Numerical Methods and Programming

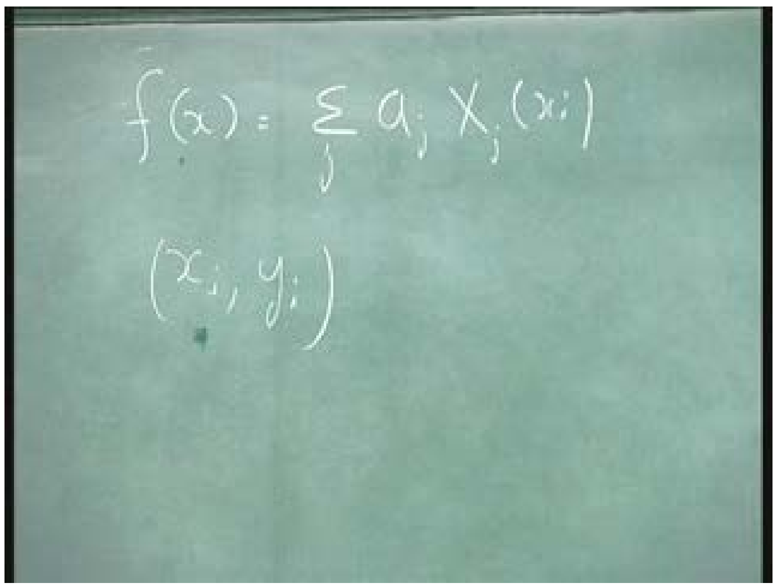
P. B. Sunil Kumar
Department of Physics
Indian Institute of Technology, Madras

Lecture - 15

Data Fitting: Non-linear fit

The last class, we were looking into modeling a data using a set of non linear functions but not having any adjustable parameters that is a set of fixed non-linear functions and we wrote down the function as of this form, that is sigma j a_j x_j of x_i.

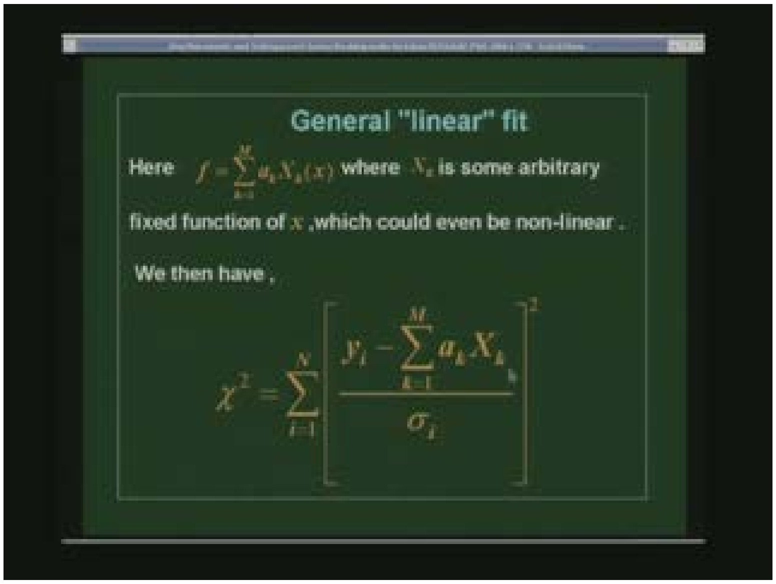
(Refer Slide Time: 02:30)



So this x_j of x_i is could be linear or non-linear functions but fixed functions in the sense there are no adjustable parameters and we have our idea was to model this, data the given data by adjusting this parameter a_j. So I will find the a_j values such that, this is a best fit to a given set of data points or in other words, we have to choose a_j such that all given data points would pass through the curve represented by this within a distance within a distance plus or minus delta y, so we had the other given data points as x_i and y_i. So that was our data points and this is the function which you fit through that.

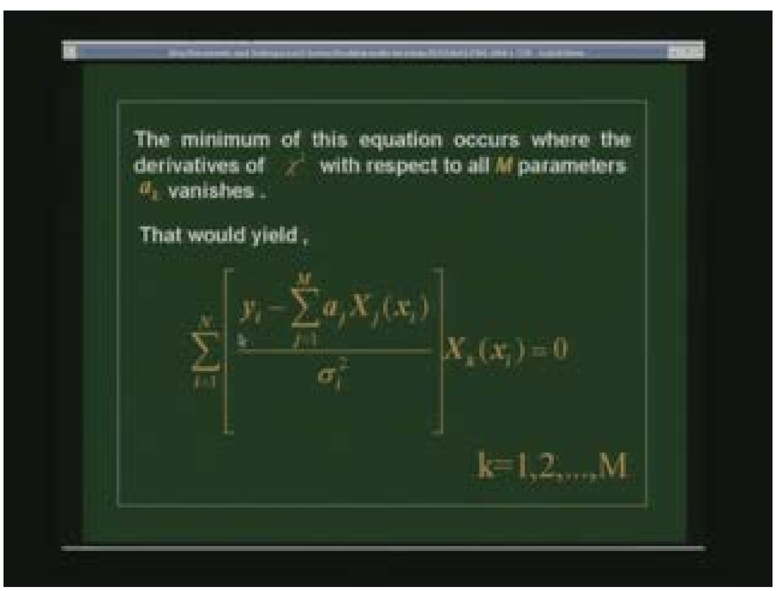
So that is, and then for that purpose we would again be define a function called chi squared and which is the difference between this function a_k x_k, x_k is the function of x, x_i's and then, these are the adjustable parameters and sigma i, you may remember is the reliability of the data x_i and y_i, for this error in the y_i for a given x_i y_i. So now given this set of this function and then we could determine this a_i value a_k values by minimizing chi squared with respect to a_k, so that is what we were doing.

(Refer Slide Time: 02:34)



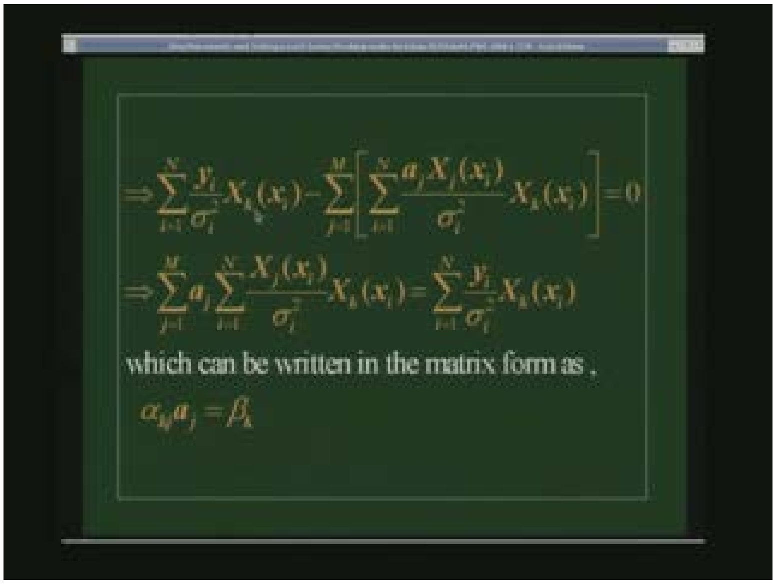
Okay when we write down, when we minimize psi squared with respect to a_k, so we would get an equation of this form a linear equation of this form linear in a_j and which is y_i minus a_j x_j summed over j divided by sigma i squared multiplied by x_k is equal to 0 and while we get, if there are m such parameters that is, this is the sum of m functions and then the m parameters, j goes from 1 to m and then we will have m such equations which has to be solved simultaneously to get the 5 functions together the values of a.

(Refer Slide Time: 03:16)



And it will not serve that now, it is not so difficult because it is a linear function and so we can write y_i x_i by sigma by i squared is equal to sigma j a_j x_j of x_i. So that is what we be would did in the last class

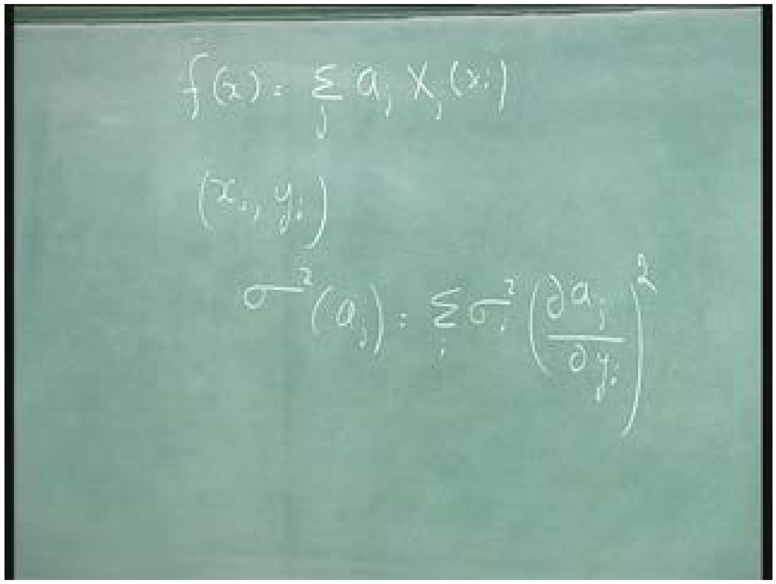
(Refer Slide Time: 04:12)



So we wrote by y_i by sigma i squared and x_k x_j going from 1 to M sigma i going from 1 to N, a_j x_j of x_j sigma i squared into x_k of x_i and now, I can write this, rearrange this terms and write it into this forms and then, you would see that this is same as this and then you see that, you have a set of coefficients multiplied by this function here that is i summed over i that is for all data points x_j x_k by sigma i squared into a_j right x_j x_k, i squared sigma i squared into a_j is equal to y_i by x_k by sigma i squared or and you can see that, this is a linear equation, we can write that in the metrics form and then you would get as alpha k_j a_j is equal to beta k. Your alpha k_j is now, x_j x_k by sigma i squared summed to a_i going from 1 to N which is the number of data points that you have and beta k is y_i sigma i squared into x_k of x_i.

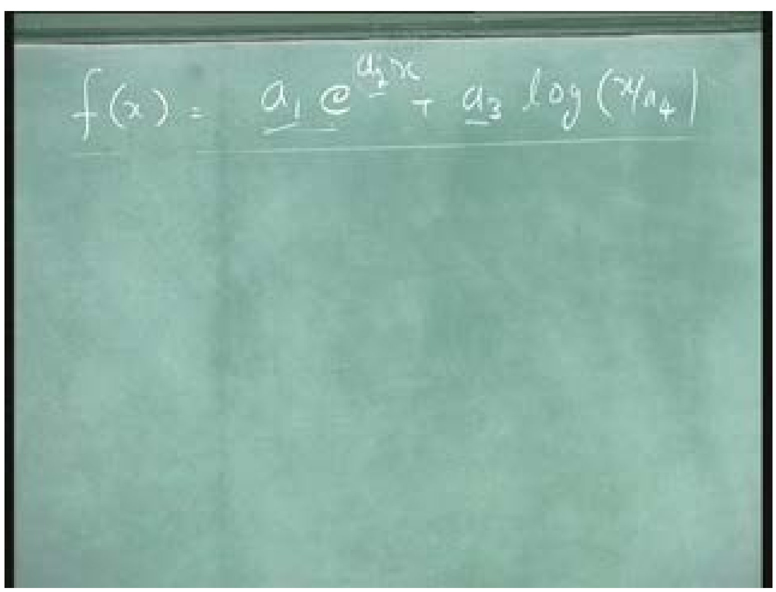
So beta k is a function, so y_i while alpha k_j is our only functions of x_i's, okay that is what we had and we could solve this equation and obtain a_j and we state that, we saw a program which does that for when there are two adjustable parameters in the general case, where there are M adjustable parameters this would be an M by M metrics and this will be column vector of size m and this will be column vector of size m that is what we have seen, and we also saw that, we could determine the errors in the a_i by defining sigma, by defining a sigma is for this functions that is, we could determine sigma is a_i as this form that is, we could write sigma squared a_j as sigma i, this is all we could write that down and since, we know the a_j is in terms of the y_i.

(Refer Slide Time: 06:46)



We can determine all n this functions also that is, what we have seen in the last class and you wanted to extend this model to a extend this discussion into now general non-linear models that is now we go from functions of this form where we had fixed non-linear functions multiplied by coefficients we could go into a non-linear function itself that means now you would want to model functions of the form f of x is equal to let us say I am just writing an example as a_1 e to the power of a $_2$ x or and then plus a_3 log of x by a_4 etcetera.

(Refer Slide Time: 08:51)

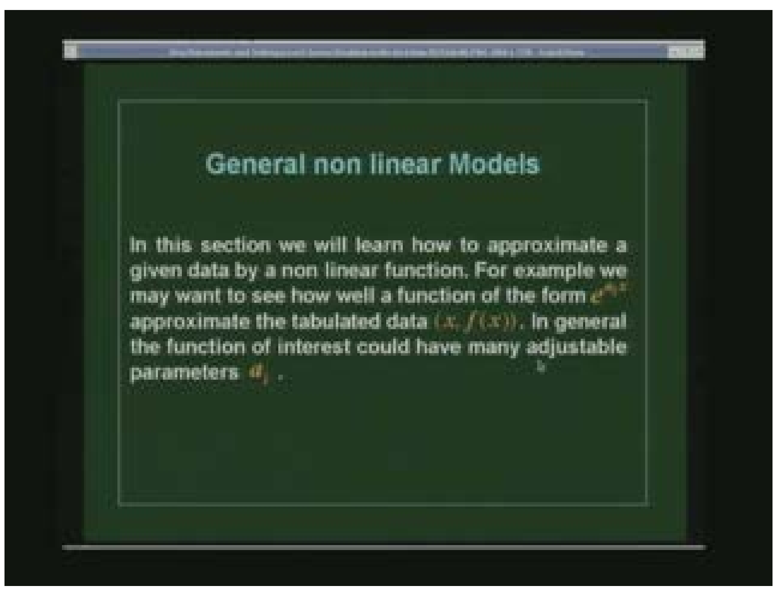


So I could write some function of this form and I want to fit in this function into a data. So important thing to note is, that in this particular case or because this coefficients for example, this is now appearing as multiplicative factor to this function while this is sitting in the exponential of this quantity and, so these quantities could have completely different scale, so different dimensions okay.

