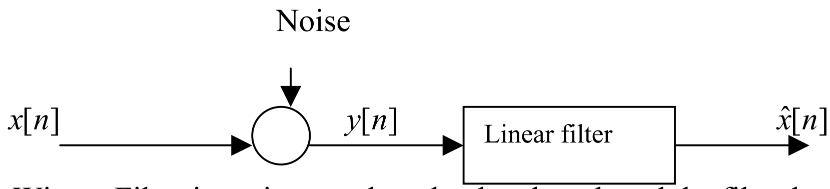
Kalman Filtering

To estimate a signal x[n] in the presence of noise.



- FIR Wiener Filter is optimum when the data length and the filter length are equal.
- IIR Wiener Filter is based on the assumption that infinite length of data sequence is available.

Neither of the above filters represents the physical situation. We need a filter that adds a tap with each addition of data.

The basic mechanism in Kalman filter is to estimate the signal recursively by the following relation

$$\hat{x}[n] = A_n \hat{x}[n-1] + K_n y[n]$$

The whole of Kalman filter is also based on the innovation representation of the signal. We used this model to develop causal IIR Wiener filter.

Signal Model

The simplest Kalman filter uses the first order AR signal model

$$x[n] = ax[n-1] + v[n]$$

where $v[n]$ is a white noise squence.

The general stationary signal is modeled by a difference equation representing the ARMA (p,q) model.S uch a signal can be modeled by the state space model

$$\mathbf{x}[n] = \mathbf{A}\mathbf{x}[n-1] + \mathbf{v}[n]$$

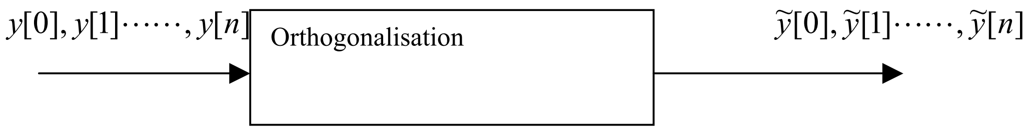
And the observations can be represented as a linear combination of the 'states' and the observation noise.

$$y[n] = \mathbf{c}'\mathbf{x}[n] + w[n]$$

Equations (1) and (2) have direct relation 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.

Our analysis will include only the simple (scalar) Kalman filter

The Kalman filter also uses the innovation representation of the stationary signal as does by the IIR Wiener filter. The innovation representation is shown in the following diagram.



In the above representation $\tilde{y}[n]$ is the innovation of $y[n]$ and contains the same information as the original sequence.

$$\begin{aligned}\tilde{y}[n] &= y[n] - \hat{E}(y[n]/y[n-1]............, y[0])) \\ &= y[n] - a\hat{x}[n-1] \\ &= y[n] - x[n] + x[n] - a\hat{x}[n-1] \\ &== y[n] - x[n] + ax[n-1] + v[n] - a\hat{x}[n-1] \\ &= v[n] + ae[n] + v[n]\end{aligned}$$

It can be easily shown that $\tilde{y}[n]$ is orthogonal to $\tilde{y}[n-1], \tilde{y}[n-2],.......\tilde{y}[0]$.

LMMSE estimation of $x[n]$ based on $y[0], y[1]\cdots\cdots, y[n]$, is same as the estimation based on the innovation sequence , $\tilde{y}[0], y[1],....\tilde{y}[n-1], \tilde{y}[n]$. Therefore,

$$\hat{x}[n] = \sum_{i=0}^{n} k_i \tilde{y}[n]$$

Where $k_i s$ are obtained by a way similar to we find the Fourier series expansion .

$$\begin{aligned}x[n] &= \hat{x}[n] + e[n] \\ &= \textstyle\sum_{i=0}^{n} k_i \tilde{y}[i] + e[n] \\ k_j &= Ex[n]\tilde{y}[j]/\sigma_j^{\,2} \quad j = 0,1..n\end{aligned}$$

