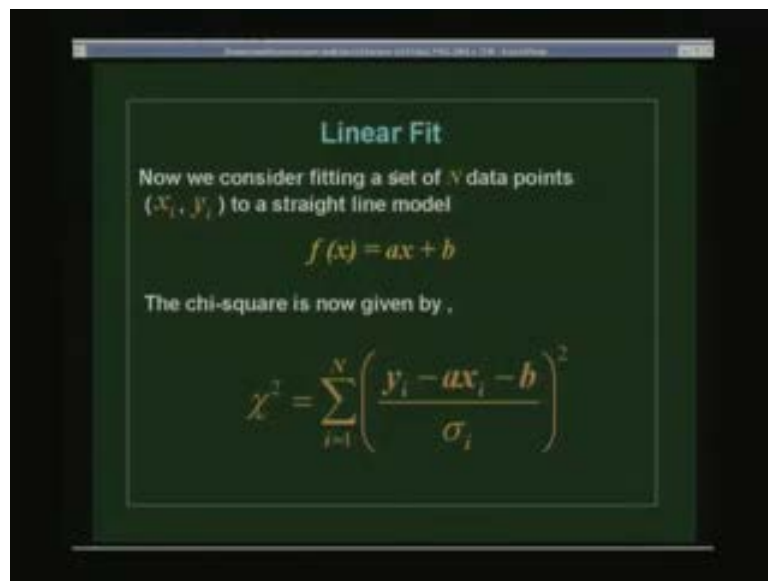


Numerical Methods and Programming
P.B. Sunil Kumar
Department of Physics
Indian Institute of Technology, Madras

Lecture No. # 14
Data Fitting: Linear Fit

Today we will continue on the methods of modelling data. So, last class we saw how we could fit a straight line through a set of data points given. So, we will continue on that and error in this fitting, and how to quantify how good is your fit or your model for a given set of data points that is what we will continue again today.

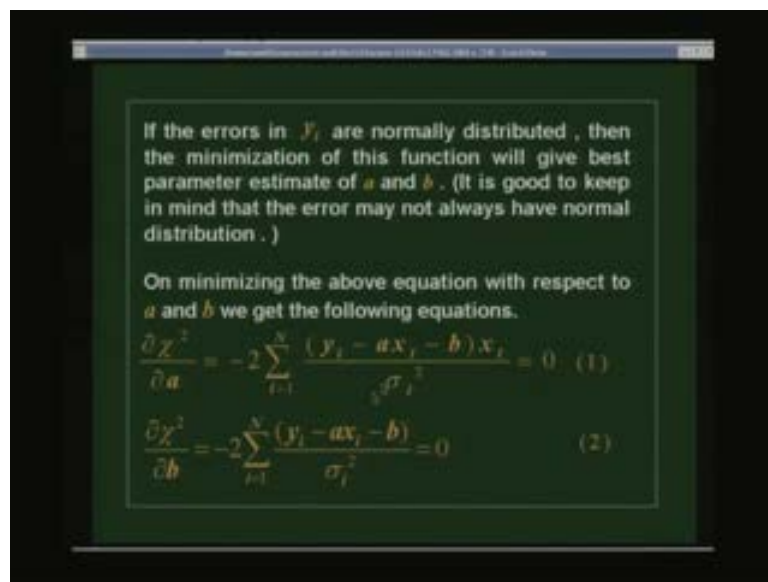
(Refer Slide Time: 01:50)



So, we were looking at linear fit that is what you are doing in the last class. So, there is we had a set of N data points and which has labelled x_i and y_i on through which we had fit is straight line and of this type ax plus b and the idea was to determine the parameters a and b such that this is a best fit to this data. So, remember we have fixed this form of the function all we do in fit is to determine the values of the parameters. So, that we used what is called we had a definition for a function called chi square which is the difference between the data points given and the function square. So, that is y_i minus f of x_i divided by σ_i .

So, σ_i represented the measurement of the reliability of each of this data points x_i and y_i other we define this function as y_i minus f of x_i divided by σ_i square summed over all data points and we said that minimizing this. Our idea is to minimize this function with respect to a and b and that would be, that is what we consider as a best fit. If this function is true the presentation of the set of data points then they should be chi square should be very small and we show that chi square of order of n minus m , m being the number of in this case two is n minus 2 is a good means is a good fit that is what we looked at and the way we did was to write down the derivatives of chi square with respect to a and b the two parameters and equating them to zero that is minimizing the chi square with respect to a and b and then we get a set of equations which we can solve to get a and b .

(Refer Slide Time: 03:52)



So, we had this couple of this equation this two equations which we can solve to get a and b and solve this we determine we wrote down some definitions for some symbols forremember here y_i x_i and σ_i are known to us what we do know is a and b .

(Refer Slide Time: 05:22)

$$\frac{\partial Z^2}{\partial a} = -2 \sum_{i=1}^N \frac{(y_i - ax_i - b)x_i}{\sigma_i^2} = 0 \quad (1)$$

$$\frac{\partial Z^2}{\partial b} = -2 \sum_{i=1}^N \frac{(y_i - ax_i - b)}{\sigma_i^2} = 0 \quad (2)$$

We write these equations in a more convenient form by using

Let $S = \sum_{i=1}^N \frac{1}{\sigma_i^2}$, $S_x = \sum_{i=1}^N \frac{x_i}{\sigma_i^2}$, $S_y = \sum_{i=1}^N \frac{y_i}{\sigma_i^2}$

$S_{xx} = \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2}$, $S_{yy} = \sum_{i=1}^N \frac{y_i^2}{\sigma_i^2}$, $S_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2}$

$\Delta = SS_{xx} - (S_x)^2$

So, we can write sigma i going from 1 to N y i by sigma i square as one quantity which we call y i into x i by sigma i square which we call Sxy and then sigma i going from 1 to N x i square x i into x i by sigma i square which we call Sxx and sigma x i by sigma i square which we call Sx. And similarly here we had sigma y i by sigma i square which we call Sy and sigma x i by sigma i square which we saw is Sx and one over sigma squared which we call one over sigma i square which recall S and with these definitions which has summarized here and substituting these definition into this set of equations we could get a set of equations in terms of S which is Sxy plus a Sxx plus b Sx equal to zero and minus Sy plus a Sx plus b S equal to zero.

(Refer Slide Time: 05:32)

From (1) and (2) we get the following:

$$-S_{xy} + aS_{xx} + bS_x = 0$$
$$-S_y + aS_x + bS = 0$$

which would give,

$$a = \frac{SS_{xy} - S_x S_y}{(SS_{xx} - S_x^2)} \quad \text{and}$$
$$b = \frac{S_x S_y - S_{xx} S_y}{(S_x^2 - SS_{xx})} \quad (3)$$

Equation (3) gives the best-fit model parameters a and b .

So, these two set of equation ns we could solve for a and b and write a and b solutions as thus. So, now I mean once we have written it in this fashion you can see that if you given the values x_i and y_i is easy to compute these quantities a and b and we could get the best fit parameters a and b. So, that is what we could do and then we said that... Now, let us get error estimate for this object. So, now, we have the parameters a and b and now we want to get an estimate of the error. So, for this in these parameters a and b.

