Optimal Filtering

Lecture Notes

Ahmet Ademoglu, *PhD*,
Bogazici University,
Institute of Biomedical Engineering

Problem Statement

Recovering a desired signal $d(n)$ in noisy observation $x(n)$

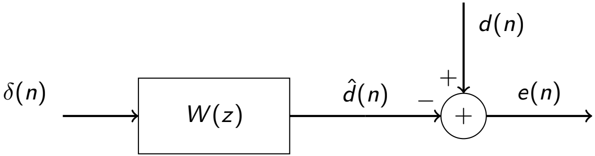
$$x(n) = d(n) + v(n)$$

assuming that $d(n)$ and $v(n)$ are WS jointly stationary random processes.

Wiener Approach

$$\xi = E\left\{|e(n)|^2\right\} = E\left\{\left|d(n) - \hat{d}(n)\right|^2\right\}$$

1. **Filtering** : Estimating $d(n)$ from the past values of $x(n)$.
2. **Smoothing** : The same as filtering but the filter is noncausal *i.e.* the $d(n)$ is estimated not only with the past but also the future data.
3. Prediction : $d(n) = x(n+1)$, estimating $x(n+1)$ using its past and present values with a filter $W(z)$.
4. Deconvolution : $x(n) = d(n) * g(n) + v(n)$ with $g(n)$ being a linear unit impulse response of a shift invariant filter which distorts $d(n)$ and is to be recovered by $W(z)$.



$$W(z) = \sum_{n=0}^{p-1} w(n) z^{-n} \qquad \hat{d}(n) = \sum_{n=0}^{p-1} w(l) x(n-l)$$

$$\xi = E\left\{ |e(n)|^2 \right\} = E\left\{ \left| d(n) - \hat{d}(n) \right|^2 \right\}$$

$$\frac{\partial \xi}{\partial w^*(k)} E\left\{ e(n) e^*(n) \right\} = E\left\{ e(n) \frac{\partial e^*(n)}{\partial w^*(k)} \right\} = 0, \; \frac{\partial e^*(n)}{\partial w^*(k)} = -x^*(n-k)$$

$$E\left\{ e(n) x^*(n-k) \right\} = 0, k = 0, \ldots, p-1$$

Orthogonality Principle

$$\sum_{l=0}^{p-1} w(l) r_x(k-l) = r_{dx}(k) \qquad k = 0, 1, \ldots, p-1$$

where $r_{dx}(k) = E\left\{d(n)x^*(n-k)\right\}$

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \ldots & r_x^*(p-1) \\ r_x(1) & r_x(0) & \ldots & r_x^*(p-2) \\ \vdots & \vdots & \vdots & \vdots \\ r_x(p-1) & r_x(p-2) & \ldots & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ \vdots \\ w(p-1) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ \vdots \\ r_{dx}(p-1) \end{bmatrix}$$

Wiener-Hopf Equation

$$\mathbf{R}_x \mathbf{w} = \mathbf{r}_{dx}$$

Minimum Mean Square Error

$$\xi_{min} = E\left\{e(n)d^*(n)\right\} = r_d(0) - \sum_{l=0}^{p-1} w(l) r_{dx}^*(l) = r_d(0) - \mathbf{r}_{dx}^H \mathbf{R}_x^{-1} \mathbf{r}_{dx}$$

Filtering

$$x(n) = d(n) + v(n)$$

Since signal and noise are assumed to be uncorrelated;

$$r_{dx}(k) = r_d(k) \text{ and } r_x(k) = r_d(k) + r_v(k)$$

Wiener-Hopf Equation

$$[\mathbf{R}_d + \mathbf{R}_v]\,\mathbf{w} = \mathbf{r}_d$$

$d(n)$: AR(1) process with $\mathbf{r}_d(k) = \alpha^{|k|}$ and $0 \leq \alpha \leq 1$, $x(n) = d(n) + v(n)$

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \end{bmatrix} \quad r_{dx} = r_d$$

$$\begin{bmatrix} 1 + \sigma_v^2 & \alpha \\ \alpha & 1 + \sigma_v^2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix}$$

For $\alpha = 0.8$ and $\sigma_v^2 = 1$,

$W(e^{j\omega}) = 0.4048 + 0.2381 e^{-j\omega}$ (LP Filter)

