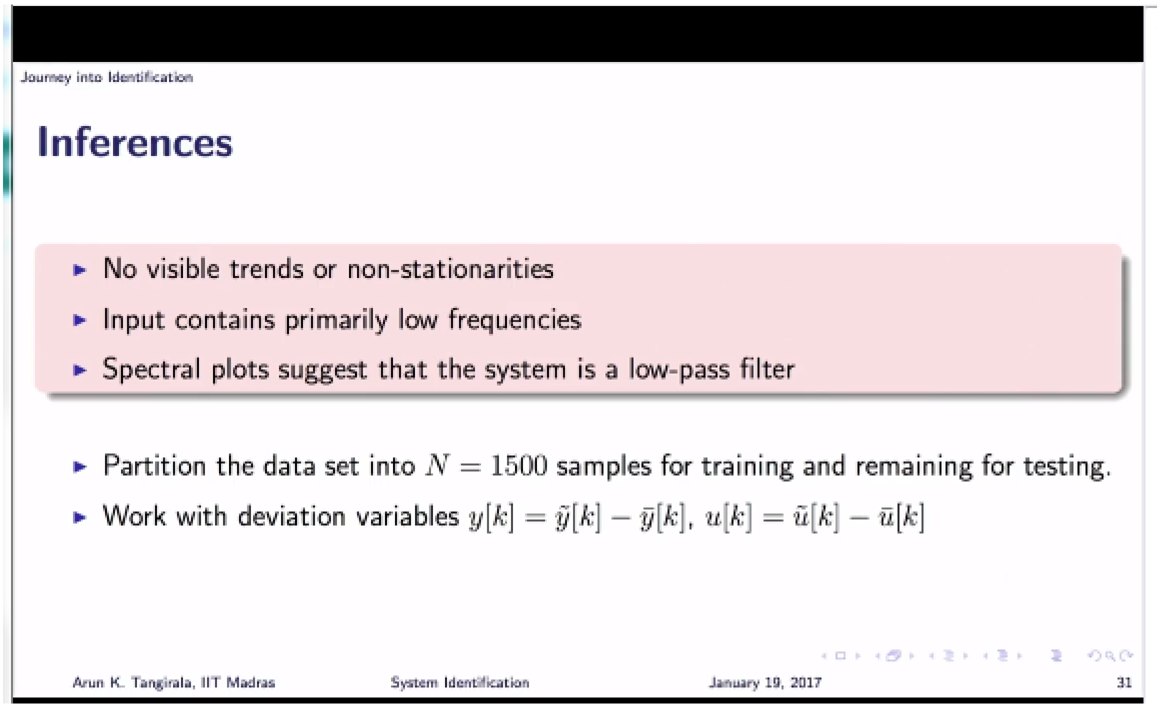
CH5230: System Identification

Journey into Identification Case Studies 10

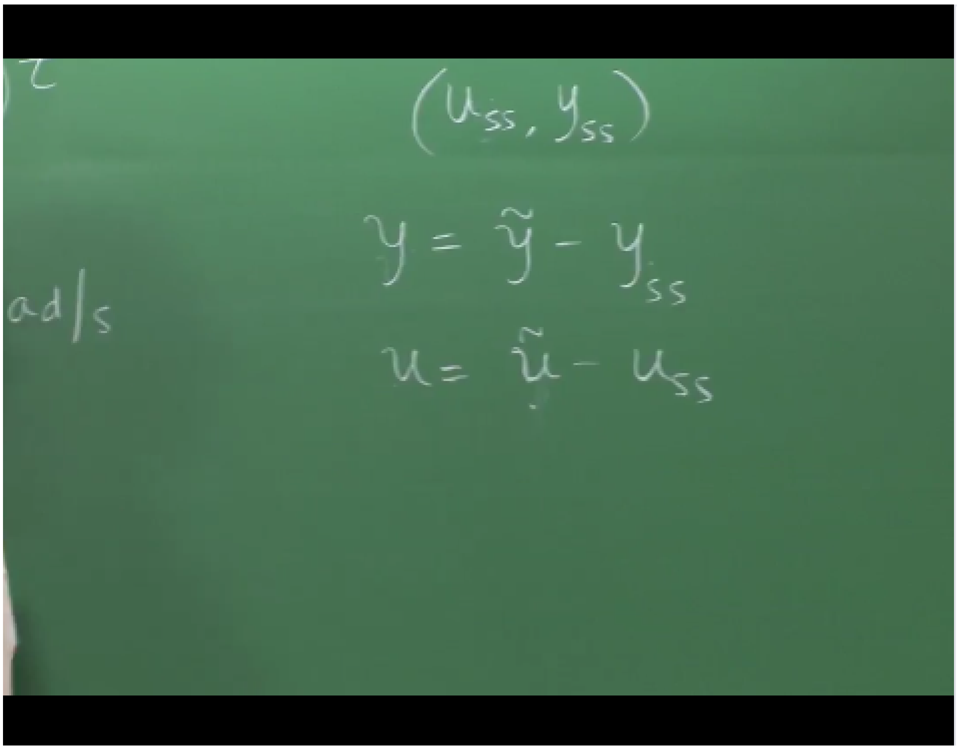
So we have already talked about these inferences. The one thing that we will do before we march ahead is that, now I have 1500 observations from the experiment.

