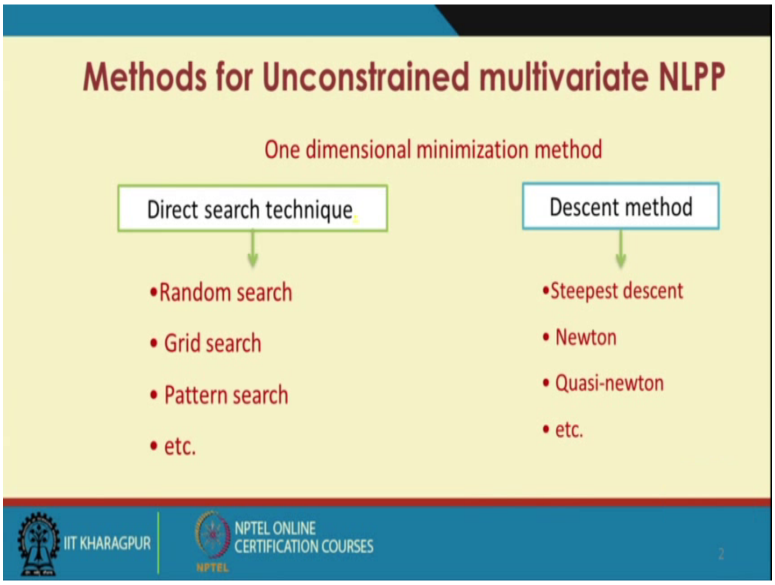
Constrained and Unconstrained Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 41
Multivariate Unconstrained Optimization- I

In continuation to my previous class on single variable unconstrained optimization today I am going into advanced to it that is multivariable unconstraint optimization. Here we are dealing the optimization problems where the problems are of unconstrained type, but the objective function that is non-linear in nature, but with in which involves more than 1 variables together. Now we need to optimize the function after that.

(Refer Slide Time: 00:54)



Now, for solving this kind of unconstrained multivariate non-linear programming problem there are a several methodologies available, but that is a multi dimensional minimization method, but here you see whenever we are considering any non-linear function, function may be continuous may be take discontinuous in the domain of definition. When it is continuous then our differential method that is the classical optimization techniques that is differentiation of function that works well, but when the function is not continuous in nature in that case that differential coefficient calculation fails, because there is discontinuity in the domain of definition.

In that case what we do we adopt searching technique there, but as you have learnt in single dimension my, single dimensional non-linear programming whatever searching technique is have you have learned those are not really working here because you might be understanding here we are dealing with a surface. If there is a function of 2 variables then in one axis one variable in other axis another variable, in the third dimension the function moves that way we are getting a 3 dimensional picture that is why instead of having the function in the curved nature we get the surface. That is why once we are getting the surface if we just try to implement the searching techniques you have learned for the single objective problem those are not working; there is a separate set of methodologies available for multivariate non-linear programming problem.

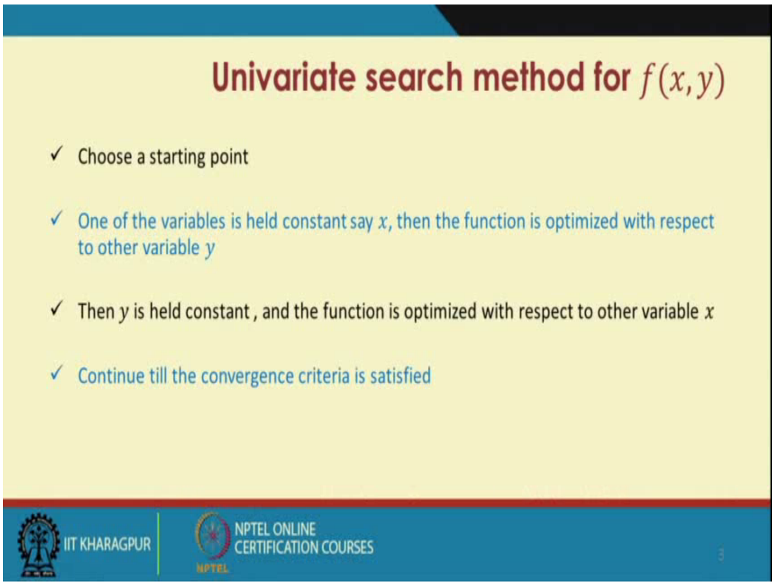
All the methodologies can be categorized into 2 types 1 is the direct search technique and another 1 is the descent method basically, direct search technique does not depend on the differentiation of the function whereas, the descent method is very much dependent on the differentiation of the function. That is why it is easily understandable; descent methods are not applicable when the functions are discrete in nature.

