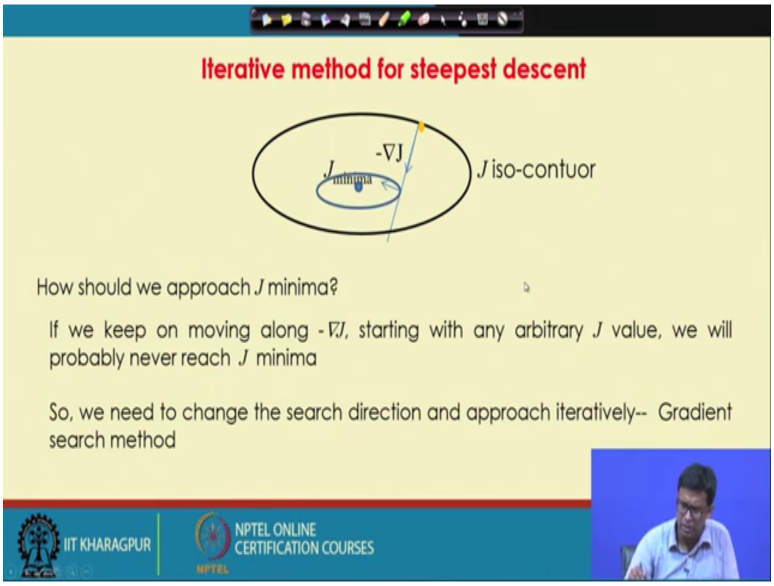
Matrix Solvers
Prof. Somnath Roy
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 44
Krylov Subspace Methods for Linear Systems

Welcome. So, we are discussing about the steepest descent method. In steepest descent method what we have seen is that solving Ax is equal to b for an Ax is equal to b is equivalent to find out J minima with J is equal to x transpose Ax minus x transpose b.

(Refer Slide Time: 00:23)

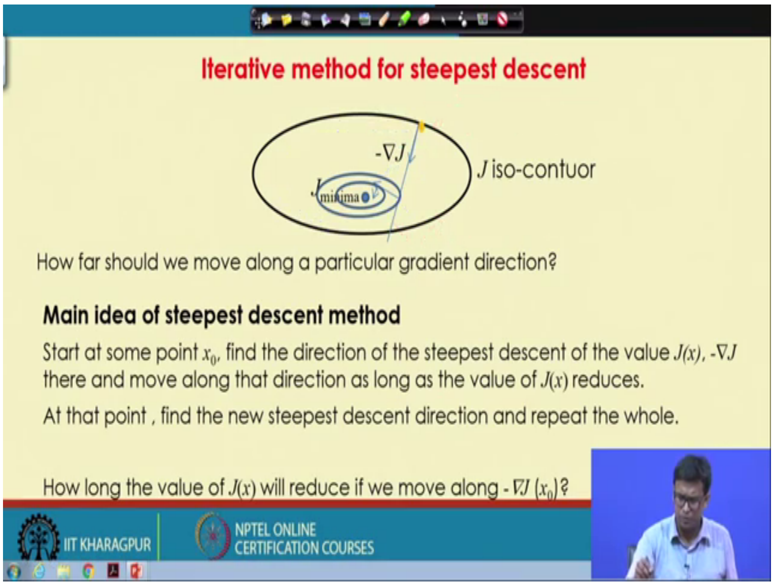


And provided A is if and only if A is symmetric positive definite matrix; so, here in solving a matrix equation is substituted by finding out minima of a particular functional. Now, why we are doing this? Because we have discussed that the direct iterative methods like Gauss Seidel and Jacobi has certain limitations when to come out of those methods and explore, the other ways of iteratively solving x is equal to b. So, what we started here is that solving x is equal to b is same as if A is as a symmetric positive definite matrix, solving Ax is equal to b is same as finding minima of the functional J x is equal to x transpose Ax minus x transpose b.

