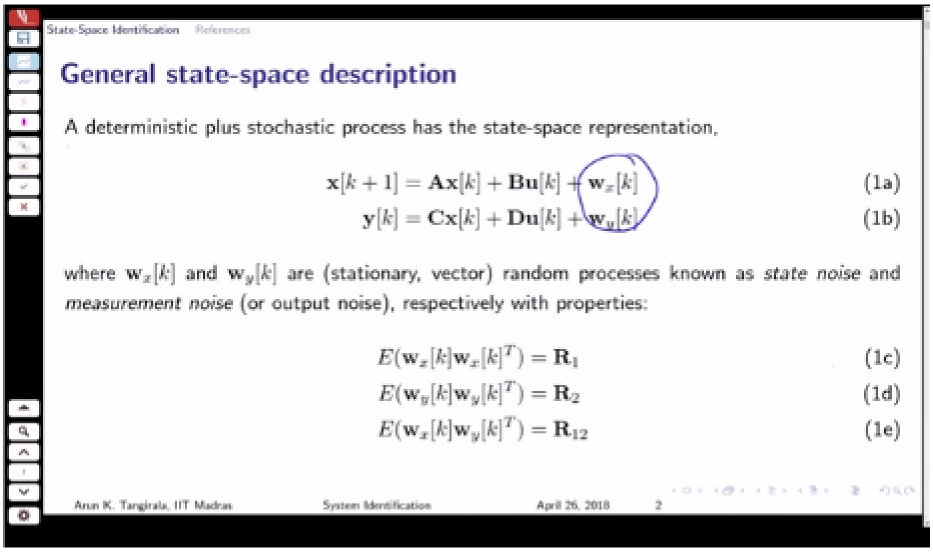
CH5230: System Identification

State-Space/Subspace Identification

Part 2

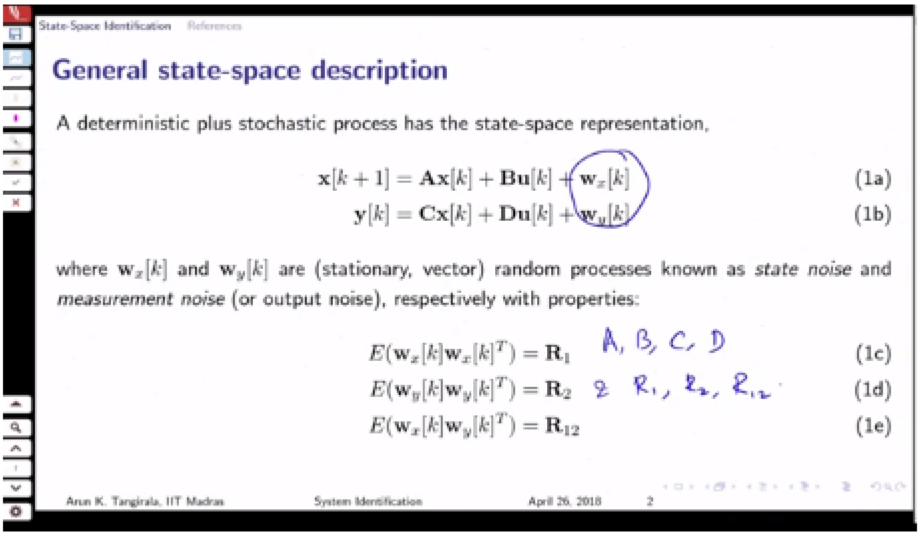
So, welcome to the lecture on Stage Space Identification. We started our journey on this topic yesterday in the previous lecture, and of course, now will plunge into the details today. So, what we were introduced to in the previous lecture is this general states space model that we had looked at, where both the deterministic and stochastic effects are represented. And as you remember, said, the key thing is this.

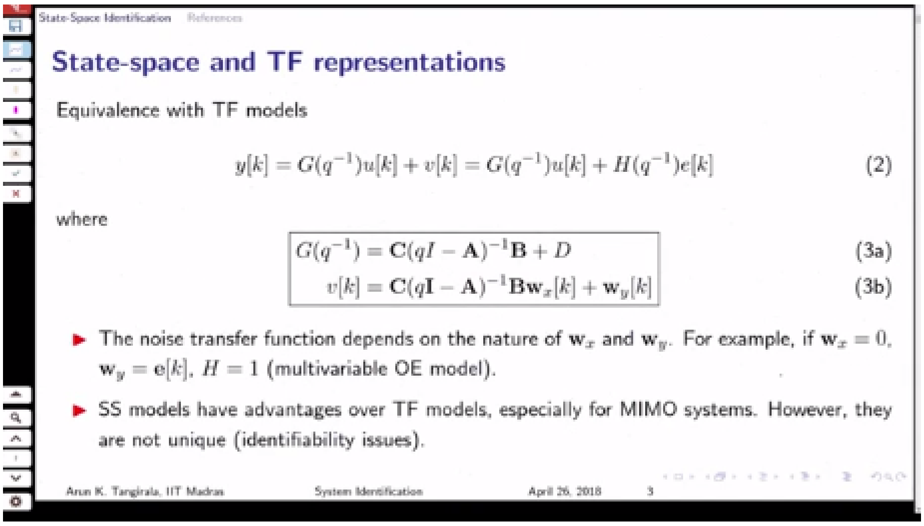
(Refer Slide Time: 00:49)



So, the key difference between the deterministic and the stochastic version is the appearance of this stochastic terms or signals in the state n measurement equation, and as I had pointed out yesterday, now the description consists of A, B, C D, and these three matrices, R1, R2, R12. In many situations, R1, R2, R12, maybe 0 that is the process and measurement noise maybe uncorrelated.

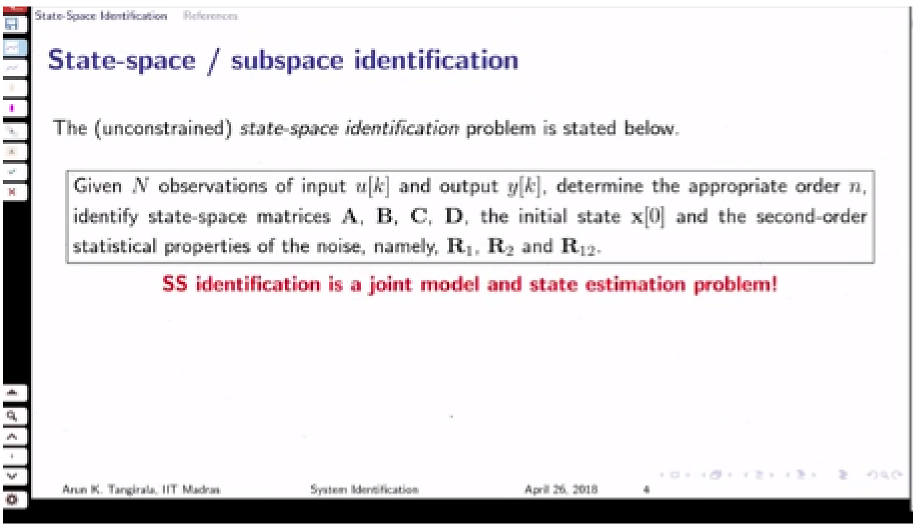
(Refer Slide Time: 01:17)





