Computational Techniques

Prof. Dr. Niket Kaisare

Department of Chemical Engineering

Indian Institute of Technology, Madras

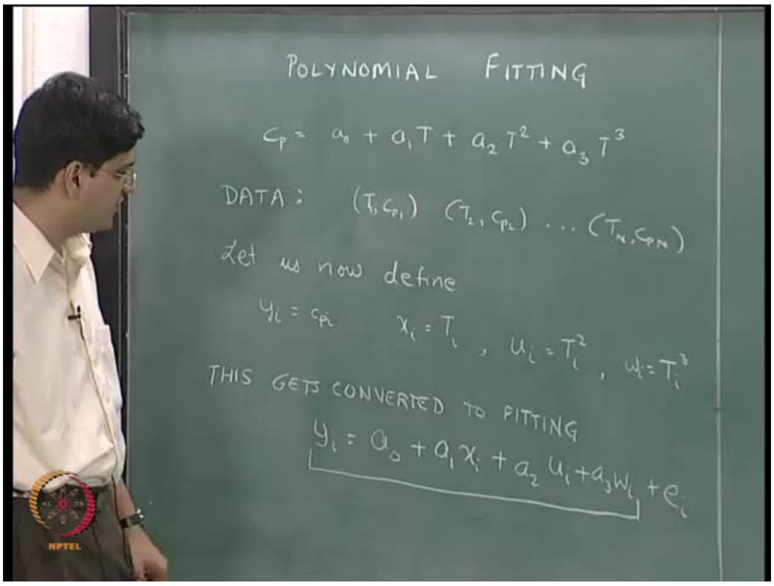
Module No. # 05

Regression and Interpolation

Lecture No. # 03

Hi and welcome to lecture 3 of module 5, where we are looking at regression and interpolation. In the previous 2 modules, we introduced what we meant by regression, what we meant by interpolation. And then we looked at various different ways of linear interpolation, started off with fitting a straight line y equal to a 0 plus a 1 x and then extended that to multi-linear regression, where y in general will be a function of more than one variables. The example that we took was a 0 plus a 1 x plus a 2 u plus a 3 w and we saw, that the overall equations that we get, follow a certain pattern and we can reduce the problem of finding a 0, a 1, a 2 and a 3 and so on.

(Refer Slide Time: 01:40)



To the problem of solving n linear equations in n unknowns, that was the first method that we looked at; the second method that we considered after that, was a matrix based method, where we wrote our equation as y equal to the linear function plus error and then amounted to a least square problem; what we are going to do today is look at polynomial fit and fitting of some functional forms, which are essentially non-linear functions, but which we can linearize them and then go ahead and use linear or multi-linear regression in the same form.

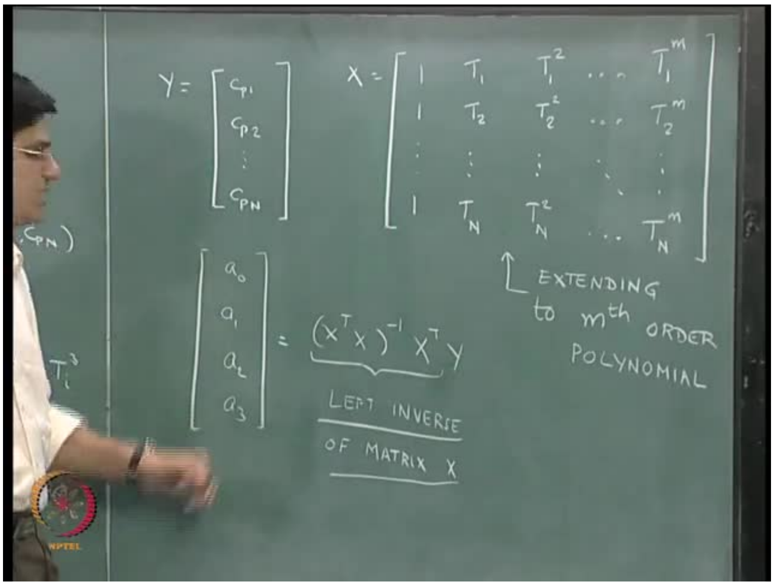
So, starting off with this polynomial fitting; an example of this would be the specific heat C p is written as a function of temperature and it why might be a polynomial function of temperature, we can write this as a 0 plus a 1 T plus a 2 T square plus a 3 T to the power 3; so this is a polynomial fit that we want to obtain; we want to obtain a 0, a 1, a 2 and a 3, such that the error between the C p computed by the model using this value and the C p value that is obtained from the data is minimized.

So, in this case, the data that we have is going to be C p 1 or rather I should write T 1 C p 1, because we have been writing it in the form x, y where x is the independent variable and y is the dependent variable; so we have T 1, C p 1 T 2, C p 2 and so on up to T n, C p N; so this is the data that we have; so, let us now define y i as nothing but C p I, x i is going to be nothing but T i u i is nothing but T i square, and w i is nothing but T i cubed.

So, with this definition y i as C p I, x i as T i u i as T i square and w i as T i cubed, so what we can say is that, the overall equation that we now want to fit, so its amounts now fitting the functional form a 0 plus a 1 x plus a 2 u plus a 3 w; in this case, x, u and w are not 3 independent variables as we had in the previous lecture, but instead x is going to be temperature, T u is T square and w is T cubed.

Once we write it in this form, we can then use the same ideas of multi-linear regression and we can then go ahead and obtain the overall solution and this is going to work as well as a general multi-linear regression of that form as well.

