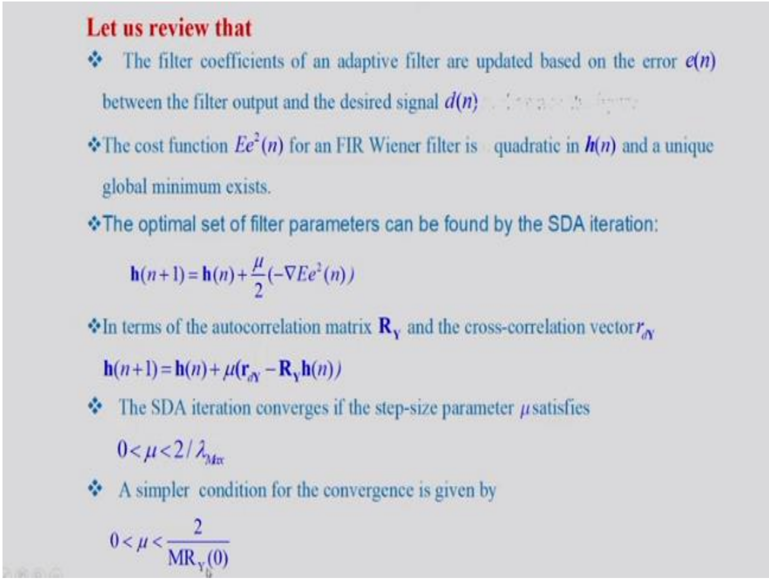
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology – Guwahati

Lecture – 26
Adaptive Filters 2

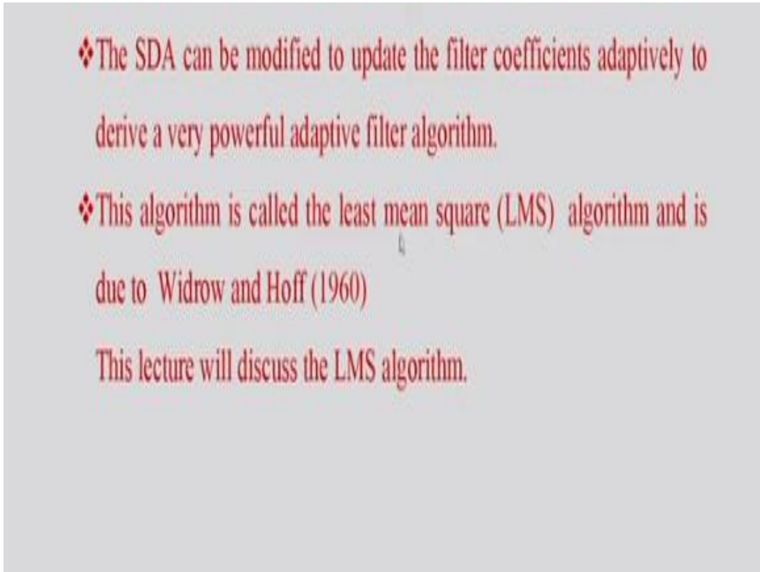
(Refer Slide Time: 00:42)



Hello students welcome to this lecture on adaptive filters. Let us review that the filter coefficients of an adaptive filter are updated based on the error en between the filter output and the desired signal dn. The cost function E of e square n for an FIR Wiener filter is quadratic in hn and the unique global minimum exist. The optimal set of filter parameters can be found by the SDA steepest descent algorithm iteration which is given by h of n + 1 = hn + mu/2 times negative of the gradient of E of e square n this is the steepest descent iteration.