Power Spectrum of $d(n)$ is $P_d(e^{j\omega}) = \frac{0.36}{1.64 - 1.6\cos\omega}$ and

$\xi_{min} = r_d(0) - w(0) r_d^*(0) - w(1) r_d^*(1) = 0.4048$

Power of the filtered signal $d^{'}(n) = w(n) * d(n)$

$E\{|d'(n)|^2\} = E\left\{\mathbf{w}^T \mathbf{R}_d \mathbf{w}\right\} = 0.3748$

Noise Power $E\{|v'(n)|^2\} = \mathbf{w}^T \mathbf{R}_v \mathbf{w} = 0.2206$

At the input : $E\{|d(n)|^2\} = 1$ and $SNR = 10\log_{10}\frac{1}{1} = 0$

At the output: $SNR = 10\log_{10}\frac{0.3748}{0.2206} = 2.302$ dB 2.3 dB increase by Wiener filtering

Linear Prediction of AR(1) process with $r_d(k) = \alpha^{|k|}$

$x_d = x(n+1) \longrightarrow r_{dx}(k) = E(x(n+1)x^*(n-k)) = r_x(k+1)$

$$\hat{x}(n+1) = w(0)x(n) + w(1)x(n-1)$$

$$\begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha^2 \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

$\hat{x}(n+1) = \alpha x(n)$ and $\xi_{min} = r_x(0) - w(0)r_x(1) - w(1)r_x(2) = 1 - \alpha^2$

Prediction with Noise: $y(n) = x(n) + v(n)$

$$\hat{x}(n+1) = \sum_{k=0}^{p-1} w(k)y(n-k)$$

Wiener-Hopf Equation

$$\mathbf{R}_y\mathbf{w} = \mathbf{r}_{dy}$$

$\mathbf{R}_x + \sigma_v^2\mathbf{I} = \mathbf{r}_{dy} = r_x(k+1)$

$$\begin{bmatrix} 1+\sigma_v^2 & \alpha \\ \alpha & 1+\sigma_v^2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha^2 \end{bmatrix} \longrightarrow \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 0.3238 \\ 0.1905 \end{bmatrix}$$

Multistep Prediction of $x(n)$ with $r_x(k) = \delta(k) + (0.9)^{|k|} \cos(\pi k/4)$

One step ahead prediction

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix} \longrightarrow \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 0.3540 \\ -0.1127 \end{bmatrix}$$

$$\hat{x}(n+1) = 0.3540x(n) - 0.1127x(n-1) \quad \xi_{min}^{(1)} = 1.7747$$

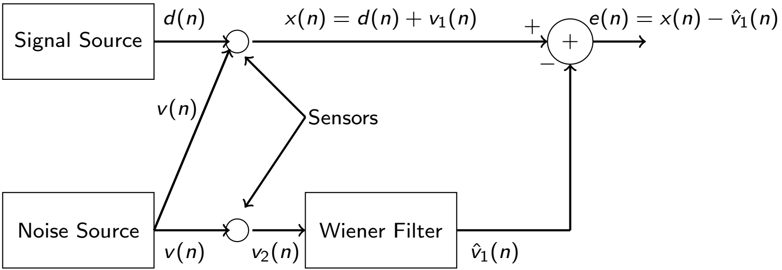
Three step ahead prediction

$$\begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} r_x(3) \\ r_x(4) \end{bmatrix} \longrightarrow \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} -0.1706 \\ -0.2738 \end{bmatrix}$$

$$\hat{x}(n+3) = -0.1706x(n) - 0.2738x(n-1) \quad \xi_{min}^{(3)} = 1.7324$$

Why $\xi_{min}^{(3)} < \xi_{min}^{(1)}$?

