NPTEL

NPTEL ONLINE COURSE

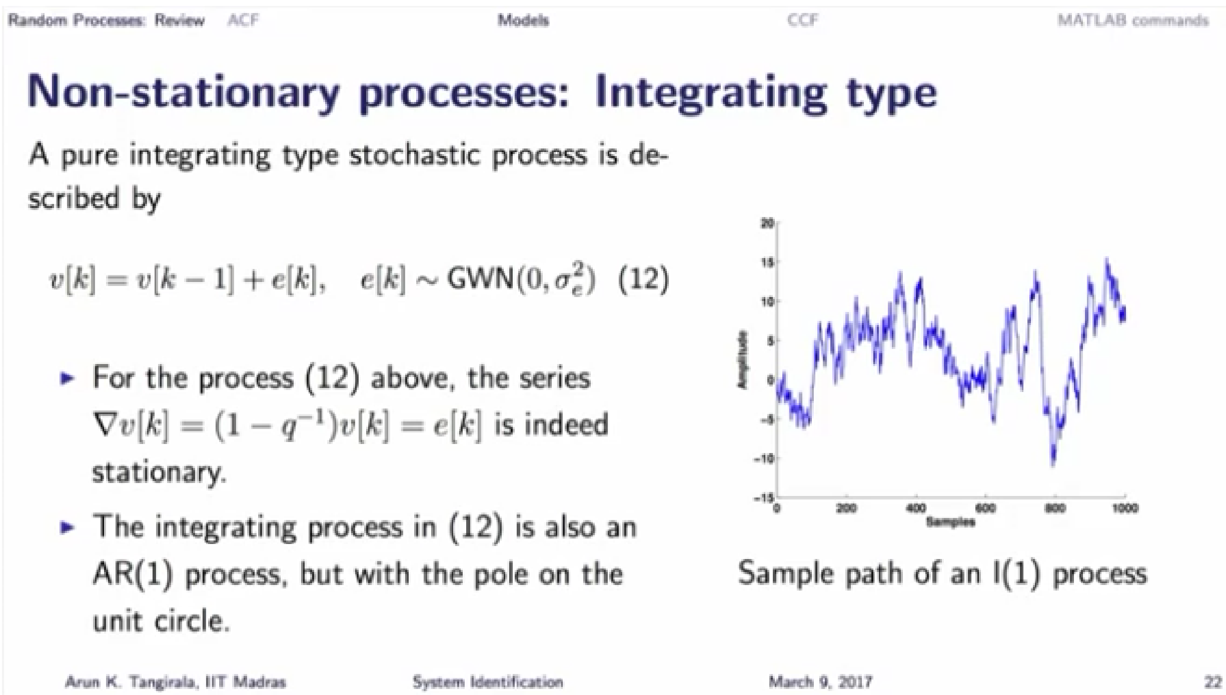
CH5230: SYSTEM IDENTIFICATION

RANDOM PROCESSES: REVIEW 8

ARUN K. TANGIRALA

**DEPARTMENT OF CHEMICAL ENGINEERING,
IIT MADRAS**

(Refer Slide Time: 00:20)



Now let's move on and quickly talk about non-stationary processes, and then cross correlation functions that will close the curtains on the review of correlation functions, and in general time domain representations of random processes.

Now until now we have discussed or reviewed stationary processes and we said when I have a stationary random process whose spectral densities satisfies some conditions which are not stated yet, I can think of the representation where the series can be represented as a white-noise passing through and LTI system.

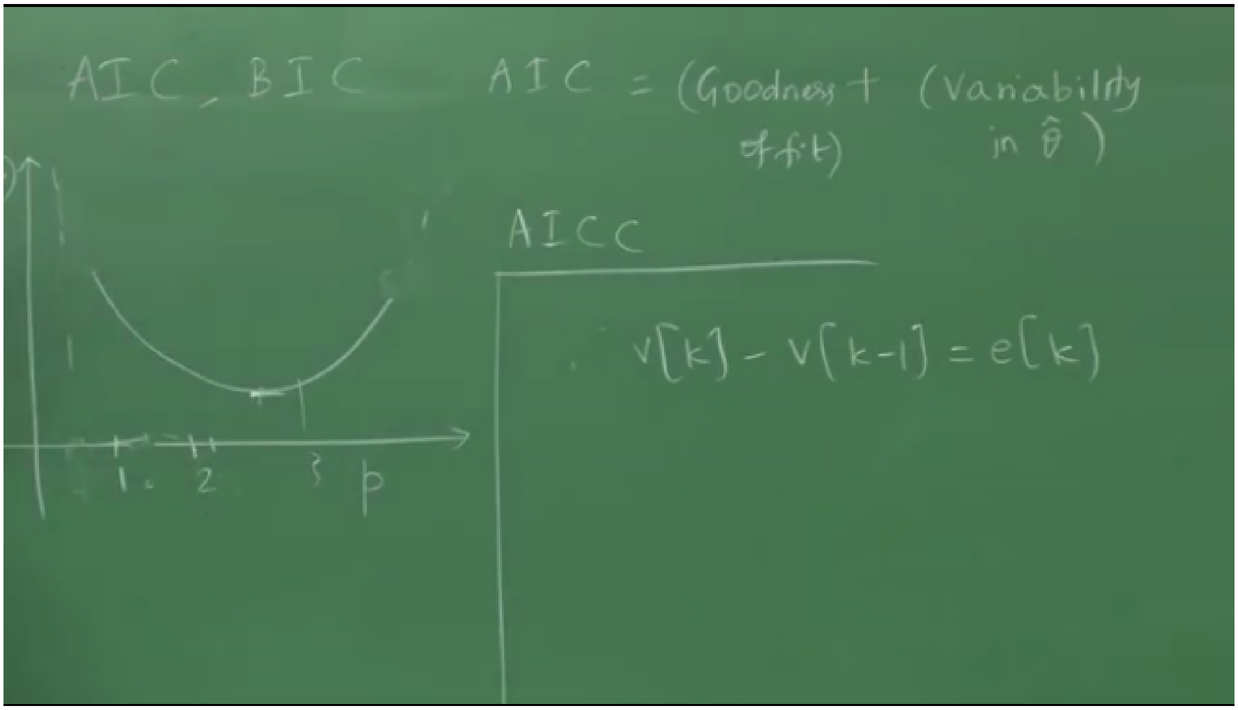
Very often of course you run into processes that are not stationary, so what do you do about it? In general non-stationarities are of different kinds instabilities are of different kinds in the deterministic world.

We don't discuss unstable systems for say, in the random process world non-stationarities in general can occur out of two types, deterministic type non-stationarities and stochastic type non-stationarities, what we mean by deterministic non-stationarities when you have a trend present in the series, it could be a linear trend or they could be a sinusoidal trend and so on, those are called trend type non-stationarities, we do not worry about them in this course. Then you have stochastic type non-stationarities, remember when we say non-stationarity, non-stationarities can occur as I said of the deterministic type, but if you approach it from a statistical property view point, mean can change with time or variance can change with time or both can change with time, PDF itself can change with time so there numerous possibilities.

The deterministic types that I just mentioned are generally discussed in the context of mean type non-stationarity, mean non-stationarities, when the local level is changing with time, right, but when the local level is changing with time it need not be just due to deterministic reasons, it could be stochastic reasons also, and one such reason, one such stochastic phenomenon which leads to non-stationarity is that of the integrating type or the random walk behavior, alright, and I'm showing you on the screen what is known as a pure integrating type process, it is acting similar to the integrated in your deterministic world, except that the driving force is now stochastic signal, otherwise you can see there is a single pole on the unit circle, you can think of the integrator as an AR1 with the pole on the unit circle.

