

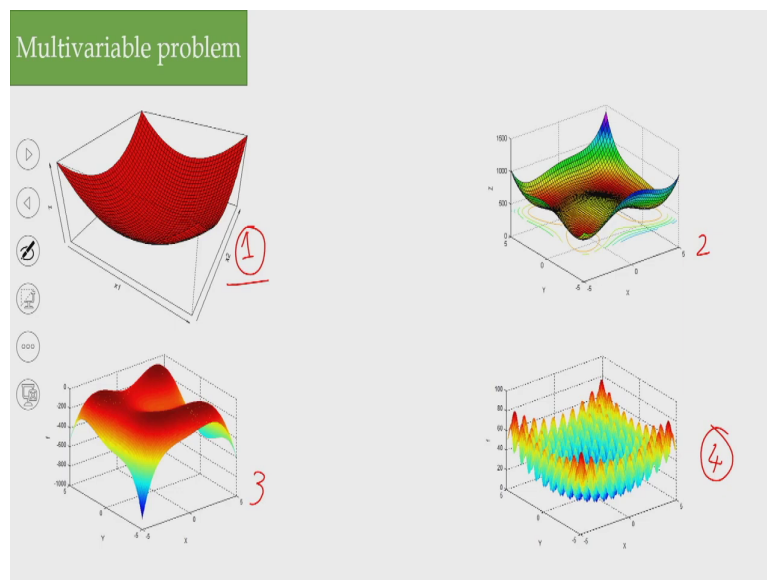
Optimization Methods for Civil Engineering
Dr. Rajib Kumar Bhattacharjya
Department of Civil Engineering
Indian Institute of Technology, Guwahati

Line Search Methods for Multi-variable Problems
Lecture - 10
Convex Function

Welcome back to the course on Optimization Methods for Civil Engineering. So, in the last class we have discussed the search technique for single variable optimization problem and we discuss about mainly region elimination technique and before that we have applied the bracketed algorithm to bracket the optima and after that we have applied mainly region elimination technique like golden section search technique, then interval huffing technique, then bisection technique and also we have discussed Newton Raphson method.

So, you can also apply Newton Raphson method. So, that is to find out the optima of a single variable optimization problem. Now, in today's class we will discuss about multivariable functions. So, we will basically see how you can search the optimal point of a multivariable function.

(Refer Slide Time: 01:41)



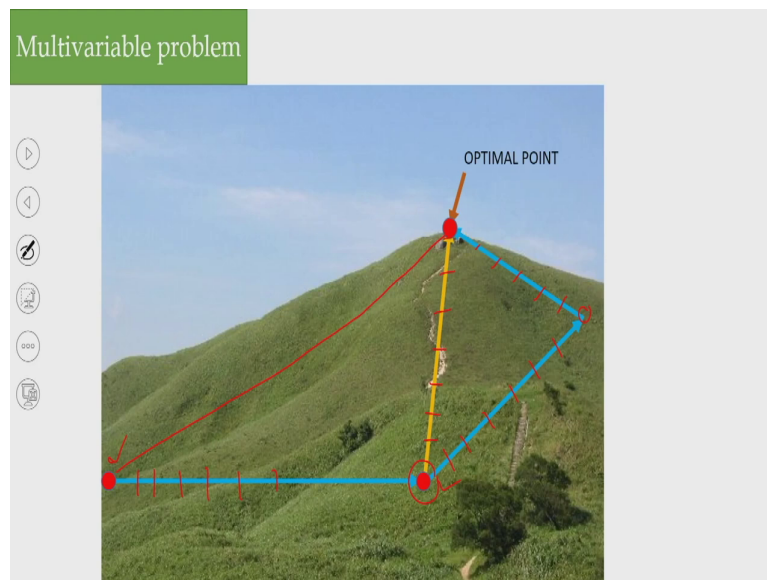
So, here I have shown you two variable functions. So, if I say this is function 1, this is relatively a easy function. So, there is only one optimal solution and I can also say that this particular function is a convex function. Then if I say that this is function 2 that is the second function and this is little bit complicated function because we have 4 optimal solution of this particular function within this search space.

And 3rd one if I say this is 3rd one, this is also relatively a difficult function because in this case also we have more than one optimal solution. The fourth one is a very complex function there are several optimal solution you can say the local optimal solutions of this particular problem and one of them is the global optimal solution. Now whatever line search method we will discuss. So, we will basically assume that the problem is of this kind ok so; that means, the function is convex ok.

And for the convex function. So, we will apply the line search technique here. So, we will discuss few line search technique how we will solve this particular problem and you can also apply the algorithm on other functions, but you may need some trial to get the optimal solution, but certainly I can say if you apply the problem on the 4th function.

So, it will be quite difficult to solve this particular problem using this algorithm. So, let us assume that our function is of first kind that is the convex function and we have only one optimal solution of the problem and let us discuss the algorithms.

(Refer Slide Time: 03:49)



Now, this algorithm can be explained using the hill climbing problem. Suppose, this is the optimal solution or I can say this is the maxima of this particular function or particular hill and I would like to find out this particular optimal point. Now suppose you were here suppose, you were here at this particular location. So, you were here and what you will do

basically? So, I can go along this direction I can go along this direction and I can do the line search. Suppose, if I go along this direction then my problem is a 1 dimensional problem.

Basically what I will do? I will give 1 step, then another step, another step like that and once I will reach the optima and I can declare that this is my optimal solution. Once you are going along a particular direction. So, your problem is now not a multivariable problem, but it is a single variable problem along that direction. Now if you are taking suppose this is the point if you are taking and I do not know which direction I should go then suppose if you are going along this direction initially then in this direction the problem is a single variable problem.

So, I can give one step one step like that and I can find out this is the optimal point in this particular direction. Similarly, suppose next direction is this and in this direction also this is a single variable problem and you are getting the optimal solution somewhere here and maybe I can go along this direction and I can finally, find out.

So, therefore, the idea of solving a multivariable problem is to convert the problem to a single variable problem and then you apply the line search technique you have already learned that is Golden-Section search method then Interval Huffing method, Newton Raphson method and Bisection method. So, any method you can apply.

