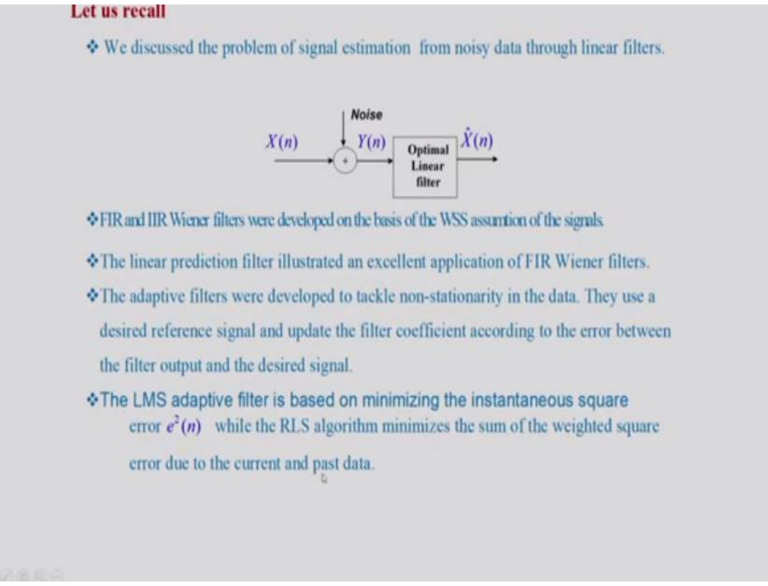
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology - Guwahati

Lecture – 35
Kalman Filter - 1

(Refer Slide Time: 00:43)

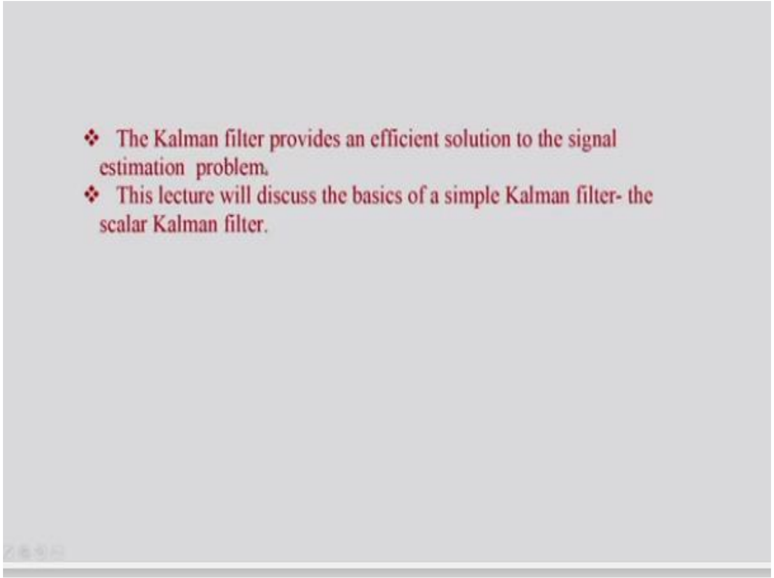


Hello students, welcome to this lecture on Kalman filter, let us recall; we discussed the problem of signal estimation from noisy data through optimum linear filters, so this was the problem; Xn is the original signal and it is contaminated by additive noise, we assumed white noise and the resulting observation is Yn and when we filtered by optimal linear filter we get the estimate of the signal X hat n.

FIR and IIR Wieners filters were developed on the basis of the WSS assumption of the signal, we assumed that Xn and Yn are jointly WSS and on the basis of that, that FIR and IIR Wiener filters were developed. The linear prediction filter illustrated an excellent application of FIR Wiener filters; we discussed the linear prediction problem, then linear prediction and prediction error filters, a faster algorithm for their implementation.

And one efficient realization of the linear prediction error filter, the adaptive filters was developed to tackle non stationarity in the data, they use a desired reference signal DN, we call it and update the filter coefficients according to the error between the filter output and the desired signal. The LMS adaptive filter is based on minimizing the instantaneous square error, e square n, while the RLS algorithm minimizes the sum of the weighted square error due to the current and past data.

(Refer Slide Time: 02:47)



The Kalman filter provides an efficient solution to the signal estimation problem; this lecture will discuss the basics of a simple Kalman filter that is the scalar Kalman filter.

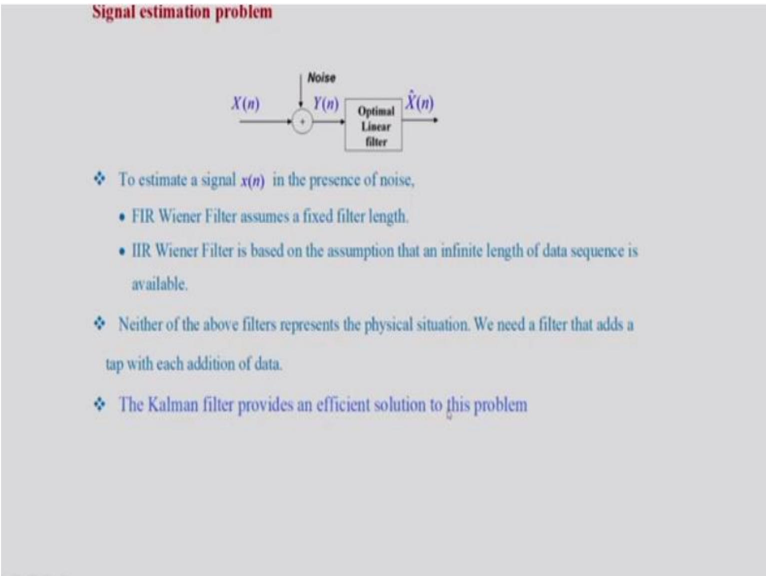
(Refer Slide Time: 03:06)

First we will give a very brief history; Kalman filter was invented in 1960 by R. E. Kalman, its origin can be traced to Karl Friedrich Gauss's least squares method, so we discussed about this method in an earlier lecture. It was first derived as a discrete filter, Kalman and Bucy extended the filter to the continuous version in 1961, this is the version commonly used in control systems engineering. Kalman filter is widely used in statistical signal processing.

(Refer Slide Time: 03:57)

Here is the photograph of R.E. Kalman, he was one of the best control theorists, many of the modern control concepts like state space representation, state estimators, observability etc., are associated with him.

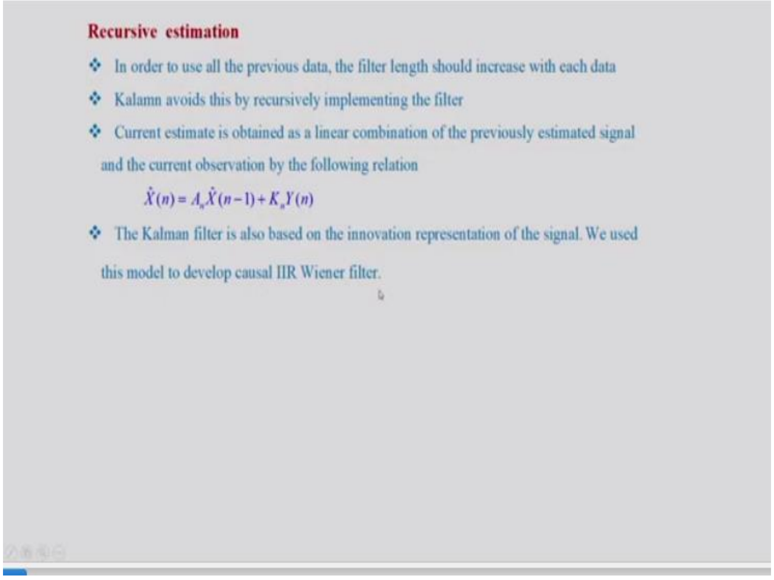
(Refer Slide Time: 04:17)



Let us again see the signal estimation problem, we have this original signal Xn and this is the noisy signal and after optimum linear filtering, we will get the estimate of the signal. To estimate the signal Xn in the presence of noise, FIR Wiener filter assumes a fixed filter length, so in the case of FIR Wiener filter, the filter length is fixed. IIR Wiener filter is based on the assumption that an infinite length of data sequence is available.

Because it is infinite impulse response IIR, so that way it makes the assumption that and infinite length of data is available, neither of the above filters represent the physical situation because data is neither infinite nor it is a fixed length, we need a filter that adds a tap with each addition of data. So, a new sample comes then filter length will be updated by one. The Kalman filter provides an efficient solution to this problem.

