Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 78
Lecture 35A - Introduction to Estimation Theory 1

Let us get started. We just got started into estimation towards end of the last lecture; now for the next few lectures it is going to be very interesting journey. Maybe it is first time that a lot of you are introduced to parameter estimation. So, once again you will be faced with a lot of new terminology and so on. But I always say this that every engineering discipline should have a course on estimation, because ultimately it is not just engineers. In fact, today you know I saw some recent statement by some other leading biology saying that it is better to read a study mathematics then study biology is what at least it says for biology it is always a good idea first to study mathematics and then to biology.

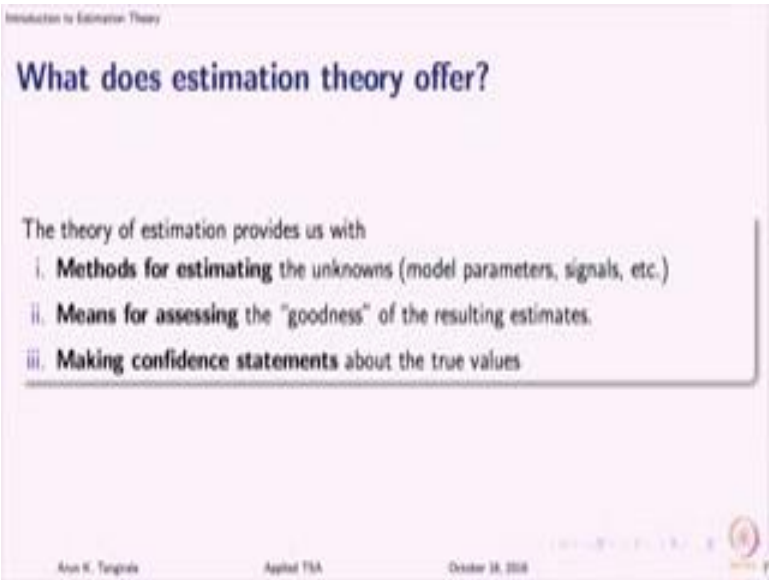
So, that a big change in prospective in attitude and so on. So, not only engineering, but I would say every discipline. If you want to be doing well in your carrier you should go through a course on estimation, because you are going to be faced with data at some point in time which you will process further to estimate something. And estimation theory has a lot to offer lot of interesting exciting ideas enormous number of contributions have come to this field; of course, statisticians initially but then later on from people in econometrics and then gradually from engineers and so on.

So, there is a lot to be learnt we will not be obviously, be able to do full justice to the subject in this course, but what we learn at the salient aspects that are necessary at least for time series analysis. At some point in time I do plan to of an online course at least a twenty hour course on parameter estimation so that everybody is familiar with the basics of a estimation. I wish I had studied this course on I was an undergraduate or even a grad student as a grad student I never went through a formal course on estimation. But instead of that sense I think all of you are very fortunate to be sitting through at least lectures on parameter estimation.

So as I said in the last class what estimation theory has to offer to us is not only methods for estimating, but also methods for assessing the goodness of the estimate, and then making some confidence statements. And this also tells you what you are expected to be doing in estimation. Generally, our thinking is when we sit down to estimate some parameters our job is done as soon as I obtained some estimates. The simplest example is fitting a straight line. If I am fitting a straight line, high school thinking is some data is being given I draw scatter plot fit a straight line that to I do not do anything these stays with a click o f a mouse the software does the fit for me, I do the kind of a (Refer Time: 03:22) and say well here are the slopes and intercept estimates see you live after that is it. And this is what our thinking is when it comes to parameter estimation.

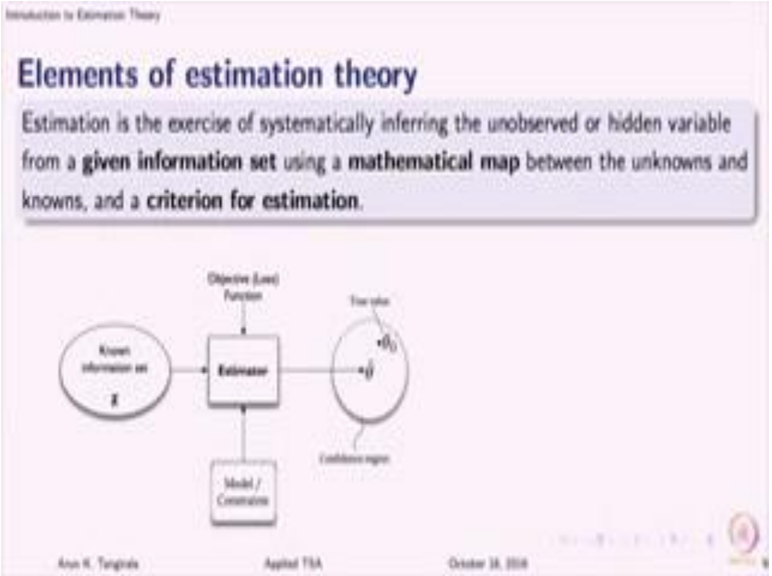
But there is so much more that one as to do. And some of the key things that one has to do in estimation are. Apart from providing what are known as point estimate, so whatever we calculate are called point estimates. We have to be able to say what is the uncertainty shrouding the estimates. In the sense whatever estimate I am providing will be an error. What is the magnitude of error that you aspects to see in this estimates? And as I said that last one making confidence statements.

