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

Lecture No 15
Programming Inverse Methods Using Normal Equations

(Video Starts: 00:21)

In this video, we will go over the code for solving the inverse conduction problem using all the methods that we have seen so, far. I will also be introducing you to how to draw parity plots and we will see the major advantage of using the normal equations approach, rather than the sort of handcrafted approach, that we did for solving the linear regression problem.

So, I am going to start with the standard inverse conduction problem within the slab that we looked at. So, I am actually copy pasting within this the file that you see here is, what is known as an mlx file you can see that on top of it imhtv3 code.mlx I will be sharing with it. The NPTEL team will tell you we have to sort of download this file for this week just so, that you can play with it yourself.

This is what is known as a live script within MATLAB this is the equivalent of the Jupiter notebook that you would find within a python script which if you are used to using python uh you would find this IPython notebooks or Jupiter notebooks. This is the MATLAB equivalent. The major advantage of it is you have these code sections which are separate and then you can write your own notes there.

Something that I do for my own research or my own work is I do stuff like, I would for example if I have some handwritten notes. let us say I have some handwritten notes for this class maybe later on within this course I will introduce you to this sort of mix also. You can actually take the handwritten notes take a photograph and put it within the code. So, this is useful because you can just write a code at the bottom and you have your handwritten notes.

Here I have not done that but you can see one small example of this, this is sort of a screenshot of the question that I had given you um for the inverse conduction within a slab problem I have just taken a simple screenshot that can be interspersed with code which is here. So, you can write formula you can write formally in latex format you can see I have a latest format here that lets me write the equation.

So, this is something that I typed by hand this is not a screenshot but this is also something that was tap type by hand. But you can take screenshots you can take photographs of your written code or written hand notes and just put them maybe you can cut paste some portions of a book for the formula put that there just. So, that you can code it a little bit more easily I think for practical coding using a Jupiter notebook in case you are using python or a MATLAB notebook is what I would like to call it they call it a live script that is all of this is extremely useful.

So, I have this here, the description of the problem is here the same old problem we have six thermocouples within a slab and kept at 0.01, 0.02, 0.03, 0.04, 0.05, 0.06 and the temperatures are given here and based on that we basically wish to find out what the we wish to estimate the heat flux as well as the boundary temperatures. Now as we know this basically boils down to three tasks.

Our tasks are basically the following. you first have to model the forward problem from the physics of the problem. Then you have to solve the inverse problem, as we saw this solution of the inverse problem is basically finding out the parameters of the forward model. So, the unknown parameters you wish to find out in the least square or in an optimal sense using linear regression.

And then you post process you know maybe you determine the quantities of interest uh using the determined parameters and the forward model. Additionally, you could also do something like let us do this, you could additionally do things like visualize the solution. So, you can have the physical plot or you can have the parity plot etcetera. So, we will do this to here in this.

So, the forward model as you remember how we determine the forward model, was to solve the physical problem which gave us $T = ax + b$ or the way we took it we called $\hat{y}$ as the predicted value of T or the model for T and we wrote that as $w_0 + w_1 x$ and basically, we want to find out the optimal values of w_0 and w_1, as we have been doing over the last couple of weeks and this is given by minimizing the cost function.

The cost function is the gap between the model and the actual predicted value, the actual ground truth squared. Now we used to current approaches one was a direct formula. So, the direct formula which we determined by analytically calculating $\frac{\partial J}{\partial w_0}$ setting it to 0, $\frac{\partial J}{\partial w_1}$ setting it to 0. And we had written if you go back to your notes, I had written two alternate expressions for this.

The expression I am using right now is the most popular one which you will see, rather than the averages approach also which I had told you basically get w_0 and w with this formula. So, I am first going to show you how the direct formula is coded in. first of course you have to write down the data I hope I have written this down correctly. So, 0.1 through 0.06, I wrote down x.

This portion $X = X'$, X Prime is just a transpose. So, this was these looks like a row but I transformed it into a column by writing this. This Prime here just represents a transpose in MATLAB. same thing here, data here is 15.46 through 11.42, 15.46 to 11.42 up here and I have just given those 6 y data points and again $Y = Y^T$; m here is the number of data points.

