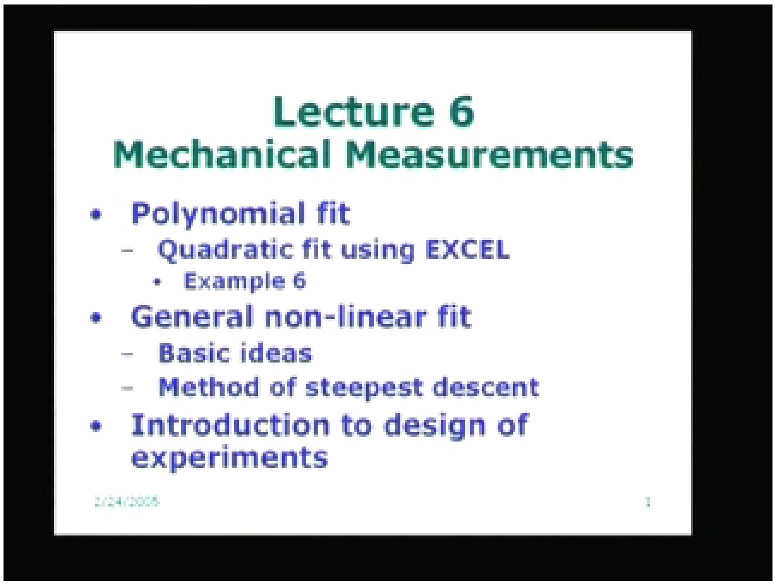
Mechanical Measurements and Metrology
Prof. S. P. Venkateshan
Department of Mechanical Engineering
Indian Institute of Technology, Madras
Module - 1
Lecture - 6
Regression Analysis (Continued)

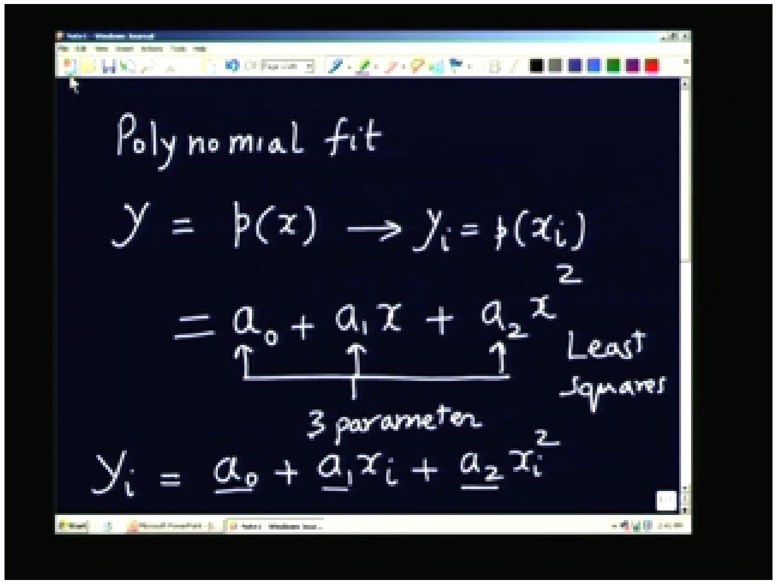
So this will be lecture no. 6 on mechanical measurements. Towards the end of the last lecture, we were talking about nonlinear fit and in particular, we talked about the polynomial fit. So what I intend to do in the present lecture is shown in the slide.

(Refer Slide Time 1:22 min)



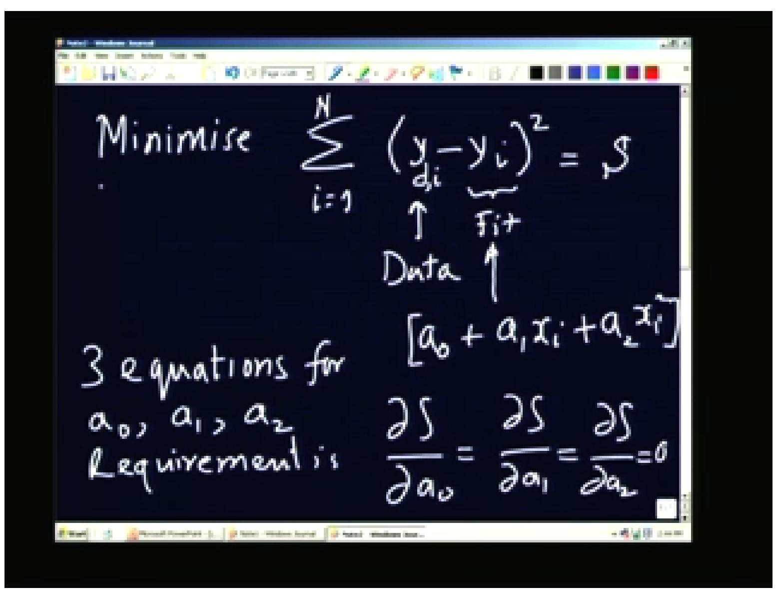
Take an example of quadratic fit, and before taking the example through EXCEL, I will actually work out a few steps on the board so that we understand what we are talking about. I will actually derive or show how the normal equations are written and then I will follow it up with example 6, which talks about the quadratic fit for a particular data set. Having discussed that, we will consider the general nonlinear fit in some detail. As I mentioned in the previous lecture, it will not be very detailed. It will only give us a sketchy description of how to go about doing a nonlinear fit. And I will just discuss in brief the so-called method of steepest descent, which is one of the optimization methods that can be used for nonlinear regression. Time permitting, I will give a brief introduction to the experiment design, and of course, it will be taken up in more detail in the subsequent lecture. So just to recapitulate what we have done (Refer Slide Time 4:40 min), we were talking about a polynomial fit, which means that I have a nonlinear relationship between variable y and variable x, I am just assuming one variable here, for the sake of simplicity. So I am thinking in terms of y being given as a polynomial of x. A particularly simple example will be quadratic, which will be given by a_0 plus a_1 x plus a_2 x square.

(Refer Slide Time: 5:45)



The three parameters which characterize this fit are: a_0, a_1, and a_2. So the expression which I have hired, y equal to p of x, will be used when it comes to the application of this model as y_i equal to p of x_i. That means I can elaborate it and write as y_i equal to a_0 plus a_1x_i plus a_2x_i square. So the problem consists of determining the constants a_0, a_1, and a_2, by using the method of least squares because that is common for all regression models. So let us just elaborate on this little more.