So that is will have some application in the analysis which you are going to do okay this is a_1 and a_2 or in general this parameters could have completely different dimensions okay, a_1, a_2, a_3, a_4 or we can have completely different dimensions. So now we want to, what we our aim would be to determine this adjustable parameters $a_{1,}$ $a_{3,}$ a_4 etcetera. Okay for, such that this is the best representation for a given, for best function, for a given data points or again to repeat the whole thing what I was saying before that is we have to determine a_1, a_2, a_3, a_4 such that all of the set of data points for any value of x would pass or would go through the function f of x or be near to the function f of x, so that is what we want to determine.

So again, we will follow the same procedure as before, we would determine a psi squared and we minimize the chi squared with respect to a_1, a_2, a_3, a_4 that is what the general procedure would be, the general procedure would be to again write down the function, the chi squared function and minimizes this functions, this chi squared with respect to a_1, a_2, a_3, a_4. So we will see how you proceed with this.

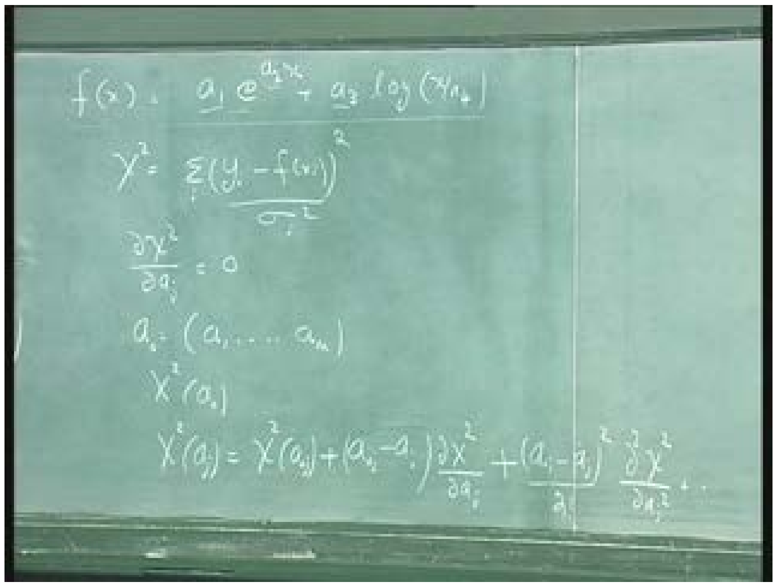
(Refer Slide Time: 08:52)



So now, so now to summaries the sets in our, we will look at how to fit in a non-linear function for given set of data points and given set of tabulated data by a non-linear function ,for example e to the power of a_1 x, a_2 x etcetera. Okay, so now, the procedure we do for this is as same before.

So we could determine chi squared as sigma i, y_i minus f of x_i the whole squared by sigma i squared that is what we have been doing and then what we would want is now, this f of x_i now as all the adjustable parameters, so we would want to determine del psi squared by del a_j equal to 0 and use the set of equations to determine.

(Refer Slide Time: 13:48)



So we would, now in general it turns out, it turns out that because this function is now non-linear in the parameters a_i with the set of equations we get from this, okay are also non-linear, so now we have a set of non linear coupled equations, so now that will become much more difficult to solve than set of linear equations which we had here. So we will follow a slightly different procedure from what we have used in here or to solve this equation. So that is to do a iterative approach.

So you will try to get the values of a in this functions, okay by an iterative method that is we will come with a set of guess values for a's and then, we will iterate around that values that is, you say that we will assume we have a set of a_i okay such that psi square is minimum for that set. Okay, let us say we call them as a, a so now, "a" is vector which contains a_1, a_2 up to a_m. Let us say or we will come up with set of guess values for all this parameters and we will write down chi squared of "a". Okay let me call that a_0 and we will assume that this the guess value that is the chi squared we obtained is kind of is close to it is minimum value.

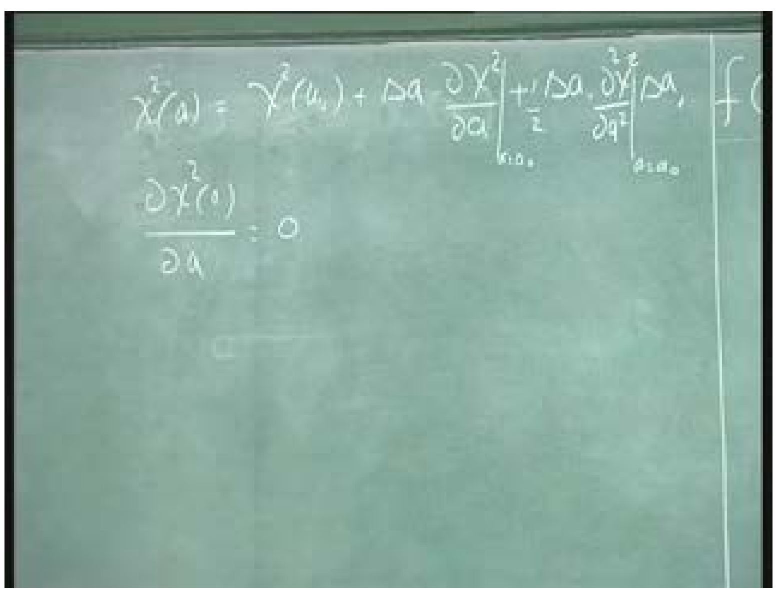
So now remember, chi square is now a function of "a" because f of x is a function of "a", so now what our idea is to minimize chi square with respect to "a". So what we are saying is that, we assume we make a gas for "a" values, we will assume that this gas is quite good that is what the for the set of values for a as assumed psi squared is quite close to it is minimum or del psi square by del a_j are close to 0 that is the derivatives of chi squared with respect to a's are very small.

In that kind of scenario, we can we can expand chi squared around that value of a_0 or we say that chi squared of "a" that is for some value around a_0 can be expanded as chi squared of a_0 plus del. Okay, I will write it for each of this functions, okay a $_{0j}$, okay for each of these elements.

I could write this as, etcetera. So you could, I could still expanding it for one parameter. I can just expand it in this fashion that is I am assuming that I have the chi squared values such that, the a_0 values I have guessed here. Okay such that, so this be a_{01}, a_{02} for a 0 being my guess value. Okay a such that, it is very close to its minimum, that is its derivatives are very close to 0's and in that case I could expand it in this fashion in a Taylor series. So I just do a Taylor series expansion or in general case for all the a's, I could write this expansion in the following form.

Okay, I would write this expansion as chi squared of a, a is now a vector, okay you write as chi squared of a_0 or plus delta a which is also a vector del psi squared by del a and plus etcetera. So I could write it in this fashion, so basically what I am trying to say is that if, if I say that my guess is good, okay then, if that is close to, if that is psi squared is close to minimum then I can taylor expand this function okay and then get a new value of psi square and I am going to say that this chi squared value, okay is my minimum is, my new minimum or I will taylor expand by function chi squared function around this a_0, okay such that I go into the minimum, I demand, I go into the minimum that is, this is the new minimum and I will determine my delta a values that is a_0 minus a of the values such that, this is the minimum.

(Refer Slide Time: 02:30)



I can do, I can always do that and then, if this is my initial guesses are good and that would mean that del chi squared the new a divided by the del a which is the same as del chi squared by delta a is equal to 0. So remember, all this values are evaluated at a equal to a_0, all the derivatives are evaluated at a equal to a_0.

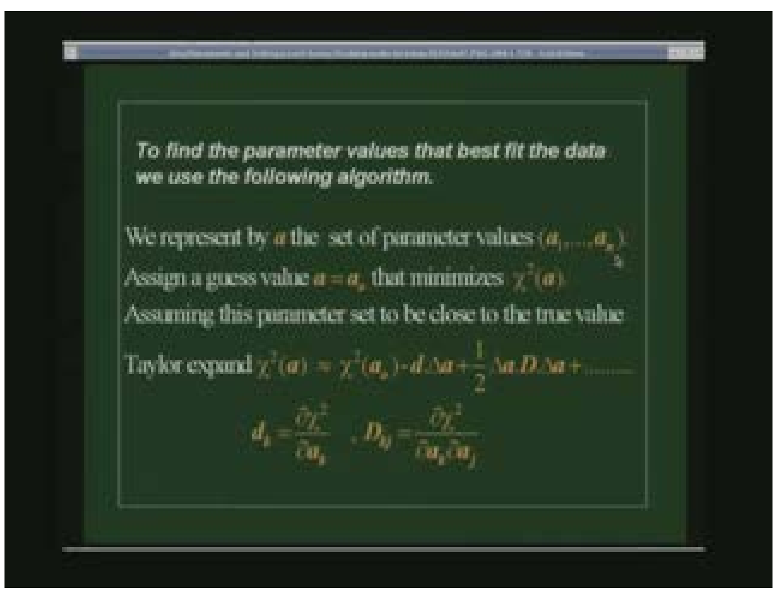
So del chi square by del a would simply mean del chi square by del a equal to a $_0$ plus delta a into del chi square, by so that is a_1 over 2. Okay del psi squared by del a squared, so that is what I would get from this. So to repeat what I am trying to say is that I make instead of solving this equations directly, I go through a iterative approach, Okay that is I make some initial guesses for my a's and I would expand my psi squared function around that initial guess and then I would write a expression like this, it is a taylor expanded okay and then I can say that my new function which I obtained. Okay the new psi squared value I obtained okay, that is around psi squared a, not ok is a minimum that is del psi squared by del a is then is equal to 0.

Okay, so that will from this I can differentiate this right hand side by a equal to 0 so that gives me an equation del chi square by del a plus delta a into del psi square by del a squared del square chi squared del squared chi squared by del a squared and that that I can solve that equations for my delta a's and thus determine the new a's.

So once I determine the new a's, I go back and again check that it is actually a minimum. Okay, when it is not a minimum, I will assume that as my new guess value, okay then I will take that as my new guess value and then continue this whole procedure. So that is what the method we will be following.

So, let me summarize that here. Okay so, we represent by a set of parameters this is a vector a here, this is a set of parameters a_1, a_2 up to a_n all right and then, we see that my guess value is a equal to a_0 and I assume that, this guess value minimizes psi squared a okay.

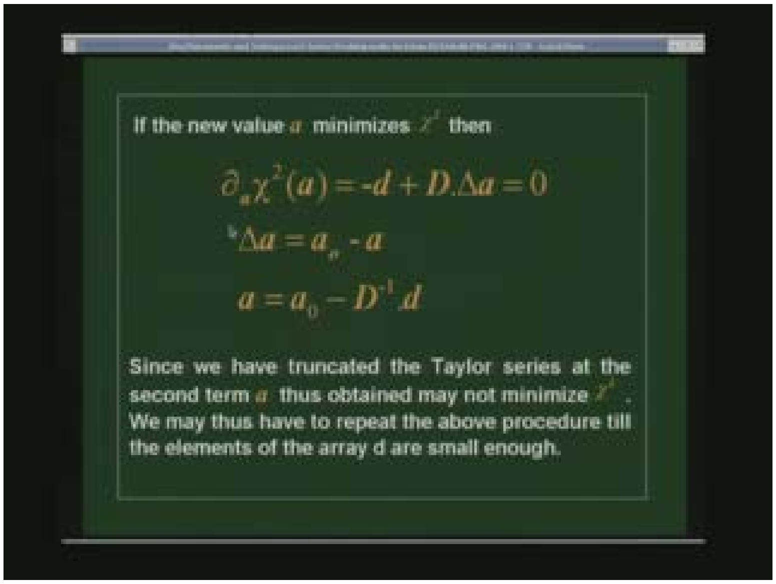
(Refer Slide Time: 17:13)



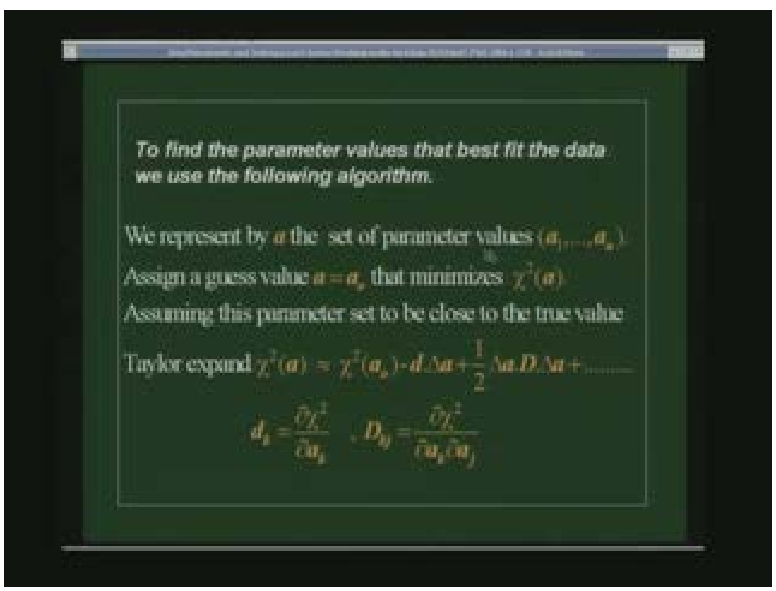
So now, of course this is a guess. So it will not be, it may not be true. Okay, in that case I say if it is not actually true and then I can expand my chi squared function around that a_0

value okay. So I could write it as del chi squared chi squared a_0 delta a delta D del squared, this is now, now you can see here this will be the first derivative, this d here will be the first derivative of chi squared with respect to whatever a k functions we are using. So then this D would be the second derivative of chi squared that is, what I have written on the board that is, the expansion of the Taylor expansion of chi squared around the value a_0 which is my guess function.

