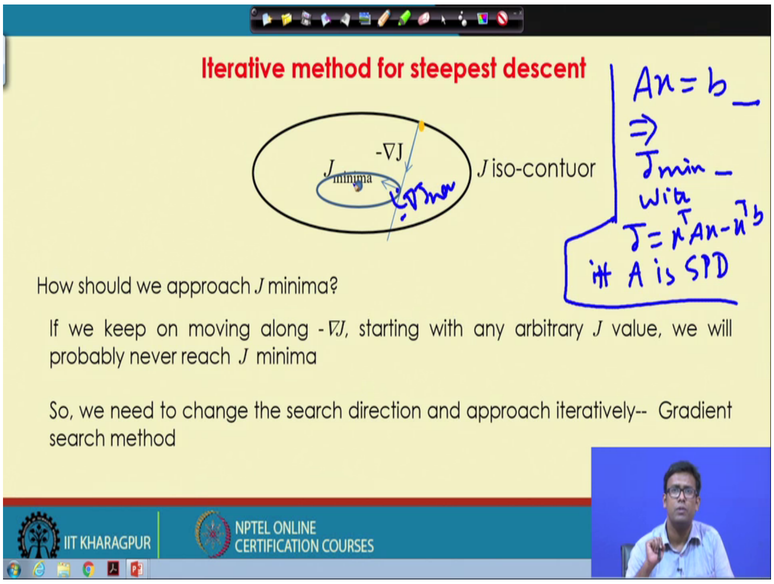
Matrix Solvers
Prof. Somnath Roy
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

Lecture – 37
Steepest Descent Method: Gradient Search

Welcome. So, a discussing about the Steepest Descent Method.

(Refer Slide Time: 00:22)



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 and provided A is if and only if s symmetric positive definite matrix.

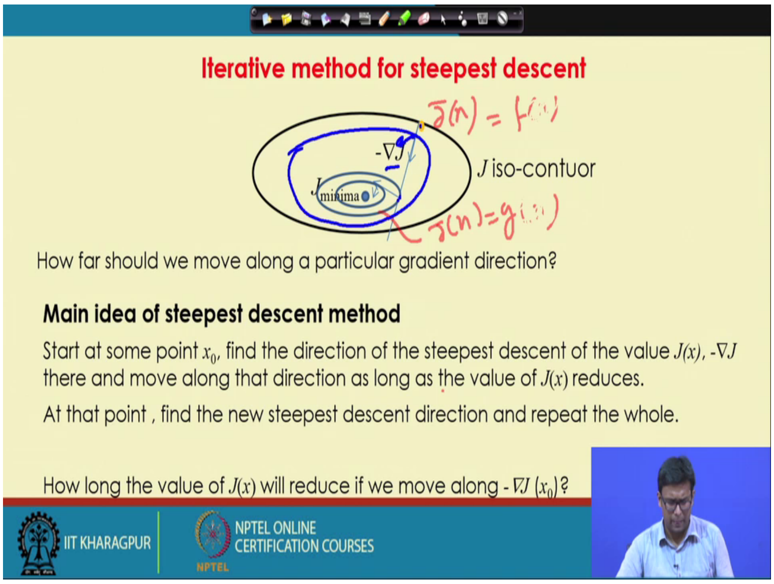
So, here in solving a matrix equation is substituted by finding out minima of a particular functional. Now why we are doing this? Because we have discussed that the direct iterative methods like Gauss Seidel or Jacobean certain limitations when to come out of those methods and explore the other ways of iteratively solving x is equal to b. So, what you started here is that solving x is equal to b is same as if a is as SPD Symmetric Positive Definite matrix solving x 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 minima 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 jx; 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 you start in 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 cen 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 will draw the J iso contour again. This another iso contour and go drop a minus grad J. So, this is a again a minus grad J nu at this particular location and move along that you know we will keep on changing the search directions and finally, approach the minima of J. So, we look into the detail of this method.

