Chapter 3

■ **Optimal Filter Theory and Applications**

**References:**

■ B.Widrow and S.D.Stearns, *Adaptive Signal Processing*, Prentice-Hall, 1985

■ S.M.Stearns and D.R.Hush *Digital Signal Analysis*, Prentice-Hall, 1990

■ P.M.Clarkson, *Optimal and Adaptive Signal Processing*, CRC Press, 1993

■ S.Haykin, *Adaptive Filter Theory*, Prentice-Hall, 2002

Optimal Signal Processing is concerned with the design, analysis, and implementation of processing system that extracts information from sampled data in a manner that is 'best' or optimal in some sense. Such processing systems can be referred to as optimal filters.

Basic Classes of Optmal Filtering Applications

1.**Prediction**: use previous samples to predict current samples
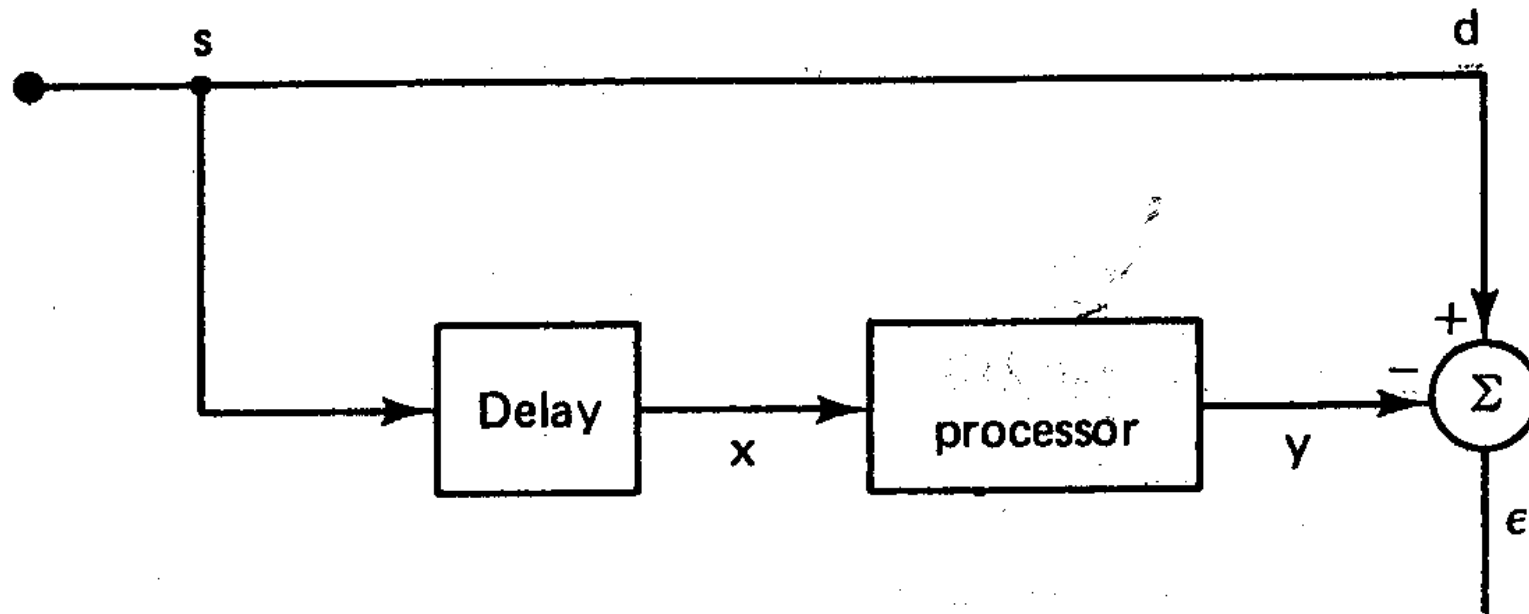
- <u>Speech Modeling using Linear Predictive Coding (LPC)</u>

Since speech signals are highly correlated, a speech signal $s(n)$ can be accurately modeled by a linear combination of its past samples:
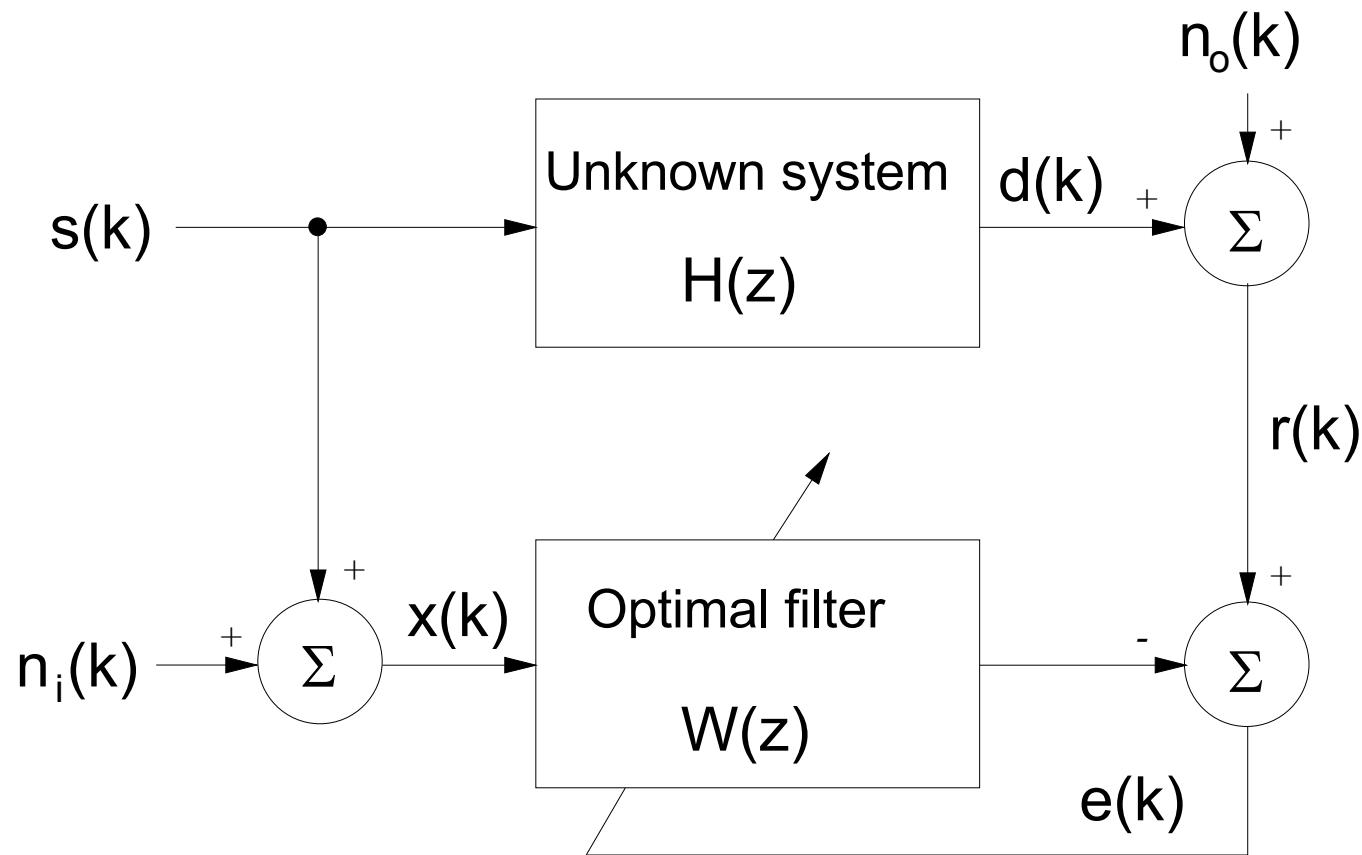
$$s(n) \approx \hat{s}(n) = \sum_{i=1}^{P} w_i s(n-i)$$

