Statistical Signal Processing

Lecture 7: Optimal Filtering, Infinite Impulse Response (IIR) and Applications

Carlas Smith & Peyman Mohajerin Esfahani

Delft Center for Systems and Control

TU Delft

Delft University of Technology

Highlights from Lecture 1

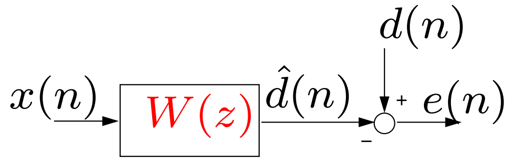
Parseval: For $x(n), y(n)$ a series of complex numbers:

$$\sum_{n=-\infty}^{\infty} x(n)y^*(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega})d\omega$$

Optimal IIR filtering: Part I

1. Recap FIR Optimal Wiener filter - The Orthogonality Condtion
2. A generic framework
3. The IIR Wiener filter
4. The filtering Problem
5. The deconvolution problem

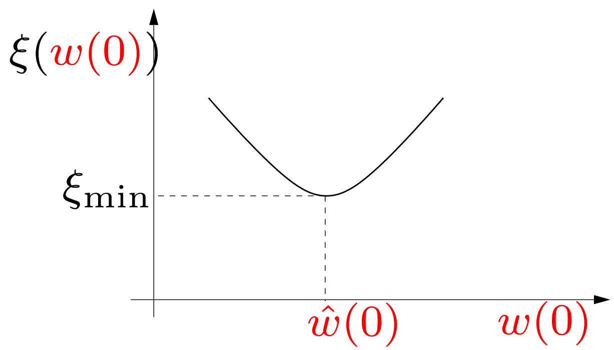
The orthogonality principle



$$e(n) = d(n) - w(0)x(n) - w(1)x(n-1)$$

Then the necessary (and sufficient) condition to minimize $\xi(\mathbf{w}) = E[|e(n)|^2] = E[e(n)e^*(n)]$:

$$\left[\begin{array}{c} \frac{\partial \xi(\mathbf{w})}{\partial w^*(0)} \\ \frac{\partial \xi(\mathbf{w})}{\partial w^*(1)} \end{array}\right]_{\mathbf{w}=\hat{\mathbf{w}}} = 0$$



Using the expression for $\xi(\mathbf{w})$ this equals:

$$\boxed{E\left[\begin{array}{c} e_{\min}(n)\frac{\partial e^*(n)}{\partial w^*(0)} \\ e_{\min}(n)\frac{\partial e^*(n)}{\partial w^*(1)} \end{array}\right]_{\mathbf{w}=\hat{\mathbf{w}}} = E\left[\begin{array}{c} e_{\min}(n)x^*(n) \\ e_{\min}(n)x^*(n-1) \end{array}\right] = 0} \qquad \text{(O.C.)}$$

Interpretation of the orthogonality principle?

Consider the error equation:

$$e(n) = d(n) - w(0)x(n) - w(1)x(n-1)$$

Here the entries of $\mathbf{x}(\mathbf{n}) = \begin{bmatrix} x(n) \\ x(n-1) \end{bmatrix}$ are called the *“regressors”* of the “minimum-variance” estimation problem that aims to minimize:

$$E[|e(n)|^2]$$

TU Delft

Interpretation of the orthogonality principle?

Consider the error equation:

$$e(n) = d(n) - w(0)x(n) - w(1)x(n-1)$$

Here the entries of $\mathbf{x}(\mathbf{n}) = \begin{bmatrix} x(n) \\ x(n-1) \end{bmatrix}$ are called the *"regressors"* of the "minimum-variance" estimation problem that aims to minimize:

$$E[|e(n)|^2]$$

Then the (O.C.) reads: $E[e(n)x^*(n)] = 0 \quad E[e(n)x^*(n-1)] = 0$

This means that the RPs $e(n)$ and each regressor $x(n)$ and $x(n-1)$ are

Interpretation of the orthogonality principle?

