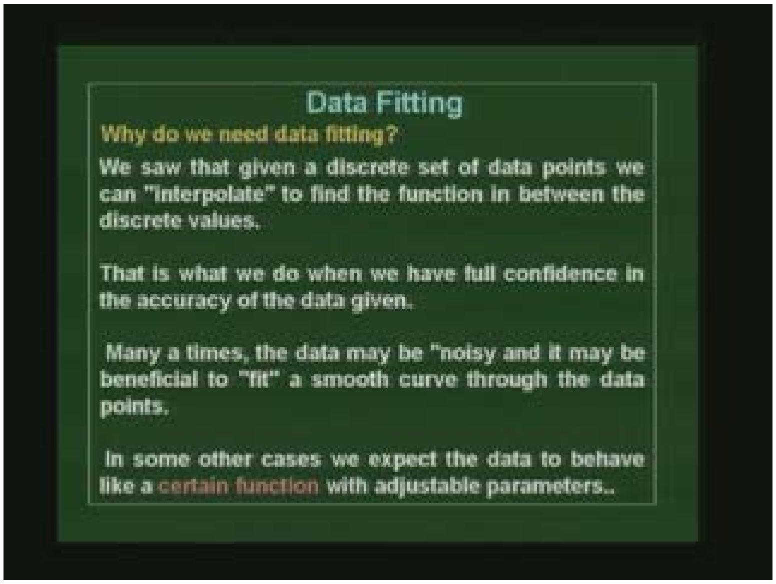
Numerical Methods and Programming
P. B. Sunil Kumar
Department of Physics
Indian Institute of Technology, Madras
Lecture - 13
Data Fitting: Linear Fit

Today's discussion on numerical methods and programming, we will focus on the modeling data in using a fit. So last time, we discussed modeling data using interpolation.

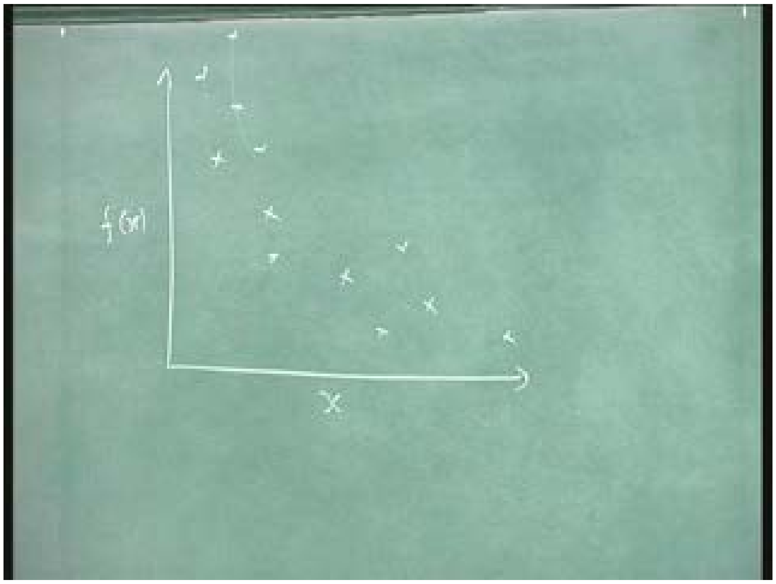
(Refer Slide Time: 01:37)



So that is we do, so if, we have a set of points this is what we did last time, if you have a set of points, let us say a function of x, variable x and then if you have a set of points and then we would say that if the, we have got some set of points like this is some measurement and if you have full confidence on the, on this data that is if you know for sure that these data points are reliable all this variations in the data points are actually reliable are actual variations and then we would have what is called an interpolation.

So then, we would have an interpolating data which would just go over all this data points. So that is what we saw in the last few classes that how to construct a line which goes through all these data points but let us say, in many cases we do not have, we do not believe that these fluctuation in the data is actually real and all we want is an average line which goes through all these points. So in that case, we use what is called data fitting. So we have a functional form which we believe represents the correct data and we want to just fit that function form into this set of points.

(Refer Slide Time: 03:08)



So that is what, we would be using we will be discussing in today is lecture, how do one do that there are two occasions, one is this data is not very reliable and so all these fluctuations may not be real fluctuations there will be some measurement error. So the reliability of each of these points has some finite, and an another case is that we know the functional form many cases that happens we have the, we know the functional form and we want to know, how that data would fit in. So as we saw in the case, that we could have for example, if you have some curve like this we could have, if you have a polynomial interpolation we would have some curve which goes like that, that would be the a polynomial kind of interpolation.

If we believe that our function should go through all this points, that is what we use if you actually believe that these fluctuations are real and if you want to evaluate of, if you want to know what is the function value at any point in between here or here for example, what is the function value and then we would use what is called the interpolation but, as I said some cases, we know that this is not supposed to be fluctuating and we know these are due to some measurement error or some numerical error, if the data comes from simulation or numerical calculation or it is because of some measurement error, if it comes from some experiments. So and we believe and all we want is just to get a smooth curve which goes through all these points, so here is a smooth curve which goes through all these points and that is all we are interested in, we are not actually interested in the actual fluctuations of the data. So how do we do that? So that is what we would do with the fit.

(Refer Slide Time: 05:12)

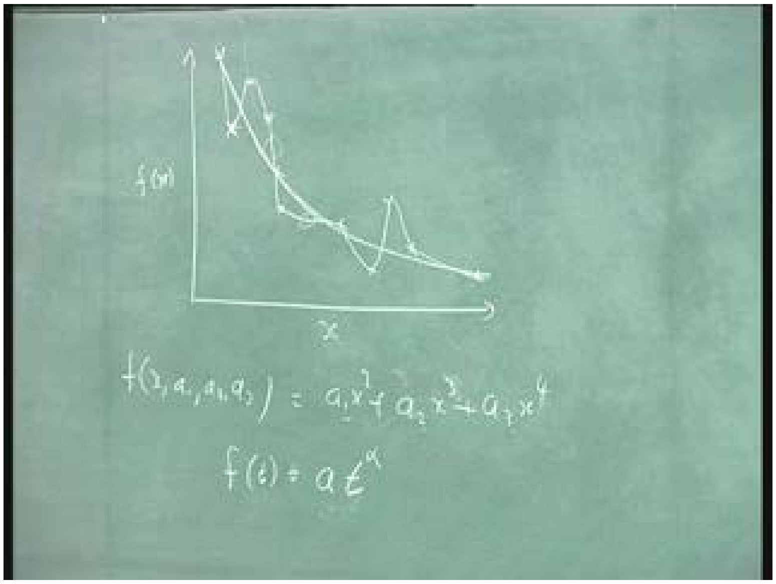
So to summarize, so if you have a given set of discrete, given set of discrete points. so we can either interpolate or we could fit. So that is what the two methods are and now we are looking at is the fit, we looked at the interpolation and now, we look at the fit and use mainly two cases mainly when we do not have full confidence in the accuracy of the data given and another case that we know that it should look like a certain function with adjustable parameters. So, if you know that the data should actually look like a certain function with adjustable parameters then, this f of x let us say, this f of x actually has some adjustable parameters a_1 a_2 a_3 etcetera.

These are adjustable parameters of this and then we could actually try to get this curve by adjusting this parameters and try to get the best curve which goes through all these points or a function which best represent this set of data points. So that is the idea of doing a fit to get this function, get a function with adjustable parameters and by adjusting those parameters, we want to get the function which fits best to this data, which is the best fit to this data, that is what we want to look at and the question is how do we get that and what is the quantity which tells us that something this is the best fit. So that is what we would be looking at into this lecture. So here is an example, let us say that there is a quantity eta which behaves which we know as eta. We know it behaves like some constant times or eta is proportional to t to the power of alpha and we only interested in basically what the alpha is this is a very common occurrence in science.

So that we have some measurement eta some quantity which behaves which varies with time or temperature or whatever some variable as eta to the power of, eta goes as t to the power of alpha. So now, if you do a typical measurement for eta or typical simulation calculation for eta you would get a very fluctuating curve. So then, from that now, if you could plot now from that we need all, we need to do or all we are interested in the experiment is to actually extract this alpha. So that is what I said. So you could have as I said this function could be some you know a_1 x square plus b_2, a_2 x cube plus a_3 x to the power of 4 or something like this, some function like this, some polynomial function like this and we are interested only in a_1, a_2 and a_3 values.

We know it is a polynomial. Let us say, we know it is a polynomial and all we are interested in is a_1, a_2, a_3 values. We were not actually interested in getting the full functional form of the, we are not interested in getting the representing the fluctuations and another case which I said here is that, we say f of t is equal to, let us say we are interested in some a t to the power of alpha. We are interested in that and we are only interested in this exponent alpha, we are not interested in this whole variation of this quantity. So in that case how do we do this. That is the case which we look at it.

