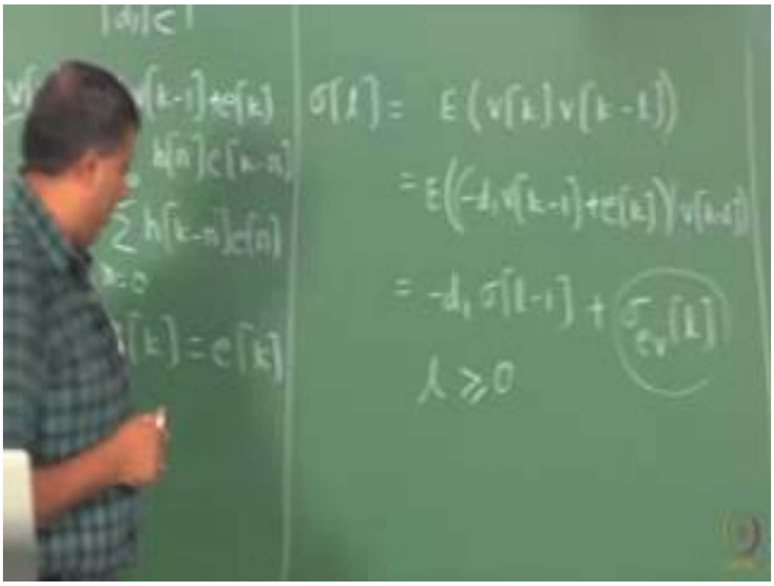
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 35
Lecture 15B - Partial Autocorrelation Functions

Anyway now, how do we proceed from here?

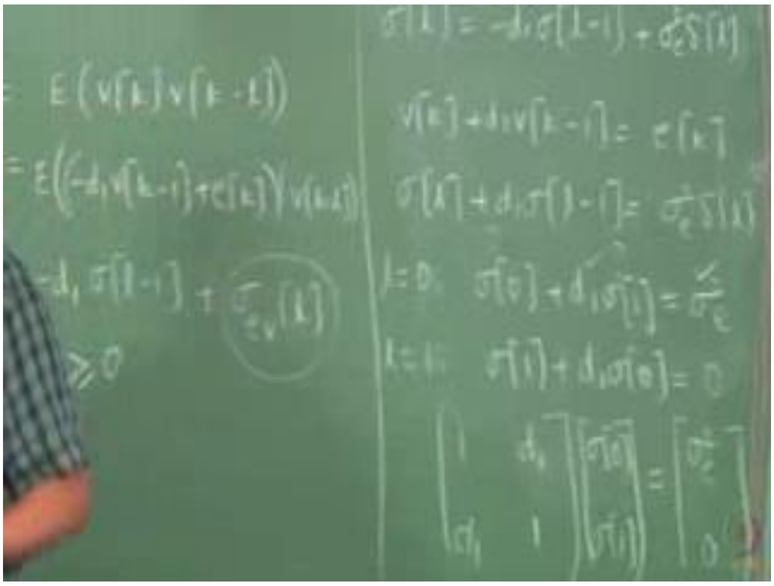
(Refer Slide Time: 00:14)



I want what do I want? I want this sigma at lag 0 and lag 1 and lag 2 and so on. So, what is the next step? Logical step from here, what is that I mean the equation is given it is like solution to the difference equation right, what else is that to do? You just substitute no, no you just substitute l equal 0 write 1 equation; write l equal I mean then again for at l equals 1 you write another equation and see if your able to solve. Essentially, now it boils down to solving at difference equation and that is what again you should remember for auto regressive processes, one ends up solving a bunch of difference equations or you can say the single difference equation here.

I do not know how many of you have studied at least the basics of solving difference equation, it is not much different from solving differential equations right, you set up the characteristic equation and then you should be given some initial conditions and then you should be able to evaluate your smiling, differential equation itself is forgotten it happens many things you forget.

(Refer Slide Time: 01:36)



So, what do we do here, substitute l equal 0; at l equal 0 we write the equation sigma 0 plus d 1, sigma at minus 1 what is sigma at minus 1? Sigma 1 very good; a symmetry saves us equal sigma square e, now I need to know sigma 0 to get sigma 1 therefore, let us see if equals 1 if which is something informative and fortunately it does what you get on the right hand side 0.