Consider the error equation:

$$e(n) = d(n) - w(0)x(n) - w(1)x(n-1)$$

Here the entries of $\mathbf{x(n)} = \begin{bmatrix} x(n) \\ x(n-1) \end{bmatrix}$ are called the

"regressors" of the "minimum-variance" estimation problem that aims to minimize:

$$E[|e(n)|^2]$$

Then the (O.C.) reads: $E[e(n)x^*(n)] = 0 \quad E[e(n)x^*(n-1)] = 0$

This means that the RPs $e(n)$ and each regressor $x(n)$ and $x(n-1)$ are orthogonal.

Solving the Multi-step Prediction using (O.C.)

The cost function we seek to optimize is:

$$\xi(\mathbf{w}) = E\Big[\Big|d(n+\alpha) - \mathbf{w}^T\mathbf{x}(\mathbf{n})\Big|^2\Big]$$

with $\mathbf{x}(\mathbf{n}) = [\ d(n)\ \ d(n-1)\ \ \cdots\ \ d(n-m+1)\ ]^T$ and $\mathbf{w}^T = [w(0)\ \ w(1)\ \ \cdots\ \ w(m-1)]$. Then the error signal $e(n)$ for $m=2$ equals:

Solving the Multi-step Prediction using (O.C.)

The cost function we seek to optimize is:

$$\xi(\mathbf{w}) = E[\left|d(n+\alpha) - \mathbf{w}^T \mathbf{x}(\mathbf{n})\right|^2]$$

with $\mathbf{x}(\mathbf{n}) = [\ d(n) \ \ d(n-1) \ \ \cdots \ \ d(n-m+1)\ ]^T$ and $\mathbf{w}^T = [w(0) \ \ w(1) \ \ \cdots \ \ w(m-1)]$. Then the error signal $e(n)$ for $m = 2$ equals:

$$e(n) = d(n+\alpha) - \begin{bmatrix} d(n) & d(n-1) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix}$$

and the (O.C.) reads for $m = 2$:

Solving the Multi-step Prediction using (O.C.)

The cost function we seek to optimize is:

$$\xi(\mathbf{w}) = E\Big[\Big|d(n+\alpha) - \mathbf{w}^T\mathbf{x}(\mathbf{n})\Big|^2\Big]$$

with $\mathbf{x}(\mathbf{n}) = [\, d(n) \;\; d(n-1) \;\; \cdots \;\; d(n-m+1) \,]^T$ and $\mathbf{w}^T = [w(0) \;\; w(1) \;\; \cdots \;\; w(m-1)]$. Then the error signal $e(n)$ for $m = 2$ equals:

$$e(n) = d(n+\alpha) - \begin{bmatrix} d(n) & d(n-1) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix}$$

and the (O.C.) reads for $m = 2$:

$$E[e(n) \begin{bmatrix} d^*(n) \\ d^*(n-1) \end{bmatrix}] = 0 \Leftrightarrow E[\begin{bmatrix} d^*(n) \\ d^*(n-1) \end{bmatrix} \Big(d(n+\alpha) - \begin{bmatrix} d(n) & d(n-1) \end{bmatrix} \begin{bmatrix} \hat{w}(0) \\ \hat{w}(1) \end{bmatrix}\Big)] = 0$$

These are the Wiener-Hopf equations for this problem.

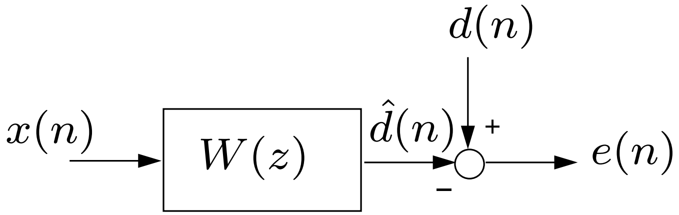
TU Delft

Optimal IIR filtering: Part I

1. Recap FIR Optimal Wiener filter - The Orthogonality Condtion
2. **A generic framework**
3. The IIR Wiener filter
4. The filtering Problem
5. The deconvolution problem

A generic problem formulation

The generic problem deals with the (optimal [a]) estimation of one signal (denoted by $d(n)$) from another signal (denoted by $x(n)$).

