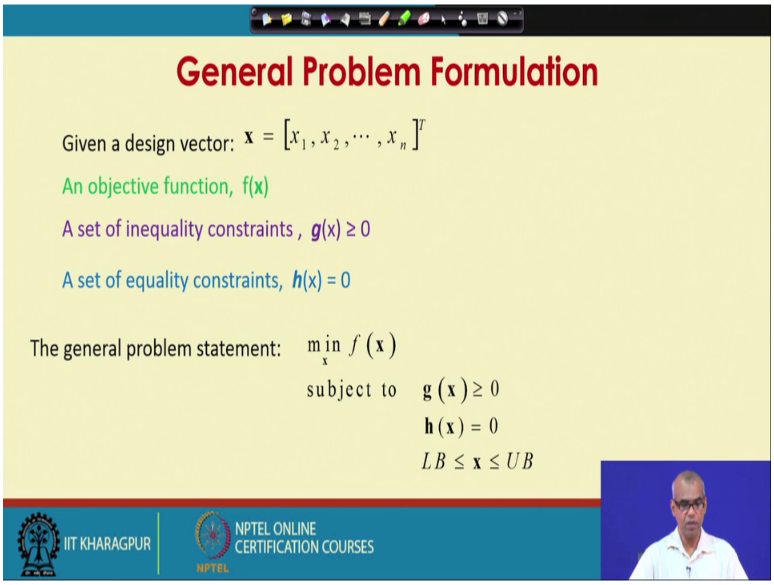
Optimization in Chemical Engineering
Prof. Debasis Sarkar
Department of Chemical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 09
Optimization Problem Formulation (Contd.)

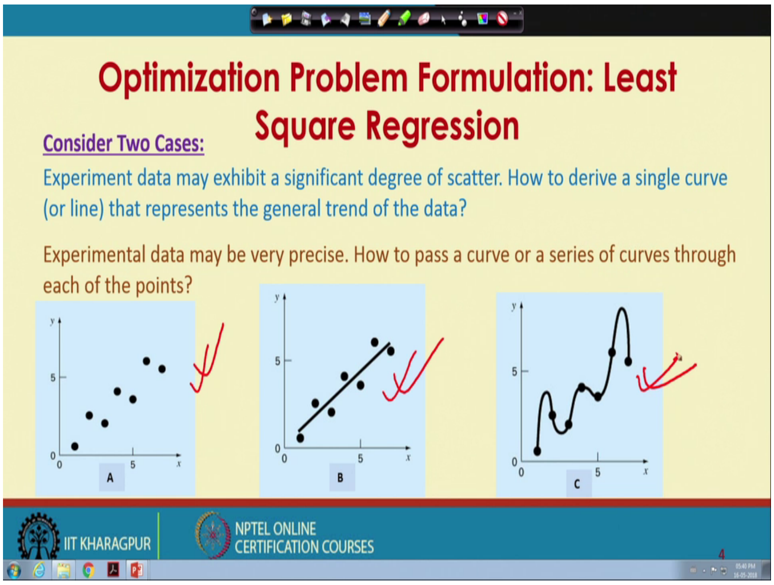
Welcome to lecture 9 of week 2. In this week we are talking about Optimization Problem Formulation. In the previous class we have seen the examples of optimization problems that were all linear in objective function as well as equations. So, they are all linear programming problems, we have also seen integer programming problem. So, in today's lecture we will first talk about one important topic known as linear regression.

(Refer Slide Time: 00:58)



So, this is a quick review of the general problem statement for an optimization problem, we have the objective function f x, you have a design vector X with n elements in it you have constraints 2 types equality consonant and inequality constants, see you want to minimise function f x by finding out suitable design vector x such that the inequality constant g x greater or equal to 0, satisfied set of equality constant h x equal to 0, satisfied and the bounce on decision variable X are also satisfied.

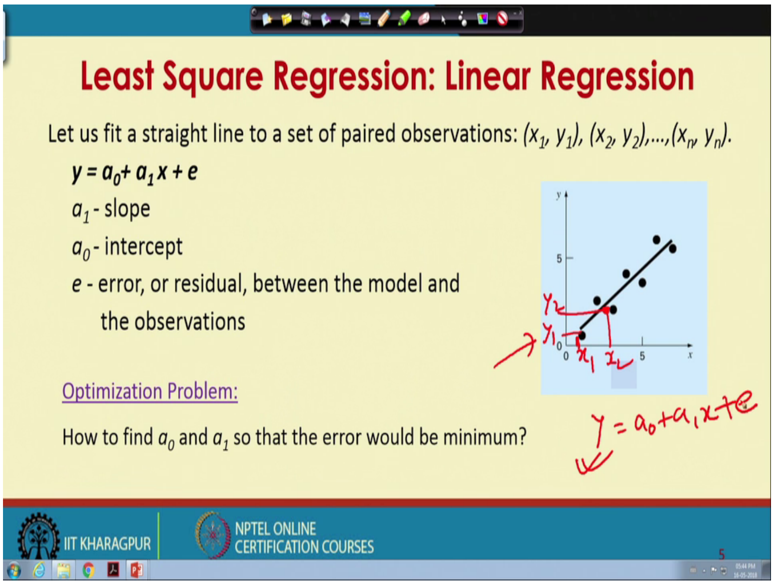
(Refer Slide Time: 01:53)