(Refer Slide Time 7:25 min)

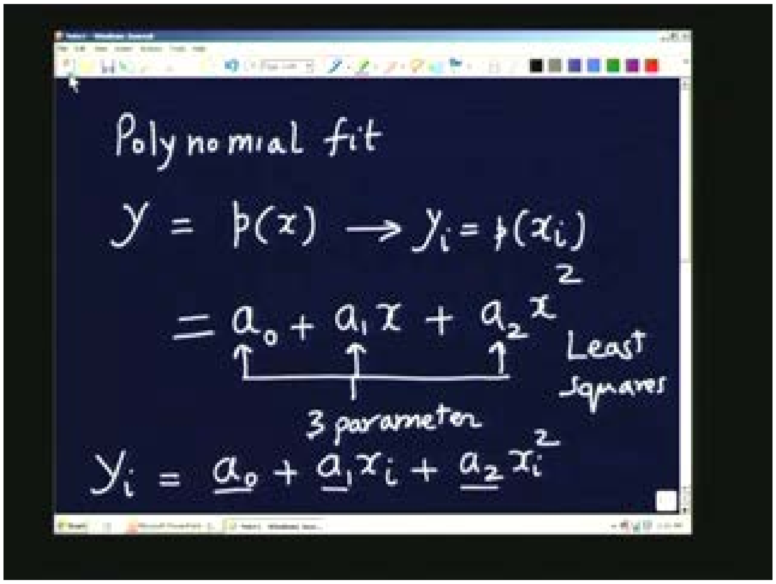


So what I will do is, I will try to minimize the sum of squares, which is given by sigma i equal to 1 to N y minus y_i whole square. This will be data, comma i, and this will be y_i, which is the fit.

This is the fit and all I have to do is to substitute for y_i expression from the previous page. See here. We have y_i equal to a_0 plus a_1x_i plus a_2x_i square. I just have to substitute here into this one, and will say this will be a_0 plus a_1x_i plus a_2x_i square. So the procedure is very similar to what we did when we were doing the linear regression model. If you remember, we were taking the partial derivative with respect to each parameter in succession and then writing the appropriate equation, which makes the partial derivative 0.

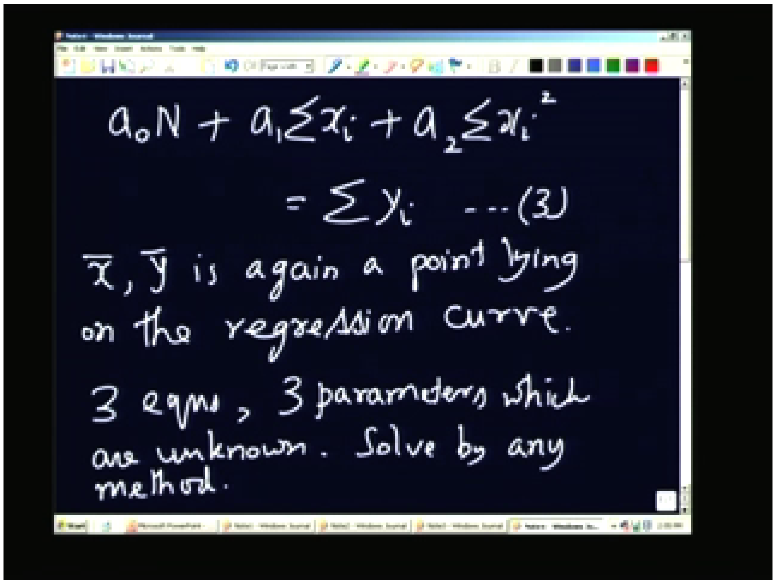
For example, in the case of linear regression, we had two such equations. In this case we will have three equations, which are required for finding out three parameters a_0, a_1 and a_2. Suppose I call this quantity as some capital S, my requirement is *doh* S by *doh* a_0 equal to *doh* S by *doh* a_1 equal to *doh* S by *doh* a_2 equal to 0. That is the requirement for the minimization of this. That means this minimization requires these steps. So if we do that, you can immediately see that instead of two long normal equations, we are going to get three normal equations. Let us just see what the normal equations are, by rating down the expressions for *doh* S by *doh* a_0 or any one of them.

(Refer Slide Time 4:40 min)



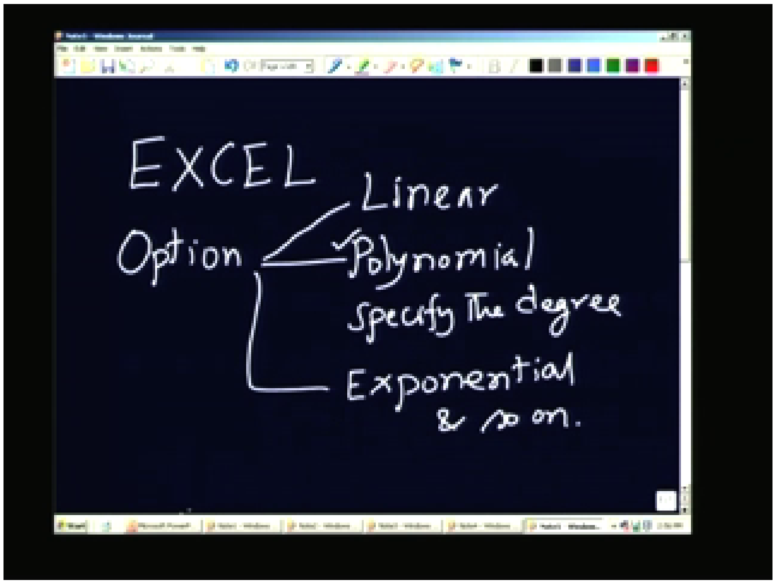
For example, I can take *doh* S by *doh* a_2; this will be sigma i equal to 1 to N, 2 into y_{di} minus a_0 plus a_1x_i plus a_2x_i square multiplied by x_i square. I am now going to differentiate with respect to a_2 because I have taken differentiation with respect to a_2 here. That gives you x_i square. If you want you can put the whole thing in a flower bracket because the entire thing here is dependent on x_i. Of course negative sign can be introduced here. So you see that now if I rewrite in a simple slightly different form, as is clear to us from earlier treatment, you will have a_0 sigma x_i square. From here to here, the product of these 2 plus a_1 sigma x_i cube plus a_2 sigma x_i to the power of 4, this equal to sigma y_{di} multiplied by x_i square. We will refer to this as equation no. 1. It is very clear if you remember what we did earlier when we had only 2 equations. This equation was absent in that case. You will also notice that the summation involves 4^{th} power of x_i, then the cube of the x_i and the square of the x_i. This is the equation no. 1 in this case. So by recapitulating what we did earlier,w e can simply write the next 2 equations and I would like the student to find out how to get them. So the next equation will be a_0 sigma x_i plus a_1 sigma x_i square plus a_2 sigma x_i cube equal to sigma $y_{di}x_i$. I can call this as equation no. 2. You will notice that a_2 is multiplied by instead of x_i to the power of 4 with x_i cube; so 1 power less than this, and this is also 1 power less, and this is also 1 power less. So this is x_i square; it becomes x_i in the case of the second equation, so this is x_i square and this is x_i.

