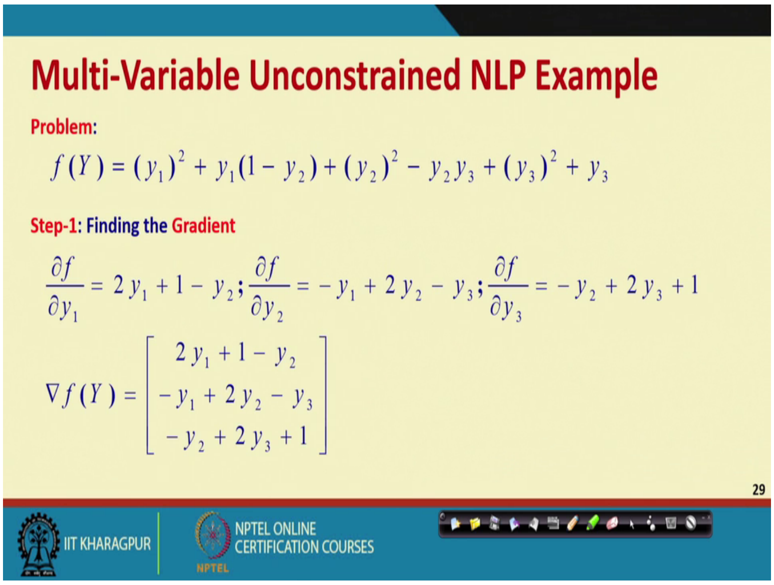
**Selected Topics in Decision Modeling**
**Prof. Biswajit Mahanty**
**Department of Industrial and Systems Engineering**
**Indian Institute of Technology, Kharagpur**

**Lecture - 24**
**Solving Unconstrained NLP**

So, in our course Selected Topics in Decision Modeling, we are now in our 24th lecture that is Solving Unconstrained non-linear programming problems. So, that is our lecture that is solving unconstrained non-linear programming problem.

(Refer Slide Time: 00:36)



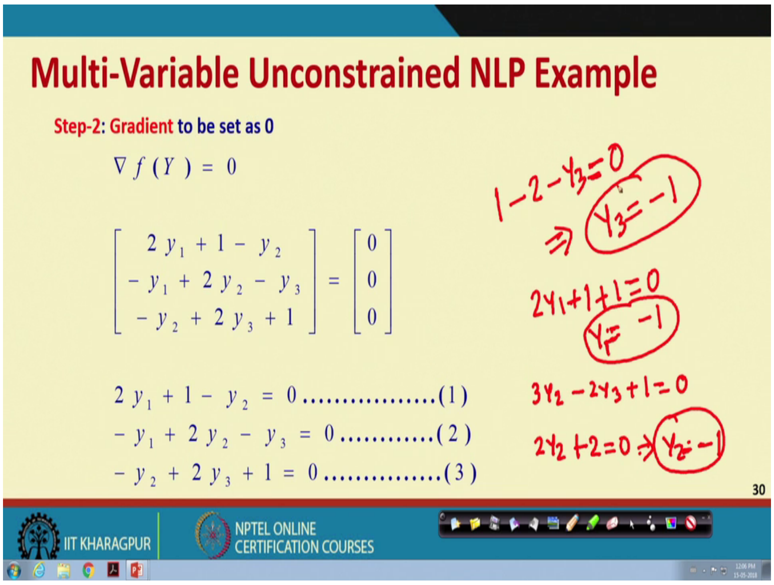

Now, you know in our previous lecture if you recall, we have seen that whether so, in the multivariable unconstrained NLP problems, we need to find out the gradient. We put the gradient equal to 0 and from those gradient, you know we also find out the Hessian matrix, then we check whether the Hessian matrix is a positive definite or negative definite. And based on that, we try to find out the maxima or minima of this particular function right so, based on the nature of the Hessian matrix. How do we find out the extrema maxima and minima, by putting the gradient equal to 0.

So, let us take some examples and see what kind of difficulties can come while following this procedure. So, here is a function before us that is f Y equal to y 1 square plus y 1 into 1 minus y 2, plus y 2 whole square minus y 2 y 3 plus y 3 square plus y 3. So, if we if we find the gradient that is del del y 1 for the function, then the what are the terms which will be useful here. So, as you can see that is the first term that will be useful, the second term will be useful and that is all. So, those are the terms that will be useful. So, in the first term what you get is 2 y 1 and in the second term 1 minus y 2. So, that is the del del y 1 of a del f del y 1.

So, the second term again you know what are the terms that will be useful that is you know because del y 2. So, it is the second term, which will be which will be required the second term, then the third term, the fourth term that is all. So, you know the second term will give us minus y 1, the next term 2 y 2 and here minus y 3. That is very simple partial differentiation and for the del del y 3 again you know we see the only the last few terms that is the last 3 terms will be useful. So, you know again you see that this term this term and this term they will be useful. So, it becomes minus y 2 plus 2 y 3 plus y.

So, when you when you put them all we get the gradient matrix right del f Y, 2 y 1 plus y 1 minus y 2, minus y 1 plus 2 3, y 2 minus y 3 minus y 2 plus 2 y 3 plus 1. So, that is the gradient right. So, after the gradient what we do? We must have the next that is the gradient should be set to 0.

(Refer Slide Time: 04:04)




So, that is what we have done here the gradient is set to 0, that is the first term basically it means that set the 3 terms equal to 0 separately right. So, if you if you set this terms you know separately equal to 0, then you know how do you find out the values of y 1, y

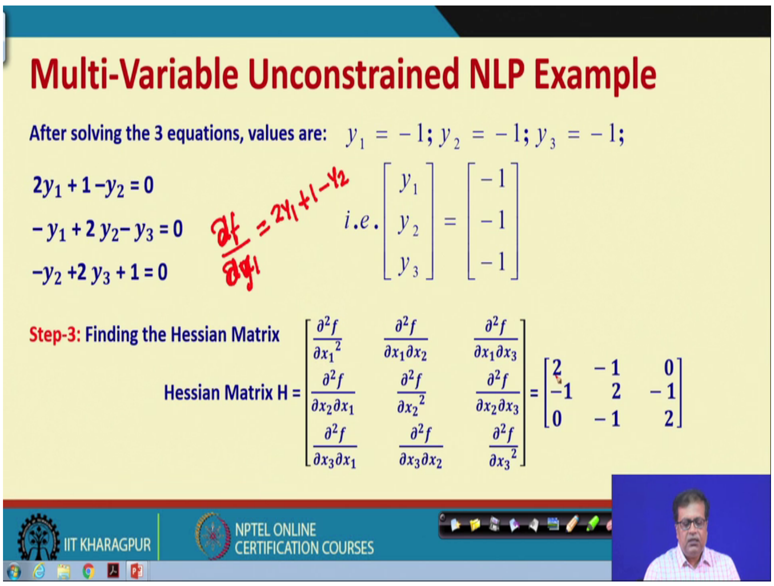
2 and y 3. So, it is a very simple problem really all we have to do is solve set of simultaneous equations, not much challenge here.