(Refer Slide Time: 18:30)

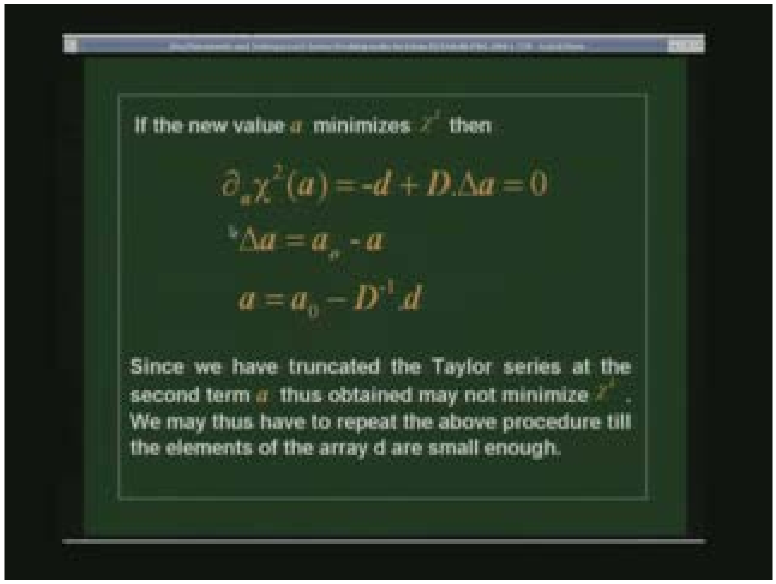


(Refer Slide Time: 18:40)

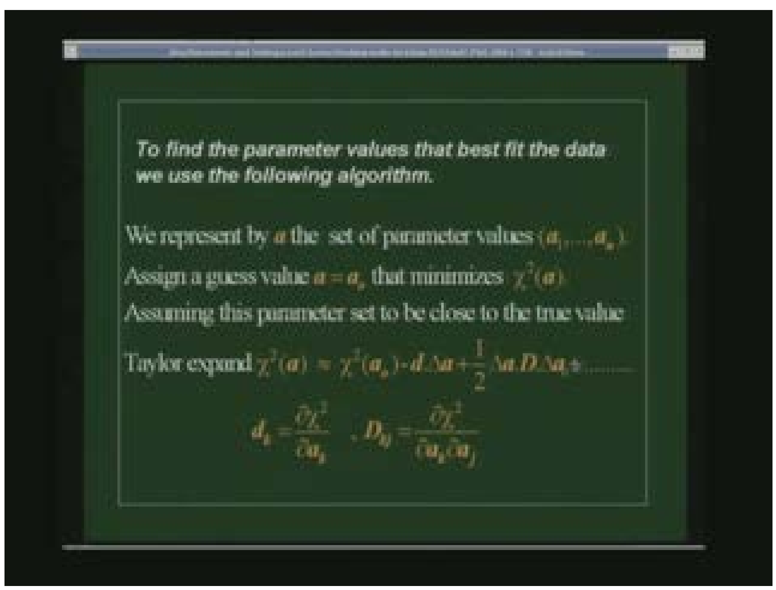


So I have this metrics elements dk_j and the element dk, now what I am going to say is that this expansion gives me a new chi squared value and which is now the minimum okay, that would imply that del psi squared by del a should be 0 because, my new chi squared is the minimum that is, it should be 0 and that would, if from this, by differentiating this I would get that as minus d plus D dot del a is equal to 0 okay. So I have written that as delta a is a_0 minus a and a.

(Refer Slide Time: 18:44)

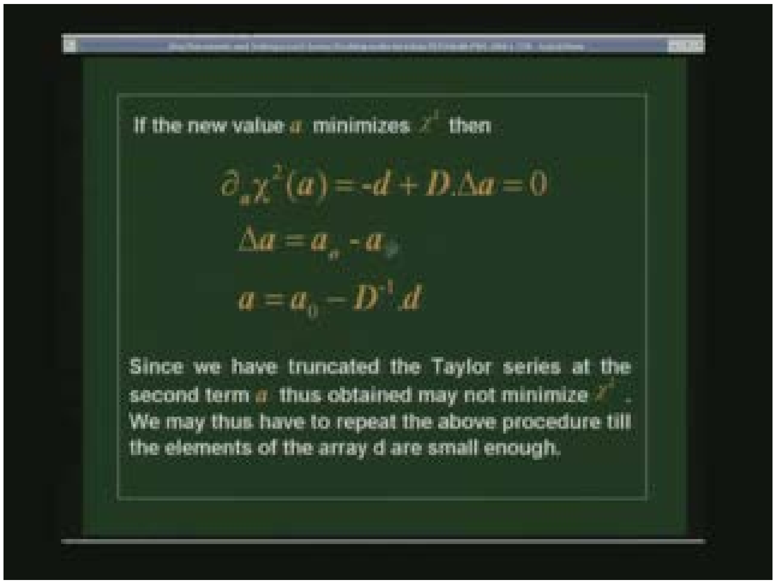


(Refer Slide Time: 19:06)



So that means that now, from this I can solve this equation. Now, you can see that we are back to our discussions in the last class that is for the general non-linear functions, so we have a set of linear equations now, so of course this taylor expansion here is terminated at the order 2 that is why we have a linear equation and if you do not terminate it at order 2, we will not have a non-linear equation but, our idea is to get a linear equation.

(Refer Slide Time: 19:23)

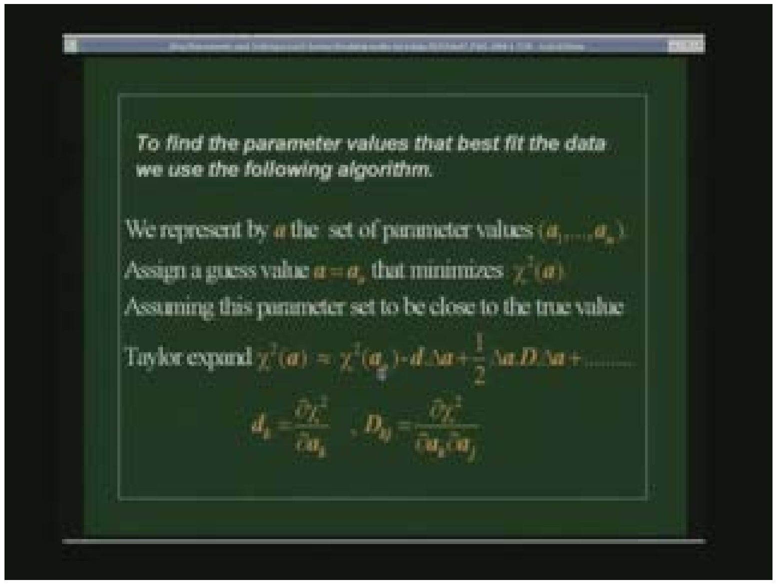


So, we terminate it at order 2. And because, we terminate it at order 2 and this the resulting linear equations when you solve you can get the delta a's and then once we have the delta a's, we can always get the a values because we know the a_0 values and from that we can obtain the a values.

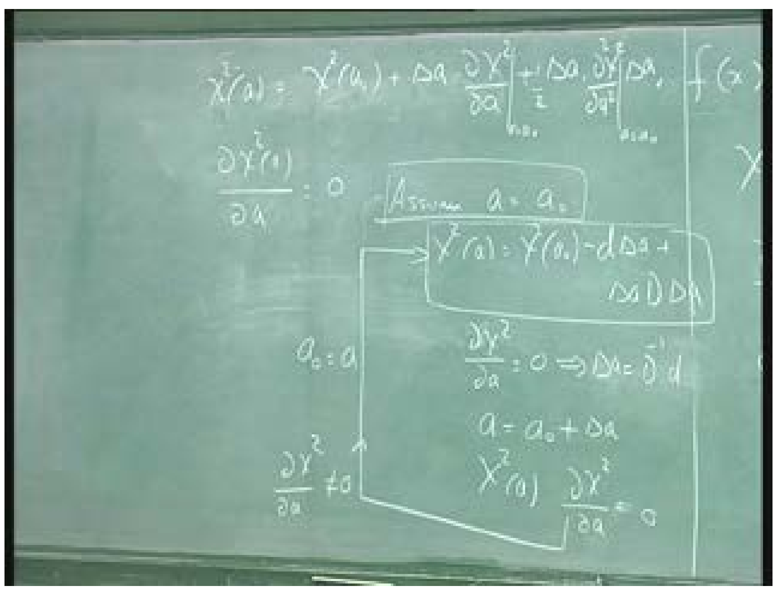
So once you determine in short, when you are determining the metrics d and a, we can get a new set of a values. So now, because we have made the approximation already that is we have terminated it order 2. So what you get may not be the true minimum again, so now if it is not the true minimum then, what we do this is go back to this equation again now, replace the new a by what we got as the new a as a_0 and continue with this equation again.

So what the changes would be that we have to again determine this d and this little d here and the capital d that is this metrics element if you would again determine at the new value. As I said here, this derivatives when we take okay there is a chi squared by del a_j's and del squared chi squared by del a_j square, this determine a this, these derivatives are determinant a equal to a_0 okay. So in short what will be the summary is that we have, we have assumed, we assume a's okay, a equal to a_0 right. We compute chi squared a as expansion around chi squared of a_0 plus we wrote it as minus d times delta a plus delta a D delta a etcetera okay.

(Refer Slide Time: 20:02)



(Refer Slide Time: 23:45)

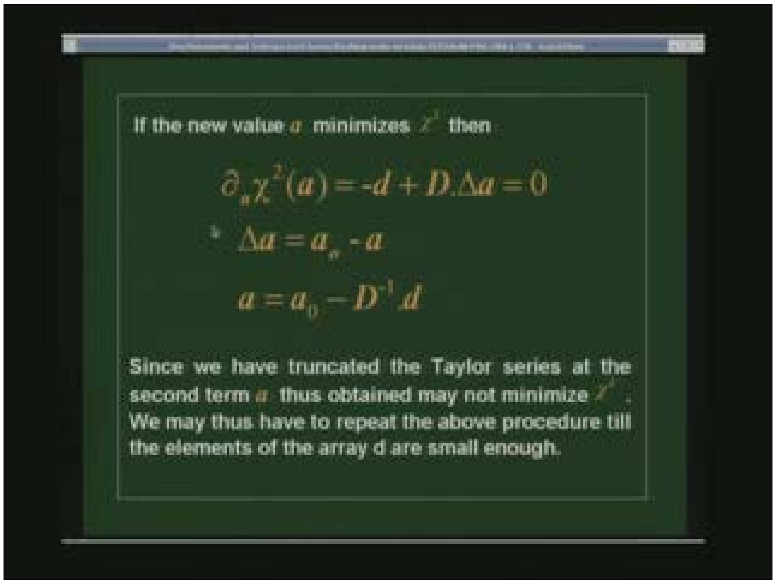


For these elements d and capital D were given by, so the elements of d towers d_k d_j is del chi square by del a_j and elements d_{jk} was del chi square by, we can expand that so that, so that is part number 2, okay and then, so first step and then we had this step in which we expanded this function, so that determines d and capital D as little d and then we can write delta a. So from this say del chi squared by del a equal to 0, so that would give us a delta a is equal to we get as minus d inverse capital D delta a as D inverse into d that is what we would get, and then from that we can determine a.

So that, so we will get new a's okay, some new a as some a not plus delta a so that is what we go and now this new a, if this does not minimize again we will determine chi squared a and we see del psi squared by, so from this, we will determine chi squared a and del chi squared by del a. Now if this is not 0 that is basically since this is same as this element d, okay if this is not 0 then we will go back from here, into this by replacing a not equal to a, this values of "a" will replace into a not we will go back here and then again compute not here we will again compute this quantity and now all the determinant all the elements now, that is the derivatives as to be now the elements of D and little d has to determine at the new value a, and then we continue this iterative procedure till the elements till this goes to 0, okay that is what we want to do.

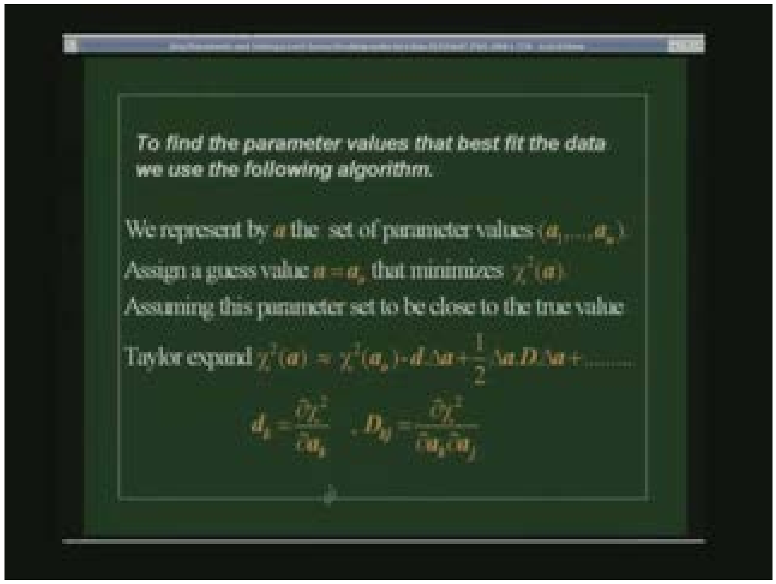
So we want to goes, we want to reach till this is 0 if, if we go back here if del psi squared by del a is not equal to 0, then we will go back here ok in this path again here and if it is 0 we stop the problem there.

