

**NPTEL**  
**NATIONAL PROGRAMME ON**  
**TECHNOLOGY ENHANCED LEARNING**

**IIT BOMBAY**

**CDEEP**  
**IIT BOMBAY**

**ADVANCE**  
**PROCESS CONTROL**

**Prof. Sachin Patwardhan**

**Department of Chemical Engineering,**  
**IIT Bombay**

**Lecture No. 12**

**Development of Control Relevant Linear**  
**Perturbation Models (Part 2)**

**Sub-topics**

**Issues in Model Development (contd.) and**  
**State Realizations of Transfer Function Models**

So we have been looking at prediction error method and I will just take a quick review again of where we stand now.

(Refer Slide Time: 00:26)

Automation Lab  
IIT Bombay

## Prediction Error Method

Given data set  $Z_N = \{y(k), u(k) : k = 1, 2, \dots, N\}$

Model:  $y(k) = G(q, \theta)u(k) + H(q, \theta)e(k)$

Optimal one-step predictor

$$\hat{y}(k | k-1) = H^{-1}(q, \theta)G(q, \theta)u(k) + [1 - H^{-1}(q, \theta)]y(k)$$

One step prediction error is defined as

$$e(k, \theta) = y(k) - \hat{y}(k | k-1, \theta)$$

Parameter Estimation by Prediction Error Method

Find  $\theta$  that minimizes objective function

$$V(\theta, Z_N) = \frac{1}{N} \sum_{k=1}^N e(k, \theta)^2$$

CDEEP  
IIT Bombay

We have this data set  $y$  and  $u$  okay you want to develop a model which is in general of this form  $y$  is smaller with respect to the known inputs typically the manipulated inputs and  $x$   $q$  is model with respect to unknown measured unknown disturbances. Now  $u$  here can include measured disturbances if you make some simplifying assumptions so  $u$  is what I that is why I kept keep saying known inputs okay so if you have measure disturbances there are ways of modifying this model to we have to make certain simplified assumptions of course.

But because the input manipulated inputs that go out of a computer it quite constant the disturbances that you measure are not  $p_2$  is constant even if you measured disturbance say for example you have some you know feed water use for cool drink some system and if you are measuring the feed water temperature feed water temperature throughout that I keep changing because atmospheric temperature keep changing even if you measure it truly speaking it is not  $p_2$  is constant.

But you are measuring fast enough you know you can make an assumption that there is  $p_2$  is constant and model so that  $u$  here could be known inputs or it could be manipulated variables which are  $p_2$  is constant, this part is everything that is not explained by the known inputs that is the only correct interpretation of this so this component actually captures unknown disturbances okay it captures measurement errors it captures errors because of approximations you are actually having a model which is linear model the true system might be easy in general in non linear very rarely our linear system is linear perfectly linear.

So this is something that captures everything that is unknown everything that is not capture by this component okay the way we modeling is to use one step I had predict studies so we develop this one step higher predictor as a part o exercise you will be act5ully developing more such predictors so developing one step higher predictors for different simple forms is part of the exercise that we have going to do tomorrow.

And we estimate this solution error okay this method is called as prediction error method because we minimize the sum of the square of prediction errors okay, this is prediction error this is  $y_k$  is the measurement at instant  $k$  okay why  $k - 1$  is sorry  $y^k$  is given  $k - 1$  is prediction of  $y$  based on measurements available up to  $k - 1$  that is the okay. So this notation we are going to use throughout the course  $k$  given  $k - 1$  means prediction of  $y$  using measurements available or up to time  $k - 1$   $\theta$  is the model parameters that you need to estimate.

And then I was saying that we minimize some of the square of errors we minimize the variance okay there is nothing will nobody scotch you from minimizing some of absolute error so you can do that okay or minimizing maximum error over the that is infinite now you can some other function in ge3neral to norm as some special problems which I am going to discuss today, why this two norm is so important why we can get some insides in to parameter estimation if you happen to use two norm okay that is why we want to use two norm.

So this method is called as prediction error method so what is the other method what are the other approaches to do system identification modeling there is one more method which is based on projections okay and this method is known as sub space identification method it has become very popular in last 10 to 15 years it is just based on projections ideas of projections. So nice thing is you know you can just use simple matrix projections to come up with the model.

So whereas here you are use non linear optimization since we are using non linear optimization here it is very important at you give a good guess if you do not give good guess okay how to give a good guess okay right now you know lanes tool box is doing it for us when you give data it use you model okay, there is never any problem for two models one is ARX model other is SOIR model which we have been looking at in the these two models are very easy to identify from data.

And that is why they are very popular in the industry but the trouble with these models is that you need a large number of parameters, so you will need these large data sets so you need to know your content they explain for a longer time which is a loss of production so the models which are easy to develop have some trouble associated with it okay, models which are difficult to develop have some other trouble you know the trouble is shifted from the experiments long experiments to difficulty in solving but probably difficulty in solving okay is easier to deal with than longer experiment, longer experiment means loss of production okay which means it is money okay.

But difficulty in solving a problem is offline okay you can collect data and do some tricks to meet the problem give a good guess and so on, so if you ask me what should you do whether you should go for ARMAX or box Jenkins model or ARX model I would say you should go for ARX model or go for ARMAX or box Jenkins model okay and try to give a good guess for example even if you have small data you first try to create or identify an ARX model which will be bad you know because the data size is small but you can use that to create a good guess for your you know ARMAX or box Jenkins model and then proceed so if we use your knowledge intelligently you can actually plan your experiments very well and save money that is important okay now let us get into the properties of the model.

(Refer Slide Time: 08:22)

**PEM: Parameter Estimation**

$\hat{\theta}_N = \underset{\theta}{\text{Min}} V(\theta, Z_N)$

Typically, the resulting parameter estimation problem is solved numerically using

- (a) Nonlinear optimization
- (b) Gauss Newton Method

If it is desired to emphasize certain frequency of interest, then, we can minimize

$$V(\theta, Z_N) = \frac{1}{N} \sum_{k=k_1}^N [\varepsilon_r(k, \theta)]^2$$

where  $\varepsilon_r(k) = F(q^{-1})\varepsilon(k)$

$F(q^{-1})$  represents a band pass filter

SEPTES 2/17/2012 System Identification CDEEP IIT Bombay 121

So I talk reason about these steps in the model development and we also said something about model structures selection, then I told you about this one particular thing here is that I mean what I have introduced to you just tip of the eyes just you know beginning what, what is the system identification actually you should not just stop at my notes I have uploaded two more documents in model I do not know how many of you seen them.

One is slides by professional why particularly professional is a well known that authority in this area there is a very nice statement book I have mentioned the book here systematic identification I personally I do not feel it is book for beginners okay the other than book which I have mentioned here is soderstrom and soica that is the better book for beginners the way he introduces his much more easily.

But Loons book is you know one of the standard reference anytime you get outs bout anything in systematic deification you can always go back to eliminate will see that he has discuss it you only realize that such a question of such problem exist may be after one or two years but this tool box mat log tool box has been actually written by per volume so there is a compressed version of this lecture notes about 30 40 pages which also uploaded there.

So there are two things one is presentation slides of a workshop he conducted university in 2004 and a second one is you know is a convinced lecture notes or systematic defecation I think there are some 50 60 pages so those are few who are going to use this techniques in future where all

these techniques are used everywhere I mean from you know somebody decides to go in to finance and then you know wants to do share price modeling you can do that it is time series.

It is a you can modulate as a stationery or non-stationery process you can review it as a stationery you can find a transfer function even by white noise that you know tells you about how the share price is we have a model for share price fluctuation and then what is  $k$  here you can take today's price tomorrow's average price yesterday average price question is can I predict if I have a model why do we develop models.

We can do predictions we can forecast okay so that part will come little later when we start using this models for forecasting now you know when you have this  $y - \hat{y}$  what is this  $\epsilon$ ,  $\epsilon$  is  $y - \hat{y}$  okay now you may want to suppress certain frequencies that have there in  $y - \hat{y}$  but say  $y$  as measurement errors which are very high frequency.

You know that these are not going to be useful in modeling you can filter those errors using a low pass filter or band pass filter and that is why you can actually have an objective function in which you minimize filter error and not the directly innovations or error itself this, this signal  $y - \hat{y}$  or this  $\epsilon$  is many times called as residuals model residuals it is also called innovations okay.

