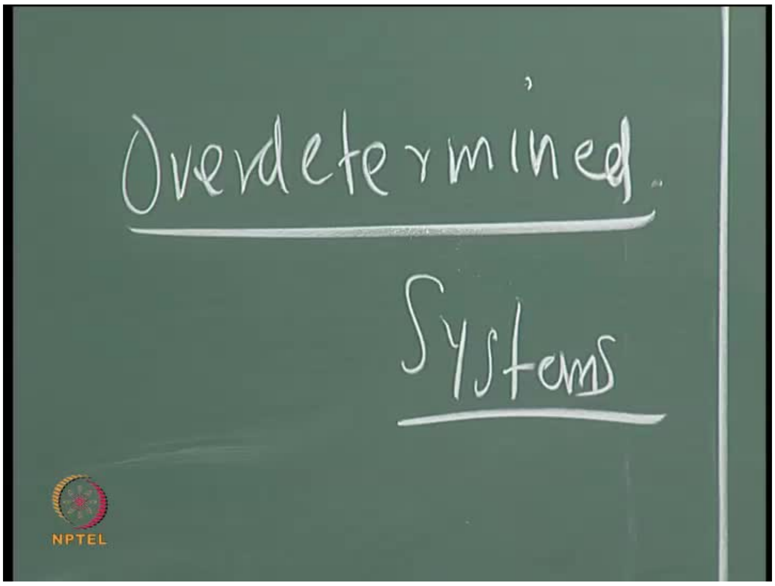
Design and Optimization of Energy Systems
Prof. C. Balaji
Department of Mechanical Engineering
Indian Institute of Technology, Madras

Module No.# 01
Lecture No. # 16

Best Fit

We will continue the discussion on curve fitting and regression. I told you that often times we get only data points which are discrete; for example, the performance of equipment, turbo machinery equipment or something like that, but we are interested in getting a functional relationship between the independent variable and depending variable so that we are able to do system simulation, and eventually we are able to do optimization. But if it is a calibration or something, we want to have an exact fit where we want the curve to pass through all the points. Then, we discussed various strategies for exact fit, like Newton divided difference polynomial, Lagrange interpolation polynomial; we also had a problem on the Lagrange interpolation polynomial in the quiz, very simple and straight forward problem, where I gave density; rather the reciprocal of the density; so, specific volume of stream and so on.

(Refer Slide Time: 01:40)



Overdetermined

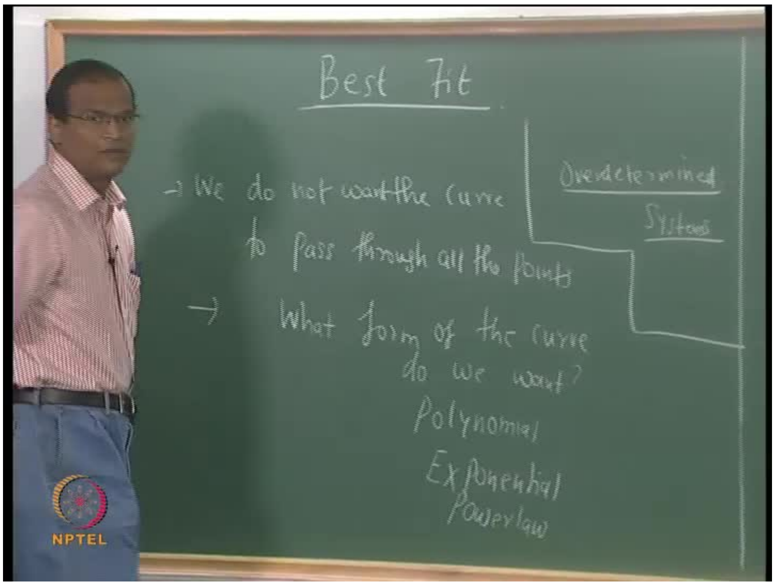
Systems

These are basically exact fits where you have only few points, but the measurements are very accurate; so, you want an exact fit. But there are several cases where the measurements are error prone. Point number 1 - there are too many points, there are too many points and it is very unwise, it is very unwise on your part to come up with polynomial which is 9th degree or 12th degree or 14 degree which is keep on oscillating.

So, these are basically called as over determined systems. What is an over determined system? Over determined system is a system which has far too many data points compared to the number of equations required to regress a particular form of the equation. For example, if you know that the relationship between, if you know that the relationship between enthalpy and temperature is linear for example, let us not worry about the pressure, if you know the relationship between the enthalpy and temperature is linear, you can put h equal to a plus b into temperature. So, if you have two values of temperature, for these two values of temperature, if you have the enthalpy, you are home; you get both a and b; but we have 25 or 30 values of temperature; for this 25 or 30 values, we have the enthalpy, all of which are associated with some error or the other.

Now, this is an over determined system because any two, any two, any pair of temperature and enthalpy will give you a value of a and b. Among these values of a and b, which is the most desirable? We do not know. Therefore, you have come out with the strategy of how to handle a over determined system where the number of data points is too many. I mean basically you cannot you just do not want to solve the system of simultaneous equation and reduce it to trivial case of take two pairs and get a and b because the whole thing is error prone. So, you want to, but you want to take care of all the points and so on. Therefore you have to evolve strategies for best fit.

(Refer Slide Time: 03:27)



Best Fit

→ We do not want the curve to pass through all the points

→ What form of the curve do we want?

Polynomial

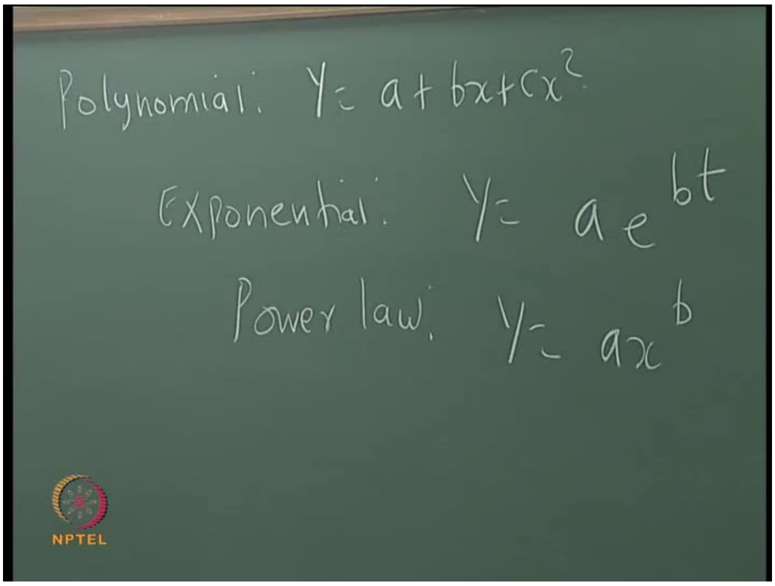
Exponential

Powerlaw

Overdetermined System

So, best fit: We do not want the curve to pass through all the points; fine, that is a reasonable, reasonably good description of what we want to do. So, what form? Polynomial, exponential, power law, or any other polynomial you know.

