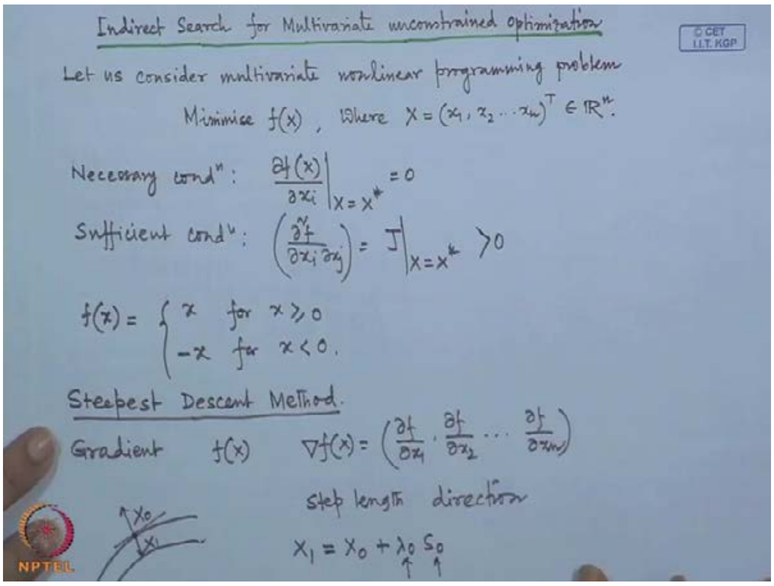
Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology Kharagpur

Lecture - 31
Unconstrained optimization techniques: Indirect Search Method

Today's topic is indirect search method for multivariate optimization problems. Now, let us consider a multivariate non-linear programming problem, and we are considering the unconstraint optimization problem that is why the problem can be stated as minimize effects where, x equal to X equal to x 1, x 2, x n.

(Refer Slide Time: 00:44)



And this is the, these are the points in the nth dimensional space. Now, we need to solve this problem I discussed about the direct search method. Today my discussion will be on indirect search method. Now, unconstraint optimization problems are very much needed for some complex situation in engineering discipline. And also the unconstraint optimization problems are also very much important, because important for the constraint optimization problem. Because in the constraint optimization problems sometimes it is needed that we need to analysis the behavior of the objective function. In that case, the unconstraint optimization technique can give some idea about the movement of the objective function; that is why in general the unconstraint optimization problems are important in decision science.

Now, if this is the problem for us minimization of effects. And there is no only the may be the for the decision variable the bounds may be given to us. Now, in this situation now we need to just solve it. Generally, as we know for locating minimum as we have learned that the necessary and sufficient conditions we can use for it. As we know the necessary condition would be Del of f (X) divided by Del X I; and as we know this values would be equal to 0 if X star is the optimal solution.

And, through this necessary condition we are finding out the stationary points where the optimum may lie; after that we go for the sufficient condition that is the second order derivative we need to do; that is Del 2 f by Del x i Del x j and we form the Jacobean for it. And we just see at the optimal point if the Jacobin value is positive definite if this matrix is positive definite. Then we say that the corresponding optimal solution X star is the minimum solution for the given unconstraint optimization problem.

But the situation is not so simple always. For example, we are having one function f (X) where, f (X) is equal to x for x greater than equal to 0 and f (X) equal to minus x for x lesser than 0. For this problem as we see the function is not differentiable at the minimum point, because; minimum of this function lies at the point x equal to 0. And at this point the function is not differentiable; that is why, if we apply the necessary and sufficient condition in this situation, it will not work.

That is why in general, we are having different techniques for solving unconstraint optimization problems using the differential using the derivatives without using the derivatives. That is why the whole method we can categorize into 2 parts; one is the direct search algorithms which are non gradient methods where we do not use the derivatives or in direct search method there that is the gradient methods. And there we are using the derivatives of the objective functional for the unconstraint optimization problem.

Now, in the direct search method there is very popular method is there; that is the steepest ascent or steepest decent method; this method is the very simple in understanding. And since this is the gradient method it uses the gradient of the function at each and every this is the iterative process. That is why each and every iteration this gradient in method, indirect search method rather in specific steepest decent or ascent method uses the gradient of the function at different points.

Now, this steepest decent ascent method; this is the generally, the direction method since we do not use the derivatives that is why these are all zeroth order of algorithm. And but; the indirect search methods are either first order or second order algorithm depending on involvement of the first order or second order derivative in the process. Now, let me start with the simplest method; that is the steepest decent method. And that is for applicable for the minimization problem. Later on I will move to the more popular method, conjugate gradient method, filches reviews method etcetera later on; that is why let me start my discussion with the steepest decent method. That is applicable for the unconstraint optimization problem in specific the minimization.

Let me just tell you the basic principle of the steepest descent method. As I said this is the gradient method, this is the indirect search method; it uses the concept of gradient. As we know in vector analysis gradient of scalar function and we know gradient of the scalar function f (X) is the function for us. And if x is n dimensional point then we can define the gradient of f (X) as Del f partial derivative of f with respect to x 1, partial derivative of f with respect to x 2 and in this way get a tappel. And that is corresponds to gradient of the function, scalar function f. Now, as we know the gradient direction; if this is the surface of the function, gradient direction at any point would be the normal to this point, because; this is the vector.