Noise Cancellation



Wiener-Hopf Equation

$$\mathbf{R}_{v_2}\mathbf{w} = \mathbf{r}_{v_1 v_2}$$

$$r_{v_1 v_2}(k) = E\{x(n)v_2^*(n-k)\} - E\{d(n)v_2(n-k)^*\} = r_{xv_2}(k)$$

$d(n) = \sin(n\omega_0 + \phi)$, $v_1 = 0.8v_1(n-1) + g(n)$ and $v_1 = -0.6v_2(n-1) + g(n)$

$\hat{r}_{v_2}(k) = \frac{1}{N}\sum_{n=0}^{N-1} v_2(n)v_2(n-k)$ and $\hat{r}_{xv_2}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)v_2(n-k)$

```
% Noise Cancellation
N=200;
phi =0;
n=transpose([0:N-1]);
w0=0.05*pi;
d = sin(n*w0+ phi);
g=randn(N,1);
v1 = filter(1,[1 -0.8],g);
v2 = filter(1,[1 0.6],g);
x = d + v1 ;
rv2=conv(v2,flipud(v2))/(N);
rxv2=conv(x,flipud(v2))/(N);

p=6;
Rv2 = toeplitz(rv2([1:p]+N-1,1));
w = Rv2\rxv2([1:p]+N-1);
d_h = real(ifft(fft(w,N).*fft(v2)) );
p=12;
Rv2 = toeplitz(rv2([1:p]+N-1,1));
w = Rv2\rxv2([1:p]+N-1);
d_h1 = real(ifft(fft(w,N).*fft(v2)) );
subplot(2,2,1); plot([x d ]);
title('signal and signal + noise')
subplot(2,2,2); plot([x d x-d_h ]);
title('6th order Wiener filtered signal, signal, signal + noise')
subplot(2,2,3); plot([x d x-d_h1]);
title('12th order Wiener filtered signal, signal, signal + noise')
subplot(2,2,4); plot([x d x-d_h x-d_h1 ]);
title('6th and 12th order Wiener filtered signals, signal, signal + noise')
```

Noncausal IIR Wiener Filter

$$H(z) = \sum_{n=-\infty}^{\infty} h(n) z^{-n}$$

$$\xi = E\{\left|e(n)^2\right|\},\ e(n) = d(n) - \hat{d}(n) = e(n) - \sum_{l=-\infty}^{\infty} h(l)x(n-l)$$

$$\frac{\partial \xi}{\partial h^*(k)} = E\left\{e(n)x^*(n-k)\right\} = 0\ ;\ -\infty \le k \le \infty$$

$$\sum_{l=-\infty}^{\infty} h(l) r_x(k-l) = r_{dx}(k)\ ;\ -\infty \le k \le \infty$$

$$h(k) * r_x(k) = r_{dx}(k) \longrightarrow H(z) = \frac{P_{dx}(z)}{P_x(z)} \text{ or } H(e^{j\omega}) = \frac{P_{dx}(e^{j\omega})}{P_x(e^{j\omega})}$$

$$\xi_{min} = r_d(0) - \sum_{l=-\infty}^{\infty} h(l) r_{dx}^*(l) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_d(e^{j\omega})d\omega - \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega})P_{dx}^*(e^{j\omega})d\omega$$

Wiener Smoohing of $x(n) = d(n) + v(n)$ when $d(n)$ and $v(n)$ are uncorrelated.

$$r_x(k) = r_d(k) + r_v(k) \longrightarrow P_x(e^{j\omega}) = P_d(e^{j\omega}) + P_v(e^{j\omega})$$

$$r_{dx} = r_d(k) \longrightarrow P_{dx}(e^{j\omega}) = P_d(e^{j\omega}),$$

$$H(e^{j\omega}) = \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})} \qquad \xi_{min} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_v(e^{j\omega})H(e^{j\omega})d\omega$$

$d(n)$ real value of AR(1) process with $P_d(z) = b^2(0)/(1 - \alpha z^{-1})(1 - \alpha z)$

$$x(n) = d(n) + v(n)$$

$v(n)$: white noise with variance σ_v^2

$$H(z) = \frac{P_d(z)}{P_d(z) + P_v(z)} = \frac{b^2(0)}{b^2(0) + \sigma_v^2(1 - \alpha z^{-1})(1 - \alpha z)}$$

$\xi_{min} = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} P_v(e^j\omega)H(e^{j\omega})d\omega = \sigma_v^2 h(0)$

For $b^2(0) = 0.25$, $\alpha = 0.5$, $\sigma_v^2 = 0.25$

$$H(z) = \frac{2(0.2344)}{(1 - 0.2344z^{-1})(1 - 0.2344z)} \longrightarrow h(n) = 0.4960(0.2344)^{|n|}$$

$$\xi_{min} = \sigma_v^2 h(0) = (0.25)(0.4960) = 0.1240$$

Without filtering $d(\hat{n}) = x(n)$ and mean square error
$E\left\{|e(n)|^2\right\} = \left\{|v(n)|^2\right\} = 0.25$