Of course if the poles are outside the unit circle that's it, I mean your series is going to explore that is you know beyond doubt, so is this a fiction? Is this just a segment of someone's imagination or do such phenomena actually exist, and the answer is this phenomena do exist in the sense there are many, many phenomena out there which can be modeled as integrating type processes, they may not be model as pure integrating type processes, but there is an integrating component present in the process.

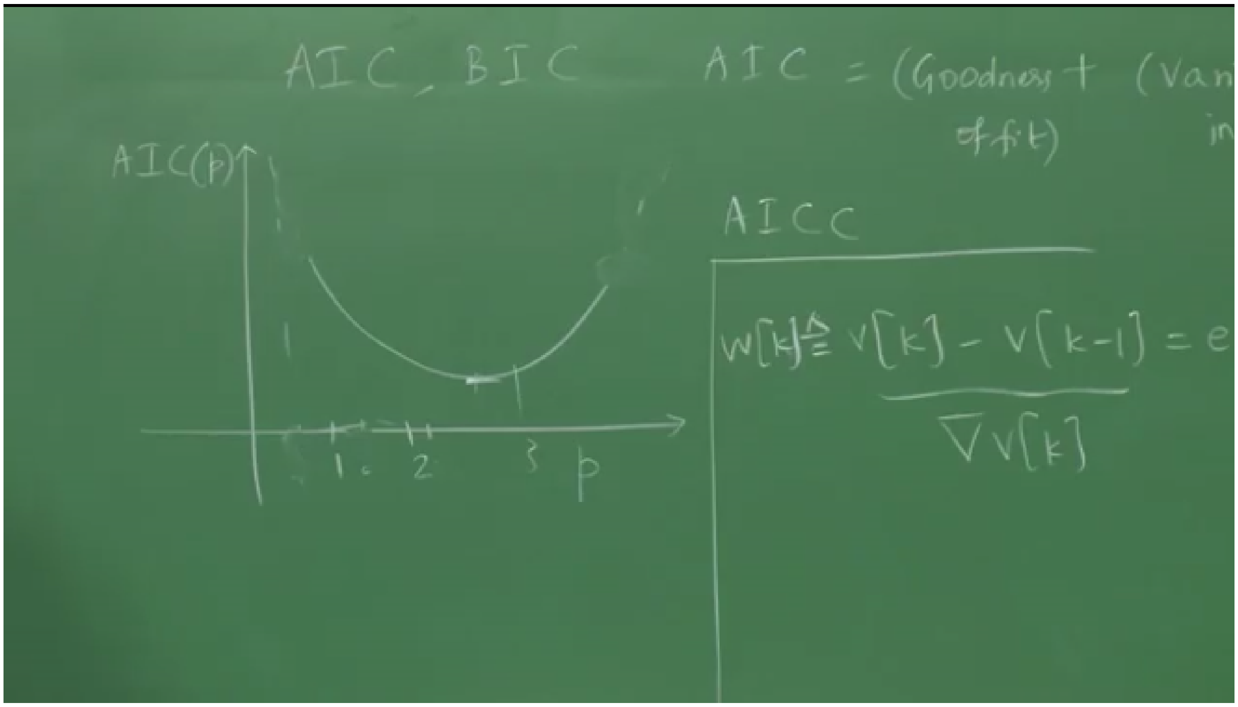
So for example you can have an integrating process on top of an ARMA process, such processes we call as ARIMA processes where you have an ARMA process but then there is also an integrating effect, so how does this model come about? Go back to the pure integrating process, if you look at the generating equation, it says VK= VK-1 + EK, now I can rewrite this by saying VK – VK-1 = EK, which means if I difference the given series that is purely integrating I can produce a stationary segment, right, so I have VK – VK-1 = EK, that's another way of writing the pure integrating type process.
(Refer Slide Time: 04:37)



So as I said this integrating type process is generally dealt with under the mean changing, that is the local levels are changing with time but not in a deterministic way. If you had a trend then you can say the mean is a linear function of time or a quadratic function of time and so on, here the mean change can is being modeled, although I don't show that to you here, but if you sit through the time series analysis videos you will see that the mean for this process can be actually modeled as a stochastic function of time.

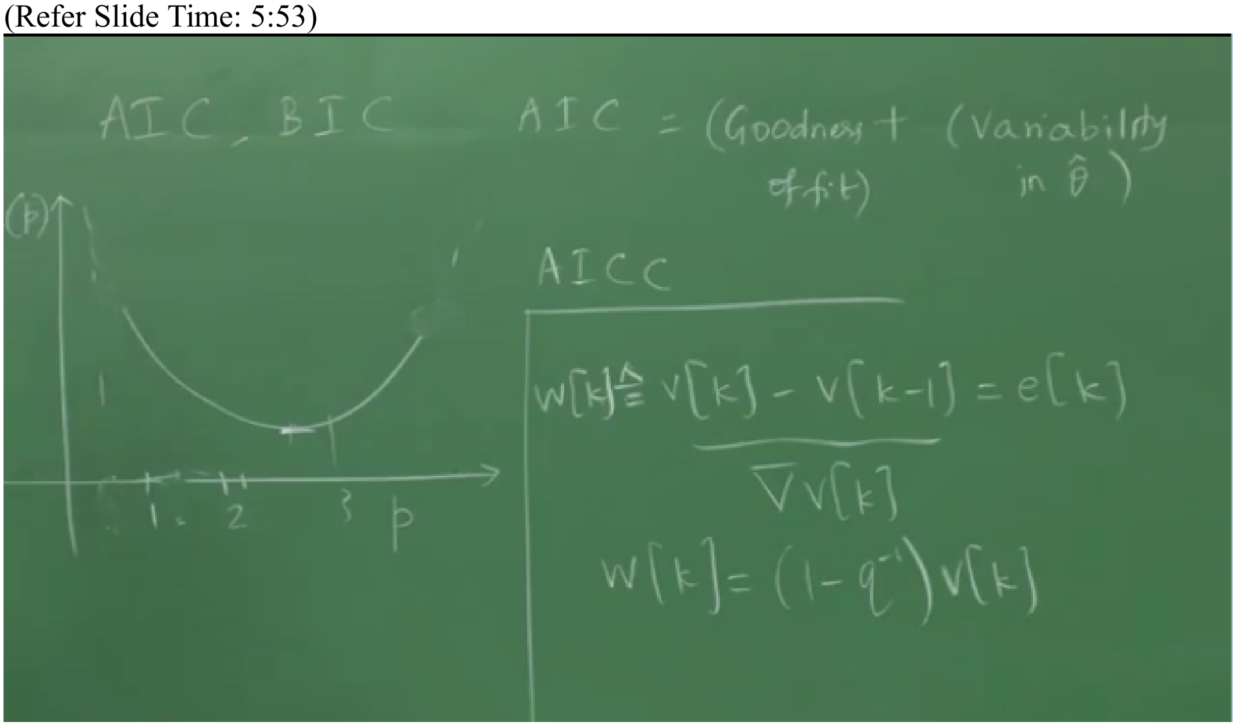
Here if I were to define a new series which is VK – VK-1 and typically we represent this with nabla,

(Refer Slide Time: 05:25)



so I have here a nabla VK which is telling you that you're differencing the series ones, this is called a differencing operation, and from the filtering view point differencing operation, so if you look at the transfer function here between W and V, what would that be? Sorry, 1-Q inverse VK, right, 1-Q inverse would be the transfer function operator.