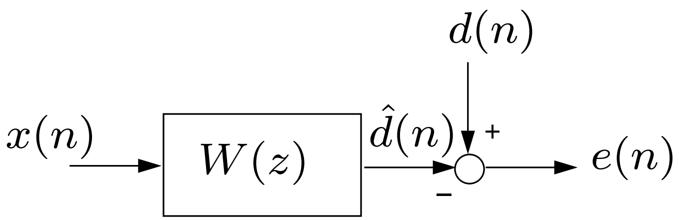


- *Filtering:* estimate $d(n)$ from $x(n) = d(n) + v(n)$.

[a] as specified by the error criterium on $e(n)$

A generic problem formulation

The generic problem deals with the (optimal) estimation of one signal (denoted by $d(n)$) from another signal (denoted by $x(n)$).

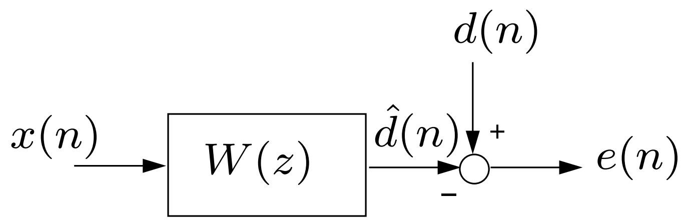


- *Filtering:* estimate $d(n)$ from $x(n) = d(n) + v(n)$.
- *Prediction:* estimate $d(n+\alpha)$ from $x(n), x(n-1), x(n-2), \cdots$

as specified by the error criterium on $e(n)$

A generic problem formulation

The generic problem deals with the (optimal   ) estimation of one signal (denoted by $d(n)$) from another signal (denoted by $x(n)$).



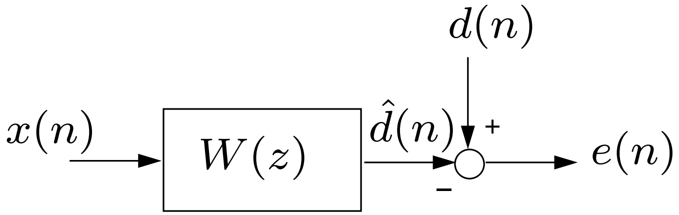
- *Filtering:* estimate $d(n)$ from $x(n) = d(n) + v(n)$.
- *Prediction:* estimate $d(n+\alpha)$ from $x(n), x(n-1), x(n-2), \cdots$
- *Deconvolution:* estimate $d(n)$ from $x(n) = g(n) \star d(n) + w(n)$

---

as specified by the error criterium on $e(n)$

A generic problem formulation

The generic problem deals with the (optimal) estimation of one signal (denoted by $d(n)$) from another signal (denoted by $x(n)$).



- *Filtering:* estimate $d(n)$ from $x(n) = d(n) + v(n)$.
- *Prediction:* estimate $d(n + \alpha)$ from $x(n), x(n-1), x(n-2), \cdots$
- *Deconvolution:* estimate $d(n)$ from $x(n) = g(n) \star d(n) + w(n)$
- *Noise cancellation:* estimate $v_1(n)$ from $v_2(n)$ and subtract it from $x(n) = d(n) + v_1(n)$.

---

as specified by the error criterium on $e(n)$

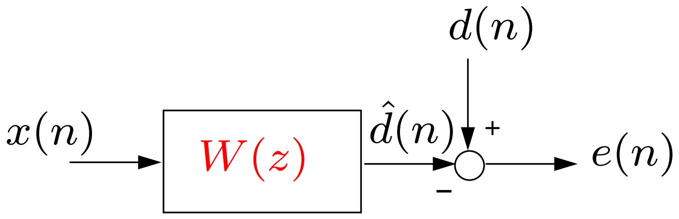
Optimal IIR filtering: Part I

1. Recap FIR Optimal Wiener filter - The Orthogonality Condtion
2. A generic framework
3. **The IIR Wiener filter**
4. The Denoising Problem
5. The deconvolution problem

The IIR Wiener filter problem

Let the filter $W(z)$ be LSI with a double sided impulse response.

