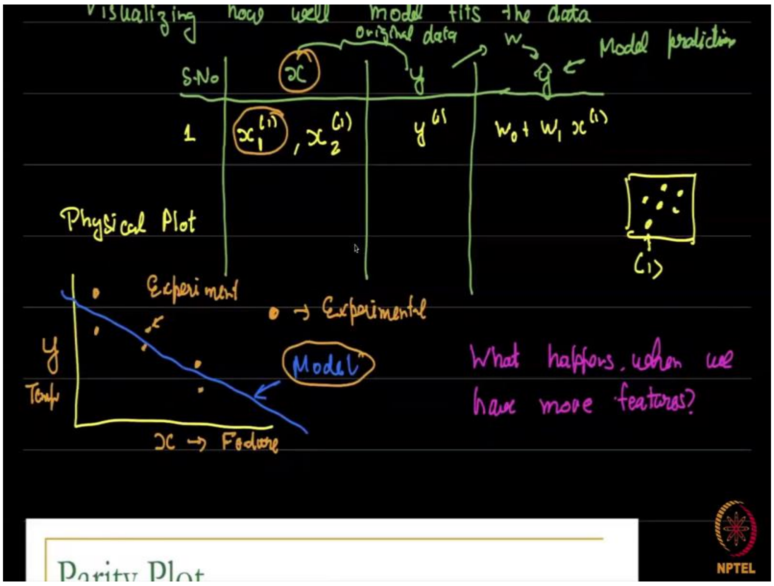
Inverse Methods in Heat Transfer

Prof: Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology, Madras

Lecture No 14
Parity Plots

(Refer Slide Time 00:21)



In this short video, we will look at this idea of parity plots. This is a way of visualizing, how well our model fits the data. So, for instance, we have this; what we have been doing So, far. we have this serial number, we had some original data, we have all this being the truth, this is the original data. And we use the original data to solve for some parameters and we use this parameter back in order to make a model prediction.

For example, the linear model that we have been seeing so, far. In this case typically we have this x at the first data point and y at the first data point. So, if it is a simple linear model. Now in this case, we have of course as we saw in great detail in the last couple of videos something on this side. now typically you would plot the usual what I will call the physical plot, which we will also see in some detail in the next video is you plot the original data, the original data could be something of this sort.

So, these points are the experimental points and then you draw the fit. The fit here is some line of some best fit line typically let us think of this mind. So, so this is the model and this is the original data experiment. So, visually the only way for us to see , how well something does is to compare, you know, how close this model line looks to the experimental data now we also had a numerical measure which is our goodness of fit for this but if we start here notice what is there on the x and y axis.

So, on the x-axis is what I will call the feature, in this case just x a single variable and on the y-axis is whatever we are plotting against. So, for example in this case, in case of thermocouples, this could be the temperature. But the problem that occurs is what happens when we have more features. what do I mean by more features for example you could have x location and y location.

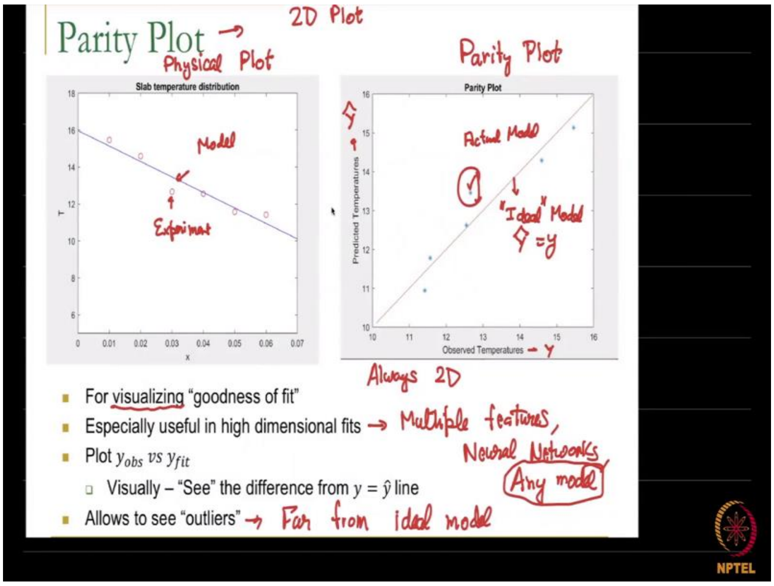
So, we will call it x_1 which is x location x_2 of the first thermocouple point. So, we are now thermocouples on a plate as I said in the last video and you have all these let us say in this case, I have drawn five thermocouples. So, this could be the first point and all we are trying to find out are what are the x location and the y location of this point. So, those two go as input.