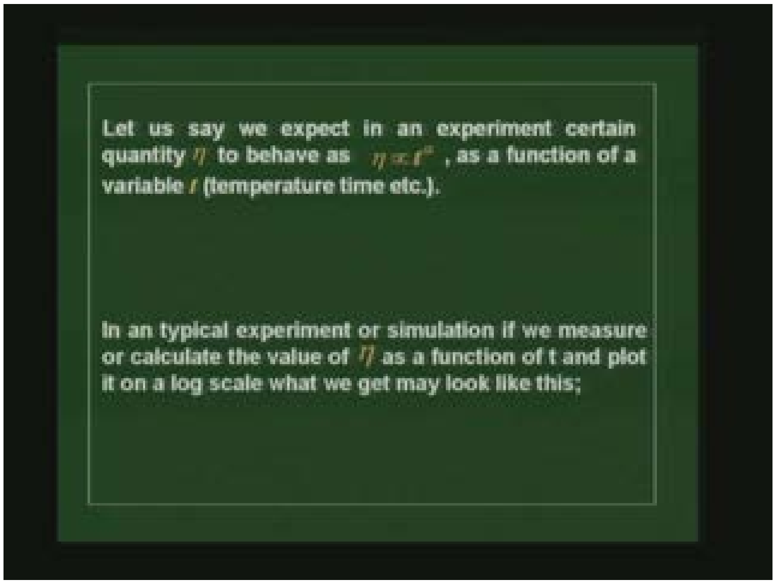
(Refer Slide Time: 08:49)



So, if you actually do a measurement and what do we find is that in an actual measurement this function might look like you know a set of scattered data points. So I have taken f of t as or the function which I call eta in that particular case. So I we are taking this function eta as a t to the power of alpha then, I am taking the log of that on either side. So that would be equal to some log of a plus alpha log of t. So and I am only interested so, I know that this goes like if I plot log eta versus log t, so I take log eta versus log t then, that is the straight line this is a constant.

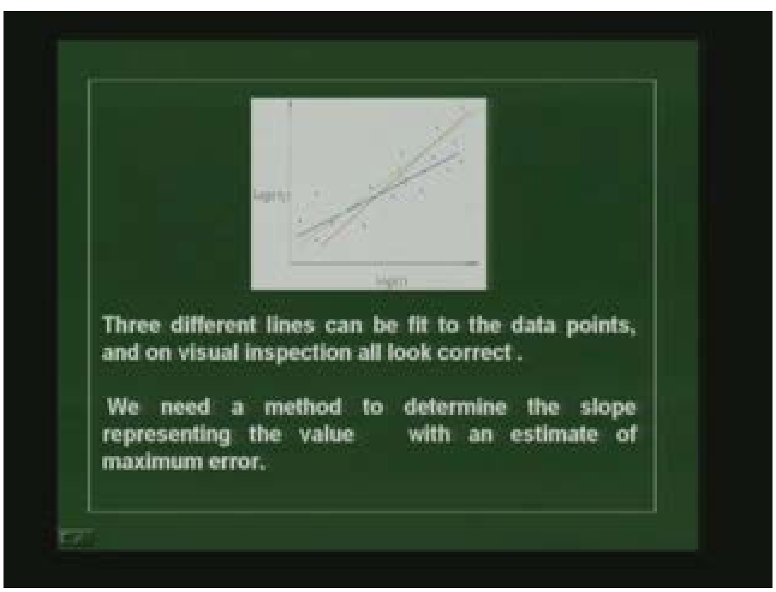
So if I take log eta and plot, if I take log eta on the y axis and log t on the x axis, I would get it as a straight line that is what I would expect but, in a typical experiment what I would get is a set of scattered data points. So I will get a lot of scattered data points and I know that it should behave as a straight line.

(Refer Slide Time: 08:54)



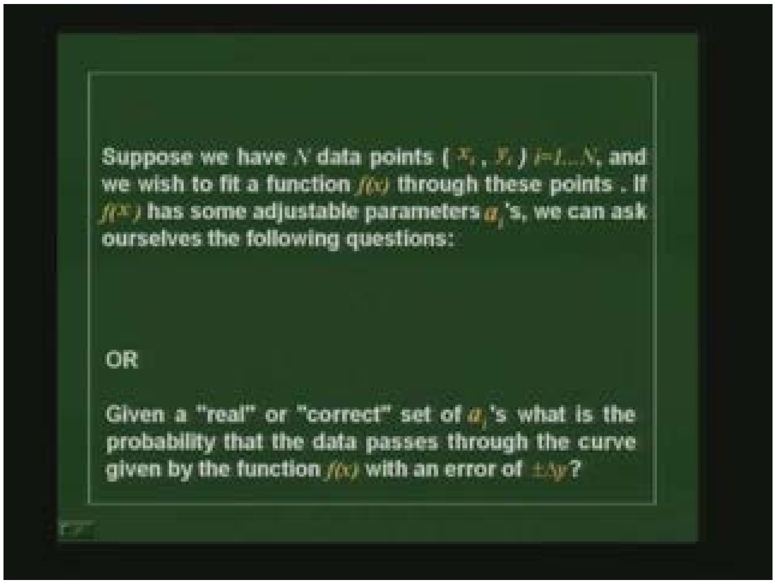
So I try to fit a straight line through this, I try to draw a straight line and then what you see is that by visual inspection I cannot do this because there are, I have drawn 3 different straight lines which all looks more or less going through this points. So that is not a good way to do it. So we should have a correct, some quantitative way of measuring this that which is the correct line which goes through this set of points. So that is what we should be trying to estimate. So we should try to estimate an error and then we should to minimize the error, so some quantity called a fitting parameter a fitting error, so that is what we would be looking at.

(Refer Slide Time: 10:42)



So we will try to now quantify this a bit and then do this. So let us move, let us say that we have a set of data points x_i and y_i.

(Refer Slide Time: 10:58)

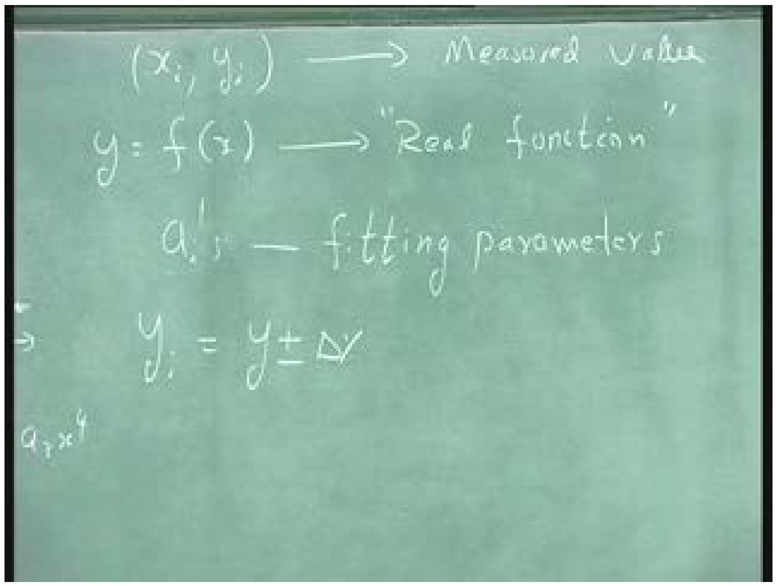


So y_i is, so x_i is the independent parameter, independent variable and y_i is the measurement. So it is a function of time, this is a function, this is cubic time temperature or something and this is some quantity which we would measure. So then, we need to fit some let us say, some curve f of x through this. So the curve which you want to fit is f of x, so if everything is true and if what the fit is there is no fluctuation and everything is correct then what you would get is, y equal to f of x but, y is equal to f of x is the real curve which we expect to get. So, but the measurement or simulation data gave us x_i and y so for each, so what we got y of x_i, y_i is the measurement and x_i is the variable which we changed and now, we want to fit in a curve through this and let us say we had have adjustable parameters a_i.

So our adjustable parameters are a_i, so that is the, so this is the real function, let us say that is the real function and that is the measured value and we have a_i are the fitting parameters. So a_i is the fitting parameter that is what we have so, we have some function, some variable set of data points and we want to fit in a function y equal to f of x type through this data points and a_i are fitting parameters, so the question again to formulate is that, what is the values of a_i which best represents this data.

The function is given to you, the function is already given to you. So that is nothing to determine there except this parameters a_i. So what are the values of a_i which best represents this data or another way the question would be is that, if I write this function with a set of a_i, what is the probability that the measured values will be close to that function that calculated value plus minus delta y? So these are two questions you can ask one is given this data points what are the values of a_i which best represents this function which best represents this set of data points or we could ask that I give you the full function with a set of a_i and what is the probability that these measured values are y_i, what is the probability that y_i which is f of x plus or minus some given delta y, some small delta y what is the distribution of this delta y that is the two questions which we can ask that is what we should try to do.

