Adaptive Signal Processing
Prof. M. Chakraborty
Department of Electrical and Electronic Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 26
Gradient Adaptive Lattice

I just start at the point, where we left yesterday, because I was doing things in a hurry. So, let me redo that you remember yesterday, we discussed that any random process can be recomposed in to two components. One is purely deterministic would be that is purely predictable and another is purely non-deterministic component, deterministic is one, where any sample can be predicted, exactly from it is all past samples.

Typical examples were sinusoidal signal with a random phase or random amplitude or both, but the purely non-deterministic is one, where any samples cannot be predicted exactly from it is entire past. So, the prediction will have at least 1 0 variance, that is finite positive variances and then I said that in this course, we will be discussing only the non-deterministic random processes.

For non-deterministic random process spectra is continuous, for deterministic ones, I have said that, they will consist of lines, each line corresponds to one sinusoidal component. So, deterministic compo part actually can be shown summation of several sinusoidal components like that, so we will be considering this purely non-deterministic component that was old decomposition.

And that further said that in this case, the purely non-deterministic process x n can be written as though it is generated by driving causal and stable LTI system with impulse for tension by a white sequence say u n. And then what you said is this, we use this for parametric spectral estimation, we said such as the output power spectral density will then be what, transform function, mod square times, input variance, where variance is just a constant.

Because, it is white, so that means for this kind of processes, I mean if I have to do power spectral estimation, I have to just identify that LTI system in posses for h n. But, again that has got infinite numbers of coefficients h 0, h 1, h 2 dot, dot, dot up to h infinity, which we cannot do. So, what you do, you can go for rational approximation, that is this n model, we try to approximate by a rational model with 0 initial conditions.

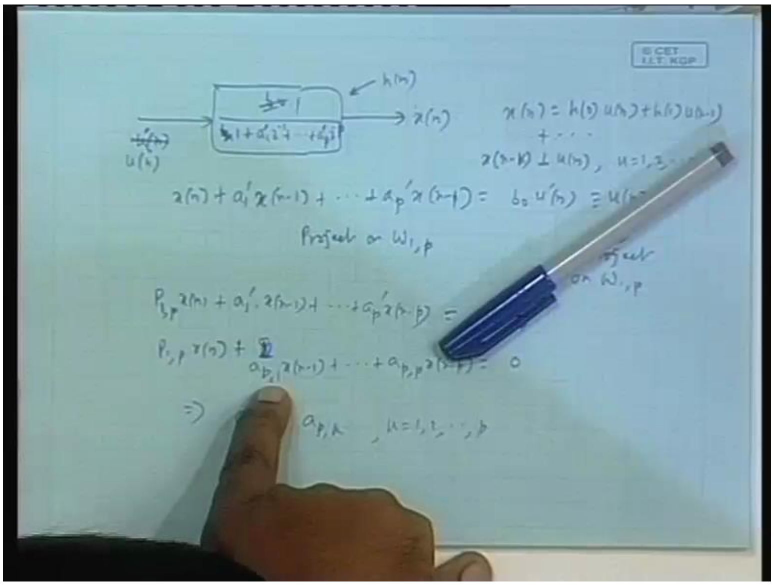
Of course, as I told you, unless you get 0 initial conditions no system remains LTI and that general is where approximate in z domain in z by a rational from like b 0 plus b 1 0 plus dot, dot plus b q with inverse q divided by 1 plus a 1 0 inverse plus dot, dot, dot plus a p 0 plus p, this is all there in notes, I am not writing it here. In time domain, it makes a constant coefficient different equation, this model is a pole 0 model, it has both poles and zeros, it is called also ARMA mode and ARMA means auto regressive moving average.

The left hand side, which where you have got x n, a 1, x n minus 1, a 2, x n minus 2, that is this is a autoregressive part, auto means self regression means this you are taking the past x n minus 1 x n minus 2 all this things means autoregressive part. Right hand side is called the moving average part, right hand side constitute of the non-linear estimation problem. Because, the coefficients b 0, b 1, b 2 they are unknown also, the input quite simple my samples u n minus 1, they are also unknown, so product of two unknowns which is why, it is a non-linear product and this is the ARMA model.

There is a another model called MA model, rational model, where you forget the poles, only zeros, that kind of transfer function than, it is called MA moving average model. Where, the output x n is nothing but b 0, u n plus b 1, un minus 1 plus dot, dot, dot plus b q, un minus q, this is just a MA model moving average model, here two unknowns are multiplied b's and u's, un's the samples are unknown.

So, what are the b's, so it is again a non-linear are estimation problem, we now only avoid them and because the algorithm is not guarantee convergence or estimation will come out. They are not very, what to say, they are not best estimation errors are large and all that. The one, which has really been a great success actually is the autoregressive modeling AR model, in the AR model, you have got all-pole-model a numerator coefficient can be b 0 like this.