(Refer Slide Time: 04:09)



Where, what do you think is a region for the truth, where do you think the true value is? So, we will go from the journey of point estimation to so called interval estimation. That is how a journey will go on. So, what will do is right now get introduced to the basics of estimation, and in a basic estimation theory; sorry in a basic estimation exercise what are the elements that participate in any estimation theory you will see these elements.

(Refer Slide Time: 04:40)

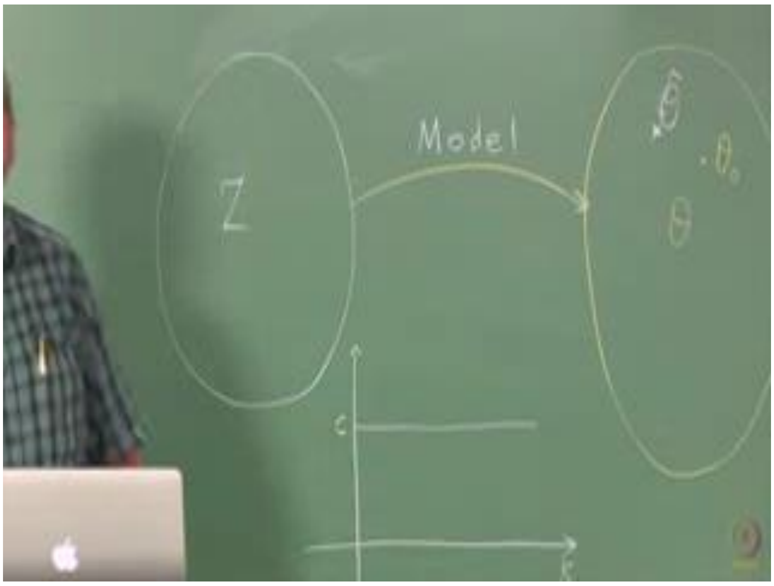


As you see on the screen that is a schematic. The first element in a estimation is data itself. All of us hopefully understand what is estimation? It is a method of systematically inferring the unknowns taking into account the uncertainties in the data, and also taking into account the uncertainties in any other information that you may have. So, if you see on the left we have what is known as known information and this known information should not be thought of us only data, it includes any other prior information you may have of the parameters or of the process everything constitutes this known information. But by and large when we say information the minimal thing that you are looking for us data, by data I mean observations at least in a time series context.

Given these observations which is the food for estimation no data practically no estimation. So, data is the food for estimation and we will come back to that statement a bit later. Apart and once this food is made available this food has to be digested, you need some support system to digest this. Ultimately, if you look at the analogy of we consuming food the role of the digestive system is to convert the food to useful energy; that is what is the role of the digestive system.

Here we have instead of a digested we have estimator what does this estimator do it takes in the food which is a data and also to other pieces of inputs, to other inputs I would say: one input is so called the objective function.

(Refer Slide Time: 06:34)



So, idea is as follows: I am here in the data world or in the information observation world, so this is z and I am supposed to be and let us say this is world of parameters. So, this is your theta, we will denote the parameter vector by theta. And I am supposed to go from here to this world. And I do not know where the truth is, I mean I do not know what the true value of parameter is. And the role of the estimator is let us say the true value is here theta naught, this is the true value. The role of the estimator is to take me to this route.

Now, the unfortunate part is I will never be able to reach the truth from finite samples. So, that is the first better fact about estimation. I may get very close to the truth, but I will never be able to reach the truth unless in some very academic situations. In all practical situations I will always be away from the truth. So, ultimately I will probably be around somewhere here. So, let us call this as this is the estimate theta hat; all this notations that we are using are more or less conventions that are used in estimation theory.

So, I end up here. And of course, this where I end up depends a lot on something; that is what we are going to discover. What is it that is going to determine where I end up in the parameter world? The truth is only 1 for now until we hit Bayesian estimation. We will assume that the truth is fixed. So, you may wonder what is this then; this is a space of all possible values. Theta naught is fixed that is the truth and estimator will take me somewhere, whereas theta hat this destination is unknown to me.

