How to find Optimal Filter?

Our goal is to \( h[0], \ldots, h[K - 1] \) to minimize the mean squared error (MSE)

\[
\underset{h[i]}{\text{minimize}} \quad \mathbb{E} \left[ \left( X[n] - \hat{X}[n] \right)^2 \right]. \tag{5}
\]

What is MSE?

\[
\mathbb{E} \left[ (X[n] - \hat{X}[n])^2 \right] = \mathbb{E} \left[ \left( X[n] - \sum_{i=0}^{K-1} h[i] Y[n-i] \right)^2 \right].
\]

How to minimize MSE?

The first order derivative with respect to \( h[i] \) is

\[
\frac{d}{dh[i]} \mathbb{E} \left[ (X[n] - \hat{X}[n])^2 \right] = \frac{d}{dh[i]} \mathbb{E} \left[ \left( X[n] - \sum_{j=0}^{K-1} h[j] Y[n-j] \right)^2 \right].
\]
How to find Optimal Filter?

Setting this to zero:

\[
0 = \frac{d}{dh[i]} \mathbb{E} \left[ \left( X[n] - \sum_{j=0}^{K-1} h[j] Y[n-j] \right)^2 \right]
\]

\[
= \mathbb{E} \left[ 2 \left( X[n] - \sum_{j=0}^{K-1} h[j] Y[n-j] \right) Y[n-i] \right]
\]

we have:

\[
\mathbb{E} [X[n]Y[n-i]] = \mathbb{E} \left[ \sum_{j=0}^{K-1} h[j] Y[n-j] Y[n-i] \right]
\]

This is

\[
R_{X,Y}[i] = \sum_{j=0}^{K-1} h[j] R_Y[j+i], \quad i = 0, 1, \ldots, K - 1.
\]
\[
\tilde{x}[n] = \sum_{i=0}^{k-1} h[i] y[n-i]
\]

\[
\min_{h[i]} \left\{ E \left( \tilde{x}[n] - x[n] \right)^2 \right\}
\]

\[
\min_{h[i]} \left\{ E \left( \sum_{i=0}^{k-1} h[i] y[n-i] - x[n] \right)^2 \right\}
\]
\[
\frac{d}{dh[i]} \left\{ \mathbb{E}\left( \ldots \right) \right\} = \frac{d}{dh[i]} \mathbb{E}\left\{ \left( \sum_{i=0}^{K-1} h[i] Y[n-i] - X[n] \right)^2 \right\}
\]

\[
= \mathbb{E}\left\{ \frac{d}{dh[i]} \left( \sum_{i=0}^{K-1} h[i] Y[n-i] - X[n] \right)^2 \right\}
\]

\[
= 2 \left( \sum_{i=0}^{K-1} h[i] Y[n-i] - X[n] \right) \frac{d}{dh[i]} \left[ \sum_{i=0}^{K-1} h[i] Y[n-i] - X[n] \right]
\]

\[
\frac{d}{dh[i]} \left\{ h[0] Y[n-0] + h[1] Y[n-1] + \ldots + h[K-1] Y[n-(K-1)] \right\} - X[n]
\]
\[ O = 2 \sum_{j=0}^{K-1} \left( E\left( Y[n-j] Y[n-i] \right) h[j] \right) - 2X[n] Y[n-i] \]

\[ E\left[ X[n] Y[n-i] \right] = \sum_{j=0}^{K-1} E\left[ Y[n-j] Y[n-i] \right] h[j] \]

\[ R_{XY}[i] = \sum_{j=0}^{K} R_{Y[i-j]} h[j] \]

\[ (n) - (n-i) \]

\[ X(n) \quad Y(n-i) \]
How to find Optimal Filter?

Write out the equations for \( i = 0, \ldots, K - 1 \), we have this equation:

\[
\begin{pmatrix}
R_{X,Y}[0] \\
R_{X,Y}[1] \\
\vdots \\
R_{X,Y}[K-1]
\end{pmatrix}
= 
\begin{pmatrix}
\vdots & \vdots & \ddots & \vdots \\
R_Y[K-1] & \cdots & R_Y[1] & R_Y[0]
\end{pmatrix}
\begin{pmatrix}
h[0] \\
h[1] \\
\vdots \\
h[K-1]
\end{pmatrix}
\]

This equation is called the Yule-Walker Equation.

So what?