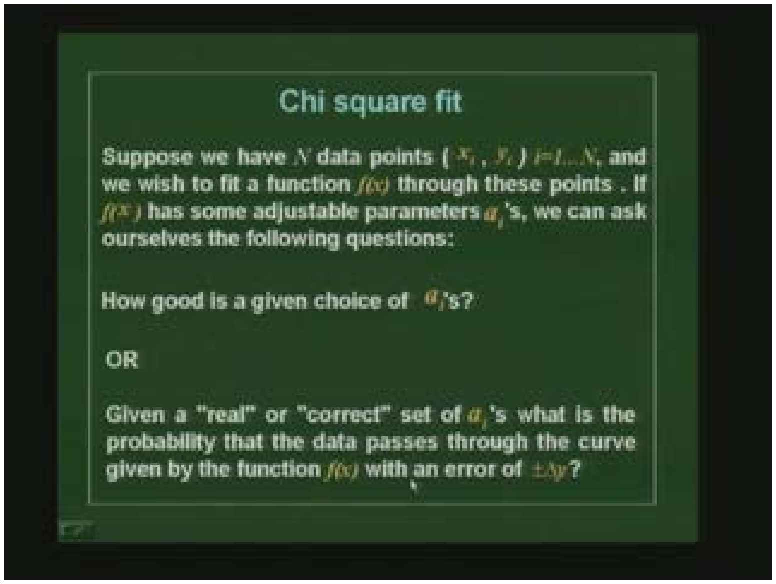
(Refer Slide Time: 14:25)



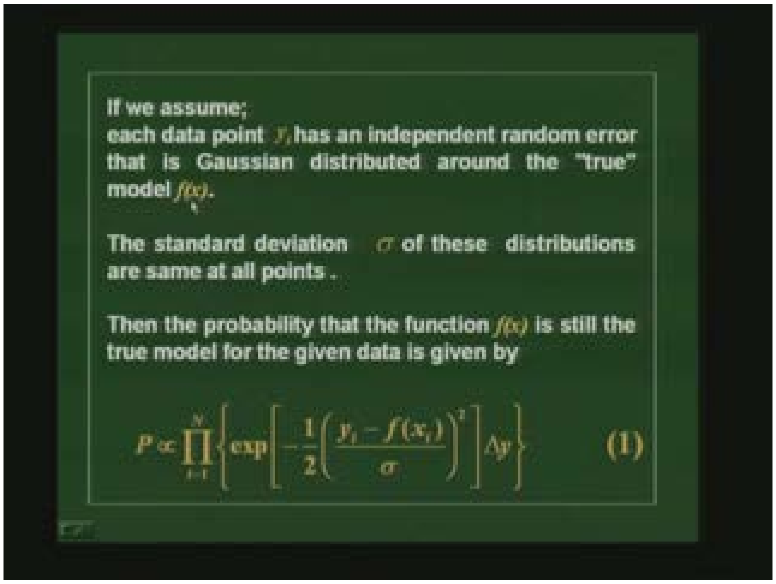
So, again to summarize, so we are given a set of n data points x_i and y_i and we want to fit a function f of x through this set of data points, we have some adjustable parameters a $_i$. so the question we are asking is, if you are given a set of a $_i$, so there are two questions which we could ask, so that is how good is a given choice of a_i that is, if I had given this functions this values x_i and y_i and an adjustable parameter a $_i$ we choose a set of a_i how good they are or in another words, what is the a_i which best represents this data, if I put it inside this function f of x or given a set of a_i, if given a set of a_i. Suppose that is correct set of a_i, suppose it is given to you a set of a_i then, what is the probability that the data passes through the curve given by the function f of x with this set of a_i with an error of plus or minus delta y. So these are the two precise questions which we can ask. So with these two questions we can actually formulate keeping these two questions in mind, we can formulate a method of doing a fit and now, this fit would be called chi square fit and we will see what that chi square fit means in a few minutes.

So now, so what we are saying is the following that, if we assume that each data point y_i has an independent random error which is gaussian distributed around the true model f of x. So what does it mean, so we have this model, this is a model so which I call real function or that why I put it in quotes, it is real function in the sense it is a model this is what we believe the correct function is and we now assume that the measured values since, this is the real function the measured values that is y_i are distributed around y in a gaussian manner what I mean by that is, if I take if I actually plot the probability of y_i the probability of y_i and then I plot it as a function of y_i for different values of y_i.

(Refer Slide Time: 15:52)

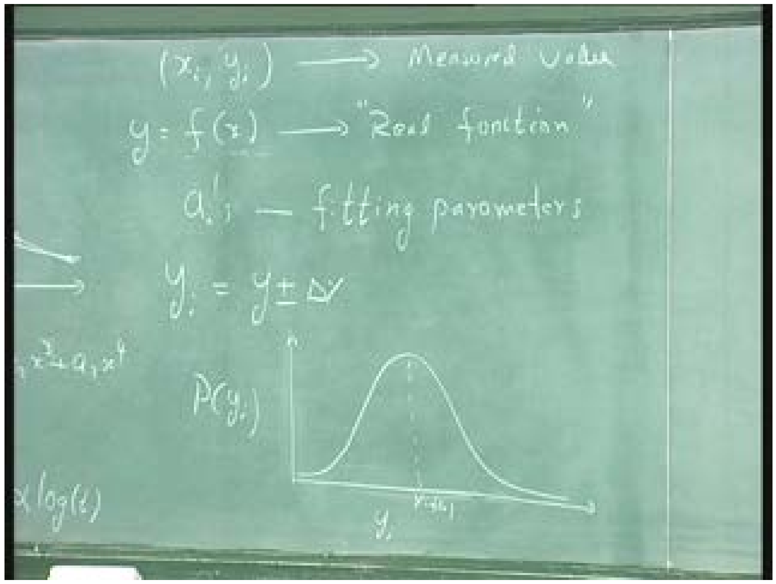


(Refer Slide Time: 16:10)



So this is, so that is this axis I plot the probability and then, if I put y_i here, y axis and then I would get some distribution like this, some distribution like this some distribution like that, this is what I would get and the peak of that distribution would be y equal to f of x. So that is what I mean by this that is it is gaussian distributed this is what I would assume.

(Refer Slide Time: 17:33)



I would assume that an independent random error is what happens, when I do the measurement that is the true function is actually this is the real model for the function model for the data and when I, so I should have got for a given x_i, I should have got f of x_i as my y_i value but I will not get I will get some value y_i and that measured value and because it is a random error, by random error it will always be gaussian distributive, that is the assumption which we are going to make. So it is always gaussian distributive or normal distributive, that is what we are going to say.

So, remember this is the, what I am plotting here is the distribution of the y_i values, if I take let us say x_i represents temperature and this y_i represents some quantity some measured, some value which I can compute using conductivity, something resistance something. So that is the temperature resistance and then I fix temperature and I make many measurements of resistance of some object a wire and I plot the value of resistance I get for different, the number of times I obtain a one particular resistance as a function of resistance, if I plot that I would get a distribution like this, where the peak is, with the peak corresponding to the real value.

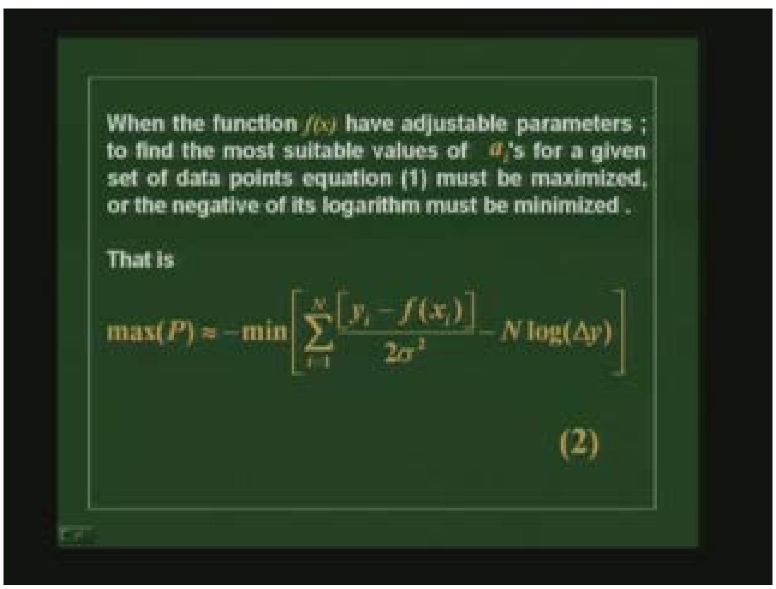
So it is distributed around the real value in a symmetric gaussian way, so that is the assumption. The reason for this assumption is that, we believe this error in this data is some random error and by central limit theorem, what is called central limit theorem, we should always get a gaussian distribution if you make enough measurements. So that is what the assumption here is that is the first assumption which we make when we go ahead and do that the fitting the first assumption that we make is that the error in the measurement of y_i is a random error and it is gaussian distributed around the true value which is f of x.

