Exercise 3:

P_1: Let $x(n)$ be an AR(2) process, i.e.,

$$x(n) + a_1^0(n-1) + a_2^0 x(n-2) = w(n).$$

- Derive the one-step forward linear predictor of order p and the corresponding mean-square prediction error in terms of the auto-correlation function $\gamma_x(m)$ for $p = 1, 2$, respectively.
- With $a_1^0 = -1$, $a_2^0 = 0.6$ and $\sigma_w^2 = 1$, compute $\gamma_x(m)$ for $m = 0, 1, 2, 3, 4, 5$. Then specify the predictors obtained above.

Solution:

- *Optimal one-step forward linear predictor:*

$$\hat{x}(n) = -a_1 x(n-1) - a_2 x(n-2) - \cdots - a_p x(n-p)$$

Denote $e(n) = x(n) - \hat{x}(n)$ as the prediction error. The optimal predictor is the solution to $\min_{\{a_k\}} E[|e(n)|^2]$. According to Theorem 2, one has

$$E[e(n)x^*(n-m)] = 0, \ 1 \le m \le p$$

which leads to (i) the optimal coefficients

$$E[e(n)x^*(n-m)] = 0 \ \Rightarrow \ E[\{x(n) + \sum_{m=1}^{p} a_m x(n-m)\}x^*(n-m)] = 0$$

that is,

$$\gamma_x(m) + a_1\gamma_x(m-1) + \cdots + a_p\gamma_x(m-p) = 0, \ , \quad 1 \le m \le p$$

and (ii) the minimum MSE: $E_p^f = \min_{\{a_k\}} E[|e(n)|^2]$

$$E_p^f = \ E[e(n)e^*(n)] = E[e(n)\{x^*(n) + \sum_{k=m}^{p} a_m^* x^*(n-m)\}]$$

$$= E[e(n)x^*(n)] = E[\{x(n) + \sum_{m=1}^{p} a_m x(n-m)\}x^*(n-m)]$$

$$= \gamma_x(0) + \sum_{m=1}^{p} a_m \gamma_x(-m)$$

1st order ($p = 1$):

$$\gamma_x(1) + a_1\gamma_x(0) = 0 \ \Rightarrow \ a_1 = -\frac{\gamma_x(1)}{\gamma_x(0)}$$

and

$$E_1^f = \gamma_x(0)[1 - |\frac{\gamma_x(1)}{\gamma_x(0)}|^2]$$

2nd order ($p = 2$):

$$\begin{bmatrix} \gamma_x(0) & \gamma_x(-1) \\ \gamma_x(1) & \gamma_x(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = - \begin{bmatrix} \gamma_x(1) \\ \gamma_x(2) \end{bmatrix}$$

and $E_2^f = \gamma_x(0) + a_1\gamma_x(-1) + a_2\gamma_x(-2)$.

In fact, $a_1 = a_1^0, \quad a_2 = a_2^0$. Why ?

- For the 2nd order AR $x(n)$, one has

$$\gamma_x(m) + a_1^0 \gamma_x(m-1) + a_2^0 \gamma_x(m-2) = \sigma_w^2 h(-m)$$

where $\frac{1}{1+a_1^0 z^{-1}+a_2^0 z^{-2}} = \sum_{m=0}^{+\infty} h(m) z^{-m}$. See Eqn. (32) of the notes.

Since $h(0) = 1$ and $h(m) = 0$ for $m < 0$, with $m = 0, 1, 2$

$$\gamma_x(0) + a_1^0 \gamma_x(-1) + a_2^0 \gamma_x(-2) = \sigma_w^2$$

$$\gamma_x(1) + a_1^0 \gamma_x(0) + a_2^0 \gamma_x(-1) = 0$$

$$\gamma_x(2) + a_1^0 \gamma_x(1) + a_2^0 \gamma_x(0) = 0$$

Since for real-valued $x(n)$, $\gamma_x(-m) = \gamma_x(m)$. Therefore,

$$\begin{bmatrix} 1 & a_1^0 & a_2^0 \\ a_1^0 & 1 + a_2^0 & 0 \\ a_2^0 & a_1^0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_x(0) \\ \gamma_x(1) \\ \gamma_x(2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \sigma_w^2$$