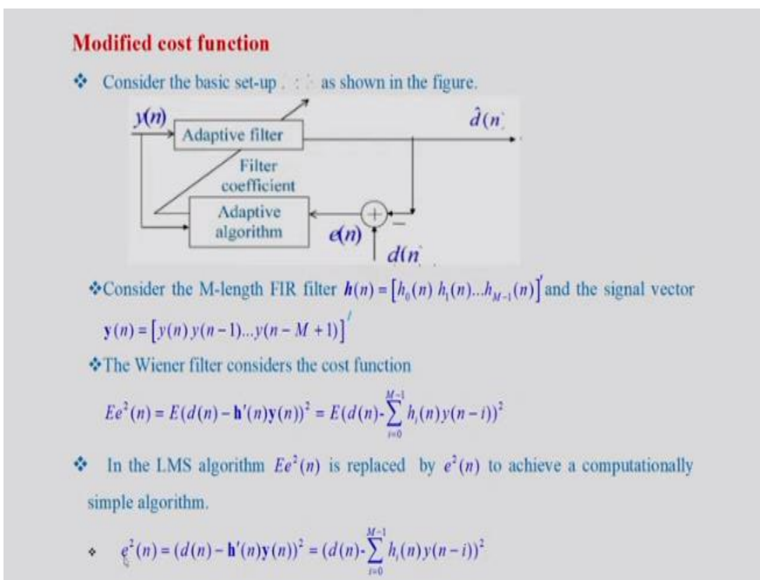
So iteration in the direction of negative of gradient in terms of autocorrelation matrix RY and the cross correlation vector a small rdY this updating can be written as h of n + 1 = hn + mu times rdY - RY into hn matrix. This SDA iteration converges if the step size parameter mu satisfy this relationship 0 is < mu < 2 / lambda max where lambda max is the maximum eigen value of the RY matrix. A simpler condition for the convergence is given by mu lies between 0 and 2/M into RY0 where M is the length of the filter.

(Refer Slide Time: 02:39)



The SDA can be modified to update the filter coefficients adaptively to derive a very powerful adaptive filter algorithm. This algorithm is called the least mean square LMS algorithm and is due to Widrow and Hoff 1960. This lecture we will discuss the LMS algorithm we will see how the cost function is modified for LMS algorithm.

(Refer Slide Time: 03:08)



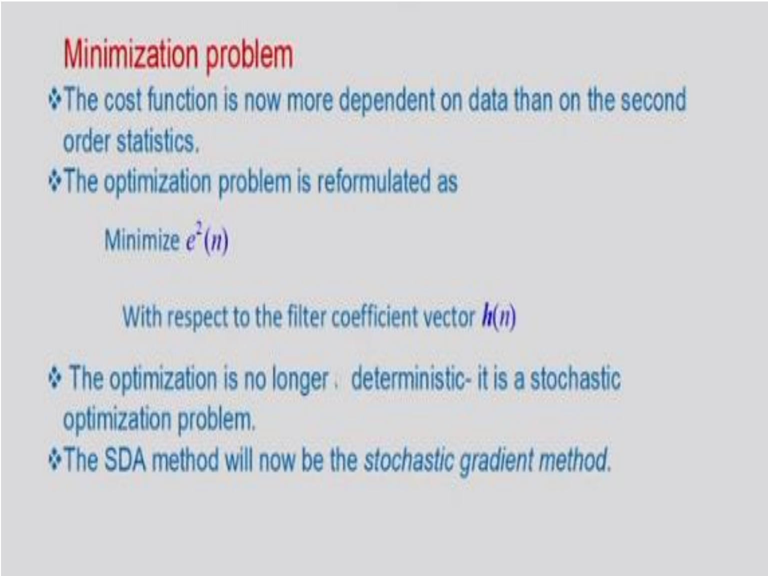
You will see how the cost function has been modified for LMS algorithm. Consider the basic setup as shown in this figure here yn is the input to the adaptive filter and its output is d hat n which is compared with this desired signal dn and the error is used to update the filter coefficient according to the adaptive algorithm and those filter coefficients are used in the adaptive filter.

We will consider an M length FIR filter hn that is filter coefficients are given by h0n,h1n up to h m - 1n and the signal vector yn comprising of yn, yn - 1 up to yn - m + 1.

The Wiener filter considers the cost function expected value of e square n that is the mean square l error that = E of dn - this is the filtered output is transpose n into yn whole square and this we can write as E of dn - summation h ai n into yn - i i going from 0 to n - 1 whole square. In the LMS algorithm E of square n is replaced by e square n to achieve a computationally simple algorithm.

So modification is very simple just E of e square n you are replacing it by E square n. Therefore the cost function is given by e square m = dn - h transpose n into yn whole square that = dn - summation h in into yn - i i going from 0 to m - 1 whole square. So this is the cost function for the LMS algorithm.