So, quick quickly if you can see if you if you add the double of the second term with the first term, then you know this y 1 gets cancelled then we get you know 4 that is 3 y 2, 3 y 2 minus y 3 3 y 2 see that is what we are doing we are adding the double of the second term. So, that is 4 y 2 minus y 2 plus 3 y 2 minus 2 y 3 plus 1 equal to 0. So, what we have do we done? We have doubled this and added with the first equation. So, it becomes 4 y 2 minus y 2 3 y 2 minus 2 y 3 plus 1.

Now, this n these third equation you can also add. So, if you add then what do we get? We get 2 y 2 minus. So, this one is plus. So, this will be plus also 1 and 1 2. So, this will be 0. So, what you is the value of y 2? It becomes minus 1 right. So, very quickly we got y 2 equal to minus 1. So, you know using them. So, minus of minus 1 is again plus 1 so, 2 y 1 plus 1. So, from the first equation we get 2 y 1 plus 1 again plus 1 equal to 0. So, y 1 is also minus 1 right.

So, we get y 1 is minus 1 y 2 also minus 1. So, put them in the second equation. So, minus of minus 1 is plus 1 and plus 1 minus 2. So, 1 minus 2, minus y 3 equal to 0, that is what we can write from the second equation. So, that gives y 3 also equal to minus 1 right very simple just this set of simultaneous equations you can solve and we get y 1, y 2, y 3 all equal to minus 1.

(Refer Slide Time: 07:22)

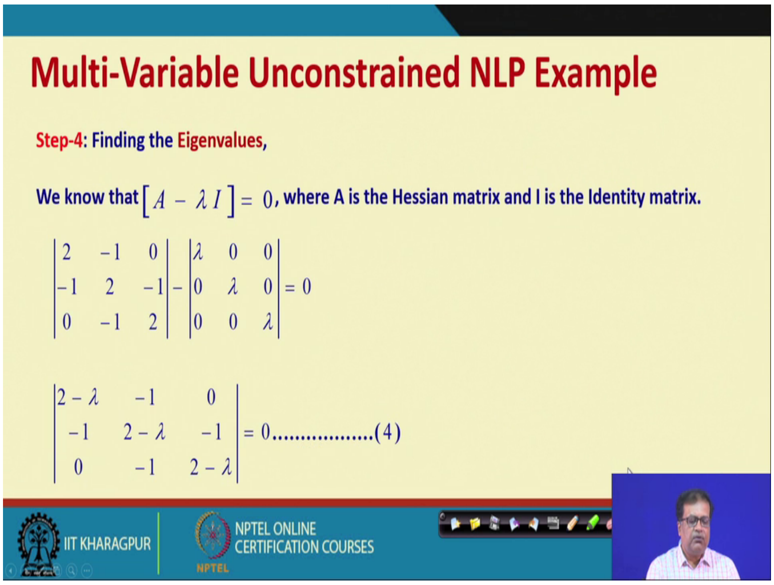


So, that is exactly what we have seen here those 3 simultaneous equations as solved and we get y 1, y 2, y 3 that is equal to minus 1 minus 1 and minus1.

Now, additionally you know we have to find out also the Hessian matrix right. So, you can you can do that as an exercise I am not doing it here; supposing we find the Hessian matrix of those you know things. So, basically you know this is our gradient so, that gradient if you really do you know say actually these are the gradient.

So, please look that del f, del f del not d del f del x 1 equal to 2 y 1, plus 1 minus y 2. So, if you if you take one more time sorry not x 1, y 1. So, if you take one more time del del y 1 this is written x 1, but they really should be all x y 1. So, if you if you take them then it should be the value should be 2 right. So, that will be our Hessian matrix. So, that should be 2 right. So, that is what it is. So, actually this kindly correct this. So, in so, that is what we get that that is going to be our Hessian matrix.

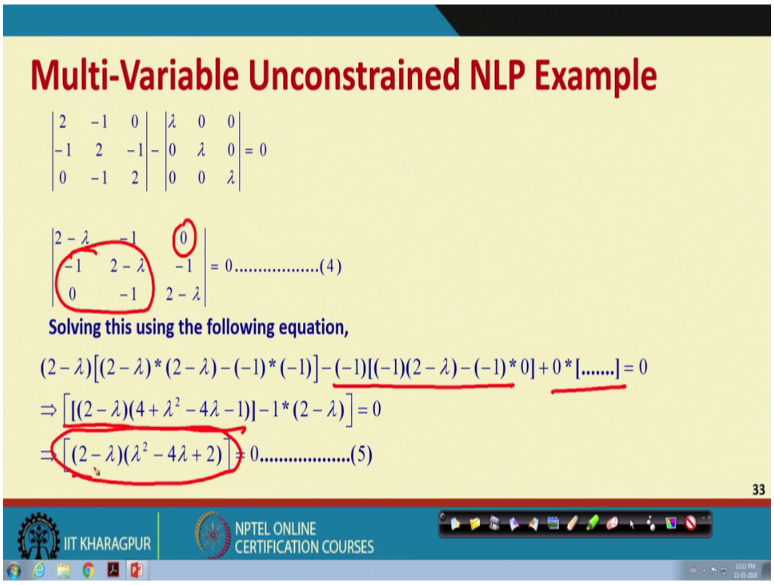
(Refer Slide Time: 08:54)




Now, once we have found out the Hessian matrix let us try to find out the eigen values. So, eigen values again as this is our Hessian matrix. So, what we have to do is, we have to you know A minus lambda I equal to 0, that is our characteristic matrix and this characteristic matrix has to be a characteristic equation.

That characteristic equation if you really put then we get the Hessian matrix minus lambda 0 0, 0 lambda 0 0, 0 lambda equal to 0. So, we get this matrix 2 minus lambda minus 1 0, minus 1 2 minus lambda minus 1 0 minus 1 2 minus lambda equal to 0 this is alright. So, that should be equal to 0 to find out the eigen values.

(Refer Slide Time: 09:55)




Now, what it is? Now, you know this is this one. So, how the multiplication happens? It should be then 2 minus lambda that is this term into 2 minus lambda into 2 minus lambda minus minus 1 into minus 1. So, basically that is how we proceed. So, we take this one and this one.

So, 2 minus lambda so, this is this one and 2 minus lambda into 2 minus lambda minus minus 1 into minus 1 that is the first term. The second term what will be our second term? The second term will be starting with this and it should take these terms, but please remember there is a co factor that should be a negative value. So, negative should be thing also. So, that is what has been done in the case of the second term.