Now, I will start my classes with the direct search technique I will give you some idea how really we are dealing with the direct search technique. Now there are again different techniques available, 1 is the random search, grid search, grid walk then pattern search that is a pawls method there are a few methodologies like that. I will not cover all the methodologies in detail, but I will give you few, I will show you few methodologies how it is working for a multivariate non-linear programming problem one by one and in the descent method there is another set of methodologies available, starting from that steepest descent or steepest ascent method, Newton method, quasi Newton method etcetera.

Now today I am starting with the direct search technique that is the most popular technique is the search that I will show you first. Now here the random search technique random search what it does, it just exhaustively search the entire domain and it considers the points randomly within the domain to get to find out the minimum value, but how to generate the random points in the domain, we use the random number generated. Through the random number generator we consider different points on the space from their we are finding out a list of functional values with us and from there we are finding out the minimum or maximum according to our desire.

But one thing is there that as you know I have already explained you in my initial classes that, function may have global minimum or maximum or a local optimum that is why all the searching techniques whatever I will discuss now, these are not really very much efficient to find out the global optimal value. Not only that there is a curse of dimensionality as well because if the dimension of the problem increases if the number of variables increases of a function that is why you could understand finding out the random points from a space where n number of variables are moving in different directions in different ranges in different scales and generating random numbers for those that is really a difficult task. But how really searching techniques are walking I will they give you some idea of it and I am starting my class with the univariate search method.

(Refer Slide Time: 05:58)



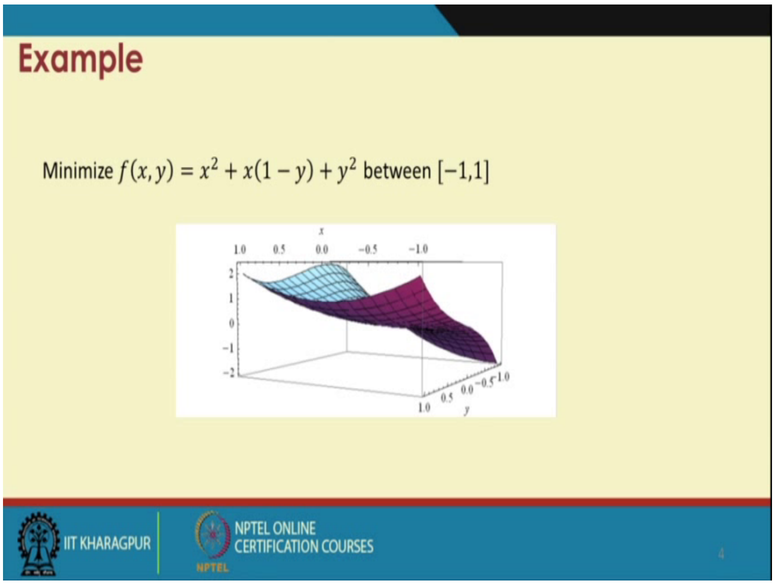
The popularity of this method is there because of the easy calculation because whenever we are considering a multivariate non-linear programming in the univariate search method what we do, we consider 1 variable at a time keeping other variables constant. That is why if there is a function f x y what we do we start from a point in the domain of definition of the function, after that we just move 1 variable at a time say x I will say whether x in the right hand direction or in the left hand direction, it what direction function is increasing or decreasing.