(Refer Slide Time: 23:46)

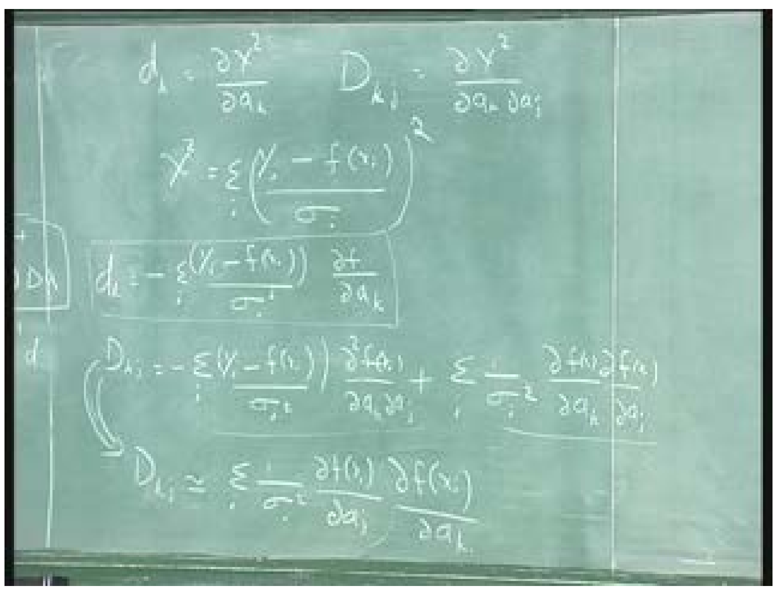


So that is the, that is the iterative approach which we are going to do. So that this summarize again, so we had del psi square and we could determine the delta a from that and from there, we could determine a as a not plus delta a which is given this and then we would repeat this procedure till we get this elements d which are actually the which are actually the first derivatives of del psi square, okay we are going to 0 because if, chi squared is actually a minimum, then the first derivative should go to 0 that means the metrics elements d should go to 0 that is, what we want to determine. So now the point is, that we have to determine here del psi square by del a's or this functions. Okay, so that a now we saw D and little d as first derivatives and the second derivatives. So let me write that down again.

(Refer Slide Time: 24:02)



(Refer Slide Time: 29:04)



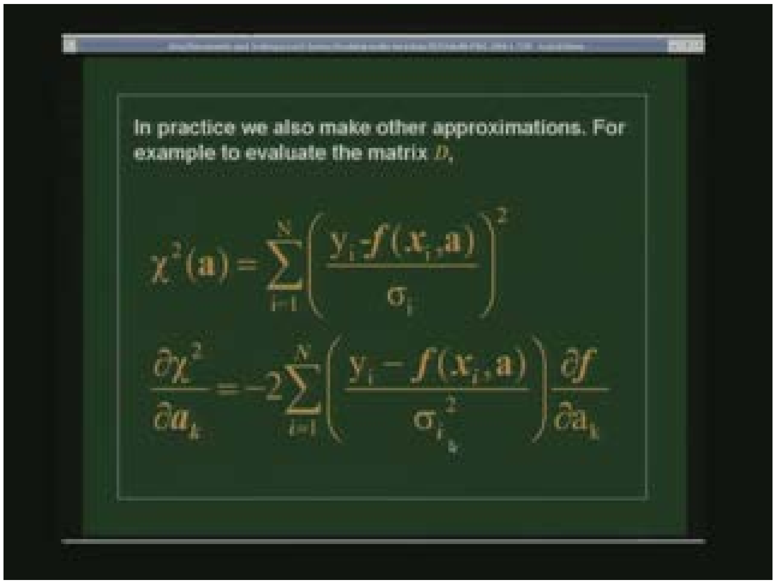
So in this whole scheme, we had d_k as del chi squared by del a_k and we have these D_{kj} as del chi squared by del a_k del a_j and, we had psi squared as y_i minus f of x_i divided by sigma i the whole squared sigma i, that is what we have. Okay, so now we can determine now f so that would mean that d_k's, y_i minus f of x_i divided by sigma i squared sum over i into minus into del f by del a_k with a minus sign that is what we get from this right and d_{kj}'s now are the derivatives of this. So the two terms in this minus sigma i y_i minus f of x_i divided by sigma i squared into del squared f by del a_k del a_j that is one term and then we would have plus 1 over okay, sum over i 1 over sigma i squared del f by del a_k del f by del a_j okay.

So we had these two functions .Okay, so now we have the metrics elements are given by this, we know that f as the function of x_i, so f are functions of x_i's and a_i this, we know that, okay that is what we have. So now, any way we are going to do, what we are doing is an approximate scheme in the sense. We have terminated this expansion in the second term or in the second order term.

So and we are doing a iterative procedure, so we could make further approximations to simplify things because, when we are doing a scheme an approximate scheme we could also approximate this function by only this term because, numerically it is difficult to determine it is there will be more errors in the second order derivative schemes derivatives okay.

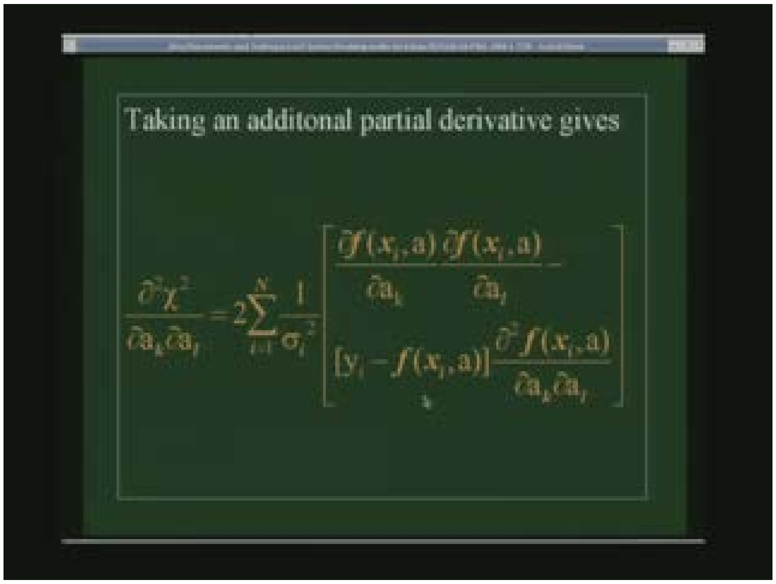
So instead of using this second order derivative of the function at every time you would just compute the first order derivative and assume that, D_{kj} is, we will approximate it by, we will just approximate this function by saying that it is sigma i 1 over sigma i squared del f of x_i divided by del a_j, okay that is what we would do. So then, we will use this an approximation okay, so we will now we are not going to use the full second derivative instead of that we will just use this derivative this has the approximation to this whole second derivative and the first derivative is given by this that is what we are going to do. I guess basically, we have to only determine the first derivative and the product of the first derivative would determine, we will determine the metrics elements for the capital D_{kj} and this little d_k will be determined by the first derivative multiplied by the Y minus f of x_i by sigma i squared.

(Refer Slide Time: 29:06)

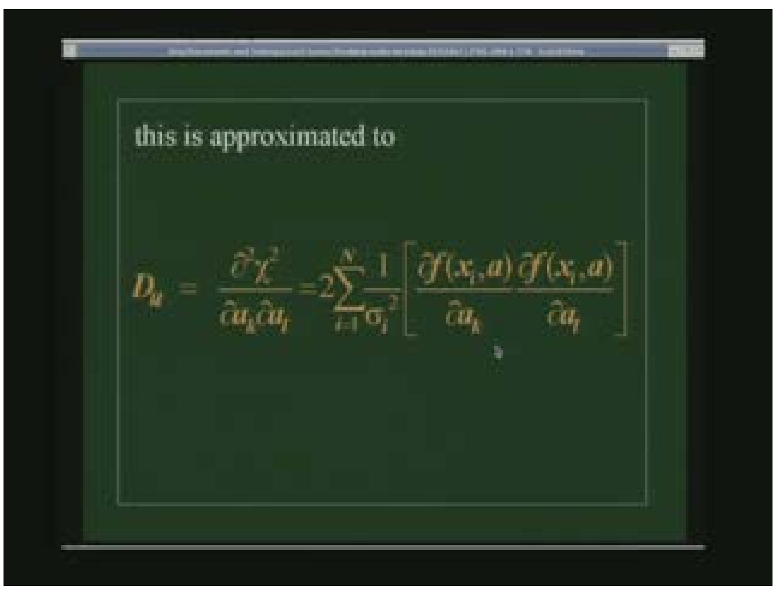


So that simplifies the calculation further ok and this is justified because any way, our idea is to finally get this to 0 right, so that is our aim and we are following a iterative scheme ok and we would see that in a implementation that the errors introduced by the second derivative will be much larger than the approximation we make by doing this. Final answer, of course will not depend on this approximation because the final answer is correct provided this goes to 0. So that is what the scheme would be and that is what we are going to do, so we have chi squared here and we have the first derivative of del psi squared by del a k as this okay.

(Refer Slide Time: 29:13)

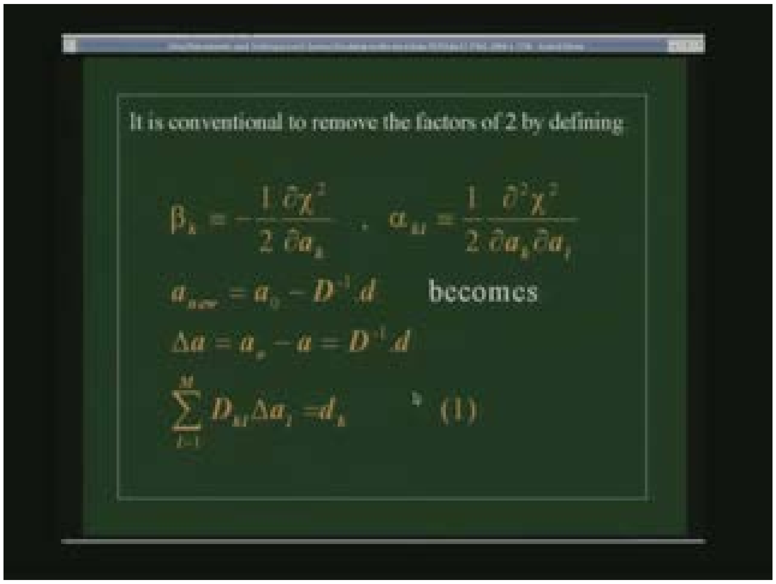


(Refer Slide Time: 29:23)



And the second derivative as I said just now there are two terms so one is the product of the first derivative and second one is the second derivative of this function so what we do is we will make an assumption that the second derivative is only given by this, this function that is del square f by del a_k del a_j term we will throw away. So now the it is that the this metrics d_{kl} is approximate as a product of this first derivatives so that is the summary of the procedure ok so on once we have that and then now we have the metrics elements, this is d_{kl} and and the little d and then we have this metrics elements we make them we will write them as 1 by 2and 1 by 2 here just to absorb that factors of two into this equations, okay and then we can write it in a simple metrics form as a new as a 0 minus D inverse d that is what we will have to do.

(Refer Slide Time: 29:51)

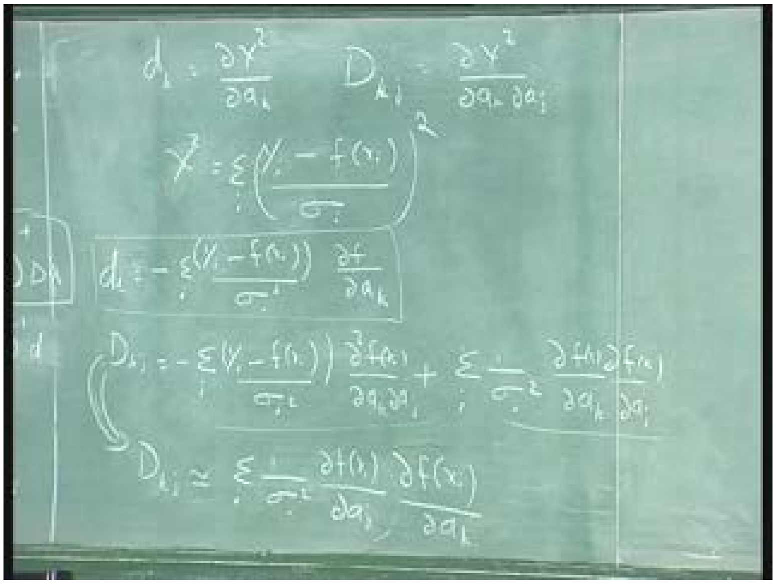


So once we have this a new as a $_0$ minus D inverse into d, we can solve this equation and we will get back, we will once we solve this equations, we have the new chi square a value and we can determine the chi squared value. So, now in an implementation that is that is what all it is, ok it is very simple so we determine the once we have function form we can determine the chi square that means we can determine d_k's and d_{kj}'s, ok and then write it into this equation and determine the new a values that is the "a" new value would be given by "a" old value plus the delta "a" which we determine as the as D inverse d and then you would assume that we will follow this iterative procedure we would finally go into and dk equal to 0.

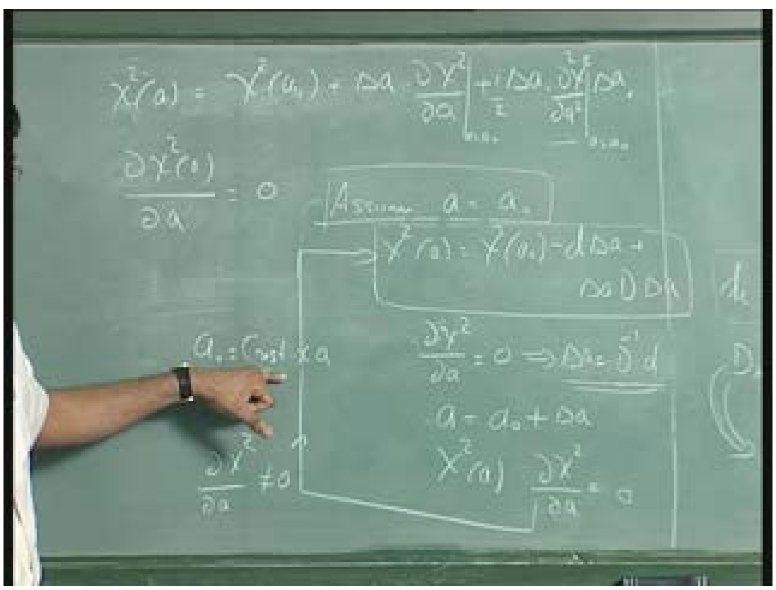
But in practice, we will see that this does not happen all the time and because we would have an initial guess our initial guess, if this is really far away from the real value, the our iteration will not curve converge it could go into in fact away from the minimum because we really do not know the chi squared behaves as the function of "a" in the path of "a" we are following.

