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

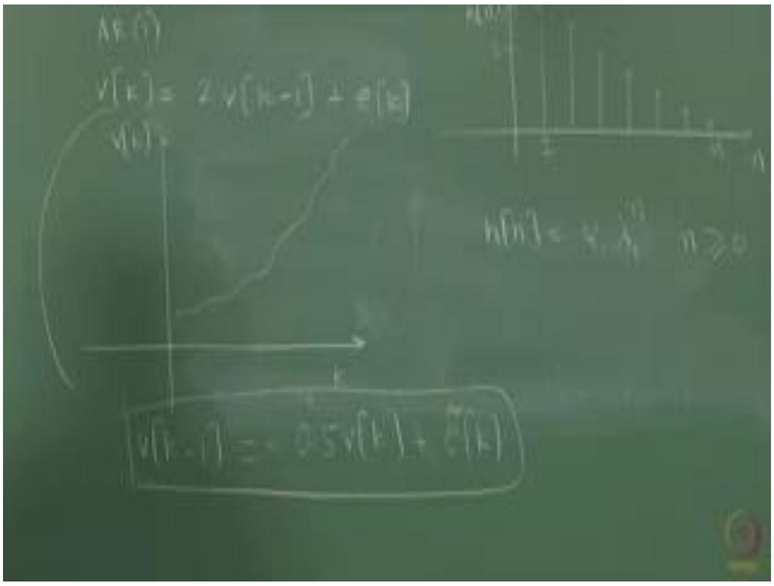
Lecture – 47
Lecture 20B - Models for Linear Stationary Processes 11

The transfer function operator form, now it is fairly clear to us how the autoregressive model is bond, it is bond by assuming that the impulse response decays according to a power law, whether we stated explicitly or not henceforth you should always assume that the movement you are fitting an autoregressive model to a series you are implicitly saying that the impulse response for that series for that process is decaying according to a power law.

We will meet another statement as a consequence of this using AR model. So, let us move on and discuss the 2 features of autoregressive model, one is this stationarity aspects which we have more or less I will discuss, but I make a formal statement and there is something more to discuss about stationarity and the second aspect is the ACF itself what you know how does a ACF look like, we have seen that earlier, but we will formalize now will spend a bit of a time with Yule Walker equations.

The transfer function operator form is fairly straight forward, it has a denominator form it is in a rational form, but only the denominator is a polynomial in queue inverse in contrast to the moving average form and now let us turn to a stationarity again, we have already stated I am going to skip this formal statement, but let us study an important aspects of stationarity and understand another dimension of this stationarity issue with AR processes what we are saying is any ARP process is stationary if and only if the roots of the characteristic equation are outside or inside the unit circle depending on whether you looking at z inverse or z in terms of z inverse so that is fine.

(Refer Slide Time: 02:15)



Does it mean now that any process, for example if take a process an AR process, let us say is 2 v k is 2 v k minus 1 plus e k, this is an AR1 process, does it mean that this process is non stationary? Yes or no?

Yes, right? This is I am asking such simple question, you will be thinking, it must be a trap somewhere as per the statement it is, is not it and it is true, if you were to simulate this, you would see an explosion of v k, it runs away. So, it is not stationary, you do not even have to think about the expectations and so on, just look at v k is just becomes unbounded not in the deterministic, in the sense not as much as a deterministic signal would go, but it will definitely become unbounded. Now it turns out that although this process is stationary in the sense of the theorem that we have given, they are non stationary in the sense of theorem that I have given, it can still be given a stationary representation. What we meant by v k running of is, if you were to plot v k versus time, you would see some kind of growth like this right from the that u point it is non stationary, it turns out that I can give a stationary representation to this process if I am going to budge a bit on this notion of causality.

