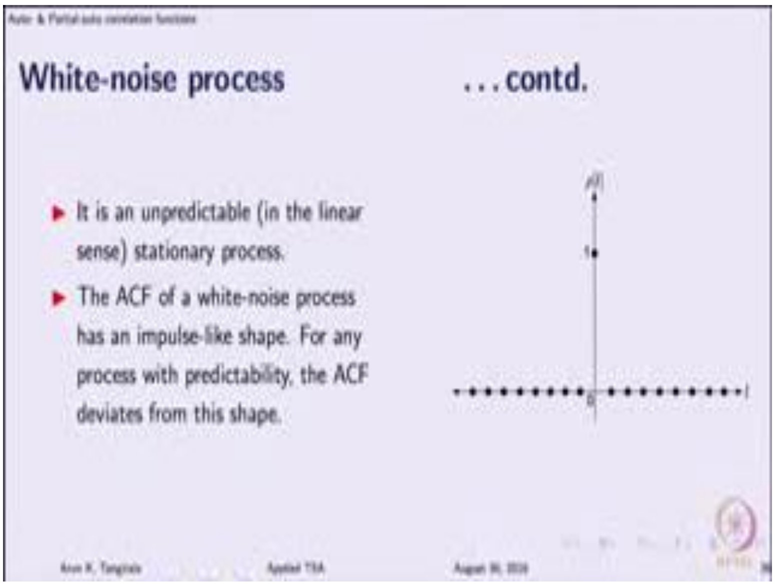
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 29
Lecture 13A - Autocovariance & Autocorrelation Functions-4

So, let us recap what we learnt in last class just towards the end. We of course, went over the definitions of auto covariance function and autocorrelation function and the primary need for turning to the autocorrelation function is to test for predictability within a series. Because that is a first step in time series analysis at least time series modeling. And the second inevitable step is to define a bench mark process, which is unpredictable in a linear sense and that is what we call as a white noise.

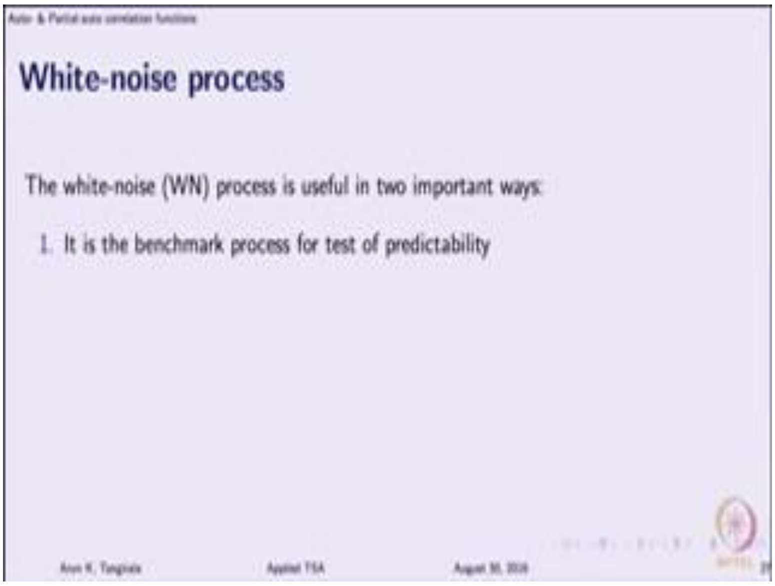
But this white noise as I mentioned in the last class has several purposes to its conception one of course, it is a bench mark processes what we mean by bench marking is when I am given a series for modeling, whatever series it may be; the first test that I conduct is whether the series is predictable or not, only when the ACF tells me that it is predictable then I proceed. So, we are implicitly bench marking every time series against this ideal white noise process whose ACF looks like this.

(Refer Slide Time: 01:24)