(Refer Slide Time 12:48 min)



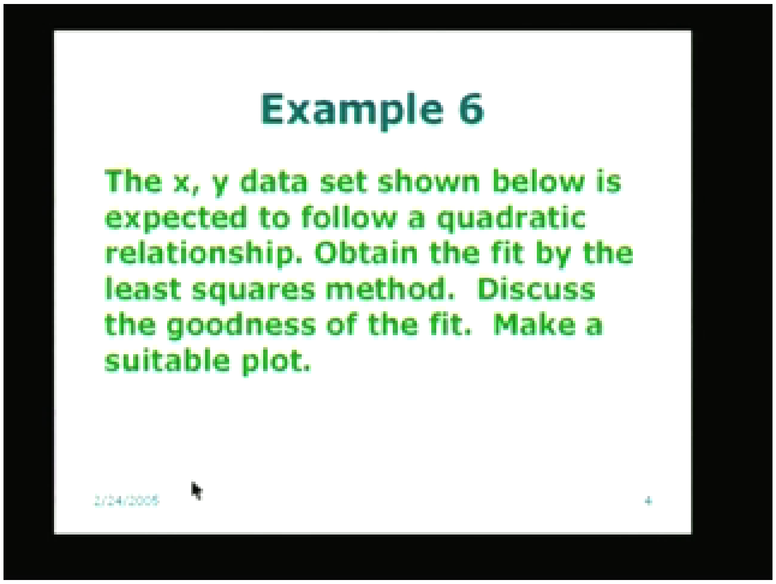
So, similarly we can write down the third normal equation, which is obtained as a_0 into sigma 1 or is equal to N times a_0 plus a_1 sigma x_i plus a_2 sigma x_i square equal to sigma y_i. Actually, this equation has a simple meaning that, y bar x bar is again a point on the regression curve. So x bar y bar is again a point lying on the regression curve. This is a quadratic in this particular case. So we have 3 equations and 3 parameters, which are unknown, and we can easily solve for this, by using Cramer's rule or by successive elimination. We can do that, solve by any suitable method. Of course, this is how one would get the answer by actually working out the problem.

(Refer Slide Time 14:08 min)



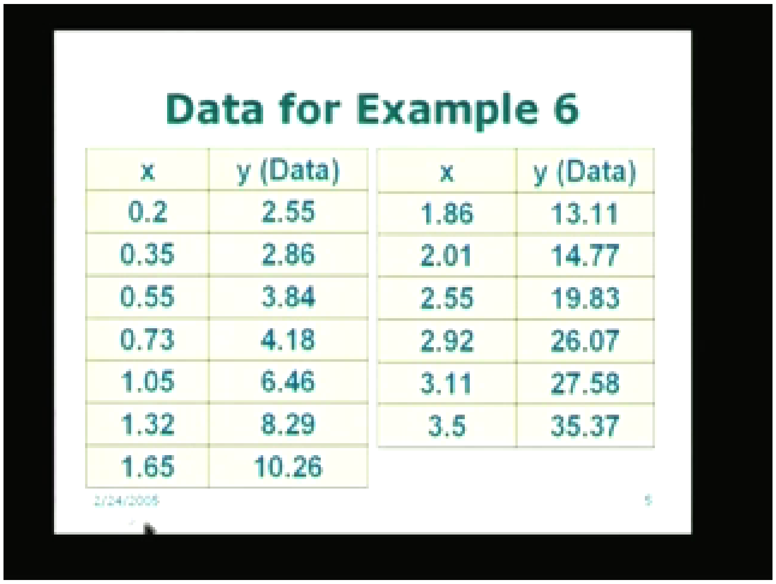
However, in EXCEL, as we indicated in the last lecture, there is an option. When we go to the chart option, one of the options is going to be—in fact we had large number of options—in fact to recapitulate you can have a linear, polynomial. When I say polynomial, I can specify the highest power or the degree of the polynomial. We can also have exponential moving average and so on. So one just has to use this option and whatever calculation procedure I showed is automatically performed by the EXCEL itself and I am going to get the answer straight away. So in order to appreciate this, what I will do is I will go back to the slide which depicts some of these. So let me look at the particular slide, which gives some information about that. I will just recapitulate whatever I wrote on the board, so the polynomial fit is easily done using analytical method by formulating the problem as an optimization problem, deriving the 3 normal equations. We can obtain summation, which we worked out on the board or I can use EXCEL, which was the built-in routine for this case. What I am going to do is in the next two or three slides, I am going to show how this can be easily done by EXCEL.

(Refer Slide Time 15:21 min)



So let me take a typical example. The x, y data set is shown in the form of a table and we expect a quadratic relation between x and y. We are expected to obtain the fit by least squares method and discuss the goodness of the fit. And if you remember, in the last lecture when we were talking about nonlinear fit, we said that the appropriate parameter which tells us about the goodness of the fit is the index of correlation. In the case of linear fit, we call it the correlation coefficient and we are able to derive it by looking at the slopes of the 2 regression lines and looking at how close they are to each other by taking the ratio of these slopes. In the case of the nonlinear fit, such a simple interpretation is not possible. So instead of that, what we are doing is compare the error or the spread in the data with respect to regression line or regression curve and compare it with the spread in the y with respect to its own mean. And in fact we come out with the index of correlation as the appropriate parameter. So that is what I am going to show in this case. In fact (Refer Slide Time 15:21 min) the index of correlation automatically comes out as an output of the EXCEL calculation sheet.

(Refer Slide Time 17:04 min)

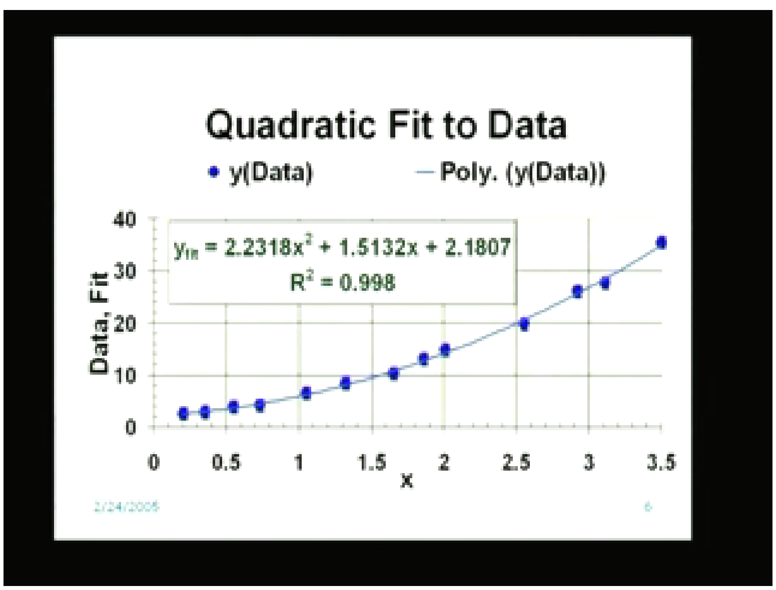


Data for Example 6

x	y (Data)	x	y (Data)
0.2	2.55	1.86	13.11
0.35	2.86	2.01	14.77
0.55	3.84	2.55	19.83
0.73	4.18	2.92	26.07
1.05	6.46	3.11	27.58
1.32	8.29	3.5	35.37
1.65	10.26		

Let us look at the example in the EXCEL sheet. So I have just taken the EXCEL sheet, copied the appropriate numbers on to the power point, so that we can see the numbers more clearly and more legibly. So the data is given with x ranging from 0.2 to 3.5 here, and y is the data, which goes from 2.55 to about 35.37 and already it has been made clear to us that we can expect the quadratic relationship between x shown in the first column with y shown in the second column. So let us look at the sketch before or plot before we go on to the goodness of the fit.