Now, we are with the question that how to find out minimum of a function. So, what do you have observed that, if we take any arbitrary value of the primary variable x, the function and define the function J x; J x reduces first test at that particular location in the direction of minus grad J. So, we will move along minus grad J and try to find out what is the minima of J. And what we have done is that we have drawn the iso-contour of J, and a normal to that is the direction of minus grad J started moving along that. And this line J minimize some way apart from this line. If only if the iso-contour is a circle the any normal would have taken me to the center, now it will not take me to the sign to the minimal value not to the minima. So, it is it will miss the minima.

So, we will go up to certain distance and then we will draw the J iso-contour again, this is another iso-contour. And go drop a minus grad J. So, this is again a minus grad J new at this particular location and move along that. In a way we will keep on changing the such directions and finally, approach the minima of J. So, we will look into the detail of this method now.

(Refer Slide Time: 02:49)



We started with one particular value one particular J iso-contour and there is a J minimum moved along one along the minus grad J reached somewhere and then got the new minus grad J and move along that. And then again he reached another iso-contour and then move along minus grad J. So, somewhere will reach it. The question is that how far should we move in one particular direction, what is the amount that we should move here? What is the amount that we should move here? For fastest convergence or for reaching the J minima fastest way.

See if I move out of this up to here this is tangential right, after this J is actually increasing. Here minus grad J is 0 this line is parallel to the J iso-contour. So, if I move along this particular line, J is not increasing here. Now J will keep on increasing. So, we can go to a minima local along one particular such direction, and then J will keep on increasing. So, if we move along this will never reach J minima. These are probably nearest point to J minima. Again we go here and somewhere we have to stop otherwise, we will miss this line will go away again we will come back to if we keep on going this if we keep on going this again I will come back to the old J value; so, I cannot do that. Somewhere I have to stop and make a new search. Find out the new minus gradient J and move in that direction.

And the main idea of the steepest descent method is that, start at some point x 0 find the direction of the steepest descent at this at that point of steepest descent of J x which is minus grad J there. And move along that direction as long as J x reduces. And at that point when J stop producing so, along this line move in a direction where J is reducing; when J will stop reducing you change the side as such direction. At that point when just J is not reducing any further, find the new steepest descent direction and repeat the process; that means, again calculate minus grad J and move along that direction. Again see where how long it is reducing, when J will stop producing. Because anyway J is bound to radius if minus grad J if grad J is non-zero along minus grad J, J will reduce.

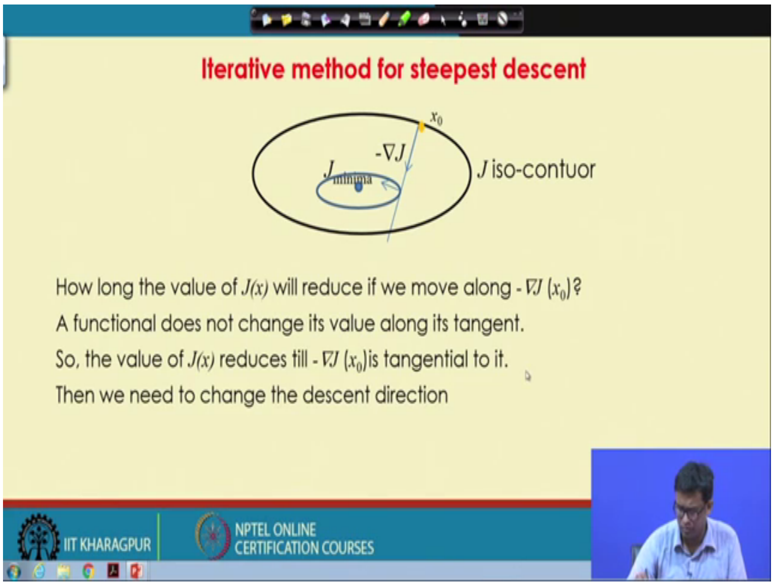
See how long minus grad J is reducing, why J will not keep on reducing along minus grad J? Now if I look here say J there is some value J x something J x is equal to f of x. Now if I and I calculated minus grad J here. When I came here the iso-contour of J is J x is a different function J x is equal to g of x or rather instead of J x is f as a particular constant value J x is equal to say g. Now grad J is not same here. So, if I move along this line this is the direction which it is anyway perpendicular to this particular iso-contours, and this is the direction in which J is reducing fastest. When I came here J when I just move away from this line minus grad J at that point is different.

