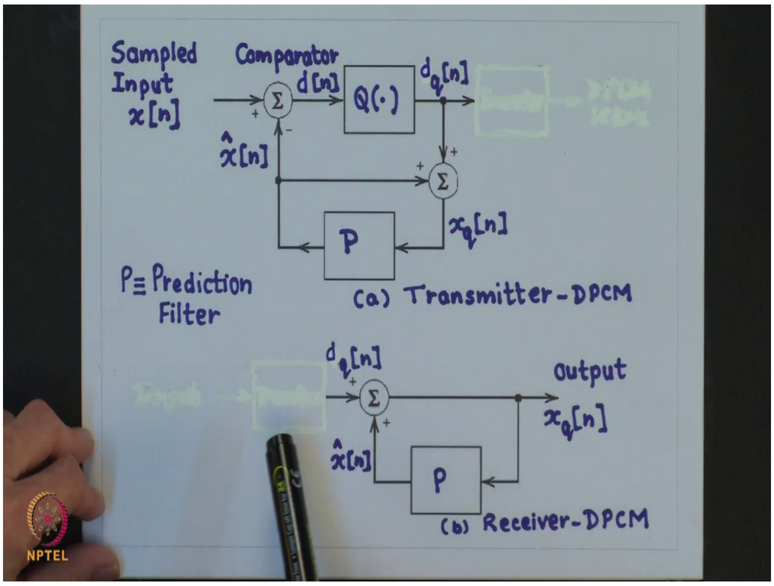
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture – 31
DPCM – II (Linear Prediction)

In differential pulse code modulation popularly known as DPCM, we transmit some kind of a different sequence rather than the original sequence itself. Now we have shown that the quantization noise associated with the sample of the original signal or the original sequence is same as the quantization noise associated with the samples of the different sequence. We have also seen that quantization noise is directly related to the variance of the dynamic range of the input to the quantizer.

Now, we have also studied that the different sequence has variance and dynamic range, which is much lower than the variance and dynamic range of the original sequence. So, DPCM gains advantage in this sense. Now how much is the reduction of this variance will depend how well you can predict the current sample from the previous reconstructed sample. So, let us try to formulate this problem of prediction in a mathematical framework, and the analytical solution which we will get to this problem will help us to understand some of the popular used approaches in the design of the predictor.

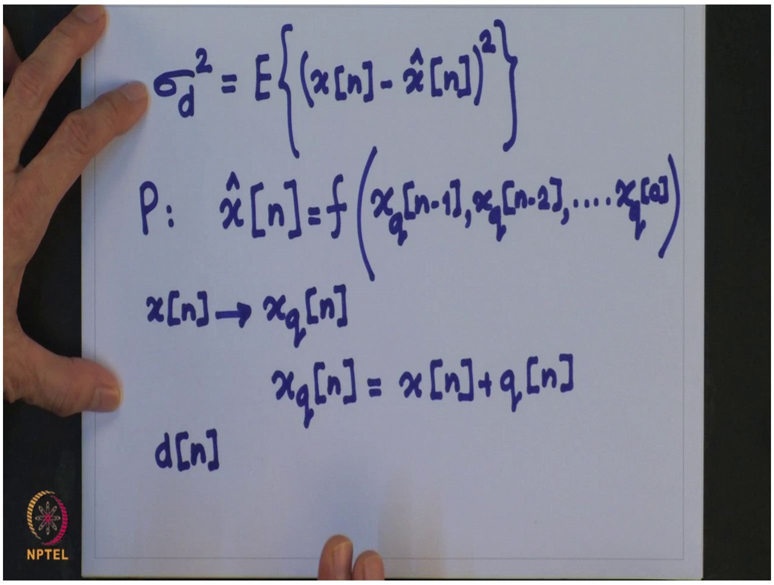
(Refer Slide Time: 02:18)



So, let me just show you this right which we had studied last time also. So, this is a transmitter end of the DPCM and this is the receiver end of the DPCM and we have this block which we call it as a prediction filter block, which is there both of the transmitter end and at the receiver end. And what we are interested is basically to get a good predictor so, this variance of this difference sequence can be substantially reduced.

Now, when we are trying to find out the optimal predictor, our criteria for optimization would be to minimize the variance of the different sequence.