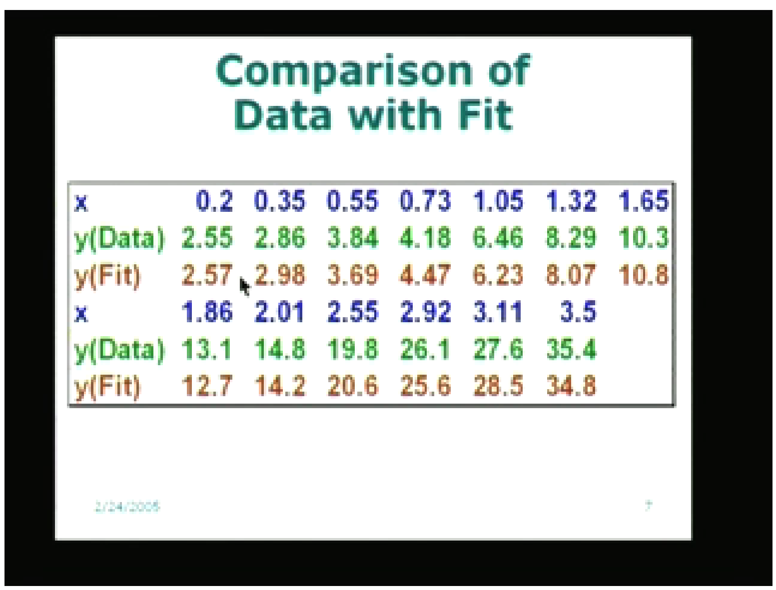
(Refer Slide Time 18:04 min)



So what I have done is I have used the EXCEL program to calculate the appropriate or choose the proper trend line and in this case, the trend line was a polynomial. When we choose the polynomial as the trend line we can show like that polynomial. Poly stands for polynomial so this is automatically showing that we have chosen a polynomial. And I have given the title Quadratic Fit to Data. That means I have chosen a polynomial with degree equal to 2. Actually what one can do is, we can, if no information has been given to you or if you do not know what is the proper thing which we should use, we can try out different fits using the trend line options available in the EXCEL program and then find out which is the best form by your stand; you look at it and then decide which is the best and then choose that.

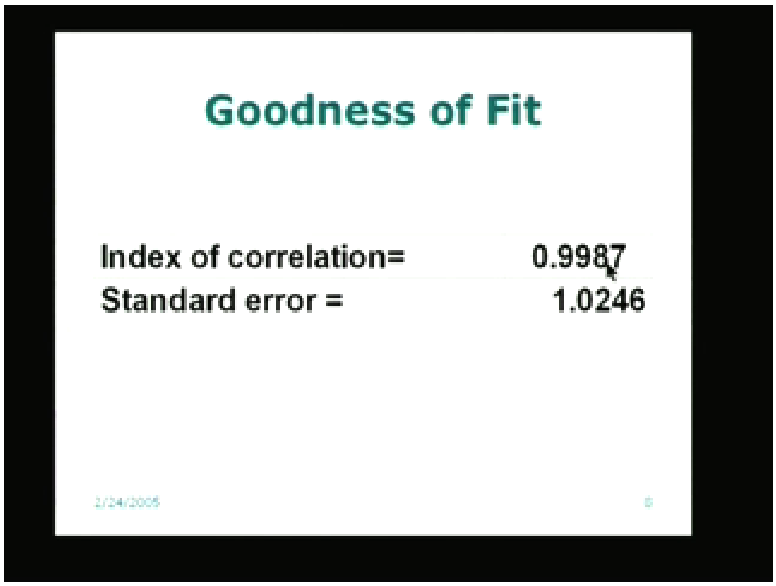
So in this case because the information was given that we are expecting a quadratic fit or a polynomial of degree 2 we have directly chosen the polynomial. So the fit is given by the formula 2.2318x square plus 1.5232x plus 2.1807. This is a_0, this is a_1 and this is a_2. I have written in the reverse order, and you see that the coefficients of all the terms a_0, a_1 and a_2 are roughly of the same order. That means all the terms are equally important in this particular example. This quantity, this as well as this, and if you remember the maximum value of x is 5, say 3.5 and x square will correspond to about a little more than 10. Therefore this will be a dominant term for a large value of x, and for small value of x, of course the other 2 terms or in fact the first term is going to be dominant and then slowly the second term becomes important and the third term is going to be most important when you go to very large values of x along the x axis. I have also plotted the error bars on this; before we do that, I will look at the how good the fit is and then come back to that sketch.

(Refer Slide Time 20:34 min)



So I have compared the data and fit, by actually evaluating the values of the fit, values at the corresponding values of x given in the table, and I have made a table which gives you x values in the first row. The second row consists of the data that has been given to us. Of course I have rounded off the y values calculated by the fit again to 2 decimal points so that we have consistent comparison between the 2 data and the fit and you can see that if you take a look at any of them the difference between these 2, 2.57, 2.55, here 8.229, 8.07 and if you go here, 35.4, 34.8, the difference between these 2 values, the one which is shown in green and data values shown in brown color, these differences are small indicating that the fit must be a good fit. And in fact it is further amplified by looking at the R squared value. As I indicated, the R squared value also comes out from the EXCEL program. It directly gives you the index of correlation and that is given by 0.998, which is a value very close to unity and therefore obviously it means that the fit is sufficiently good for our purposes.

(Refer Slide Time 21:18 min)

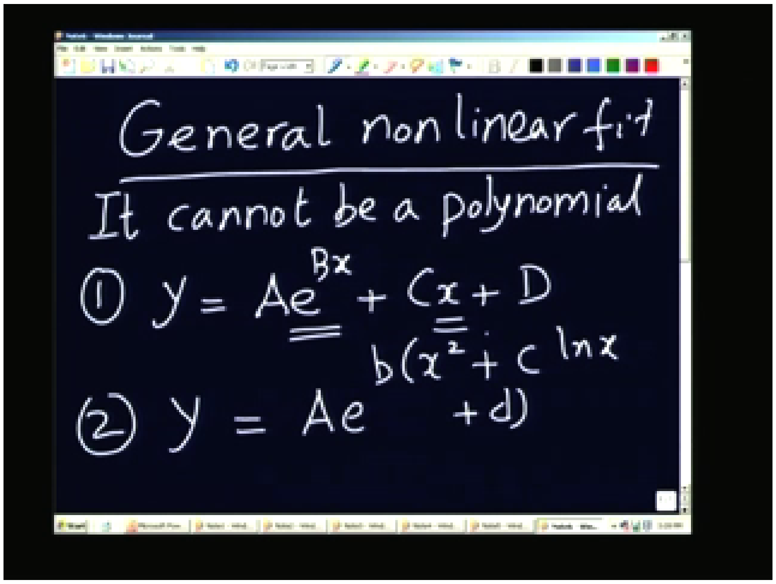


And in fact I have myself calculated the coefficient of correlation by using the definition given earlier and the index of correlation is in fact nothing but a square root of the R squared value. And the square root of the number which is less than 1 will always be more than the other, so this is slightly bigger, we are getting almost 0.999 index of correlation. The other one was 0.998. And in fact I can also calculate the standard error. How do I calculate the standard error? I am going to calculate the difference.

