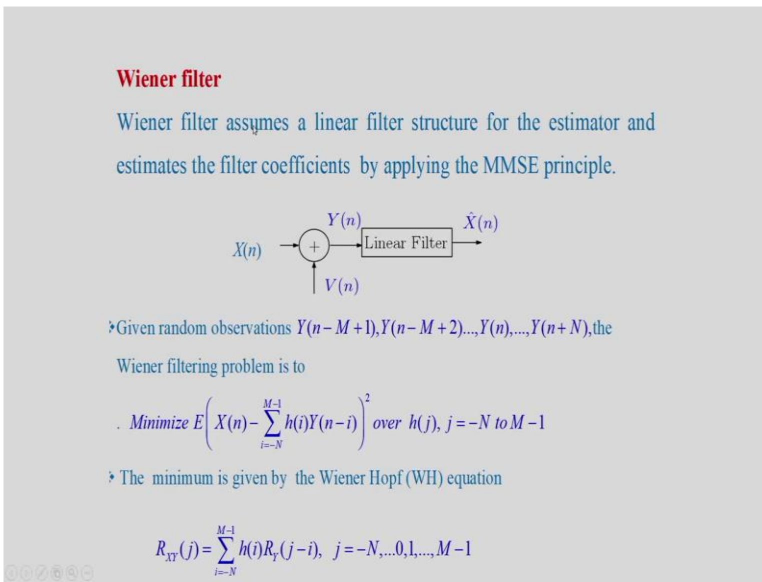
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology – Guwahati

Lecture - 39
Review 2

Hello students will to this review lecture to in this lecture I will briefly review the signal estimation by optimal linear filters. We discussed optimal linear filters then adaptive implementation adaptive filters and Kalman filters, we will briefly discuss those.

(Refer Slide Time: 01:02)



Wiener filter assumes a linear filter structure for the estimator and estimate the filter coefficients by applying the minimum mean square error principle. So, Wiener filter already is a linear filter structure this is Xn is the signal Vn is some additive white noise Yn is the observed signal it is passed through this linear filter we get the estimated signal. This estimated signal should be optimum under the minimum mean square error criterion with this linear filter structure.

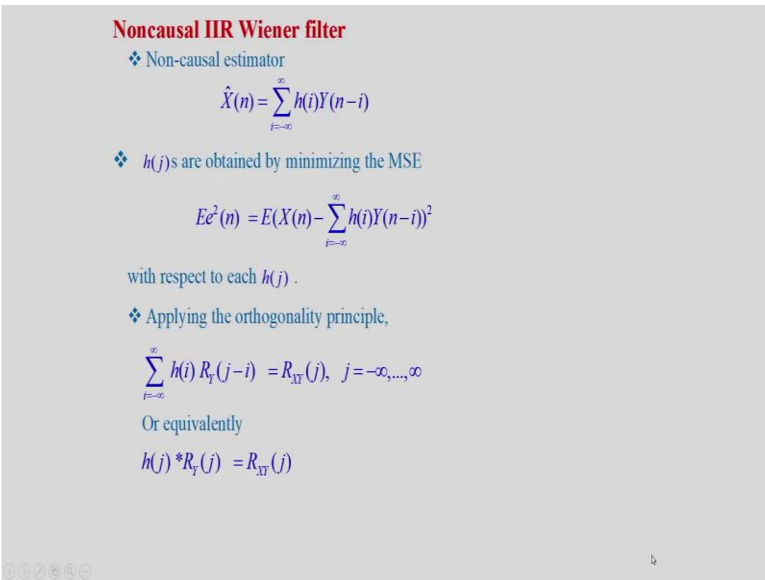
Given random observations Y n - M + 1 Yn - M + 2 up to Yn and then upto Y n + N so we have a block of signal from time instant n - M + 1 to n + N. So, for this duration signal is available our minimization problem will be now minimize E of Xn - this filter output that is summation hi into Y n - i Yn is input signal i going from - n to M - 1 whole Square this we have to minimize over hj j going from - n to M - 1 and we derived a Wiener Hopf equation corresponding to this if we differentiate with respect to E sub h i it is E sub hj we get Weiner Hopf equations given by Rxy j is equal to summation hi into R Y j - i i going from - n to M – 1.

