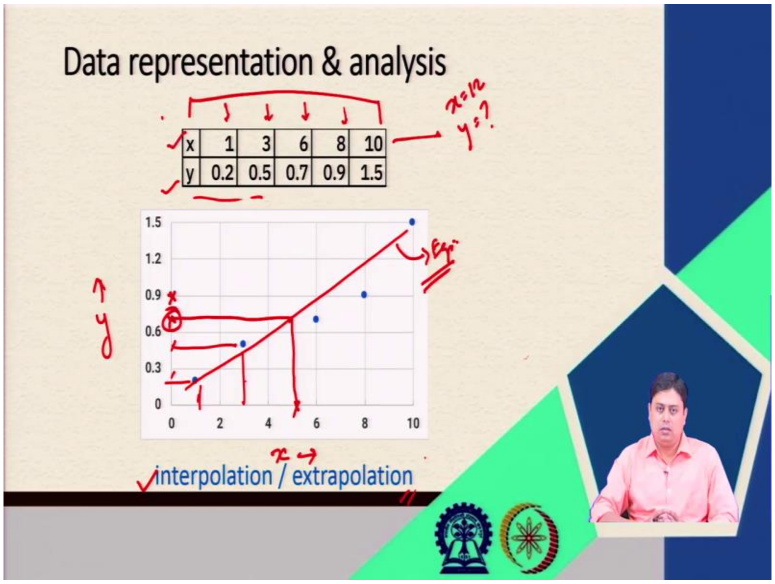
Material and Energy Balance Computations
Prof. ARNAB ATTA
Department of Chemical Engineering
Indian Institute of Technology, Kharagpur

Lecture –03
Introduction to Engineering Calculations (Contd.,)

Hello everyone, welcome to the another class of Material and Energy Balance Computations. We are in module one, that is the introduction to engineering calculations processes and process variables. So we are in the third lecture, in the last 2 lectures we discussed about units, dimensions, significant figures, dimensionless numbers or say dimensionless quantity and the requirement of dimensional homogeneity in a valid equation.

And today we will see data representation or, say, the data estimation for a certain type of calculation.

(Refer Slide Time: 01:04)



Say, for example, when we are measuring something or we are calculating estimating by experiment some values of some parameters, for that case here we have an independent variable which is x, say based on x we have estimated certain values of y. Now we typically tabulate those data like that is written here now the point is that here we have taken or estimated or measured those values of y for a certain condition of x or a certain point in x.

Therefore, this is not an analytical solution or the analytical representation of the data. Now analytical representation of the analytical equation means when you have the continuous value of dependent parameter based on the independent value. Now, in this case, we have instead of continuous values we have some discrete values at a certain location. If we plot it in a rectangular plot, we see that the data scatters like this kind of pattern that is shown here.

So all the 5 data points are plotted here, this is the x coordinate, and this is the y. So for each and every value of x, we have a certain value of y. Now we are in a problem statement such kind of experiment was done, and certain values were tabulated for some material properties. At a certain temperature, the viscosity of this liquid has been reported like the value has been reported in y. But your problem statement involves a different value of a temperature for which the measurement has not been done.

So, in that case, what do we do or how do we estimate the liquid viscosity or density or say whatever the dependent parameter based on the temperature that we have understood in this problem. How do you estimate or calculate those values? Now, this data representation and analysis would guide us in that way. So the easiest way is that look at the data scatter like in a rectangular plot. We plot those values.