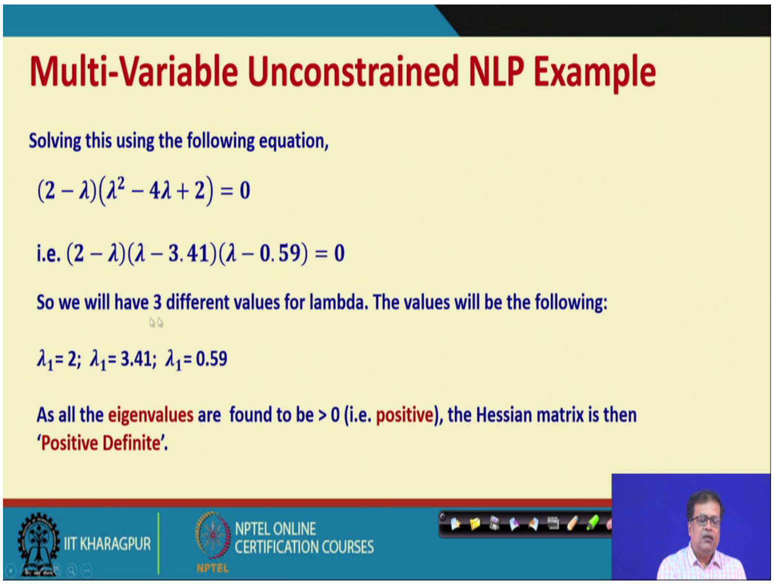
So, all these are first term, this is first term. Now, in the second term you see what we have done, we have taken the minus of minus of that is very important minus 1 into that negative that that is that usually come for the second term. Minus 1 into minus 1 into 2 minus lambda minus minus 1 into 0 right and in third term what will be the third term? The in case of third term we should take this one and this one. So, plus 0 into dot dot dot because minus 1 into minus 1 minus 0 into 2 minus lambda, but there is no need to write it, because anything multiplied by 0 will be 0 only alright.

So, when you combine them all, then we get you know 2 minus lambda can be taken as command and all of these terms. So, you know 2 minus lambda into 2 minus lambda. So, it will be 4 plus lambda square minus 4 lambda minus 1, which is coming from here. So, that is minus 1. So, all of these are first term, then in the second term also you know this is our second term if you if you see that second term. In the second term if you take out 2 minus lambda then this term is 0, then we really have a minus 1 here. So, all that minus 1 will come inside.

So, this is 3 minus 1. So, it will become 2 and 2 and 2 minus lambda can be taken component right. So, finally, we get this kind of equation, 2 minus lambda can be taken component right. So, finally, we get these kind of equation 2 minus lambda, lambda square minus 4 lambda plus 2 equal to 0.

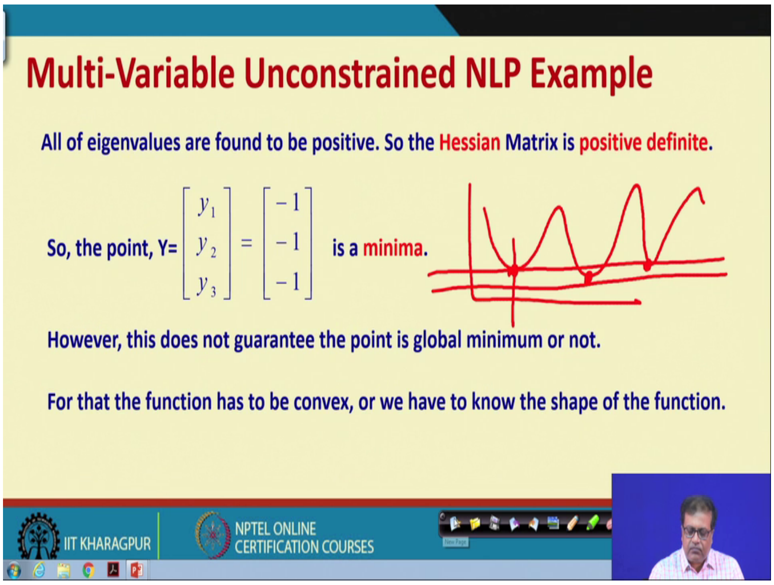
Now, this problem was simple so; obviously, we could easily get into this kind of forms, but in all problems it may not be so, easy, but then I know we have to find out the lambdas by some way or the other, and that is a major challenge for such kind of problems. So, we should remember that right.

(Refer Slide Time: 13:09)




Now, once we find out, now you know this can be factorized the idea is that this into this is2. So, 4 should be divided in such a way that the multiplication is 2. So, that is 3.41 and 0.59. So, we have 2 minus lambda, lambda minus 3.41, lambda minus 0.59 equal to 0. So, we have 3 different values of lambda and those values are 2, 3.41 and 0.59 right. So, we have now the 3 lambda values. So, what is your conclusion about the eigen values? The eigen values are you know all found to be positive. So, since the all the eigen values are positive, the Hessian matrix be then positive definite right we have a positive definite Hessian matrix and the conclusion therefore is that the point y 1, y 2, y 3 minus 1 minus 1 minus 1 is a minima right.

(Refer Slide Time: 14:12)

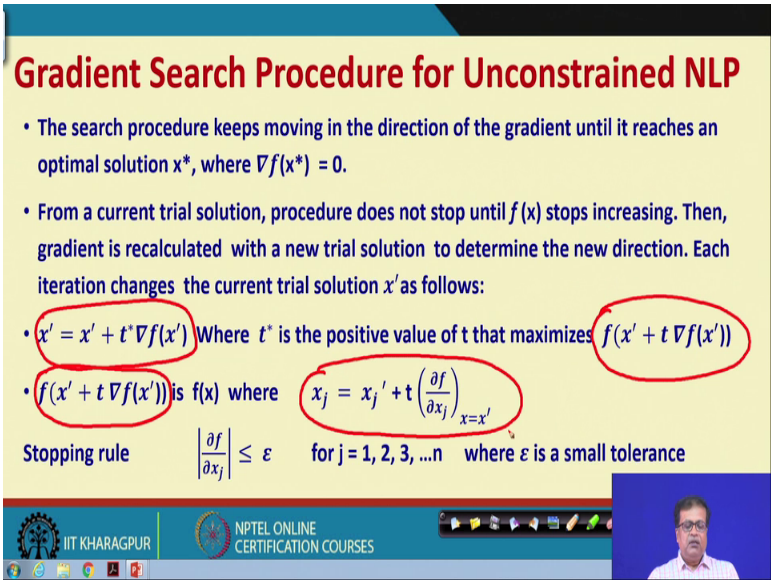



So, minima has been found, but is it a global minima? We may not know because the function has to be convex or we have to know the shape of the function right. So, that also is additional task that is required only thing we found a minimum, but it can be local minima as well right.