Fortunately now, we do not have to proceed further, we have a couple of equations here 2 equations and 2 unknowns; now you should remember this scenario, at this moment we are given the model and we are interested in the theoretical auto covariance functions, but in reality I am actually given the series, from where I estimate the s e v f from where I estimate the model parameters. So, there its reverse incident, right now we are looking at the forward case I am given the model I would like to know what its auto covariance signatures are, but in practice I am not given the model, I am out to actually estimate the model. So, I start with the series, estimate the auto covariance and if I have to use the method of moments approach, I would assume that the estimate satisfy this equations as well; in which case do I have enough equations to estimate the model? Yes or no. So, what is that now?

So, the question is if I am given the series from where I estimate the s e v f and if I assume that the estimates satisfy this theoretical equations do I have enough set of equations to estimate the model? And what we mean by model is not just the coefficients also sigma square e. So, it is useful for both purposes, this set of equations are use its useful for obtaining theoretical auto covariance as well as estimating the model given the other. Here we are given d 1 and sigma square e, we know this and we would like to know what are these, these are the set of equations which are known as the Yule walker equations named after Yule and walker to independently discover, it is independent discovery is only a matter of the pass this station, you can sit at different places and claim in early 1900 or may be pre internet era or free mobile era.

That yes you have clay you have work done this independently. Nowadays you have talked about dependence already so let us not get in to that. So, you it is very you have to be together and then show in of evidence that you have worked down this jointly fine. So, that is it is. So, you get your equations how do we set up your equations now in a matrix form, how do you write your equations depends on what you want to know. What I mean by that is in this case I am given the model and I would like to know the auto covariance's, but tomorrow if tomorrow does not mean Friday alone not just on the Tuesday.

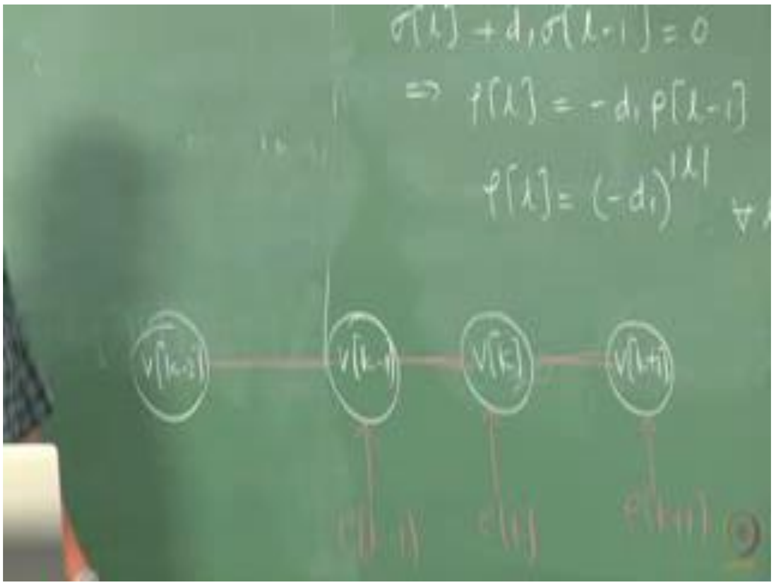
But tomorrow means any other day if and given the auto variances and I have to estimate the model, then these entries will change; write now what are the unknowns, Sigma 0 and sigma 1 right. So, how do I fill up these matrices now? What does the fact matrix contain here the big one?

Student: (Refer Time: 05:24).

Very good 1 and d 1, d 1, 1 very good and then I have here sigma 0 sigma 1 and here I have sigma square e and 0. So, straight away you can arrive at the answer right, I hope you know how to compute the determinant of a matrix. So, what is sigma 0? Sigma square e by 1 minus d 1 square; exactly the answer that we got through the auto covariance generating function.

And sigma 1 is minus d 1 times sigma 0; that is the second equation is telling me right what about sigma 2 how do I get it? Just go back to the generating equation and keep recursively generating your auto co variances. In fact, this itself tells us. In fact, I do not have to go through this to generate the auto correlation function, I have to go through this if I have to work out the auto covariance's, if I only need the auto correlation then I can straight away go back to this equation and write here; since you need only auto correlations I can rewrite that equation.

