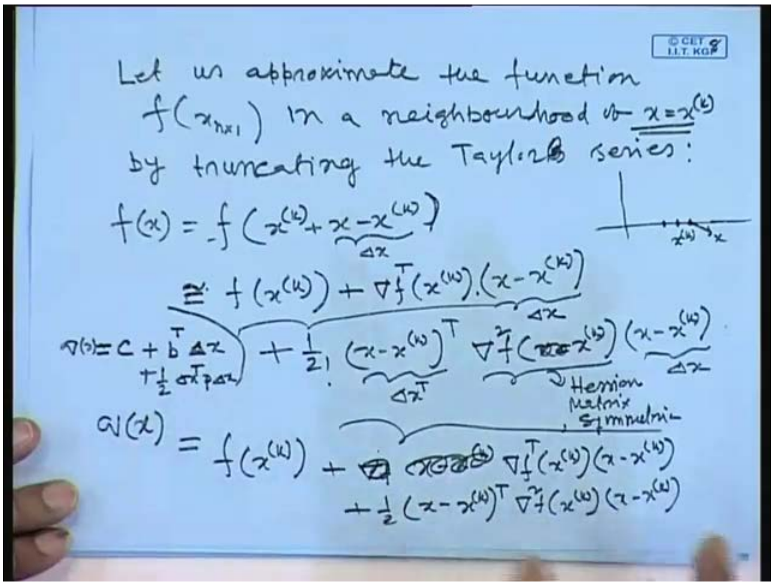
Optimal Control

Prof. G. D. Ray
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 7
Solution of Unconstrained Optimization Problem Using Conjugate Quadient Method and Networks Methods (Contd.)

So, last class we have discussed the unconstrained optimization problem using the numerical techniques. First we have discussed what is descent method? And then we have written the algorithm for the steepest descent method. That means, if you move from one point to another point the function value will decrease and what is the necessary condition for decent direction of the function that we have established. And we have seen the steepest descent method is very slow convergence and this convergence rate is in linear order, in order of 1. Then we have seen to improve convergence of this algorithm.

(Refer Slide Time: 02:09)



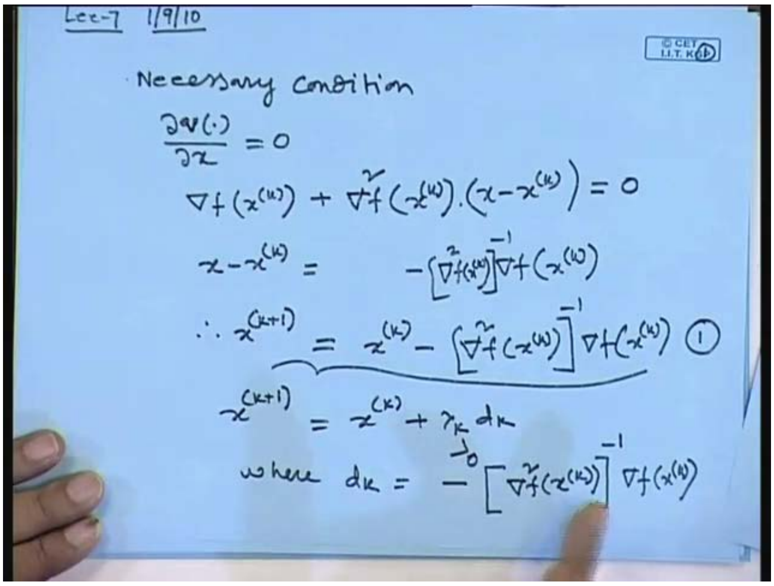
And you have considered the conjugate gradient method, that is based on the previous information of disintegration and presented information of disintegration, combination of these two information we have used it and we have found that there is a, rate of convergence that is improved. But the order of this convergence is not too less than 2, then we have discussed what is called the newton's method for solving the unconstrained optimization problems.

And if you recollect we have just considered this function, let us approximate the function f of x, in a neighbourhood of x is equal to x superscript k, means at kth point the value of x is x superscript k. And that function is what is called approximated by using Taylor series function. Then we have written x of k equal to f of x equal to f of x superscript k plus x minus x superscript k. So, this is the delta x part up from the x at the kth point what is the value of x, from there is the partavation is delta x. So, we are, the Taylor series expression we made it and kept it up to second order partially derivative terms, up to second order we have done and kept it.

Then, if you look at this expression, it is this, this function f x is nearly equal to the three terms we have considered. That is nothing but a quadratic form of a polynomial is that this function, one can write into this form c, this is a constant. Because, you know the function value at k integration, this is c plus this is also known, you can consider known b transpose and this is x minus delta x this is unknown. But x of superscript k is known, but x is unknown, so it is convenient delta x plus half then this you can write it, this you can write it, delta x transpose this is a hessian matrix and that is a symmetric matrix, that you write p and delta of x.

Then you see this is a quadratic form, that function which is f of x nearly equal to, can be approximate with a quadratic function because higher terms or the third order terms of the Taylor series expansion we have neglected. So, this function we are writing, we are equal to q of x we are writing so now q of x is equal to f of superscript, this plus this term plus this term. So, we have discussed up to that point now, q of x is known in the sense except this function, we have to minimize that means, we have to differentiate this with respect to x, x is a vector which elements are, that vector elements are x1, x2, dot dot xm.