So we may not, may be our initial guesses may not be close enough to the real minima in that case, the minima of chi squared. So in that case, we may not go into that minima by this procedure, so now let us look at what we are doing, we are doing as to get the new delta a's, we are using the second derivative or that is what is called as the hcm of the function, okay that is what we are using.

(Refer Slide Time: 30:26)



(Refer Slide Time: 35:12)



So to get the new delta a, we are using the second derivative of metrics, okay we are using the second derivative of metrics, that is what we do it to get to the new function, new values of a. So we could which we have approximated as the product of the first derivatives but when it goes away from the equilibrium from the minimum that is, if our iterative scheme is not converging that is one test of convergence would be the elements of d_k right if the elements of d_k are not going towards 0 but going away from 0. In that case, what we could do is not use, so we could use a new a new guess, so what we do is we will just abandon those guess and going to a new guess value that is one approach and then we try our luck on whether we are converging or not. Another method is to say that, ok I will just take this old a value and multiply it by some constant and go into a new value so that is another scheme.

Instead of making a completely new guess, I have the what ever value, I have obtained from this solution of this equation, I take this a and instead of using that a completely as I am saying we are equating this a equal to a the first equation, what I get as new a compute that psi squared of a from that and del squared x squared by del a or I evaluate this d_k metrics again and then see your elements are going towards 0, if they are not going to 0 then I have to, I am supposed to go back here and if they are not 0, I go back here but, if they are going away from 0, then what I should do is instead of writing a_0 equal to a, I would write this as now a_0 equal to some constant times a and then put them back.

So that is the next technique, ok so these are numerical techniques that is, if every things are all right if your initial guess is fine then everything should converge smoothly and if it does not converge and if you see that if your derivatives are going away from 0 instead of going towards 0 and then you take this new a values which you obtained multiply it by some constant and then give it back here again. Instead of giving it a_0 is equal to a, I say now a $_0$ is constant times a.

So this is a technique which we can use to put the, to get the new value so then question naturally arises is, what is the constant value, so before I go into that I would like to summarize this once again. So, we said that, we have the chi squared function which we written like this and then we determined the d_k and the D_{kj} as the second derivative and the first derivative of the psi squared function and we said that we make an approximation into the second derivative, instead of the full derivative of the chi squared function, we will take it as the product of the two first derivatives divided by 1 by sigma i squared and then we taylor expand the psi squared around the guess value and from this taylor expansion, we say that taylor expanding, I can go to the actual assumption leads me to the function del psi square by del a should be 0 at the new value a. I guess that should be 0 at the new value a, okay that defines del psi square del a should be 0.

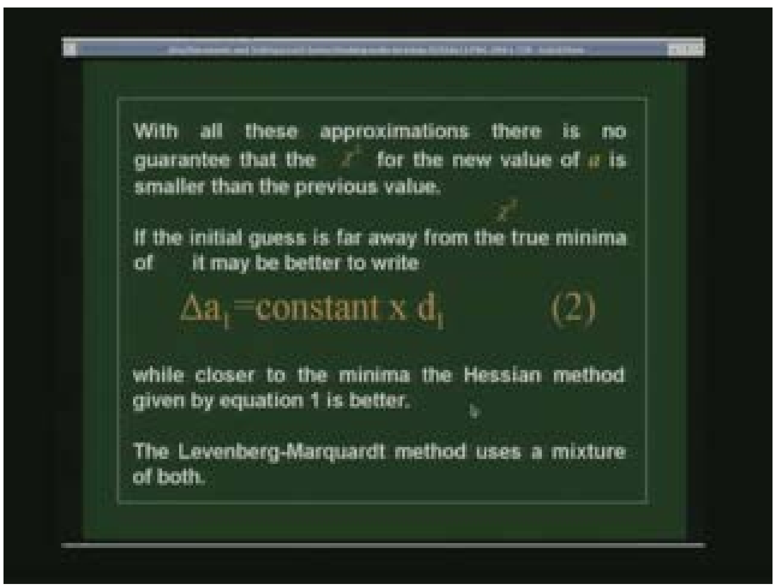
So that from this equation then, I can get an expression of this sort, okay that is I can determine the delta a's and from that delta a's from that I can determine the new a's and I say that if that is, and because I have made all these approximations this new a's may not be the real minima in that case, I would go back and do the iteration but a technical question comes what happens, if this new value a obtained it takes me away from the minima than closer to the minima.

How do I know that, I know that by computing del chi squared by del a if del psi squared by del a is not going towards the 0 then that means I am going away from the minima in that case I should, instead of using the new a value completely as a_0, I should use the new a value multiplied by a constant as a_0. so that is the summary. So now, the question is what is the constant, which I can use to multiply? So we see that that is why I said in the beginning that, this a’s this function f which contains a or different a’s of different dimensions ok one could be dimension of length other could be dimension of time and so they are all different dimensions.

So you cannot use the same constant to multiply all the a’s. So because they have different dimensions, ok so and the scales could be very different. So it is not correct to use the same constant you multiply all the a’s. So we need for different a’s a is a as I said as a column vector now, so it has contained all the a_1, a_2 up to a M parameters. So we need a different constant to multiply the a’s think, how do we do that, okay so one way is to do is to take that if you look at this function ok this metrics here ok and we will see that the diagonal elements of this metrics which is this del f by del a_j del a_j whole square ok has the dimension of a_j’s.

So to determine the new a_j what we are doing is to take the this metrics d and find the inverse of that and multiply it into this first derivative metrics. So, what we see is that when it is not converging, we need to multiply it by a constant. So now problem is what to do? How do we determine the constants because each of the a’s are different dimensions some have length some have time etcetera okay.

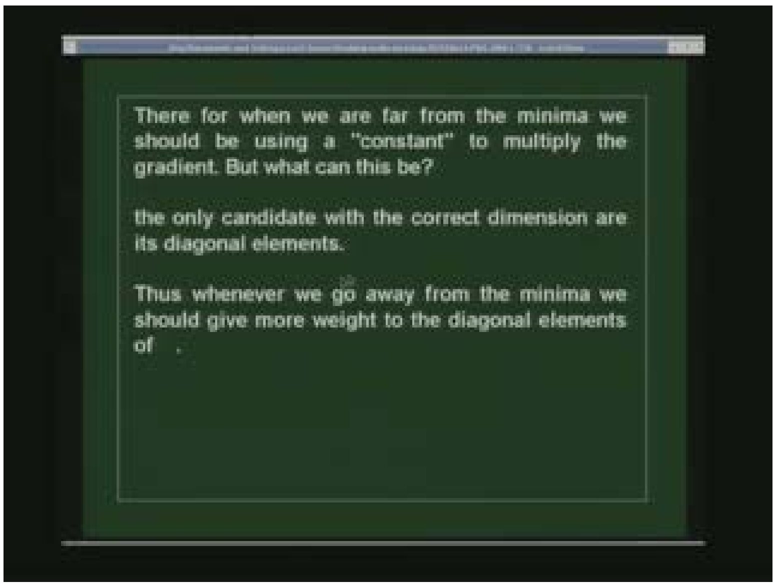
(Refer Slide Time: 39:04)



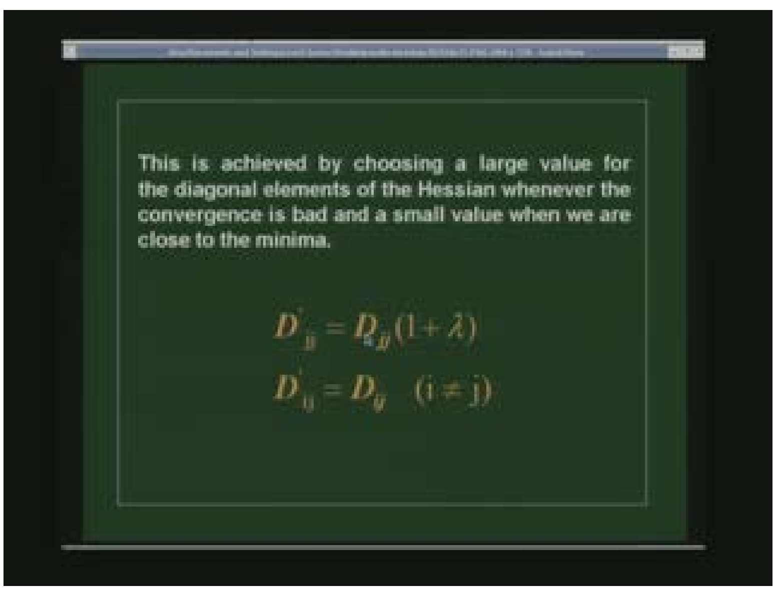
So one way to determine that as I said is to take this d and look at the diagonal elements of that ok and give that diagonal element more weight age, so that is one way of doing this instead of multiplying by a constant here and this constant is different for different a’s so, what we do is to take this metrics d here and this D_{kj} metrics and give more weight age to its diagonal elements. So instead of using D_{kj} metrics straight away here, we will give more weight age to its diagonal elements and then multiply it by the d so that is equivalent to doing this.

So if it is converging, we could use give all the elements of the metrics the same way and if it is not converging, we would give more weight age to its diagonal elements that is what we would be following. So that is what I meant by doing this here Levenberg marquardt method is this exactly that one that is in which we determine the hcm that is this element this metrics D_{kj}'s and then we would give more weight age to its to ts diagonal elements when, when the chi square is not converging or the chi squared is not going to the minimum. Okay, so that is the, we use this constant, we have to use the diagonal element of the of the metrics okay.

(Refer Slide Time: 39:34)



(Refer Slide Time: 39:42)



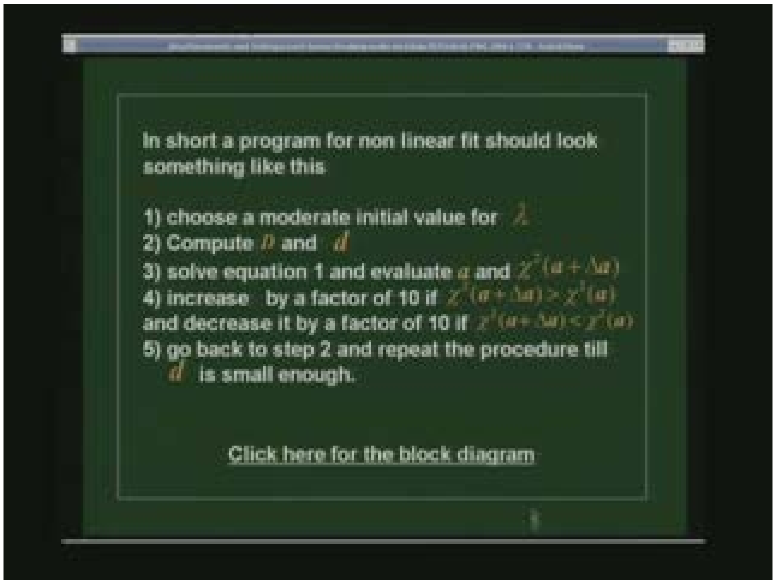
So in other words, we would write now the metrics as the diagonal elements of the metrics we will write as 1 plus lambda into D_{jj}. D_{jj} is the diagonal metrics, diagonal elements of this metrics d. So now we will rewrite those metrics elements as 1 plus lambda and we put lambda equal to 0 when everything is converging and lambda equal to large number, if things are not converging, so when we have a additional parameter lambda here, okay which is in the iterative scheme.

So we see are we going towards the del chi square by del a equal to 0 and if we are going towards 0 then, we would put lambda equal to 0 and if we are going away from that then we put lambda equal to large values say 10 or something like that and then you would go back and then in the step, if it improves then we decrease the lambda again. So we decrease the lambda in general we start with lambda equal to let us say one, because we do not know our assumption initial assumption how good it is.

So, let us say it starts with lambda is equal to one and if we are going towards the, if you are converging we keep decreasing the lambda and if you are diverging, we keep increasing the lambda. So that is the general scheme which would follow. So that is been summarized here.

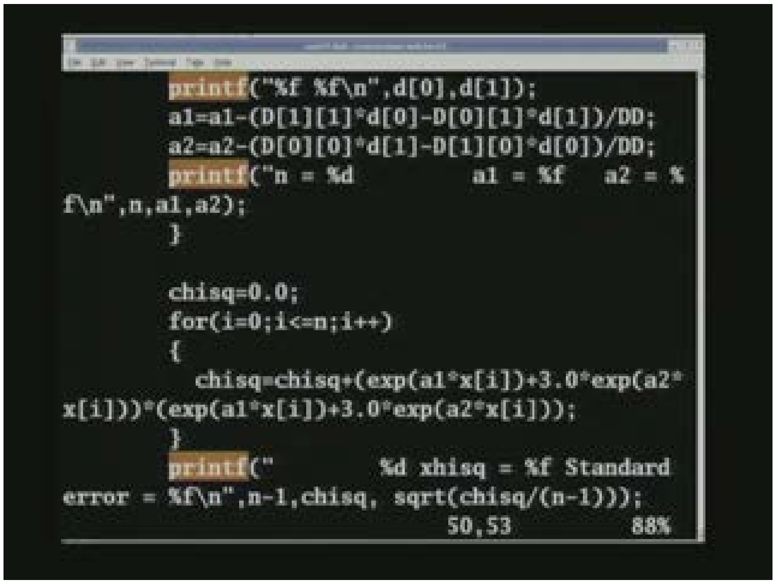
So new this method of choosing them lambda values we choose a moderate initial value for lambda the same lambda equal to one and then we would compute the metrics D and the little d and now we will solve those equation for delta a and then we will determine the chi squared at new delta a new a value that is a plus delta a right.