(Refer Slide Time: 05:06)



This one Instead of a 1 to a p let me call them a 1 prime to a p prime for some reason. this is driven by this, I did yesterday also some process say un, probably a prime n is a white noise, 0 mean white noise, white process means what, every two samples are uncorrelated and therefore, they are orthogonal, this is your x n, this is the model.

In this case in time domain the equation is of course, this model means you have to give 0 initial conditions. In time domain this means x n, because x n in to 1 x n a 1 prime x n minus 1 plus dot, dot, dot a p prime is b 0 in to u prime n, but b 0 in to u prime n, you call it un. Because, u prime n is white b 0 in to u prime n is also a white, only the variance changes variance in any case even here is unknown, I do not know the variance of u prime n.

So, observe b 0 in to that and call it a white process, we will find out the variance of un, that variance times transfer function square, that is z replaced by e to the power j omega mod of that square in to variance, that will be output power spectral density right, this called the AR model. Autoregressive model is a linear estimation problem of because on this side you have got the known in data for x n, only unknown are the parameters, so unknown in to known, it is a linear problem.

Then, I said that suppose such a model is really true, this is true model, that x n is indeed generated by a model like this or model order selection also is correct. Not only the model, it is told that x n is a AR process, such process I mean any random process generated by a AR model will be called a AR process. So, is it is not only told, that x n is a AR process also the order also is given, that I choose order up to p, suppose that order is give you, that order is also correct; that means, it is a true model.

If this be the case, then I will show as I did yesterday in a hurry, that this AR parameter coefficients, this filter parameter coefficients. They are nothing but what you get, if you do linear prediction of x n from it is past p samples, that linear combiner coefficient combining x n minus 1, x n minus 2 up to x n minus p, those coefficients are nothing but this coefficients. To see that, suppose I take the left hand side and project it, orthogonally on the space, project means orthogonally project, project on w 1 comma p.

So, this also will project on w on w 1 comma p and assuming this model to be causal and stable that is an assumption, I am making that is mode model is causal and stable. So, it is impulse response, if it is h n; that means, x n is nothing but convolution between h n and u prime n or rather h n and un, you can say in this model or h n, you can drop this b 0, you can make this now equal to 1 and you can now call it un and h n, so x n is h n convolved with u n.

Since, it is causal and stable you have to start at h 0 in to u n plus h 1, u n minus 1 plus dot, dot, dot. But, U N is an orthogonal sequence, u n minus 1, u n minus 2, they are uncorrelated white. So that means if you say x n minus 1, my claim this is orthogonal to u n, why because x n minus 1 dependent on what, u n minus 1, u n minus 2, u n minus 3 all that.

So, that is orthogonal to u n because u n is orthogonal to all the past samples u n minus 1, u n minus 2 u n minus 3, because u n is a white process. Similarly, In fact, I should say now, x n minus k is orthogonal to u n, where k can be 1, 2, dot, dot, dot, so if you project u n on the space spent by on the w 1 comma p, what is w 1 comma p space spent by x n minus 1 up to x n minus p. Now, u n is orthogonal to all of them that mean u n is an orthogonal to this space, so; that means, projection error is 0 and u n itself is the error projection is 0 and u n, itself is the projection error.

So, projection here is 0, now it is perpendicular to do the space mathematically also, you can see. If you are consider a external vector say x in a vector space v is a vector space, w is a sub space, small p is a vector lying outside w, but inside the v can be orthogonally projected on w. So, v is nothing but a summation of projection and the error, if v itself is orthogonal and then there is Pythagoras theorem easily seen, norm square of projection plus norm square of the error equal to norm square of b.

If b itself is orthogonal on w, then that projection will be 0, we can easily apply that Pythagorean logic here. So, and projection is 0, that mean this side if I project I get 0, 0 mean 0 random variable mine you and this side I get, p 1 comma p x n, because I am projecting it on w 1 p plus projection is a linear operator. So, projection of the summation is nothing individual projections summed.

So, projection of this guy say a 1 prime, because of linearity, it goes outside, it is a constant times projection of x n minus 1 on to w 1 comma p, which is x n minus 1 itself that will give the minimum possible error between the 2, which is the 0 here. So, you will get this and same for others, but we know that p 1, p x n, it is what plus that is if you write this way a p 1, all notation projection is nothing but, you remember I said that in summation and I put a minus sign in the coefficients.

Linear combination of x n minus 1, x n minus p, there was a negative sign in the coefficients I used, so you bring them bring that to the left hand side then it will become plus. Now, if you just take this 2; obviously, if I say by comparing a 1 prime is this, a p prime is this, but mathematically, if you subtract from this 1, from this also get subtracted.