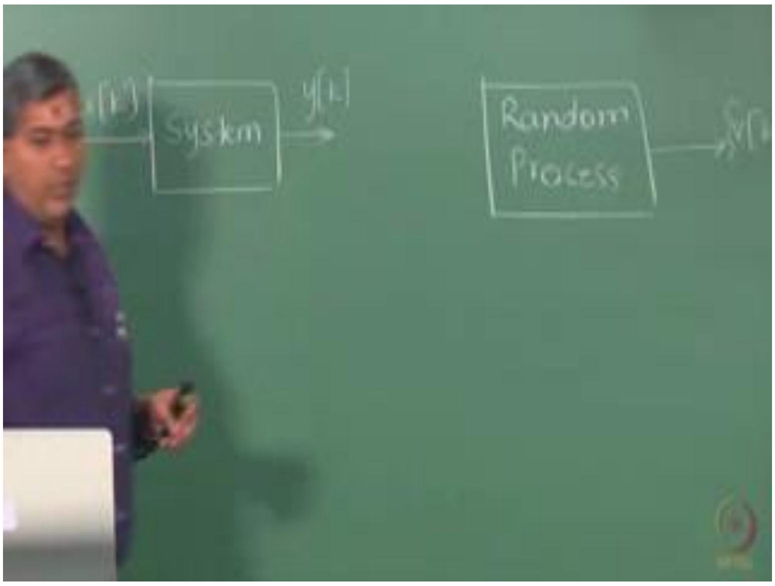
It is an impulse like shape in fact it is a unit impulse to be precised.

(Refer Slide Time: 01:34)



And that is a first purpose of a white noise process, there are two other purposes that the white noise process serves and the second one that we talked about is that it servers as this ideal residual that I desire in a good model. So, one is before modeling I bench marked the series against the white noise process just by examining the ACF and then post modeling, I access the model that have built by examining the ACF of the residuals. So, these are 2 of 3 purposes there is the third purpose which is that it serves as the fictitious input that we had indicated none of the introductory lectures, I have drawn systematic if you recall.

(Refer Slide Time: 02:20)

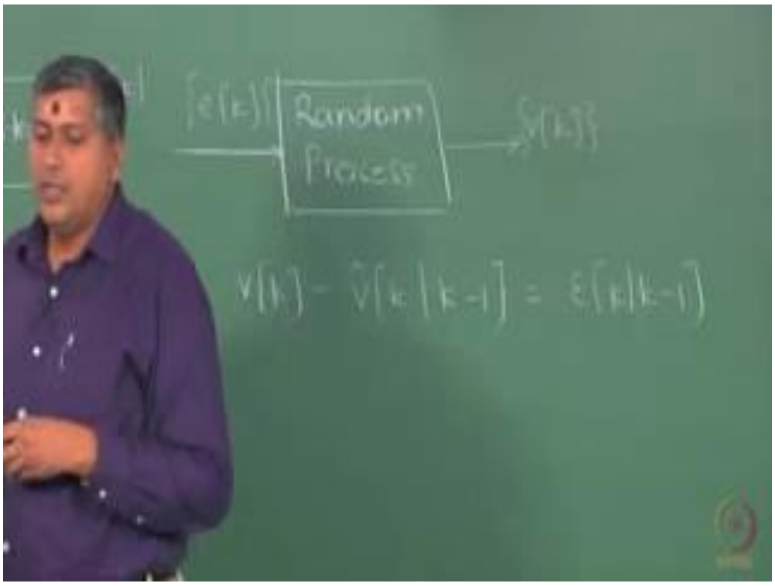


Initially when we begin with random process, we say that it produces a signal random signal, but we do not know what signals are that causes excite this random process. What are the causes I do not know, there are many I do not either know it or I know but I cannot measure them. When that is not such a great thing to in great situation to be in, when you want to build a model for time series that we mean by time series is some kind of evolution equation, just to keep thing simple. And this evolution equation as engineers typically when we look at typically look at physical systems, we are used to this kind of thinking that there is some system which produces some response y; discrete time signal when perturbed by some input or some probe signal; this is how generally we are tuned to thinking of systems there must be a cause producing a response, person just cannot be a really a responding just like that. It may be external or internal there must be cause.

We are used to this kind of systems where there is an external cause generating a response y, but when it comes to random processes the situation is that we only know that it is responding, but we do not know the causes are. Now when you look at history of time series analysis or random processes it is not that in the early years of thinking Wiener or Kolmogorov or the other researchers attempted to bring things into this frame work. Today when you look back in (Refer Time: 04:18), yes it look like whether the explicitly stated it or not essentially the attempts were bring the random process world or random process into this kind of frame work.

So, how did they end up doing that it? As I said it was not necessarily by way of imagination that you could bring the random process into the system world; their initial efforts were always in the direction of forecasting prediction. So, the entire effort was in building some kind of a predictor for v given all the information up to k minus 1; this is that is how you read this v hat of k given all the information; information meaning the series up to k minus 1, how can I now forecast, how can I build mathematically a predictor for v which takes into account all the past information and then came the rest of the theory.