So, you see sometimes if the if the nature of the plot is of this type, suppose we found this point; obviously, it is a minima, but look here there are other minimas as well and you know if you look at this, then this point is you know having a lower value. So, these minima is could be a local minima and may not be a global minima right. So, that has to be remembered fine.

So, that is how we found out by with the help of the gradient and the Hessian matrix, and the eigen values of the Hessian matrix we are able to solve the unconstrained problem.

(Refer Slide Time: 15:38)




But then there could be the difficulties what are they let us see. See what really happens that the gradient we have seen, it was not easy to really put those gradient values and put two 0's and then finding Hessian matrix and all those things you know. So, there should be a procedure, where it even if you know we are not able to put del f x equal to 0, you know you know there should be a method by which we can approach and we can do with the help of gradient search itself. So, that is the technique that is called the gradient search procedure for unconstrained NLP.

The search procedure keeps moving in the direction of the gradient until it reaches an optimal solution x star, where the gradient equal to 0 right. So, from a current trial solution, procedure does not stop until fx stops increasing right. So, supposing this procedure is shown for a maximization problem.

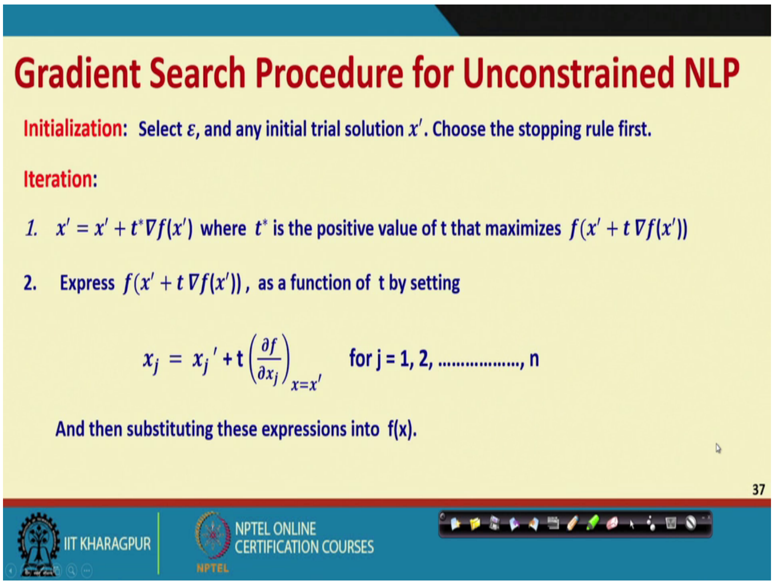
So, supposing we have a maximization problem and we find a gradient and then we know that if we go on the direction of the gradient you know and at a at a given point, that we have got a gradient value and you know we see the fx is increasing then we recalculate the gradient for a new trial solution, for there is a value called x dash.

Where x dash equal to x dash plus t times the gradient; where t star is the positive value of t that maximize the function at that point right. So, it is 3 times the gradient if you add with the current you know trial solution value, the functional value at that point we maximize is it alright.

So, that is how we do and these you know these f x dash, x dash is the current trial plus the t times the gradient is nothing, but the fx where see please remember this is a multi valued. So, x is not one variable, it is a number of variables. So, where individual values xjs or xj equal to xj dash plus t times del f del x j at x equal to x 1 and where do you stop when the gradient is lower than for each individual gradient of individual x i js are less than equal to epsilon where e is a small tolerance.

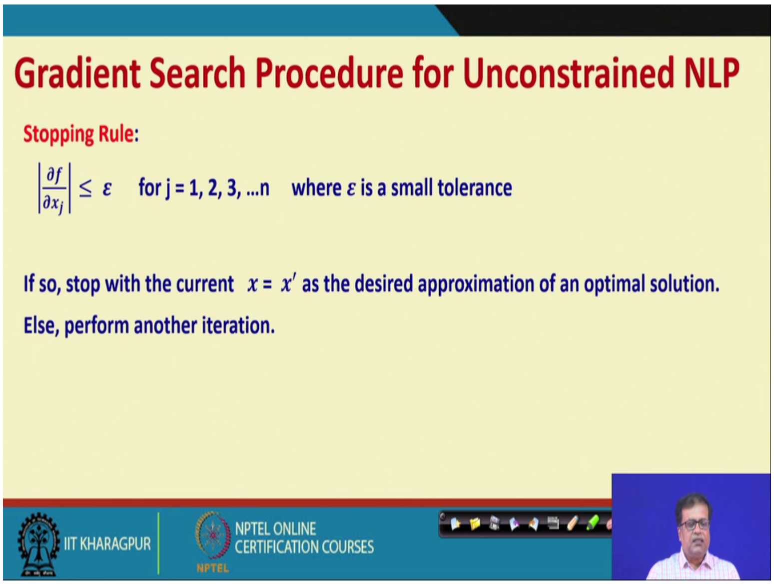
So, it is a standard gradient search procedure, basically we find a current trial solution and we find out the gradient and we add the gradient with regard to the t and see the functional value at that point, and that we maximize to find t and then you know change the xj's in that direction and make it our new trial solution. So, that is the procedure we it will be very clear once you see a problem, then we can we revisit these once again if we have some doubt.

(Refer Slide Time: 19:15)




So, initialization select an epsilon small value, a trail solution x dash and choose the stopping rule. Iteration x dash equal to x dash plus t star the gradient at x dash, where t star is the positive value of t that maximize the functional value at that x. Now, express these you know this functional value as a function of t by setting, xj equal to xj dash plus t del f del x j at x equal to x dash for each x and then substituting these expressions into fx. So, that is how the procedure goes.

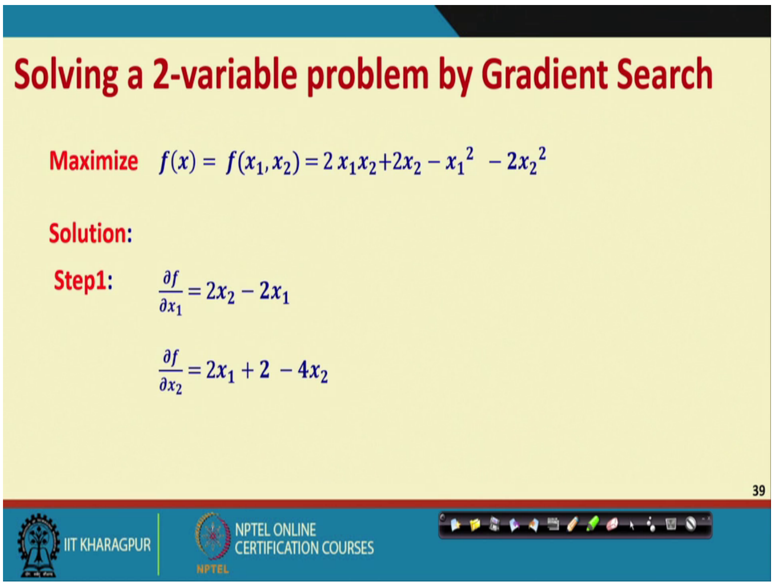
(Refer Slide Time: 20:06)




So, how it is let us see and; obviously, we have the stopping rule that the gradient that the that the partial derivative, for each variable should be less than equal to epsilon. And if so, stop with the current x equal to x dash as the desired approximation of an optimal solution else perform another iteration. So, that is the method.