So, we will first talk about least square regression. This is an important topic and we will see how this least square regression is posed as an optimization problem, consider 2 cases you have done an experiment. Experimental data may exhibit a significant degree of scatter how do derive a single curve line that represent the general trend of the data. Consider another case where you think the experimental data is very precise, then how do you pass a curve or a series of curves through each of these experimental points.

So, you have let us say these experimental data Y versus X where you vary X measure Y. So, X is an independent variable Y is dependent variable. Now, if you consider that there is significant degree of scatter how do you fit let us say a straight line or best fit straight line that represents the general trend of the data as shown in this figure. Or if you think the data is very precise and you want to pass a curve through each of these data points, how do you do that? You will require least square regression for this. So, we will first learn what is least square regression? And our interest is how do you pose least square regression as an optimization problem.

(Refer Slide Time: 04:27)



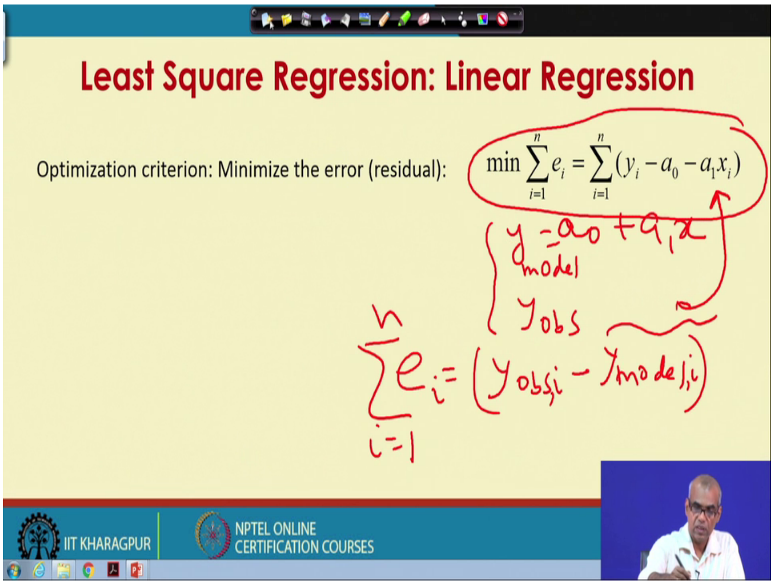
So, let us say we have n observations or more correctly let us say we have n paired observation like, X 1 Y 1, X 2 Y 2, X 3 Y 3, up to X n and Y n. So, when independent variable was X 1, the response or dependent variable was Y 1, when independent variable was X 2 the dependent variable was Y 2 so and so forth we have a set of n paired observations. We first want to fit a straight line through these data points, so an equation of a straight line is Y equal to m x plus c, where m is the slope you know and c is the intercept here we are writing as say y equal to y 0 plus a 1 x as the equation.

Now, if I want to represent this n paired observation by an equation like Y equal to a 0 plus a 1 x , the equation Y equal to a 0 plus a 1 x may not exactly represent the data points there may be some error. So, I say the observation Y is a 0 plus a 1 x plus error e. So, here a 1 equal to slope, a 0 equal to intercept and e equal to error or residual also known as residual between the model and the observation. So, the model is Y equal to a 0 plus a 1 x.

So, the data points that I have is let us say this is x 1 and this is y 1, this is may be x 2 and for this is y 2 so on and so forth. So, we want to feed Y equal to a 0 plus a 1 x, so if I put the value of these x, I can compute this y from this model that may not exactly match with this actual observation. So, there may be some error which is taken care of by this e here. So, this e is error or residual between the model and the observation.

So, the optimization problem is how do I find out a 0 and a 1 so, that the error would be minimum? If the equation for the straight line perfectly represents the observation the error e will be 0. That may not be possible in actual practice, so would like to have as low error as possible. So, we have to find out a 0 and a 1, such that the error e is minimum so that is the optimization problem formulation.

