Applied Time-Series Analysis

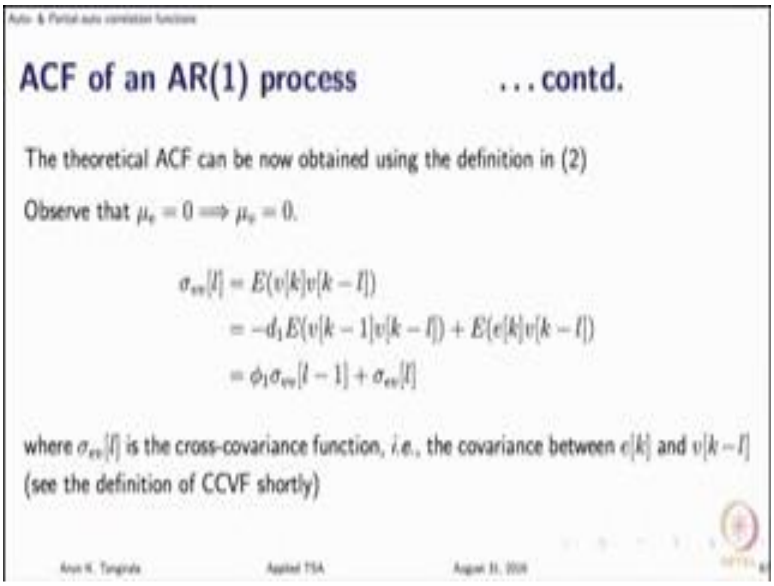
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 34
Lecture 15A - Autocovariance and Autocorrelation Functions-9

Good morning. Let us actually continue our discussion on the Autocorrelation Function of auto regressive Processes. Through our quick discussion and the derivation based on using the autocovariance generating function we were able to show that the autocorrelation function of an auto regressive process diced on exponentially unlike that of a moving average process, in which case the ACF diced on abruptly right after the lag equalling the order of the process.

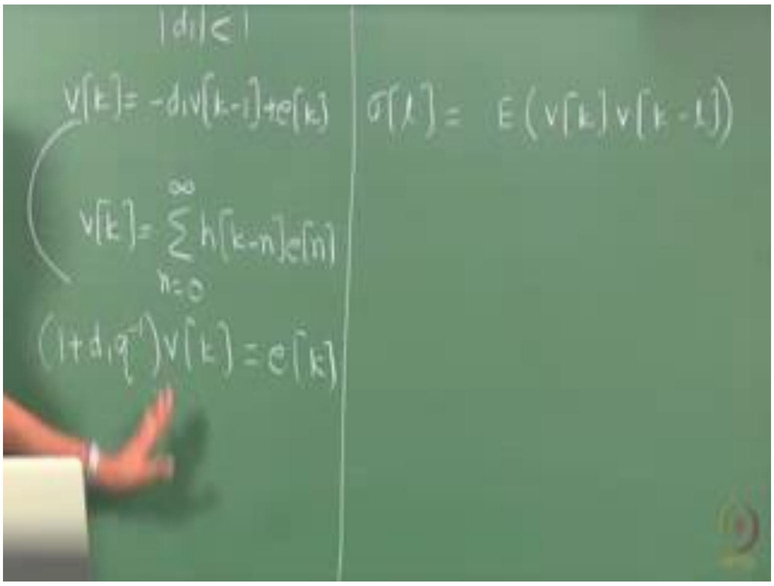
So, what will do now is we will actually derive the ACF in the traditional way. The autocovariance generating function method is in principle suited for all linear random processes, but when it comes to auto regressive processes it is better to stick to the traditional method wherein we will derive set of equations that are very classic in time series analysis; these are known as Yule Walker equations. We have I briefly mention those few lectures ago. The autocovariance generating function is ideally suited for moving average kind of processes.

(Refer Slide Time: 01:31)



So, let us look at this screen now where I show you how to derive the autocovariance function of an auto regressive process. In particular, for the example of AR 1 now I have written it this way, but there is an implicit assumption that the mean is 0. We should first establish that the mean is 0 and then only write this way.