And a 1 prime minus this, a p 1 into x n minus 1 plus a 2 prime minus a p 2 x n minus 2 plus dot, dot, dot, a p prime minus a p p in to x n minus p equal to 0, those are linearly independent samples so; obviously, this coefficient is 0. A prime 1 equal to a p 1, a 2 prime equal to so that way you understand that a k prime becomes equal to a p k, k can be 1, 2 up to p. So, this AR models coefficients are nothing but the linear predictor coefficients, this coefficients are called linear predictor coefficients.

Now, I said that model must be causal and stable and that that factor, I have used here and that that factor I have used here under that only using that only I got this relation. So, this is nothing but linear predictor coefficient, but how can I justify that this linear is a causal and stable.

That I proved yesterday, that if you do linear prediction, then the linear prediction predictor polynomial is always causal and stable. For any order, I proved that yesterday, do you remember using P.P. Vaidhyanathan's proof, whichever be the order p th, order p plus minus 1th order, p plus 2 th order. Any of this polynomials, will give rise are minimum phase polynomial roots lying within the unit circle.

So, this part is a linear predictor polynomial, but this has now gone to the denominator, so; that means, root here with a poles they are lying within the unit circle that is guaranteed from linear prediction theory, which I proved yesterday. So; that means, this model is such, it has all poles lying within unit circle in z plane, which means it is causal and stable

So, to identify or to do AR modeling or to identify equivalency to identify this parameters, what we have to do is simply take x n and generate these coefficients. If there is a pth order model, generate this coefficients a p 1, a p 2, a p p those will be these model coefficients. How to generate the coefficient, we have seen this approach already by using lattice, that lattice algorithm in lattice algorithm, we are updating sigma p f square sigma p b square delta p's and all those and there again a p I, v p j they also are updated.

We started with the giving statistics of x n, using the statistics, you may not be given x, you may be giving a data record of x n. So, by sample average you find out this autocorrelation values estimates give that 2 a lattice algorithm, lattice means not the lattice filter, because I am not interested in the online operation of the lattice filter, I am only interested in lattice algorithm means recursive computation of sigma p square delta p's and those a p and b p coefficients, go up to the order p, because this is the model order p.

And you get, those coefficients those will be nothing but the corresponding AR model coefficients and that will be guarantee corral linear stability, because that amounts to a filter with roots within unit circle, that is what we proved last time. By now, all these things, we are doing now; this is my approach that is do linear prediction use the orthogonal projection of treatment.

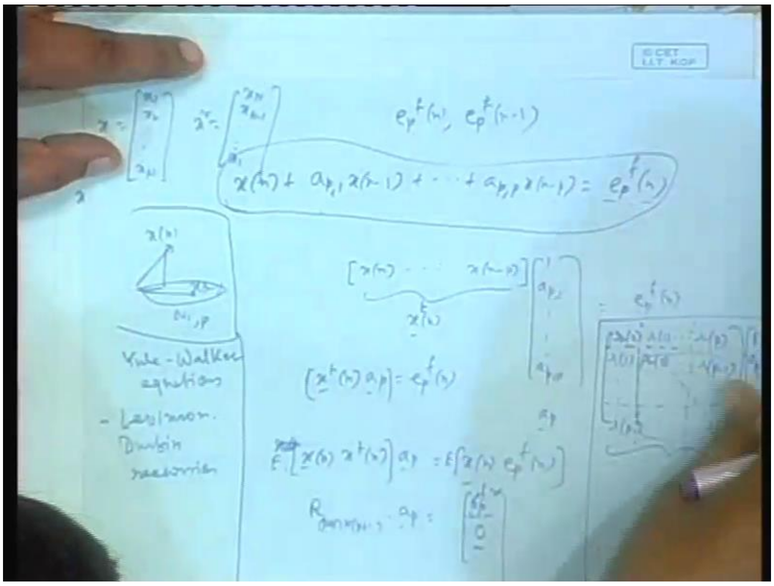
Using vector space concepts, get all lattice forward prediction error, backward prediction everything and link AR modeling with this by this theory. And therefore, we obtain the AR parameters from just those coefficients, but in book also there is another approach where, but I mean the approach is for those who are not familiar with vector space theory.

They are they find out this coefficients by you known developing an equation called Yule Walker equation solved by some of you know Levinson Durbin algorithm. I will just narrate that because otherwise, you may get confused in future since some algorithm you may wonder as to what is this. This is nothing but an equivalent way of obtaining those sigma p square deltas P’s and those p and b p parameters.

Now, if you do that if you think there is no AR model, even when though there is no AR model I can have an equation like this where u n is simply e p f n and these are the predictor coefficients, there is the more general way of writing. In the case of AR model, what is the difference in the case of AR model, can you tell me, this is always true, suppose I am doing pth order linear prediction. Then x n and if these are the predictor coefficients, then this is always equal to e p f n, e p f n is orthogonal to x n minus 1 to x n minus p.

