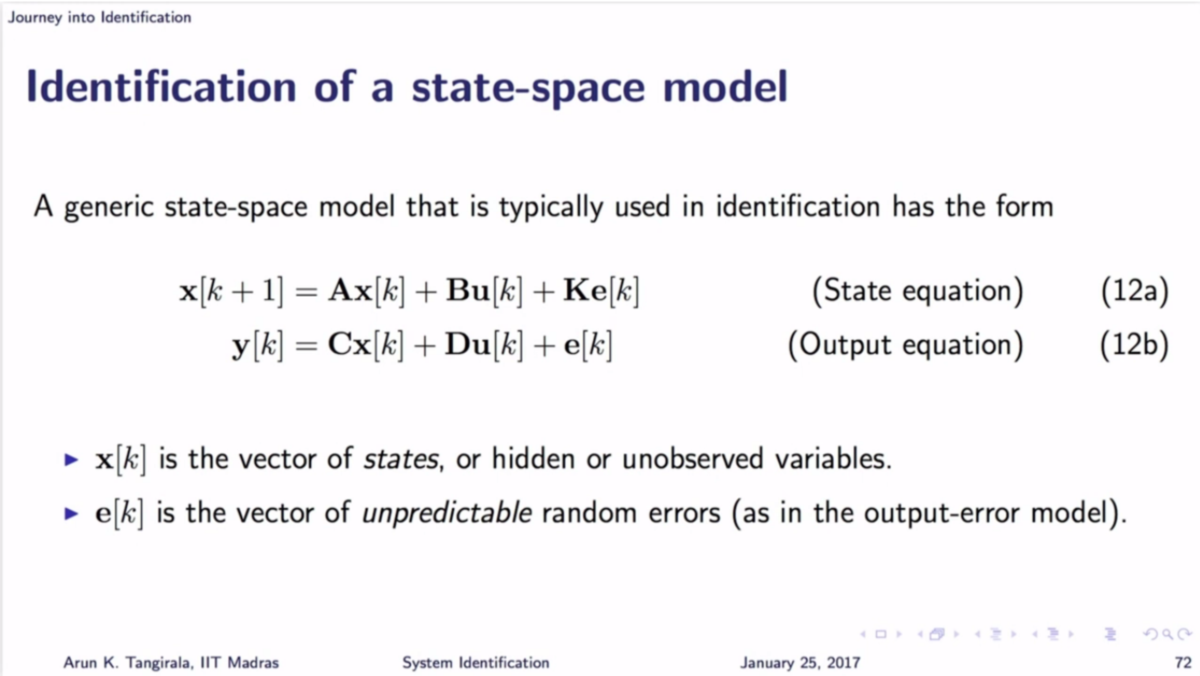
CH5230: System Identification

Journey into Identification 17

Okay.So let's move on quickly and wind-up this case study with a discussion on Identification of the State-Space Model. Now I know that I've not introduced state-space models to you, but just as a curtain-raiser a typical state-space model, generic state-space model will look like what you see on the screen. The state-space modelrelies on the notion of what is known as states. And when the time comes I will explain to you what is the notion of a state and describe many different viewpoints of a state. But for now think of states as some hidden variables that you are not able to observe directly, but there are those variables that describe the process. Okay. And there are many different ways of choosing the state.

(Refer Slide Time: 01:01)



