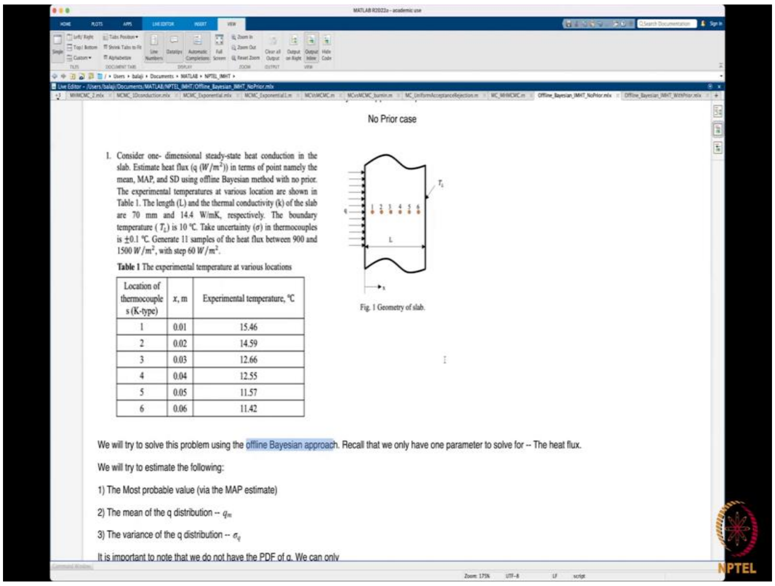
Inverse Methods in Heat Transfer
Prof. Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology - Madras

Lecture - 40
Offline Bayesian Estimation -- MATLAB Demo

(Refer Slide Time: 00:19)



What I will do. Now is to show you a MATLAB code which will again be uploaded in your respective folders by the NPTEL team. But this MATLAB code will just repeat what I just showed in theory in the previous video. So, the idea is very simple we have the same old problem the slab the given data. Now I am going to show without prior one case and I will show you another case with the prime just like we did in the last video.

So, when we try to solve this problem using the offline Bayesian approach as I said the offline Bayesian approach means, you automatically take samples and you just sample the distribution function there. So, we have only one parameter to solve for. So, we are just solving for the heat flux. Now what we will be trying to do is to estimate the following the most probable value which is the peak of the distribution.

We also want the mean of the q distribution, which is given by the expectation if you remember. So, q_1 will basically be the expectation of q via the PDF and of course we can also find out the variance of the q distribution. Now remember we do not actually have originally the full PDF of q. So, what we can do is to we can sample the probability, I will come back to this point in the next video.

(Video Starts: 02:14)

And in fact, you can only sample the numerator of the probability if you remember well the denominator from the Bayesian is actually missing. So, there is a sigma there, there is an integral there, which we do not calculate. So, we do this in steps as I had written in the previous video you calculate you create the original data set basically give all the given data here.

So, for example the physical problem parameters are thermal conductivity is given, the length is given, the temperature at the right end is given. Now we make some measurements and let us. Now include that data if you remember plus minus 1 is the uncertainty in the thermocouples, this I have written as Sigma underscore M where m stands for measurements then the thermocouple locations x and the thermocouple measurements t. So, these are the measurements N_m is the number of measurements that we have made. Now next we actually create the sampling points.

Now this is where we will differ when we come to the Metropolis Hastings Markov chain Monte Carlo method in this case, we are simply creating offline sampling points. So, we create a sample a static sample. So, if you look at q_s you can see, it is simply a linear step between 900 and 1500 we have 11 samples. So, let us just run this code sequentially till now. So, I will just run it just so, that we can deal with this step by step.

So, if I step this you can now see what q_s is. So, the sample q_s in fact we can write this down is 900, 960 up till 1500, x of course is the locations and T is the temperatures that we have measured so far. Now here is the key step, this is basically the Marco chain step. If you want to see it that way basically what we do is we have these ground measurements T_i but for each q. So, if I step through 1 through NS for each q, I will actually generate the actual temperature that we are predicting and find out the differences.