You can see this in the formula also, m is the number of data points. MATLAB allows this you give a $\vec{x}$ it will calculate m; m would have been calculated to be six. I will show it to you once we run the program. Sx, I have put the terminology that capital S represents sum. So, Sx is basically you can calculate sum of x, there is an internal function in MATLAB which does this Sy is sum of y; Sx2 is obviously sum of x^2.

If you put a dot, it basically means every element of x^2 and gives a new Matrix. So, I can dot represents element wise multiplication or division etcetera. you can do this or power you can do all those operations. So, x dot square if I do x square it cannot do x square, because x is a column Matrix you cannot do a column Matrix versus multiplied by a column Matrix because the Matrix Sciences will not match.

So, same here you have the column Matrix x the column vector y you can do dot star y means $x_1 y_1$, $x_2 y_2$ would be the output and then I take sum of that. So, I have sum of xy, you can see that above there then you have $\sum xy$ this, then is the denominator which you can see at the bottom of both these expressions for w_0 and w_1. I have declared a denominator, which is m times $\sum x^2 - (\sum x)^2$.

So, this is the denominator of both the expressions. The numerator of the x_0 or w_0 expression is $\sum y \sum x^2 - \sum x \sum xy$. Similarly, the numerator of w_1 is m times $\sum xy - \sum x \sum y$. So, these two are calculated and at this point I evaluate w_0 as numerator 0 by denominator and w_1 + numerator 1 by denominator.

So, we will just run this section and here you get a display at the bottom you can see 15.97 which we would have calculated even earlier is w_0 and minus 83.91 is w_1. And using this, you can now calculate all the other terms which we already did in the previous videos. For example, determining T_0 can be done by doing $w_0 + w_1.0$ etcetera. So, if you wish you can find out T_0 as $w_0 + w_1$, you can just call this as x_0 is 0 the location 0.

And if you want T_0 multiplied by x_0 similarly x at the right end is 0.07 you can now find out $T_L = w_0 + w_1 x_L$. we can display both these values T_0 and T_L, we had calculated these values earlier also. sorry I should have put a bracket here. So, you we found this out even earlier 15.97 was the predicted temperature at the left and 10.10 is the predicted temperature at the right.

We could also find out minus K times w_1 which will be $-k\frac{dT}{dx}$ etcetera. So, all these things we have done earlier. but the more interesting thing is to calculate the same w_0 and w_1 in a different way, which is the normal equations approach. So, the normal equations approaches are remember we just did it $X^T XW = X^T Y$, this is the equation that we got in the last couple of videos we had derived this and we want to set this up remember what x is x is the design Matrix.

So, if I come here and we can look at what x looks like, you can see here I will just write it out here. So, you can see one and then 1 and then x_1 and 1 and then x_2 and so on so forth. So, this is just every row here is one data point with 1, 1 appended at the beginning that is what I am doing here a bunch of ones, how many ones there are m ones because m is 6 as you can see here m is 6 and you write six ones and then you put an x on the next column that is what we did.

And then we can evaluate y. y is just the right-hand side just the temperatures that we had given here. I should come here and I will just run this section and we will run this section also. So, when I come here, I have the LHS and let us just remember as you saw in the last video should be a two by two Matrix I had calculated that the last time. So, $X^T X$ is the LHS, remember Prime represents transpose.

So, $X^T X$ is the LHS and $X^T Y$ is there RHS and you are now just solving a two-by-two matrix. Now this thing here LHS is backslash this is a MATLAB way of writing for inverse of LHS multiplied by RHS but this is more efficient. So, it uses internal routines like Gauss elimination etcetera. So, never use in unless really required even within MATLAB especially, when you have large matrices, it is inadvisable and inadvisable to actually use inverse of LHS.

Instead, it is better to use LHS and put a black backslash. Put a backslash RHS it simply means an effective way of calculating LHS multiplied by inverse of LHS multiplied by RHS. So, if you do that you see you get exactly the same results as before. see earlier we had here 15.97 and 83.9143, same thing here 15.97 and minus 83.9143. you might suspect that some other digit later on might not work.