(Refer Slide Time: 04:41)



Polynomial: $Y = a + bx + cx^2$

Exponential: $Y = a e^{bt}$

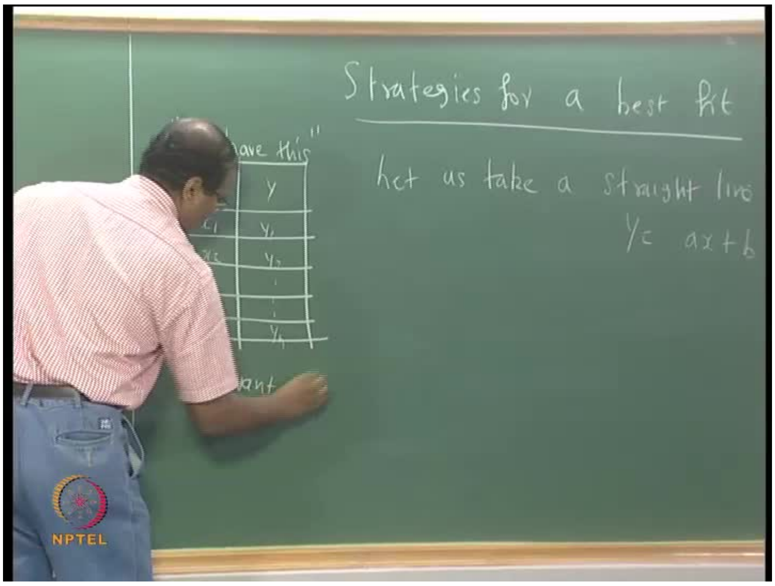
Power law: $Y = ax^b$

Write the polynomial form; could be Y is equal to a plus bx plus cx square; it is general depiction; you can have higher order polynomial also, or you can have an exponential form Y equal to a e power to the bt which is typically what happens. If you have an initial value problem, concentration decreasing in time or the population is changing with time. We modeled for assignment one, a problem, where limited resources and unlimited resources and so on, or you can have a power law form where Nusselt number or Nusselt number or screen friction coefficient is equal to a into Reynolds number to the power of b, and so on. So, there could be various forms like this.

But how, who will tell, who will tell, who will tell us what is best way to get a, b, c for the polynomial or a and b for the exponential, or a and b for the power law? We have chosen a form, alright; what is the strategy you would employ or use to get this a and b because it is a over determined system. We have only two, we need to get only two constants a and b, but there are [FL] data points and they are not highly accurate. Minimize the got error; why spare of the error? Though (( )) some mathematical form okay; what is so great about it?

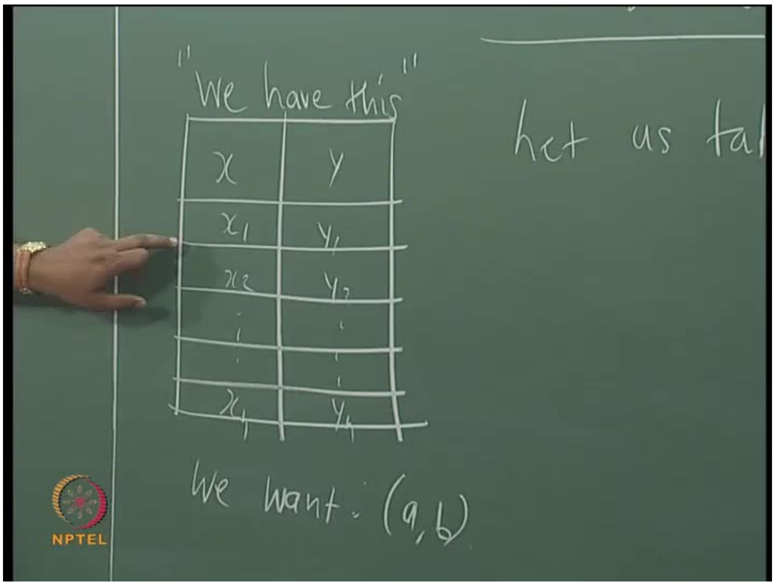
When I am provoking you; that does not mean that what you are saying is wrong or something. Are you getting the point? We have to propose some strategy now. We just cannot solve simultaneous equation and go home because we have a over determined system. So, we have to discuss what are the strategies for a best fit?

(Refer Slide Time: 07:05)



So, for an exposition, for an exposition on a strategy for a best fit, let me take a straight line Y is equal to ax plus b.

(Refer Slide Time: 08:26)



"We have this"

x	y
x_1	y_1
x_2	y_2
$\vdots$	$\vdots$
$\vdots$	$\vdots$
x_n	y_n

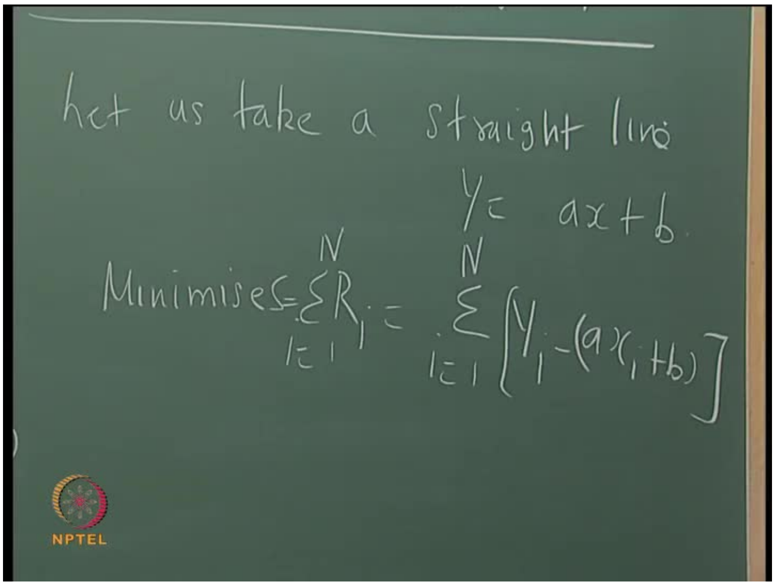
We want : (a, b)

Let us ta

What is the problem we have? Please remember, we have this; we want (a, b). How do you get this? You will either do a fluent computation or you write your own program, or you do your own experiment in the laboratory. What could be this x and y? x could be Reynolds number; y could be screen friction coefficient; that x could be anything; x could be temperature; y could be enthalpy. We are looking at a simple relationship; I mean we are looking at a simple one variable problem. Now, for this, we need to a,b need to determine a, b; it is over determent system, for the benefit of the people would came late, it is a over determined system because we have n number of points where n far exceeds a, b.

If you have just two points, you can simultaneously solve for a and b, and say that Y is equal to ax plus b, but that is not going to solve the problem. Because it will locally fit a line between two points, but we want to find the best line which is the good representation for the complete data set because eventually with this line, we are going to do system simulation and optimization and so on.