But, now suppose it is not only linear prediction, I am also I have also the AR model, so for the case of AR model also have the same equation, what is the differences between just linear prediction an AR modeling. In the case of linear prediction, u n is e p f n which is orthogonal only to x n minus 1 to x n minus p not to x n minus p minus 1 not to x n minus p minus 2. But, when it is AR model, this thing which is not only e p f n and additionally, it also the input, that is orthogonal to all the past of x, because of this relation.

(Refer Slide Time: 18:34)



This is the thing, this equation is always true, some noise process, this is always true, this is orthogonal to x n minus 1 to x n minus p, but if I only answer the linear prediction problem and there is no guarantee that x n itself is a AR process, that it also satisfy the AR model. Then, I cannot say e p f n is orthogonal to x n minus p minus 1, x n minus p minus 2, x n minus p minus 3 that I cannot say. But, if additionally it state that x n also is an AR process, that we have seen p th order AR model, p th order AR process.

Then, we have seen this is nothing,, but this turns out to be what these are the predictor coefficients, also AR model coefficients and therefore, this is nothing but this prediction error is nothing but then input white process. But, that is orthogonal to because of that causality of that system and all that is orthogonal to not only to x n minus 1 to x n minus p, orthogonal to all the past.

So, linear prediction problem this is actually to start with it is a linear prediction problem e p f n is always orthogonal to x n minus 1 to x n minus p. Additionally, it is also real model, if x n is an p th order, AR process e p f n is orthogonal not only up to x n minus p, but to in all past samples of x. Because, in that case this is transferred to be the AR model input white process, you can see one thing, one more difference.

Suppose, you are solving one linear prediction problem, there is no guarantee that e p f n and e p say f n minus 1 they are orthogonal, because what is e p f n is x n, you are projecting on what w 1 comma p and this is error. And, what is e p f n minus 1, somewhere in this space is your x n minus 1, that is projected on another sub space, that consist of x n minus 2 up to x n minus p plus 1 more element. So, there is no guarantee that these two projection errors will be mutually orthogonal.

So, it is not a necessary a white sequence, that e p f n minus 1, e p f n minus 2, they are white, not necessary in a general linear prediction case they are not, but in additionally, if an AR model, it is given that x n also is a p th order. AR model then, we have this additional thing about this e p f n and these coefficients, coefficients are they AR model coefficient, e p f n is not only linear prediction error, it is also the AR model input.

And therefore, it is white sequence it is orthogonal to not only x n minus 1 to x n minus p as is valid in the linear prediction case, but is orthogonal to all past x. The additional things, that come up in the AR model case are what all is valid here, this is always true for linear prediction, this is orthogonal to space span by x n minus 1 to x n minus p, e p f n minus is orthogonal to space span by x n minus 2 up to x n minus p minus 1, there is no guaranty that these two will be initially orthogonal, because the two spaces are different in general.

But, I said that, I have just proved, that if x n is given to be a p th order AR process, that is there is indeed a p the order AR model driven by a white input, then we have proved that AR parameters are this predicted parameters. This part is the predictor negative of this is the predictor, original minus the predictor is the error and this end is nothing but input white process.

So, this e p f n will be a white sequence, further this will be orthogonal not only to x n minus 1 to x n minus p as this is valid only in linear prediction case. It will be orthogonal to even past value, further past values of x n, x n minus p minus 1, x n minus p minus 2 all up to x minus infinity. So, even in the general case, so you do not consider AR model, now we are just considering the linear prediction problem in that.

Also, we are solving linear prediction problem, we did not assume an AR model, here also I am not assuming it is an AR model, I only linked it up with AR modeling. So, you understand, what is the relation between AR model and this, where in indeed AR model is valid by solving linear prediction problem, you can get the AR model parameters and e p f n. This prediction error, then satisfy some property, it is a white sequence in that case, all those things come up.

But, now I consider this linear predictor problem, because if I solve this, if I can get this coefficients either by lattice predictor or by a new method, which I said just now Yule Walker Equation Levinson Durbin equation. These coefficients, if I get it is enough because, if there is an additional AR model assumption made for condition giving this coefficients itself, this coefficients themselves will be by AR model coefficients, that I have proved.

So, it is better we start, we just look for linear prediction problem, only we do not assume any AR model thing, we find out a technique to obtain this coefficients. We already know how to obtain the coefficients by our lattice method, the other method which is given in books and especially for those have who are not interested in vector space that is how we will just consider again. We will not take it to the end, I will only start give the basic equation basic information that is all.