(Refer Slide Time: 12:12)

Automation Lab  
IIT Bombay

## MIMO System Identification

- **ARX Model:**  
Method for ARX parameter identification can be extended to deal directly with multivariate data
- **OE / ARMAX / BJ Models:**  
Typically, an  $n \times m$  MIMO system is modeled as  $n$  MISO (Multi Input Single Output) systems

$$y_i(k) = G_{i1}(q)u_1(k) + \dots + G_{im}(q)u_m(k) + H_i(q)e_i(k)$$

$i = 1, 2, \dots, n$

MISO models are combined to form a one MIMO model

**Input excitation**

Inputs can be perturbed sequentially or simultaneously

CDEEP  
IIT Bombay  
124

RIPTER  
2/17/2012  
System Identification

So modeling error whatever you want to call it residual is very, very commonly used model residuals yeah,  $\epsilon$  we stop when you  $\epsilon$  is white noise but not  $\epsilon$  has a filter how can I filtered signal given by white it must be colorless  $\epsilon$  yeah, yeah so whether by minimizing  $\epsilon$  is will you will be able to find  $\epsilon$  which is white by probably choosing order you should be able to find  $\epsilon$  that is my you know should be possible to whether it is guaranteed always is I have to go back and check.

I think it should be possible to find out just because you are minimizing the filter value does not mean yes, you can go back to  $\epsilon$  here see here you are minimizing the filtered value but that filter dose not enter your model anywhere okay filtered see that filtered value is only use to certain frequencies so has to emphasize particular frequencies in your model okay.

So actually hen you use that model the way it will enter probably in your model is afterwards if you want to use the model you probably have to use filter  $f$  and filter  $u$  okay so it will translate to filter  $f$  and filter  $u$ , no  $\epsilon$  is cannot be white noise so by definition  $\epsilon f$  is a correlated noise there is a transfer function multiplying  $\epsilon$ ,  $\epsilon f$  is cannot be white noise  $\epsilon$  can be white noise okay.

No, no what I was what I disrupt here question last time was if you minimize the subjective function is it possible to get an  $\epsilon$  which is white noise I think it should be possible to get but what are the guarantees and all that we have that it should correspond to the system band width and the system band width is something which you want to engineer you know which is the knowledge which should know that in what band width I should which band width is relevant for my control which band width I should cut off.

That is why this choice of this filter here can play a very critical role in identification so this modeling in some sense is not completely that box when you chose this filter you have to have okay or when you chose a model order you have to have some idea about so in some sense it is grey box modeling you are no, no so you have to go and do some preliminary experiments with the system.

Now you look at the data and do some spectral analysis you cannot do it just like that so as I mentioned multiple input multiple output systems ARX models can be very easily adapted can be change ARX modeling scheme for multiple input multiple output system trouble be the ARX is large number of parameters other possibility is that output error are max typically they are developed as multiple inputs in single output models.

So if we have, if we have a system which has two inputs and two outputs you develop two models for output 1 and 2 inputs output 2 and 2 inputs and then you club them together I am not going to talk about how do they club together today okay.


(Refer Slide Time: 16:17)



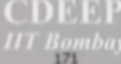
Automation Lab  
IIT Bombay

## Frequency Domain Analysis

- Time domain formulations of parameter estimation problem
  - Useful for carrying out parameter estimation
  - Does not provide any insight into internal working of optimization problem
- Frequency domain (power spectrum) analysis
  - Based on Fourier transform of auto-correlation and cross correlation function of signals
  - Powerful tool for analysis (analogous to use of Laplace transforms in linear control theory)
  - Provides insight into various aspects of optimization formulation
  - Can be used for perturbation signal design and estimation error analysis


 2/17/2012

System Identification

  
 171

Now even if I want to do this part very briefly I would at least three four lectures or three lectures probably but I am going to just talk about it very, very briefly I am not going in to details if you are want to know more about this I have notes here of the end appendix, this appendix here this appendix here I have explained.

You know basis for this analysis that frequency analysis the analysis sis to get inside into how parameters you know into these two things one is bias error and a variance level and what is bias error and variance error I will talk about it I will talk about the final expression which are derived after all this analysis okay.

Now if you for the time being take engineers approach and say that as well there us a derivation which is true let me concentrate on the final result and let me see how I can use it analyze it your two understand how to plan my experiment if I can do that, that is enough okay because ultimately given a few understand all the derivations finally how to use that particular result to plan a experiment is more important.

Not how you add with the derivation okay so as a engineer I am more interested in final result which I can use to analyze so I am going to directly talk about final result so this is power spectrum analysis and this is based on the Fourier transform the auto correlation and cross correlations so very, very powerful to lend here you are able to do this because you are using too long.

Because Fourier transform you can take and talk about you know interpretation in the frequency space because you are working with Hilbert's space where two norm is you know available to you and you can move back and forth between different reference time domain and frequency domain reference and then or view points.

(Refer Slide Time: 18:32)

**Asymptotic Bias and Variance Errors**

In reality, true model order is not exactly known, data length is finite and data contains unmeasured disturbances / noise. This results in two types of errors in estimation

$$G(q) - G(q, \hat{\theta}_N) = [G(q) - G(q, \theta^*)] + [G(q, \theta^*) - G(q, \hat{\theta}_N)]$$

$[G(q) - G(q, \theta^*)]$ : structural or bias error induced by fact that model set is not rich enough to exactly characterize the plant

$[G(q, \theta^*) - G(q, \hat{\theta}_N)]$ : noise induced or variance errors due to unmeasured disturbances / noise

$$\left. \begin{array}{l} \text{Total Error} \\ \text{of Estimation} \end{array} \right\} = |\text{Bias Error}| + |\text{Variance Error}|$$

System Identification

And so skipping this long story short we just want to look at these two what kinds of errors that can occur it okay when you identify models one data okay there are two types of fundamental that two types of errors for the time being actually I would see the three types of errors but the third one which comes because of approximation of a non linear system with the linear system.

Let us ignore that right now let us assume that true plant is perfectly linear okay under that situation what are the errors that can occur one error is that plant is perfectly linear I got data I do not know what is the order of the plant so I choose some guess you know second order transfer function, third order function, fourth order transfer function this is my guess.

And ultimately I am going to use some criteria like kind information criteria and mix some column which model is good okay so I do not know what is the truth okay so one error so what I have done here is that let us say  $G$  is represents  $g$   $\theta^*$  represents the model transfer function that you are estimate from the data and  $Gq$  is the true is the true this is so I am putting this  $\theta^*$  and  $\theta^*$  in between okay.

So this difference, this difference is between you know what I would say the structural bias, the structural bias come because that true model is different okay true model has let us say 7 parameters and I am using a model with two parameters yesterday we had we had seen one problem okay.

We are trying to identify using one parameter so similar situation that is because you did not know what is true order you guess okay so now first type of error is a structural error that true differential equation is 10<sup>th</sup> order you modeling at the third order so that structural error so whatever you do a third order differential equation cannot imitate behavior of the 10<sup>th</sup> order equation okay.

You can make them bring them close but 10<sup>th</sup> order is 10<sup>th</sup> order and 3<sup>rd</sup> order is 3<sup>rd</sup> order, 3<sup>rd</sup> order cannot imitate 10<sup>th</sup> order beyond the point so this is first thing that you have to know that is the structure of the error when you are identifying the model second error comes suppose you know the structure okay.

What is the other type of error other type of error come because of variance error, variance error are because of the data length okay you know that you know that if you take infinite data you will get perfect model okay but you cannot taking infinite data you cannot run the test for quadrations for infinite time you have to stop you have to take finite data so finite data will give rise to errors which are called as variance errors.

So this second type of errors are introduce because of the variance errors okay so total error and estimation is a combination of bias error and variance error bias error comes because of structure and this match variance error comes because of limitations finite data okay they also come of course because of unmeasured disturbance as a noise.

But the two things are tightly and you can if you want to reduce the influence of noise on the estimation you better take last data so will see what the relationship is just looking at a final expressions.

(Refer Slide Time: 22:39)

Automation Lab  
IIT Bombay

## Bias Error: Concept

Real systems are of very high order and model is always chosen of lower order  
**Thus, bias errors are always present in any identification exercise**

Classic Example in Process Control

Process Dynamics:  $G(s) = \frac{1}{(10s+1)^2}$

Identified FOPTD model:  $\hat{G}(s) = \frac{1}{(50s+1)} e^{-36s}$

CDEEP  
IIT Bombay  
129

2/17/2012 System Identification

Let me again go over this idea of bias error what is this bias error concept? In process control this is very, very popular model this model okay this model form if I go back.