(Refer Slide Time: 02:51)



Now, we started with one particular value one particular J iso contour and there is a J minima moved along one along the minus grad J reached somewhere and then got the nu minus grad J. And move along that and then again he reached another iso contour and then move along minus grad J 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 contoured. 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 search direction and then J will keep on increasing. So, if we move along this will never did J minima this is an probably nearest point to J minima. In we go here and somewhere we have to stop otherwise will means this line will go away again we will come back to if we if we keep on going this if we keep on going this again I will come back to the over 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. That start at some point x 0 find the direction of the steepest descent at this value at that point of steepest descent of Jx, 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 mover in a direction where J is reducing when J will stop producing you change the side as search direction at that point.

When J is 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 well how long it is reducing when J will stop producing because anyway J is bound to reduce if minus grad J if grad J is non0 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 you know 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 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.

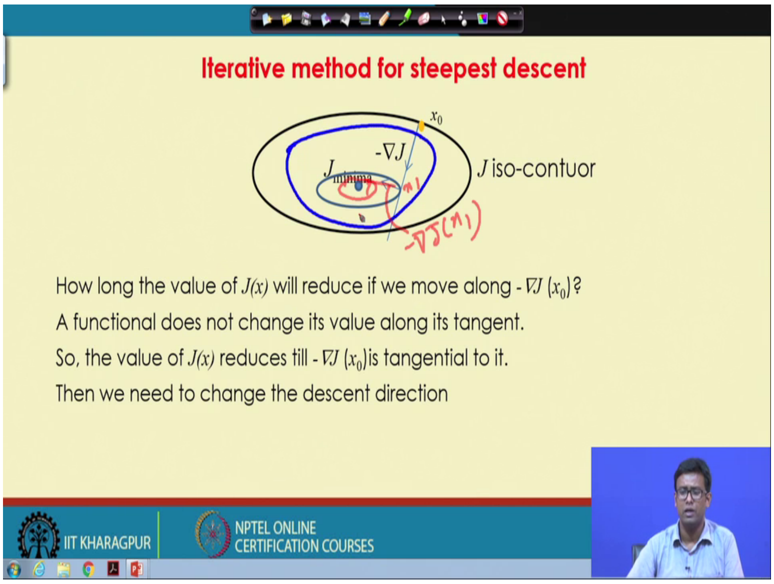
Now, grad g is not same here. So, if I move along this line this is this is the direction which is any way perpendicular to this particular iso contour and this is the direction in which J is reducing first test. When I came here J when I just move away from this line minus grad J at that point is different. So, J is not reducing fast testing that line anymore only in this point J was reducing first test.

Now, J is if the reduction rate of reduction of J will slow down, because minus grad J is different if I go 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 first step 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 gate 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 is 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 pal 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 reduced the same question.

(Refer Slide Time: 08:16)



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 fast steepest way, but still it is reducing, but here it is not reducing any far that. So, you have to find out how long value of J x will reduce. Then they are up to then because I started with one particular value of J I am trying to reach a minima. And this is an iterative method I started with some guess 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 minimum value of J? Because if a is symmetric positive definite matrix in the way I J define I define J? J minimize the solution of x 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 reduce reducing till I reach J minima ideally in if I can go in any direction I should reach after I reach a minima value will increase.

If value is if value of J is increasing I will never reach J minima. I am going in a certain I started with on 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 reach minima I will reach the least value I will see if value of J is reducing I have probably we 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, you 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 is value along it is tangent. So, when I will go along this direction as soon as the grad J becomes tangential to 1 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.