(Refer Slide Time: 01:59)



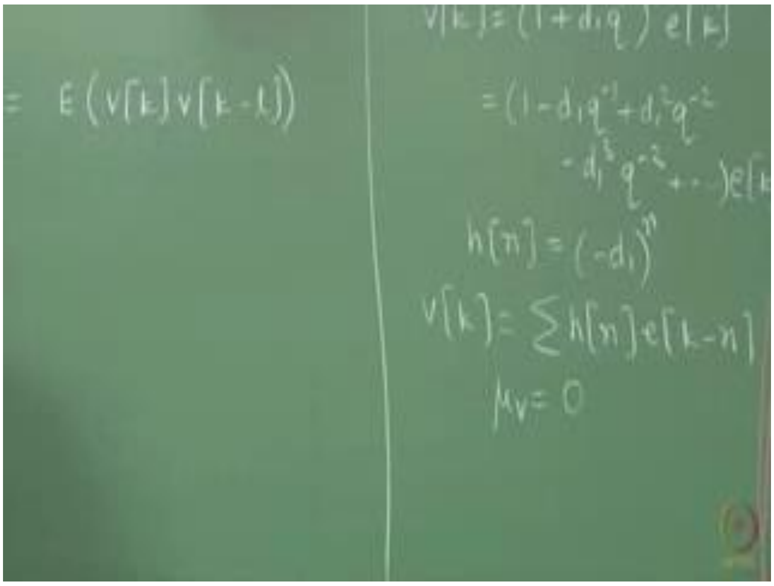
So, let us go back to the process itself which we know is described this way. Now I have chosen minus d 1, there is no particular reason at the movement later on you will realize that it is nice to have the same kind of science in the numerator and denominator of the transfer function operator and that is why I have chosen minus d 1, but different text books may used different notation.

So when you look at this expression here, it is not so obvious that the process v k should have 0 mean. First of all it is not so obvious as to why it should be stationary for all values of d 1. Of course, we have restricted ourselves to the case where the magnitude of d 1 is less than 1, and now we have to figure out what the mean is. First of all will have to guarantee that when d 1 is less than 1 in magnitude it is stationary, and then may be proceed to the calculation of mean. It may be a mistake to assume that expectation of v k and v k minus 1 are the same up front. I know e k is stationary, but there is no guarantee necessarily that this is stationary as well. And this is a habit that you should get in to; you should not assume by default that any equation of this form will necessarily generate stationary v k.

So, one of the ways to arrive at the mean of v k is to recast this into the general linear random process form. What is the general linear random process form? This is a convolution equation that we had written couple of lectures ago and running from like; let us restrict ourselves to casual processes. So, this is the convolution equation that we had written. Obviously, the AR 1 equation the governing equation is not a directly in this form. How do we catch this equation into this form? There are two different ways one is to write recursively; to recursively substitute for v k minus 1, the other is to use a transfer function operator approach or just a shift operator approach.

So, this we know can be written as 1 plus d 1 q inverse v k, operating on v k produces v k-this is we know. Now may be for the benefit of people who cannot see below this I am just going to write here. So, this is the equation that we have; we can now use a long division approach to arrive at this form, although that is not the technically correct way of doing it but it works. The technically correct way of arriving at this equation is a lot more regression, one has to be careful. You know you have to start off with this assumption and then equate the coefficients on both sides and so on, but will not do that.

(Refer Slide Time: 05:36)



What will do is we will just use the approach that works so that we can write v k as 1 plus d 1 q inverse rise to minus 1 operating on e k. As I will repeat this is not the perfect way of doing it technically the correct way of doing it but it works, you are not really committing a claim by adopting this method but it is ok; it is still not be perfect way of doing it. Anyway, so what to be obtain here? When I carry out a expansion series expansion of this what is the series that we obtain; 1 minus d 1 q inverse plus d 1 square q inverse square and so on right minus d 1 cube q rise to minus 3 and so on operating one e k. So, it is never ending series operating on e k.