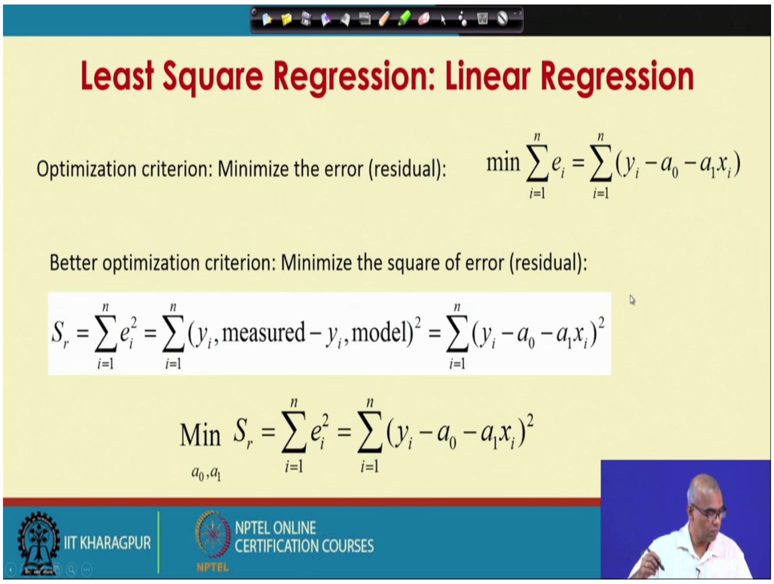
(Refer Slide Time: 08:05)



So, the criterion or the objective function can be written as you minimise error, which is y i minus a 0 minus a x note that y equal to a 0 plus a 1 x is my model. So, I have y model I also have y observations. So, the error is basically y observation minus y model. Now if you have n number of observations, so you can find out for each of these observations each of these n observations. So, note that this is what I have written here this y model is nothing, but a 0 plus a 1 x, so it is being written as y i which is the observation minus a 0 minus a 1 xi.

So, for each experimental point you have if you have n experimental point, you have n error values you have to sum up all these errors and you want to minimise that error sum. So, this can be an objective function or optimization criterion. Now here there is a problem the problem is the error sometimes can be positive it can be negative also. So, positive and negative errors either may cancel each other out or can represent a lower value of the error which is actually not.

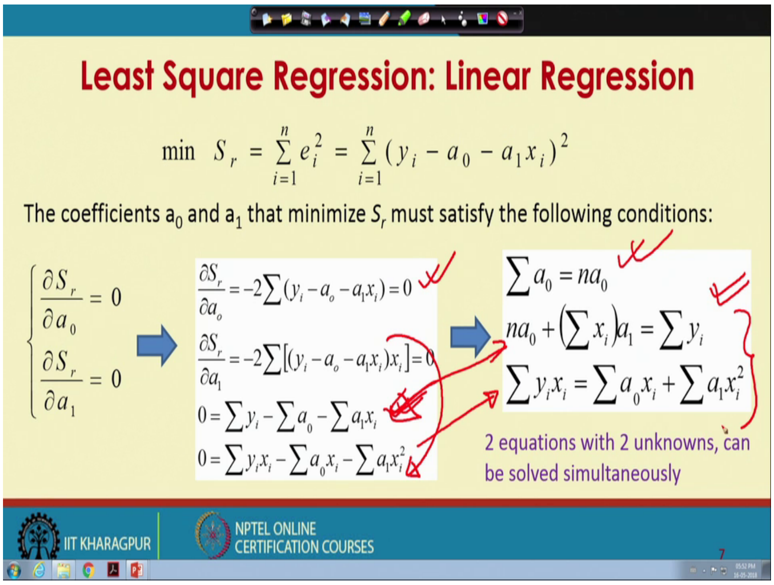
(Refer Slide Time: 10:33)



So, a better optimization criterion will be the square of error or the square of residual where instead of taking error as y observation or y measured minus y model, we take y measured minus y model whole square for each of these observations. So, we call this let us call this S r which is square residual square of residual or a square of error. So, better optimization criterion will be you minimise sigma is square summed over all I all I means all observations.

So, find out a 0 and a 1 such that sigma e i square, which is sum for all values of i equal to 1 to n, where n is the number of experimental data points is a better optimization criterion. So, S r which is sigma i equal to 1 to n e i square, which is represented by this expressions which is nothing, but y observation minus y model whole square is being considered as objective function.

(Refer Slide Time: 12:03)