(Refer Slide Time: 05:20)



But let me just give you quickly some preface to what we are going to do discuss today and tomorrow and of course, for rest of the course, where we will encountered average AR modules, ARMA modules and many such modules.

As to why were this is a entire thing come from of assuming this random process is actually generated sorry existed by a white noise sequences. This is what we will be lead to by starting out on a discussion on this. As I said; this was not the initial effort, the initial effort was to develop a forecast we had such that the differences between the truth which is whatever was observed at k and the forecast is something. Let us call that for now as epsilon on k given k minus 1; we call this as a 1 step by ahead prediction error; this is the one step ahead prediction because you are standing at k minus 1, you have all the information and you are making prediction of what happens at k like pretty much in a game in chess; we do that. So, here also we have the same story and this is the one step ahead prediction error.

In order to understand how this comes about starting from this, let us ask a question what is it that I desire of this epsilon that is I am; my goal is to forecast we, but in some particular manner there is something that I want of this forecast such that the prediction error has some characteristics. What kind of characteristics I want to prediction error to have; any idea, we do not how we are going to build this forecast;

let us not to worry about that of now. In whatever method we forecast v, what is it that I desire of the prediction error.

Student: (Refer Time: 07:32).

There should be no correlation; at least in the linear world, it should be white noise in a more stricter sense it should be IID; the Independent on Identically Distributed process, but the bottom line is that there should be no predictability in epsilon. In a linear world we use the term correlations, so we say that epsilon should be such that if I generate the series epsilon; at every k it should have the characteristics of a white noise process that is what we mean by point number 1 it is a bench mark for the test of predictability.

(Refer Slide Time: 08:07)

So, we can actually replace this epsilon on by e k; this is what I ideally want whether we achieved this or not will see; this is what I would decided to have, so that is point number 1. Point number 2, now is let us turn our attention now is concerning this we had second point is concerning is forecast itself. We said earlier regardless of the way I forecast, the residual or the prediction error should have white noise characteristics and that is what we said earlier also test for your module that you build.

Now let us ask briefly, how we would like to forecast, obviously there are numerous ways in which I can process the past to make a forecast if what you mean by numerous ways is; numerous mathematical, operations numerous models that I can build there are in difference ways which I can do that. One of this simplest ways do this is; remember I am given information in the past v k minus 2, v k minus 1 and so on and I am standing here and then of course, v k minus 3 and so on.

So, one of the simplest ways generating such a forecast which is very intuitive what is it?

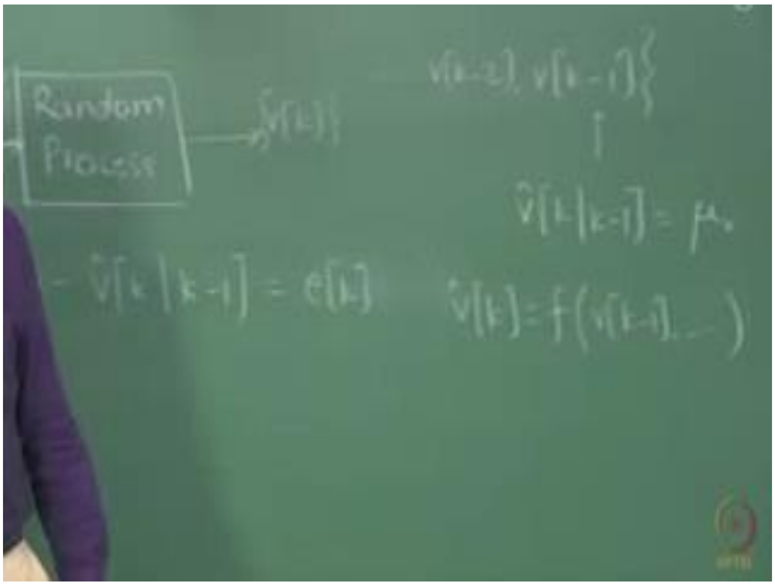
Student: (Refer Time: 09:34).

Well that I cannot compute; I am not going to a have that is theoretical expectation is fine that is a best conditional expectation is a best predictor, but we have already argued that conditional expectations are very difficult compute, what is other way?

Student: (Refer Time: 09:52).

Average; average of what? Average of the series.

