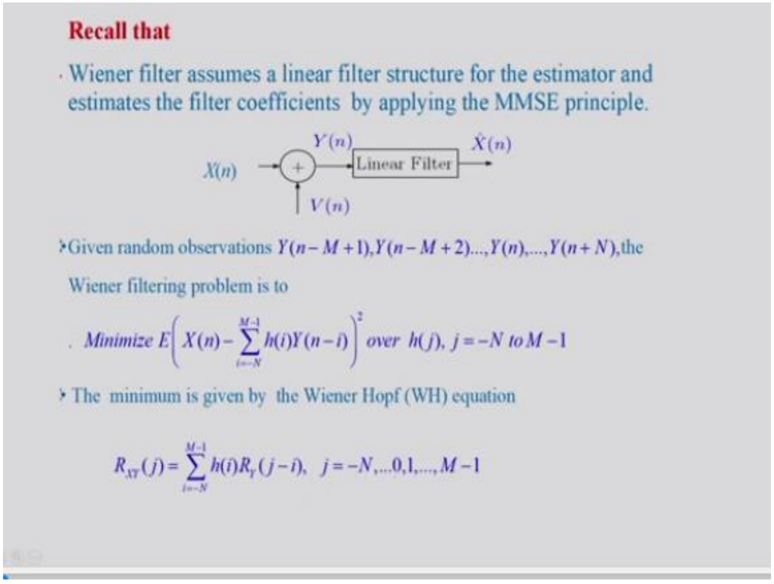
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology - Guwahati

Lecture – 18
FIR Wiener Filter

(Refer Slide Time: 00:41)

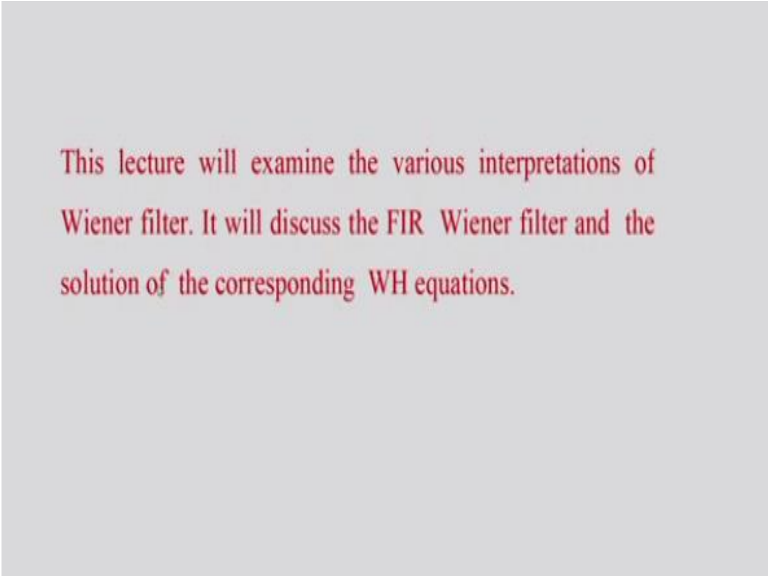


Hello students, welcome to this lecture on Wiener filter, recall that Wiener filter assumes a linear filter structure for the estimator and estimates the filter coefficients by applying the MMSE principle, this is the signal model; Yn is equal to Xn + Vn, then it is pass through a linear filter to obtain the estimate X hat n and we have to determine the parameters of this linear filter.

Given the random observations Y n - M + 1, Y n - M + 2 up to Yn, then up to Yn + N, the Wiener filtering problem is to minimize the mean square error, this is the mean square error over all filter coefficients hj; j going from - N to M – 1, this is the Wiener filtering problem and the minimum is given by the Wiener Hopf equation that we established in the last lecture.

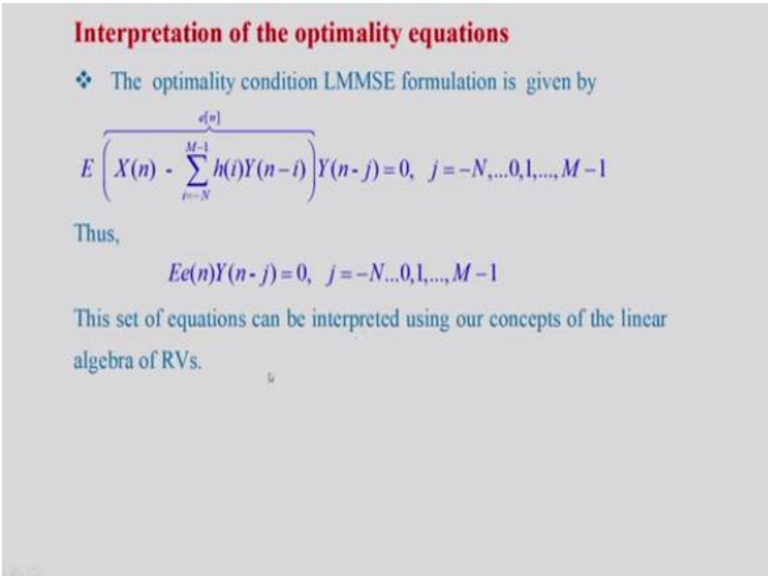
This is the relationship; Rxyj that is the cross correlation is given by the convolution of the impulse response sequence that is the filter coefficients and the autocorrelation function, so that way Rxyj is equal to summation hi into Ry j – i; i going from - N to M - 1 and for j equal to - N up to M – 1.