(Refer Slide Time: 05:47)



So, if you do this one that our necessary condition for this function to be minimized, for the function this, this quadratic function is to be minimized, our necessary condition we have derived earlier. If you recollect this one, our necessary condition for the function q to be minimized is del q, del q dot means x function of with respect to x is equal to this. So, if you see this one that mean I am differentiating this function with respect to x. So, this is a constant so this derivation, derivative of this function will be 0. So, this is constant now, we have differentiating this with respect to x.

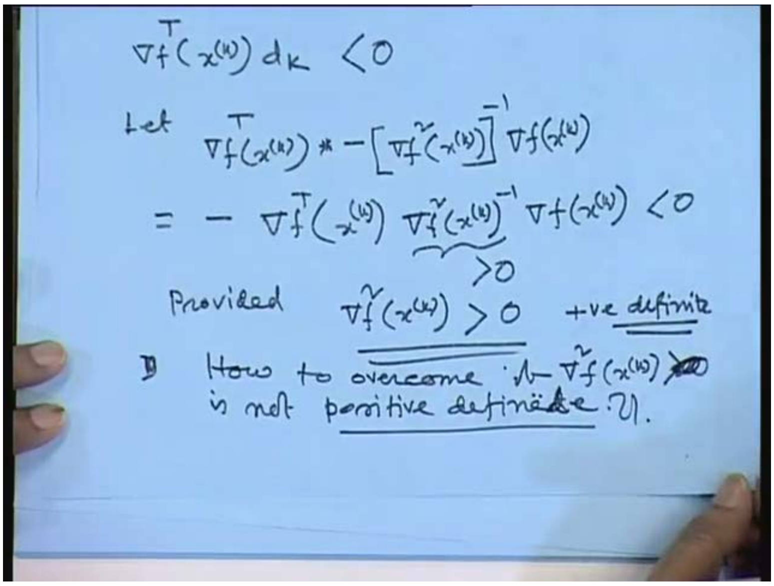
So, the results we have shown you earlier, if you have a f of x is equal to b transpose x, the derivative of f of x is nothing but a b. So, if you can say this is b transpose x, if you differentiate this with respect to x, this will be nothing but a b. b means delta f transpose, we have considered b transpose, that results if we apply, then this term will become b and this term it will come half is there so twice p into x. So, our results if you differentiate this one, it will come that f x superscript k, that gradient of this function at x is equal to x superscript k, I mean kth point plus delta square f x superscript k, this into delta x is equal to 0.

So, is nothing but a, if you see what we are doing is nothing but a the gradient of q is assign 0, I am finding out the roots of this polynomial, whatever this polynomial is coming this. So, if you now see this one what is x, therefore x minus x superscript of k is equal to, you can write, if you take this right side, that minus delta f x super script k this and this is the hessian matrix, which is a square matrix of dimension n cross n, what is n? n is the number of variable or decision variable involved in the objective function or cost function f of x. So, this inversion is del square f x superscript k whole that inverse. So, next is we can write it therefore, from kth point to another point we have moved so it is, it is, we can write it instead of x we can write, k plus 1th point.

This value is updated with knowledge of kth that point minus, see that is the hessian matrix at kth point, you compare then take this inverse of that one into del of x, x superscript k. So, this is our equation number 1, which every instant, every each iteration this x value is updated like this way, but look at this expression this nothing but it is similar to our Newton Raphson method, which is finding out the roots of the polynomial or any equation. It is the something like this way, x superscript k is equal to in general, if a function is given, how to find out the roots of the function? It is kth iteration value x minus f of x divided by f dash of x, if it is a, that x is a single variable it is, then it is. But here x, I am finding out the roots of gradient of this q.

So, this expression is now updated like this way. So, one can write it more general form that one, is like this way, if you say those, I can write it that x k plus 1 is equal to x of k plus lambda k into d k. And this is nothing but a, what is called newton's method, this is called the newton's method, when is that x variable is updating with this one. So, this is the newton's method, this can written into more general format, superscript k lambda k, lambda k value is greater than 0. And this is dk that means, the descent direction, descent direction to where, dk you can write it where? dk is equal to minus the second derivative of the function at x is equal to x superscript k. You find out and it takes the inverse, then add f of x k. So, this can be easily proved that, this is and dk, it can be that it is the descent direction, how to, what is the condition that the function value from one point to another point, if you move it? The function value should decrease, how to prove it?

(Refer Slide Time: 10:52)



You multiply it by dk, our condition we have derived if you recollect, we have condition we have derived it, the delta f x super script k transpose into dk must be 0. If it is in the descent direction of this function that we know dk, let us see whether this, let us see what is this value of that one. This is the condition that function value that, this condition, function value will decrease if we move from one point to another point. If it is, this condition is satisfied, then this function value will decrease from kth point to k plus 1th point, when you kth point to k plus 1th point. So, let us see this function is what and delta square and delta k we have just defined, we have got it, this is the delta k.

So, it is a minus delta k is minus, minus the delta k is minus, if you write it, delta f square x of this, the assignment is square into what is this, that your delta f of see, this delta f of x k. So, delta f of x superscript k this is multiplied, let us see what is this one. So, minus, minus sign is this one, delta f transpose x superscript k this, that f square x of superscript k, this inverts delta f x superscript k. So, now see this will be a, this is a quadratic form, this will be a negative when this matrix, the hessian matrix must be positive definite, then only this will be negative. So, this condition, this condition that gradient transpose into that dk will be less than 0, provided this will be less than 0 provided delta f hessian matrix of this function f x kth point, this one is greater than 0 means, positive dependent.