So, J is not reducing fastest in that line anymore, only in this point J was reducing fastest. Now J is the reduction rate of reduction of J will slow down, because minus grad J is different if I grow a contour here another iso-contour here, that will probably have a different minus grad J direction. So, that drop is in a different direction. So, it is not the direction where J is reducing faster anymore.

However, we will still keep on continuing that, will come here where J is not reducing any further. Why because, grad J is tangential to this grad J this particular value of grad J evaluated here is tangential to this iso-contour. So, if we change J little bit if we move little bit along grad J, J will not change. J is constant, if we come out of it that then J if we move further J was increasing. So, I have to stop here and now I will evaluate grad J at this point, because here J is not reducing any further and make a new search.

And geometrically it is well explainable that you move along one particular direction see that this direction vector is getting pair tangential to a new iso-contour, you stop here take another turn. But how to find out mathematically? So, the question is how long the value of J will reduce if we move along minus grad J. Where should we stop? How long will this value reduce?

(Refer Slide Time: 08:13)



The same question I have started with an x 0, how long the value of J x will reduce if we move along minus grad J evaluated at x 0? This I am moving along this minus grad J over this is evaluated at x 0.

So, if I come to any other contour, sorry, if I think of any other contour minus grad J is evaluated here in a different way. So, it is not probably the way where it is reducing in the fastest way, but still it is reducing. But here, it is not reducing any further. So, you have to find out how long value of J x will reduce. Then they are up to then I started with one particular value of J I am trying to reach J minima. This is an iterative method I started with some gaze value of x where there is some value of J. I took a method which will take me to the I am trying to get an method which will take me to the minimum value of J. Why I am trying to take a method which will take me to a minimum value of J because, if A is symmetric positive definite matrix, then the way I J define I define J, J minimize the solution of Ax is equal to b.

So, I am trying to reach J minima starting from any arbitrary value of J. I will see the value of J is reducing, the value of J will keep on reducing till I reach J minima ideally. In if I can go in any direction, I should reach after I reach J minima value will increase. If value is if value of J is increasing I will never reach J minima, I am going in a started with own particular value of J moving in some direction.

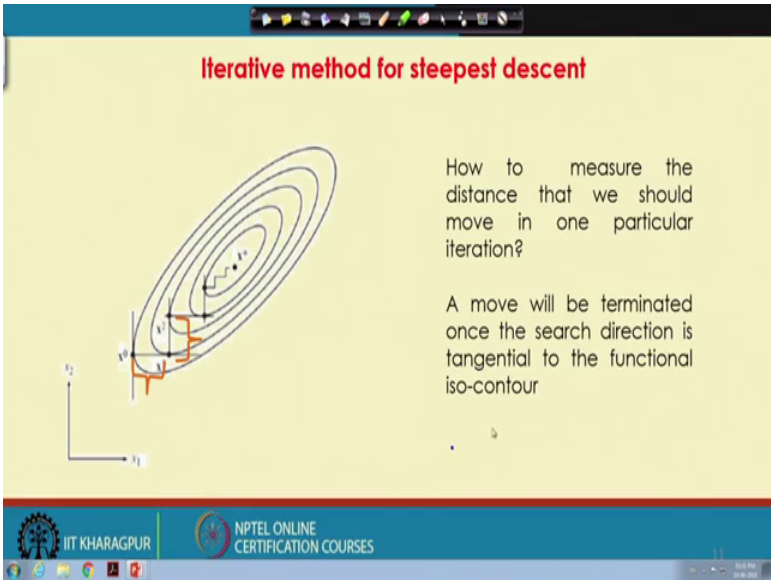
If value of J is increasing I am not reaching minima I am going away from the minima. So, I iteratively I will start with one particular value of J move along one direction, work till value of J is reducing. If I have reached minima, I will reach the least value. I will see if value of J is reducing I probably have reached not reached minima, but it is stop reducing. So, if I still move along that particular direction I will never reach minima. So, have to stop that point and then change check the search direction.

