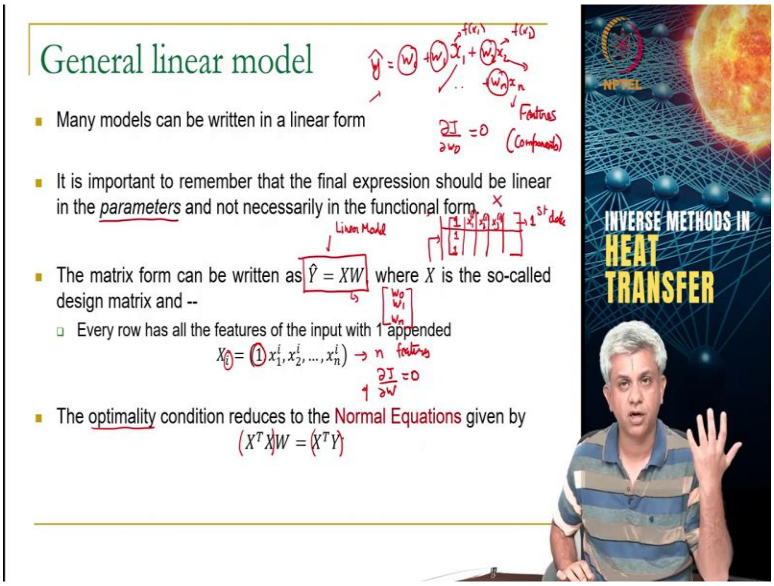
Inverse Methods in Heat Transfer
Prof: Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology, Madras

Lecture No 17
Summary of Week 03

(Refer Slide Time: 00:32)



Welcome back. This is just a summary of what all we saw in week three and what its significance is. we started if you remember with trying to generalize the linear model that we had last week. We saw that many models can be written in a linear form the general trick is that if you have any model and it can be written in the form $\hat{y} = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n$.

And in fact, this could be some other function of x_1 also and this could be some function of x_2 also any of these cases is linear provided it is linear in the parameters. So, linear in the parameters means it should be linear in w_0, w_1, w_2 and w_n. So, that is where we started that is what we call a generalized linear model. Now what was important was even if we have a generalized linear model if you want to solve keep on solving $\frac{\partial J}{\partial w_0} = 0$ etcetera.

And all these equations we saw last week that is cumbersome, even for a quadratic model with just three parameters it got painful. but you can write a general Matrix form provided you have you do a few tricks which we did this week. So, we had this x which was a matrix which as I have written here. every row is basically sort of a row in an Excel sheet. This contains all the input vectors appended with a 1 in the beginning and you would have something like x_1, x_2, x_3 where x_1, x_2, x_3 are features.

These are called features are another way to think of it is these are components of the input. I gave you the example this could simply be Cartesian components for example x is a location vector and if you are in 3D then x_1, x_2, x_3 would literally be the x y z coordinates or x could be something more complicated like temperature and pressure or temperature and heat flux and you are measuring something else.

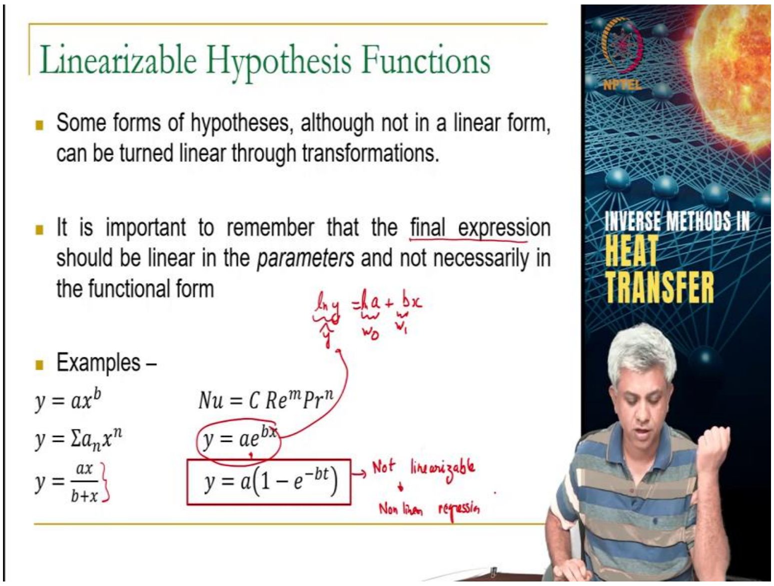
Now if you have something of that sort in general since we are doing the general case here, in general it is just going to have multiple features and we store each data point. So, again think of these as each of one corresponds to a single data point. So, the first data point will have multiple input features and the ith data point will have n features with one appended here to in order to make the model work.