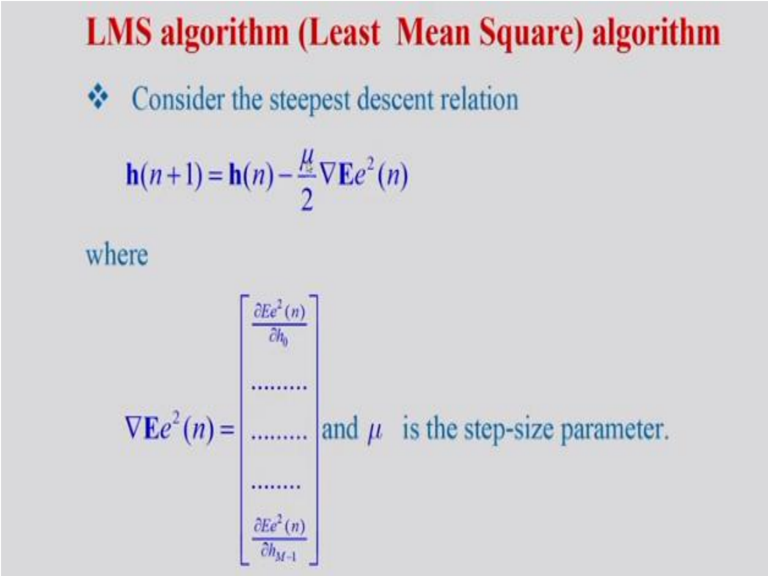
(Refer Slide Time: 05:33)



The cost function is now dependent on the data rather than the second order statistics because we know that in the case of Wiener Hoff solution, the solution is in terms of the second order statistics that is cross correlation matrix and autocorrelation vector. The optimization problem is reformulated as minimize e square n with respect to the filter coefficient vector hn. This is the LMS optimization problem.

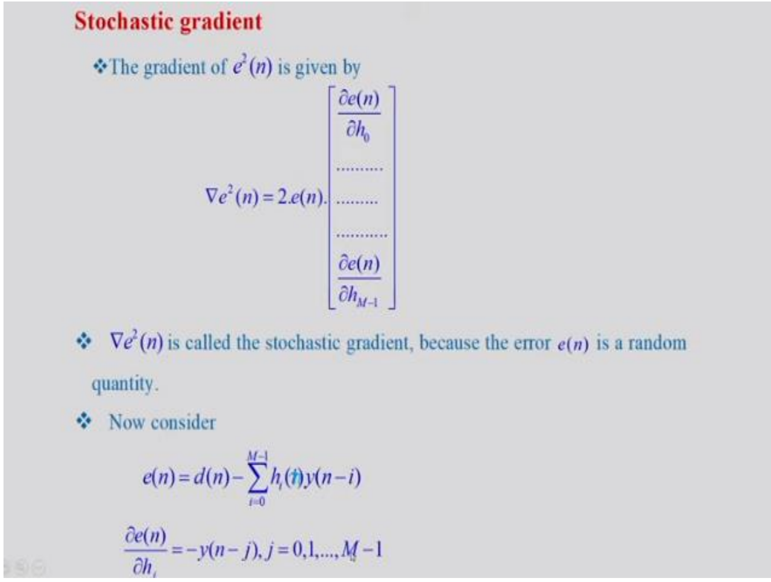
The optimization is no longer deterministic it is a stochastic optimization problem why because e square n it is dependent on the data which are random. This SDA method will now be called stochastic gradient method steepest descent algorithm will be used here but it will be called this stochastic gradient it is the gradient of a random quantity.

(Refer Slide Time: 06:44)



We will discuss the LMS algorithm consider this steepest descent relation that is iteration h of n + 1 = hn – mu/2 into gradient of E of e square and here 2 is used because we will get 2 terms because of this gradient both of them will cancel. Where that gradient of E of e square n is given by this vector first component will be del del is not E of e square n like that last component will be del del is m - 1 of E of e square n and mu is this step size parameter. Therefore for the convergence of the iterations we have to choose this new parameter properly.

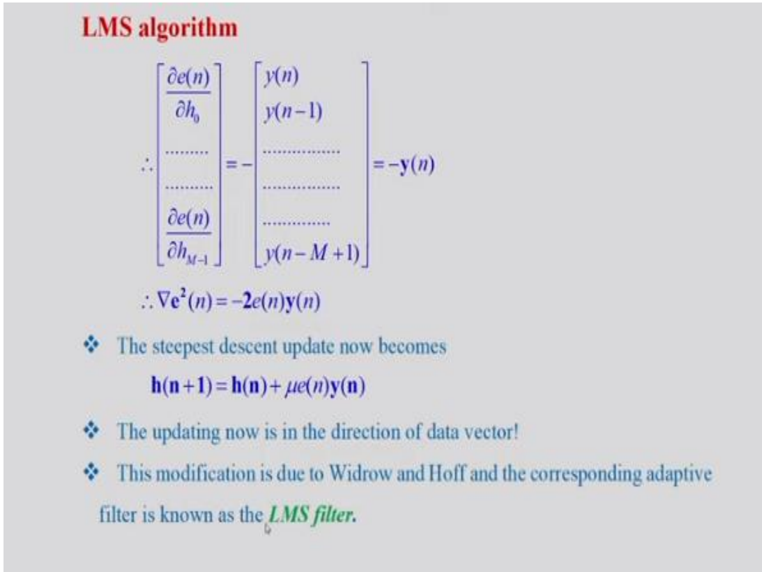
(Refer Slide Time: 07:50)