If you don't follow what a state is remember that we use the term steady state or we say state of mind, we say, and "That person's state of mind doesn't look good today." Right? What do we mean there that means there is some variable qualitative variable that is describing the characteristics of the mind at that time? You are not observing the state correctly you're observing the consequence of that. And the consequence or you can say the manifestation of state is in the form of measurement y. So what we observe is y?What we are typically unable to observe is x. Sometimes it can happen depending on how the measurements and states are related. For example if Cis 1. Let us take the case where C is 1 or identity and D is 0 and let us say, let us ignore the noise for the moment then y becomes x. It is possible that you are observing state itself, but that those are, those situations are typically very small in number. Generally you would observe a function of the state, a leaner function of the state, scale version I would say.So the ABC D take care of the deterministic part and the K which we will learn later on as a Coleman game and e together take care of the stochastic part.Together with partly with one of the other matrixes there. So for now let us say that for the liquid level case study I would like to identify a state-space model like this. Why I won't identify state-space model, what are the advantages, we shall discuss later on. This is yet another description of the system which offers some advantages. We will not discuss the advantageous part, right now. Now, what is the state-space identification problem? The same story, whatever is given to me has not changed. What is given to me in an identification problem?What do I have with me? Data that's all.That doesn't change.The food does not change. It is still the input output data. So in these equations here, what I know are u and y. In a general situation u and y are vectors, but in our case u and y are scalars. Right? And the goal is to. Now look at how many things we have to estimate in a state-space model A B. What else? C D E. I mean almost all English alphabets looks like, but ABCD E K and then straight away jump to x. All right. That's a lot. That's a lot. The other thing that we should remember is that of identifiability. We spoke of identifiability,right? And I have mentioned at that point in time that state-space models are not identifiable. Which means there is no unique state-space representation? You should not think that when I say state-space models are not identifiable you cannot estimate them, no.What this means is that doesn't exist a unique answer. What does this mean? I can actually if I have a state-space model AB C D K, I can create another state-space model by a change of states. Right? The simplest is to scale x by factor of alpha. And I can come up with another solution. This clearly tells me that the state-space model is not unique. Okay. We'll keep aside the unique part we just want some state-space model. Doesn't matter, I'm not searching for a unique answer. So as we said the goal is to estimate ABC D K and x. We may not worry so much about x.In fact, I am more worried about ABC D and K.In that process anyway x is also estimated but will not talk about it. All right.So, one of the major decisions that one has to make in identifying the state-space model and that is only decision that you have to make is how many states, right? If you take a country like India we know how many states it has. How many do we have?

Twenty eight.

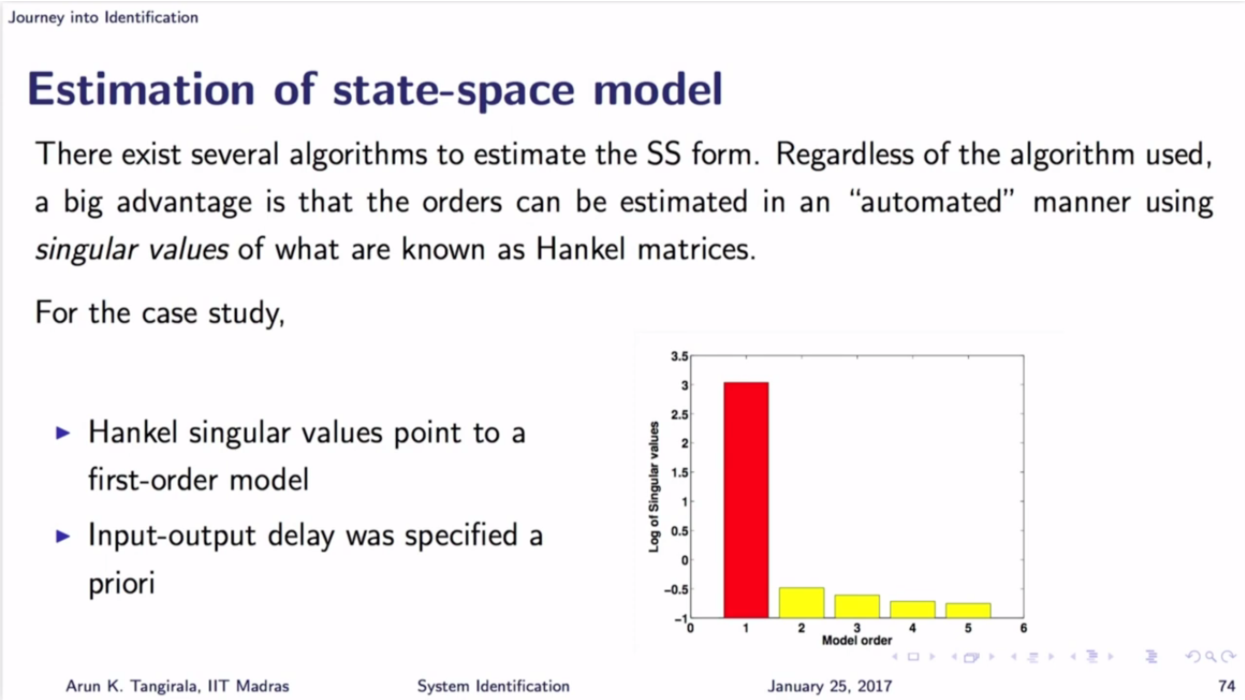
Sorry.

Twenty eight?

Twenty nine.