(Refer Slide Time: 03:10)

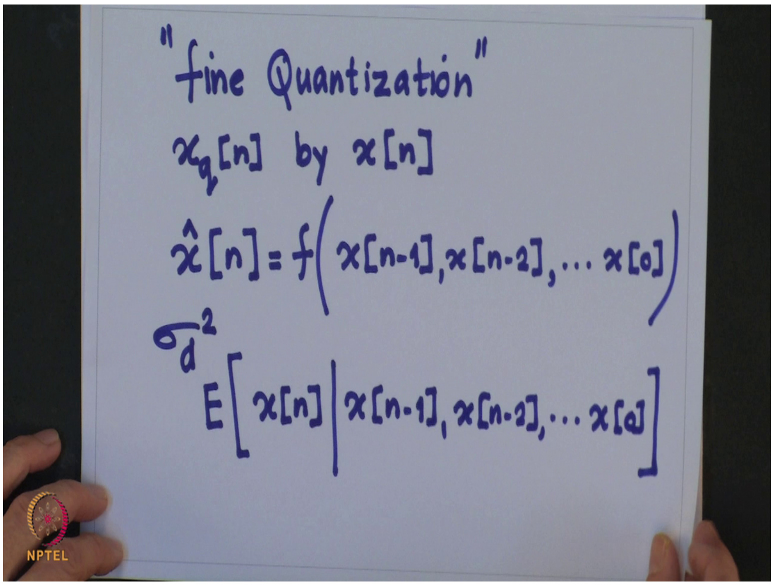


This is your input sample and this is the predicted sample squared error and expectation basically will give you the mean value of this squared error. And we said that our predictor is nothing, but it is a function of the past reconstructed values. Now design of a good predictor is essentially the selection of this function f that minimizes the variance of the difference sequence. Now the problem with this formulation is as follows.

The reconstructed sample of x n is x q and this xq n is equal to x n plus quantization noise. So, this quantization noise depends on the variance of the difference sequence. Thus by picking the function f we will affect the variance of the different sequence, which we in turn will affect the reconstruction x q which then affects the selection of f. This coupling makes an explicit solution extremely difficult for even the most well behaved source and most of the real sources are far from well behaved. So, the problem becomes computationally intractable in most applications.

So, we will try to find the solution to this problem by making an assumption and that assumption is known as fine quantization assumptions.

(Refer Slide Time: 06:20)

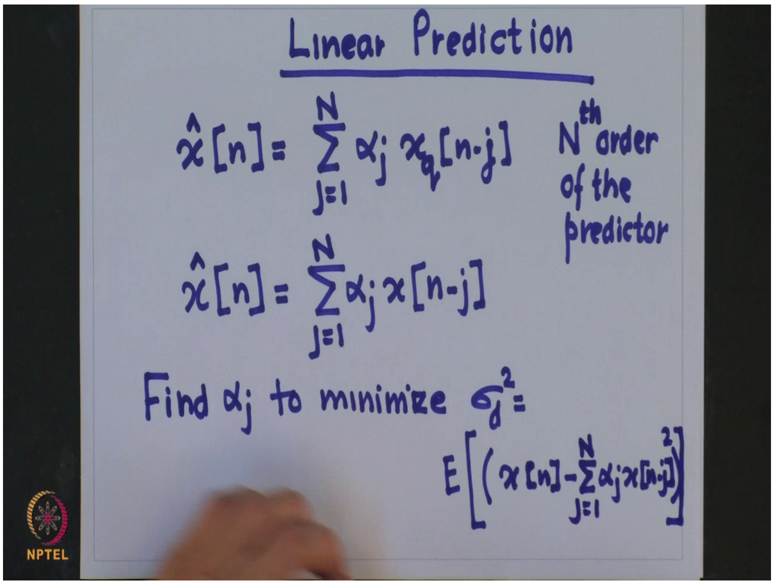


This assumption states that we will assume that your quantizer step size is so small, that we can replace the reconstructed value of the sample by the value of the sample itself correct. In that case your x hat which is the predicted value will become functions of the past original samples and not the past reconstructed samples. Now once the function has been found using the past original samples, we can use with the reconstructed values x q to obtain x hat.

Now, from the study of random processes, we know that for a stationary process the function that minimizes the variance of the difference is the conditional expectation of the sample x n given the past samples. Now unfortunately this condition of stationarity is generally not true in a practical applications and even if it was true, finding the solution as conditional expectation is difficult because this involves a nth order conditional probabilities.

So, we will simplify our predictor design by assuming that we are interested in what is known as linear prediction.