(Refer Slide Time: 06:35)

$\sigma_a^2 = \sum_i \sigma_i^2 \left(\frac{\partial a}{\partial y_i} \right)^2$

$$a = \frac{SS_{xy} - S_x S_y}{(SS_{xx} - S_x^2)}$$
$$= \frac{SS_{xy} - S_x S_y}{\Delta}$$

So, for that we use the definitions of we call this parameters that errors in the fit parameters as σ_a and σ_b . So, that is what we had used and we said that σ_i we had this definition for propagation of error and which we said that if there is an error and a parameter or floating point x then how does that error propagate into the determination of value for a function f of x and we had the definition for that and we use the same definition in determining σ_a and σ_b . So, with using that definition we could write σ_a as σ_i which is now the error in the value y_i . So, this the x_i y_i and this is the error in the determination of y_i itself and then could say $\frac{\partial a}{\partial y_i}$ that is what kind of definition which we have written before. So, in this case we would define this as $\sigma_a^2 = \sigma_i^2$ similarly would write an expression for σ_b^2 . So, that is what we have here is goes to [] summarized here.

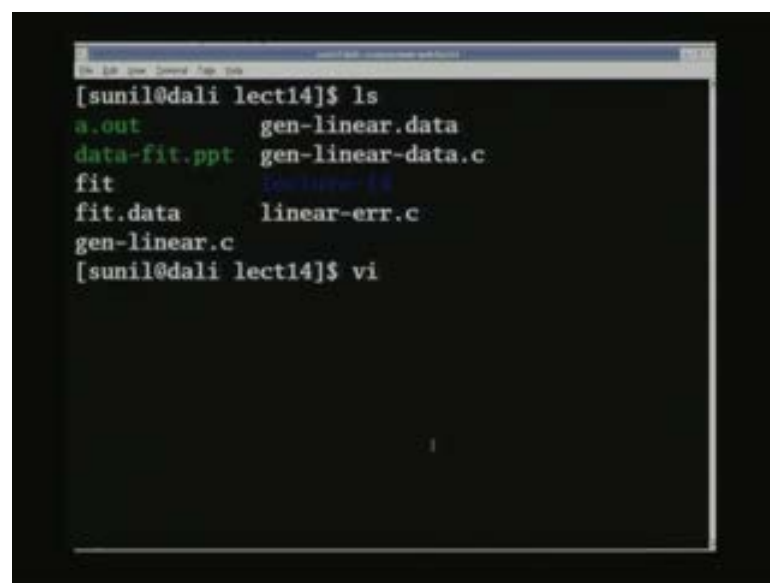
So, we had in general any function f would have the $\sigma_i^2 = \left(\frac{\partial f}{\partial y_i}\right)^2 \sigma_i^2$ the partial derivative f with respect to y_i into σ_i which is the error in the thing itself. So, here instead of that now we have the order determinant the error in the a . So, then we would determine that as the error in the y 's and multiplied by the dependence of the derivative of a which is now consider a function of y and derivative of a with respect to y remember if you look at this definitions you can see that this a is a function of y that because S_{xy} and S_y etcetera contains y_i . So, a is a function of y similarly b is a function of y and we know what is the reliability of the value y_i . So, using that we can actually determine the error in σ_a^2 . So, that is...

So, with that we complete the determination of a b and it is error and then we could write down σ derivatives in form like this. From this definition of a and b I can derivatives in this form and substituted back into the σ_a^2 and σ_b^2 values and then we could get the determine the error. So, when we will see that σ_a^2 then simply when submit the substitutions you will find the σ_a^2 is simply S by Δ and σ_b^2 is S_{xx} by Δ where Δ is our S_{xx} square is this simply this quantity S into S_{xx} minus S_x square. So, from this from we have this expression for a as S into S_{xy} minus S_x into S_y divided by S_{xx} minus S_x square and now this are functions of y . So, we could determine a as derivative of a with respect to i and then I can write. So, this is written as $S S_{xy} - S_x S_y$ divided by Δ .

So, that is a definition of Δ and then I can write from this I can write using now I will find the derivative of this with respect to this a with respect to i and then substitute that

here and then I can write finally, that sigma a square as simply S by delta that is what you would get S that is a function which you writing here S by delta and similarly sigma b square as Sxx by delta and now this is we would see. So, now, from define this two derivatives as has this two errors .We can also define what is called as standard deviation and if you do not know the sigma values we could get that we could also get an estimate called standard error .The standard error is chi square by n minus two. So, we will complete our linear regression of the linear fit using these three obtaining a b sigma a square sigma b square and the sigma y x which is the root of chi square by n minus two.

(Refer Slide Time: 12:09)

A terminal window screenshot showing a file listing command and an editor command. The prompt is [sunil@dali lect14]\$. The ls command lists files: a.out, data-fit.ppt, fit, fit.data, gen-linear.c, gen-linear.data, gen-linear-data.c, and linear-err.c. The vi command is entered at the end.

```
[sunil@dali lect14]$ ls
a.out          gen-linear.data
data-fit.ppt  gen-linear-data.c
fit           linear-err.c
fit.data      linear-err.c
gen-linear.c
[sunil@dali lect14]$ vi
```

So, that we complete the fit to a linear function and we would just see an implementation of this gen-linear program that is what you would look at now. You just look at this implementation in a program.

(Refer Slide Time: 12:16)

```
for(i=0;i<=n;i++)
{
    chisq=chisq+(a*x[i]+b-y[i])*(a*x[i]+b
-y[i]);
}
printf("        %d xhisq = %f Standard
error = %f\n",n-1,chisq, sqrt(chisq/(n-1)));

printf("        sigma_a = %f sigma_b
= %f\n",S/D,Sxx/D);

}

"linear-err.c" 44L, 924C      44,0-1      Bot
```

So, here I try to implement this whole idea into this program. So, we defined this is c program. So, we have an array x and y here I use an array x and y to store the data which we have which for to which we have to set `y` the linear function and this data is stored in some file called fit dot data.

(Refer Slide Time: 12:20)

```
#include<math.h>
#include<stdio.h>
main()
{
    int i,j,n;
    float x1,x[25],y[25],df,df1,x1,a,b;
    float S=0.0,Sx=0.0,Sy=0.0,Sxx=0.0,Syy=0
.0,Sxy=0.0,D;
    float chisq;
    FILE *FP,*fp1;

    fp1=fopen("fit.data","r");
    i=0;
    while(!feof(fp1))

1,1-8      Top
```

So, then and I define here as floating points the functions the value function S Sx Sy Sxx Syy and Sxy and d which is my delta. So, this is this defined as floating point numbers and then I open a file here which is been declared as a file pointer here fp1 I open that

file and I read using this f scan function all the x_i and y_i values and then... So, then I get the total number of points in that and for that i do the sum over i going from 1 to n 0 to n in this case 0 to n minus 1 in this case of all the data points to get is S S_x S_y and S_{xx} that is what I am doing here.

So, S is simply one over σ_i^2 and σ_i since in this data the σ_i 's are not specified in that kind of circumstances we assume σ_i to be equal to one all data points are equally reliable and we put σ_i equal to one and then S is now which is $\text{sig}[ma]$ summed over σ_i^2 summed over i will become just simply summed over i summed over one, because one by σ_i^2 is now one. So, S is this sum over i one and the S_{xx} is now x_i^2 divided by σ_i^2 becomes S of x which becomes just simply sum over i x_i^2 .