Causal Wiener Filter

$$\hat{h}(n) = x(n) * h(n) = \sum_{l=0}^{\infty} h(k)x(n-k)$$

$$r_{dx}(k) = \sum_{l=0}^{\infty} h(l) r_x(k-l); 0 \leq k < \infty$$

Let's assume that $g(n)$ is the Wiener Filter for unit variance noise $\epsilon(n)$

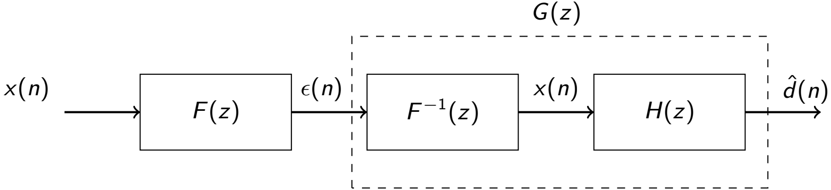
$$r_{d\epsilon}(k) = \sum_{l=0}^{\infty} g(l) r_\epsilon(k-l); 0 \leq k < \infty$$

Since $r_\epsilon(k) = \delta(k)$ $g(n) = r_{d\epsilon} u(n)$ and $G(z) = [P_{d\epsilon}(z)]_+$
If $x(n)$ is a random process with $P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$ and $Q(z)$ is minimum phase then $F(z) = 1/\sigma_0 Q(z)$ is a whitening filter.

$$P_\epsilon(z) = P_x(z) F(z) F^*(1/z^*) = 1$$

Since $Q(z)$ is minimum phase, $F(z)$ is causal and stable whose inverse is also causal and stable.

$$G(z) = F^{-1}(z) H(z)$$



$$r_{d\epsilon}(k) = E\left\{d(n)\epsilon^*(n-k)\right\} = E\left\{d(n)\left[\sum_{l=-\infty}^{\infty} f(l)x(n-k-l)\right]^*\right\}$$

$$r_{d\epsilon}(k) = \sum_{l=-\infty}^{\infty} f^*(l) r_{dx}(k+l) \longrightarrow P_{d\epsilon}(z) = P_{dx}(z)F^*(1/z^*) = \frac{P_{dx}(z)}{\sigma_0 Q^*(1/z^*)}$$

$$G(z) = \frac{1}{\sigma_0}\left[\frac{P_{dx}(z)}{Q^*(1/z^*)}\right]_+$$

$$H(z) = F(z)G(z) = \frac{1}{\sigma_0^2 Q(z)}\left[\frac{P_{dx}(z)}{Q^*(1/z^*)}\right]_+$$

$$\xi_{min} = r_d(0) - \sum_{l=0}^{\infty} h(l) r_{dx}^*(l) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - H(e^{j\omega})P_{dx}^*(e^{j\omega})\right] d\omega$$

$x(n) = d(n) + v(n)$, $v(n)$ is unit variance noise and is uncorrelated with $d(n)$

$d(n) = 0.8d(n-1) + w(n)$ with $\sigma_w^2 = 0.36$ and $r_d(k) = 0.8^{|k|}$

$$\begin{aligned} P_{dx}(z) &= P_d(z) = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} \\ P_x(z) &= P_d(z) + P_v(z) \\ P_x(z) &= P_d(z) + 1 \end{aligned}$$

$$P_x(z) = 1 + \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} = 1.6\frac{(1-0.5z^{-1})(1-0.5z)}{(1-0.8z^{-1})(1-0.8z)}$$