So, I am going from the city of data to a city of unknowns and I do not know where I am supposed to go. Now it sounds very philosophical like this guy he was broken up in the movie and we say ticket [FL] I do not know where ever we go and so on. So, something like that, so I do not know wherever the estimator takes me I will go. So, I do not know unknown destination, I do not know the truth but yet I aspect the estimator take me close to the truth. So, that is how the bit we are to begin with. Yet the problem is solved. Now this estimator actually establishes facilitates this theta hat for me and what connects these two worlds is the model. In some sense it establishes a connection between the known and unknowns. Unless this bridge exists I will never be able to reach that other city.

So the role of the model, when I say model here it could be anything; it could be any equation whatever it is, it is some kind of mapping between the knowns which is a data and unknowns which is a theta. Now once the bridge as been constructed as you can see there is a model there, there may be other constraints also accompanying this model that you know you cannot go anywhere this bridge may only end up at some places for example, when you are estimating let say average temperature or some other averages some other parameters that may be constraints are parameters are non negative valued and so on. For example, if you are estimating spectral densities, we know spectral densities are non negative valued so we should guarantee, we should incorporate that information into the estimation exercise and so on. So, these are all called constraints. It is a part of your model.

So, the model establishes a bridge between the knowns and the unknowns. And now it is we need one more person. Once we have gone into this unknown city we need a guide, we need a guide that will take us. If you think of it you gone to the guide and say I do not know where I am suppose to go there is something called truth, but take me a very close to the truth. Do not try this out in practice its human beings, person I think you very weird, but in estimation you are allow to do that.

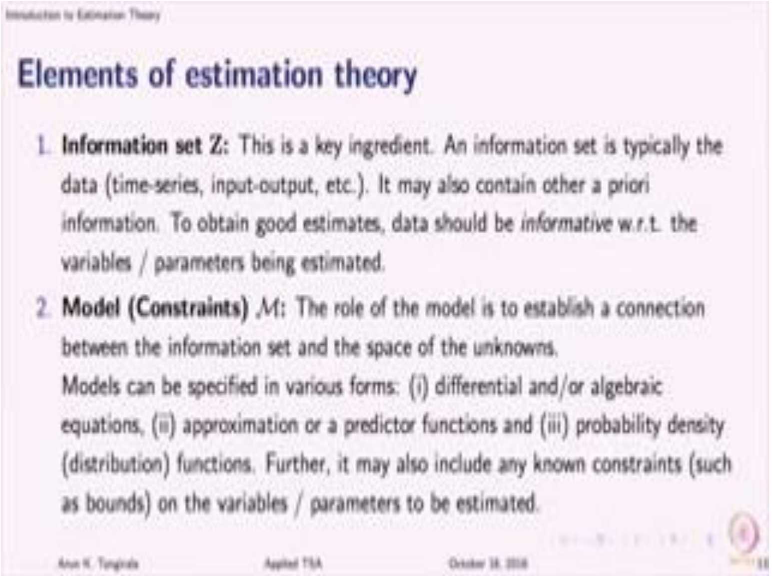
So, this guide that will take you very close to the truth is your objective function, itself that is now you want to reach very close to theta naught. What is close you have to define mathematically. You need a metric for proximity; you need a metric for distance sorry, or closeness. And that metric can be Euclidean norm, one norm, Mahalanobis whatever measure of distance that you have you can use. And there are many more variances, but right now you can think of squared distance is being a very common metric for the distance itself.

So, the objective function will guide me so in such a way that I reach as close to truth as possible. Now at the moment it sounds a bit strange because I do not know the truth how will the objective function or how will the optimization problem. You can now see that there is an optimization problem, because once you get into this there are multiple solutions and among the multiple solutions the optimizer is suppose to pick the one that is optimal in some sense without knowing the truth. So, at this stage thinks are a bit ill posed because I do not know the truth how can it actually take me very close to the truth.

But, when we go through a simple example thinks will become lot more clearer. That is what is actually practiced is slightly different from what I have just said, but ultimately achieve this job of taking us close to the truth. So, these are the basic elements of the estimation exercise. And when you put together everything that is when you put together the objective function the model and any uncertainty assumptions that you have on the data everything and then you solve the optimization problem you end up with some kind of a formula to be crude some kind of an expression some formula to which you feed the data outcomes theta hat; that formula that expression is called the Estimator. So, it is like your digester it digests everything and then produces what you need. And, this estimator can have any form; it can have a very simple form to the most complicated form may be sometimes close form expression also do not exist.