(Refer Slide Time: 05:38)



So, let us see what we got in the multi-linear regression case for this particular model; for this model, what we had is the matrix Y was defined as y 1 y 2 and so on up to y n; Y recall is nothing but C p; so, we have C p 1 C p 2 and so on up to C p N, that is going to be our vector Y. Our matrix X, if you recall from the previous lecture, the first column of the matrix X was 1 1 1 1 repeated n times; the second column was x 1 x 2 x 3 up to x n; third column u 1 u 2 u 3 up to U N and fourth column w 1 w 2 w 3 up to W N using the relationship that we have just written few moments back. We can write X as 1 1 and so on up to 1, we have n number of 1's over here; next is going to be T 1 T 2 and so on up to T n; T 1 square T 2 square and so on up to T N square and T 1 cube T 2 cube and so on up to T N cube.

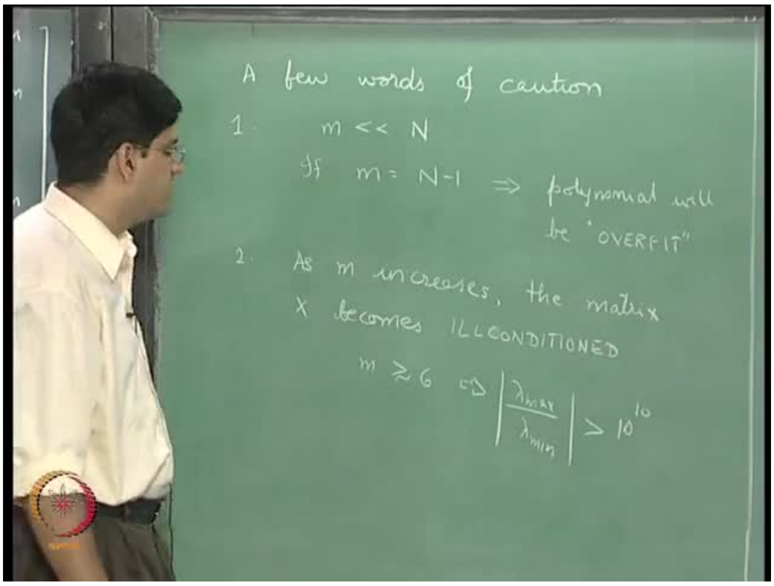
So, this is going to be our Y; this is going to be our X (Refer Slide Time 05:38); and our a 0 a 1 a 2 and a 3 are going to be nothing but X transpose X inverse multiplied by X transpose multiplied by Y; this particular term X transpose X inverse X transpose is known as the left inverse of matrix X.

So, this is what is the function form, this is how we go ahead and do the polynomial regression; for in general we have n data points and we want to fit an mth order polynomial.

So, to extended from this particular case to a general mth order polynomial, if we need to extended all, we need is additional columns in this particular matrix; so, if extending it to n to mth order polynomial, in that case I will erase this particular column also.

So, we will have T 1 square T 2 square up to T 2 to the power n, T 1 cube T 2 cubed up to T N to the power 2 and so on up to T 1 to the power m T 2 to the power m and so on up to T N to the power m and this is what we will get; if we have to extended to a general mth order polynomial in temperature, what the first thing that is necessary in a case like this is that, m has to definitely be less than N, you cannot have m to be equal to or greater than N; in general, m has to be much lesser than N for this to work, if we do not have that particular condition satisfied, we will do this particular left inverse of the matrix will not exist and if the left inverse does not exist, there will not the a least square solution a 0 a 1 a 2 up to a m plus 1; keep in mind that, the number of coefficients that we are going to find through this procedure are going to be one more than the order of this equation, why because we have the first coefficient as a 0, so we will have a 0 a 1 up to a m, that makes total of m plus 1 coefficients that that we need to obtain through this least squares procedure.

(Refer Slide Time: 10:18)



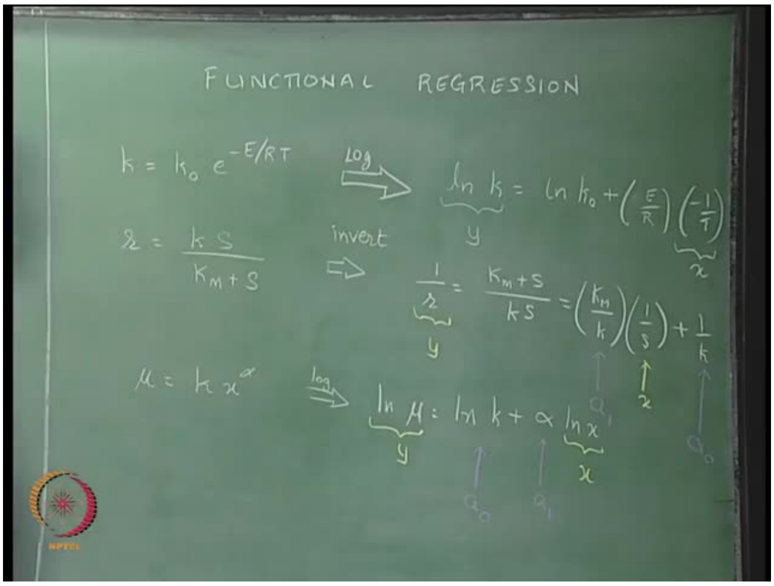
So, this is what is known as the polynomial regression. A few words of caution for polynomial regression, first thing is the polynomial fit that that you want to obtain; if you have the number of data points as capital N, the order m has to be in general much less than the value of N in order for you to have a good confidence on the values of the coefficients that we have obtained; of course, if m is of the order of N or m is 1 less than N in that particular case what we will end up having is, over fitting of the polynomial.

So, if polynomial will be over fit, which basically means that, you can perhaps use this as one way of doing interpolation, but in order to get a regression fit, this is a very poor way of getting a regression fit; in general, to fit an mth order polynomial, my rule of thumb that I use essentially is N, it should be about at least 2 or 3 times greater than the value of m.