(Refer Slide Time: 09:59)



If you want to use a average, so you say that v hat of k given k minus 1 the best forecast is the average itself, but this is nothing but unconditional expectation; it is just see you knew when there is no correlation in the series this is fine. But when there is correlation; that means, there some commonality, there is a some pattern that exists among the series intuitively itself, we know that the average is a suboptimal prediction and we have shown that when I am given another random variable. The best predication conditional expectation; here v k is a random variable that I want to predict given all the bunch of the variables from minus infinity upto k minus 1, so conditional expectation is a best one. So, working with mu is going to be suboptimal and we will discuss more than when we formally discuss prediction theory, but let us move on and ask what would be an intuitive way of developing forecast; let us see what you think.

Student: Weighted average.

Weighted average, what do you mean by weighted average of what?

Student: Of the observation.

Of the observations?

Student: (Refer Time: 11:20).

Fine, so you believe that there is some correlation is lying down; any answer from the other hall. In other words let me be more clear I am seeking some mathematical function that is I would like to write here v develop a forecast which is some function of the past.

Student: (Refer Time: 11:51).

Representing v k as a linear function ok.

Student: (Refer Time: 12:00).

Good. So, any I think the other person also regression, from the other hall is there any answer or suggestion.

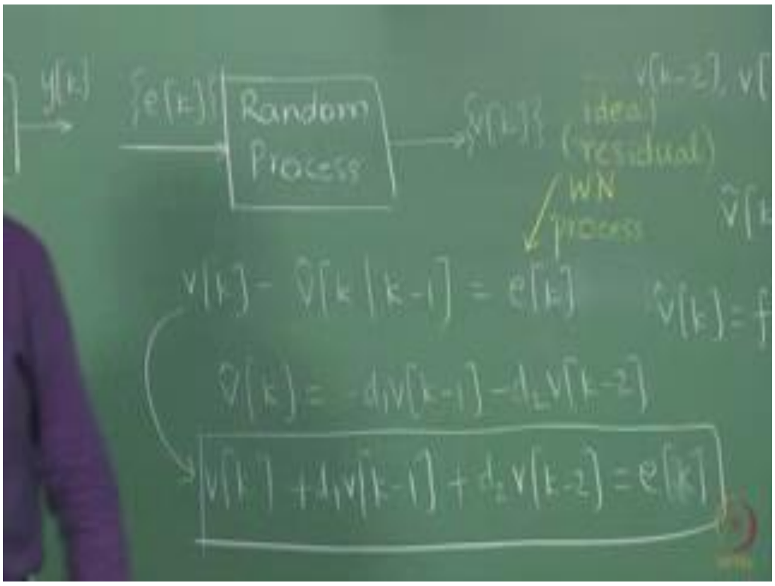
Student: (Refer Time: 12:17).

Of passed observations.

Student: (Refer Time: 12:26).

So I think that is also all most similar to what is being proposed suggested here by here by few students; any other answer good. So, I think you are on the likes of Wiener because Wiener also proposed something similar of course, Kolmogorov thought in a different way.

(Refer Slide Time: 12:51)



But essentially the idea is for example, just as an example this is not the only way you can forecast but going by the suggestions that were given out we could think of using just the two passed observations, what was meant by moving average is I stand at k minus 1, I mean if I want to stand at k again ask; what is the forecast; I am looking at k forecast at k and look at two observations in the past and generate a linear combination or the weighted average whichever way you want to look at. Why not 3? Yes you can. Why not 4? Yes you can it is just an example, in general you can consider all the past, but that depend on the process, we do not know whether all processes required us to work with infinite past as I said so there are so many things that have expiry date, I mean we may have learnt, so many things in our school, but we do not remember the details.

So, everything has a certain memory here we assume just for the sake of discussion that is this particular process has a kind of memory of two; we say some order second order process no when we do this something interesting comes out, whether you and the argument applies to other order as well as, next rewrite this equation here by subsisting for v k was of course, by default if I do not write anything it is one step ahead prediction. So, do not get confused here, so let us substitute and then and we get something interesting to answer, so you have this equation already for the process.

So, look at how things have evolved just out of common sense; we are said that we want to make a forecast and we want to build a forecast such that the residual whatever is left over, something has to be left over because it is a random signal, they cannot predict accurately and that left over should have the characteristics of white noise and then we said as an example, if I forecast this way then I can rewrite this equation in this form. What is that equation that you see in the box, what is the name to such equations in mathematics?

Student: (Refer Time: 15:36).

Difference equations, recursive equation; the general name that the we run into is difference equation form and it is a discrete term analog of differential equations and now how do you view e k as. So, is there a name one technical name is input; is there any other name that is used.

Student: (Refer Time: 16:12).