Since $x(n)$ is real $P_x(z) = \sigma_0^2 Q(z)Q(z^{-1})$, $Q(z) = \frac{(1-0.5z^{-1})}{(1-0.8z^{-1})}$ and $\sigma_0^2 = 1.6$

$$H(z) = \frac{1}{\sigma_0^2 Q(z)}\left[\frac{P_{dx}(z)}{Q(z^{-1})}\right]_+$$

$$\frac{P_{dx}(z)}{Q(z^{-1})} = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)}\frac{(1-0.8z)}{(1-0.5z)} = \frac{0.6}{1-0.8z^{-1}} + \frac{0.3}{z^{-1}-0.5}$$

$$H(z) = \frac{1}{1.6}\frac{(1-0.8z^{-1})}{(1-0.5z^{-1})}\frac{0.6}{(1-0.8z^{-1})} = \frac{0.375}{1-0.5z^{-1}} \longrightarrow h(n) = 0.375\left(\frac{1}{2}\right)^n u(n)$$

$$\xi_{min} = r_d(0) - \sum_{l=0}^{\infty} h(l) r_{dx}(l) = 0.375$$

Noncausal Filter

$$H(z) = \frac{P_{dx}(z)}{P_x(z)} = \frac{P_d(z)}{P_x(z)} = \frac{0.36/1.6}{(1-0.5z^{-1})(1-0.5z)} \longrightarrow h(n) = \frac{3}{10}\left(\frac{1}{2}\right)^{|n|}$$

$$\xi_{min} = r_d(0) - \sum_{l=-\infty}^{\infty} h(l) r_{dx}(l) = 0.3$$

Causal Filter $\quad h(n) = 0.375\left(\tfrac{1}{2}\right)^n u(n)$

$$\hat{D}(z) = H(z)X(z) \longrightarrow \hat{D}(z)/X(z) = H(z) = \frac{0.375}{1-0.5z^{-1}}$$

$$\hat{d}(n) = 0.5\hat{d}(n-1) + 0.375x(n) = 0.8\hat{d}(n-1) + \underbrace{0.375}_{Gain}\underbrace{\left[x(n) - \overbrace{0.8\hat{d}(n-1)}^{\hat{x}(n)}\right]}_{innovation}$$

Causal Linear Prediction

$$\hat{x}(n+1) = \sum_{k=0}^{\infty} h(k)x(n-k)$$

$$d(n) = x(n+1) \rightarrow r_{dx}(k) = r_x(k+1) \rightarrow P_{dx}(z) = zP_x(z)$$
$$P_x(z) = \sigma_0^2 Q(z)Q^*(1/z^*)$$

$$H(z) = \frac{1}{\sigma_0^2 Q(z)}\left[\frac{zP_x(z)}{Q^*(1/z^*)}\right]_+ = \frac{1}{Q(z)}\left[zQ(z)\right]_+$$

$$\begin{aligned}\left[zQ(z)\right]_+ &= z(1 + q(1)z^{-1} + q(2)z^{-2}+, \dots) \\ &= q(1) + q(2)z^{-1} = q(1) + q(2)z^{-1} + q(3)z^{-2} + \dots = z\left[Q(z) - 1\right]\end{aligned}$$

$$H(z) = \frac{1}{Q(z)} z\left[Q(z) - 1\right] = z\left[1 - \frac{1}{Q(z)}\right]$$

$x(n) = 0.9x(n-1) - 0.2x(n-2) + w(n)$ and $w(n)$ is unit variance white noise

$$P_x(z) = \frac{1}{A(z)A(z^{-1})} = \frac{1}{(1-0.9z^{-1}+0.2z^{-2})(1-0.9z^{1}+0.2z^{2})}$$

$$H(z) = z\left[A(z)-1\right] = 0.9 - 0.2z^{-1}$$

$$d(n) = \hat{x}(n+1) \rightarrow \mathcal{Z}^{-1}\left\{H(z)X(z)\right\} = 0.9x(n) - 0.2x(n-1)$$

ARMA Process : $x(n) + 0.8x(n-1) = w(n) - 0.6w(n-1) + 0.36w(n-2)$

$$P_x(z) = \frac{(1-0.6z^{-1}+0.36z^{-2})(1-0.6z^{1}+0.36z^{2})}{(1-0.8z^{-1})(1-0.8z)}$$

$$H(z) = z\left[1 - \frac{1}{Q(z)}\right] \rightarrow Q(z) = \frac{(1-0.6z^{-1}+0.36z^{-2})}{(1-0.8z^{-1})}$$

$$H(z) = \frac{0.2+0.36z^{-1}}{1-0.6z^{-1}+0.36z^{-2}}$$

$$\hat{x}(n+1) = 0.6\hat{x}(n) - 0.36\hat{x}(n-1) + 0.2x(n) + 0.36x(n-1)$$

Kalman Approach

$$y(n) = x(n) + v(n)$$

Wiener solution requires

1. $x(n)$ and $y(n)$ must be jointly WS stationary processes.
2. The filter cannot be turned on at time $n = 0$ *i.e.* the observations are assumed to be available for $k \leq n$.