The second problem is as m increases, the matrix X becomes yield conditioned; what we mean by the matrix becoming yield conditioned is that inverse of the matrix, there is possibly going to be a lot of errors associated - with the - with the inverse of the matrix. In other words, what yield conditioning really means is that, the largest Eigenvalue of the matrix X transpose X is several order of magnitude greater than the smallest Eigenvalue of X transpose X; when we are inverting a number, for example, if you are to invert a number say 1000, when we inverted, that number becomes 0.001; on the other hand, if you are inverting a number say 10 to the power minus 5, when we invert that, the inverse becomes 10 to the power 5; as a result, the small numbers in X transpose in some ways, again I am using pedagogical liberties over here, but the small numbers in X transpose X are the small Eigenvalues; X transpose X becomes large Eigenvalues in its inverse.

As a result, small errors in those eigenvalues appear as very large errors when you try to invert the matrix; because of this when we try to invert a particular matrix, we have to ensure essentially that the largest Eigenvalue divided by the smallest Eigenvalue should not be a very large number; usually, what will happen? Let say, if m becomes greater than or equal to and again I am using an approximately equal to sign, say m becomes greater than or equal to 6, the matrix becomes fairly yield conditioned, in that lambda max divided by lambda min starts becoming greater than 10 to the power 10, at under these conditions, if you try to invert the matrix X transpose X, an inversion algorithm that we are trying to use should give a warning that this particular matrix is going to be yield conditioned and we may not be able to relay on the results that we obtained from this matrix.

(Refer Slide Time: 14:58)



So, that is essentially what I wanted to cover about polynomial regression. Next, we go on to functional regression and the idea over here is the same idea that we used while fitting - the rate law - the Arrhenius rate law, for example, when we had the rate constant k equal to k 0 e to the power minus E by RT; we took logarithm on both sides and when we took the logarithm, we actually got l n k, I will just write it again k equal to k 0 e to the power minus E by RT; when we took the logarithm, we will get l n of k is l n k 0 plus minus E by R multiplied by minus 1 by T. So, in this particular case our l n k was our y and minus 1 by T was our x.

So, when we plotted, essentially l n k against minus 1 by T, we got this particular curve as a straight line with the y intercept as l n k 0 and the slope as the dimensionless activation energy E by R, actually not dimensional less activation energy, it is a activation energy E by R.

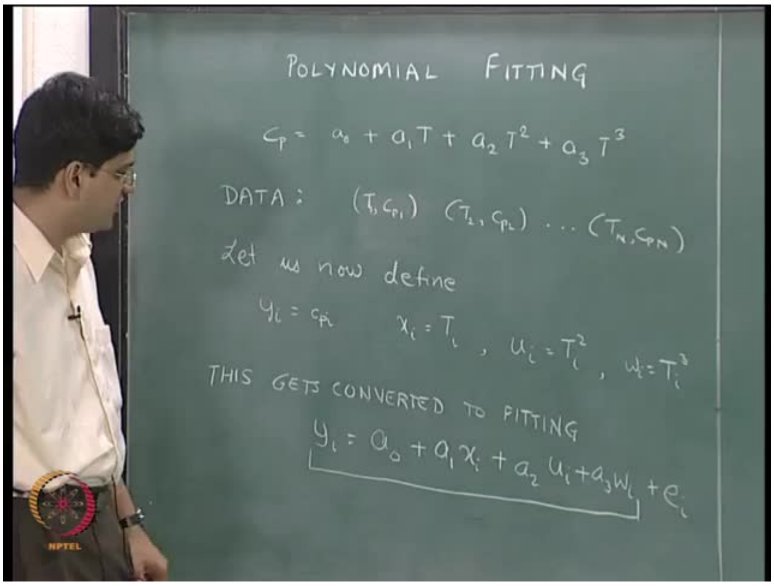
So, that is what we got when we took logarithm over here; another example comes from biological systems, enzymatic kinetics follow, what is known as the Michaelis–Mentens kinetics and over there the rate constant k is going to be or the rate r is going to be given by some constant k multiplied by the substrate S divided by K m plus S, which is K m is going to be a saturation constant and S is the concentration of the substrate that we are interested in and when we invert this particular expression, we will get 1 by r is going to be equal to K m plus S divided by k plus S sorry divided by k S not k plus S, which we can write this down as K M divided by k multiplied by 1 by S plus 1 by k. So, in this particular expression 1 by r is our y, 1 by S is going to be our x, this particular term that we have K m divided by k is going to be our a 1 and 1 by k is going to be our a 0 (Refer Slide Time 14:58).

So, when we try to do a linear regression or try to fit a straight line between 1 divided by r as the y axis and 1 divided by S as the x axis, 1 by rate constant k is going to be our - intercept - y intercept and the slope is going to be nothing but the saturation constant K M divided by the rate constant k.

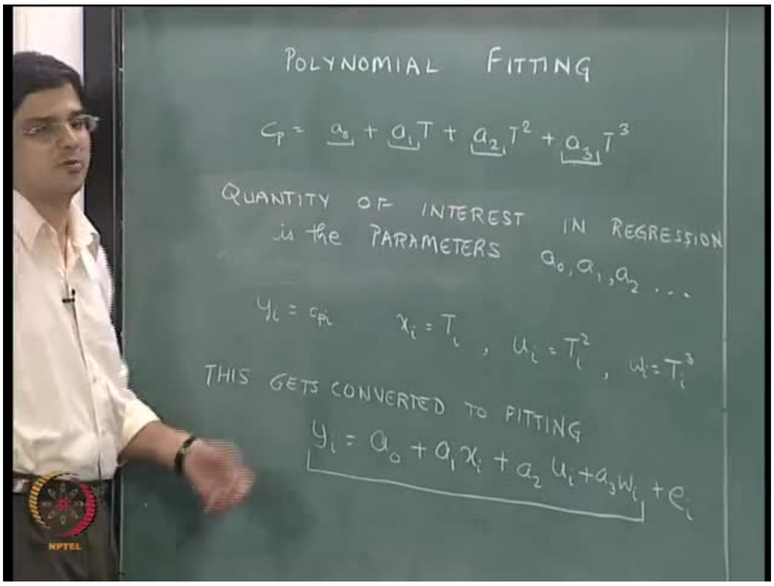
So, that would be another example of functional regression - regression of any particular functional form; third example that we can look at is a power law type of a model, let say we have a model which of the form mu equal to k times x to the power alpha, we can then again take logarithm of that and then we will have l n of mu is going to be equal to l n of k plus alpha times l n of x; so, l n of mu becomes our y; our l n of x becomes our mu x; our alpha is the slope a 1 and l n k is the intercept a 0; so, this is this is what we get.