Source anything else; forcing function, have you heard of forcing functions? It forces a system to respond like I force you to respond by asking you a question; straight to you. So, there are different names, you can think of when you look in this equation the interpretation of e k changes, compared to the interpretation that we had here. At this stage here we said well this should be a white noise process and the interpretation was that it should be this ideal residuals. But the same equation when we substitute for the forecast expression, gives us the different interpretation to this white noise; which is that it is somehow forcing this process v k to respond and that this process is governed by a difference equation in fact, precisely it is called a stochastic difference equation; the forcing function is the stochastic signal. Unlike our regular deterministic you know in the deterministic world we encounter differential equations where the forcing function are deterministic could be a sinusoid or some mathematical function, but here the forcing function is stochastic signal.

Giving us the hint at the kind of models that we will end up looking at, in fact, precisely this is called the auto regressive process of order 2 because it regresses on to itself and it relies on two observations in the past, it has a memory of two at least direct memory of two; it does remember its entire past, but you when you look at it in the face of it only remember it only needs to remember two observations or its two states in the past.

So, at this stage here we can think of this as an input or a forcing function and that is what gives us this interpretation. Earlier there was no e k and suddenly now we seem to have brought in an input, but let us look at the bottom line; where is this e k coming from; is it external? This white noise that we are referring to the e case whether in this equation or this equation, is it coming is it external to the process what do you think?

Student: (Refer Time: 18:56).

What you think, is it external?

Student: (Refer Time: 19:10).

Could be, how?

Student: (Refer Time: 19:17).

But at this moment we are saying there are no, remember the frame work that we set out to discuss ruled out any exogenous effects. No trends nothing said will take a pure random process seemingly driven by itself, this e k where has it come from, it has come from here. It is an integral inevitable component of v k of that signal itself, it is not an external signal at all, it is that part of the signal v k which cannot be predicted and that is what constitutes a random signal.

In a random signal whether there is a predictable component or not, there is always an unpredictable component. If that were not to be exists then we are contradicting ourselves, so this white noise that we are referring to here has take birth from v k, it is internal to v k. What we have done now for modeling purposes or for prediction purposes; took out that part and kind of you know taking the tail of it and kind of attaching itself.

So, it is a self exited process that is another name for random processes, there are many ways of describing a random process; many terms that you can say one of the terms that you will find commonly is e k's endogenous to the process. It is not exogenous, it is not coming from somewhere; that is the difference all the between this is this system and this process here. In this schematic clearly the input is external to the system, where as in this schematic although it looks similar in terms of the drawing, this is not external, it is not exogenous; it is a part of this process that we have kind of pulled out for mathematical purposes, for forecasting purposes.

Of course there is a huge difference; this input is it can be deterministic whereas, this cannot be. It is usually deterministic, but can have some error in it randomness in it, but this is by definitions (Refer Time: 21:33) and it has some specific statistical properties. This is the basic idea in time series model; all along we have said you do not know what the causes are; now we are saying that there is a cause, but this cause is internal; it is endogenous itself exited. Is it true; it is not, if I look at the stock market price or the temperature atmospheric temperature or any other signal for which I am planning to build a time series model, it would be not correct to say that it is a self driven process. So, it is purely a mathematical or statistical abstraction, but it works.

So, do not try to now relate this e k to some external cause; you will not find it like they describe in many news journals and papers and so on, when a market crashes and say some shocks occurred across the market, they are also not able to figure out what the shocks are, they may give you some reasons, but those shocks are this shocks wave here internally. But again may be they are talking of external shocks, but we cannot at least in the univariate random process. In a multivariate random process; for example, when I am predicting relative humidity using temperature, then you can think of this kind of a frame work.

So, in this in such a situation why the response would be relative humidity u; would be temperature, but the point is both are random signals because both are observed variables. The temperature is observed, I am not middling going and middling around of the temperature and performing an experiment, I cannot; fortunately no human being can, it would be crazy doing that.

So, the temperature is an observed variable and the relative humidity is an observed variable; both are random signals, but both are observed. Then you can use some of the systems theory, you know frame work to handle that and of course, the estimation theory frame work as well, but there is another complication in many situations where we given a bunch of variable, I do not know which is the cause, which is the effect that comes into your vector time series molding. So, there is a difference between multivariate time series analysis and vector time series analysis.

Typically when we say multivariate, you will know the cause and the effect but in a vector time series frame work; clearly we may not know which is a cause, which is effect. So, we are jointly modeling that; we do not get into such things in the course we mostly confine ourselves to this or may be a bivariate case.