(1) Properties of Wiener Filter.

Last time:

\[ y(n) = u(n) + e(n) \]

\[ d(n) \] - desired response

\[ e(n) = d(n) - y(n) \]

\[ y(n) = w^H u(n) = \sum_{k=0}^{\infty} w_k^* u(n-k) \] - output

Optimization criterion:

\[ J(w) = E[e(n)^* e(n)] = E[F_x e(n)^* e(n)] \]

Result: The filter is minimized in this sense if

\[ w = R^{-1} P \] where

\[ R = E\{ u(n) u(n)^* \} \] - correlation matrix

\[ P = E\{ d(n) d(n)^* \} = E\{ [u(n) u(n-1) \ldots u(n+k)] d(n+k)^* \} \] - cross correlation vector

The minimum of the error is

\[ J(w = w_{opt}) = 6\sigma^2 - P^H R^{-1} P \]

\[ 6\sigma^2 \] - variance of the desired system response

\[ P^H R^{-1} P \] - portion of variance explained by filter output

Principle of orthogonality:

Suppose that the filter is operating with optimum weights.
Consider.

\[ e(n) = d(n) - y(n) = d(n) - w^H u(n) \]

Taking complex conjugate of both sides,

\[ \bar{e}(n) = \bar{d}(n) - \bar{w}^H \bar{u}(n) \]

Proceeding with \( \bar{u}(n) \) and take expected value.

\[ E \{ \bar{e}(n) \} = E \{ \bar{d}(n) \} - E \{ \bar{w}^H \bar{u}(n) \} = E \{ \bar{u}(n) \} E \{ \bar{w} \} \]

Cross correlation of input and even reactor.

\[ E \{ \bar{u}(n) \cdot \bar{d}(n) \} = P \] (cross correlation of input reflex and deni)

\[ E \{ \bar{u}(n) \cdot \bar{w}^H \bar{y} \} = P \] (correlation between input and input process)

\[ E \{ \bar{u}(n) \cdot \bar{e}(n) \} = P - P \cdot W \]

When the filter is operating in its optimum state \( \omega = \omega_{opt} = P^{-1} P \)

\[ E \{ \bar{u}(n) \cdot \bar{e}(n) \} = P - P \cdot P^{-1} P = P - P = 0 \]

Therefore.

\[ E \{ \bar{e}(n) \} = 0 \Rightarrow \text{error is zero at time } n \text{ is orthogonal to samples of signal taken in time instants before } n. \]

\[ E \{ u(n-k) \cdot \bar{e}(n) \} = 0 \] - orthogonal sequences for \( k = 0, 1, 2, \ldots, M-1 \)
How should we interpret the above result

\[ y(n) \rightarrow \stackrel{\text{filter}}{\longrightarrow} \hat{d}(n) \rightarrow e(n) \]

\[ H(n-1) \]

\[ H(n-H+1) \]

\[ \hat{d}(n) \text{ - predicted on the basis of "knowledge" about } d(n) \text{ that can be extracted from } y(n) \]

When the filter is operating in open-loop position all knowledge is used. Therefore \( \hat{d}(n) \) is all that could have been predicted. \( e(n) \text{ is part of } d(n) \text{ that can not be predicted from } y(n) \). Therefore \( e(n) \) and \( \hat{d}(n) \) are orthogonal. The orthogonal on a component by component basis.

**Corollary of the Principle of Orthogonality**

Consider

\[ E_y y(n) e(n) = \sum \omega^* y(n) e(n) \]

\[ = \sum \omega^* E_y y(n) e(n) \]

When the filter is operating in open-loop point \( \omega = P^{-1} P \) and

\[ E_y y(n) e(n) = 0 \]

Therefore

\[ E_y y(n) e(n) = 0 \]

How should we interpret the above result.

\( y(n) \) is produced as a linear combination of \( u(n), u(n-1), \ldots, u(n-H+1) \). Therefore \( y(n) \) cannot have any additional information about \( d(n) \) that is not in \( y(n) \). Since
\( c_{ni} \) is orthogonal to every element of \( \{a_{ji}\} \) and to the linear combination of their \( \omega_{jk} \)s as well.