So, the problem here is that, I do not have any specific algorithm for solving a multivariable problem, but what we are doing basically. So, we are converting the problem to a single variable problem then we have the algorithm for solving single variable problem then we are applying that one and we are getting the optimal point.

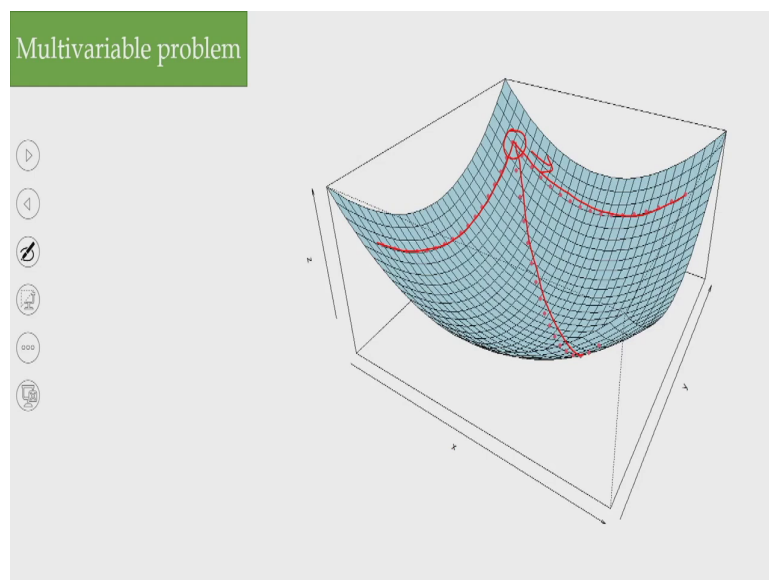
So, idea here is that, you have to get a direction you have to get a direction and suppose if I say I am here right now, I am here and somehow if I am getting this particular direction and if I go along that direction I will get the optimal solution.

So, therefore, you have to get a direction and you go along that direction, find out the optimal point, then you change your direction then you go in other direction and find out the optimal

point. So, that way after some iteration you will reach the optimal solution of that multivariable problem ok this is the idea.

So, here as I said that actually we are solving a single variable problem using the region elimination technique or other gradient based technique to find out the optimal solution of the function. So, this is the basic idea. So, we are converting the problem to a single variable problem.

(Refer Slide Time: 07:12)



So, let us see a function. Suppose, this is a function and here if I take a point here ok. So, if I take a point and I can take a direction suppose the direction is initially in x direction; that means, in this particular direction.

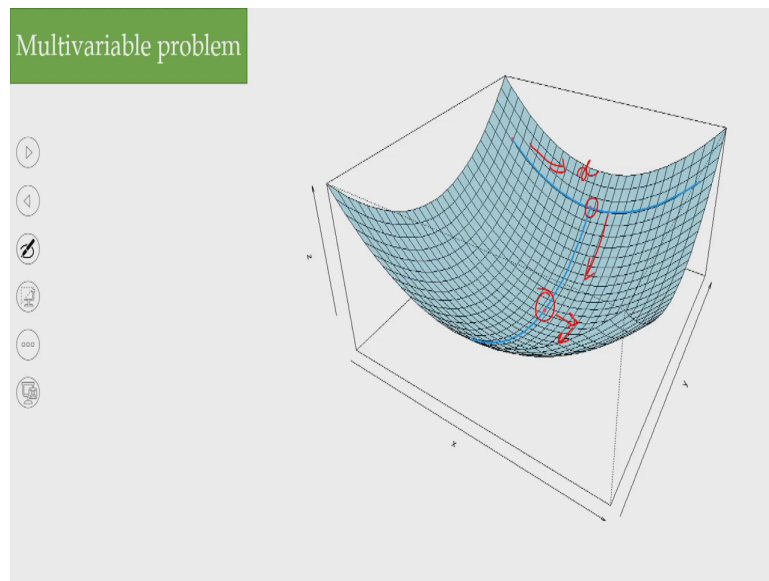
So, I can convert my problem means, if I get the direction to a single variable problem; that means, I would like to find out the optima along your direction. So, I can search it ok. So, I can search it along this particular line and then your problem is a single variable problem and you can find out the optima.

So, you can do the line search and once you are converting your problem to a single variable problem. So, you can apply the algorithm already you have learned to find out the optimal solution along that direction. I can go also along this direction; that means, in y direction. So, this is also a line search along this particular direction and similarly I can go along this direction also ok. So, this is the direction.

So, now question is that, I have to fix a direction to convert my problem to a single variable problem ok. So, this is you have to do that one. So, now, how to solve this problem? So, as I have explained that, suppose a simple algorithm is that you go along x direction initially and then you go along y direction.

So, basically what I will do that, I will go along x direction find out the optimal solution along that direction and then I will go in y direction, then I will find out the optima in that direction. So, if I am not getting the optimal solution again I will repeat this procedure.

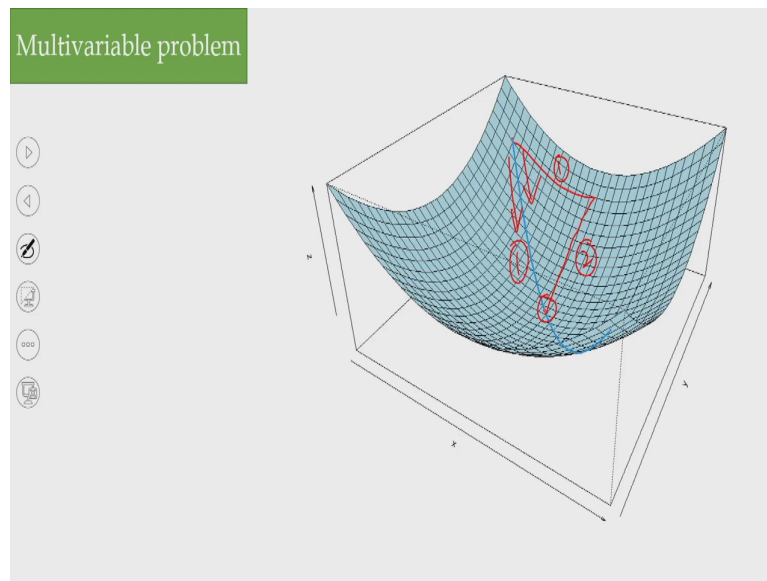
(Refer Slide Time: 09:05)



But in this case I will get the solution in two iteration. So, first I am going along x direction. So, this is along x direction. So, this is along x direction and the optimal solution is somewhere here ok. So, this is the optimal solution. Now what you have to do? You have to go along y direction ok.