Once we could see that function is increasing in certain direction if my desire is to find out the maximum of the function I will move through that path, but on the other hand if we see the function is decreasing in another site then if we are desiring to find out the minimum of the function we will move in that way. What is the advantage? Advantage is that if I change x and y together there an infinite number of directions all together, within the space, but once I am changing 1 variable at a time that is why my options are left only I can move in the right or I can move in the left all right, that is the thing we are employing in the univariate search technique.

That is why if I just see the basic algorithm for univariate search we start with a starting point, one of the variable is held constant say x then the function is optimized with respect to the other variable say y. I can do the other way, I can keep why as constant and I can change x as well then in the next I will do the reverse process.

Either I will make y as constant, I will just change the x in that way I will just move and I will do repeated iterations there and after that we will see that with method will converge because somewhere we will get the local optimum value there we will stop that is the basic idea.

(Refer Slide Time: 08:26)



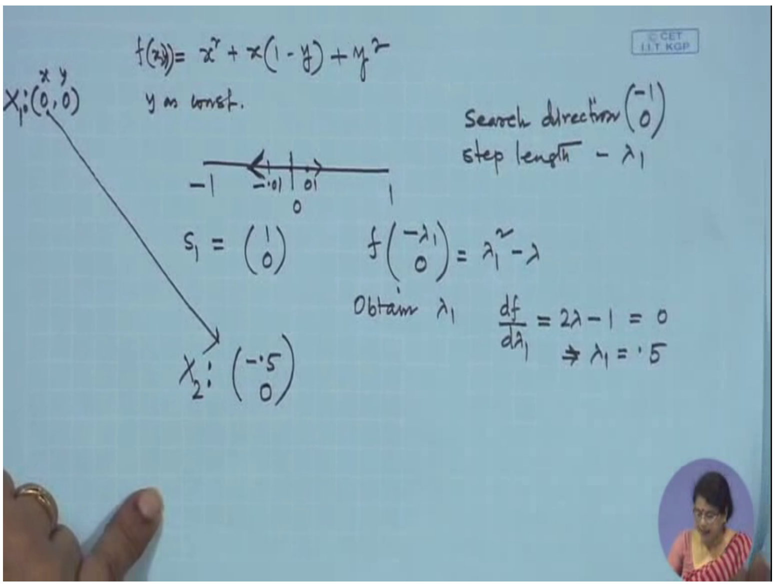
Now, for doing that thing I have taken 1 example first you say the example we are trying to minimize f xy where x square plus x into 1 minus y plus y square and the function will we need to find out the minimum value within the range minus 1 and 1.

Now, if you see the pattern of the function, pattern of the function is this one I have drawn this pattern in the mathematica. Now x and y, z are all moving and we could see that that several minima occurs, function is not really unimodal within the range whether the function is unimodal or not it is very difficult to find out all the time ok.

Now, if this is the thing now how really we can find out the solution for this problem, we need to find out a local minimum. Now if you are interested to find out the global minimum then you have to repeat the process by starting from different points within the range, I can start from minus 1, minus 1 because this is a whole square all right. Now the x is same actually this is written I should write x in changing from minus 1 to 1, y is changing from minus 1 to 1 that so we are getting a square of it I can start from minus 1 and minus 1 taking both an x and y I can state, I can start from minus 1 a 1, I can start from 1 when, I can start from 0 0 from anywhere I can start.

Depending on my starting position convergence will be reached quickly or not that depends on that, because in the search process if the local minimum is there very nearby to the, my guess point that is my starting point then quickly we will reach to the optimal solution. Otherwise I need to find out I have to repeat the process by starting from different base points and we will have a list of local optimum from there we can select the global (Refer Time: 10:39) that is there, but I will show you now how to reach to a local optimal point for this function at least all right.