We discussed the FIR Weiner filter FIR Wiener filter here we have to minimize filter is X hat n is equal to summation hi into Y n - i i going from 0 to M - 1 this is the M type a fire filter we considered and the filter parameters were obtained by applying the orthogonality principle what was that E of signal minus the filtered output that is hi into Y n - i i going from 0 to M – 1, so this is the error, this error is orthogonal to data the Y of n - j that is equal to 0.

So from this, this is the orthogonality condition from this we will get that R XY of j is equal to summation hi into R Y of j - i i doing from 0 to M - 1 so R X sub j is equal to summation hi into R Y of j - i i going from 0 to M - 1 that is true for j is equal to 0, 1 up to M - 1 because these are our data point. So, that way this Weiner Hopf equation we derived for a fire filter. And similarly minimum mean squared error also we can find out MMSE minimum minimum mean square you know that is given by E of e square n and this is same as E of X n - summation hi Y n - i i going from 0 to n – 1 this is the error into X of n.

So because remaining data are orthogonal to this error so that way we got it as R X of 0 - summation hi into R Y of i i doing from 0 to M - 1 so this is the minimum mean square error.

(Refer Slide Time: 05:55)

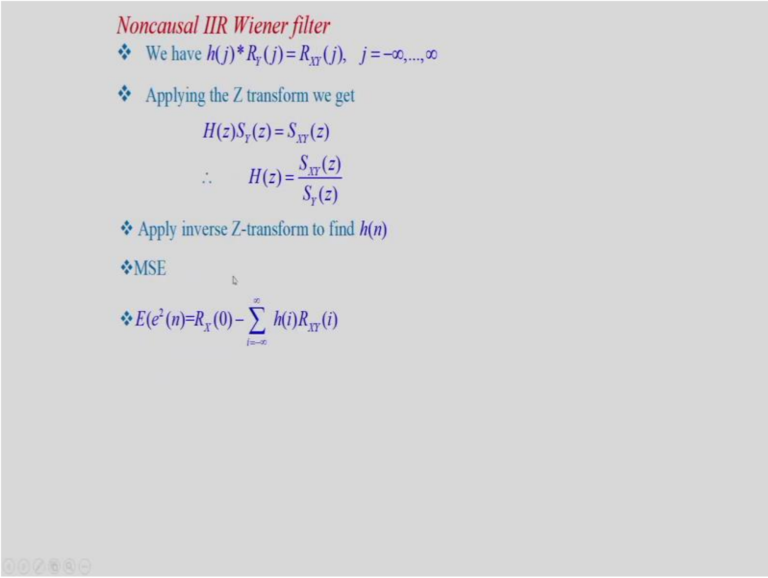


We also considered it non-causal I our Wiener filter in that case impulse response sequence is infinite therefore X of n is equal to summation hi into Y n - i i going from minus infinity to plus infinity. So, in this case the mean square error will be E of Xn - summation Delta into Y n - i i going from minus infinity to plus infinity whole Square and this mean square error we have to minimize with respect to each hj.

Again we apply the orthogonality principle, what is orthogonality principle? Error is orthogonal to data E of Xn - summation hi into Y n - i i going from minus infinity to plus infinity this is the error is orthogonal to Delta Y of n - j j going from minus infinity to plus infinity so this must be equal to 0. So, this is the orthogonality condition and this orthogonality condition will give us that is R XY of the j E of Xn into Y n- j that will be R XY of j that will be equal to this summation hi into Eof Y n - i into Y n - j that will be R Y of j – i.

So that we summation hi into R Y of j - i i going from minus infinity to infinity that will be equal to R XY of j this is the Wiener Hopf equation which is obtained by applying the orthogonality principle and here we have two sequences hi is known to be a two-sided sequence because impulse response is from minus infinity to plus infinity and autocorrelation function is also a two-sided sequence therefore we can perform convolution and then solve this problem in that from domain.