(Refer Slide Time: 5:53)

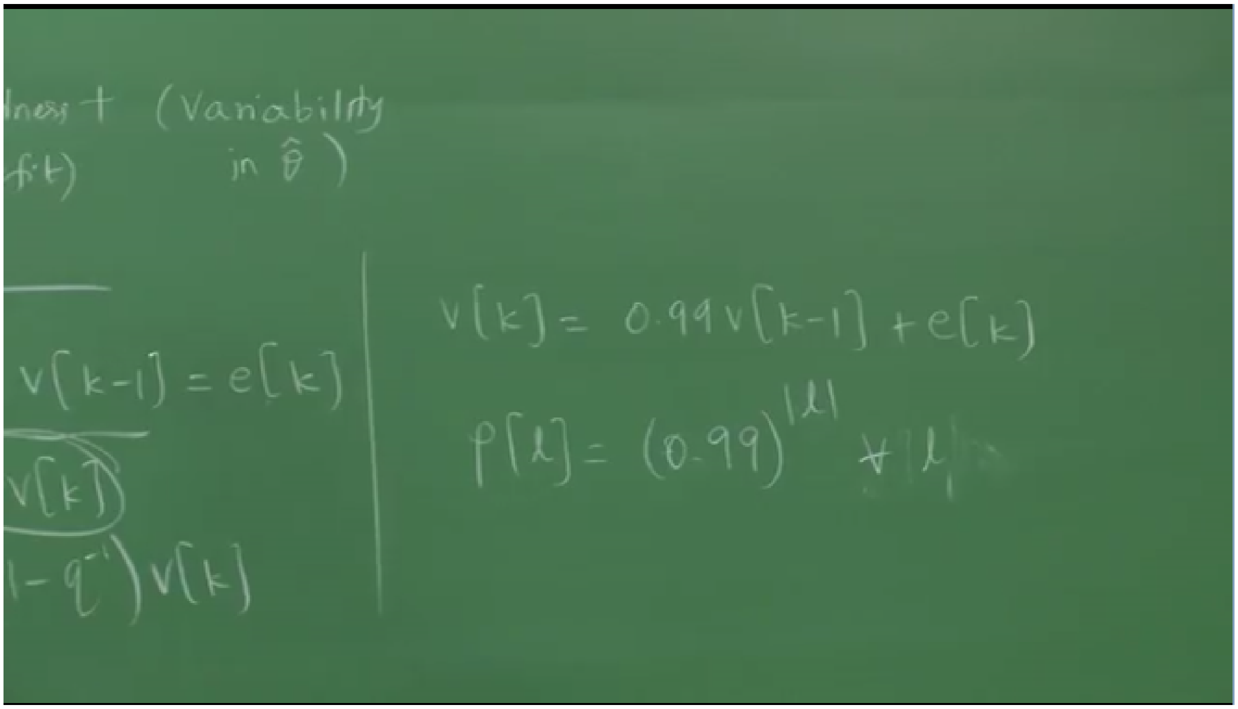


And if you were to draw the bode plot of 1- this you know equivalent of 1-Q inverse, you will find that it is a high pass filter, which means differencing amplifies high frequencies, and that make sense because when I construct the difference, that is successive differences of any series of two observations, what I'm looking at is if I generally move from one observation to another observation, the fluctuations are the one that are actually going to defer, make the two observations look different, so what you are doing by differencing is you are removing the commonality and you are highlighting the fluctuations, so that's one way of understanding why the differencing operation amounts to high pass filtering, that means you are letting only the high frequencies go through. In other words WK will contain largely the high frequency components of VK.

On the other hand if I had averaged, let us say I constructed a series by averaging two successive observations, averaging is the smoothing operation from the filtering view point it's a low pass filtering operation, that means it's going to iron out any high frequency that is rapid fluctuations, and only going to give you a smoother one. From a filtering view point again if you were to draw the bode A plot for the transfer function of an averaging operation you will see that it has low past characteristics, and this is important and later on when we talk of modeling systems which have a non-stationary kind of noise, alright.

So coming back to the point you see that the difference series is stationary, of course it is white, but white means also stationary, so in general you don't have necessarily a purely integrating process, in general instead of this W being driven by EK, W can be driven that is this being driven by EK, it can be driven by another stationary process, it need not be just EK, and that is the idea behind ARIMA models, what you are saying in ARIMA model is the D times difference VK is been driven by another stationary process, it need not be white, if it is white yeah that also is taken, that can be accommodated in this model, so we're just generalizing the idea, you are saying that there is an integrating effect, and there is also stationary effects there, if there is no stationary and only integrating and of so that a single differencing produces a stationary sequence than we say D is 1 that is the number of times you have to difference a series as one.

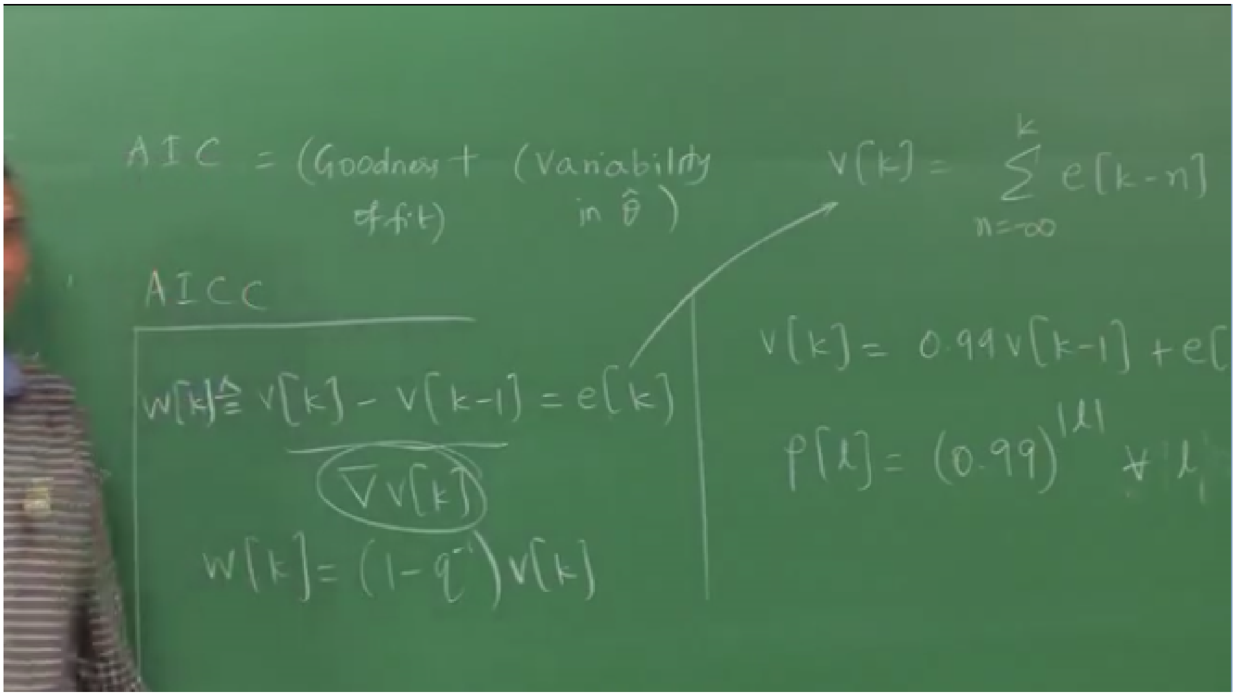
If there are double integrating effects then you will have to difference twice and so on, how do I know that there are integrating effects in a process, this is the first thing that question that comes to our mind, given a series how do I figure out? Now the answer to that of course there are two ways of looking at it, doing that one is to turn to the ACF, other is to use the formal unit, so called unit root test, so going back to the pure integrating process here, if you look at the integrating process I said it's an AR1 model, now I cannot use the ACF expression that I have been using until now, because that ACF expression is valid only for stationary processes, so instead let us assume that I have a process that is nearly, which has a pole very close to the unit circle, nearly integrating, okay, this we know from our discussion has an auto-correlation of 0.99 raised to mod L, right for all L, this we have already derived in a class, right,
(Refer Slide Time: 10:08)