(Refer Slide Time: 23:01)

CDEEP  
IIT Bombay

CLASS / Slide

$$y(s) = \frac{k_p}{Ts+1} e^{-\theta s} u(s)$$

FOPTD

$$y(s) = \frac{k_p}{\tau^2 s^2 + 2\tau\zeta s + 1} e^{-\theta s} u(s)$$

SOPTD

$$y(k) = \frac{\beta_1 q^k + \beta_2 q^{-2}}{1 + \alpha_1 q^k + \alpha_2 q^{-2}} q^{-d} u(k)$$

CDEEP  
IIT Bombay

And show you here there are two models form which are very, very often use in chemical process control one is y or I will write it in terms of Laplace domain you can convert it into time domain or in to discrete time Laplace domain this model is  $ke/Ts+1 e^{-\theta s}$  okay this is called as

FOPTD, F is first order okay, O is first order okay first order with the time delay okay and I think P is for single time process pole okay is the first order with time delay model okay.

You have one time constant okay and you have again time delay and time constant very simple model okay many times useful to approximate high order system you have some distillation column which is 100 order system you do not want to model a 100 order system you model it has this the order model which is very, very of course it will be us the other model which is very popular is called as FOPTD so this is  $kp/T^2 S^2+2 TGS +1*e^{-\theta s}$  this is the second order + time delay I think P stand for plus second order+ time delay FOPTD.

So this is SOPTD and this is first order+ time delay model okay so these models are very, very popularly used these are low order models first order or second order trying to use one pole or two poles to approximate you know the system which is high order okay of course when you convert this into discrete form this will be  $y_k = \beta_1 + q^{-1} + \beta_2 q^{-2} / 1 + \alpha_1 q^{-1} + \alpha_2 q^{-2}$  this SOPTD model when you convert it into discrete form you will get the second order difference equation.

This is second order differential equation this is second order difference equation okay so there inter convertible but again here you are trying to everything using  $\beta_1, \beta_2$  and  $\alpha_1$  and  $\alpha_2$  and then this is given the true system might be very high order you tend to use a small order model see suppose you have a system which is multiple input and multiple output between each input and output pair you tend to assume a model which is of this form or this form.

Because you know overall order of multiple input multiple output starts going up okay we will see that how it happens so often this smaller dimensional form is convenient okay but now what is the trouble? Why it is and then there is in the books all process control you will find special methods to identify these models this  $kp/\tau$  and  $\theta$  from some step change and all that okay.

So this has been very a popular method of modeling where is the trouble so let us look at a scenario where you have 8<sup>th</sup> order transfer function this is  $1/10s+1$  okay I did this identification exercise and I dissolved to model it as a first order+ time delay this is my time constant and this is the time delay this combination approximates this transfer function okay.

(Refer Slide Time: 27:30)

**Bias Error: Concept**

Automation Lab  
IIT Bombay

Real systems are of very high order and model is always chosen of lower order

Thus, bias errors are always present in any identification exercise

Classic Example in Process Control

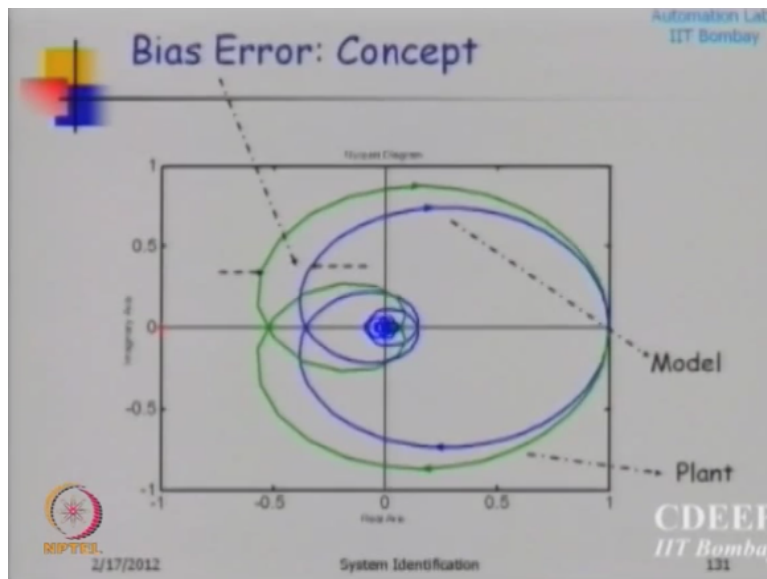
Process Dynamics:  $G(s) = \frac{1}{(10s + 1)^3}$

Identified FOPTD model:  $\hat{G}(s) = \frac{1}{(50s + 1)} e^{-3s}$

IPTES 2/17/2012 System Identification CDEEP IIT Bombay 129

And typically how do you check in process control it is typically you check the step responses seem to much pretty well okay the again is correct okay that is see this I blue line is the approximation and this green line is the true plant and you will say well this is not bad optimization okay it is matching quite okay in a step responses moving to compare the frequency responses you see what is the problem?

(Refer Slide Time: 27:57)



Okay frequency response do not match okay frequency response are matching in this region right this is the low frequency region because there is the good match in the low frequency region you know you see good match in the step response but if you inject a signal which is high frequency okay.

Then there will be significant mismatch between the model behavior and the plant behavior now suppose it happens that when you are operating the plant the frequencies of the plant lie in this region then there is the big mismatch okay mode is wrong and you will use model for doing all kinds of things you use model for control so you are using a wrong model in this frequency region okay.

Now that there are two solutions one solution is not to use first order model okay use 8<sup>th</sup> order model but in this case I know the order is 8 in near plant suppose the order is 100 I am going to use 100 model you know identifying 100 time constants of a 100<sup>th</sup> order model will be difficult it is not so easy task okay, so I do not want to use 100 model I am not comfortable with it. I want to use still first order model, or 2<sup>nd</sup> or 3<sup>rd</sup> maybe you know not a very high order model okay.

Then the question is for this particular approximation the good match is in this region but I know the real one frequency is in this region, can I shift this match from here to here, I do not mind if there is a mismatch here but I want good match you understand what I am saying? So what I am trying to say now let us be practical that the real system is high order.

I am always going to develop a low order model so there is always going to be this structural mismatch between the truth and the model. Now in which frequency band you want the mismatch to be low and in which frequency band you want the mismatch to be high that is your choice okay but how do you decide, how do you analysis that? That is where this frequency domain comes into the picture.

(Refer Slide Time: 30:29)

Automation Lab  
IIT Bombay

## Bias Error: Interpretations

$$\lim_{N \rightarrow \infty} \text{Bias} = \int_{-\pi}^{\pi} \left[ \phi(e^{j\omega}) - \hat{\phi}(e^{j\omega}) \right]^2 \Phi_{\epsilon}(\omega) \cdot \phi_{\epsilon}(\omega) \frac{1}{|\hat{H}(e^{j\omega})|^2} d\omega$$

- Bias distribution of  $|\phi(e^{j\omega}) - \hat{\phi}(e^{j\omega})|^2$  in frequency domain is weighted by Signal To Noise Ratio
- Input spectrum can be chosen intelligently to reduce variance errors in certain frequency regions of interest

MPTES 2/17/2012 System Identification CDEEP IIT Bombay 133

This derivation which I am presenting on this one page would probably require a lot of time to go through if you start doing it but is a lot of thing you have to accept finally we will just look at the result okay. You just go back into the appendix and try to see whether you can understand the derivations. Everyone on here is with me on this expression how do you get error between, see you have to.

Yeah we have this expression now substituting for the truth and for the estimate okay doing some algebra okay which I am not going to do right now here okay. You can show see what is what are the terms here, just tell me what is h? Estimate of when we use  $\hat{\cdot}$  what is the convention? What is  $\hat{g}$  estimate of the truth okay.

And when I say g it is the truth okay, so I have expressed the prediction error in terms of 4 things what are the 4 things, true g true h okay, estimated g, estimated h do again with some algebra just believe this okay. you can see what is y is function of true g true h what is  $y \hat{\cdot}$ ? Function of



estimated  $g$ , estimated  $h$  and what is epsilon difference between 2 okay, just believe me in that you can do some algebra and get this expression okay.

How I am going to use this? Now what is the variance, how do you estimate the parameters by minimizing the variance okay is everyone with me on this, I am minimizing the variance and estimating the parameters. Now I am taking finite data length right, when I actually do estimation on the parameters I am taking finite data length.

But doing analysis with finite data length is very difficult you have to do analysis by taking limiting case, so what is my limiting case? As  $n$  tends to  $\infty$  okay, what is this quantity variance, what is  $R_0$ , Variance of the signal, variance of epsilon okay? Any doubt upto this point okay, now is the trick or trouble, so you can using this Parseval's theorem you can interpret you can interpret this quantity.