So, there are a lot of functions of interest to chemical engineers, especially they arise in rate expressions of various definite forms and under those conditions we can actually do some straight forward manipulation of those equations; in order to get those equation in a functional form which would still be linear in a modified parameter, whereas it can be non-linear in x, x is the independent variable and that is essentially a very important point when we talk about linear regression; for example, all these functional forms if we look at, they are reduced to y equal to a 0 plus a 1, x y equal to a 0 plus a 1, x y equal to a 0 plus a 1 x; x as such is a non-linear function of the independent variable; y itself could be a non-linear function of the dependent variable; but in either case whether they are non-linear functions or not, finally the functional form that we are going to use in order to do the fitting should be a linear functional form of the type y equal to a 0 plus a 1 x plus a 2 u plus a 3 w and so on and so forth, if we go on back to what we looked at when we talked about the polynomial fitting part.

(Refer Slide Time: 21:43)

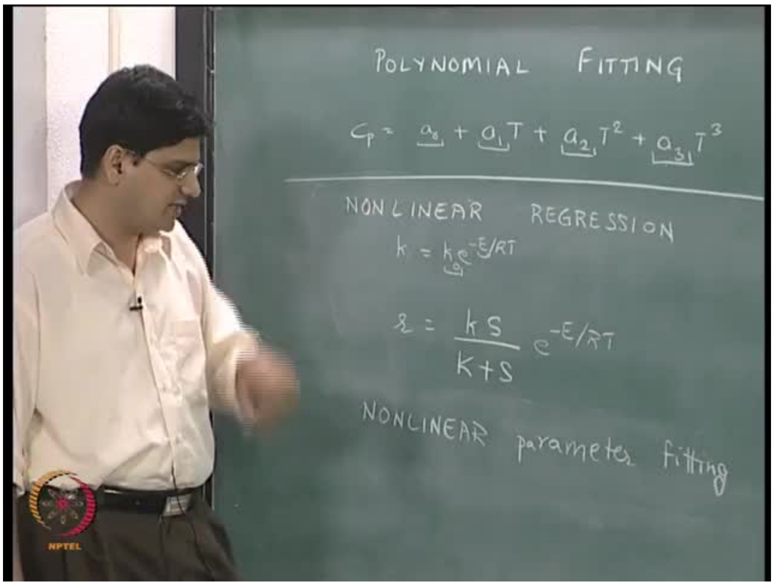


(Refer Slide Time: 22:45)



So, this is what we did with our polynomial fitting and one of the questions – that - that often get us, but this is not a linear equation, so how can we do linear fit? It is not linear in terms of the temperature T, however in terms of the parameters a 0 a 1 a 2 and a 3, this is indeed linear, what does linearity essentially mean? Linearity basically means that, there is either multiplication of that particular parameter by a scalar quantity or a known constant quantity or there is just simple addition, there are no sign terms, no exponential terms, no power terms and so on and so forth, in the quantity that we are interested in and when we talk about regression or curve fitting the quantity that we are interested in regression is the parameters, so these are the parameters that we are interested in; so, our overall expression has to be linear in parameter space.

(Refer Slide Time: 24:01)



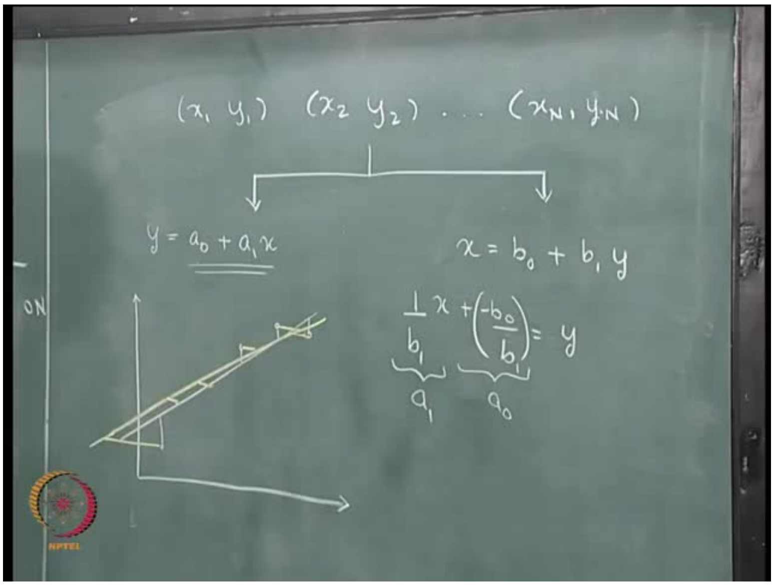
So, if we look at the parameters a 0 a 1 a 2 a 3, there is no term a 0 square or a 1 square or a 2 square or something to the power a 0 a 1 a 2 so on and so forth; as a result of - this - this polynomial fitting that we have is still a linear fitting; however, there are other certain conditions where the fitting may not necessarily be linear and the cases where, for example, the fitting are non-linear, of course the obvious example where the fitting is non-linear is indeed, if we write k equal to k 0 e to the power minus E by RT and we are interested in finding out k 0 and e without having to wanting to go through that transformation by taking a logarithm; if we do not do the transformation indeed this is a non-linear example, its only after the logarithm transformation that we will get this particular equation as say a linear equation.

