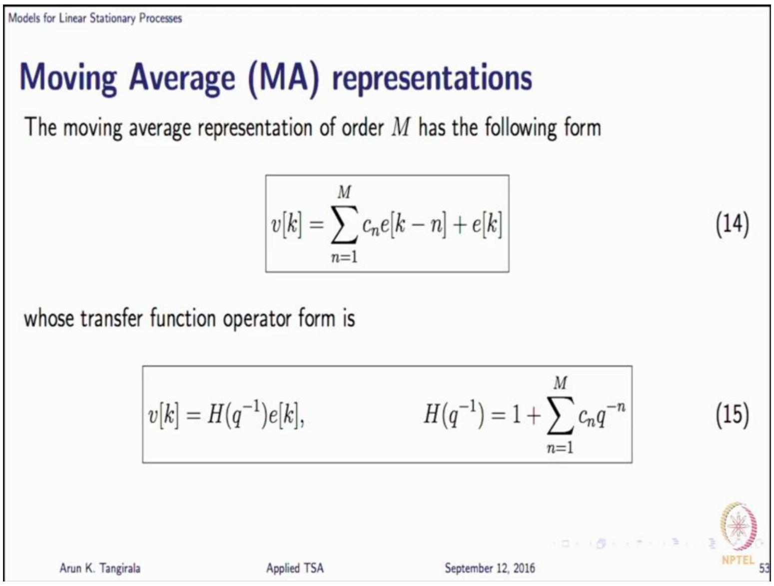
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 44
Lecture 19B - Models for Linear Stationary Processes-8

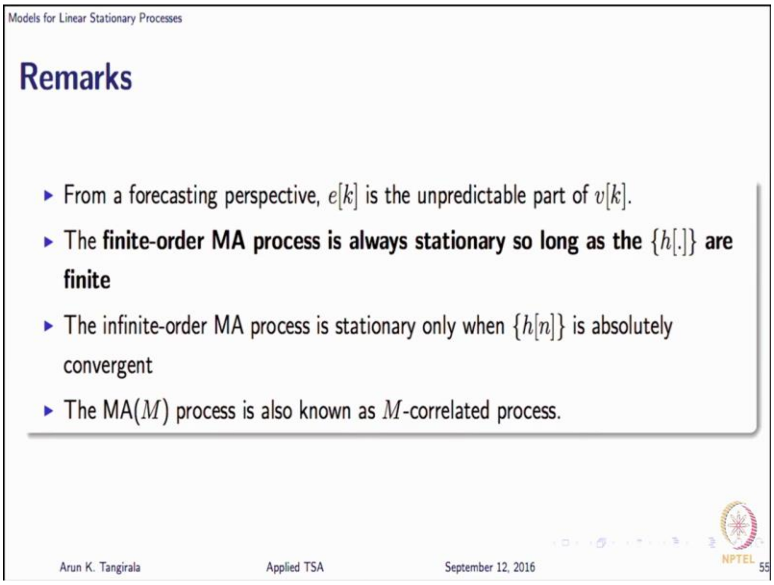
We will skip the second option for now and discuss the moving average model first. So, on the screen you see the general expression for a moving average model of order M.

(Refer Slide Time: 00:22)



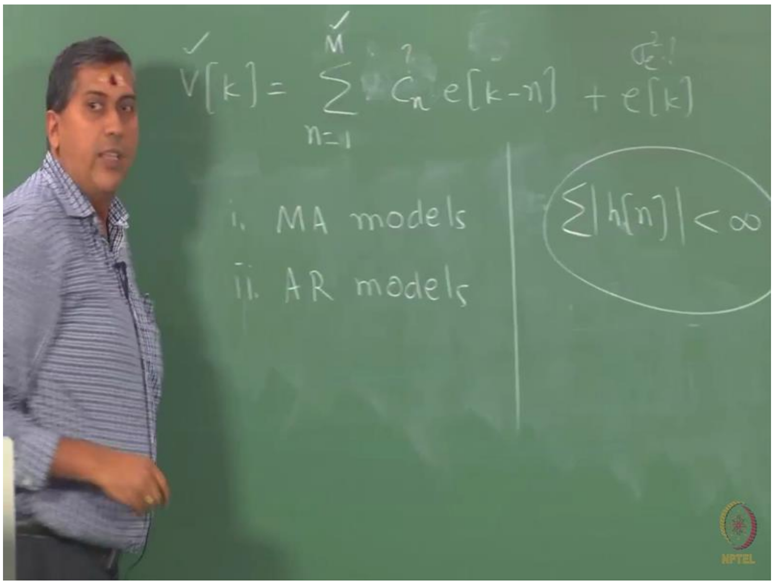
And as you can see that there are n terms now and there m unknowns plus sigma square e, in total you have m plus one unknowns to be estimated from the date and again here you should keep in mind the prospective is that the time series is purely being explained by the shock waves alone not by it is past. At a later stage we will show the equivalence between this representation and the counter path the auto regressive model, where the time series is explained purely in terms of it is own past will show that they are equivalent, but right now let us we can focus on the moving average model and the transfer function operator is fairly straight forward to see.