(Refer Slide Time: 06:45)



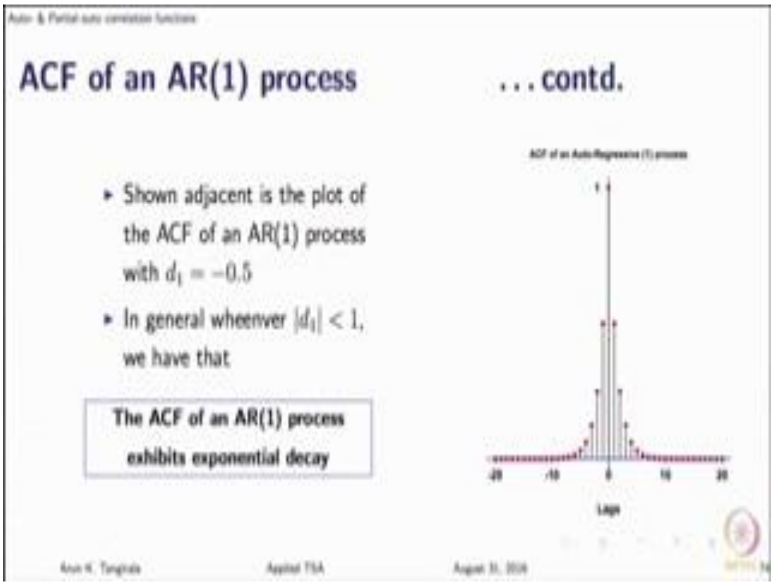
I do not need the auto correlation lag 0 by definition it is 1. So, I can focus 1 non 0 negative non negative lags, 0 therefore I get rho l as what you I get as rho l; minus d 1 rho l minus 1.

But we know that l equal 0, what is a value of auto correlation? 1 all though this equation you have written for l greater than 0, but I know already rho of 0 is 1. So, I can use that and therefore, write this as rho l, I can start from l equals 1 and then write for general l so that do not forget to write this model right. In fact, this is true for all l, the same answer that we arrived at yesterday. So, in general now the procedure for arriving at the auto variance or the auto correlation function is to write the Yule walker equations and then solve. So, for example, if I want the auto covariance's, you can see because it is AR 1 I had to set up two equations; if it is AR 2 a set up 3 equations, because I would need sigma at 0 1 and 2.

Extended this argument for an ARP process, let p is the order I would have to set up p plus 1 equation and solve for the auto covariances from lag 0 to p and then p plus 1 onwards you can use the generating equation to recursively compute. That is the general procedure for computing the auto covariance well is straight forward, you should remember never the less that you will be working with Yule walker equations. So, as I said is Yule walker equations or also useful in estimating the model parameters, which we will revisit in estimation theory when we learn how to estimate models we will revisit this Yule walker equations. In fact, today Yule walker equations is believed to belong to the class of this method of movements, method of movements assumes that the estimated movements satisfy the same set of theoretical equations that theoretical movement satisfy.

Whether that works for all classes for models will see, but I can tell you that this Yule walker equations are very good or this method of moments is very good for is at estimating auto regressive models and not so good at estimating moving average models; what I mean by this is yesterday we when through a small exercise of recovering the model given the auto correlations of an Ma 1 process correct, that is also method of moments approach, but you can show later on that it yields so called in efficient estimates; that means, estimates with large errors as compare to Yule walker equations for AR, which generate efficient estimates of the auto regressive model under fairly relax conditions. So, I think we can now skip and I this is just a procedure that we went through and this is these are set of equations and here you have on the screen graphical display of the auto regressive process of order 1 with the coefficient d 1 said to minus 0.5 and you can see there it diced down exponentially.

(Refer Slide Time: 10:31)



But what is clear is from this exponential d k I cannot figure out what is the order and that is the subject of discussion for the next 10 minutes, it is a fairly straight forward thing to resolve, since we have already studies partial correlations, but before I proceed further let me tell you this exponential d k characteristic of an auto regressive process is true for any general auto regressive process; that means, not just AR 1, it is a characteristic of ARP process. How do you justify this? Well remember that the auto variance function of an auto regressive process is governed by the difference equation there right now for those of you who are familiar with the theory of difference equations as I said earlier you start of your characteristic equation root.