Now, do you see that we have obtained this form of an equation, is it clear; so we can now write it in the summation form. Now of course the other way of writing this is remember we can also write this as h n e k minus n that is the summand itself we can write as h n e k minus n. So, if you were to write in this fashion what do you see as h n? Now you just have to compare the equation that you have there with the summation; minus d 1 rise to n, is it correct? Everyone is in agreement with this right fine. So, v k has been recast into the standard convolution form with the coefficients given as above here. Now the condition of stationary if you recall for a linear random process requires that the coefficients h b absolutely convergent.

What does that translate to on d 1? Not d 1 less than 1 let us exactly the requirement. So, at least they have proved that if mod d 1 less than is less than 1 you will get stationary and also that is a requirement, in fact we have proved both ways. Now what can you say about the mean of v k given that mod d 1 is less than 1. So, we have guaranteed now v k stationary either now you can go back to this equation here and assume that the mean of v k and v k minus 1 are the same. Am I right you can assume that both are the same, what is answer that you arrive at for 0? Alternatively, you can simply evaluate expectation here and your summing up 0 mean random variables, you should expect 0 mean right. So, these are two different ways of arriving at the answer, but the n result is mu v is 0.

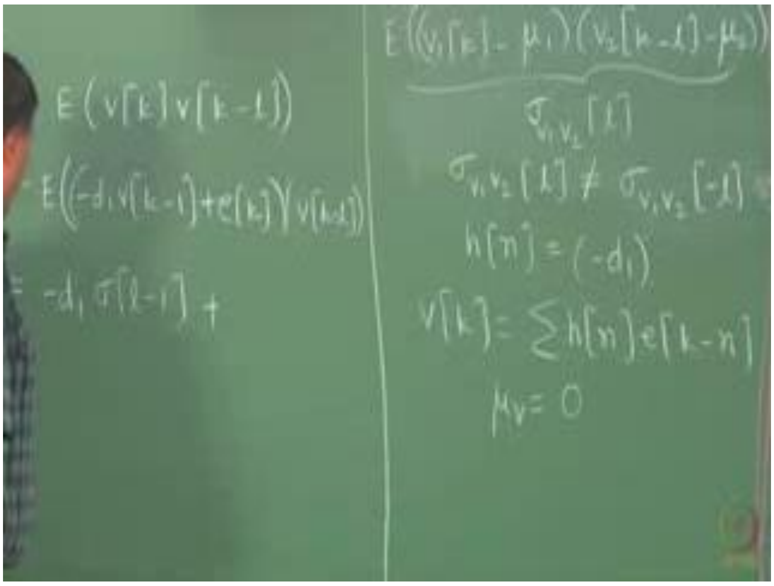
Now we are safe to begin with this expression here. So, there was so much behind that expression here, that is why one has to be really careful in now time series analysis because very quickly things can actually slip out of here hands if you do not verify certain things. Fine, now we proceed further as I have shown on the screen so you see that in place of v k, what I have done is actually I have replaced v k with the governing equation.

(Refer Slide Time: 09:49)

In other words we have written v k as minus d 1 v k minus 1 plus e k times v k minus l, all I have done is our just replaced v k with that. I have not replaced v k minus l for a reason, you know if you start substituting for v k minus l as well then the equations become a bit messy. So, we will only replace v k with its governing equation so that you get minus d 1 sigma at l minus 1; remember they expectation of v k minus 1 times v k minus l how many observations apart are there, they are l minus 1 observations apart or instants apart sorry.

Plus you have a cross term here, now the second term is not any autocovariance that we have seen before it is the, what is it? It is a cross variance good. So, it is a cross-covariance between e k and v k minus l. Now we have not defined formally cross-covariance, but that is no big deal we can straight away define what we mean by cross-covariance.

(Refer Slide Time: 11:10)



So, when I have two sequences v 1 and v 2 or two you can say discrete time random signals then the cross-covariance between them is defined as expectation of v 1 k minus mu 1 times v 2 k minus l minus mu 2. So, this is your sigma v 1 v 2 at lag l; this is the covariance. Again at the heart of this is covariance, this is nothing mu here except that now we are correlating observation at the k-th instant of one series and another observation of another series located at k minus l. We assume that v 1 and v 2 are jointly stationary, that is a very important requirement otherwise your cross-covariance is not function of lag alone; it would be a function of k and k minus l. So, that is an implicit assumption, but a very important assumption.