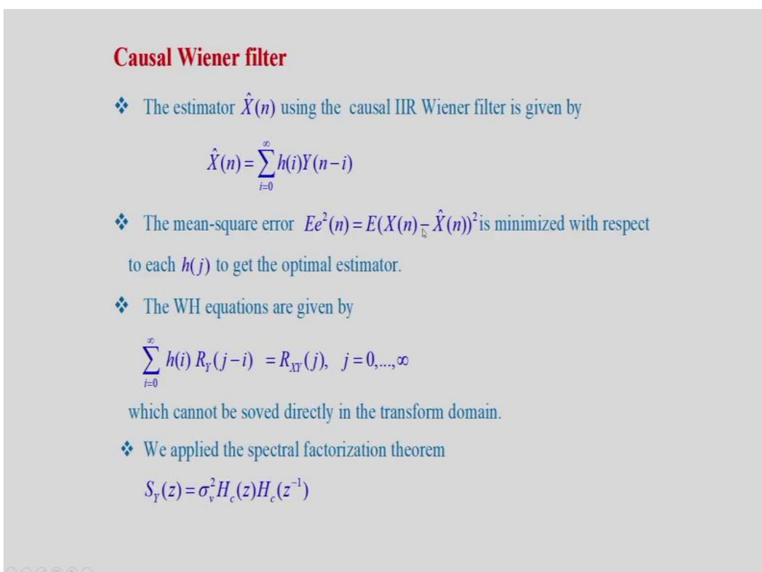
(Refer Slide Time: 08:17)



So noncausal IIR Wiener filter is therefore given by hj controlled with R Y j is equal to R XY j j are going from minus infinity to plus infinity then applying that transform we will get this is that H sequence transfer function is Hz here S yz is equal to S XY z, so from this we get that it transfer function of the noncausal Wiener filter is given by Hz is equal to S XY z divided by S Y z that cross power spectral density divided by the power spectral density of Y.

And once we have Hz we can apply inverse transform to point hn. Now how to find out the MSE again same orthogonality principle we can apply that is E of MSE is equal to E of en is Xn - summation hi into Y n - i i going from minus infinity through infinity into Xn because remaining part of the error will be orthogonal to this part so because the remaining that apart will be their Z part will be orthogonal to this error therefore only this expression will come and that is equal to R X of 0 first term and remaining term will be summation hi into R XY i i going in from minus infinity to infinity that is R X of 0 - summation hi R XY cross-correlation i going from minus infinity to plus infinity.

(Refer Slide Time: 10:22)

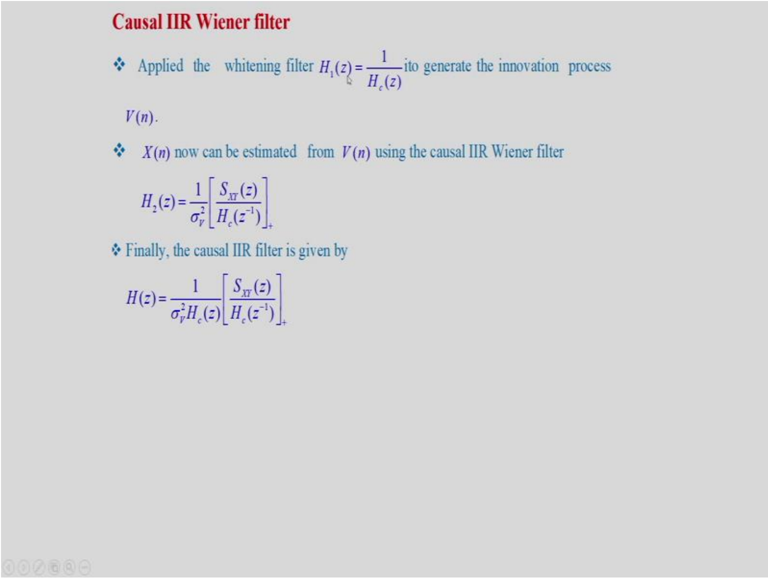


Then we discussed Causal Weiner IIR filter in this case the estimator X of n is given by X of n is equal to summation hi into Y n - i i going from 0 to infinity because the filter is causal this filter sequence is defined from i is equal to 0 to infinity so that way it is causal. And the mean square error is given by E of e square n that is E of Xn - X hat n whole square and this is to be minimized with respect to all hi’s.

So Weiner Hopf equation again here is given by E of error is orthogonal to data that is Xn - summation hi Yn - i i going from 0 to infinity into error is orthogonal to data Y of n - j j is equal to 0, 1 up to infinity so that way we will get first term will be R XY j and then remaining term will be summation hi i going from 0 to infinity into R Y j – i. So, if I take the difference between this and this will get j – I, so that was same as hi into R Y j - i i going from 0 to infinity is equal to R XY j. So, this is the winner Hopf equation corresponding to Causal IIR Weiner filter.

So now you know that here is hj j is equal to 0 to infinity is a right-sided or causal sequence but RY here is 2 sided sequence RY is two sided sequence therefore direct convolution is not possible so therefore which cannot be solved directly in the transform domain. And now we will apply the spectral factorization theorem to the power spectral density of Y that is Sy z is equal to sigma V square into Hc z into Hc z inverse so that way we will use this relationship to get it whitening filter.