So, once you know the method, as I said I will revisit the method may be one more time after we see how a problem is solved.

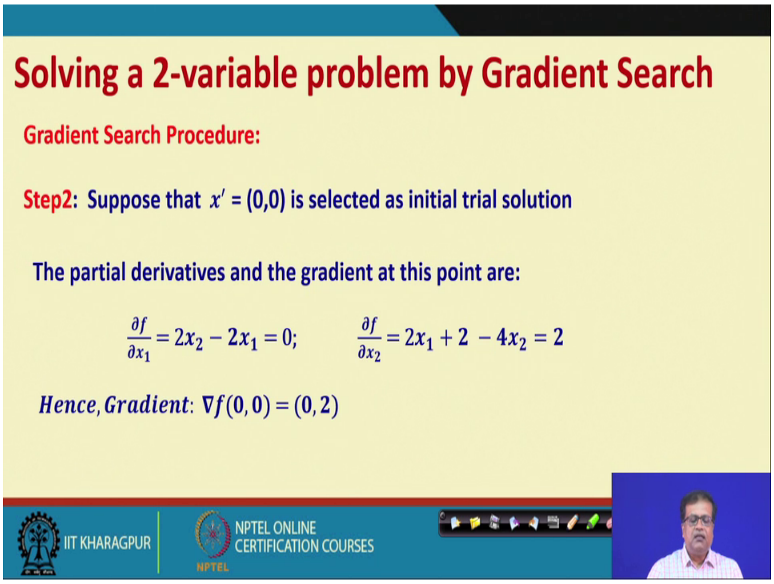
(Refer Slide Time: 20:37)



So, supposing we need to maximize a function, which is f x which is where x is basically 2 variable. So, x 1 and x 2 and the function is 2 x 1 x 2 plus 2 x 2, minus x 1 square minus 2 x 2 whole square right. So, if you really want a you know to find this problem and details further you can see the Hillier Lieberman book you know the operations research where the problem is solved as well. So, I have taken it from the Hillier and Lieberman book.

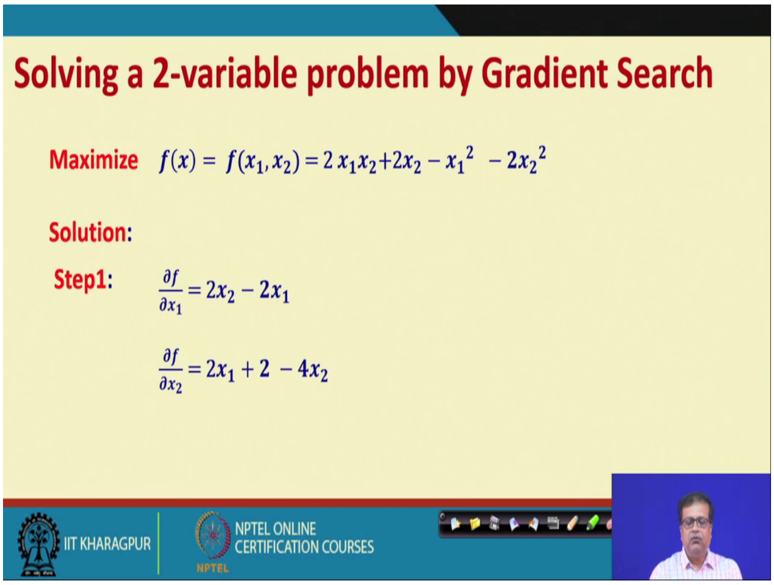
So, what we do? We first find out the gradient gradients or the del del x 1 and del del x 2 of the function. So, what will be the del del x 1 of the function? You know very clear it should be 2 x 2 minus 2 x 1 and what is del del x 2 of the function? 2 x 1 plus 2 that will come from here and from here minus 4 x 2. So, that is the first step find the gradient right. So, these 2 are obtained. So, this part is very clear now, go to the next step.

(Refer Slide Time: 21:56)



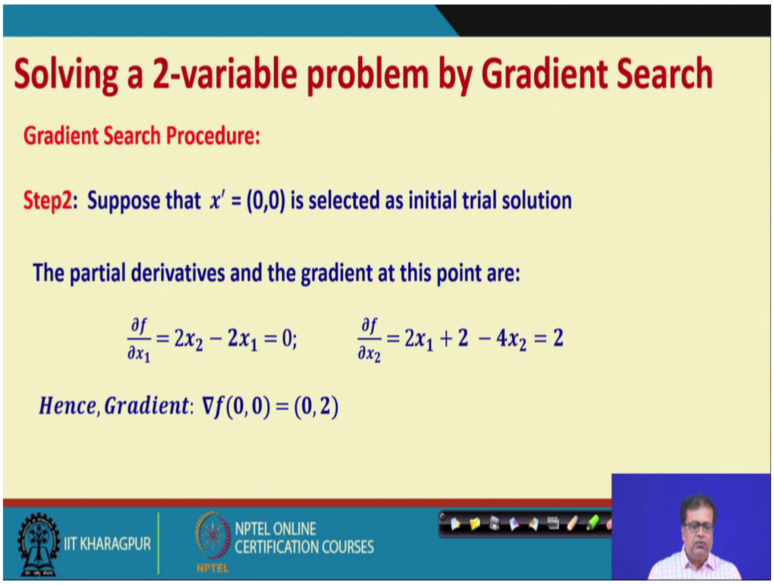

So, after that what we do, supposing we arbitrary start at a point 0 0, we start at a origin.

(Refer Slide Time: 22:07)




Because see the if you look at the function, the value of the function is 0 at the origin right because x 1 0, x 2 0 the functional value is 0 and if we are going to get a positive value, we assume that we have a positive value which is higher than the 0 value right.