(Reference Slide Time 00:27)



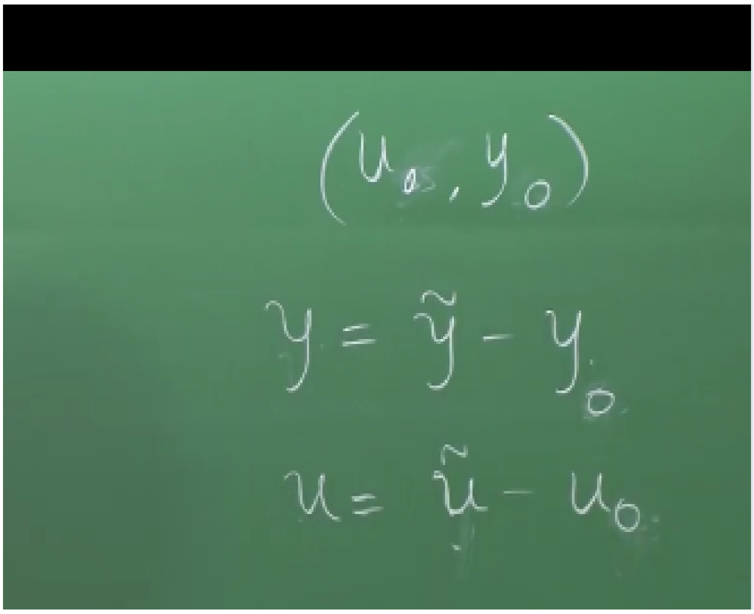
I say samples but, you know, it's better to use a word observations, because in statistics a sample itself is a collection of observations. So that's why it can be confusing to a few. I partition these 1500 observations into two data sets, it's like I have a big question bank, question and answer bank as an instructor, and I want to use a portion of that for training the students, assignment problems and so on, and another portions for testing for exams and so on. Exactly the same story here, I'm going to use a part of this data for training the model, and then another part for cross validating the model. How should I partition it? How many observations should I keep in the training and in the test and so on, there are no definitive answers to those questions. But one guideline that you can keep in mind is that, you should by and large be fair to the model, that is, try and include features in the training data set over which you want the model to learn very well and has similar features in the test data. Maybe some additional features but not too much that is absolutely not seen within the model at all. So, for example, you should not systematically include only low frequency content in the training data, and then have all the medium and other frequencies in the test data. Then your model can fail miserably because your model has not seen, right? It's like, I have trained students how to answer in detail, given sufficient time how they answer, but I have not trained the students in answering quickly, which the coaching centres actually are very good at, right? They will train you how to answer quickly, but unfortunately, they don't necessarily train you in thinking, that you have to inculcate because they have their own agenda. As an individual, you will have to go back to your desk and say, "Well, you know, now I have a lot of time, let me develop my conceptual abilities." There is no room for that in the training session in a coaching centre. But here there's nothing like that. You have to be fair to the model and judiciously divide the data into training invested[2:44] and there is a body of literature that gives you more guidelines and so on. For now, we will keep things simple. That is one point. The second point is, remember, I said, we are going to build a model that explains changes around an operating point. Therefore, it's good to work with what are known as deviation variables. That is, the process is at some steady state. If you were to say, let's call this as u ss and y ss. So this is the steady state operating point, which I may know or I may not know, it all depends. But what we are going to work with is, in fact, what we call as y later on would be if y tilde is absolute output, then I'm going to work with y tilde minus y ss and likewise for u, u tilde is absolute value. And I'm going to work with the deviation from steady state.

(Refer Slide Time 3:45)



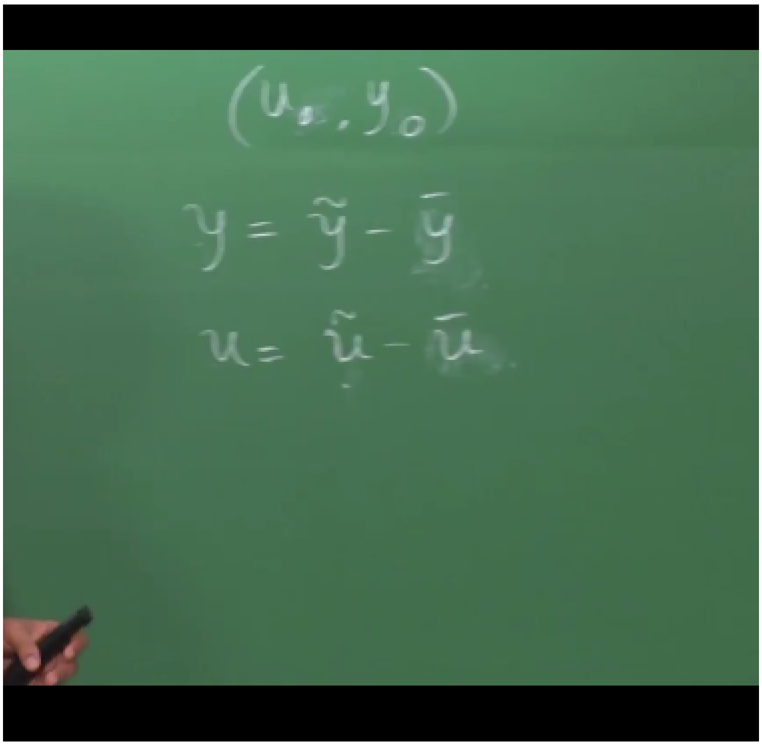
Now, there are many different approaches to this steady state. In fact, a more general way of saying this is, I will use some reference point need not be steady state, u 0, y 0. And then the choice is yours what you use as this reference point.