(Refer Slide Time: 09:53)



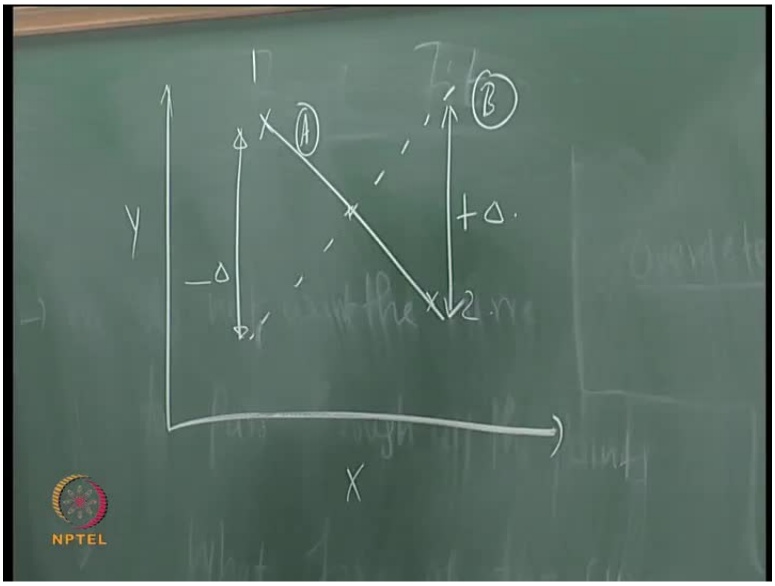
Let us take a straight line

$$y = ax + b.$$

$$\text{Minimise } S = \sum_{i=1}^{N} R_i = \sum_{i=1}^{N} \left[y_i - (a x_i + b) \right]$$

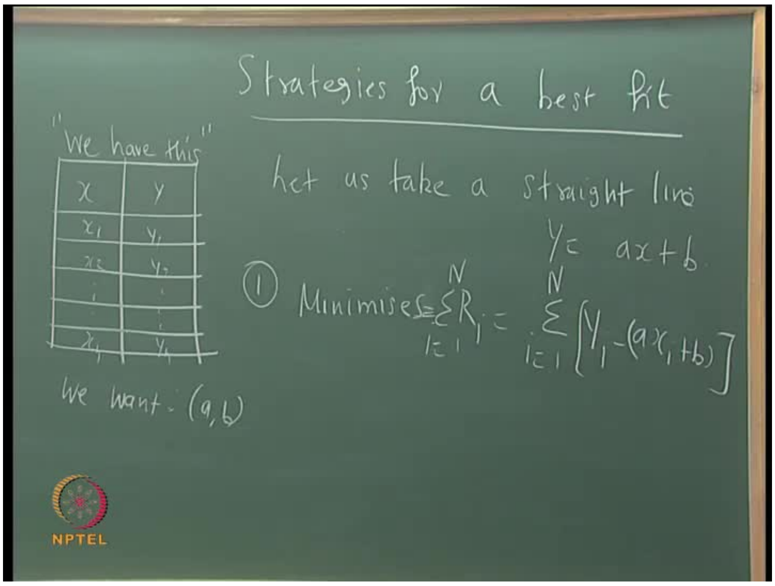
So, the first is minimize R, rather let us call it S its equal to sigma… is it okay? I am not saying whether it is right or wrong. People will try to come up with the simplest possibility right? Try to minimize the residue. What is the problem with this? What is the problem with this?

(Refer Slide Time: 10:31)



Let us take just two points for example. I have two points 1 and 2; common sense tells me that I should join them by a straight line. Suppose, I take something here; I propose this line now. This is line A; this is line B. If you propose this line, so here it has a deviation of minus delta; it has a deviation of minus delta from the original line, but here it as a deviation of plus delta from the original line. This minus delta and plus delta cancel out and then this will be give exactly the same value of the S of some of the residues as oppose to the correct line; are you getting the point? In fact, any line other than the vertical, any line other than the vertical for this case of two points will reduce the sum of residues, s to 0. So, you will not get a unique line; so, associated with non-uniqueness.

(Refer Slide Time: 11:45)



Strategies for a best fit

"We have this"

x	y
x_1	y_1
x_2	y_2
⋮	⋮
	⋮
x_N	y_N

We want: (a, b)

Let us take a straight line

$$y = ax + b.$$

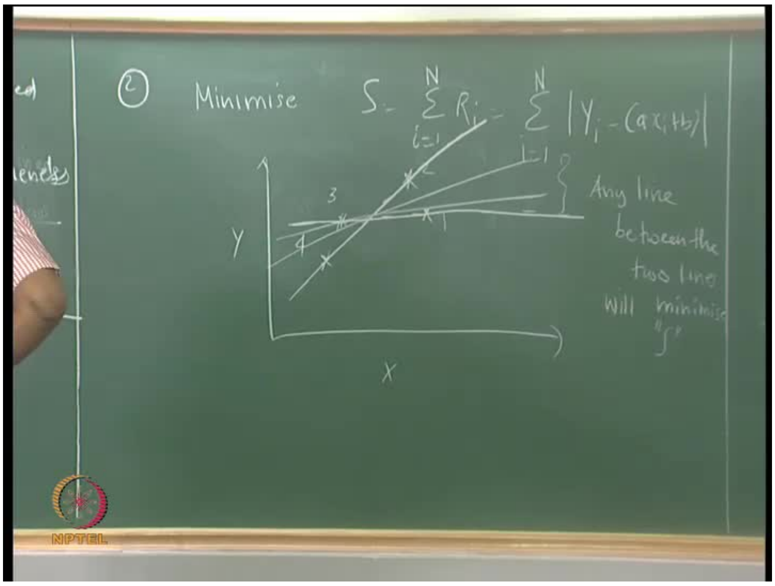
① Minimise $S = \sum_{i=1}^{N} R_i = \sum_{i=1}^{N} \left[Y_i - (a x_i + b) \right]$

Instead of looking at mathematically from a commonsense perspective, large negative errors may be compensated by large positive errors. It will accept that also be because it is not worrying about the modulus or its not worrying about the square or any other thing.

So, you can deviate from as many points as you can deviate merrily, so along as minus error is balanced by the plus error, it is alright; so it will lead to non-uniqueness. So, this is not the best strategy for, this is not the best strategy for… Please remember that we think, we think it is all silly and all that, but 150 years back, 200 years back, there were people who were working on all this. What is so obvious to us today, it is obvious because you have a book; you have the information somewhere. But before regression was developed, some people have to slog it out. Who was the fellow who finally figured out regression? Who was the guy who finally figured out? There are only 3,4 people know, one of them.

No, No. Gauss, Gauss. So, I do not know why these guys had nothing better to do. So, they keep, they work in so many fields and keep hunting us during night with new theorems.