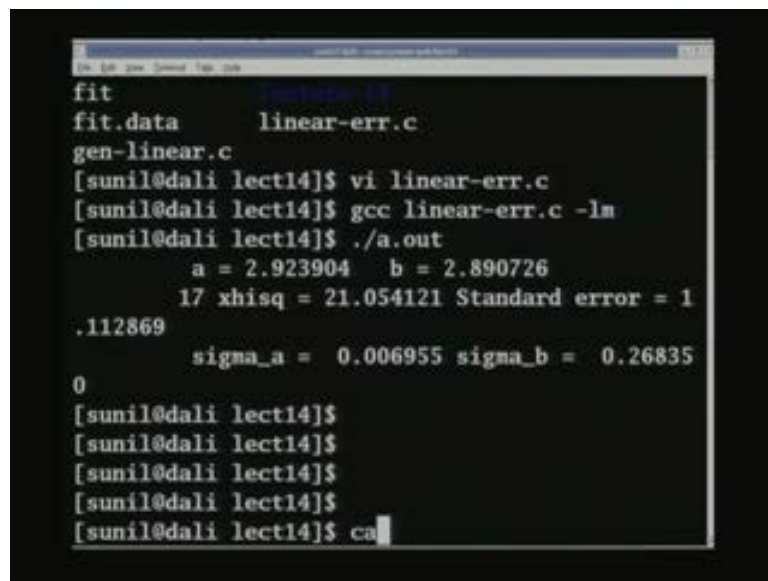
So, this for loop that is this for loop grants from 0 to n and then keeps sum in x_i y_i and x_i^2 into x and then y_i into y and x_i^2 into S_{xx} . So, that gives you all the S_x S_{xx} S_y S_{yy} and then I could compute the d which is the delta which is s and S_{xx} minus S_x^2 and I could compute a and b once I compute it a and b and then I can compute the chi square now. So, I can compute chi square here since $(())$ So, the chi square now is the sum over i going from 0 to n minus 1 of the function value. So, the function value which is $a x_i + b - y_i$ whole square. So, let us multiply by itself. So, that is a chi square.

So, I having determine a and b and determine what is the chi square now x I know that I determine a and b by minimizing chi square. I determine a and b this function a and b by minimizing chi square that is the way I get this expressions then I need to know what is the chi square value what is the minimum chi square is and that is what I am computing here. So, I am just after determining a and b here I determine the chi square value as a of $a x_i + b - y_i$ whole square that is $a x_i + b - y_i$ into $a x_i + b - y_i$ and then I print out using this printer function on to the screen the chi square value and the standard error. Now, the standard error is now as you can see is just simply given by square root of chi square by n minus 1. So, there is why my standard error S .

So, that I print out the standard error I print out the chi square value and the n minus 1 itself and then I could have the σ_a 's and σ_b 's and σ_a is now given by S by delta and $S \sigma_b$ is given by S_{xx} by delta. So, that will give as a reliability of a and

b how good the function S. So, that is what we are going to do and we will just run this program and see what we get. So, we will compile this and run it . So, that gives us a as two point nine two three and b as two point eight nine the values which I got for a and b and the chi square n minus 1 is seventeen the number of data points now is about is eighteen and then the chi square which is of and the two eighteen data points and we have two sitting parameters.

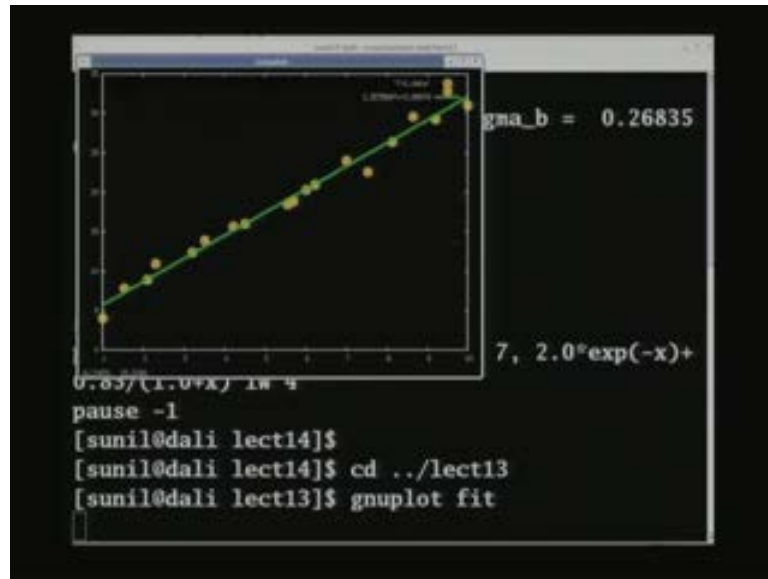
(Refer Slide Time: 17:47)



```
fit
fit.data      linear-err.c
gen-linear.c
[sunil@dali lect14]$ vi linear-err.c
[sunil@dali lect14]$ gcc linear-err.c -lm
[sunil@dali lect14]$ ./a.out
      a = 2.923904   b = 2.890726
      17 xhisq = 21.054121 Standard error = 1
      .112869
      sigma_a = 0.006955 sigma_b = 0.26835
0
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$ ca
```

So, the chi square is around twenty-one which is not too bad and your standard error are one point one one and sigma a that is a reliability of a is point zero zero six nine and b is point two six that is what we get from this and we could see how the data would be fitting into this we will say see that by fitting that line into the... That is a fit which we have here.

(Refer Slide Time: 18:06)



So, we have the data points spread around here and that is our linear fit using $ax + b$ function that is what we see with this I think.

(Refer Slide Time: 18:22)

```
Linear-err.c
#include<math.h>
#include<stdio.h>
main()
{ int i,j,n;
  float x1,x[25],y[25],df,df1,xl,a,b;
  float S=0.0,Sx=0.0,Sy=0.0,Sxx=0.0,
        Syy=0.0,Sxy=0.0,D;
  float chisq;
  FILE *FP,*fp1;
  fp1=fopen("fit.data","r");
  i=0;
```

So, now, we go back and to look at the little more general fit form of this. So, we will now extend that is what we mean by that if you will extend the function which we were using as.

(Refer Slide Time: 18:33)

```
while(!feof(fp1))
{
    fscanf(fp1,"%f" "%f",&x[i],&y[i]); i++;}
fclose(fp1);
n=i-2;
for(i=0;i<=n;i++)
{ S=S+1; Sx=Sx+x[i]; Sy=Sy+y[i];
  Sxx=Sxx+x[i]*x[i];
  Syy=Syy+y[i]*y[i];
  Sxy=Sxy+x[i]*y[i];
}
D=S*Sxx-Sx*Sx; a=(S*Sxy-Sx*Sy)/D;
```

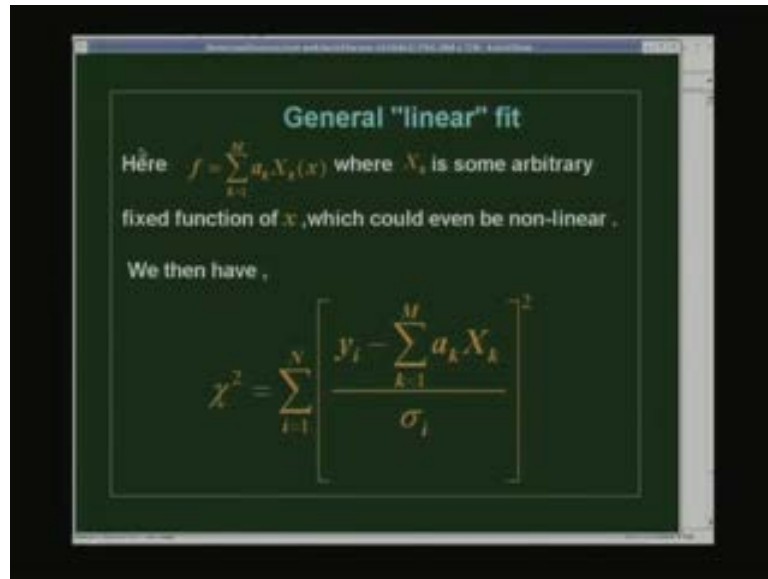
So, for we using trying to fit functions of the form f of x is equal to ax plus b that is f of x plus ax plus b .