Every vector is a having the magnitude and the direction; that is why the gardening magnitude can be very easily, can be we can get through the norm of this vector. But the direction of this gradient would be always in the, at any point. If I just draw a tangent then the gradient tangent plane in the surface then the gradient would be in the normal direction. And we know gradient is the direction in which the function increases quickly; that is why if we just move from one point to the other point to the gradient direction. And since this the optimization problem we are always excepting the improvement of the functional value; that is why if we just move through the gradient direction very quirkily we can reduce to the optimal solution.

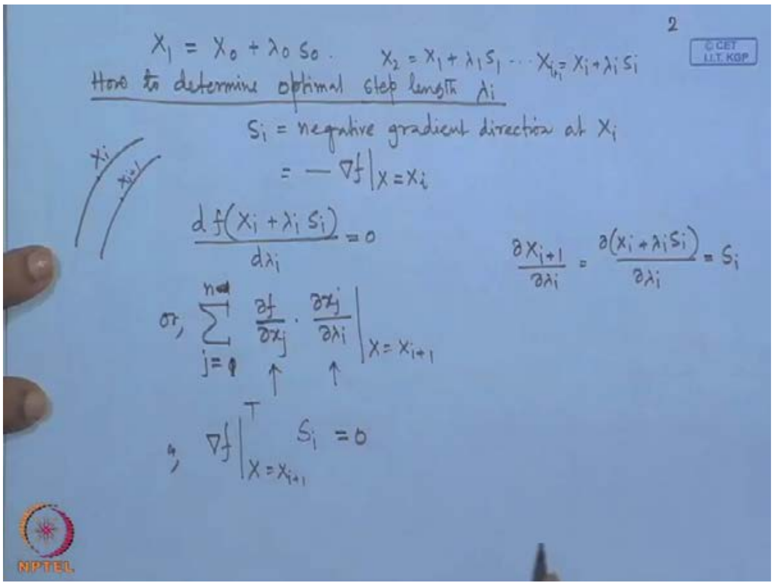
But since the problem is minimization problems, problem that is why if we just move in the gradient direction then the functional value instead of decreasing it will increase further; that is why what we do in the maximization problem. We use the gradient direction as far the movement of the respective iteration process. And but in the minimization problem we just consider the negative direction of the gradient as the best direction for improvement of the functional value; that is why; this is the direction for the maximization problem. And we will just move through the negative direction in the minimization problem, at this point we will see the functional value. We will move to the next, through this direction in the negative direction in the minimization problem.

And, with the some step length we will reach to the other, another point and this is the another level surface of the function, scalar function. And in this way we will just try to reach to the optimal solution. Now, in the process it is very clear, if we just discuss the minimization of the problem, it is very clear that if we just move in the negative to the gradient direction. The next question comes; how far should we walk; shall we walk to the optimal solution directly? That means, is it that within 1 iteration we will reach to the optimal solution that is not so really; that is why there is concept of step length.

There is a concept direction in this in direct search method and we will discuss; we will consider the corresponding step length and corresponding direction in each and every iteration. That is why my next discussion would be specific the steepest descent method. Where, we will find out the optimal step length and the direction has already been fixed for the steepest descent method. Since, this is the minimization problem; we are using the steepest descent method, we will move through the negative gradient direction. And we will use the steepest ascent method for maximization problem. And we will move through the positive direction of the gradient; that is why let me start the steepest descent process. We will start from guess point. This is the guess point for us, this is x naught.

And, from X naught we will move to the point that is X 1, what should be the point x 1? X 1 would be X naught plus lambda naught S naught. That means, in the first iteration; we are considering lambda naught as the step length; that means, with this step we will move in the direction of S naught. And that is fixed here in the negative direction. That is why; let me just in general, let me just give you the idea of the steepest descent method first. Then I will come to the algorithm of the steepest descent method the next.

(Refer Slide Time: 11:10)