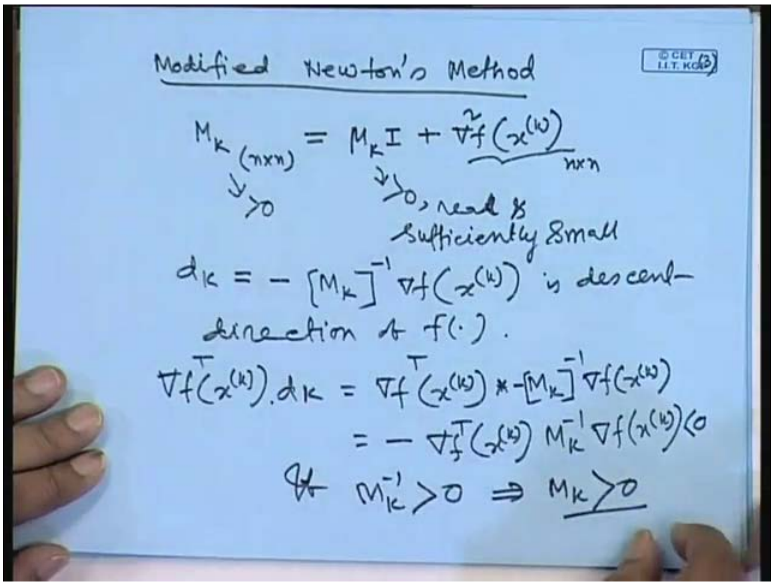
If this is positive dependent, if p is a positive dependent matrix, it inverse is also positive dependent matrix, that means this is positive dependent matrix. So, this is the condition implied. So, next condition is if this is fails, this is not satisfied this one, that means we are not approaching in the descent direction. So, and we are not approaching to the optimum value of the function. So, and which in turn, it indicates that we are, that convergence is not guaranteed until and unless that hessian matrix of the function at kth iteration must be greater than 0 means positive definite, greater than 0, means positive definite. So, what is the drawback of this one, algorithm? Drawback of this newton's method is there, in newton method how to update this one, x k plus 1 is equal to x superscript k plus lambda k dk.

So, this is the each iteration, it has to update this one and it will go in the descent direction of the function provided that, hessian matrix of the function at each duration must be greater than 0, mean positive definition. Then it will move, it will, we will go to the optimum means minimum value of the function, we are approaching to the descent direction, each iterations, that is the condition. So, our drawback of the, this method, that this must be a, one thing should be positive, one thing. Another thing if you see, when we are finding out the x k plus 1 updating this one, you need the inversion of a hessian matrix.

So, there is a inversion is involved and that matrix dimension is done by n because since our decision variable x has a dimension here. So, this is a n by n matrix inversion you have to do each iterations. So, that is the one drawback and another drawback, the hessian matrix or second derivative of the function must be, what is called positive definite. So, these two are the drawbacks of this algorithms. So, suppose that this that function, hessian matrix or the second derivative of function at each iteration does not satisfy this condition, then how to overcome these problems?

So next is that what is called, how to overcome if, how to overcome if delta square, second derivative of the function at each iteration x superscript k or hessian matrix of this one, inversion, if this one is not positive definite. So, how to overcome if this is, if this is not positive definite? So, how to overcome this one?

(Refer Slide Time: 17:08)



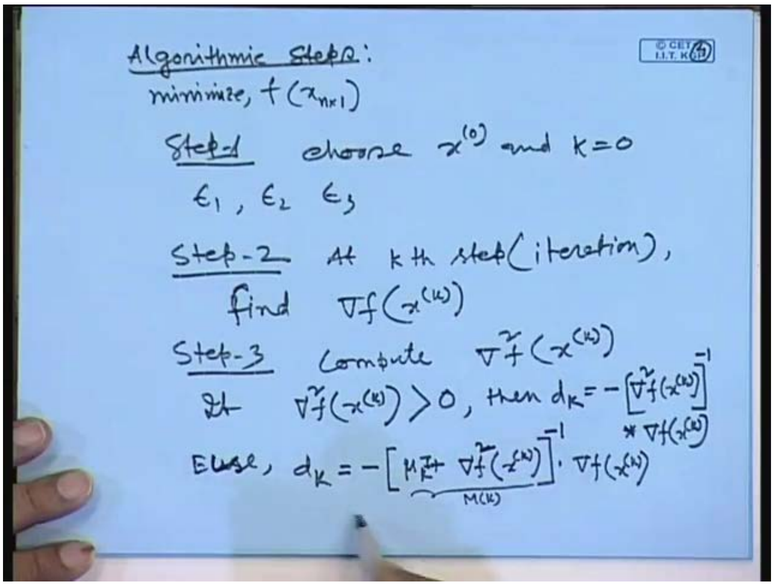
So, one of the, that is called the modification of newton's method for finding out the minimum value of unconstrained optimization problems using. So, our modified, you can say modified newton's method while solving the unconstrained optimization problems. So, let us define a new matrix of same dimension of the hessian matrix, whose dimension, I am just writing n cross n, matrix dimension n cross n. And this n cross n means, there is a number of variables are n x1, x2 dot dot xn. So, hessian matrix dimension is n by n, and then naturally m should be an n by n. So, that I am just adding a matrix, which is diagonal matrix with the hessian matrix and this matrix, this matrix dimension n cross n, this matrix dimension is n cross n.