Now the gradient of e square n is given by gradient of e square n = 2 times en into this vector del en del h0 to del en del h M-1. Because it is a function of a function so first, we have to consider a derivative of e square n that is twice en then derivative of en with respect to each of the parameter. A gradient of e square n is called a stochastic gradient because the error is a random quantity because the n is a derived from the random data therefore it is a stochastic quantity and corresponding gradient of e square n is stochastic gradient.

Now consider the error en that = dn - the filtered output, filtered output is summation its in into y and - i i going from 0 to n – 1. This is the error signal now taking the partial derivative del en del hj = this negative sign is there we are taking the partial derivative with respect to h there so only yn - j will be there. So therefore del en del hj = - y of n – j, j going from 0 to m – 1. So, that way partial derivative of en with respect to hj is given by negative of yn – j.

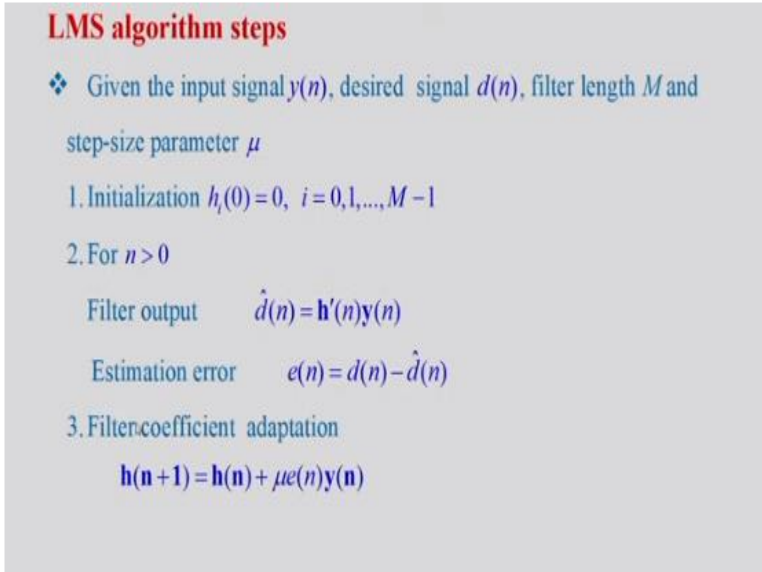
(Refer Slide Time: 09:50)



The partial derivative factor is now given by negative of this vector yn yn - 1 up to yn - m + 1 that = - yn vector. Therefore gradient of e square n will be - 2 times en into yn vector from this relationship variant of e square n = 2 times en into this partial derivative vector. So from that we will get that gradient of e square n = - 2 times en into yn. This steepest descent update now becomes h of n + 1 = hn + mu times n into yn.

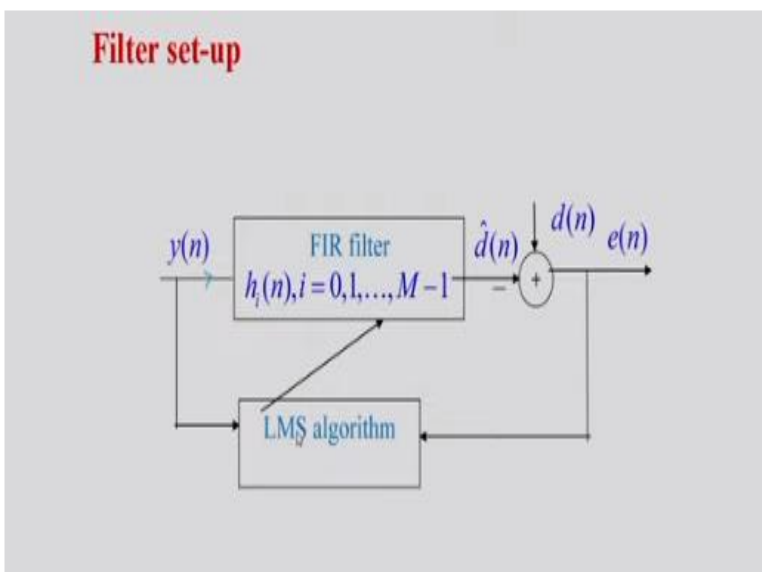
Now updating is in the direction of yn instead of in the direction of the gradient of the mean square error. So the updating is in the direction of the data vector. This modification is due to Widrow and Hoff as we have told earlier, and the corresponding adaptive filter is known as the LMS filter.