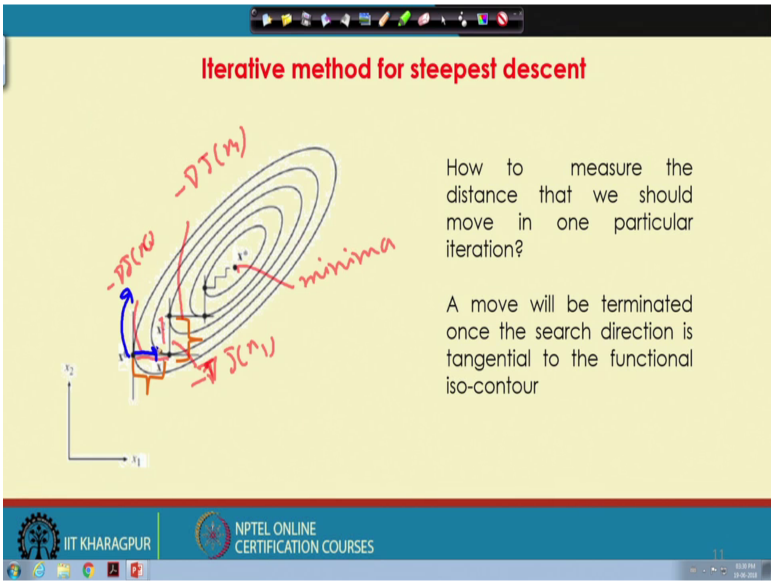
So, x is a multivariable function. I started with one particular value of x I can think of a 2 2 d functionalities and ellipse if they are elliptical conclude x 0 square plus y 0 square is is equal to constant, where f; f of x is equal to x 0 x square plus y square.

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 a 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 differently like. 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 where 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 long 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: 13:47)



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 gradient of J x 1 this direction this direction is minus gradient of J x 2 minus, you come to x 3 and change the direction and so on you approach the minimum value this is the minima.

So, starting from any arbitrary value this will take us to the minimum 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 to variable functions we can extend it maximum to the third r 3, but this geometric concept is applicable for a higher order and Rn things.

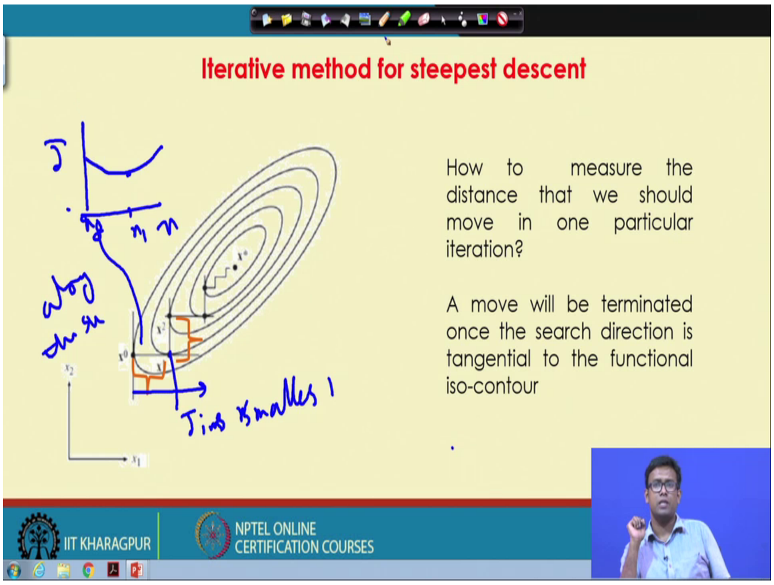
However, once we finish this discussion and propose the final algorithm from strips or steepest descent, we should check that algorithm why is that converge or geometrically we can see that starting from any x 0 y 0 it will take me to the minima. In Rn the algorithm also should converge that is starting 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 this 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 we will start search with a particular size 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 will 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 1 particular iso contour there will stop. After that J will keep on increasing.

So, I have to start with a search direction and check when this 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 nu J x will be 0, the these that a these the nu, nu search direction and all search directions are perpendicular to each other themselves assume.