You can transform this quantity limiting quantity  $z \rightarrow \infty$  this quantity you can convert into a frequency domain okay. Now I am going to use this expression for converting into frequency domain okay and if I convert I will finally get this particular term okay. So minimizing this quantity, see I have taken spectrum of epsilon which is given by this quantity here, spectrum of epsilon.

Which is given minimizing what is see I have taken spectrum of epsilon which is given by the quantity here, spectrum of epsilon is when you is when you have transfer function you can estimate the spectrum of by putting  $q = j\omega$  okay they are putting  $q = j\omega$  you can estimate the details are given in the appendix for this go back and check, so I can relate this variance with this, I can convert into spectral domain.

And what is this 5, spectral density of  $\epsilon$  that is given by this quantity why this quantity? It is coming from here okay now let me explain looks very complex when you see for the 1<sup>st</sup> time, but now why it is useful why am I saying that this going to be useful. Now look at it, look at things here, what is this quantity? What is  $e^{j\omega t}$  this is frequency response of true.

What is this? Frequency response of estimate, so I am saying that difference between this frequency response and response is weighted by spectrum of inputs okay, so what is the consequence? So this difference is weighted by see if you are doing optimization okay, in

optimization if you have some of the square of certain terms okay. Some terms have higher weight age and some are lower.

What is the tendency? Tendency of the optimizer, wherever there is the higher weight age it will try to reduce that term more and more, where ever there is lower weighted it will not bother about optimizer, see because if it tries to change that variable the objective function does not change, you get my point. See if you have an optimization problem which as different components some components have higher weighted, some components have lower weighted.

Now the tendency of the optimizer see wherever there is higher weighted optimizer will try to bring you know minimize that component is more sensitive to that component. Now how this difference see actually behaves in the estimated model depends upon how you choose this frequency spectrum, okay so actually this frequency spectrum shipping can be used to shift this difference to different zones.

(Refer Slide Time: 38:12)

Automation Lab  
IIT Bombay

## Bias Error: Interpretations


$$\lim_{N \rightarrow \infty} \left[ \sigma_N \right] = \int_{-\pi}^{\pi} \left| \hat{\theta}(e^{j\omega}) - \hat{\theta}(e^{j\omega})^2 \right|^2 \Phi_u(\omega) \cdot \Phi_v(\omega) \frac{1}{|\hat{H}(e^{j\omega})|^2} d\omega$$

- Bias distribution of  $\left| \hat{\theta}(e^{j\omega}) - \hat{\theta}(e^{j\omega})^2 \right|^2$  in frequency domain is weighted by Signal To Noise Ratio
- Input spectrum can be chosen intelligently to reduce variance errors in certain frequency regions of interest

For Output Error model (i.e.  $H(q)=1$ ),

$$\lim_{N \rightarrow \infty} \left[ \sigma_N \right] = \int_{-\pi}^{\pi} \left| \hat{\theta}(e^{j\omega}) - \hat{\theta}(e^{j\omega})^2 \right|^2 \Phi_u(\omega) d\omega$$

Thus,  $\hat{\theta}(e^{j\omega}) \rightarrow \hat{\theta}(e^{j\omega})$  if model is not under-parameterized


2/17/2012
System Identification
CDEEP  
IIT Bombay  
133

See let us go back to I have these 2 signals input signals, and what is there corresponding power spectrum this is their power spectrum, so what is the meaning of this? This power spectrum as

low power at low frequencies it has high power at middle frequencies okay. Whereas this signal as high power at low frequencies and almost no power at middle and high frequencies okay, now if I plan using this signal.

Then let us go back to this which as high content at low frequency okay then optimizer will work in such a way that this difference is small at low frequencies okay. So the model is good the frequency is matching in the low frequency region okay and it does not bother about matching the frequencies responses in high frequency region why? Because spectrum this integral is weighted by this spectrum right.

This spectrum I slow at high frequencies and middle frequencies so the tendency minimizing the optimization function in the time domain would implicitly do reduce frequency domain mismatch at low frequencies. This is the insight which this equation gives okay this is not possible for one norm or infinite norm this is possible with 2 norm, why 2 norm? We can convert this and get into the frequency domain do this analysis.

Even if I derive this equation finally I am going to say this is the important part of it, how do you derive it this equation does not bother about it right now okay. So it tells you this frequency domain expression tells you how to plan your experiments, very critical. See if I use this signal, if I exit the plant using this signal then it emphasis middle frequency and this high frequency and low frequency are not so important.

What would happen is this right now this model match is very good here, it will shift from here to here, and there will be mismatch at low frequencies. If I use white noise I will get, white noise as the entire frequencies okay. As I told you that white noise using for perturbation is only in computer simulations, you cannot do it in reality okay. So even though it is ideal it is not practical okay.

So that is the problem, since it is not practical you are force to make a frequencies choice, what is the frequencies choice you made? What is the frequencies range of your interest? How does it influence the parameter estimates and the frequencies response that is given by this expression? This expression tells you that I can shape this difference by using shaping the input spectrum, if that is the only message that I want to take there is nothing more even if you understand the derivation finally you have understand this.

I can shape this difference, this interpretation is possible only because of theorem, only because of using two norms Fourier transform and then you can interpret this into frequency domain and say that well. So how I plan my input okay is can be understood through this analysis. That is why Luks book filled with frequency domain analysis along with time analysis. So this entire complex expression the take home message is only this.

The input spectrum can be chosen intelligently to minimize difference between the frequencies response of the truth and the model in certain okay, now what is the effect of adding that filters if you add that filter and do all the calculation the spectrum will come here. I talked about this filter okay it will turn out that this filter okay. It will turn out this filter is another way of shaping, see one way is to shape the inputs okay.

Other ways to choose is shaping filter spectrum of shaping filter will appear in this equation, if you do that filtering the spectrum of that signal will appear in this expression and so if you have not choose inputs correctly you can choose shaping filter correctly and try to main the error by choosing the signal. So there are tricks which you can see only when you go to this frequencies domain it is not possible to see this.

(Refer Slide Time: 43:49)

**Variance Errors**

Asymptotic variance of estimates using PEM are

$$\text{Var}[\hat{G}(e^{j\omega})] = \frac{n}{N} \frac{\Phi_y(e^{j\omega})}{\Phi_u(e^{j\omega})}$$

$$\text{Var}[\hat{H}(e^{j\omega})] = \frac{n}{N} \Psi(e^{j\omega})^2$$

n: Model Order    N: Data Length

CDEEP  
IIT Bombay  
124

The other thing is done in Luke book in detail is that variance error, so what you can show is that variance of estimate frequencies response, variance of response okay. You can think of it in terms

of possible error band in the frequency response. Variance will tell you what possible error band this related to these two terms,  $n$  here is number of model parameters okay.  $N$  is the data length.

$N$  is the data length okay and this is noise spectrum and this is input spectrum okay this ratio of noise spectrum to input spectrum is called as noise to signal ratio okay. Just look at this expression and tell me how will you make variance error small? One way is to choose this  $n$  use less number of parameters if you are using large number of parameters better chooses large  $N$ , apart from this you have one more parameter that you can manipulate.

What is  $\sigma_v$  here,  $v$  is the noise okay, and noise spectrum is there in the plant, what is in your choice, input spectrum. So I can choose input spectrum in such a way that noise to signal ratio becomes insignificant, so if the signal dominant over noise this ratio is small variance error is low okay. If my input spectrum dominates over the noise spectrum then when you do this modeling and perturbation people will talk about.

Signal noised ratio is other way around  $5d$  is signaled noised ratio, so whichever the way I mean sometimes people use noise to signal ratio. So you have to make as small as possible or signal to noise ratio as large as possible okay signal should dominant and this insight does not come looking at some of the square of errors, this will come only when you look at expression.

That is why frequency domain it is quiet important when it comes to, so how do you reduce the variance errors, by choosing large data length okay, by choosing correct signal noised ratio to be large  $\phi_u / \phi_v$  is large and then, so all this analysis is extremely important while getting a good model. Now I am just giving you the final bits for what is useful all scenarios is extremely important in developing good model.

You cannot blindly use tool box as which are available now, unless you understand these entire theory okay and these just to you about this, there is lot more to this. You can get whenever you want, so you can shift match. See the idea is that you have live with the bias because the real world problems are very high order.

See the take home message from this is that real world problem is actually high order this is always going to which is low order. So the bias is going to be there, bias is part of identification. So all that you can do is to shift the emphasis from which frequency band you want good match

in which frequency band live with mismatch okay. How do you do that while choosing input spectrum.