In that case if we write w itself as a matrix w_0, w_1, w_2 and w_n as an n cross 1 Matrix you can then write this equation which is the general linear model. And if you write it as this General linear model, in fact if you would have seen an advantage of doing. So, while we were writing our code you can write the $\frac{\partial J}{\partial w}$ equations, the optimal equations as the normal equations and it reduces to a system of equations which is $X^T X$.

This is the Matrix you are solving or this is the Matrix which sits on the left-hand side, multiplied by w is the right hand side, which is x-transpose y. we will see an extension of this next week when we go to non-linear regression. So, we will see that this equation is not just what it appears it has certain other subtleties which I briefly alluded to this week but will see a more General case next week.

In any case we saw that using this kind of approach actually easily let us go from linear to quadratic to polynomial to many types of models and this can be easily programmed in in a compact fashion. So, that is the advantage of having a general linear model.

(Refer Slide Time: 04:33)



Now apart from a model which looks linear to start with, you can have models that do not look linear in the parameters. I am not talking about linear in you know the features but linear in parameters. So, you could have terms like e^w okay e power the parameter you are solving for if the parameter you are solving for is a you could have terms like e^a as we will see shortly or something power a parameter etcetera.

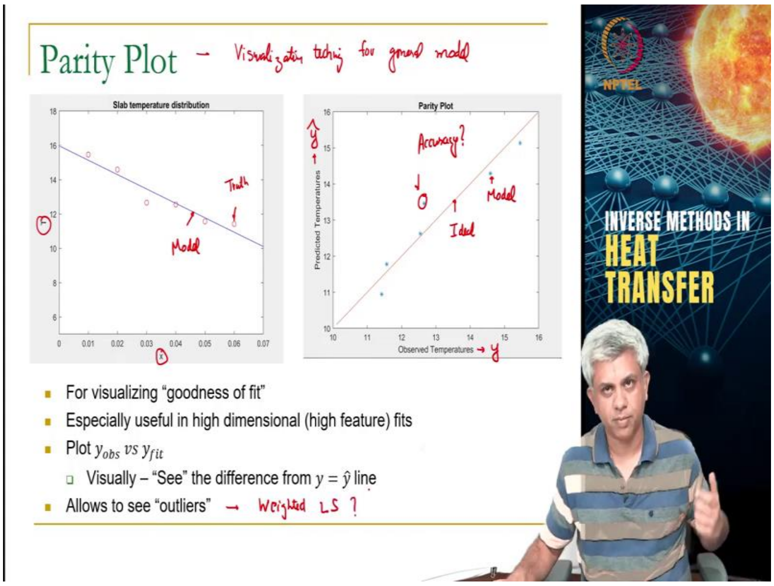
So, all these things are not linear in the parameters, but we can make them linear and the important thing is the final expression. Again, I am just repeating the statement from the previous Slide the final expression should be linear in the parameters it need not even be the initial expression you just have to transform it into a form. So, here are some examples we looked at this week. many of the examples can simply be linearized by taking a log, for example, if you have let me take this example this can be converted into a linear form by taking log.

So, you have ln $a + bx$ as you can see this you can call $\hat{y}$ this you can call w_0 and we I literally did the same example earlier but this is just to jog your memory that you can actually turn something that looks non-linear into linear. As I also told you within this week, this form can occur within fins we are taking simple cases. So, for example a fin with infinitely long condition basically meaning that it is much longer in comparison to its thickness in comparison to its width.

So, in this case you can actually use this approximation and we will in fact at least we will see some exercise problems etcetera, as we go on in the course in such cases. You could also have things that do not look linear here and if you carefully handle it as I showed you this week you can turn it into something linear. However, there are cases that are not linearizable. This does not mean that every single case is linearizable. there are some cases we will come to which are non-linearizable.

And this is where next week we will switch to non-linear regression to handle such cases. So, here is an unsteady again an unsteady convection case in some sense, which is non-linear which is you have to handle with ah non-linear regression.

(Refer Slide Time: 06:58)



The next idea we looked at was a parity plot. This as I told you in the beginning of this week, is basically a visualization technique for a general model. What do I mean by that the entire effort this week was to extend it to cases that look more than just $w_0 + w_1x$. So, if you have x_1, x_2, x_3, x_n xn you cannot do this kind of visualization which is just x versus y or x versus your predicted variable.