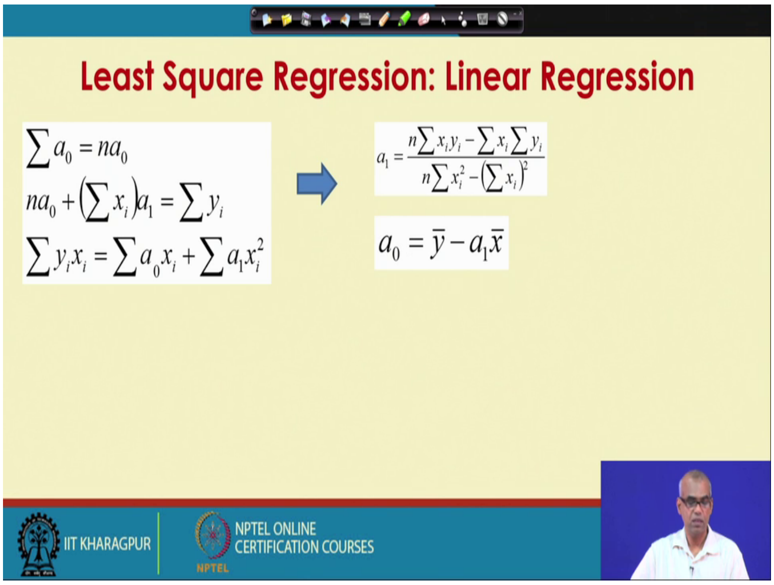


Now, the coefficients a 0 and a 1 that minimises S r or the sum of squared error can be found out by taking the derivative of this S r with respect to a 0 and a 1 and setting those values equal to 0. So, I have to minimise S r by finding out a 0 and a 1, I have an expression for S r which is given here, now the values of a 0 and a 1 that minimises S r must satisfy the derivative of S r with respect to a 0 equal to 0 and derivative of S r with respect to a 1 equal to 0.

If you take this derivatives and set that equal to 0 you will get 2 equations as shown note this is nothing, but this and this is nothing, but this. So, by taking derivative of S r with respect to a 0 and taking derivative of S r with respect to a 1 and setting those 2 derivatives equal to 0 we get 2 expressions as shown if you rearrange this you write like this you have 2 equations with 2 unknowns a 0 and a 1 and you can solve easily simultaneously these 2 equations.

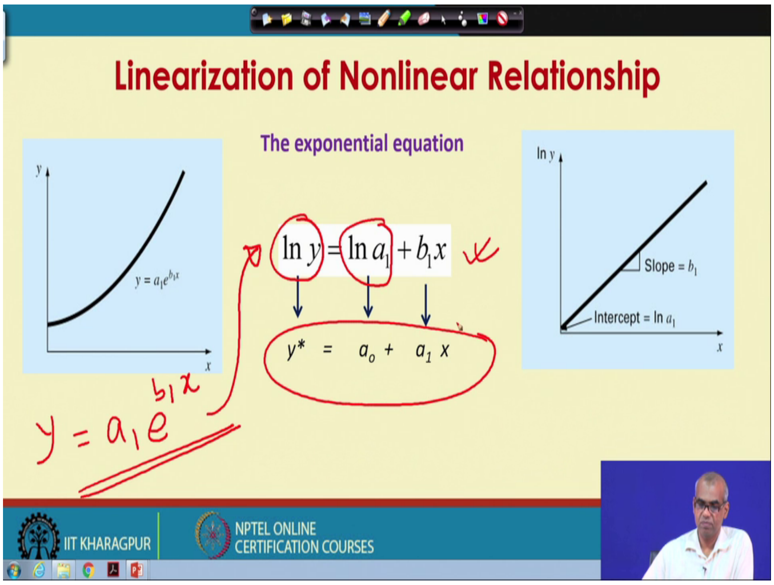
Look at this equation this is same as this equation sigma a 0 the sigma is the sum over all n value, so sigma a 0 is nothing, but n a 0. So, this equation can be written as n a 0 plus sigma x i in to ai equal to sigma y i, similarly this equation is nothing, but this equation which is arranged as this. So, we have these 2 equations and we have 2 unknowns a 0 and a 1. So, 2 equations 2 unknowns you can always solve simultaneously and find out the values of a 0 and a 1.

(Refer Slide Time: 16:07)



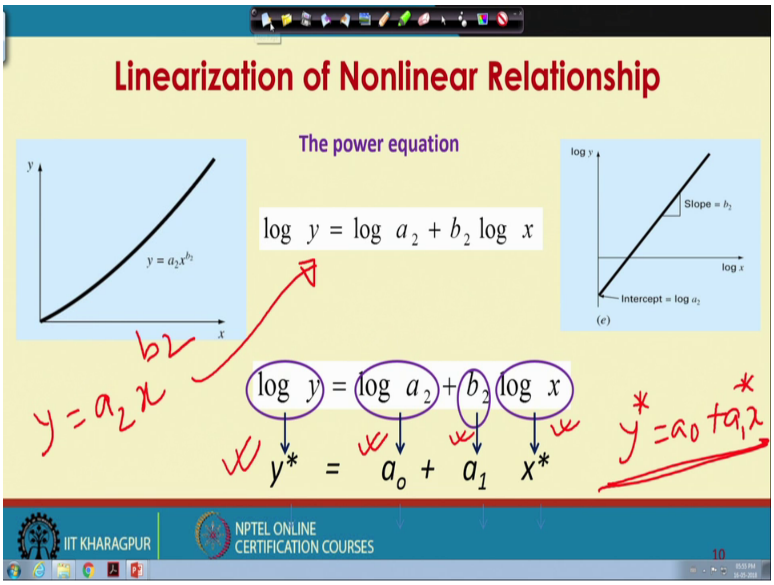
So, if you find out the value of a 0 and a 1 the expression that you will get as follows, so this is how we can find out the equation for the best fit straight line. Sometimes linearization of non-linear relationship is possible, let us consider the exponential equation y equal to a 1 into e to the power b 1 in to x. So, I have the relationship a 1 into e to the power b 1 in to x. So, the relationship between y and x is shown by this equation. So, this is a non-linear equation.