For AR 1 what is the characteristic what is the root of the characteristic equation? What is the characteristic equation by the way of that difference equation correct? So, what is the root? Minus 1 might even its only two possibly your trying out, why is there so much silence, what is the root of the characteristic equation? Common you should be able to tell me what is the root of the characteristic equation, you should give you (Refer Time: 12:15) why there is so much silence, all the brilliant people that are sitting here are taking time to think of what is root of the characteristics equation there? Minus d 1, is only 1 root I hope you realize that; if I had an ARP process I would have a characteristic equation p roots and if those p roots are lambda 1 to lambda p then we know from the theory of difference equation it is not much difference from the theory of differential equations.

In differential equations we also solve use a characteristic equation roots method, where we write the generic solution right in a continuous time case that is in the differential equation case we would write c 1 e e to the lambda 1 t, plus c 2 e to the lambda 2 t and so on.

In the difference equation case the concept is the same, but the form of the solution is different you would write c 1 times lambda 1 rise to k, plus c 1 time c 2 times lambda 2 rise to k in this case l. So, you would write here sigma l as a generic form as some c 1 times lambda to the power of l mod l because its symmetric, that is because there is only 1 characteristic root; if there are p characteristic roots then you would write sigma c I lambda I rise to model and those c I s are determined by a set of known conditions. So, without going further into that, all you can see is that the generic auto covariance solution to the auto covariance function of an auto regressive process has an exponential form.

So, exponential d k; when we talk of exponential d ks in discrete time do not always thing e to the power of something, in different in the discrete time or discrete index domain exponential typically means something power the index. So, that is what we mean you have an exponential d k in general as well. Now let us ask how do we figure out what is the order of the auto regressive process given the auto correlation function I cannot look at the exponential d k and say I know that it should be of AR 1 its going to be very tough right, to be able to answer whether it is AR 1 or AR 2 and so on. On the other hand for the moving average process it was straight forward.

So, to answer that question we have to go back and ask why do I see and exponential d k, what is the exponential d k suggesting? If you look at the a c of plot carefully, what is it telling me? It says at lag 2, at lag 1 I can straight away see by looking at the generating equation this or here I have written as well, it is pretty clear that there is you should expect a correlation between V K and V K minus 1. So, I can explain why there is a correlation at lag 1. How do we explain the correlation at lag 2? I do not see that in equation right. So, the correlation there it is showing at lag 2 can be explain; however, when it comes to guessing the order, what do we mean by order here, what do we mean by first order, does it mean that V K minus two does not affect v k?

So, what is the meaning of order? If you crack that them the problem is solved and we can straight away go to the solution.

Student: (Refer Time: 16:09).

How many?

Student: (Refer Time: 16:10).

Can you put it in a better way using how many differences?

Student: (Refer Time: 16:16).

Imagine if we have trouble how much trouble lawyers should have yes ok.

Student: (Refer Time: 16:31).

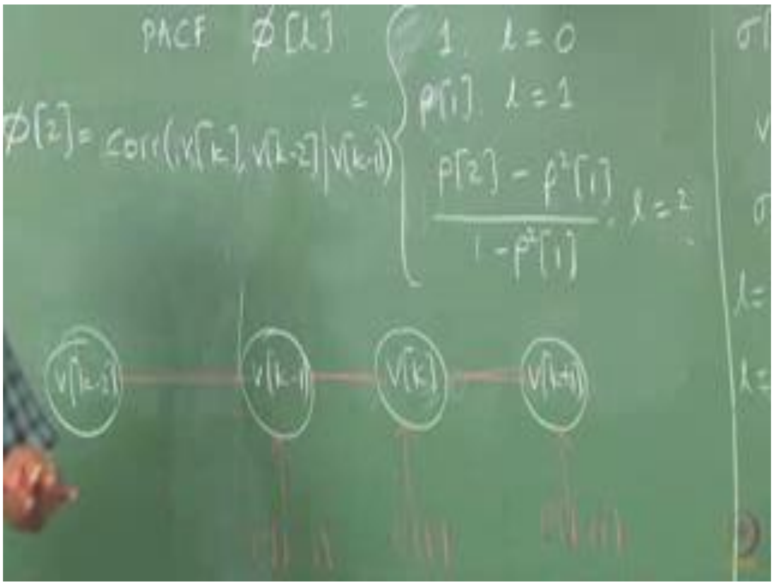
Number of differences no, that should be dangerous, this is even if you difference ones will not get white noise that is only for a purely integrating process, any other try attempt to define the order, anybody from the other room. So, let me help you define. So, let us draw a graphical representation of what is happening in the auto regressive process. So, that it becomes easy to see what is happening. So, here you have V K minus 1 right and then of course, it is an infinite chain there of observations and let say here I have V K minus 2. So, what is happening if I have to draw graphically the lines here, I see that V K minus 1 effects V K right from the equation and V K should also influence V K minus V K plus 1 and if I were to go back in time V K minus 2 also influences V K minus 1 and further at each instant we have this innovation as we call the innovations is nothing, but the creative thing that is happening which is white noise.