(Refer Slide Time: 02:26)



This lecture will examine the various interpretations of Wiener filter; it will discuss the FIR Wiener filter and the solution of the corresponding Wiener Hopf equations.

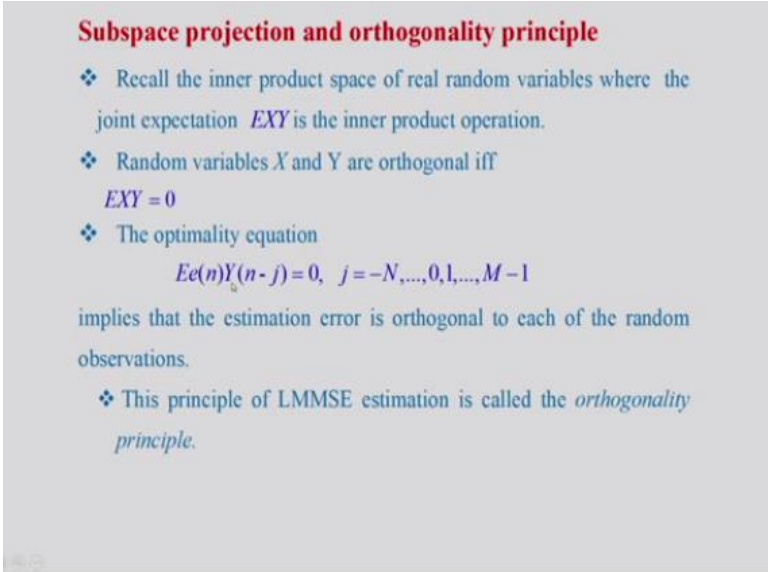
(Refer Slide Time: 02:40)



Let us interpret the optimality conditions for Wiener's filter; the optimality condition for LMMSE formulation is given by this expected value of this error into Y n - j that is equal to 0, for j is equal to – N up to M – 1, thus we can write, this is the error part; E of en into y n - j is equal to 0 for j going from - N to M – 1. Now, this set of equations can be interpreted using our concepts of linear algebra of random variables.

V; so depth are the random variables can be interpreted as vectors in a vector space, so those interpretations will help us.

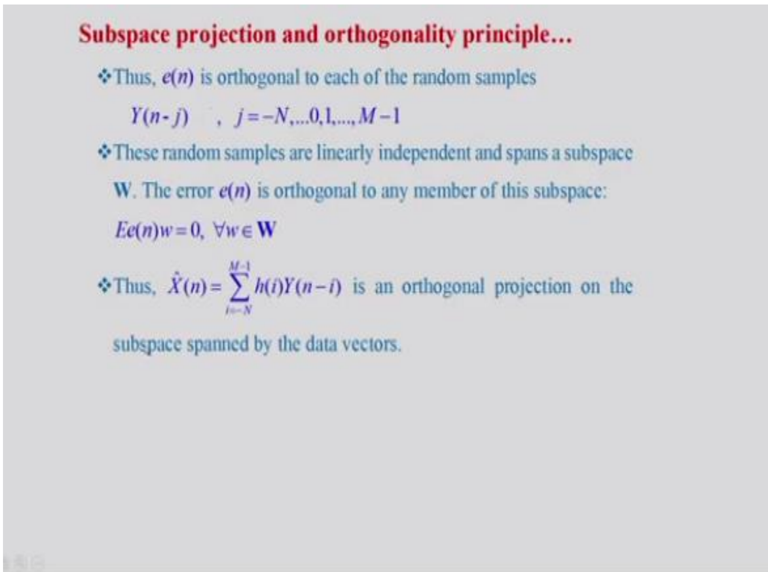
(Refer Slide Time: 03:41)



Recall the inner product space of real random variables, where the joint expectation EXY is the inner product operation, EXY defines an inner product, random variables X and Y are orthogonal if and only if E of XY is equal to 0, this is the definition. Now, the optimality equation E of en into Y n - j equal to 0, implies that the estimation error is orthogonal to each of the random observations.

For example, e of n is orthogonal to Y of n - 1 etc., this principle of LMMSE estimation is called the orthogonality principle that is error is orthogonal to each of the observation random variable.