So, this is the y direction and if you are getting the solution in the y direction. Then I can repeat this process again I can go along x direction and then y direction, but in this case I got the solution of this problem. So, this is the solution of this problem and so, I will not repeat my iteration. So, here in 2 iterations I am getting the solution. So, initially I have search along x direction and then y direction, but I can also do I can search along this direction itself.

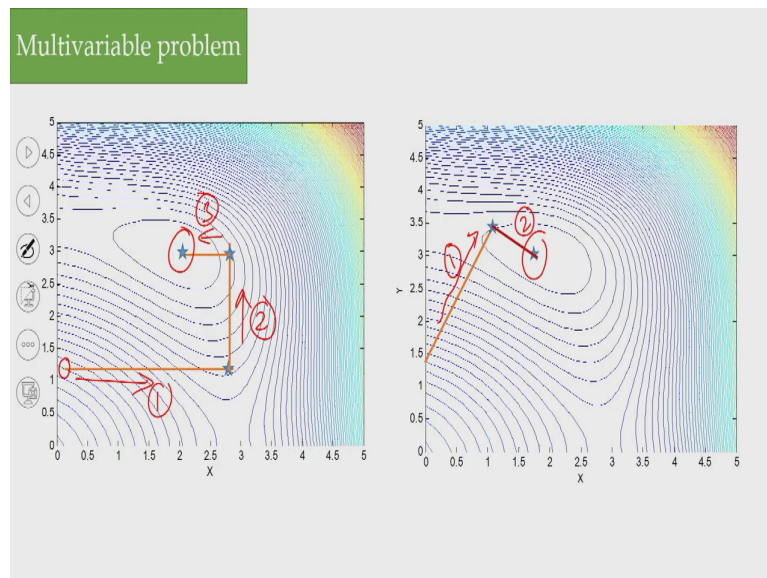
(Refer Slide Time: 09:55)



So, if I know that if I go along this direction, I will get my optimal solution in one iteration ok. So, this is the solution of this problem. So, therefore, that your solution process or you can say that iteration number of iteration will depend on how you have chosen your search direction. So, as I said that if I go along x direction then y direction. So, I need one iteration and two iteration. So, I need two iteration.

But if I go along this direction in single iteration I am getting this particular solution. So, therefore, there are several algorithms to fix the direction and that we will basically discuss in this particular lecture ok. So, today we will mainly discuss about how to fix this direction ok. So, whether you will go along x direction, y direction or whether you have some other algorithm or other method to get the direction ok. So, that we will discuss in today's class.

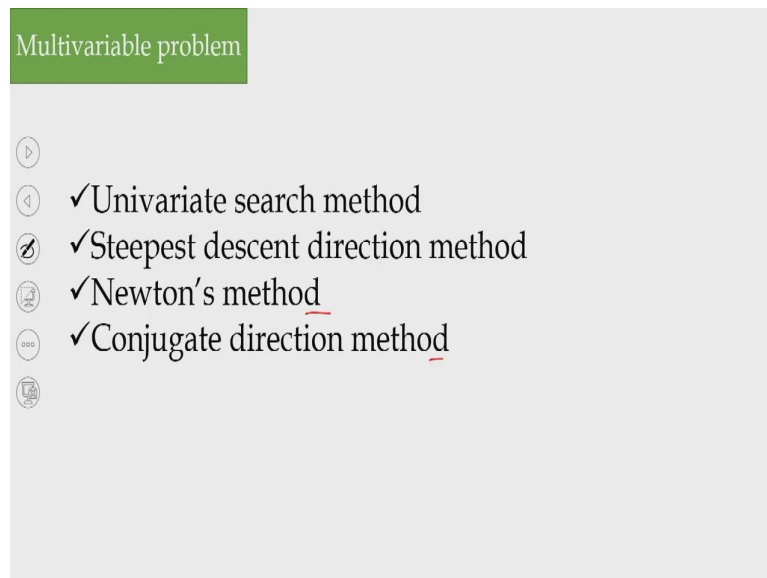
(Refer Slide Time: 11:05)



So, as I said the idea is that. So, in this particular problem optimal solution is somewhere here and you have start your algorithm from here. So, you are going along x direction, then along y direction, then again x direction something like that.

So, here probably in 3 iteration I will get the optimal solution of this particular problem. Now, the same problem if I go along this direction. So, this is 1st iteration and this is 2nd iteration and probably in 2nd iteration I will get the solution. So, therefore, how to choose the direction ok.

(Refer Slide Time: 11:45)



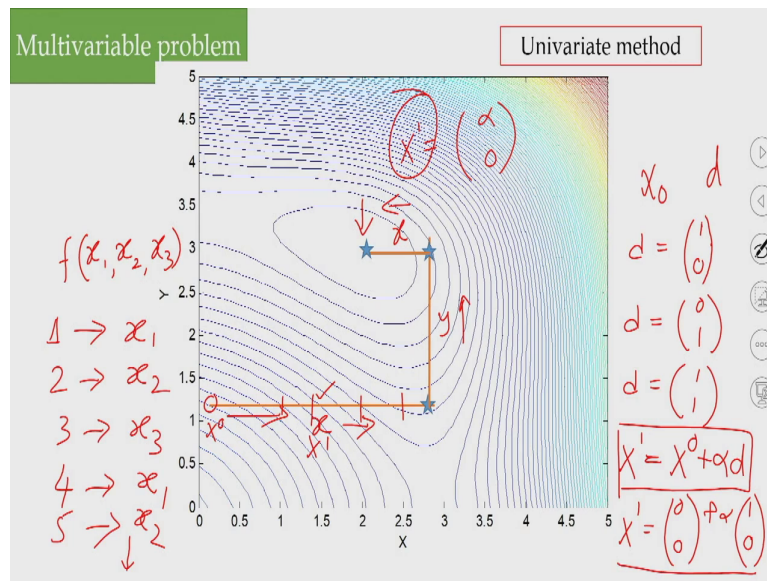
Multivariable problem

- ✓Univariate search method
- ✓Steepest descent direction method
- ✓Newton's method
- ✓Conjugate direction method

So, that let us see. So, we have several methods you can say, these are Univariate search method. So, one is Univariate search method, one is steepest descent direction method, then third one is Newton's method and there is another one conjugate direction method there are some other method also.

