EXCELing with Mathematical Modeling
Prof. Sandip Banerjee
Department of Mathematics
Indian Institute of Technology Roorkee (IITR)
Week – 12
Lecture – 57 (Estimation of Parameters-I)

Hello welcome to the course EXCELing with Mathematical Modelling.

Today we will be discussing about an important aspects of mathematical models, which is estimation of parameters.

So, to estimate we have many methods, which we use and which runs in the background.

One of them is method of least squares.

Let me quickly explain the theory that will be used to estimate the parameters.

So, let us do it with help of a straight line, it will be easy to understand.

So, I have some points and I want to fit a curve to this.

So, I choose a point which is somewhere here and I name it $P(x_i, y_i)$.

So, this is the x axis and the y axis and we take the equation of the form, say,

$$y = m\,x + c.$$

So, here m and c are the parameters which I need to estimate using these data points.

So, I choose our data point and I name it as (x_i, y_i).

So, if I draw this line, so this is N, this is P and this is M.

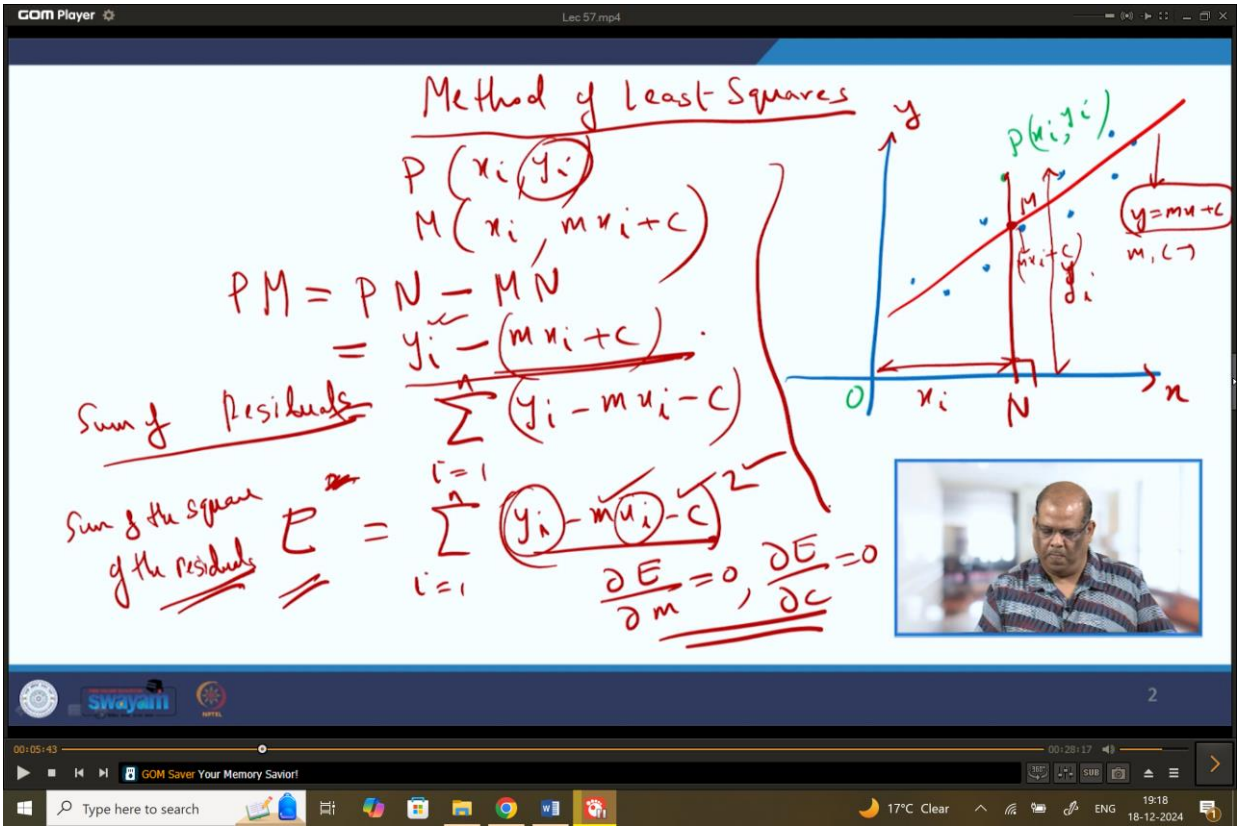
So, this perpendicular this intersects the straight line at this point both of them have the same x coordinate which is x_i and the error which you get of the y value is this is your actual y_i and at this point it satisfies this equation of the straight line.