Let $x(n), d(n)$ be WSS with mean zero.



Then the estimate $\hat{d}(n)$ is given by:

$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) = w(n) \star x(n)$$

The optimality to find the coefficients $w(\ell)$ is expressed by minimizing the mean-square error:

$$\xi(\mathbf{w}) = E[|e(n)|^2] = E[\left|d(n) - \hat{d}(n)\right|^2] = E[\left|d(n) - w(n) \star x(n)\right|^2]$$

The solution to the IIR Wiener filter problem

THEOREM: Let the conditions stipulated in the previous slide hold, let in addition $\{\mathbf{x}(\mathbf{n}), d(n)\}$ be jointly WSS and the following power and cross-spectra be given:

$$P_{\mathbf{x}}(e^{j\omega}) > 0 \quad P_{d\mathbf{x}}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{d\mathbf{x}}(k)e^{-j\omega k}$$

then the estimate of the signal $d(n)$ from $x(n)$ is derived from

$$\hat{W}(z) = \arg\min_{W(z)} \xi\Big(W(z)\Big) \quad \text{and given as} \quad \hat{D}(z) = \hat{W}(z)X(z)$$

with the optimal filter $\hat{W}(z)$ given by,

The solution to the IIR Wiener filter problem

THEOREM: Let the conditions stipulated in the previous slide hold, let in addition $\{\mathbf{x}(\mathbf{n}), d(n)\}$ be jointly WSS and the following power and cross-spectra be given:

$$P_{\mathbf{x}}(e^{j\omega}) > 0 \quad P_{d\mathbf{x}}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{d\mathbf{x}}(k) e^{-j\omega k}$$

then the estimate of the signal $d(n)$ from $x(n)$ is derived from

$$\hat{W}(z) = \arg\min_{W(z)} \xi\Big(W(z)\Big) \quad \text{and given as} \quad \hat{D}(z) = \hat{W}(z)X(z)$$

with the optimal filter $\hat{W}(z)$ given by,

$$\hat{W}(z) \quad = \quad P_{d\mathbf{x}}(z)P_{\mathbf{x}}(z)^{-1} \quad (\text{"The Wiener-Hopf equations (WH)"})$$

The solution to the IIR Wiener filter problem

THEOREM: Let the conditions stipulated in the previous slide hold, let in addition $\{\mathbf{x}(\mathbf{n}), d(n)\}$ be jointly WSS and the following power and cross-spectra be given:

$$P_{\mathbf{x}}(e^{j\omega}) > 0 \quad P_{d\mathbf{x}}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{d\mathbf{x}}(k) e^{-j\omega k}$$

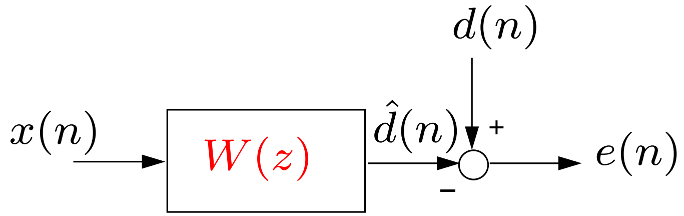
then the estimate of the signal $d(n)$ from $x(n)$ is derived from

$$\hat{W}(z) = \arg\min_{W(z)} \xi\Big(W(z)\Big) \quad \text{and given as} \quad \hat{D}(z) = \hat{W}(z) X(z)$$

with the optimal filter $\hat{W}(z)$ given by,

$$\begin{aligned} \hat{W}(z) &= P_{d\mathbf{x}}(z) P_{\mathbf{x}}(z)^{-1} \quad (\text{"The Wiener-Hopf equations (WH)"}) \\ \xi\Big(\hat{W}(z)\Big) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} P_d(e^{j\omega}) - P_{d\mathbf{x}}(e^{j\omega}) P_{\mathbf{x}}(e^{j\omega})^{-1} P_{d\mathbf{x}}^{*}(e^{j\omega}) d\omega \end{aligned}$$