(Refer Slide Time: 16:29)



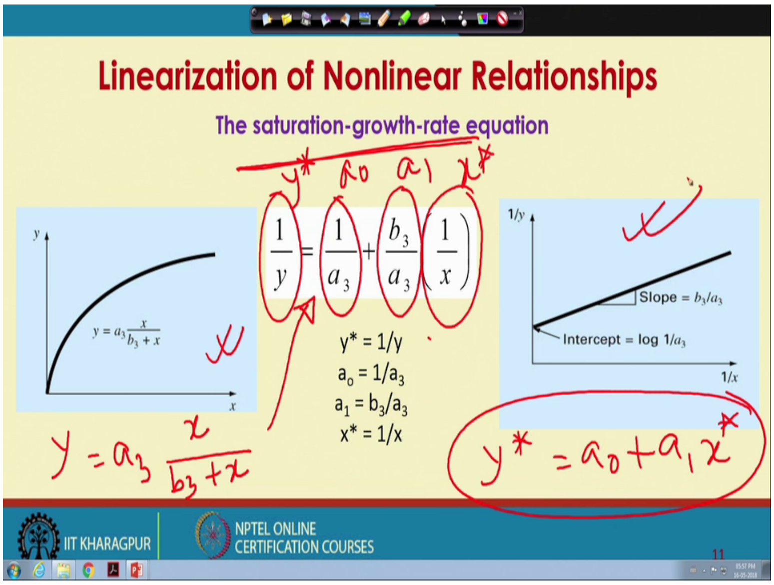
So, how do I apply the concepts of linear regression look if you take log of this equation you get this, log of y equal to log of a 1 plus b 1 into x. Now if I say ln y equal to y star, ln a 1 equal to a 0 and b 1 equal to a 1. This equation can be written as y star equal to a 0 plus a 1 x. So, now, whatever we have learned for the linear regression in last few slides can be readily applied.

(Refer Slide Time: 18:12)



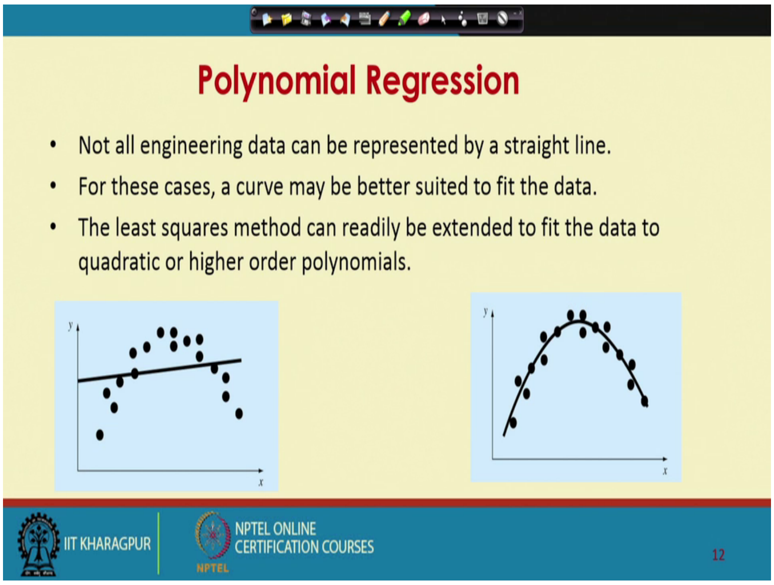
Similarly, if you have power equations such as y equal to a 2, x to the power b 2 can I apply the linear regression? Again yes if you take log you see you get log y equal to log a 2 plus b 2 into log x. Now if I consider log y as y star, log of a 2 as a 0, b 2 as a 1 and log x as x star, then I have y star equal to a 0 plus a 1 x star. So, this is the linear equation and whatever we learnt for linear regression can be readily applied.