So, how long will move along grad J? Till the value of J is reducing along that direction. How long value of J will reduce if we move along grad J? It will not reduce forever because, after a particular step it will keep on increasing, because this J line is probably not grad J line is probably not taking me to the minima. A functional does not change it is value along it is tangential tangent. So, when I will go along this direction as soon as the grad J become tangential to one particular iso-contour of J the value of J will further not reduce. A functional does not change along the tangential direction.

So, if value of J reduces, till if value of J reduces the value of J will reduce till minus grad J x 0 is tangential to J x evaluated at a different level. So, x is a multivariable function. I started with one particular value of x. I can think of a2 d function where it is and ellipse if they are elliptical conclude x 0 square plus y 0 square is equal to constant, if f of x is equal to x x square plus y square. So, I started with a one x 0 y 0 and evaluated grad J there. I moved along that, but along that particular line. Now reached a point where this line has become iso-contour to another x square plus y square is equal to b or not x square I started with x square plus y square is equal to a contour moved along grad x, now I reached another line where x square plus y square is equal to say c square another different line.

And this line is now tangential now these are contour is now tangential to grad J. So, I will stop here and change the direction. So, the value of J x reduces till minus grad J x 0 which is evaluated here. J will reduce till this is tangential to this particular contour. Then we need to change the descent direction. So, descent means were coming down to the minimal value. So, I will find out the new grad J here; which is the new value say this is x 1. So, this is minus grad J x 1, I will find out the new value here and move along it. And again I will how far should I move when I will see that this is again tangential to another particular iso-contour, maybe there is another iso-contour like this. And this line is tangential to that I will come here, and then I will evaluate again minus grad J and move in that direction and that is how I will approach it.

(Refer Slide Time: 13:48)



So, the method is pictorially like this. They start with one particular x 0, start with one particular x 0, move along a direction. See it is being tangential to the functional here, then you change the direction move along this find x 2. And so, this is this is like minus J x 0. This is like minus J minus radiant of J x 1 this direction. This direction is minus gradient of J x 2, you come to x 3 and change the direction and so on you will approach the minimum value. This is the minima.

So, starting from any arbitrary value this will take us to the minimal value. I am trying to explain this geometrically, and geometric explanation is only restricted to R 2 to 2 variable functions. All these curves are for to 2 d curves, 2 variable functions we can extend it maximum to the third R 3. But this geometric concept is applicable for a higher order and R n things.

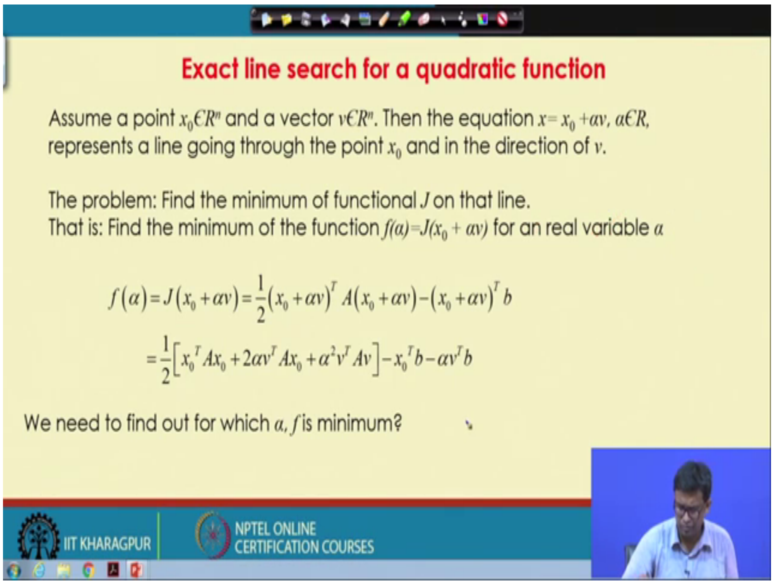
However, once we finish this discussion and propose the final algorithm from strips or steepest descent, we should check that algorithm wise that converge or geometrically we can see that starting from any x 0 y 0 it will take me to the minima. In R n the algorithm also should converge that is starting from any x x 0 value it should converse to the right solution and you should approach the J minima; will check it once we are done with this discussion.