(Refer Slide Time 4:07)



You can use it as a steady, you can use steady state as a reference point if that's what you're going to use a model for. Or you can use some other reference point and so on. And sometimes, you may know this reference point ahead of time. And sometimes, some other times, you may not. Here, we will assume that the steady state is not known and we'll just replace with a simple estimate, as this sample mean of the output and input that I have obtained.

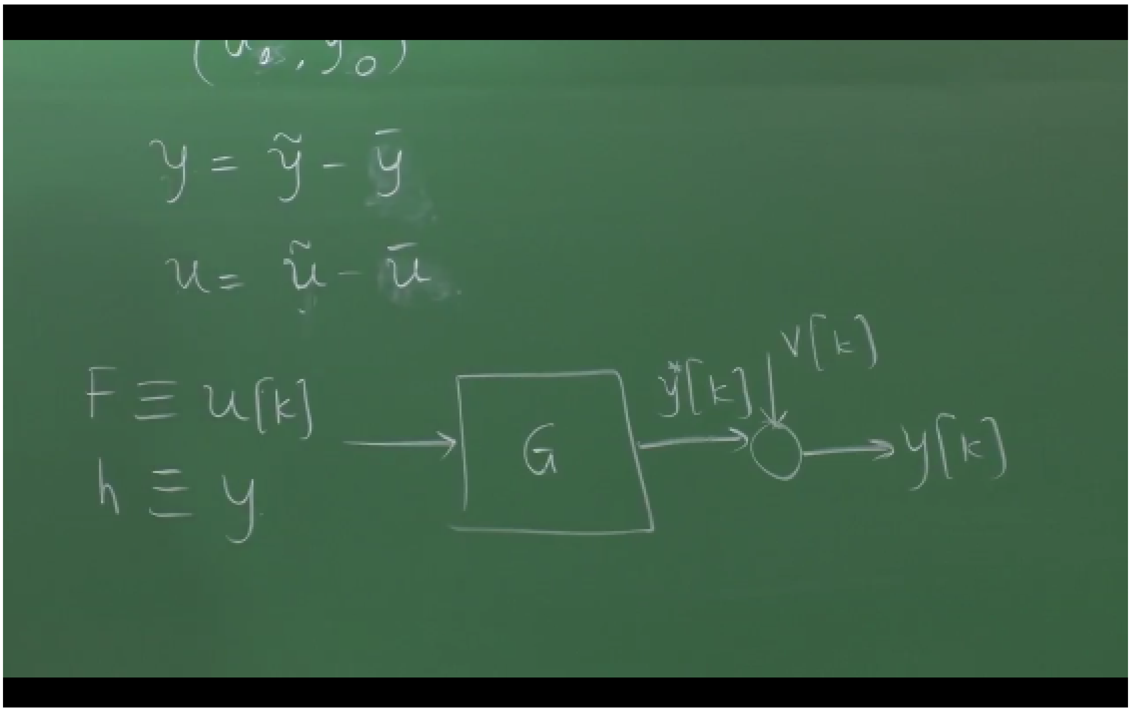
(Refer Slide Time 4:38)



Essentially, the average of the input is my reference point, average of the, likewise, average of the output is my reference point for y. It's very simple, but let me tell you, again, one can actually go deeper into this and there are some open ended questions here. How to estimate these optimally, and so on? Can I estimate y 0 and u 0 together with the model parameters and so on? Let's not get into those complications. Alright. So, I hope clear. We are partition the data, we are going to work with deviation variables.

What is the next step? What is the next logical step? I want to build a model, right? So, we need to have an idea of what we mean by model. In terms of the symbolic quantities I have flow which is denoted by u, and level which is denoted by, assume these are deviation variables now, by y. So we want a model that relates this input uk to yk. But remember, we are going to have only a corrupted version of y star. So, this is the G.

(Refer Slide Time 6:00)



What I'm given is the input and the measured output data. It's a very, there's a very important point to note here. We are seeing only the output is known with error, input is known free of error. This is the classical case in System Identification. There is another large body of literature which is present in identification, econometrics, biomedical engineering and a lot of other domains where the input is not known, I mean, is known only with error. It's not known accurately. You would not have the true input with you. You would have only a measured input with you. That's the case. Suppose I want to build a model between temperature and relative humidity for the atmosphere, right? It cannot be cast in the classical system identification setting because both, which would be the input and which would be the output, by the way, if you were to build a model between temperature and relative humidity? Thermodynamics, physics, whatever, what does it tell you? You're not so sure. That's not good.