(Refer Slide Time: 13:43)

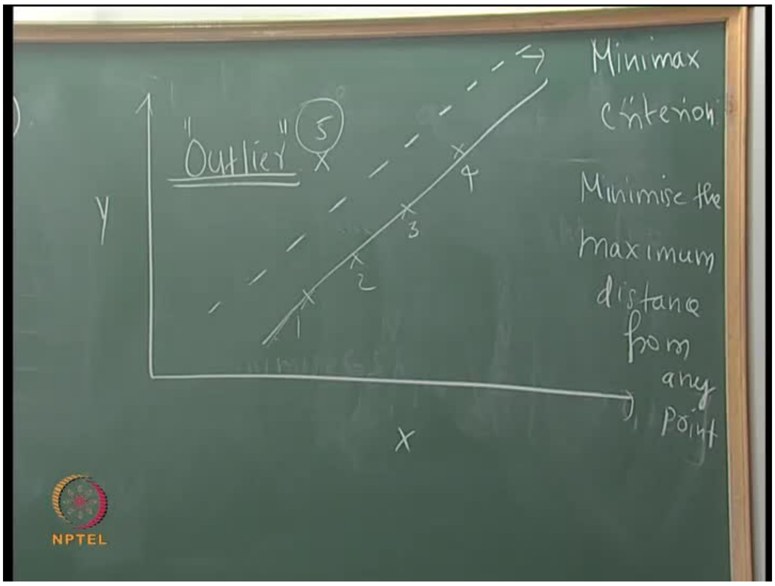

② Minimise $S = \sum_{i=1}^{N} R_i = \sum_{i=1}^{N} |Y_i - (aX_i + b)|$
3
4
2
1
Y
X
Any line between the two line will minimise "S"

So, the strategy number 1, minimize fails; strategy number 2, there we get a little better minimize. See, I am putting a general notation; S is the sum of the residue, S. S is the terminology or a simple frequently used in statistics. I can again put. Shall we do this? Looks better. I am minimizing the modulus; large negative and large positive cannot cancel out each other.

What is the problem with this? Let us take a hypothetical example. What could be the story? For example, let me draw this. Let me connect this to, let me connect; is it okay or you want me to connect to the straight line and do the, I means, solve it line? So, there are four points. So, I have joined points 2, 4 with the line and points 1,3 with the line. Any line which is occurring between these two lines will try to minimize this. There could be so many lines which will satisfy this. You can put it in Excel and try to sort it out. So, the line like this this, for any line, or alternative I can say, I can come out with many lines which will try to minimize this which are line between these two.

My whole point is unless I go home with unique value of a and b, I cannot rest in peace because I am looking for a b pair which will best represent the data. I am not trying to look at, I am not interested in doing research on trying to find out how many possibilities of a and b there, whether there exist (( )), uniqueness; I am not getting into a mathematical proof. Basically, we are engineers; we are looking at data analysis; we are looking at a perspective of trying to get best representation for the data. So, this also does not work.

(Refer Slide Time: 16:58)



Now, we are left with I have a nice set of data: 1,2,3,4; I have a nice set of data 1,2,3,4 which is like this. Unfortunately, I have one fellow like this. I also measured some thermo couple; it came. Common sense tells me that this should be the line. That should be the actual line, but I am not caring about 5 if I do that. So, one strategy would be I do not want point 5 to get angry. I will choose a line like this. On what basis did I choose this line? This line minimizes the maximum deviation from any point. This line minimizes; so, it is based on what is call mini max criterion. This mini max criterion basically comes from decision theory. People who are studying management or war might have studied. Has somebody studied about decision theory? People talk about zero sum game, stock market for example is zero sum game.

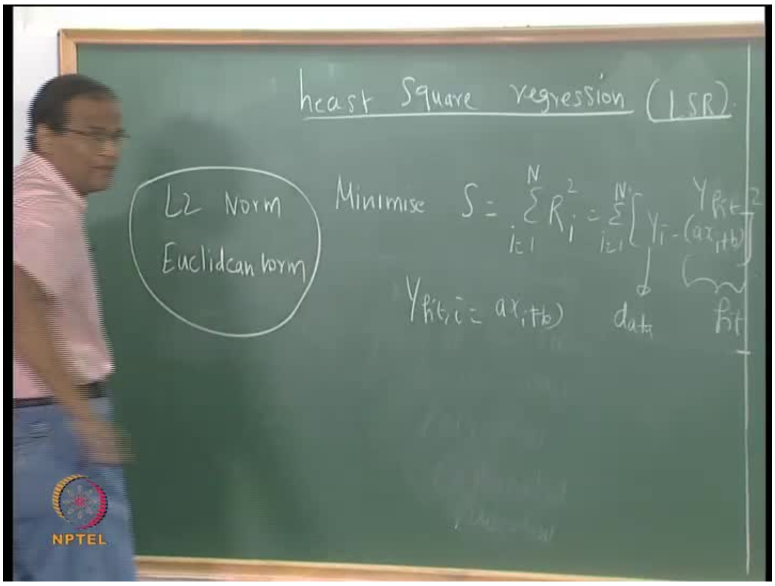
Zero sum game, basically if you keep on getting profit, somebody going to get loss because it is not a productive economy; we are not trying to make Hero Honda motor cycles into selling or something. Basically, you feel that, today reliance share if you buy a 10 rupees more, its worth for a you and this thing; it is more worthy and all that; somebody does not feel so and so your management people will model games; I mean game based on game theory and so on. So, the min max is based on maximizing the minimum gain and minimizing the maximum loss or something; so, you play such that you do not lose maximum or you at least ensure a minimum gain.

They talk about other things; for example, Prisoner's dilemma. Have you heard about the Prisoner's dilemma? So, two people are caught; two thief are caught; only one fellow has actually committed the crime; they are put in separate jails and all that; during interrogation, if both keep quiet, they can be set free, but if one says against the other, he will get a 5 years and then it will lead to some hypothetical situation, and other fellow keep… So, there is a, you can, you can come out with what… what is the strategy that each of them has to adopt? For example, if it is game and so on, but generally you will try to (( )) of this. When you (( )), actually it is not beneficial to either of you; that is what it says; if both of them keep quiet, it is much better decision and so on. So, if you go to Wikipedia and look at Prisoner's dilemma, there are so many variants to this and people have written papers and all that. So, basically, I think these are some decision theory. So, they look at complex moves which the human mind makes.

And as for as we are concerned, we are trying to minimize the maximum distance from any point, maximum deviation from any point, but let us not forget that this point, this point is a rank outsider. There is something fundamentally wrong with this problem. Statistically this problem is called an outlier. It is an outstanding point; it is an outlier. So, this mini max criterion unnecessarily tries to give too much importance to an outlier. See, it happens even in a meeting. When you are trying to discuss with your friends, if there are 8 people, one fellow shouts too much, even if he is un reasonable, we will try to pacify because other people are keeping quiet.

The most vocal person will be heard, will be heard, though that may not be the most reasonable view; this we often see in live. So, in statistics, this gives too much importance to outlier; so, we do not want that. So, what we will do is suppose we show it to a prof, what I will advise my student is to remove the point 5; I am not saying, quietly rub it off; we should not do that. What I will say is go and check; go back and do the experiment 2 times or 3 times. And then, 99 percent of the time there was some mistake associated with that point because suddenly we will not get a new physics or new heat transfer or new fluid mechanics; something is going up and down; there is something wrong with that point; either you did not allow for steady state to take place or something happened; there is a fluctuation in the voltage at that point in the time, or you made an error in reading the I mean the meniscus Nanometer; there should be some problem like this.

