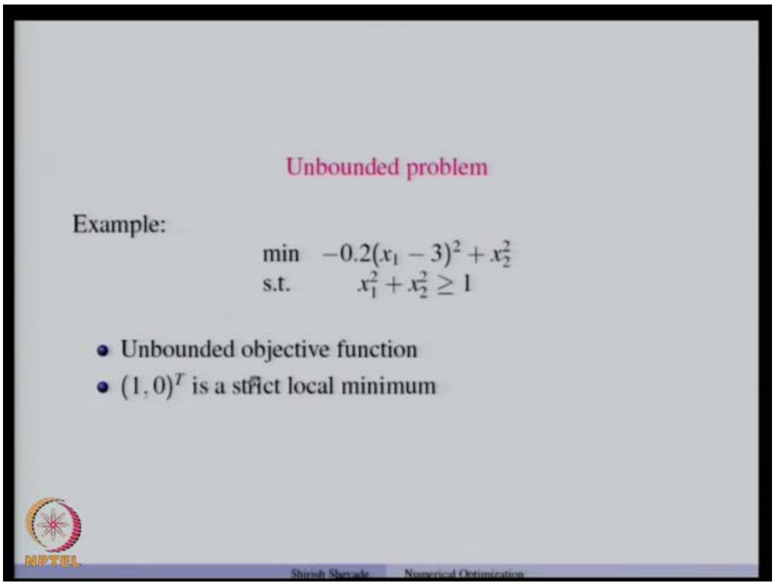
The 1st order and second order optimality conditions for unconstraint optimization problems. So, you take the derivative of this and equate it to 0 and that will give us the solution. You can check that that gives us y star to be 0. And, therefore x star x 2 star equal to 0 and x 2 star equal to 0 which means that x 1 star equals to 1. So, by converting appropriately into unconstraint problem we got the same solution for the given problem 1 comma 0. So, it is very important to understand that reformulation has to be done in a careful way.

(Refer Slide Time: 27:12)



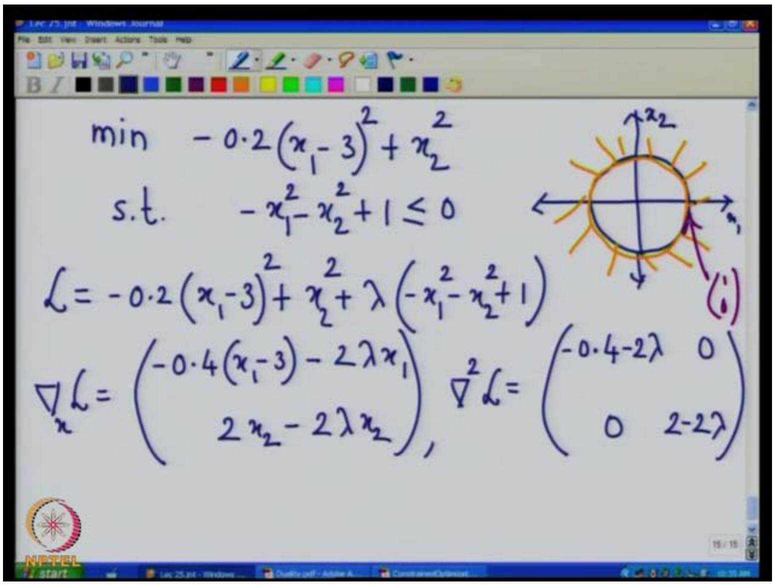
So, we formulated the problem like this. And, we saw that 1 0 is a KKT point with mu star to be minus 4 by 5. And, I also satisfy the second order conditions with respect to that (( )) matrix. And, therefore this 1 0 is a strict local minimum and a reformulation has to be done in a careful way.

(Refer Slide Time: 27:38)



Now, let us look at the unbounded problem which is given here. So, minimize minus point 2 into x 1 minus 3 square plus x 2 square subject to the constraint that x 1 square plus x 2 square greater than or equal to 1. Now, this quantity is a nonnegative quantity, this quantity also nonnegative quantity but this quantity is multiplied by a negative quantity. So, if you make x 1 very large; then that quantity will be multiplied by a negative quantity. And, therefore truly the minimum does not exist for this problem but does this local minimum exists and that is what we want to answer. So, as I said that objective function is unbounded. And, let us find out whether this point 1 0 is a strict local local minimum or not.

