Principle of Orthogonality (Method 2) Continued

Wiener-Hopf Equations

From orthogonality principles, we have

$$E[u(n-k)e_o^*(n)] = 0$$

where e_o is the error which results when using the optimal filter. Substituting for the error signal, we have

$$E\left[u(n-k)\left(d^*(n) - \sum_{l=0}^{\infty} w_{ol} u^*(n-i)\right)\right] = 0,\ k = 0,\ 1,\ 2,\ldots$$

where w_{ol} is the lth coefficient of the optimal filter. Expanding and rearranging, we have

$$\sum_{l=0}^{\infty} w_{ol} E[u(n-k)u^*(n-l)] = E[u(n-k)d^*(n)],\ k = 0,\ 1,\ 2,\ldots$$

or the system of equations a.k.a. Wiener-Hopf equations

$$\sum_{\iota=0}^{\infty} w_{oi} r(i-\kappa) = \pi(-\kappa)$$

where

- $r(l-k) = E[u(n-k)u^*(n-l)]$ is the autocorrelation function of the filter input for lag $l-k$
- $p(-k) = E[u(n-k)d^*(n)]$ is the cross-correlation between the filter input $u(n-k)$ and the desired response $d(n)$ for lag $-k$.

For finite dimensions, the Wiener-Hopf equations can be written in matrix form as

$$\mathbf{Rw} = \boldsymbol{\pi}$$

with the optimal solution given by the Wiener filter

$$\boxed{\mathbf{w}_o = \mathbf{P}^{-1}\boldsymbol{\pi}}.$$

Canonical Form of the Error-Performance Surface

From our first method of solution we have

$$J(\mathbf{w}) = \vartheta_{\mu\iota\nu} + (\boldsymbol{\omega} - \boldsymbol{\omega}_o)^H \mathbf{P} (\boldsymbol{\omega} - \boldsymbol{\omega}_o)$$

where it is clear that if $\mathbf{w} = \mathbf{w}_o$ we have minimized the MSE. The vector $\mathbf{w} - \mathbf{w}_o$ is often called the misalignment vector and $J(\mathbf{w} - \mathbf{w}_o)$ represents a translation of the paraboloid to the origin. We rewrite using the diagonalized form of $\mathbf{R}$

$$J(\mathbf{w}) = \vartheta_{\mu\iota\nu} + (\boldsymbol{\omega} - \boldsymbol{\omega}_o)^H \boldsymbol{\Theta\Lambda\Theta}^H (\boldsymbol{\omega} - \boldsymbol{\omega}_o)$$

where $\mathbf{Q}$ is the matrix whose columns are the eigenvectors of $\mathbf{R}$ and Λ is the diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{R}$. We define

$$\nu = \mathbf{Q}^H(\mathbf{w} - \mathbf{w}_o)$$

as the rotated (by **Q**) and translated (by $\mathbf{w}_o$) parameter vector and rewrite the MSE in canonical form

$$
\begin{aligned}
J &= \vartheta_{\mu\iota\nu} + \nu^H \Lambda \nu \\
&= \vartheta_{\mu\iota\nu} + \sum_{\kappa=1}^{M} \lambda_\kappa \nu_\kappa \nu_\kappa^* \\
&= \vartheta_{\mu\iota\nu} + \sum_{\kappa=1}^{M} \lambda_\kappa \left|\nu_\kappa\right|^2
\end{aligned}
.
$$

The canonical form represents the error surface in an uncoupled form in the sense that each component of the gradient vector of the MSE w.r.t. to $\boldsymbol{\nu}$ is a function of a single parameter, i.e.

$$\nabla_\nu J = [2\lambda_1 \nu_1 \quad 2\lambda_2 \nu_2 \quad \mathrm{L} \quad 2\lambda_M \nu_M]^T .$$

Thus the vector $\boldsymbol{\nu}$ constitutes a principal axis of the error-performance surface. It can also be shown that $\boldsymbol{\nu}$ is an eigenvector of **R**. (The rotated misalignment vector is a result of being projected on the eigenvectors' direction). Therefore we conclude that the eigenvectors of **R** define the principal axes of the error-performance surface.