(Refer Slide Time: 23:11)

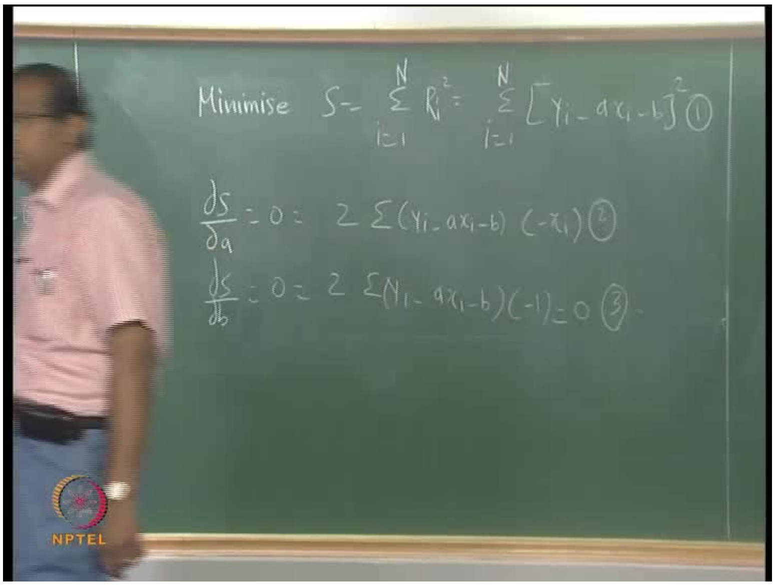

Least Square regression (LSR)
L2 Norm
Euclidean Norm
Minimise $S = \sum_{i=1}^{N} R_i^2 = \sum_{i=1}^{N} [Y_i - \overbrace{(ax_i+b)}^{Y_{fit}}]^2$
data
fit
$Y_{fit,i} = ax_i+b)$
NPTEL

So, with all these strategies having failed, so we try to see if we minimize the square of the differences between the data and the fit. If that is minimized, we can get a good handle on this; this called LSR; this like your YSR. So, there is a LSR; so, least square regression.

So, what is the least square regression? So, for the linear case, we will say minimize S sigma R i squared. What is this? This is data. This is the fit. What does it mean? Y it is equal to ax plus b if you have two points, you can exact you can get the values of a and b right away. If you have 200 points, you will have so many combinations of a and b. Now, we are trying to fit one global value of a and b for which best represents the data. So, whenever you have a procedure by which we determine a and b and you substitute ax plus b, this will be different from Yi data. So, the difference between the Y i data and then this is Y fit, right? Y fit of i is equal to ax i plus b. So, the difference between the Yi data minus Yi fit is basically called a residue. So, the residue is whole square and you are trying to minimize the sum of those residues. So, this is the least; so, you are minimizing the sum of the residue in a least square sense because the square take care of all your negative and positive and all this, and this also will helps you in getting a unique value of a and b for this. This basically, as Deepak said, this is basically what is called the Euclidean norm. It is called Euclidean norm; it is called L2 norm. It has got very interesting intuitive physical interpretations or very it is quite insightful and all that, which we will discuss in next class.

But now, suppose we have a data y versus x, what do we actually get as a and b? Differentiate the norm; two partial differentiation; do dou x by dou a equal to 0; do dou x by dou b equal to 0; get a set of simultaneous equations; solve for a and b. The next 5 minutes I will take an attendance. Please do it and then I will work it out on the board; please do this exercise, differentiate the.. Ashuthosh, you have to start doing. So, differentiate x with respect to, partially with respect to a and partially respect to b; equate it to 0; please note that sigma will keep coming through out and the total number of data point is n, small n.

(Refer Slide Time: 27:01)

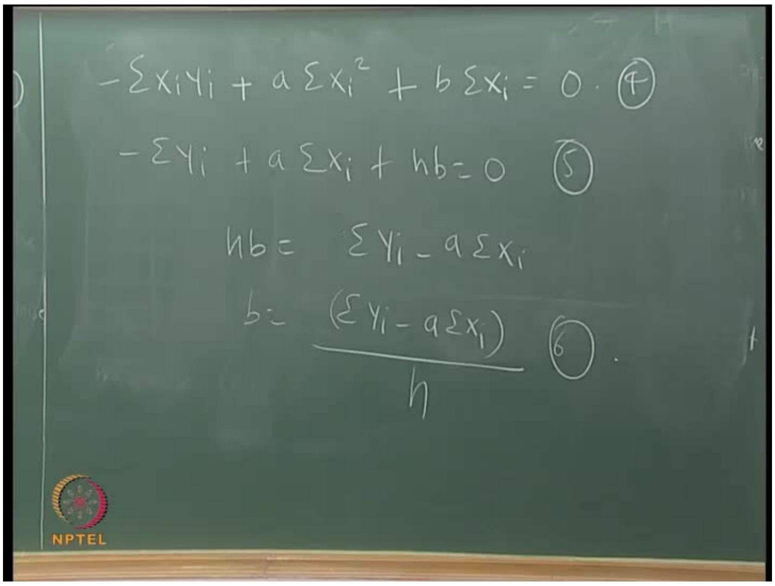


Minimise $S = \sum_{i=1}^{N} R_i^2 = \sum_{i=1}^{N} [y_i - ax_i - b]^2$ (1)

$$\frac{\partial S}{\partial a} = 0 = 2 \sum (y_i - ax_i - b)(-x_i) \quad (2)$$

$$\frac{\partial S}{\partial b} = 0 = 2 \sum (y_i - ax_i - b)(-1) = 0 \quad (3)$$

(Refer Slide Time: 28:27)



$$-\Sigma x_i y_i + a\Sigma x_i^2 + b\Sigma x_i = 0 \cdot \textcircled{4}$$

$$-\Sigma y_i + a\Sigma x_i + nb = 0 \quad \textcircled{5}$$

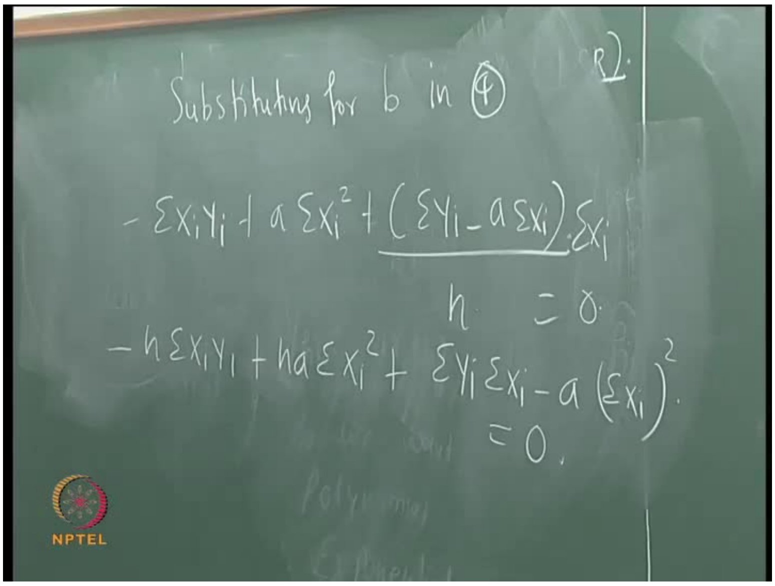
$$nb = \Sigma y_i - a\Sigma x_i$$

$$b = \frac{(\Sigma y_i - a\Sigma x_i)}{n} \quad \textcircled{6} \cdot$$

Correct? Is it okay? 1, 2, so, you can leave the 2; I may make mistakes - minus, plus; please let me know if it is mistake free; Kousthub, is it clean? Now, go ahead and solve for… I just tried to solve for b first; looks easy, but b has a; we will solve for b and then substitute for b in the first equation; then, get a and then we will say that if you have done so much with that, we can get the other one also. So, how to do this? So, what you want do know? Next step?