So, this value is $y = mx_i + c$. So, the value of P is I mean the coordinates of $P(x_i, y_i)$ whereas the coordinates of $M(x_i, mx_i + c)$

So, now if I want to see the error which is $PM = PN - MN = y_i - (mx_i + c)$.

So, this gives you the residual actually.

So, the next thing what you do is you take the sum of the residuals.



So, this is true for the i[th] point if I consider are n number of points this is

$$\sum_{i=1}^{n}(y_i - mx_i - c)$$

Then what we do is we make it as an error square.

Instead of taking sum of the residuals, we take sum of the square of the residuals. So this is going to be

$$E = \sum_{i=1}^{n}(y_i - mx_i - c)^2$$

This is the residual, this is the square of the residual and then I sum it from 1 to n and we name it as E. So you have the error square and then you minimize this error with the help of our known calculus that you find the derivative with respect to the parameters put them equal to 0 and calculate the value of m and c in terms of yi and xi.

So, this is the theory in behind what we will be doing now.

However, we will be using a software, I mean in this particular case Microsoft Excel or otherwise you have to write a code and the code will follow this algorithm.

So let's see that how we can estimate the parameters of unknown functions using Microsoft Excel.

So I have X value and the Y value.

So the X values are 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22.

These are the X values and the Y value is 244.5, 200, 161.8, 145, 134.5, 122.7, 109.2, 110.8, 109.8, 111.2 and 105.9.

So let me quickly increase the font size and bold.

So the X values, the Y values and let me quickly plot them.

You get something like this.

So while looking at the graph, you have to predict a function.

So, say somebody say okay let me fit a parabolic curve because it looks like a bit of parabola here.

So, I will put $y = a + bx + cx^2$

So what you have to do now is you put the value of a, b and c here.

Some arbitrary value you have to put.

Say I put them to be 1, 1, 1.

Now so I assume that let the curve be of the form $y = a + bx + cx^2$ and I calculate this value.

So this is equal to a which is a constant plus b again by x plus c multiplied by x square get some value.

So, the next thing is let me quickly remove the grid lines.

So, you click here then right click and then select data.

So, this select data you click this add and the you can write this as best fit then go to the X select them enter go to the Y and select them and enter.

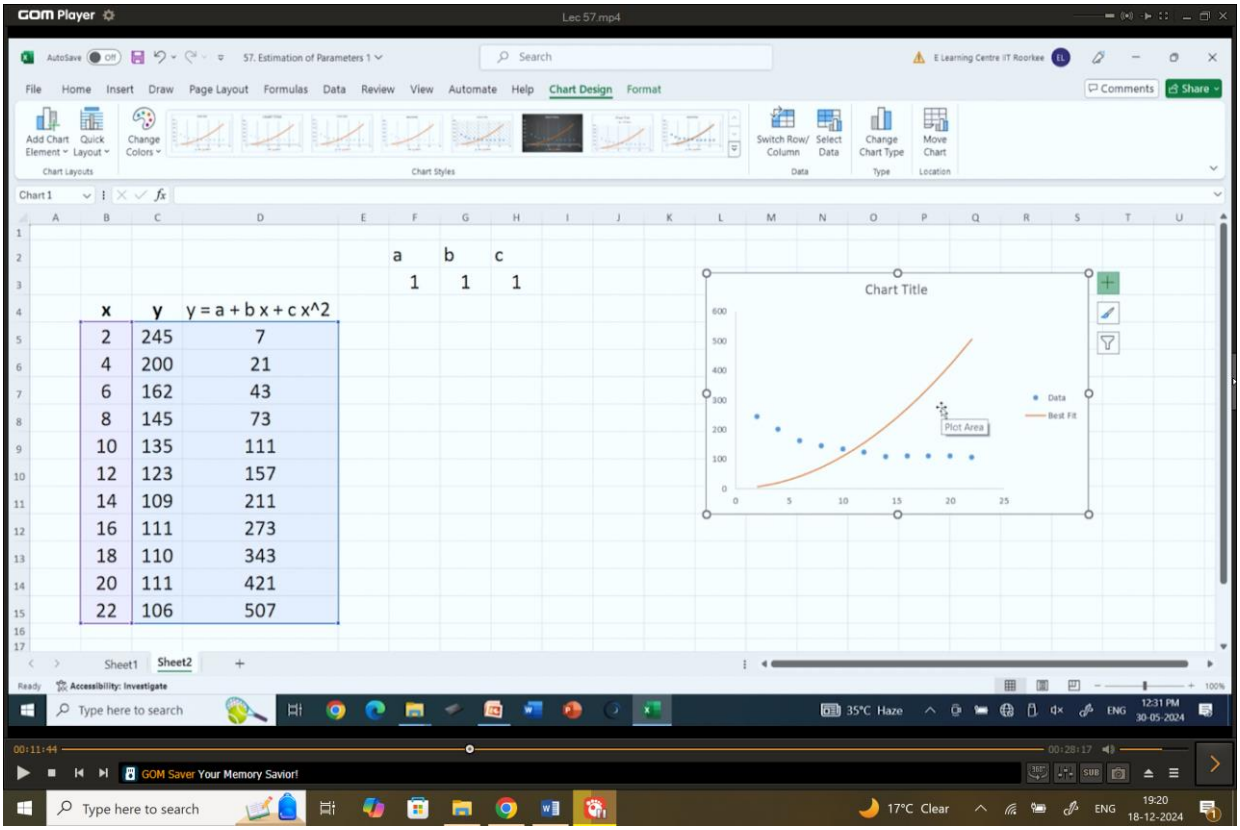
So, I got something and show legends the series one I can go to data short design select data series one edit and name it as data