How to measure the distance that we should move in one particular direction? So, this distance x 0 to x 1 x 1 to x 2, how to measure these distance that we should move along one particular of minus grad J. A move will be terminated the idea is that, that a move will be terminated once the search direction is tangential to the functional iso-contour. So, we will start with a search will start search with a particular search direction; which is based on minus grad J x 0, based on the value of x 0 the search direction, based on the value of x 0 this search direction has been taken. This direction this search direction is tangential to 1 particular iso-contour, then the move is terminated.

So, when should again we will stop which iso-contour because you can have infinite number of iso-contours there. We should stop in that particular iso-contour after which this value J keep on increasing. Or this is the iso-contour at which we are stopping; that means, where this search direction is tangential this will go with the search direction check where this direction is being tangential to on particular iso-contour there we will stop. After that J will keep on increasing.

So, I have to start with a search direction and check when the search direction is being search direction is we will give me some grad J x 0; when this grad J x 0 is being tangential to 1 J x that is grad J x 0 dot grad J x will be 0 new J x will be 0. The this that this the new search direction and all such directions are perpendicular to each other.

(Refer Slide Time: 17:46)



I assume and also another factor is that, if we go along this particular line, if we go along this particular line locally J is; so, if I smallest, not say minima. Or if I draw say J versus x along this line along this line. So, this is x 0 and this is x 1, I will see J is locally mini local minima is along this line. At this point J is locally minima, because here J is not increasing any further and then J will keep on increasing.

So, I will get a local minima on that particular line. So now, I can probably search it in that idea also; that I will go along one particular line and see when J is minimum along that line. Instead of finding out a global minima will find out local minima along one particular search direction; where J is minimal, I will change my search direction. Why J is minimal there? Because J is being like grad J is the direction in which J is changing. Grad J is tangential to the iso-contour; that means, over that iso-contour, J is not changing along the iso-contour J is fixed. So, at that at that point grad J is not giving us any change of J, because it is tangential to the iso-contour of J. So, this is the local minima of J 2.

So, I have to find out the local minima of a function along a particular line. So, we will first express the function along a particular line, how the function changes along the particular line. Assume a point x 0 which belongs to real number coordinate of or dimension in R n and a vector v on that same space R n. And then the equation x x is equal to x 0 plus alpha v, alpha is a real number represents a line going through the point x 0 and in the direction of x 0 plus in the direction of v.

So, this is actually very straightforward. I have a point x 0 and a direction vector v. Any point here any point x here can be represented x 0 plus alpha v. The and this distances I will find that sense. Now the problem is, find the minimum of the functional of J on that line; that is, that at this point J will also be I will get J of x 0 plus alpha v. And this v here is gradient of minus J.

So, I have to find out along this line where J of x 0 plus alpha v is minimum. Now if v is fixed given the point x 0 v is equal to minus gradient of J at x 0, or let me write it down somewhere else. So, we got v is equal to minus gradient of J x 0. That is a direction along which we are doing the search. This is the direction v, direction vector v.

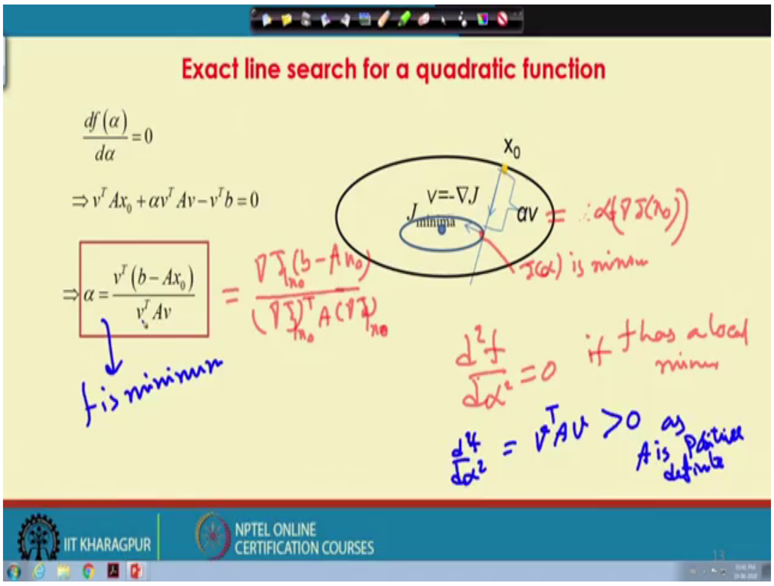
So, I started from here, v is fixed v is the particular direction which is fixed. I started from here moving along v. J is evaluated as J of x 0 plus alpha v. X 0 is fixed because that tells me what is the direction from which I have started, v is fixed. So, the only variable is alpha. So, depending of the value of alpha, I should get J minima somewhere in the this line. The local J minima will be where for one particular alpha where del J del alpha is equal to 0.