And we also discuss the transfer function equivalence and talked about the relation between G, the plant model and A, B, C, D, which we know from our previous lectures, early lectures on LTI systems and as far as the noise model is concerned, we said that one cannot obtain an expression between A, B, C, and the noise covariance matrices, because it simply depends on the nature of Wx and Wy. For the multivariate OE model, Wy, sorry, Wx is 0, and Wy is simply a vector of white noise sequences. In other words, H equals 1. Because v[k] turns out to be e[k]. In fact, you can go back to the other descriptions, namely Armax, Arx, OE, sorry B-J models and ask what would be the corresponding descriptions. But we'll do that a bit later. We'll work out one such exercise a bit later when we discuss what is known as the innovations form of the state-space model. This is not the only way to describe a deterministic-plus-stochastic system in state-space framework. There is something called innovations form that involves Kalman gain which we will talk about a bit later, after we have reviewed the concepts of Kalman filter. Now, of course, we also talked about the advantages that stage space models have over TF models, right. If you recall, mainly to advantages, the joint estimation in an identification problem is best, done in the state-space framework. And MIMO systems are also nicely handle in the states space domain.

(Refer Slide Time: 03:05)

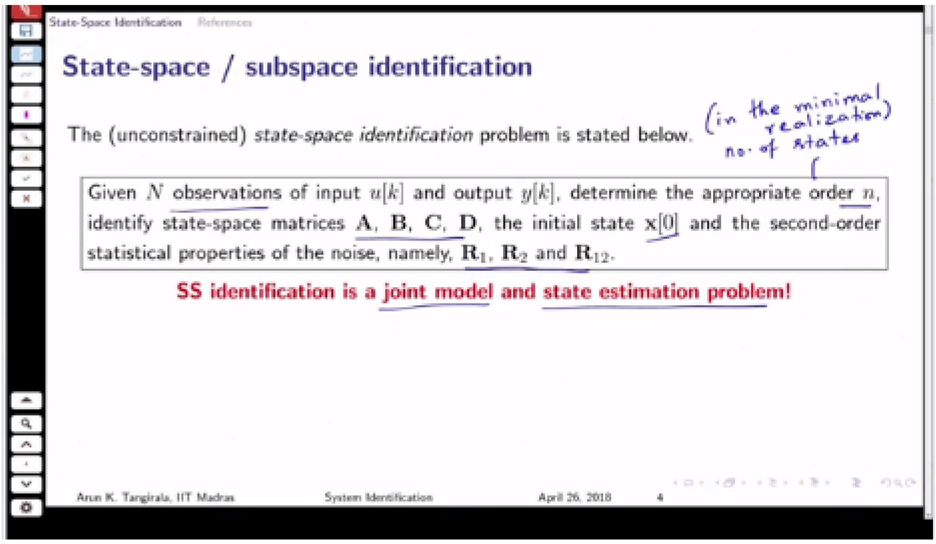


So what is this problem of state-space identification that we are interested in? So that's the statement that is given in the box here, given n observations of input u and y. Look at the number of things we are required to estimate. We are supposed to estimate the order, which means the number of states.

And although, I write number of states here, you should think of this as a number of states in the, sorry, in the minimal realization. We have spoken of what is a minimal realization in our early lectures. We'll also give a definition of minimal realization later on after we have reviewed the concepts of controllability and observability. So we need to figure out how many states are required and then identify the state-space matrices, the initial state, the initial condition of the system. And the second order statistical properties of the noise, namely the covariance matrices. So, so many things have to be estimated, in contrast to, what we have seen in the transfer function or input-output framework, where the information given is the same input-output data but the deliverables of the problem are different.

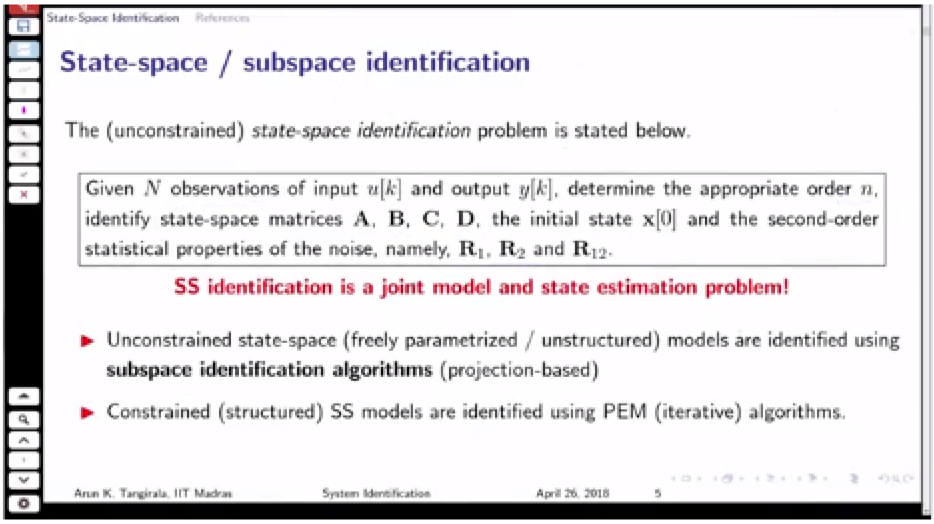
There, we are interested in the numerator and denominator polynomials of G and H and sigma square. E. That's about it. Maybe, in some situations, initial condition y[0], but here it seems to be a lot more and what is important to remember is, we are estimating the state-space matrices and you should be familiar with the dimensions of these matrices. A is a square matrix and so on, and the dimensions of the other matrix depends on the number of states, inputs and outputs. So, in general, as you look from look at the problem statement, the state-space identification is a joint model and state estimation problem.

(Refer Slide Time: 05:11)



So, we are not only estimating A,B,C D, but we are also estimating the initial condition x[0]. Once i know the initial condition x[0] and also the statistical properties of the noise, then I can compute the states or rather, sorry, I can compute the estimates of the states at subsequent instance by propagating x[0] through the model and also come to the errors in those estimates using the statistical properties of the noise. So, in effect, what we are doing is, we are estimating the model and also the states jointly. And that is the reason why we have to estimate so many quantities as against in an input-output identification problem or transfer function identification problem.