(Refer Slide Time: 18:42)

```
b=-((Sx*Sxy-Sxx*Sy)/D);
printf("    a = %f  b = %f\n",a,b);
chisq=0.0;
for(i=0;i<=n;i++)
{
    chisq=chisq+(a*x[i]+b-y[i])*(a*x[i]+b-y[i]);
}
printf("    %d xhisq = %f Standard error = %f\n",n-1,chisq, sqrt(chisq/(n-1)));
printf("    sigma_a = %f sigma_b = %f\n",S/D,Sxx/D);
}
```

So, we will now look at little more general case that is we will try to fit functions of this

(Refer Slide Time: 18:59)



one that is some we could say that sigma i a i some functions of x i .Now the a of x,x could be a capital X which I have used is could be any function of x i that is need not be a linear function it could be even non-linear function could be sin of x cos of x or exponential x all the parameters are here not inside this function. This is a linear combination of many non-linear functions is possible, but it is still as for as a is concerned is a linear this f of x is concerned is still linear in a. So, that is general linear fit. So, now, the question is how do we fit and how do we fit such a function into a set of data points next step to do after go a simple function like this. we could go explain... a little more complicated function of this form. So, that is this could be some polynomial it could be any non-linear function this one as but we do not have any parameters inside this function to fit all the parameters are outside of this function sum of all such functions.

So, we will see the general case of that. So, we will go ahead and define function the general function as this one there are now M such function there is a m parameter [] we got. So, we saw we did earlier also the number of parameters to be fitted is called M this m. So, that we have M such functions called X k's. So, we have a k's as the parameters and X k's as the function and this function is completely known and what we need to determine is this a. So, now, some arbitrary function which is completely known. So, now, we can determine our chi square write down our chi square as before there is y i minus f of x i divided by sigma i whole square that is what we suggestion earlier.

So, you remember chi square was written as $\sum_i (y_i - f(x_i))^2$ divided by $\sum_i 1$ the whole square that is what we have written,, but now our $f(x_i)$ as become that. So, it will become $\sum_i (y_i - \sum_k a_k X_k(x_i))^2$ divided by $\sum_i 1$ the whole square. So, now ,where is become our chi square function and now we can minimize this chi square function with respect to all a_k 's that is what we kind to look and you will minimize this chi square function with respect to a_k 's and then we would write it as when you write that you will get $\frac{\partial \chi^2}{\partial a_k} = 0$ that is what we would write and that would give us now No Audio:22:38-22:56 and we would get here as .So, that is what you would get.

So, you would chi square and you will say chi square is given by this and then I would minimize chi square by each of these parameters a_k . So, that would give I this function. So, that will give $\sum_i (y_i - \sum_k a_k X_k(x_i))^2$ divided by $\sum_i 1$ square into $X_k(x_i)$ this particular k .So, this goes from 0 to $M - 1$ or 1 to m . So, the number of parameters which we have and this is what we will get and now this is the equation which would I will solve. So, we would have for each of the k 's each of the k 's that is going from 0 to $M - 1$ we will have that many number of equations and that is what we would have to solve and that is what we would have summarized here again. So, summarize that here. So, we will have $y_i - \sum_k a_k X_k(x_i)$ divided by $\sum_i 1$ square X_k as the number of... that many number of equations where k running from 1 to M .

So, we used to 1 to m here that is what we have. So, now, I can expand that and then write this equation in this form. So, that I will write as [] I will just multiply this out right and then write just to avoid confusion we will use j here [] and we use and then you would write this as $\sum_i (y_i - \sum_k a_k X_k(x_i))^2$ that is a first term and then i will have $\sum_j a_j X_j(x_i)$ going from 1 to M $a_j X_j(x_i)$ into $X_k(x_i)$. So, that is what we would have. So, we have this is the equation to solve. So, that equal to zero is the equation which we will have to solve for these functions again you can see that we will have. So, this is equal to zero is the equation. So, this equal to that is the final form of the equation. So, that we will write in the final form.

So, before I just summarize. So, what we did was we wrote the chi square as now in this form which is same as before that is $y_i - f(x_i)$ by $\sum_i 1$ whole square we minimize each of this a_k 's and then you would get this equation now with which can be written as $\sum_i (y_i - \sum_k a_k X_k(x_i))^2$ is equal to $\sum_i (\sum_j a_j X_j(x_i) - y_i)^2$

into X_k of x_i that is what we have and we would call ... now this is the form which we would have to solve and we have to solve this function for the a_j values that is what. So, we have an equation of this form and we can solve them for the a values. So, we will use we will represent this by one symbol and we would represent this by one symbol and then write to this matrix form as d equal to D times a . So, now, d is a column vector which contains these functions and D this is now a M by M matrix which contains $X_j X_k$ elements and a is again a column vector which contains M elements that is what we would have. So, let me summarize this once again here. So, that is written here.

(Refer Slide Time: 27:18)

$$\Rightarrow \sum_{i=1}^M \frac{y_i}{\sigma_i^2} X_k(x_i) - \sum_{j=1}^M \left[\sum_{i=1}^M \frac{a_j X_j(x_i)}{\sigma_i^2} X_k(x_i) \right] = 0$$

$$\Rightarrow \sum_{j=1}^M a_j \sum_{i=1}^M \frac{X_j(x_i)}{\sigma_i^2} X_k(x_i) = \sum_{i=1}^M \frac{y_i}{\sigma_i^2} X_k(x_i)$$

which can be written in the matrix form as ,

$$\alpha_{kj} a_j = \beta_k$$

So, I have this equation that y_i by $\sigma_i^2 X_k$ and x_i minus σ_i going from 1 to M σ_i going from 1 to N j 1 from 1 to M and i going from 1 to M and $a_j X_j$ by $\sigma_i^2 X_k$ equal to zero that is what we get by minimizing the expression and then we just wrote this as σ_j equal to 1 to M σ_j going from i going from 1 to N a_j into $X_j X_k$ by σ_i^2 on the left hand side and this on the right hand side that is what we have done and then i represented this by a matrix which actually I would call the α here and then a_j β_k this is what

(Refer Slide Time: 28:32)

The chalkboard shows the following steps:

$$\frac{\sum y_i X_k(x_i)}{\sigma_i^2} = \sum_j a_j \sum_i X_j(x_i) X_k(x_i)$$

$$d_k = a_j \sum_i X_j(x_i) X_k(x_i)$$

$$d_k = \alpha_{jk} a_j$$

$$d = \alpha a$$

So, beta k is now simply beta k for each of this k's beta k is given by y i by sigma i square and X k and a k j is a j into X j by sigma i square into x i summed over i that is what this function is. So, we will write that explicitly here. So, what I done there is to write this sum over i X k for each of this k you will get this d k each of this k. So, you will have **1 to m** k goes from 1 to m you remember. So, k... So, that is M such column vector elements in the column vector and now this D is actually.