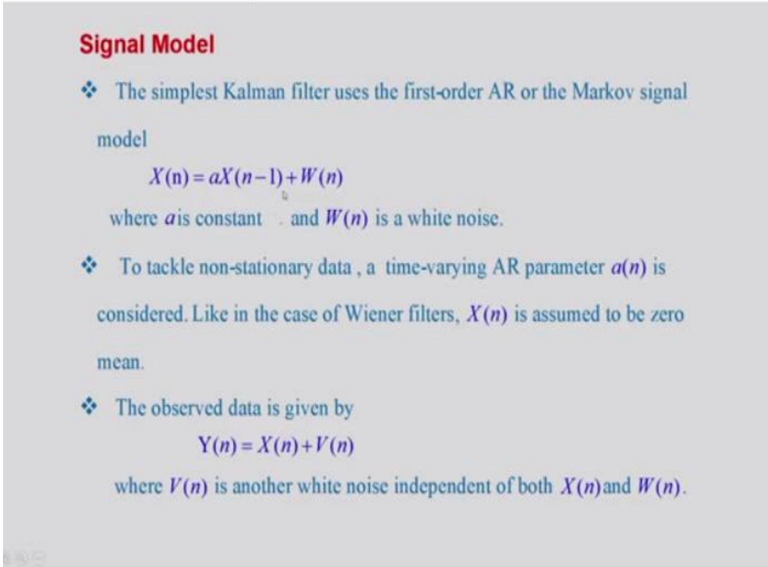
(Refer Slide Time: 05:45)



So, such a filter is based on recursive estimation in order to use the previous data, the filter length should increase with each data, Kalman avoids this by recursively implementing the filter, so the filter is recursively implemented. Current estimate is obtained as a linear combination of the previously estimated signal and the current observation by the following relationship; we will see how this relationship is obtained.

X hat n is equal to An times X hat n - 1 + Kn into Yn, the Kalman filter is also based on the innovation presentation of the WSS signal, we used this model to develop the causal IIR Wiener filter, here also that innovation representation of WSS signal will be used.

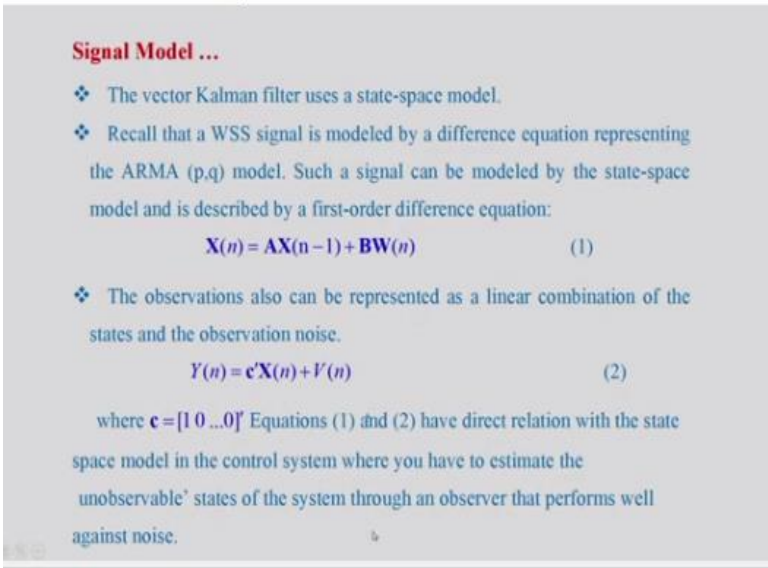
(Refer Slide Time: 06:54)



First, we will discuss the signal model; the simplest Kalman filter uses the first-order AR autoregressive or Markov signal model, Xn is equal to a times Xn - 1 + Wn, where a is a constant and Wn is a white noise, so Wn is a white noise, a is a constant, to tackle non stationarity in data, a time varying AR parameter an is considered, we can use this coefficient to be time varying, so that non stationarity of the data can be tackled.

Like in the case of Wiener filter, Xn is given to be of 0 mean, so that in all our signal estimation case, we assume the signal to be of 0 mean. The observed data is given by Yn is equal to Xn + Vn, where Vn is another white noise independent of Xn and Wn, so this is the measurement noise actually, this noise is independent of either Xn or Wn here.

(Refer Slide Time: 08:14)

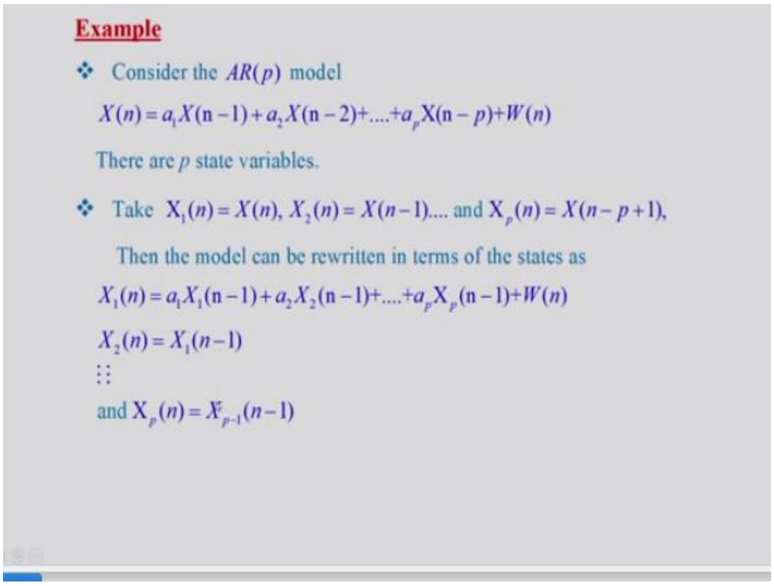


The vector Kalman filter uses a state space model instead of this simple model, Xn is equal to a Xn - 1 + Wn and Yn is equal to Xn + Vn, vector Kalman filter uses a state space model. Recall that a WSS signal is model by a difference equation representing the ARMA pq model, such a signal can be modelled by the state space model and is described in terms of first-order difference equation.

Suppose, in ARMA pq, it was a pth order difference equation now, it will be converted into a first-order difference equation in terms of matrices, so this is the Xn vector is equal to A times Xn - 1 vector + B times Wn, so this is the state space model. The observations also can be represented as a linear combination of the states and the observation noise or measurement noise that way Yn we can write as c transpose Xn + Vn, where c is equal to 1, 0, 0 transpose.

So, it is a vector comprising of first element 1, rest of the element 0, equations 1 and 2 have direct relations with the state space model in the control system, where you have to estimate the unobservable states of the system through an observer that performs well against noise, so there also Kalman filter comes.

(Refer Slide Time: 10:10)

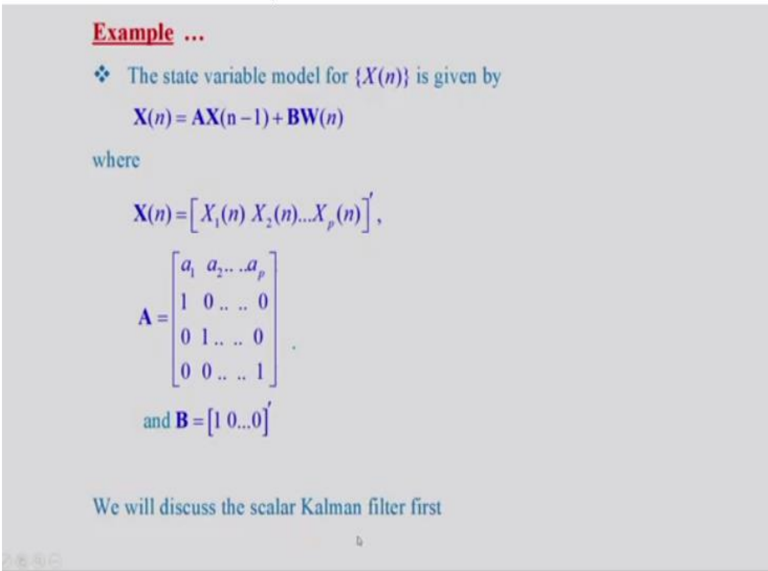


We will consider one example; consider the ARp model, so here Xn is equal to a1 times Xn - 1 + a2 times Xn - 2 + up to ap times Xn - p + Wn, so this is the pth order autoregressive model, where Wn is the white noise. We will take X1 because it is a pth order difference equation, there will be p state variables and we will take suppose X1n, the first state variable as Xn similarly, X2 n is equal to X of n - 1 like that, Xpn that is the pth state variable is X of n – p + 1.