Twenty nine. So the order of India, the order of a state-space model is actually the number of states required to describe the system. In fact, I'll give you a more correct statement later on when we understand the concept of a minimal realisation, right? So the order of the system is a number of states required to describe the process in some minimal way. I can always have more states than necessary. So if you take India, it's a twenty ninth order system. Plus you have auxiliary states as well, right? Which call as union territories. So the order of our country has changed since the times I have gone to school and since the times perhaps you have also gone to school. It is becoming increasingly higher order system which means difficult to manage, right? Obviously higher order systems have more inertia. So now coming back to the liquid level case study, which is relatively innocuous one?We will not guess the number of states that is a beauty in state-space identification algorithms. I do not have to guess the order like we did for the input output case. There exist algorithms which can give me reasonable estimates of the order. And don't think that there is. I don't have to do anything out. Then I can just fit a state-space model. What I meant earlier when I said I don't have to guess the, a good initial guess is generated by the algorithm itself, which you may have to refine further. So for this liquid level case study we implement that algorithm which throws out what are known as singular values of some Matrix. Let us not worry about that. And by looking at the singular values we can actually determine what order is appropriate for a given system.They said,do not worry right now, of what matriceswe are computing the single values there is something called a Hankel matrices, which is constructed from data by the way.

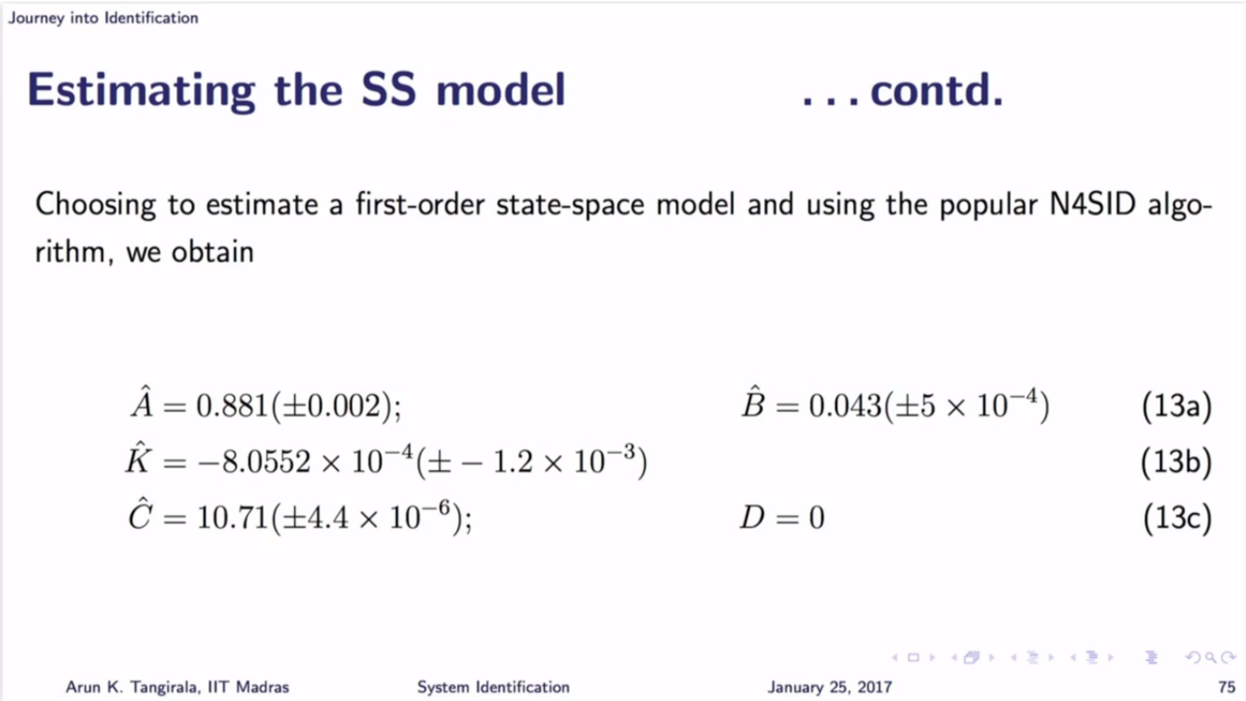
(Refer Slide Time: 07:27)