(Refer Slide Time: 29:50)

The slide contains the following content:

$$\frac{\partial Z^2}{\partial a} = -2 \sum_{i=1}^n \frac{(y_i - ax_i - b)x_i}{\sigma_i^2} = 0 \quad (1)$$

$$\frac{\partial Z^2}{\partial b} = -2 \sum_{i=1}^n \frac{(y_i - ax_i - b)}{\sigma_i^2} = 0 \quad (2)$$

We write these equations in a more convenient form by using

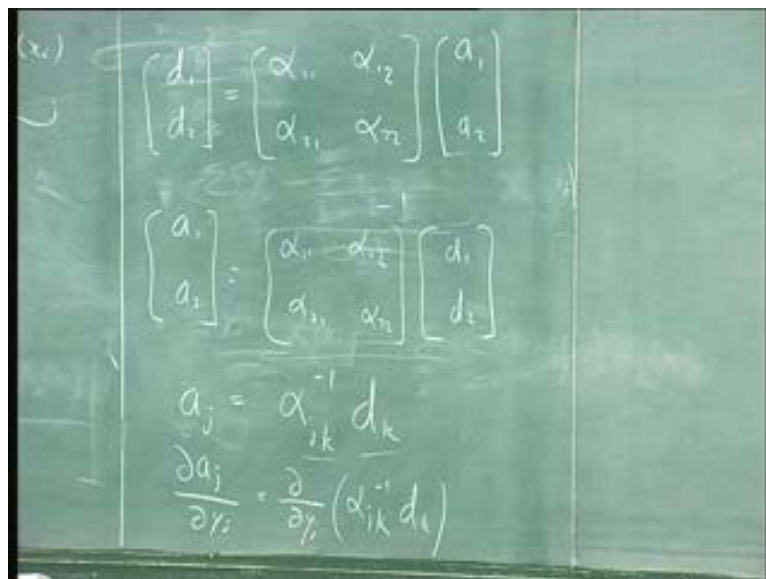
$$\text{Let } S = \sum_{i=1}^n \frac{1}{\sigma_i^2}, \quad S_x = \sum_{i=1}^n \frac{x_i}{\sigma_i^2}, \quad S_y = \sum_{i=1}^n \frac{y_i}{\sigma_i^2}$$

$$S_{xx} = \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}, \quad S_{yy} = \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2}, \quad S_{xy} = \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2}$$

$$\Delta = SS_{xx} - (S_x)^2$$

So, in this case we can write this I will write it as a j and then sum over i X j of x i into X k of x i that is what we written and that is what I would call as α_{jk} . So, that is written as that is what d k s for each of this elements or in the matrix I would just simply write that as alpha a this is a column vector this is a this is a M by M matrix and this is a column vector, that is the equation which we have to solve .Once an often we solve those equations we can... So, we can write that as and then we can word that think and then solve it. So, that I could simply invert this a j's and then and write it in this form and. So, once you have that and then it is easy to program this particular object for you did know how to invert matrices. So, in the present case we will just do this for a two parameter modal. Two parameter model this will be just two by two matrix. So, we have here a simple case of a **column vector multiplied by** column vector equal to a matrix times by another column vector.

(Refer Slide Time: 30:37)



$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \alpha_{11}^{-1} & -\alpha_{12}^{-1} \\ -\alpha_{21}^{-1} & \alpha_{22}^{-1} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$$a_j = \alpha_{jk}^{-1} d_k$$

$$\frac{\partial a_j}{\partial r_i} = \frac{\partial}{\partial r_i} (\alpha_{jk}^{-1} d_k)$$

So, for k is where m is two; that means, you have only a 1 and a 2 this would be a very simple equation we will write that as... we could write this as d 1 d 2 that is to α_{jk} is equal to and we will have alpha 1 1. So, that is simple. So, we can write this as a 1 a 2. So, all in other words as you can easily see this is equal to... (No Audio:31:00-31-09) So, inverse of this matrix multiplying d 1 d 2. So, in a general case that is when you have M such M parameters to determine there is a j's are M by M is case to determine and then this should be an M column vector and will be an M by M matrix and this will be an M column vector .

So, then you would need to know some good programs to know how to invert these matrices we will see that in the next classes how to do, how to write a program to invert matrices. Before the time being just to a two parameter case which is... then we know how to do this how to add this inverse of the matrix by hand. So, we can write this inverse of this matrix and then we can multiply these two matrices to get the solution and once the search what that is got I will be trying to do now. So, then once we have the a's and the b's we can again write down the sigma a square and the what is sigma in general sigma a j square as we have written before there is del a j by del y i and we can write this expression here and then we would now we will see what you would get for that. So, we know that a's are given by... So, we would have from this expression we have a j as alpha j k inverse and dk. So, and since the both these two quantities are functions of y i I can write del a j by del y i as del by del y i of alpha j k inverse dk I can do that, because these are functions of y i. So, I can write that and once I have that I know by sigma a square is simply sigma i square into.

So, sigma sigma a square is sigma i square into this function square that is what we have saw just in the general linear fit and all those definitions use through for this particular k's also. So, that is what here. So, sigma i square into the derivative and the derivative now gets into this form and that then we can write that this function. So, let us go back and see that again. So, we had alpha j k's as this. So, that is our alpha j k function which is this one. So, that is what we have to call as alpha j k. So, we have the function d written as this.

So, we will write back we will write this as dk that is alpha j k this thus particular thing some over i and this summed over i is what we call dk. So, this is a function of y and this is not a function of y. So, I can write this simply as del by alpha j k inverse del by del y of d k and this to compute that is what we would see we call that alpha g k inverse I call that error as C j k that is a general definition so that and then we can do a little bit more of algebra on this status we could actually write down the explicit form of that and finally, it turns out it is just sigma square a j is C j j just the diagonal elements of the inverse matrix. Now, this is easy to see we will just do that few steps here. So, we will try to work out this sigma a square completely.

(Refer Slide Time: 35:38)

$$\begin{aligned} \sigma^2(a_j) &= \sum_{i=1}^n \sigma_i^2 \left(\frac{\partial a_j}{\partial y_i} \right)^2 \\ &= \sum_{i=1}^n \sigma_i^2 \left[C_{jk} \frac{X_k(x_i)}{\sigma_i^2} \right] \left[C_{il} \frac{X_l(x_i)}{\sigma_i^2} \right] \\ &= \sum_{k,l} C_{jk} C_{il} \sum_i \frac{X_k(x_i) X_l(x_i)}{\sigma_i^2} \\ &= \sum_{k,l} C_{jk} C_{il} \alpha_{lk} \\ &= C_{jj} \end{aligned}$$

So, sigma of a j square this is what we have to write and we write this as sigma i going from 1 to n $\sum_{i=1}^n$ and write sigma i square also del a j divided by del y i the whole square that is what we will write. We saw that del a j by del y i is given by this function that is alpha j k inverse into dk divided differentiate with respect to y i. We also saw the alpha j k's are not functions of y, dk is a function of y. So, I can write it as alpha j k inverse into del by y of dk. So, here summed over k is \sum_k . So, this actually is the sum over k. So, I not written that explicitly here, but there should be a any repeat of in this is mean then there is the sum over that.