Now we also assume that, the standard deviation of sigma of these distributions are the same at all points for the time being we make this assumption that is, what we are saying is now if I change this x of my temperature to some other value and so I would change the y_i ,if I change x_i, I would change y_i, if I change x_i will change y in the true model also, so and the distribution around that value was also similar that is the same width and that is also similar that is another assumption we make.

It is not a very hard and fast rule and we do not need to make this assumption and we can do without this assumption but to simplify things in the beginning, let us make an assumption that this distribution is the same irrespective of what the value of x we choose, we choose another value of x the distribution will shift but it is only a shift and the functional form will remain the same that is what we assume. So that is what has been given here in short, we are saying that this probability for each of this x_i, let us forget this for the time being each of this x_i the probability is simply given by this, so the probability that.

So in short, what I am trying to say is that this probability is given by p of y or y_i will be is equal to exponential minus half, so f of x which I call y minus y of i, so that is what I am going to write divided by sigma square, that is what I wanted to say, so that is the function value which I wanted to get or the probability that I would find the y_i value between plus or minus delta y is given by this multiplied by delta y this multiplied by delta y will give me the probability that in a given measurement, I would get a value which is plus or minus delta y around the true value y that will be given by this.

(Refer Slide Time: 24:02)



So remember, the question here is i give you a set of, I give you a real model f of x, now the probability that my measurements y passes through or is close is within a distance delta, plus or minus delta y from the real value f of x is given by this. Now given a set of n points the probability that all this n points are within a distance plus or minus delta y from the real value f of x_i is given by this. So p is equal to, p is proportional to i going from one to n of this, so n data points of this. So that is the function.

So now, what we need to do is this f of x as we said, now the full probability is what we are writing now here as-, the full probability which we are interested in is p proportional to phi, that is the product function phi 1 to n. So I will say this is as f of x_i, so that is the real value and that is the product which we are interested in and this is what we want to minimize. So when I am given this, so now if I am given this data points x_i and y_i what I want to do is I want to now find out this function f as adjustable parameters a_i. I want to find out those set of values a_i such that this is minimal.

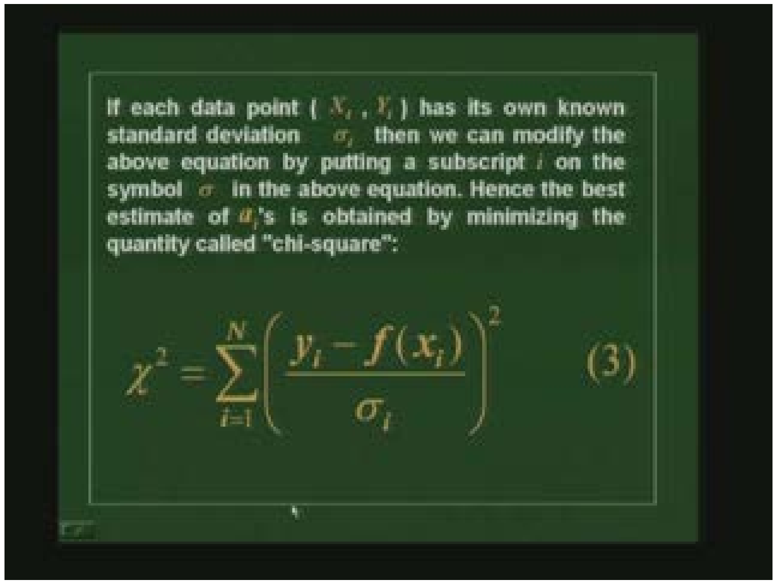
So that means I want to, this is the maximum that means I want to maximize the probability that it actually goes through the real function y_i that is what I wanted to do or I want to actually make the negative of this logarithm minimized, negative of this function minimized. So, I want to basically make sure that this is maximum that is why, I said negative of this logarithm maximized that is what we want to look at this.

So we want to make sure that this goes through-, the probability that it goes through a value plus or minus delta y from the real thing, real function should be maximized that means this minus of this quantity in the bracket which I am writing here should be-, we want to make it minimum. So that is what we wanted to do, so that is what we have written down here so what you want to do is to do this. So, maximum probability would mean that this quantity here, this whole quantity which is basically saying that I want to make this, whole thing in the bracket a minimum.

So exponential of this quantity with this minus sign here, so when you minimize this this function would be maximized, so that is what we wanted to do, so maximum probability would mean minimizing this whole quantity so and then we remember this f have all the adjustable parameters a_i, so what we want to do is we have to take this function, we want to take this function and minimize this with respect to the a_i this is not a function if a_i. So basically what we want to do is minimize this particular function that is sigma i going to 1 to n y_i minus f of x_i which also contains the a_i divided by 2 sigma square. We want to minimize this function with respect to a_i. So that is the essence of what we call the chi square fit.

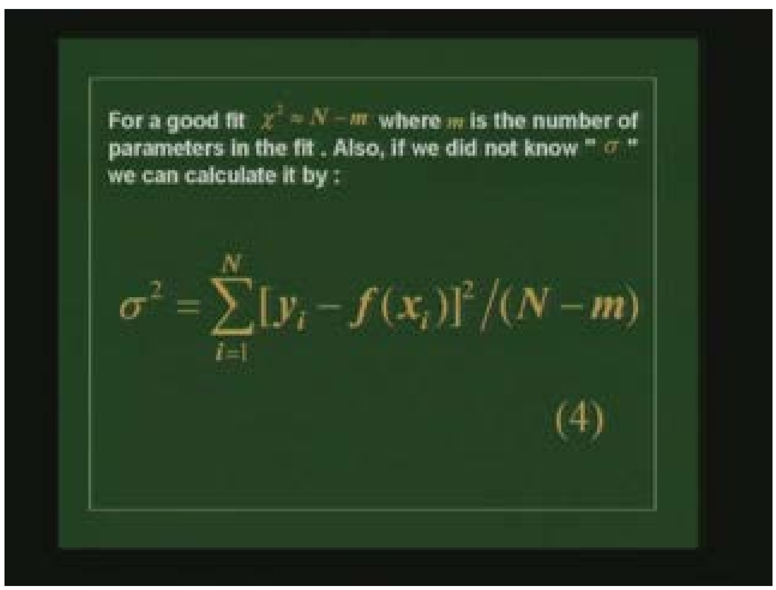
So now we are going to define the psi square as just that quantity in the bracket so basically this is my chi square. So, this quantity is what I am going to call as the chi square. So you can see that there is a chi square which I have written here, so it is, that it is that, every point x_i y_i now we could assume, now we could get rid of even the particular assumption that we made that each point x_i y_i had the same error, the same standard deviation we can get rid of that and we can say that each point x_i and y_i had a standard deviation sigma i and then I could define this now as y_i minus f of x_i divided by sigma i the whole square i going from 1 to n and that is what my chi square would be, what is called the psi square and now what I have to do is minimize this psi square with a function of a_i and that is called the psi square fit.

(Refer Slide Time: 26:58)



So we would call that psi square fit and you would see this almost everywhere in science and engineering. So now we can minimize this you said I can take this chi square.

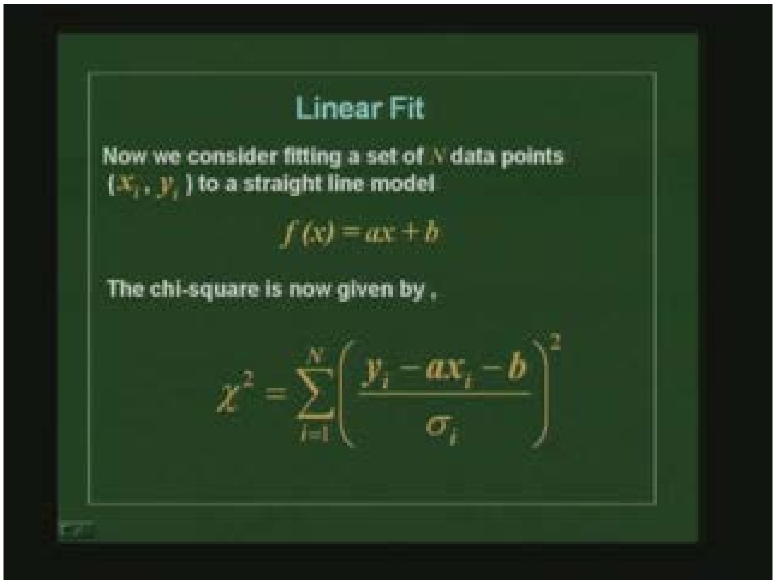
(Refer Slide Time: 28:01)



So what I am going to do is I am going to take this chi square which I defined as f of x_i minus y_i the whole square divided by sigma i square that is what my chi square is now f contains all the a_i and I am going to minimize this so basically I would do del by del a_i of chi square is 0 to minimize this function for all a_i and I would find out what the a_i values are. So now and then I find some a_i values and I put it back into this function and I can calculate this psi square, so how do I know which is, it is a good fit or not what is the kind of thumb rule, so the thumb rule is that for a good fit we should have psi square of the order of n minus m, where n is the number of data points and m is the number of adjustable parameters in the fit, so that is the chi square rule.