And let us make this best fit change chart type and best fit let us make it a line.

Okay, so do not panic that you get a very very bad fit.

So this is a data and this is the curve.

So the next thing is you have to change these values so that you get close to this.



So as you can see I have to lift this.

This is somewhere here 1 and this is somewhere around 250 it seems.

So if I make this 250, okay, it does not work.

So you have to play with this.

If I make this 250, okay, it goes here somewhere and then I have to bring this down.

So, if I want to bring this down say minus 10 okay it comes a bit minus 20 okay it is going down and then I have to bring this one also down say 100 okay it

So you have to play this a bit.

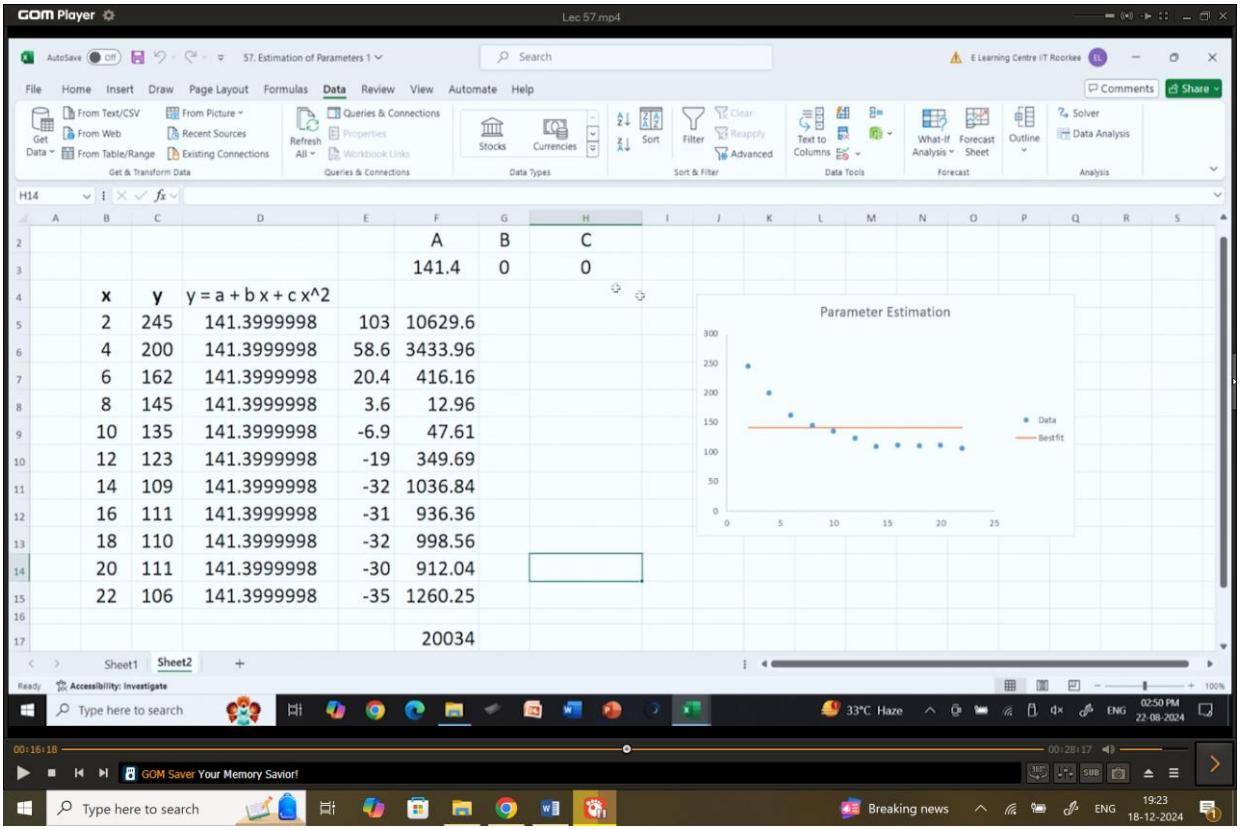
The idea is that you bring this curve bit close to these values and then what you have to calculate is the square of the residuals which is this minus this.