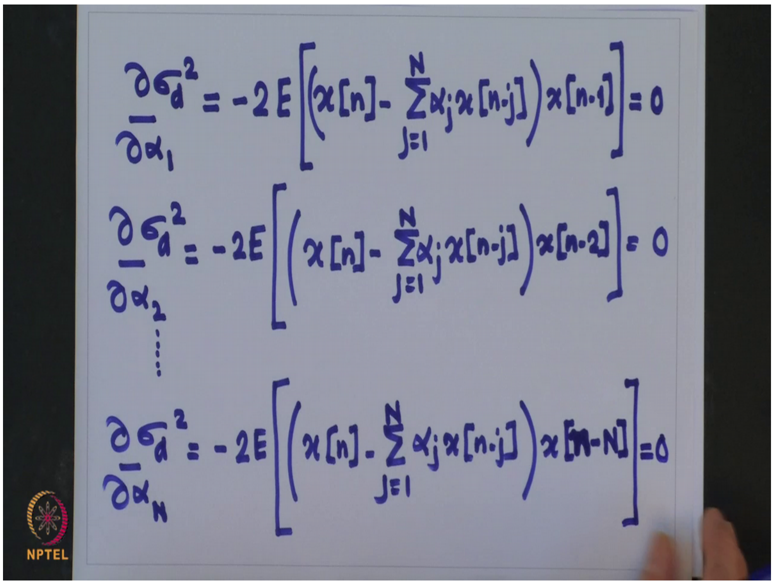
(Refer Slide Time: 08:54)



In linear prediction what we are supposed to do is that, find out those values of these weights called alpha j. So, the problem is you find the predictor using the linear combinations of the past reconstructed values and we use capital N to be the number of past samples. So, this N is also known as Nth order of the predictor. So, original problem has been put in the linear framework like this and if we assume that we have fine quantization, then our problem reduces to finding alpha j such that the variance of the different signal using this linear combination gets minimized. So, find alpha j, j equal to 1 to n to minimize the variance which is equal to expectation of j equal to 1 to capital N.

Now, we will assume that, this sample sequence comes from a wide sense stationary process and that condition is satisfied I will be able to calculate the autocorrelation and cross correlation of the sequences so let us do that.

(Refer Slide Time: 12:00)

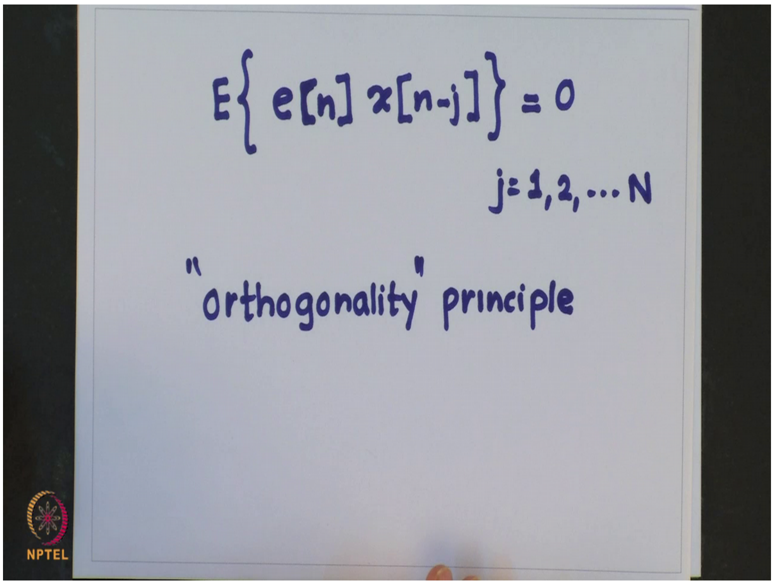


So, we will take the derivative of this partial derivatives with respect to each of the coefficient. So, when I take it to with respect to alpha 1 I will get this value the first term.

It is important to note that differentiation and the expectation operators are linear. So, in most of the practical cases, we can interchange the order of this operation and this is what has been done here. First the partial derivative has been taken and then the expectation, this will be equal to 0, taking differentiation with respect to alpha 2 would give you and then we are to take the differentiation of this differentiation of this with alpha 2 will be 0, here all the terms will vanish except corresponding to alpha 2 and from there you will get only this term left out this will be equal to 0. So, this way if you proceed, this is the last this is small n minus capital N means, this is equal to 0 this quantity out here is nothing, but the error in prediction.