So remember, that we if we did not know sigma i and if we only knew, if we knew the real function f of x_i, f of x and the measured values y of i, y_i then we could have actually computed the sigma as in this fashion, y_i minus f of x_i whole square sigma i going from 1 to n divided by n minus m. So that would be our standard deviation, by standard definition. So now the question is, now we want to do this fitting, so we will take a particular function and then do this fit, so now we need to know what is the functional form of this, if we want to do this.

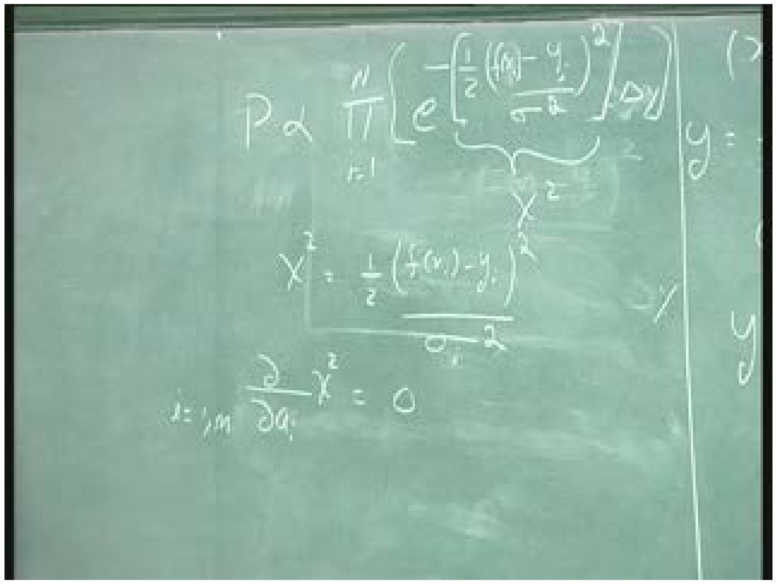
(Refer Slide Time: 29:38)



So we just so far, we said some general model f of x contains some fitting, some function, some parameters a_i and we can define a chi square like this from the measured value and the real function divided by sigma i square and I want to, I can minimize this chi square with respect to all the a_i, alpha i going from 1 to m. So I could minimize this m is the number of fitting parameters which I have and I could write this equation and I could solve this equation for the a_i and then, I would get a model.

Now, so far we have not said what f of x is so f of x could be a non linear function, a linear function it could be any of this. So now we will take a specific example and then do this specific set of functions, so first we will do is linear function. So we will assume that, f of x is a linear function of x and then we will actually solve this model and see what is the, how do determine the a values are.

(Refer Slide Time: 30:42)



That is what we are going to do now. So basic philosophy is now just this to do this, so we have taken here a set of values f of x_i, so we have f of x as a_x plus b, we took this f of x as a_x plus b and now we can define the chi square as that and then we can minimize this function that's what we are going to do.

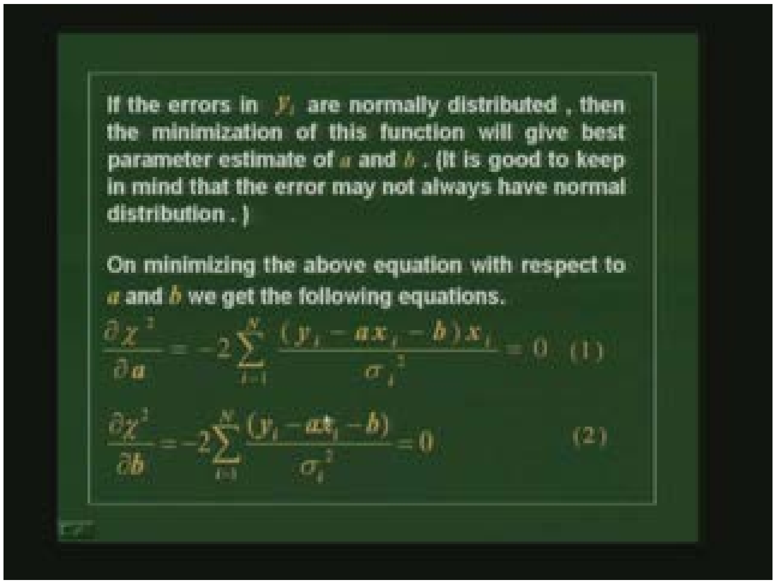
So we said f of x is a x plus b, so that is our function and we can define now the chi square would be then equal to y_i minus a x minus b the whole square divided by sigma i square that will be our function sigma, actually this two does not matter. So that is what you would have and then, we could just minimize this, so we can minimize del psi square by del a that is what first thing we should do and then a x_i and then we have to do del chi square by, so if you do this we would get sigma y_i minus x_i divided by that is not.

That is what we get if I just um minimize with respect to a and then I have another function which will be del chi square by del b so that will again sigma i, y_i minus sign. So that is the quantity which you would get.

So you want to equate these 2 to 0, if you do this and if you can solve this 2 equations, so you have to sum over I going from 1 to n all the data points 0 to n minus 1 or 1 to n and we want to solve these 2 set of linear equations in it is linear in a and b. So it is the set of linear equation we want to solve and we want to get a and b that is what we are going to do.

So that is summarized here so we just minimize with respect to a and b and we got these functions which is written about here so it is y_i minus a, x_i minus b by sigma i square into x_i and this two does not matter actually and then y_i minus a, x_i minus b by sigma i square and then now we will just, what we will do is we would come with some parameters to, we have to first compute y minus y_i by sigma i square x_i by sigma i square and y_i into x_i by sigma i square x_i square by sigma i square 1 by sigma i square and we sum each of them from 1 to n similarly, here and then we would set, write it in terms of a set of parameters on this fashion.

(Refer Slide Time: 34:19)

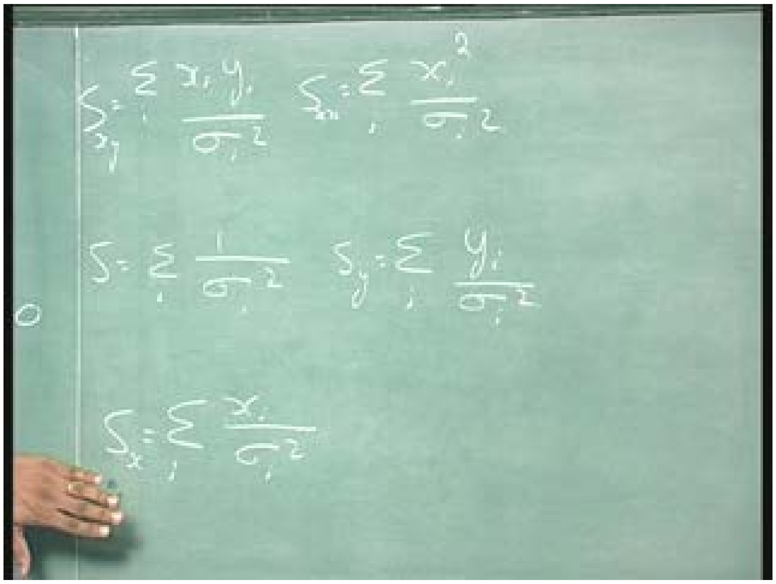


So now that is given here, so this equations which we had now are, we will write just concentrate on this, so this is, we would come with some new set of parameters which is s as sum over i from1to n 1 by sigma i and s_x as you know x_i by this is sigma I, so we will write all this functions here, so we would use y_i, so we want to have sigma i x_i y_i by sigma i square. We want to compute that and then, we want to compute sigma i x_i square by sigma i square and we want to compute sigma i 1 over sigma i square and we want to compute sigma i y_i by sigma i square, so and sigma i x_i by sigma i square.