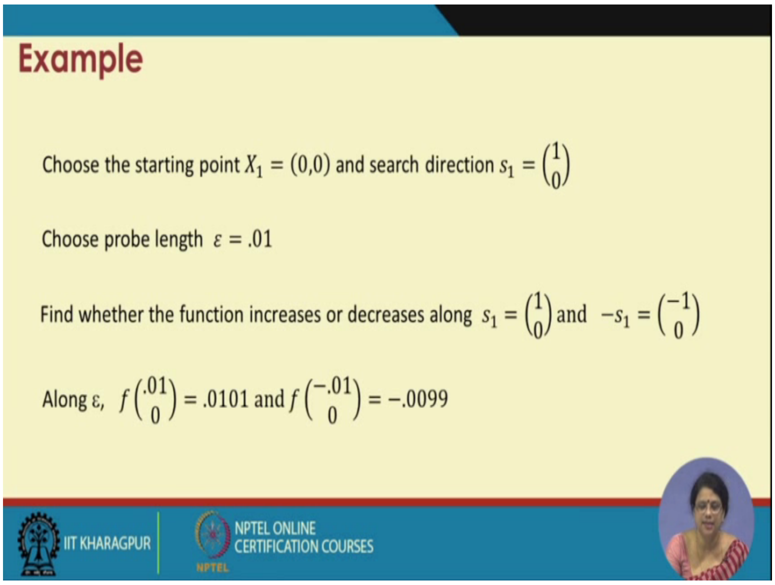
(Refer Slide Time: 10:55)



Now, let me start the process. Now you just see this is my function x square all right. Now my starting point is 0 0, this is my first point all right. Now if I start from 0 0 this is x, this is y, I can keep x as a constant I can change y or I can keep y as a constant I can change x, all right. If we change x keeping y as constant, what is happening you see x is, this is there. Now if x is I am starting from x equal to 0 I can move this way or I can move that way, I will see if I move this way what is the functional value whether the functional value is increasing or not and if I just move this way, what is the functional value increasing or not I have to see that one. That is why you see there are 2 things are involved one thing is that the search direction in which direction I will do the searching and the second thing is that what is the step length of searching, shall I jump a lot or shall I move very with a small delta, delta that is where 2 things are involved your search direction and the step length.

You see I have considered the search direction 1 0 all right and let me consider a step length that is 0.01 all right.

(Refer Slide Time: 12:48)



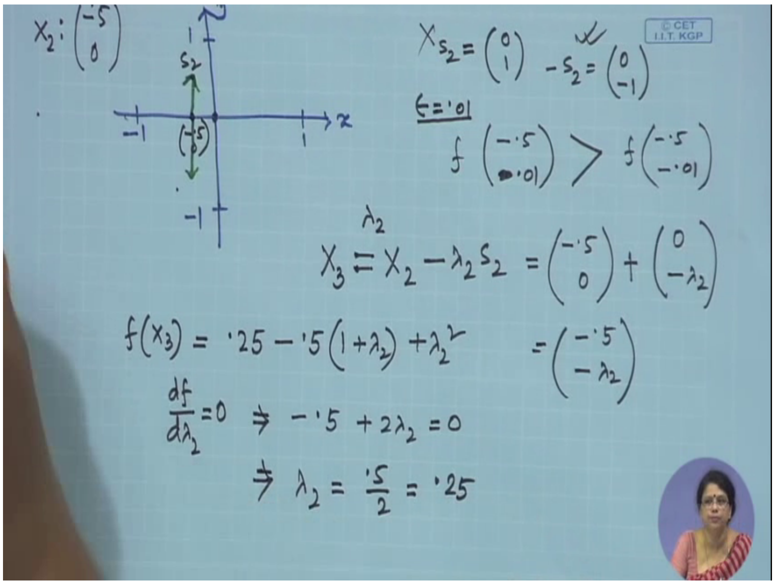
If we just do that, we see that the functional values at f 0.010; that means, my search direction was, my search direction was 1 0 all right and my step length say this is here I have considered epsilon. Now this is taken as 0.01 then if I find out the functional value at 0.010; that means, I am moving from here to 0.01 only all right by keeping y as constant 0 we could see the functional value is 0.001 all right and what is the functional value at 0, 0 that was 0 only, that is why functional value is increasing in the right hand side.