(Refer Slide Time: 19:24)



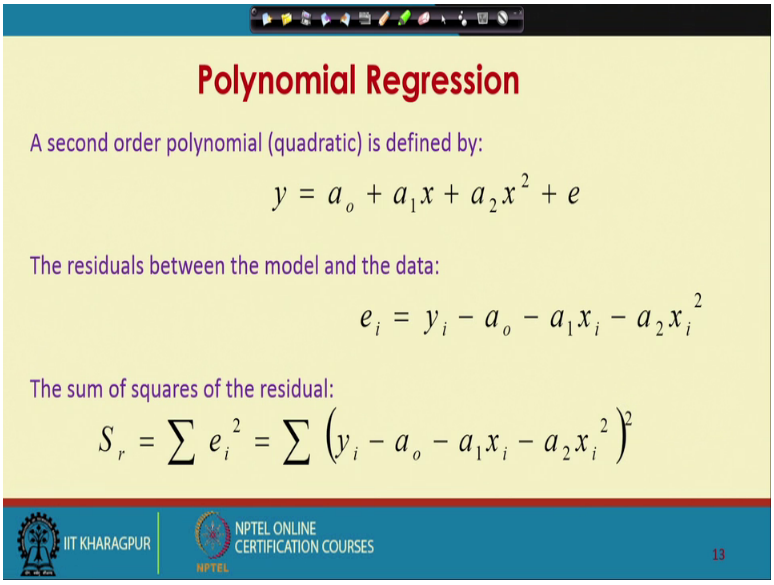
Similarly, the saturation growth rate equation such as y equal to a 3 into x by b 3 plus x can also be linearized as shown, the saturation growth rate expression can be written as 1 by y equal to 1 by a 3 plus b 3 by a 3 into 1 by x. If I consider 1 by y as y star 1 by a 3 as a 0 b 3 by a 3 as a 1 and 1 by x as x star, I have y star equal to a 0 plus a 1 into x star which is an equation of straight line again the linear regression can be applied readily. So, look how this non-linear curve after taking log after taking this transformations or rearrangement can be represented by a straight line.

(Refer Slide Time: 21:05)



Now, not all engineering data can be represented by a straight line, if you want to attempt to fit a straight line sometimes you will see that you will get a very poor fit because it may not be possible to represent those data by a straight line, for such cases may be a curve is the better option. So, a curve may be better suited to fit the data the least square method can readily be extended to fit the data to quadratic for higher order polynomials. For example if this is the kind of data you have between X and Y and if you want to fit a straight line you will get a very poor fit whereas, a quadratic is the better fit here for second order polynomial.

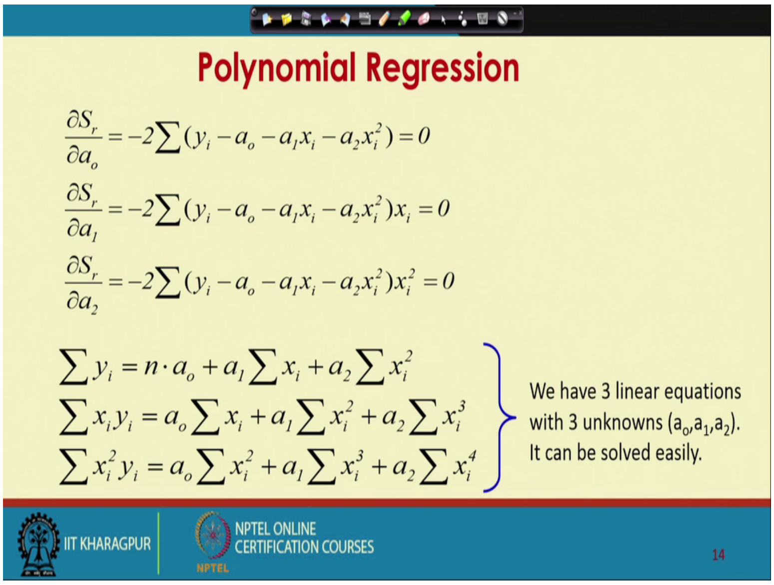
(Refer Slide Time: 22:14)



So, we will talk about polynomial regression; let us talk about second order polynomial or quadratic. A second order polynomial or quadratic is defined as y equal to a 0 plus a 1 x plus a 2 x square and if you want to model, the data given data by a quadratic or second order polynomial we will write like y equal to a 0 plus a 1 x plus a 2 x square plus that error or residual. Note that this term is now additional this is the quadratic term, if we drop this term it becomes a linear equation.

So, a second order polynomial for quadratic expression is due to the presence of a 2 x square term, the residuals between the model and the data e is observation i th observation minus i th prediction from the model. So, y i minus a 0 minus a 1 x i minus a 2 x i whole square, so that becomes the ith residual. So, again would like to minimise the sum of squares of residuals. So, a r equal to sigma e i square summed over all values of i i equal to 1 to n. So, this is the objective function which needs to be minimised.