We have actually said earlier when we said h of n is alpha 1 lambda 1 power n no alpha lambda, 1 satisfying your usual condition, but importantly we said n greater than or equal to 0, all the way I am not straightened, it was simplicity, assume that h is 0 for negative times correct, if I am willing to satisfy scarify that assumption, if I am willing to let non causality creep into my model then I can give a stationary representation to this seemingly non stationary process in the causal world that is strange, is not it. In the causal world this processes is non stationary what we mean by causal world is in the impulse response is 0 at negative times, but when I look at it backwards, in other words you can rewrite this equation as v k minus 1 equals, how do you write this minus 0.5 v k, assume some e tilde k, it would be minus 0.5 e k, sorry there would be no minus here, 0.5 e k, minus 0.5 e k, but we can observe that minus 0.5 into e k and create a new sequence and think of this as e tilde; e tilde is nothing but minus 0.5 e. If I were to give you this model, the first thing you observe is non causal, but it is stationary. So, an explosive series can be given a stationary representation and that is the point here, there is a connection between causality and stationarity when it comes to autoregressive models and that the example that I have worked out more in a symbolic passion on the screen for you where I instead of 0.5, I have d 1. So, the point here is a non stationary causal or an explosive causal model or causal series can be given a stationary representation, but with non causality.

Is this model any use to me? What do you think? Is it of any use to me in forecasting? It is of no use and I am very, it is just for my fun is just to confuse you further, but having said that there are a set of problems where this kind of model is actually useful in what is known as back casting, have you heard of back casting? We have all, we start off forecasting, but there is something called back casting when would be back casting useful, any idea, any application?

Student: (Refer Time: 07:11).

Sorry.

Student: (Refer Time: 07:12).

When forecasting is not giving me the right thing, turn to back casting, turn your back countering.

Student: (Refer Time: 07:24).

Correct, very good. So, if you want actually extent that it is not exactly that when I have missing observation, for example, right and I have data, historical data, I can use back casting to recover the missing data. So, back it is not that is non causal models are of no interest at all, now you have to change your opinion the non causal models are indeed useful many times. In fact, many a times we use the non causal models in estimation there are 3 different problems, one is that prediction which is forecasting then there is that of filtering, what is the difference between prediction and filtering? Prediction is given all the information up to k, I would like to estimate the signal at k plus 1, k plus 2 and so on, commonsense thing, filtering is I am given all the information up to k, meaning measurements up to k and I will not get an estimate of the truth at k just because I am given measurements, I am not given the truth remember there is always going to be this error and I will give you an a very simple example that you can relate to, but let me discuss the third problem.

The third problem is that of smoothing, what is smoothing? Smoothing is sorry, is the exercise of obtaining an estimate of the truth at k using the information to the left and at means you would use k plus 1 k, k minus 1 or maybe even more essentially you look to the right and to the left and say this is where I am right, sometimes when we get up from sleep, deep sleep, you do not know where you are, it happens to almost everyone. So, your days really if you look to one side, you still not able to gage where you are, but you would look around and say now I know my GPS code position, Abin has figured out what my coordinates are, our more practical example is when we go and watch movies where actually fans volume is louder, typically then the movie volume you are trying to actually get an idea of what this dialog is, at that moment or you have to go out and may be get a popcorn or something to eat and you come back with Bollywood movies it is excellent, actually you should really happy with Bollywood movies because there is such an element of predictability in Bollywood movies that that you have, it is a problem with Hollywood movies.

First of all, with Hollywood movies, you have to really shut off everything you cannot even talk to the person who came with you; you cannot pay attention to your child, nothing because the guy is talking so silently as is there will not enough light, on top of it they would be whispering and you wondering what on earth is going on and one has to watch 10 times and it is unpredictable and so depends if it is James Bond movie, it is predictable, but by enlarged Hollywood movies have a problem, but Bollywood movies; you should be really thankful that it is predictable in the sense that even if you miss a few minutes, you can actually recover that. Based on what you heard before you left the hall and what you heard after you enter the hall that is smoothing. Many a times seriously you think that actually these problems are alien, no they are not. we are using it every day; only difference is we are looking at formularization of these concepts that is all.