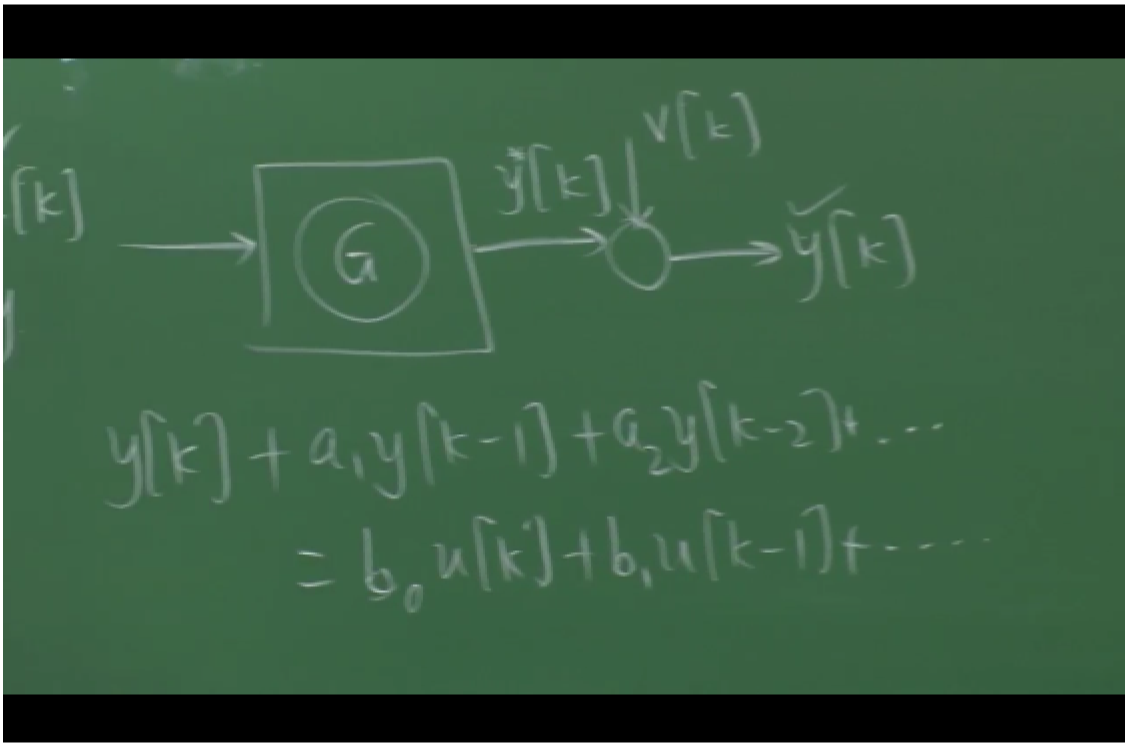
Temperature--

Temperature would be 1. Okay. I thought you're having temperature just listening to this question. Okay. So, correct. So, we would think of temperature as being the input, as being the cause, and the relative humidity being the effect. Correct? Now, unlike here, I can't go and change the temperature.

Thankfully, nobody can change the atmospheric temperature. It has to change on its own. So I only observe the temperature changes, and I only observe the relative humidity changes, and then build the model. The moment I say, I observe, both the input and output are known an error. That is called the EIV case, errors-in-variables case, that's a different kind of framework, and we don't talk about it in this course, 99.99% we don't talk about it, I do mention it briefly occasionally. That's about it.

So, remember that we assume throughout the course that the input is known without error, unless stated otherwise. Whereas, the output is always assumed to be known with error. Now the goal for us is given u and y, we want to get a model for G. And in doing so, we have to admit the fact that there is stochastic term v. Right? So, the focus is always on G, while acknowledging that there is a stochastic term. Now, what do we mean by a model for G? Some kind of a mathematical equation, right? Because this is a discrete time system, we think of a difference equation, because we are building a dynamic model. If it was a continuous-time system that we're looking at, then we would think of a differential equation. But here, it's a discrete time system. So we hope to develop eventually some kind of a difference equation for, a description for G. But to get to that, remember, a difference equation, suppose I'm thinking of this kind of a difference equation, y k plus a 1, y k minus 1plus maybe a 2, y k minus 2 and so on. I don't know how many terms to include. Likewise, on the right hand side, there could be b0 u k, if I'm looking at linear systems, b1 u k minus 1, I don't know.

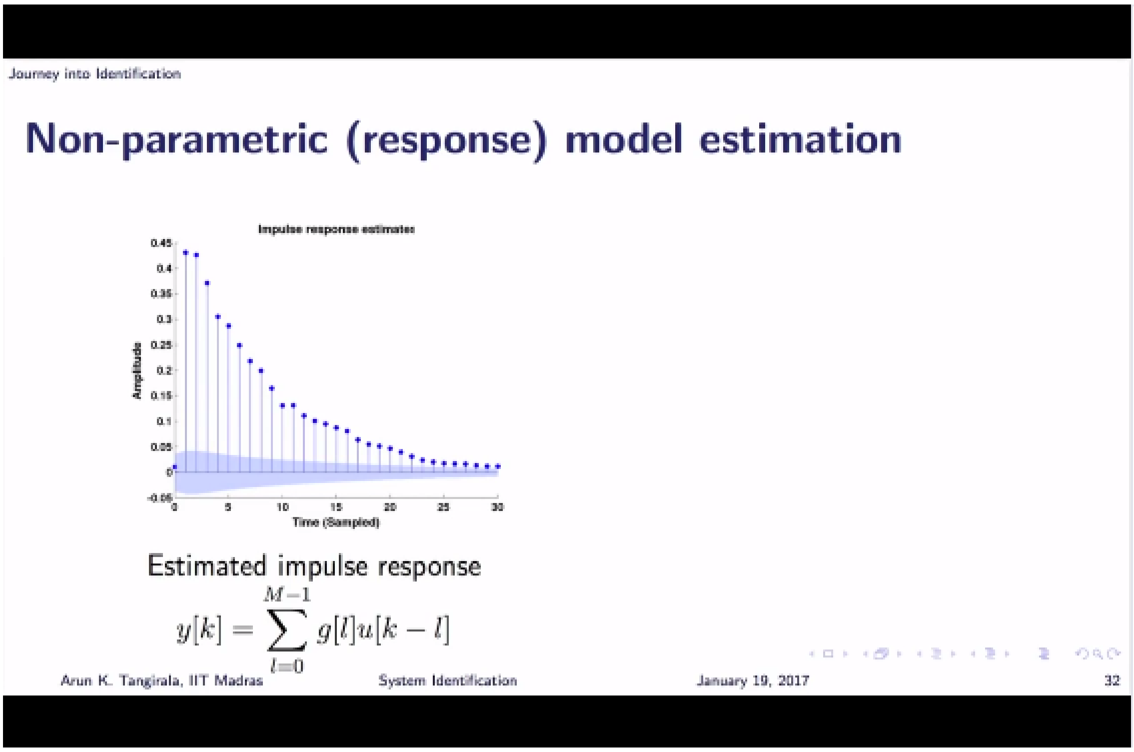
(Refer Slide Time 9:44)