(Refer Slide Time: 24:08)

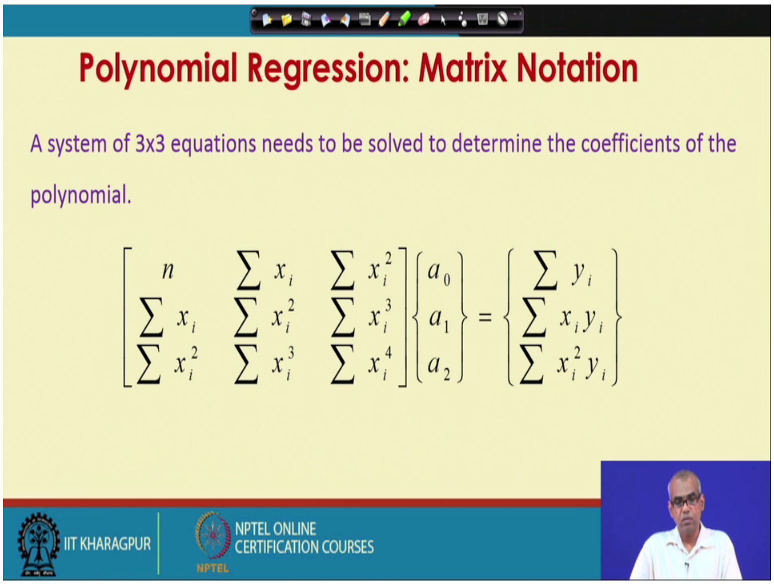


Again we can find out the value of a 0 a 1 and a 2, now we have 3 parameters to be found out a 0 a 1 and a 2 by taking the derivative of S r with respect to a 0 a 1 and a 2 and setting those derivatives equal to 0. So, that will lead to 3 equations so we will have 3 equations we will have 3 variables it is possible to solve simultaneously to get the values of a 0 a 1 and a 2.

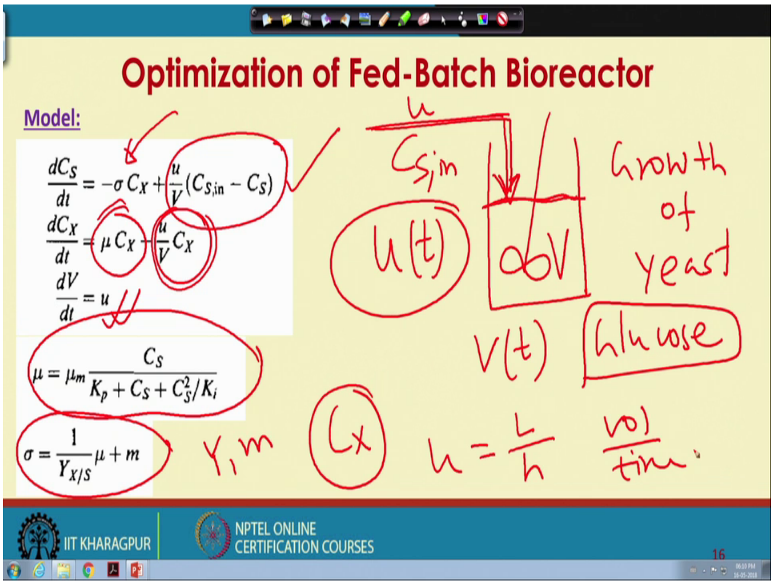
So, you take the derivative of S r with respect to a 0 you get this equation set that equal to 0 this is derivative of S r with respect to a 1, set that equal to 0 this is derivative of S r with respect to a 2 set that derivative equal to 0. If you do that you will get 3 equations are shown. So, you have 3 linear equations with 3 unknowns note here X is not unknown here the unknowns are a 0, a 1 and a 2, x was x are the known values for the independent variables. So, you have 3 linear equations with 3 unknowns a 0 a 1 and a 2 and it can be solved easily.

So, the extension of linear regression to polynomial regression is straightforward as we have shown for second order polynomial for higher order polynomial also you can very easily extend this only thing is number of equations will increase. So, for second order polynomial for quadratic we have seen that we have we are getting 3 equations and 3 unknown variables. So, it is possible to write those set of equations in a compact manner using matrix notations.

(Refer Slide Time: 26:06)



(Refer Slide Time: 26:44)



Now, after talking about linear regression and polynomial regression we will talk about now another optimization problem formulation and this is also a different class of problems we consider the optimization of a fed batch bioreactor. So, what is fed batch bioreactor the fed batch bioreactor is similar to semi batch bioreactor where you have a feed stream continuously comes into the bioreactor of fermenter, but there is no product that is being taken out. So, in a batch process there is no feed in there is no feed out all the reactance are dumped in the beginning in a continuous process there is feed in there is feed out in a semi batch process there is only feed in, but there is no feed out.

So, a fed batch bioreactor or fed batch fermenter is like semi batch operation it is a semi batch operation, consider a fed batch bioreactor where we are growing a microorganism let us see we are growing yeast. So, yeast is a microorganism with commercial interest it produces ethanol, now we have taken substrates now yeast will grow on substrate let us say substrate is glucose. So, you start with some initial amount of substrate that is glucose and also some amount of microorganism yeast then you feed the glucose stream.