If we assume this state variables, then this equation Xn is equal to a1 Xn - 1 + a2 into Xn – 2 like up to ap into X of n – p + Wn, this relationship we can write as this X1n that is the first state is equal to a1 times X1 of n - 1 because X1 n is equal to Xn + a2 times X2 of n – 1, where X2 n is nothing but X of n – 1, so that way this will be X2 of n - 1 similarly, the last term we can write as ap into Xp n – 1 + this Wn.

So, this way the first state equation can be written in terms of other states like this and here we see that all are first order difference and according to the definition, X2n is equal to X1 of n - 1 like that up to Xp of n is equal to X of p – n – 1, so that way all these state variables are represented as a first order difference form.

(Refer Slide Time: 12:44)

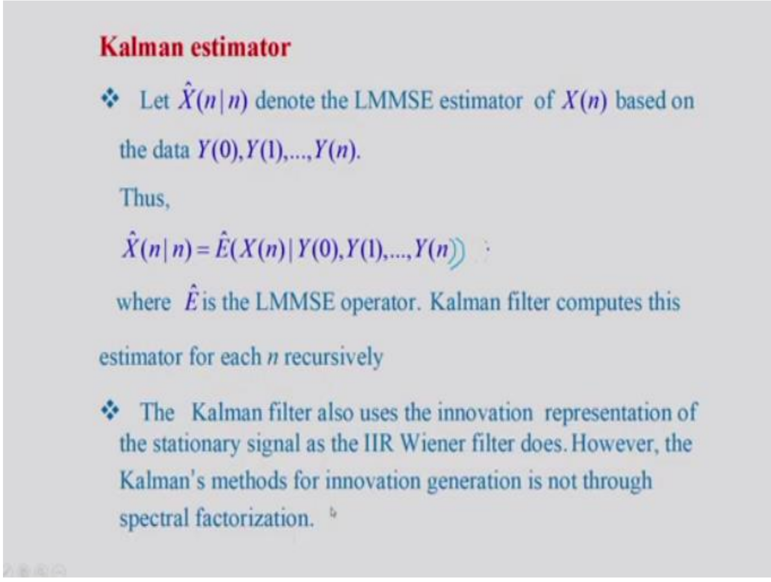


So, this we can write in a matrix; Xn that is the state vector comprising of X1n, X2n up to Xpn, so this can be written as Xn is equal to A times Xn - 1 + B times Wn, where A will be now; first row will be a1, a2 up to ap and next row will be 1, 0, 0, next will be 0, 1 up to 0, like that and last row will be 0, 0, last element 1 and similarly, B can be written as 1, 0, 0 up to 0, transpose of this vector, so that way this is the B vector.

So that way we can write now, Xn is equal to A times X of n - 1 + B times Wn, where B is a column vector with first element as 1 because there is only 1 Wn terms in this; in the ARp model and similarly, A matrix will be given by this; first row will be given by the that is first state which is expressed as the linear combination of all other states and where a1, a2 up to ap are the coefficients.

And similarly, from the other representation of the states that is X of 2n is equal to X1 of n - 1 like that so, this relationship we are writing by giving 1 in the corresponding position, so that way this is the A matrix and this is the B matrix. We will discuss this model; this type of model will be used in vector Kalman filter but first we will discuss the scalar Kalman filter.

(Refer Slide Time: 14:40)



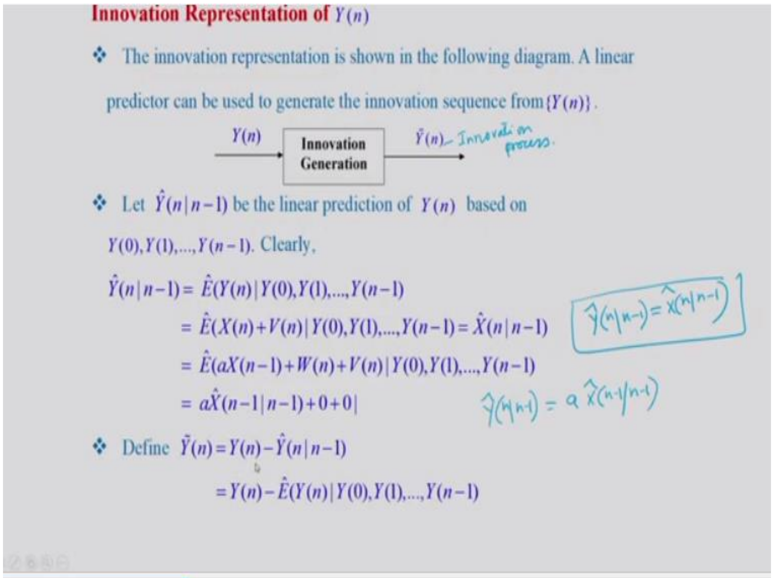
The Kalman filter uses all the available data that way let X hat n given n, this is the notation we will be using to denote the estimator, so this means that X hat n given n means, this estimator is based on observation till time n, so let X hat n given n denote the LMMSE; linear mean square error estimator of Xn based on the data Y0, Y1, up to Yn, so this is the estimator.

Thus, X hat n given n is E hat Xn given Y0, Y1 up to Yn, where this symbol E hat is the linear minimum mean square operator. Normally, E of Xn given Y0, Y1 up to Yn that is the conditional expectation but since we are using the linear minimum mean square error estimation, therefore E hat denote the LMMSE operator, thus X hat n given n is E hat Xn given Y0, Y1, up to Yn, so where E hat is the LMMSE operator, to denote linear minimum mean square error operation, we have used E hat.

And if suppose, signals are Gaussian, then E hat will be equal to simply E that is the conditional expectation, the Kalman filter is also uses the innovation representation of the stationary signal as in the IRR Wiener filter, so we will derive the Kalman filter on the basis of innovation representation but other derivation of Kalman filters are also available however, the Kalman's method for innovation generation is not through spectral factorization.

We know that in IRR Wiener filter, we got the innovation process through spectral factorization but here the approach is different.

(Refer Slide Time: 17:12)



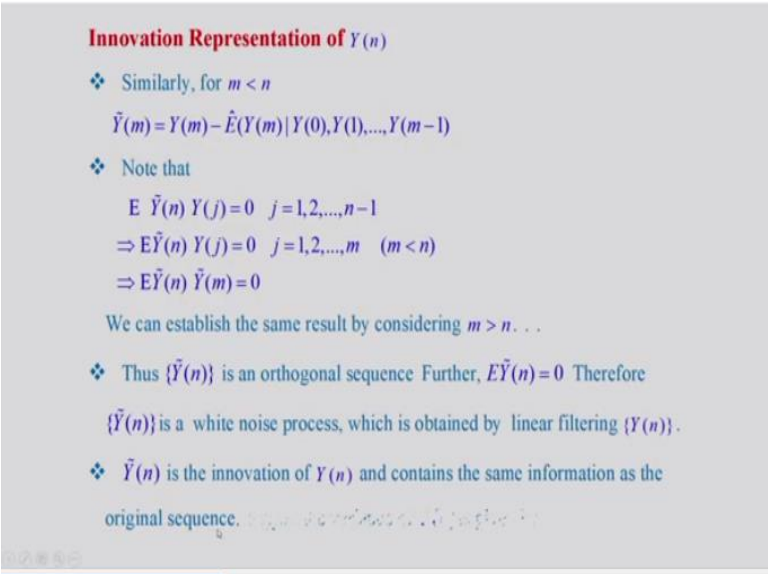
The innovation representation is shown in the following diagram; this diagram, a linear predictor can be used to generate the innovation sequence from Yn, here this innovation generation process is a linear predictor, so this side Yn is the input and Y tilde n is the innovation sequence, this is the innovation process. Let Y hat n given n - 1 with the linear prediction of Yn based on all past data that is Y0, Y1 up to Y n - 1 clearly, Y hat n given n - 1 is equal to that is LMMSE estimation of Yn given Y0 Y1 up to Y n – 1.

Now, this Yn we can write as Xn + Vn, therefore this expression will be equal to E cap of Xn + Vn given Y0, Y1 up to Y n – 1, this linear estimation of noise because it is uncorrelated given the data will be 0, therefore this will be simply the linear prediction of Xn given data up to n – 1, so this is one important relationship therefore, Y cap n given data up to n - 1 is equal to X hat or X cap n given n – 1, so this relationship is important for us.