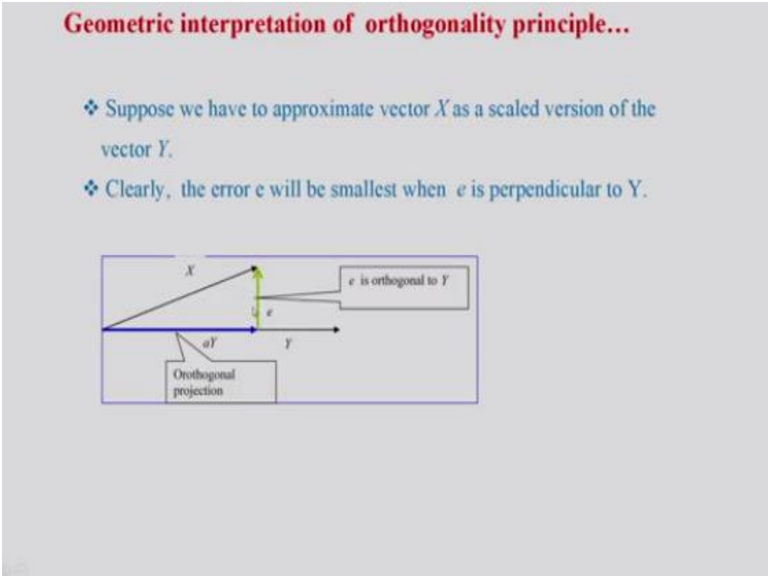
(Refer Slide Time: 04:44)



Thus en is orthogonal to each of the random samples Y n – j; j going from – N to M – 1, this random samples are linearly independent and spans a subspace W, all the random samples when say subspace W, the error en is orthogonal to any member of this subspace, E of en into W will be equal to 0 for all W belonging to depth subspace capital W, thus X hat n which is given by this expression that is linear combination of Y n - i is an orthogonal projection on the subspace spanned by the data vector.

So, this is important observation X hat n that is the Wiener estimation is an orthogonal projection on the subspace spanned by the data vectors.

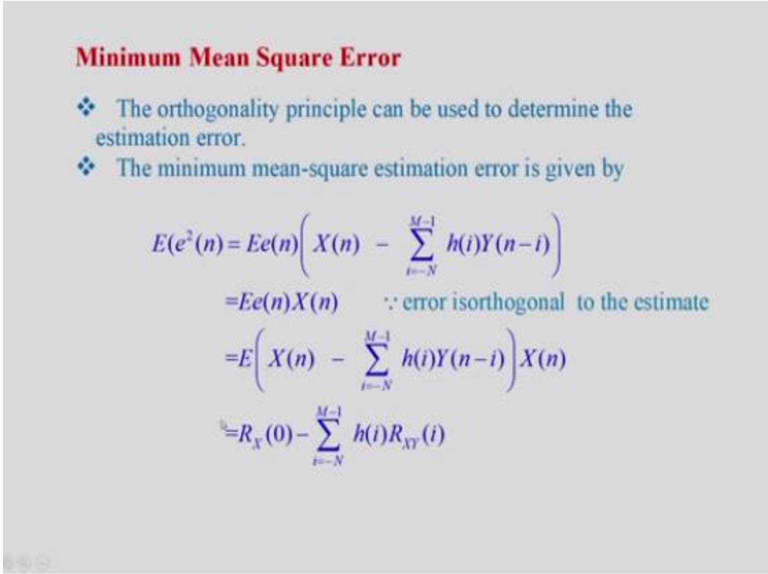
(Refer Slide Time: 05:52)



Let us try to have the geometric interpretation of orthogonality principle; suppose we have to approximate a vector X as a scaled version of another vector Y, so this is a vector Y, this is vector X, we want to approximate X as the scaled version of Y that is we have to find aY, where a is a scalar. Now, if we drop a perpendicular, then this part will be the orthogonal projection and this error will be minimum.

So, the error e will be smallest, when e is perpendicular to Y, so that is the orthogonality principle, so we have to select this scalar multiple of Y in such a way that error is orthogonal to the data vector, this is the orthogonality principle.

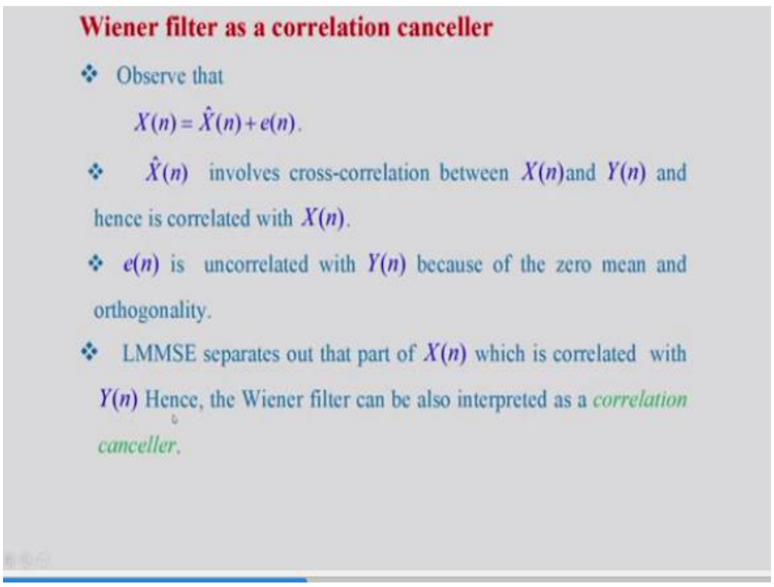
(Refer Slide Time: 06:51)