Similarly

$$\begin{aligned}x[n-1] &= \hat{x}[n-1] + e[n-1] \\ &= \textstyle\sum_{i=0}^{n-1} k_i' \tilde{y}[i] + e[n-1] \\ k_j' &= Ex[n-1]\tilde{y}[j]/\sigma_j^{\,2} \quad j = 0,1..n-1 \\ &= Ex[n]/a\sigma_j^{\,2} \quad j = 0,1..n-1 \\ \therefore k_j' &= k_j / a\end{aligned}$$

Again,

$$\hat{x}[n] = \sum_{i=0}^{n} k_i \tilde{y}[i] = \sum_{i=0}^{n-1} k_i \tilde{y}[i] + k_n \tilde{y}[n]$$

$$= a\sum_{i=0}^{n-1} k_i' \tilde{y}[i] + k_n \tilde{y}[n]$$

$$= a\hat{x}[n-1] + k_n (y[n] - \hat{E}(y[n]/ y[n-1]............, y[0]))$$

$$= a\hat{x}[n-1] + k_n (y[n] - a\hat{x}[n-1])$$

$$= (1-k_n)a\hat{x}[n-1]) + k_n y[n]$$

where $\hat{E}(y[n]/ y[n-1]............, y[0]))$ is the linear prediction of $y[n]$ based on obseravtions $y[n-1]............, y[0])$ and which is same as $a\hat{x}[n-1]$.

$$\therefore \hat{x}[n] = A_n \hat{x}[n-1] + k_n y[n]$$

with $A_n = (1-k_n)a$

Consider the estimator

$$\hat{x}[n] = A_n \hat{x}[n-1] + k_n y[n]$$

The estimation error is given by

$$e[n] = x[n] - \hat{x}[n]$$

Therefore $e[n]$ must orthogonal to past and present observed data .

$$Ee[n]y[n-m] = 0, m \geq 0$$

We want to find A_n and the k_n using the above condition.

Case 1

$e[n]$ is orthogonal to current and past data. First consider the condition that $e[n]$ is orthogonal to the current data.

$$\therefore Ee[n]y[n] = 0$$

$$\Rightarrow Ee[n](x[n] + w[n]) = 0$$

$$\Rightarrow Ee[n]x[n] + Ee[n]w[n] = 0$$

$$\Rightarrow Ee[n](\hat{x}[n] + e[n]) + Ee[n]w[n] = 0$$

$$\Rightarrow Ee^2[n] + Ee[n]w[n] = 0$$

$$\Rightarrow \varepsilon[n] + E(x[n] - A_n \hat{x}[n-1] - k_n y[n])w[n] = 0$$

$$\Rightarrow \varepsilon[n] - k_n \sigma_w^{\ 2} = 0$$

$$\Rightarrow k_n = \frac{\varepsilon[n]}{\sigma_{w^2}^{\ 2}}$$

Case 2

Secondly we consider the case that consider the condition that $e[n]$ is orthogonal to the pasr data.

We have

$$Ee[n]y[n-m]=0, m>0$$
$$\Rightarrow E(x[n]-A_n\hat{x}[n-1]-k_n y[n])y[n-m]=0$$
$$\Rightarrow Ex[n]y[n-m]-A_nE\hat{x}[n-1]y[n-m]-k_nEy[n]y[n-m]=0$$
$$\Rightarrow R_{xx}[m]-A_nE\hat{x}[n-1]y[n-m]-k_nR_{xx}[m]=0$$
$$\Rightarrow (1-k_n)R_{xx}[m]-A_nE\hat{x}[n-1]y[n-m]=0$$
$$\Rightarrow (1-k_n)R_{xx}[m]-A_nE(x[n-1]-e[n-1])y[n-m]=0$$
$$\Rightarrow (1-k_n)R_{xx}[m]-A_nR_{xx}[m-1]=0$$
$$\Rightarrow A_n=(1-k_n)\frac{R_{xx}[m]}{R_{xx}[m-1]}$$
$$\Rightarrow \therefore A_n=(1-k_n)a$$

Therefore the recursive relationship is

$$\begin{aligned}\hat{x}[n]&=(1-k_n)a\hat{x}[n-1]+k_n y[n]\\ &=a\hat{x}[n-1]+k_n(y[n]-a\hat{x}[n-1]\end{aligned}$$

We have to estimate $\varepsilon[n]$ at every value of n.
How to do it?