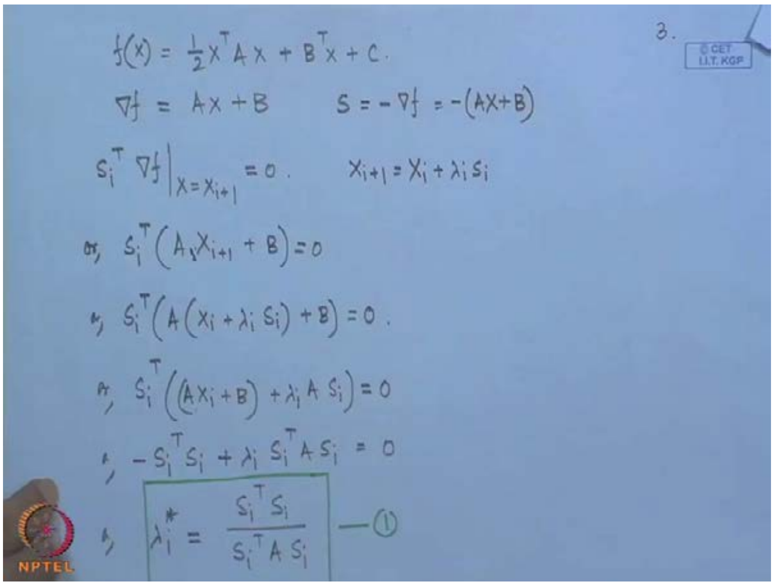


Now, what is the task for us? Let me just write it down X 1 is equal to X naught plus lambda naught S naught. Now, my question is that how to find optimal step length lambda naught? Now, let me consider the ith step; that means, X 1 is moving to X 2, X 2 is moving to X 3 in this way we are proceeding further and further. X 2 would be X 1 plus lambda 1 S 1 in this way we are moving through to X I; X i would be X i plus 1 would be is equal to X i plus lambda i S i. Now, we are giving the general procedure to find out the optimal step length lambda i at ith step. Now, this is the position of the point X i. Now, we want to move to the next X i plus 1, we know S i is the negative gradient direction at X I; that is why S i would be minus grad f at X is equal to X i. Now, for finding out lambda i we will find out the functional value at point X i plus lambda i S i.

And, we will consider that point as that value of lambda as optimal where, it is equal to 0. Actually, we are finding out that lambda i which gives the minimum value for function at the point X i plus 1. That is why; if we just write in this way, what we get in the next? We get since; there are n number of this is point in n dimension. That is why; the we can just write in this way summation j is equal to 0 to n minus 1 del f. Let me start with 1 to n del f by del x j into del x j by del lambda i. And we will find out this value at point X is equal to X i plus 1. Because we have consider the function at point X i plus 1, this is X i plus 1.

Now, if we just consider Del (X i) plus 1 divided by lambda i, what is the value for this? This would be; Del (X i) plus lambda i S i Del of lambda I; that is why; we will get this is equal to S i. If we just use this fact here, there are 2 terms; one term is the gradient of f and this is another term that is S i. Because just now we got this value we are considering at point X i plus 1; that is why we have considered the partial derivative with respect to lambda i at point X i plus 1. If this is so from here directly we can write that grad (f) at point X is equal to X i plus 1 T S i must be is equal to 0 alright.

(Refer Slide Time: 15:26)



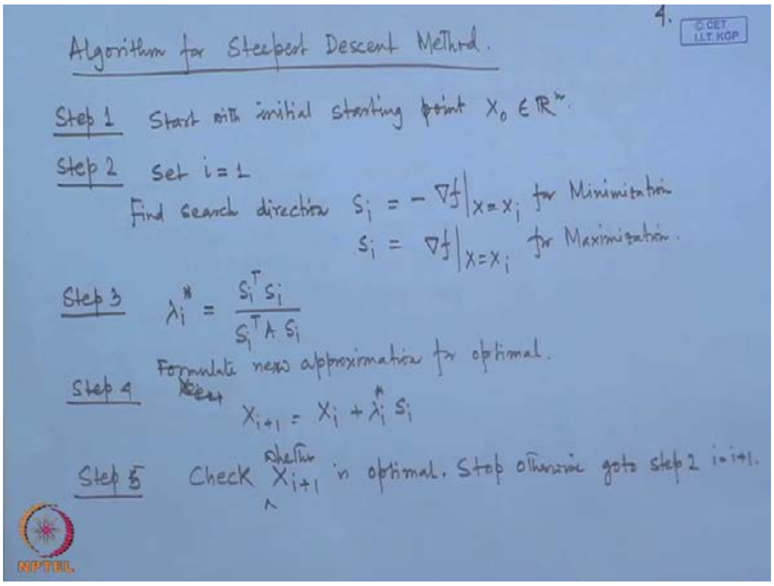
Now, let us consider a function; a quadratic function let me consider in this fashion; f (X) is equal to half X T A X plus B T X plus C. We may consider A as the positive definite and the symmetric matrix in this quadratic function. Then if this is so then the function we will convex. And if the function is convex then whatever optimal solution through the steepest descent method we will that would be the global optimal. Because as we know through the Kuhn tucker condition through the convex programming problem; that any convex objective function if we just find out the optimal solution that would be the corresponding global optimal instead of the local optimal.

Now, if this is so what is grad f? Grad f would be is equal to A X plus B. Now, if this is so then whatever function we got previously that is this one; we just take the transpose of it. Then we get a S i T grad f X is equal to X i plus 1 equal to 0. Now, grad f at X is equal to X i plus 1 means; in place of X we will put X i plus 1; that is why let me write it down, S i T grad f as X sorry not here, X i plus 1 plus B is equal to 0. Now, what is X i plus 1? Just now, we have same X i plus 1 is the we are considering the of the after the ith step; that is why this is equal to X i lambda i S i. What is my intension? My objective is to find out the optimal value of lambda i, which minimizes the function f at point X i plus 1. And that is why we have taken the first order derivative with respect to lambda i.

Now, let me just put the value of S i here itself and we are getting A X i plus lambda i S i plus B is equal to 0 or if we just further simplify then we are getting S i T A X i plus B this is one part. And another part would be lambda i A S i equal to 0 and this value is equal to grad (f) at point X I; that is why we that is rather grad of f is equal to X plus B. And we are considering the steepest descent method, the direction that is the negative direction. That is why S can be considered as minus of grad f, that is why this is equal to minus of A X plus B. Because; we need this S value is in this expression. Then here we are getting that S i T, this is the grad this is the S value at point X i with the minus sign. That is why we can put in this way, plus lambda i is the constant and S i T A S i equal to 0.