Now, this orthogonality principle can be used to obtain the estimation error because the minimum mean square error estimation error is given by this expression; E of e square n, then that we can write E of en into en which can be expressed like this, Xn minus summation hi Y n – i; i going from - N to M – 1. Now, note that this part is orthogonal to error, so error is orthogonal to the estimate, this estimate.

Therefore, this will be simply E of en into Xn now, if I expand this en that is Xn - summation hi Y n – i; i going from – N to M – 1, then we get this will be E of and this expression into Xn and if I write now, Xn into Xn, if we take the expected value we will get Rx of 0, similarly we will get the cross correlation function here by taking the expected values, so that way E of e square n will be equal to Rx of 0 - summation hi Rxy i; i going from – N to M – 1.

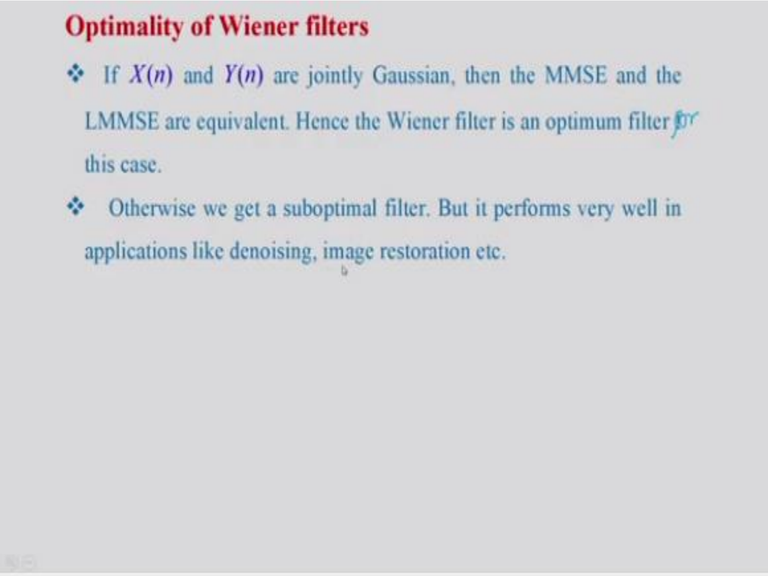
(Refer Slide Time: 08:22)



So, the orthogonality principle can be used to derive the mean square error of estimation, we will interpret Wiener filter also as a correlation canceller, observe depth Xn that is the unknown signal is the estimated signal plus error, X hat n plus en now, X hat n involves cross correlation between Xn, Yn and hence X hat n will be correlated with Xn, en is uncorrelated with Yn because of the zero mean and orthogonality.

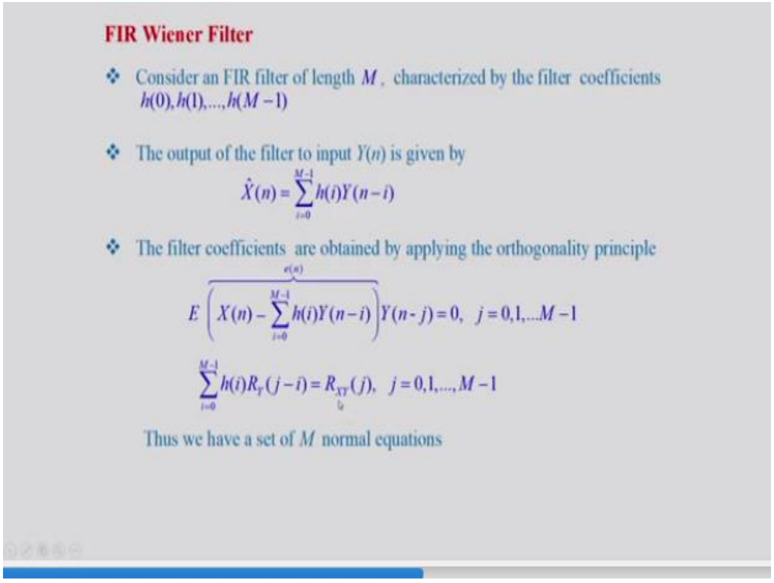
We know that en is orthogonal to Yn but Yn is 0 mean and consequently, en is also 0 mean, therefore orthogonal and 0 mean means that en is uncorrelated with Yn, so this is very important that en is uncorrelated with Yn, therefore LMMSE separates out a part of Xn which is correlated with Yn, hence the Wiener filter can also be interpreted as a correlation canceller, whatever correlated part is there in the signal that will be cancelled out.