(Refer Slide Time: 17:51)

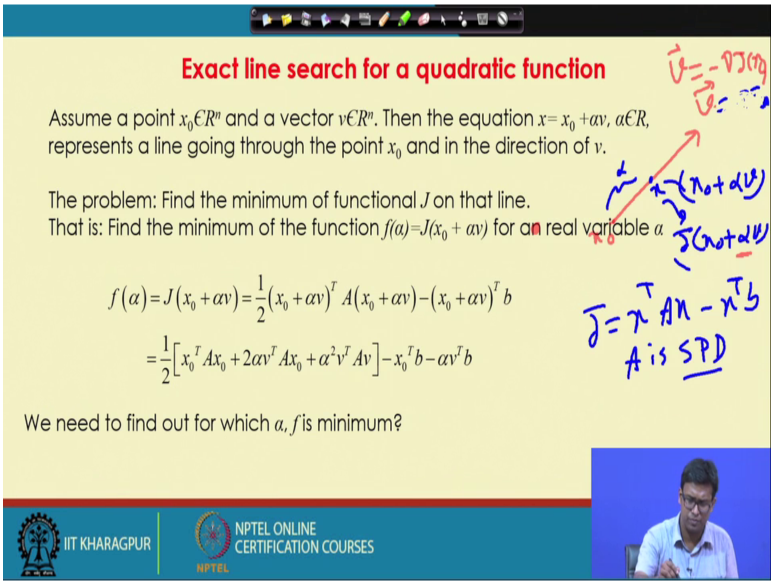


And also another factor is that, if we go along this particular line if we go along this particular line locally J is. So, if I smallest not say minima or if I draw say J versus x along this line along this line.

So, this is x 0 and this is x 1. I will see J is locally mini local minima is along this 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 saturation where J is minimal I will change my saturation.

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 control of J. So, this is the local minima of J 2.

(Refer Slide Time: 19:29)



So, I have to find out the local minima of a function along a particular line. So, 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 a real number coordinate of all dimension in rl and a vector v on that same space Rn. And then the equation xz 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 way 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 on 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 that tells me what is the direction from which I have started v is fixed.

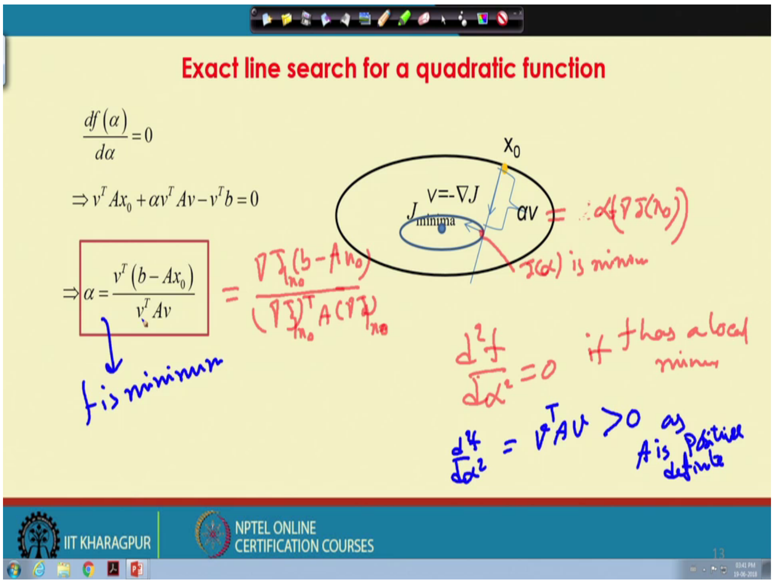
So, the only variable is alpha. So, depending of the value of alpha I should get J minimum somewhere in the this line. The local J minima will be where for one particular alpha where del J del alpha is equal to 0. That is find the minima of the function f of alpha J which is J of x 0 plus alpha v as 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 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 b sorry A x 0 plus alpha v transpose into a into x 0 plus alpha b minus x 0 plus alpha v transpose into 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 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 df d alpha is equal to 0.

(Refer Slide Time: 24:03)



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 f v and this is v transpose minus v transpose v.

So, this way 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 directions 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 I will find to 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 ax 0 by gradient of J transpose a into gradient of J or rather this is that evaluated rate x 0 this is evaluated an x 0 this is evaluated an 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 them minima. Again because we are saying that J has a local minima here it is also important to show that d 2 fd 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 equal 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.