If I go back to the previous slide, I am going to take the difference between these 2 quantities, the green and the brown values, square them and add all the points which are given here and I am going to divide it by n minus the total degrees of freedom, which is n minus 3 because there are 3 parameters, which have been determined by using the normal equations. Therefore I am going to divide it by not 13. Instead of 13, I am going to use 13 – 3 is equal to10. Therefore that is going to give me the expected error or the standard error and it comes to 1.0246, I can round it off to 1.03 and what I have done is in the plot, I have shown the plus or minus that value as the appropriate error bar for the values. That means the fit values, which represent the data, if the errors are within that region, which is shown by the error bar, will give you comparable results. That means within that limit you can use the fit as a representation of the data. So this is something about the quadratic and the polynomial fit. Now let us look at the general nonlinear fit problems. The difficulty with respect to the general nonlinear fit problem is that it may be represented by an equation which is difficult to handle. Of course any fit will have certain parameters, which is going to characterize that and in fact, in a few minutes, I will show you some of the nonlinear fit, which cannot be handled by actually deriving the normal equations and so on.

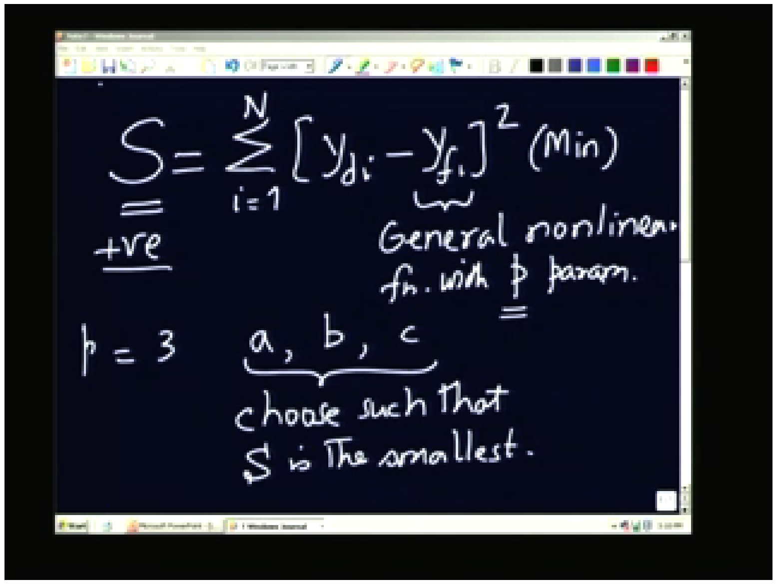
So, only under such condition we need to go about solving, using some method that actually optimizes or minimizes the sum of the squares of the error. That is what I mean by an optimization problem. Let me just go to the board and try to look at some of these.

(Refer Slide Time 26:20 min)



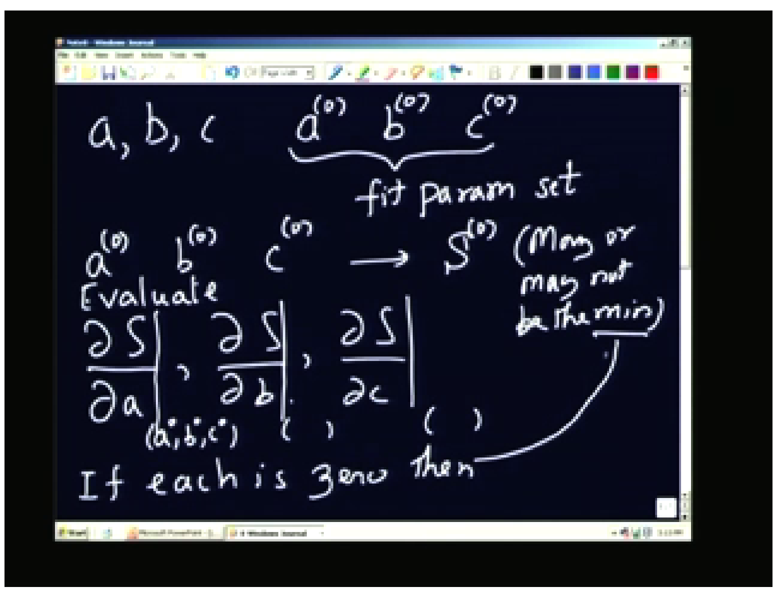
Let me just take general nonlinear fit, which is an expression that cannot be written in a form that resembles the quadratic or a cubical form. It's not a polynomial, so it cannot be a polynomial. Because if it is polynomial, we do not need to worry about the general procedure. Let me just give one or two simple examples: (1) y equal to A e to the power of Bx plus Cx plus D. This is an exponential term, this is a linear term, the combination of exponential and linear. If I write down the normal equation it will be difficult to do that. So this is one case which is not amenable to this simple method. The second expression, for example, may be I should put one more constant here. We can put d here, c here, so you see that in this case—let me just change it—I didn't use 1 or 2 more constants, for example. So these 2 expressions are what we can classify as general nonlinear fit and it is difficult to use the method which was used in the case of polynomial function. Just to develop the method, let us look at what we are trying to do.

(Refer Slide Time 28:06 min)



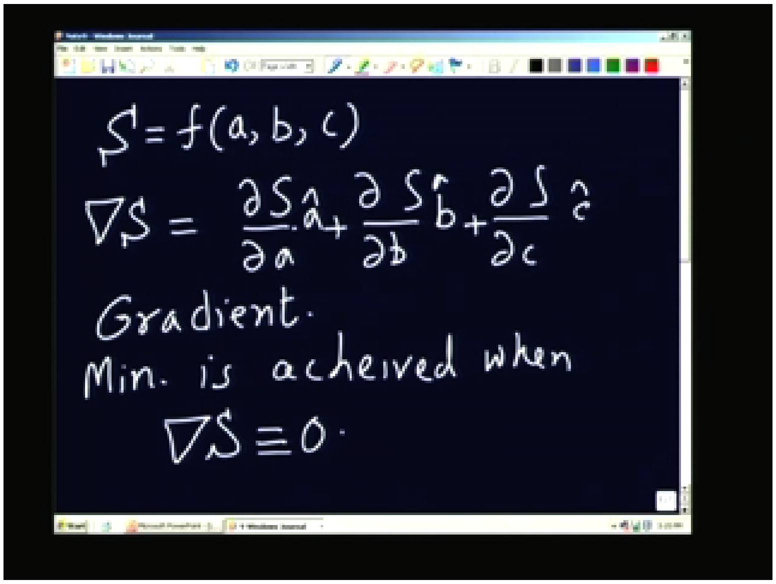
So the general problem is that I have got the sum of squares as usual, i equal to 1 to whatever n number of data. So y data minus y_i I will say fit i whole square. So we want to minimize this, and this is a general nonlinear function with p parameters where p may be 2, 3, 4; whatever the number. So what I have to do is I have to find out the values, for example, if p is equal to 3, that means I will have a, b, c, choose such that or find such that S is the smallest. So before we proceed, let me just look at this quantity S here, because it is the sum of squares it is always positive. So suppose I make a sketch of S for example, you remember that I have a, b, and c, I have got a b and c, so if I use values that is $a^{(0)}$, $b^{(0)}$, $c^{(0)}$ arbitrarily; I am just doing $a^{(0)}$, $b^{(0)}$, $c^{(0)}$: 3 different constants, which describe the fit and therefore we call it the fit parameter.