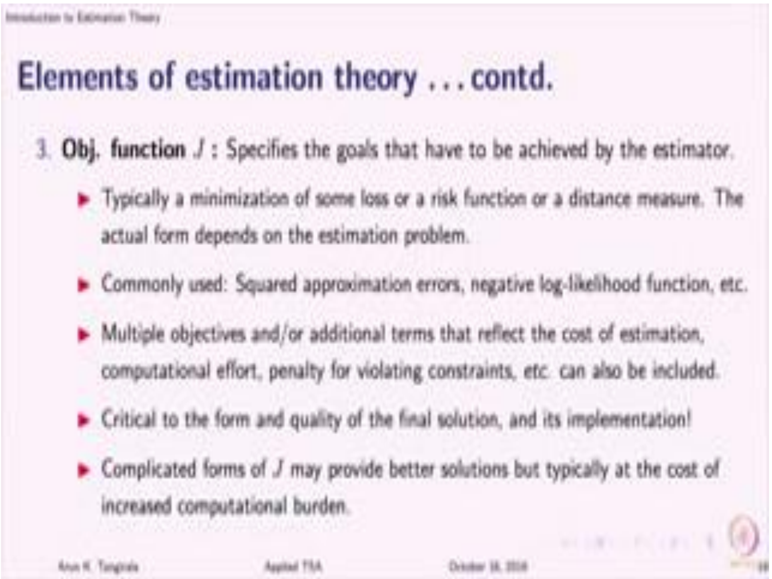
And, on the other hand you also have this privilege of saying I want to have a formula that looks like this, you can pre impose. You can say that ultimately I want an estimator to be linear for example, or I want the estimator to be locally linear whatever you can presuppose and then figure out what expression you get and so on. So, there are several things that you can do a lot of variations come in. What we will do is will go over a simple exercise that will give you a feel of what is typically involved in an estimation exercise. And this is perhaps the simplest example that you can thing off and we will keep revisiting that example time and again.

(Refer Slide Time: 14:21)



So, I am going to skip lot of this slides I have talked about what we mean by information set. What we mean by model.

(Refer Slide Time: 14:23)



And, I briefly talked about objective function. In the example that is are coming I will show you what impact the objective function can have on the nature of the estimation formula that you end up with; on the form of the estimator that you end up with. So, I have already said that some of the commonly used objective functions are squared distances. We will of course, also look at negative log like likelihood function and so on.

Of course, also said just now that estimator is essentially that mathematical device, it is some formula or expression that computes the estimates using the information the model and the objective function. And you should also now start equating the term estimator with what is known as filter. Ultimately what are you doing, you are taking data is a raw material and you are really extracting the juice out of it and filtering out what is necessary and living the rest. What you are filtering out is theta, whatever you want.

When we talk of estimation problems please do not think of only parameter estimation, a bit later we will talk about different types of estimation problems. There are also signal estimation problems which are very classical problems, and then there is a state estimation problem; that is another hot area is being there for about 50- 60 years. So, there are different types of parameter estimate sorry estimation problems of which parameters estimation is one of the oldest once that is have been studied.

So, we will talk about this filtering, prediction and so on a bit later. Let us get to the example and this example is that of estimating a constant signal embedded in noise. There are two different ways in which this example can be thought of: from a signal processing view point we can say that there is a signal which is constant and I am measuring this signal. So, the here is the signal this is time let us say- theoretically it is constant; I know its constant because I have ensure that the process conditions are such I am suppose to get a constant, but when I measure it I have measurement noise corrupting this therefore what I have with me is some kind of noisy reading of this and from this noisy reading that is the measurement I am suppose to estimate that constant. It is a classic problem that is studied in signal processing.

The statistical prospective of the same problems is estimation of mean of a random signal. We can say that I have a random signal and I want to estimate the mean. Both lead to the same thing, but look at it in this simple example we have present a two different types of estimation problems; what are those two problems? The same problems we have given at one flavor from a signal processing view point and another flavor from a statistical view point. From a signal processing view point what type of estimation problems are we are studying? Signal estimation we are there is a constant signal that I am estimating, from statistic viewpoints parameter estimation.

So, you see signal estimation and parameter estimation are equivalent; in general also it is true. But it is the nature of the problem and the origin of this problems that have actually resulted in lot of different kinds of literature that you see methods have evolved, but ultimately they need like all the reserve meet at the sea all this problems meet and they go and meet at the confluence where they merge with the big estimation theory.

So, signal estimation problems and parameter estimation problem are equivalent. There is a third prospective, this is from a state estimation view point. For those of you who are not familiar with state space modeling a state is some kind of an unobserved variable that gives you an idea or gives you a complete picture of the system dynamics but it is not observed, something else is observed but not the state. That something else contains the information about the state. So, from a state estimation view point what you are not observing is a constant signal; that is a state. And you are observing is a measurement.

So, this is also a state estimation view point, you can turn this into a state estimation problem. And that is the beauty of this problem. It is so simple but it has the flavors of all different types of estimation problems; parameter estimation, signal estimation, state estimation. And then it also serves us in illustrating the impact of objective function the change of objective function and how the estimate itself changes.