when you plot this autocorrelation function, how they'd look like? Very slow decay, it will take really long, so as in the limit as D1 approaches unity, you would see that ACF wouldn't decay at all, so try doing this, go back to your MATLAB, simply simulate a series an integrating process.

Now at this point maybe your sim routine, I don't know in MATLAB whether the sim routine checks for stationarities in your model, for example if you take R, it may not simulate non-stationary processes, that means if you include an integrator in your model and you say simulate such a process it will refuse to do it, so you may have to turn, use other ways of simulating this.

Now one of the simplest ways of simulating an integrating process is to rewrite this in a different way, the same process can be rewritten in a different way, do you see that?
(Refer Slide Time: 11:23)



Right, so keep just recursively substituting for VK-1, VK-1 would be VK-2 + EK-1 and so on, of course I have ignored the initial condition of the process, I have set V at –infinity was 0, this tells you also why this process is called an integrator, because it is summing up all the shock waves right from the beginning of the universe up to this point, we don't have data when you simulate we don't have the white noise process on the beginning of the universe, but at least you pretend that whatever you have the starting point is the beginning of the universe, and then you simulate simply using a function called cumsum or whatever, essentially it is just cumulatively summing it up, you don't have to do much, all you have to do is generate white noise process, construct a cumulative sum that is your simulated integrating process.

Simulate that way and compute ACF and see how the ACF looks like, see when you apply the ACF routine what you are doing is you're estimating ACF, you may ask is this estimation procedure valid, using the formula that is used in XCOV or you know that is used in general ACF estimation, there is no, nothing wrong with it you can use any estimation expression with anything, but whether this estimate can be used to say something about the true process is always a question. It is true when you for the purely integrating process, when you use the standard ACF routine or XCOV routine to estimate the ACF you cannot necessarily say that whatever you see is the feature of the underlying process because the underlying process is actually strictly speaking non-stationary, but ignore you assume that you have generated a near integrating process, so you can do that also, you can generate instead of generating this, you can actually generate this, you can use 0.999, the routine may not object.

It is almost same as the integrating process, so look at ACF it will decay slowly. What would the PACF tell you? Why there is so much silence, what would the PACF tell you? In general what is PACF tell you for an AR process, why are you thinking so much? How would that PACF look like? So beyond lag 1 that the estimate should be insignificant which is the case, correct, for near integrating process the PACF would look like, would have the same structure as a PACF of an AR1 process, and the first coefficient, there will be only one coefficient which will be closed to unit, so that's another way of ascertaining whether integrating effects are present, but this is for purely integrating process.

Suppose I have an ARIMA process, how does the ACF look like? Will the ACF again decay very slowly, what do you think? It should, because ACF of the stationary part decays, what dominates, what continuous to prevail is ACF of the integrating process, that will still dominate, so you will still see a slow decay, but not as slow as for a purely integrating one, because the stationary part pulls it down, right, to a some extent so you will see a decay but then that decay will be slow, it will be not as slow as for a purely integrating one, so that is another feature of an integrating effect present in your model.

From a system identification view point you should ask the question, why should I worry about integrating effects or near integrating effects, see if the pole is close to unit circle, and not outside the unit circle, then it is stationary, so why should I break my head on this, so suppose even going back to pure time series modeling, suppose this is the process I have generated the process this way, that true process is strictly stationary, at least second order, on the second order view point, then why should I worry about it, it's a stationary process, I should be able to estimate it very well, any answer why I should really worry about near integrating processes? Error in what? So when I fit a model, and you should try this out as well, simulate again for the same series that you are simulated after examine the ACF you fit as if it is an AR1, you fit an AR1 model and look at the parameter estimate, the point estimate as we call the single value that it gives you may still be within the unit circle, maybe it depends on the algorithm but most likely it should be.

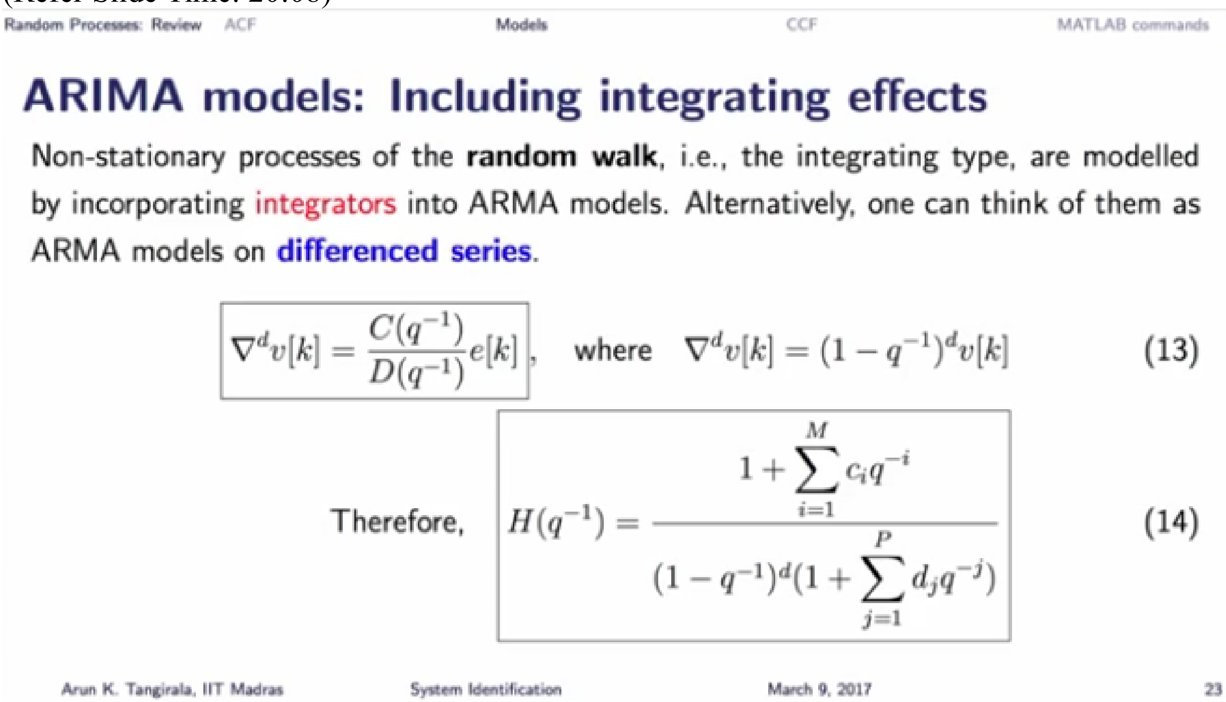
However when you look at the error in the parameter estimate and eventually when you construct so called confidence interval, the confidence region for the true parameter can include poles outside the unit circle, which means you are saying one of the possible models is a non-stationary one for a process that is stationary, in other words these parameters are so closed to the unit circle that estimation, estimating them precisely becomes difficult, that is when we say precession we are talking of the error in the parameter estimate.