So there are different ways of handling this if you look carefully this h is actually a shaping filter for this mismatch, comes here so this is signaled noise ratio this is signal and this is noise spectrum it appears here. So signaled noise ratio has shapes, so many times people says noise modeling is not because you want to identify disturbance model is noise modeling is because you want good model.

So noise modeling is the way of shaping this term okay that is the idea so let us move on this brings to the end on lectures on systematic identification hopefully will solve the problems and when you start doing actual simulation that is when you learn much more about this than just these lectures. So let us go back now we want to go to control okay.

(Refer Slide Time: 49:46)

Automation Lab  
IIT Bombay

## State Space Realization

Consider 3<sup>rd</sup> order SISO T.F.

$$y(k) = G(q)u(k) = \frac{b_3q^3 + b_2q^2 + b_1q}{q^3 + a_2q^2 + a_1q + a_0} u(k)$$

Problem: Derive state space model of form

$$x(k+1) = \Phi x(k) + \Gamma u(k)$$
$$y(k) = Cx(k)$$

such that  $C(zI - \Phi)^{-1}\Gamma = G(q)$

2/17/2012  
System Identification

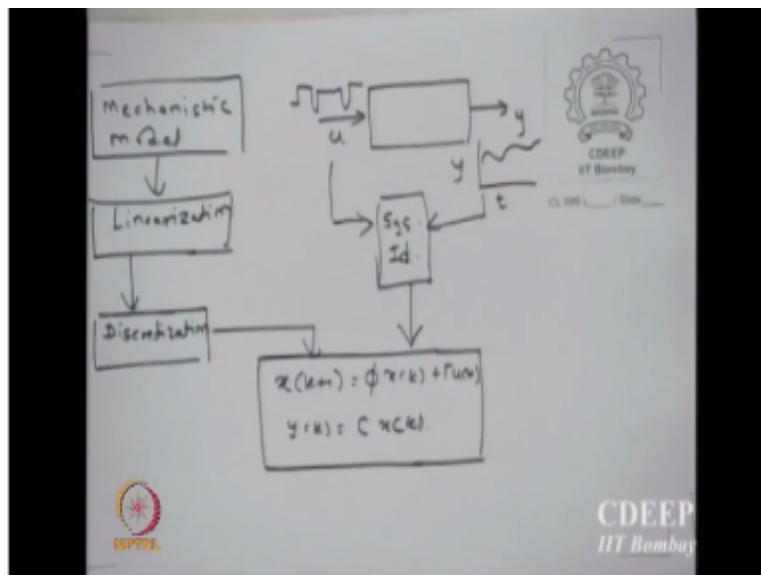
CDEEP  
IIT Bombay  
149

I have spent almost 40% of my time in talking about modeling well in reality when you go to a plant 70% of your time will go in modeling 30% will go in control, once you have good model you are done you are there so now what we will be doing is mostly algebra and that is much easier because now you are in the once you translate the relate into a model which is nice linear difference equation you are in the world of linear algebra you can do all kinds of things okay.

Now as I said there are two viewpoints domain in control one is transfer function view point the other one is state place view point and I do not want to profess one or the other I belong to the state place point so I like to work with state place it is all legal algebra and simple linear algebra of matrixes so I am going to convert my module back into state space form I like state space form so I want.

I identified this transfer function okay we will take about how to deal with ek and all that right now let us look at one transfer function single input single output how do I convert this into this standard form that is my question okay so afterwards I am not going to bother about how did you get this standard form we have this standard form

(Refer Slide Time: 51:31)



It is quite likely that I have a mechanistic model then I did linearization and then I did discretization and then I got this model okay or you know I had this system to play with these are the inputs these are the outputs I introduce some fluctuations I recorded the output as a function of time and then using this you know input and output data and using some system identification tool I come up with this okay I do not care finally how do I come up with this model.

I could have I identified ARMAX model you know model output error model whatever you choose finally I am going to convert into this form I am going to work with it okay so which root you came to this form afterwards is not important okay so we could come to this form any okay that is not going to matter so I am going to talk about one possible way of doing realization other

possible ways are given in the notes I will explain one okay which is little complex to understand first we will try to see whether we can cover first two today.

(Refer Slide Time: 53:23)

Automation Lab  
IIT Bombay

## State Space Realization

Consider 3<sup>rd</sup> order SISO T.F.

$$y(k) = G(q)u(k) = \frac{b_2q^2 + b_1q + b_0}{q^3 + a_2q^2 + a_1q + a_0}u(k)$$

Problem: Derive state space model of form


$$x(k+1) = \Phi x(k) + \Gamma u(k)$$

$$y(k) = Cx(k)$$


such that  $C(qI - \Phi)^{-1}\Gamma = G(q)$

Canonical realizations

- Observable canonical form
- Diagonal canonical form
- Controllable canonical form


2/17/2012

System Identification



But doing this is not so difficult it is pretty easy what the aim is? Aim is to choose five  $\gamma$  and  $z$  in such a way that  $c$  times  $z_i^{-5}$  inverse  $\gamma$  is same as or not 0 should be  $q$  okay I want to choose  $c$   $\Phi$  and  $\gamma$  in such a way that this  $G(q)$  is the transfer function that is my aim there are different ways of infinite ways of doing it there are no finite number of ways I will prove that also but I am going to talk about some you know in a popular ways and why they are popular also become clear after sometime when you start doing controller developing.

(Refer Slide Time: 54:09)



Automation Lab  
IIT Bombay

## Controllable Canonical Form

Introducing intermediate variable  $\eta(k)$  as

$$\{q^3 + a_1q^2 + a_2q + a_3\}\eta(k) = u(k)$$

$$y(k) = \{b_1q^2 + b_2q + b_3\}\eta(k)$$

we have

$$\eta(k-3) = -a_1\eta(k-2) - a_2\eta(k-1) - a_3\eta(k) + u(k)$$

Defining state variables

$$x_1(k) = \eta(k-2); x_2(k) = \eta(k-1); x_3(k) = \eta(k)$$

we get three first order difference equation

$$x_1(k+1) = -a_1x_1(k) - a_2x_2(k) - a_3x_3(k) + u(k)$$

$$x_2(k+1) = x_1(k)$$

$$x_3(k+1) = x_2(k)$$

System Identification

CDEEP  
IIT Bombay

2/17/2012

So this first form is called as the controllable canonical form okay what I have done here is I have introduced one pseudo variable here you see this pseudo variable  $\beta$  okay I will read it in this equation I have rewritten this equation okay by introducing an intermediate variable which is this  $\rho_k$  okay same equation so I am saying that this operator operating on is effect of  $u_k$  entering the system and  $y_k$  is same transfer function just a trick introduced one more term in between okay now this particular first equation.

Okay is equivalent to this difference equation just check this is  $q^3$  so I will get  $nk+3$  this is  $q^2$   $nk+2$   $q^1$   $nk$  and this is okay yeah all three are negative here so they should be positive because when they come on left on side they will be positive I have taken them on the right hand side so that is why they are negative is everyone let me on this right okay so I am going to define three state variables okay  $x_1(k)$  is  $\rho_{k+2}$   $x_2(k)$  is  $\rho_{k+1}$  and  $x_3(k)$  is  $\rho_k$  now those are few have done numerical methods we have done something.

Similar if you remember converting high ordered differential equation .n sort differential equations right this is equivalent thing for the difference equation okay equivalent for difference equation is this okay this term I am calling as  $x_1$  this term I am calling as  $x_2$  this term I am calling at  $x_3$  okay three states I have defined okay what is this term  $x_1$  yeah this is this term is  $x_1(k+1)$  will be  $\rho_{k+3}$  that is you see this okay I am going to use that next see here what is done  $x_1(k+1)$  is nothing but  $\rho_{k+3}$  okay.

So goes this equation is written in terms of  $x_1$ ,  $x_2$ , and  $x_3$  okay so now do you see I have converted a third order differences equation into 3 first order equations I convert it okay a third order differential equation into three first order equation in general  $n$  third differential equation can be converted into  $n$  first order equations okay is this transformation okay see.