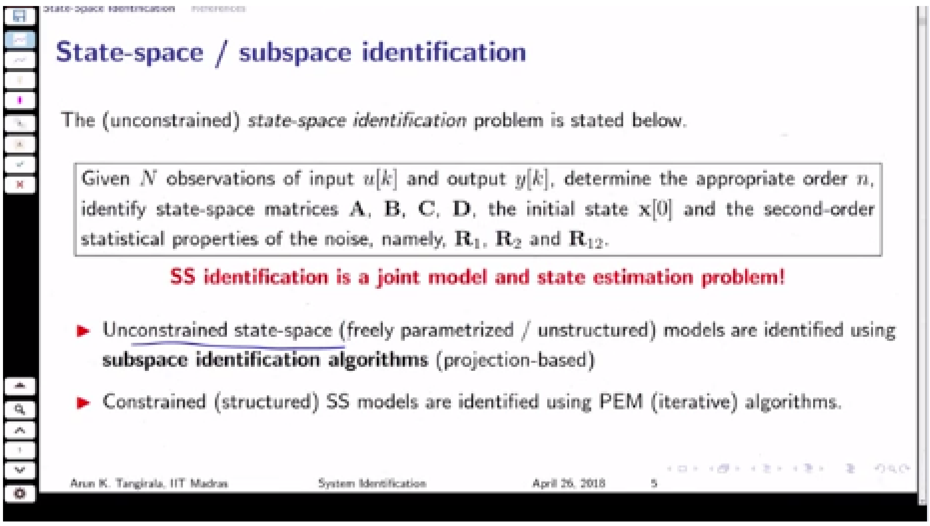
(Refer Slide Time: 06:06)



Now there are two scenarios when it comes to state-space identification. One is the unconstrained scenario, where there are no constraints whatsoever on the matrices A, B, C, D. Remember, one of the main demerits or the criticisms that states space models have is that they are not unique, which means they're not identifiable, unless some structural constraints are imposed on A, B, C, D, and we have talked about this at length before, whereas you generally don't run into such issues necessarily in the input-output domain.

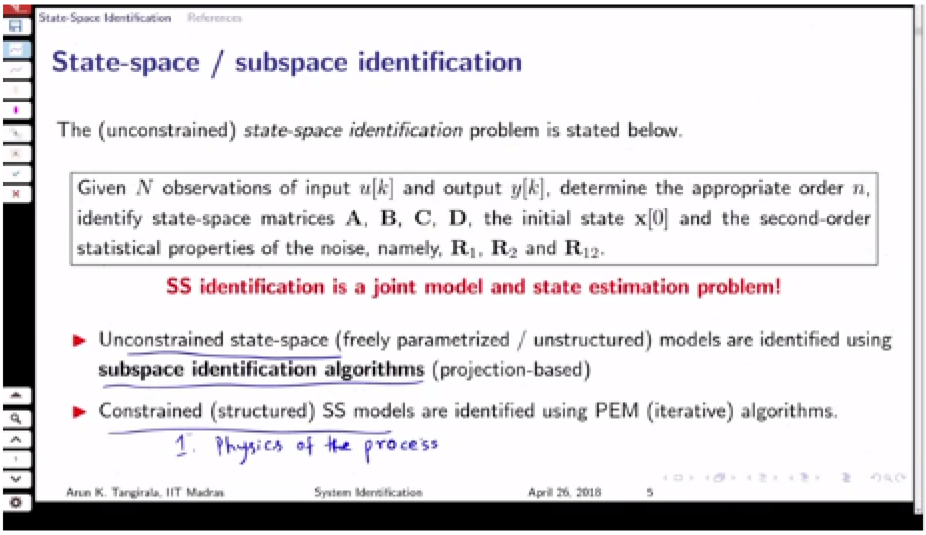
Of course, there is also identifiability requirements that the input-output models have to satisfy. But those are much milder and you can always assure that they're satisfied and you can obtain an identifiable say input-output description. Whereas the state description, by default, when there are no constraints placed on any of this matrices, they are not identifiable. And when you don't place any such constraints on A, B, C, D, matrices or even the noise covariance matrices, then we call that as a freely parameterized. You can say it's a non-parametric state-space model or you can also call it as a freely parameterized state-space model.

(Refer Slide Time: 07:31)



Essentially, there is no structure and all the parameters of the, all the increase of this matrices have to be estimated. And that is where the subspace identification algorithms come into play. The subspace identification algorithms are linear algebra based algorithms. Although, I say linear algebra, it has been shown that there they have an optimization equivalent. That means, they are optimal in at least squares sense, but you don't see that in the explicit formulation of the subspace identification algorithms, which is what will spend quite a bit of time on. Then the second scenario is a constraint case, where, I impose constraints on the state-space matrices. And how do I know what constraints to impose? Well, these constraints may be driven from two considerations. One consideration is from the physics of the process. When I say physics, please treat as chemistry, biology, everything. So, whatever knowledge I have of the process tells me that I should use, I should impose this constraints on A, B, C, D. For example, I might have derived a first principles model and the first principles model tells me what should be or should A, B, C, D, look like? How should they look like?

(Refer Slide Time: 08:50)

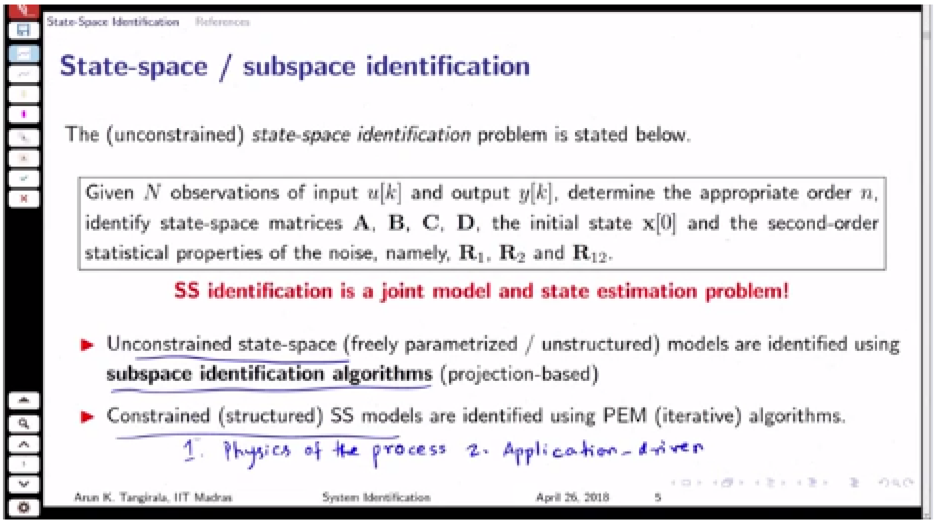


And then I only estimate the non-zero entries of these matrices. The other source of these constraints, that is, how do I know what constraints exist on A, B, C, D, is application driven, right. For example, I may be interested only in a diagonal form of A. But regardless of how these constraints are being imposed, the important thing to remember is unless, you impose sufficient number of constraints on A, B, C, D, identify ability cannot be guaranteed. Right. We have discussed also in the class, an example. So, suppose I, impose only some constraints on A, let us say that A should be diagonal, does it mean that I have a unique state-space model need not be, right. You can always go back to your transfer function representation and ask, what should be the number of constraints and the type of constraints that you have to impose, so as to make the state-space model identifiable. And generally when you impose such structural constraints, not driven necessarily by physics of the process, we call them as canonical models.