(1) Solution via The orthogonality condition (OC)



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

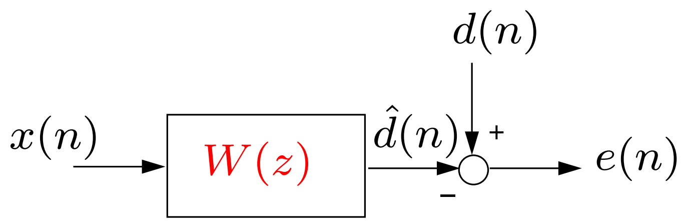
The OC reads: $\boxed{E[e(n)x^*(n-k)] = 0 \;\; \forall k : -\infty < k < \infty}$

$\Rightarrow$

$\Rightarrow$

$\Rightarrow$

(1) Solution via The orthogonality condition (OC)



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

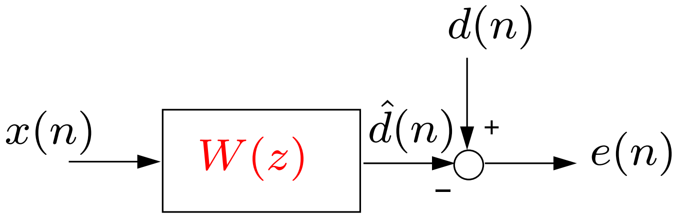
The OC reads: $\boxed{E[e(n)x^*(n-k)] = 0 \ \ \forall k : -\infty < k < \infty}$

$$\Rightarrow \quad E[d(n)x^*(n-k)] - E[\sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)x^*(n-k)] = 0$$

$$\Rightarrow$$

$$\Rightarrow$$

(1) Solution via The orthogonality condition (OC)



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

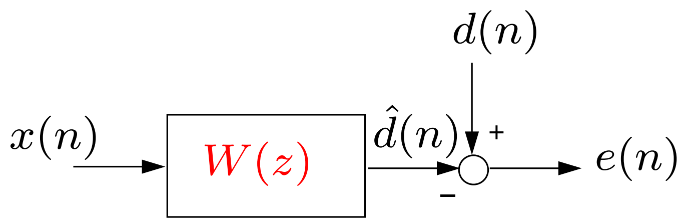
The OC reads: $\boxed{E[e(n)x^*(n-k)] = 0 \;\; \forall k : -\infty < k < \infty}$

$$\Rightarrow \quad E[d(n)x^*(n-k)] - E[\sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)x^*(n-k)] = 0$$

$$\Rightarrow \quad r_{dx}(k) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)r_x(k-\ell) = 0$$

$$\Rightarrow$$

(1) Solution via The orthogonality condition (OC)



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

The OC reads: $\boxed{E[e(n)x^*(n-k)] = 0 \quad \forall k : -\infty < k < \infty}$

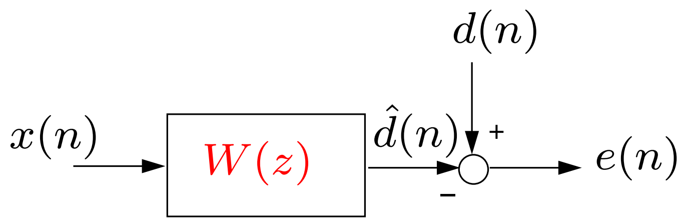
$$\Rightarrow \quad E[d(n)x^*(n-k)] - E[\sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)x^*(n-k)] = 0$$

$$\Rightarrow \quad r_{dx}(k) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)r_x(k-\ell) = 0$$

$$\Rightarrow \quad r_{dx}(k) - \hat{w}(k) \star r_x(k) = 0$$

$$WH \quad \Rightarrow \quad \boxed{P_{dx}(e^{j\omega}) - \hat{W}(e^{j\omega})P_x(e^{j\omega}) = 0 \Rightarrow \hat{W}(z) = P_{dx}(z)P_x(z)^{-1}}$$