Now, you just take the reverse process. Now the reverse is that instead of going this way I am going the other way, that is why my search direction is minus 1 0 all right. Now again I am taking the step length in the left direction as 0.01 then by point would be minus 0 1 0 all right. I am seeing an again the functional value what is the first change in the functional value if I just look at the values we could see that functional value is increasing in the right and decreasing in the left, that is why what should be my desired direction? My desired direction all of us will be in the left because I am searching for the minimum of a function all right, but still you see we have taken the probe length rather the step length initially as 0.01 that should not be, because step length I have taken in a in an ad hoc way we should we should have said certain mathematical logic how much step length we should take.

So, that quickly we can reach to the optimal solution it may happen instead upon 0 1 if I could have taken 0.02 then we could reach in the same way, the functional pattern is being mentioned through that that is why whether to take 0.01 or 0 2 there is a guideline to it. That is what I am saying that I am moving in the left is with a step length lambda 1, what it is, which is variable all right then I can say let us let me find it out the functional value as minus lambda 2, lambda 1 0 all right; that means, my star search direction is minus 1 0 I have seen the functional value is decreasing in that direction and step length I have considered as a variable that is lambda 1.

Since it is the in the left direction I am considering f of minus lambda 1 0, if it is substitute here what I am getting lambda 1 square minus lambda all right. Now what is my aim, my aim is to find out that obtain lambda 1 which optimizes f all right; that means, the first order derivative of f with respect to lambda 1 must be 0. What we do? We will consider df by d lambda 1 this is coming as 2 lambda minus 1 is equal to 0, if I consider this we are getting lambda 1 is equal to 0.5 that is why it is clear that from this point I will move to the next point as minus 0.50 that is why I am starting from x 1 0 0 I am moving here this is my next point x 2 all right.

(Refer Slide Time: 16:54)



If I want to explain this one in the 2 dimensional space, just you see what is happening this is the x, this is the y, direction all right my range is from minus 1, 1, 1, minus 1 all right. What we do is, we did we started from here we checked everything and we could see that if I just jump to the next this point which point is this point is minus 0.50 then my searching would be optimum all right.

Now, after reaching here what I will do, as I say that in the last iteration we took x as a con x as a variable y as a constant here, we are doing the reverse, we are considering y as a constant y as a variable x as a constant if we keep x as a constant what is happening or c we can move. If I just stays there why either we can move this way or we can move this way is it not all right because I kept x as a constant if I move this way what is my search direction, my search direction would be 0 1 all right this is the s 2 and if I move the reverse way then it would be minus 0, minus 1.

Let me see again whether the functional value is increasing or decreasing, we could see that if I move this the in this direction; that means, with a, with a small probe length epsilon is equal to 0.01 again. Again I am not doing any optimization here to find out what should be the optimum step length. So, that I could reach the solution very quickly that, but if I even do not do, if I just consider a small shift of wipe on 0 1 then my initial point was x 2 when where I am reaching that was minus 0.50 then my here my point would be minus 0.5 minus sorry plus 0 1 all right. I will see the functional value here and I will see another in the lab in the lower part minus 0.5, minus 0.01 all right.

If we just calculate the values I have seen this value is lesser than this that is why it is we are showed that this direction is the best direction to get the minimum this is not, clear. Again the problem comes if this is. So, again the problem comes I am not moving this way I am moving this way from here to here I am moving, but what would be the optimal step length, that is again the question comes.