Student: ((Refer Time: 24:21))

x n minus 2 up to x n minus p minus 1, so there is new sub space and we can say so, yes, yes and there is no guaranteed these two headers are mutually orthogonal and such in general there is no guaranteed. Find out, what would be the orthogonal just, what would be the property of x and so that, those would be orthogonal just for your fun or practice just take it up back home? Point this orthogonal, praise this sequence also in orthogonal just do it, I can see there is geometric inside and that is not difficulty actually.

Anyway, so that alternate method, I just start upon here, that is called Yule Walker, Yule Walker also pronounced as Yule-Walker equations. Levinson Durbin recursion, we have done this Levinson Durbin already, without knowing that we are solving a set of equations, you remember I obtained a p plus 1 comma I all those coefficient from a p comma I and b p comma j that is Levinson Durbin recursion.

Actually, but here people in text books in on especially they approach this entire thing from just some algebra point of view. But, see this equation now, this equation I have on my purpose is to find out this coefficients, now e p f n is orthogonal to x n minus 1 to x n minus p.

So, suppose I write like this and for the time being in order to make life simple let us assume everything real no complex data, then again everywhere I have to be possessed and all just complexes. So, suppose I write like this I am doing this way, because I want to save time, this vector I will call x n transpose and this is I will call a p vector this is equal to e p f n.

Suppose, this equation is x transpose n, a p is e p f n, I multiply this left hand side by x a vector, how to name that vector, I multiply this by say, I am doing many things in one step. Because, I really want to be precise here multiply this by this a p. Now, I take an expectation, so obviously, you will work on this a p is not random, data is random, a p is not random, because of stationarity thing, that is our assumption.

Remember, why a p is not random or a p is constant and all that because of stationarity assumption that I made, because we never forget that, that comes for fundamental optimum filtering thing. This E working on this, this will be R and again R independent of n, because of stationarity thing, R what size p plus 1 in to p plus 1, after all this vector has got how many terms p plus 1 terms, So, our p plus 1 in to p plus 1 in to a p.

And these vector, here you see is a cross correlation thing, x n has got all those components. Now, if you leave aside the top most component in x n, take the second and downward, all the correlation terms will be 0, because e p f n is orthogonal to x n minus

1 to x n minus p. So, you will have things like this a 0 vector, I am not writing the size, you can make it understand, what the size will be and first term will be what x n in to e p f n and than expected value.

That is inner product between x n and e p f n and that I can write as inner e p f n and e p f n, you remember I can shift that projection operator and all that I am not doing it please. So, this is sigma p f square say f square I do not know here, since I am not doing that is whether I do not know whether sigma p f is a sigma p b are same, so we have written this is the form this is called Yule walker equations.

This equation mind you this equation is a bit I mean it is not in the usual form because this is a matrix type a p is an unknown quantity first, I do not know input, I do not know this prediction error, I this quantity is to known it. Remember in lattice also, we did not do this, we are updating it, we do not know this so, this is unknown and then zeros and a p.

So, this is not a usual ax is equal to b kind of equation it is just but, actually it is because, what is happening is the form of the equation, if you see how many are you familiar with Levinson Durbin. Here, nobody in inequality digital communication course, people have people to teach, anyway this is your matrix R, since only one process, I am not giving the subscript R x x and all that because that will consume too much of space.

R means, I am refine to process x n only, I am dropping the subscript x x everywhere, r 0, r 1 dot, dot, dot, dot, r p then again r 1 real. So, I am not putting r minus 1 or r 1 star and all that you understand n a r 1, r minus 1 they are same, so this r minus 1 which is r 1 I told you about topics, but this is in the context of optimum filter, this is the correlation transpose matrix.

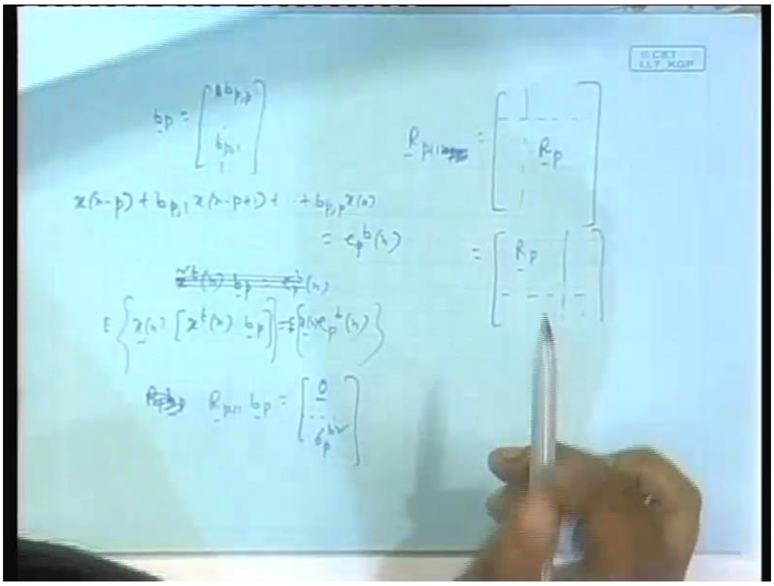
This transpose matrix is beautifully exploited here in this Levinson Durbin in to 1, a p 1 dot, dot, dot app equal to an unknown quantity sigma p f square 0 vector, what is the meaning of this equation, the way to look at this equation is this first, forget about the first row, you forget about this row times, this column will be gives rise to 1 equation, this row time, this column equal to sigma p f square.