With $a_1^0 = -1$, $a_2^0 = 0.6$ and $\sigma_w^2 = 1$, one has $\gamma_x(0) = 2.5641$, $\gamma_x(1) = 1.6026$, $\gamma_x(2) = 0.0641$ and hence for $m \geq 3$

$$\gamma_x(m) = -a_1^0\gamma_x(m-1) - a_2^0\gamma_x(m-2).$$

Calculations show

$$\gamma_x(3) = -0.8947, \quad \gamma_x(4) = -0.9359, \quad \gamma_x(5) = -0.3974$$

1st order ($p = 1$):

$$a_1 = -\frac{\gamma_x(1)}{\gamma_x(0)} = -0.6250, \;\; E_1^f = \gamma_x(0)[1 - |\frac{\gamma_x(1)}{\gamma_x(0)}|^2] = 1.5625 > \sigma_w^2$$

2nd order ($p = 2$):

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -\begin{bmatrix} \gamma_x(0) & \gamma_x(1) \\ \gamma_x(1) & \gamma_x(0) \end{bmatrix}^{-1} \begin{bmatrix} \gamma_x(1) \\ \gamma_x(2) \end{bmatrix} = \begin{bmatrix} -1.0000 \\ 0.6000 \end{bmatrix},$$

which satisfy $a_1 = a_1^0, \;\; a_2 = a_2^0$, and as expected,

$$E_2^f = \gamma_x(0) + a_1\gamma_x(1) + a_2\gamma_x(2) = 1 = \sigma_w^2.$$

P_2: Look at the noise canceller depicted in the following figure, where the measurable signal $d(n)$ has the desired signal $s(n)$ and an additive noise $v(n)$, which is uncorrelated with $s(n)$. The second measurable signal $x(n)$ is assumed uncorrelated with the zero-mean $s(n)$ but correlated with $v(n)$.

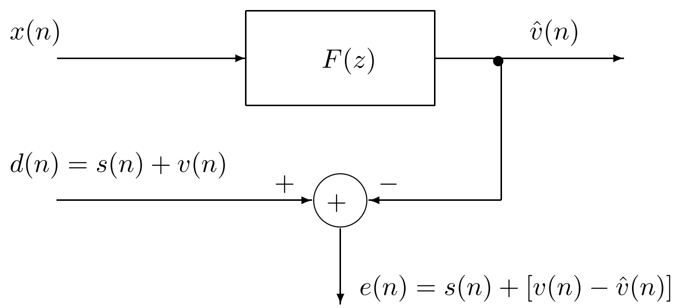

Figure 4: Block diagram of a noise canceller.

The Wiener filter is designed to to minimize $E[e^2(n)]$. Show that the Wiener filter can be obtained by minimizing $E[|\hat{v}(n)-v(n)|^2]$. Find out the optimum FIR Wiener filter of order N.

Solution: The Wiener filter is obtained from

$$\min_{F(z)} E[e^2(n)]$$

where $e(n) = d(n) - \hat{v}(n) = s(n) + [v(n) - \hat{v}(n)]$. Since the zero-mean $s(n)$ is uncorrelated with $v(n)$ and $x(n)$ (hence $\hat{v}(n)$), it is uncorrelated with $\Delta v(n) = v(n) - \hat{v}(n)$. Therefore, $\gamma_e(m) = \gamma_s(m) + \gamma_{\Delta v}(m)$, and particularly with $m = 0$

$$E[e^2(n)] = E[s^2(n)] + E[|v(n) - \hat{v}(n)|^2].$$

Knowing $s(n)$ has nothing to do with $F(z)$, one has

$$\min_{F(z)} E[e^2(n)] \quad \Leftrightarrow \quad \min_{F(z)} E[|v(n) - \hat{v}(n)|^2].$$

Let $F(z) = h_0 + h_1 z^{-1} + \cdots + h_N z^{-N}$, then

$$\hat{v}(n) = h_0 x(n) + h_1 x(n-1) + \cdots + h_N x(n-N)$$

According to the orthogonality principle,

$$E[e(n)x(n-m)] = 0 \quad \Leftrightarrow \quad E[\{d(n) - \hat{v}(n)\}x(n-m)] = 0$$

for $m = 0, 1, 2, \cdots, N$, that is

$$\gamma_{dx}(m) = h_0 \gamma_x(m) + h_1 \gamma_x(m-1) + \cdots + h_N \gamma_x(m-N).$$