So, let us do that. So, I have just created these matrices if you remember P_1, P_2, P_3 will come back to this point but if you look at the simulation. So, let me write this down here just so, that you can see this. So, we are at the first step, you notice that the actual temperature is these six values, whereas our value of q we are taking the first sample we chose q equal to 900 and that predicted this new set of temperatures, which is what I call the simulated temperature.

Now once you do that you can actually calculate, what the difference between what we predicted and what the truth is. So, you can see so, this is the difference we had seen this in the last video also. Now once you have this, you can actually create S and let me write down S is now 6.286, what this represents is the gap or the sum square between what I predicted which is T Sim and what the ground truth was.

So, this is one single number, this is of course the sum of the errors or sum square of the errors. Now P_1 remember is the numerator, this is the numerator of the PDF of q. So, this is defined by if you remember actually the noise of q. So, minus half times y hat minus y Square, divided by now notice the sigma M Sigma m is the error in each of the thermocouples so that is what we are doing. P_2 let us come here.

Now P_2 is used in order to calculate the mean remember expectation is calculated as whatever value we are interested in in this case q, multiplied by the probability distribution integrated. So, that is what we actually do so, instead of having P sorry instead of just having the PDF you have q times the PDF. Now I have defined a new quantity which I call P_{3a}. So, let me show that, this I did not write in the last video. So, let me write it down here.

Now this is simply q Square. Now why is this useful? We will use the property that expectation of or variance of q is expectation of q square minus expectation of q the whole Square. So, this formula you remember. Now why is this useful? this is useful because when I want to calculate P_3 notice this property P_3, P_3 requires q_1, q_1 requires me to actually find out sum of P_1, P_1 and I do not have it yet I will have to run through the whole iteration in order to calculate sum of P_1.

So, I do not want to wait to do that. So, instead of that I will simply sum I will find out expectation of q Square then I will subtract expectation of q the whole square and that will give me the variance. So, I am using that formula that we had derived earlier on in this course. So, so P_1 is just the PDF this one is the numerator of the mean or the expectation and this is numerator of expectation of q Square.

So, that is what P_1, P_2, P_{3a} are so, these are the quantities which we have calculated. So, if we keep on doing that and we continue let me just stop and I will run it once more. So, if I come here, I have calculated all these quantities. Now you can see if I write down what P_1 is, it is mostly 0 except for a couple of values which is really small values as you saw value, we were calculating it before also.

Similarly, you can see P_2 only two values light up and you can check P_3 also for P_{3a} and only two values light up here as well. After this we calculate the denominator which is the normalizing constant we can. Now calculate the posterior PDF and we can calculate the mean value and we can calculate the variance. So, notice I have calculated the variance as expectation of q square minus expectation of q the whole Square.

So, that is how I have calculated it and sq is the standard deviation in q, which is the square root of the variance. So, that's what is calculated. Now finally I plot the PDF that I have calculated and you can see that here. So, this actually is the PDF. So, you can see point to point, it actually Peaks somewhere if I see this examine this value it Peaks at around 1260 and you give this chart PDF.

Now notice again as I had mentioned in the last video, we are not sampling at many points we are just sampling at these 10, 11 points. Really speaking the PDF could be something really crazy. But what we tend to do is to actually move on and sample it deeper at these points that we have already calculated. Now we can also see the values of q_m, that we obtained I got 1258.7 or 1258.2. I think we got 1258.17 the last time let us just check.

Yeah, it is 1258.17 we can also check sq which is the standard deviation which is about 10.30. So, this these can serve as priors for our next iteration. So, let us do that remember that in the previous case I had now added all of this information. So, now you can see. Now we are going to move between 1200 and 1300. we are going to focus on the region where something important is happening.