(Refer Slide Time: 13:28)



So, that we apply the whitening filter H on z is equal to 1 by Hz to generate the innovation process so we apply suppose this is my Yn pass it through 1 by Hc z that is the innovation filter or whitening filter to get the innovation sequence. And now using this innovation sequence Xn now can be estimated using Vn and this was shown this and the corresponding filter is given by

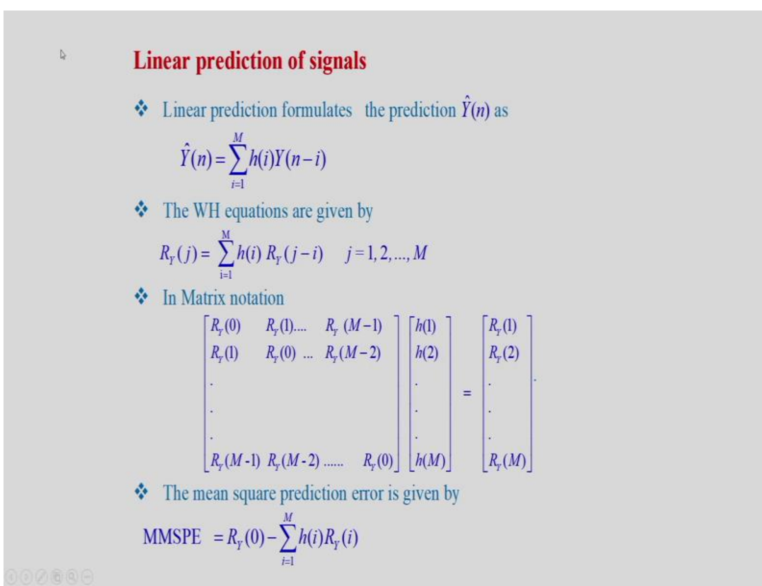
Hz is equal to 1 by Sigma v square into the causal part of S XY z divided by Hc z inversion this is the Wiener filter to estimate Xn from the white noise sequence.

And finally we got the Causal IIR Wiener filter as the Cascade of this filter 1 by Hc z and then into H2 z, so Vn is passed through is 2 n this is P n Vn is passed through this H2 z that is the filter we have already obtained then we will get this estimator Hz and that way it is cascade of two filters one is 1by Hc z and of course this sigma v square is also there 1 by sigma v square into; so, spectral factorization gives the whitening filter H1 z is equal to 1 by Hc z. And this generates innovation sequence Vn's of variance sigma v square.

And this innovation sequence is used or used to estimate X of n from Vn so that way the transfer function of this filter H2 z which is optimum is given by this expression 1 by sigma v square into the causal part of S XY z divided by Hc z inverse. So, this is the Weiner estimator of Xn from the Vn sequence. Now once we have this Hz part we can cascade it to 1 by Hc z to get the combined Wiener filter.

Therefore Weiner filter will be Cascade of this and this and it is given by 1 by sigma v square Hc z into S XY z divided by Hc z inverse and then we have to take the causal part.

(Refer Slide Time: 16:37)



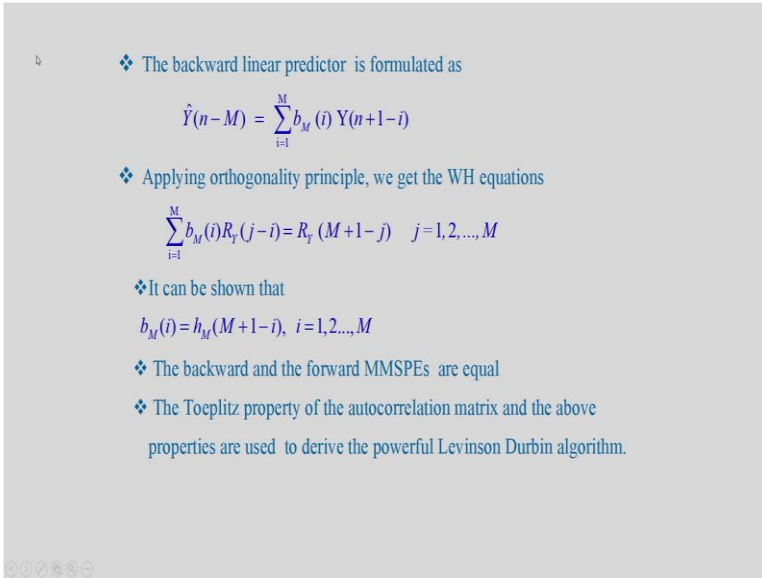
We considered one application that is linear prediction of signals. The linear prediction problem was formulated as follows Y hat n is equal to summation hi into Yn - i i going from 1 to M so this is the linear prediction or it is forward prediction problem. Again we apply the orthogonality condition that is E of here actual data is Yn and we are predicting Y hat n that is orthogonal this is the error is orthogonal to date the data is y of n - j because it is M they are going from 1 to M because we are using only past data up to n – M.