So, what I did, if it fails that this quantity is not a positive definite matrix, why it requires a positive definite matrix? That d k, descent direction whether we are moving to the descent direction function or not, that we have seen, the condition must be satisfied, what condition? The gradient of the function multiplied by dk must be less than 0, which in turn implies that gradient, what is called hessian matrix of this one, should be positive definite. Suppose it fails this one, then you add a diagonal matrix with this one and where mu k is a very greater than 0 and real quantity greater than 0, real and sufficiently small, quantity greater than 0, real and sufficiently small this mu k.

So, even if it is negative define matrix, I am adding with some positive quantity and diagonal elements is this one. In other words you can say, if it is a negative definite matrix you multiply it by both side by z, any vector z transpose z, it will be negative. You multiply it by mu both side z transpose z so it will be positive. So, positive and negative in turn it may, what is called give you the m of k positive definite matrix. So, this is the idea of that one. So, this will make that m k is greater than 0 so what is our now then dk in place of the hessian matrix inversion, I will just write it minus mk. Now our modified is hessian matrix is like this, m k inverse gradient of x of superscript k of this one, is a descent direction of f.

So, this will ensure, this will ensure, this will ensure that f of our necessary condition that, when we move from one point to another point, that our descent direction, whether we are moving to the descent direction or not. And this condition will ensure, this into dk will be, will be see what is this one? Del f superscript k transpose and dk, I will write it dk value is here, dk value. So, it is a minus 1 minus then you write m k whole inverse, then gradient of that. Now again minus is there, I can write minus del f transpose x of superscript k mk minus 1 del f x k, that quantity will be less than 0, that quantity will be less than 0 means, negative. If mk is, mk inverse is positive quantity because ((Refer time: 22:02)) minus, this implies mk will be positive definite.

(Refer Slide Time: 22:36)



So, the mk suppose it fails it is a not positive definite matrix, then I am adding a small quantity of, along with this one, then we made it this mk positive definite, to avoid that problems if it is a negative definite. If it is a negative definite, the second derivative of the function, then the algorithm will not converge.

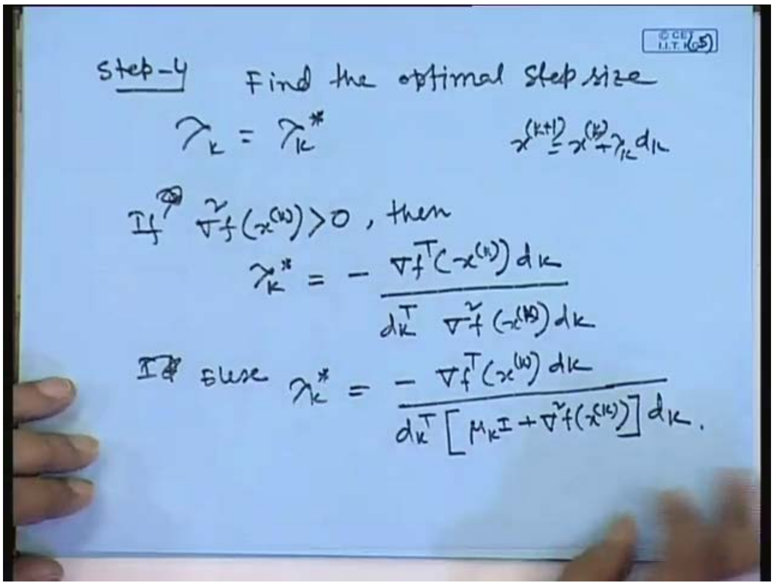
So, we can write this now, our algorithmic steps or newton's method for to solve the, what is called the unconstrained minimization problem. So, algorithms or optimization problems, steps, it is same as earlier what we have discussed first it is ((Refer time: 23:06)) method, then we have discussed conjugate gradient method, exactly it is same way we will write it first. Our problem is if you recollect, our problem is minimize this, minimize f of x. So, our step, first step choose the initial guess x of 0 and let our iteration starts at 0th iteration, k is already 0, then epsilon 1, epsilon 2, epsilon 3 are all positive real quantity, but very small quantity, very very small quantity.

And this is nothing but a tolerance for stopping criteria of the algorithms. So, is the tolerance for stopping the, what is called iterative process, this is the things. Step 2, once you know this one x of 0 is considered immediately I can find out what is called, gradient of this function, immediately I can find out the gradient of this function. Let us assume I am finding out the gradient of the function at kth iterations or kth step, kth iteration. So, you find out the gradient, find or determine the gradient at this because I need that information when will find out the descent direction of the function. So, x of k this you calculate of this. Next step, step 3 once you know this one.

So, I have to check it whether this is, I have to check, compute the hessian matrix of the function at x is equal to x superscript k, I mean kth iteration, what is the value of the decision variable value at this point you compute. Now, check if this gradient of what is called the hessian matrix of the function f x at kth step, this is greater than 0 means positive definite matrix. Then dk you update descent direction will be minus hessian matrix of that one, of the function inverse multiplied by gradient of this function x superscript k. So, this dk you find out. Suppose it does not satisfy, you try else, you try dk is equal to minus, this is mu k plus that our, what is this one, you have to consider? Gradient of f square, you see this one, what I am writing mk is modified mu I plus delta square hessian matrix of this one.

So, our this one is mu I, sorry I missed I, mu I that this square, that is the second derivative of the function at kth step, this you do it that one, then take inverse that is nothing but a mk. This whole thing is our mk if you recollect, this is mk whole thing inverse into that our f x superscript of k. So, this part, this inversion because if it is less than, if it is not greater than them means positive definite, then if you try with this one the algorithm will not convert. So, we are trying with adding with this one a positive quantity mu and then taking, updating the dk is like this way. And we have shown this one, if you multiplied by gradient transpose of this one condition for descent direction is satisfied, if it is a positive definite, this. So, dk we got it now.