So, likewise, the other example was the saturation type of an example, where we saw r equal to kS divided by K plus S; in this particular equation, again we had this as a linear form, which can be converted into a linear form essentially by doing certain transformations.

However, instead of this, if we let say for argument say we had this particular equation as, say e to the power minus E by RT, in a case like this, it is going to be difficult or in this particular case it is not going to be possible for us to - perhaps - split this into a functional form, which is going to be linear in the various parameter space; sometimes it is not possible to do that and when that is not possible, then we will need; if you want to fit, we want to do a non-linear parameter fitting, then we need to go to non-linear regression or non-linear regression is non-linear parameter fitting.

So, non-linear parameter fitting will be used when the model that we have developed is going to be non-linear in the parameter space; it may be linear or non-linear in the input outputs space, but it has to be non-linear, it cannot be linear in parameter space; if it is linear in parameter space, of course, we will go ahead and use the linear fitting - in - in this particular case and we will look at the gauss Newton method for non-linear regression in a short while.

(Refer Slide Time: 27:12)



So, let us say we have the data x 1 y 1 x 2 y 2 and so on up to x N y N, what we have seen so far is we try to fit a linear form of the form y equal to a 0 plus a 1 x; the other possibility is to fit the form x equal to b 0 plus b 1 y; the question that one would ask is, this particular form and this particular form, since both are linear, they are to be equivalent; so, the question is whether a linear regression technique of this form is going to give us the same straight line or not; for example, let us just rearrange this by subtracting by b 0 and dividing by b 1; so, we will get x minus b 0 divided by b 1 equal to y and comparing it y equal to a 0 plus a 1 x, what we get is a 0, is nothing but minus b

0 by - b sorry - b 1, I will just write this as plus minus b 0 by b 1, that is nothing but our a 0 and 1 by b 1 equal to b 1.

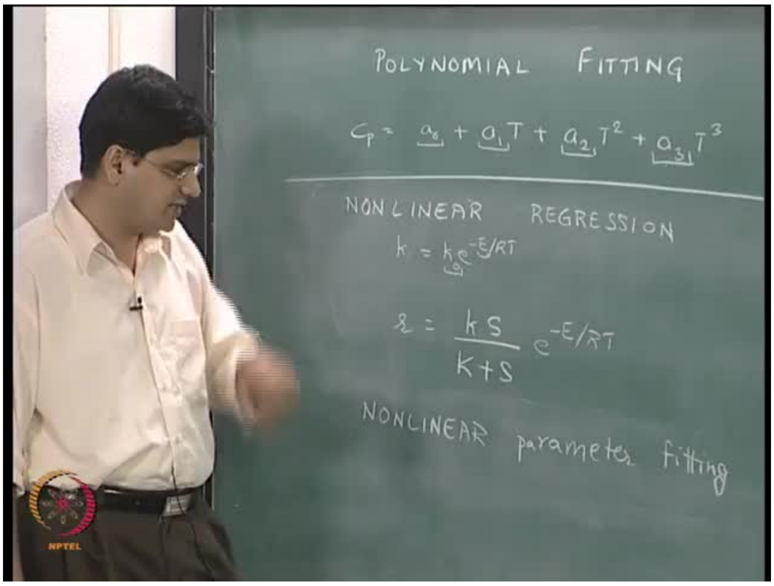
So, the what I am seeing is this; is let us say, we have data as before and let say we got this particular curve as y equal to a 0 plus a 1 x (Refer Slide Time 27:12) , then we try to fit another linear form x equal to b 0 plus b 1 y, what we see over here is the with a straight forward division of this type b 0 by b 1 multiplied by minus 1 and 1 by b 1; we will get this to be converted into y equal to a 0 plus a 1 x.

Now, the question is, is the curve y equal to a 0 plus a 1 x obtained from this particular method and the lines not the curve - the lines I am sorry - obtained from this particular method are they going to be one and the same line or not and the answer to that question is, it depends on what we use as our functional form in order to get this particular regression fit, remember the functional form that we use in order to get; so, this was the model and then we had certain errors, the errors in this particular case we assumed y i was a 0 plus a 1 x i plus e I; in this particular case, it is going to be x i equal to b 0 plus b 1 y i plus e I; so, in this case, what we are assuming was that the errors are in x and in this case we are assuming that the errors were in y.

So, what I mean by this is, let us say this is our best fit curve shown over here, then what we are actually minimizing? As we had seen - in the - in one of the previous lectures in the module is, we are minimizing these vertical distances between this straight line and the data this is what we are actually minimizing; but if we have to fit the curve x equal to b 0 plus b 1 y, it is not the vertical distances that we are minimizing, but its indeed the horizontal distances that we are minimizing.

So, let us say, this is going to be the best fit curve and this is the best fit curve because we are minimizing these horizontal distances; so, these are the horizontal distances that we are trying to minimize or under certain cases we would definitely get our a 0 equal to minus b 0 by b 1 and a 1 equal to 1 by b 1 to be same as a 0 and a 1 over here; but in general, for most cases, that is not going to be true; so the question is, whether you need to use a functional form y equal to a 0 plus a 1 x or whether you use x equal to b 0 plus b 1 y depends on what you know about the overall system, for example, if we go back to this particular example of trying to fit, C p as a polynomial form with respect to temperature.