And, what we get from here? We get from here, lambda i is equal to S i T S divided by S i T A S i. That is why what we are getting here, we are getting that is the optimal value of lambda. If we consider the negative gradient direction at point X i that is my S i, if I just consider S i T S i divided by S i transpose A S i this value, this is a constant value A is the matrix of the given size. This if it is a quadratic 2 by 2 matrix would be there then that would be constant. This is the constant and this ratio can be consider at the ith level at the optimal step length. We will use this fact in the steepest descent method, we will use this lambda i star value in the steepest descent method. Let me consider this as one, that is why we can write down the algorithm in the next for the steepest descent method here it is. Let me use the other page.

(Refer Slide Time: 20:50)

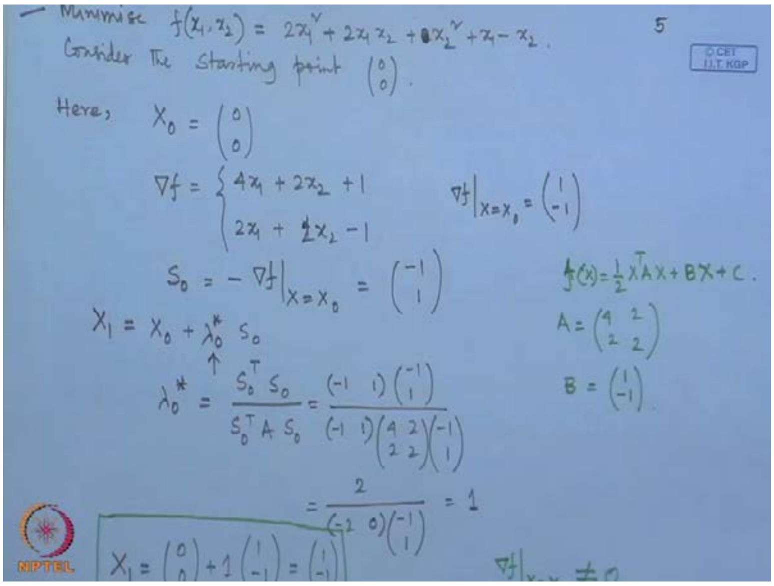


Let me write down the algorithm for the steepest descent method. Step 1 would be start with a guess point X naught the, this is the iterative process; that is why as I said we will start from X naught move to X 1 with the step length lambda 1 star. Then we will go to the X 2 with the step length lambda 2 star. And at each point we will consider the negative direction as S i, because this is the minimization problem that is why let me write down the first step as the start with the initial guess point. And that is more important if the starting point is far from the optimal point. Then the number of iterations would be more that is why starting point selection is very important for the iterative process as we know start with the initial starting point guess point X naught starting with this is in R n.

Then, we will move to the next step; we will find out the gradient of the function at point X naught. And we will consider X naught as the negative of the gradient direction that is why we will let me start the iteration we i is equal to 1. Then we will find the search direction at i is equal to 1 we will just increase the value of i in the respective case. And find the search direction at i as minus of grad f at point X is equal to X I; this is for the minimization problem. Let me write down the algorithm for the general non-linear unconstraint optimization. And grad f at X equal to X i this is for the maximization problem.

Now, once we are getting S I, our next task is to find out the next point from the X naught to X 1. That is why; we know S 1 we want the lambda 1; that is why; we will find out lambda 1, lambda i star is equal to from the equation 1; S i T S i divided by S i T A S i. Next step 4; find out the next point X i plus 1. Formulate new approximation for the optimal solution X i plus 1 is equal to X i plus lambda i star S i. Then step 5 check for optimality, there are certain procedures for checking optimality of the new point check whether, stop otherwise go to step 2 with the consideration i is equal to i plus 1. This is the whole algorithm for us. Let us use this algorithm for problem for solving unconstraint non-linear programming problem.

(Refer Slide Time: 25:23)