Mu 1 is the mean of the first signal and mu 2 likewise is a mean of v 2. Now here the ordering of subscripts does matter, you have to be very careful because there is essence of direction here. Through this we are asking how much does v 1 k influence v 2 k minus l or vice versa, we do not know that direction but certainly we know that we are looking at the k-th instant of one series and k minus l th instant of another series; l can be again here positive or negative or even 0. When l is 0 it is nothing but cross-covariance, your classical cross-covariance that we have discussed earlier.

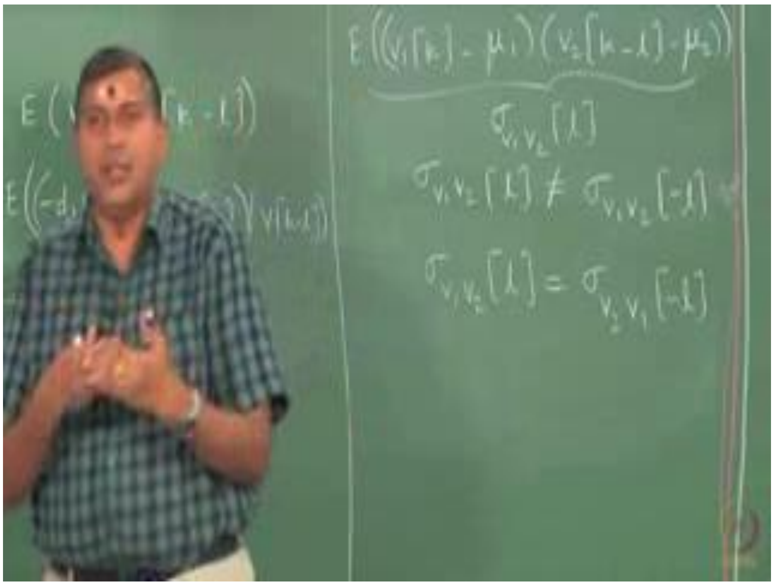
And obviously, the cross-covariance function is an a symmetric function; that means, the way v 1 at this instant influences v 2 at a later instant or even is influenced by the past of v 2 is not necessarily the same way when you change the direction of analysis. So, what we mean by this is; sigma v 1 v 2 at lag l is not equal to necessarily sigma v 1 v 2 at minus l, whereas the autocovariance is symmetric. So, that is a prime difference between the autocovariance and cross-covariance.

Again you should not forget the interpretation of cross-covariance. What we are studying through cross-covariances, how much does one signal linearly influence another signal may be now or at a later instant or is influenced by another series in the past. And it is an extremely useful measure; you will find this as one of the most widely used statistical measures in signal analysis. There is no field where cross-covariance function has not been applied and there are many applications in the classical application is in the estimation of delays.

In radar signal processing we know the basic principle right. There is a signal does transmitted at one station and the signal actually hits the object that is flying they are and then is transmitted back. Assuming under ideal conditions you know there is no that both the transmitted and the received signal travel with the same velocity. So, you look at the difference in the time of arrival time of time at which you have transmitted and the time at which you have received the signal; knowing the delay right you can actually and the distance that is probably travel you can know the velocity of the flying object or if you know the velocity or the flying object you can figure out how far the object is.