(Refer Slide Time: 09:42)



Now, question is whether the Wiener filter estimator is an optimum estimator, if Xn and Yn are jointly Gaussian, then the MMSE and the linear MMSE are equivalent because that we have established in earlier lecture, hence the Wiener filter is optimum for the Gaussian case, otherwise we get a suboptimal filter but it performs very well in applications like denoising, image restoration etc., though it is a suboptimal filter because of its simplicity, it is very useful in practical applications.

(Refer Slide Time: 10:27)

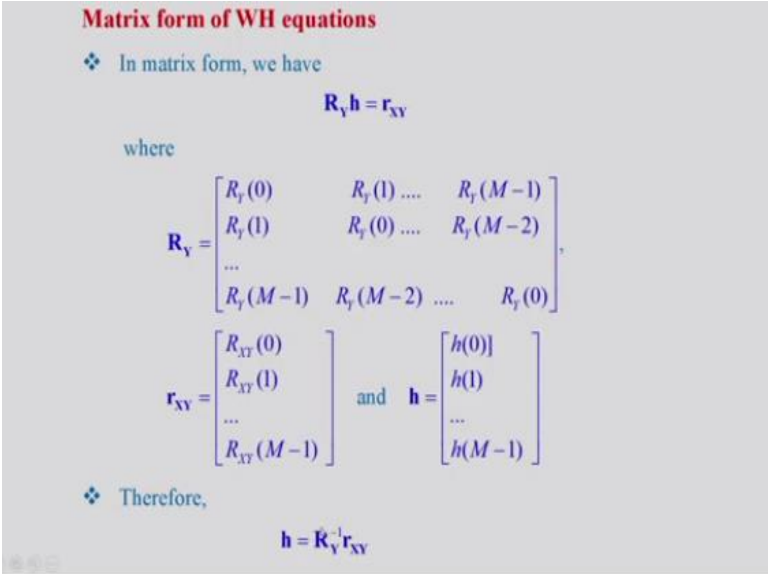


Now, we will discuss FIR Wiener filter. Consider an FIR filter of length M characterized by the filter coefficients h0, h1 up to h M – 1, there are M coefficients, so that is a FIR filter of length M. The output of the filter to input Yn is given by this relationship convolution relationship, X hat n is equal to summation hi Y n – i; i going from 0 to M – 1, this is the filter output.

And now, we have to find out these filter coefficients; the filter coefficients are obtained by applying the orthogonality principle, what is that; error is orthogonal to data, so E of Xn minus summation hi Y n – i; i going from 0 to M - 1 that is the error into Y of n - j will be equal to 0 for j is equal to 0, 1 up to M – 1, so this will give now the Wiener Hopf equations that is if we take the expectation now, we will get summation i going from 0 to M - 1 of hi into Ry j - i will be equal to Rxy j, where j is equal to 0, 1 up to M – 1.

So, these sets of equations are the M normal equations or Wiener Hopf equations and we have to solve these equations to find out the filter parameters.

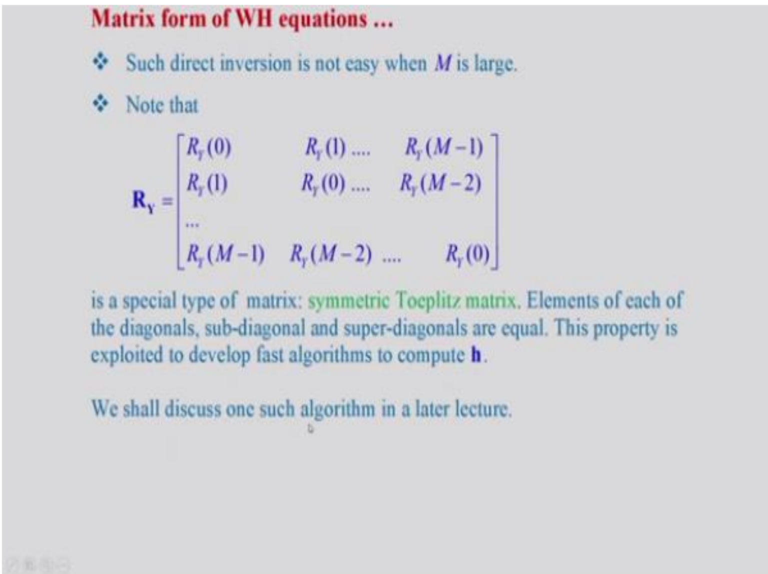
(Refer Slide Time: 12:19)



In the matrix form, we have R1h is equal to small rxy, this is the autocorrelation matrix of Y and this is the filter vector, this is the cross covariance vector, thus Ry is equal to this matrix, Ry0, Ry1, etc., up to Ry M - 1 similarly, second row will be Ry1, Ry0 up to Ry M - 2 and so on similarly, rxy is the cross covariance vector and its elements are Rxy of 0, Rxy of 1 and so on up to Rxy of M – 1.