So, we will do that, but we will add a prior the prior will be 1258.17 as mu and we had chosen Sigma of 0.1 or 10 of μ_P. Now there was an error in the previous video that I had shown in the values which I will correct right. Now not of μ_P and σ_P but while calculating the PDFs I made a small error, we let it be because the purpose of the last video was just to show you the procedure.

I will correct that the values are not of great importance as far as we are concerned because these are toy problems. we still get reasonable numbers. Now you notice almost exactly everything else is the same, I have included the priors here, μ_P is 1258.17 and σ_P have chosen as 10 percent of μ_P this one should be 1200 to 1300. we have 11 samples that also remains the same, everything else is the same, except we have the small change. this P_1 the basic probability distribution now depends on two parts.

As I had explained in detail in the last video, the first part is the data loss which is pi minus T simulation square and the second part is the loss due to the prior. So, in some sense you can say that a is the data loss and b is prior loss. So, that is this means how far is my model from the current data and this asks the question how far is my model from the prior information that I already have. So, we use all this.

So, for example I mean again I give this kind of detective story example, but let us say there is a detective story and you know you come somewhere a detective comes and sees a murder. Now all the data points towards let us say somebody that you really know I mean you really know some friend or relative of yours and all the data points there, but your prior information is that that person could not do some such thing.

So, these two things are in conflict what prior data says and what current data says, if those who are in a conflict you have to find out the mean between the two, you have to find out some way of optimizing both these. So, remember I said I made an error in the previous video I had missed this factor of two in the previous thing. So, some of the numbers were wrong in this case with the with one extra prior.

So, we have added this prior and there was an error in the previous video. So, let please do in case you are going by those numbers please do not get confused if you calculate them, you will find them off by a little bit anyway, I have corrected that here in the board um again P simulation is exactly what we had derived in the last video minus q by kx plus q L by K plus

PL and everything else remains the Same the only thing that changes here is this power of s this s.

Now becomes a plus b and a is the data loss and b is the prior loss. we will look into this multiple loss terms later on when we come to the machine learning portion also, especially when we come to physics informed neural networks. Again, if you see here the entire code is practically the same, if I compare the code here sorry the code here with the code here the only extra thing is I have an extra line with a and b defined.

So, a is now the data loss and b is now the prior loss and S is a plus b and I have put a factor of half because it is common for both of that and this again remains simply just Sigma n Square. So, include this everything else remains the same and we can now happily run this and you will see we have a much more better-informed network here. Let me rerun this because so now you can see we are getting things in greater detail here.

Of course, the peak is at around 1250, I think I had given 1242 or some such thing or 1240 in my previous video. But that was because of an error in update you can also see that the new mean has changed the new mean is now 1252, instead of the prior which was 1258. So, it has come down a little bit with the new information and the new standard deviation is around 9.66. So, here a little bit surer despite the large uncertainty in the previous things.

So, now one thing we can try playing with is this σ_P, I can reduce σ_P and say I am a little bit more certain about this and I can say σ_P is let us say 10 and if I run that, oh I think I ran it incorrectly, let me just run back once again. So, let us say I gave σ_P of 20. So, you can see what happens if you run σ_P with 20 you are fairly certain about the prior. So, the prior gets weighted very highly and it does not move very much it just gets stuck at around 1258 or 1260.

So, if this is why this is one of the reasons why we take 10 percent of μ_P here and give some leeway so, how we want to compute the posterior and then you get a little bit more detailed procedure. So, we will take this 10 percent as a general rule of thumb in creating posterior distributions and using prior distributions. If you use a lower σ_P you will get sharper posteriors if you use a higher σ_P you will get a little bit more elaborate posterior as you can see here.

So, here we saw a simple coding example using the offline Bayesian. In the next video I will simply give you an idea though not in too much detail about the sort of state-of-the-art method, which is Metropolis Hastings Marco chain Monte Carlo. So, I will see you in the next video, thank you.

(Video Ends: 18:57)