(Refer Slide Time: 11:23)



You can get these types of LMS filter. Given the input signal yn, the desired signal dn the filter length M and the step size parameter mu these are the input to the algorithm. Initialization hi0 = 0 for i = 0 up to m - 1 all the filter coefficients are initialized to 0. Now this is the updating for n>0 filter output is d hat n that = h transpose and into yn. So yn is passed through the adaptive filter that output is d hat n estimation error en = dn - d hat n. Now the filter coefficient will be adapted as h of n + 1 = hn + mu en into yn vector.

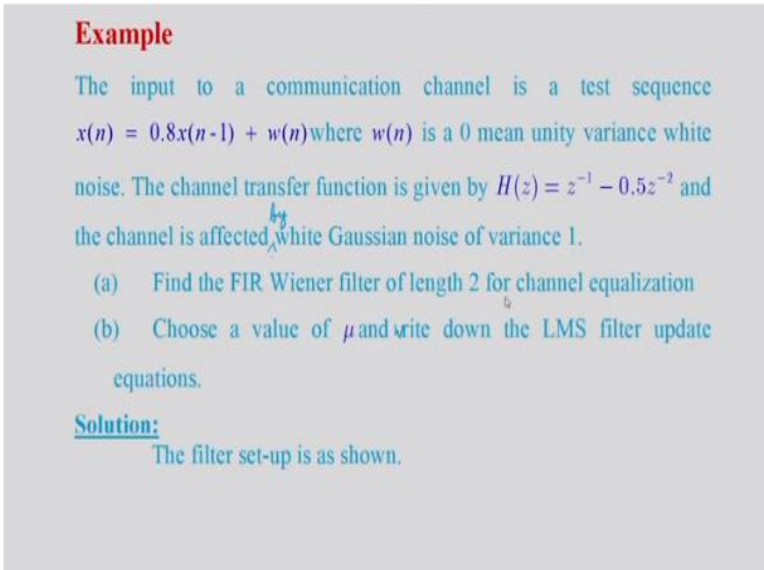
(Refer Slide Time: 12:35)



Now we describe the LMS algorithm based adaptive filter set up here yn is the input this is passed through the FIR Wiener filter of step length m - 1 then we will get the estimate of the desired signal and this is the desired signal which we have to decide. So this is positive this one is negative so that difference is the error signal this error signal is passed to the LMS algorithm and they are the filter coefficient will be updated according to the error and the filter coefficient will be passed over to the FIR adaptive filter.

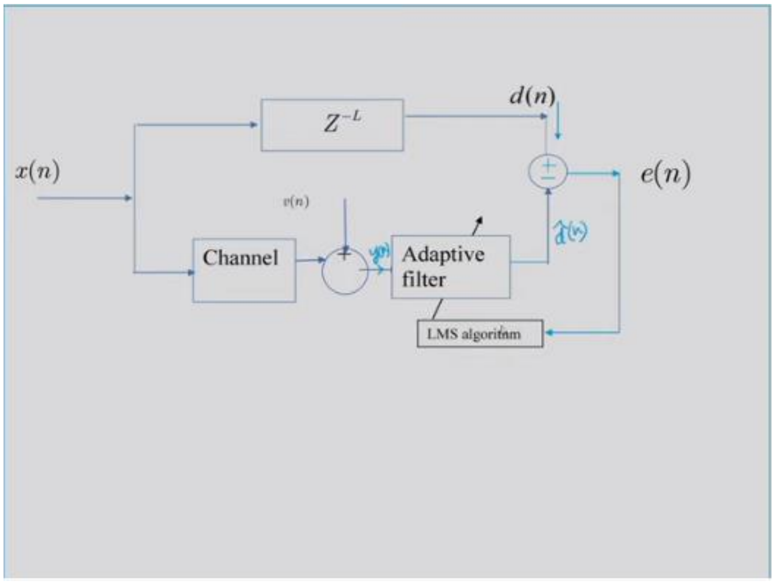
So that way this LMS algorithm has 2 input one is the data itself because the time vector is needed there, and this side is the error signal.