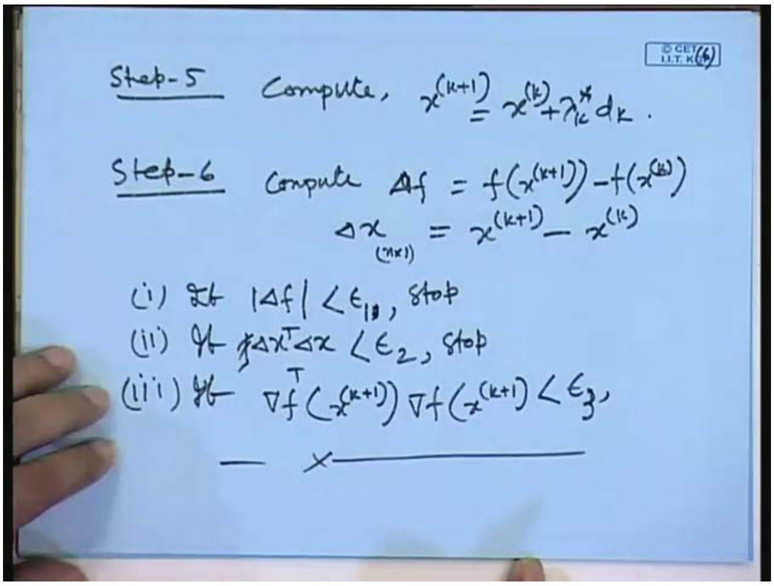
(Refer Slide Time: 27:55)



Once you know dk of this one, immediately I can find out step 4, up to this step 3, find the optimal step size lambda k is equal to lambda k star and that we already know, if you can recollect this one, that we know already, how you need to find out that one. So, that means lambda is equal to lambda k, how you have done it? You know dk now, by this time you can find out x superscript k plus 1 is equal to x k plus lambda k and dk, you know dk, you know xk. So, in the function you put it x is equal to x superscript k plus 1. So, that function will be now function of lambda k, now question is what should be the choice of lambda? Now it is a single variable function, what should be the choice of lambda k so that, function below will decrease as small as possible in the descent directions, that is our problem.

So, if you differentiate this one with respect to lambda assign to 0, then you will get it the lambda value star is, we have done it earlier also. Now our question is like this way, if gradient, not gradient that hessian matrix of the function at kth iteration this is greater than 0, means positive definite then your, then your lambda star, lambda k star is equal to minus lambda f transpose x of superscript k, this into dk, divided by that dk, this. So, I know dk then you update that one, but if it is, if it is less than, if it is not true, else if this greater than 0 mean positive definite matrix, then update like this way. Else lambda k transpose star is equal to minus delta f transpose x of superscript k dk but already it will be a change of that one. So, dk transpose it will be a mu k I plus delta square f super script k, this is k, that into dk, that is the things. So, this way you have to update the optimum size of lambda, you can get it.

(Refer Slide Time: 31:22)



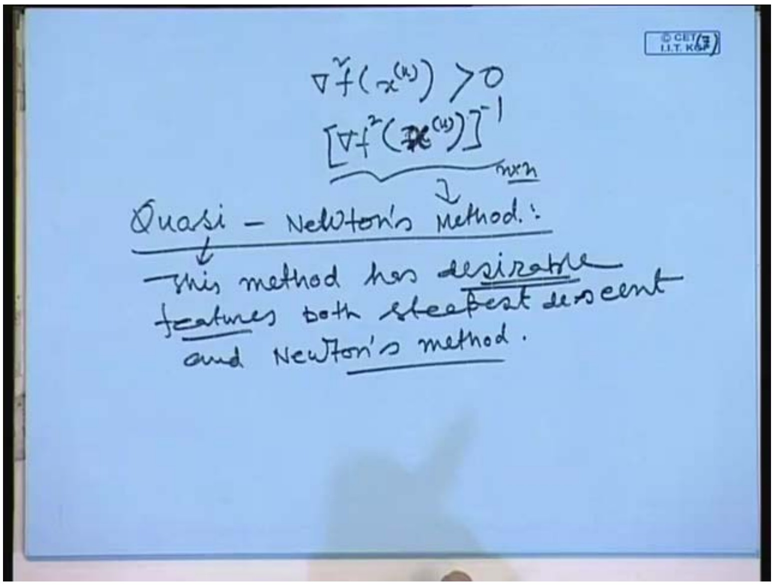
Once you get it, optimum step size then immediately I can find out what is the updated value of the decision variables. The step 5 compute x superscript k plus 1 x superscript k plus lambda k star, just we have to obtain that one in to dk. So, next is check the, with the tolerance whether iteration will be stopped or not, this one each after each iteration we have to check that one. So, compute df and df we have denoted by f of superscript k plus 1 function value at kth iterations minus, the function value the k plus 1th iteration, function value of kth iteration, the difference is delta f. And delta x this is a scalar quantity, but this is a vector of dimension n cross 1 and that is the difference of value in at two points at kth iteration decision variables value and k plus 1th decision variable value is minus kth iteration, two successive iteration value difference.