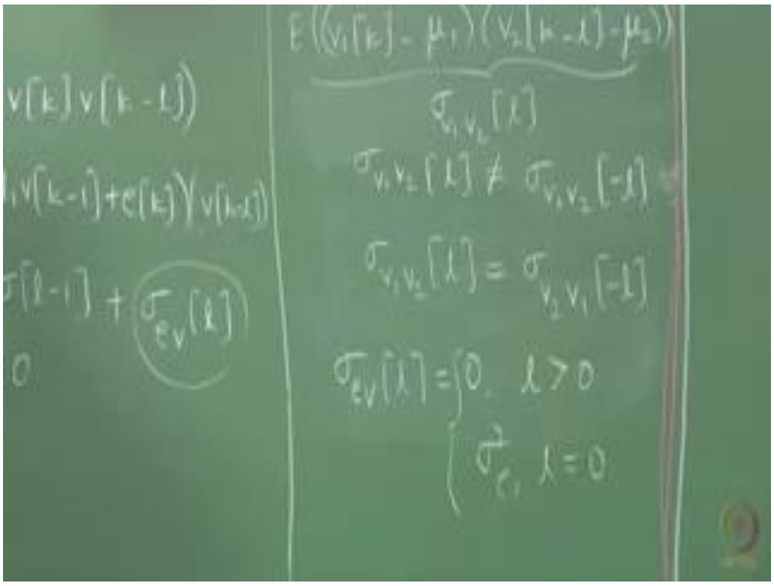
So, we can do many things, but this is one of the crucial measures used in radar signal processing. Not only radar signal processing everywhere where you are interested in estimating delays. More over it finds a lot of use in system identification where we estimate impulse response coefficients and so on. We will not get into that now.

(Refer Slide Time: 15:29)



So, the other property that is of interest to us is sigma v 1 v 2 at lag l is sigma v 2 v 1 at minus l right, and you should be able to see that. So, v 1 v 2 measures the covariance between v 1 at there is at lag l, measures the covariance between v 1 at k and v 2 at k minus l; whereas v 2 v 1 looks at the influence between v 2 k and v 1 at k minus l, obviously there should be kind of a conjugate symmetry. So, keep these in mind that always co-variances are measuring linear influences.

(Refer Slide Time: 16:19)



Coming back to the problem here now we can write this as sigma e v at lag l. And when you right sigma e v at lag l do not forget the interpretations and so on; you should keep in mind the interpretation as well as the definition. It is essentially a measure of how much e k, how much of commonality linear in a linear sense exist between e k and v k minus l.

Now, before we proceed further we will restrict ourselves and say that it is sufficient to restrict ourselves to non-negative lags, because we are looking at autocovariance it is sufficient to theoretically evaluate over one side of the lag axis. So, let us restrict ourselves to non-negative lags. Now the one of the main obstacles is this term here. So, once we get read of this or once we know how to handle this you can proceed further. What you thing is the answer, sigma a v at lag l what is the cross-covariance between e k and v k minus l since we are restricting ourselves to non-negative lags; v k minus l is either now or in the past, that is the horizon that we are looking at.

So now, you have to ask; what is the cross-covariance between a shock wave at this instant e k, and the generated signal at this instant or in the past. If l is 0?

Student: (Refer Time: 18:00).

Two answers; 0 for all lags, greater than or equal to 0, we want to revisit your answers. Anybody from the other hall?

Student: (Refer Time: 18:15).

And does anybody want to answer what is this value for they are one process at all lags greater than or equal to 0. No answer from there. Any from here, any other answer; so have you revisited your answer one of you said that it is 0 at all lags greater than or equal to 0; what is the difficulty in arriving at this answer I would like to know. Question is; what is the influence you can think of it as influence you can think of it has commonality whichever where you want to look at.

Let us look what is a commonality between e k it helps a lot, between e k and v k minus l lot of times I see students actually getting into an infinite loop here. So, you start off with e k and v k minus l and then you write e k as v k plus d 1 v k minus 1 and then you start writing further equations or you keep substituting for v k minus l recursively and then you says sir there was not enough time to answer the question.

Student: (Refer Time: 19:44).

But, your write just always in mathematics there are this similarities which are they to trap you. So, for l greater than 0 your statement is correct right. What about l equal 0?