So, I actually calculated this. I calculated the vector w minus, w is my new vector which is the normal one and w_0, w_1 were the old method. So, I did this error and the error was calculated by the norm. Remember I had talked about the norm in the last thing, all it does is it takes the square root of the error 1 square + error 2 square + error 3 square and you see that the difference is somewhere in the 13th a decimal place and this is only because of round off.

Machines do not calculate exactly. if you have a calculator, you can calculate it only to six decimal places or sorry eight decimal places or ten decimal places. Similarly, this has a certain amount of accuracy and this is called round of error basically there is no difference between doing it in the normal equations approach or in the approach that we have done here. So, you can see that but, I will first show you the plots. but there is a major advantage to the normal equations approach which I will show you I have separated out the code in another place but before that let me just show you a plots.

So, we have done that let us come here. So, I am going to plot the results. So, if you plot the results if you just look at the physical plot, this is the physical plot of course it is blown up it would have been a little bit smaller or the Gap would seem to be smaller in the case of the things that I showed in the slides but this is basically the physical plot. Again, you can see the blue dots here are the experimental truth.

And the red line here is the model prediction and the red line and the model is somewhat close to the blue dots, but how close how far is something we cannot easily determine for which you need the R square metric. So, the line here is just plot. So, notice this $\hat{y}$ is our model prediction, I should actually change this to just. So, that it looks better. So, yeah, so, $\hat{y} = w_0 + w_1 x$, which I have written here this is the model prediction and then I have plotted x versus y-hat.

So, notice what is being plotted in a physical plot x versus y or $\hat{y}$ is what is being plotted here. Now we had our R square metric. metric to determine how good or how bad the fit is remembering what R square was $R^2 = 1 - \frac{S_r}{S_t}$, where is S_r, is sum of residuals and S_t is the total residual basically the Gap from the mean.

So, Sr is calculated here y is of course our ground truth or experimental truth and y hat is our model some of that again notice the dot. this dot is there because we want to do $(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2$ S1 and. So, forth $\bar{y}$ is the mean value or the average value and now you can easily calculate this as $S_t = (y - \bar{y})^2$ and $R^2 = 1 - \frac{S_r}{S_t}$.

And you can see this value here we had calculated it. I might have said 0.9174, again that is due to round off I was I might have been taking just three four decimal places. Anyway, you know that this is an around 0.917 and this is a pretty good fit. So, $R^2$is 0.917. Now let us look at the parity plot. Physical plot will show x versus y, but the parity plot which is what we are interested in, will show y experimental versus y model that is y ground growth versus y model.

Let me see you can change this. So, we also plot a 45-degree line which is the ideal model the ideal model would give you exactly $y = \hat{y}$ and we want to plot that too. So, the plot the reference ideal line which is $y - \hat{y}$. this is the model points. Remember as I said for a parity plot, all your plotting is y versus $\hat{y}$ notice this. So, now you get this here y versus $\hat{y}$you can see the scattered plots.

The blue dots here are y versus $\hat{y}$. for example, as we saw earlier here approximately our prediction I should have put the x and y labels, let me quickly put those here. So, x label is we have the ground truth experimental or observed data versus y is the model prediction. So, this is what we notice here. So, what we notice is there is some scatter here which is far off, from the ideal line, the red line on the other hand uh you can see here is basically the ideal line. So, this ideal line it differs from how we are predicting that you can see that at certain points for example this second blue point that you can see on your screen seems far away from the ideal line.

So, you might conceivably say it is an outlier what should be thought of as an outlier and what not is something we will discuss when we come to the probability portions of this course, but as of now it looks far away.

So, this is at least a point worth investigating further and this kind of insight comes only within a parity plot. again, if you had this in 3D it would still look similar because we are only plotting y versus $\hat{y}$ here and y always reminds a single variable and $\hat{y}$ are always still remains a single variable even if you have three-dimensional temperature or whatever plot. So, you can always draw this plot and we will exploit this later on in the course.

Now I want to show you the advantages of using a normal equation approach. So, I am going to take another code. So, this code is a purely normal equations code. I have not done that w_0, w_1, x exact expressions directly here. But let us see how we can exploit it. So, it is the same code as before. So, let us read on the entire code, all I have done is removed the previous portions. you can still see the same results as before 15.97 minus 83.91 and you have these plots here, the parity plot R^2etcetera.