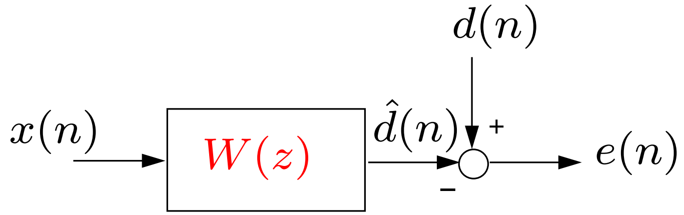
(2) The optimal value of the criterium $\xi_{\min}$



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

$$\xi_{min} = E[e_{min}(n)d^*(n)]$$

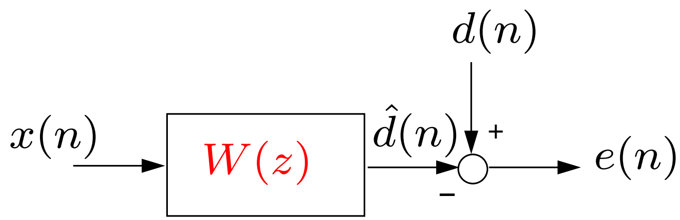
(2) The optimal value of the criterium $\xi_{\min}$



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

$$
\begin{aligned}
\xi_{min} &= E[e_{min}(n)d^*(n)] \\
&= E\Big[\Big(d(n) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)\Big)d^*(n)\Big]
\end{aligned}
$$

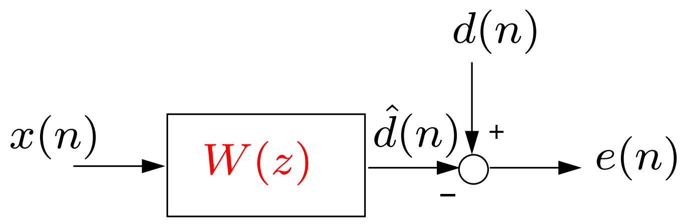
(2) The optimal value of the criterium $\xi_{\min}$



$$\begin{aligned} e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\ \xi(\mathbf{w}) &= E[e(n)e^*(n)] \end{aligned}$$

$$\begin{aligned} \xi_{min} &= E[e_{min}(n)d^*(n)] \\ &= E\left[\left(d(n) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)\right)d^*(n)\right] \\ &= r_d(0) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)\underbrace{E[d(n)x^*(n-\ell)]^*}_{r_{dx}^*(\ell)} \end{aligned}$$

(2) The optimal value of the criterium $\xi_{\min}$



$$
\begin{aligned}
e(n) &= d(n) - \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) \\
\xi(\mathbf{w}) &= E[e(n)e^*(n)]
\end{aligned}
$$

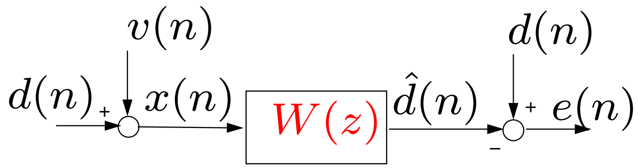
$$
\begin{aligned}
\xi_{min} &= E[e_{min}(n)d^*(n)] \\
&= E\Big[\Big(d(n) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)x(n-\ell)\Big)d^*(n)\Big] \\
&= r_d(0) - \sum_{\ell=-\infty}^{\infty} \hat{w}(\ell)\,\underbrace{E[d(n)x^*(n-\ell)]^*}_{r^*_{dx}(\ell)}
\end{aligned}
$$

$$
\boxed{\xi_{\min} \overset{\text{Parseval}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[P_d(e^{j\omega}) - \hat{W}(e^{j\omega})P^*_{dx}(e^{j\omega})\right] d\omega}
$$