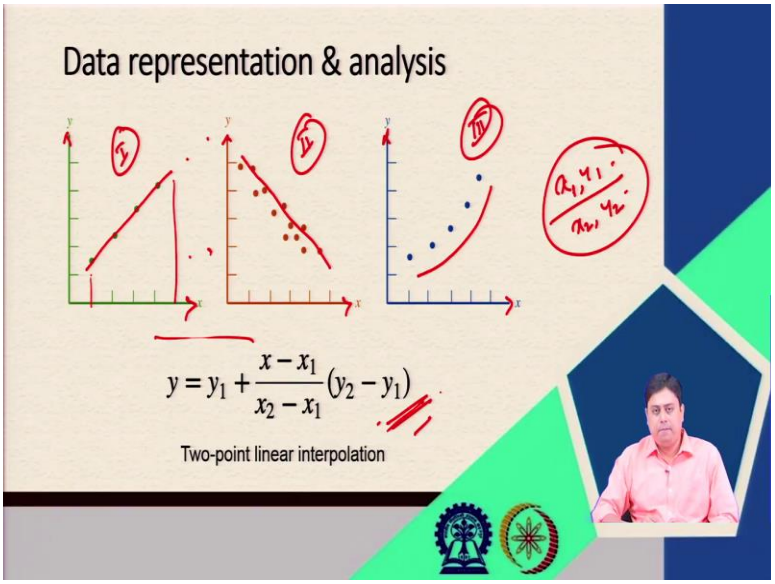
We see the trend, how it is moving, and now, if we see that scatter, we can connect all the data points with a reasonably fit straight line. Then we can always write the equation of this straight line, and we consider that all the data points fall nearer to this straight line with some error but to have and reasonable estimate, we use the equation of that straight line to have the value of unknown y for a known x.

So, in this case, we have now an unknown value that we have to calculate for say x = 5, then on that straight line, we can find the value of the dependent parameter. If we know the equation of this statement and there that we can derive from the fitting of these data points, the best fit line. So, that means if we have an intermediate point in this range of x say from 1 to 10. We have measured the value of y at 1, 3, 6, 8 and 10, but what about the 0.5 that I should the example.

Or say what about x = 7, what will be the value of y? Those requirements are classified under interpolation. That means we have to interpolate these data range in order to find a particular value of x with which is within the range that is mentioned here. It is sometimes also may be required that we calculate the value of y for a particular value of x, which is beyond this data range. Say for x is equals to 12, what would be my y?

That is called the extrapolation of the data range. So this interpolation or extrapolation of these processes is done by this data analysis and its proper way of representation. Now, as I said that one of the easiest ways is to represent this scatter by a straight line because it is easier to write the generic equation of a straight line which is y = mx + c, where m is the slope of the line and c is the intercept.

(Refer Slide Time: 07:45)



So say we have this kind of scatter plot or data points somehow we have achieved. If I mention that this is case 1, this is case 2 and this is case 3. So for case 1, we see that with the variation of x, the variation of y is quite linear, and we can easily fit or we can easily take all the points in a straight line. In case 2 what we have, we have very discrete points at very close range, within a close range, we have very discrete points and that was scattered.

In this case, also we can possibly draw a best fit straight line which falls in between all the points, or most of the points can be joined by this straight line. But in the third case, we see that there is curvature exist, the trend is not a straightforward linear, but it has a nonlinear trend. So, which means in these 2 cases, we can have the linear fitting of the discrete points, and in this case, we have nonlinear data fit.

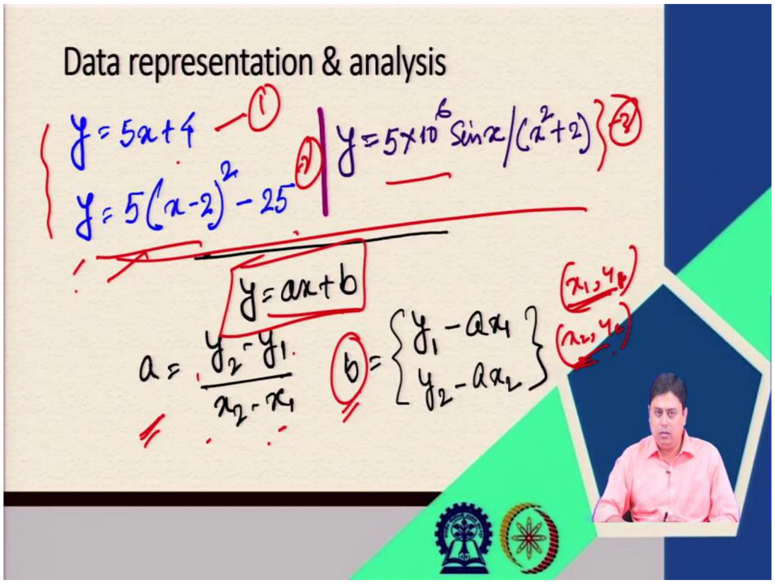
There are several methods that exist, which we will not go into the details of this, this comes under the purview of little mathematics for this nonlinear fitting and linear fitting, but a couple of easiest solutions I will tell you today. So, one of the easiest ways is the two-point linear interpolation. We try to fit any set of data with a straight line to have an estimate. If the trend is linear, that is estimate would have better accuracy and precision.