To make sure that my confidence regions include stable models or stationary models for processes that have poles very close to unit circle it is like on the border, then you have difficulties, you can think of this as ill conditioning, so these are special, always borders are a problem in any border case is a problem, whether they are countries or parameter estimates or filtering, even in filtering you say at the ends of the signal filtering there is a problem, even when I'm looking at solving an assignment just one hour before due light day you are closed to the border there is an issue, so everywhere borders are an issue and so estimation is no exception, therefore when you find out that there are these processes that you are looking at which have near integrating effects, then you should maybe take special steps.

One of the remedies that is suggested is integrate, now difference is series so you will say this for this process it looks like there is a near integrating effect, you will not be able to generally distinguish, of course you can conduct a statistical test, but generally speaking you suspect that there is a pole either on the unit circle or very close to unit circle and you say let me for safety of the model, fixed that pole on the unit cycle, so what you do? You difference a series and then work with a differenced series, but this remedy should be this asthram as I say, this is one of the astras, should be use carefully, why? Because when you difference the series you are going to amplify noise. You are going to amplify the noise, here it’s time series but we’ll deal ultimately with the response of data which will have both input and noise effects, and if you suspect that there are integrating effects, then when you difference the series to get rid of the, to model of the integrating effects then you have to be careful, and we’ll talk about that a bit later, but this is all something that you should be aware of upfront.

To summarize ARIMA models are extensions of ARMA models to processes that have random walk type behavior, and this random walk are integrating type behavior is seen in many, many processes, I’ll not go into detail, but you can just look up the resources on the net, if the history of this random walk process starts off with brownie and motion, which then was investigated by Einstein and then others came along,

(Refer Slide Time: 20:08)



okay, and I think Einstein is the one who gave the name random walk.

So looking to somehow this resources, but the fact is your, by constructing the ARIMA model you are forcing at least one pole on the unit circle, maybe you have to force more than one, it depends on the number of integrating effects present in the series.

So let’s move on,

(Refer Slide Time: 20:35)

Cross-covariance function (CCVF)

The cross-covariance function is a generalization of the ACVF to two series. It measures the linear dependence between two samples of a series.

Cross-covariance function

The CCVF between two random processes $\{y[k]\}$ and $\{u[k]\}$ is defined as cov$(y[k], u[k-l])$

$$\sigma_{yu}[k,l] = E((y[k] - \mu_{y[k]})(u[k-l] - \mu_{u[k-l]})) \tag{15}$$

For stationary processes, the CCVF is only a function of the lag

$$\boxed{\sigma_{yu}[l] = E((y[k] - \mu_y)(u[k-l] - \mu_u))} \tag{16}$$

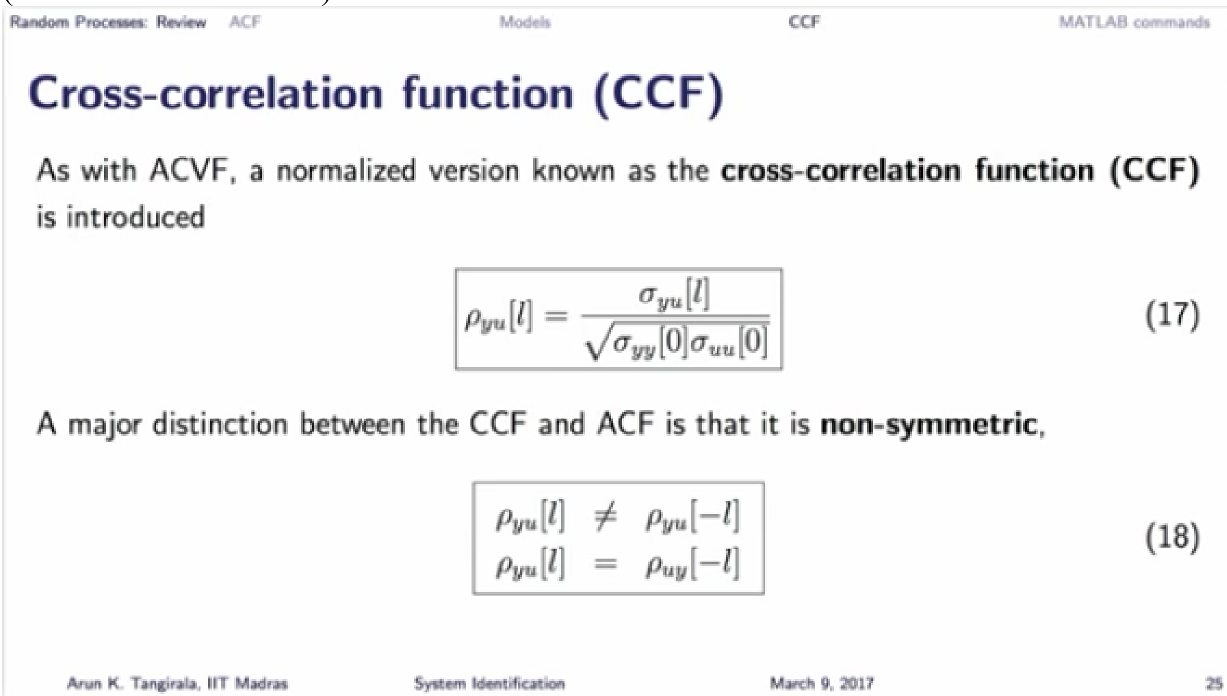
so that closes the noise models that we wanted to look at, the class of stochastic models that we wanted to discuss, for stationary processes that ARMA models for non-stationary, particularly of the integrating type only, we look at the ARIMA models. There are other types of non-stationarities which we have not discussed, for example a very common type of non-stationarity is the one which is variance type non-stationarity where you encountered terms like heteroscedasticity, that means the variance, variability changes with time, that PDF itself may remain the same, it maybe Gaussian, uniform or whatever, mean also may remain the same, but variance changes with time.

Again you have to ask, do such phenomena exist? Yes, you take earthquakes or ECG type series and so on you will see that the variability keeps saying, even many economic processes, econometric process actually exhibit heteroscedasticity, in fact heteroscedasticity is most prevalent in you know, econometric phenomena.

We do not talk about it extensively or a lot in this course, so let's conclude now with the discussion on cross covariance function because this is one of the other tools that we'll use extensively. The cross covariance function as we know is an extension of the notion of auto-covariance now to two series, the purpose is the same to see if one series has anything to offer with respect to the, for predicting the other series in a linear way, because at the heart of this is again is the covariance measure.

Now obviously when I'm looking at two series it need not be the case that UK, suppose I have two series here let us say you know Y and U, here we are going to assume that Y and U are stochastic, just for the sake of discussion, it need not be that UK instantly effects YK, and that is why we are looking at lagged covariance as well, because there could be a delay, right, so and there could be dynamics also, so it continuous to effect UK not only instantly effects but also continuously effect for a while, we want to make all such inferences, therefore we introduce this cross covariance function which helps us in many ways, in fact in model validation the cross covariance function or the cross correlation which is a normalized measure is extensively used, we have already seen that in the liquid level case study,

(Refer Slide Time: 23:08)



we said there are two requirements of a good model in system identification, one is that the residuals should be divide of effects of input, which is the test or which is the criterion for the goodness of G, the plant model, what is the second criterion on the residuals?