where $\{w_i\}$ are known as the LPC coefficients. Techniques of optimal signal processing can be used to determine $\{w_i\}$ in an optimal way.
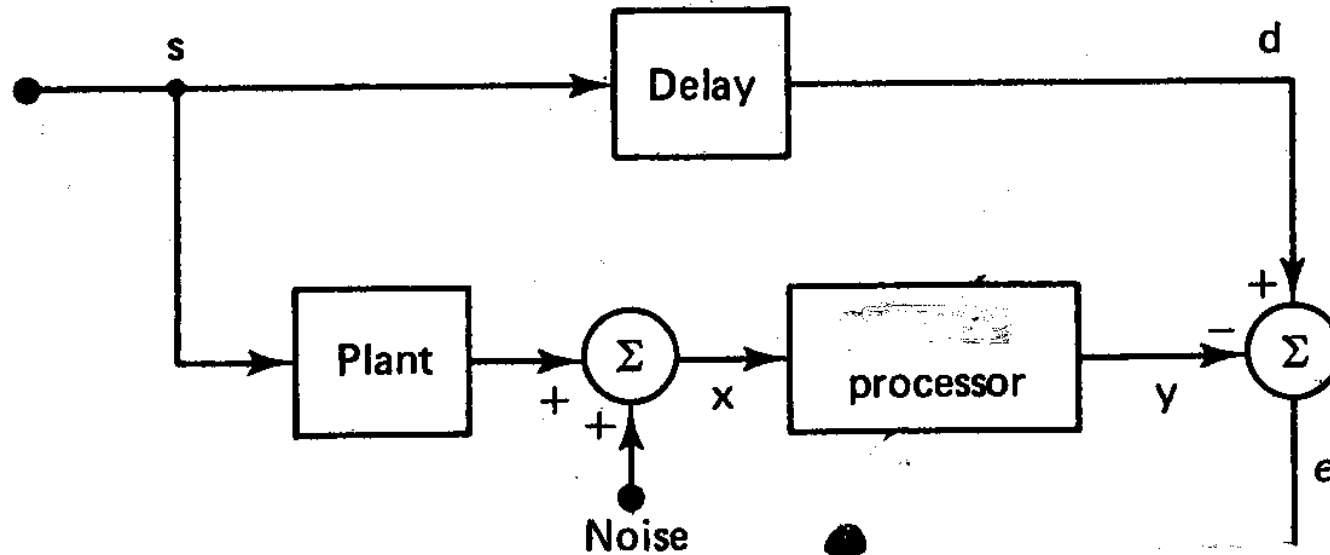
2. **Identification**
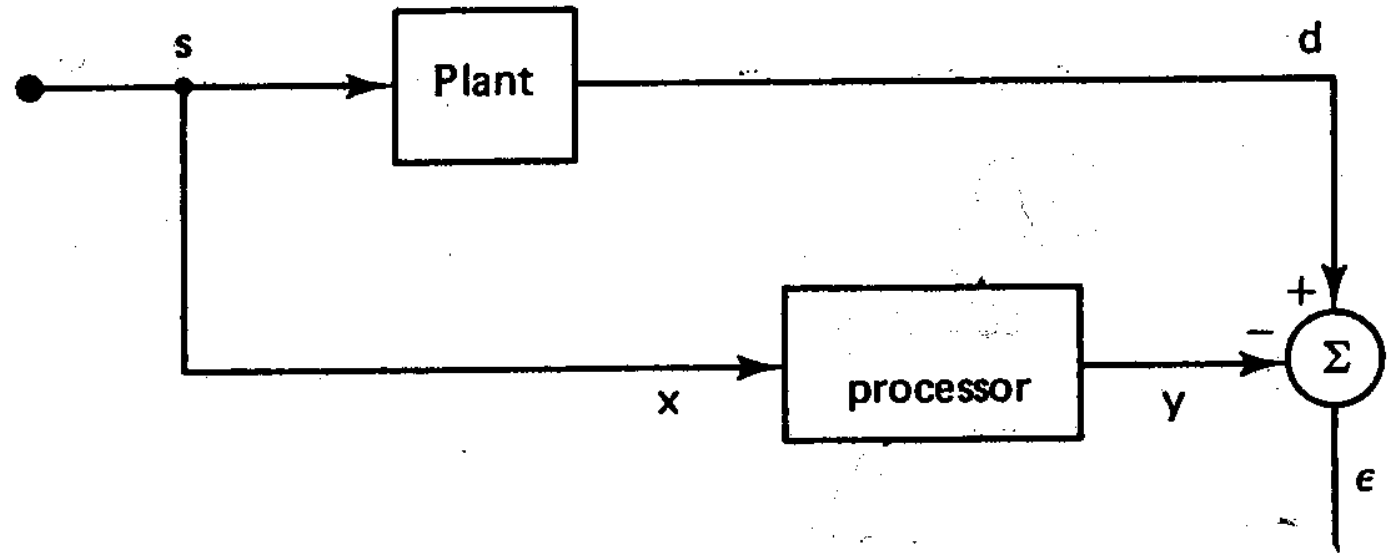
- **System Identification**

Given noisy input $x(k)$ and/or noisy output $r(k)$, our aim is to determine the impulse response of the unknown system $H(z)$ using $W(z)$

3. **Inverse Filtering**: find the inverse of the system
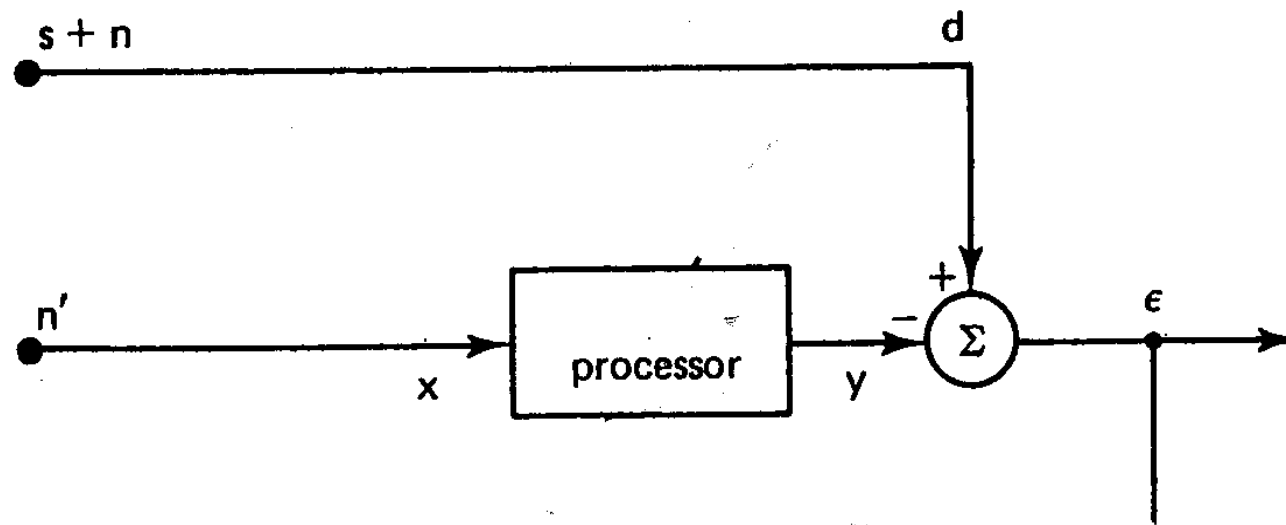

- Signal Recovery

Given a noisy discrete-time signal:

$$x(k) = s(k) \otimes h(k) + w(k)$$

where $s(k)$, $h(k)$ and $w(k)$ represent the signal of interest, unknown impulse response and noise, respectively. Optimal signal processing can be used to recover $s(k)$ in an optimal way.