Human brain has been doing all of this for ages since time in memorial many a time you see in the movie theater the movie dialog being predicted by the fan like [FL]. This standard acting, know that may be a dialog is changed now. So, model has to be updated that is only difference that dialog probably that predication applies to a class of movies until 70s or probably 80s, but beyond that they where the hero calls out of the villain, the model has changed. So, if you have seen enough movies then your model is updated and then you use a model switching depending on the era in which the movie was made and you predict that is predication based on what you have heard, you predicting what is going to happen.

Filtering is you have heard it, but there are so many fans shouting around you that you are not able to get the clear picture, yet based on what you have heard already and whatever you little you have heard you guy trying to figure out what the exact dialog is.

The person sitting next to you may be your family member is asking you what is it that they said then you immediately reconstruct and say yeah this is what perhaps it may and then you confirm that when as the movie goes on, yes, I was right or maybe you correct it he says no, no that is what actually he said, but it is that correlation that is allowing you to make that; imagine if you are movie was right noise seriously they would there would be no fun at all you, you have to sit really tight. So, you should really thank may be 1 of you can send thank you mail to the film makers saying that please do make movies that have high element of correlation in them, that is you take the ACF, there is very slow d k.

Anyway, coming back to the point, non causal models are useful for back casting subject is if it made, if you establish connections with real life, it is real fun. So, henceforth we will work only with causal models and it assumed. In fact, in many test that is what is also stated, the movement we said causal AR model, it assume it stationary you do not have to, again one does not have to explicitly state that it is stationary unless you know it is stated explicitly, but by default causal autoregressive model implies stationary model.

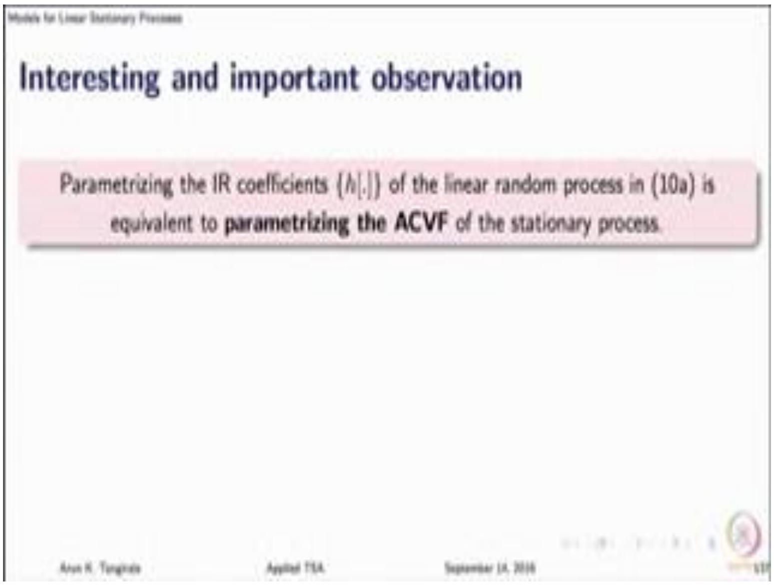
If it is not then it would be a non casual I mean you can build a non casual stationary representation. So, in a sense it is always possible to construct stationary autoregressive representation it is a matter of whether you are willing to tolerate non causality if you are not then you will work with only casual models which is what we are going to work with nice. So that now settles a stationary issue of autoregressive models.

Now we move on quickly to the ACF, we have seen this before we know that the ACF has an exponential d k and we have also seen how to, learned how to derive the ACF for an autoregressive process by writing the Yule Walker equations and we are just going to go over that as a refresher. We know that the auto covariance function of an autoregressive process satisfies same difference equation as the process itself and when you write down those equations you are left to this bunch of equations known as Yule Walker equations. Generally we separate we write an equation separately at lag 0 and where the right and side you have sigma square e and then at all nonnegative lags the right hand side is 0 and it is just a different equation form pretty much looking like the process itself.