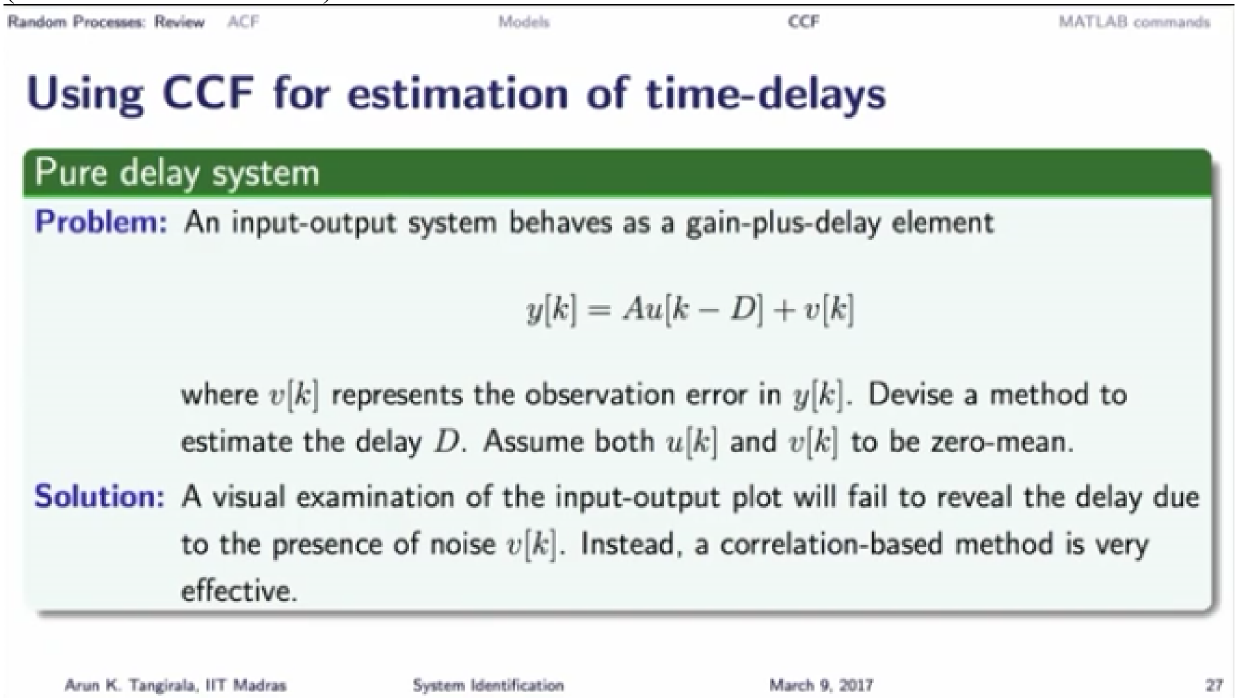
That is what we have just mentioned, what is the second one? See there are two models that you are fitting in society G and H, if your G is good you've estimated G sufficiently well, it's adequate, then the criterion that we just mentioned applies, the cross correlation between the residuals and the inputs should be insignificant, theoretically 0.

What is the second requirement? I'm sorry, yeah, of what? Auto-correlation of the residues, the second one is the test on H, where you look at auto-correlation of residuals like we did in yesterday's example, right, in yesterday's example there was no input effects, it's pure time series modeling, so both cross correlation and auto correlation play very critical role in model assessment, you should just get this embedded in your minds that the first step after you build a model is this check of goodness of G and goodness of H, if there is no G you're only testing for H, where you're only going to use auto-correlation functions that is one crucial role that cross correlation function plays.

The other application of cross correlation function which is perhaps even more prominent then the one that we just discussed is in the estimation of delays in radar signal processing everywhere the cross correlation function is used in estimating delays.

And the third use is in the estimation of impulse response functions, so I'll just show you couple of examples here, particularly with respect to delay estimation,

(Refer Slide Time: 25:17)



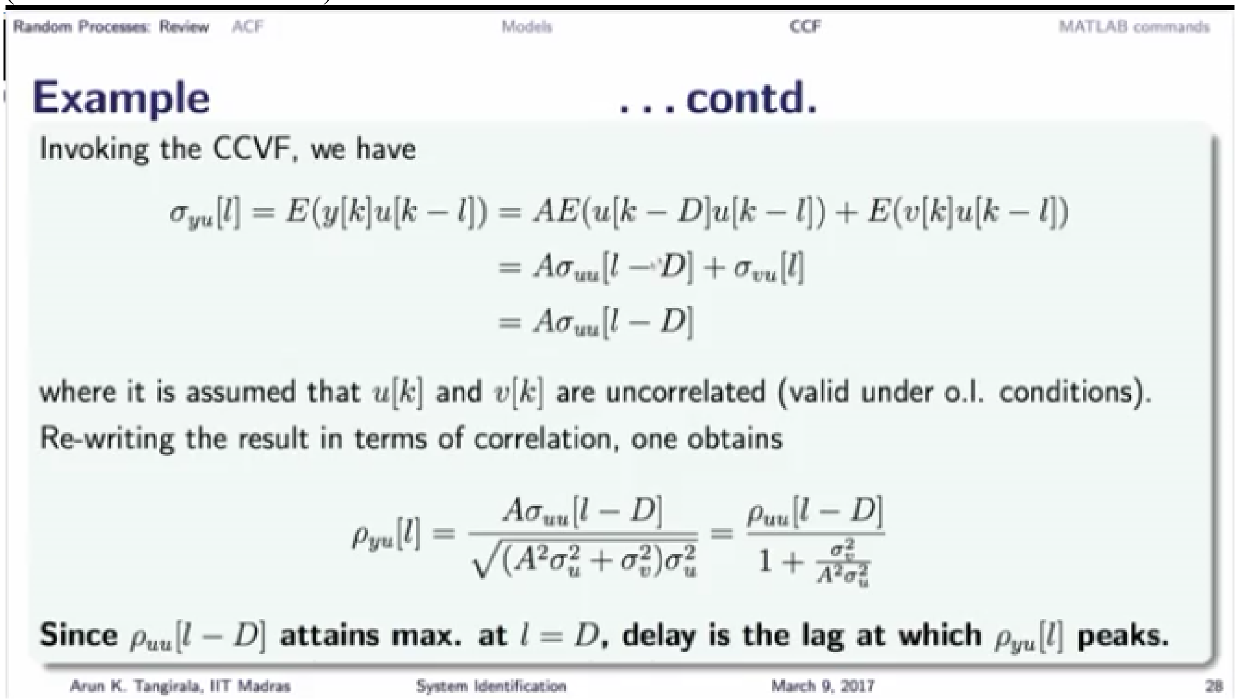
how the cross correlation function is used in an impulse response estimation will talk about that a bit later, but this is one of the most ubiquitous uses of cross correlation function, that is in delay estimation, so let's look at the simple example here, I have a pure delay system, alright, and the generating process is given here, I have YK = A times you know UK – D + VK, there are no dynamics here, A times UK – D that is your G, deterministic process, so the deterministic process is a pure delay one, I'm assuming that there is some observation error, you can say this VK again as usual contains sensor noise and effects of unmeasured disturbances and so on. For simplicity we'll assume U and V to be 0 mean, and we'll also assume open looked conditions without saying explicitly, but I thought I'll just mention that because it's important to make that assumption.