Unfortunately, right hand side is an unknown quantity, there do not look in to that equation for the time being considered from second row onwards, what is it?

One times, that is if you have this part only and forget about this element, you got zeros and I told you how to do matrix in to vector multiplication. Remember one times, this column plus this sub matrix times, this vector is equal to all 0. So, take this column to the right hand side, so this sub matrix times this unknown vector equal to a column on the right hand side, so you solve that I am assuming these entire matrix are invariable; that means, positive definite.

So, the correlation matrices in general not negative definite, but I am assuming positive definite, so they are all similar in invertible and all solved. So, once you find the coefficients put that back here, then carry out this row times column, you get the unknown quantity on the right hand side, this is the way to go about.

(Refer Slide Time: 33:16)



Now one thing you can see this R matrix is such R matrix, R can be written like this I can take out this part. So, this sub matrix is nothing but auto correlation matrix of one order less, so r 0 to r p minus 1 same thing is propagating. So, this is we write like this, some elements here I am not writing them, some elements here and here, but essential thing is it is also instead of p plus 1 to p plus 1. Let me call it just p plus 1, because too much of space for consumed here p plus 1 into p plus 1.

When, I say p plus 1 where R is capital, so it is a matrix p plus 1 is p plus 1 in to p plus one. So, this is p at the same time, I can have a partition is forget about this dots, you can also partition like this. Same thing you will get R p and some elements here, some elements here, this is this is this comes due to the conflicts nature of the matrix.

Same sub matrix is sliding down same sub matrix has come over here on top of it there I get symmetric also though symmetric is not really needed very much here.

Non-symmetric case also, I can apply Levinson Durbin equation, if it is done in the AR model in the AR estimation part of ARMA modeling problem. Then, this matrix is not symmetric matrix, but probably and that is enough for Levinson Durbin to up now they go about there is this, that you are solving a equation like this, why not also solve another equation, which is a backward prediction equation, what is a backward prediction equation, that if I have this I can write as x tilda n transpose, I will tell you what is tilda n times b p vector as e p b n.

My definition is giving any vector for x as say x 1, x 2, dot, dot, dot say x n, x tilda means, we just reverse the order, x tilda means just x n, x n minus 1 dot, dot, dot x, first will become last will become first that way. So, you understand or you know what is x n, x n vector starting at xn up to xn minus p, I have made it tilda application. So, it is x n minus p to x n and then row of that b p vector, b p vector is my definition is b p vector or may be let us not do tilda application.

Here, I can observe that here only b p comma p tilda application I will do later, 1 this may be b p vector and then I take this x n, I multiply. So, here also I multiply x n and here also you take expected value, so x n into x transpose n, that is R b p, R is R p plus y, R p plus 1, this is p plus 1 again in to b p. Now, this vector e p b n is what x n minus p projected on all the p future terms and the space span by future term say x a, x n, x n minus 1 up to x n minus p plus 1 and the error.

So, that error is orthogonal to whom, this error is orthogonal to top component next component, next component, except for the last one. So, we will have 0 on top and last value is x n minus p, that times e p b n and expected value inner product between x n minus e p b n which again by shifting that projection operator. And all, I will not do it here, you know that turns out to be inner product of e p b n with itself, so this is sigma p b square.

So, this Levinson Durbin says that you solve instead of solving one equation like this 1 equation like this solve both the equation simultaneously and recursively for various orders, That is what we do in lattice, what we do in lattice, this is just at that algebraic version, that is if you take out that vector space treatment the inner vector space meaning of that entire just look at it from equation solving point of view.

Then, you have that you simultaneously solve the forward and backward prediction error problem order recursively. If I assume I will assume that say solution for p minus 1, p th order is known for these equation also, these equation also, then how to solve this and this equation. Then, suppose p minus 1 p th order for the matrix should have been R p that is known to us, how to use, how to solve for p th order.