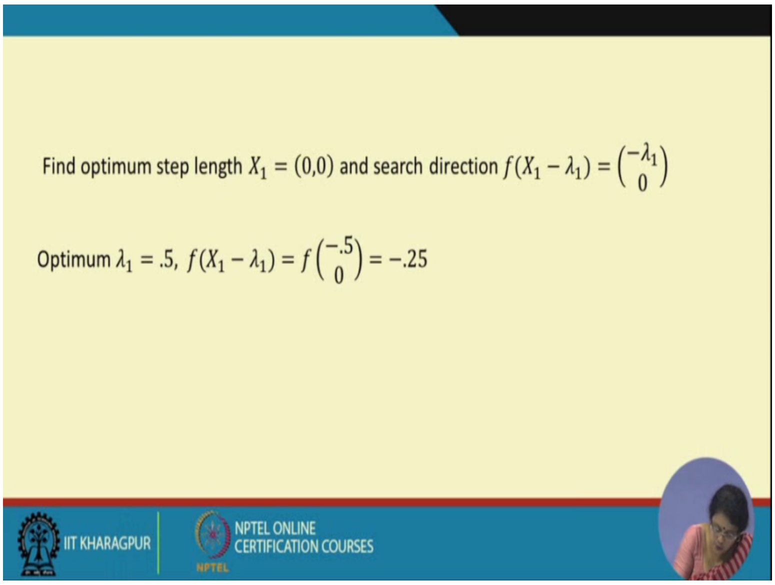
Let me consider the optimum step the step length as lambda 2 in this case, lambda 2 if I consider as a variable then in the next what we will do we will find out the point x3 which is is equal to x2 minus lambda 2 s2 all right. Because if s 2 if I am going in the positive direction it should be x 2 plus lambda 2s2, but here we are considering the negative direction that is why x 2 minus lambda 2 s 2 all right.

Now, in this point, what is the point is coming if I consider this is minus 0.50 this is the direction is 0 minus 1 that is why 0 minus lambda 2 all right, if I consider this point is coming, I cannot take both has a minus that is why if I can see there minus here I should take plus here, all right minus lambda 2.

Now, let me find out the functional value at x3, that is a function of minus points with my a function of lambda 2 rather then what would be the function, I will get the function x square plus x into 1 minus y plus y square, again I am doing the first order derivative and I am equating to 0. What we are getting? We are getting minus 0.5 plus 2 lambda 2 is equal to 0, which implies lambda 2 is equal to 0.5 by 2 that is is equal to 0.25, clear.

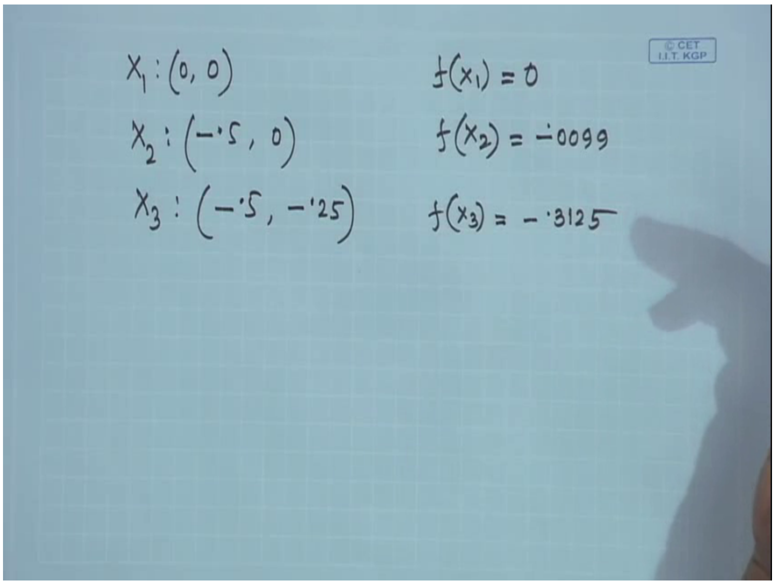
Now, if we just move in this way, then what should be my x 3 in the next if I just write down then my initial guess was 0 0.

(Refer Slide Time: 22:25)