I don't know how many times to include. In fact, on the right, the system may have a delay, we don't know. Then I don't know what's the first term on the right hand side. So, there are several decisions that I have to make before I go ahead and fit a model. One approach is to simply assume some orders, some delay and proceed, but that's not at all recommended. That is the blind users approach.We will take the enlightened users approach. Okay. As a learned users approach, which is to systematically discover these pieces of information and come up with a very good model. So the first step is, to do what is known as a non-parametric analysis, where without assuming any of this. I don't care really what is the order, what we mean by order is, how many past outputs are involved in the difference equation, or what is the delay, or how many terms on the right hand side, that means, input dynamics. Nothing, I don't make any assumption. I just do a plain analysis of the system only assuming that the system is linear and time invariant, that's all.

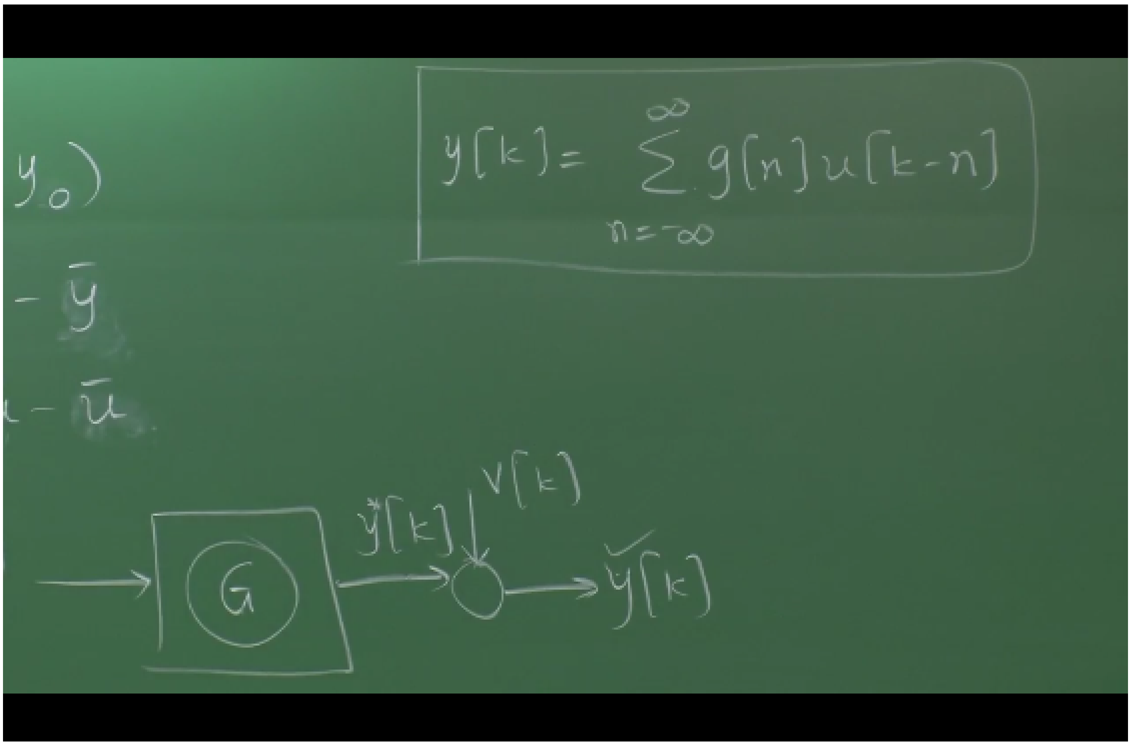
And the first step in such an analysis is the estimation of impulse responses. Why?

(Refer Slide Time 11:05)



Because LTI systems, linear time-invariant systems, for those of you who have had some exposure to linear systems theory, you may recall that LTI systems are described by the so called convolution equations. We will, of course, go through this convolution equation in detail next week or maybe hopefully tomorrow onwards.