Denote

$$\bar{h} = [h_0 \;\; h_1 \;\; \cdots \;\; h_N]^{\mathcal{T}}$$

$$\bar{\gamma}_{dx} = [\gamma_{dx}(0) \;\; \gamma_{dx}(1) \;\; \cdots \;\; \gamma_{dx}(N)]^{\mathcal{T}},$$

then

$$\Phi_x \bar{h}_{opt} = \bar{\gamma}_{dx} \quad \Leftrightarrow \quad \bar{h}_{opt} = \Phi_x^{-1} \bar{\gamma}_{dx}$$

where Φ_x is the $(N+1) \times (N+1)$ auto-correlation matrix of $x(n)$.

P_3: If the noise canceller depicted in Figure 4 above is an adaptive FIR filter. Derive the corresponding LMS algorithm and analyze its convergence behavior.

Solution: The Wiener filter is obtained from

$$\min_{F(z)} E[e^2(n)]$$

The adaptive noise canceller is to solve the above problem on-line. The LMS algorithm is

$$\bar{h}(n+1) = h(n) - \mu \frac{de^2(n)}{d\bar{h}}|_{\bar{h}=\bar{h}(n)}.$$

Noting $e(n) = d(n) - h_0 x(n) - h_1 x(n-1) + \cdots - h_N x(n-N) = d(n) - \bar{x}^{\mathcal{T}}(n)\bar{h}$, where

$$\bar{x}(n) = [x(n) \quad x(n-1) \quad \cdots \quad x(n-N)]^{\mathcal{T}},$$

one has $\frac{de^2(n)}{d\bar{h}_k} = -2e(n)x(n-k)$

$$\frac{de^2(n)}{d\bar{h}} = -2e(n)\bar{x}(n)$$

and hence the adaptive LMS is

$$\bar{h}(n+1) = \bar{h}(n) + 2\mu[d(n) - \bar{x}^{\mathcal{T}}(n)\bar{h}(n)]\bar{x}(n).$$

Convergence analysis

Denote $\Delta\bar{h}(n) \stackrel{\triangle}{=} \bar{h}(n) - \bar{h}_{opt}$. The above LMS yields

$$\Delta\bar{h}(n+1) = \Delta\bar{h}(n) + 2\mu[d(n)\bar{x}(n) - \bar{x}(n)\bar{x}^{\mathcal{T}}(n)\bar{h}(n)]$$

and hence

$$E[\Delta\bar{h}(n+1)] = E[\Delta\bar{h}(n)] + 2\mu\{E[d(n)\bar{x}(n)] - E[\bar{x}(n)\bar{x}^{\mathcal{T}}(n)\bar{h}(n)]\}.$$

Assume that $\bar{h}(n)$ is uncorrelated with $\bar{x}(n)$, then

$$E[\bar{x}(n)\bar{x}^{\mathcal{T}}(n)\bar{h}(n)] = E[\bar{x}(n)\bar{x}^{\mathcal{T}}(n)]E[\bar{h}(n)]$$

and therefore,

$$E[\Delta\bar{h}(n+1)] = E[\Delta\bar{h}(n)] + 2\mu\{\bar{\gamma}_{dx} - \Phi_x E[\bar{h}(n)]\}.$$

Noting $\Phi_x \bar{h}_{opt} = \bar{\gamma}_{dx}$, one then has

$$E[\Delta \bar{h}(n+1)] = E[\Delta \bar{h}(n)] - 2\mu E[\Delta \bar{h}(n)]\} = (I - 2\mu\Phi_x)E[\Delta \bar{h}(n)].$$

As shown in Problem 2 of Exercise 2, $E[\Delta \bar{h}(n)]$ goes to zeros when $n \to +\infty$, that $\bar{h}(n)$ converges to $\bar{h}_{opt}$ statistically, if all the eigenvalues of $I - 2\mu\Phi_x$ are absolutely smaller than one.