Now if delta f absolute value of this one is less than epsilon stop it, it indicates the function value is not changing at all. Then 2, if delta x here I cannot write that mod because it is the f is scalar I can write mod, now I have delta x is a vector, either you write the iteration norm or you write the distance square delta x of delta x, it is a scalar quantity. Now, this will be a, this is epsilon 2, this is epsilon 1, stop, this indicates that decision variables value are not changing. And third criteria one can use it, what is called that our gradient function value, that has for gradient of f transpose at k plus 1th iteration and delta f of this one iteration.

This indication is what? This is nothing but a gradient of function is the vector that means, we are finding out the slope of the f at x is equal to x 1, keeping all the variables fixed. Similarly, gradient of this, that is the function derivative of the function at x is equal to at, what is called, x is equal to x 2, point all other points x1, x2, x3 dot dot other points are remaining fixed means, it is a vector, gradient is also a vector. Vector is a column vector multiplied by row vector. So, if this quantity is less than epsilon means, if it is very small that means, we have, we are approached to the minimum value of this function, where the slopes are gradients are almost 0, that is epsilon 3, that, it indicates that we have reached to the convergence.

So, this is the, our algorithm steps for modified or the, our conventional newton's method. So, next is that, I told you that, that is a difficulty is, if the hessian matrix of the function is not positive definite matrix. Then we have modified that matrix by adding with a scalar quantity mu k, multiplied by identity matrix of the same size of a same matrix, this is one way of doing. Another method of doing is there, I mean you have mentioned that the draw backs of the what is called, the newton's method.

(Refer Slide Time: 35:44)

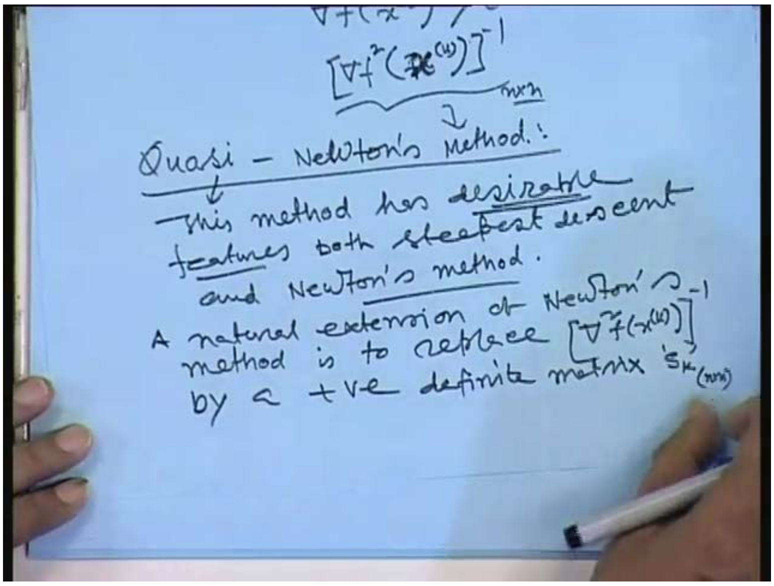


One drawback is that hessian matrix or second derivative of the function must be positive definite, that means, that one should be positive definite. If this is a positive definite, this means we are approaching to the descent direction of the function. If it is not that means, we are not approaching to the descent direction, away from the descent direction that means, that, the what is called, this each iteration to the algorithm will diverse, if this is not satisfy this one. Another disadvantage of newton's method is, we have to take each iteration the inverse of a hessian matrix. So, this, that is x superscript k, this hessian matrix inversion, we have to take each iteration and competition burden is involved here is, much, if it is a number of variables is more.

So, it is a time consumed for determining the search directions, you can say time consuming because the dimension of the matrix is this and inversion we are taking. So, it is nothing but a time consuming to find the direction of this search, the function. So, next is what is called, we will take a Quasi-newton's method, this is similar to newton like method, newton similar to newton method. This Quasi-newton's method takes the advantage of what is called, the two, the steepest descent information is taking and newton's method what is called, information also taking.

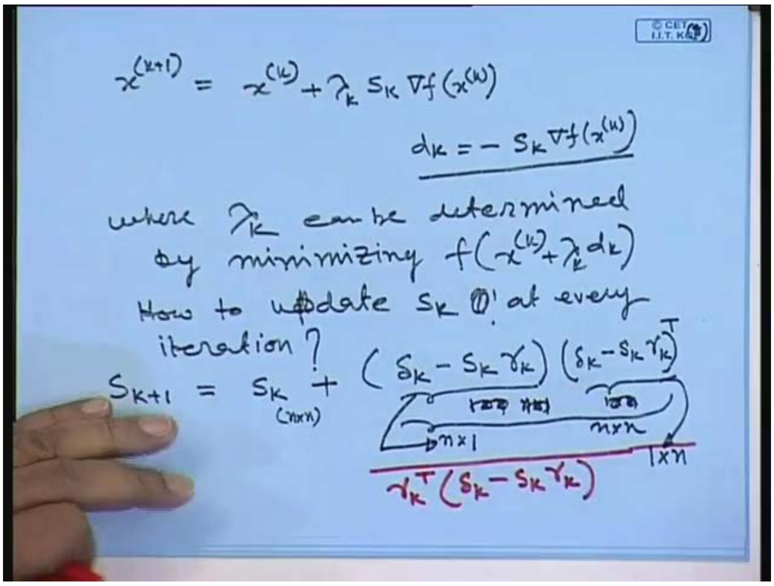
In newton's method we see hessian matrix is there, inversion you have to do it. So, it is the quasi this one, this method you can write it, this method has desirable features, both steepest descent and the newton's method. But here you will see that, it does not take directly the inverse of this matrix, inverse of hessian matrix, it finds the inversion, in place of taking the inversion, it finds the iterative method to avoid the inversion of assigned methods. So, what is this method will see this one. So, a natural extension of that newton method is from this method, you can extend it decreasing the inversion of a hessian matrix by some matrix.