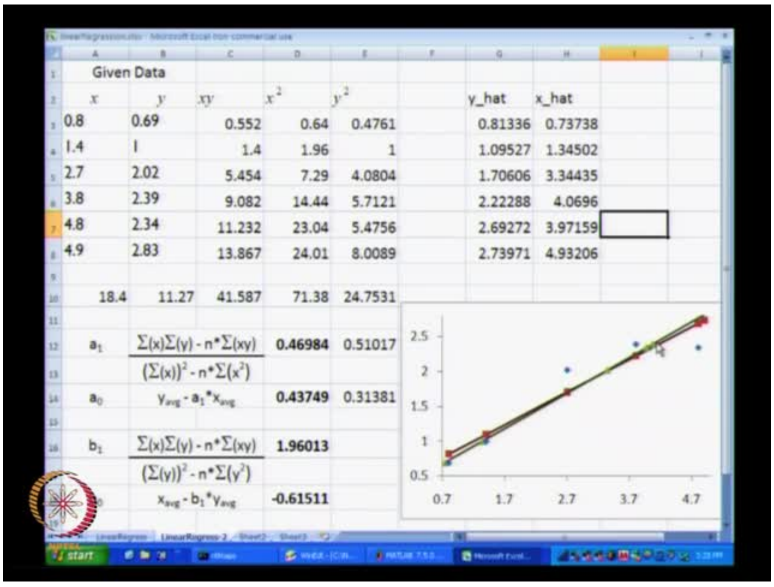
(Refer Slide Time: 32:31)



What we get is, essentially there are errors in measurement of the specific heat C p; there are errors in the measurement of temperature; so, there are errors both in x and y; in that case, if we minimize the errors in x versus we minimize the errors in y, we are going to get different results.

Again we will take up the example that we had worked in the previous lecture and see what results we will get if we are going to do regression of the form x equal to b 0 plus b 1 y.

(Refer Slide Time: 33:07)



So, this is what we had done in the previous lecture, we were given the data x and y, this was the data that we got and then we had obtained summation of x, summation of y, then we computed xy and summation of xy; we computed x square and this (Refer Slide Time 33:07) and then we obtained a 0 and a 1 using these formulae; this is the a 0 and a 1 that we had obtained in the previous lecture, I have just made them bold phase.

Now, what we are going to do is just interchange x and y data, because now x is going to be the dependent variable and y is going to be the independent variable and as we had done before our b 0 and b 1 will have almost the same form, the only difference is wherever we have x that will be replaced by y, wherever we had y that will be replaced by x.

So, we will have this as indeed, the numerator is indeed going to be summation of x multiplied by summation of y divided minus n times summation of x multiplied by y; if you replace essentially x with y and y with x, you are going to get the same expression; however, in the denominator what we will have is, we will have summation of y the whole square minus n times summation of y square, that is going to be different in this case and b 0 is going to be nothing but x average minus b 1 multiplied by y average; this is going to be our value b 0. So, those are the two main differences really that we will have in this particular example; so, I will just drag it upward so that it is all visible in a single.

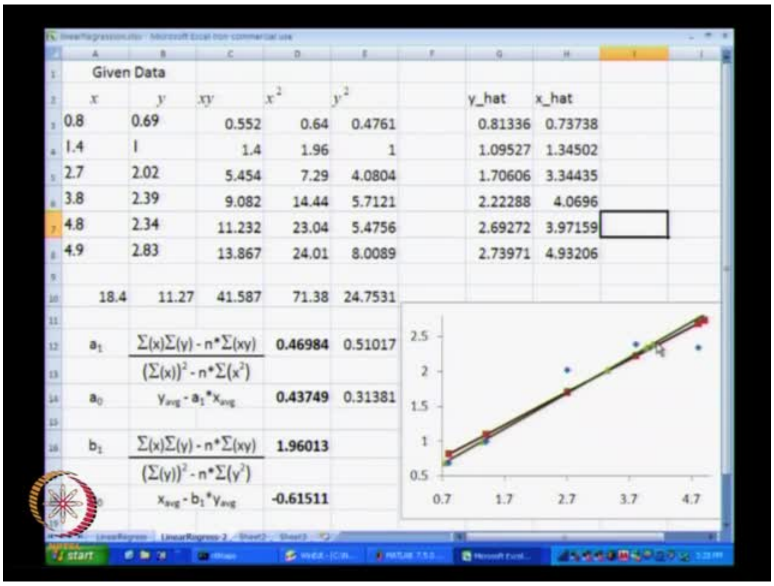
So, this is what we get for b 1 and b 0; so, for b 1 and b 0, we will really need is xy, we will need or x 0 and now we will also need our y square, because instead of x square we will be using y square in this particular expression; so, y square is going to be nothing but this to the power 2; I will just increase the font, so that it is all visible for us.

So y square is going to be this value to the power 2 and I am just going to drag it all the way, so that all the data is covered and we will get the summation and then we will just save it over here; y hat remember is nothing but our y that is obtained from the model given x from the model a 0 plus a 1 x.

So, same way now we are going to try to find b 1 value; b 1 value is nothing but summation of, this is not visible again so I will just do this very quickly and now we are good to go. So, now we have the expression for b 1; b 1 expression is summation of x which is this value multiplied by summation of y which is this value minus n which is 6 times summation of xy, which is this particular value.

So, A10 is summation x; B10 is summation y; C10 is summation of xy and 6 is our n, that is our numerator divided by the denominator and the denominator in this case is summation of y the whole square, this is summation of y, B10 is summation y whole square minus n, which is 6 multiplied by summation of y square, that is the value we have recently found and that is what our b 1 is and b 0 is nothing but x average; so, I will just change the font size here again, because it is not visible; so, this is nothing but x minus b 1 times y whole divided by n summation of x minus b 1 multiplied by summation of y whole divide by n which is n was 6.