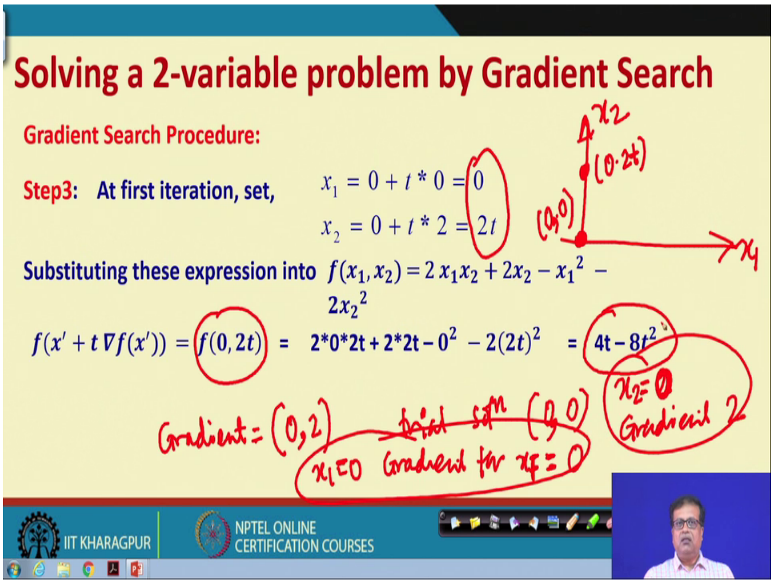
(Refer Slide Time: 22:31)




So, that is how we start at the origin.

Now, already found that the partial derivatives and their gradient at this point can be obtained as at this point 0 0, if you put that is that these are our del f del x 1 and del f del x 2. So, if we put x 1 equal to 0 and x 2 equal to 0, then the first value becomes 0 and the second value becomes 2. Because x 1 0 and x 2 0 that remains 2 these becomes 2. So, the gradient at this point is 0 2 right. So, we have found the gradient at that given point 0 0 is alright. So, we found the gradient right now what we do with this gradient.

(Refer Slide Time: 23:20)

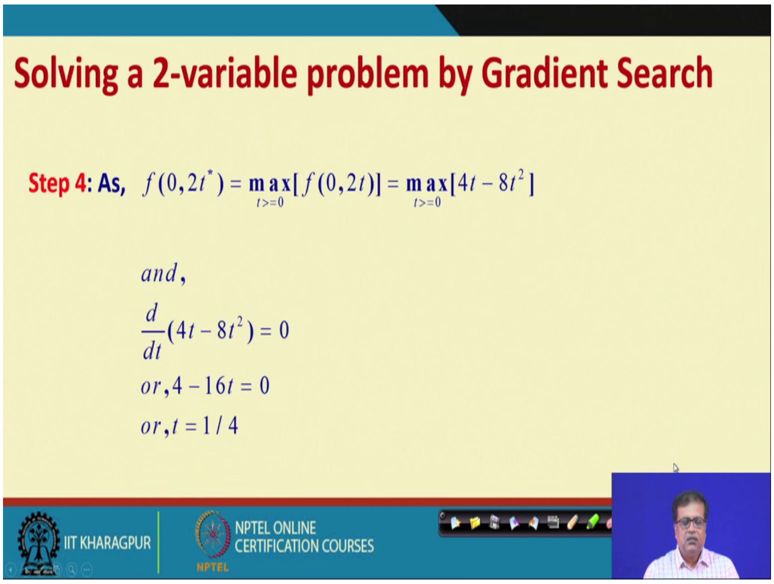



Now, what we do see look here this is our x 1 and x 2. So, please recall please if you if you recall what was our gradient. Gradient was 0 2 right so, 0 for x 1 and 2 for x 2. Now, the point we started the trial solution is 0 0 right. So, basically x 1 equal to 0 and gradient for x 1 is 0 right. So, gradient for x 1 equal to 0 and x 1 equal to 0, similarly x 1 equal to 2 and gradient x 1 equal to x 2 equal to 0 and gradient is 2. So, these are the things.

Now, if we move in the direction for t distance, then what will be the new value of x 1 and x 2 in terms of t? It should be 0 plus t star 0 that is 0 for x 1 and x 2 will be 0 plus t star 2 because gradient is 2 equal to 2 t. So, 0 and 2 t are the new values for x 1 and x 2 and what is the functional expression? This is our functional expression. So, if these functional expression value we find out for 0 and 2 t, that is our new functional value. So, what it will be? 2 into 0 into 2 t plus 2 into 2 t minus 0 square minus 2 into 2 t whole square. So, these comes to 4 t minus 8 t square.

See look here what we found. So, graphically if you want to know then we had a point. So, this side is x 1, this side is x 2 right. So, we are here, this is our initial trial point 0 0. Now from 0 0, we have come to you know a new point which is 0 2 t. Why? Because the gradient is 0 2, 0 for x 1 so, we are not moving in that direction and 2 for x 2. So, we know that if we move in this direction our functional value is going to go up, and the new functional value in terms of t is 4 t minus 8 t square that is all we have found right. So, I hope you understand.

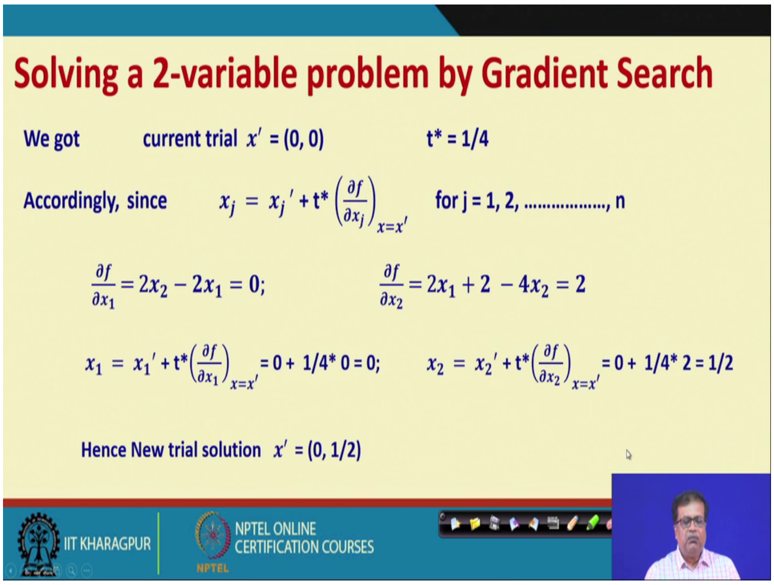
(Refer Slide Time: 26:47)




So, now with that knowledge, we try to maximize these value the functional value because that is all we try to do how do we do? We differentiate. So, as we differentiate this 4 t minus 8 t square we find the 4 minus 16 t equal to 0.

So, we find a value of t which comes out to be 1 by 4. So, you see how much we move to the direction is you know is 0 in the x 1 direction and 2 t in the y x 2 direction. How much is 2 t? T comes out t star comes out to be 1 by 4. So, that is essentially exactly what happens right. So, t comes to be 1 by 4.