(Refer Slide Time: 01:21)



Now, from a forecasting prospective once again e k is your unpredictable component that keeps flashing at you and now we talk about the conditions or the impositions that we have to place on the model, so as to guarantee stationarity, but since we have starting anyway from a linear random process and throwing of all those coefficients that we assume to be zero we should not except any major requirements coming into force as for as stationarity is concerned, if you go back to the general linear random process what was a requirement for stationarity?

(Refer Slide Time: 02:06)



We said that this sequence here the impulse response coefficient sequence should be absolutely convergent. Now, that is readily satisfied now we replace this with C m that is readily satisfied, it is a finite number of terms it is satisfied as long as the coefficients are bounded and it is a very very very mild assumption you should keep this in mind. The moment we move from this infinity long the stationary linear random process to now this finite impulse response model, the stationarity requirement are not so binding.

In fact, now you can view the linear random process that you have been seeing as a moving average process of infinite order right, the earlier one that we had written you can write it as m a infinity, that is how it is referred to as in the literature many times, this is m a m and the linear random process that we have been looking at is m a infinite, but when you move from this finite impulse response to infinite impulse response model you have to worry about this restriction where as it is not present here it is, but you know it is a easily satisfied.

So, just to as a (Refer Time: 03:33) MA process of order m is also known as m correlated process, for reasons you know already the ACF of this process dies exactly after lag m. So, the correlations only last up to lag m and therefore, it is called m correlated process. So, we will quickly breeze through the ACF of m a process and the reason we are doing is because for a given process I would like to know whether an MA model is suited.

The series is not going to speak to me and say here is the model, I have to figure out. I have to calculate some measures or matrix and figure out what is the nature of the underlined process and we know already at least from theory that if the ACF of the series dies exactly after some finite lags, a possible model is an m a model of order m. So, good which means we not only have a model with us which is workable. That means, I can estimate finite number of known, finite number of constants and sigma square e, but also a signature that tells me when this model is warranted, that is a most important thing making decisions on a model structure the type of model and so on is what actually it is a way a lot of your time in time series modeling; all the other things there is data cleaning, but once your data is clean most of the time is spent on determining the type of model and a lot of times you do end up with multiple models computing for each other then you have to conduct an election right and take a pole, you do not ask people around that is not what I mean do not do that during exam, but they there are information theoretic measures. For example, that will allow you to do that we should allow you to it allows you to determine which model is suited when you have a bunch of models completing for a given process.

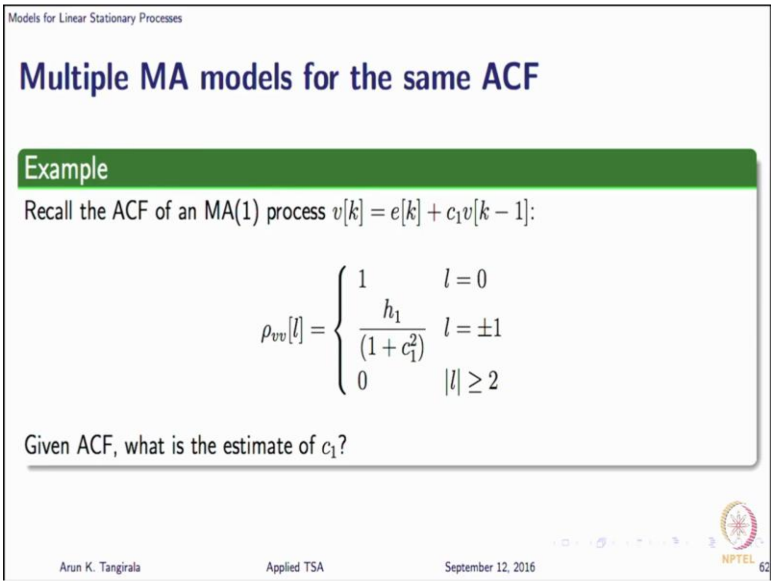
But right now let us say that at least now we have this ACF telling us whether an mam model is suited or not right and just we have seen this already, I am just going through this for the sake of completeness the ACF and MA 1 process we have seen many a times and probably a lot of times thanks to your quiz, so we will skip and we have also proved through the use of autocovariance generating function that the ACF dies of exactly after m lags.