Figure 3.2 (Widrow and Sterns)

Wiener Filter Example

In this example we wish to equalize or undo the effects of a channel (ie. telephone, cable, airwaves, etc.) on a signal of interest. Therefore, let us build a length 2 Wiener filter whose input is $u(n)$ (output of the channel) and whose output is an estimate of $d(n)$ (signal that went into the channel) that is optimum in the mean-square error sense.

Figure Overall system problem

The desired response $d(n)$ is modeled as an order 1 AR process

Figure 5.5a

The communications channel is modeled as an IIR filter whose output is corrupted by noise

Figure 5.5b

We assume $v_1(n)$ and $v_2(n)$ are uncorrelated. To build the length 2 Wiener filter proceed with the following steps:

1) Get the 2 × 2 correlation matrix, **R** of $u(n)$, i.e. $\mathbf{R} = E[\mathbf{u}(n)\mathbf{u}^T(n)]$
2) Get the 2 × 1 cross-correlation vector, **p** between $\mathbf{u}(n)$ and $d(n)$, i.e. $\mathbf{p} = E[\mathbf{u}(n)d(n)]$
3) Compute $\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}$

Step 1: We note that $x(n)$ can be thought of as being generated by the cascade of $H_1(z)$ and $H_2(z)$ with input $v_1(n)$. We have

$$\begin{aligned} H(z) &= H_1(z)H_2(z) \\ &= \frac{1}{(1+0.8458z^{-1})(1-0.9458z^{-1})} \\ &= \frac{1}{1-0.1z^{-1}-0.8z^{-2}} \end{aligned}$$

or equivalently

$$\begin{aligned} x(n) + a_1 x(n-1) + a_2 x(n-2) &= v(n) \\ a_1 &= -0.1 \\ a_2 &= -0.8 \end{aligned}.$$

We note that $u(n) = x(n) + v_2(n)$ and thus [assuming $\mathbf{u}(n)$ real]

$$\begin{aligned} \mathbf{R} &= E[\mathbf{u}(n)\mathbf{u}^T(n)] \\ &= E[(\mathbf{x}(n)+\mathbf{v}_2(n))(\mathbf{x}(n)+\mathbf{v}_2(n))^T] \\ &= E[\mathbf{x}(n)\mathbf{x}^T(n)] + 2E[\mathbf{x}(n)\mathbf{v}_2^T(n)] + E[\mathbf{v}_2(n)\mathbf{v}_2^T(n)] \\ &= \mathbf{R}_x + \mathbf{0} + \mathbf{R}_2 \end{aligned}$$

since $x(n)$ and $v_2(n)$ are uncorrelated. From our earlier work in AR models we can show

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

Since $v_2(n)$ is a white noise process of zero mean with $\sigma_2^2 = 0.1$ we have

$$\mathbf{R}_2 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}.$$

Therefore

$$\mathbf{R} = \begin{bmatrix} 1.1 & 0.5 \\ 0.5 & 1.1 \end{bmatrix}.$$

Step 2: We have

$$p(-k) = E[u(n-k)d(n)]$$

and since our processes are WSS and real, $p(k) = p(-k)$. We also have

$$x(n) = 0.9458x(n-1) + d(n).$$

Now,

$$\begin{aligned} p(k) &= E[\upsilon(\nu-\kappa)\delta(\nu)] \\ &= E[(\xi(\nu-\kappa)+\varpi_2(\nu-\kappa))(\xi(\nu)-0.9458\xi(\nu-1))] \\ &= E[\xi(\nu-\kappa)\xi(\nu)] + E[\varpi_2(\nu-\kappa)\xi(\nu)] + E[\xi(\nu-\kappa)(-0.9458\xi(\nu-1))] + . \\ & E[\varpi_2(\nu-\kappa)(-0.9458\xi(\nu-1))] \\ &= \rho_\xi(\kappa) + 0 - 0.9458\rho_\xi(\kappa-1) + 0 \end{aligned}$$