Let us say with feed rate u and the concentration of the glucose of substrate in this feed stream is let us say C s in the volume of the fed batch bioreactor is V note that V is the function of time and V is changing with time because of the feed in there is of course, a maximum volume restriction of the fed batch bioreactor let us consider that, so what are the processes that are taking place the microorganisms will consume the glucose and will multiply; that means, there will be growth of microorganisms.

So, that is one phenomenon that is taking place growth of microorganisms, another phenomena that is taking place is the consumption of substrate by the microorganism. So, there will be reduction of substrate if you do not feed in there will be reduction, but if you feed in you can regulate the concentration of the substrate or the concentration of the glucose in the bioreactor.

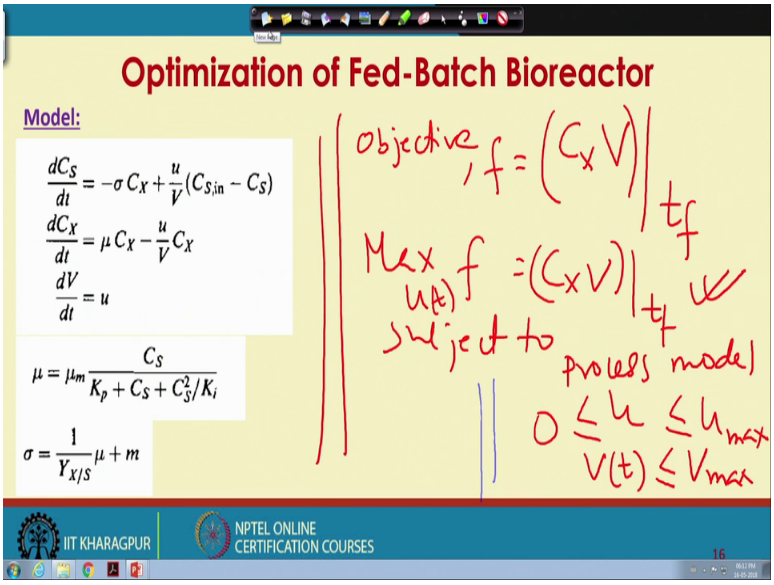
Now, this microorganism may grow at a particular substrate concentration at a highest rate, let us say the growth rate will be maximum at a particular substrate concentration. So, we would like to regulate u we would like to change the feed rate with time such that the substrate concentration in the fed batch bioreactor is maintained at a value, which maximises the growth rate of the bioreactor growth rate of the microorganisms.

The third phenomena is taking place is the volume is changing due to feeding the substrate stream. So, to model this system I have to write down 3 must balance equation. The first balance equation represents the change of concentration of the substrate let us say glucose with respect to time substrate is consumed by the bioreactor. So, minus sigma into C x, sigma is the substrate consumption kinetics which is given by nu by Y, Y is the yield parameter and m is the maintenance parameter.

So, Y m are known parameters and mu represents the growth rate expressions, so they are like kinetic expressions. Now, feed stream due to the presence of the feed stream the substrate concentration can be increased substrate concentration is increased substrate or the glucose is coming into the bioreactor. So, that is being taken care of by this, now C x let us say is the concentration of the microorganisms. So, there is growth of the microorganisms and because of the dilution there will be decrease in the concentration because of the growth there is increased and because of the dilution there is decreased.

So, rate of change of concentration of the substrate is mu C x minus mu my u by v into C x and the volume is changing due to feed in. So, dV dt equal to the feed rate u, where u is volumetric feed rate. So, it is like say liter per hour or any volume per time unit so, d V dt equal to u rate of change of volume is equal to feed rate. Now, what is the objective objective may my objective function may be I want to maximise the growth of the microorganisms.

(Refer Slide Time: 34:03)



So, I want to get the maximum quantity of yeast at the after the reaction for after the fed batch fermentation process is over. So, C x is the concentration of the microorganisms and v is the volume of the fed batch bioreactor. So, C x times v gives me the amount of microorganisms. So, this is the quantity that I want to maximise at the end of operation which is t f represents the time at which the fermentation process is completed. So, objective function let us say f equal to C x v evaluated at final time t f.

So, I want to maximise f which is C x into V at tf subject to the process model and there is restriction on the feed rate u which is cannot be less than 0 can be 0 cannot be less than 0 and cannot be more than ones specified value may be depends on the maximum capacity of the pump. Similarly your V at any time must be less than the maximum volume of the fed batch bioreactor so that there is no overflow so this is the problem formulation for a fed batch bioreactor.

So, here I want to find out u the feed rate which is a function of time which maximises this objective function and the solution must satisfy these constants. So, with this I want to stop the discussion of fed batch bioreactor and we want to stop this lecture here