There is a relationship between these two variables the relationship between these two variables that is captured through these  $x_{2k+1}$  will be where written correctly there is a error here okay now just check is it okay  $x_{2k+1}$  will be  $k+2$  so that is  $x_1$   $x_{3k+1}$  will be so I have three difference equation in place of one third order it is not coming here is this okay

(Refer Slide Time: 58:23)

Automation Lab  
IIT Bombay

### Controllable Canonical Form

The above equations can be rearranged as

$$x(k+1) = \begin{bmatrix} -a_1 & -a_2 & -a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$

$$y(k) = [b_1 q^2 + b_2 q + b_3] r(k)$$

$$y(k) = b_1 r(k+2) + b_2 r(k+1) + b_3 r(k)$$

$$y(k) = [b_1 \ b_2 \ b_3] x(k)$$

Above form can be easily extended to develop state realization for single input multiple output (SIMO) system with a common denominator polynomial

CDEEP  
IIT Bombay  
101

IAPTEL  
2/17/2012  
System Identification

Then I mentioned you know I just convert this into what is called as controllable canonical form okay I will just rearrange these three equations these bottom three equations I have rearranged into matrix if you get this particular form which is called construable canonical form and I want to now get  $y=cx$  okay is it is very easy because  $y_k = b_1 q^2 + b_2 q + b_3$  so if you multiply you will get this whi8ch is same as this is  $x_1 k$  this is  $x_2 k$  this is  $x_3 k$  so I got  $y=cx$  okay  $y=cx$  the next one I wanted to go back and read and tell me if there is an error.

(Refer Slide Time: 59:17)

Automation Lab  
IIT Bombay

## Observable Canonical Form

$$y(k+3) - a_1 y(k+2) - a_2 y(k+1) - a_3 y(k) = b_0 u(k+2) + b_1 u(k+1) + b_2 u(k)$$

$$y(k+3) + a_1 y(k+2) + a_2 y(k+1) - b_0 u(k+2) - b_1 u(k+1) = -a_3 y(k) + b_2 u(k)$$

Defining  $x_1(k) = y(k)$  and

$$x_1(k+1) = y(k+3) + a_1 y(k+2) + a_2 y(k+1) - b_0 u(k+2) - b_1 u(k+1)$$

System Identification

CDEEP  
IIT Bombay  
152

In another form called as observable canonical form the derivation more complex.

(Refer Slide Time: 59:23)

Automation Lab  
IIT Bombay

## Observable Canonical Form

$$x_2(k+1) = y(k+2) + a_1 y(k+1) - b_2 u(k+1)$$

↓

$$x_2(k) = y(k+1) + a_1 y(k) - b_2 u(k)$$


which can be rearranged as

$$x_1(k+1) = -a_1 x_1(k) + x_2(k) + b_1 u(k) \quad \dots(3)$$

The equations (1), (2) and (3) can be rearranged as

$$x(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} x(k) + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} u(k)$$

$$y(k) = [1 \quad 0 \quad 0] x(k)$$



2/17/2012

System Identification

CDEEP

IIT Bombay

154

Just go back and read this derivation okay it is not there is nothing fundamentally you know difficult to understand it is an algebra to get another way of state realization what I am going to get if I do this way of state realization.

(Refer Slide Time: 59:44)

IIT Bombay

## Observable Canonical Form

$$x_2(k+1) = y(k+2) + a_1 y(k+1) - b_2 u(k+1)$$

↓

$$x_2(k) = y(k+1) + a_1 y(k) - b_2 u(k)$$


which can be rearranged as

$$x_1(k+1) = -a_1 x_1(k) + x_2(k) + b_2 u(k) \quad \dots(3)$$

The equations (1), (2) and (3) can be rearranged as

$$x(k+1) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_1 & 0 & 0 \end{bmatrix} x(k) + \begin{bmatrix} b_2 \\ b_1 \\ b_1 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x(k)$$



2/17/2012

System Identification

CDEEP

IIT Bombay

154

I have explained the steps here I will get this I am just growing through it very, very quickly yeah I get this form this particular form is called as observable canonical form here we will see that the matrixes are now just change there you know one is transport of the other c and v seem to your stage is placed but, there are different ways of getting state realization there is no unique way of coming up with this state realization that is a important message in different ways you can get state realization all of them will have same transfer function.

(Refer Slide Time: 01:00:22)

Automation Lab  
IIT Bombay

## Non-uniqueness of State Realizations

Consider a square non-singular matrix  $\Psi$

Defining new state vector  $\eta(k)$  as

$$\eta(k) = \Psi x(k) \Rightarrow x(k) = \Psi^{-1} \eta(k)$$

Substituting for vector  $x(k)$  in state space model

$$\eta(k+1) = [\Psi \Phi \Psi^{-1}] \eta(k) + [\Psi \Gamma] u(k)$$


$$y(k) = [\Psi^{-1} C] \eta(k)$$

Defining new matrices

$$\tilde{\Phi} = [\Psi \Phi \Psi^{-1}] ; \tilde{\Gamma} = [\Psi \Gamma] ; \tilde{C} = [\Psi^{-1} C]$$

we have state dynamics in terms of transformed state

$$\eta(k+1) = \tilde{\Phi} \eta(k) + \tilde{\Gamma} u(k)$$

$$y(k) = \tilde{C} \eta(k)$$


CDEEP  
IIT Bombay  
155

MPTTL  
2/17/2012  
System Identification

All of them will have same transfer function in fact you can do any transformation of this state by multiplying by an invertible matrix and you can show that is also so if I get this and if I multiply both sides by invertible matrix okay I will get another you know state space form that is very much possible and even that new space form will have same transfer function transfer function is invariant state realization can be variable okay.

So this simple thing tells you that there are you can transfer into some other form where this intermediate state variable is  $\beta$  and pretty much the transfer function will be the same so you will not get a different transfer function so the realization of a transfer function into a space you got unique there are infinite possible ways you can get it normally we get it using this two forms controllable canonical form observation canonical form because that are very easy to construct okay but you can multiply that form with the invertible matrix you will get another realization all of them will have same transfer function that is the key thing.

(Refer Slide Time: 01:01:47)

Automation Lab  
IIT Bombay

## Non-uniqueness of State Realizations

Consider a square non-singular matrix  $\Psi$

Defining new state vector  $\eta(k)$  as

$$\eta(k) = \Psi x(k) \Rightarrow x(k) = \Psi^{-1} \eta(k)$$

Substituting for vector  $x(k)$  in state space model

$$\eta(k+1) = [\Psi \Phi \Psi^{-1}] \eta(k) + [\Psi \Gamma] u(k)$$

$$y(k) = [\Psi^{-1} C] \eta(k)$$

Defining new matrices

$$\bar{\Phi} = [\Psi \Phi \Psi^{-1}] ; \bar{\Gamma} = [\Psi \Gamma] ; \bar{C} = [\Psi^{-1} C]$$

we have state dynamics in terms of transformed states

$$\eta(k+1) = \bar{\Phi} \eta(k) + \bar{\Gamma} u(k)$$

$$y(k) = \bar{C} \eta(k)$$

CDEEP  
IIT Bombay  
155

MPTCL  
2/17/2012  
System Identification

Any one of them, you take one realization  $c\gamma$  another realization  $c\psi\gamma\psi$  all of them will have same transfer function okay that you can show because this difference matrices are relative to invertible transformations okay if they relate to invertible transformations you can be sure that the it is the same the realization of a so what is the meaning of these states which you get here  $x$  here the physical there is no physical meaning that you can attached to this  $x$  see I got this state.

But vector here  $x_{k+1}$ =this matrix into  $x_k$  plus this vector into  $u_k$  what is the real physical signal here  $u_k$  what is on the real physical signal  $y_k$  okay this  $x_k$  has no physical meaning it is a mathematically construct which helps you to put everything into one standard point okay so we are going to use this  $x$  okay because it is convenient to put into this form and then work with it okay linear difference equation models very well understood.



(Refer Slide Time: 01:03:12)

Automation Lab  
IIT Bombay

## Non-uniqueness of State Realizations

Note that both realizations have identical transfer functions, i.e.,

$$\begin{aligned} \bar{C}[qI - \bar{\Phi}]^{-1}\bar{\Gamma} &= C\Psi^{-1}[qI - \Psi\Phi\Psi^{-1}]\Psi\Gamma \\ &= C[qI - \Phi]^{-1}\Gamma \end{aligned}$$

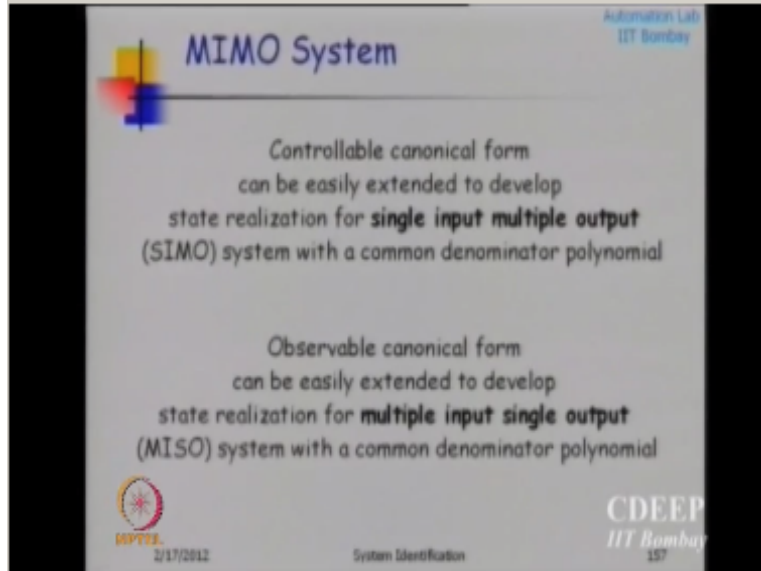



2/17/2012 System Identification 106

Recognitions are non unique there can be infinite ways of realizing the same transfer function into different  $\gamma$   $c$  matrices all of them will give you identical transformation that is the message yeah I have come good question.