(Refer Slide Time: 33:07)



So, this is going to be our b 0 and the above value was our b 1; let us find out what value of a 0 and a 1 that we get over here; a 1 is nothing but 1 by b 1 equal to 1 divided by b 1, which is 0.51 which is different from the previous value that we had obtained and a 0 is nothing but negative b 0 by b 1, so a 0 - I will just increase the font size here again - equal to negative b 0 divided by b 1 and this is the value of a 0 and a 1 that we will get corresponding to the fitting, instead of x versus y we have now fitted y versus x; so for the given y data, what is going to be x hat? Again as usual we will just increase the font size before continuing. So, we have x hat over here; x hat is nothing but x hat equal to b 0 plus b 1 multiplied by y I; so x 1 hat is nothing but b 0 plus b 1 multiplied by y.

Remember what we did in the previous lecture, because the terms b 0 and b 1 are not going to change when we drag the particular equation downwards, that is why we will put dollar signs over there; dollar signs basically mean that, as we drag this particular column D18 which is our b 0 value and D16 which is b 1 value remains the same, the only value that is changing as we drag this particular column downwards is going to be the b 3 value, that is going to change.

So, I drag this particular line downwards and this is what we get as x hat. So, now, let me add that data to our original data, so we will add this data, this first guy is going to be x hat, the second is going to be again our x hat value and the y axis is going to be the y values.

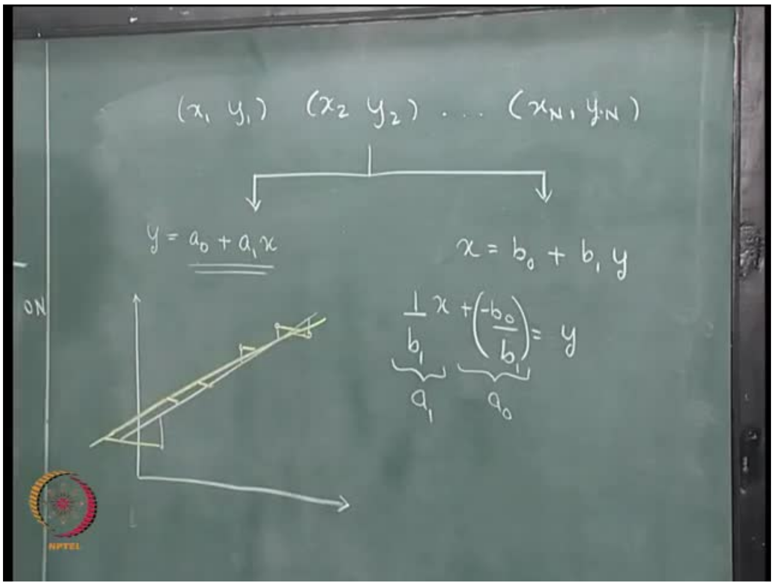
So, these are essentially our green data points, I will select the green data points and format data series and I will just zoom in format axis and I will just zoom into that particular axis by taking values of x as from going from 0.7 to going to 5. And likewise, the y axis also we will zoom in the reason why I just want to zoom in and show you essentially is to see that the two regression methods are indeed going to give us different results; so, we have 2.8; so, this red curve that I have highlighted now is the curve that we had fitted the data y hat equal to a 0 plus a 1 times x hat and this green curve that I have highlighted now is the data that is fitted, is x hat equal to b 0 plus b 1 y.

So, as you can see these two lines are indeed two different lines, what happens in case of the red line is, we have this minimization of these vertical distances between the data point and between the actual line, whereas in case of the green line, the minimization is between the horizontal distances over here between the blue data's and the green data points.

So, this is the two different cases or two different ways for the same data the linear regression is behaving - one case when the error is in x variable; the another case is when the error is in y variable.

We will go back to the board and we will take up the same chart, we will look at again and we will see what else can actually be done and this goes into what is known as multi variable data regression type of an idea.

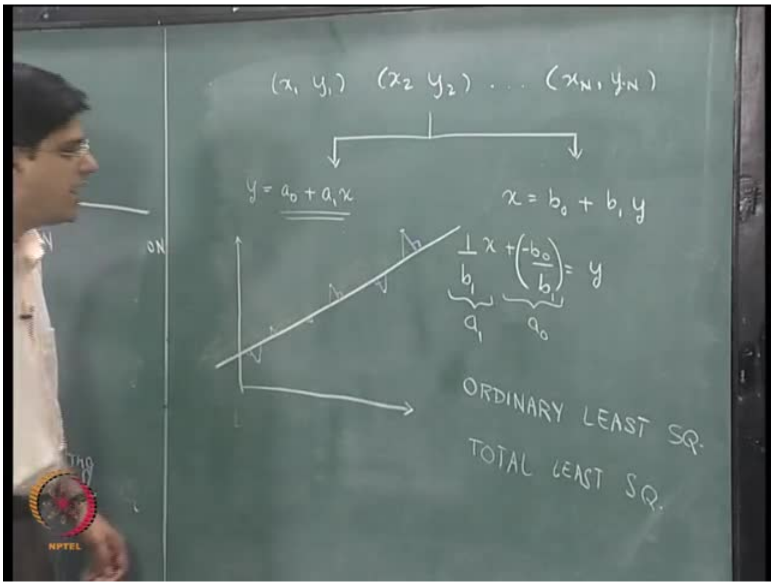
(Refer Slide Time: 45:05)



What we have done so far; now if you want to go into advanced data analysis stuff, which we are not going to cover, but just a very quick two minute summary of that is, what we have covered so far is what is known as ordinary least squares.