(Refer Slide Time: 40:56)

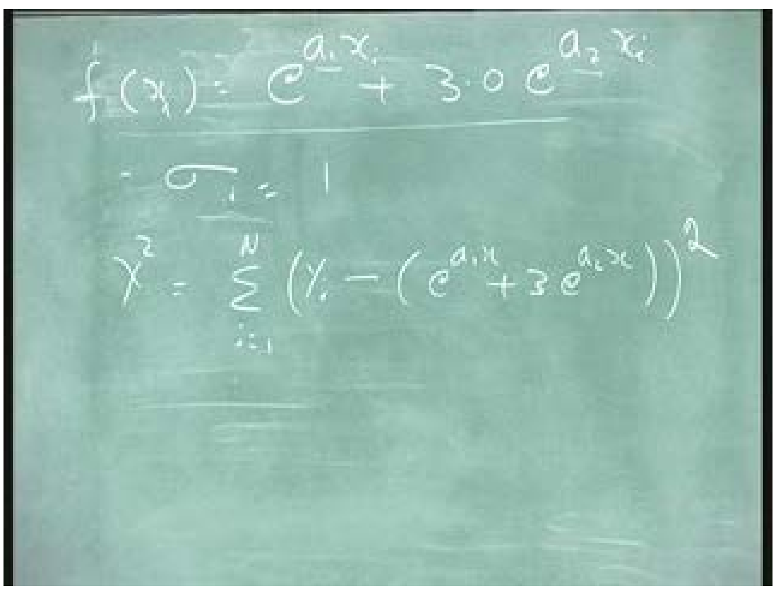


We solve this equation d, d inverse, d inverse d to get the new delta a and from the delta a, we can compute the new value of a or new value of chi square that, psi square a plus delta a. Now, if psi square a plus delta a is going away from the minimum that is, it is going it is increasing as upon the earlier value then, we would increase the lambda value by a factor of 10 and if it is decreasing, we will decrease it by a factor of 10. So that is the scheme which we do and we will go back and repeat the process unless this is the start.

(Refer Slide Time: 42:28)



(Refer Slide Time: 44:08)

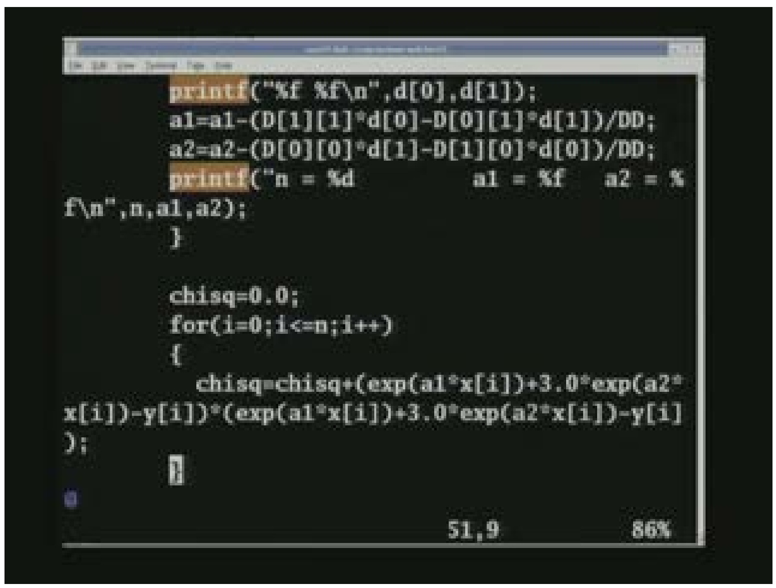


So that is a, that is a way we do, if lambda is, if our scheme is not converging, so we will now see an implementation of a this code in a this, this algorithm in a program. We will first see the implementation without the lambda and we will see the implementation of the lambda after that. So here we are trying to fit, if a given data point or a set of data points through a function of this form that is exponential a_1x plus 3 times exponential a_2 of x that is the function which you are trying to fit. So I will write that function here.

So here, we trying to fit a function of this form, that is exponential a 1 times x_i or this is f of x_i plus 3 times exponential a 2 x $_i$. So, I chosen this function, so that again we allow only 2 parameters to fit. So our metrics would only be 2 by 2 metrics. So it is easy to solve without actually going through a numerical technique to invert the metrics because we will use this, this as an example and what our idea would be to determine a 1 and a 2 from this by minimizing chi square.

So we will write chi squared as 1 over, we have all the sigma i as 1 again and then we would write chi squared as sigma i 1 to n y_i minus the whole squared and then, we would determine all the the metrics d and capital D and we would implement this in this program, that is what we are doing here so this is the chi square.

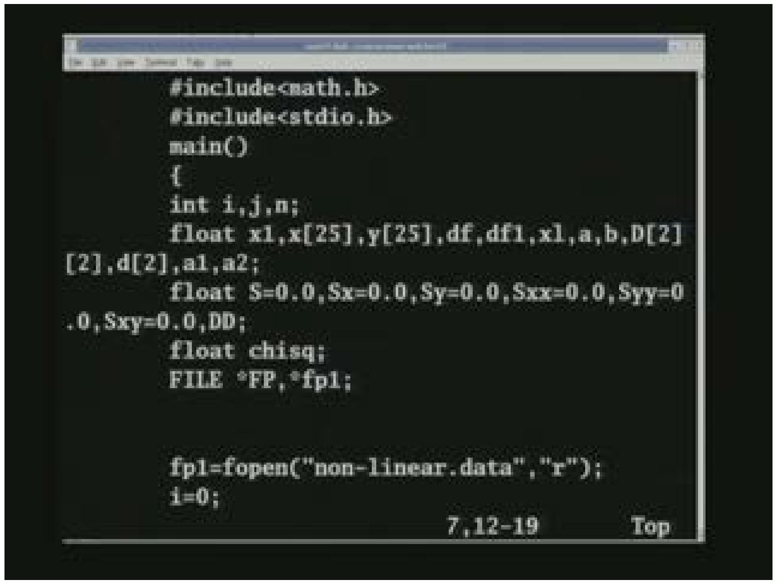
(Refer Slide Time: 44:09)



psi square is exponential a_1 x plus 3 exponential a_2 x square. So I multiply that again by this and minus y of i the psi square is minus y of i. So I should substitute minus y of i. Okay, so psi squared is “a” exponential a_1 x plus 3 exponential a_2 x, a_2 x minus y of i.

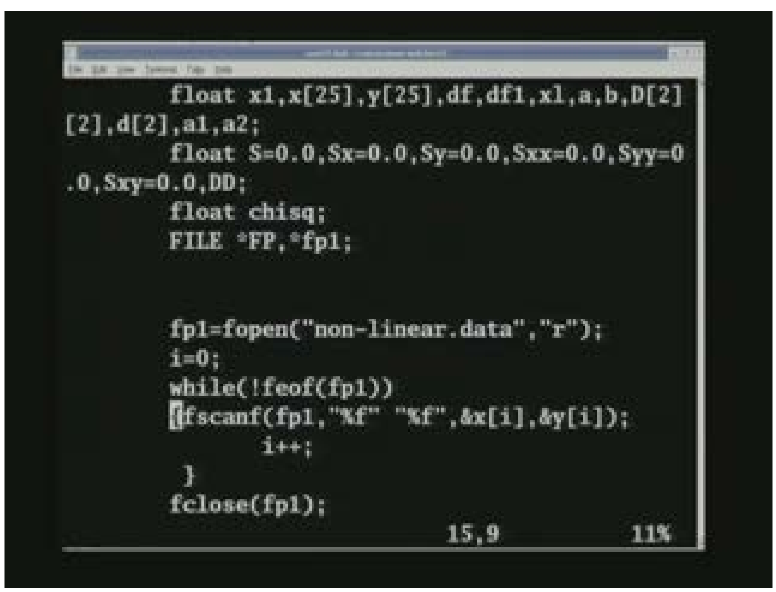
So that is what we have to determine is the a 1 and a 2 values from this, okay so now begin with we have a similar program to that a similar to the earlier ones that is we have the x of x in array x and y okay in which we would store or we will read of the elements of the data points given, if the data points are now given in the non-linear dot data this this function this file so we will read of that file to get our x and y values till the end of the file.

(Refer Slide Time: 44:52)



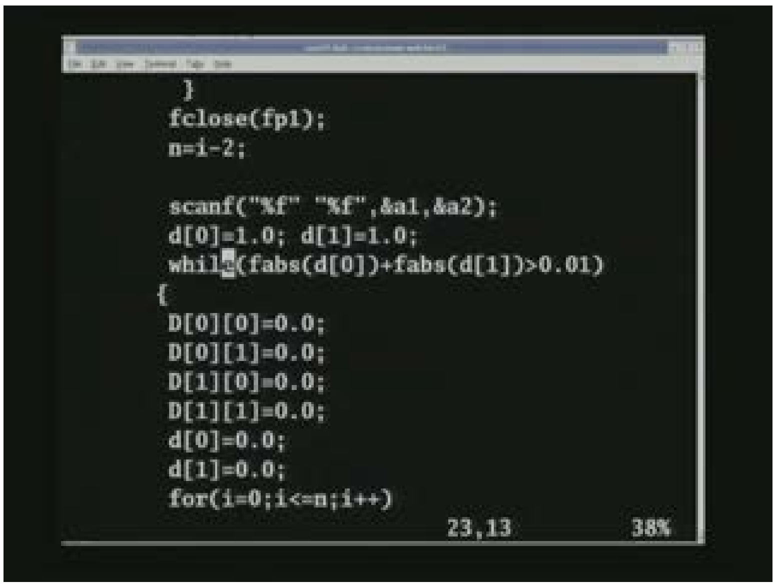
So remember, this is while loop here and this while loop will read it till the file end of file statement is seen. So we do not have to give the number of data points it will read off by itself and the number of data points are then equal to i minus 2. Okay that is what we read of, so we have all the data in this file and then we would the program would kind of look for the input for a_1 and a_2 values that is the initial guesses for the a_1 and a_2 values okay over the function e to power of a_1 x plus 3 e to the power of a_2 x. So a_1 and a_2 are the ones you have to determine or the idea of this program would be is to determine a_1 and a_2 but we would give a_1 and a_2 a guess value right.

(Refer Slide Time: 45:15)



And then, so it initialize, we now, we have to determine the D metrics and the little d metrics and we will so this is the start of the of the loop, iterative loop okay, we first, we have to initialize this d metrics and the little d metrics which is the first derivative of the chi square and we will continue this till this elements of this d metrics that is the first derivative of chi squared goes to 0 right okay, that is what while loop doing. So that is, this the start of the iterative loop, so the iterative loop will continue till sum of the elements of this d metrics goes to 0 right, so that is the, the so I have taken the sum of the absolute value of the elements in this metrics. I cannot say numerical program should go to zeros given some tolerance here the tolerance is ".01" ok you could change this tolerance to the desired value.

(Refer Slide Time: 46:12)

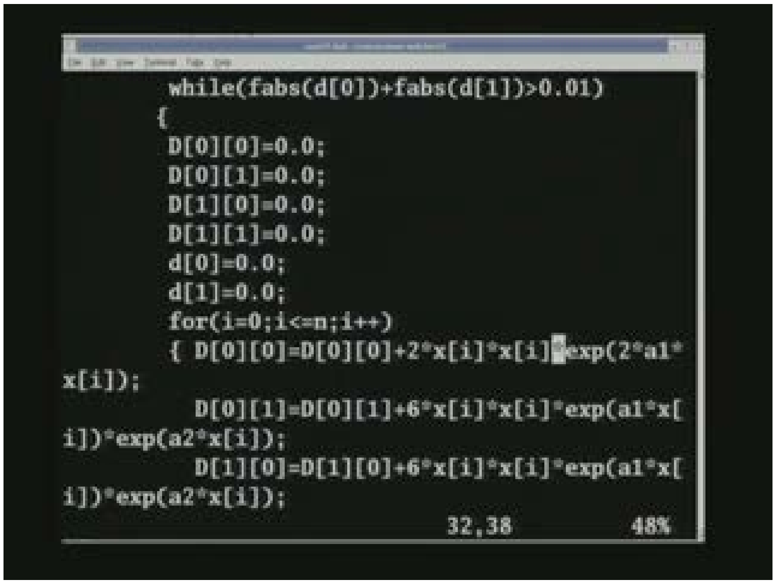


So here, so now while loop that is the iterative loop starting, okay now, this iterative loop starts from here and then keeps continue it, continue it keeps continuing or it continues up to of the elements of the d metrics goes to 0 and I determine that by computing the absolute values of the element the sum of the absolute values of the elements.

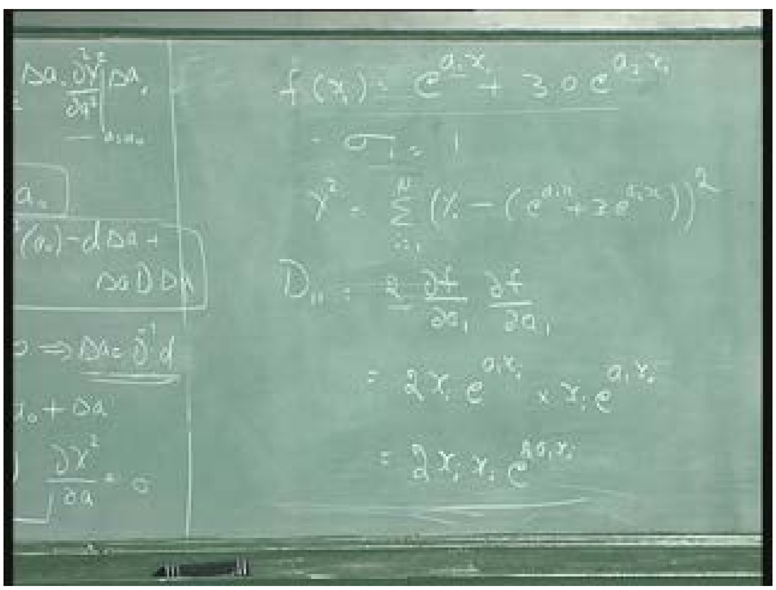
So that is what my, it should be greater than if it is greater than ".01" then the loop should continue if it goes less than .01, then it will stop that is the that is the which I have given and this tolerance is something which is determined by the user. Okay, so now when we initialize the d metrics and the hcm and then we would compute that values. Okay we are computing the sum over i and the product of the two derivatives, ok so that is what we are doing it is the product of 2 derivatives and the factors of the 2 incorporated because we wanted to write it as simple delta a as d inverse d.