(Refer Slide Time: 01:03:34)





**MIMO System**

Automation Lab  
IIT Bombay

Controllable canonical form  
can be easily extended to develop  
state realization for **single input multiple output**  
(SIMO) system with a common denominator polynomial

Observable canonical form  
can be easily extended to develop  
state realization for **multiple input single output**  
(MISO) system with a common denominator polynomial

2/17/2012 System Identification CDEEP IIT Bombay 157

So let me come to multiple input multiple systems okay what I will show you is that if I work with multiple input multiple output system and if I work with state space models final form is the same irrespective of whether it is multiple input multiple output single input for a SIMO transfer function it is a scalar transfer function MIMO transfer function matrix is very complex business okay here everything is same whether it is SIMO or BIMO or SIMO or whatever okay the state space model will be just look same point finally okay.

(Refer Slide Time: 01:04:15)

Automation Lab  
IIT Bombay

## Realization for SIMO Model

Consider SIMO T.F. model

$$\begin{bmatrix} y_1(q) \\ y_2(q) \\ y_3(q) \end{bmatrix} = \frac{1}{q^3 + a_2q^2 + a_1q + a_0} \begin{bmatrix} b_{11}q^2 + b_{12}q + b_{13} \\ b_{21}q^2 + b_{22}q + b_{23} \\ b_{31}q^2 + b_{32}q + b_{33} \end{bmatrix} u(k)$$

Then, Controllable Canonical state realization for above  
SIMO system can be written as follows

$$x(k+1) = \begin{bmatrix} -a_2 & -a_1 & -a_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} x(k)$$

System Identification

CDEEP  
IIT Bombay  
150

IMPTES  
2/17/2012

So let us say I have let us consider a situation where you have three output and one input okay so I construct a state realization for this which looks like this there are three outputs  $b_1, b_2, b_3$  for each one of them will appear in the  $c$  matrix joist go back and think about it is not see there are three parallel lines all of them have same denominator okay here you have  $b_1$  to  $b_{13}$  here you have  $b_2$  to  $b_{23}$  and so on okay so all that will happen is that this matrix instead of becoming single output it will become multiple output all three are modeled here okay by the form mathematically looks same  $\phi\gamma c$  okay.

(Refer Slide Time: 01:05:22)

Automation Lab  
IIT Bombay

## State Realization of MIMO T.F.

Consider 2 x 2 T.F. model

$$y(k) = \underbrace{\begin{bmatrix} G_{11}(q) \\ G_{21}(q) \end{bmatrix}}_{y_{u_1}(k)} u_1(k) + \underbrace{\begin{bmatrix} G_{12}(q) \\ G_{22}(q) \end{bmatrix}}_{y_{u_2}(k)} u_2(k)$$

↓

$$y_{u_1}(k) = \begin{bmatrix} G_{11}(q) \\ G_{21}(q) \end{bmatrix} u_1(k)$$

with controllable canonical  
state realization

$$x(k+1) = \Phi^{(1)} x^{(1)}(k) + \Gamma^{(1)} u_1(k)$$

$$y_{u_1}(k) = C^{(1)} x^{(1)}(k)$$

CDEEP  
IIT Bombay  
129

2/17/2012      System Identification

Any of these are completed decimation so I do not belong to some of the I do not like graphic terms now what about 2.2 if you have two inputs two outputs what you do okay I can develop one model here okay I can develop with respect to u1 I can devolve a state realization okay with respect to u2 I can develop another state realization and I can put them together and create a big state realization.

Okay this is okay I have two things and then I am going to stack x1 and x2 see go back here and see what we have done we have spilt this there are two additive parts one is effect of u1 on y other is effect of u2 on y for this component we have developed one state realization this is yu1 that is effect of or contribution to y due to u1 that is given by this state contribution of u2 to y is given by the state plus model and this plus this is nothing but white gate okay.

(Refer Slide Time: 01:06:41)

Automation Lab  
IIT Bombay

## State Realization of MIMO T.F.


Defining an augmented state vector  $X(k)$  as

$$X(k) = \begin{bmatrix} x^{(1)}(k) \\ x^{(2)}(k) \end{bmatrix}$$


we can write MIMO state realization as

$$\begin{bmatrix} x^{(1)}(k+1) \\ x^{(2)}(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \Phi^{(1)} & [0] \\ [0] & \Phi^{(2)} \end{bmatrix}}_{\Phi} \begin{bmatrix} x^{(1)}(k) \\ x^{(2)}(k) \end{bmatrix} + \underbrace{\begin{bmatrix} \Gamma^{(1)} & \bar{0} \\ \bar{0} & \Gamma^{(2)} \end{bmatrix}}_{\Gamma} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$X(k+1) = \Phi X(k) + \Gamma u(k)$$



System Identification



So I am going to use this to define a combined state vector  $x_1$  and  $x_2$   $x_1$  coming from  $u_1$  model  $x_2$  coming from  $u_2$  model okay I can just combine it I can just stack you know this state space model is mini state space models can be stacked into big state space model this form is again the same whether it is multiple input multiple output single input multiple output whatever finally form is same it is not different.

So my algebra can be only on this particular form I do not care if I start working with matrix description if I start working with polynomial you know matrix description algebra is very control design becomes very easy well you had advantage you can work with frequency domain but here we have computational advantage okay you can do everything in time domain.

(Refer Slide Time: 01:07:38)

Automation Lab  
IIT Bombay

## State Realization of MIMO T.F.

Consider  $2 \times 2$  T.F. model

$$y_1(k) = \delta_{11}(q)u_1(k) + \delta_{12}(q)u_2(k)$$

with state realization

$$x^{(1)}(k+1) = \Phi^{(1)}x^{(1)}(k) + \Gamma^{(1)}u_1(k)$$

$$y_1(k) = C^{(1)}x^{(1)}(k)$$

and

$$y_2(k) = \delta_{21}(q)u_1(k) + \delta_{22}(q)u_2(k)$$

with state realization

$$x^{(2)}(k+1) = \Phi^{(2)}x^{(2)}(k) + \Gamma^{(2)}u_2(k)$$

$$y_2(k) = C^{(2)}x^{(2)}(k)$$

Defining an augmented state vector  $X(k)$  as

$$X(k) = \begin{bmatrix} x^{(1)}(k) \\ x^{(2)}(k) \end{bmatrix}$$

we can write

MIMO state realization as

$$X(k+1) = \begin{bmatrix} \Phi^{(1)} & [0] \\ [0] & \Phi^{(2)} \end{bmatrix} X(k) + \begin{bmatrix} \Gamma^{(1)} \\ \Gamma^{(2)} \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} C^{(1)} & 0 \\ 0 & C^{(2)} \end{bmatrix} X(k)$$

CDEEP  
IIT Bombay  
162

System Identification

2/17/2012

So I can just finally get the same form mathematical form remains same I have just given the same thing for MISO models and observable canonical forms if you have MISO models you can develop observable canonical forms which will look like this okay and then same thing you know you consider a two cross two system and then you develop two models you combine them into one big model finally will look like the same so techno message is state space form state space realization is you know will look same irrespective of number of inputs number of outputs you know finally mathematically form I just have to deal with one equation afterwards.

(Refer Slide Time: 01:08:25)

Automation Lab  
IIT Bombay

## Example: Four Tank System

ARMAX Model (4th Order)

$$A(q)y_1(t) = B_1(q)u_1(t) + B_2(q)u_2(t) + C(q)e_1(t)$$


$$A(q) = 1 - 0.6236q^{-1} - 0.8596q^{-2} - 0.0758q^{-3} + 0.568q^{-4}$$

$$B_1(q) = 0.08324q^{-1} + 0.02757q^{-2} + 0.02681q^{-3} - 0.1214q^{-4}$$

$$B_2(q) = 0.004045q^{-1} + 0.03261q^{-2} - 0.01841q^{-3} + 0.0201q^{-4}$$

$$C(q) = 1 - 0.4695q^{-1} - 0.8017q^{-2} - 0.1065q^{-3} + 0.4855q^{-4}$$

Loss function 0.0243707



2/17/2012

System Identification

CDEEP  
IIT Bombay  
103

Life is simple so four tank model so AXMAX model we can look at it as one output and two inputs what are the two inputs  $u$  and  $e$  are the two inputs okay I can look at the model with two inputs.