(Refer Slide Time: 28:43)

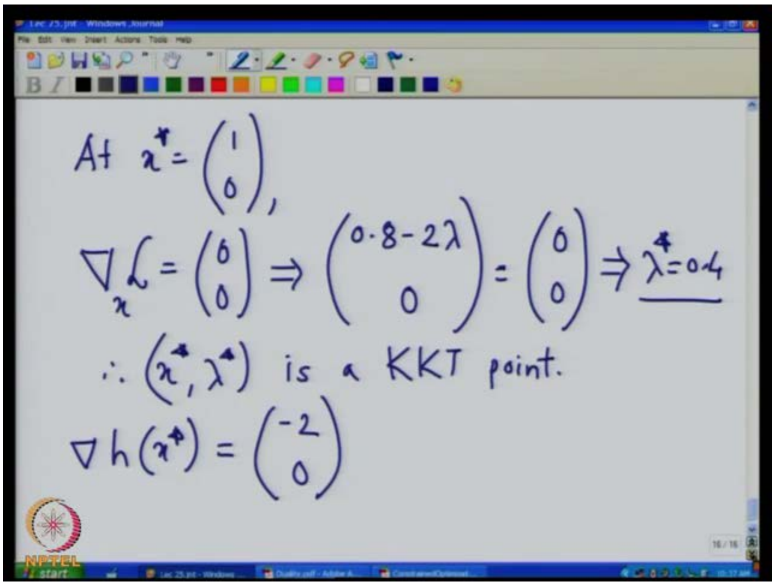


So, the problem we want to solve is minimize minus 0.2 into x 1 minus 3 square plus x 2; square subject to minus x 1 square minus x 2 square plus 1 less than or equal to 0. So, the constraint was x 1 square plus x 2 square greater than or equal to 1. we just wrote it in form h x less than or equal to 0. Now, if we draw the constraint set. So, we have the 2 axis x 1 and x 2 and suppose that this is the circle whose equation is x 1 plus x 2 square equal to 1. Then, what we are interested in a region where x 1 plus x 2 square is greater than or equal to 1. So, that means we are interested in the region on the circle or outside the circle; we are interested in this feasible set.

And, clearly the objective function is unbounded. And, therefore the minimum does not exist but let us see whether there is a local minimum at this point at 1 comma 0. Now, let us write down the Lagrangian as our usual procedure. So, Lagrangian is minus 0.2 in to x 1 minus 3 square plus x 2 square plus lambda into minus x 1 square minus x 2 square plus 1. So, lambda is the Lagrangian multiplier corresponding to the inequality constraint; and clearly this lambda has to be nonnegative.

Now, if you take the gradient of l with respect to x and that will be minus 0.4 into x 1 minus 3 minus 2 lambda x 1 ad 2 x 2 and minus 2 lambda x 2 this is the gradient vector. And, the matrix (( )) matrix will be minus 0.4 minus 2 lambda; then the derivative of this with respect to x will be 0. And, that derivative of this with respect to x 2 is 2 minus 2 lambda. So this is going to the (( )) matrix. And, we are in particular interested in finding out whether the point 1 0; this point is a strict local minimum or not.

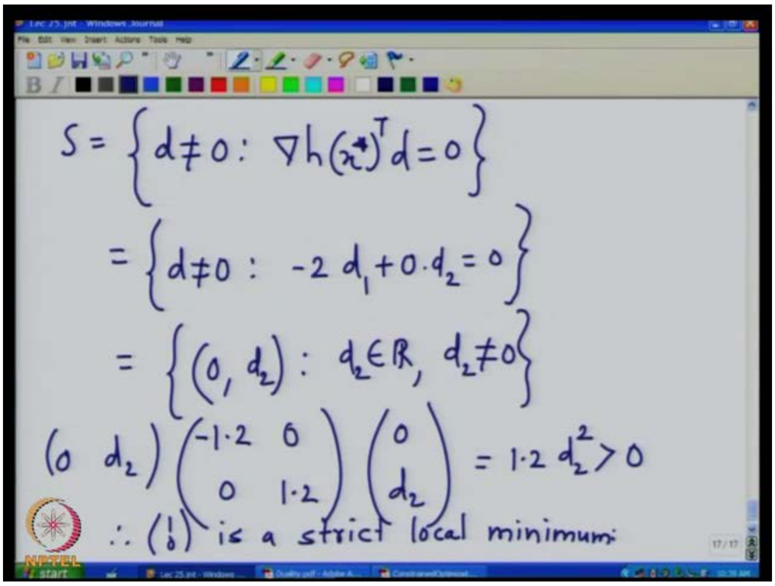
(Refer Slide Time: 32:20)