And this is the vector representing the filter coefficients, there are M filter coefficients, so these M filter coefficients are represented by this vector, we can find h by inverting this relationship Ryh is equal to small rxy, therefore h will be Ry inverse into small rxy vector, so this inverse matrix we have to find out.

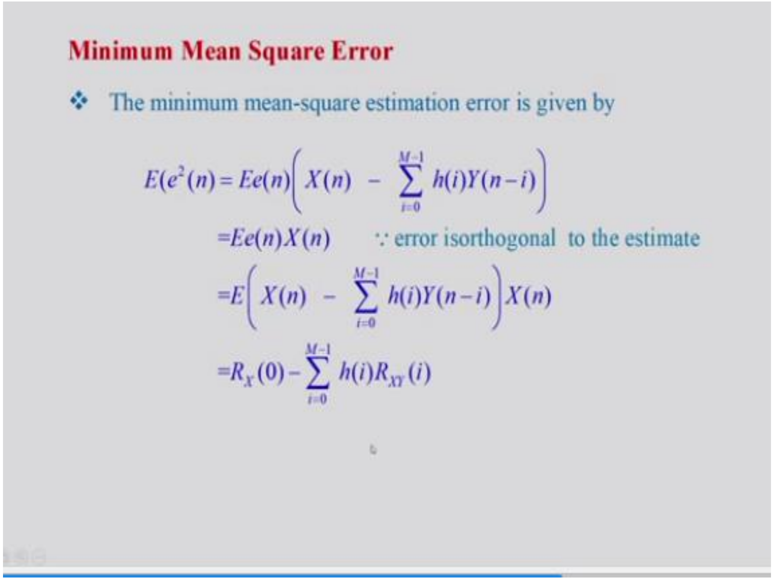
(Refer Slide Time: 13:34)



Such direct inversion is not easy when M is large because I know that Ry is an M by M matrix, so if M is large, inversion will be very complicated and now, let us examine this Ry matrix we see that all its diagonal elements are Ry0 similarly, this super diagonal elements will be Ry1, all elements will be Ry1 similarly, this is also all elements along this sub diagonal will be also Ry1.

So, that way, Ry is a special type of matrix what is known as the symmetric Toeplitz matrix, elements of each diagonal is same and elements of each super diagonal is same like that elements of each sub diagonal will be also same, this is not only Toeplitz but it is symmetric also, this one is Ry of 1, here also Ry of 1, so that way this is a symmetric Toeplitz matrix and this property is exploited to develop fast algorithm to compute h.

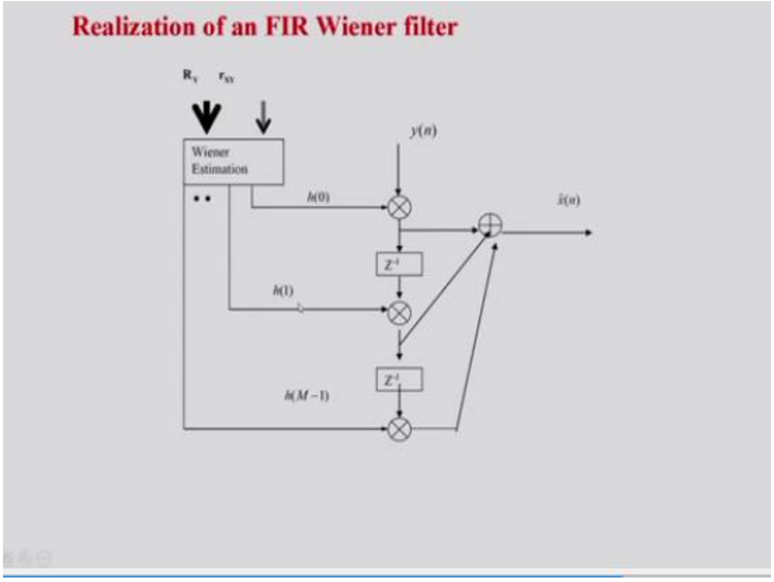
(Refer Slide Time: 14:47)



We shall discuss one such algorithm in a later lecture, now this minimum mean square error of FIR filter, so that we have to find out, the minimum mean square estimation error is given by this expression; E of e square n that is E of en into X n minus this estimate and using the orthogonality principle as I shown earlier, it will be E of en into Xn, so therefore the mean square estimation error is equal to E of en into Xn.