So, I am not discussing here, but I will mainly discuss these four method and these four methods we are doing these four methods will give you the direction ok. So, this is basically how you will get the direction and how we will perform the iteration. So, these four method we will explain. So, they have different concept.

(Refer Slide Time: 12:26)



Now, the first one is as I said this is a Univariate method; that means, you are selecting one variable at a time; that means, in this case what you are doing? You are going along x direction; that means, you are searching your optimal point suppose this is the starting point you can say this is x_0 and initially you will go along x direction; that means, you have chosen only one variable.

So, if you have X and Y variable; that means, initially you will go along x direction then second time you will choose another variable that is y. So, you are going along Y direction then again X direction. So, this method is known as Univariate method; that means, what you are doing? You are going along one variable at a time basically. Suppose, if your function is f equal suppose you have 3 variables ok x_1, x_2, x_3 . So, in the 1st iteration you will go along

x 1 direction, in the 2nd iteration you will go along x 2 direction, in the 3rd iteration you will go along third direction.

And then if you are not getting your optimal solution in the 4th iteration again you will go along x direction, in the 5th iteration you will go along x 2 direction ok. So, that way I will continue; that means, at a time. So, I will search my optimal point in one direction that is either x, x 1, x 2, x 3 and if you have suppose more variables that you have x 4 also, then in the 1st iteration it is along x 1 direction, 2nd iteration x 2, 3rd iteration x 3, 4th iteration x 4 and then I will repeat that one.

But in this particular function we have only two variables ok. So, that is x and y. So, therefore, I will go along x direction initially and then y direction, then I will repeat along x direction then y direction something like that. So, at a time I will go along one variable. So, therefore, this particular method is known as Univariate method. So, concept is very simple what you are doing? You are taking x initial point; that means, I will search my algorithm from x naught and what you need? You need a direction.

Now, how to define a direction? Now if I say that direction is 1 0 direction is 1 0 so; that means, what you are doing? You are going along x direction. Now, if I say that direction is 01; that means, you are going along y direction ok then I can also say the direction is 1 1; that means, you are going along 45 degree line. So, let me explain that one. Suppose, I would like to find out a particular point here suppose this is my x naught and this is a particular point I can say that this is your x 1 ok.

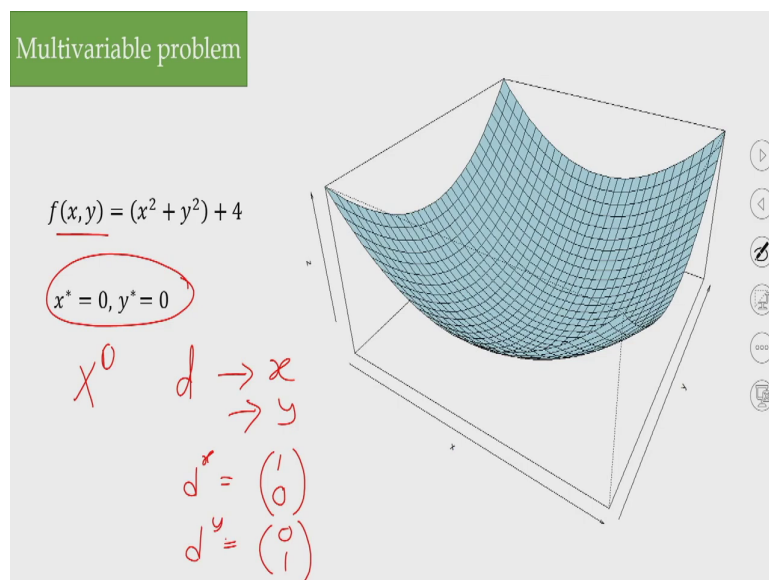
Now, what is x 1? So, X 1 I can write it like this that X naught plus alpha d ok. So, I can write it like this X 1 equal to X naught plus alpha d. Now what is X 1? So, this is suppose my x naught is 0 0 ok. So, this particular x naught is 0 0 and then this is alpha and d is 1 0. So, d is 1 0; that means, what is X 1 here? So, I am just writing here the X 1 equal to.

So, I can write that this is alpha and this is 0 is not it? So, this is alpha and 0; that means, if I write that alpha equal to 1, then I will get suppose this point. If I write alpha equal to 1.5, then

I will get this point then if I write alpha equal to 2 then I will get this point 2.5 this point. So, something like that now if I put this particular X 1 ok.

So, in the original function my original function will be a single variable function; that means, the function I can express that particular function in terms of alpha. So, therefore, the problem will be a single variable problem ok. So, let me explain this thing. So, I will solve one problem, then I will explain how this algorithm will works.

(Refer Slide Time: 17:28)



Let us take this particular problem. So, the function is f of x, y equal to x square plus y square plus 4 ok. The solution is that x equal to 0 y equal to 0. So, this is the optimal solution of this particular problem. So, this is a very simple function and I would like to find out the optimal solution of this particular function. So, what I will do? I will take a initial point you can say X naught and I need a direction.

So, in this case I will apply the Univariate method; that means, the first direction will be along x direction and second will be along y direction ok. Now what is x direction? That you can say suppose this is d x and this is 1 0; that means, there is no sense in the y direction and what is d y? d y direction and this is 0 1 direction. So, basically I will apply these 2 directions. So, initially I will go along x direction then y direction ok.