So a canonical state-space model is nothing, but a structured states space model and there are many such, we have discussed, for example, controllability form of state-space model, controllability canonical form, observability canonical form, controller canonical form, and so on. Diagonal or Jordan canonical form. That canonical essentially means a structure and one has to show that a particular canonical form that you have leads to an identifiable state-space model. That means, there is only one state-space model that has a structure for the given LTI system. And there is a very nice result to that effect. You can refer to Ljung's book who showed where they show that the multivariable for the general multivariable system the observer ability canonical form is always identifiable. So if you want to really get a state-space model that is unique, the summary is that you have to impose a structure and sufficient structural constraints have to be imposed. But if you want a state-space model, not the state-space model, then you can simply work with subspace identification algorithms to get to state-space model in some basis.

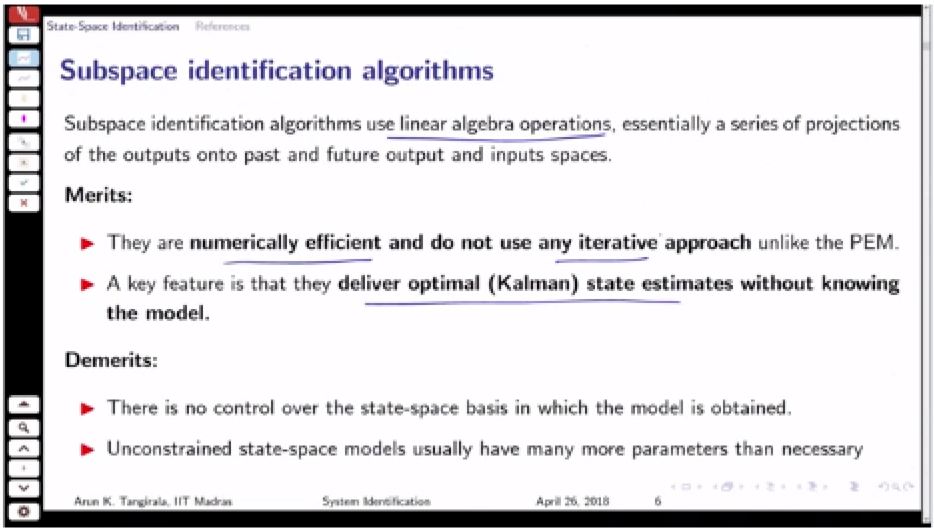
And in many applications, that is enough, we may not need a state-space model in a particular basis. So, you have to figure out which one is a requirement. As far as our lectures are concerned, and today and the next lecture, we will focus on subspace identification algorithms largely and then conclude with a discussion on the structured state-space models. All right.

(Refer Slide Time: 11:47)



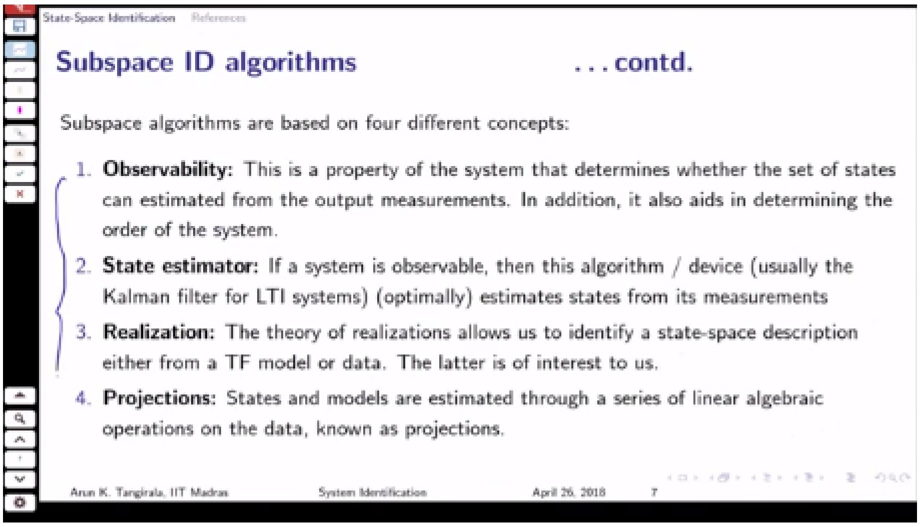
So what are the subspace identification algorithms which we are going to spend some time on? Well, essentially, as I said, they use linear algebra, linear algebra operations, which are a series, involve a series of projections of the outputs onto past and future output and input spaces. At the moment, it may sound vague to you, but more details will come much later. And we will have to learn this in part, because subspace identification algorithms involved not only linear algebra operations, but more importantly certain core concepts based on observability, Kalman filter when state estimation and so on, which we will learn in part. So essentially, we'll peel one layer after the other. One of the reasons why the subspace ID algorithms are popular is because they're numerically efficient. And they do not use any iterative approach. Like the PEM approaches, right? You know that the PEM approach in general can lead to a nonlinear least squares problem where an iterative method is required. And also, that they deliver optimal state estimates. Remember, these identification algorithms are not only supposed to give me the model, but also the state estimates. And the state estimates that these subspace ID algorithms deliver are optimal state estimates in the Kalman sense, right.

(Refer Slide Time: 13:08)



The beauty is, the regular Kalman filter that you learn in many of the advance control courses and so on. As you know, require the specification of a model, whereas the subspace identification algorithms do deliver this a Kalman state estimates without the knowledge of the model, when we say without the knowledge of the model, without the use of specifying the model. The model is implicitly derived from data. So we call these states has numerical Kalman filter states. Or you can say that there is a numerical Kalman filter underlying the subspace ID algorithms. The reason for calling numerical is obvious, because, this Kalman filter is estimates are derived just from data alone. It is not your model based Kalman filter. There are demerits, of course, as we have said there is no control over the state-space basis. That means, it will give you a state-space model which is not necessarily unique and it also suffers from the same demerits as your non-parametric models, which is that there are many more unknown is unnecessary. I may be unnecessarily identifying 50 unknown as compared to maybe 10 or 12, 20 or 15, or much more, much lesser than 50. If I know the structure of A, B, C, D. Obviously, that means you need some prior knowledge that you have to pump in. Anyway. So let's not worry about a structured stage space models for now. We'll focus on unconstrained states space models. As I had said, a short while ago, the subspace ID algorithms involved four different concepts.