So, at x star to be 1 0; if we equate the Lagrangian the gradient of the Lagrangian to 0; which implies so what we have is we substitute x 1 equal to 1 and x 2 equal to 0 in this the gradient of the Lagrangian. And, what we get is 0.8 minus 2 lambda and 0 equal to 0 and this implies lambda star to be 0.4.

So, this quantity is greater than 0. So, the point x star lambda star is the KKT point. Therefore, x star, lambda star is a KKT point. Now, we have to see whether it satisfies the second order condition and for that purpose what we need is the derivative of the gradient of the constraint. So, this is our constraint and we write it as h x equal to 0. So, the gradient h x is minus 2 x 1 minus 2 x 2 and that will be evaluated at 1 1 0. So, gradient h x star gradient h x star will be minus 2 x 1 minus 2 x 2 and x 1 is 1 and x 2 is 0, so it will be minus 2 0. Now, let us write down the set s.

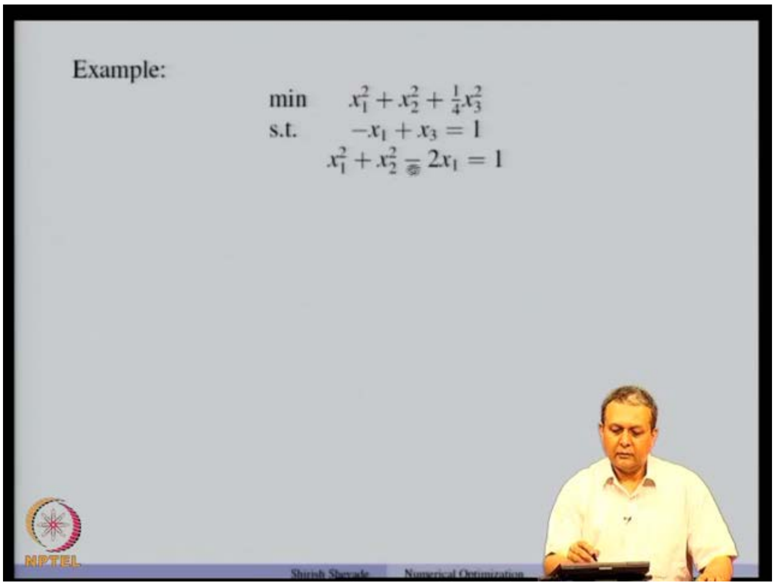
(Refer Slide Time: 34:28)



S, is the only constraint that is active is x 1 square plus x 2 square equal to 1. So, let us write down set d with a set of all nonzero directions such that gradient h x star transpose d equal to 0. And, that is nothing but the set of all this nonzero such that minus 2 d 1 plus 0 into d 2 equal to 0. And, this is nothing but a set of so this implies that d 1 has to be 0, d 2 can be any value.