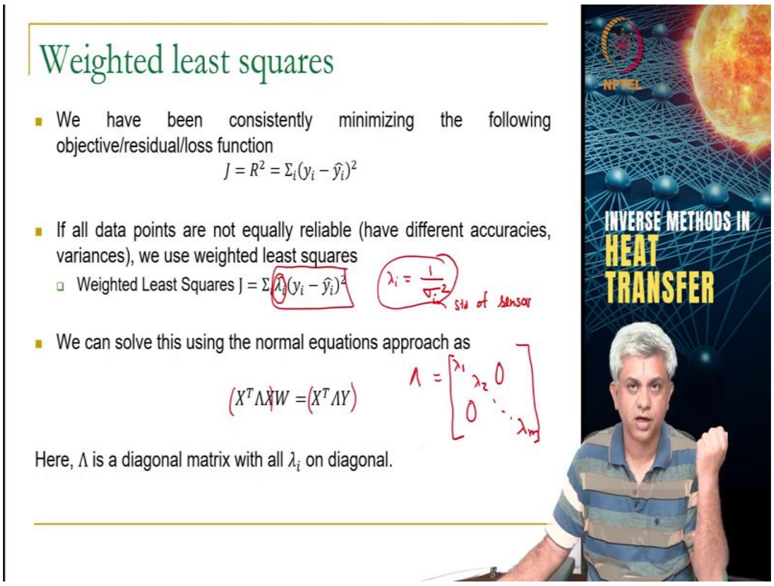
Because x is now multiple variables now how do you handle such a case now instead of plotting your features versus the output, you just plot your y versus your $\hat{y}$ which is always just single or at least you have a maximum of three or four predicted variables. So, you just have 3, 4- 2D plots um the idea here is to plot this ideal line and this line shows how well our model performs.

So, you see in fact in some sense it is switched in this case the Dots here represent a reality and the lines here represent the model and it is kind of flipped in how a parity plot works. So, in this case the ideal is a 45 degree line and the model depending on how far it goes from here you can then start identifying things that are really far off and start examining whether your sensors are working properly or whether their accuracies are desirable or not.

And in case it looks like you have given equal weightage to lines, which you should not have given equal weightage, we will come to weighted least square. So, that is one way to approach weighted least squares is to notice that some points seem to be too far off from our model and if we are confident about our theory, we can then start examining is the accuracy of the sensor good or not.

So, if the accuracy is not good and each point has a different accuracy then we can start thinking whether we should reweight each accuracy. So, as I have mentioned here it is useful for high dimensional or high feature cases you plot y versus y-hat and you can visually see the difference and allows you to see outliers and perhaps you can move to weighted least squares if required I mean if we find something off.

(Refer Slide Time: 09:28)



So, the final topic we did this week was weighted least squares. So, consistently we have been minimizing ah $(y - \hat{y})^2$ but however just like I mentioned in the previous slide, let us say that if all data points are not equally reliable. So, these have different accuracies or you plotted a normal you know least square line and then you observed as I showed you last time that there is some outlier you start examining.

And it turns out let us say that the thermocouple is old; its error has increased, standard deviation has increased something of that sort in that case you can put weighted least squares.

Weighted least squares have this extra factor here and this corresponds to $\lambda_i = \frac{1}{{\sigma_i}^2}$ and sigma here is the standard deviation of the sensor. So, suppose you make let us say a known temperature you measure it 100 times and you see a certain standard deviation within the sensor you measure it and say ok.

So, you have to calibrate your sensors find out your thermocouples. let us say and then you say ok the variance is so, much. And then if not all of your thermocouples are equally inaccurate or equally accurate you will have to weight them differently. The solution in the normal equation sense which I just wrote I did not derive it but it is something it is possible for you to derive it.

I derived the scalar form of this week but you can derive this one is

$$X^T \Lambda X W = X^T \Lambda Y$$

where Lambda is a diagonal matrix. So, Lambda will have Lambda 1, Lambda 2 up till Lambda m if there are m sensors and off diagonal is all 0s and remember Lambda itself is 1 over Sigma Square. So, if Sigma is very low it means it is a very highly accurate, low variance, like it always gives you very pinpoint results.

Then it means that such results should be weighted highly it means you should I in fact showed you an example ah this week of what happens when you weight a sensor highly the line sort of biases itself towards it. So, and the line will bias itself away from a sensor that is not so accurate. So, that is what we did, I hope. This summary kind of gels together what we did this week and the overall theme.

Next week we will move on to the case of non-linear regression things that cannot be handled with simple linear models. So, that is it for this week I hope you enjoyed this. I hope you do the exercise problems there are a couple of exercise problems main exercise problems this week one deals with weighted least squares and one deals with the normal equations sorry the unsteady conduction case where you have to linearize and then do a linear regression. So, I hope you enjoy this week's exercises see you next week thank you.