You can say that you can easily solve for a and b, but I want an expression so that it is easy for you to use it exam. It should be there in your notes. So, what do you want to do? Shall we write for a or b? You want to write for b? okay. It is not very good because it has a; it is not fully done yet.

(Refer Slide Time: 31:14)



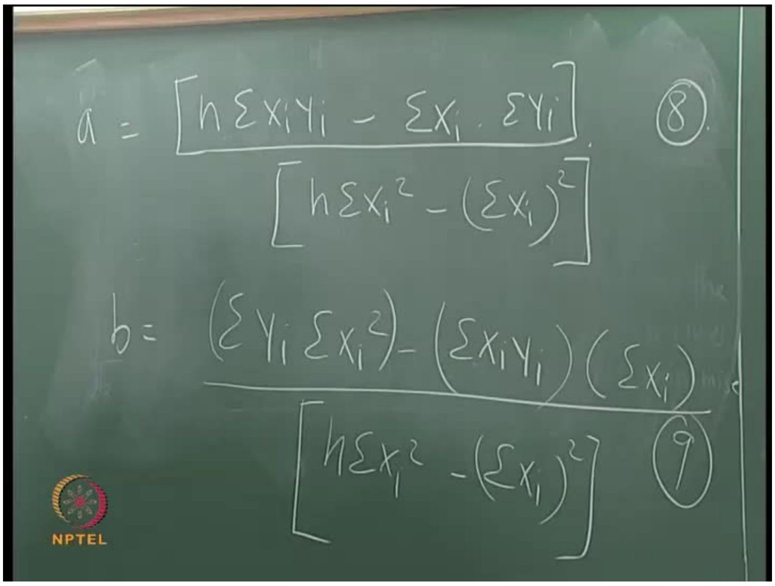
Substituting for b in $\textcircled{4}$

$$-\Sigma x_i y_i + a\Sigma x_i^2 + \frac{(\Sigma y_i - a\Sigma x_i)\Sigma x_i}{n} = 0$$

$$-n\Sigma x_i y_i + na\Sigma x_i^2 + \Sigma y_i \Sigma x_i - a(\Sigma x_i)^2 = 0.$$

So, you can substitute for b in 4. Inter sigma xi; is it correct? Is this okay? Okay. So, minus, plus - alright? Please note, please look at the board carefully; sigma x i squared is far different from the sigma x i squared. I hope you are able to get the difference. So, a equal to… please tell me.

(Refer Slide Time: 33:08)



$$a = \frac{[n\Sigma x_i y_i - \Sigma x_i . \Sigma y_i]}{[n\Sigma x_i^2 - (\Sigma x_i)^2]} \quad \textcircled{8}$$

$$b = \frac{(\Sigma y_i \Sigma x_i^2) - (\Sigma x_i y_i)(\Sigma x_i)}{[n\Sigma x_i^2 - (\Sigma x_i)^2]} \quad \textcircled{9}$$

Sigma?

Student: Sigma x i sigma y i.

Student: n sigma x i square sigma x i whole square

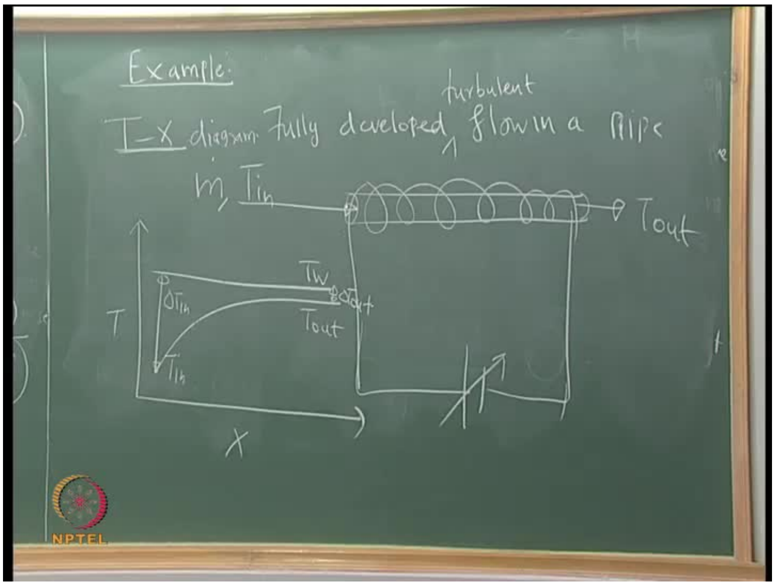
6, 7, b equal to?

Student: sigma y i sigma x i square into sigma yi x i

Student: sigma x i y i? by n sigma n sigma x plus same thing.

Please try to do it yourself and do not copy directly from the board because in the exam in the quiz we are going to do; you are going to use this; you will not have time to do dou x; you have seen my paper once. Now, you will not have time to check whether what you have done is right or not. If you are taking the formula, if it is incorrect, there is the propagation of error. So, please ensure that this formula is right. It is very tricky. This square you can miss a square; you can miss a minus sign; you can miss a n; so, be careful about this. So, you have got the values of a and b; unique values of a and b. So, this is the best representation of the data. Fine? What these mean and whether we have really, whether we have really conquered the problem? Whether there is a substantial improvement in our understanding for the representation by trying to force y equal to ax plus b and all that we will evaluate in a short while. Anyway, so we got a and b at least.

(Refer Slide Time: 35:35)



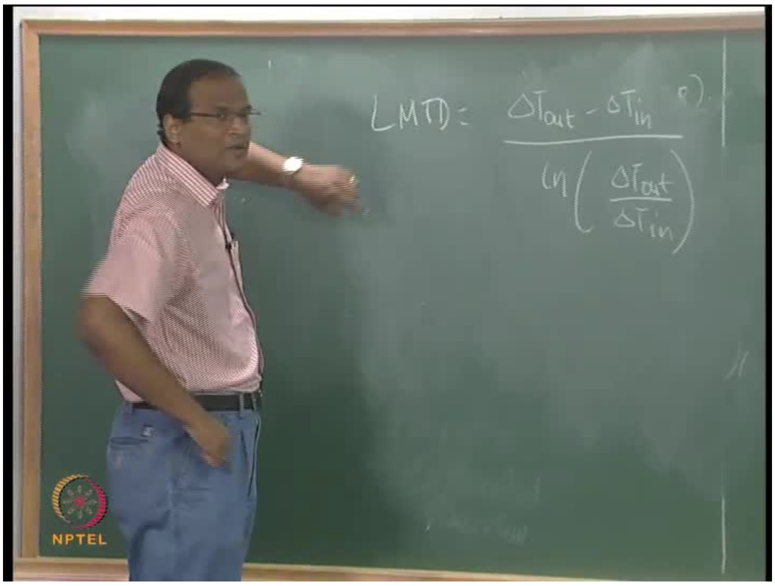
So, let us start with simple example. Let us say, we have a fully developed flow in a pipe. I have taken a pipe like this. I have a fully developed turbulent flow. I send water at the rate of m dot. I can measure the temperature at the inlet. I can measure the temperature using a thermometer or thermocouple or what have you. Now, I can wrap it p with the electrical heater. I can weigh the power. It is also possible for me to individually control so that so I have constant temperature on the outside. Correct?