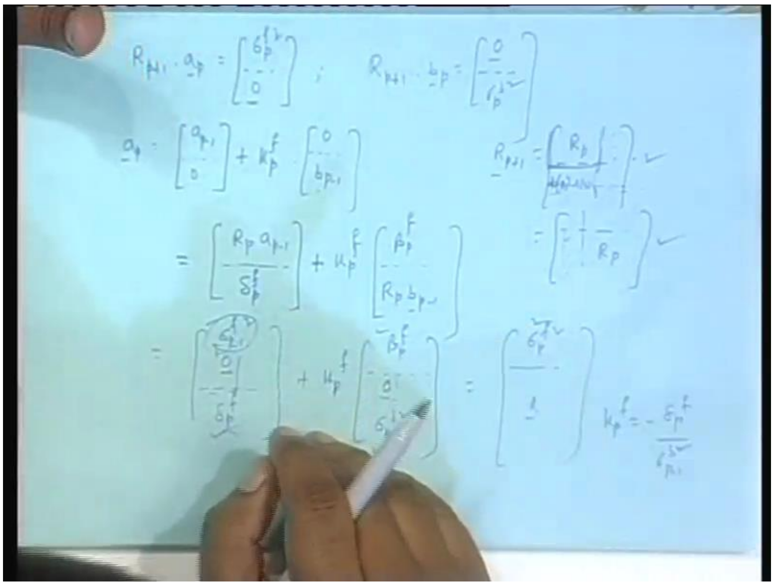
So, the way they go about it and again you know I will just give that indication and stop because, we will basically get the same equation and I am doing it. So, that and at least I am doing the formality of Levinson Durbin with the purpose that, when see books or come across Levinson Durbin, you are not you do not get confused thinking that it is something else and all that you see the formulation of this.

So, now I say that there are two equations, one is you can now write R p plus 1in to a p is equal to this form these equation, another is these two equations and I have got that this decomposition of R. Now, start with any of them, suppose you are solving this and remember you have got all zeros here, only 1 nonzero component and all zeros. Then, people say that suppose that unknown solution a p can I write it like this a p minus 1, I am just constructing the solution.

And see, I will give you the logic for this kind of and some coefficient k p minus 1 in to b p minus 1, where this is an unknown to be found out, what is the logic for this kind of construction, if I put that here this equation in this first equation R p plus 1 in to a p, a p is this is the equation to be solved, 0 vector if you put that here if you replace a p by this.

By the way, just a minute I was making one mistake, it is not, let me do it over separate page, this is dimensionality problem, A p has got how many coefficients 1 and p coefficients, a p minus 1 how many coefficients on top 1 n and p minus 1 coefficients. So, two vectors are not same, I have to put a 0 to make them same, one is of length, one higher than the other.

(Refer Slide Time: 42:14)



Our basic equations, I am writing in to a p is sigma p f square by 0, this is one equation and another one was in to b p, a p was 1, than a p coma 1 and dot, dot, a p coma p that suppose, I am saying this is to found out. I know it is value further p minus 1 th order a p minus 1 suppose I know that, I am saying that suppose I construct a solution like this and this is equation, it will have only one solution, because this is the invertible matrix ax equal to b a invertible means solution is unique a inverse b.

So, I have constructed a solution, if it qualifies to be solution, this is the only solution, because this equation has unique solution. So, I am just trying to construct a solution, I will you will see immediately that this way you can construct a solution, K p sample error k p f. Sometimes, 0 b p minus 1, suppose I put like this and I want to put it on the left hand side, right hand sides this is unknown essentiality is will I get a block of zeros, if I put it here.

So, first R p plus 1 in to this, what was the break up to R p plus 1, I did it in the previous page R p and some terms and also R p and some terms, you remember this one, either this top square matrix is sub matrix is R p and then one row, one column or first row, first column, then R p. Now, if I substitute these for a p here, R p plus 1 will work on this for the first that time, suppose I use this form, that mean R p in to a p minus 1 plus this elements will multiply the 0, this row times this, the last element is forgotten, because of the 0.

Second row times, this last element again multiplied by 0, so last element has got no impact, it is only R p times a p minus 1, that is the previous Yule Walker equation for p minus 1 th order. That is here, you have got R p, a p minus 1 and this row times this, there will be some element, I am done with this part times this I am not with last row times, this will be some quantity, but that quantity I can work out.

Because, I know these terms, I know this position, wells I know a p minus 1, this is what manually you have to compute. In fact, this is the thing which you do in the delta p composition, even in the delta p composition, I told you there is no recursion possible, you have to carry out the sum, this is what is equivalent to that those things will come up. So, here, actually if you really put those terms R p, what are the terms here R p, R 0, dot, dot, dot.

Those terms times a p minus 1, but my purpose is not to get the exact figures and all, I am just giving you the philosophy used in because of solving this, because I have already solved it, by lattice method I am only telling you what they do. See you get a known quantity here, we call it delta p f, fine this is the known quantity, what is the unknown quantity inner product of row times, that is, this row times this column, but this is known.