Now, we can further write because Xn is a times Xn - 1 + Wn and this Vn is Vn only, so that way this will be X; E cap of a times X of n - 1 + Wn + Vn given data Y0, Y1 up to Y n - 1 and this quantity is the estimators for X of n – 1, so that way given all data up to Y n – 1, we are determining X of n - 1 that way it is estimator of X of n - 1 given data up to n - 1 plus; again this is a quite noise, this will be 0 and this will be 0.

So, therefore we can further write that Y cap that predicts on Y cap n given n - 1 that is also equal to a times X hat n - 1 given data up to n – 1, that way this is the estimator of the signal at instant n – 1. Now, we will define the innovation like this; Y tilde n is equal to Yn - Y hat n given n – 1, so that way this is equal to Yn - E cap Yn given Y0 Y1 up to Y n – 1, so that prediction error is the innovation sequence.

(Refer Slide Time: 20:52)

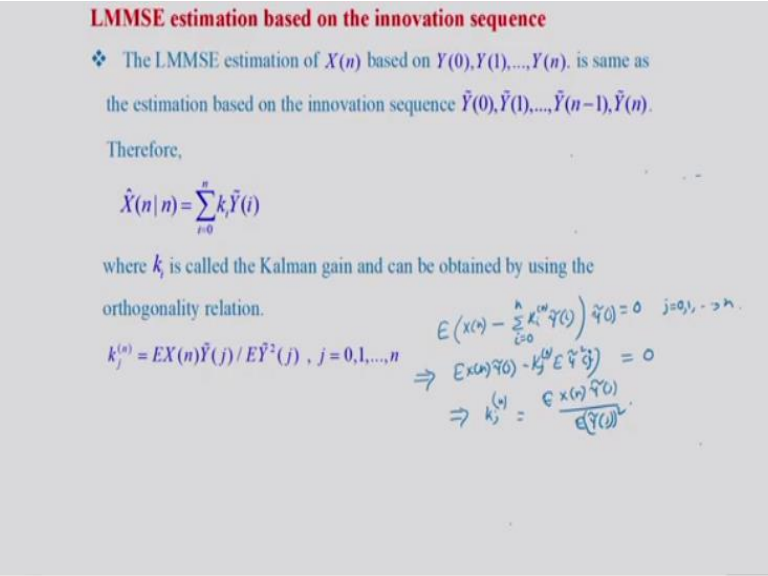


Similarly, suppose for any m < n, we can consider Y tilde m will be equal to Y m - summation E hat of Ym given Y0, Y1 up to Y m – 1, so this is the mth order predictor error. Now, because m is less than n, we can show that E of Y tilde n into Yj is equal to 0, for j is equal to 1 to up to n - 1 because error is orthogonal to data, so all first data it will be orthogonal.

Similarly, Y tilde n will be orthogonal to Yj because it is up to n – 1, so if we take some m which is less than n, then also this relationship will be true, therefore E of Y tilde n into Yj will be equal to 0, for j is equal to 1, 2, up to m now, this Yj we are considering m Yj’s and their linear combination is Y tilde m, therefore E of Y tilde n into Y tilde m will be equal to 0. What does it means? This Yn and Ym are orthogonal, we can establish the same result by considering m greater than n also, thus Y tilde n is an orthogonal sequence furthermore, E of Y tilde n is equal to 0 because all signal we are considering 0 mean only.

Therefore, this Yn process is a white noise process, which is obtained by linear filtering Yn because the generation mechanism is through linear prediction. Y tilde n is the innovation of Yn because after linear operation we are getting this Y tilde n, which contains the same information as the original sequence, Y tilde n is the innovation of Yn and contains the same information as the original sequence, so this innovation sequence is obtained by passing Yn through a linear filter.

(Refer Slide Time: 23:33)



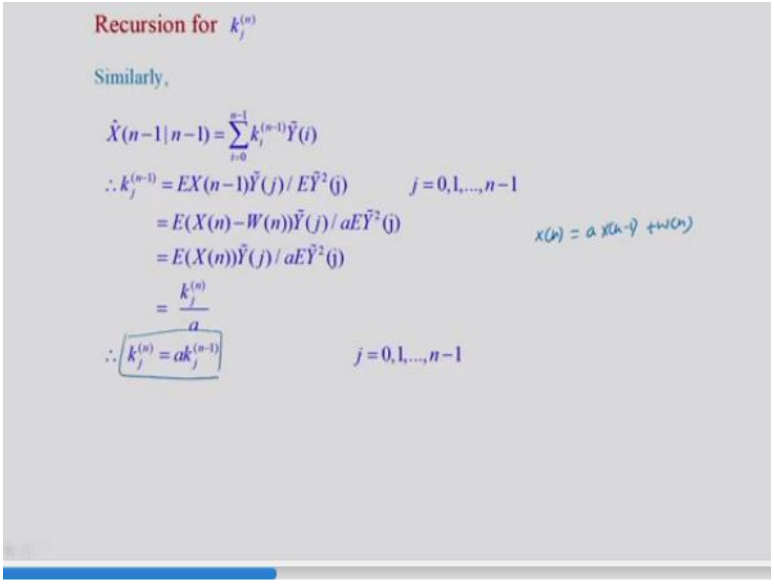
Now, let us see the LMMSE estimation based on the innovation sequence, we have got an innovation sequence, we will see the LMMSE estimation of Xn based on Y0, Y1 up to Yn is the same as the estimation based on the innovation sequence Y tilde 0, Y tilde 1 up to Y tilde n and thus we can write the estimator X hat n given n as summation Ki Y tilde i; i going from 0 to n, where Ki is called the Kalman gain, we will see later on why it is called Kalman gain.

And can be obtained by using the orthogonality relationship, these coefficients can be obtained by using the orthogonality relationship and this is the relationship we get, Kjn is equal to E of X n into Y tilde j divided by E of Y tilde square j, for j equal to 0 up to n, we will see how this relation can be obtained because error is orthogonal to data, therefore E of Xn - summation Ki that is nth instant into Y tilde i; i going from 0 to n into this is the error and that i is Y tilde j that must be equal to 0, for j is equal to 0, 1 up to n.

And this is equal to implies that E of Xn into Y tilde j minus; now because this Yi sequence is orthogonal, there will be only 1 term that is Kjn into E of Y tilde square i, so that must be equal to Y tilde square j that must be equal to 0 from which we will get Kn is equal to that is

Kjn is equal to E of Xn into Y tilde j divided by E of Y tilde j whole square like this, so this is the relationship.

(Refer Slide Time: 26:34)

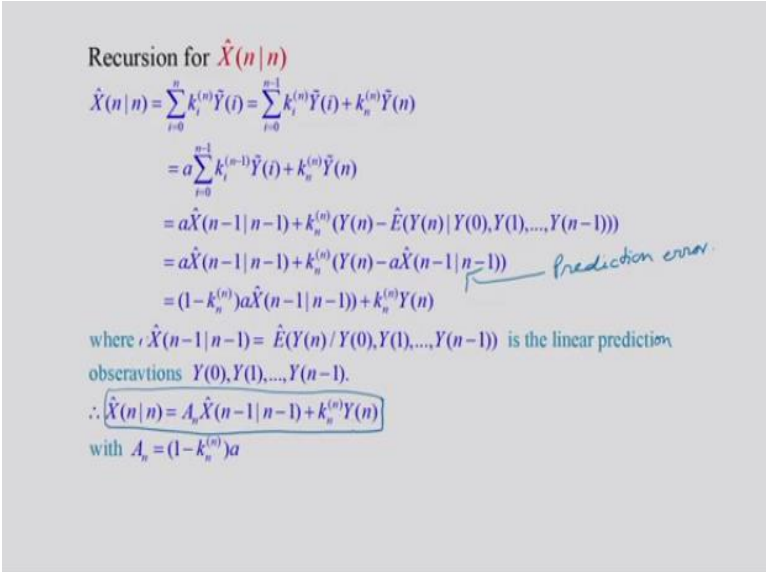


Now, let us see how we can get a recursive relation for Kjn that is the Kalman gain, so now let us see how to get the recursive relation for the Kalman gain Kjn, we have; we will first write X that is estimator at time n - 1 X hat n - 1 given that up to n – 1 is even by this relationship; summation Ki n – 1; i going from 0 to n - 1 Y tilde i and by the same manner, we can find Kj n - 1 that will be equal to E of X n - 1 into Y tilde j divided by E of Y tilde square j, this is for j equal to 0, 1 up to n – 1.