(Refer Slide Time 31:01 min)



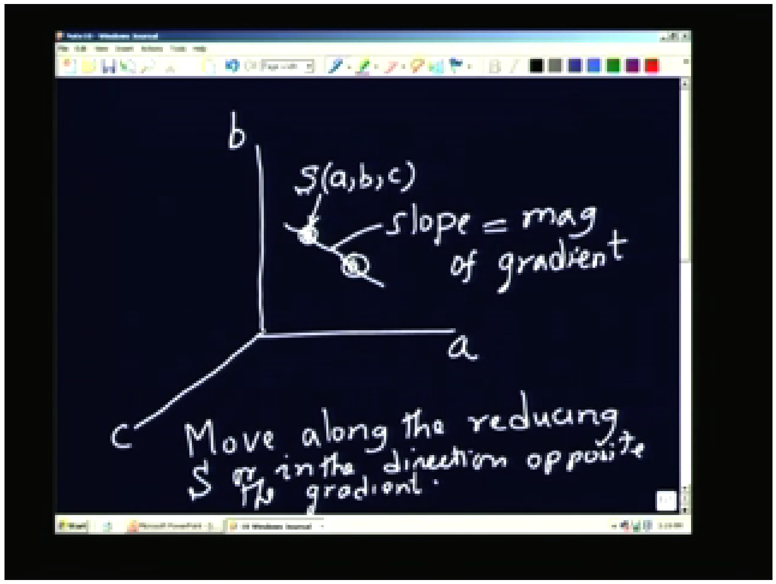
I will call it fit parameter set. That means I am looking at 3 values: a, b, and c at the same time and what is its effect on the value of S. So I can say that if I have $a^{(0)}$, $b^{(0)}$, $c^{(0)}$, as the 3 values, then it will give you some $S^{(0)}$. This may or may not be the minimum. So we will say, may or may not be the minimum. How do we verify whether it is a minimum or not? All I have to do is to find out the partial derivatives and compare with 0, if each one of them is 0, then of course we can say that S has become a minimum therefore S is of minimum value. So we have to evaluate the partial derivative with respect to a, b and c at these assumed values; so I will put a bar here and say $a^{(0)}$. So I am evaluating the derivatives with respect to a, with respect to b and with respect to c. If each is 0 independently, then it is a minimum. That means I have to verify the value of the partial derivatives at each stage. So let us look at these partial derivatives little more closely and look at a Taylor expansion, for example, y is the function of a, b and c. I am talking about the parameter instead of y, we should use the value of S here. So S is a function of a, b, c. After adding all the terms I am getting this, therefore now S has a dependence on a, b, and c values. For example, if it is $a^{(0)}$, $b^{(0)}$, $c^{(0)}$, S equal to some $f(a^{(0)}, b^{(0)}, c^{(0)})$.

(Refer Slide Time 33:10 min)



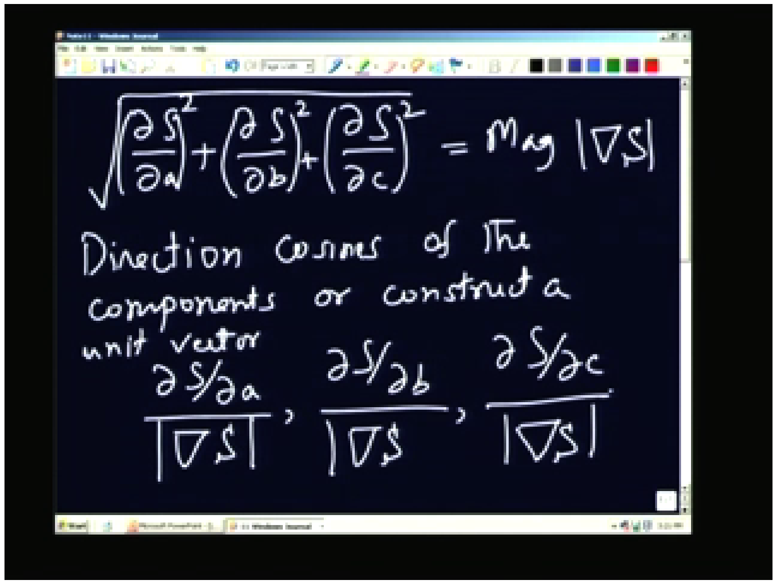
So I can find out the slope or the gradient of this quantity. This will contain 3 components, which are nothing but *doh* S by *doh* a, *doh* S by *doh* b, *doh* S by *doh* c and I will in fact write this along this direction of a. So I will say unit vector along the direction here and here, so I will say this is the gradient. So the requirement for a minimum is achieved when it is equal to 0. That is what we are looking for. We will also look at this gradient in slightly more detail now. What does the gradient indicate? It indicates the partial derivatives with respect to the 3 parameters, so I am talking about the parameter space. May be I can even make a simple sketch to show that, which I will do in the next time sheet.

(Refer Slide Time 36:14 min)



Suppose I plot a here, b here and c here. Each a, b, c combination is going to give one value of S and let us put that as the value here. Now the gradient is giving me the slope of the variation of S with respect to the 3 values a, b and c. So it will be a line which may be passing through these 2 points for example. The line is a transition to the surface, which is characterized by S(a, b, c). I have the point, which is S, and the gradient is a vector which gives the direction in which it is varying by the largest amount and therefore I can show the direction like this. Of course it can be going in the direction along which it is going to change and the slope is given like this. The slope of this line is the magnitude of the gradient. Slope equal to magnitude of gradient. We know that the slope or the gradient is directed in the direction in which the maximum amount of change will take place. What I will do is I will go in the direction of reducing value of S. So I will move along the reducing value of S or in the direction opposite the gradient. Let me explain by showing what is being done.

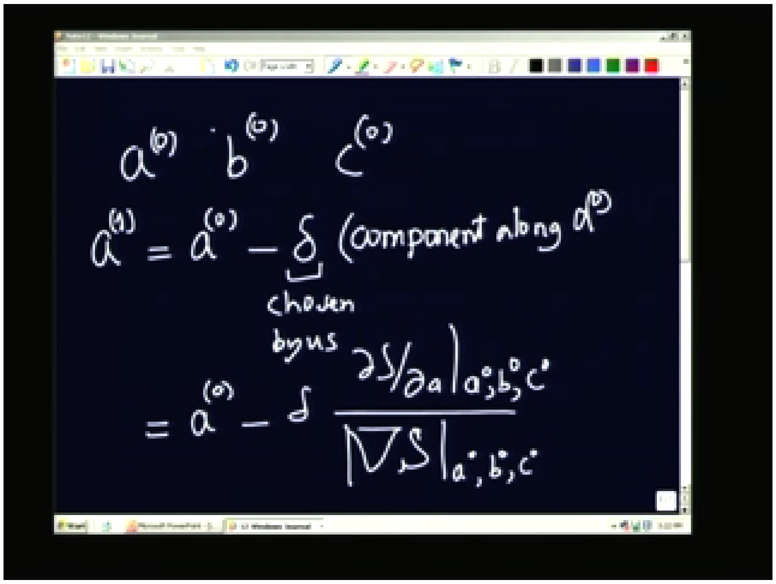
(Refer Slide Time 38:21 min)



We have already calculated the 3 components. Of course, there is a problem with more number of parameters, there will be more number of this, and in fact I can calculate the sum of the squares of these 3 and take the square root. This actually is the magnitude. From this magnitude, I can also obtain the direction cosines of the components or we can construct a unit vector given by the following: *doh* S by *doh* a divided by this magnitude, I will say magnitude of delta S.