$$x(n) = \sum_{k=1}^{p} a(k)x(n-k) + w(n)$$

$$\mathbf{x}(n) = [x(n), x(n-1) \cdots, x(n-p+1)]^T$$

$$\mathbf{x}(n) = \begin{bmatrix} a(1) & a(2) & \cdots & a(p-1) & a(p) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w(n)$$

$$y(n) = [1, 0, \ldots, 0]\, \mathbf{x}(n) + v(n)$$

State Space Model

$$\begin{aligned}\mathbf{x}(n) &= \mathbf{A}\mathbf{x}(n-1) + \mathbf{w}(n)\\ \mathbf{y}(n) &= \mathbf{C}^T\mathbf{x}(n) + \mathbf{v}(n)\end{aligned}$$

$\mathbf{A}$: State transition matrix
$\mathbf{w}(n) = [w(n), 0, \ldots, 0]^T$: noise vector
For an AR(1) process, optimum estimate of $\mathbf{x}(n)$ using all measurements up to time n are

$$\hat{\mathbf{x}}(n) = \mathbf{A}\hat{\mathbf{x}}(n-1) + \mathbf{K}\left[y(n) - \mathbf{c}^T\mathbf{A}\hat{\mathbf{x}}(n-1)\right]$$

For a nonstationary process

$$\mathbf{x}(n) = \mathbf{A}(n-1)\mathbf{x}(n-1) + \mathbf{w}(n) \qquad \mathbf{y}(n) = \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n)$$

$$E\left\{\mathbf{w}(n)\mathbf{w}^H(k)\right\} = \begin{cases}\mathbf{Q}_w(n) & k = n\\ \mathbf{0} & k \neq 0\end{cases} \qquad E\left\{\mathbf{v}(n)\mathbf{v}^H(k)\right\} = \begin{cases}\mathbf{Q}_v(n) & k = n\\ \mathbf{0} & k \neq 0\end{cases}$$

$$\hat{\mathbf{x}}(n) = \mathbf{A}(n-1)\hat{\mathbf{x}}(n-1) + \mathbf{K}(n)\left[\mathbf{y}(n) - \mathbf{C}(n)\mathbf{A}(n-1)\hat{\mathbf{x}}(n-1)\right]$$

Definitions
$\hat{\mathbf{x}}(n|n)$: Best linear estimate of $\mathbf{x}(n)$ given the observations $\mathbf{y}(1), \mathbf{y}(2), \ldots, \mathbf{y}(n)$.
$\mathbf{e}(n|n)$: Corresponding state estimation error
$\mathbf{P}(n|n)$: Covariance of error $\mathbf{e}(n|n)$

$$\begin{aligned} \mathbf{e}(n|n) &= \mathbf{x}(n) - \hat{\mathbf{x}}(n|n) \\ \mathbf{e}(n|n-1) &= \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1) \end{aligned}$$

and

$$\begin{aligned} \mathbf{P}(n|n) &= E\left\{\mathbf{e}(n|n)\mathbf{e}^H(n|n)\right\} \\ \mathbf{P}(n|n-1) &= E\left\{\mathbf{e}(n|n-1)\mathbf{e}^H(n|n-1)\right\} \end{aligned}$$

Two step process for state estimation:
1. Prediction

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n-1)\hat{\mathbf{x}}(n-1|n-1)$$

and its error

$$\mathbf{e}(n|n-1) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n-1)\mathbf{e}(n-1|n-1) + \mathbf{w}(n)$$

If we assume that $\hat{\mathbf{x}}(n-1|n-1)$ is an unbiased estimate of $\mathbf{x}(n-1)$,
$\rightarrow E\{e(n-1|n-1) = 0\} \rightarrow E\{e(n|n-1)\} = 0$

$$\mathbf{P}(n|n-1) = \mathbf{A}(n-1)\mathbf{P}(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n)$$