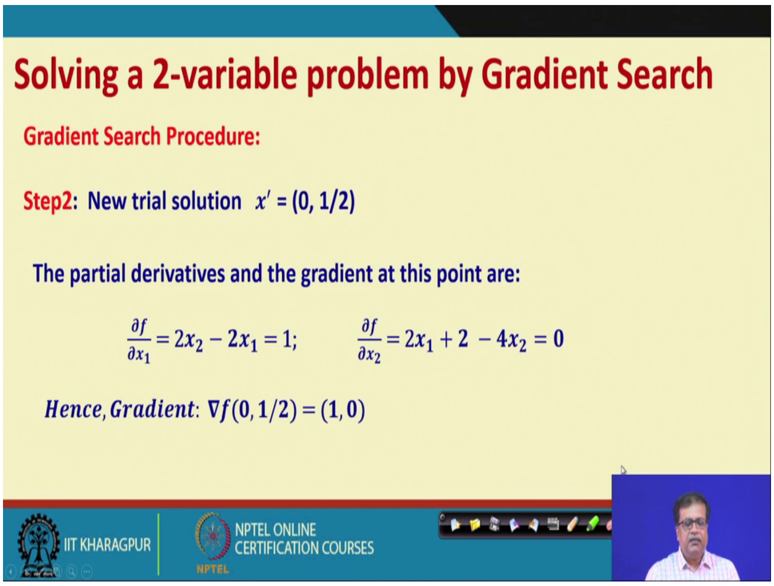
(Refer Slide Time: 27:37)




So, we got current trial 0 0 and t star equal to 1 by 4. So, accordingly since x j equal to x j star plus t into the you know the partial derivative at x equal to x dash, and we had 0 and 2. So, x 1 become 0 and x 2 becomes half right.

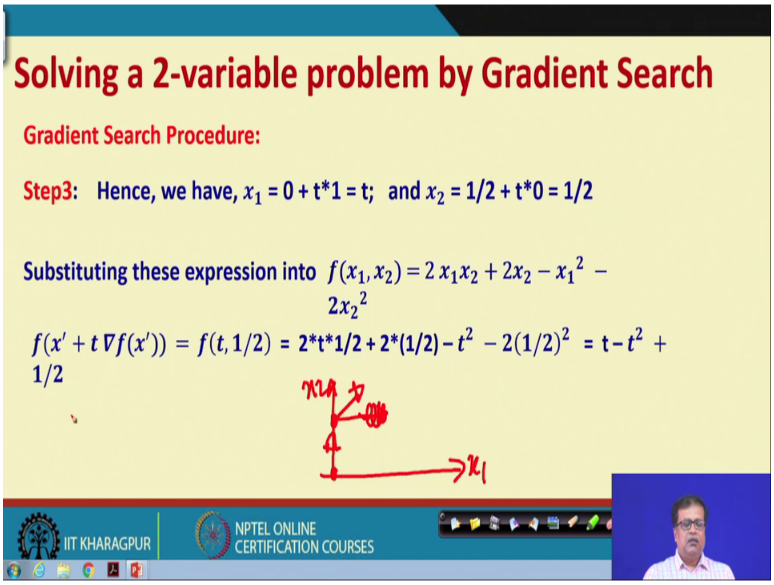
So, it is clear that you know one fourth is t and it has moved 2 t, that what we saw earlier. So, new point has to be 0 and half right. So, that is our new trial solution 0 and half. So, with that 0 and half we again iterate.

(Refer Slide Time: 28:23)



So, new trial solution 0 and half so, again we find partial derivatives. So, we know those partial derivative values. So, this time if I put x 1 equal to 0 and x 2 equal to half, then the partial derivative with respect to x 1 is 1 and partial derivative with respect to the x 2 is 0. So, the new gradient will now become 1 0 that is at that point. So, this time we are not moving towards x 2, we are now moving towards x 1 alright.

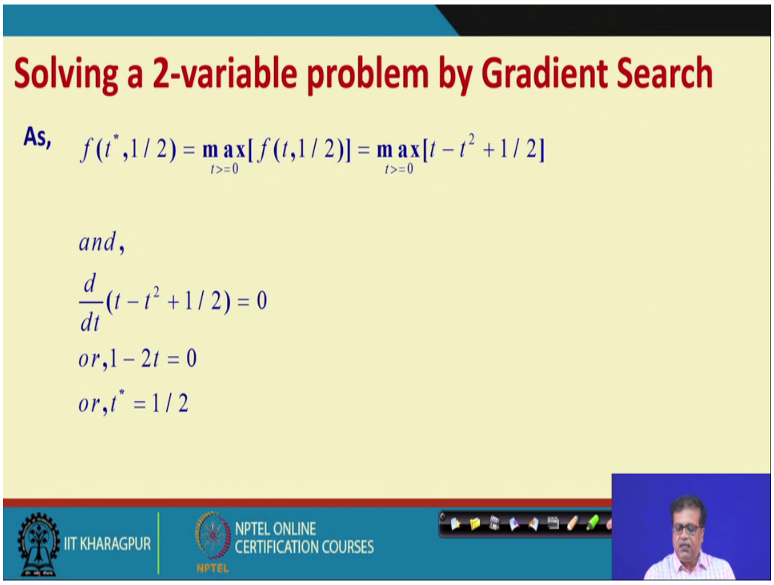
(Refer Slide Time: 29:03)




So, now we have x 1 is 0 plus t star 1 t, and x 2 remains at half. So, what happens if you recall? You know we started at origin, we have come to a point now we are going in this direction not earlier we have come in this direction this is x 1 direction, this is x 2 direction now we are moving in the x 2 direction because the although you know little bit on x 2 also because x 2 is half right.

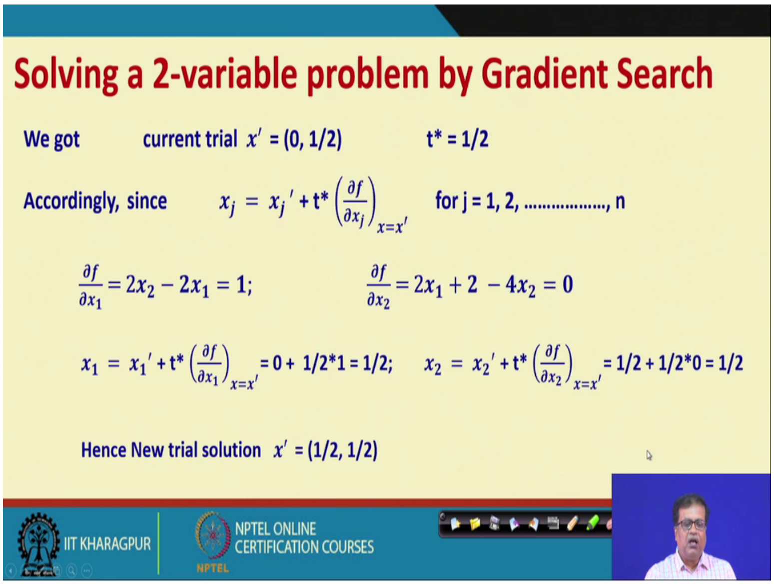
And not dependent on x 2 so, it will move slightly this direction, not this direction alright. So, if you substitute this expression, then the functional value becomes you know after calculation t minus t square plus half. So, this is going to be our new functional value.