So, I am sending cold water inside a pipe, the outside of which is either maintained at a particular temperature or it is spread with electrical heating so that the job here, so this is essentially a heat exchanger which is tries to heat the incoming water. Now, it is possible for me to change the velocity. When I change the velocity, in order to maintain the temperature of the out of all constant, I can change the settings. So, now, which means for every change in velocity, I can change the heater setting and I can keep the temperature constant. So, I can repeat this experiment; that means, when I change the velocity, I change the Reynolds number and then I am changing the heat transfer rate. So, I am seeing how the heat transfer rate varies with the fluid velocity.

This is essentially what we do in convective heat transfer. If this pipe where assumed to be at a particular temperature, then the pipe wall temperature is like this and then this temperature goes up like this. This is called a Tx diagram of the hate exchanger; temperature length diagram, T x diagram.

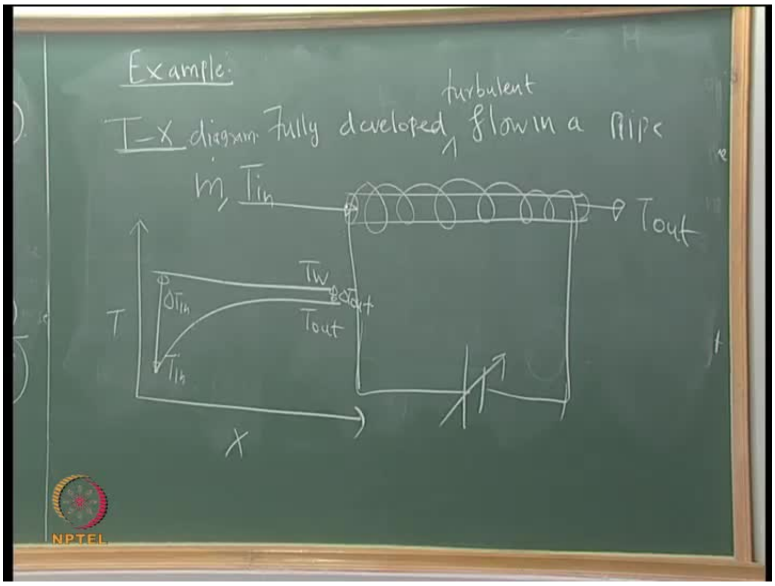
When you do this, so, this is delta T in and this is delta T out; I am surreptitiously discussing one question in the quiz. So, you because one is a straight line and the temperature distribution is exponential, you cannot use an arithmetic mean temperature difference. Therefore, it becomes contingent upon you to introduce a log mean temperature difference. What is a log mean temperature difference?

(Refer Slide Time: 38:45)



So, basically, here it is delta T in and here it is delta T out. So, the log mean temperature difference is delta T out minus delta T in divided by log of…. The logarithmic the mean temperature difference is the temperature difference at the one end of the heat exchanger minus temperature difference at the other end of the heat exchanger divided by the logarithm of the ratio of these two temperature differences.

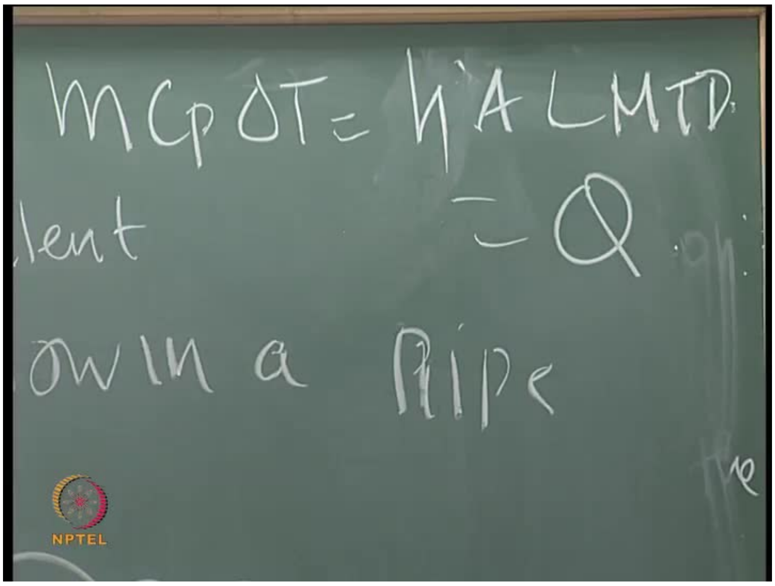
(Refer Slide Time: 35:35)



It cannot be equal to arithmetic mean temperature because the capacity for the fuel to pick up the heat will keep on decreasing because the potential difference decreases. So, it is the law of nature that since it decreases, then it will have an, it will have a saturation curve. When there is a saturation curve, if you take T in and T in plus T out by 2, and put, mark some point and take this difference, that is a very bad approximation. In several cases, that is called the AMTD - arithmetic mean temperature difference, which does not work. So, essentially, you can do experiment like this.

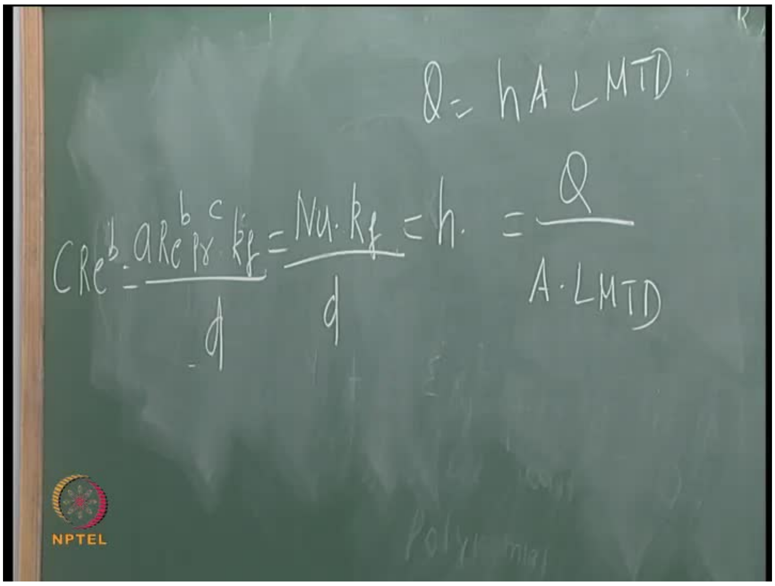
Suppose, instead of water you can change the fluid; if you change the fluid, an additional parameter which enters the problem is a Prandtl number. So, this is an experimental approach to determinate the heat transfer coefficient. So, how are you going to do this?