4. **Interference Canceling**: Remove noise using an external reference

- <u>Interference Cancellation in Electrocardiogram (ECG) Recording</u>

In biomedical engineering, the measured ECG signal $r(n)$ is corrupted by the 50Hz power line interference:

$$r(n) = s(n) + i(n)$$

where $s(n)$ is the noise-free ECG and $i(n)$ represents the 50Hz interference. An external reference for $i(n)$ is another 50Hz signal.

Problem Statement for Optimal Filters
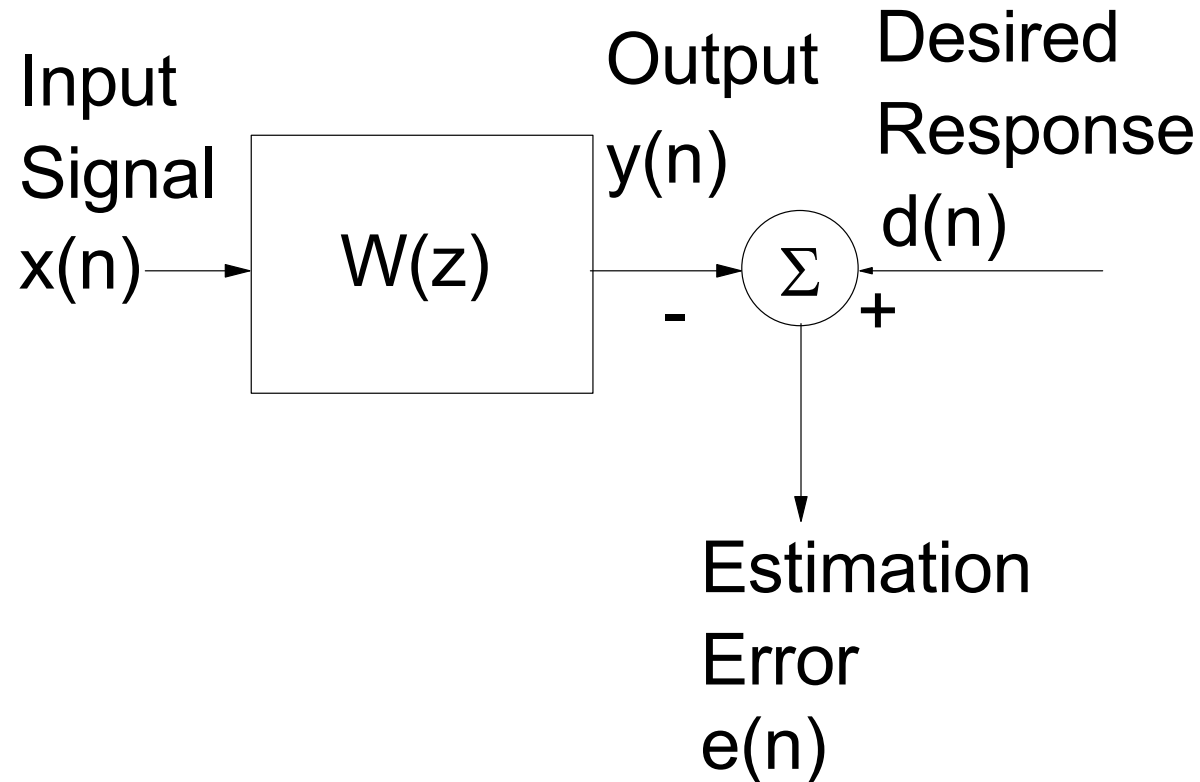
Given the input $x(n)$ and the desired response $d(n)$, we want to find the transfer function $W(z)$ or its impulse response such that a statistical criterion or a cost function is optimized.

Some common optimization criteria in the literature are:

1. *Least Squares* : find $W(z)$ that minimizes $\sum\limits_{n=0}^{N-1} e^2(n)$ where N is the number of samples available. This corresponds to least-squares filter design.

2. *Minimum Mean Square Error* : find $W(z)$ that minimizes $E\{e^2(n)\}$. This corresponds to Wiener filtering problem.

3. *Least Absolute Sum* : find $W(z)$ that minimizes $\sum\limits_{n=0}^{N-1} |e(n)|$.

4. *Minimum Mean Absolute Error* : find $W(z)$ that minimizes $E\{|e(n)|\}$.

5. *Least Mean Fourth* : find $W(z)$ that minimizes $E\{e^4(n)\}$.

The first and second are two commonly used criteria because of their relatively small computation, ease of analysis and robust performance. In later sections it is shown that both viewpoints give rise to similar mathematical expression for $W(z)$.

Q.: An absolute optimization criterion does not exist? Why?

Least Squares Filtering

For simplicity, we assume $W(z)$ is a causal FIR filter of length L so that

$$W(z) = \sum_{i=0}^{L-1} w_i z^{-i} \qquad (3.1)$$

The error function $e(n)$ is thus given by

$$e(n) = d(n) - y(n) \qquad (3.2)$$

where

$$y(n) = \sum_{i=0}^{L-1} w_i x(n-i) = \underline{W}^T \underline{X}(n),$$

$$\underline{W} = \begin{bmatrix} w_0 & w_1 & \cdots w_{L-2} & w_{L-1} \end{bmatrix}^T$$

$$\underline{X}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots x(n-L+2) & x(n-L+1) \end{bmatrix}^T$$

The cost function is

$$J_{LS}(\underline{W}) = \sum_{n=0}^{N-1} e^2(n) = \sum_{n=0}^{N-1}\left[d(n) - \sum_{i=0}^{L-1} w_i x(n-i)\right]^2 \qquad (3.3)$$

which is a function of the filter coefficients $\{w_i\}$ and N is the number of $x(n)$ (and $d(n)$).

The minimum of the least squares function can be found by differentiating $J_{LS}(\underline{W})$ with respect to $w_0, w_1, \cdots, w_{L-1}$ and then setting the resultant expressions to zero as follows,

$$\frac{\partial J_{LS}(\underline{W})}{\partial w_j} = \frac{\partial}{\partial w_j}\left\{\sum_{n=0}^{N-1}\left(d(n) - \sum_{i=0}^{L-1} w_i x(n-i)\right)^2\right\} = 0, \qquad j = 0,1,\cdots,L-1$$

$$\Rightarrow 2\sum_{n=0}^{N-1}\left(d(n) - \sum_{i=0}^{L-1} w_i x(n-i)\right)(-x(n-j)) = 0 \qquad (3.4)$$

$$\Rightarrow \sum_{n=0}^{N-1} d(n)x(n-j) = \sum_{n=0}^{N-1}\sum_{i=0}^{L-1} w_i x(n-i)x(n-j) = \sum_{i=0}^{L-1} w_i \sum_{n=0}^{N-1} x(n-i)x(n-j)$$