Because we are considering the estimator at instant n - 1 now, I know that X of n - 1 is equal to Xn – Wn divided by a from the model Xn is equal to a X n - 1 + Wn, we can write X of n - 1 is equal to Xn - Wn divided by a, so we will get this relationship Kj n - 1 is equal to E of Xn - Wn into Y tilde j divided by a by E of Y tilde square j. Now, Wn is independent of Y tilde j because this is, this involved data up to n – 1, they going from 0 to n – 1.

But Wn is the noise at instant n therefore, it is uncorrelated with all previous data, therefore this term will become 0, so therefore we will only have E of Xn into Y tilde j divided by aE of Y tilde square j, so this is the expression. Now, we know that E of Xn into Y tilde j divided by E of Y tilde square j that is Kjn that we derived here, so that way Kjn will be equal to a times Kj n – 1, for j is equal to 0 up to n – 1.

(Refer Slide Time: 29:39)



So, this is the relationship between the Kalman gain at different instant now, let us see how we can get the recursion for the estimator, we have X tilde n given n that is the estimator is equal to summation Ki n into Y tilde i; i going from 0 to n and this is equal to summation Ki n Y tilde i; i going from 0 to n - 1 + Kn n into Y tilde n, so that we are taking the last term out. Now, we can substitute this Ki n is equal to a times Ki n – 1.

So that this relationship; Kj n is equal to a times Kj n – 1, so therefore we will get this is here, from here we will get a times summation Ki n - 1 into Y tilde i; i going from 0 to n - 1 + Kn n times Y tilde n and we know that this is the estimator at instant n – 1, so therefore this will be a times X hat n - 1 given n - 1 + Kn times now, this is the prediction error, so that way E of that is Kn n into Yn minus; that prediction error is given by this E cap that is the LMMSE estimator of Yn from Y0, Y1 up to Y n – 1.

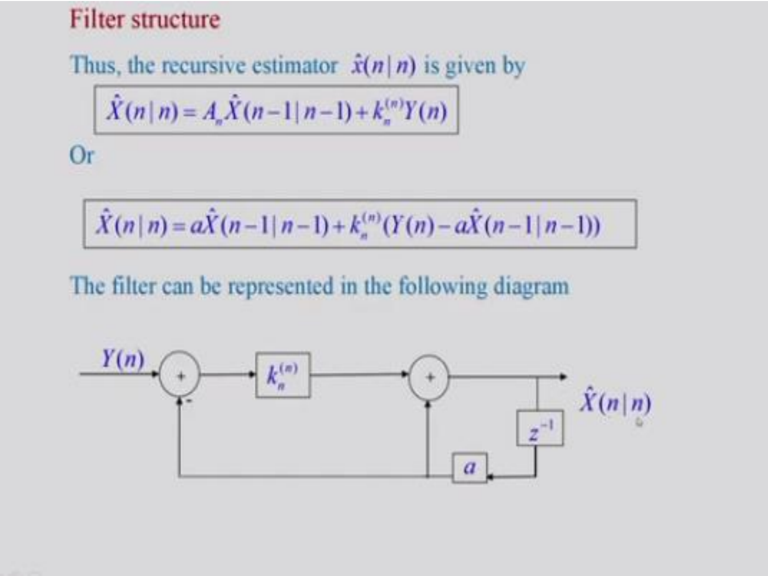
And this we have already observed that this is equal to a times X hat n - 1 given n - 1 that we have derived earlier, so that way this X cap n given n that is the estimator at instant n is equal to a times the estimator at instant n - 1 + Kn n times Yn - a into X hat n - 1 given n – 1, so this is the relationship we obtain that is current estimator is related to the previous estimator and if we take out this X hat n - 1 given n - 1 here, so we can write here 1 – Kn n into a times X hat n - 1 given n - 1 + K n times Yn.

So that way this is the recursive relationship, we see that this part is related to previous estimator and this is the part which is obtained from the current data, so current data Yn is multiplied by the Kalman gain vector that is Kn n, so previous estimator weighted by this number and then plus this current data weighted by this number, then we will get the current estimator.

So, this is the relationship that is current estimator is either we can write like this; a times X hat n - 1 given n - 1 previous estimator + Kalman gain times the estimation that prediction error, this is the prediction error, this part is prediction error and that we can write in terms of data itself that is 1 - Kn n into a times X hat n - 1 given n - 1 + Kn n times Yn, so this is the that gain vector; Kalman gain is weighing the data.

So that way we will see the importance of Kalman gain in this estimation, thus we have this relationship which we described earlier that the current estimator is a linear combination of past estimator and current data.

(Refer Slide Time: 34:35)

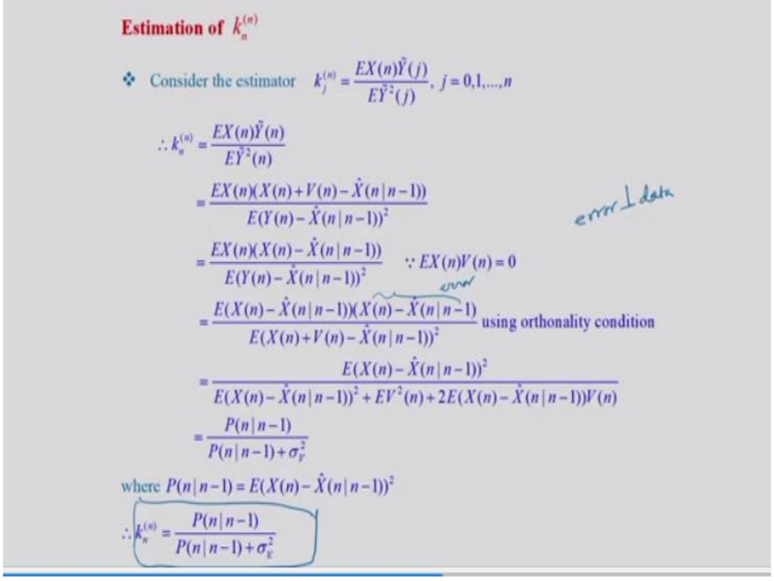


So, we have a filter structure now that is current estimator that is filter output is An times previous filter output plus Kn times Yn or current estimator that is filter output is equal to a times X hat n - 1 given n – 1, this is the delayed version of the filtered output plus Kn n times that is Kalman gain multiplies Yn - a X hat n - 1 given n – 1, this is the prediction error, so that way we have the structure.

Suppose, this is my filter output; filter output is here and this is Yn, so the filter output will be delayed, multiplied by a and then this will be subtracted from Yn and that will be scaled by Kn n and then edit to the delayed version of the estimator, the past estimator that is X hat n -

1 given n - 1 multiplied by a, these 2 will be added and then we will get the current estimation, so this is the filter structure for the scalar Kalman filter.

(Refer Slide Time: 36:10)



Let us see estimation of Kn n, we have the estimator Kj n is equal to E of Xn into Y tilde j divided by E of Y tilde j that we have obtained earlier, so therefore Kn n will be equal to E of Xn into Y tilde n divided by E of Y tilde square n. Now, this Y tilde n we can write that is observation Xn + Vn – the prediction, so prediction is X hat n given n – 1, so that way Kn n will be E of Xn into Xn + Vn - X tilde n given n – 1.

This Xn and Vn are independent, so Vn will not be there so, we can simply write this is equal to E of Xn into Xn – X hat n given n – 1, so this is the relationship we get. Now, this is the expression for mean square prediction error because we can write here Xn – X hat n given n – 1, then also this part is orthogonal to this error part, therefore the result will be same, so instead of this expression now, we can write this as E of Xn - X hat n given n - 1 into this part is there, Xn – X hat n given n – 1.

So, this is the result we got because of the relationship error; error is orthogonal to data now, in this case data is, this error is this part is error and this estimator is a linear combination of data, therefore this error will be orthogonal to this, so that way if we add this part, then the expression will remain the same as this that is the idea. So, therefore what we obtain is that; that is Kn n is equal to E of Xn – X hat n given n - 1 whole square, so this is one term.