Now in the case of two-point linear interpolation, this equation is known to us? It is basically joining the straight line with two x and y values. For example, (x1, y1) and (x2, y2), with these 2 points, we draw a straight line, and that would have an expression or the equation of this one. i.e., we are taking 2 points, we are drawing a straight line, and then within the range of these data points, we can find any values of the dependent one based on the independent parameter.

So this is the interpolation, now the same equation we can use it for extrapolation, but there if the data points are not in a linear fashion and we do not know what would happen outside that range, that is why the use of this equation for extrapolation should be done with very caution. You have to be very cautious while using this equation to use it for extrapolation because, with confidence, you know that within the range, the trend is linear.

But beyond this range, what is the trend of the data points should be? We have no idea until and unless we have the experimental measurement. But still, if you use this one and fortunately property comes to be in the linear fashion after this range as well or outside the range as well, then the accuracy of estimation of the prediction would be much better.

(Refer Slide Time: 12:43)

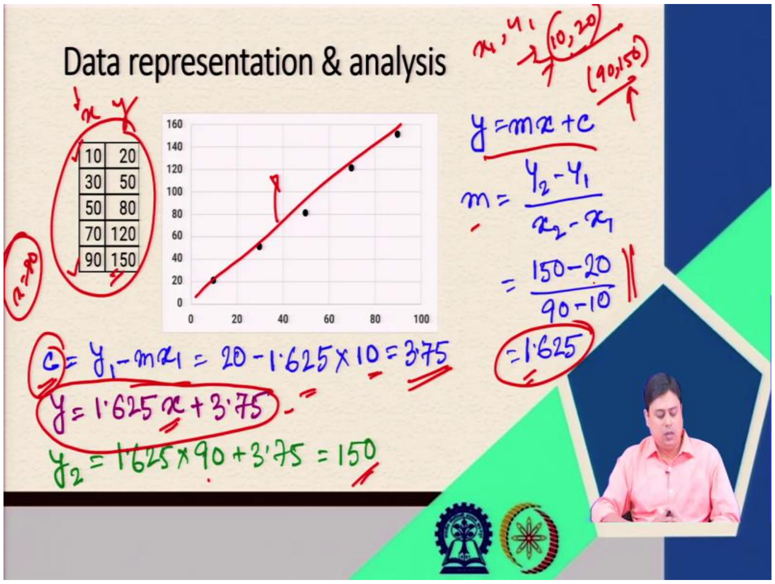


So the easiest one that I said is a straight line fit. Say, for example, we can have this kind of a analytical solution for a certain problem. So there are three equations I have shown here, 2 and 3. So in these 3 questions, there are different variations. The first one is linear, the second one and the third one are the nonlinear ones. So, if these are the analytical solution that means you have a concrete equation for the relation.

Then you can use it to estimate the value of y for any given x, or from the data fit, we have to derive such expression which fits that nonlinear trend by this type of expression. There can be several different types of nonlinear expressions. Now they say the straight-line fit is the easier or the easiest among all the process. So the straight line that I said will have a generic expression of $y = ax + b$ where a is the slope, and then once you know the slope based on the two given data points for (x1, y1) and (x2, y2).

We can calculate the value of intercept b based on any one of these sets, either (x1, y1) or (x2, y2). So, once we use any one of these data set or data point with the combination of 'a' in this expression, we can easily find the value of b and then that we also can verify by the substitution of another data set to cross-check whether the calculations are right or not. Then we have the expression, and we try to interpolate for whatever the value of the independent parameter to know or to estimate the value of the dependent quantity.

(Refer Slide Time: 15:19)



So, for example, if I try to elaborate this further that somehow we have got this data point, where this is x, this is y. We have measured some parameter y based on independent variable x and this I have plotted here. How do you estimate the value of y at x = 80. So we see that if I try to fit this clearly, we can see it is a visibly a straight line fit would work for us.