- Yule-Walker equation is a \( K \times K \) system, which is typically very easy to solve.
- We do not need to know the channel from \( X[n] \) to \( Y[n] \).
- We only need to know \( R_{X,Y}[k] \) and \( R_Y[k] \).
Linear Predictive Code

In LPC, we want to design $h[k]$ such that it can encode the input signal $Y[n]$.

\[ Y[n] \rightarrow \text{LPC} \rightarrow \hat{X}[n] \]

Let us give a structure to the estimated signal $\hat{X}[n]$:

\[ \hat{X}[n] = \sum_{k=0}^{K-1} h[k] Y[n - k]. \]

We want $\hat{X}[n]$ to stay as close to $Y[n]$ as possible. So we need to solve

\[ \min_{h[i]} \mathbb{E} \left[ \left( Y[n] - \hat{X}[n] \right)^2 \right]. \]

That is, we substitute $X[n]$ by $Y[n]$ in Yule-Walker.
\[
\min E[(Y[n] - \sum_{j=0}^{K-1} h[j] Y[n-j])^2]
\]

\[
\hat{x}[n] = \sum_{j=0}^{K-1} h[j] Y[n-j]
\]

\[
\sim Y[n]
\]

\[
\hat{x}[n] \sim Y[n]
\]

\[
Y[n-2], Y[n-1]
\]
Linear Predictive Code

In this case, we have

\[ R_Y[i] = \sum_{j=0}^{K-1} h[j] R_Y[j \oplus i], \quad i = 0, 1, \ldots, K - 1. \]

which is equivalent to

\[
\begin{pmatrix}
R_Y[0] \\
R_Y[1] \\
\vdots \\
R_Y[K - 1]
\end{pmatrix}
= 
\begin{pmatrix}
\vdots & \vdots & \ddots & \vdots \\
R_Y[K - 1] & \cdots & R_Y[1] & R_Y[0]
\end{pmatrix}
\begin{pmatrix}
h[0] \\
h[1] \\
\vdots \\
h[K - 1]
\end{pmatrix}
\]

This Yule-Walker equation is a Toeplitz system, which is easy to solve. In industry, people often use the Levinson-Durbin algorithm.
How to Find $R_Y[k]$ and $R_{X,Y}[k]$?

By Mean-Ergodic Theorem, we can approximately find

$$R_Y[k] \approx \frac{1}{N} \sum_{n=0}^{N-1} Y[n + k] Y[n]$$

$$R_{X,Y}[k] \approx \frac{1}{N} \sum_{n=0}^{N-1} X[n + k] Y[n].$$

These two results hold because we assume $X[n]$ and $Y[n]$ are W.S.S. assumption. Therefore, the mean Ergodic theorem applies and makes the statistical average equivalent to the temporal average.
Orthogonality Condition

Remark: The condition

$$\mathbb{E} \left[ \left(X[n] - \hat{X}[n]\right) Y[n - i] \right] = 0$$

is called the orthogonality condition. It is called "orthogonal" because the residue $X[n] - \hat{X}[n]$ is orthogonal to the measurement $Y[n - i]$. 
\[ R_{X,Y}(i) = R_{Y,Y}(i) \]

\[
\begin{bmatrix}
1 \\
\vdots \\
1
\end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix}
\begin{bmatrix} h[0] \\
\vdots \\
h[k-1]
\end{bmatrix} \\
1
\end{bmatrix} = A^{-1} \begin{bmatrix}
\begin{bmatrix} h[0] \\
\vdots \\
h[k-1]
\end{bmatrix}
\end{bmatrix}.
\]
Quantization

\[
\begin{bmatrix}
e_{[0]} \\
e_{[1]} \\
\vdots \\
e_{[N-1]}
\end{bmatrix}
\]

\[
\begin{bmatrix}
e_q[0] \\
e_q[1] \\
\vdots \\
e_q[N-1]
\end{bmatrix}
\]

\[
\begin{bmatrix}
h[0] \\
h[1] \\
\vdots \\
h[k-1]
\end{bmatrix}
\]

\[
\begin{bmatrix}
e_{[0]} \\
e_{[1]} \\
\vdots \\
e_{[N-1]}
\end{bmatrix}
\]

\[
\begin{bmatrix}
e_q[0] \\
e_q[1] \\
\vdots \\
e_q[N-1]
\end{bmatrix}
\]

\[
\begin{bmatrix}
h[0] \\
h[1] \\
\vdots \\
h[k-1]
\end{bmatrix}
\]

\[
\begin{bmatrix}
x[n] \\
\end{bmatrix}
\]

\[
Y[n]
\]

\[
x[n]
\]

\[
e[n]
\]

Filtering command.

\[
sig.\ 1\ filter(B, A, x)
\]
$$E[n] = \hat{Y}[n] \left( 1 - \sum_{k=1}^{K} a(k) \right)$$

$$\frac{E(2)}{Y(2)} = \frac{(1 - \sum_{k=1}^{K} a(k) z^{-k})}{1}$$

$$\frac{Y(2)}{E(2)}$$

$$B = \begin{bmatrix} 1 & -a(1) & -a(2) \ldots \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & -a(1) & -a(2) \ldots \end{bmatrix}$$

$$B = \begin{bmatrix} \vdots & 1 & \ldots \end{bmatrix}$$