That is find the minima of the function f of alpha J which is J of x 0 plus alpha v as x 0 and v are constant for our real variable alpha. Depending on the for, sorry, depending on the value of alpha this will have a minima. This becomes only a function of alpha now. So, f alpha is equal to J of x 0 plus alpha v is equal to half of x 0 plus J x; if we remember the old functional term J is equal to x transpose Ax minus x transpose b. And A is symmetric positive definite SPD matrix.

So, J of x 0 plus alpha v is half of x 0 plus alpha v transpose x 0 plus alpha v sorry A x 0 plus alpha v transpose into A into x 0 plus alpha v minus x 0 plus alpha v transpose to b. Which is half of x 0 transpose Ax 0 plus 2 alpha v transpose Ax 0 plus alpha square v transpose alpha v. So, if we break it down we will get this expression, minus x 0 transpose b minus alpha v transpose b.

Now, we need to find out an alpha for which this f is minimum. So, v is fixed A is fixed b is fixed x 0 fixed alpha is only a function of f where f is minimum d alpha d alpha d d f d alpha is equal to 0.

(Refer Slide Time: 24:04)



So, we got d alpha d f alpha d alpha is equal to 0, or if we go back to the old form we take a derivative with respect to alpha which will this is 0. This will be 2 v transpose Ax 0, 2 and half will cancel out. So, v transpose Ax 0 and this will be 2 alpha 2 and half will cancel out alpha v transpose Av and this is v trans minus v transpose v. So, this will d alpha d alpha will give me v transpose Ax 0 plus alpha v transpose v minus v transpose b is equal to 0.

So, alpha is equal to v transpose b minus Ax 0 by v transpose f v the this carries out here. Therefore, this is the distance, this is the amount alpha that we should go alpha v we should go alpha v along one particular direction till we get the local minima from which the changed search direction should be changed. Or I will start with one particular value x 0, and go along alpha v is equal to alpha into minus grad J evaluated at x 0, alpha into sorry, this is alpha into this. So, I will go alpha into grad a gradient of J x 0 in this particular direction and here, I will get J alpha is minima. So, I should change the direction.

Again I will calculate what is the gradient of J here and based on which I will calculate what should be the alpha. So, as v is gradient of J, this can be also written as gradient of J into b minus A x 0 by gradient of J transpose A into gradient of J. Or rather this is that evaluated at x 0 this is evaluated at x 0, this is evaluated at x 0. So, again when I will come here I will reevaluate the gradient of J and reevaluate the alpha and go along that direction. I should reach the minima.

Again because we are saying that J has a local minima here, it is also important to show that d 2 f d alpha square is equal to 0; if f has a local minima. This is the global minima and this is the local minima along that particular line. And then we can see that if we find d 2 f d alpha square, this will be v transpose Av we differentiate it. So, this is 0 this is v transpose Av and this is 0, v transpose Av. And this is always greater than 0 as A is positive definite. We started with A is a positive definite matrix. So, we will always get at this look this point f is minimal. This is always a minima of f.

So now we know the distance that we should move to reach the J minima. And then in next session we will see how can we implement in a code. Or how can we finally, write an algorithm which will solve b minus Ax, considering it to be a problem of finding J minima.

Thank you.