Now from the theory of difference equations, as I have even said earlier, you can write a generic solution, I am assuming for just sake of discussion that the roots of the characteristic equation, now you see the ACF has a same characteristic equation, the difference equation for ACF has a same characteristic equation as that of the process. So, roots are the same, the only difference is h; the constants that appear in h when you have sigma, alpha i's and so those alpha i's different now so that is why I have betas now. But otherwise the form of the solution does not changed what does tell us, now it reveals us something very nice earlier we said parameterization of impulse response coefficients according to the power law implies autoregressive form of a certain order. Now we learn another thing by parametrizing the impulse response coefficients in a particular way we are also implying that, what are we implying now? That the ACF also has a certain parameterization I go to the statement and I will come back to the Yule Walker equation just found though bring up that statement on the screen and then will come back talk about the p ACF and so on.

Parametrizing the impulse response coefficients assumes is equivalent to parametrizing the autocorrelation function itself.

(Refer Slide Time: 16:28)



Whether you are doing it explicitly or not that is a consequence; that means, when you look at the ACVF of a series and you are fitting an autoregressive model you are implicitly saying the ACF actually decays according to a certain law and now you can also see this for the MA model as well. In the MA model we assume that the impulse response coefficients die down abruptly after time instant m that there also that assumption translated to the same situation ACF diced down exactly after lag m. So, whatever I did on the impulse response coefficient is more or less the same on ACF as well.

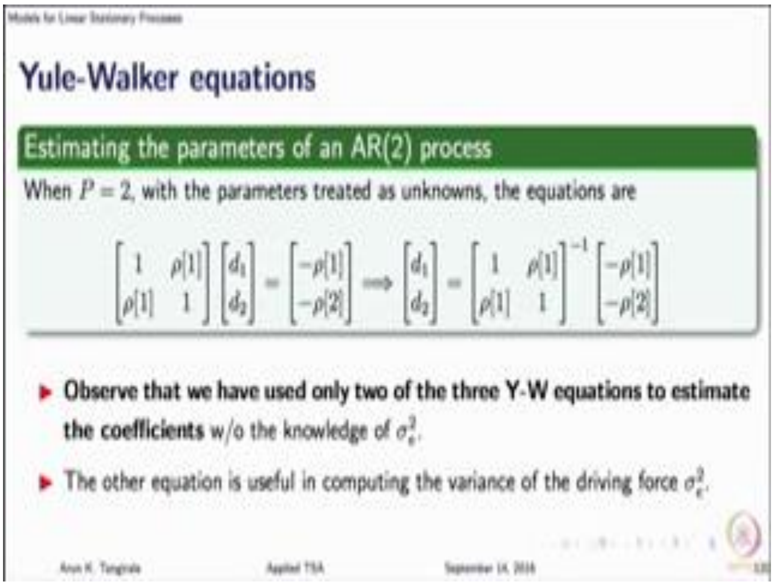
So, whether you assume parameterization of impulse response coefficients or parameterization of ACF or a particular model they are all equivalent how; which do you consider is starting point is your choice, you can start from 1 and then end up with the others and going further, we know that these spectral density is related to the autocorrelation function through the Fourier transform and by virtue of that relation parametrizing the ACF also amounts to parametrizing this spectral density, see parametrizing means what that we are actually expressing the function in terms of those parameters that is what parameterization means.

In effect you are AR models are actually parameterized linear random models for linear random processes and, so or moving average models, but in the moving average model there is no parameter per say all you are assuming that is that the impulse response coefficient go to 0 or the ACF goes to 0 or the ACF go to 0 or that the spectral density has a certain form and so on of course, we have not thought about spectral density, but it is waiting for us in the next room. So, we will board that vegan pretty soon.

So, that is a point you have to remember the moving average whether it is moving average or autoregressive model whatever you do there are 2 other consequences - one is parametrizing the impulse response other is parameterization of the ACF and the third one is of course, parameterization of spectral density. Instead of rhyming it as parameterization you can think of it that your, it this way you are assuming a certain shape and a mathematical equation for that shape either for the IR coefficients or the ACF or the spectral density.