This Hankelmatricesof which we are computing the singular values is purely constructed from data. No other information is procured in any manner. What do you see on the screen here is, I will just zoom this for you further, is a plot of singular values in fact log of singular values versus the model order you can say that is, how many singular values are significant and so on. So you can see that thesingular value there is actually significant compared too few other singular values. I can have more singular values plotted there depending on what I ask from the algorithm.At the moment I say give me the top few singular values. And the way to determine a suitable order is to look at the most significant singular values. So if you look at the plot. The most significant ones are the first one itself. If it was a second order system you would have seen the second, the second singular value also being so-called significant in quotes. When the system is not corrupted by noise and the system is linear you can show theoretically that the number of non-zero singular values is exactly equal to the order of the system. So that is the premise on which this algorithm works. I will repeat. Under noise free conditions and if the system is LTI the number of non-zero singular values will be equal to the order of the system.But we have noise in the system. Of course, we also have some mild non-linearity. Even if I were to ignore the mild non-linearity I have some noise there. So the noise actually contributes to the non-zero nature of the theoretically what was supposed to be zero singular values, right? In other words we typically attribute the small valued singular values to either noise or non-linearity and the most significant one of the linear part of the system. To date there exists no clear cut algorithm that will be able to see through the noise and say yeah, you know, even though the data has noise in it, I am able to see through. By the way we're working on such algorithms and we have some success but I think very soon we'll be able to see full success on that. So it's one of the hot problems in state-space identification.

So here we use what are known as subspace identification algorithms. Why it is called sub-space we'll worry about later on. So the lesson message for us is that I am able to figure out the order by looking at the singular values and it clearly tells me the underlying system is for start-up. In fact, it is giving me the order not only of G and H, but also H put together everything. Remember there are two subsystems. There is the deterministic part and then there is a noise part, right? So this plot is giving me the order, put together. So it's a first-order system. Which we know actually the liquid level system is a first-order now. And if I were to identify now the first-order state-space model these are the estimates that I would get.

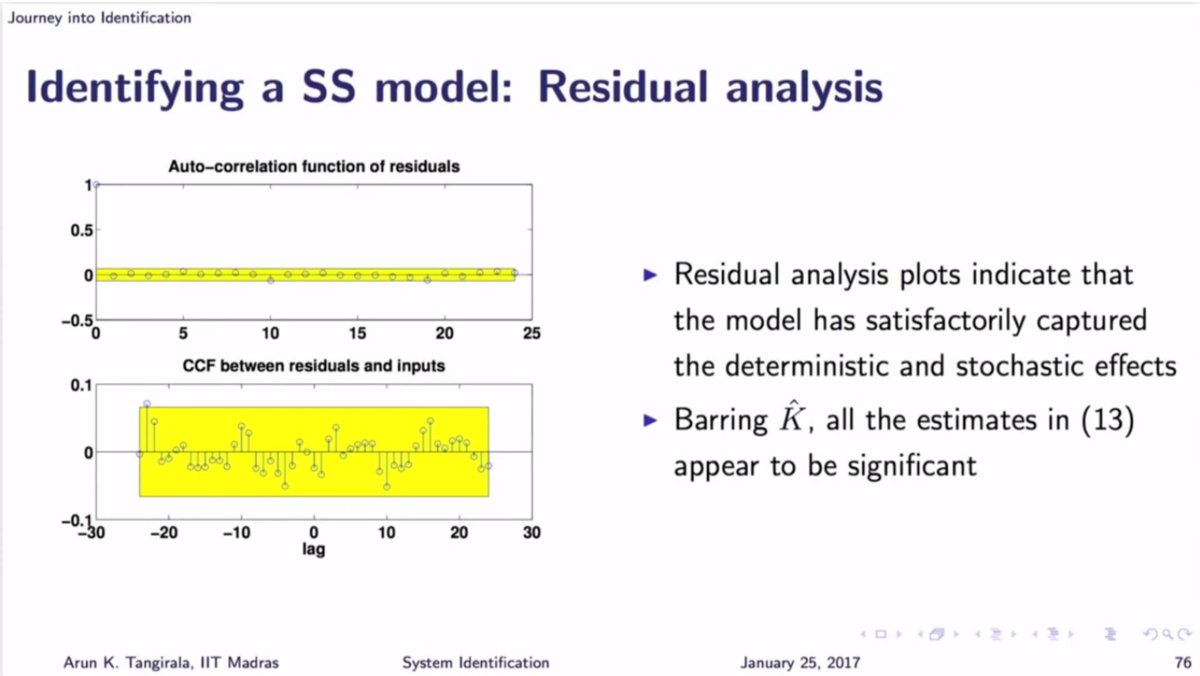
(Refer Slide Time: 10:54)