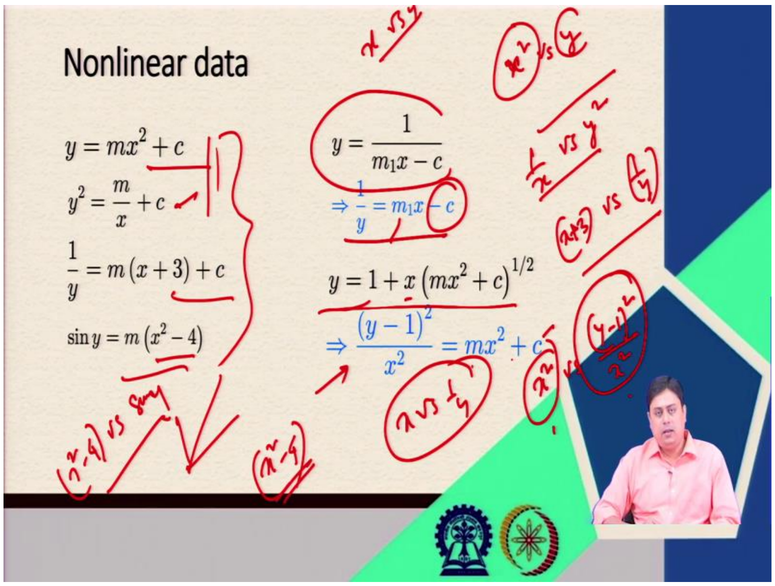
So we try to write the generic expression of a straight line, which is y = mx + c, m is the slope, (y2 – y1) / (x2 – x1). We take the 2 endpoints i.e., (x1, y1) is our 10 and 20 and (x2, y2) is 90 and 150. We take these two data points, and we find the slope of this straight line. Once we have it, with one of the data sets or one point that is 10 and 24, which we have used this and the slope that we are calculated m, we find the value of c, like this.

So we substitute y1 and x1, y1 is 20 for x1 is 10 and the slope and we get the value of c as 3.75. So, which means the line connecting or say the best fitting all these data points will have an expression. In fact, let us not talk about the best fitting here with the two-point linear interpolation. The straight-line equation here is that y = 1.625x + 3.75, which means we have an equation or expression for this line.

To, check whether the expression is fine or not, whether our calculation stands or not. We have used while calculating the value of c the first data point that is 10, 20. We have not used 90 and 150. So we use that in this expression, that we now replace here instead of x we use 90 to see what is the value of y and it correctly shows that if you replace here x value as 90 the value of y becomes 150 which is already there, that means and it should be if the calculations are right because we use those 2 values while calculating the slope of this line.

This is how we cross-check, our result as well at the same time we write the expression. A quick step will tell us whether these calculations are right or wrong.

(Refer Slide Time: 19:18)



And then comes the nonlinear data. So here I have written couple of expression which is clearly not linear, variation of x and y. So, all these expressions if we try to set it the way that I mentioned earlier that we try to forcefully fit it into a straight line, what would be the case? How do you do that? The thing that we can do, say for the first expression here $y = mx^2 + c$. If we plot x^2 versus y, then eventually it would give us a straight line with the slope m and intercept y.

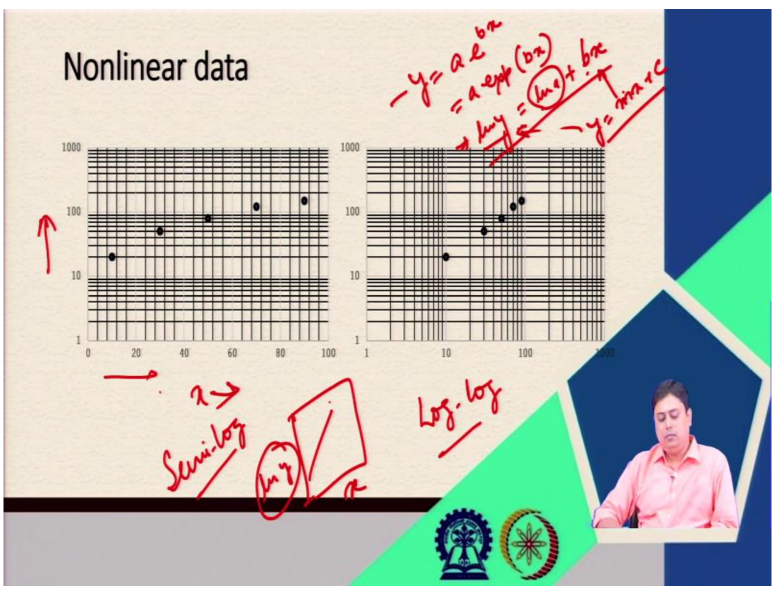
For the second equation similarly, if we plot 1/x versus y^2 it would be again a straight line with slope m and intercept c, but all these are the nonlinear expressions. But our data representation is such a way that we are making it as a straight line. Similarly in this expression 1/y = m (x + 3) + c. So, in this case, our selection of x would be that instead of x versus simple y plot in a straight line, a linear curve would have been a straight line.