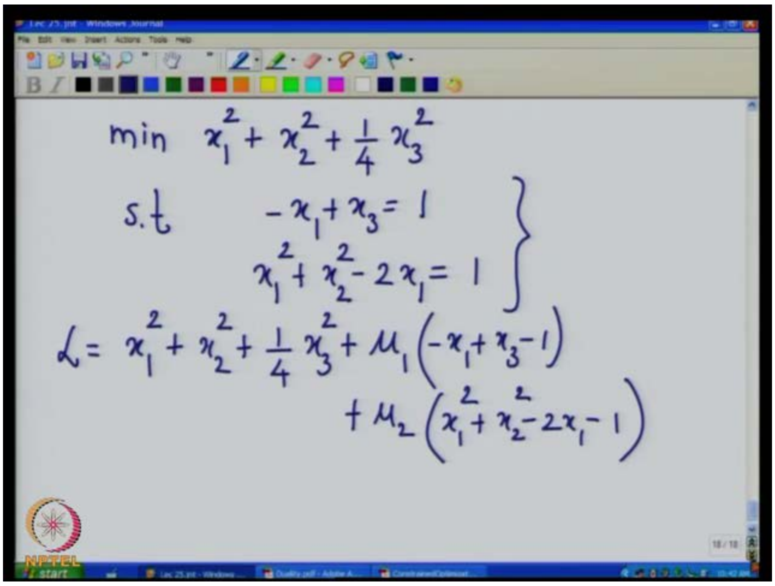
So, set of all 0 comma d 2 such that d 2 belong to R and d 2 not equal to 0. So, let us take vector from the space. So, let us take a vector 0 d 2 itself and then see whether the (( )) positive definite or not. So, the (( )) evaluated at x star lambda star will be this multiplied by 0 d 2. So, d transpose (( )) of Lagrangian into d; where d comes from the space s space (( )) s. Now, note that this is a 1 dimensional space; so it is enough to take only 1 vector here. And, this is nothing but time 1.2 d 2 square and since d 2 is not 0 this quantity is greater than 0. And, therefore 1 0 is a strict local minimum. So, although the function is unbounded; this point is a strict local minimum of the given problem.

(Refer Slide Time: 36:48)



Now, let us take this example and solve it with respect to the equality constraints which are shown here.

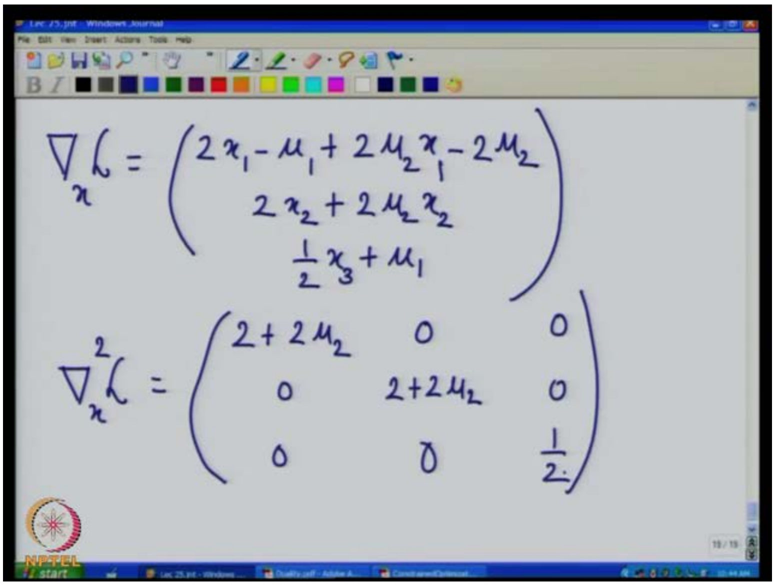
(Refer Slide Time: 37:08)



The problem is to minimize x 1 square subject to minus x 1 plus x 3 equal to 1 and x 1 square. So, note that we have 2 equality constraints here and both have to be active at a optimal point. So, let us write down the Lagrangian to be the function. So, for the first constraint let us assume that the Lagrangian multiplier is mu 1 and for the second

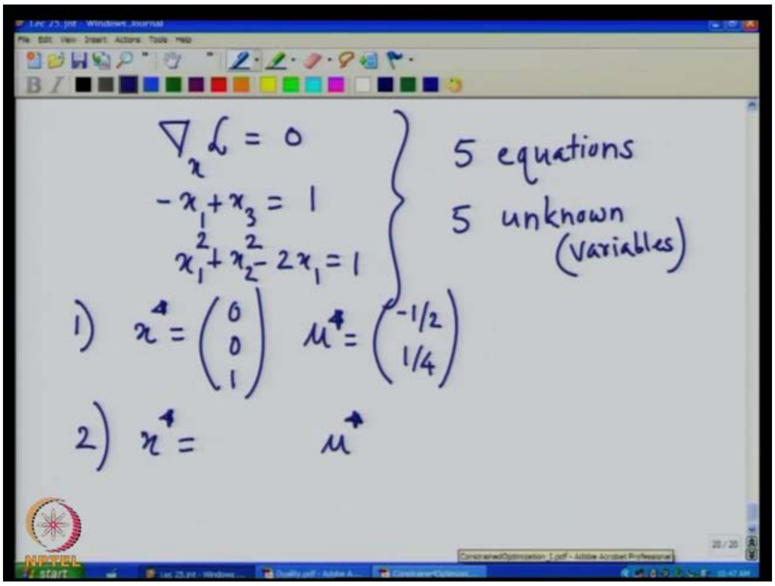
Lagrangian multiplier is mu 2. So, we have minus x 1 plus x 3 minus 1 plus mu 2 into x 1 square plus x 2 square minus 2 x 1 minus one.