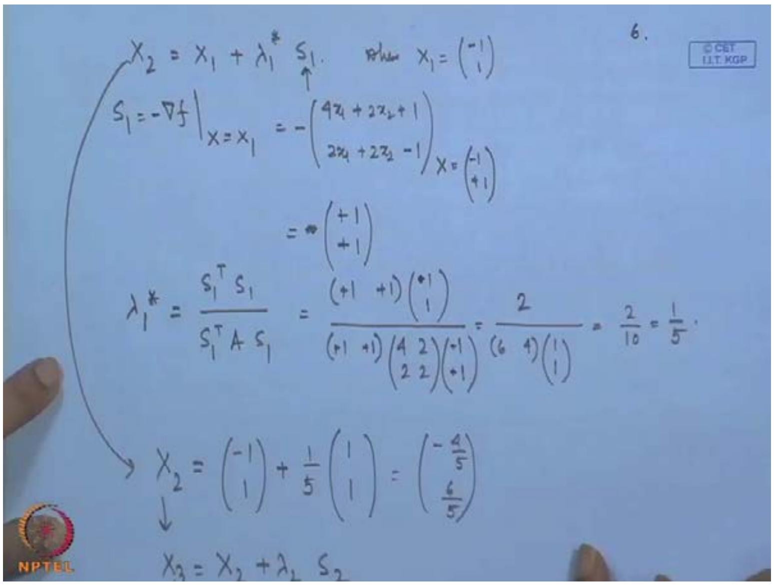
Let us consider one problem; minimization problem, minimize f (x 1, x 2) is equal to 2 x 1 square plus 2 x 1 x 2 plus 2 x 2 square plus x 1 minus x 2. Now, it is given that consider the starting point as 0, 0. We do not know the optimal lies that is why let me start with the most conventional point that is 0, 0 point. Then I am moving to step, next step. Here, certainly X 0 that is the initial starting point is 0, 0. What is grad f? That is a gradient value of this function; this is equal to 4 x 1 plus 2 x 2 plus 1. Because we are considering this one, the function the partial derivative of f with respect to x 1. What about with respect to x 2? It would be 2 x 1 plus 4 x 2. Sorry, we are not considering x 1 here, 2 x 2 square here and only we are considering x 2 square 2 x 2 minus 1. Now, then what would be S naught? S naught would be negative gradient direction at point X is equal to X naught. That is why; if this is so just put 0, 0 in this value. Then we will get S naught as grad f at X is equal to X naught would be equal to 1 and minus 1; that is why S naught would be negative to that minus 1, 1.

Now, we will move to the next; that is X 1 point from X 0. And with this consideration lambda 0 star S naught. How to find lambda 0 star? As we have just learn that the lambda 0 star would be is equal to S 0 T S 0 divided by S 0 T A S 0. Now, let me consider few things here, that is here A is equal to if we consider the function f (X) in the form of half of this is the quadratic function of half of X T A X B X plus C. If this is so then A would be is equal to 4, 2, 2, 2 this matrix is the symmetric matrix. And corresponding B would be is equal to 1 minus 1. Then we will use this fact here therefore, S 0 T would be is equal to (minus 1, 1) (minus 1, 1) divided by (minus 1, 1) 4 2, 2, 2 and S naught again. If we just get this value we are getting it as 1 plus 1 2. This is equal to minus 4, plus 2, minus 2, 0 and this is (minus 1, 1). That is why; we are getting the lambda naught star value as is equal to 1.

Once we are getting it, we are getting x 1. What is my x 1? x 1 would be (0, 0) plus 1 into 1 minus 1. That is why; this is equal to (1, minus 1) this is my next approximation. I do not know how many iterations I need to reach to the optimal solution; we can check whether, at this point the function is optimum or not. How to check it? Now, one check there are few criteria for checking optimality, but one of the most popular criteria is that we will just find out the grad f value at point X is equal to X 1. Now, grad f at point X is equal to X 1 if I just put this is 4 minus 2 and this the whole value is coming a 3 and something.

That means it is not equal to 0; if this is not equal to 0 then minus grad f will not be again 0. Then what the implication of this? It implies that there is a possibility for further improvement of the functional value; if we see at any point in the process of iteration the grad f function is coming 0. That means, we have to stop the iterations there itself, because there is no possibility to proceed further because functional value will not increase further if I just move.

(Refer Slide Time: 31:30)

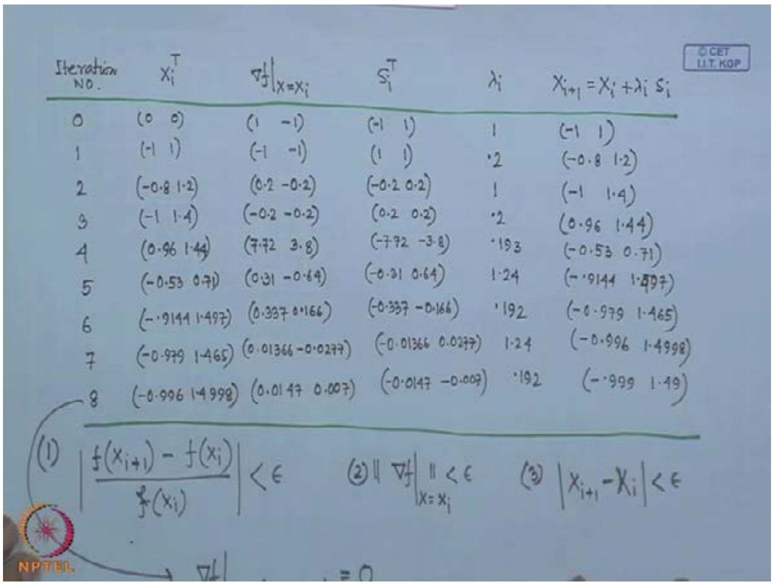


Now, X 1 we got that is why let us move to the next point that is my X 2. What is x 2? Again, X 2 would be is equal to X 1 plus lambda 1 star S 1. First a fall we need to find out S 1. What is S 1? S 1 is the grad f; at point minus of grad f at point X is equal to X 1. Now, just now we got the grad f value and that is 4 x 1 plus 2 x 2 plus 1 that is why 4 x 1 plus 2 x 2 plus 1 and here, it is coming 2 x 1 plus 2 x 2 minus 1. If I just move at the, if I just write down at X is equal to (1, minus 1) then this value will come as 4 minus 2; that is 2 plus 1. Now, we are considering X 1 as equal to X naught plus lambda 1 into S 1. I am sorry, X 1 is equal to lambda X naught plus lambda naught into S naught. Lambda naught is lambda naught star is coming one; what about S naught? S naught is coming minus 1 plus 1. That is why; it is coming minus 1 plus 1 that is my next S 1.