So, in ordinary least squares what we are doing is, we are minimizing the vertical or horizontal distances, or horizontal or vertical error between the best fit line and between the data points; An alternative to this is, what is known as total least squares or error in variable methods and so on and so forth. There are various different methods that use the statistical properties of the data, that we have in order to get a better fit of the overall curves and simplest way of looking at it is, rather than minimizing the vertical distances, we can minimize the perpendicular distance of the particular straight line for the best fits the data and the data point.

(Refer Slide Time: 46:24)

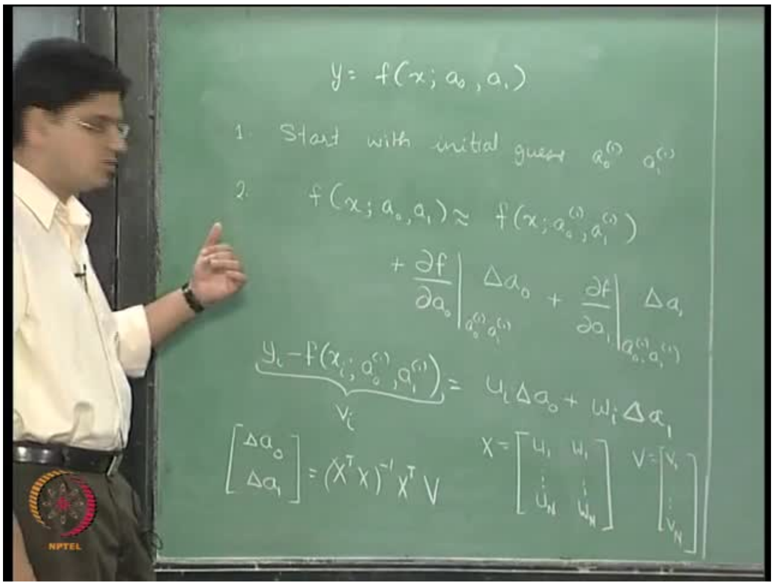


So I will just erase this (Refer Slide Time 46:24) and draw the various points all over again and where I am just widely spacing these points so that and let us say, this is the best curve line.

So, what we have seen when we are trying to fit a 0 y equal to a 0 plus a 1 x; what we have done over there is, we are minimize the vertical distances that I am showing as thin white lines and we fitted x equal to b 0 plus b 1 y, we minimize the horizontal distances; another alternative is, if there are errors in both x and y variables, what we may want to do is to minimize the perpendicular distance of the data point from the line, that we have drawn over there and some of the advance techniques allow you to - actually - do this kinds of a least squares minimization.

So, that essentially covers the linear regression part. In the final few minutes of this particular lecture, I will just talk about non-linear regression.

(Refer Slide Time: 48:11)



And let us say that, we want to do a non-linear regression, when the equation is of the form say y equal to some function f of x comma a 0 comma a 1; in this particular case what we will do is, we will use a similar idea that we use earlier, for example, in Newton-Raphson's method or any of the earlier methods, what we will do is, we linearize the equation f with respect to a 0 and with respect to a 1.

So, we will start off with initial guess. Next, we linearize the function f with respect to a 0 and with respect to a 1; so, f of x, a 0, a 1 can be written as f of x a 0 1 a 1 1 plus partial f by partial a 0 computed at a 0 1 a 1 1 multiplied by delta a 0 plus partial f by partial a 1 computed at a 0 a 1 multiplied by delta a 1; and we are ignoring any of the higher order terms over here.

So, y is going to be nothing but equal to this particular value; we will take this value to our left hand side and what we will get is this, is y i minus f of x I, x I; a 0 1 a 1 1 is going to be equal to dou f by dou a 0 computed at a 0 a 1, we will call that our u i and this we can call this as say v i.

So, v i is y i minus f of x i a 0 1 a 1 1, u i is nothing but partial f by partial a 0 computed at the current value and w i is again this particular expression and from this, we will essentially we able to get delta a 0 delta a 1 is nothing but X transpose X inverse X transpose V, where V is v 1 v 2 up to V capital N, x is u 1 u 2 up to u capital N, w 1 w 2 up to W capital N.

So, our capital X u 1 up to U N; w 1 up to W N and our capital V is nothing but v 1 and so on up to V N; so the overall procedure is to start with initial guesses a 0 and a 1, then compute delta a 0 and delta a 1 using this particular linear regression; based on delta a 0 and delta a 1 we will get a 0 2 as a 0 1 plus delta a 0 a, 1 2 is going to be a 1 1 plus delta a 1 and so on and keep repeating until there is no more change in the delta a 0 and delta a 1 value.

So, that essentially ends our the third lecture of this module and with the third lecture of the fifth module, we have finished talking about regression. From the next lecture onwards, I will talk about interpolation. In regression, what we did is, we started off with linear regression in one variable moved on to multi-linear regression. We looked at two different methods of doing multi-linear regression, both of this methods are exactly equivalent to each other, only it expressions are slightly different from each other. Next, we look at the regression for polynomial functional form polynomial regression, after that we covered functional regression; then we took up an example where we compared the regression of the form y equal to a 0 plus a 1 x with the regression of the form x equal to b 0 plus b 1 y; although those equations could theoretically be for equations for the same line, they are inter convertible into each other. We saw that the two different ways of doing these regression; in the first case, we minimize the vertical distances when we plot y versus x; in the second case, we minimize the horizontal distances when you plot y versus x. And finally, very briefly we covered the non-linear regression part just involves linearizing the overall regression module with respect to the regression coefficients a 0 and a 1.

So, that is essentially what I intended to cover in regression; for the most part, it is linear regression and the reason why we spoke about non-linear regression as well as total least squares idea is to motivate some of you to go ahead and read up these advance topics if you have more interest in these topics. From next lecture onwards, we will start talking about interpolation and we will start off with polynomial interpolation.

Thank you.