(Refer Slide Time: 38:37)



Now, let us write down the Lagrangian the gradient of the Lagrangian to be so you will see that the gradient of the Lagrangian is 2 x 1 minus mu 1 plus 2 mu 2 x 2 2 mu 2 x 1 minus 2 mu 2. Then, the second component is with respect to x 2 and that is 2 x 2 plus 2 mu 2 into x 2. And, the third component is with respect to x 3 and that is half x 3 plus mu 1. Now, the (( )) Lagrangian will be a 3 by 3 matrix and that will be so we have 2 plus 2 mu 2; and there are no terms in this which involve x 2 in x 3. So, the other components are 0 and since it is a symmetric matrix this components are also 0. Then, let us look at the second term that is the derivative with respect to x 2 is 2 plus mu 2. So, we have 2 plus mu 2 and there is no term involving x 3. So, in the second component and the third component we have half. So, this is going to be the (( )) matrix of the Lagrangian.

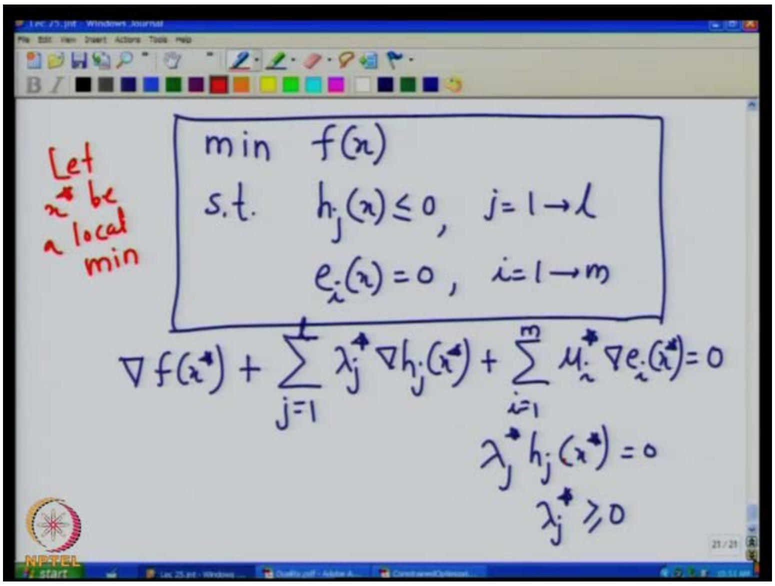
(Refer Slide Time: 40:44)



So, what we need to do is that we need to find out x such that the gradient of x vanishes and the 2 constraints are satisfied because those are equality constraints. So, they need to be satisfied. So, the other equality constraint is x 1 square plus x 2 square minus 2 x 1 equal to 1. So, in all we have 5 variable x 1, x 2, x 3 and then mu 1 and mu 2. And, we have 5 equations. So, we have 5 equations and 5 unknowns.

So, they are 5 equations and 5 variables. And, so you can check that 1 solution of this is x star to be 0 0 1 and mu star be minus half and 1 fourth. So, both the constraints are active at this point and 1 can write down the KKT conditions for this. And, the other point is so we can find out the other x star and the other mu star? And, write down the conditions and check the KKT points for this. So, for we have seen the constraint optimization problems. So, in particular we were looking at the problems of the type.