(Refer Slide Time: 13:40)



Let us consider one example the input to a communication channel is a test sequence given by xn = 0.8 xn - 1 + wn where wn is a 0 mean unity variance white noise. This channel transfer function is given by Hz = z inverse - 0.5 z to the power – 2. So this is an FIR filter and this channel is affected by white gaussian noise of variance 1. Find the FIR Wiener filter of length 2 for channel equalization choose a value of mu and write down the LMS filter update equations. First, we will show the filter set up.

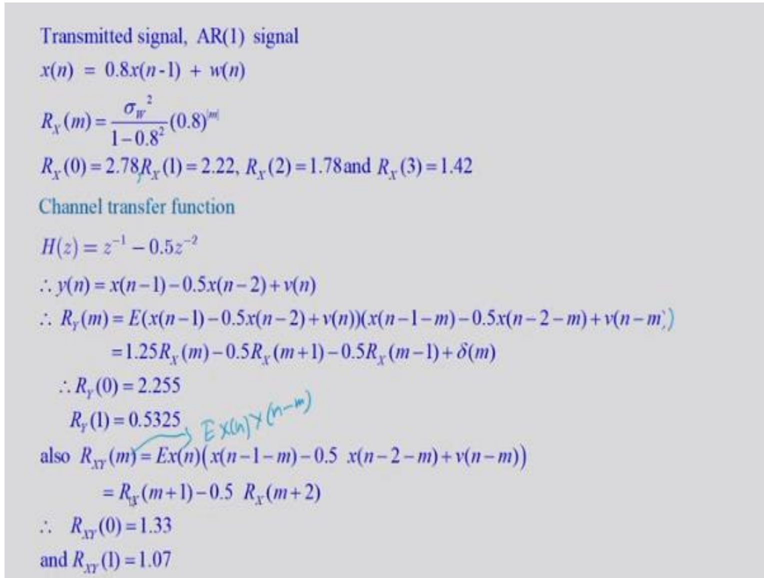
(Refer Slide Time: 14:33)



The filter set up is like this the test sequence xn is passed through a delay to the receiver so that where dn is the test sequence, but it is delayed by some amount. Now the test sequence is passed traditional good transfer function we know and then there is a addition of noise so this vn noise is added here and the output is passed to the adaptive filter. The adaptive filter will estimate d hat. So d hat n will be estimated, and the error will pass over to the LMS algorithm.

So this is this channel equalization problem here this is the test sequence this is passed through the receiver as it is but there is a communication delay Z to the power – L that is L unit of time delay then this test signal is passed through the channel. There is addition of gaussian noise, and the output is yn and this is filtered by the adaptive filter of length 2. The estimated signal is d hat n and the corresponding error is sent to the LMS algorithm for updating the filter coefficients.

(Refer Slide Time: 16:09)



Let us solve the problem transmitted signal is an AR1 what is this xn = 0.8 xn - 1 + wn. Recall the AR model in our earlier lecture we know that autocorrelation function which was derived from the Yule Walker equation is given by RX of m = sigma w square divided by 1 - 0.8 square into 0.8 to the power m and here sigma w square = 1. So if we carry out the calculation, we will get RX of 0 = 2.78, RX of 1 = 2.22, RX of 2 = 1.78 and RX of t = 1.42 like that.

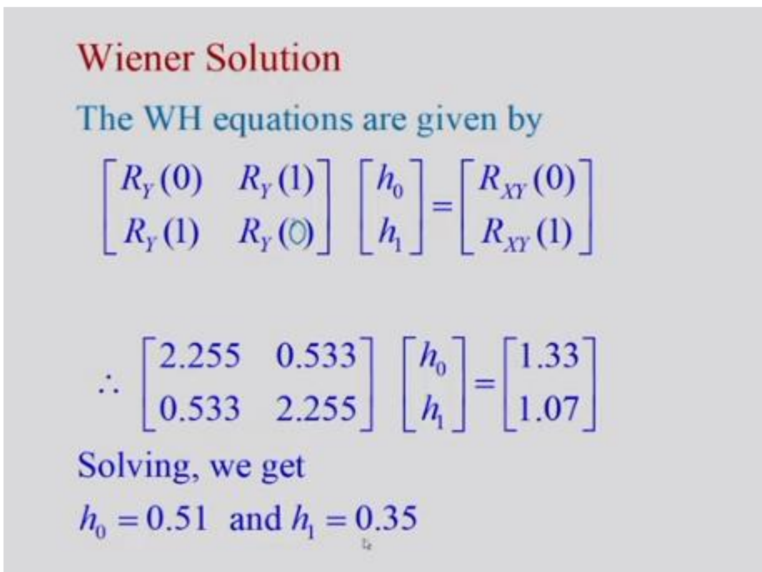
Now channel transfer function is given by Hz = z to the power - 1 - 0.5 z to the power - 2 this is an FIR filter. Therefore we can write output of this filter that is xn is this this is passed through Hz so output will be yn = x of n - 1 - 0.5 xn - 2 + bn. So this is the output of the FIR filter. Therefore RYm will be = e of yn into yn - m and this will be = e of x of n - 1 - 0.5 xn - 2 + bn into x of n - 1 - m - 0.5 times x of n - 2 - m + v of n – m.