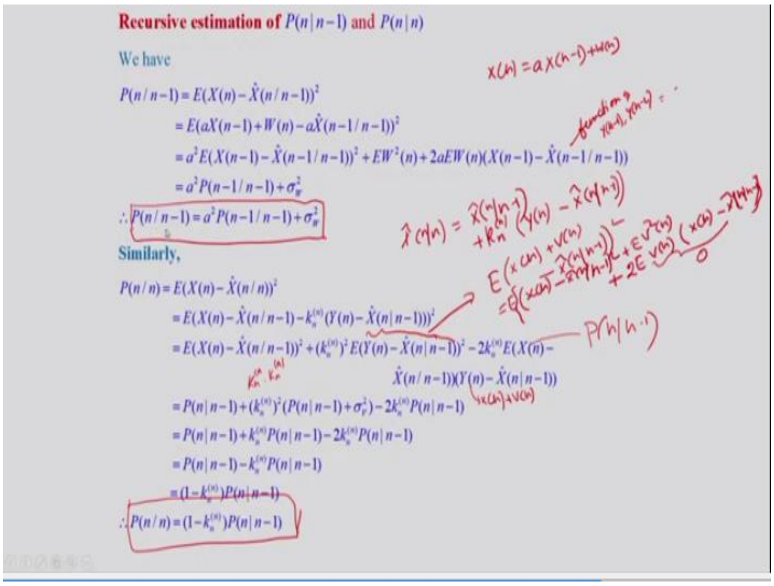
In the denominator, we have E of Yn – X hat n given n - 1 whole square, so now similarly we can write Yn is equal to Xn + Vn and minus this quantity, so that way denominator will be E of Xn + Vn – X hat n given n - 1 whole square. Now, after taking the expectation we will have E of this part we can take together, E of Xn – X hat n given n - 1 whole square and then E of V square n will be there plus twice E of Xn - X hat n given n - 1 into Vn.

So that way we are expanding this expression now, we observe that Vn is orthogonal to this; Vn is that is the noise; noise is orthogonal to all data, this Vn is the noise which is orthogonal to Xn and this X hat n given n - 1 these are linear combination of previous data, therefore Vn will be orthogonal to this vector also, therefore this term will become 0. So, what we will have therefore, in the denominator, we will have E of Xn – X hat n given n - 1 whole square + E of V square n that is the denominator.

And this term now, we will denote this term by P of n given n – 1, this is the mean square prediction error, you see this is Xn and this part is the prediction based on data up to n - 1 that way it is mean square prediction error, denominator also this is mean square prediction error, same mean square prediction error plus E of V square n is sigma V square, therefore what we get is that Kalman gain at instant n is equal to P of n given n – 1.

That is mean square prediction error based on data up to n - 1 divided by mean square prediction error + sigma V square, so this is one relation which we will be using, to determine the Kalman gain we need the previous mean square prediction error and the variance of the noise.

(Refer Slide Time: 41:37)



Now, let us see how we can get this mean square estimation error, this P n given n, let us find the mean square prediction error P of n given n - 1 that is equal to E of Xn – X hat n given n - 1 whole square. Now, according to the signal model, Xn is equal to aXn - 1 + Wn and this estimator X hat n given n - 1 is equal to a times X hat n - 1 given n - 1 that is the estimator at n - 1 multiplied by a.

Therefore, we can write P of n given n - 1 is equal to E of a Xn - 1 + Wn - a X hat n - 1 given n – 1, now this a we can take common and there will be X n - 1 - X hat n - 1 given n – 1, so that way we can write a square times E of X of n - 1 – X hat n - 1 given n - 1 whole square that is the first term, then Wn, so E of W square n, then there will be cross terms, 2 times a into E of this Wn into X of n - 1 - X hat n - 1 given n – 1.

So, that way 2a into E of Wn into X of n - 1 - X hat n - 1 given n - 1 now, this first term it is the mean square estimation error at instant n – 1, therefore this is P of n - 1 given n - 1 and E of W square n is equal to sigma W square. Now, let us examine this expression, so if we consider the signal model Xn is equal to a times X of n - 1 + Wn and this Wn is independent of Xn - 1 or Xn - 2 etc.

So that way, X Wn will be independent of all past values of Xn, similarly Wn will be independent of all possibilities of data also, so therefore E of Wn into X of n - 1 – X hat n - 1 given n - 1 now, Wn is independent of Xn – 1, this term will be equal to 0. Now, this estimator is a linear combination of past data only, it is a function of past data only, therefore this Wn will be independent of this factor also.

So, this is a function of Yn - 1 Yn - 2 etc., so that way this will be also independent of Wn, so this part will contribute 0 therefore, therefore we will have P of n given n - 1 is equal to a square times P of n - 1 given n - 1 + sigma W square, so this is the one recursive expression for mean square prediction error. Now, let us examine P of n given n, mean square estimation error.

This is equal to E of Xn - X hat n given n whole square now, writing X hat n given n, so this is; we write in terms of that Kalman filter expression and that is equal to that prediction part X hat n given n - 1 + that Kalman gain multiplied by that innovation Yn – X hat n given n – 1, so we write like this therefore, you have Xn – X hat n given n whole square can be written like this.

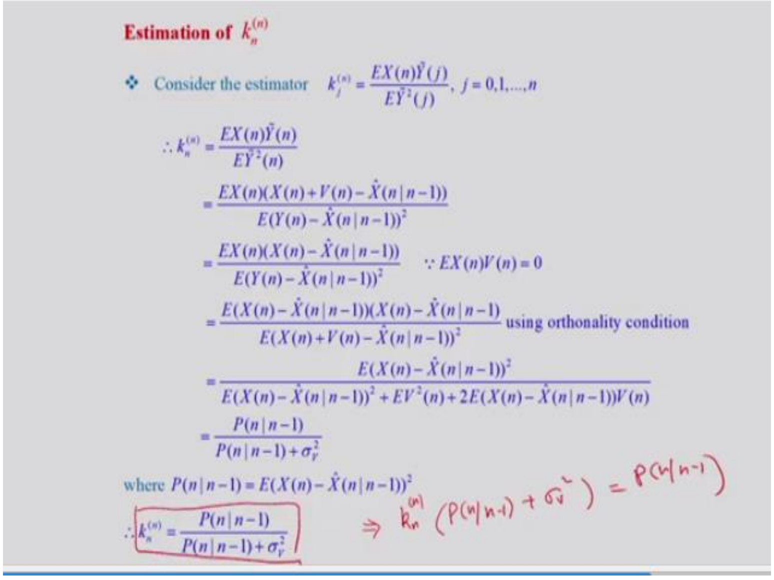
Now, we will combine this part, so Xn - X hat n given n - 1 whole square, one term will be there, Kn n whole square into E of Yn minus rest of the term; Yn – X hat n given n - 1 whole square, so this is the square term + Kn n whole square into E of Yn – X hat n given n - 1 whole square, so this part minus, now twice every term will come, so minus twice Kn n into E of Xn – X hat n given n - 1 into this part Yn – X hat n given n – 1.

Now, obviously this part is the mean square prediction error, P of n given n - 1 and this term will be E of Yn is equal to Xn + Vn - X hat n given n - 1 whole square, so this we can write as that is equal to E of Xn - X hat n given n - 1 whole square, 1 term then, plus E of V square n + twice E of Vn into Xn – X hat n given n – 1, now this part Vn is independent of Xn and Vn is independent of all past data, so that way Vn is independent of X hat n given n - 1 also.

Therefore, this part will result in 0, so what we will have is; this is mean square prediction error and this is sigma V square, so we can write it as Kn n whole square into P of n given n - 1 + sigma V square now, we will examine this term, twice Kn n into E of Xn - Xn given n - 1 into Yn - X hat n given n - 1 and here this Yn we can write as Xn + Vn, so what we will have is that E of Xn – X hat n given n - 1 into Xn + Vn – X hat n given n – 1.

Now, this Vn part will be independent of this Xn and X hat n given n – 1, so contribution of Vn will be 0, therefore we will simply have E of Xn - X hat n given n - 1 into Xn - X hat n given n – 1, so that way it will be E of Xn - X hat n given n - 1 whole square which is P of n given n – 1, so that way this expression will be equal to P of n given n – 1. So, we will have this quantity, so first term is P of n given n – 1, first term is P of n given n - 1 + Kn n whole square into P of n given n - 1 + sigma V square - twice Kn n into P of n given n – 1.