In fact you should have noticed something I've used x here in the state-space model and I have also use this notation earlier, right? In the input output models, there we have used x to denote the noise free response, which we don't observe. So you can think of the noise free part that I don't observe as a state. I am just trying to help you connect both these models.Anyway, so what I asked for the algorithm is an estimate of, the estimate of a first-order state-space model. Why do I say “a” because there are many first-order state-space modelsthat I can build. It's not unique. And it comes back with these estimates for A B C D and K.Remember that this D that we saw in the equation earlier, it represents the direct effect of input on output. What we mean by direct effect is it bypasses the initial part of the system. For example if I had a liquid level system like this, where I'm not modelling.Let's say relation between a Fi and H, but probably a modelling a Fi and F0, and if I had a bypass here. So this is my Fiand alpha Fi goes here and 1 minus alpha Fi is bypassing the system. And if I were to build a model between Fi and FO not H, then what would happen is that.
This is the initial part of the system. That means it has some dynamics to it. And this we call as the direct feed-through. So if there was a direct feed-through then the input would instantly affect the output. In your state-space model there is naturally builtin one delay. If that were the case then D would be 0. Inverse, if there D stands for Direct, you can think of it this way, D for direct or direct feed-through term. If there is a feed-through term then D would be non-zero. Typically in the systems that we look at thereare feed-through terms but you have to watch out. You have to look at the system that we had model. So in this case D being 0 makes sense. And then we have ABC and K. I have given you the values of the errors in the estimates, but please take them off from the slide. Do not even trust those error values.Because remember the truth is not fixed. There is no unique state-space

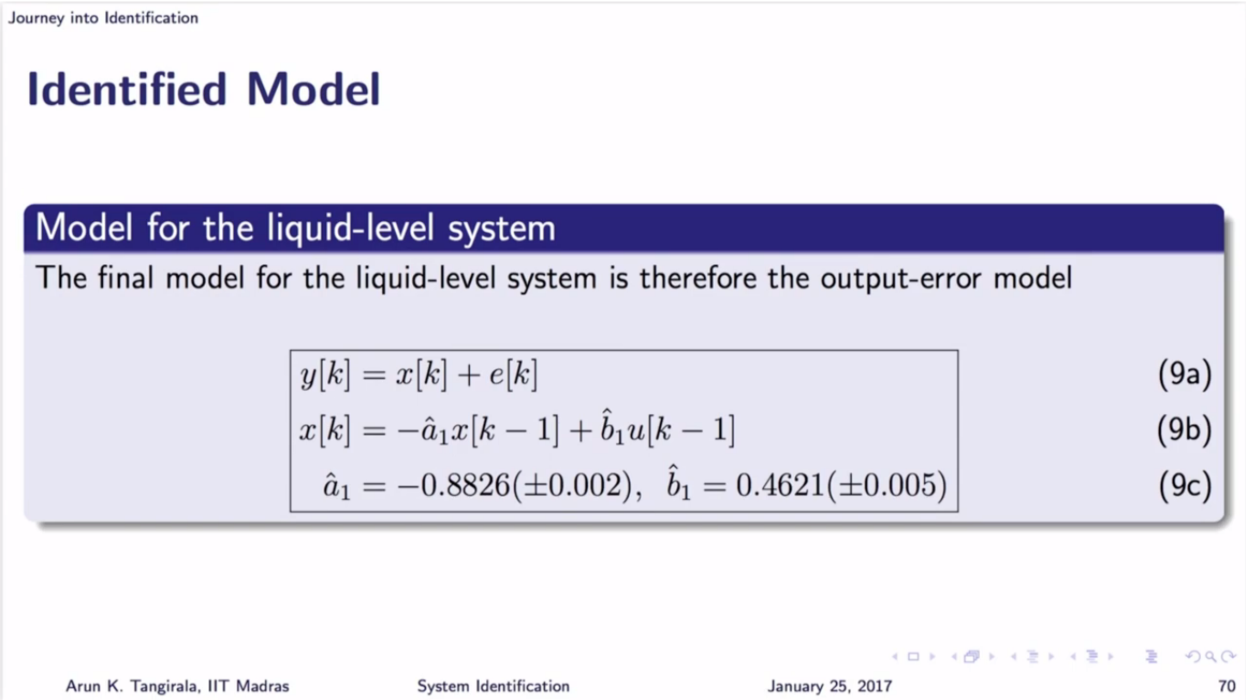
model then it doesn't make sense to talk about errors at all. So just look at the point estimates here. It's kind of obvious compared to ABC, K extremely small. So we could guess that maybe K should not be included at all. I'm not arguing this based on the errors remember that I'm reporting. I'm just saying numerically if you look at the estimates K looks very small. So if I were to omit K, this residual analysis is telling me that this model is good. I'm going to skip that.