So we want to write, so we put in the factors of 2 in here and then this is the product of the first derivatives of the functions. So this is d_0 would be just a product del f by del f by del a_1 multiplied by del f by del a_1. So that is simply x of i into x of i exponential 2 a_1 of x of i, we have the function of this form and we said that the metrics element D_{11} would be then 2 times del f by del a_1 into del f by del a_1. 2 coming because, we want to write this as simply d inverse d, so that is, that would give us as your 2 times x_i into del f by del a_1 will be again, e to the power of a_1 x. So e to the power of a_1 x_i and into x_i e to the power of a_1 x_i or I wrote this as 2 times x_i into x_i that is x_i squared e to the power of 2 a_1 x_i.

(Refer Slide Time: 47:54)



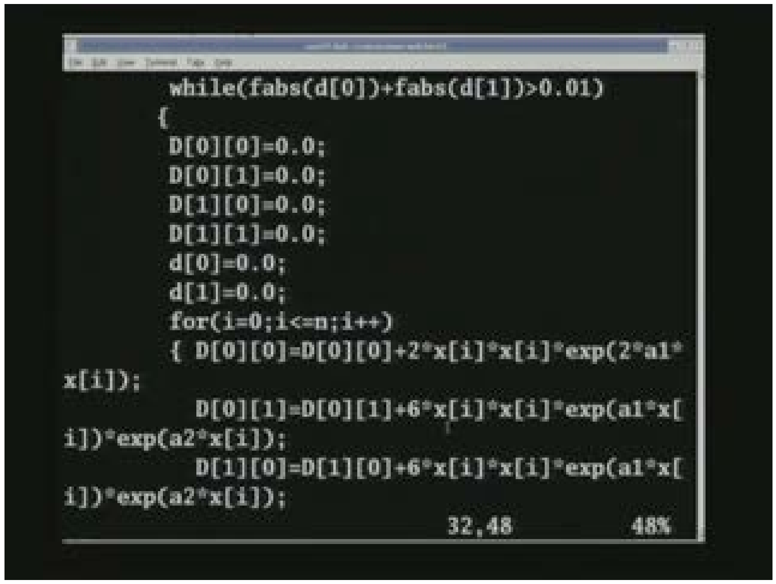
(Refer Slide Time: 49:29)



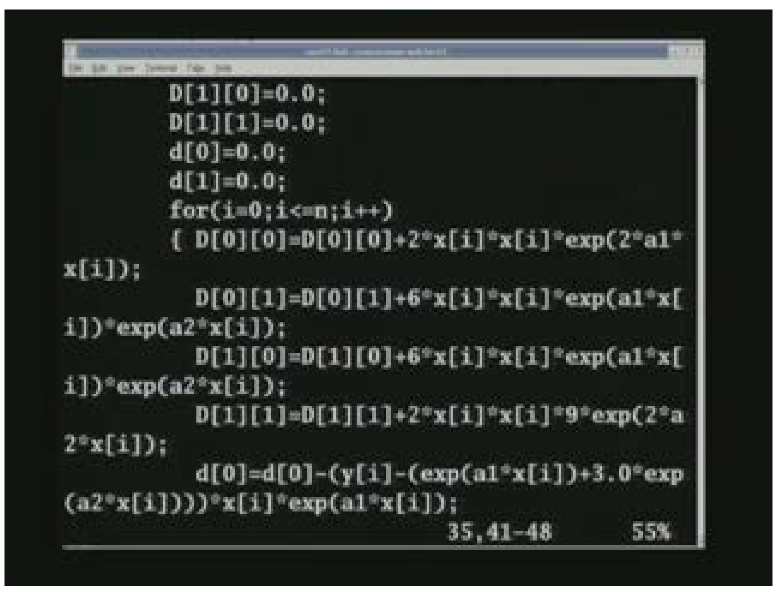
So similarly, other elements I can write now and that is what has been put in here okay so, x_i into x_i e to the power of 2 into a 1 x_i and D_{01} would be now the multiplied by 2 times x_i into the e to the power of a_1 x_i into x_i into e to the power of a 2 x_i because, now

D_{01} is the a_2 the second part of the function and D_{01}, D_{10} are same they are symmetric and then you have D_{11} which will be now 2 times x_i into x_i into 3 times exponential a_2 x into 3 times exponential a_2 x that is 9 times exponential 2 a_2 x that is what is here.

(Refer Slide Time: 49:30)



(Refer Slide Time: 49:52)

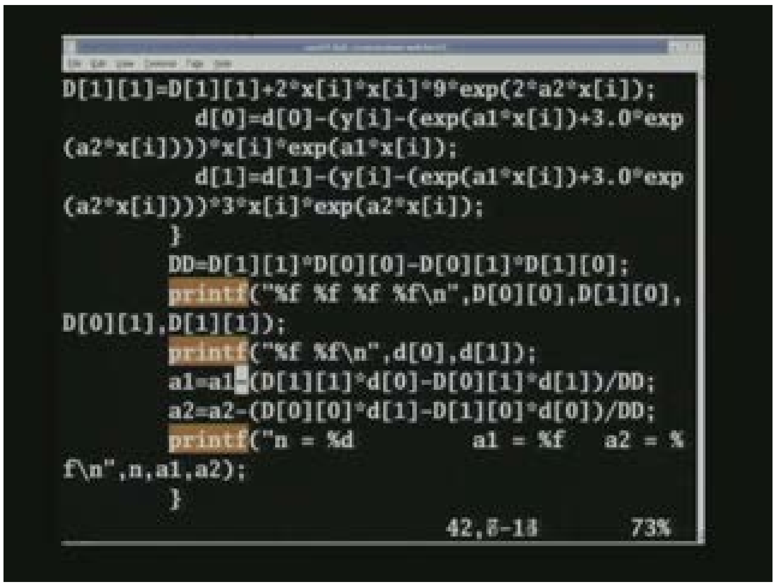


Okay, so that is the metrics elements as we determined, okay the d_{11} is 2 times x_i, 9 times exponential 2 a_2 x_i. So that is what and then we have the d_0 which is y_i minus exponential a_1 x_i plus 3 times exponential a_2 x_i that is y_i minus f of x_i into the derivative of the function that is x_i into exponential a_1 x_i, the derivative of the function with respect to a_1. So that is x_i into exponential a_1 x_i and similarly, d_1 would be y_i minus f of x_i

multiplied by the derivative of the function with respect to a $_2$, now that is 3 times x of i exponential a $_2$ x_i .

So now, we have the metrics and then we would compute the inverse of the metrics as, we have done for the a general linear fit. So we have the determinant of the metrics and then I take the, I would compute that the inverse of the metrics as the adjoin divided by the determinant and then from which using that I can determine now a_1 and a_2 values.

(Refer Slide Time: 50:53)

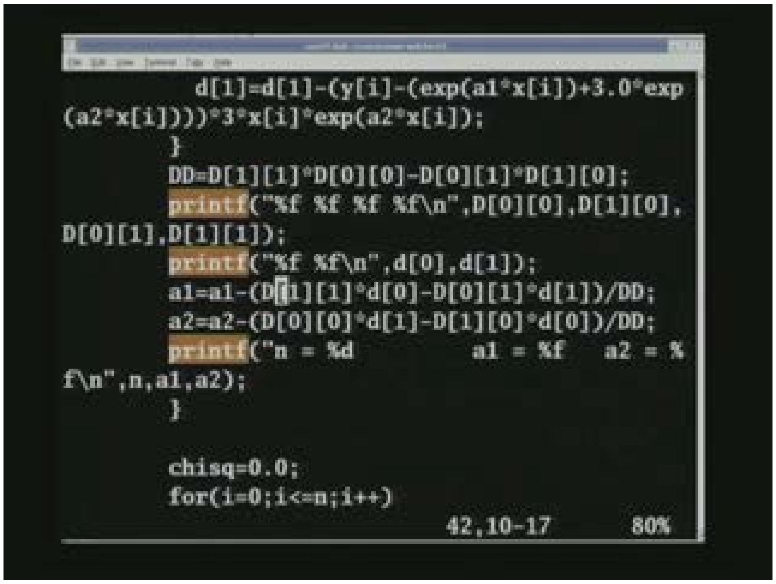


So by multiplying the metrics with the d inverse with d such that with capital D inverse with d, I can determine the delta a and then I name a new value a 1 as a 1 minus delta a as we just saw. So a $_1$ minus delta a is implemented this a_1 and a_2 are the new values. So the old value minus the product of metrics, okay gives me the new values of a_1 and a_2. So then I would determine now.

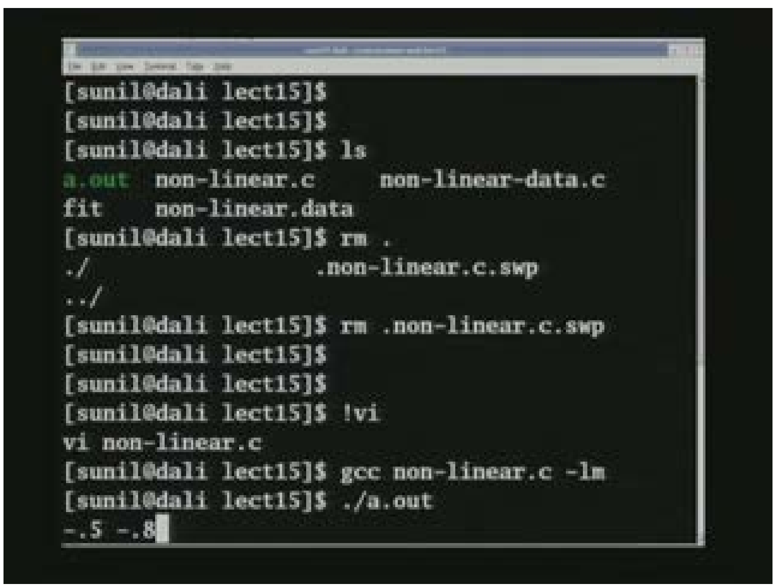
Now I will once again a_1 and a_2 values here, I can I also have the d values here and here I will just check whether these d values satisfy my condition of the while loop that is the absolute value of d_0 plus d_1 is less than ".01" if it does not satisfy then this loop will continue with this newly determined a_1 and a_2 values as the new values and I compute again the d_0 metrics and the d, the capital D and the little d metrics and we continue the iteration all the way.

And once that condition is satisfied that is the value of the, so once this the absolute the sum of the absolute values of the d's are less than ".01" then it comes out of this thing this loop ok and writes the new chi square value and the standard errors that is what this program shall do so we will look at this program.

(Refer Slide Time: 51:30)

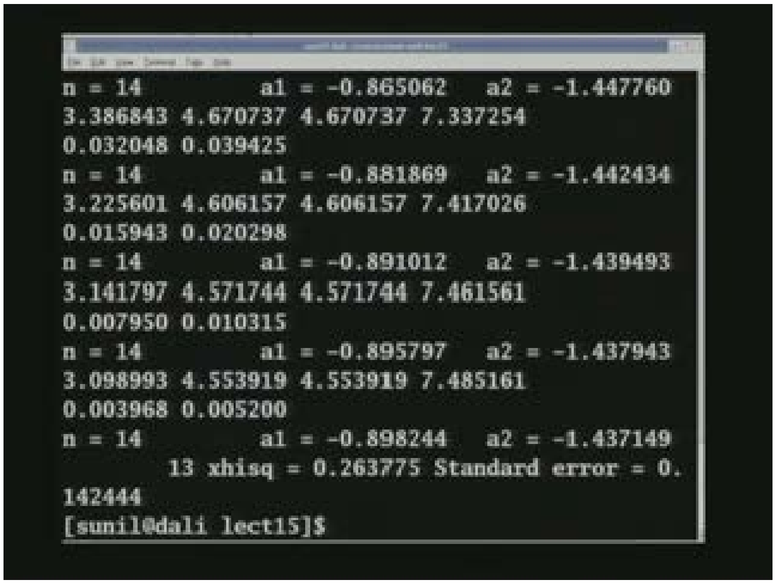


(Refer Slide Time: 52:53)

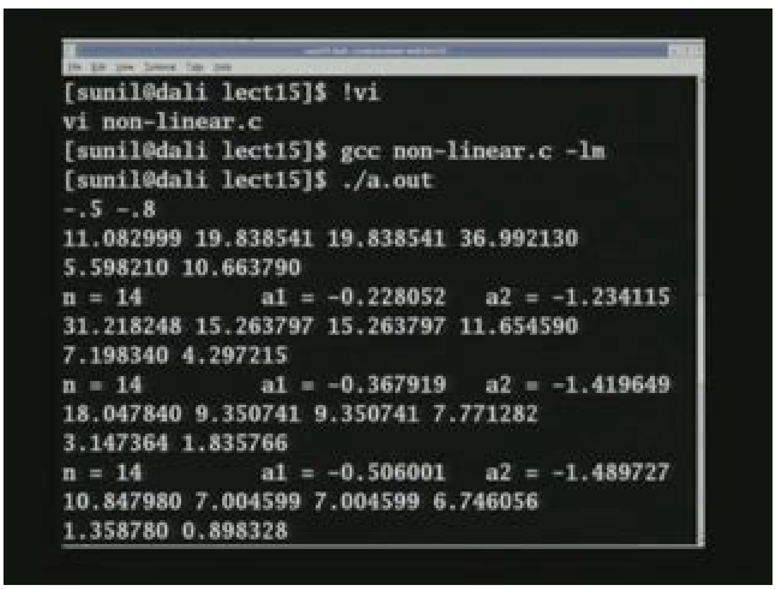


So we will compile this and then run this program and since, it is waiting for some input values to be given so give some input values as minus ".5" and minus ".8" okay, so it computes that thing and it gives me as minus ".898" and minus. minus "1.43 " as my as a solution finally okay, so it went through some number of iterations ok and then it comes out obviously it did not complete this in 1 iteration it set it in many iterations. Okay, so and it, but it converged this program okay and it started with a_1, a_2 values as ".22" and "1.23" and then ".36" "`.4" etcetera.