Consider

$$\begin{aligned}\varepsilon[n]&=Ex[n]e[n]\\ &=Ex[n](x[n]-(1-k_n)a\hat{x}[n-1]-k_n y[n])\\ &=\sigma_x{}^2-(1-k_n)aEx[n]\hat{x}[n-1]-k_nEx[n]y[n]\\ &=(1-k_n)\sigma_x{}^2-(1-k_n)aE(ax[n-1]+v[n])\hat{x}[n-1]\\ &=(1-k_n)\sigma_x{}^2-(1-k_n)a^2Ex[n-1]\hat{x}[n-1]\end{aligned}$$

Again

$$\begin{aligned}\varepsilon[n-1]&=Ex[n-1]e[n-1]\\ &=Ex[n-1](x[n-1]-\hat{x}[n-1])\\ &=\sigma_x^2-Ex[n-1]\hat{x}[n-1]\end{aligned}$$

Therefore,

$$Ex[n-1]\hat{x}[n-1]=\sigma_x^2-\varepsilon[n-1]$$

Hence $$\varepsilon[n]=\frac{\sigma_v^2+a^2\varepsilon[n-1]}{\sigma_v^2+\sigma_w^2+a^2\varepsilon[n-1]}\sigma_w^2$$

where we have substituted $\sigma_v^2 = (1-a^2)\sigma_X^2$

We have still to find $\varepsilon[0]$. For this assume $x[-1] = \hat{x}[-1] = 0$. Hence from the relation
$\varepsilon[n] = (1-k_n)\sigma_x{}^2 - (1-k_n)\sigma^2 Ex[n-1]\hat{x}[n-1]$

we get
$\varepsilon[0] = (1-k_0)\sigma_x{}^2$

Substituting $k_0 = \dfrac{\varepsilon[0]}{\sigma_{w^2}{}^2}$

We get $\varepsilon[0] = \dfrac{\sigma_x{}^2\sigma_w{}^2}{\sigma_x{}^2 + \sigma_w{}^2}$

The Kalman filter algorithm can be rewritten as follows:

Given: Signal model parameters a and σ_{v}^2 and the observation noise variance σ_{w}^2.

Initialisation $\hat{x}[-1] = 0$

Step 1 $n = 0$. Calculate $\varepsilon[0] = \dfrac{\sigma_x{}^2\sigma_w{}^2}{\sigma_x{}^2 + \sigma_w{}^2}$

Step 2 Calculate $k_n = \dfrac{\varepsilon[n]}{\sigma_{w^2}{}^2}$

Step 3 Input $y[n]$. Estimate $\hat{x}[n]$ by

$\hat{x}[n] = a\hat{x}[n-1] + k_n(y[n] - a\hat{x}[n-1]$

Step 4 $n = n+1$.

Step 5 $\xi[n] = \dfrac{\sigma_v^2 + a^2\varepsilon[n-1]}{\sigma_v^2 + \sigma_w^2 + a^2\varepsilon[n-1]}\sigma_w^2$

Step 6 Go to Step 2

Problem Given

$x[\mathrm{n}] = 0.8x[\mathrm{n}-1] + v[n] \qquad \mathrm{n} \geq 0$

$y[n] = x[n] + w[n] \qquad \mathrm{n} \geq 0$

$\sigma_V^2 = 0.16,\ \sigma_W^2 = 1$

Find the expression for the Kalman filter equations at convergence and the corresponding mean square error.