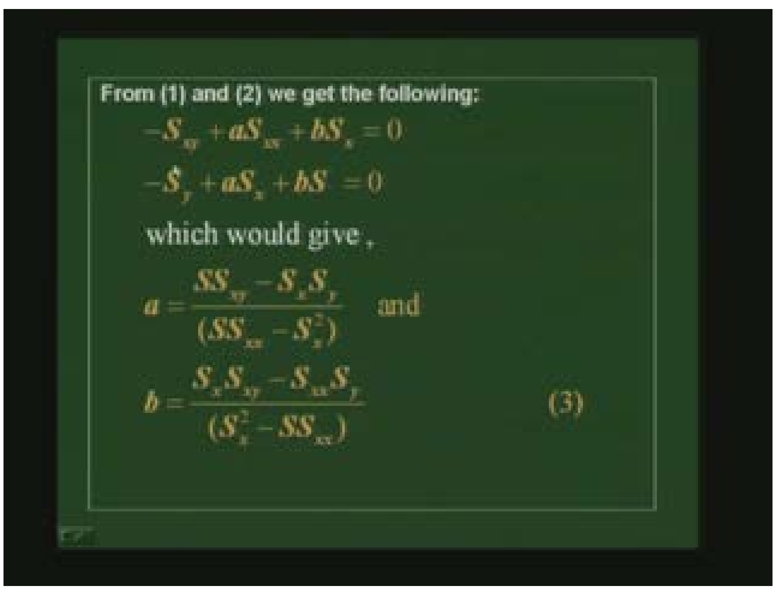
So all, that finishes all the quantities which we compute, we will first do that that will be the algorithm first we will do this x_i y_i by sigma i square because that are all known to us and x is square by sigma i square x_i by sigma is square and y_i by sigma i square and 1 over sigma i square and we call this as s_{xy} and we call this s_{xx} and this as s and this as s_y and this as s_x. So now I can rewrite these set of equations in terms of these parameters.

So I can write down I can use these this functions and then plug them in into my main equation and write it, put it in this equation and write it in a different fashion. So I will just substitute all that and then I will get a set of equations which are given by this. So if I just do that in terms of s_{xy}, s_{xx}, s_x, s_y, ss_x and s, I can write it like that also two linear equations and I can solve them and get a and b that is the algorithm. So it is easy to compute in that way.

(Refer Slide Time: 36:19)

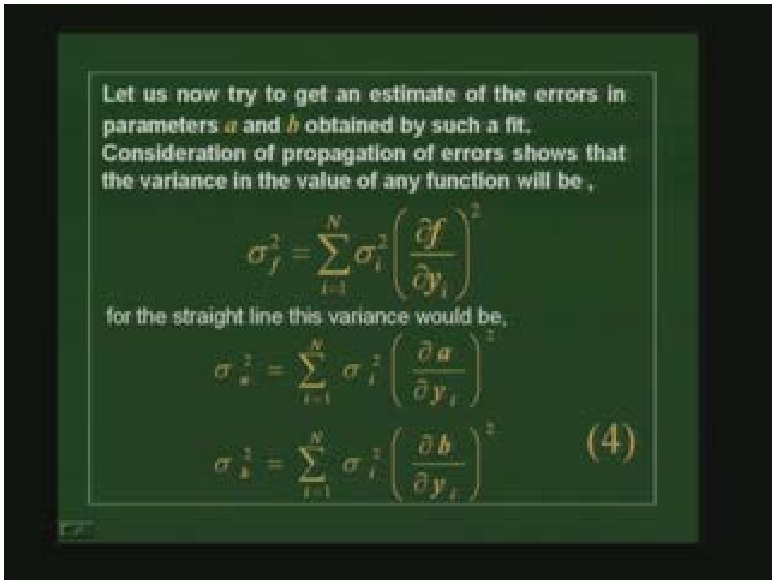


(Refer Slide Time: 36:44)



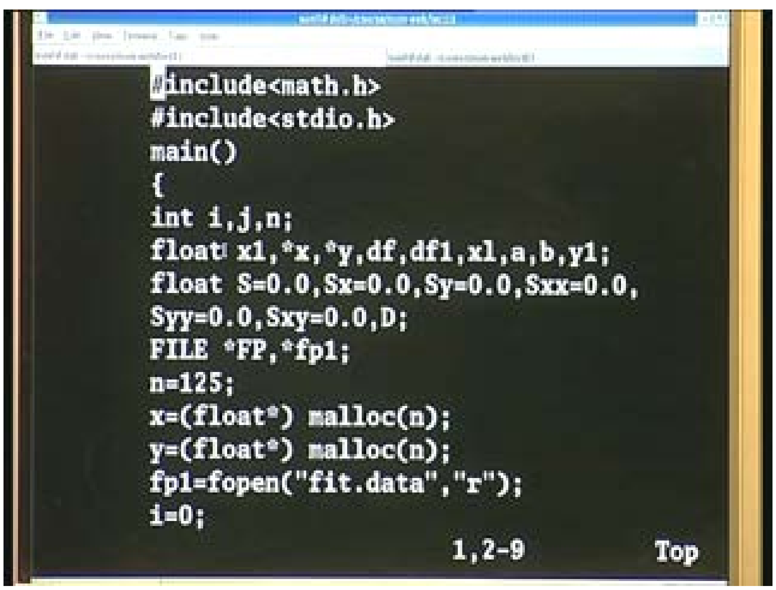
So for a general linear fit, we can compute it using these two linear equations and I can solve for that and then I get a and b as s into s_{xy} minus s_x s_y divided by s into s_{xx} minus s_x square and the same denominator here and s_x s_{xy} minus s_{xx} s_y divided by s_x square minus s into s_{xx}. So I have defined a and b in that way, so I can solve this equation and then get a and b. So then you know once you have this functions a and b then the next question would be what is the, how reliable this a and b is and so we need to define what is the error or what is the deviation in the a and b itself is so that we would get that would be, that is what we would next go into but, before we go into that we will actually see the implementation of this particular form of linear fit in a program.

(Refer Slide Time: 38:00)



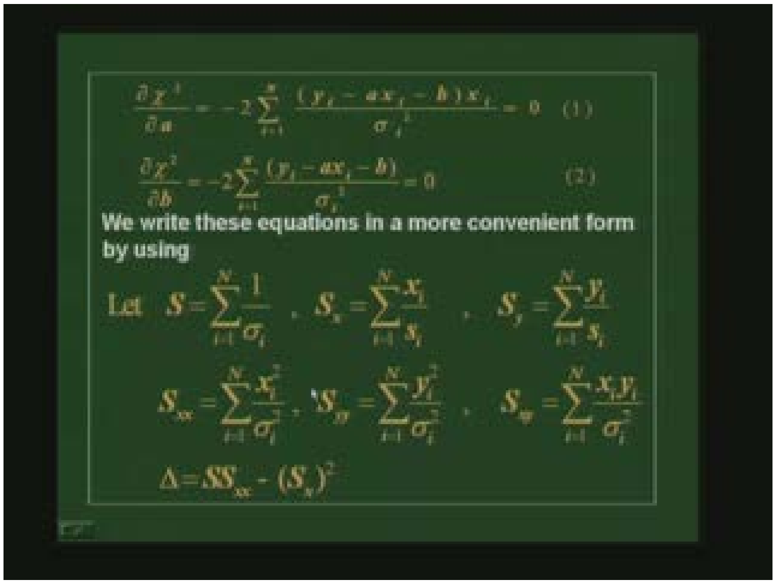
So that is what we will do next, we will now look at the implementation of this fitting a linear function into a given set of data as we go along we will also pay attention to the programming details.

(Refer Slide Time: 38:16)



So here is the program which would implement the following. So we have a set of function values given at x_i and y_i. So instead of an array this program I am using a pointer x and y and then as I said in the case that what we want to fit is a function f of x is equal to a_x plus b, so we had defined our psi square as y minus a_x minus b whole square divided by sigma i square and we have to minimize this chi square with respect to a and b and we now, and then we get this set of equations and to program this equation, so to write an algorithm for this solving this we first define y_i by sigma i, 1 over sigma i squares and y_i by sigma i squares and x_i by sigma i square etcetera in this set of functions s, s_{xy}, s_{xx}, s_{yy}, s_{xy} etcetera.

(Refer Slide Time: 39:26)



And then, so we want to compute these quantities first so that is what we are trying to do first so here is that code so we have defined them as all of them as floating point numbers and we set they are all, set equal to 0 to begin with and then, now we want to read this x and y the set of points x_i and y_i from some file data files I call that data file as fit dot data and then I can open that file. so I open that file using this function pointer the file pointer fp 1 but now I want to read that into an array which is x and y.

So before I can put it into x and y which is a pointer here I had to allocate memory to that, memory for that pointers and that is what I am doing here using this malloc function so, floating point, so I am allocating memory to that using float star malloc of n, so that is what I am doing here. So we have gone through this before, so I am just showing the implementation of that in a real program. So here is the x and y being allocated some memory and then I use scanf function, the fscanf function after opening this file I use scanf function to actually read this the values, so the function values are tabulated in what is called fit dot data, so here is in this form I have tabulated this function values the x_i and y_i.