So, all this equation suggests that the error is orthogonal to all those components which are being used for linear prediction.

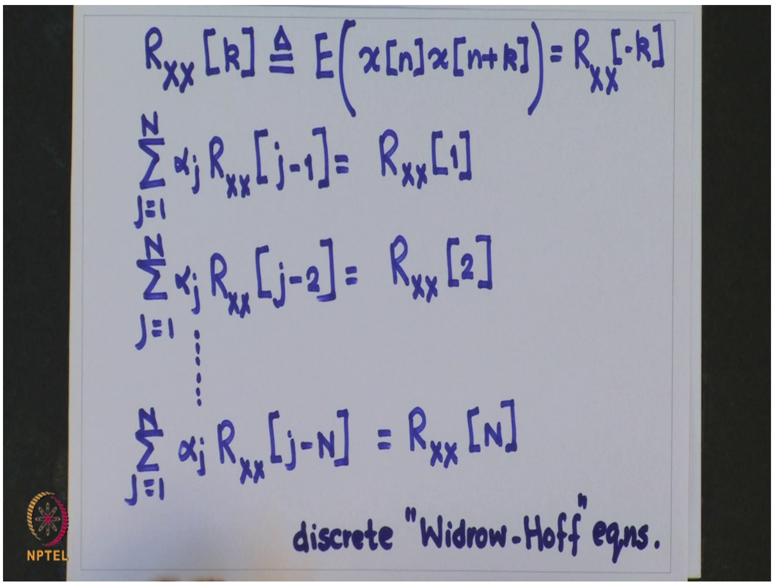
(Refer Slide Time: 15:18)



So, what it means is that, expectation of the prediction error and all the components which are being used for linear prediction, this is equal to 0 for j equal to 1 to up to n. This is known as orthogonality principle.

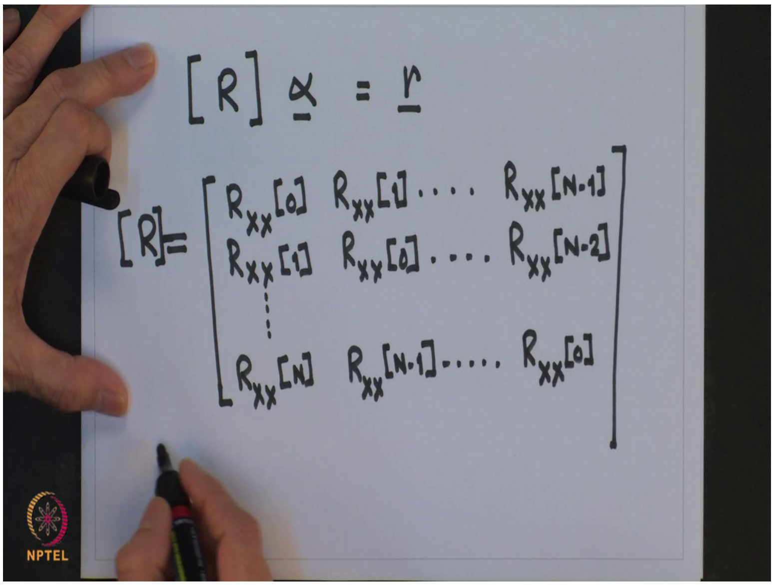
So, whenever the error is orthogonal through the components, which are being used for linear prediction, then you will get the minimum mean square error. Now taking the expectations here this will get expectation e will go inside we can rewrite this equation. But before we rewrite this equation let me define the autocorrelation as follows.

(Refer Slide Time: 16:46)



I am assuming that all our samples in the sequence are real, this is a property of the autocorrelation. Autocorrelation is possible because of the assumption of wide sense stationary. So, all these equations which we have n number of equations can be re written as follows. The last equation which will have; see this set of equations are also known as discrete "Widrow-Holf" equation.