In this case instead of x, this would be x + 3 versus 1/y, this plot would give us the straight line of the linear fit. Now, in this case, the fourth one, the whole sin y would be plotted as y are in the y coordinate y-axis and this $x^2 - 4$ is our x-axis. So it would go through the (0, 0) because it is of y = mx, it would pass through the origin. The plot of $x^2 - 4$ v/s sin y. Sometimes we have to rearrange the expression.

In order to force it to fit into this straight line, for example, this one $y = 1 / (m_1x - c)$. In this case, we rearrange the expression like this, so that now we can clearly see that if I plot x versus 1 / y, it would give a slope of m_1 with the intercept of – c. A more complicated expression is this one that $y = 1 + x\,(mx^2 + c)^{1/2}$. This also can be rearranged in this way.

That y – 1 we take the 1 in the left inside y – 1 then, we make the square on both the sides with the square and then we divide by the x component here so that it becomes now $mx^2 + c$. So in this case, what do we have to plot? We have to plot x^2 versus $(y - 1)^2 / x^2$. This is the whole y and this is the whole x. If we plot these 2 on the x-axis and y-axis, will have a straight line with slope m and intercept y.

(Refer Slide Time: 23:56)



The other thing is that instead of now all the plot that we are mentioned here is on a rectangular plot. But this nonlinear data if we plot it; so here the situation is we have this kind of complicated parameter on the x-axis and then we say this is also a much-complicated parameter on the y axis. Fine, now instead of that what we can do that instead of that for the nonlinear curve for the other popular method is to plot the values in the logarithmic scale.

This is the representation of certain scattered data to show you 2 different logarithmic scales or 2 different log plots. On the left-hand side, if you look at it here, the x-axis is in the rectangular scale, but the y-axis we have in the log scale, the values and this major and the minor axis are shown in order to show this gradual increment. In the x-axis, we have an equal interval. Here, we have 1, 10, 100, 1000. So, such plot where we have one axis as the logarithm scale and the other axis as normal scale, the rectangular scale we call this as a semi-log plot.

But when we have both the axes in logarithmic scale, we call that that has log-log plot. So the utility of this is that say we have an expression $y = a^{bx}$. If you take the natural log on the both sides so we have ln y = ln a + bx. So now we see that we can represent this as mx + c. If you try to have a similarity between these 2 expression where our y is ln y, m is the b and c is ln a.

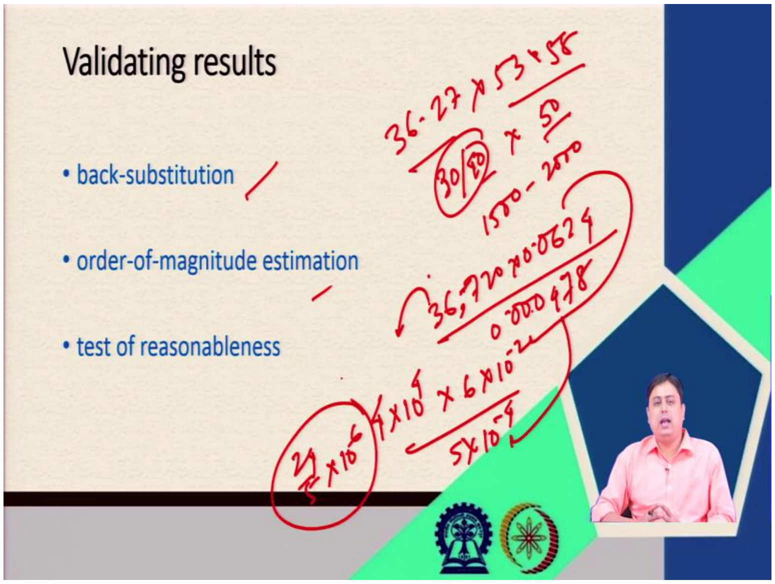
Which means, if we plot this expression in a rectangular plot where the y-axis is ln y and x is the normal scale then we can have a straight line that will have an intercept of ln a and the slope of b, that means in a rectangular plot we have to plot in this case which is x and ln y, and then this expression can give us a straight line representation. But remember this is the nonlinear form and we are trying to enforce it to fit in a straight line manner for over calculation to be easier.