So I have them here. So I am just reading that off in this using this fscanf function, since they are floating points I use percentage f to read them to read them into two variables x_1 and y_1 which I declared as floating and I store them in this arrays x_i and y_i which is now I store them because, I have allocated memory into that x and y, so that is important we need to allocate memory first and then I store them into that, so I just print that out when I run the program I just print it out on the screen the x_i just to see that everything is all right and then I have the red values and then I again print them the x_i and y_i here so, just to show you that what we have read is exactly the same as what has been put in just to help us in debugging the program if there is any bugs.

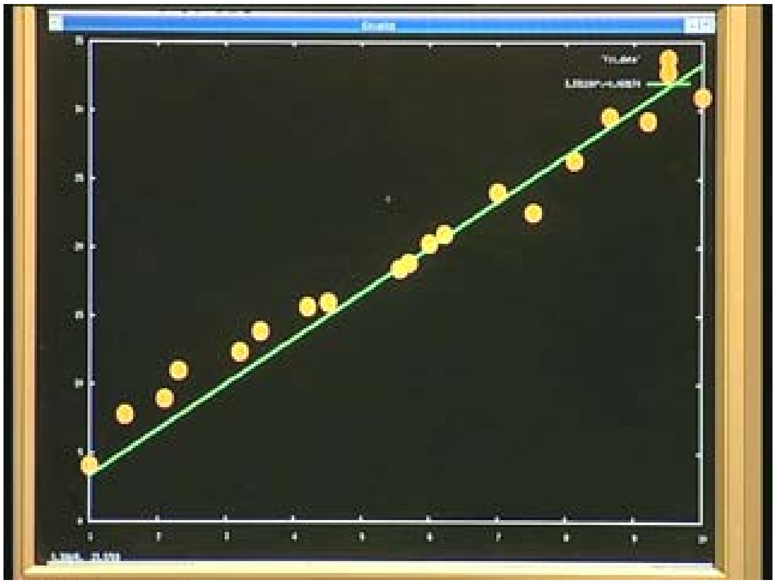
So now that is a little programming detail then, we go into the calculation of s, s_x, s_y. So we have initialized all of them into 0 then, we compute the sigma in this data given to us sigma has not been given, so we assume all the sigma to be 1 all the sigma is the same because, we just assume them to be 1. So we could any number in this thing sigma all of them are the same. So we assume them to be 1 and then we just-, so i just put them as that itself, so we can just put s as s plus 1 s as 1. So I just have to sum over all the i's, this is 1 and then we have s_x and s_y so s_x being, since sigma is 1. I just said s_x is x of i. So remember s_x is x of i by sigma i.

So this is sigma I, s_x by sigma i but then sigma i is 1. So it is just x s_x is x. Similarly, s_y is y_i and s_{xx} is x_i into x and s_{yy} is y_i into y_i and s_{xy} is x_i into y_i. So this is just summing over all the points, remember we have to initialize them first before you do this sum and we have done in here in the declaration itself. Then we compute the denominator so first we need to, this is the solution which we get, we know that we substitute all that we get an equation like this and the solution is this. So we compute the denominator of this function and then that is first and then we can compute a and b, so I computed the denominator here and I can compute a and b like this function so that is what we will do and then we will look at the fit of this function.

So that is what we will just run this program now. So we compile that and we run this program and we get the a and b values as, so this is just printing out the x_i and y_i values and the a and b values are "3.333" and ".069174" that is the a and b values which we have got and we could plot this. So I just plot that now and that is the plot, so this is the value and so that is the again the circular points here are the values which are tabulated and the straight line going through is my fit for the given value a_i b_i. So then, we can actually, so we have a and b and so we have a straight line.

So given for this spread of values with same sigma i's, we said all of them are all the points here have equal reliability and then we have this straight line which goes through this points, which is the best fit to this point, best chi square fit to this point the set of points, that is what we have now next question obviously is what is the error and how good is my fit, so that is the question which you would ask now we have an estimate for a and b we have got a and b, now can we can we try to estimate errors in the parameter a and b given, obtained by such a fit. So then, we have to recall something which we learned in the beginning lectures that if there is a function which is a function of x, let us say and if there is an error in x given by some delta x and then how does that error in x propagate to the function value itself.

(Refer Slide Time: 45:14)

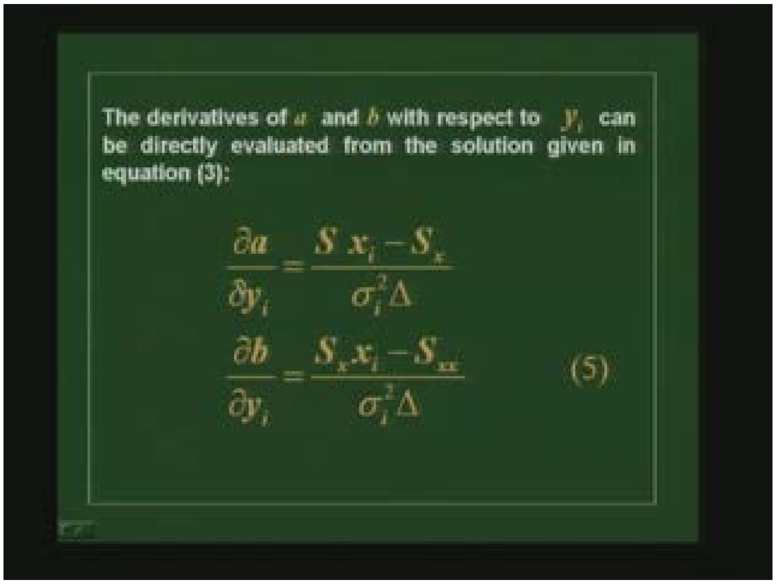


That is something which we learned before and we said that if there is a function of x and x is a floating point representation in that context we had learnt that. So there is an error in the representation of x which is given by delta x, let us say and then the delta f that is the error in the function value f because of the error in x is given delta f is given by delta x into derivative of f with respect to x that is what we have seen so same thing which we can use here we are looking at what would be the sigma in the function and if given some function f given that the variables had an error sigma i and then we would say that it is you know sigma f square is sigma f would be given by sigma i into derivative of that thing, we are taking square the here but derivative of that so sigma f would be sigma i into derivative of that function with respect to i.

so here we are looking at the errors in a and b, a and b obtained now as functions of x_i and y_i that is what it is finally. So the question is what is the spread in the y_i how does that affect or how does that give us a spread in a that is what will be given by sigma a square, so sigma a square is then sigma i square which is the spread in the y_i's multiplied by derivative of a with respect to y whole square. Similarly, the spread in the b sigma b square would be the spread in the i, a spread in the y which is sigma i square into derivative of b with respect to y square. That is what we would get as sigma a square and sigma b square. So remember, we can actually if you go and look back, look at this you see that a and b are actually represented as s, s_{xy} and s_x, s_y etcetera are these half functions of y_i.

So these are functions of y_i, so we can actually compute the derivatives of a with respect to i and derivative of b with respect to i from these functions and it is possible to compute this values of sigma a square and sigma b square. So that is, I am having so because, we can actually evaluate derivative of a with respect to i from that expression which we just saw and second derivative of b with respect to i and so using this we can actually sigma a and sigma b.

Refer Slide Time: 48:36)



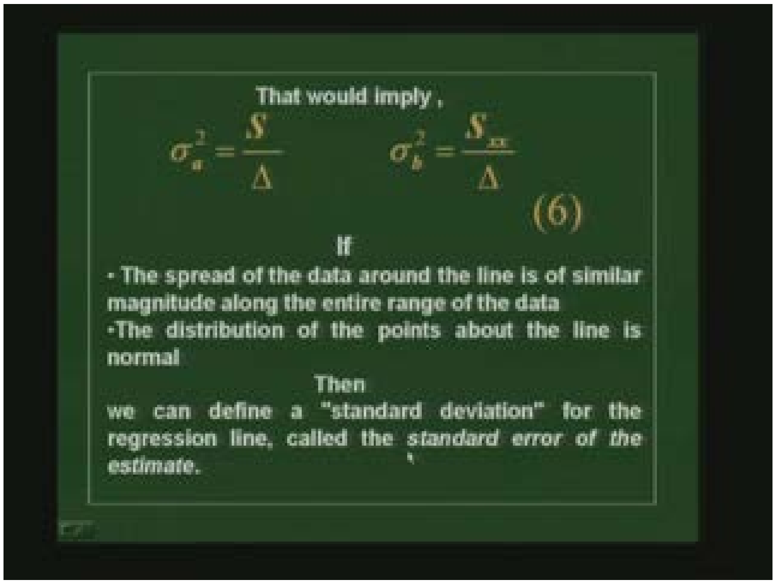
So now, if you do that computation then we would find that sigma a square is simply given by s over delta or delta is that s_{xx} minus s_x square which we just saw so and sigma b square is s_{xx} by delta. So we can actually not only get estimates for the values for a and b we can even make estimates for the error in a and b in sigma a square and sigma b.