(Refer Slide Time: 51:42)

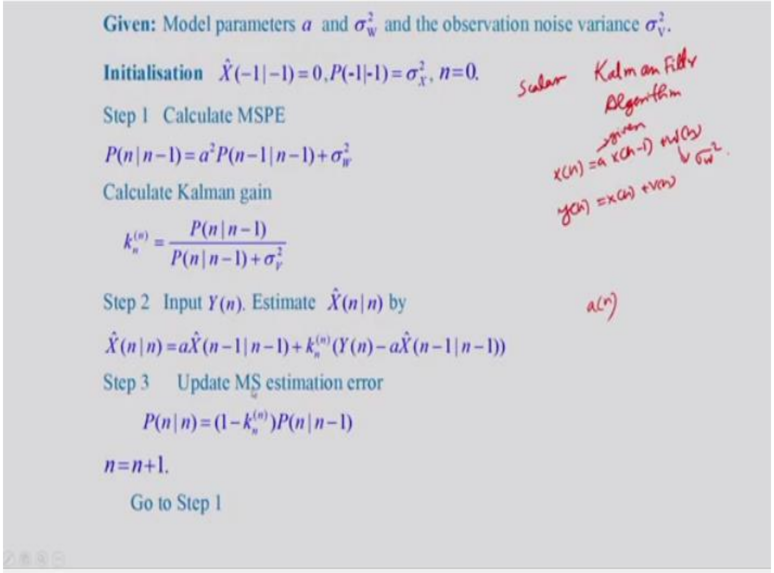


Now, we examine this expression, so Kn is equal to P of n given n - 1 divided by P of n given n - 1 + sigma V square, from this relationship we can get that Kn n into P of n given n – 1 + sigma V square that must be equal to P of n given n – 1, so that way here we will write, there are Kn n square is there, 1 Kn n will keep n into P n given n - 1 + sigma v squares, so this entire expression will be simply equal to P of n given n – 1.

Because this term we are writing as Kn n into Kn n into these things, so that term will be simply equal to P of n given n – 1, so therefore we have P of n given n - 1 + Kn n into P of n given n - 1 - 2 Kn n into P of n given n – 1, so 1 Pn given n - 1 will get cancelled, so we will have only 1 term here, so therefore this will be P of n given n - 1 – Kn n into P of n given n - 1 and this we can write like this, 1 – Kn n into P n given n – 1.

Therefore, mean square estimation error can be expressed in terms of mean square prediction error by this relationship that is mean square estimation error is equal to 1 - Kalman gain into mean square prediction error. So, we have this relationship now that is recursion for P n given n - 1 Pn given n and then that is the update equation for P of n given n - 1 and P of n given n and also we have examine that relationship for Kn n that is Kalman gain.

(Refer Slide Time: 54:33)



So, these all these relations will be used now to; in the algorithmic form, we will now describe the scalar Kalman filter algorithm, here given the model parameters a that is signal model is Xn is equal to a Xn - 1 + Wn, so a is given, then this for noise, this is a 0 mean white noise and its variance sigma W square is also given and the similarly, observation model is Yn is equal to Xn + Vn.

So, Vn that is observation noise is given to be 0 mean and variance sigma V square, so these are given, we have to initialize the estimator, so that way X hat - 1 given - 1 is equal to 0 because we do not use any estimator below, so that way it will be 0 and under that situation that is mean square prediction error, P of - 1 given - 1 will be equal to sigma x square is the variance of the signal.

And we can use any positive number here, we start with n is equal to 0 now, first step will be calculate the mean square prediction error that formula we know P of n given n - 1 is equal to a square times P of n - 1 given n - 1 that is the estimation error of the previous step plus sigma W square. Once this mean square prediction error is found out or calculated, we can calculate the Kalman gain that is given by Kn n is equal to P of n given n - 1 that is given here divided by P of n given n - 1 + sigma V square.

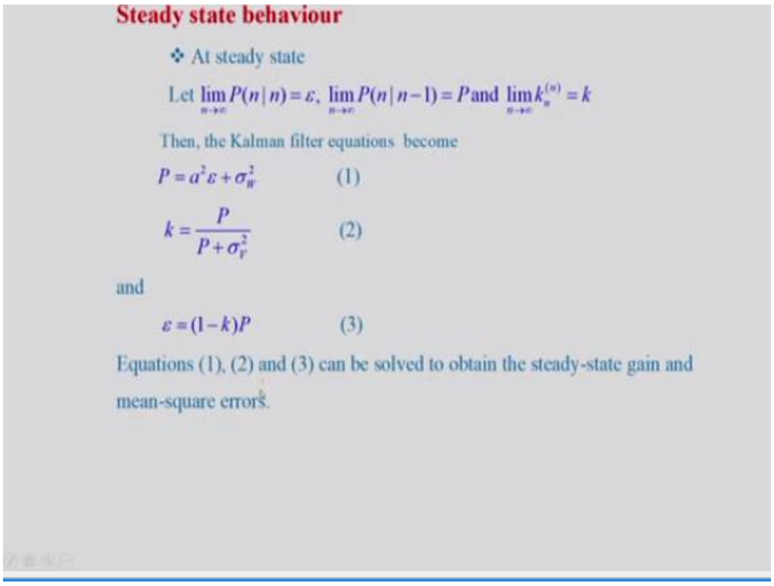
So, we will determine the Kalman gain now, we will take the input; input Yn now, we can estimate X hat n given n by this relationship that is X hat n given n is equal to a times X hat n - 1 given Xn - 1 that is the estimator at instant n - 1 which is multiplied by a here plus

Kalman gain Kn n into that innovation part is Yn - a times X hat n - 1 given n – 1, so this is the innovation, this is multiplied by the Kalman gain and added to the predicted value.

Now, we will update the MS estimation error that is P of n given n is equal to 1 – Kn n that is Kalman gain into P of n given n - 1 mean square prediction error because this value is found out here that will be used to update the mean square error and then we will have n is equal to n + 1 and will go to this step, we will again calculate the mean square prediction error using this relationship in that way, this Kalman filter will go on.

Here, we have taken constant a but to tackle non stationarity, we can take a as an as a function of n, in that case, an will be also input to the algorithm at this stage, so along with Yn we have to give the value of an also, so that way we have seen that Kalman filter recursively estimate the signal using a model that is the AR1 model or Markov model and under stress annuity condition of the signal, this algorithm will converge.

(Refer Slide Time: 59:19)

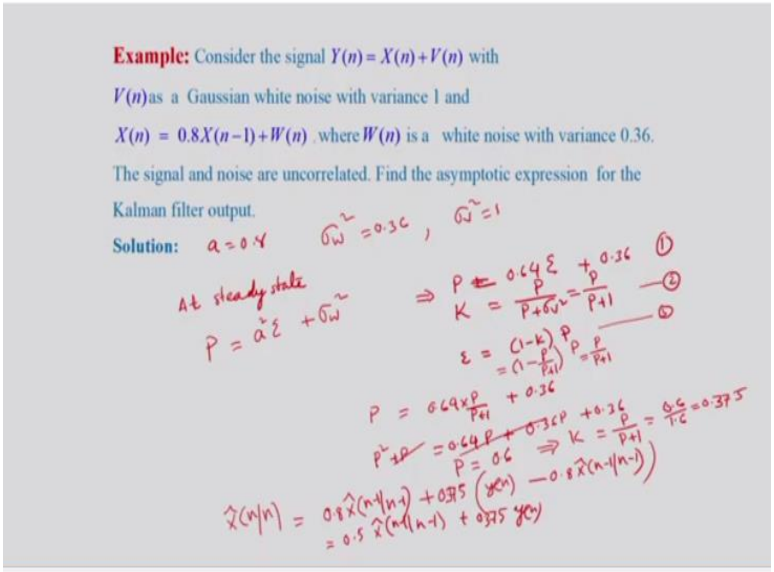


And now, let us see this steady state behaviour; at the steady state let limit of P n given n that is the mean square error that is epsilon suppose, limit of P n given n - 1 that is the mean square prediction error that is equal to suppose P and the Kalman gain limit of Kn n as n tends to infinity is K, then we will have the limiting form of the 3 Kalman filter equation.

First one is we know that P of given n - 1 is equal to a square times P of n - 1 given n - 1 + sigma W square and as n tends to infinity, this quantity will converge to P and this quantity will converge to epsilon, so that way we will have first one is P is equal to a square times epsilon + sigma W square, so from here, P that is the limit of this is P into a square limit of this is epsilon + sigma W square that is the first equation.

Then, Kalman gain as n tends to infinity that will be equal to K and this one is by our definition this is P and this is also P + sigma V square, so that way second equation will be K is equal to P by P + sigma P square. Now, third update equation P n given n, so that way limit of this will be mean epsilon; epsilon is equal to 1 - K into P, so that is the third equation, epsilon is equal to 1 - K into P.