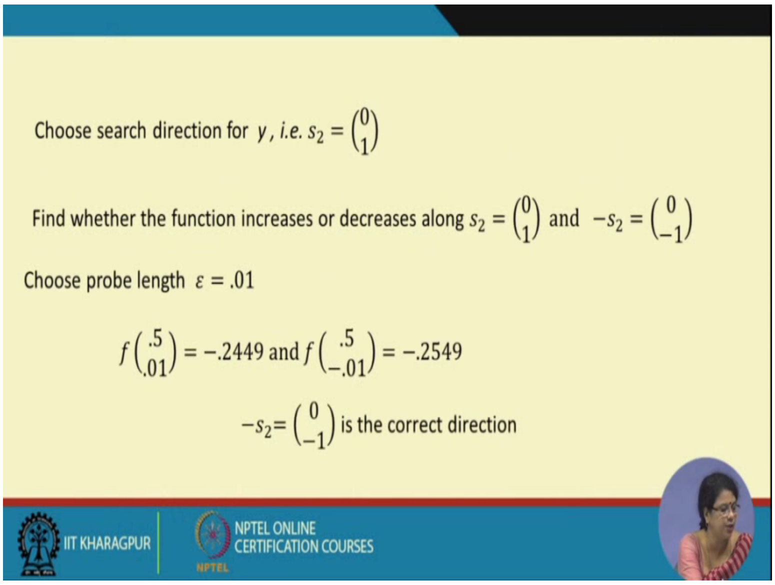
My next point was minus 0.50 my next point is coming minus 0.5, minus 0.25 all right.

(Refer Slide Time: 22:32)

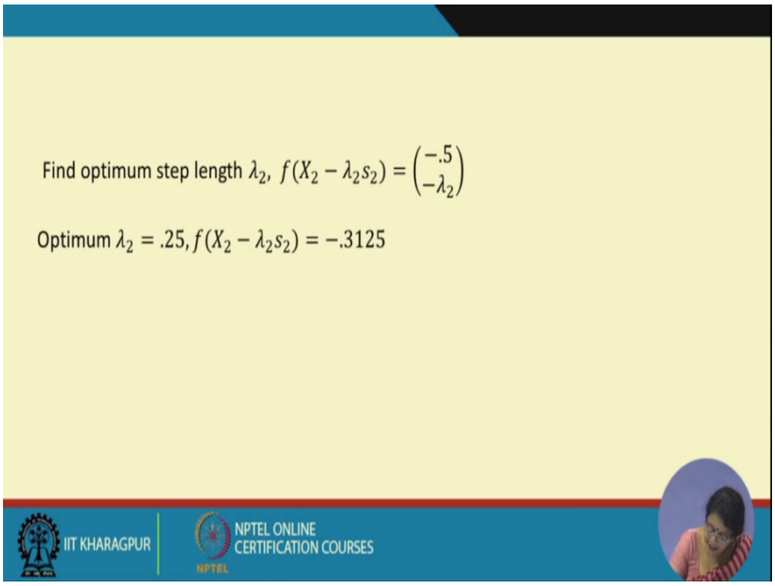


Now, let us see the functional value at this point, these are the functional value as I was saying to you and we could see.

(Refer Slide Time: 22:53)



(Refer Slide Time: 22:55)



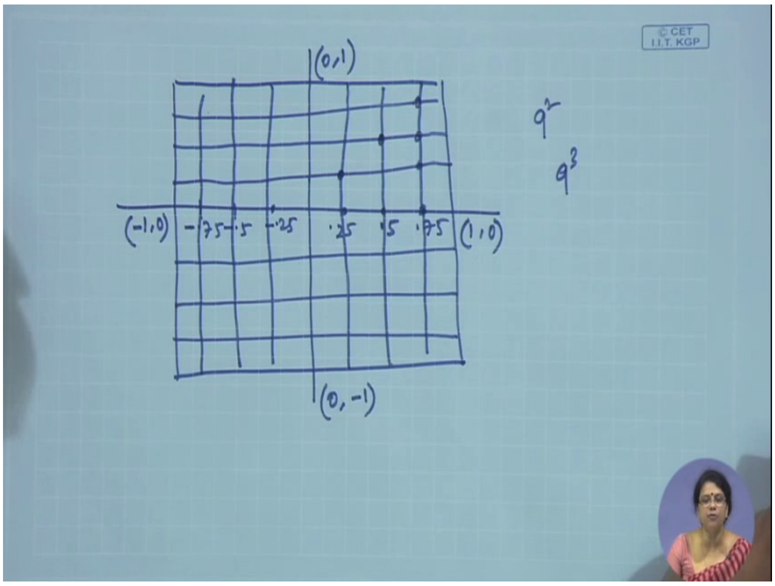
The functional value at x3 is coming as minus 0.3125, but previously the functional value was it was 0 it must be more than that, that was 0.0099 like that and now we are getting much better value ok.