Optimal IIR filtering: Part I

1. A generic framework
2. The IIR Wiener filter
3. **The denoising Problem**
4. The deconvolution Problem

Denoising signals



$$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \;\; \forall n, \ell$$

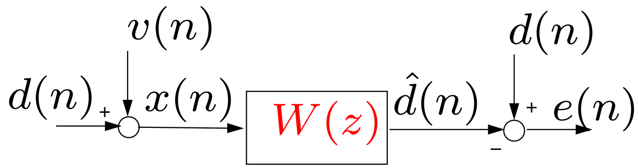
$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell) x(n-\ell)$$

$$\hat{W}(z) = P_{dx}(z) P_x(z)^{-1}$$

$$\begin{aligned} r_{dx}(k) &= E[d(n)x^*(n-k)] = E[d(n)d^*(n-k)] = r_d(k) \\ r_x(k) &= E[(d(n)+v(n))(d^*(n-k)+v^*(n-k))] = r_d(k) + r_v(k) \end{aligned}$$

$$\Rightarrow \hat{W}(z) = \frac{P_d(z)}{P_d(z) + P_v(z)}$$

Denoising signals



$$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \;\; \forall n, \ell$$
$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell)$$

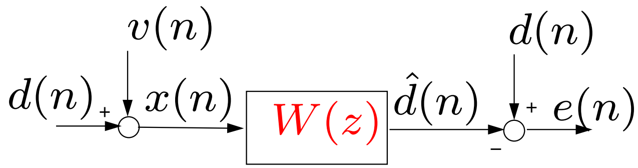
$$\xi_{\min} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \hat{W}(e^{j\omega})P_{dx}^*(e^{j\omega})\right] d\omega \text{ and } \begin{cases} \hat{W}(z) = \frac{P_d(z}{P_d(z)+} \\ r_{dx}(k) = r_d(k) \end{cases}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}P_d^*(e^{j\omega})\right] d\omega$$

$$=$$

$$=$$

Denoising signals



$$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \ \ \forall n, \ell$$
$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell) x(n-\ell)$$

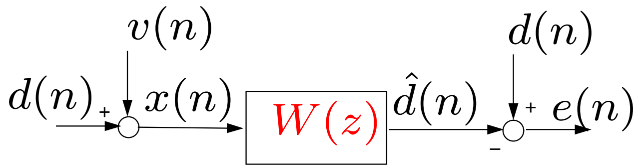
$$\xi_{\min} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \hat{W}(e^{j\omega})P_{dx}^*(e^{j\omega})\right]d\omega \text{ and } \begin{cases} \hat{W}(z) = \frac{P_d(}{P_d(z)+} \\ r_{dx}(k) = r_d(k) \end{cases}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}P_d^*(e^{j\omega})\right]d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{P_d(e^{j\omega})^2 + P_d(e^{j\omega})P_v(e^{j\omega}) - P_d(e^{j\omega})^2}{P_d(e^{j\omega}) + P_v(e^{j\omega})}\right]d\omega$$

$$=$$

Denoising signals



$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \ \ \forall n, \ell$

$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell) x(n-\ell)$

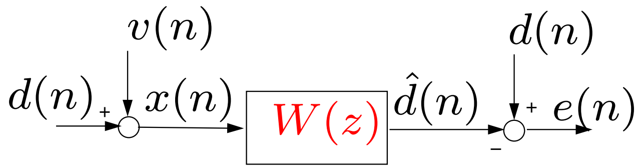
$$\xi_{\min} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \hat{W}(e^{j\omega})P_{dx}^*(e^{j\omega})\right] d\omega \text{ and } \begin{cases} \hat{W}(z) = \frac{P_d(z}{P_d(z)+} \\ r_{dx}(k) = r_d(k) \end{cases}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}P_d^*(e^{j\omega})\right] d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{P_d(e^{j\omega})^2 + P_d(e^{j\omega})P_v(e^{j\omega}) - P_d(e^{j\omega})^2}{P_d(e^{j\omega}) + P_v(e^{j\omega})}\right] d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_v(e^{j\omega})\frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega}}\right] d\omega$$