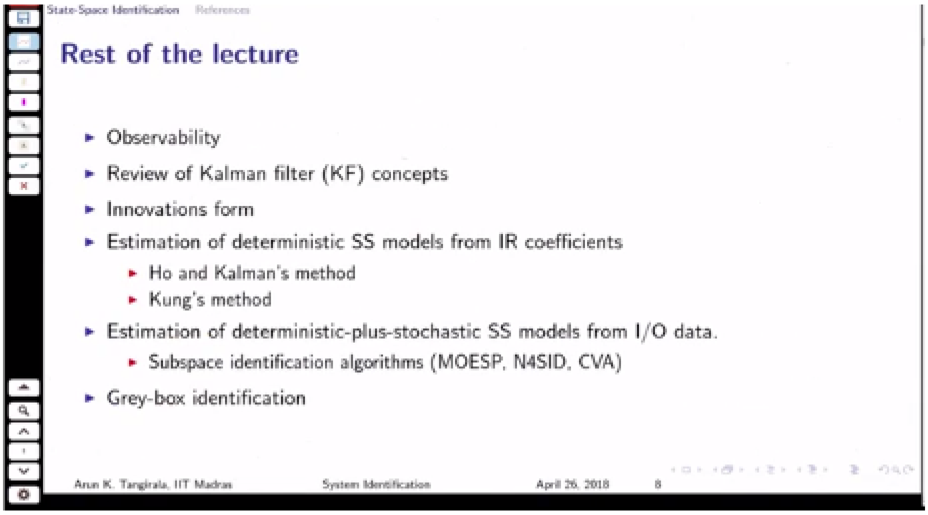
(Refer Slide Time: 15:07)



One is this notion of observability, okay. The other involves the concept of state estimation or a state estimator. And the third concept is this notion of realization. This realization, in some senses, similar

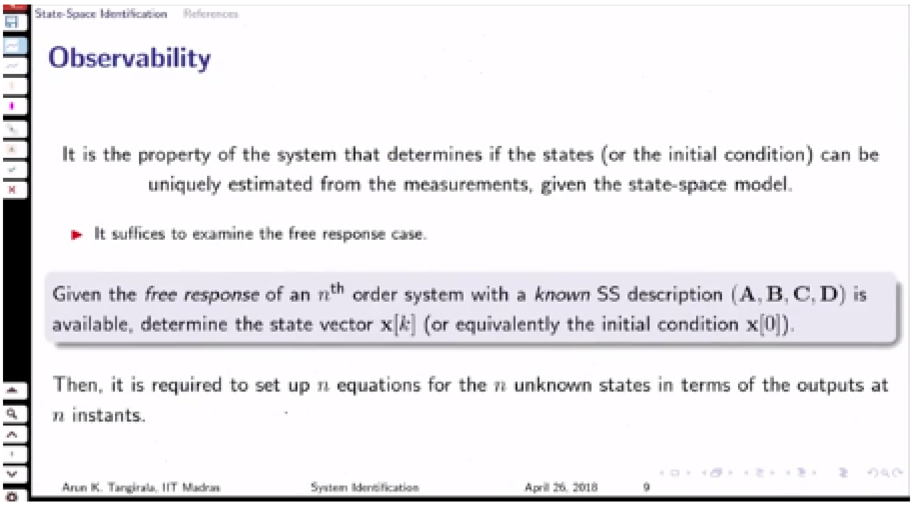
to the notion of realization that you've seen random processes, but is different. What we mean by realization in state-space literature is one description, one state-space description for a LTI system. And then you have this concept of projections. So the notions of observability and realization offer the theoretical basis and to a large extent state estimation also, but the observability and realization concepts offer the theoretical foundation and the state estimation projections take care of the data driven or data based analysis and inferences that the subspace ID algorithms do for you.
So, we will learn very briefly each of these concepts in detail, right?

(Refer Slide Time: 16:15)



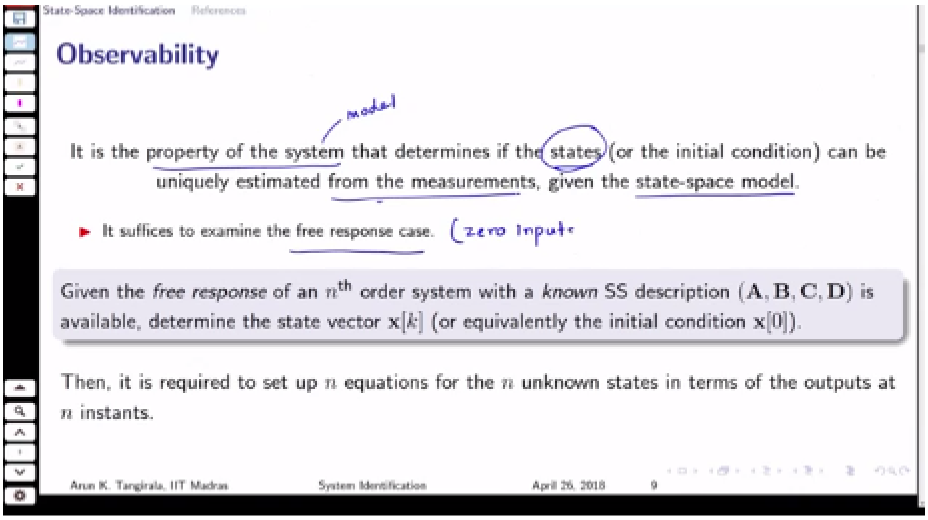
So the rest of the lecture is organized as follows. We'll first look at observability, then review Kalman filter concepts. And, although, I say observability here, we will also look at controllability. Okay? And we'll review Kalman filter concepts and then talk about the innovations form briefly, I mentioned this earlier. Then, we will learn how to estimate deterministic states-space models from IR impulse response coefficient. This is necessary both from a pedagogical and historical viewpoint, because if you look at the history of subspace ID algorithms, one of the first algorithm to appear was by Ho and Kalman. Who showed how you could identify a realization of a state-space model from the impulse response coefficients and the noise free conditions and noisy conditions would, was largely due to Kung who extended the idea of Ho and Kalman's to the noisy case, and then, having understood that we are in somewhat better position to deal with estimation of deterministic-plus-stochastic models from arbitrary input-output data. So you see there's big jump here. We are moving from deterministic to deterministic-plus-stochastic. And we'll move from impulse response coefficients to input-output data. And that is where we'll run into these algorithms called MOESP, multivariable output error, state-space algorithms or enforce it or a CVA and so on. And then, finally conclude with Grey-box identification. So there is a bit of a journey for us ahead. But we'll learn in parts and therefore things will be okay.