If I consider S 1 here, X 1 is equal to minus 1 plus 1. If this is so X equal to minus 1 plus 1. In the next point would be X 2 point, X 2 point would be is equal to x 1 plus lambda 1 star into S 1. Now, we need to find out the value of S 1 here. And for finding out the value of S 1 that is the negative gradient direction at point X equal to X 1. That is why; this is the gradient value at X equal to X 1 minus 1 plus 1; we get the value as minus 4 plus 3 that is minus 1. And this is again minus 1 then once we are getting X 1 what about lambda 1 star? Lambda 1 star would be S 1 T S 1 divided by S 1 T A S 1. If I just write down the value here, A value would be 4, 2, 2, 2 this we are getting from the objective function minus 1 minus 1.

One thing should be mentioned here, this is minus of grad f. That is way; this is minus plus 1, plus 1. If we get it that would be 2 here and here, it would be 6, 4, 1, 1. That is why; it is coming as 2 divided by 10 equal to 1 by 5. Once we are getting lambda 1 star then from here we can get X 2 as X 1. That is (minus 1, 1) lambda 1 star that is 1 by 5. And next is my S 1, S 1 we are getting it as, (plus 1, plus 1) and it is coming minus 4.5, 4 by 5 and 6 by 5. This is my X 2. What I will do next? Again, I will move to X 3, X 3 would be is equal to X 2 plus lambda 2 S 2. And again we will find out S 2 first. Once we get S 2 we will get lambda 2 in this way we will proceed.

(Refer Slide Time: 35:50)



Iteration No.	X_i^T	$\nabla f\|_{X=X_i}$	S_i^T	λ_i	$X_{i+1} = X_i + \lambda_i S_i$
0	(0 0)	(1 −1)	(−1 1)	1	(−1 1)
1	(−1 1)	(−1 −1)	(1 1)	·2	(−0.8 1.2)
2	(−0.8 1.2)	(0.2 −0.2)	(−0.2 0.2)	1	(−1 1.4)
3	(−1 1.4)	(−0.2 −0.2)	(0.2 0.2)	·2	(0.96 1.44)
4	(0.96 1.44)	(7.72 3.8)	(−7.72 −3.8)	·193	(−0.53 0.71)
5	(−0.53 0.71)	(0.31 −0.64)	(−0.31 0.64)	1.24	(−·9144 1.497)
6	(−·9144 1.497)	(0.337 0.166)	(−0.337 −0.166)	·192	(−0.979 1.465)
7	(−0.979 1.465)	(0.01366 −0.0277)	(−0.01366 0.0277)	1.24	(−0.996 1.4998)
8	(−0.996 1.4998)	(0.0147 0.007)	(−0.0147 −0.007)	·192	(−·999 1.49)

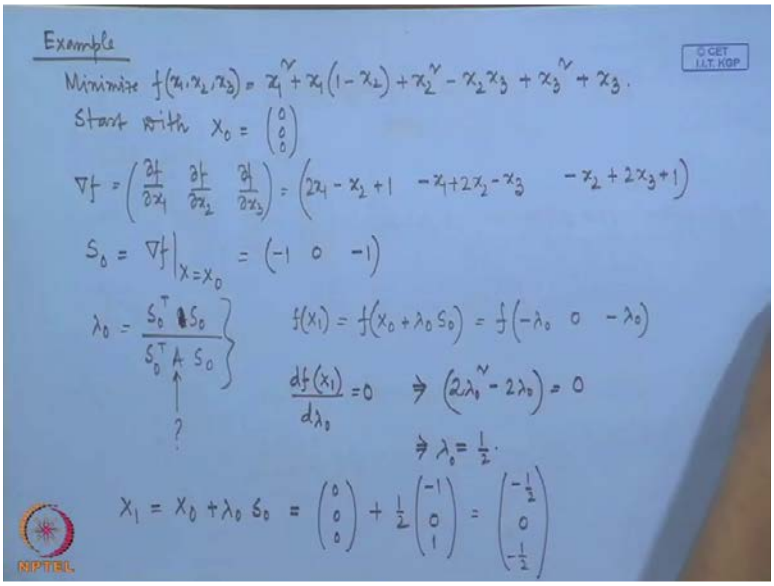
If I just write down the entire value together, then we see, we need 8 iterations to just reach to the optimal solution. Now, this is the first that I have just now solved, we are reaching to X 1 (minus 1, 1) that is my X 1. And these are all the gradient value that is negative gradient value corresponding lambda i. And the next point that is equal to X i plus lambda i into S i, this is my X 2. X 2 is coming minus 4 by 5 6 by 5. Just now, we got that is why; this minus 0.8 and 1.2. In this way, we are moving again, through the gradient, negative gradient corresponding lambda i. If we just proceed the whole thing, what are the things we can observe from this table that; if we just look at the X i values here, as we see that almost we are coming to 0.9 for the last 3 iterations.