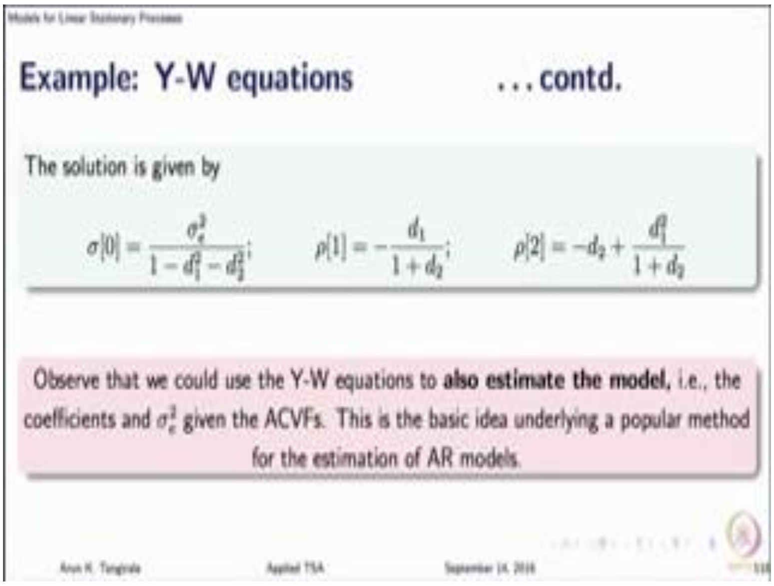
That is something that you should remember, let us get back to the ACF part. So, we know we have seen these Yule Walker equations before.

(Refer Slide Time: 19:10)



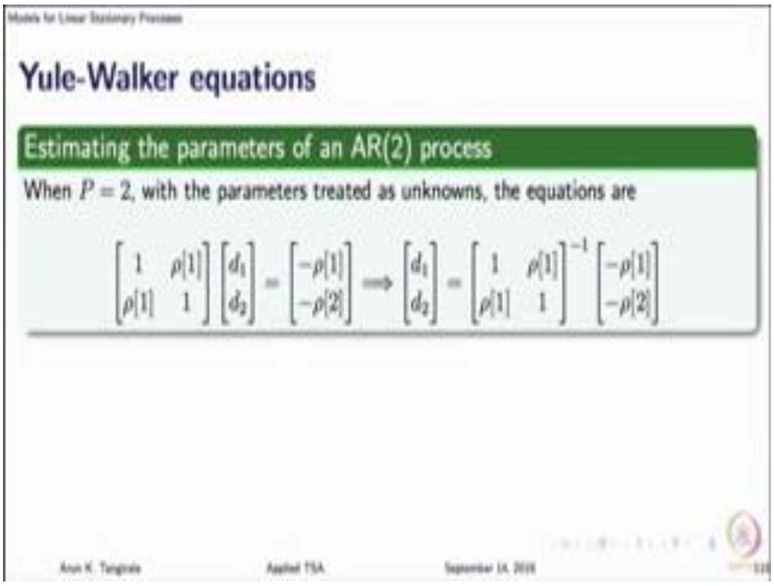
And I am just showing you some sample calculation as to how you can solve this Yule Walker equation for getting you are ACF.

(Refer Slide Time: 19:19)



So, for example, for an AR2 process, if you solve the Yule Walker equations, this would be the expressions for auto covariance at lag 0 1 and 2. Practically the use of Yule Walker equations they are more popular for estimating model parameters and we have talked about this you can use it for 2 purposes either given model I can calculate the theoretical ACF as I have done right now for you.