Denote

$$\underline{\hat{R}}_{dx} = [\hat{R}_{dx}(0) \quad \hat{R}_{dx}(-1) \quad \cdots \quad \hat{R}_{dx}(-L+2) \quad \hat{R}_{dx}(-L+1)]^T \tag{3.5}$$

where

$$\hat{R}_{dx}(-j) = \sum_{n=0}^{N-1} d(n)x(n-j), \qquad j = 0,1,\cdots,L-1$$

and

$$\underline{\hat{R}}_{xx} = \begin{bmatrix} \hat{R}_{xx}(0,0) & \hat{R}_{xx}(-1,0) & \cdots & \hat{R}_{xx}(-L+2,0) & \hat{R}_{xx}(-L+1,0) \\ \hat{R}_{xx}(0,-1) & \ddots & & & \\ \cdots & & \ddots & & \vdots \\ \hat{R}_{xx}(0,-L+2) & & & \ddots & \\ \hat{R}_{xx}(0,-L+1) & \hat{R}_{xx}(-1,-L+1) & & & \hat{R}_{xx}(-L+1,-L+1) \end{bmatrix}$$

where

$$\hat{R}_{xx}(-i,-j) = \sum_{n=0}^{N-1} x(n-i)x(n-j)$$

In practice, for stationary signals, we use

$$
\hat{\underline{R}}_{xx} = \begin{bmatrix} \hat{R}_{xx}(0) & \hat{R}_{xx}(-1) & \cdots & \hat{R}_{xx}(-L+2) & \hat{R}_{xx}(-L+1) \\ \hat{R}_{xx}(1) & \ddots & & & \\ \cdots & & \ddots & & \vdots \\ \hat{R}_{xx}(L-2) & & & \ddots & \\ \hat{R}_{xx}(L-1) & \hat{R}_{xx}(L-2) & & & \hat{R}_{xx}(0) \end{bmatrix} \tag{3.6}
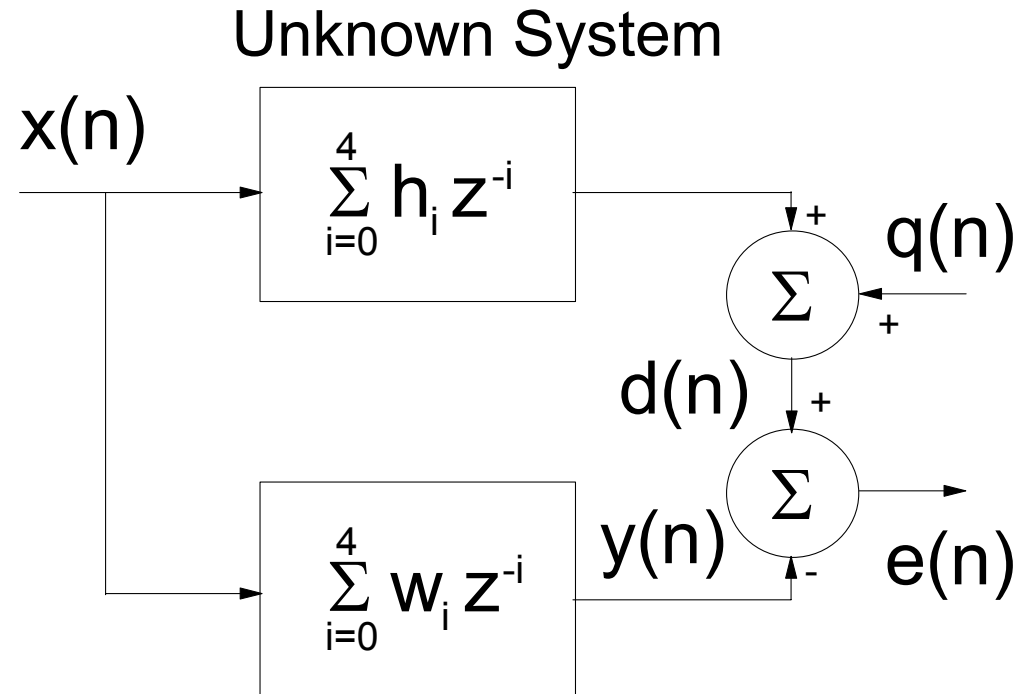$$

where

$$
\hat{R}_{xx}(i) = \sum_{n=0}^{N-1-i} x(n)x(n+i) = \hat{R}_{xx}(-i)
$$

As a result, we have

$$
\hat{\underline{R}}_{dx} = \hat{\underline{R}}_{xx} \cdot \underline{W}_{LS}
$$

$$
\Rightarrow \underline{W}_{LS} = \left(\hat{\underline{R}}_{xx}\right)^{-1} \cdot \hat{\underline{R}}_{dx} \tag{3.7}
$$

provided that $\hat{\underline{R}}_{xx}$ is nonsingular.

Example 3.1

Unknown System

$$x(n)$$

$$\sum_{i=0}^{4} h_i\, z^{-i}$$

$$+ \quad q(n)$$

$$\Sigma$$

$$+$$

$$d(n) \quad +$$

$$\Sigma$$

$$\sum_{i=0}^{4} w_i\, z^{-i} \quad y(n) \quad e(n)$$

$$-$$

In this example least squares filtering is applied in determining the impulse response of an unknown system. Assume that the unknown impulse response is causal and $\{h_i\}=\{1,2,3,2,1\}$. Given N samples of $x(n)$ and $d(n)$ where $q(n)$ is a measurement noise.

We can use MATLAB to simulate the least squares filter for impulse response estimation. The MATLAB source code is as follows,

```
%define the number of samples
N=50;
```

```
%define the noise and signal powers
noise_power = 0.0;
signal_power = 5.0;
```

```
%define the unknown system impulse response
h=[1 2 3 2 1];
```

```
%generate the input signal which is a Gaussian white noise with power 5
x=sqrt(signal_power).*randn(1,N);
```

%generate R_xx

```
corr_xx=xcorr(x);
for i=0:4
for j=0:4
R_xx(i+1,j+1)= corr_xx(N+i-j);
end
end
```

%generate the desired output plus noise

```
d=conv(x,h);
d=d(1:N)+sqrt(noise_power).*randn(1,N);
```

%generate R_dx

```
corr_xd = xcorr(d,x);
for i=0:4
R_dx(i+1) = corr_xd(N-i);
end
```

%compute the estimate channel response

```
W_ls = inv(R_xx)*(R_dx)'
```

$$\hat{\underline{R}}_{xx} = \begin{bmatrix} 251.6413 & 24.4044 & 36.4998 & -11.8182 & 4.5115 \\ 24.4044 & 251.6413 & 24.4044 & 36.4998 & -11.8182 \\ 36.4998 & 24.4044 & 251.6413 & 24.4044 & 36.4998 \\ -11.8182 & 36.4998 & 24.4044 & 251.6413 & 24.4044 \\ 4.5115 & -11.8182 & 36.4998 & 24.4044 & 251.6413 \end{bmatrix}$$

$$\hat{\underline{R}}_{dx} = [390.8245 \quad 658.8827 \quad 873.3125 \quad 616.0732 \quad 334.9508]$$

$$\underline{W}_{LS} = [1.1095 \quad 2.0151 \quad 2.8291 \quad 1.8501 \quad 0.8176]^{T}$$

When N is increased to 500, we have

$$\underline{W}_{LS} = [0.9975 \quad 1.9943 \quad 2.9927 \quad 1.9942 \quad 0.9828]^{T}$$

When N is increased to 5000, we have

$$\underline{W}_{LS} = [1.0000 \quad 1.9981 \quad 2.9976 \quad 1.9972 \quad 0.9984]^{T}$$

When $N = 500$ and the noise power is 0.5 (SNR=10 dB), we have

$$\underline{W}_{LS} = [1.0158 \quad 1.9826 \quad 2.9728 \quad 1.9773 \quad 0.9925]^T$$

When $N = 500$ and the noise power is 5.0 (SNR=0 dB), we have

$$\underline{W}_{LS} = [1.0900 \quad 2.0138 \quad 2.9484 \quad 2.0249 \quad 1.0591]^T$$

It is observed that

1. The estimation accuracy improves as N increases. It is reasonable because as N increases, the accuracy of $\hat{\underline{R}}_{xx}$ and $\hat{\underline{R}}_{dx}$ increases due to more samples are involved in their computation.