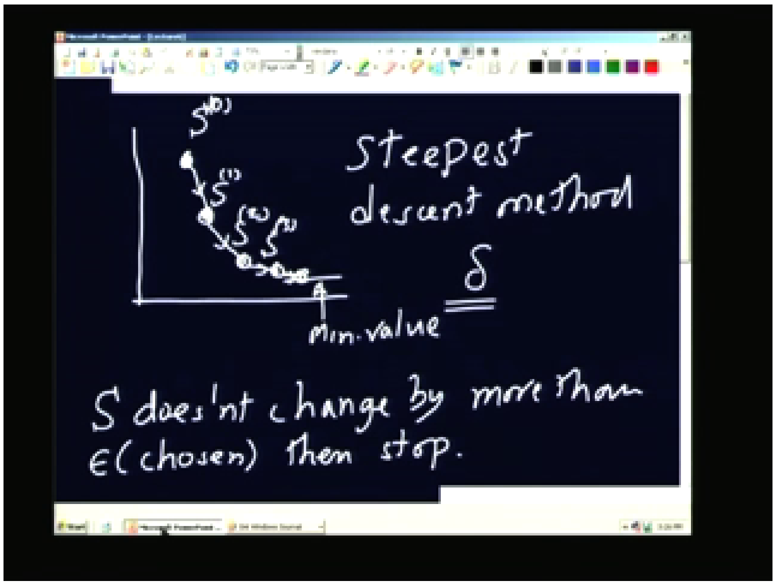
We have the 3 components of the unit vector, which is parallel to the gradient vector. So if I take the step I know that S is not a minimum because this magnitude should be 0. Because it is not 0, I know that it is not the maximum. So I want to go to the next point by taking a step, which is proportional to the component multiplied by certain small value, which I can choose and multiply it with the negative number because I want to go in the direction of decreasing gradient; therefore let us look at that.

(Refer Slide Time 40:04 min)



So what I will do is I start from $a^{(0)}$, $b^{(0)}$, $c^{(0)}$. Next I may take the step such that I will have $a^{(1)}$ is equal to $a^{(0)}$ – some small quantity, I will call it δ, which is chosen by us, multiplied by the component in the direction of unit vector component along $a^{(0)}$. This will be given by $a^{(0)} - \delta$ into *doh* S by *doh* a evaluated at a^0, b^0, c^0 divided by magnitude of S. Of course these are also evaluated at the same point.

(Refer Slide Time 43:45 min)



If I do that I am going to expect the following. I have to write such expressions for the others also. So $a^{(1)}$ similarly $b^{(1)}$; here I will have $b^{(0)}$ – δ times, same δ. δ is the same for the 3 increments because they have got to be parallel to that vector. You should not multiply by 3 different numbers. The delta might be the same for the 3 because I am going to move parallel to those quantities. Let us just look at a one-dimensional line just to see what is happening. So this is the point where the value of S is $S^{(0)}$ .I took a step in the direction of decreasing $S^{(0)}$ and therefore next I will come here. I have just exaggerated and as the next step I would do the exact same thing. This will be $S^{(1)}$,it will go on like this. So it is approaching the minimum value. So what I am doing is I am going in the direction exactly opposite to the gradient vector and therefore it is moving along the steepest or the fastest chain as far as the value of S is concerned, and therefore this is called the steepest descent method. I have indicated only for one parameter; this can be done for any number of parameters. You are taking small steps each time and moving towards smaller and smaller values of S and of course you stop somewhere when it doesn't make sense to do any more.

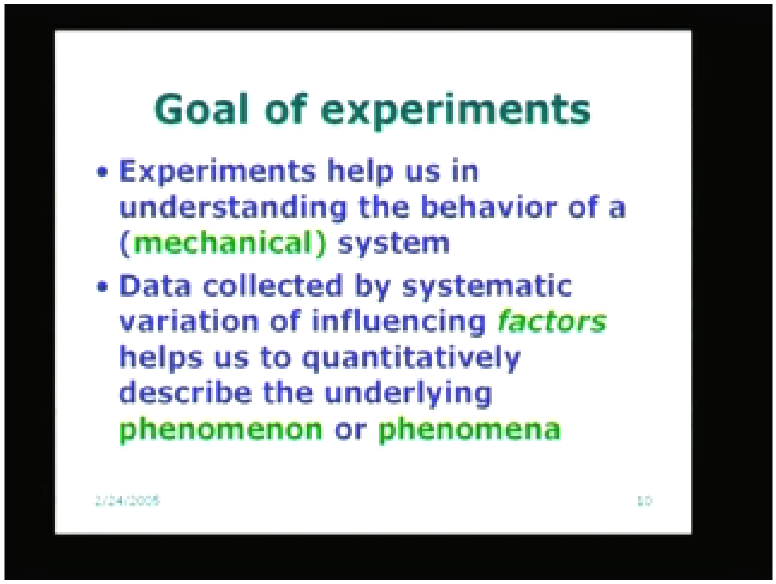
For example, if S doesn't change by more than some epsilon chosen by us, then stop. Of course in this method if you take too large a step, it may actually miss and it may start giving trouble as you reach the minimum point. Therefore initially, we may have to do a little bit of experimentation, find out the suitable value of the delta. Initially we may take slightly larger delta but once we close in on the minimal value then we should start reducing the value of the delta and obtain this. So the whole point of this discussion has been to indicate that the general optimization or general regression using a nonlinear model is possible by using a simple technique of minimizing by using an optimization technique like the steepest descent method. It may or may not work always and therefore there are better methods and so on, but we just want to understand briefly what we are trying to do. So in case you don't want to spend too much time and effort in doing this, you remember in linear regression and the polynomial exponential and so on, you are using EXCEL, which has all the built-in programs.

We can also use other programs which have built-in routine for doing nonlinear regression. For example, there is a program called Axum, which can do this, whatever I explained in the previous slides, whatever I wrote down on the board; the procedure of steepest descent and so on. These are built-in programs available in Axum and therefore you can use Axum for doing that or a program like sigmaplot can be used. There you can specify any complicated expressions in the regression model and then it will be able to do that. There is also a program called NLREG or Non Linear Regression available as a free download from the Net.

You can go to the Net and freely download this particular program and use it for doing nonlinear regression. So at this juncture, we will assume that we have learned enough about regression and of course one learns regression by actually collecting data, making plots, fitting curves and so on, which is an experience one has to learn by actually doing. Whatever we have given is only a sort of an introduction to the area of regression. There is lot more to it and the amount of time we have in this particular course is limited; therefore we have to stop here.

So what we will do is, now having understood something about measurement principles something about errors and their characterization, we talked about the specific properties of the error and then we proposed what we called the least squares principle and then based on that, we have been able to do the regression analysis, either linear quadratic, or nonlinear and so on. So this is the background, which is required by anybody who wants to do experiments and wants to present data in a proper form acceptable to other people who are working in the same field or who are going to be using your data; you must know all these things and use it carefully. So with this background let us look at what other things we have to do in conducting experiments and I am just going to give a brief introduction to the area of experimental design and in the next lecture am going to cover it in more detail.