For us right now what is more important days mapping the ACF with the model parameters; we know that the ACF tells us not only the fact whether MA model is suited or not, but also what is the order at least theoretically. Now having determined the order m, now before I begin I do not know what is m? Whether this model is suited or not; so, I look at the ACF and if the ACF supports my postulate that there is m a m model or it gives me substantial evidence then I postulate this model and it ACF is lot at that stage I do not m, I do not this constants and I do not known sigma square e right this is also unknown, but once I look at ACF I also figure out what is m.

So, now the job in hand for me is to estimate this parameters and sigma square e therefore, I need to figure out, whether there is first of all a mapping between ACF and the model parameters and if that mapping is unique? Here again I have to raise that uniqueness issue because may be when I sit down to estimate this parameters, I may run into this non uniqueness issue, there may be multiple solutions we have fixed only one layer of non uniqueness there can be several layers of non uniqueness struggling in. So, there is a second layer of non uniqueness that we may run into. So, we have discussed this in a somewhat quick manner in the last class, but let us spend a bit more time on it and also formalize those results and the invertibility condition.

So, if we were to go back again to this MA 1 model, we have already seen that there are two solutions.

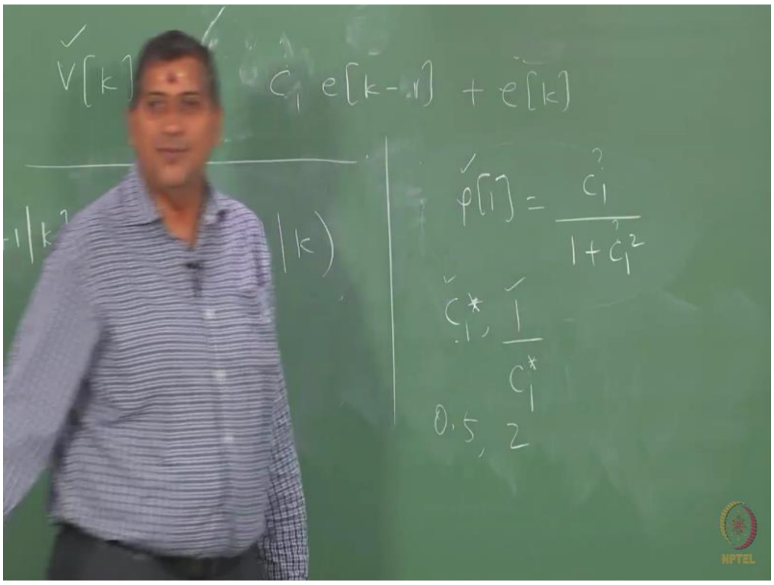
(Refer Slide Time: 08:16)



Then I want to estimate the model. So, I want to know what is C 1 given rho; we know that there are two solutions to this, it is not that always I will use this equation to estimate C 1, do not get this into your mind, but this is one of the ways to do it. Now what about other methods do they make use of this relation between ACF and the model parameters? Not explicitly, but whatever you do whatever you try to do ultimately you will run into multiple solutions for example, later on when we learn estimation theory and we at a point at some point we will learn how to estimate moving average models formally. So, this is also formal way of estimating the moving average model, but using least square approach for example, this is not a least square approach; this is called of method of moments where I am forcing the moments that I am estimating to satisfy the theoretical equation that is the idea and method of moments.

But when we learn least square, there also we shall learn that we would run into a non-linear least square problem not our regular linear least square, which as a unique solutions and non-linear least square problem has multiple solution. So, is problem of multiple solutions will continue to (Refer Time: 09:39) us whenever we deal with moving average models. The nice thing about moving average models is stationarity is not an issue; that means, I do not have to worry about placing any restrictions on this coefficients, but somewhere I pay price and here is where one of the places where I pay a prices where I start feeling the heat.