(Refer Slide Time: 18:41)

Multivariable problem

Univariate method

A multivariable problem can be converted to a single variable problem using the following equation $x^{t+1} = x^t + \alpha d^t$

Eq. No. 1 $f(x, y) = (x^2 + y^2) + 4$

$X^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $X^1 = X^0 + \alpha d$

$X^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 + \alpha \\ 1 \end{pmatrix}$

Putting in equation (1)

$f(\alpha) = ((1 + \alpha)^2 + 1^2) + 4$

Taking first derivative

$f'(\alpha) = 2(1 + \alpha) = 0$

$\alpha^* = -1$

$X^1 = \begin{pmatrix} 1 + \alpha \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$X^2 = X^1 + \alpha d$

$X^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 + \alpha \end{pmatrix}$

Putting in equation (1)

$f(\alpha) = (0 + (1 + \alpha)^2) + 4$

Taking first derivative

$f'(\alpha) = 2(1 + \alpha) = 0$

$\alpha^* = -1$

$X^2 = \begin{pmatrix} 0 \\ 1 + \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

So, let me solve this particular problem. So, as I said that I can write x t plus 1 which is equal to x t plus alpha d t ok time. Now, in this case the objective function is f of x, y equal to x square plus y square plus 4 ok.

Now, I have taken X naught equal to 1 1; that means, I would like to start my search from 1 1 ok from this particular point I would like to start the search process ok and direction initial is

1 0; that means, you are going along X direction and as I said that X_1 equal to X_0 plus αd ok.

So, if I put it that this is my X_0 and this is α direction is 1 0. So, I am getting 1 plus α ok. Now, if you are putting this f basically I am putting in X_1 ok. So, f of X_1 , then X_1 I can write in terms of α . So, my function will be $f(\alpha)$ now. So, this is 1 plus α^2 plus 1 ok; that means, X_1 is 1 plus α and y is 1 plus 4.

So, now, you just see. So, I have converted this particular problem to a single variable problem now my problem is in terms of α only ok. So, now, I can apply my region elimination technique to find out the optimal value of α and optimal value of α will give you the optimal solution along X direction ok.

So, optimal value of α will give you the optimal solution along X direction. So, here I will not apply the region elimination technique, but when once you are writing your algorithm or you are writing your program, you can apply the bracketed algorithm. So, you can bracket the optima and after that you apply region elimination technique or Newton Raphson method or other method to find out the optimal value of α , but as I am showing the hand calculation. So, what I will do? I will take the derivative; that means, I will apply the necessary condition. Necessary condition is that first derivative should be 0 ok.

So, if I apply the necessary condition then it will be twice 1 plus α equal to 0 and if I solve it then I am getting α^* , α^* equal to minus 1 and therefore, the x_1 ok that the optimal solution along X direction ok from the starting point is 1 plus α plus 1 now α equal to minus 1; that means, the optimal solution along X direction is 0 1. So, I am getting this particular solution and this is you can say this is the optimal solution or I can put a star here for optimal solution.

So, that X_1^* is 0 and 1. So, now, I am getting the optimal solution along X direction ok now what is the next step? That next step is you go along y direction ok. So, now, I would like to find out what is X_2 ; that means, I would like to find out the optimal solution along y direction now and in this case I will start my algorithm not from X_0 , but from X_1 ok

so; that means, this is your X_1 I am going to X_1 that is X direction now you go along y direction and find out X_2 ok. So, this is the idea.

Now, my starting point is X_1 . So, now, if I put it that X_2 equal to the starting point is X_1 and that is 0.1 plus α now I am going along y direction. So, y direction means 0.1 is not it? So, y direction is 0.1 and therefore, X_2 you are getting 0.1 plus α ok. So, now, if I put in equation 1 that is I can put f of X_2 and that is in terms of α I am I can write that is f α and which is equal to this is X equal to 0 and y equal to 1 plus α^2 plus 4 .

Then to find out the optimal value of α ok. So, I am taking again derivative first derivative this is the necessary condition and if I solve it I am getting α^* equal to minus 1 and if I put the value of α^* in X_2 . So, I am getting 0.0 and this is the optimal solution of this particular problem ok.

So, now, this function was a very simple function and in the second iteration itself I got the optimal solution, but when you are writing your algorithm. So, you have to check every suppose iteration you have to check whether you are getting the optimal solution or not.

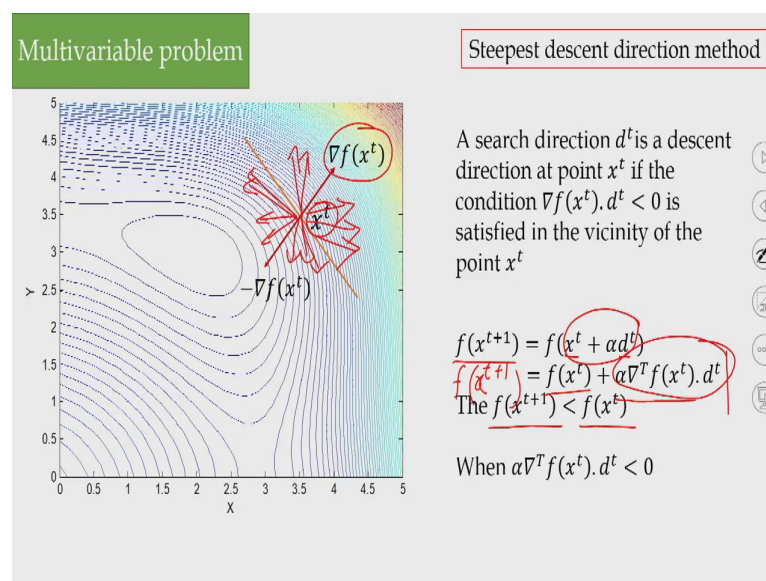
So, how you will check? So, you can take the derivative value if derivative value is sufficiently small ok it will not be 0 , but it will be smaller I suppose 10 to the power minus 3 minus 6 ok minus 12 . So, whatever precision you need. So, you define that one if derivative value is sufficiently small. So, in that case you will stop your iteration. So, in this case. So, I can put this one in the first derivative of the objective function and I can check whether derivative is 0 or near to 0 or not.