2. The estimation accuracy improves as the noise power decreases.

Example 3.2

Find the least squares filter of the following one-step predictor system:
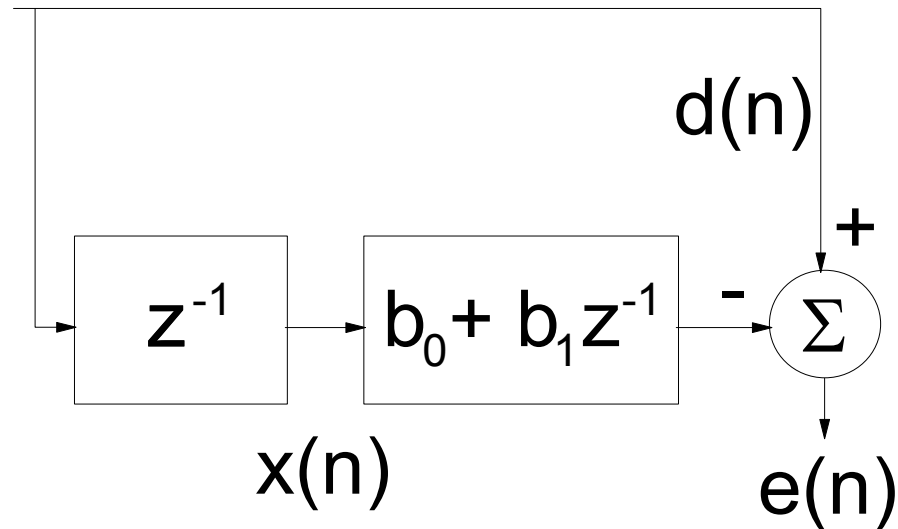
where

$$s(n) = \sqrt{2} \sin(2\pi n / 12)$$

$$x(n) = s(n-1) = \sqrt{2} \sin\left(\frac{2\pi(n-1)}{12}\right)$$

$$d(n) = s(n)$$

Given $d(n)$, the aim is to find b_0 and b_1 in least squares sense.

The MATLAB source code is as follows,

```
N=50;                              %define the number of samples

n=0:N-1;
d=sin(2.*pi.*n./12);               %generate d(n)
x= d(2:N);                         %generate x(n) from d(n)
d=d(1:N-1);                        %keep lengths of x(n) and d(n) equal

corr_xx=xcorr(x,'unbiased');       %unbiased estimate of correlation
for i=0:1
    for j=0:1
        R_xx(i+1,j+1)= corr_xx(N-1+i-j);
    end
end
```

corr_xd = xcorr(d,x,'unbiased');
for i=0:1
**   R_dx(i+1) = corr_xd(N-1-i);**
end
W_ls = inv(R_xx)*(R_dx)'

The result is:　　　　　　W_ls　=　[1.7705　　　-1.0440]

$N \rightarrow 5000$　　　　　　$\Rightarrow$　　　W_ls　=　[1.7324　　-1.0004]

$N \rightarrow 500000$　　　　　$\Rightarrow$　　　W_ls　=　[1.7321　　-1.0000]

The optimal b_0 and b_1 can be shown to be $[\sqrt{3} \quad -1]$

$\because$　for a real tone:

$$s(n) = 2\cos(\omega)s(n-1) - s(n-2)$$

$\Rightarrow$
$$s(n) = 2\cos(2\pi/12)s(n-1) - s(n-2)$$
$$= \sqrt{3}s(n-1) + (-1)s(n-2)$$

Wiener Filtering

The cost function to be minimized is

$$J_{MMSE}(\underline{W}) = E\{e^2(n)\} \tag{3.8}$$

Following the derivations in the least squares filter, the minimum of $J_{MMSE}(\underline{W})$ is found by

$$\frac{\partial J_{MMSE}(\underline{W})}{\partial w_j} = \frac{\partial}{\partial w_j}\left\{E\left\{\left(d(n) - \sum_{i=0}^{L-1}w_i x(n-i)\right)^2\right\}\right\} = 0, \quad j = 0,1,\cdots,L-1$$

$$\Rightarrow 2E\left\{\left(d(n) - \sum_{i=0}^{L-1}w_i x(n-i)\right)(-x(n-j))\right\} = 0 \tag{3.9}$$

$$\Rightarrow E\{d(n)x(n-j)\} = E\left\{\sum_{i=0}^{L-1}w_i x(n-i)x(n-j)\right\} = \sum_{i=0}^{L-1}w_i E\{x(n-i)x(n-j)\}$$

Assume $d(n)$ and $x(n)$ are jointly stationary, we have

$$R_{dx}(-j) = \sum_{i=0}^{L-1} w_i R_{xx}(i-j), \qquad j = 0,1,\cdots,L-1 \qquad (3.10)$$

Define

$$\underline{R}_{dx} = [R_{dx}(0) \quad R_{dx}(-1) \quad \cdots \quad R_{dx}(-L+2) \quad R_{dx}(-L+1)]^T \qquad (3.11)$$

and

$$\underline{R}_{xx} = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \cdots & R_{xx}(L-2) & R_{xx}(L-1) \\ R_{xx}(1) & \ddots & & & R_{xx}(L-2) \\ \vdots & & \ddots & & \vdots \\ R_{xx}(L-2) & & & \ddots & R_{xx}(1) \\ R_{xx}(L-1) & R_{xx}(L-2) & \cdots & R_{xx}(1) & R_{xx}(0) \end{bmatrix} \qquad (3.12)$$

As a result,

$$\underline{R}_{dx} = \underline{R}_{xx} \cdot \underline{W}_{MMSE}$$

$$\Rightarrow \underline{W}_{MMSE} = (\underline{R}_{xx})^{-1} \cdot \underline{R}_{dx} \qquad (3.13)$$

provided that $\underline{R}_{xx}$ is nonsingular.