Now the interesting thing is this now suppose I change my features back here and say I do not want just a linear model I want a quadratic model. All I need to say is it does not depend only on 1 and x but I also have an x^2. the moment you have x square the magical thing is, this now so, let us come here just to show you how this works. now let us write x on the right-hand side.

So, now you see it has three columns the first is the constant column you can think of it as x power 0 the next is $x^1$and the third is x^2. So, you can see it has Square the second column. Now LHS y still remains the same LHS has now become $X^T X$. So, now if I write out LHS you will see this is now a 3 cross 3 Matrix. this is just very small do not assume that the last element is 0.

This is just a very small element here. So, this is now a three cross three matrix and RHS is still a 3 cross 1 matrix and if I solve for w this automatically recognizes that now you are going to have three features or three basically coefficients w_0, w_1 and w_2 and the function as you would have seen in the previous videos is basically $w_0 x_1$ or $w_0 . 1 + w_1 . x + w_2 . x^2$.

So, it has automatically calculated these coefficients without me explicitly writing the model because this assumes it is a linear model and that the three features are a constant, x and x^2. As I said you could have made it $\sin x$ and it would have given an entirely different model. Now if we come here and we plot the data, I am going to do the whole thing now. Now notice our model prediction is a curve.

So, the red line here is a curve where that is because it is a quadratic curve. it kind of seems to fit better but we put no extra effort as far as the formulation of the problem is concerned. This is the power of the normal equations approach; we really did not recalculate the coefficients we just calculated all the expressions one shot. You can see that it kind of fits better, however we know that this is not the physics of the problem later on towards the fact end of the course when we come to neural networks we will discuss this in greater detail.

How to determine models when we do not know the underlying physics of the problem. but you can see that the experimental truth is a little bit farther away. but you can see the curve for the model prediction. R^2 is higher 0.96. I make a parity plot. Notice this the parity plot does not show occur because what we are comparing with is just a straight line, we are just comparing with the line that $y = \hat{y}$.

Whereas there is some amount of variance and this seems genuinely lesser than before. but that does not mean this is necessarily a better model, the model was the experimental data points were generated from a linear fit. But just because something looks better does not mean, it is more physical but this would fit the data better right now. Now if we go back and by chance, I add an extra feature x cube I can do that too.

If I do that now let us look at LHS. LHS is now a four by four again these are not 0s. these 0s which appear here are just some very small numbers. So, I am just going to continue here now notice I have four predictions $w_0 + w_1 x + w_2 x^2 + w_3 x^3$. So, we have four sets of coefficients and again a curve here R^2 is approximately the same so, this is not necessarily a better fit.

And you can also see a parity plot, which gives you a maybe again some points are better predicted some points are not so better predicted. but the advantage that I want to emphasize is that you can actually put anything here any so, called non-linear feature here the map is always linear in the W's and it is automatically calculated using this normal equations approach.

For example, I do not know how well this will perform, but just for the sake of argument let us say we put a e^x. So, this model will be $w_0 + w_1 x + w_2 e^x$, what is the best fit let us just run it and check this does not look too bad actually, if you look at it. So, this one looks a little bit better we can also try sine x or cosine x or something of that sort. so, some amount of curvature.

Let me just clear the whole thing up and then redo the run and you can see I mean sign has captured some little bit of it. Now a couple of things I want to emphasize here, is that the normal equations approach is easily generalizable. So, if you are doing assignments, I would recommend that you write a normal equation approach or at least you can copy the code that I have and you can do several assignments using this.

The second thing is notice that it is irrelevant, what we put in the model here. This simply says what uh that coefficient multiplies with. So, basically you are going to get $w_0 + w_1 x + w_2 \sin x$ is the model. So, nothing stops us the terms that we cannot have been terms like sine of w_2 or $w_1 w_2$ those terms we cannot have this is basically linear in w can be non-linear in x.

So, please remember what linearity means. In the next video we will look at a couple of other variants of linear regression. all of them fall roughly within the same framework. We will also look at weighted linear regression which changes the normal equations a little bit, but we will see all that in the next video, thank you.

(Video Ends: 28:26)