(Refer Slide Time: 14:05)



If I omit K that means I'm assuming that the states are not affected by noise. Go back to the state-space model. By assuming K to be 0, what am I saying that. K there is no noise in the States. There is only what we call as measurement noise. Remember that whenever K is 0 in this what is known as these innovations far more state-space model. We are saying that the noise is only affecting the measurement not the states. In fact you can show that the moment you assume K to be 0, we are actually fitting what is known as an output error model. Okay. So you see now where we are being led to the same conclusion that we drew earlier. When we looked at input output models we said output errormodel is better suited for this process. Take the data and fit a state-space model as well.And you will find that you cannot set K to be 0. You will need K. So whenever K is not 0. Then you, that means, you're looking at coloured noise case.If you look at from an input output perspective.Whenever Kis 0 you're looking at an output error case that you can show theoretically as well. So what we are led to is an output error model. As you can see straight away x of k plus 1 is something times x k plus something times uk. The only difference is earlier we wrote x off k is something times x k minus 1 plus something times u k minus 1, but otherwise it remains the same. And then y k is, you know, the c hat times xk plus ek. Again, please completely disregard the error values that have reported. These are not calculated in a theoretically rigorous sense, so just ignore them. Now, what we have is a first order state-space model corresponding to the output error structure. All right.Now,if I want to compare the model that I have obtained earlier. What is a model that I have obtained earlier, Input output model. x k is minus a1,x k minus 1 plus b1, uk minus 1 and y is x plus e,right? This is what is a model that we have. Suppose I want to compare the statespacemodel that I have obtained with this one, then I can force further in this state-space algorithm, identification algorithm. I can ask for a state-space model such thatc is 1. See this effectively means I'm forcingc to be 1. If I want this kind of a structure, why am I doing that just to see if the algorithm that I'm using for state-space identification, gives me the same estimate as I get from the output error routine.

(Refer Slide Time: 16:50)



In fact, whenever I force any of the matrices to take on a particular value or a particular structure. We say that we are building what is known as a structured state-space model. Okay. So what we are doing towards the end as a conclusion is building a structure state-space model. And when I do that this is the state-space model that I obtain. And you must be now familiar with the estimates that we have obtained earlier.The estimates of the coefficients a and b are pretty much same as what we have a obtained earlier with the output error routine,right?

(Refer Slide Time: 17:36)

Journey into Identification

Identifying a structured state-space model

One could further force the SS model to have a structure similar to the OE model. Forcing $C = 1$ (in addition to $K = 0$) produces the estimate

$$x[k+1] = 0.8826(\pm 0.002)x[k] + 0.4621(\pm 0.005)u[k] \quad (15a)$$

$$y[k] = x[k] + e[k] \quad (15b)$$

Arun K. Tangirala, IIT Madras   System Identification   January 25, 2017   78

So this is an alternative way of arriving at a model for the given system. What is the prime difference between this approach and the input output approach? What is the major difference? Any answers. What did we have to do earlier? In the input output case where we explored output error an equation error case. What did we do there and what else did we do something different in the state-space identification? What is the prime deference, man? Why is there such silence in the class? Prime difference is in order determination. I didn't have to struggle much. I still have to subject the model to residual analysis as I showed you earlier. The top plot is the articulation plot of the residuals and the bottom plot is a cross correlation plot. But in fitting a preliminary model, I have very good leads. Whereas, I had to rely on step response. My reading of the step response, guess that it is a first-order. And then go through all of that, right? But here in a state-space identification approach I don't have to really break my head so much at least in the preliminary estimate. Please do not think that the singular values can be fully trusted on. They give you a very good lead, so it saves you a lot of headache, lot of time.Once you come up with a good estimate from the singular values you still have to go, subject the model through the standard residual analysis test. And then see if the guest one is appropriate. Why you have to do this, because remember you are making some assumptions on what is significant and what is insignificant you are taking your calling the shots. And maybe your understanding or interpretation of what is significant or insignificant is perhaps not correct. Okay.So remember that. That is a prime difference and of course, the other prime difference isidentifiability. There's no unique state-space model.So this case study has given you. I would say a very good and elaborate overview I would say not details a very broad overview of what you will get to see in any system identification exercise. And I've highlighted a lot of important facets of what you would see in identification. We will talk about response based descriptions which I'm not going to go with this only slide I want to show.We will talk of response based descriptions where we will talk of convolution.