In the first case, we are getting 0.9 minus 0.91, in the second minus 0.97, in the third we are getting the last one rather, last but one and previous to that minus 0.99. That is why;

we should have some process through which we can check that where we can stop our iteration. There are 3 alternative methods are there; we can adopt any one of that. We will see that the functional value in the last, in the step and the previous to that point. And if we just see the ratio is very small then we can stop our iteration this is one process. Second process is that we will just see the gradient of the function at point X is equal to last one say, X i. If we see the gradient value is very small, if it is 0 that is nice enough. We should stop the process, but; even it is very small, very near to close to 0 then also we can stop the process. Because we can see that there is no possibility for further improvement.

And, there is another third process is also there. That is we will see the X i value, if they are very close then also we can do. Actually, we are considering here the norm of delta f, because; delta f is a vector. If see the norm of that value is coming very small then only we stop. And here we can see that after the 8th step the functional value at point X is equal to, as we know at point minus 1 and 1.5 this gradient value is 0; that is why this is almost minus 1 and this is 1.5, thus we are stopping our iterative process. Now, this one 2 dimensional problem, we consider let me just move through the 3 dimensional problem.

(Refer Slide Time: 39:21)



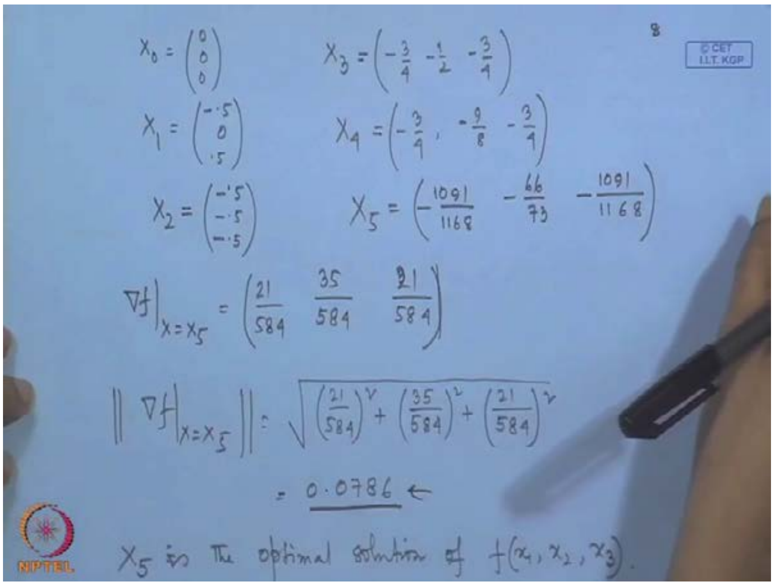
The same steepest descent method we will use there. The example we are considering; minimize f x 1 x 2 x 3 is equal to x 1 square plus x 1 into 1 minus x 2 plus x 2 square minus x 2 x 3 plus x 3 square plus x 3. Now, it is written that start with x naught is equal to 0, 0, 0. Let us apply the process here, as we see that if we use that process where, we have used the matrix A in the quadratic problem. But here to construct the matrix is very difficult. That is why; if we just adopt that process for lambda that lambda i star is equal to S i T S i T divide by S i T A S i. S i is very easy to get, but; getting A would be difficult here for 3 dimensional problem. That is why; I will show you how to handle this problem in a different way.

Now, let me start the process; what we want? We want that X naught should move to X 1 and while X 1 would be is equal X naught plus lambda naught S naught. That is why; let us first find out the grad of function f. This would be is equal to Del f by Del x 1, Del f by Del x 2 Del f by Del x 3. And this would be 2 x 1 minus x 2 this is, there is 1 x 1 plus 1. And del f by del x 2 that would be is equal to minus x 1 plus 2 x 2 minus x 3. Partially differentiating f with respect to x 2 then the next point would be with respect to x 3. It would be minus x 1 plus 2 x 3 plus 1 alright. No, there is a term; that is minus x 1 x 3 must be there then only it will work. x 1 square 2 x 1 minus x 2 plus 1 minus x 1 plus 2 x 2 minus x 3 minus x 1 plus 2 x 3 plus 1. Sorry, this is, this should not be here alright, and it should be minus x 1 only.

If this is the grad f function then let us move to S naught, S naught is equal to minus grad f at point X is equal to X naught. If this is so we are getting from here X naught is 0, 0, 0. That is way; it should be minus 1 0 and here it would be minus 1. Now, once we get S naught, as we learnt previously; we should get lambda naught is equal to S naught T S naught divided by S naught T A S naught. But here, finding out A would be very difficult that is way will not use this process. What we will do? We will find out the functional value at point X 1 that is f of X naught plus lambda naught S naught. That is equal to f of minus lambda naught 0 minus lambda naught. And we will find out that d f of that at point X naught must be is equal to 0.

If we just equate that was my basic principle for the steepest descent method. We have adopted that thing in our algorithm if this is so then; we are getting the value as equal to 2 lambda naught square minus 2 lambda naught equal to 0. And this gives me the value for lambda naught is equal to half. Once we get lambda naught equal to half then we will go to X 1. X 1 would be is equal to X naught plus lambda naught S naught that is 0, 0, 0, half. And S naught would be minus 1 0 minus 1. That is why; my point would be minus half 0 minus half this is my x 1.