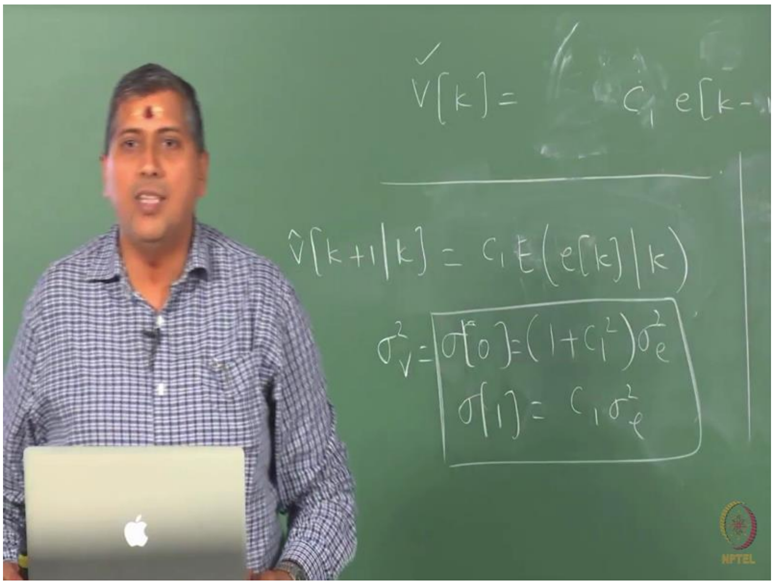
(Refer Slide Time: 10:04)



So, now I have two solutions as we discussed in the last class, if C 1 star is a solution, then 1 over C 1 star is also solution right, two solutions are present both are candidates and I have to decide now which one to pick, and as we discussed in the last class the criterion that we use for selecting this model here is forecasting because ultimately I am going to use this model for forecasting and I should select a model that will give me stable forecast, but it is hard to actually think of it is straight away, you cannot actually look at this and say you know I know there must be an issue with one of the things I have one may give me stable and other one may not give me an stable forecast, it is not so straight forward, it is not as intuitive as I am making it sound.

One requires some formal thinking may be it could have happened that, someone try to use the other solution which is not so called invertible and god beaten up and then come back to the desk and figure out you know there is a possibility that one of this models can give me unstable forecast.

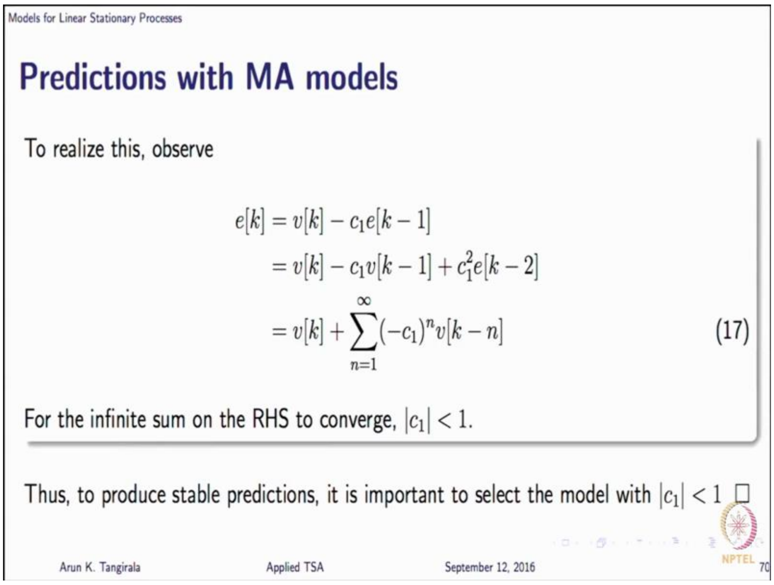
(Refer Slide Time: 11:16)



But you know may be there something else that is also happen in any case, now we can ask what would happen if I were to use this model for forecasting? So, if I am given information up to k and I want to make a prediction at k plus 1 then what would I do I would actually go back to my model because I chosen to work with MA 1 model, the forecast is C 1 e k strictly speaking this should be expectation of e k given information up to k, which is not 0, if you have any questions on that you should ask the unconditional expectation of e k is 0, but the conditional expectation of e k given all information up to k it is not a prediction we call that as an estimate or a filter, this is not zero because v k and in fact v k minus one both will contain information about e k.

So, I can leverage on that and make a better estimate get a better estimate of e k as compare to the unconditional expectation. So, the question is now how do we recover this? And that is what we discussed in the last class, how do I recover e k from the given information and whether it makes a difference? Where If I use C 1 start or 1 over C 1 star and as we have seen in the last class we had one prospective, but let us go through 2-3 different prospective on recovering e k from v k, this is a simple way what you see on the screen is a very simple manipulation of the difference the equation that we wrote. Your model is v k equals C 1 e k minus 1, plus e k this is your model and what we want is an estimate of e k. So, we rewrite this and just do some algebraic (Refer Time: 13:24) there.