Relationship between Least Squares Filter & Wiener Filter

When the number of samples $N \to \infty$ and if ergodicity holds, i.e.,

$$\lim_{N \to \infty} \{\frac{1}{N} \hat{R}_{dx}(-j)\} = \lim_{N \to \infty} \{\frac{1}{N} \sum_{n=0}^{N-1} d(n)x(n-j)\} = R_{dx}(-j) \qquad (3.14)$$

and

$$\lim_{N \to \infty} \{\frac{1}{N} \hat{R}_{xx}(-i,-j)\} = \lim_{N \to \infty} \{\frac{1}{N} \sum_{n=0}^{N-1} x(n-i)x(n-j)\}$$
$$= R_{xx}(i-j) = R_{xx}(j-i) \qquad (3.15)$$

the least squares filter is equivalent to the Wiener filter, i.e.,

$$\underline{W}_{MMSE} = \underline{W}_{LS} \qquad (3.16)$$

Properties of the Mean Square Error (MSE) Function

The MSE function $E\{e^2(n)\}$ is also known as *performance surface* and it can be written in a matrix form:

$$E\{e^2(n)\} = E\left\{\left(d(n) - \sum_{i=0}^{L-1} w_i x(n-i)\right)^2\right\} = E\left\{\left(d(n) - \underline{W}^T \underline{X}(n)\right)^2\right\}$$

$$= E\{d^2(n)\} - 2E\left\{\left(\underline{W}^T \underline{X}(n) d(n)\right)\right\} + E\left\{\left(\underline{W}^T \underline{X}(n) \cdot \left(\underline{W}^T \underline{X}(n)\right)^T\right)\right\} \quad (3.17)$$

$$= E\{d^2(n)\} - 2\underline{W}^T E\{(\underline{X}(n) d(n))\} + \underline{W}^T E\left\{\left(\underline{X}(n) \cdot \underline{X}(n)^T\right)\right\} \underline{W}$$

$$= E\{d^2(n)\} - 2\underline{W}^T \underline{R}_{dx} + \underline{W}^T \underline{R}_{xx} \underline{W}$$

1. The elements of w_i in $E\{e^2(n)\}$ appear in first degree and second degree only. This means that $E\{e^2(n)\}$ is a *quadratic* error function and thus it is *unimodal*, i.e., there is only a unique (global) minimum and no local minima exist. (However, it is only true for FIR filter but not true for IIR filter).

An example for $L = 2$ is shown below:
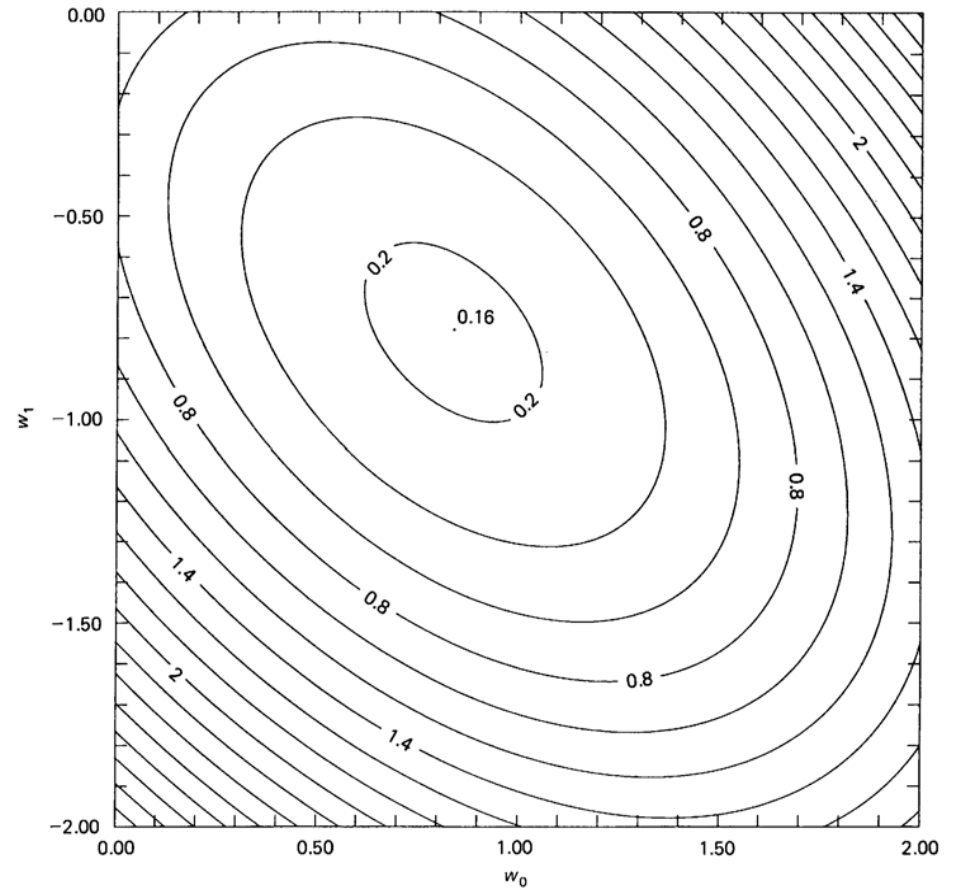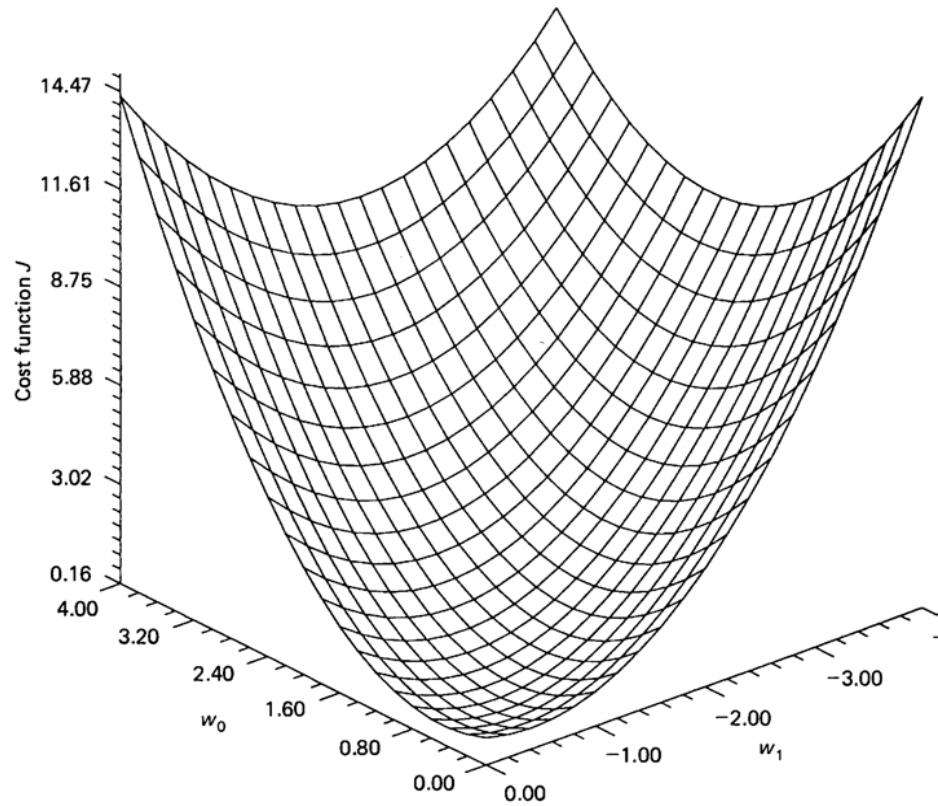
2.The minimum of $E\{e^2(n)\}$ is obtained by substituting $\underline{W} = \underline{W}_{MMSE}$:

$$\varepsilon_{min} = E\{d^2(n)\} - 2\underline{W}_{MMSE}^T \underline{R}_{dx} + \underline{W}_{MMSE}^T \underline{R}_{xx} \underline{W}_{MMSE}$$

$$= E\{d^2(n)\} - 2\left(\underline{R}_{xx}^{-1}\underline{R}_{dx}\right)^T \underline{R}_{dx} + \left(\underline{R}_{xx}^{-1}\underline{R}_{dx}\right)^T \underline{R}_{xx}\left(\underline{R}_{xx}^{-1}\underline{R}_{dx}\right)$$

$$= E\{d^2(n)\} - 2\underline{R}_{dx}^T(\underline{R}_{xx}^{-1})^T \underline{R}_{dx} + \underline{R}_{dx}^T(\underline{R}_{xx}^{-1})^T \underline{R}_{dx}$$

$$= E\{d^2(n)\} - \underline{R}_{dx}^T \underline{R}_{xx}^{-1} \underline{R}_{dx} = E\{d^2(n)\} - \underline{R}_{dx}^T \underline{W}_{MMSE}$$