**Wiener-Hopf equations**

**Denoted by principle of orthogonality**

\[
E_f \{ -u(n-k) e^{j\omega_{0} m} \} = 0 \quad k = 0, 1, 2, ..., N - 1
\]

\[
E_f \{ u(n-k) \left[ d(n) - \omega_{0} y(n) \right] \} = 0
\]

\[
= E_f \{ u(n-k) \left[ d(n) - \sum_{p=0}^{N-1} u(n-p) \omega_{0} \right] \}
\]

\[
= E_f \{ u(n-k) d(n) \} - \sum_{p=0}^{N-1} E_f \{ u(n-k) u(n-p) \} \omega_{0} = 0
\]

Consider individual terms:

\[
E_f \{ u(n-k) d(n) \} = p(-k) \quad \text{Cross correlation between input signal and desired signal for lag } -k
\]

\[
E_f \{ u(n-k) u(n-y) \} = r(y-k) \quad k = 0, 1, 2, ..., N - 1
\]

Therefore

\[
p(-k) = \sum_{p=0}^{N-1} \omega_{0} r(y-k) = 0, 1, 2, ..., N - 1
\]

\( x \) are referred to as Wiener-Hopf equations or more often normal equations.
Let us rearrange Wiener-Hopf equations in matrix form.

\[ P(0) = \sum_{\nu=0}^{N-1} w_{0\nu} \cdot r(\nu) = \begin{bmatrix} r(0) & r(1) & \cdots & r(N-1) \end{bmatrix} w_0 \]

\[ P(-1) = \sum_{\nu=0}^{N-1} w_{0\nu} \cdot r(\nu-1) = \begin{bmatrix} r(-1) & r(0) & \cdots & r(N-2) \end{bmatrix} w_0 \]

\[ P(-H+1) = \sum_{\nu=0}^{N-1} w_{0\nu} \cdot r(\nu-H+1) = \begin{bmatrix} r(-H+1) & r(-H+2) & \cdots & r(0) \end{bmatrix} w_0 \]

\[
\begin{bmatrix}
P(0) \\
P(-1) \\
P(-H+1)
\end{bmatrix}
= 
\begin{bmatrix}
r(0) & r(1) & \cdots & r(N-1) \\
r(-1) & r(0) & \cdots & r(N-2) \\
r(-H+1) & r(-H+2) & \cdots & r(0)
\end{bmatrix}
\begin{bmatrix}
w_0 \\
w_1 \\
w_{-H+1}
\end{bmatrix}
\]

Therefore for optimum weights

\[ P = R \cdot w_0 \]

Same as condition that we have derived assuming that

\[ w_0 \] minimizes error performance surface

Of course from Wiener-Hopf equations we obtain

\[ w_0 = R^{-1} \cdot P \]

Canonical form of the error performance surface

We have defined error performance surface as:

\[ f(w) = E[f(E[n] \cdot e(n)^2)] = E[f(e(n))^2] \]
Recall 
\[ c(n) = d(n) - \omega^H u(n) \]

Therefore

\[ f(\omega) = E f(\omega) - \omega^H y(n) \] (d(n) - \omega^H u(n)) \times y =

\[ = E f(\omega) - \omega^H y(n) \] (d - \omega^H u(n)) \times y =

\[ = B d^2 - \omega^H P - P^H \omega + \omega^H P \omega \] We have used this before to derive \( \omega_{opt} \)

Recall \( \omega_0 = P^T P \)

\[ f(\omega) = B d^2 - \omega^H P^T P \cdot \omega - P^H P \omega + \omega^H P^T P \cdot \omega \]

\[ = B d^2 - P^H P \cdot \omega + (\omega - P^T P) \cdot P \cdot (\omega - P^T P) \]

\[ = B d^2 - P^H P \cdot \omega + (\omega - \omega_0) \cdot P \cdot (\omega - \omega_0) =

\[ f(\omega) = \omega_0 + (\omega - \omega_0) \cdot P \cdot (\omega - \omega_0) \] \( \Rightarrow \) Tell us lot about the simchae of

\[ \nabla f(\omega) \bigg|_{\omega = \omega_0} = \omega_0 = B d^2 - P^H P \cdot \omega_0 \]

\[ (\omega - \omega_0)^T P (\omega - \omega_0) \geq 0 \] due to positive definite property of \( P \)

\[ \nabla f(\omega) \bigg|_{\omega = \omega_0} = \omega_0 = B d^2 - P^H P \cdot \omega_0 \]

\[ (\omega - \omega_0)^T P (\omega - \omega_0) \geq 0 \] due to positive definite property of \( P \)

\[ \nabla f(\omega) \bigg|_{\omega = \omega_0} = \omega_0 = B d^2 - P^H P \cdot \omega_0 \]

Since penalty surface is quadratic \( \Rightarrow \) there is a unique solution

\[ \therefore \] there is one and only one optimum solution

Minimum of the penalty surface is **global**