(Refer Slide Time: 13:26)



And straight away see that e k is v k plus sum infinitely long summation. So, what this tells you in the first place is to recover the shock wave that occurred at kth instant exactly what you need what does it tell you?

Student: (Refer Time: 13:46).

Yes you need the entire history you need the infinite past. In fact, it should tell you this is actually a very classical problem that you see even physics with big bang theory, what happen at the time origin? What is it that generated what we see today to be able to actually come up with what really triggered the universe; you need that entire history do we have that we do not. So, it is people will keep making a guesses some thousand PhDs will be generated, then somebody else will come a and say no that is not possible I have a better evidence he keep guessing [FL] no problem, at least it is keeping people busy right. So, it will continue to keep people busy because you do not have any infinite history.

So, this is only a theoretical expression which is usefully in practice I do not have infinite history right. So, we will talk about how to deal with the lack of infinite history and so on at a later time when we specifically focus on forecasting; at this moment what is clear to me is that if I want even if I am given the infinite history to generate a stable forecast right, what do you mean by stable forecast is the recovery equation should not result in an unbounded sequence because e k is stationary, that is a minimum requirement and for that minimum requirement to be satisfied, the C 1 has to be less than 1 in magnitude right. The straight away tells me which I should choose if C 1 star is a solution and that happens to be less than 1 in magnitude then I discard one over C 1 star.

But what is puzzling to us is even though there is a series that must have come out of one over C 1 star, I still have to work with C 1 star in a sense let us fix some values here. So, that we do not talk symbolically, suppose the solutions are 0.5 and 2, suppose I have used over the process is actually running with two the value C 1 being value of 2, I still cannot work with that value. It is a bit strange because I have not working with the truth in other words I will generate a series out of an MA 1 process using C 1 value of 2 and I will give you the series you cannot even though you find the 2 is a value a along with 0.5 you cannot work with 2 you have to work with 0.5; why is that, why are we running into this kind of a very enigmatic situation although you have manage to figure out, what I have used in my simulation you are still force to use another value. In other words you will not be able to figure out whether I use 0.5 or 2 that is another issue.

But even if you have figured out that is in our set the true value has come up, you are being force to use another value for generating stable forecast what is a reason here, what is the prime reason I am being force to do this?

Student: (Refer Time: 17:23).

Sorry.

Student: (Refer Time: 17:31).

You are already meditating man at an angle there is a tomorrow people invited you are already meditating practicing me I do not know what is that asana called (Refer Time: 17:47) asana that is do not worry what is the reason?

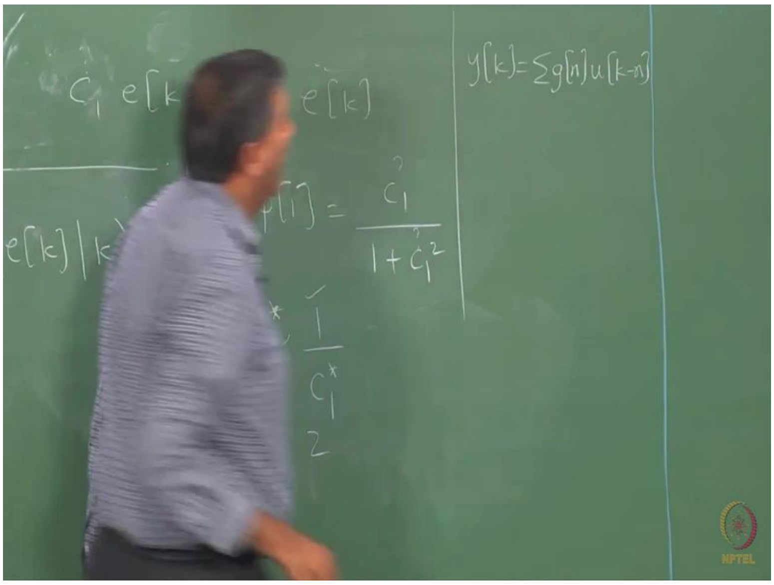
Student: Absolute convergence of.

Absolute convergence of what?

Student: (Refer Time: 18:02).

Let me actually would you let me ask you a different, but a related question would you run into the situation.