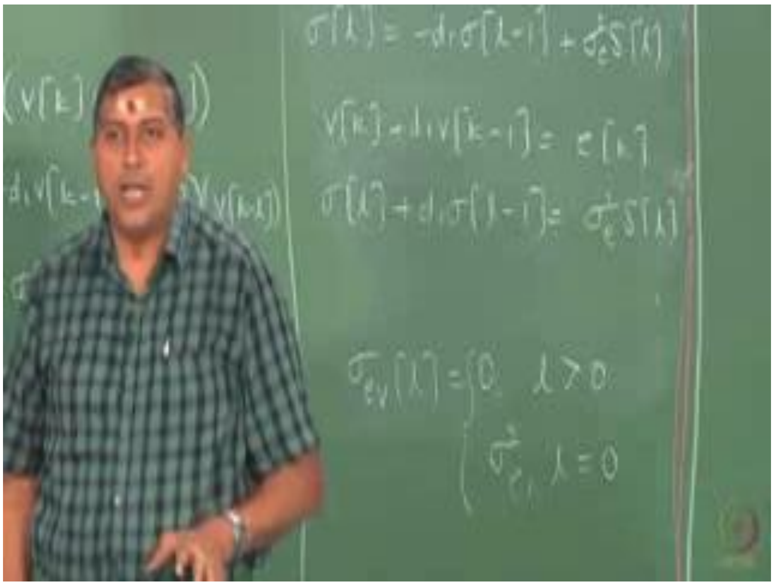
Student: (Refer Time: 20:06).

V k contains influence of v k right, very good, so that see observation that you want to make if you want get out of that infinite loop. You have to ask what is v k minus l made up of. It is made up of all the past effects or l greater than 0 all the past shock waves and by definition e k is uncorrelated with its own past. So, we can straight away right that sigma e v at lag l is 0 for all lags greater than 0. You have to be careful; you cannot make this statement for negative lags.

So, now you see why we smartly restricted ourselves for non-negative lags. And at l equal 0 what is the answer? At l equal 0 you looking at the covariance between e k and v k, but by our own assumption that is a generating equation v k is made of its past plus e k. This we have now determined to be uncorrelated with e k. So what you get? Expectation of e k times v k would be simply sigma square e, very good.

So, this is how with clever reasoning you sort out things and proceed further. Now, we know that the sigma e v at lag l is this.

(Refer Slide Time: 21:40)



Therefore I can write and equation here as for l greater than 0 as sigma at lag l, I am going to drop the of course I have not use the subscripts. So, we can proceed minus d 1 sigma at l minus 1 plus we can say that it is sigma square e delta of l, where delta is your Kronecker delta function. Of course one sided, even you can use two sided does not matter, but this equation itself is valid for l greater than or equal to 0. We will keep remaining ourselves of that.

What is the role of this a Kronecker delta at lag 0 it is a unit function unit impulse, at lag 0 it assumes a value of 1 and 0 at all over lags. Now, very interestingly if you compare the generating equation with the equation that we have here; so let me rewrite sigma l plus d 1 sigma at l minus 1 is sigma square e times delta of l. You see a very strong similarity between the generating equation for v k and the generating equation for the autocovariance. That is the feature or the hall mark of auto regressive processes that is a big advantage. It is very easy to compute autocovariances of auto regressive processes; because once you are given the generating equation you can straight away write the generating equation for the autocovariance.

Although, I am only showing this for AR 1 process you can extend is argument to other auto regressive processes as well. You can see if it was an AR 2 you would arrive at the same equation. But have a problem is not yet completely solved we still need the expression for the autocovariance. So what we do now how do we proceed from here. So, as the first observation which is very important that the generating equation for the auto variance is the same at least the homogeneous part is the same as the one for the process itself, only the forcing functions are different. And the other prime difference is this is the stochastic difference equation, while this is the deterministic difference equation you should be careful because sigmas are deterministic quantities.

So, the forcing function for the process itself is white noise, whereas the forcing function for the autocovariances what is it; it is the autocovariance of the white noise itself right. In fact, on the right hand side why have we come up with this because it is a autocovariance of course we are looking at non-negative lags, but if you keep that a side momentarily what you see essentially is white noise is forcing the process to generate v k, it is not a covariance is forcing the process to generate its autocovariance. We do not see this kind of necessarily a property for moving average, it was completely different there.