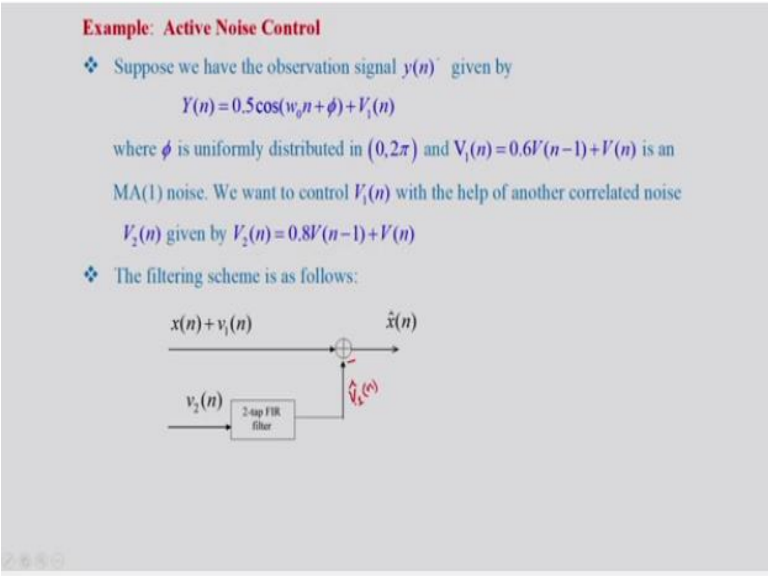
Now, I can expand en that is Xn minus summation hi into Y n – i; i going from 0 to M - 1 and then multiplied by Xn, if we simplify this will become Rx of 0 minus summation hi into Rxy i; i going from 0 to M – 1, this is the expression for mean square estimation error.

(Refer Slide Time: 15:48)



The FIR Wiener filter can be realized as follows, we have the autocorrelation matrix Ry and the cross correlation vector Rxy as input to the Wiener filter estimator and the estimator will compute the filter coefficients, h0, h1 up to h M – 1, this is the input signal yn, so this input signal yn will be multiplied by h0, then the delayed part will be multiplied by h1 and so on and all these outputs will be summed up and we will get X hat n.

(Refer Slide Time: 16:34)

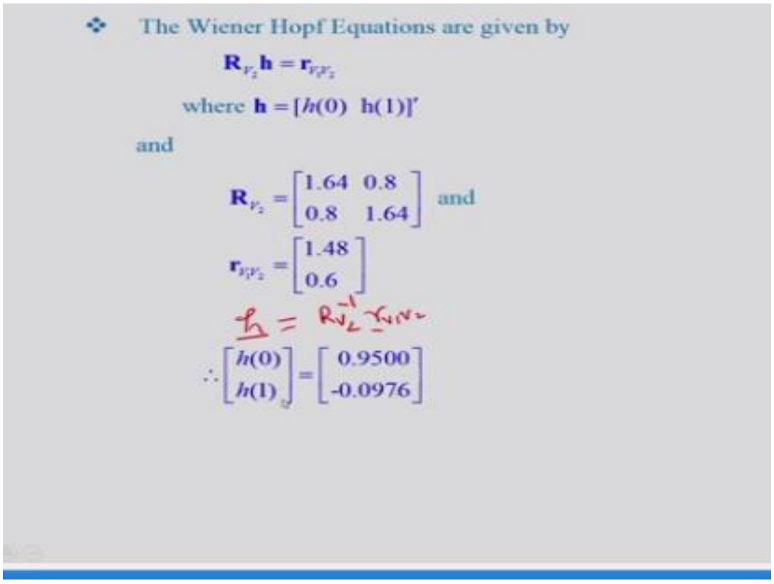


So, this is the realization of FIR Wiener filter, let us consider one example, this is a very important application what is known as the active noise control. So, if we have to cancel the noise in a recording room, then this active noise control can be applied. Suppose, we have the observation signal yn that is given by Yn is equal to 0.5 cos omega 0 n + pi + V1n, there is a noise in the signal, where pi is uniformly distributed in the interval 0 to 2 pi.

And V1 n is equal to 0.6 Vn - 1 + Vn, so this is known as a moving average model, MA 1 noise, so the signal Yn is a cosine function plus some noise, we want to control V1 n with the help of another correlated noise V2 n which is given by V2 n is equal to 0.8 Vn - 1 + Vn this is the V2n; V2n and V1n are correlated noise now, using the 2 tap FIR Wiener filter we will try to estimate V1 and from this input because they are correlated, therefore we can estimate V1 hat n here.

And this V1 hat and that estimated value that will be added here, so that it is positive and this is negative and this part will be cancel by this V1 hat and this is the active noise control, so that ultimately we will get the noise free part of this noisy signal, so this is the filtering scheme.