In this case I got the solution, but if I am not getting the solution, then I have I can find out X_3 . Now X_3 will be X_2 plus α now in this case I will go along X direction ok. See after getting X_3 . So, you can calculate X_4 ok if you have not reach the optimal solution or if you are yet to get the optimal solution then you have to repeat this then this will be X_3 plus α . Now you go along y direction that way I will repeat ok repeat the iteration.

So, you just see that every iteration you are doing you are solving a single variable optimization problem; that means, you are doing a line search along a particular direction. So, your problem is a single variable problem. So, now, this particular problem we have solved using Univariate method; that means, at a time we are going along one variable ok. So, in this case we have two variables. So, that is why initially we are going along X direction and then y direction. So, that way we are repeating our iterations ok.

So, X direction then y direction then again X direction then y direction. So, you have to repeat this process until you are not getting the optimal solution and as I have mentioned how you will check? So, you can calculate the derivative after each iteration and just check whether the derivative value is sufficiently small or not ok. Once it is small in that case you can stop your iteration.

(Refer Slide Time: 26:31)



Now, let us discuss another method that is Steepest descent direction method ok. So, in this method. So, what we are doing? Suppose you are here ok. So, you are at x^t then you calculate the derivative value ok. So, you calculate the first derivative value ok ∇f at x^t . Now, this method is that you go along the Steepest descent direction what is descent direction? The descent direction is that.

So, let me explain a search direction d^t is a descent direction at a point x^t if the condition ∇f at x^t dot d^t is less than 0 is satisfied in the vicinity of the point so; that means, what you are doing? Suppose, if I go along this direction, if I go along this direction, if I go along this direction, if I go along this directions my function value will reduce ok.

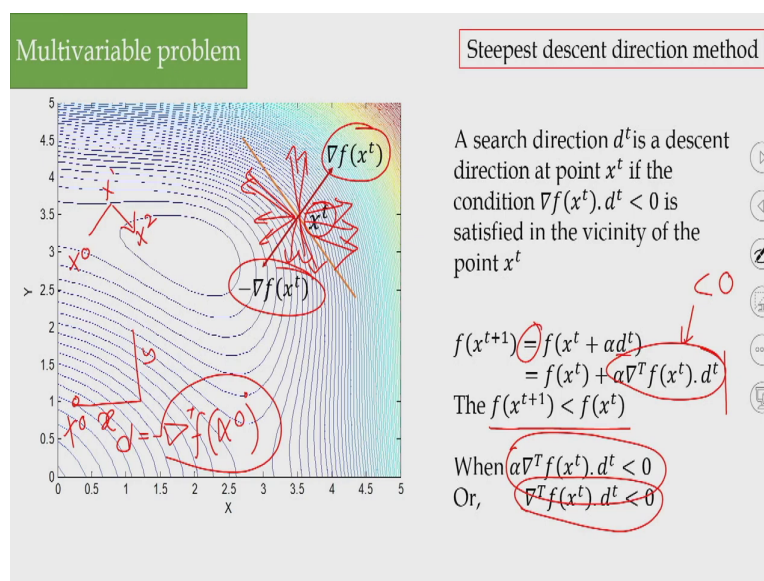
So, my function value will reduce therefore, if you are going along that direction. So, if you take any direction in this side your function value will reduce and you can say that direction is a descent direction. Now, if I go along this direction your function value will not decrease your function value will increase therefore, that particular directions are not descent direction. So, I can define a search direction d^t is descent direction at a point x^t , if the condition ∇f at x^t dot d^t less than 0 is satisfied in the vicinity of the point x^t .

I can prove that one. So, I can write f of $x^t + 1$; that means, it is another point and if I go along that direction d^t from x^t and I will get the value of f of x^t ok. So, I can write f of $x^t + 1$ which is equal to f of $x^t + \alpha d^t$ ok. So, I am going along I am going along direction d^t from x^t . Now, if I apply Taylor series and if I take up to the first derivative then I can approximate f of $x^t + 1$, f of $x^t + 1$ I can approximate it like that which is equal to f of f at $x^t + \alpha \nabla f$ at x^t dot d^t ok.

Now, my direction will be descent direction if f of $x^t + 1$ is less than f of x^t is not it? So, the direction because you are finding this $x^t + 1$. So, $x^t + 1$ is in the direction d^t . Now, that f of $x^t + 1$ will be less than f of x^t if this particular value is less than 0 is not it? This particular value is less than 0.

So, let me explain it again let me remove this all if the direction d^t is a descent direction in that case what will happen? f of x^t plus 1 will be less than f of x^t . Now, when it will be because this is a equality sign, if this portion that is $\alpha \nabla f(x^t) \cdot d^t$ if it is less than 0 then the direction will be the descent direction ok. So, if it is less than 0 then in that case f of x^t plus 1 will be less than f of x^t ok.

(Refer Slide Time: 30:43)



So, therefore, this particular portion should be less than 0. So, I can say that if because alpha is constant. So, if this particular value; that means, the first derivative at $x^t \cdot d^t$ is less than 0. So, in that case the direction whatever direction you are going you are going along d^t direction, that direction will be a descent direction.

So, now in the Steepest descent direction method. So, what we are doing? We are going along the negative of the gradient and negative of the gradient will give you the Steepest descent

direction; that means, if you are going along that direction that reduction in function value will be maximum ok.

So, therefore, in the Steepest descent direction method we are going along the negative of the gradient ok. So, this method is very simple. So, in the Univariate method. So, what we have done? So, we have gone along x direction then y direction along x direction then y direction. But in this case what you will do? Suppose this is your starting point and this is your X_{naught} and now what you will do? You will calculate the derivative ok.

So, you will calculate the derivative at X_{naught} ok. So, derivative at X_{naught} . So, that you will calculate and now d will be now d will be negative of this gradient is it clear? So, what we are doing? This method is again simple method we are calculating the direction using the gradient information at that particular point. So, we will calculate the gradient and we will go along the negative of that gradient ok.