So, in this notation any repeat of in this is always mean that sum over k, but anyway write that explicitly here. So, we substitute all that here. So, then I can write this function as sigma this as sigma i sigma i square now del a j by del y i whole square. So, and now this is what we call C jk this inverse we call now C jk. So, we have it as C jk into del by del y i of dk. So, now, that comes out to be. So, we remember what dk is. So, what we wrote as dk is this function there is y i by sigma i square into X k of x i. So, the derivative of this dk with respect to y i is simply X k of x i by sigma i square.

So, that is what the dk would be that is what would be writing it here. So, we have this function here as sigma X k ,X k by sigma i square that is what you will have that is what you would have to write now, C jk into that is inverse of that and X k that is X k by sigma i square. So, that is one and this is square this. So, we have to square that. So, that

will be one more that is now we will call... So, now, this is summed over k . So, now not to have any confusion we use the next domain X as l . So, then we will write it as X^l of x^i divided by σ_i^2 . So, I repeat this once again. So, I writing σ_i^2 into δ_j by δ_j whole square and we saw that δ_j by δ_j is given by this summed over k as α_{jk}^{-1} which I now call C_{jk} δ_j of δ_j and δ_j by δ_j of δ_j is simply X^k by σ_i^2 that is what we have now. Now, we want to square this. So, I write $\sum_k C_{jk} X^k$ by σ_i^2 whole square.

So, that is product of two functions. So, that is $C_{jk} X^k$ by σ_i^2 into $C_{jl} X^l$ of x^i by σ_i^2 . So, this is the product. So, this is l and k are dominant this is which means there is the sum over k here and there is a sum over l there that is why we have. Now, we can simplify this and write this as $C_{jk} C_{jl}$ and one σ_i^2 cancels we have X^k of x^i and X^l of x^i divided by σ_i^2 again there is a sum over k and sum over l and there is a sum over i all of them inside this.

So, there is a sum over i function, there is a sum over i X^k of x^i X^l of X^j divided by σ_i^2 and there is a sum over k here and there is a sum over l , but we know that that $X^k X^l$ by σ_i^2 was a definition of our α_{jk} which is in ways of the C 's. So, by... So, now, this is $C_{jk} C_{jl}$ and sum and then we have now you will have C_{kl}^{-1} here actually that is what we will have C_{kl}^{-1} would just get that as C_{jk} that is what you would get because I can do $C_{jk} \alpha_{jk}$ summed over k l I will just simply get sorry $C_{jl} \alpha_{jk}$ summed over for j for every j I would gets this as just δ_{jk} and then I could write that as C_{jk} into δ_{jk} . So, I will get it as C_{jj} that is what is summarized here. So, that is what we would summarize here. So, I have written it as $C_{\sigma_i^2}$.

(Refer Slide Time: 41:21)

$$\text{or, } \sigma^2(a_j) = \sum_{i=1}^N \sigma_i^2 \left[\sum_{k=1}^M C_{jk} \frac{X_k(x_i)}{\sigma_i} \right]^2$$

$$= \sum_{i=1}^N \sigma_i^2 \sum_{k=1}^M C_{jk} \frac{X_k(x_i)}{\sigma_i} \sum_{l=1}^M C_{jl} \frac{X_l(x_i)}{\sigma_i}$$

$$\sigma^2(a_j) = \sum_{k=1}^M \sum_{l=1}^M C_{jk} C_{jl} \left[\sum_{i=1}^N \frac{X_k(x_i) X_l(x_i)}{\sigma_i^2} \right]$$

$$\sigma^2(a_j) = C_{jj}$$

So, that will have two functions here sigma i square into C j k X k of x i by sigma i square sigma l over 1 to M C j l into X l by x i by sigma i square. So, now, this product of this I write it as C j k C j l into this .Now, this is a inverse of this C matrix. So, I would when I do sum over this I just get the diagonal elements of this matrix. So, what is the summary, the summary is that once I got this alpha matrix which is X k X l by sigma i square and I have the inverse of that if I have the machinery to compute the inverse of this matrix then the diagonal elements of that matrix tells me the fitness of the parameters a j this is I have two balls in one shot.

(Refer Slide Time: 42:20)

The inverse matrix $C_{jk} = [\alpha]_{jk}^{-1}$ is closely related to the probable uncertainties of the estimated parameters or a_j 's.

To estimate these uncertainties we calculate ,

$$a_j = \sum_{k=1}^M [\alpha]_{jk}^{-1} \beta_k = \sum_{k=1}^M C_{jk} \left[\sum_{i=1}^N \frac{y_i X_k(x_i)}{\sigma_i^2} \right]$$

So, I can determine the a_j 's by looking by getting this matrix $C_j C_j^k$, because that is simply given by a_j 's are given by C_j^k into β_k where, β_k was given by y_i into X_k of x_i by σ_i^2 and then diagonal elements of that matrix would give me σ_i^2 of $\sigma_i^2 a_j$ or the fitness of the a_j functions then we could now use this function to actually determine the parameter for fit of the form given by here.

(Refer Slide Time: 42:48)

Problem 2: Fit a function of the form $f(x) = ax^c$ and estimate in the error in the parameters: a, c

1.000000	1.767047	6.000000	1.024910
1.500000	1.203254	6.500000	1.005924
2.000000	1.217474	7.000000	0.947200
2.500000	1.265928	7.500000	0.914435
3.000000	0.976787	8.000000	0.742393
3.500000	1.046207	8.500000	0.914230
4.000000	1.073715	9.000000	0.918245
4.500000	0.704091	9.500000	0.589067
5.000000	1.004790	10.000000	0.914338
5.500000	0.947944		

So, I were use a linear function of this, because you could or if you have α followed known then I can use this is still looks like non-linear function, but it is still a linear fit, because its only for a if you are fitting only for a if you are fitting for a and α we could still use a linear fit in this particular model ,because I could take a log of this and I write it as $\log a + \alpha \log x$ and then in the log plot it is still a linear function we will come to that in a minute we will before that we will just see implementation of this particular function, this particular form of general non-linear fit in a in a program. So, that is what we would just look at now.

(Refer Slide Time: 43:41)

```
0
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$ cat fit
set pointsize 4
plot 'gen-linear.data' w points 7, 2.0*exp(-x)+
0.85/(1.0+x) lw 4
pause +1
[sunil@dali lect14]$
[sunil@dali lect14]$ cd ../lect13
[sunil@dali lect13]$ gnuplot fit

[sunil@dali lect13]$
[sunil@dali lect13]$
```

So, here we have a function called I would there is a code general linear dot c. So, here is a program. So, this program would be actually we will use again we have let us say set of data points x and y have it. So, and then we determine the same. So, that and the data file is in general linear dot data.