(Refer Slide Time: 18:05)



So let's start with the notion of observability, and we briefly spoken about it earlier. Now it's the time to discuss a bit more in detail. What is essentially observability, it is a property of the system. In fact, more so, it is also a property of the model as well. Okay. So, whether this model allows, this model is such that or the description is such that you can estimate or determine the states which are usually hidden from us or their initial conditions from the measurements and given the model. So, what is given to us, the measurements and the model, and what is to be known? The values of the states. So this is the object of interest, okay. So, that is, so observability will tell me if it's if a model is observable or system is observable, then that means I can estimate the states, given the model and the measurements. So let's look at the free response case. That means we will turn off the inputs without any loss of generality.

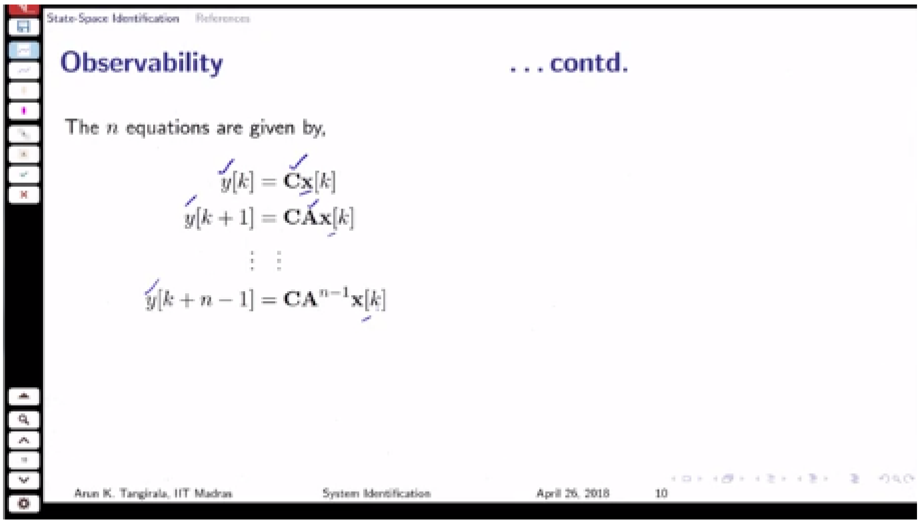
(Refer Slide Time: 19:12)



Because if inputs are known, I can always accommodate it. It's much easier to look at the free response case. So the problem statement is as follows. Given the free response of an Nth order system with a known state space description, that means a model is given. And right now, we look at deterministic systems. We'll not worry about deterministic-plus-stochastic. The objective is to determine the state vector x[k] or equivalent the initial condition. If I know the initial condition, I can let it propagate to the model to determine the states. So from a linear algebra perspective, you have n unknown, because let's assume that x[k] is some real valued vector and it lives in an n dimensional space. Sometimes we may use nx also. Right? So x[k]as escape belongs to this, sorry. All right. Or you can say simply that this is an n by 1 vector of real valued numbers. So obviously since there are n unknowns, I have to set up n equations.

There's no doubt about it. And how do you do that? Any idea? What is given to me? There are two things that are given to me, right? One is the model, other are the outputs or measurements y. So naturally, I will turn to the output equation, right? So I turned to the output equation y [k] equals e x[k]. And let us say, I'm interested in estimating x[k] at any instant. Then, I would, I can even rely on past measurements. It doesn't matter. I don't have to write future equations, but since [k] is generic, it's all right. So I write n equations involving the outputs and states, that means the knowns, remember. I know this and I'm also know C, I know A and I know all these measurements.

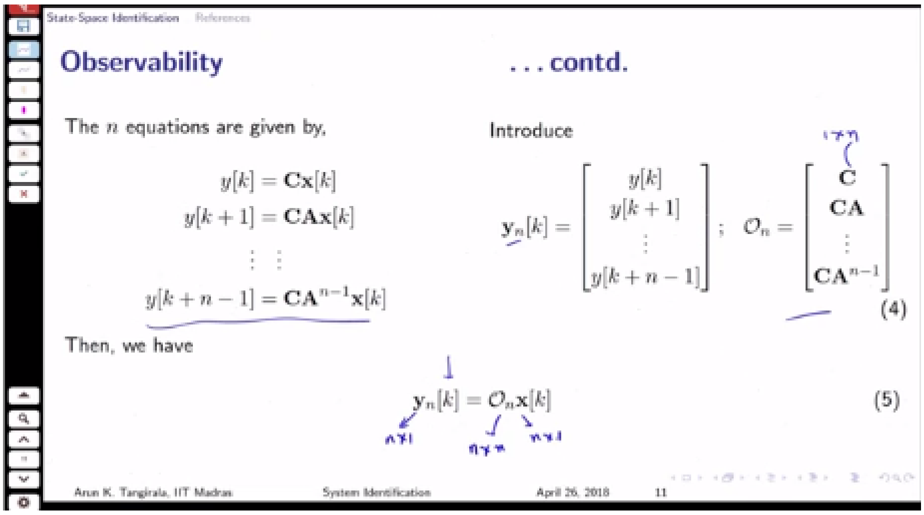
(Refer Slide Time: 21:18)



So if you see there are n questions here and n unknowns and if I introduce these two vectors, Yn and y now is a bold face, it's a vector of dimension n, where simply I stacked the left hand sides of the equations and this matrix O subscript n, this is a calligraphic O which has C, CA, stacked one below the other, up to C, A to the N minus 1. I can write these equations in a nice vector form. Y[k] is Onx[k]. So remember, this is of dimension n by 1. What is the dimension of O n? By construction, what is the dimension of O n? Why does it take so long? What is the dimension of C. Right now assume that you're looking at a SISO system. 1 by n by n, and how many rows do I have? So O is, O n is n by n. And x of course is n by 1.

So you have a square set of equations and from linear algebra, we know that a unique solution exists, if, O n is of full length, correct? That is essentially translated to observability condition. That is exactly the observable condition. I will be able to uniquely estimate the states, if and only if, the observability matrix O n is a full rank.

(Refer Slide Time: 23:00)