So that way from n - 1 to n - M therefore this is the orthogonality relation we get here yn - summation Y hat n into Y n - j is equal to 0 you know it is orthogonal to data and from that we directly get the Weiner Hopf equation or normal equation and this is given by RY j is equal to summation hi into RY j - i i going from 1 to M and this is there are M equations corresponding to M values of j. And in matrix notation this equation we write like this.

This is the autocorrelation matrix into coefficient vector is equal to the autocorrelation vector. So, this autocorrelation matrix we see that it has very attractive property it is symmetric and it is topless. For example this sub diagonal will be same this time it will be same like that, so this is a symmetric topless matrix into A specter is equal to R Y vector. And similarly the mean square prediction error also can be obtained by applying the orthogonality principle that is MMSPE minimum mean square prediction error is given right E of that is error is Yn - summation hi into Y n - i i going from 1 to M into yn.

And the remaining part of error that that will include yn - 1 yn - 2 etcetera they are orthogonal to this error therefore their contribution will become 0 and we get this on the minimum mean square prediction error is equal to E of Yn - summation hi into Yn - i i going from 1 to M into y n so remaining term will contribute 0. So, that way it will be equal to R Y 0 - summation hi into RY of i i going from 1 to M - 1 so this is the expression for minimum mean square prediction error.

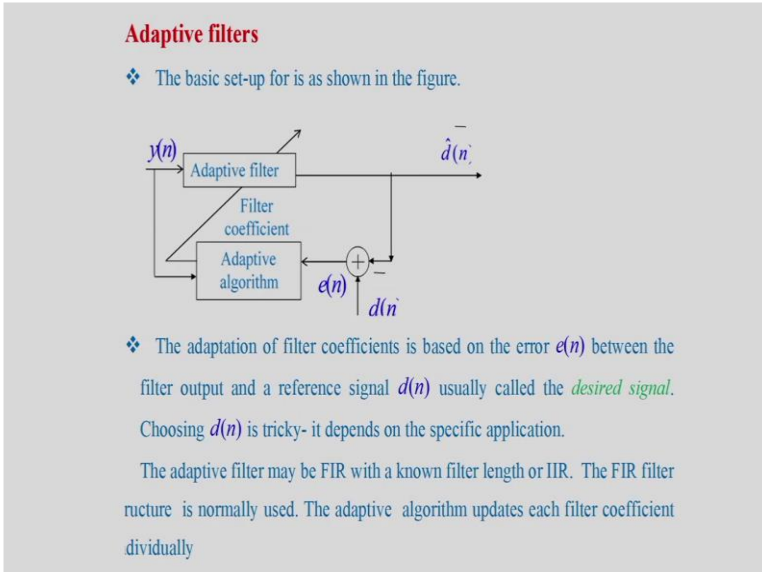
(Refer Slide Time: 20:16)



Similarly we also discussed the backward prediction problem here given Y n Yn - 1 up to Y n - m + 1 we predict Y hat n - M so Y hat n - M is summation b Mi this is the backward prediction coefficient into Y of n + 1 - i i going from 1 to M this is the backward prediction problem and here also we apply the orthogonality again error is orthogonal to data and that we applied there are M data y n yn - 1 up to Y n - 1 + M so that way M that are there and corresponding to that we will have this Weiner Hopf equations.

And from this Wiener Hopf equation and linear prediction Weiner Hopf equation or forward prediction Weiner Hopf establish an important result that is backward prediction coefficient at instant i is same as forward prediction coefficient h M M + 1 - i so this is the relationship and also we established at backward and forward minimum in mean square prediction errors are equal. So, we use this result and also this symmetric topless structure of this matrix to derive the famous Levinson Durbin algorithm that we have discussed in detail.