(Refer Slide Time: 43:46)

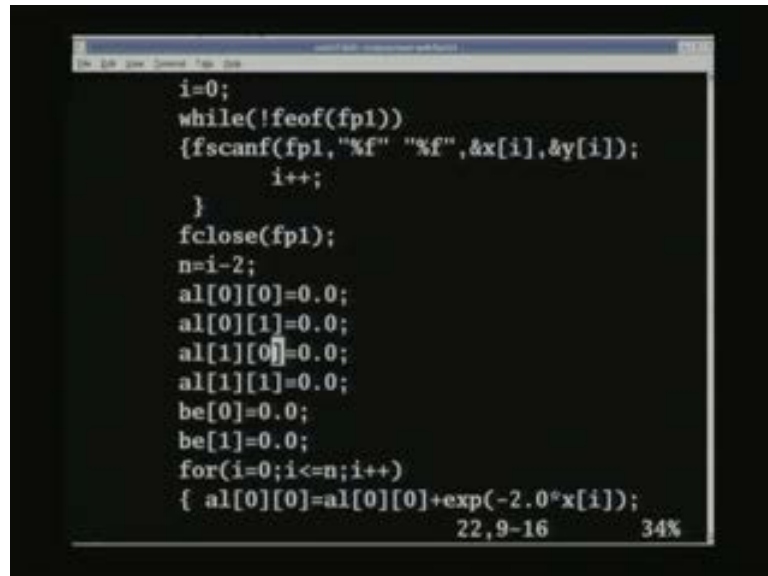
```
plot 'gen-linear.data' w points 7, 2.0*exp(-x)+
0.85/(1.0+x) lw 4
pause -1
[sunil@dali lect14]$
[sunil@dali lect14]$ cd ../lect13
[sunil@dali lect13]$ gnuplot fit

[sunil@dali lect13]$
[sunil@dali lect13]$ cd ../lect14
[sunil@dali lect14]$ ls
a.out          gen-linear.data
data-fit.ppt  gen-linear-data.c
fit
fit.data      linear-err.c
gen-linear.c
[sunil@dali lect14]$
```

So, there is a data file like this and then I would read of this data file which gives me x and y i as we have done in the case of linear fit. So, we (x_i) of i and y of (y_i) and then

we now determine the matrix elements. Now, the matrix elements I call them as you know a alpha l. So, this a matrix elements. So, alpha l a l here.

(Refer Slide Time: 44:29)



```
i=0;
while(!feof(fp1))
{fscanf(fp1,"%f" "%f",&x[i],&y[i]);
  i++;
}
fclose(fp1);
n=i-2;
al[0][0]=0.0;
al[0][1]=0.0;
al[1][0]=0.0;
al[1][1]=0.0;
be[0]=0.0;
be[1]=0.0;
for(i=0;i<=n;i++)
{ al[0][0]=al[0][0]+exp(-2.0*x[i]);
  22,9-16 34%
```

So, matrix elements are all initialized to begin with and. So, beta elements are also initialized. So, both the column... So, if this a and b here are my d functions and alpha functions. So, as we said we this is the a and alpha and the b we can see that it is a two by two matrix 0 0 0 1 1 0 1 1 and. So, this is the column vectors are of two elements I got two elements and we trying to determine a function of the form I will show the function of the form given by a exponential minus x of i divided by 1 plus x of i. So, let us I will write down this function here. So, which trying to set because we trying to fit some function of the form are using these matrix equations we are trying to fit in function f of x i as a 1 exponent to the power of minus x i plus a 2 divided by 1 plus x i.

(Refer Slide Time: 45:42)

$$f(x) = a_0 e^{-x} + \frac{a_1}{(1+x)}$$

$$X_0(x) = e^{-x}$$

$$X_1(x) = \frac{1}{1+x}$$

$$a_{00} = e^{-2x_i}$$

$$a_{11} = \frac{1}{(1+x_i)^2}$$

$$a_{01} = a_{10} = \frac{e^{-x_i}}{(1+x_i)}$$

So, as we can see this as my two $x \times x$ functions now this is what. So, let me call X_0 would be now exponent to the power of minus x_i , X_0 of x_i and X_1 of x_i is now 1 by 1 plus x_i . So, that is a kind of function I am trying to fit. So, I have a set of data points and I believe that this data should be a combination of these two functions. So, I wrote a linear combination of these two functions and I am trying to fit this χ^2 determine a a_0 and a a_1 that is what the program is doing here. So, I have to first compute the matrix elements which is \sum_i remember which is the matrix elements are given by $\sum_i \chi^2$ by $\frac{\partial \chi^2}{\partial a_i}$ and $\frac{\partial \chi^2}{\partial a_j}$.

So, that is alpha matrix elements α_{ij} that is what I am trying to do here. Now, this is the 0 th element of $0,0$ that is alpha in this case α_{00} that will be $\frac{\partial \chi^2}{\partial a_0}$ into $\frac{\partial \chi^2}{\partial a_0}$. So, in this case it is just exponential minus x_i into exponential minus x_i . So, that will be just simply exponential minus $2x_i$. So, we can write down these matrix elements from this. So, this element first element would be some $\frac{\partial \chi^2}{\partial a_0}$ would be now sorry $\frac{\partial \chi^2}{\partial a_0}$ would be given by $\sum_i e^{-2x_i}$. So, remember the matrix element we have the matrix elements as thus here that is what we want to write down.

So, when you have $k=0, j=0$ is a χ^2 product of these two χ^2 's in this case again is one. So, the matrix elements are given by x_0 into x_0 that is a_{00} we can write that. So, a_{00} is x_0 into x_0 . So, that is x_0 is that is one into x_0 that is equal $\sum_i e^{-2x_i}$ and

similarly a 1 1 would be x^{-1} into x^{-1} . So, that be 1 by 1 plus x^{-1} the whole square. So, will have a 1 0 which is equal to a 0 1 equal to e^{-x} divided by $1+x^{-1}$ that is a functions we have. Now, those functions are put in here.

So, we have a 1 0 and a 0 1 both are given by exponential minus x^{-1} to be $1+x^{-1}$ and a $1+x^{-1}$ and then a 0 0 as exponential minus $2x^{-1}$ a 1 1 as $1+x^{-1}$ into $1+x^{-1}$ we can see that here that is $1+x^{-1}$ into $1+x^{-1}$ this $1+x^{-1}$ whole square. Now, the beta function is given by y^{-1} into x^{-1} . So, that is y^{-1} into exponential minus x^{-1} of i and then y^{-1} into exponential y^{-1} divided by one over x^{-1} plus 1 over plus x^{-1} again now, these are the two beta values now we have to determine the inverse of this matrix, before that we know inverse is the adjoin divided by the determinant. So, I determine the determinant of this matrix here first determine the inverse of this a matrix. So, we need to get alpha in versions which I call C_j 's here which are the inverse of these matrix a 0 0 a 1 0 a 0 1 a 1 1 matrix. So, I want the inverse of that.