So, you get these values you calculate them then you square the errors that is this residuals and then the sum of the squares.

We now have to minimize this sum of the residual square for which we use solver.

So I click this solver and this window opens here.

So we have to minimize this sum by changing the variables a b and c we keep this same and just click solve and okay so if you look into this figure you see that it is a very very bad fit and that has happened because the function which we have chosen it's not an appropriate function that will match this data that is given.



So, we have to choose an appropriate function such that you can estimate the parameter values and get a proper fit.

So, I can choose the function in the form y equal to a into exponential of minus b x plus some constant.

So, I have the parameters a, b and c. So, I can get an initial value of these parameters and generally we keep it simple let us use them to be 1 1 1.

So, once you get the values, now, I will calculate so this is the x values this is the y values the data and this is the one we try to fit a curve to this data.

So this is we have assumed it in the form of an exponential function.

So this is will come with some sort of intuition but looking at this behaviour of the data points you have to guess what kind of function may fit into this particular data may be wrong in 1 and 2 but then in the third case you will be able to you know basically you have to choose the correct kind of function to get a good fit.

So here we calculate the initial value first by taking the values of a, b and c and then slowly we will manipulate this a, b and c to get a good fit.

So this is equal to a multiplied by E exponential minus B multiplied by the X values plus C. Now, the A B C's are constants.

So, I will put constants values here just by putting a dollar sign here and I get the values okay.

Next I can change this to parameter estimation, okay, then you come here give a right click so next I have to add these values here so there are two ways either you you just highlight all this and get the value again i mean get graph again or you just right click and you will get select data you select data you go to add

Give a name I put it best fit get the X values which is these values and get the Y values which is these values.

Okay so the one which you get here is reasonably bad but do not panic you have to play with this ABCs such that this goes you know close to this value.

So the very first thing is I have to lift it this value somewhere here.

So let me just change this say to 100, okay, there is a bit lift, let us lift this one, if I make this also 100, okay, so it goes up.

Now, I have to put this a bit, see if I make it 10, okay, it becomes straight, so this value has to be less, say 0.5, moves up 50.6 moves down.

So, let us make them 0.4 going up this if I put it a little more say I make it 250 and it goes up.

So, basically you have to play this a bit so that this gives a little this curve go close to these points.

So, if I make this as 300 yeah a little more it is going so like that.

So, let us make it 0.3, yeah.

So, this looks like now that this curve is going and fitting this.

But what are the best combination of A, B and C's?

So, that we have to calculate using this method of least squares.

So, this is the value of Y value which you got from the data and this is what is the predicted value.

So, this is residuals and this will be equal to the data minus the value calculated from the equation and you get something.

So, this is all the residuals and then you calculate the square of the residuals.

So, this square and then comes the sum of the residuals.

So, I can write sum of the square of the residuals change the form to 16.

So, this is equal to sum of all these residuals.

So, you got the sum. Now, according to the theory I have to minimize the sum, okay.

For that we use solver. So, you go to data and you will see there is a solver here.

If you do not see it here, you go to file, you go to option and then you go to add-ins and here it is Excel add-in, you click go.

And there you will see solver add-in, so you have to check this box like this. And the moment you do that, so let me uncheck this box and I click OK and you see the solver is gone.

So, again what you have to do is file, go to this option, go to this add-in, then this Excel add-in.

There are other options also if it is not there, you click to Excel add-in then go and here you see the solver add-in, you check this box and click okay and the solver will come. So, next you click in this solver and a window like this is open.

So, what is your objective? set objective.