(Refer Slide Time 47:37 min)

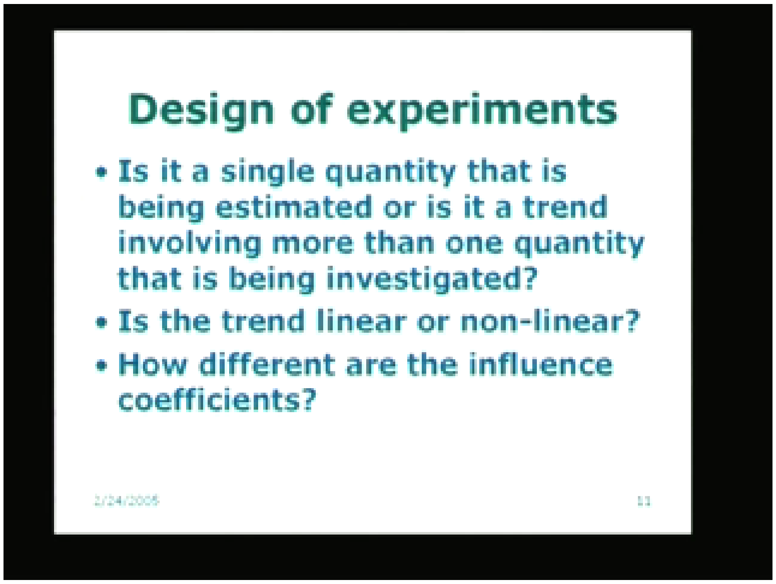


So what is the goal of the experiment?
In fact we already touched on it in the very first lecture. We talked about the need for goal of experiments and I am just recapitulating here, probably in more detail. So you can see from the slides that experiments help us in understanding the behavior of the mechanical system. I have put the mechanical system in brackets; we are only interested in the mechanical system in this particular course, it is equally applicable to any area of work in which you may be doing your experiments. So mechanical system is only one subset of all the possible areas in which one could do the experiments. What we do in our experiments is to collect the data by systematic variation. I am going to introduce some words, which are going to occur here and again in design of experiments. One of them is the influence factor. So there is a fancy name for the independent variable.

You remember? y equal to f(x); x is varied independently. x is a factor; so that means the outcome of the experiments, which is the measured value y, is influenced by the value of the factor x or the variable x, so we call it a factor. In an experiment, for example, temperature, pressure, humidity and so many different things may play a role and each one of them becomes a factor. This is just a terminology. The data collected systematically by varying the factors in a proper systematic fashion over the range of values, which we have already decided, what is it going to do? Is it going to help us to quantitatively describe the underlying phenomenon if there is one phenomenon? Or may be several phenomena are taking place in the same time. May be they are influencing each other. How to get the information about these underlying principles or phenomena, is our goal.

So what we are going to do is to do several experiments. How do we decide some of these issues? For that we have to look at what are the things we are doing. For example, in a manufacturing process if we are trying to make some machine part, you will be interested in the machine component and its quality and find out what all are the factors which are going to influence that. For example, in a typical process we may be using the machine to either modify the shape or the surface finish or whatever it is. In this process I am going to use some parameters like the speed at which I am going to run the machine and the depth of cut. So each one of these becomes a factor. The depth of cut, the speed at which the machine is running, becomes a factor. This becomes a factor and the end product, the shape of the product or the size of the product or the nature of the surface of the project depends on these factors. So what I would be interested in the point of manufacturing would be to look at the best suitable values of the parameters of these factors such that I am able to satisfy some requirements like the surface finish has to be within certain limits of the experiment.

(Refer Slide Time 50:42 min)



So it may be a single quantity that is estimated. Look at the slide; it is a single quantity that is going to be estimated. That means I am interested only in one particular process in which the outcome is one particular thing I am looking for and I want to find out how to choose the parameters of the factors such that I am going to get that. That is one kind of experiment, which is, of course, very important in mechanical engineering. Or I may be doing an experiment in fluid mechanics, where I am not interested in one particular small bit of information. I want to get the general understanding of what is happening, so I may be measuring one quantity which depends on a certain other quantity, like we were talking about regression.

Regression is finding out what is the linking relationship between two things: one which is measured or one factor which is kept constant or varied in a systematic fashion and the outcome of that experiment will be some quantity, which is of interest to us. In the case of a fluid flow problem, I may be interested in the pressure drop in a certain length of pipe, for example. It will be determined by various factors including the velocity of flow, the diameter of the pipe and so on. I would like to know what is the underlying relationship; I am not looking for one particular information, I am looking for a global information about the variation with respect to various parameters.

So the experiments we are going to do and the type of experiments we have to do will depend on some of these important things, whether you are interested in only one parameter: one quantity which you want to measure or optimize or you want to get the trend. Then if you are looking at the second kind of experiments where you want to find the trend, you would like to ask the question whether the trend is linear or nonlinear. So depending on the variations, if it is linear we can do very few experiments and get the information. If it is nonlinear, the relationship is not one to one, one straight line variation. Therefore we would like to know how undulating is the curve; how curved is the relationship, how nonlinear is the relationship. The more the nonlinearity, the more number of data we have to collect to represent that nonlinearity in greater detail.

Of course, this will become clear as we go forward in the next lecture. Then we would like to answer or ask the question how different are the coefficient influencing factors? That means, if you have a relationship or the trend between 2 variables or more than 2 variables, we would like to ask, what is the importance of each one of these? This can be important even in the case of manufacturing problem. For example, the case I gave, the surface finish may be different on 2 parameters or 2 factors, that is, the speed of the machine and the depth of that may be more influenced by one of them and less influenced by the other. So we would like to do the experiment and from the experiment find out which is the factor which has got higher influence. It also may be that there may be an interaction between these factors. If speed of the machine is very high and the depth of cut is very small or depth of cut is very high and the speed of machine is low, it may lead to different types of interactions. So the two factors may even interact at the outcome of this experiment. It may be a thing which depends on both of them simultaneously.

Therefore to answer all these questions, we would like to do the following. What are the different things available? I will just like to mention them, and in the next lecture we are going to take a deeper look into this. For example, in the case of fluid flow problems we may do a dimensional analysis and find out whether it can give some information regarding what we are looking for. So can we identify for example the dimensionless group that influences the quantity or quantities being measured. That is the question we would like to ask. We would like to find out whether it is possible and then of course, the most important thing we would like to find out is how many experiments we must perform to get the information we are looking for. This leads to the general problem of design of experiments and this involves several steps. What I am going to do is, I am going to address both the kinds of problems.

(Refer Slide Time 55:02 min)

You have one process and you want to optimize the process, find out what are the parameter factor values you must choose by simple experiments. Or you may want to get the trend among various factors and the quantity of interest to us to see how to design such experiments. Both the questions will be taken up and possibly, it’s going to take two more lectures. The subsequent lectures, 7 and 8, are going to deal with these things and that will bring us to the close of the first module of this set of lectures on mechanical engineering. Thank You.