(Refer Slide Time: 29:59)




So, with this functional value, what we can do further; we find out the functional value becomes maximize of this and if you if you differentiate again, then we find t star equal to half.

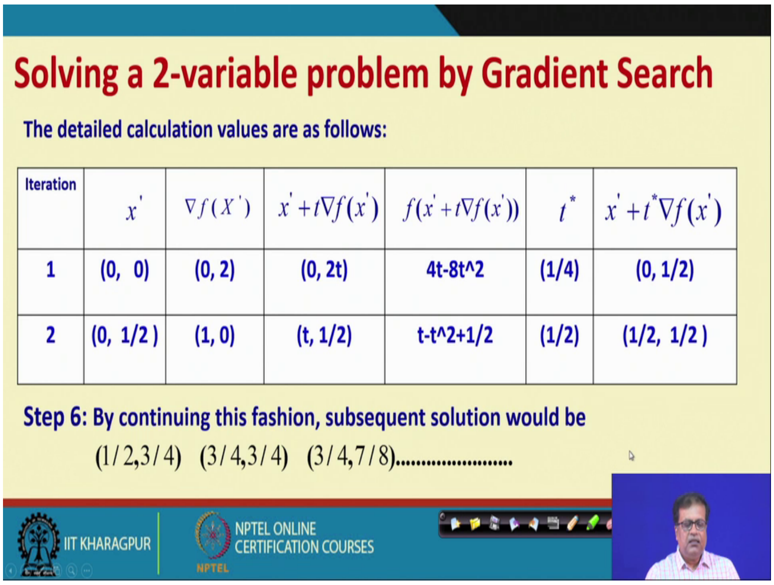
(Refer Slide Time: 30:14)




So, what is the significance of this t star equal to half? So, this is was our current trail x dash now 0 and half t star is half. So, x j values will be these are our differentials. So, x 1 will be 0 plus half into 1 that is half, and x 2 will be half plus half into 0 equal to half because at these point this is these are our the gradients right so; that means, we are not moving in the x 2 direction, although we got you know value in that directions. So, since the we are not moving. So, new point becomes half and half right. So, that is the new point that will become half and half.

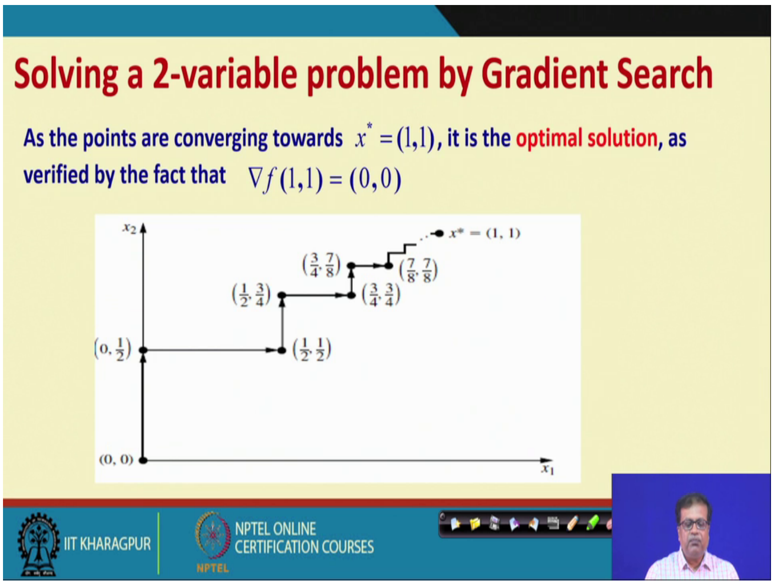
So, all of these are tabulated here in the next table.

(Refer Slide Time: 31:07)




So, initially we started with 0 0 then we got 0 and half, then from 0 and half we moved to half and half. So, if you continue in this fashion, we go to half 3 by 4, 3 by 4, 3 by 4, 3 by 4, 7 by 8 etcetera alright.

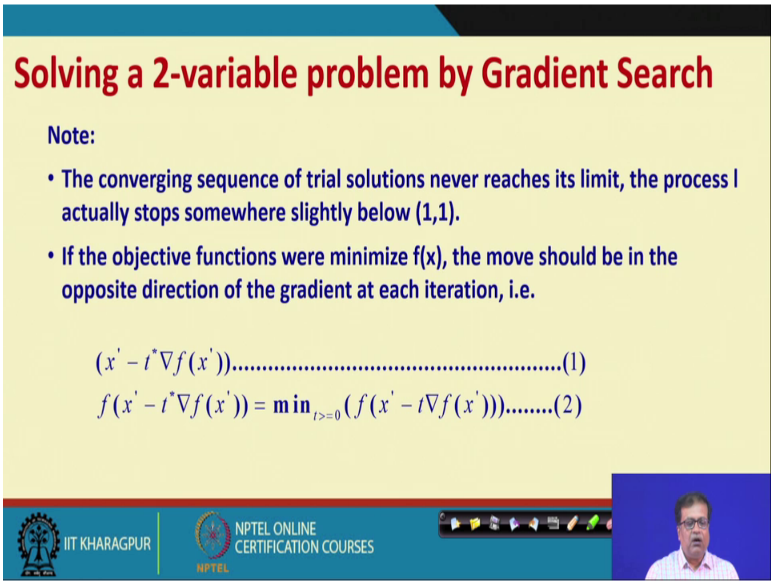
(Refer Slide Time: 31:27)



So, we go to all this points and it looks like these. So, I was showing you from 0 0 to 0 and half to half and half to half and three fourth, 3 by 4th to 7 8, then like this we keep moving and you know if you really look at the direction, we are actually moving to 1 1 it is the optimal solution, verified by the fact that del f 1 by 1 is 0 0. Because if you move any further then gradient will fall; that means, the curve will be like this right. So, that is going to be the optimal point.

So, really we may not be able to get to the optimal point.

(Refer Slide Time: 32:05)




So, the note should be the converging sequence of trial solutions never reached its limit, the process actually stops somewhere slightly below 1, 1. So, if the objective function were minimize fx, the move should be in the opposite direction of the gradient at each iteration. So, instead of plus it should be minus and instead of maximizing it should be minimizing right.

So, this is how without really computing the Hessian matrix and eigen values, if we cannot do that then from a gradient value itself we can do a gradient search right. So, in our next lecture we shall see some more numerical methods for such kind of problems right so.

Thank you very much.