Now in instead of calculating this ln value or the logarithmic of y, the utility of this log-log on the semi-log scale now can be used here because now we see that we have to plot ln y. So, in a logarithmic scale where the y-axis is already in the logarithmic scale, if you use this graph where we have the x in the rectangular plot and here y in the logarithmic scale if you directly plot this value here, it would give us a straight line fit.

So, we need not calculate further the value of ln y, if we have a semi-log graph paper to plot this. Similarly, if you say have $y = ax^b$, if we take log on both the sides we have ln a + b ln x. So, in the rectangular plot, we have to plot ln y versus ln x in order to have a straight line. But if we plot the values directly y and x in a log-log plot, this should give me a straight line fit without calculating the values of the natural logarithm of y and x or any further step.

These are the utility of logarithmic log-log graph paper on the semi-log graph paper where, if you have such kind of format nonlinear expression we can quickly plot without calculating the logarithmic of the x or the y of the independent and dependent parameter.

(Refer Slide Time: 30:38)



Now I will wrap up this lecture by telling you that how to validate our results. How to have confidence on our calculations? But we have done so many things that we have understood significant figures, we have understood this units dimensions, we are doing calculations and then we have interpolated extrapolated now we have for the calculated all the parts. So, how do we now estimate?

How do we ensure that our results are correct? There can be 3 tests. Those are kind of logical ones one is the back-substitution. This we have been doing for a while, once we calculate some value or we have got the result, we again back-substitute that into the equation to check whether the left-hand side is equal to the right-hand side or not. We check our result to have a gross estimate by the order of magnitude estimation.

That when we have certain calculations like, this multiplied by 53.58 like this what the gross value we can quickly calculate by estimating this either by 30 or 40 whichever is easier for the calculation multiplied by 50 say. Quickly we do in this case that the valued be in the range of

1500 to 2000, this kind of range we can have the value. This is called the order of magnitude estimation. So for example say we have further more complicated expression.

The more complicated expression could have been that we have $(36.720 \times 0.0624) / 0.000478$, now the order of magnitude estimation tells us that we consider this parameter with the scientific notation say 4×10^4, say instead of 36 point we say 36000. So in that case it would be 4×10^4, multiplied by this parameter. We can estimate say, by 6×10^{-2} and this parameter we estimate as 5×10^{-4}.

So that our calculations become easier, $(24 / 5) \times 10^6$. so the result would be in the order of this value and if you do it by the calculator, you see that it will definitely give you the reasonable estimate if we go by this order of magnitude estimation. And the final point is the test of reasonableness. Whatever the value you have calculated for that problem statement it should have some logical sense.

For example, if you are calculating a reactor diameter for a commercial plant and the reactor diameter comes out to be the diameter that is more than the Earth diameter. Then that means something is wrong. You are calculating, say the velocity of water flowing through a pipeline, and that velocity comes out to be the velocity that is way higher than it could have been. So, the mass of a certain component you're calculated is more than the mass of a Sun.

The velocity of water that I mentioned earlier if the velocity is higher than the velocity of light, the speed of light means there is something wrong. That is called the test of reasonableness. We do these 3 things after the calculations and we then we go back with confidence that the calculations probably are right. The calculations can be right but then the problem, the way you have handed that you will see slowly in the next couple of classes.

So, with this, I thank you for your attention and will see you with the next couple of lectures on the process variables and what are process and all these definitions. Thank you.