(Refer Slide Time: 18:33)

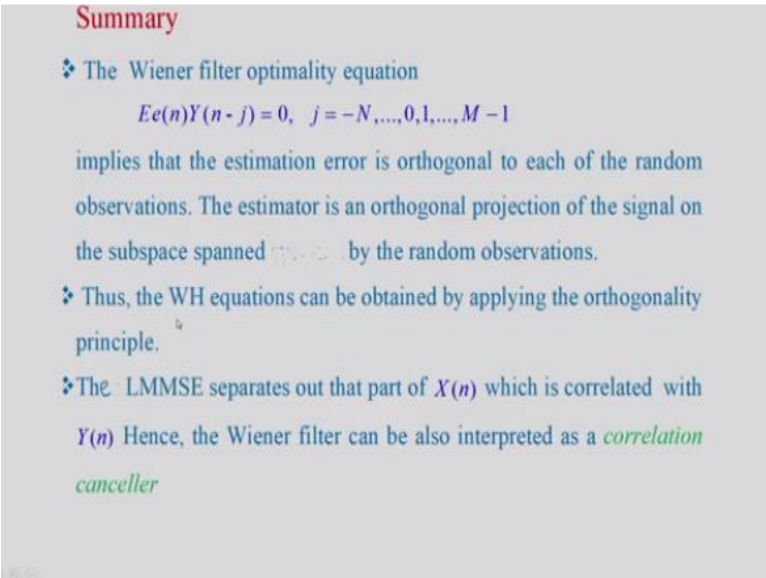


And we can find the Wiener Hopf equations that is given by Rv2 into h that is the autocorrelation matrix of v2 into h is equal to cross correlation vector rv1v2, where h is 2 length filter, so that way it is h is equal to the vector composing of h0 and h1 now, we can compute Rv2 from the model and these are given by 1.64, 0.8, 0.8, 1.64, this is the autocorrelation matrix of v2.

And similarly, rv1v2, this vector also we can compute from the model and rv1v2 is the vector comprising of 1.48 and 0.6, therefore we can find out now what will be h; h is equal to Rv2 inverse into small r v1 v2 okay, so if we apply inverse this matrix and then multiply with this and we will get this result that is h0 is equal to 0.95 and h1 is equal to – 0.976. So, now these are the filter coefficient corresponding to these 2 tap FIR filter.

And when we filter this noise sequence with this filter, we will get V1 hat n and this will cancel out V1 n, so that way we get the noise cancel output.

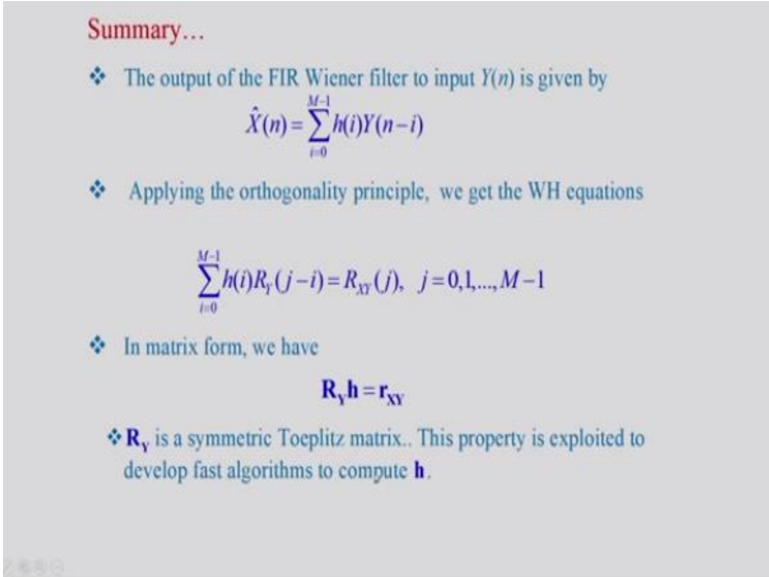
(Refer Slide Time: 20:14)



Let us summarize the lecture; the Wiener filter optimality equations; E of en into Y n - j is equal to 0 for j going from - N to M – 1, these are the optimality equations for the general Wiener filter we have considered and this optimality conditions imply that the estimation error is orthogonal to each of the random observations. The estimator is an orthogonal projection of the signal on the subspace spanned by the random observations.

Thus the Wiener Hopf equations can be obtained by applying the orthogonality principle, we have also seen that the LMMSE separates out that part of Xn which is correlated with Yn, hence the Wiener filter can be interpreted as a correlation canceller.

(Refer Slide Time: 21:19)



The output of the FIR Wiener filter to input Yn is given by this and this filter parameters or filter coefficients we have to find out using the orthogonality principle, applying the orthogonality principle we get the Wiener Hopf equations like this summation hi into Ry j – i; i going from 0 to M - 1 is equal to Rxy j for j is equal to 0 to M – 1, M equations are there and there are M unknowns.

Therefore, we can solve this set of equations to get the filter coefficients, in the matrix form we have Ryh is equal to rxy that is autocorrelation matrix into h vector is equal to cross correlation vector and we also observed that Ry is a symmetric Toeoplitz matrix and this property is exploited to develop past algorithms to compute h, thank you.