So we have to find out the expectation of this quantity and if we carry out the multiplication and then take the expectation, we will get first term will be 1.25 into RX of m because it will be e of x of n - 1 into x of n - 1 – m. So, that way we will get RX of m and this also 0.5 xn - 2 into 0.5xn-2 - m this will also give RX of m. So this will be 0.5 into 0.5 0.25 RX of m. So that way we will get here 1.25 RX of m.

Similarly there will be the cross terms so because of that we will get - 0.5 RX of m + 1 and - 0.5 RX of m - 1 and this noise is of brilliance one and they will be uncorrelated with the signal and uncorrelated with the sample of the noise also but only when m = 0 they already hold a relation and that is the variance. So that way this term will be delta m on that variance = 1. Therefore RYm will be = 1.25 RX of m - 0.5 RX of m + 1 - 0.5 RX of m - 1 + del m and we know the values of RX of m.

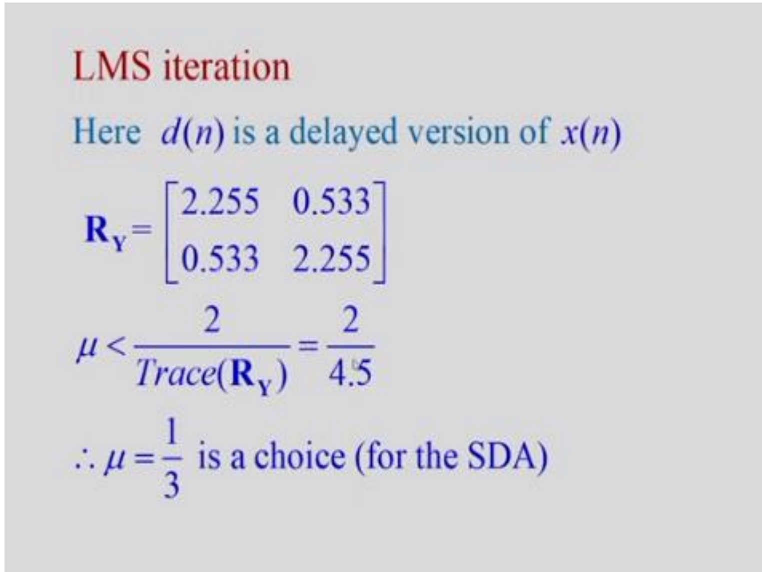
Similarly we can find out the RXYm that = e of x n into x of n - 1 - m - 0.5x n- 2 - m + v of n - m this is the yn - m this is xn. So that way RXYm we write it as this = e of xn into yn - m so that way if we substitute yn - m by this expression we will get this result so that way RXYm = RX of 0.5 into RX of m + 2 it means that RXY of 0 = 1.33 RXY1 = 1.07 we have here RY of 0 = 2.255 RY of 1= 0.5325. So we have the autocorrelation and cross correlation values so we can solve the Wiener Hoff equation.