(3.18)

As a result, the $E\{e^2(n)\}$ can be written as

$$E\{e^2(n)\} = \varepsilon_{min} - \varepsilon_{min} + E\{d^2(n)\} - 2\underline{W}^T \underline{R}_{dx} + \underline{W}^T \underline{R}_{xx} \underline{W}$$

$$= \varepsilon_{min} + (\underline{W} - \underline{W}_{MMSE})^T \underline{R}_{xx}(\underline{W} - \underline{W}_{MMSE})$$
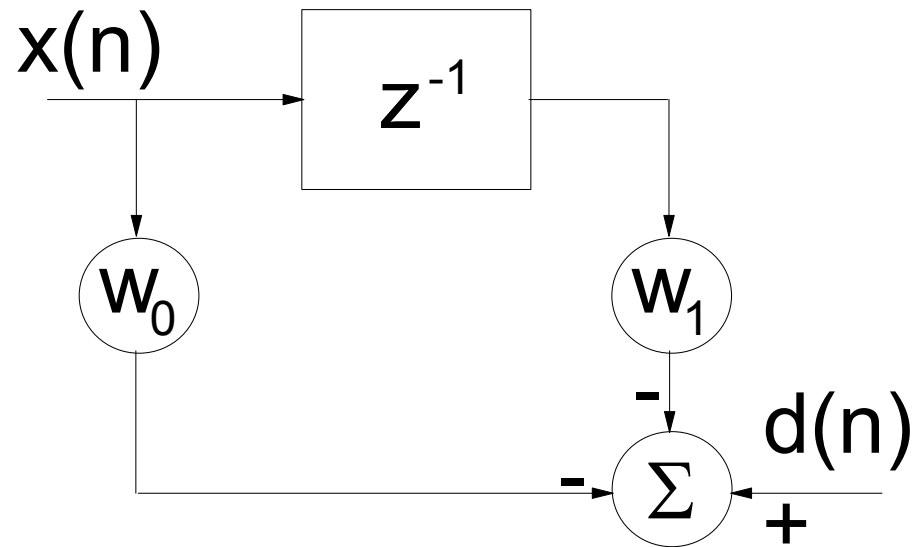
(3.19)

3.When $d(n)$ is exactly a linear combination of $x(n), x(n-1), \cdots, x(n-L+1)$,

i.e., $d(n) = \sum_{i=0}^{L-1} h_i x(n-i)$, the Wiener solution is $w_i = h_i$ for $i = 0,1,\cdots L-1$

and $\varepsilon_{min} = 0$.

Example 3.3

Determine the performance surface and the Wiener filter coefficients of the following system,

where

$$E\{d^2(n)\} = 42, \qquad \underline{R}_{xx} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \underline{R}_{dx} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

The performance surface is calculated as

$$E\{e^2(n)\} = E\{d^2(n)\} - 2\underline{W}^T \underline{R}_{dx} + \underline{W}^T \underline{R}_{xx} \underline{W}$$

$$= 42 - 2[7 \quad 8]\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} + [w_0 \quad w_1]\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

$$= 2w_0^2 + 2w_1^2 + 2w_0 w_1 - 14w_0 - 16w_1 + 42$$

While the Wiener filter weight is

$$\underline{W}_{MMSE} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Notice that the inverse of any nonsingular two-by-two matrix is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

In practice, when the $E\{d^2(n)\}$, $\underline{R}_{xx}$ and $\underline{R}_{dx}$ are not available, we can estimate them from $x(n)$ and $d(n)$ using least squares filtering method. The resultant filter coefficients are least squares filter coefficients.

Example 3.4

Find the Wiener filter of the following system:
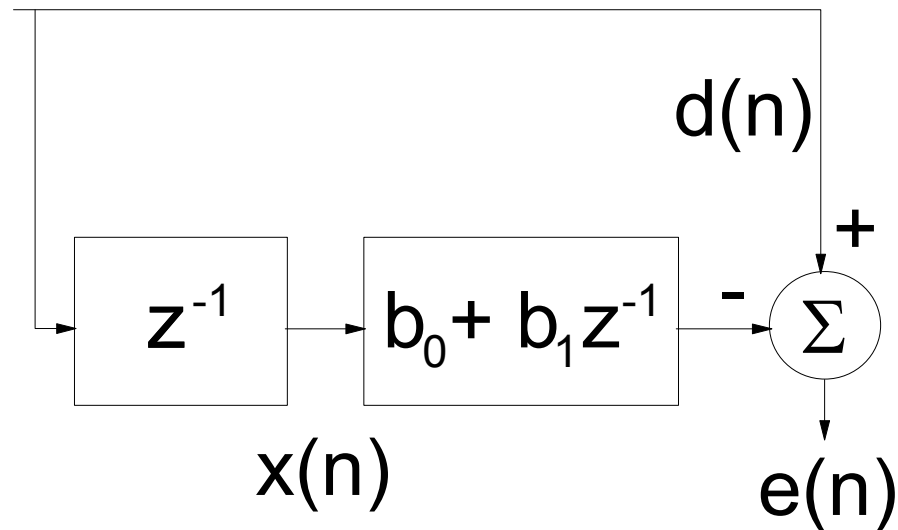
where $s(n) = \sqrt{2}\,\sin(2\pi n/12)$.

It can be seen that $x(n) = \sqrt{2}\,\sin\left(\dfrac{2\pi(n-1)}{12}\right)$ and $d(n) = s(n) = \sqrt{2}\,\sin\left(\dfrac{2\pi n}{12}\right)$.

The required statistics $R_{xx}(0)$, $R_{xx}(1)$, $R_{dx}(0)$ and $R_{dx}(-1)$ are computed as follows.

Using

$$\sin^2(A) = \frac{1 - \cos(2A)}{2}$$

and

$$\sin(A)\sin(B) = \frac{1}{2}\left(\cos(A - B) - \cos(A + B)\right)$$

then

$$R_{xx}(0) = E\left\{\sqrt{2}\sin\left(\frac{2\pi(n-1)}{12}\right) \cdot \sqrt{2}\sin\left(\frac{2\pi(n-1)}{12}\right)\right\}$$

$$= 2E\left\{\frac{1 - \cos(2(2\pi(n-1)))}{2}\right\}$$

$$= 1 + E\{\cos(4\pi(n-1))\} = 1$$

$$R_{xx}(1) = E\left\{ \sqrt{2}\sin\left(\frac{2\pi(n-1)}{12}\right) \cdot \sqrt{2}\sin\left(\frac{2\pi n}{12}\right)\right\}$$

$$= E\left\{ \cos\left(\frac{2\pi(n-1)}{12} - \frac{2\pi n}{12}\right) - \cos\left(\frac{2\pi(n-1)}{12} + \frac{2\pi n}{12}\right)\right\}$$

$$= \cos\left(\frac{-2\pi}{12}\right) = \frac{\sqrt{3}}{2}$$

$$R_{dx}(0) = E\left\{ \sqrt{2}\sin\left(\frac{2\pi(n-1)}{12}\right) \cdot \sqrt{2}\sin\left(\frac{2\pi n}{12}\right)\right\} = \frac{\sqrt{3}}{2}$$

$$R_{dx}(-1) = E\left\{ \sqrt{2}\sin\left(\frac{2\pi(n-2)}{12}\right) \cdot \sqrt{2}\sin\left(\frac{2\pi n}{12}\right)\right\} = \cos\left(\frac{-4\pi}{12}\right) = \frac{1}{2}$$