Now, the question comes we can stop the process here, but we can move further even how to move further again I will keep x as a variable y constant, I will move this way, but it may happen if this is a local minimum and if my step length is very small I will see after this functional value will never decrease, functional value will increase in the neighborhood of this point.

This is that is why we can declare that is a minimum, that is why question comes whether this point it is at all a minimum or not, local minimum or not, but there is other method as well I have checked with the great search technique that this method gives us the global minimum. Not only there are several local minimums with the same values, but this is as well the global minimum, but how really you can sense it, when you will do the mathematical process that time we will see that if I go in the right direction or left direction we will see there it there all be any improvement of functional value. That means, they are only any possibility to get any minimum value than this we have to stop the process, but we cannot.

But this is one of the univariate search method you must have seen this is a very easy method, but there is another method that is mathematically very easy, but very difficult if you just can do the calculations on your own, that is called a grid search method I am just going to tell you that grid search technique now.

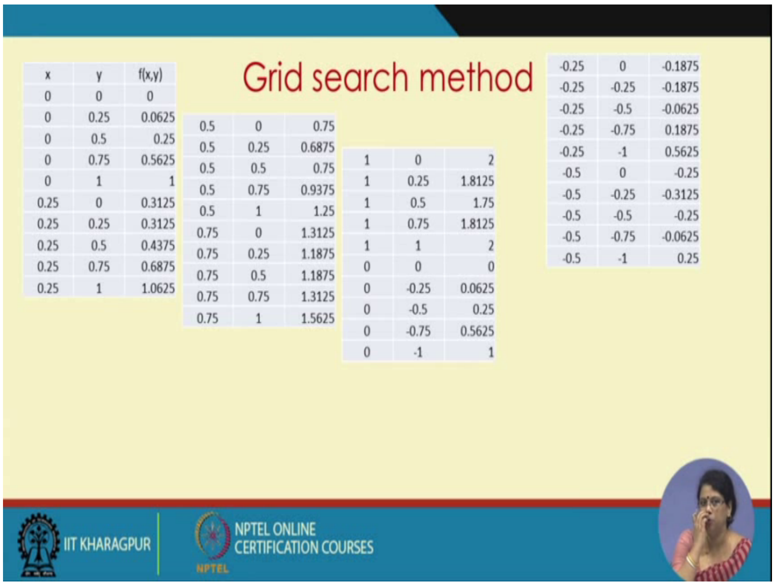
(Refer Slide Time: 25:11)



Now, this is my domain of definition I am getting this is the I am searching minimum within this range, this is my minus 1 0, 1 0, 0 1, 0 minus 1, what I can do? I can in the grid search technique what we do we just divide the whole space in different grid, with the grid points how really. Let me consider it that I am taking this 0.5 I am taking the grid point like this, here also the same thing minus 0.25, minus 0.5 minus 0.75 I can do the I can take the grid like this all right.

Similarly, I can take the grid points along y axis by considering the same way minus point to a plus 0.25, plus 0.5, plus 0.75 minus 0.25 minus 0.5 minus 0.75. Now function is moving here as I showed you the picture what I will do I will find out the functional value at each grid point. For finding out the functional value at each grid point will not be difficult for you because you can write a simple program for it first you find out you just divide partition the whole space with different points and you just see, what is the pattern of the function there? How many points you will get here, you will get total here the 9 here 9; that means, you are getting the 9 square points when we are considering 2 dimension, when we are considering 3 of dimension it will become 9 cube. If we have 4 variable 9 to the power 4, that is why you can understand if I increase if the number of variables are increasing though we could guess what sheets the pattern of the function.

(Refer Slide Time: 27:26)



But still find out the functional value is really difficult here, for this problem I found out few values of this in that way you can find out 81 points and a functional value at 81 points and you can you could see, what a how the functional values are just moving here. Now that way very nicely you can find out the local minimum or if you want you can find out the global minimum. In the next class I will continue my class with the same topic.

Thank you for today.