Now of course what you could do is you can plot the input and output one below the other and see whenever a change occurred in input how long did it take for the output to change, but that is, that approach is okay when there is no noise, when there is noise there is no way you can figure that out without ambiguity, instead when you turn to the correlation domain that's a beauty, in the correlation domain the effect of this noise is killed, how does it happen? So you look at this equation here, before I go in to the actual derivation, you look at this equation, in the observation domain noise effects, why?

Suppose I take, I correlate the output with the input, what happens? That means technically I multiply both sides with UK and take expectations, because anyway UK is 0 mean since U and V are 0 mean, Y is also 0 mean, therefore expectation of YK times UK is going to be the cross covariance.

On the right hand side what would you have? You would have two terms right, sorry, auto-correlate, so the left hand side you have cross covariance, you multiplying with UK, on the

right hand side also you're multiplying with UK, so the first term would be when you take expectations would be auto covariance of the input, right, so A times, yesterday I wrote a very important result, I said that if for a stochastic process input is driving an output, then it is as if the auto covariance is driving the cross covariance, correct, fine, so the first term is auto covariance, what about the second term? What is it? Why? Open loop, correct, it's going to be the cross covariance between the disturbance V and the input, under open loop conditions they are uncorrelated, but this has to be valid, sometimes the sensor noise may depend on the operating conditions, in which case that may not be true, sometimes there maybe feedback in which case the second term may not vanish, but will ignore all such so this is the derivation for you.
(Refer Slide Time: 28:38)

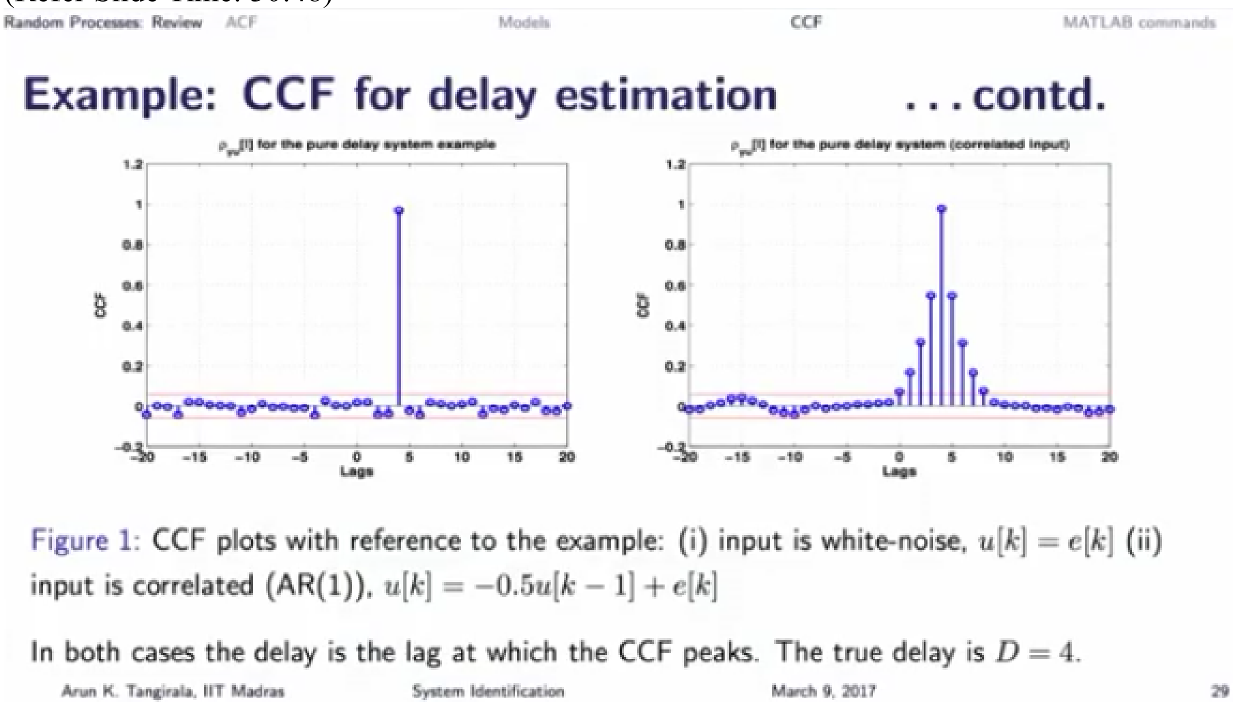


On the left hand side you have cross covariance at any lag L, instead of multiplying with UK, I multiply with UK-L, so that left hand side becomes a cross covariance function, and on the right hand side I have only the auto covariance, so what happened to the noise here? I have managed to kill the effects of noise theoretically, which is great, that means I have moved to a new domain where noise doesn't bother me, now I can go ahead with my operation, I can figure out the delay, so what do I have to do? All I have to do is plot the cross correlation, I didn't discuss so much about the cross correlation because it's again a normalized measure of cross covariance, nothing much, and I've already said cross covariance is an asymmetric function for natural reasons, so if you look at the result here the cross correlation is given by this expression for this example.

What do you notice? On the numerator you have auto correlation of the input, and the denominator has 1+ sigma square U, sigma square U/A square sigma square, sorry, sigma square V it should be, sigma square V/A square sigma square U, what is that quantity, that ratio? It's an inverse of SNR, A square sigma square U is the variance of the noise free part, sigma square V is the variance of the noise, so what does it tell you? For high SNR's, suppose

SNR is very high then the cross correlation between input and output is same as a autocorrelation of the input, but shifted by a certain amount.

Regardless of whether noise is present or not, the result that falls out is the cross correlation peaks at what lag? At lag equals delay, because the auto correlation of the input peaks at 0, therefore what I do is in practice I plot the cross correlation as a function of lag L,
(Refer Slide Time: 30:48)



alright, and then such for the peak, so this is what happens, look at the left hand side plot this is for the pure delay where I have used the white noise input for simulation, how does the auto-correlation of white noise input look like? Just an impulse, what this result says is when -- input, look at the cross correlation - - output and input, the cross correlation is going to be simply shifted ACF of the white noise. The ACF of the white noise is an impulse, and it is just shifted exactly by that amount and you just have to notice, read of that lag at which you see the peak, and that is delay, and in this example delay is 4.