So, which is my C. So, I have to determine C that is the next step in the program. So, for that I will write that as the adjoin divided by the determinant and that is what I have written here. So, here is the determinant. So, where a two by two function is easy 2 by 2 matrix is easy. So, I can write down the determinant and then and I wrote down the inverse matrix here

(Refer Slide Time: 50:19)

```

    al[1][1]=al[1][1]+1.0/((1.0+x[i])*(1+
x[i]));
    be[0]=be[0]+y[i]*exp(-x[i]);
    be[1]=be[1]+y[i]/(1+x[i]);
}
D=al[1][1]*al[0][0]-al[0][1]*al[1][0];
printf("%f %f %f %f\n",al[0][0],al[1][0]
],al[0][1],al[1][1]);
a=(al[1][1]*be[0]-al[0][1]*be[1])/D;
b=(al[0][0]*be[1]-al[1][0]*be[0])/D;
printf("n = %d      a = %f    b = %f\n
n",n,a,b);

chisq=0.0;
for(i=0;i<=n;i++)

```

36,2-9 76%

So, inverse matrix I want to determinant the inverse matrix is the adjoin of that and I use that to determine the a and b's here. So, now, this **this** step in the particular step I determine a and b as the inverse of matrix a l here and would multiplied by b e. So, that is this step is what I am doing and then I print out the values of a and b I was obtained. Once I have the values of a and b I can compute the chi square for that which is this y i minus exponential f of x i is a star exponential minus x i plus b by 1 plus x i minus y of i the whole square that is multiplied by the same function again. So, that is my chi square and then I print out the chi square and then I have the standard error which is root of chi square by n minus 1. So, I write down the standard error also. So, I have chi square and the standard error now here that is what we going to do and which is run this program once.

(No Audio: 51:33 - 51:42)

So, we have thus the matrix elements here which is **[]** sequence. So, a is comes under by two point three and b is point five seven. So, here is two point three exponential minus x i plus point five seven nine eight one two divided by 1 plus x i. So, the chi square is one point something and this standard error is point two eight nine that is what I get the chi square here. So, we will just plot now that function and see how function fits in.

(Refer Slide Time: 52:21)

```

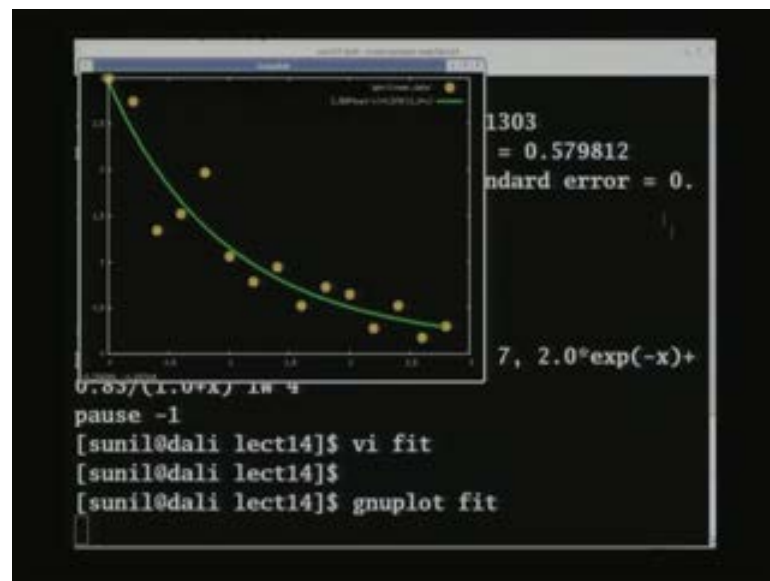
gen-linear.c
[sunil@dali lect14]$ vi gen-linear.c
[sunil@dali lect14]$ gcc gen-linear.c -lm
[sunil@dali lect14]$ ./a.out
3.025726 3.457046 3.457046 4.251303
n = 14      a = 2.363883    b = 0.579812
          13 xhisq = 1.089980 Standard error = 0.
289559
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$ cat fit
set pointsize 4
plot 'gen-linear.data' w points 7, 2.0*exp(-x)+
0.85/(1.0+x) lw 4
pause -1
[sunil@dali lect14]$

```

So, I have written a small script which is basically plotting this data points I plot this linear data that is a data given to me using points and then I will have the function which

is now what was the really substitute the value for the function either is two point three six three eight eight into exponential minus x we should have and point five seven nine. So, you should substitute point five seven nine here and two point **three six eight eight** three six three eight eight I will do that two point three six eight and one point five seven nine by 1 plus x .So, that is our function and we will just plot this function and that is what we would get.

(Refer Slide Time: 53:18)



So, that is the circular points here are the data points given to me and that is the green line here is my fit of two point three eight six eight exponential minus x **minus x exponential minus x** plus point five seven nine divided by 1 plus x.

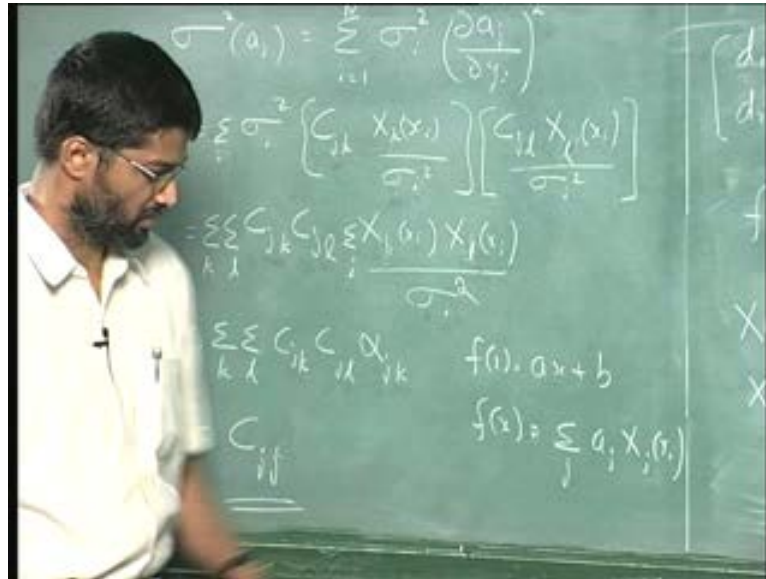
(Refer Slide Time: 53:49)

```
[sunil@dali lect14]$ ./a.out
3.025726 3.457046 3.457046 4.251303
n = 14      a = 2.363883    b = 0.579812
          13 xhisq = 1.089980 Standard error = 0.
289559
[sunil@dali lect14]$
[sunil@dali lect14]$
[sunil@dali lect14]$ cat fit
set pointsize 4
plot 'gen-linear.data' w points 7, 2.0*exp(-x)+
0.85/(1.0+x) lw 4
pause -1
[sunil@dali lect14]$ vi fit
[sunil@dali lect14]$
[sunil@dali lect14]$ gnuplot fit
```

So, this is the fit of the data along the point and then we see that this is the data fit is pretty good visually and we can see that in the chi square value and the standard error which we can get is been pretty low .So, now, after doing this general linear fit and some function f of x is equal to. So, now, we saw. So, for two types of fits that is our f of x is equal to ax plus b and then we saw functions of the form f of x is equal to $\sum_j a_j X_j$ of x where X_j of x itself can be any non-linear function and that itself is linear combination of this function and we determine this values.

So, these are two general linear functions for function fit which we solve. So, this being a very special case of the general linear fit. So, now, thus now we could go into and look at set of non-linear functions. So, that is what we should try to do .In the next lecture we would look at a general non-linear model so; that means, now we could have functions of this form. So, for we only saw functions of this two forms what we could have functions of it quite common in many experiments would be equal to have a $1 e$ to the power of minus $a 2 x$ of this form.

(Refer Slide Time: 54:08)



So, that is a non-linear function and you want to determine both a 1 and a 2 remember here we do not do that we have a non-linear function,, but we have no fitting parameters inside the non-linear function we had only the coefficient of that in this model,, but in a general non-linear fit we may have to even determine the parameters in non-linear function and that is what we would be looking at in the next class. So, we will stop here and we look at the in the next class the general non-linear models fit.