(Refer Slide Time: 01:08:44)

Automation Lab  
IIT Bombay

## ARMAX: State Realization


$$x(k+1) = \Phi x(k) + \Gamma u(k) + \zeta e(k)$$

$$Y(k) = C x(k) + e(k)$$

$$\Phi = \begin{bmatrix} 0.6236 & 1 & 0 & 0 \\ 0.8596 & 0 & 1 & 0 \\ 0.0758 & 0 & 0 & 1 \\ -0.5680 & 0 & 0 & 0 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} 0.0832 & 0.0040 \\ 0.0276 & 0.0326 \\ 0.0268 & -0.0184 \\ -0.1214 & 0.0201 \end{bmatrix} \quad \zeta = \begin{bmatrix} 0.1541 \\ 0.0579 \\ -0.0307 \\ -0.0826 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$


System Identification
CDEEP  
IIT Bombay  
104

And I can convert this into a state space form okay met lap tool box off course will give you this state space form if you give the model and say give me SS I think it is called id to SS or something these are functions it will give you the state space form okay so you will get a state space matrix see you have the states or there should be there is a mistake here okay you have two inputs one is u and other is e is e is the innovation okay there is only single output suppose you take single ARMAX model one input one output and one innovations then you can convert this into this standard form.

(Refer Slide Time: 01:09:27)

Information Lab  
IIT Bombay

## ARMAX: State Realization

$$x(k+1) = \Phi x(k) + \Gamma u(k) + L_w e(k)$$

$$y(k) = Cx(k) + e(k)$$


$$\Phi = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -a_{n-1} & 0 & 0 & \dots & 1 \\ -a_n & 0 & 0 & \dots & 0 \end{bmatrix}; \Gamma = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}; L_w = \begin{bmatrix} c_1 - a_1 \\ c_2 - a_2 \\ \dots \\ c_n - a_n \end{bmatrix}$$

$$C = [1 \ 0 \ \dots \ 0]$$

$$G(q) = \frac{B(q)}{A(q)} = C[qI - \Phi]^{-1}\Gamma; H(q) = \frac{C(q)}{A(q)} = C[qI - \Phi]^{-1}L_w + I$$

Interpretation as a State Observer

$$x(k+1|k) = \Phi x(k|k-1) + \Gamma u(k) + L_w [y(k) - Cx(k|k-1)]$$


CDEEP  
IIT Bombay

1/17/2012      System Identification      105

So this as a deeper relationship with something we are going to study later called filtering will talk about it later.

(Refer Slide Time: 01:09:38)



Automation Lab  
IIT Bombay

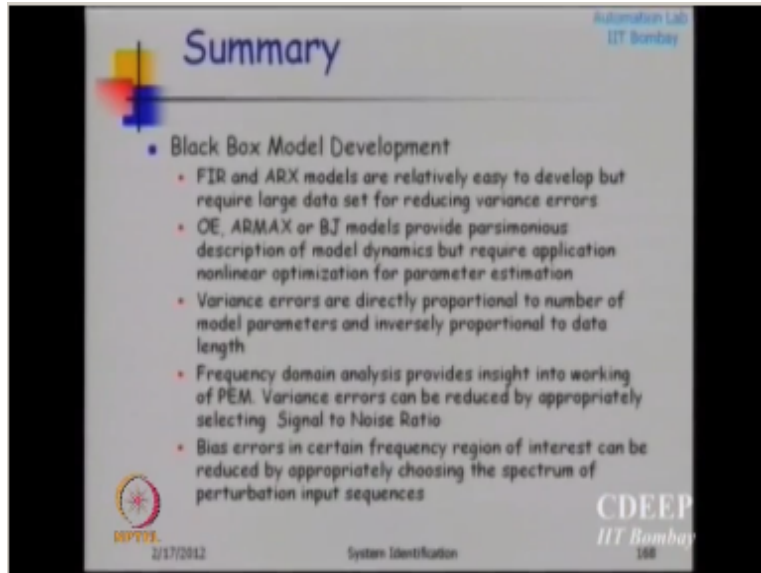
## Summary

- Grey box models
  - Better choice for representing system dynamics.
  - Provide insight into internal working of the system
  - Development process time consuming and difficult
- Black Box Models
  - Relatively easy to develop
  - Provide no insight into internal working of systems
  - Limited extrapolation abilities.
- Black Box Model Development
  - Noise modeling is necessary to be able to extract the deterministic component of the model properly
  - Prediction error method used for parameter estimation

IPTCL 2/17/2012 System Identification CDEEP IIT Bombay 167

So with this I come to the end of this lectures so we looked at different kinds of models we looked at grey box models black box models off course you have a mechanistic models nothing like it okay that is the best model if you do not have you can still develop a model completely from the data that is black box models if you can merge that too is grey box model okay and this is the most curial part in a any control project developing a good model.

(Refer Slide Time: 01:10:04)



The slide is titled "Summary" and is part of a presentation from the Automation Lab at IIT Bombay. It discusses "Black Box Model Development" and lists several key points. The slide includes logos for DPTTL and CDEEP, along with the date 2/17/2012 and the text "System Identification".

Automation Lab  
IIT Bombay

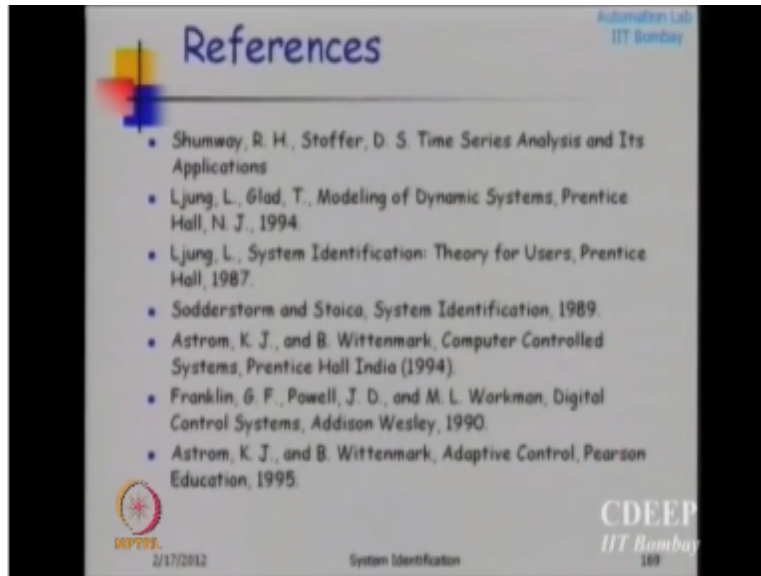
## Summary

- **Black Box Model Development**
  - FIR and ARX models are relatively easy to develop but require large data set for reducing variance errors
  - OE, ARMAX or BJ models provide parsimonious description of model dynamics but require application nonlinear optimization for parameter estimation
  - Variance errors are directly proportional to number of model parameters and inversely proportional to data length
  - Frequency domain analysis provides insight into working of PEM. Variance errors can be reduced by appropriately selecting Signal to Noise Ratio
  - Bias errors in certain frequency region of interest can be reduced by appropriately choosing the spectrum of perturbation input sequences

DPTTL  
2/17/2012  
System Identification  
CDEEP  
IIT Bombay  
168

So we have discussed all kinds of issues that are related to model.

(Refer Slide Time: 01:10:09)



Choose that are related to model development these are the some of the references that I have been using for my notes so there are not from one place a good book for beginners is this soderstrom and stoic system identification or shipway and stoffer these two are first and the fourth are very good books for beginners okay for advanced users loons book or actually I would say one three and four these are very good books for beginners even these two books Astrom and Franklin Powell so the last book and the second book sorry last book and the third book are they are very good but they are little advanced so.

**Producer**

Arun Kalwankar

**Project Manager**

M.Sangeeta Shrivastava

**Sr.Cameraman**

Tarun Negi

**Sr.Online Cameraman**

Sandeep Jadhav

**Digital Video Editor**

Tushar Deshpande

**Technical Assistants**

Vijay Kedare

Ravi Paswan