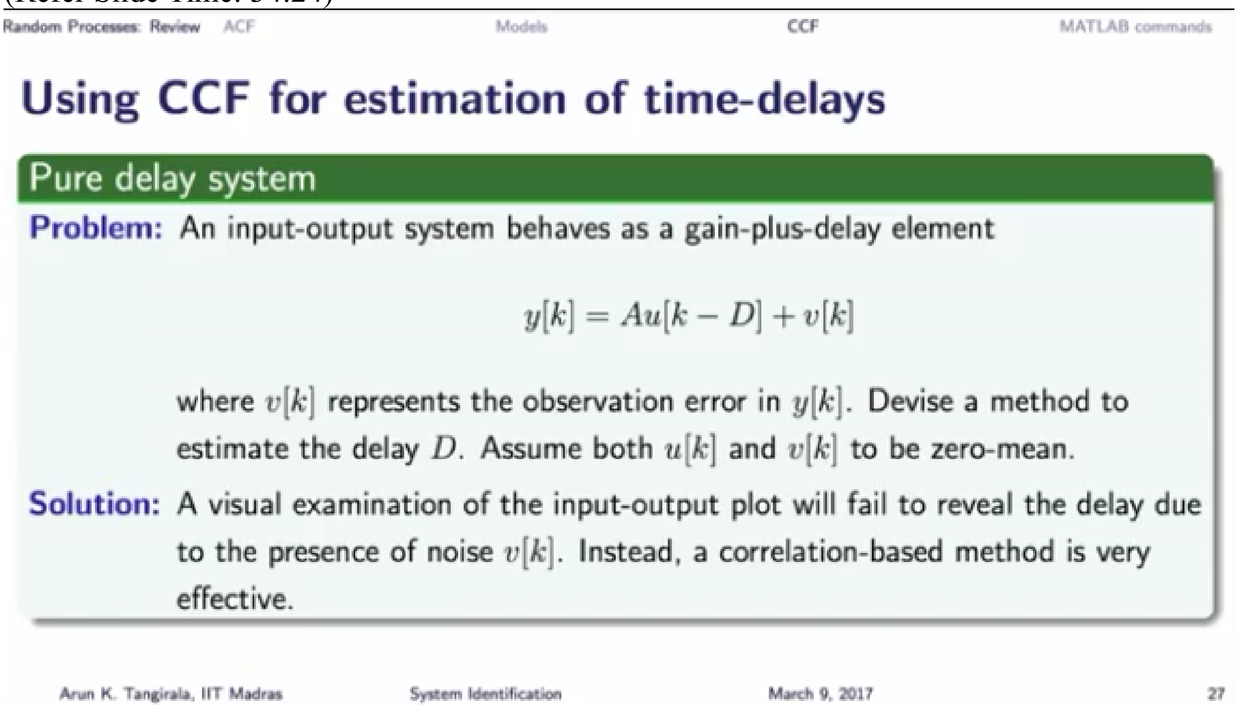
In the second case what I have done is the same process, but I have used a correlated input, how does the auto-correlation of a correlated input look like? So what I have done is, I have used an input that comes out of an AR1 process, in other words UK is not white, I have used UK as the output of an AR1 process, the most important thing for you to observe is the cross correlation is simply shifted auto correlation of the input, that's all and shift by how much? Delay, that's all, you know the input, therefore you know it's autocorrelation, you are just looking at the shift.

When we use the white noise the ACF of white noise is an impulse, cross correlation is going to be shifted impulse, and you just read of the lag at which you see the peak that is a delay, if the input is not white, but it has colored kind of characteristics, at this point also some of you must be wondering how can the input be white noise, we've been saying all along that input is deterministic in society, correct? That question is valid, but when we say input is white, input has white noise like characteristics, for example I can generate an input using my random, I say one realization of white noise is going to be my input, that's all, one realization of white noise process will have the properties of a white noise processes, so whenever we say in society input is white, the way you should read it as, the input has white noise like properties, it is not that the input is random, it has the properties of a random process.

When you look at the correlation autocorrelation of the input it will look like that one of, that realization that you have generated, so the message here is that the cross correlation of the, for a pure delay system the cross correlation is simply they shifted autocorrelation and you can use that to read off the delay.

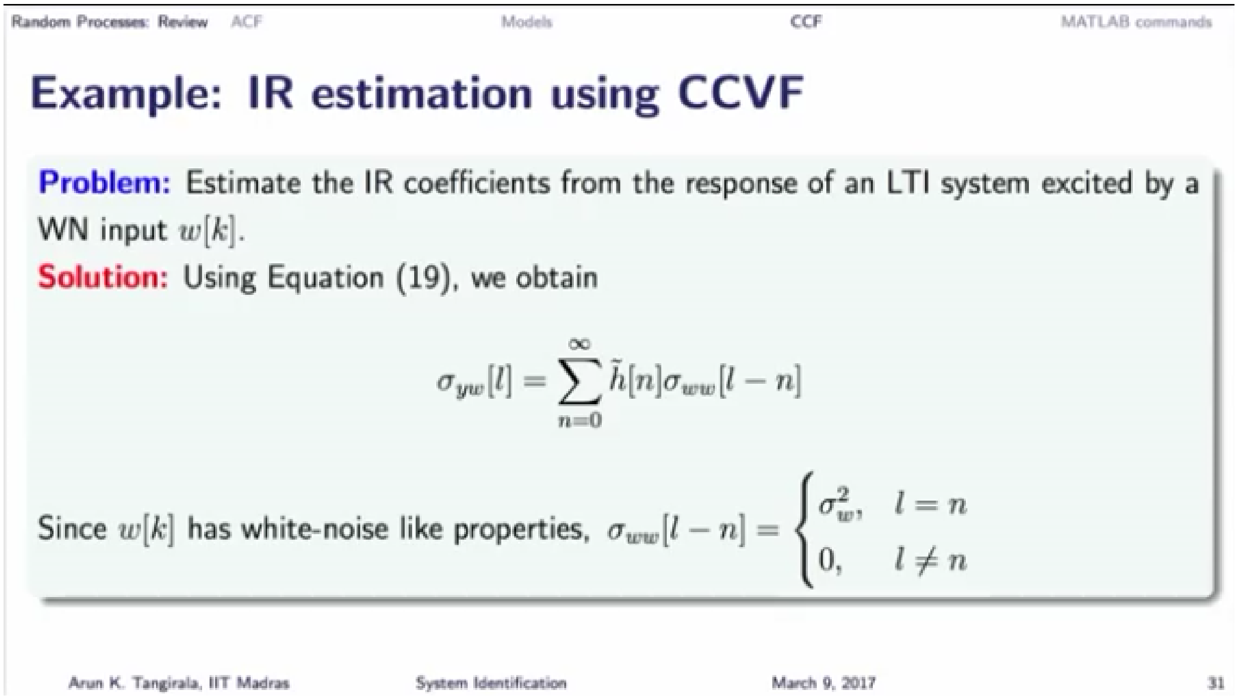
The more complicated situation is when the system is not a pure delay, but delay with dynamics, our classical you know first order + delay, second order + delay, can I still use the cross correlation to figure out to estimate the delay, and answer is yes.

There is an assignment question I think to that effect, figure that out and see if you can answer it, I have shown you for the pure delay case, suppose this was not a pure delay and there were dynamics also, for example go and work out this equation, suppose the generating equation was YK = some –A1, YK-1 + this UK-D + VK, how would the cross correlation work out too,
(Refer Slide Time: 34:24)



what would be the expression? Can you still say that the peak in the cross correlation will give you the delay or something else, but the fact is the cross correlation will give you an estimate of the delay, work that out and this brings us to the closed of the cross correlation functions, I will talk about the role of cross covariance in impulse response estimation, when we actually talk of estimating impulse response model, all the way I give you the expression here.

So when we meet in the next class,
(Refer Slide Time: 34:56)



we will talk briefly about spectral densities and then move on to the, comeback to a deterministic + stochastic world, alright. We’re already at the interface, alright, have a good weekend.

Online Editing and Post Production

Karthik
Ravichandran
Mohanarangan
Wilhelm Benjamin Anand
Sribalaji
Komathi
Vignesh
Mahesh Kumar

&

Web-Studio Team

NPTEL Co-ordinators

Prof. Andrew Thangaraj
Prof. Prathap Haridoss

IIT Madras Production

Funded by
Department of Higher Education
Ministry of Human Resource Development
Government of India

www.nptel.ac.in