As a result,

$$\begin{bmatrix} \widetilde{b}_0 \\ \widetilde{b}_1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}$$

The performance surface is given by

$$E\{e^2(n)\} = E\{d^2(n)\} - 2\underline{W}^T \underline{R}_{dx} + \underline{W}^T \underline{R}_{xx} \underline{W}$$

$$= 1 - 2\begin{bmatrix} \dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \begin{bmatrix} b_0 & b_1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$= b_0^2 + b_1^2 + \sqrt{3}b_0 b_1 - \sqrt{3}b_0 - b_1 + 1$$

Notice that $E\{d^2(n)\} = E\{x^2(n)\} = 1$. Moreover, the minimum MSE is computed as

$$\varepsilon_{min} = E\{d^2(n)\} - \underline{R}_{dx}^T \underline{W}_{MMSE}$$

$$= 1 - \begin{bmatrix} \dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix} = 1 - \dfrac{3}{2} + \dfrac{1}{2} = 0$$

This means that the optimal predictor is able to shift the phase of the delayed sine wave and achieve exact cancellation, resulting in $\varepsilon_{min} = 0$

Questions for Discussion

1. A real sinusoid $s(n) = A\cos(\omega n + \theta)$ obeys
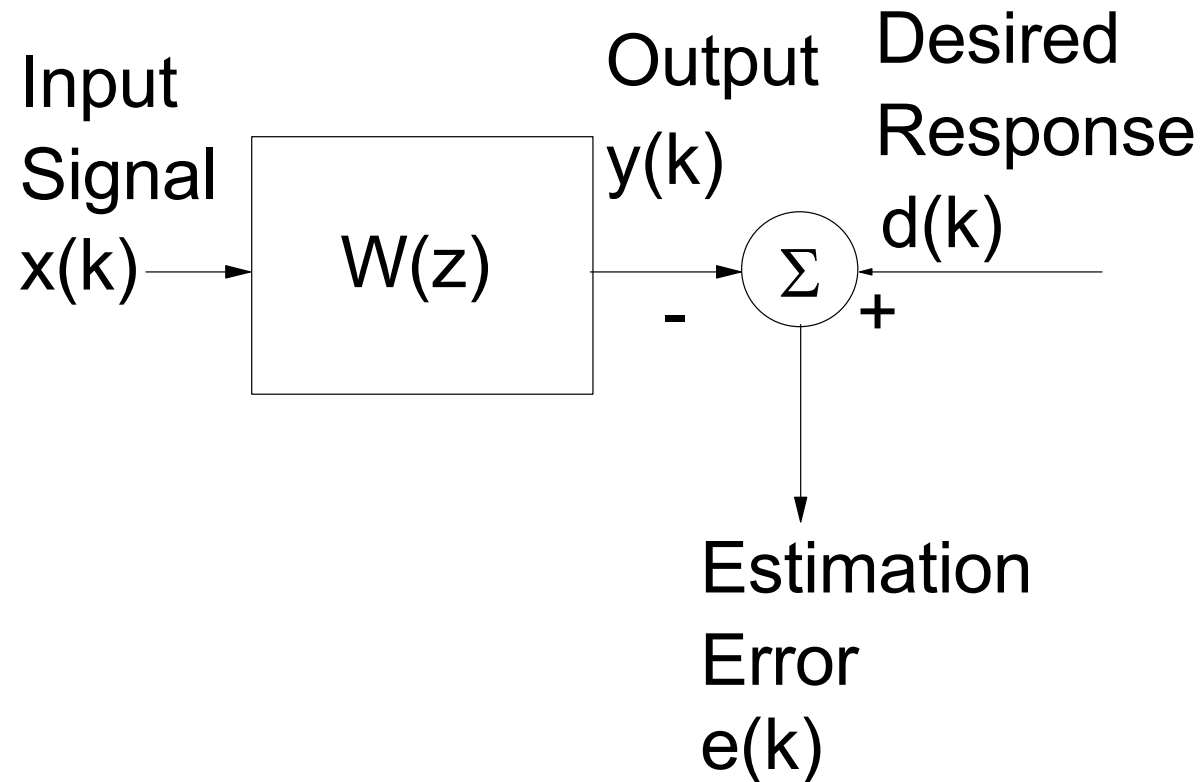
$$s(n) = a_1 \cdot s(n-1) + a_2 \cdot s(n-2)$$

where $a_1 = 2\cos(\omega)$ and $a_2 = -1$. Is $s(n)$ a 2nd order AR process?

2. Can we extend the least squares filter or Wiener filter to the general IIR system model? Try to answer this question by investigating the Wiener filter using a simple IIR model:

$$W(z) = \frac{b_0}{1 - a_1 z^{-1}}$$

That is, given $d(k)$ and $x(k)$. What are the optimal b_0 and a_1 in mean square error sense? Assume that $x(k)$ is white for simplicity.

Input
Signal
x(k) → W(z) → Output y(k) → Σ ← Desired Response d(k)

- (minus)   + (plus)

Estimation
Error
e(k)

Steps:
  (i)   develop $e(k)$

  (ii)   compute $E\{e^2(k)\}$ in terms of R_{xx} and R_{dx} only.

  (iii)   differentiate $E\{e^2(k)\}$ w.r.t. b_0 and a_1

3. Suppose you have a ECG signal corrupted by 50Hz interference:
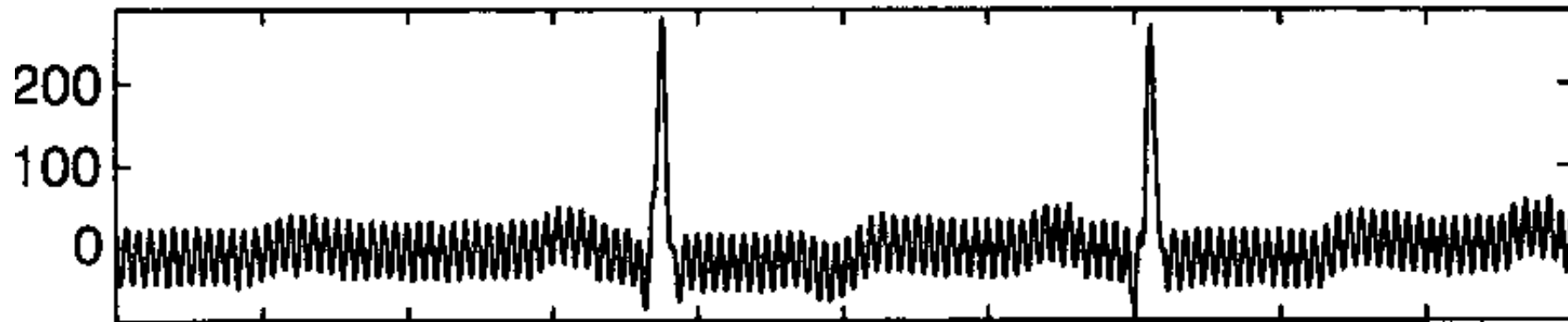
Suggest methods to eliminate/reduce the 50Hz interference.