2. Correction

Predicted $\hat{\mathbf{x}}(n|n-1)$ is corrected with the new observation $\mathbf{y}(n)$ by a *linear* and *unbiased* estimate of $\hat{\mathbf{x}}(n|n)$ which will minimize the mean square error $E\left\{|e(n|n)|^2\right\}$

$$\hat{\mathbf{x}}(n|n) = \mathbf{K}^{'}(n)\hat{x}(n|n-1) + \mathbf{K}(n)\mathbf{y}(n)$$

$$\begin{aligned}
\mathbf{e}(n|n) &= \mathbf{x}(n) - \mathbf{K}^{'}(n)\hat{\mathbf{x}}(n|n-1) - \mathbf{K}(n)\mathbf{y}(n) \\
&= \mathbf{x}(n) - \mathbf{K}^{'}(n)\left[\mathbf{x}(n) - \mathbf{e}(n|n-1)\right] - \mathbf{K}(n)\left[\mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n)\right] \\
&= \left[\mathbf{I} - \mathbf{K}'(n) - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{x}(n) + \mathbf{K}'(n)\mathbf{e}(n|n-1) - \mathbf{K}(n)\mathbf{v}(n)
\end{aligned}$$

For an unbiased $\hat{\mathbf{x}}(n|n) \to \mathbf{K}'(n) = \mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)$

$$\begin{aligned}
\hat{\mathbf{x}}(n|n) &= [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)\mathbf{y}(n) \\
&= \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)]
\end{aligned}$$

$$\begin{aligned}
\mathbf{e}(n|n) &= \mathbf{K}^{'}(n)\mathbf{e}(n|n-1) - \mathbf{K}(n)\mathbf{v}(n) \\
&= \left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{e}(n|n-1) - \mathbf{K}(n)\mathbf{v}(n)
\end{aligned}$$

$$\mathbf{P}(n|n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\,[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]^H + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}^H(n)$$

We determine the *Kalman Gain* $\mathbf{K}(n)$ by minimizing the mean square error by

$$\frac{d}{d\mathbf{K}}\xi(n) = E\left[\mathbf{e}(n|n)\mathbf{e}^H(n|n)\right] = \frac{d}{d\mathbf{K}}Tr[\mathbf{P}(n|n)]$$

$\frac{Tr[\mathbf{KA}]}{d\mathbf{K}} = \mathbf{A}^H$ and $\frac{Tr[\mathbf{KAK}^H]}{d\mathbf{K}} = 2\mathbf{KA}$

$\frac{d}{d\mathbf{K}}Tr[\mathbf{P}(\mathbf{n}|\mathbf{n})] = -2[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)\mathbf{C}^H(n) + 2\mathbf{K}(n)\mathbf{Q}_v(n) = 0$

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)]^{-1}$$

$$\begin{aligned}\mathbf{P}(n|n) &= [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1) \\ &= -\left\{[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{K}(n)\mathbf{Q}_v(n)\right\}\mathbf{K}^H(n)\end{aligned}$$

The second term is zero

$$\mathbf{P}(n|n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)$$

Kalman Filter Algorithm

1. Initialize $\hat{x}(0|0)$ and $P(0|0)$
2. For n=1,2,... compute,
   $\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n-1)\hat{\mathbf{x}}(n-1|n-1)$
   $\mathbf{P}(n|n-1) = \mathbf{A}(n-1)\mathbf{P}(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n)$
   $\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)]^{-1}$
   $\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)$
   $\mathbf{P}(n|n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)$

Estimation of an unknown constant

$$y(n) = x(n) + v(n)$$

where $x(n) = x(n-1)$ is the constant to be estimated.
$\mathbf{A}(n) = 1$, $\mathbf{C}(n) = 1$, $\mathbf{Q}_w(n) = 0$, $\mathbf{Q}_v(n) = \sigma_v^2$

$$P(n|n-1) = P(n-1|n-1)$$

$$K(n) = P(n-1)[P(n-1) + \sigma_v^2]^{-1}$$

$$P(n) = [1 - K(n)]\, P(n-1) = \frac{P(n-1)\sigma_v^2}{P(n-1) + \sigma_v^2} = \frac{P(0)\sigma_v^2}{nP(0) + \sigma_v^2}$$

$$K(n) = \frac{P(n-1)}{P(n-1) + \sigma_v^2} = \frac{P(0)}{nP(0) + \sigma_v^2}$$

$$\hat{x}(n) = \hat{x}(n-1) + \frac{P(0)}{nP(0) + \sigma_v^2}\,[y(n) - \hat{x}(n-1)]$$

1. As $n \to \infty$ then $K(n) = 0$

   $\hat{x}(n)$ reaches its steady state value.