(Refer Slide Time: 21:58)

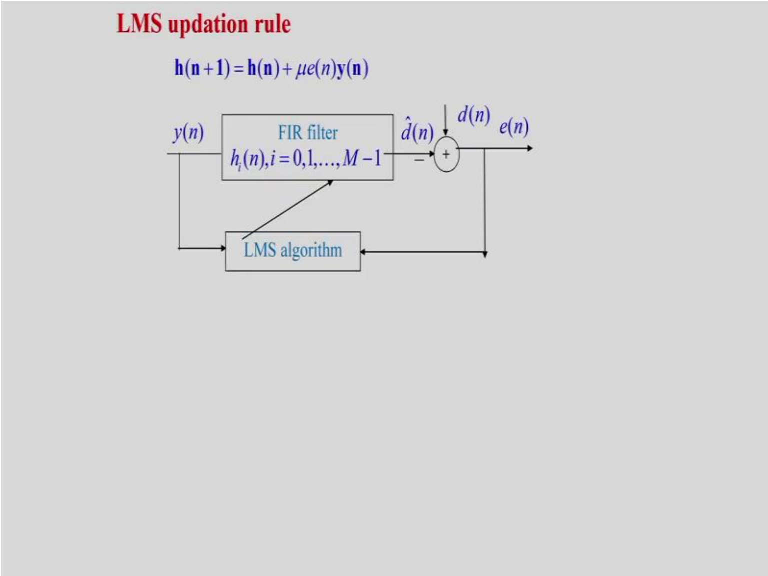


We also discussed adaptive filter the basic setup for the adaptive filter is like this, this is Yn is the input and this, this filter structure generally it is a FIR filter structure and this is the absorbed signal and we get some output here that we call it as d hat n. And there is an reference signal in the case of adaptive filter this is dn the we want to match the output of this filter with this dn so that way the dn is the desired output which we know somehow and we want to match the output of the adaptive filter with this dn.

And whenever there is no matching that error will be there that error is paid back to update the adaptive algorithm which will compute the filter coefficients adaptively. So, this adaptive filter coefficients will be computed by this adaptive algorithm based on this error en and the input signal Yn. And then this updated filter will be used again to filter this yn and to be the estimate of the desired signal.

The adaptation of the filter coefficients is based on the error en between the filter output and a reference signal dn usually call the desired signal, so the in dn is tricky it depends on this specific application. For example we showed how we can choose dn in the case of adaptive channel equalization and adaptive system identification.

(Refer Slide Time: 24:01)



We considered one basic adaptive filter what is known as LMS adaptive filter and it is updating rule was this h of n + 1 is equal to hn + mu into en into yn we get this by applying this steepest descent algorithm to minimize this cost function e square n, so minimize e square n with respect to hi n n is equal to 0, 1 up to M - 1 so this optimization problem when we apply this steepest descent algorithm we will get this updating rule.

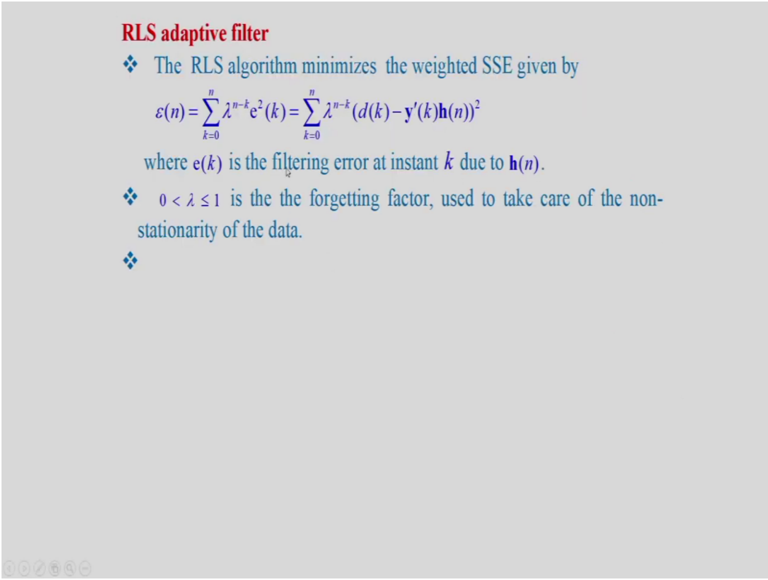
So this is the LMS updating rule so yn is the input and this is the FIR filter and then d hat n is the output and this is compared to digital signal output. This is compared with the desired signal and the error is fedback to the LMS algorithm. So, LMS algorithm update the filter coefficient by this rule it computes the error you know error is nothing but yn will be filtered at instant n with filter parameter of; so I will get d hat n that is summation hi n into yn - i i going from 0 to n - 1 so we filter the signal yn with the filter coefficient at instant n.