(Refer Slide Time: 39:13)



So, what is this? It is doing here a natural extension of newton's method is to replace the inversion of hessian matrix at kth iteration, at each iteration, that inversion is replaced this thing by a, as I told you, positive definite matrix. Say, the matrix is s k, whose dimension is n cross n, hessian matrix dimension is n cross n. So, I am replacing this one by a metric, s of suffix k means, kth iteration. Then this metric should be positive definite because in order to move in the descent directions.

(Refer Slide Time: 40:20)



So, what we are doing it here you see now. Now, x of the descent variable at the kth iteration is equal to x of k plus 1th iterations is equal to kth iteration. This is in variable plus lambda k, the step size and that descend direction will move it, that we have to find it optimal tips size, that we have shown how to find out. Then this multiplied by what is next is here in newton's method that hessian matrix inversion, in place of this one I am writing sk is a matrix, that inversion is now replaced by sk and delta f of superscript of k, gradient of this matrix. Now therefore, our dk is equal to minus sk what is called such direction is what, sk is called positive direction matrix, which is replaced by this is, this sk in place of hessian matrix inversion, I replaced by sk.

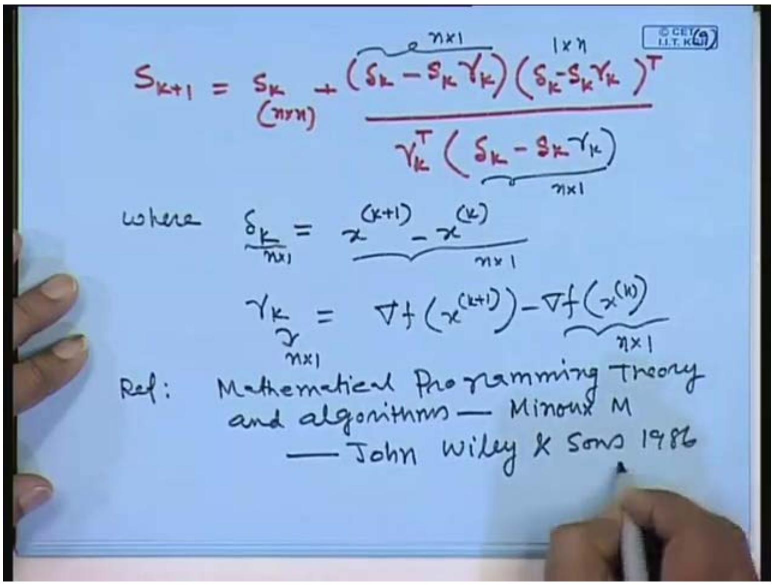
So, that f of x like this. So, this is our direction matrix so long, sk is positive definite matrix, the condition of descent direction matrix of the function is satisfied. So, where lambda k can be determined by minimizing the function lambda k, same way lambda k can be determined by minimizing x of superscript k plus lambda k then your dk this one, that way I have repeatedly we have told how to get this value. And this value, that lambda k what is optimal value of this value of lambda k so their function value is minimized at that iteration, kth duration we know how, what is the choice of the duration.

So, now xk you know it, but next iteration s, at k plus 1th iteration what should be the value of that one? You know this one, but at that instant again you have to find out the what is called, hessian matrix value at x is equal to x superscript k plus 1, I mean k plus 1th iteration. But inversion you have to take it, but that things we have to avoid that one inversion we have to avoid. So, I have to update sk by what is called, sk plus 1. So, next is, our choice is how to update sk, next question at every iteration. So, that is our question. So, if you see this one that how you are updating, I am writing the algorithms then I will show the proof this one.

So, sk plus 1 that next iteration, the value of the hessian matrix inversion that sk plus 1 at k plus 1th iteration will be sk, previous iteration value sk plus delta k, I will define what is delta k, sk then gamma k, multiplied by, mind it, this sk is what? The dimension of sk is n cross n because it is a hessian matrix inversion I replaced by sk. So, this is dimension is n cross n and this dimension, if you see this dimension is a one row and column, I will discuss how it is. Then multiply it by same matrix, sorry same vector gamma k, this is gamma k sk transpose, this is a row vector of dimension, this dimension is one row and column. And that dimension is, if this dimension is n row, this dimension is, if you see this dimension no, sorry, this dimension is n row, one column, sorry this dimension is n row one column. And this dimension is your one row n column, just a minute, one minute.

So, this dimension, now see this one, this product of this one must be a matrix of dimension n cross n. Let us see now this one because we are adding with sk. So, this should be a, how many rows are there? n rows one column, this one should be n rows one column and this will be a yours, if you see this one, this one row n column so I was correct earlier. Now what is this one? So, that I am writing, I am using this one, that whole thing is divided by, this whole thing, this is the dimension I have written it, this row vector, this column vector is divided by gamma k transpose. Then this row vector delta k minus sk gamma k.