E k plus 1; so I have e k plus 1 e k and e k minus 1 and so on and likewise here at k minus 2, e k minus 2. What we mean by order is the number of observations in the past that have a direct influence on the observation at this moment. So, if I am standing at k, I am asking how many of them have direct link with the V K. So, in this case here only V K minus 1 is directly influencing V K; V K minus 2 is indirectly influencing V K, but correlation does not know that that is the main drawback or you can say it is a drawback or it is a plus point depending on the application, correlation does not know it you have not told that there is something called direct and indirect and so on.

All your asking it to do is evaluate the correlation between two observations; it says yes their correlated, unless you tell that yes there is V K minus 1, that is possibly connecting these two, it wouldn’t know right and that is what is the issue here, correlation is measuring the direct a total effect and we have talked about earlier in the case of random variables and there we use a term confounding, that is exactly the issue that we have run in to here. So, it is no different except that now we are not talking of any random variables, we are talking of observations of a random process that is all. So, we bring that theory that we have learnt how did we resolve the issue there? Partial correlation, the same concept we apply here and instead of evaluating regular correlation, we evaluate partial correlation and see if it helps us the partial auto correlation now as we called because you want apply to the observations of the same series, should be able to resolve for us.

So, for example, here I can straight away right the partial auto correlation given all the background that we have, given all the background that have on partial correlation first of all if I have to write the partial correlation or auto correlation let us denote back by phi.

(Refer Slide Time: 20:23)



So, phi will be use to denote the partial auto correlation, this is the PACF auto correlation function, if I have to ask you what is phi at lag 0 for the AR 1 process; at lag 0 with there is no definition because auto correlation is defined. In fact, if you turn to the literature strictly speaking the PACF at lag 0 is not defined, but just to keep things consistent, we will define it has 1; we should be carefully strictly speaking there is no definition of partial auto correlation because there is no need, but just to keep things intact some of the software packages do plot the PACF at lag 0.

So, will keep this as 1 that is all it does not matter, what about at lag 1? Will focus on the non negative lags PACF also symmetric function like DACF, because it is a stationary process; so what is the PACF at lag 1? That means, you are asking what is it direct correlation between V K and V K minus 1, what is that same as ACF right. So, you have here this as nothing, but rho at lag 1; what about at lag 2? We had an expressions at lag 2 what is it that we have measuring we are asking what. So, at lag 2 let me write down here phi at lag 2 is correlation between V K and V K minus 2, given what V K minus 1 very good right and we had derived an expression for partial correlation between x and y given the confounding variable z. V K minus 1 is the confounding variable for us.

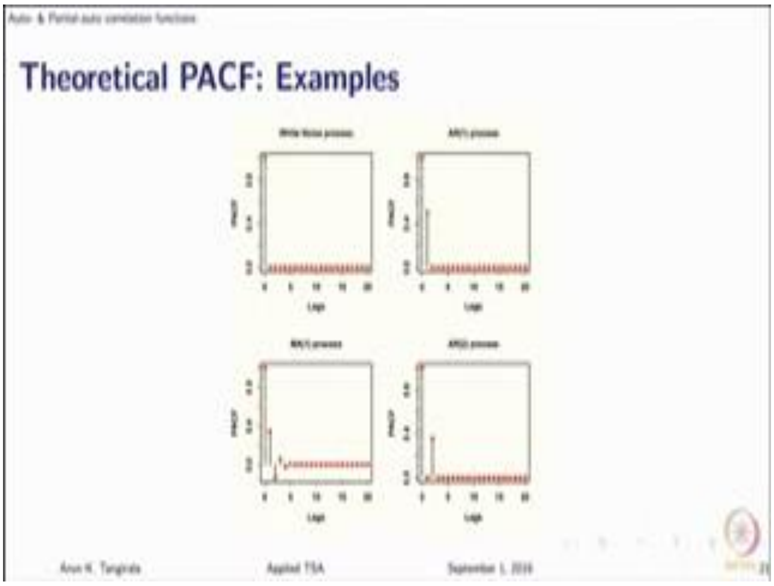
So, can you use at expression and quickly tell me what is the PACF of a random stationary process at lag 2; you had some expression right rho x y its rho x y minus rho x z times rho y z by ray. So, what is what are x and y for us here? V K and V K minus 2; that would be in the numerator you would have rho at lag 2, minus rho 1 very good by 1 minus p square very good and this is at l equals 2 and 1 can go 1 writing the PACF at lag 3 and so on, but when you go to lag 3 what are the confounding variables? V K minus 1 and V K minus 2; you would have to use a more I had given you an expression when you have a vector of confounding variables. Remember, fortunately we do not use that approach in practice. In practice there is a much more elegant way of computing the partial auto correlation function.