And then after that we will get the error, error is equal to en is equal to dn - d hat n, so this error is used to update the filter coefficient h of n + 1 that is obtained as that previous filter hn + mu into en into yn and we discussed several modification of this algorithm that is NLMS algorithm leaky LLMS algorithm that is leaky LMS algorithm, then for a efficient implementation will consider block LMS algorithm and sign error LMS algorithm.

The LMS is a very simple adaptive filter but problem is how to suit this step length parameter mu so this depends on the eigen value of the autocorrelation matrix. Suppose we are filtering it WSS signal Y n then the eigenvalues of the corresponding autocorrelation function of yn will determine this filter step length parameter. So, that way we establish some relationship that is mu should be lying between 0 and 2 by lambda max this type of relationship we establish and then also with this chosen mu the LMS adaptive filter will not convert to XL Wiener filter there will be always some excess error that excess error will also depend on the parameter mu.

And we also observed that the rate of convergence depends on the eigen value spread. If it is high convergence rate will be low and that was a serious problem in the case of LMS filter.

(Refer Slide Time: 28:38)



And to overcome this we consider the RLS adaptive filter. So, in the case of RLS adaptive filter the error is sum square error. We discussed the least square estimation principle same principle is applied here also and here the error to be minimized is sum square error weighted by factor lambda to the power n – k, so that way the error to be minimizes epsilon n is equal to summation lambda to the power n - k into e square k k going from 0 to n.

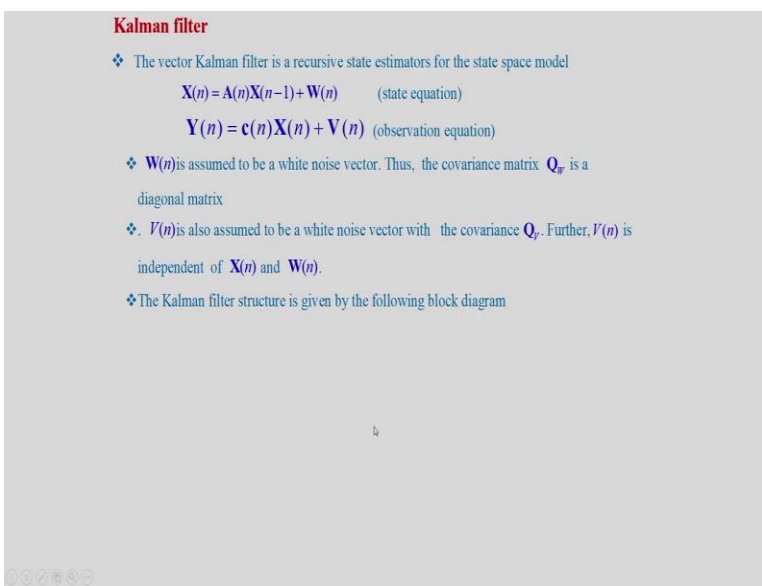
So from the beginning n is equal to 0 to current n all the errors are considered and this is given by summation lambda to the power n - k into dk - y transpose k into h n whole square k going from 0 to M and this is important that we are filtering the all the passed signal values which respect to hn current estimate of the filter and that way we are determining the error and this error are weighted by the what is known as the forgetting factor lambda to the power n - k and this lambda lies between 0 and 1.

In the case of WSS signal we can take lambda is equal to 1 and this is used to take care of non stationarity of the data. So, for example since lambda is a fraction as n tends to infinity so lambda to the power n - k for k is equal to 0 it will become lambda to the power n it will be close to 0 when n is large. So, therefore for large value of n it will give less and less weight to the previous errors that is the idea behind tackling the non stationarity of the data.

Past errors are given less with present errors are given more weight and that way we derive the RLS adaptive filter we got a corresponding normal equation by filtering that is del epsilon n into del hn that is equal to 0, so from this we got a matrix normal form of normal equation matrix form of normal equation and which is solved by the matrix inversion lemma we go get a recursive estimate or using the matrix inversion lemma.