2. If there is no *a priori* information about x then $\hat{x}(0) = 0$ and $P(0) = \infty$.

   In that case $K(n) = 1/n$ and $\hat{x}(n) = \hat{x}(n-1) + \frac{1}{n}[y(n) - \hat{x}(n-1)]$

   which can be expressed as

   $$\hat{x}(n) = \frac{n-1}{n}\hat{x}(n-1) + \frac{1}{n}y(n) = \frac{1}{n}\sum_{k=1}^{n} y(k)$$

   On the other hand, this is the recursive implementation of sample mean.

Estimation of AR(1) process : $x(n) = 0.8x(n-1) + w(n)$, $\sigma_w^2 = 0.36$

$$y(n) = x(n) + v(n)$$

where $v(n)$ is a unit variance white noise and uncorrelated $w(n)$.
$\mathbf{A}(n) = 1,\ C(n) = 1$

$$\hat{x}(n) = 0.8\hat{x}(n-1) + K(n)\left[y(n) - 0.8\hat{x}(n-1)\right]$$

$$\begin{aligned} P(n|n-1) &= (0.8)^2 P(n-1|n-1) + 0.36 \\ K(n) &= P(n|n-1)[P(n|n-1)+1]^{-1} \\ P(n|n) &= [1-K(n)]P(n|n-1) \end{aligned}$$

with $\hat{x}(0) = 0$ and $P(0|0) = E\left\{|x(0)|^2\right\} = 1$

```
Pn1n1 =1;
for i=1:10,
 Pnn1 = 0.8*0.8 *Pn1n1 + 0.36;
 k = Pnn1/(Pnn1+1);
 Pn1n1 = (1-k)*Pnn1 ;
 K(i) = k ;
end;
```

n	$K(n)$
1.0000	0.5000
2.0000	0.4048
3.0000	0.3824
4.0000	0.3768
5.0000	0.3755
6.0000	0.3751
7.0000	0.3750
8.0000	0.3750
9.0000	0.3750
10.0000	0.3750

In steady state it yields the Wiener filter:

$$\hat{x}(n) = 0.8\hat{x}(n-1) + 0.375\left[x(n) - 0.8\hat{x}(n-1)\right]$$

Estimation of AR(1) parameter by Kalman Filter : $y(n) = ay(n-1) + v(n)$, $v(n)$: white noise variance with σ_v^2

State vector :	$x(n) = a$
State transition matrix :	$A = 1$
State Noise Covariance :	$Q_w = 0$
Output Noise Covariance :	$Q_v = \sigma_v^2$
Output matrix :	$C = y(n-1)$

In state space form

$$x(n+1) = x(n), \quad y(n) = Cx(n) + v(n)$$

1. Initialize $\hat{x}(0|0)$ and $P(0|0)$
2. For n=1,2,... compute,

$$\hat{x}(n|n-1) = \hat{x}(n-1|n-1)$$
$$P(n|n-1) = P(n-1|n-1)$$
$$K(n) = P(n|n-1)y(n-1)[y(n-1)P(n|n-1)y(n-1) + \sigma_v^2]^{-1}$$
$$\hat{x}(n|n) = \hat{x}(n|n-1) + K(n)[y(n) - y(n-1)\hat{x}(n|n-1)]$$
$$P(n|n) = [1 - K(n)y(n-1)]\, P(n|n-1)$$

$y(n) = 0.4y(n-1) + v(n)$ with $\sigma_v^2 = 0.1$

```
N=3000;
a =0.4;
sigma_v2 =0.1;
v=sqrt(sigma_v2)*randn(N,1);
y=filter(1,[1 -a],v);
    %Kalman Filtering
% Initialize
x_n=0.1; P_n=0.1;
for n=2:N,
x_n1 =x_n;
P_n1 =P_n;
C=y(n-1);
K = P_n1*C/(C*P_n1*C + sigma_v2);
x_n = x_n1 + K*(y(n)-C*x_n1);
P_n = (1-K*C)*P_n1;
x(n)=x_n;
end;
close;
plot(x(2:end));grid
```