(Refer Slide Time: 47:00)



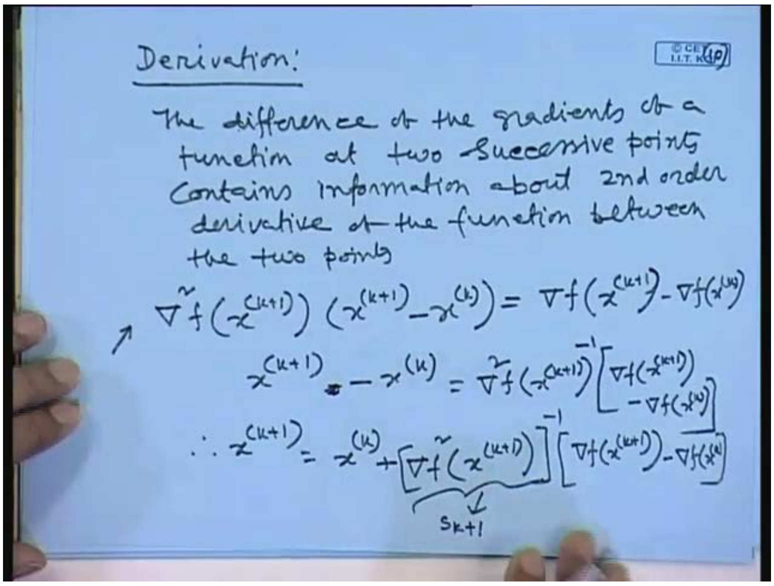
So, I repeat once again, that what you have written it here clearly that, s k plus 1 is equal to s k whose dimension is n cross 1 plus delta k sk into gamma k multiplied by transpose of that one, delta k minus sk, sk gamma k, whole transpose divided by that gamma k transpose into delta k minus sk, this is sk into gamma k. Now, I told you what is dimension of this matrix; n row one column, and what is this dimension of this one row and n column. So, we multiply by this it will be a matrix. So, matrix divided by that must be scalar quantity because you cannot divide by another matrix this one.

So, this thing I told you what is this one, that is your n row one column, n row one column and that must be a one row n column. Let us see where, what is what, delta k is what? Where, I can write delta k now it will be clear what is the dimension of this one delta k is the difference in decision variable at two successive iterations and what is the dimension, you know n row and one column this dimension. So, gamma k is nothing but a, the difference in gradient value at two successive iterations, so delta f x k plus 1 minus delta f x superscript k, this one, what is the dimension of that one? We know already, this dimension is n row one column.

So, this gamma dimension is n row one column and this dimension also n row one column. Now, you see this one, this dimension you check it, you will see this will be a matrix, I have written this one. So, this way you have to update now, what is this, if you recollect this one, that our quasi what is called, newton's method is nothing but a, it is newton’s method only the inversion of the hessian matrix is replaced by a matrix sk. That inversion and then multiplied by the gradient and this two commonly is given the, this thing is given the, our descent direction. If you see the descent direction is, dk is calculated minus the hessian matrix square, hessian matrix inversion into gradient of this function at kth iteration.

In that one we have sk is this one minus sign is there so this is the descent direction. So, I am just then once you know that s, then how to update this x, next iteration. This update is done by using this expression what is called, this expression. Now question is how you got it this relationship? So, I am just putting it this one, reference you take mathematical I will derive it, but still more study if it going to do, mathematical programming theory and algorithm, that author is Minoux M, publishers is this is a book, John Wiley and sons, 1986. So, the basic steps of that one is derivation just I will show you, how I have written this updated expressions and sk is what? Is the hessian matrix inversion is replaced by sk and how sk is updated, it is written by that expression.

(Refer Slide Time: 52:14)



So, next is derivations, we know the difference of a function at two successive point, carries the information about its first derivative. And multi variable case, multi variable function the difference of a function at two successive points, carries the gradient of this functions. Similarly, I can say difference in what is called, gradient of a function at two successive points carries the information of second derivative or carries the information of the hessian matrix. So, in short I can write it, what is difference in the difference that one I am just writing, the difference of the gradients of a function at two successive points contains information or carries information about second order derivative of the function between the two points.

This is the same as the difference in function value at the two consecutive values, carries the information about the first derivative. Next is the, I am telling the difference in what is called first derivative, value of the difference of first derivative of a function at two consecutive points carries the information of the second derivative. So, we can write it in an variable function that gradient of that f x k plus 1 is second derivative, x superscript k plus 1 minus x of k, this is equal to difference in gradient at two successive points minus f of x superscript k.

So, difference in gradient of two successive points carries the information of the second derivative of the function. In multi variable case I can multi variable functions, this difference in gradient and it is a hessian matrix carries the information of hessian matrix. And this is the, multiplied by that vector, if a single variable case you can easily understand the difference in derivative divided by because x is a similar I can divide this one. But in multi variable one I cannot divide that one. So, this is the basic step of that one. So, what we can do this now is you can see here, xk plus 1 is equal to, if you see this one x k plus 1.

I can write it x k plus 1 minus x superscript k is equal to, this you take that side, that will be a gradient, not gradient it is a hessian matrix of that second derivative of the k plus 1th iterations, inverse multiplied by that one, difference in gradient value at consecutive two points minus gradient delta f of superscript k. So, this one see. So, what is this expression now, I can write it x k plus 1 is equal to x of k plus the second derivative of function or hessian matrix x superscript k plus 1, whole inverse into del f x superscript k plus 1 minus del f x superscript of k this one. So, this I can represent by a x suffix, this I can write it x suffix k plus 1. I have subscribed kth iteration inversion I have replaced by xk. Now it is a xk plus 1 iteration, that hessian matrix inverse, I denoted by it is a sk plus 1. So, today I will stop here, I will continue next class, the continuation of this chapter will be next week.

Thank you.