Now, what is the implication of this is that if the spread of the data around the line is similar in magnitude along the entire range of the data or the distribution of the points along the line is normal, now if these two why I have set an, if here is because these two have been assumed in our actual derivation of the data.

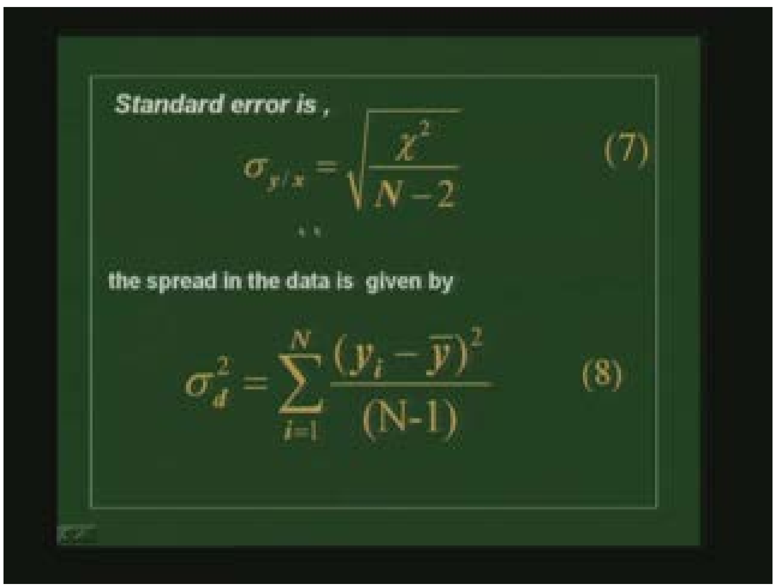
So this two have been assumed in our actual derivation of the data, so the derivation of the, that is the spread is normal the y_i is normally distributed that it is gaussian we assume that. Then we can define a, now we can define a standard deviation for the regression line that is fit for that we can define a standard deviation and it is called the standard error of estimate.

So summarize, we see that we can not only compute this a and b, we can also compute the errors in a and b, that is sigma a square and sigma b square and we can write them as a very simple function of in terms of s and delta as s by delta and s_{xx} by delta and then we can define what is called a standard error for the fit and that is given by chi square by n minus two.

(Refer Slide Time: 50:00)



(Refer Slide Time: 50:29)

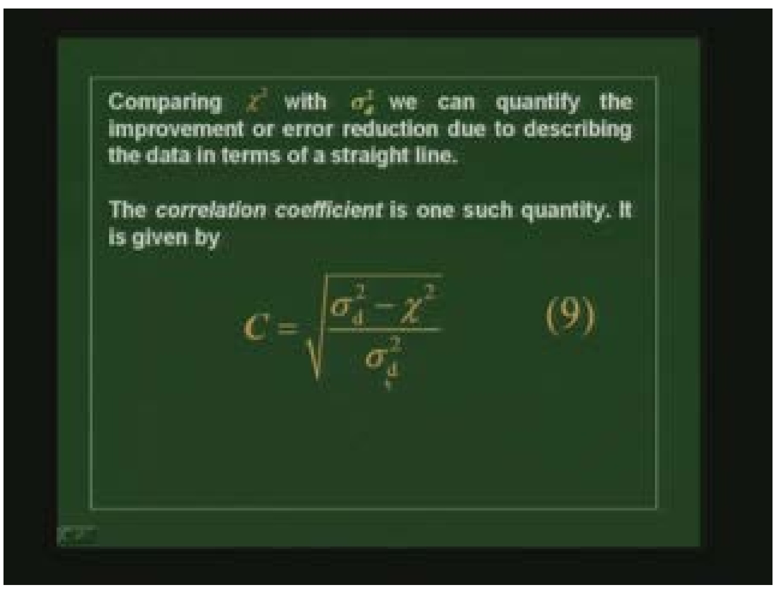


Now that is the definition of the standard error and this error is, we can define that only if the y distribution is gaussian that is for a given x value if you make measurements of y and if the y distribution is gaussian then we could define a standard error as given by sigma s chi square by n minus 2 for the fit. So psi square is something which we get from the fit and then the spread in the data, we can then we can define another quantity called a spread in the data which is y_i minus y average square by n minus i square, y average would be for a given x as we said there is a spread which is um gaussian and you take an average of that. So that is the y_i and y_i is one measured value and that gives you the measure of the spread.

So now given these two quantities, so remember sigma d here is the spread which means that for a given value of x that is fixed temperature. We make many measurements and then we take an average and we will take many measurements we get a spread, we get various different values of y_i and that y_i is we believe are gaussian distributed and that, and then we can take an average of that and then the average at y_i minus y average whole square divided by n minus 1 is what would be what we call the spread in the data.

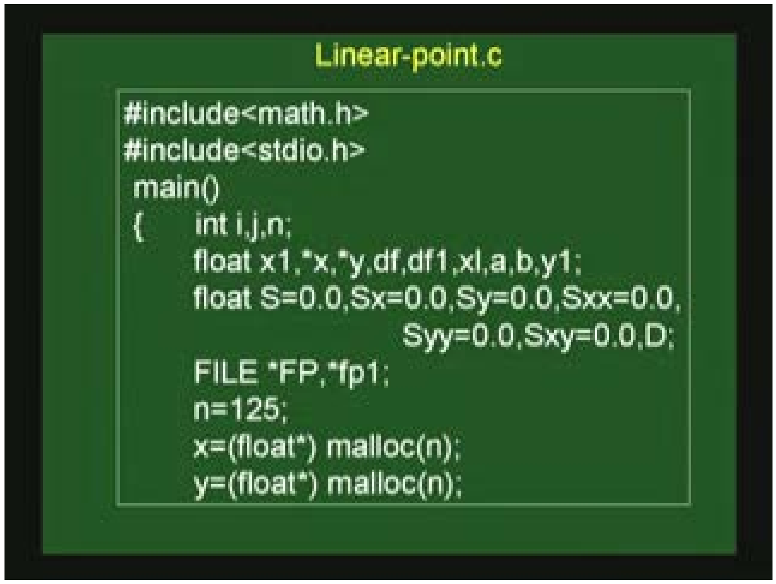
So y_i minus y average square, that is like mean square average, mean square distribution mean square y value so given this chi square and sigma d square, we can define another quantity called a correlation coefficient and that is given by sigma d square minus psi square by sigma d square. So I am just giving you a set of quantities which is computed as a standard thing, when you do a fit, so one is a correlation coefficient for which all we need to know is the sigma d square and chi square and so then we can compute sigma d square minus chi square by sigma d sigma d square and then we can also compute what is called a standard error which is chi square by n minus 2 and we can then, we can compute the errors in the a and b as sigma a square and sigma b square.

(Refer Slide Time: 52:40)

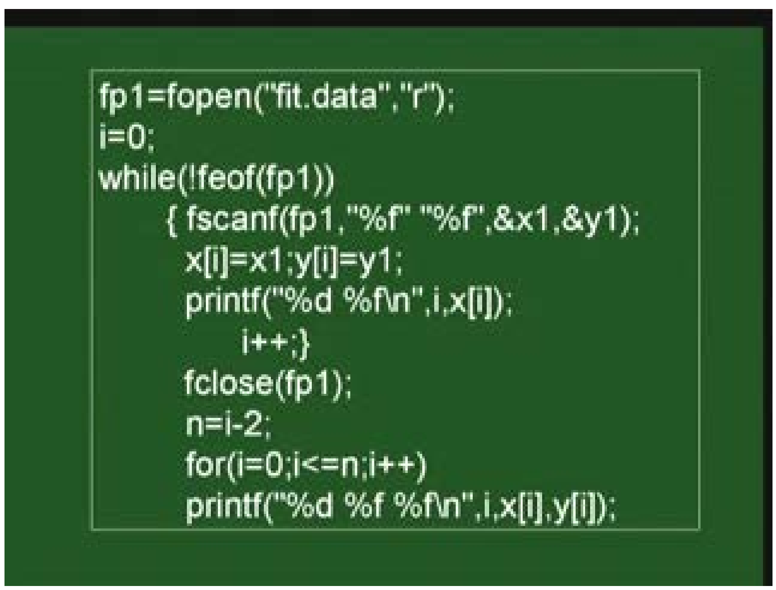


This is all apart from just computing a and b itself which is the fit to the given set of data for a function f of x f is equal to a_x plus b. So we can generalize this idea into a general linear fit so now we had just done this particular quantity that f of x as a_x plus b, we will generalize this fit this algorithm into a general linear fit in the next lecture that is what we will be looking at in the next lecture.

(Refer Slide Time: 53:33)



(Refer Slide Time: 53:48)



(Refer Slide Time: 54:04)