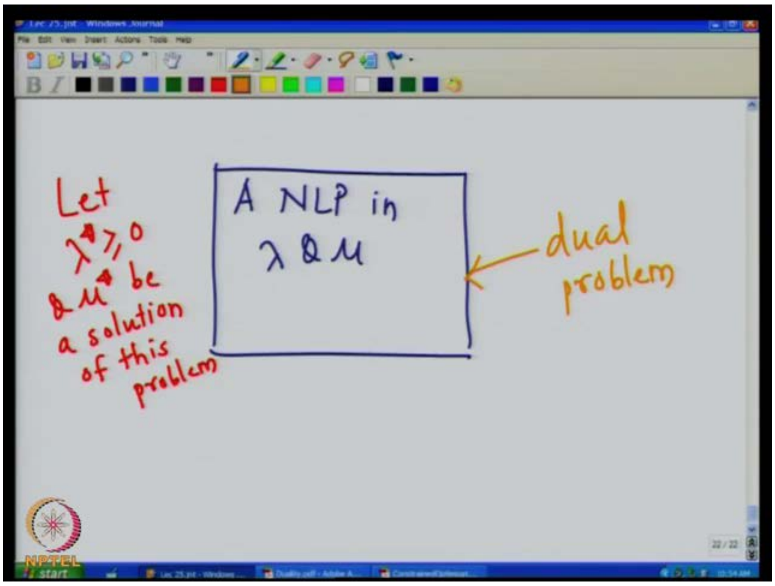
(Refer Slide Time: 43:16)



Minimize of x subject to x less than or equal to 0. If you are going from L and e i (x) equal to 0; i going from 1 to m. Now, we also wrote down the KKT conditions for this problem. In particular we saw that the gradient of the Lagrangian vanishes, so gradient f(x) star. So, there exists x star lambda star mu star such that gradient f(x) star plus sigma j lambda j star gradient x star plus. So, j going from 1 to L plus sigma i going from 1 to m, mu i star gradient e i (x) star to be 0. And, we have the complimentary slackness condition which is lambda j star x j x star to be 0 and lambda j star nonnegative. So, the Lagrangian multiplier is corresponding to the inequalities constraints are nonnegative.

The Lagrangian multipliers corresponding to the equality constraints are unrestricted in sign. So, we are interested in finding out x star lambda star and mu star we satisfy this. And, then then if the second order conditions are satisfied at this x star lambda star mu star then we say that x star is a local minimum. So, let us assume that x star be local minimum. So, our aim is to solve this problem and get x star which satisfies this conditions as well as the second order conditions. Now, to satisfy this conditions what we need is that we need x star lambda star and mu star. Now, suppose if we have some problem for which lambda star and mu star is a solution or in other words suppose we have some other problem.

(Refer Slide Time: 46:33)



As of now we do not know the nature of this problem. So, a non-linear programming problem NLP in lambda and mu. And, let us assume that this optimization problem has solution lambda star and mu star. So, let lambda star which is greater than or equal to 0 and mu star be a solution of this problem.

Now, suppose that this problem is easy to solve; then what 1 can do is that 1 can get lambda star and mu star from this. And, suppose it turns out that lambda star and mu star can be used here; and then we can get easily get solution of this problem. So, given the problem the idea is that if we can find out another non-linear programming problem involves variables lambda and mu; where lambda is are nonnegative mu is are unrestricted in sign. And, they have the same dimension as the number of constraints in the original problem like lambda will have lambda will be l dimensional and mu m dimensional but lambda is are nonnegative.

So if we have such a problem and if we can get a solution of this problem very easily. And, if that solution can be used to solve this sets of equations to find x star. Then, we directly get the solution of this problem. So, this problem which we are going to study this problem is called a dual problem. And, the problem which we had originally written is called the primal problem. Now, note that although for the primal problem there exists a dual program dual problem; but solution of the dual problem are not does not always give us the solution of the primal problem.

Only in the certain conditions the solutions of this dual problem which are lambda star mu star; they can be used to get the solution of the primal problem. So, this called the duality theory. And, in the next lecture we are going to see that under what conditions do the primal and the dual problems have the same solution? And, the advantage of the dual problem is that many a times this problems are easier to solve. And, therefore and they also have some nice structure and 1 can utilize the nice structure of dual problems; and finally, get back the solution of the primal problem. So, we are going to see those duality related ideas in the next class.

Thank you.