(Refer Slide Time: 18:14)



If you are working with deterministic models, I will not cross Lakshman Rekha the font size is adjusted accordingly; would you run into this situation with the in the deterministic world? Yes or no.

Student: Yes.

Yes.

Student: No.

No you cannot go by my reaction it will not Lok Sabha. In Lok Sabha only people go reaction (Refer Time: 18:49) the numbers reaction need not by truth by facts anybody from that hall would you run into this kind of a enigmatic situation, strain situation that you see here in the deterministic world? Who would like to answer at this wonderful hour of the day?

Student: No sir.

I was hoping he would say yes, yes or no with the in the deterministic world would you run into the situation.

Student: No.

Why?

Student: (Refer Time: 19:26).

Why?

Student: because we have (Refer Time: 19:29).

So, the cause is known; here after estimating this model I have this additional headache of going and recovering the cause also to make a forecast, whether I am doing it explicitly or implicitly you will see that this need of explicitly or this kind of an equation is not required in auto regressive models, because the auto regressive model is very nice it directly models the series as a function of it is own past which is observed, here the moving average model is modeling the series as a function of an unobserved quantity sorry unobserved variable right. So, there is this additional responsibility of recovering that unobserved cost if I have to complete my forecast.

Even in auto regressive model somehow if not that something else will (Refer Time: 20:28), life is not actually you know so nice where ever you go, it is just that you know here you deal with some situation in their when you go to the west you say those problems are not there, it is so nice and so on; what a period of time you figure out there are problems their also. So, it is a matter of what problems you are comfortable with, that is all in life is all about where you settle predominantly depends on what problems you are comfortable dealing with right here if you do not mind you know running for the bus or something like that and you love doing that or you know love facing all the kinds of problems that have you actually love bargaining and so on it is great.

When you go to the west there is no bargaining it is boring the really boring, you cannot really haggle around and so on. So, there other issues so beautiful, just you know minus 20 you cannot actually walk around with your slippers of different colors and half broken and so on, I also worn those do not worry you cannot do that right and no juggards and so on, but there are other problems likewise you see in modeling also with moving average models, you may have some issues and you may not have some issues well with are models, you will have some other issues and so on.

So, the main problem here the main reason as to why we run into the situation, where I am force to work always with that model which gives me stable forecast regardless of what has been used for generating the series is because a causes unknown. If only I knew e k then automatically things are fixed you see this here, but fortunately for us we are using an equation that does not have sigma square e, we are very happy that we got our sigma square e, but the problem makes it is way through in a different form, if you look at actually the estimation of moving average model of order 1 there are two equations right, one equation is sigma at lag 0 which is sigma square v is 1 plus C 1 square times sigma square e and the other one is at lag 1, which is C 1 time sigma square e. So, perfect I have actually two equation and two unknowns.

But in going from here to here I have gotten read of sigma square e and we are very happy about it I can estimate C 1 without knowing sigma square e, which will also be the case in auto regressive models I can actually estimate the parameters without knowing sigma square e then post parameter estimation I can estimate sigma square e. But even though I have when I manage to get rid of sigma square e, it still makes it way that is the fact that e k is unknown; imagine if e k was known right if e k was known what would happen?

Student: (Refer Time: 23:29).

Then I have to choose a C 1 that exactly satisfies the other one right, there are actually redundancy set of equations and but that redundancy will help me; correct? If I choose the work with the first one what would happen? I will get 2, but it has to satisfies the second one. So, inevitably then I will have only one value of C 1 and I will leave happily with that, it is fact that sigma square e not known which is actually causing me to I am mean forcing me to run into this situation. So, some thinking will really help you in spectral factorization we will see the same thing appearing in a different way, whatever we have spoken today that is you know why we had to fix h naught equals one and why we are running into this inverses as I mean in fact this kind of called as inverse is actually C 1 star and 1 over C 1 star and you will run into the situation with moving average modes of order two, order three does not matter while if you have one solution the other solution is an inverse. In spectral factorization when we learn the result we will see that whatever we are doing in the time domain here to estimate the parameters, actually amounts to factorizing the spectral density.

Spectral density at any frequency is a number, if you are looking at a univariate time series, gamma of omega is just a number, it amounts to factorizing that number at each omega into a product of factors although you may not see that here at all when we go through frequency domain analysis and all the wonderful things that come within then we realize that the spectral density actually is being factorized behind the scenes as we are working with ACF CR and we does not you know it is a very the simple thing to know that factorizing a number into product of two numbers is not unique. So, that non uniqueness is what is also being reflected here and vice versa. So, that is the point.