But I will worry about at a big later, very quickly like me tell you that whatever we have written here is a is the PACF for any linear stationary random process, it has got nothing to do with the underline process, but now if the underline process is AR 1 right if the underline process is AR 1 what is the PACF evaluate?

(Refer Slide Time: 24:13)

What does it evaluate for AR 1 process I do not have to evaluate for any process there is no point I mean in evaluating at lag 1; at lag 0 at sorry at lag 0 at lag 1 PACF would be minus d 1 you already know that right; what at what about at lag 2? 0 right at lag 2 at 0 and in fact, you can show that lag 3 onwards also it is 0.

(Refer Slide Time: 24:45).



So, as you see on his screen I have shown you the PACF of, of course different processes white noise on the left corner and AR 1 on the a right corner, you can see that PACF has exactly the same behaviour as ACF for moving average processes.

You can say that they are duels the way ACF works out for Ma processes, PACF works out for AR processes. In fact, if you look at the left bottom plot, I show you the PACF of MA 1, what is happening to the PACF? It is not exactly dying down after lag 1, same as we observed for ACF of auto regressive, ACF did not die down all through the nature in which the ACF die down for AR 1 is not necessarily the same way as PACF dies down for MA 1, but in both cases I do not get any information on the order. So, in practice in time series analysis, I should look at both the ACF and the PACF and then come to a decision whether I want to fit AR or MA model, the situation is even a bit more complicated later on because we may have to fit an ARMA model, in which case what do we do, will talk about that next week and the last plot that you sees for AR 2.

So, you can see the reason I chose that example is PACF at lag 1 0 as you can see, but PACF at lag 2 is not 0 and thereafter it remains at 0. So, you should be miss led by just the PACF at 1 lag being 0, you have to plot the PACF over a range of lags and make sure that you correctly determine the order. Now I will conclude the class by saying by going back to the point saying, where we said we do not have to really compute the PACF this way, we compute the PACF in a different way using a very important observation which will not prove at this point and time when you talk of auto regressive processes at that time will prove it, which is that the partial auto correlation function at any lag l is sorry that are the wrong slide; that any lag l is the last coefficient of the auto regressive model at lag l what we mean by this is.

Suppose I want to compute PACF at lag 1, I fit an AR 1 model at lag 1, what would may fitting? I solve this Yule walker equations, that is I solve for d 1 and sigma square e for the model coefficients, that is what we mean by fitting and AR 1 model at least using Yule walkers method and then look at the coefficient at lag 1; that means, in this case d 1 that or minus d 1 you can say although as I PACF its coefficient at lag 1 it is nothing, but minus d 1, when you write at and they auto regressive form that is exactly what we have here, the PACF at lag 1 is the coefficient of the last coefficient of the AR l model; when I move to ARPACF at lag 2, again I repeat the procedure I fit AR 2 model and pluck out the last coefficient of the AR 2 model.

So, what this tells me is if I have to fit an AR 2 model to an AR 1 process, I would end up getting that last coefficient as 0 and in fact, fit any AR from model of higher order the necessary; theoretically all those extra coefficients would be 0. Using this 1 can develop a computationally efficient and recursive algorithm which was done already sorry you cannot contribute much now it is a due to Durbin and Livingston; that is say algorithm that is used in the computation of theoretical PACF or even estimation of PACF that you use in R. So, as the simple exercise go back to your R and use the ARMA ACF routine, instead of asking for ACF there is an option for asking for theoretical PACF and reproduce a plots that I have there I am showing you on the screen.