So, at that derivative. So, we will go negative of this gradient. So, therefore, this method we can apply iteratively suppose you are here ok. So, you are here this is your X_{naught} you are calculating the gradient then you are going along this direction. So, you are getting X_1 , then you calculate the gradient go to the negative of the gradient then you find out X_2 something like that. So, I can change my direction and finally, I will get the optimal solution after few iterations ok.

(Refer Slide Time: 33:14)

Multivariable problem

Steepest descent direction method

Eq. 1 $f(x, y) = (x^2 + y^2) + 4$

$X^1 = X^0 + \alpha d$ $X^2 = X^1 + \alpha d$

$X^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $d = -\nabla f(X^0)$ $d = -\nabla f(X^1)$

$\nabla f(X^0) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $X^3 = X^2 + \alpha d$

$d = -\nabla f(X^0) = \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

$X^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 1-2\alpha \\ 1-2\alpha \end{pmatrix}$

Putting in equation (1)

$f(\alpha) = ((1-2\alpha)^2 + (1-2\alpha)^2) + 4$

Taking first derivative

$f'(\alpha) = 0$

$\alpha^* = \frac{1}{2}$

$X^1 = \begin{pmatrix} 1-2\alpha \\ 1-2\alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

OPTIMAL SOLUTION

So, let me solve the same problem. So, whatever we have discussed earlier that is f of x, y equal to x square plus y square plus 4 and this is I would like to find out. So, X naught is there. So, I would like to find out what is X 1 and here it is αd , but X naught is like earlier case. So, it is 1 1 and in this case direction is given by ∇f at X naught ok. So, this is the direction and if I calculate gradient at X naught. So, I am getting 2 2. So, either you can write 2 2 anyway. So, this is the direction will be negative of the gradient.

So, you can write minus 2 2 or I can also take 1 1. So, I can divide this by 2 1 1 basically you will go along the same direction. So, you can also take minus 1 minus 1 it does not matter. So, now, I am getting the direction. So, direction here is minus 1 minus 1 or minus 2 minus 2 or you can write minus 3 minus 3 also you are you will go along the same direction. So, whether you are writing minus 1 minus 1 minus 2 minus 2 minus 3 minus 3 same.

Let us solve this problem. So, now, X_1 equal to X_0 plus αd . So, d in this case minus 2 minus 2. So, I am getting $1 - 2\alpha$ $1 - 2\alpha$; that means, X equal to $1 - 2\alpha$ Y equal to $1 - 2\alpha$ put in equation 1 and so, you are getting this is X^2 this is Y^2 plus 4. Now this equation is in terms of α . Now you take the derivative ok first derivative you apply necessary condition and you can find out the optimal value of α .

So, in this case optimal value of α is half ok. So, therefore, X_1 is equal to $1 - 2\alpha$, α is half. So, you are getting 0 0. So, in this case this is the optimal solution of this problem you can see that in case of Univariate method there were two iterations ok.

So, in the X direction Y direction then we got the optimal solution, but when you are applying the Steepest descent direction method in a single iteration you are getting your optimal solution, but if you are not getting your optimal solution then what you will do basically? So, you will find you will go for the next iteration and next iteration is X_1 plus α .

So, now αd at X_1 ok direction at X_1 . So, what you will do? d will be equal to negative of the gradient at X_1 is not it? So, this is that similarly. So, I can repeat this iteration. So, X_3 equal to X_2 plus ok. So, αd now you will calculate the direction at X_2 . So, that way I can perform the iterations and I will stop this iteration and when the gradient value is sufficiently small ok. So, its near to 0 though.

So, I can check that one, again I would like to clarify here that I have taken the first derivative, but in actual case you will apply the region elimination technique or other method single variable optimization method for finding the value of α . So, this is the optimal solution of this particular problem.

(Refer Slide Time: 37:08)

Multivariable problem

Newton's method

Newton's method

Taylor series $f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \dots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(x) + \frac{h^n}{n!}f^{(n)}(x) + \dots$

$x_{i+1} = x_i + h$
 $h = x_{i+1} - x_i$

$f(x_{i+1}) = f(x_i) + (x_{i+1} - x_i)f'(x_i) + \frac{(x_{i+1} - x_i)^2}{2!}f''(x_i) + \dots$

By setting derivative of the equation to zero for minimization of $f(x)$, we have

$f'(x_{i+1}) = 0 + f'(x_i) + f''(x_i)(x_{i+1} - x_i) = 0$

$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$

Now the next method is Newton's method. Already we have derived this equation from the derivative concept ok. So, we have taken the derivative at x_i and then we basically obtained this relation that is x_{i+1} equal to x_i minus derivative at x_i first derivative at x_i divided by second derivative at x_i . Now I can apply this is your Newton's method. So, how we are getting this one? So, I can also derive this equation from Taylor series.

So, I can write the Taylor series f of x plus h ok. So, I would like to approximate the function value at x plus h knowing the function value at x . So, I can write f of x plus h f' dash x plus h square by factorial 2, f'' double dash x plus.

So, I can continue this is up to suppose f h to the power n divided by factorial n $f^{(n)}$ x plus. So, this is an infinite series. So, in this case I will apply I will use the truncated Taylor series; that

means, up to the second derivative. So, now, if I write that $x_i + 1$ which is equal to x_i plus h ok.

So, in that case I can write this is f of $x_i + 1$ which is equal to f of x_i and now what is h ? From here from here I can write that h equal to $x_i + 1$ minus x_i ok. So, I can write it h equal to $x_i + 1$ minus x_i first derivative at x_i and h^2 that is $x_i + 1$ minus x_i whole square divided by factorial 2 second derivative at x_i . Now, I would like to find out the optimal solution. So, if I apply the necessary condition; that means, I will take the derivative, I will calculate the derivative and equate to 0 for minimization of $f(x)$.

So, if I take the derivative of this particular function. So, I am getting $f'(x_i + 1)$ which is equal to 0 plus $f'(x_i)$ plus $f''(x_i)h$ plus 1 minus x_i . So, this must be equal to 0 ok. So, to find out the minima of this particular function. So, now, if I rearrange this, then I am getting this equation that is $x_i + 1$ equal to x_i minus derivative at x_i divided by second derivative at x_i first derivative at x_i divided by second derivative at x_i .