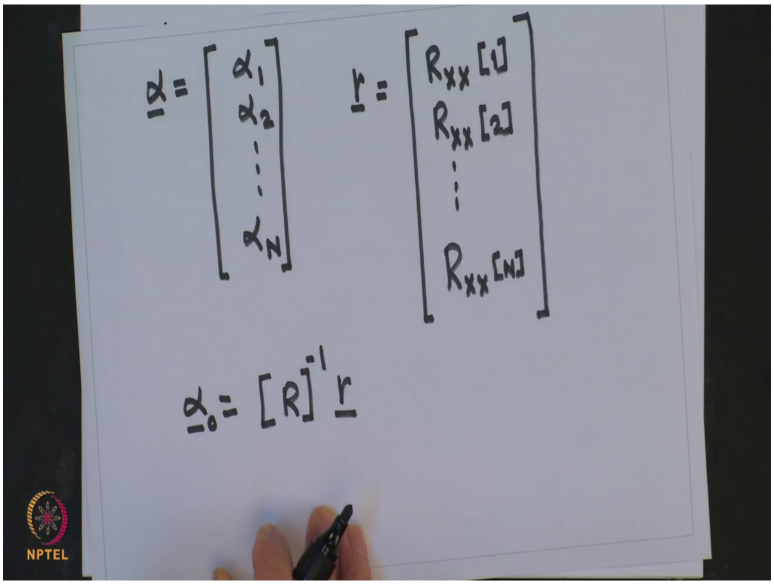
(Refer Slide Time: 19:00)



We can put it in the matrix form as follows, where your R matrix is n by n matrix and it is composed of the following elements here.

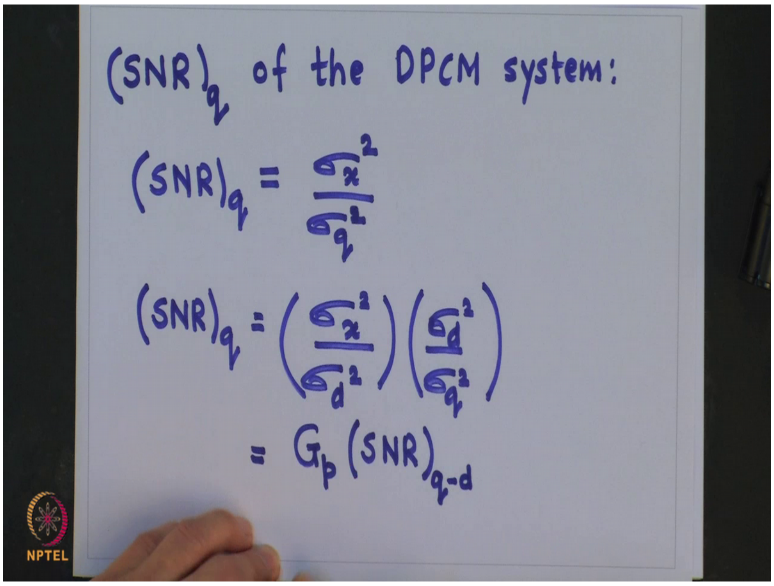
While writing this terms in this matrix, we will use the property of the autocorrelation.

(Refer Slide Time: 20:24)



This says that this autocorrelation is a symmetric function your alpha is a weight vector, your small r vector is autocorrelation elements. Now if you assume that this matrix R is invertible, then I will get my alpha optimum to be ok. So, when I take this, this will call it as alpha optimum fine ok.

(Refer Slide Time: 21:25)



Now, let us calculate the signal to noise ratio of the DPCM system, when I say signal to noise ratio here I mean the quantization noise of the DPCM system this is nothing, but the variance of the input signal divided by the quantization noise of the samples in the sequence. So, fine now this I can rewrite it as. Please note that the quantization noise associated with the samples of the original signal is the same as the quantization noise associated with the samples of the different sequence.

Now, so, this is the property of the quantizer, the difference sequence is input to the quantizer and this is the quantization noise associated with that quantizer. This quantity out here is known as the processing gain of the DPCM and is denoted by G p and this quantity out here I would denote it as the quantization noise associated with the different signal. To have a high signal to quantization noise ratio for DPCM system for any particular quantizer so, this is your quantizer property. What you have to do is try to increase the processing gain; that means, your design of a prediction filter is to maximize this ratio that is maximize your.

We have seen that the DPCM consists of 2 blocks, one is basically the quantizer block and other is the predictor block. Now we know that the quantization noise associated in the DPCM system is dictated by the quantization noise of the different sequence. So, in that case what it means that, we could decrease the number of quantization levels in the quantizer to obtain the same quantization noise, which I would get by quantizing my original sequence if that is acceptable. So, this better prediction allows you to decrease the number of quantization levels in the quantizer.

The question is how low can you have this value of quantization level. The lowest value you can have is 2. So, you have only 2 levels. So, I require only one bit. This form of DPCM where I use a quantizer of just one bit is known as delta modulation, which we will study next time.

Thank you.