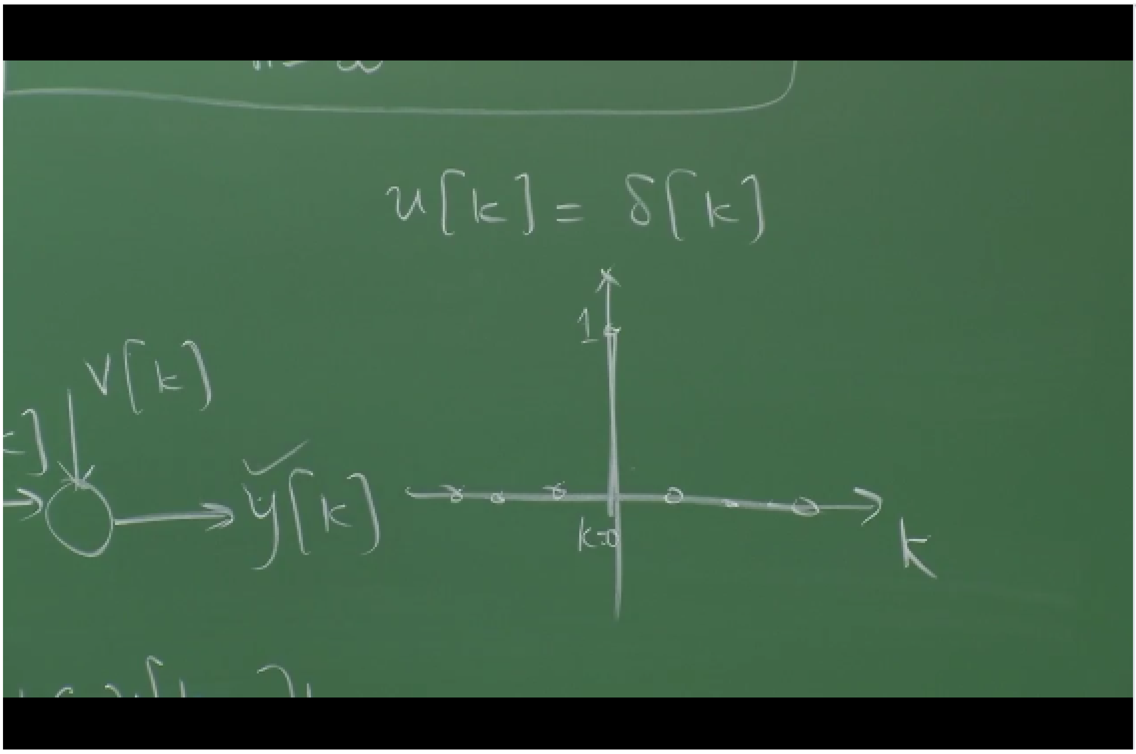
(Refer Slide Time 11:37)



But this is the, you can say, kind of, the mother of all equations that describe linear time-invariant systems. This is called the discrete time convolution equation. What does it assume or what does it say? It says that the response of a linear time-invariant system is a linear combination of the past, present, and future inputs. And those Gs that you see there are called impulse response coefficients. Why? Because if you excite the system with an impulse, a discrete time impulse, not a continuous-time, not a direct one but a chronicle. Then, so imagine that in place of u, you have a chronicle. This is how it looks like. It is 1, assume it unit impulse. It's 1 at k equals 0 and 0 elsewhere to the left and the right of the origin.

So imagine replacing, substituting for uk with this delta. What you'll be left with is a impulse response coefficient G k.

(Refer Slide Time 12:47)



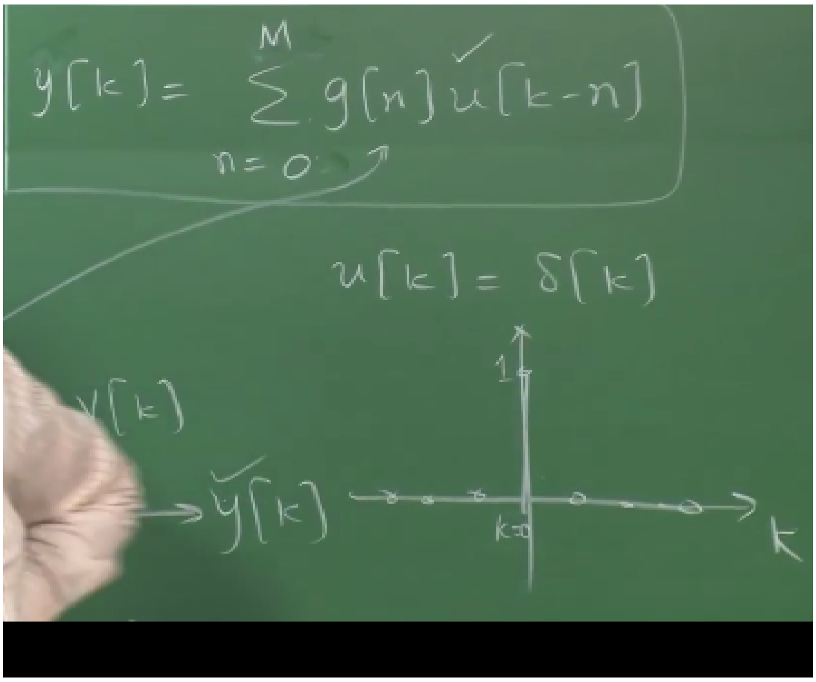
That's why the Gs are called impulse response coefficient. That is, if I punch the system and withdraw, the response that I see is impulse response. The input may last only for one instant, the input lasts only for one instant, but the response lasts for some time. Correct?