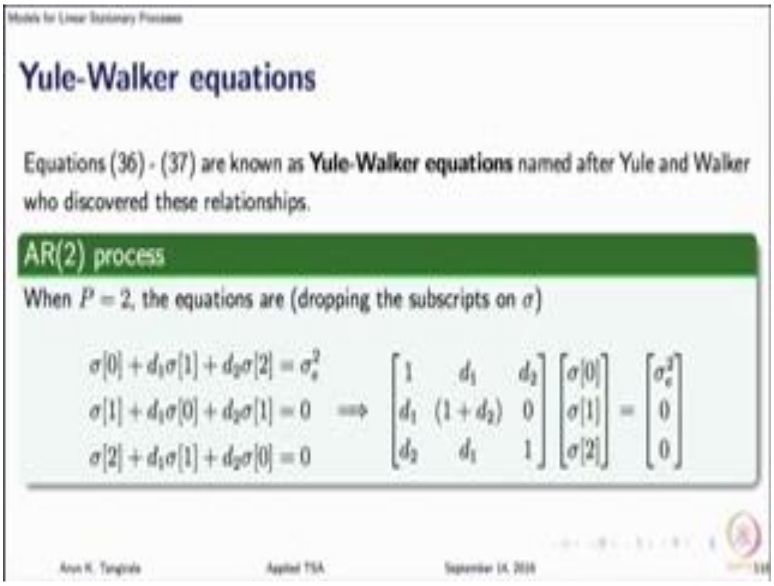
(Refer Slide Time: 19:41)



Or given ACFI can calculate the model parameters and these are the equations that you would use for an AR2 process for example, if I am given the ACFI would use these 2 equations to solve for d 1 and d 2. Now notice that for an ARP process typically we right p plus 1 equations right what we have used here to solve for d 1 and d 2 are only the equations for non negative lags sorry, for non zero lags positive lags let me put it that way.

In other words let me go back to these equations that we are written here sorry.

(Refer Slide Time: 20:21)



Here if you look at the 3 equations, I have used only the bottom 2 equations for getting my d 1 and d 2 because I can decouple the problem of estimating sigma square e and the parameters I would like to decouple right. And then go back and substitute this into the top equation to get an estimate of sigma square e that is the usual procedure. So, the Yule Walker equation gives us nicely a set of linear equations that is the beauty with autoregressive model that is the beauty that you will always you can end up with the bunch of linear equations depending on the optimize that you use. If you use Yule Walker method you will end up with the bunch of linear equations and easily solve for the parameters this was not the case for moving average models.

Why do I keep pointing out these differences between these 2 models? Because in practice one has to make a decision at the time of modeling effect give you data and you have to make a decision on which model to choose you have to be very well versed with the features of all this models in terms of how easy it is to estimate the parameters, what are the ACF signatures of this. Generally if nothing is specified in terms of which model to prefer there are certain criteria and I will talk about it in tomorrow's class when we talk about which models to choose, but its common sense the is of estimation will dictate which model I should choose.

Between AR and MA model which are easier to estimate? Autoregressive model because I have bunch of linear equations right, now the only thing that I have to guarantee is this bunch of linear equations that I am solving will give me estimates that are going to satisfy the stationarity requirement. Whereas with MA model there was no issue, but there was another issue with MA model apart from solving non-linear equations I have to choose the one that is invertible, but I am guaranteed there will always be an invertible solution in provided those conditions are satisfied and so on.

So, there are pros and cons of each of this, but fact is for autoregressive models the equations that you solve are linear and usually this is a procedure that is used and that is why in many packages many software packages including R there is a separate routine for estimating AR models there is a routine for estimating ARIMA models which is ARIMA in R, but there is a separate routine for estimating AR model which is AR and within that AR if you just go back go home and look up the help on AR you will see there are number methods listed YW, l s and mod if, m c o v, modified covariance and burgs methods. So, on YW stands for Yule Walker method it is essentially uses this equation. And default is Yule Walker; it is very easy to estimate.

We will show later on the list square methods also gives us bunch of linear equation pretty much like the Yule Walker equations almost for last sample it does not matter actually for large sample sizes both are going to give you same results. Burg's method is different the big advantage burgs methods estimating AR model is guarantee stationary solution that was way back in mid 70s at Burg's proportion. So, this general set of equations that you see for an ARP process and this is how you estimate your d and in practice you would replace the variance the auto covariance that you see in this matrices and vectors with the respective estimates that is the method of movement approach.

When we meet tomorrow will talk about very briefly the order of AR process determining the order and then also talk about the equivalence between AR and MA and talk about ARMA models as well.