So, this equation is already we have applied and this equation is known or this method is known as Newton's method. So, you can find out the optimal solution iteratively. So, I can apply in case of a single variable function I can apply this equation and after some iteration I will get the optimal solution of this particular problem. Now in this method what you have to do you have to calculate the derivative and as well as second derivative first derivative as well as second derivative.

(Refer Slide Time: 40:50)

Multivariable problem

Newton's method

Quasi Newton's method

Sometime it is not possible to have the closed form expression of the function or it may be difficult to calculate the derivatives of the objective function. In such a scenarios, the derivative of the function can be calculated

$$f'(x_i) = \frac{f(x_i + \Delta x) - f(x_i - \Delta x)}{2\Delta x}$$

$$f''(x_i) = \frac{f(x_i + \Delta x) - 2f(x_i) + f(x_i - \Delta x))}{\Delta x^2}$$

① $f(x + \Delta x) = f(x) + \Delta x f'(x) + \frac{\Delta x^2}{2!} f''(x) + \dots$

② $f(x - \Delta x) = f(x) - \Delta x f'(x) + \frac{\Delta x^2}{2!} f''(x) - \dots$

① - ② $\Rightarrow f(x + \Delta x) - f(x - \Delta x) = 2\Delta x f'(x)$

$$x_{i+1} = x_i - \frac{\Delta x [f(x_i + \Delta x) - f(x_i - \Delta x)]}{2[f(x_i + \Delta x) - 2f(x_i) + f(x_i - \Delta x)]}$$

$\Rightarrow f'(x) = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}$

Convergence $|f'(x_i)| \leq \epsilon$

So, what we do basically? The first derivative I can or second derivative I can calculate using numerical method ok. So, what we do means sometime it is not possible to have the closed form expression of the function or it may be difficult to calculate the derivative of the objective function sometime it may happen.

So, it is difficult to calculate the derivative. So, in such a scenario, the derivative of the function can be calculated using this equation. So, what I can do? I can apply these equations to calculate the derivative and this equations has again derived from the Taylor series.

So, I can write the first derivative at x_i which is equal to $f(x_i + \Delta x) - f(x_i - \Delta x)$ divided by $2\Delta x$ ok. Similarly, second derivative I can calculate the

function value at $x_i + \Delta x$ minus twice f of x_i plus f of $x_i - \Delta x$ divided by Δx^2 .

Now if I apply these two equations in the Newton's method. So, I will get finally, this equation and here this method also you can say Quasi Newton's method again I can check the convergence. So, as I said I will calculate the derivative value and this derivative value is sufficiently small.

So, in that case I can stop my iteration and I can say that I got the optimal solution of this particular problem. So, let me explain how we got this particular first derivative one. So, as you know that I can write the Taylor series that is f of $x + \Delta x$ which is equal to f of x plus Δx first derivative at x plus $\frac{1}{2}$ by factorial 2, second derivative at x plus ok.

So, I am not considering the higher order derivative. Similarly, I can also write f of $x - \Delta x$ ok. So, which is equal to f of x minus Δx first derivative at x sorry this is Δx square plus Δx square by factorial 2 second derivative at x ok. Now, if I suppose if I say that this is equation 1 and this is equation 2 and if I write 1 minus 2 ok.

So, if I write 1 minus 2 what I will get? f of $x + \Delta x$ minus f of $x - \Delta x$ 1 minus 2, then f of x and f of x will cancel out and this second term will be positive. So, what I will get? Equal to twice Δx f' of x ok then the third term will again cancel out. So, anyway I am not considering the higher order derivative. So, from here I can write it that the first derivative at x equal to f of $x + \Delta x$ minus f of $x - \Delta x$ divided by twice Δx ok.

So, I got this first derivative ok. Now, similarly what I can do? I can add these two equations. So, if I add these two equations. So, now, I am getting the first derivative. So, this is your first derivative, now I can also add these two equations; that means, 1 plus 2. So, let us see what you will get if you add it. So, I would like to rub this portion.

(Refer Slide Time: 45:43)

Multivariable problem

Newton's method

Quasi Newton's method

Sometime it is not possible to have the closed form expression of the function or it may be difficult to calculate the derivatives of the objective function. In such a scenarios, the derivative of the function can be calculated

$$f'(x_i) = \frac{f(x_i + \Delta x) - f(x_i - \Delta x)}{2\Delta x}$$

$$f''(x_i) = \frac{f(x_i + \Delta x) - 2f(x_i) + f(x_i - \Delta x))}{\Delta x^2}$$

$$x_{i+1} = x_i - \frac{\Delta x [f(x_i + \Delta x) - f(x_i - \Delta x)]}{2[f(x_i + \Delta x) - 2f(x_i) + f(x_i - \Delta x)]}$$

Convergence $|f'(x_i)| \leq \epsilon$

① $f(x + \Delta x) = f(x) + \Delta x f'(x) + \frac{\Delta x^2}{2!} f''(x) + \dots$

② $f(x - \Delta x) = f(x) - \Delta x f'(x) + \frac{\Delta x^2}{2!} f''(x) - \dots$

① + ② $f(x + \Delta x) + f(x - \Delta x) = 2f(x) + \Delta x^2 f''(x)$

$\therefore f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x))}{\Delta x^2}$

So, in this case what I am doing? I will add it I will add I will add these two. So, if you add it. So, you are getting f of x plus Δx plus f of x minus Δx ok. So, this is the left hand side. Right hand side is that twice f of x ok this is twice second term will cancel out plus minus and third term will be this is half and this is half. So, you are getting that Δx square second derivative is not it?

And the fourth one will cancel out anyway. So, we are not considering. So, from here I can write the second derivative. So, second derivative is equal to from here you can write that f of x plus Δx minus twice f of x plus this term that is f of x minus Δx divided by Δx square ok.

So, I am getting this this equation and this is the second derivative equation second derivative equation. So, if I put it finally, our Newton's method will be this one. So, I can apply this

method iteratively to get the optimal solution of the problem single variable problem. Let us stop here. So, in the next class we will discuss the Newton's method for solving multivariable optimization problem.

Thank you.