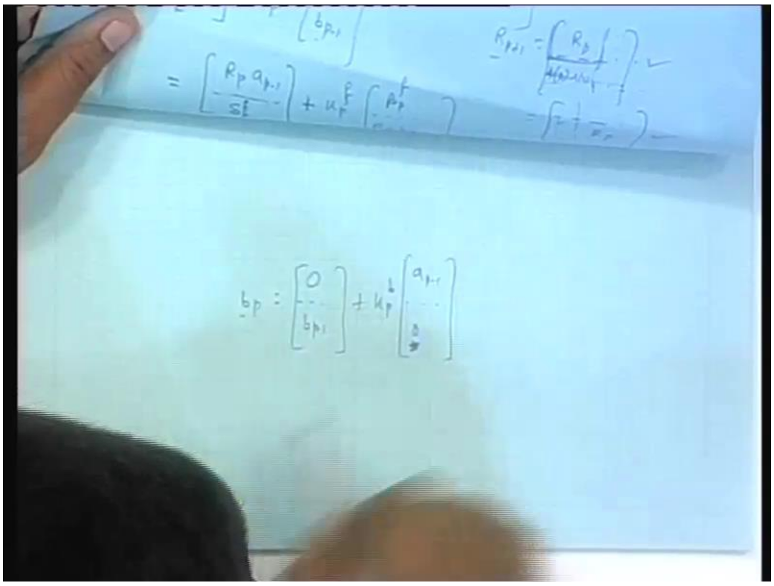
And next, one is k p f times R p plus 1, this time I will use this break up, this in to this vector. So, first row times this will be something, you can call it beta p f, so known quantity, but after that what you have first element times 0, 1 than R p. All the first elements they get multiplied by zeros, so you essentially are left with R p times b p minus 1 and these are all previous order p minus 1 th order, that Yule Walker equation.

So, this was giving rise to sigma p minus 1 f square than zeros followed by delta p f and k p f beta p f. This is the previous equation of this kind, only thing is instead of p plus 1, it becomes p this p minus 1, so this is 0 and sigma p minus 1 b square. I want this to be of this form that is a positive quantity I want this positive unknown quantity and then zeros.

Here zeros are coming, see this guy plus k p f times, this is known to be from, this is known to be from the previous order solution. Suppose, I know k p f in that case, this plus k p f in to this will be this unknown quantity that is obtained, then there is a series of zeros here there is a series of zeros 0 plus k p f time 0. Then, the rest of zeros are required, only the last row that also has to be entered in to 0, so that means, this plus k p f times this must be equal to zero, but these are known quantity that will give you k p f.

So, you solve for k p f, k p f is minus delta p f by say by sigma p minus 1 b square, that k p f you put back here, you will get this term. Similar, approaches for in this case backward prediction case, I am not doing that, similarly in the backward prediction case also similar approach exists.

(Refer Slide Time: 49:09)



That is there, I am just writing the solution there, so this b p you find out, this b p you write there as this only 0, dot, dot, dot, b p minus 1 and say k p b all 0. That is this vector remains same, earlier to this plus k p f times this and next times this plus k p b times this, I will use those factor of this decomposition of this matrix, I use the same philosophy. Evaluate k p b that at there, you can try to prove that, this delta p f square, delta p b square, they turn out to be same because of symmetric probably R matrix and all that.

Whatever be in the R matrix the first row terms and last row terms they are same, because of nature and symmetric nature. All those, that is why those will be same sigma p f square sigma p, this you can try to prove that, directly which we did by using vector space statement is a lattice.

So, this is another treatment used simply equation for been treated, but indeed here you do not get to know what is happening inside, the other thing, why I prefer that vector space treatment is back, I will now go to first adaptive version of that lattice. But after that, I will move to what is called recursive least squares not elements type, another class where that vector space formulation will be the key, that is how a similar treatment, that is why I want, I think the best formulation for an base adaptive filters and spectral is linear is that vector space treatment.

So, actually I thought that today's class will be variant adaptive lattice, which I have to be continuing tomorrow, gradient that so for the lattice filter, that we have considered that is given by some constant parameters k p for a more k p's less than 1 and all that. And k p is given by delta p by sigma p square and all that, if you know the input correlation, the values construct the lattice fine.

But suppose input statistics is changing from time to time and you do not really know the statistics, so that means like an adaptive filter, here also you have to make them adaptive. Suppose input statistics is changing from time to time and you also do not know, so you have to learn from the data and adapt going to an adaptive mechanism by using those coefficients reflection, you have to iteration on that elements like manner. So, that they will finally, converge on their actual optimal values, so that lattice is called an adaptive lattice or called gradient adaptive lattice. Unfortunately, there is no time to do, so, I will consider it tomorrow.

Thank you.