Now how would we plot in such a case, this becomes a little bit messy. your output is still one it is still temperature but your inputs are two now imagine when you have three at least in this case you could draw an x, y and a z axis and then maybe draw a surface plot. So, in this case you will have x_1 and x_2 and this would be temperature and then you would have to draw surface plot and that already becomes a little bit hard to visualize so, hard to visualize.

The fit now worst-case scenario as we will see of course in engineering problems it is not just two but you could have three now in 3D means you have to now visualize a four-dimensional plot which you cannot do at all. So, if you are doing something like I said a microchip temperature or something of that sort and you are in a 3D body with thermocouple everywhere you have x_1, x_2, x_3.

Now how do we plot we really cannot plot the fit and we cannot visualize it. So, the solution to that is the parity plot now the things to notice is what is there on the x axis on the y axis of the parity plot. it is not the location or the feature versus the output temperature it is something else. So, let me show you that.

(Refer Slide Time: 05:25)



So, now on the left here, is the physical plot. I am going to call it the physical plot or the plot of the basic experimental setup. The next one is the parity plot. So, these two-look different as you can see. Again, the physical plot is very simple on the x-axis you have the x location and on the y-axis you have the temperature. you of course have the model and this is the experiment. So, by looking at how similar or how different these two are you can notice whether the fit is good or whether it is bad.

Now let us come here, I have something else plotted here. Now what you have here is observed temperature versus predicted temperature. Another way of saying it in our language is observed temperature is just y and predicted temperature is y-hat. So, if I have a data set like this here, the physical plot is between the first two variables. it is between these two is the physical plot, whereas between these two is the parity plot what is the advantage. In case you have just one single predicted variable.

For example, if you have just uh one temperature which is predicted, then you just need one plot and it will always be a two-dimensional plot. A parity plot is always a 2D plot. Because regardless of how many dimensions your box has whether it is 2D or whether it is 3D, whatever case you have even, if you have 2D the temperature itself is just one single value. So, instead of plotting the x, y locations of this versus the temperature what you plot is what was the actual temperature versus what was the predicted temperature. Actual temperature versus predicted temperature.

That is all the parity plot does. So, here the blue stars if they are visible are this. for example, at this point the actual temperature was 11 Point let us say 5 whereas the predicted temperature was something around 11. So, you can see that there is a difference and that difference is shown through this line in the middle which is the ideal model.

An ideal model regardless of whether it is linear non-linear whatever it is, the ideal model will always be a straight line because the ideal model says that $\hat{y}$ will always predict exactly whatever is y. So, that is what a parity plot does this is the ideal model of course this remains the Same regardless of what our model is whereas this star is the actual model. So, the distance between the ideal model and the actual model tells us how well we are predicting.

Now not only that we can also find out outliers. So, let me go over this. So, the purpose of the parity plot is to visualize, goodness of it, we of course have a single number talking about goodness of it, but this one visualizes in a very clear way. This is especially useful for high dimensional fits or wherever you have multiple features. Now the other thing is, this is also useful for asset filter or not we will show you neural networks.

Any model it does not have to be a linear model whereas the goodness of fit that number itself depends strongly on it being a linear model, which I will discuss later within the course. But this is good for any model whatsoever. Why because always, this is 2D. Now you might ask suppose I do not have just one predicted variable or one measure variable, suppose I have two or suppose I have three, I am making a model for the three different variables or two different variables like uh temperatures and heat flux.

Then if you have multiple variables, you make multiple parity plots. but that is not difficult compared to saying I am making a multi-dimensional plot. You will always have only a few are predicted variables. So, you will have two, three for usually for engineering applications. we typically do not have more than three predicted variables worst case scenario four. So, you look at four different plots.

But each one of them will simply be simple 2D plot. So, it is always easy for you to see how well each variable has been predicted. This point here as I have mentioned this point about outliers at the bottom. This point here sort of looks like an outlier. An outlier is a point which is very far away from the ideal model. So, outlier is far from ideal model. In some cases, this might be because we are making poor predictions or it might be because of poor observations.

In case it is due to poor observation, you examine that sensor and you find out that it is faulty, you can throw away this data. So, this kind of thing is actually possible when you have a parity plot. So, a parity plot is a extremely useful thing, to see this difference between the y equal to $\hat{y}$ line, basically the ideal Model line and it is a useful toolkit to have within your tool bag.

So, what we will do in the next video is to actually see, how to generate parity plots as well as all the other normal equation solutions that we did in the previous videos, Thank you.