Solving for $p(0)$ and $p(1)$ we have

$$\begin{aligned} p(0) &= 1 - 09458(0.5) \\ &= 0.5272 \\ \pi(1) &= 0.5 - 09458(1) \\ &= -0.4458 \end{aligned} .$$

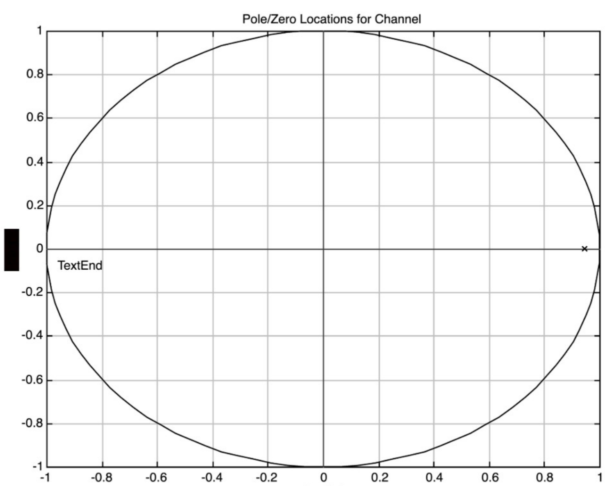
Therefore

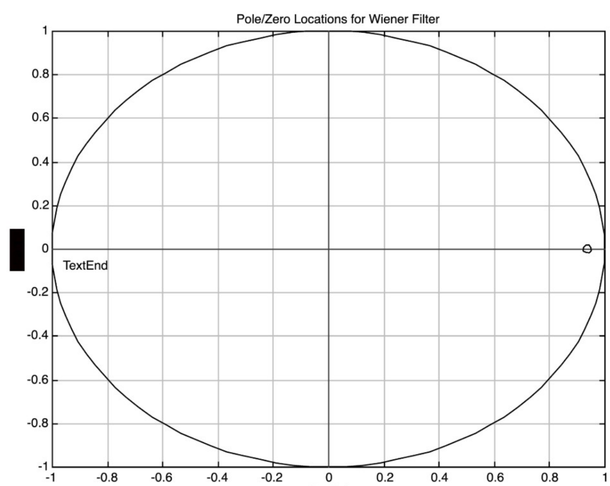
$$\mathbf{p} = \begin{bmatrix} 0.5272 \\ -0.4458 \end{bmatrix}.$$

Step 3: The Wiener filter is then given by

$$\begin{aligned} \mathbf{w} &= \mathbf{P}^{-1}\boldsymbol{\pi} \\ &= \begin{bmatrix} 0.8360 \\ -0.7853 \end{bmatrix} \end{aligned} .$$

The pole/zero patterns of the equalizer and Wiener filter are given by





Clearly the Wiener filter zero cancels the channel pole.

How good is this solution? Using the optimum (in the MSE-sense) filter (Wiener filter) the minimum MSE (MMSE) is

$$\begin{aligned} J_{\min} &= \vartheta(\boldsymbol{\omega}_o) \\ &= \sigma_\delta^2 - \boldsymbol{\pi}^H \mathbf{P}^{-1}\boldsymbol{\pi} \\ &= \sigma_\delta^2 - \boldsymbol{\pi}^H \boldsymbol{\omega}_o \\ &= 0.1579 \end{aligned} .$$

We note that our optimal filter does *not* yield zero MSE due to the channel noise we can't control. Next we note that the eigenvectors of **R** are given by

```
»[V,D] = eig(R)
V =
    0.7071    0.7071
   -0.7071    0.7071
D =
    0.6000         0
         0    1.6000
```

We compare the eigenvectors with Figure 5.7 and note that they coincide with the principal axes of the ellipses of constant MSE.