Imagine that there is a pendulum. And you just punch it. Imagine that you're able to do it like an impulse. You have withdrawn the force, but the pendulum continues to respond for a while. That is your impulse response, right? So, likewise, even, I mean, you must have had your encounters with monkeys on campus. The monkey may not stop its response, right? That is the impulse response of the monkey. But depends, that's a nonlinear system anyway. Let's not get into it. So there are many, many examples that one can give. Now, why are we turning to the [13:50 inaudible] of an equation, because this equation does not make any assumption on the [day 13:56 inaudible] on the order, nothing. It only assumes that the system is LTI. So, it's a good idea first to get, to look at this impulse response coefficients. And obviously, when you look at it from a system identification perspective, what you are given, this convolution equation can be viewed by many people in many different ways.

For a person in System Identification y and u are given, right? You have to be very careful here. What I have actually written is, for, let me be more precise here. What's actually, what I have written is for y star, for G not for y. This is for G. That means, it actually relates u to y star, but for now ignore noise. Okay. Ignore that v is present. This is just a preliminary analysis. Therefore, we shall ignore v and assume y star to be y. Anyway, this is not the final model that we are going to work with. Okay. Or you can say, it's a prediction, doesn't matter. For now, assume that there is no noise, that is how you treat it as-- we'll make a more technically correct statement later on.

No you are given y and u, Input and output data. What remains to be known? Impulse response coefficients, right? What do I do with those impulse response coefficients? Do I go and have a party? Or it's some laddu that I will eat. It is a laddu. As far as LTI system is concerned, it is a big laddu. It's a big pressure. Because once I know the impulse response coefficients, I can predict the response of the system to any other input, right? At the moment, I don't know G, I know u, so I'm out to estimate G. The day I know G, the point at which I know G, then I can use this model now to predict the response of the system to any other arbitrary input u. That is the idea with which we want to estimate G. Not only that, from the estimate of the impulse response coefficients, I can also figure out the delay. But before I do that, I have to make some modifications here for practical reasons, because one, it has infinite number of unknowns which I cannot estimate.

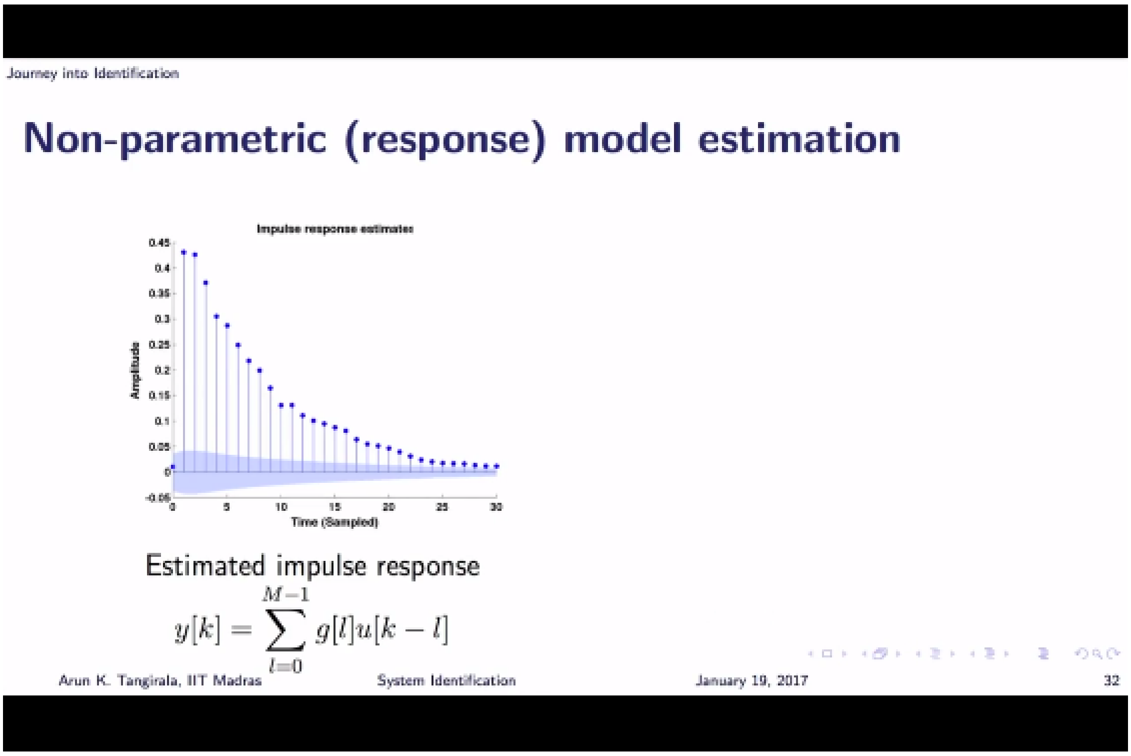
(Refer Slide Time 16:21)



And two, I have cut down also restricted the lower subscript to 0 assuming that the future inputs would not affect the present output. That means, we are assuming causality. We are assuming causal systems. The flow change at a later time, why should I expect that to affect the level now? Under open-loop conditions, that is not true at all, there is no connection between that.