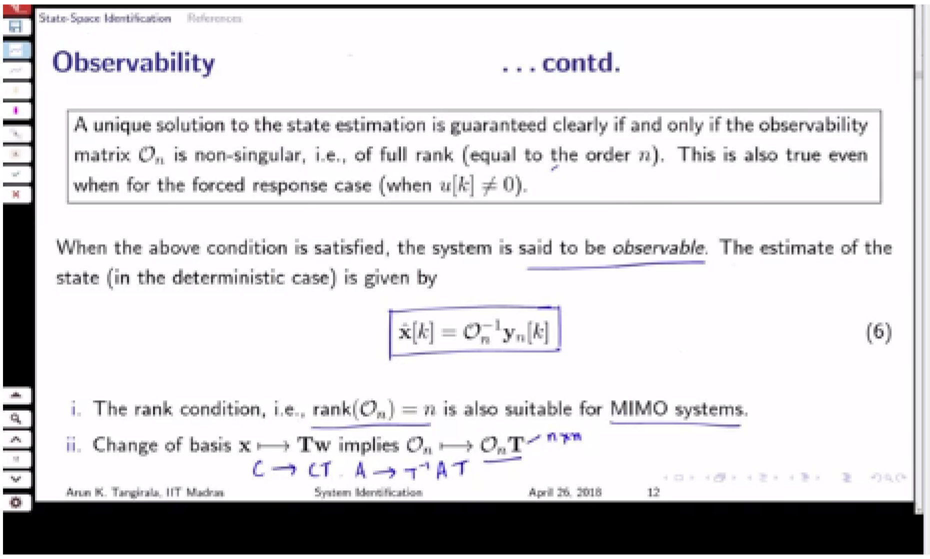
And that is what we have stated here, right? And you can apply the same condition to the first response case as well. It doesn't change. When that condition is satisfied, the system is set to be observable. So that's natural. It's intuitive, right? That C and A govern the observability. B has no role to play. Because what does B connect? The inputs to the states? How the inputs affects the states is what is determined by B. But we are not using the inputs, even if you were to use the inputs, it won't make a difference to the observability condition, what is the key other measurements. And C is the matrix that connects the measurements to the states and A is the matrix that tells me how the states are coupled to each other. So together, both make up the observability matrix and when you have an observable system, you can straightaway right an estimate of x.

Here, you estimate word is not so correct, but we'll still call it as an estimate. Because there's a deterministic system, we are assuming process to be linear, you'll get an exact solution. We'll still call it an estimate in anticipation of noisy data and stochastic situations coming up. So, this rank of O n equal to n, although we had derived for SISO case, it is also, it suit, suitable for MIMO systems as well. So you don't have to really worry about a different condition for MIMO systems. That is a nice thing about states-space models. And, also, this is an important, the second point is important.
If I were to work with a different states, the state-space basis, that is from x to w related through a non-singular transformation T, then observability matrix changes accordingly. Which means, just as a state-space model is not unique, the absorb into matrix is also not unique.

It better not be, otherwise, you will have difficulty. If you're changing states, the observable the matrix also should change, then only the states will be obtained correctly. How does it change? By a factor of T. And remember, the multiplication is post. What is the size of T? Sorry. N cross n. And O n is n by n. So there could be a confusion and these confusions typically arise only in the exam, of course, you will have slides with you, but suddenly you may think, "Oh, it is T times O n or O n times T. All you have to do is, if you're in confusion, either refer to the slides or write your state-space model and it'll be pretty clear, you know, how C changes, how A changes, right. In a change of basis, Tw, we know that C changes to C times T. And A changes to? Do you read, remember? What does A change to? Krishna? Any idea? You recall? Why does it take so long? In a in a few seconds you should be able to do that. AT, that's all.

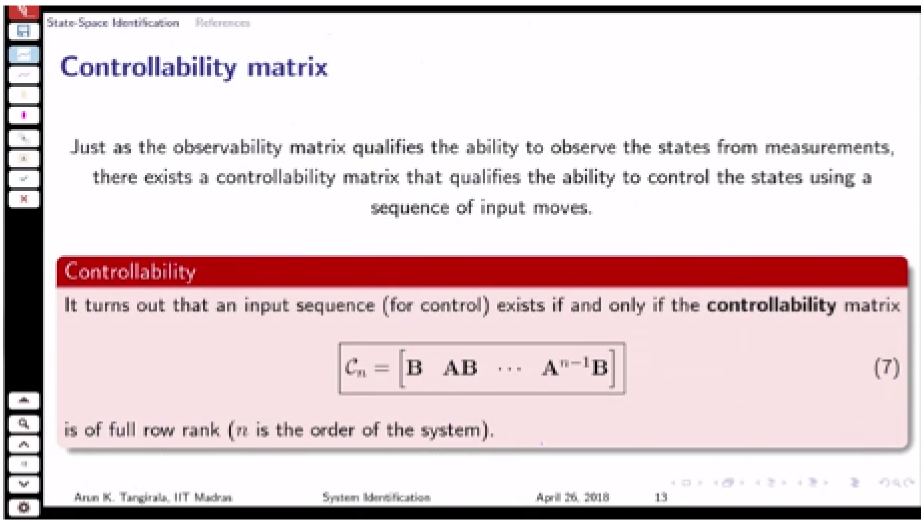
I've drawn a vertical line first, so it cannot be AT. This is our multiple choice way of figuring out what the answer is. If the first factor was A I would have drawn a slanting line. T, AT. Okay, so you plug this new C and A. Hey you should not, at this stage of the course you should be quite comfortable with this. If you don't know the answer, you should be able to derive it immediately. It doesn't take a few, more than a few seconds to do this. So you just plug in the new C and new A into your observability matrix definition, you will get this answer. Anyway, so we will make use of this condition two later on.

(Refer Slide Time: 27:22)



Let's proceed. In a similar way, you can define what is known as a controllability matrix. What is the controllability matrix? First of all what is controllability? Like observability, controllability is a property that tells you whether the system is controllable. What do you mean by controllable? Whether I can drive the states of the system from a non-zero condition to zero condition. Reachability is the property of the system which tells me whether the states of the, states can be driven from non-zero condition, one non-zero condition to another non-zero condition. Controllability specifically about being able to bring back the system to steady state or its equilibrium position. They're mostly related and the condition is the same and going by similar arguments and intuitively we should expect the controllability matrix to be only governed by B and A.

(Refer Slide Time: 28:13)



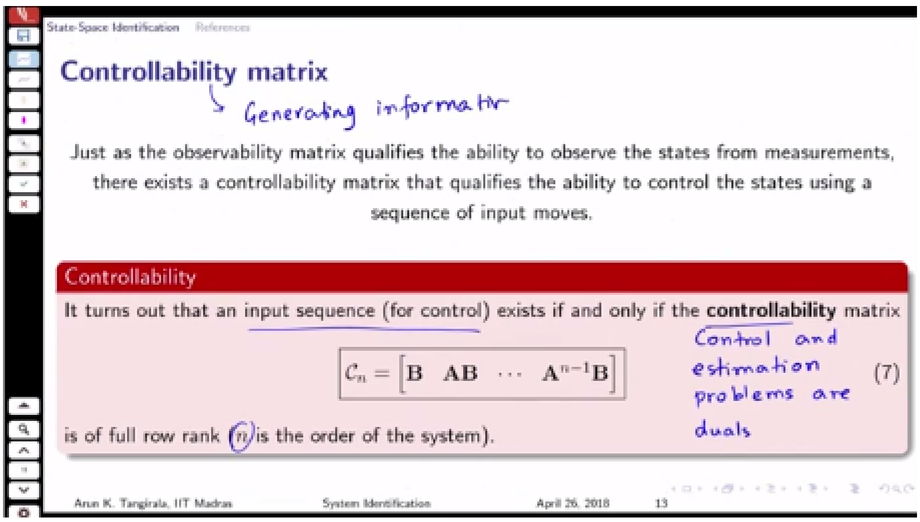
So A is central to both. It's common to both observe ability and controllability, because that tells you how the states are coupled with each other. If it is observability C comes into picture. If it is controllability, B comes into picture. Okay. And you can see nicely that the controllability matrix has a similar structure, but it has a transpose structure of the observability. And once again, this kind of confirms that control and estimation problems. After all, state information is parameter estimation as well. Problems are , what do we keep saying? Duels of each other. That duality is reflected very nicely here. Okay. So we say a system is controllable, I mean, system state-space system model is controllable which means that there exists and input sequence for control. In this case the goal is not to find the states, the goal is to find the inputs. So your inputs become the unknowns now.