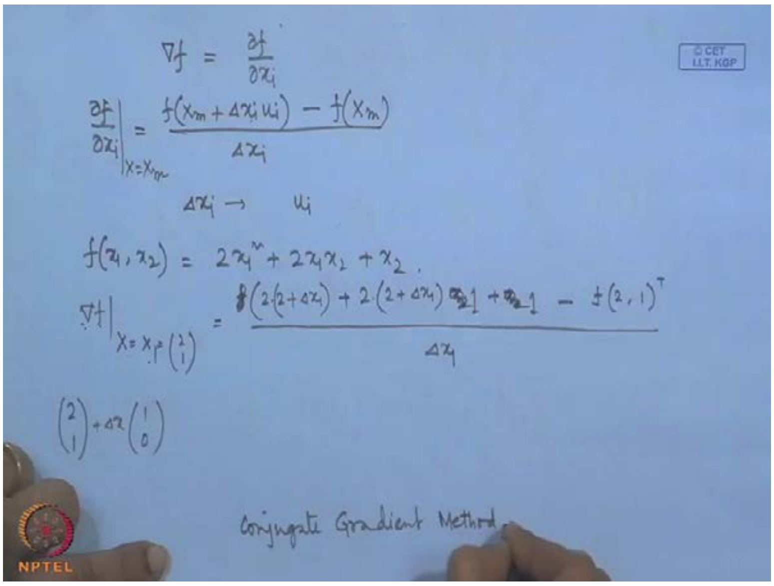
(Refer Slide Time: 45:53)



And, if I just apply the same technique in the next, what we get? We get the X 2 value. That is way; we started with the X naught that was 0 0 0 moving to X 1 that is minus 0.5 0 0.5. If we apply the same process here, we will get X 2 and that would be minus 0.5 minus 0.5 minus 0.5. And proceed further X 3 would be is equal to minus 3 by 4 minus half minus 3 by 4. X 4 would be is equal to minus 9 by 8 minus 3 by 4. And x 5 this is the last point of iteration, this value is coming minus 1091 divided by 1168 here minus 66 by 73 and here minus 1091 divided by 1168. And as I say that I will just find out the either the functional value difference in the functional value at point x 5 and x 4, we will compare it. If we see the change is very small or rate of change is very small we can stop out process.

But here, we would see that grad f at X equal to X 5 if we just calculate we are getting these values 21 point, 21 by 548 35 by 584 21 by 584. And norm of this value at X equal to X 5, norm is equal to root over this one. This value is coming 0.0786, we may consider as the small this is a very small value. That is why; we can stop out iteration at this point. And we can declare that X 5 is the optimal solution for this optimization problem. Now, that is all about the steepest descent method. But if we just discuss indirect search method in general and that there will see that steepest descent method is the most popular method, because; it is very easy to calculate the entire process. But there is another very popular method is there; that is called the conjugate gradient method. And we are moving to that point, that method in the next.

(Refer Slide Time: 49:37)



Now, one thing must be say that; if we see that, in the steepest descent method if we see that the gradient of the function we are calculating. Before going to the conjugate gradient, let me just discuss this thing also. Grad f we are considering as partial derivative, but it is not always so easy to find out the partial derivatives of the function. Because if the function is discontinues in that case; it is not easy to calculate the partial derivative of the function at point x i.

In that case, what we do? We may use the difference, finite difference method to estimate the value for the partial derivative of f with respect to x i. How we do it? We are considering f of X m plus delta x i u i minus f of X m divided by delta x i. and what is the meaning of it? We are considering the function partial derivative of function with respect to x i. Now, if this is so at point X is equal to X m, we want to find out the gradient of the function. Then what we will do? We will just change only the value for x i with a small increment delta x i u i.

Now, the small increment is delta x i. And what is u i; u i is the direction. In the x i direction that is the unique vector we are considering; that is why delta x i u I, if I just consider; let me take one example for this. Then it would be very easy for us to understand. Say, f (x 1 x 2) is equal to we are having 2 x 1 square plus 2 x 1 x 2 plus x 2. Now, we want to find out grad f at point x is equal to say, x 1. x 1 is equal to some point we are having. Let me consider at (2, 1) alright then what we will do? We will take the functional value at this point. 2 into 2 square no, not 2 square, we will consider 2 plus delta x 1. Some increment over x 1 plus 2 into 2 plus delta x 1. Instead of 2 we are substituting 2 plus delta x 1 into x 2 plus x 2.

Then, this is the functional value at point (2, 1) delta x (1, 0) clear; x 2 here 1 it is also 1, minus functional value that is f at point X 1 itself; that is way we will consider f at (2, 1) divided by delta x 1. If we consider this difference and the ratio this corresponds to the gradient of f at point X 1. And this process can be adopted by without using the differentiations of the function itself. Now, this is one process; this process can also be further improved by using the central difference method later on. But this is the case for the steepest method in the next we will move to the conjugate gradient method. And we are getting better result with lesser number of steps in the conjugate gradient method; that is all for today.

Thank you.