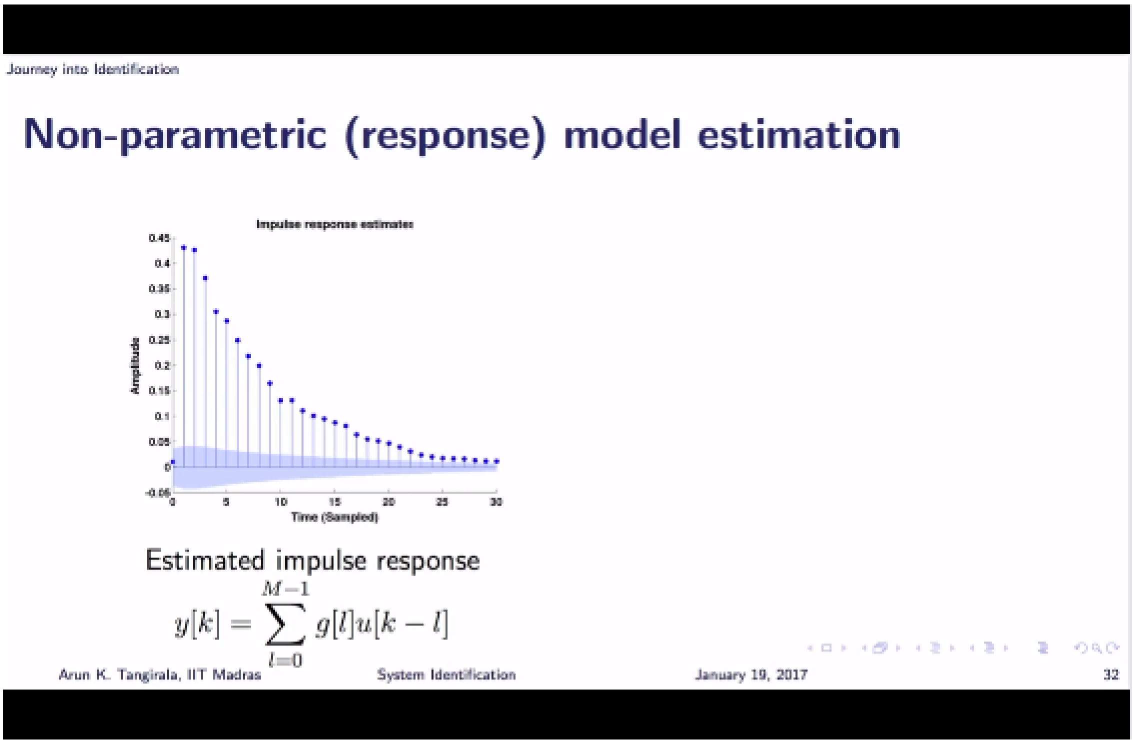
So assuming first causality, and two, for practical reasons we truncate this infinitely long convolution model to a finite impulse response model. These are called FIR models. Now I estimate this coefficients using some estimation algorithm, I've used a least-squares approach and the plot is shown here.

(Refer Slide Time 17:12)



Just spend one or two minutes and then we'll continue tomorrow. So the plot of impulse response coefficients is shown. In fact, when I do this, you remember, I'm making an approximation. You should remember that. What is a plot showing you? It is showing the estimates of G.

(Refer Slide Time 17:29)



It's essentially estimated impulse response coefficients. The y axis shows the value of the impulse response coefficient and x axis is called the lag or time. There are two things that you should notice. One, let me see if you can quickly pick the first one.

What do you notice of the impulse response coefficients from the start? What can you say by looking at the impulse response coefficients as to when does it actually start responding significantly? What do you think? Second, that is actually at time one, our time count starts from zero. The second impulse response coefficient is the most significant one. Now you have to understand this, we have estimated the impulse response coefficients. We have not experimentally generated this. Our input to the experiment was not an impulse. I could have given an impulse, but we'll talk about that later on. What I have done is, I have given a PRBS. From the PRBS, I have recovered, response to the PRBS, I have recovered the impulse response coefficients. And this plot tells me that the first impulse response coefficient that is significant occurs at lag one or time one, which means that, if I give an impulse to the system at time 0, it will not respond immediately. It will take one sampling interval to generate the first response. That means, we say, the system has one unit delay. Now what you should think is, what you should think is, why there should be a unit delay in the System, it's a liquid level system, it should respond immediately. It's not that I'm giving flow changes in Kashmir and I'm recording at Kanyakumari, right? There is not such long pipes that I'm introducing the flow change elsewhere. I'm just introducing the change here and the level is being recorded, so it should immediately respond. Why is there one unit delay you have to think, okay? And secondly, you see a band there. That band, again, is is called the significance band. Remember we talked about statistical significance yesterday. An estimate that falls within that band is considered statistically insignificant. And that is how we have come to the conclusion that the first impulse response coefficient is 0. That means, the estimate is statistically insignificant. That's how we are able to say, from lag one onwards you see a significant coefficient, right? And the third point you should observe is, the impulse response decays. What does it tell you about the system? System's characteristics, stability, right? If it was an unstable system, I ask just one question and it’ll just keeps screaming forever. Or maybe it can explode also, right? Like there are many people whom you just want to ask, what's your name or what's the time and so on. And then they keep talking to you forever in life. Well, it depends whether you want them to. But anyway, this system doesn't, and obviously, it is a stable system. So look at…