So that way we discussed the RLS algorithm and this RLS algorithm estimates the signal estimate the signal recursively. So, it is a recursive least square method.

(Refer Slide Time: 32:02)



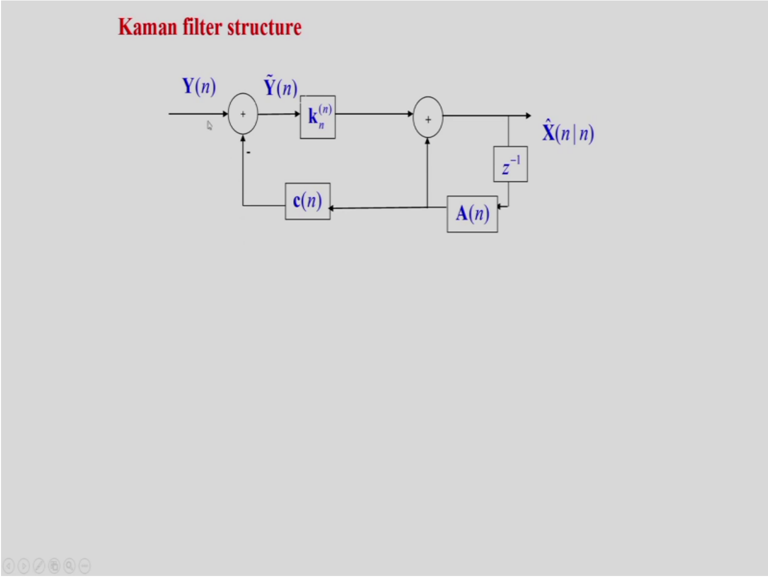
We discussed two form of Kalman filter that is scalar Kalman filter and vector Kalman filter. This scalar Kalman filter this uses the applies the model what is the model? Model is AR1 signal so Xn is equal to A times X of n - 1 + W n so W n is a white noise and A the constant which is known and it can be considered as a function of n so that non stationarity can be tackled. And observed data Yn is equal to Xn + Vn.

And we discussed that we derived it derived using the innovation representation that is get the innovation what was the innovation in Kalman filter will got it s and that is Yn prediction error minus summation hi into Yn - i i going from 0 to n - 1 so this is the prediction error and this prediction error we saw that this prediction error is an orthogonal sequence that way we generated the innovation sequence.

And we also considered the vector Kalman filter it is a recursive state estimator and uses a state space model for the signal that is Xn is equal to An into Xn - 1 + W n if we consider An to be time variant then this state space model is Xn that is the signal vector is equal to A n into X of n - 1 + Wn, Wn is noise known as process noise. And the observed data Yn is equal to Cn a matrix Cn is a matrix multiplied by Xn, Xn is the state vector + Vn.

Vn is again another noise vector which is observation noise vector which is uncorrelated with Wn and uncorrelated with Xn. So, that where Wn is assumed to be white noise vector does the covariance matrix Qw is a diagonal matrix. Vn is also assumed to be a white noise vector with covariance matrix Qv further Vn is independent of Xn and Wn. The Kalman filter structure is given by this block diagram.

(Refer Slide Time: 35:33)



So, Yn this is the input signal and this is the output signal and normally we denote it by X of n given n estimation of the signal at instant n. And this is delayed by one unit here we will get X of n - 1 given n - 1 and then if we multiplied by An matrix we will get the predicted signal. So, this is the X of n given n - 1 this is the predicted signal. Now to get the estimate we use this predicted signal in the observation equation there is a multiplication of Cn and this state vector.

Therefore this predicted state vector is multiplied by Cn and we will get this predicted input vector this is the predicted input. So, Yn we are predicting here this prediction will result error that is the prediction error this is the innovation, so this is the innovation output. And this innovation output is scaled by the Kalman gain and then this corrected value this part is the correction term and this correction term will be added to the predicted value.

So this X of n given n - 1 plus this below that is kn n times y tilde n that will be added and we will get X of n given n, so that way we see that here there is a prediction part and there is a correction part that collection part is obtained by passing the predicted signal through that output matrix multiplier and then we get the predicted input that predicted input is subtracted from the input and we will get the innovation and this innovation is used to forget that collection term.

This innovation is used to get deep correction term that way Kalman filter is a prediction character filter. I will conclude my lecture here and this is the end of the course I hope that this course gave you a background on statistical signal processing. There are advanced techniques of statistical signal processing which was not possible to cover in this course, thank you.