(Refer Slide Time: 53:02)



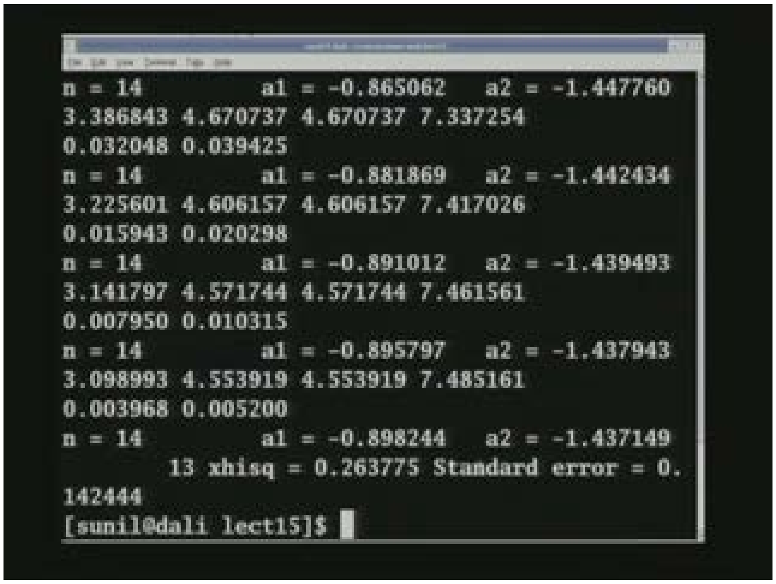
(Refer Slide Time: 53:28)



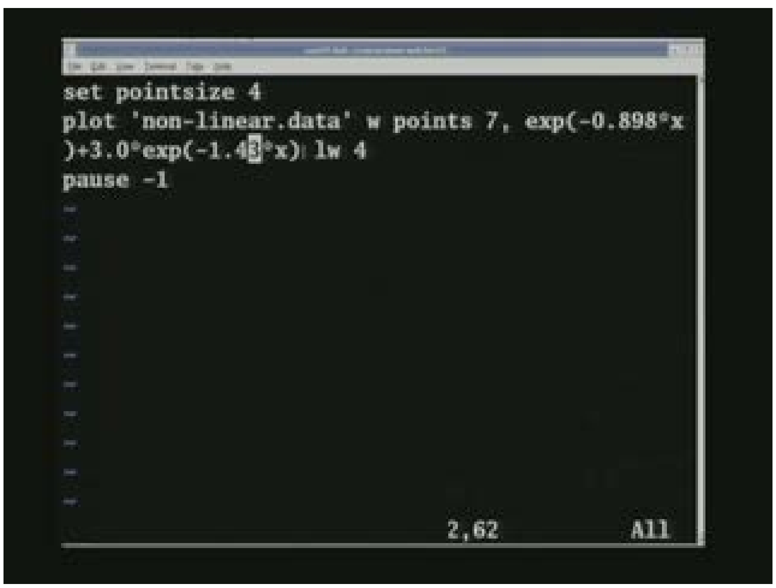
And finally it converged into this value that is minus “.898” and “1.43” and with a chi square which is “.26” with a standard error of “.14” which seems to be pretty good.

so we will look at this now we will plot the data which we had that non-linear data function in the file and we will compare that with this exponential minus “.898” x_i plus 3 times minus exponential minus “1.43” x_i function such what we will do here so minus here “.898” and “1.437” are the parameters here.

(Refer Slide Time: 53:39)

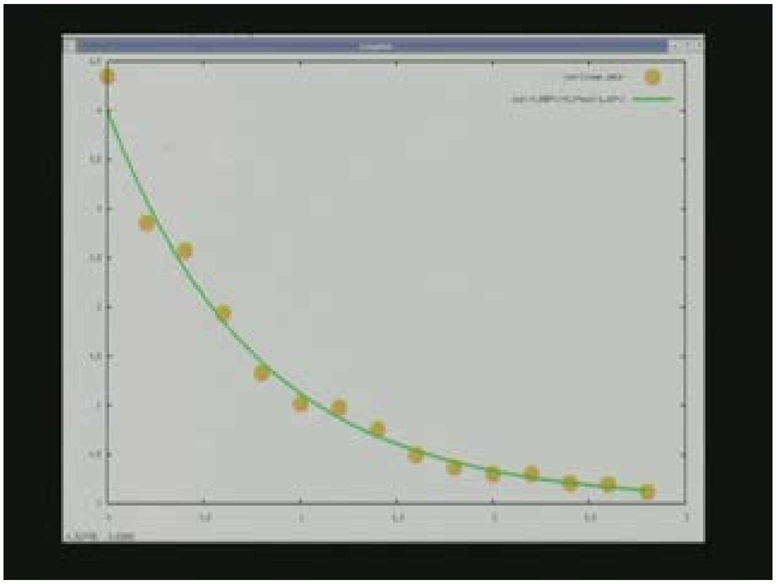


(Refer Slide Time: 54:39)



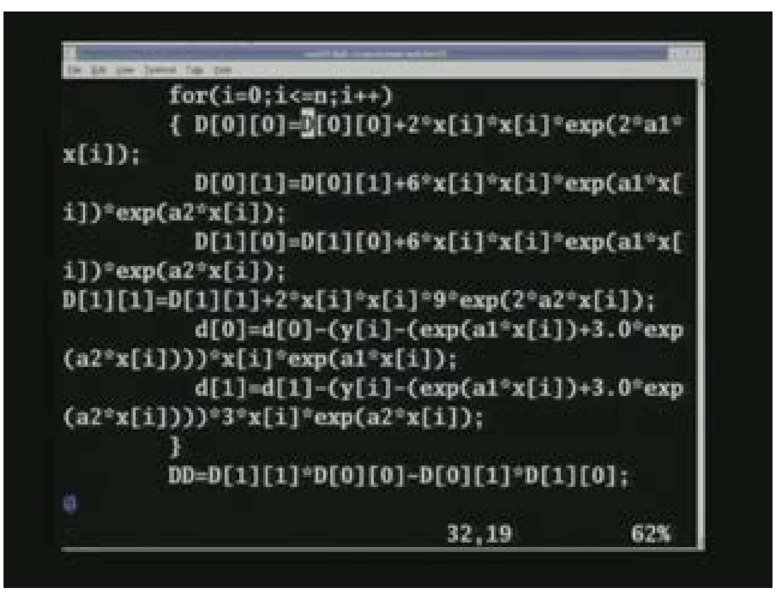
So “.898” and “1.43” as our parameters and that is the function I am putting and non-linear dot data is the data which is given to us and exponential minus “.898” x plus 3 times exponential minus “1.43” x is our function. So I will plot this function and that is what we would get. Now, in this plot, okay we can see that the the round are again data given to us and this is our fit through it and this is the best fit which we can get with this double exponential function. So now, we should now look at an example, where this may not be so simple that is this function may not converge as easily as we just now saw.

(Refer Slide Time: 54:49)

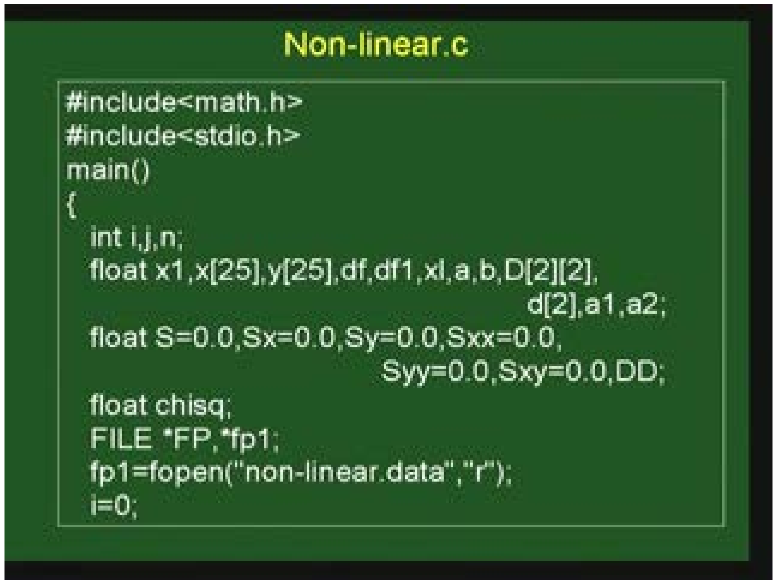


So in that case we will have to now, in the program we will have to make some modification that is we will have to now say that d_{00} instead of saying it is just that product of 2 derivatives we will have to multiply this by 1 plus lambda, ok and whenever we go away from the convergence we will have to increase the lambda and we are going to wards the convergence we have to decrease the lambda and that is the only modification in terms lambda which we have to make only the d_0 element and the d_1 element, we will have to change and include this into to accommodate the convergence the solution for the convergence problem, okay that is what we should be trying to doing this, in this program and we will see in the in the next class I will stop here.

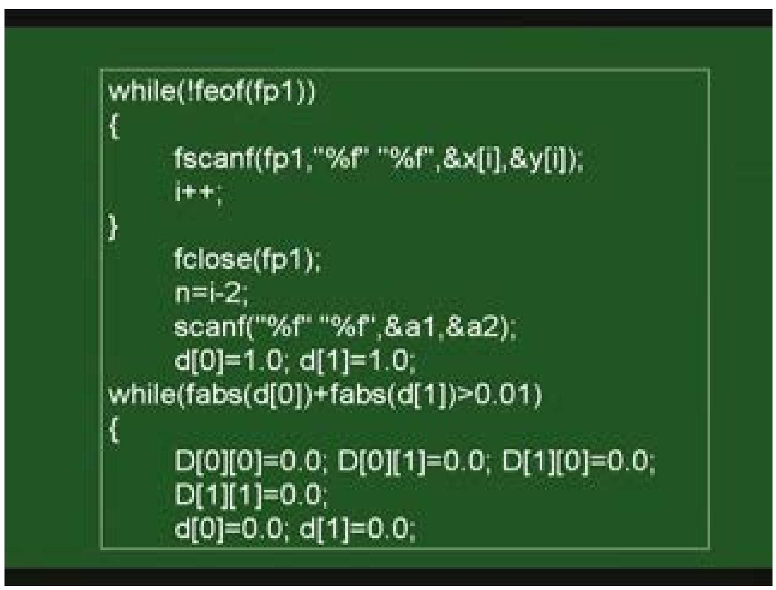
(Refer Slide Time: 55:23)



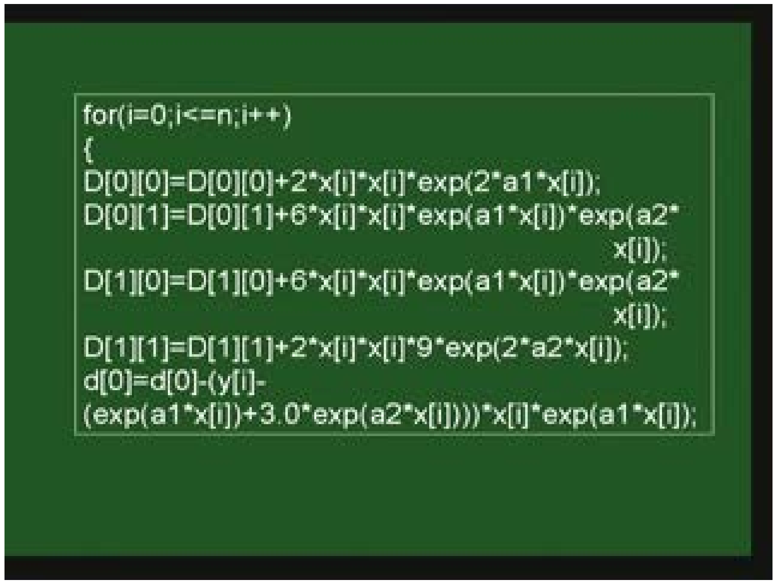
(Refer Slide Time: 56:06)



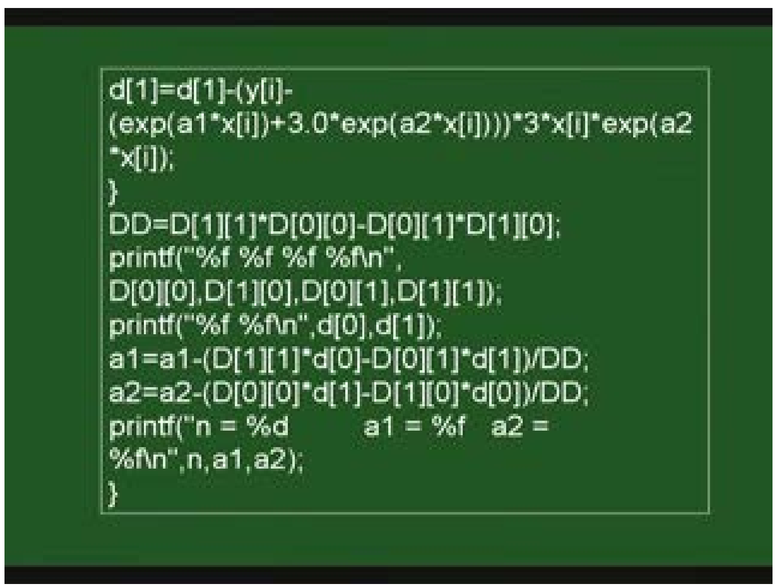
(Refer Slide Time: 56:16)



(Refer Slide Time: 56:28)



(Refer Slide Time: 56:42)



(Refer Slide Time: 56:54)