The states are given, the terminal state is given, the initial state is given. In the observability, the measurements are given and the states are unknowns. So there's a nice duality there. If an, so this, if an input sequence for control exists if and only if this controllability matrix is full rank. And again, n is out of the system. So what is the role controllability in identification? Can you think? Why are we talking of controllability? Observability is kind of clear, somewhat, but what about controllability? What role do you think it plays? Any ideas? What has my ability to move the states from one initial condition to another, maybe zero or another condition. Got to do with identification? Suppose the system is not controllable what does it mean?

Certain states are not perturbable.

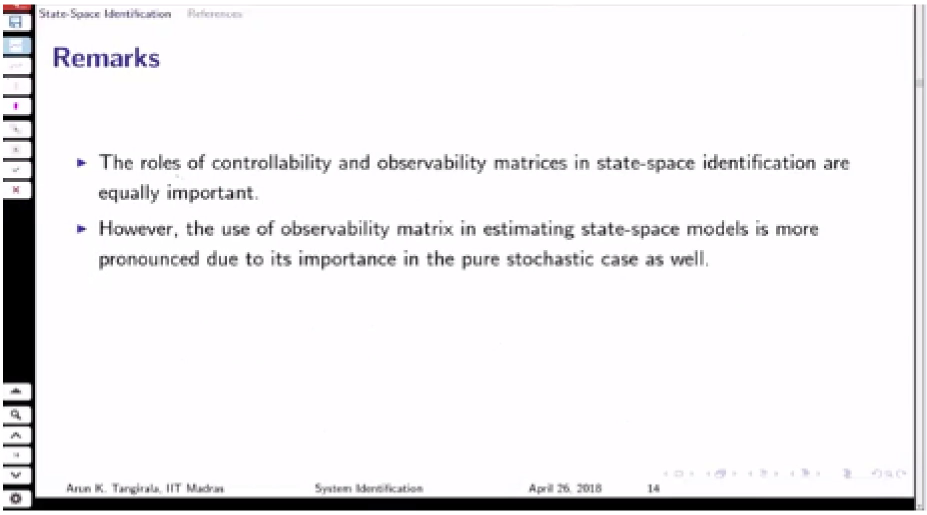
Certain states are not perturbable. You cannot pert up them. You cannot excite them. They will remain in the, they could remain in the steady state, they will never budge from their conditions. And if they don't get excited, how can you identify those modes? So it's important, so pre-experiment or during the experiment, you have to be guaranteed that your input changes are perturbing all the states of the system. Post experiment, you should be able to estimate those states. So controllability comes into play during your experiment and for informative data requirement, whereas observability comes into play for inferencing. One is for generating informative data, right? So controllability is for generating informative data.

(Refer Slide Time: 31:50)



Whereas observability is about inferring the states from this informative data. Both have to be fulfilled. Is it possible that a system is controllable but not observable, yes. And vice versa. So which means, you may be able to excite the states but you have chosen some wrong measurement. You place the sensors wrongly or somewhere there is a problem with your sensing, because observatory has got to do with how you sense and controllability about how you excite and whether you're able to reach out to all the states, right. So, both are required for, you to be able to identify a state space model. Any questions? Right.

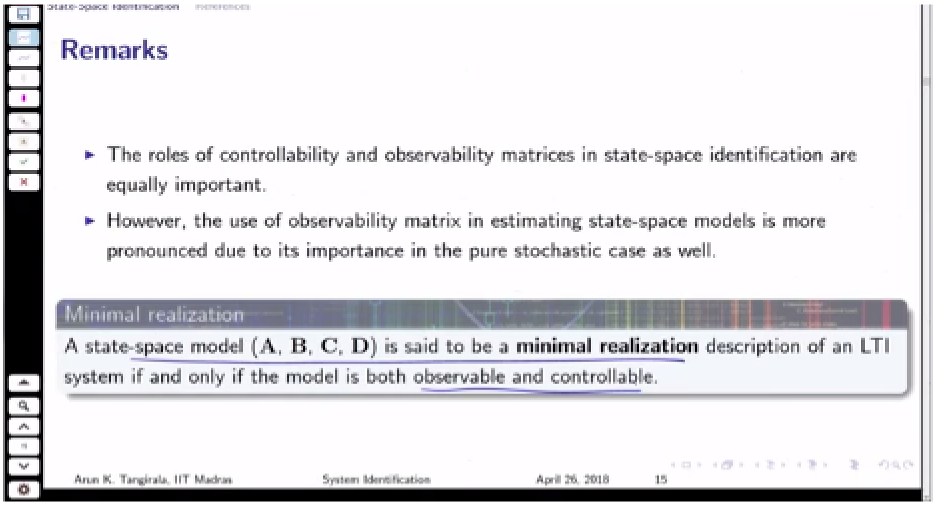
(Refer Slide Time: 32:45)



So the roles are controllability and observability are equally important, but the observatory matrix is very is more common and popular in the state space literature, because of the time series analysis case, the time series case. In time series modeling, there are no inputs. That means I don't perform any experiments. See, when I perform experiments, on my own the controllability comes into play. But if I'm not performing the experiment, someone else is doing either the system is doing by itself like nature atmospheric process, or some other individual has done it for me, or maybe routine operating data, the data is being generated, then there is not much I can do. Then, I regardless of whether an experiment, I have access to the experiment or not, my ability to infer the states from data is governed by observability. So the final step is inferencing and therefore, observability is the more commonly encountered concept in the state-space literature, at least the state space identification and state estimation literature.

Now at this stage, it's nice to now formally define what is the minimal realization? We have spoken about this earlier as well, long ago. Now, we can give a formal definition for the minimal realization. And this is due to Kalman. A state-space model is said to be minimal realization, if and only if, the model is both observable and control. So, what happens if it is not observable? Then there are states in the system which you cannot estimate and that means there is some redundancy there that you shouldn't be having, likewise for controllability.

(Refer Slide Time: 34:40)



So, this is to make sure that there is a formal way, this is a definition. This is can be also used for checking, if a given system is given state-space description is of minimal realization, right? So, keep this in mind and we'll move forward.