(Refer Slide Time: 01:01:28)



These 3 equations can be solved to obtain the steady state gain and mean square errors, we will consider one example, consider the signal Yn is equal to Xn + Vn, where Vn is a Gaussian white noise with variance 1 and Xn is equal to 0.8 Xn - 1 + Wn where Wn is a white noise with variance 0.36. The signal and noise are uncorrelated; find this asymptotic expression for the Kalman filter output.

So, here we have the signal model that is Xn is equal to 0.8 Xn - 1 + Wn that is AR 1 model with a is equal to 0.8, value of a is equal to 0.8, then Wn is a white noise with variance 0.36, so sigma W square is equal to 0.36 and Vn is also a white noise with a variance sigma V square is equal to 1, so these things are given. Now, the steady state equations will be we have already established those equations here.

So, first equation will be P is equal to a square epsilon + sigma W square, so at steady state that is we have defined this quantity P is equal to that is the steady state prediction error is equal to a square times epsilon + sigma W square that is the first equation we got here, so therefore implies that P is equal to is now, 0.8, so 0.64 into epsilon + sigma W square is 0.36 that is equation 1.

Now, second equation we have the equation for Kalman gain K is equal to P by P + sigma V square and this will be equal to P by P + 1 because sigma V square is equal to 1, so this is equation 2 and the third equation is mean square estimation error in terms of the mean square prediction error, so that way third equation is epsilon is equal to 1 - K into P, this K we have got here, this is the third equation.

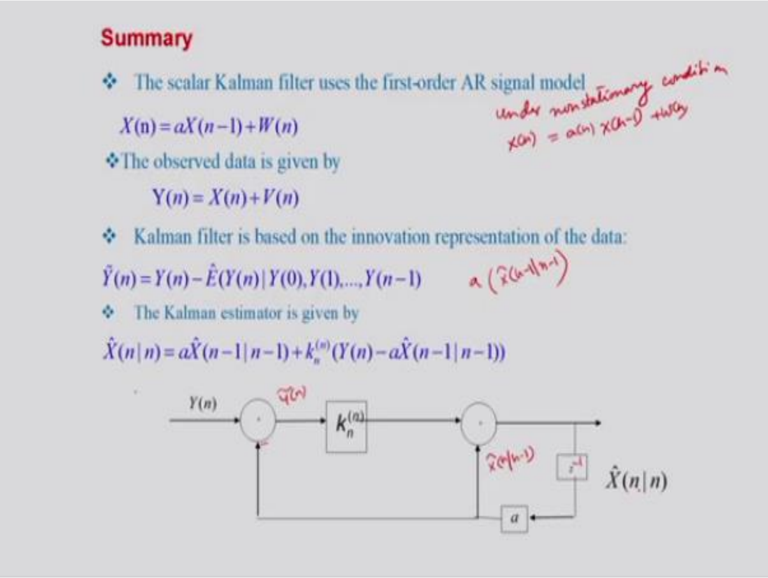
Now, we know K is equal to P by P + sigma V square that is P by P + 1, so that way this will be equal to 1 - K is P by P + 1 into P and now, this will be simply 1 by P + 1 into P, so this this is simply equal to P by P + 1, so epsilon this is mean square estimation error is equal to P by P + 1. Now, this if we substitute here, so substituting this value in equation 1, we will get P is equal to 0.64 into P by P + 1 + 0.36.

So, multiplying by P + 1, what we will have is P square into P will be equal to 0.64 P + 0.36P + 0.36, therefore this P and 0.64 P + 0.36 P that will be P only, so this P and this P will get cancel, so we will get P is equal to square root of this, it is a positive quantity, therefore this will be P is equal to 0.6 and also we have found out P, therefore K will be equal to imply that K is equal to P by P + 1, so this will be 0.6 by 1.6; 1 + 0.6, so that was 0.375.

So, we have found out the Kalman gain and we have found out value of P, so once we have the Kalman gain now, that signal update equation will be now X hat n that is we have to find out the asymptotic expression for the Kalman filter output, that is X hat n given n that is equal to that predicted value first; first is predicted that is a is 0.8 X hat of n - 1 given n - 1, this is the predicted value plus Kalman gain; Kalman gain is 0.375 into that innovation part is Yn - 0.8 into X hat n - 1 given n – 1.

And we can bring this here, so that way this part we can combine, if we combine this, this is 0.8, this one will be 0.8; 0.375 here into 0.8 and both if I add, I will get 0.5 here; 0.5 into X hat n - 1 given n - 1 + 0.375 into Yn, so this is the asymptotic equation for the Kalman filter output that estimator is given by this expression, this is the 0.5 times the previous value of the estimator plus 0.375 into current data.

(Refer Slide Time: 01:08:35)



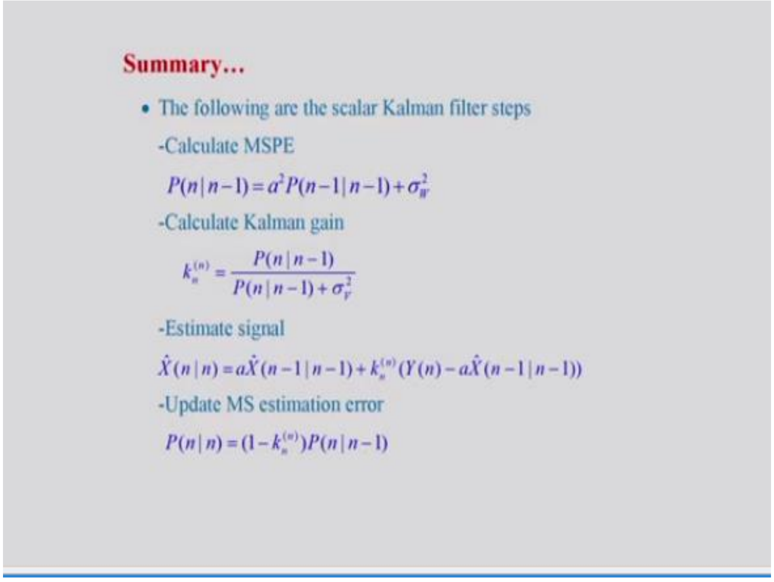
Let us summarize the lecture; this scalar Kalman filter uses the first order AR signal model, it is also known as Markov model, Xn is equal to a times Xn - 1 + Wn and under non stationarity condition, this a will be replaced by an, so under non stationary condition, Xn will be equal to an times X of n - 1 + Wn, the observed data is given by Yn is equal to Xn + Vn, where Vn is independent of Xn and Vn is independent of this Wn.

We derived the Kalman filter based on the innovation representation of the data that is Y tilde n is equal to Yn - E hat that is the linear mean square estimator E hat of Yn given Y0, Y1 up to Y n - 1 and using the model, this part we showed that this is same as a times X hat n - 1 given n – 1. The kalman estimator is given by X hat n given n is equal to a times X hat n - 1 given n - 1 that is estimator at previous instant multiplied by the Kalman gain Kn n into that is this part is the innovation Yn - a times X hat n - 1 given n – 1.

This is the Kalman filter recursion and this we presented in a block diagram like this, this is the Yn data and this is the output X hat n given n - 1 that will be delayed by 1 unit z to the power - 1 and then multiply it by a, so that way we will get the prediction a times X hat n - 1 given n - 1 that will be subtracted from Yn and we will get the innovation, this one is Y tilde n.

And this Y tilde n is multiplied by Kalman gain and then added with the predicted value, this is the predicted value X hat n given n – 1, this is the predicted below plus this is the corrected below Kn times Yn, so that will give us the estimator X hat n given n.

(Refer Slide Time: 01:11:51)



So, we established the following Kalman filter steps; first one is calculate the mean square prediction error that is P of n given n - 1 is equal to a square times P of n - 1 given n - 1 + sigma W square, so mean square prediction error is equal to a square times mean square estimation error at the previous instant plus sigma W square. Then we will determine the Kalman gain on the basis of this P of n given n – 1, so this is the expression; Kn n is equal to P of n given n - 1 divided by P of n given n - 1 + sigma V square.

Now, we will estimate the signal X hat n given n is equal to a times X hat n - 1 given n - 1 + Kn n times Yn - a times X hat n - 1 given n – 1, so this is the Kalman recursion, then we will update the MS estimation error; P of n given n is equal to 1 - Kn n into P of n given n – 1, so this is the scalar Kalman filter and it explains the theory of Kalman filter very well. Next, we will see how we can extend it to multiple signal case that is vector Kalman filter, thank you.