(Refer Slide Time: 21:28)



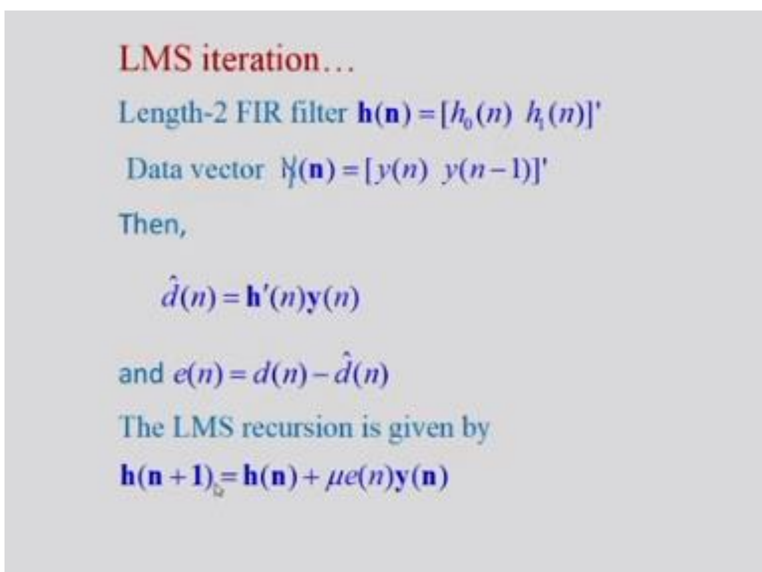
Therefore the Wiener Hoff equations are given by this autocorrelation matrix comprising of RY0, RY1, RY1, RY0 multiplied by the coefficient vector h0 h1 is equal to the cross correlation vector that is RXY0 RXY1 and if we substitute the values well get this matrix multiplied by h0 h1 is equal to this cross correlation vector comprising of 1.33 and 1.07 solving we get h0 = 0.51 and h1 = 0.35.

(Refer Slide Time: 22:17)



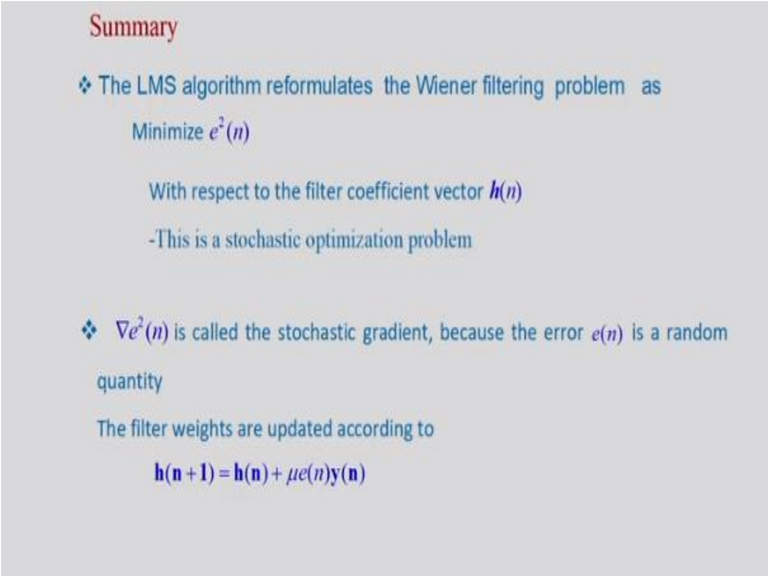
Now let us see how to get the LMS solution so we have to get dn in this diagram so that way dn is a delayed version of xn that is also input to the algorithm and RY is given by this in this case we can determine what should be the value of mu because RY is given therefore one conversion relationship is that mu which is positive should be less than 2 by Trace of RY and Trace of Ry = 2.255 +2.255 so that way it will be 4.51 and therefore mu = one third is a choice for the SDA because here we have not yet discussed how to choose mu for the LMS algorithm.

(Refer Slide Time: 23:20)



Now the LMS iteration will start length 2 FIR filter hn that is h0n + h1n transpose these are the filter coefficient and the data vector yn is given by this 2 component vector comprising of yn and yn – 1.Then the output of the adaptive filter d hat n is given by h transpose n into yn that we can compute the algorithm will compute and en is given by dn - d hat n dn already we know d hat n we have computed here therefore we can compute en. The LMS recursion is now given by h of n + 1 = hn we should have defined here + mu times en into yn; in this way, we can update the filter coefficient.

(Refer Slide Time: 24:27)



Let us summarize the lecture the LMS algorithm reformulate the Wiener filtering problem as minimize e square n instead of E of e square n it is minimize e square n with respect to the filter coefficient vector hn. So this is a stochastic optimization problem gradient of e square n is called stochastic gradient because the error en is a random quantity. The filter weights are updated according to the relation h of n + 1 = hn + mu times en into yn where yn is the data vector and en is the filtering error. We introduced the LMS algorithm it is a very simple algorithm, but the convergence analysis of this algorithm is quite difficult. We will discuss this in the next lecture. Thank you.