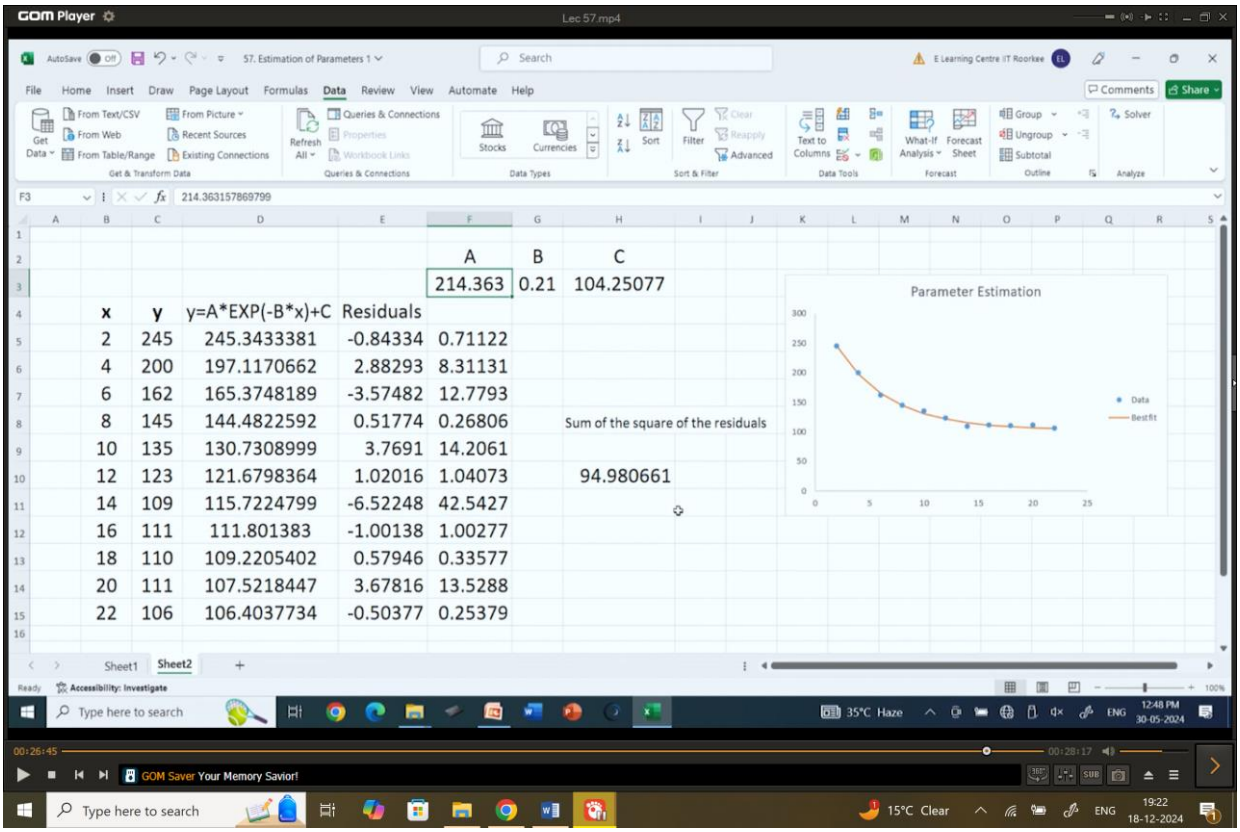
Objective is that I have the sum of the square of the residuals, I have to minimize this.

So, you click this particular cell only and it will appear here.

So, this is the sum of the square of the residuals which came here.

Your objective is to minimize these values and how you will minimize?

By changing the variable cell that means these values of a, b and c.



So, one more time set objective is where the sum of the square of the residuals is you have to minimize that by changing the variables means here it is a, b, c keep them as it is and then solve.

You will get keep solver solution you already see from here that it is now a good fit and you get the values to be 214.363.21 and 104.25077.

So, this is how you estimate the parameters of a unknown curve where ABCs are your parameters which need to be determined from the data.

So, as you see that, given a data, you first have to plot them and once you plot them you will be able to get the behaviour of the data and looking at the behaviour you have to guess the function.

Initially you may get wrong in one or two times but slowly with experience you will understand okay that this is the function which will be good for to estimate the parameters and which will fit this particular data.

So once you take that function there will be some unknown parameters and then I have shown you that you use the method of least squares to estimate the value of the parameters and in Microsoft Excel we use the solver to estimate those parameters.

So with this, we come to the end of this lecture.

In my next lecture, we will carry on with this estimation of parameters where we will be using now differential equations with content parameters to estimate those parameters.

Till then, bye-bye.