(Refer Slide Time: 40:33)



So, this mc p delta T is equal to UALMTD. In this case, as the outside temperature is fixed, I do not have to use the U because I know what the temperature is there outside. I do not need to use the concept of overall heat transfer coefficient. I will just use HA LMTD; this must be equal to the Q which is coming.

(Refer Slide Time: 40:41)


$Q = hA\ LMTD$
$CRe^b = \frac{aRe^b Pr^c\, k_f}{d} = \frac{Nu \cdot k_f}{d} = h = \frac{Q}{A \cdot LMTD}$
NPTEL
Polynomial

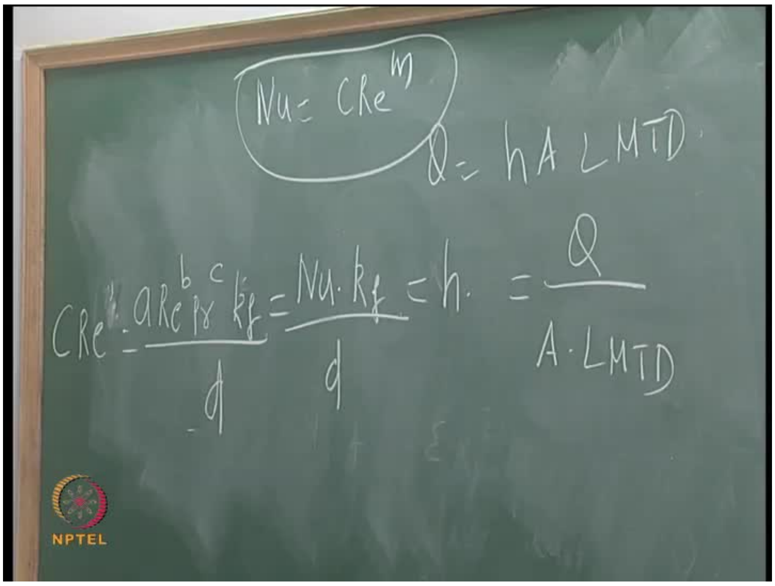
So, once I use this, I can divide Q divided by A into LMTD; this is by heat transfer coefficient. This heat transfer coefficient is equal to by Nusselt number in the thermal conductivity divided by the diameter. This Nusselt number is equal to a Reynolds number to the power of b Prantal number to the power of c into thermal conductivity divided by d. This is c Reynolds number to the power of b, where the c absorbs all the other constants.

Now, you know the link, missing links in the chain. Suppose, I just say that, in a heat transfer experiment, x is known to be a, x is independent variable and y is the dependent variable, as an engineer, we will post the problem like this. So, I have worked out something based on this and I will give you a problem. Now, you open out a tablular column and solve this problem. We will continue the solution in the next class.

In another 5 minutes, you can start off problem number 14? 16

Experiments on a fully develop turbulent flow in a pipe, experiments on a fully develop turbulent flow in a pipe under steady state conditions have been carried out; experiments on fully develop turbulent flow in a pipe under steady state condition have been carried out Experiments on a fully develop turbulent flow in a pipe under steady state conditions have been carried out with water as the medium.

(Refer Slide Time: 43:51)



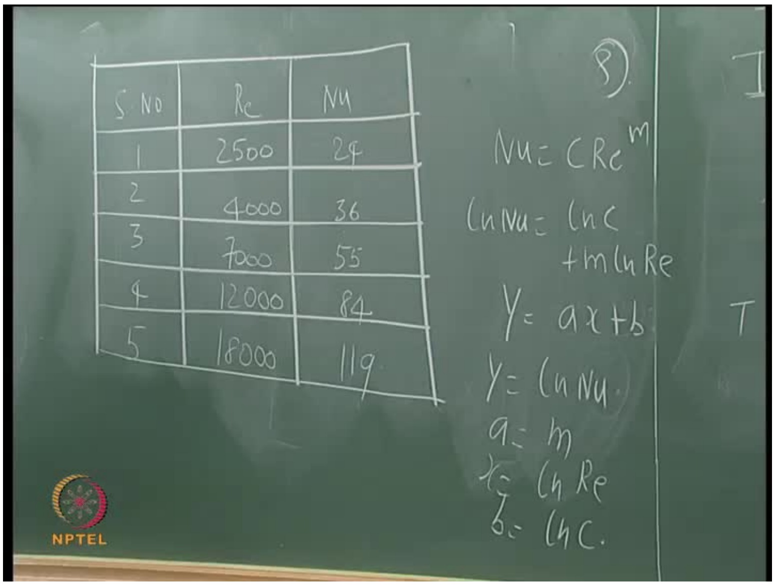
$$\boxed{Nu = CRe^m}$$

$$Q = hA\,LMTD$$

$$CRe^m = \frac{aRe^b Pr^c k_f}{d} = \frac{Nu \cdot k_f}{d} = h = \frac{Q}{A \cdot LMTD}$$

The variation of the Nusselt number with Reynolds number is given in tabular form. Assuming a power law for the Nusselt number of the form, I can say it is easy for you to have a and b, okay, of the form Nu equal to C Re to the power of m. using least square regression get the best estimates of C and m, using least square regression.

(Refer Slide Time: 45:19)



S.No	Re	Nu
1	2500	24
2	4000	36
3	7000	55
4	12000	84
5	18000	119

8)

$$Nu = CRe^m$$

$$\ln Nu = \ln C + m \ln Re$$

$$Y = ax + b$$

$$Y = \ln Nu$$

$$a = m$$

$$x = \ln Re$$

$$b = \ln C.$$

I have straight away introduced the complication. What is that? I told you how to regress only a straight line, but now I am saying it is a power law. So, what should I do first step? I take log on both sides; I take a natural log on both sides and reduce it like this, first step. Why I told you that it is important to write the formula is you have to expand this tabular column and just start putting lonRe, lonNu, and lonRe you call it as x and lonNu you call it as y; then you know what all you require; xi x a y i x square; then you open up one more row here and put sigma for all these. So, you will have sigma x I, sigma y I, sigma x i y i, sigma x i square. Then, once you have all these values, then it is plug and play; put all these values for the formulas a and b, and get a and b.

You can just start now. You can just start with the tabular column and anyway, it is going to take 20 minutes. You are going to do it in the tomorrow's class. We can we can make start now. It will take half an hour. Do not do it with your calculator. I have already told you right? If you put this 5 hours, it will give the regression or… anyway it does not in work; in exam you have to show me the table and what will do you do for a non-linear regression? Calculator will not help. Quiz 2 non-linear regression, the Gauss Newton method, where the beauty of the non-linear regression is a and b will keep changing with iteration; how nice it will be right?

Tell me how many columns you need? Serial number, Reynolds number, Nusselt number that is 3; one, then log of Reynolds number - column 4; log of Nusselt number - column 5; then x square x y and then I also have columns for y bar, y fit, y fit minus y bar whole square; I mean all that we need to statistically evaluate the performance of a regression. I will talk about this in tomorrow's class. I think it is time now. So, you will go to the next class. So, we will stop with this. So, we will solve this problem in tomorrow's class and look at the insights it offers.