Denoising signals



$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \ \ \forall n, \ell$

$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell)$

$$\xi_{\min} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \hat{W}(e^{j\omega})P_{dx}^*(e^{j\omega})\right]d\omega \text{ and } \begin{cases} \hat{W}(z) = \frac{P_d(z}{P_d(z)+} \\ r_{dx}(k) = r_d(k) \end{cases}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_d(e^{j\omega}) - \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}P_d^*(e^{j\omega})\right]d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{P_d(e^{j\omega})^2 + P_d(e^{j\omega})P_v(e^{j\omega}) - P_d(e^{j\omega})^2}{P_d(e^{j\omega}) + P_v(e^{j\omega})}\right]d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[P_v(e^{j\omega})\frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}\right]d\omega$$

Conclusion: If $v(n)$ and $d(n)$ have spectra that do not overlap, their product is $0 \ \forall\omega \Rightarrow \xi_{\min} = 0$.

Example 1 (Lecture 6 Ct'd): Denoising (real case)

Consider the AR(1) process $d(n)$ given by ($a = 0.8$):
$d(n+1) = ad(n) + r(n)$ for $r(n)$ ZMWN($\sigma_r^2 = 1 - a^2$) and let the noise $v(n)$ in $x(n) = d(n) + v(n)$ to be also ZMWN($\sigma_v^2 = 1$), then the optimal IIR Wiener filter is:

$$\begin{aligned} \hat{W}(z) &= \frac{P_d(z)}{P_d(z) + P_v(z)} = \frac{(1-a^2)}{(1-a^2) + (1-az^{-1})(1-az)} \\ &= \frac{0.225}{(1-0.5z^{-1})(1-0.5z)} \end{aligned}$$

Resulting in the value of the cost function

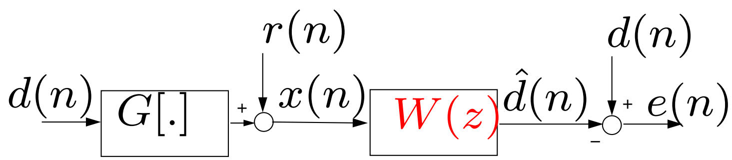
$$\xi_{\min} = \sigma_v^2 \hat{w}(0) = 0.3$$

Ex1Wf.m

Optimal IIR filtering: Part I

1. Recap FIR Optimal Wiener filter - The Orthogonality Condtion
2. A generic framework
3. The IIR Wiener filter
4. The denoising Problem
5. **The deconvolution Problem**

Wiener Deconvolution



Signal model:

$$
\begin{aligned}
x(n) &= g(n) \star d(n) + r(n) \qquad E[d(n)r^*(\ell)] = 0 \ \ \forall n, \ell \\
\hat{d}(n) &= \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell)
\end{aligned}
$$

$$\hat{W}(z) = P_{dx}(z)P_x(z)^{-1}$$

$$
\begin{aligned}
P_{dx}(z) &= G^*(1/z^*)P_d(z) \\
P_x(z) &= G(z)G^*(1/z^*)P_d(z) + P_r(z) \\
\Rightarrow \hat{W}(z) &= \frac{G^*(1/z^*)P_d(z)}{G(z)G^*(1/z^*)P_d(z) + P_r(z)}
\end{aligned}
$$

Optimal Causal IIR filtering: Part II

Highlights from Lecture 3 and 4

Example Lecture 3: Consider:

$$H(z) = \underbrace{\frac{z}{1-0.9z}}_{\text{anti}-\text{causal} \quad \left([\text{H(z)}]_-\right)} + \underbrace{\frac{1}{1-0.9z^{-1}}}_{\text{causal} \quad \left([\text{H(z)}]_+\right)} \qquad \text{ROC}\Big(H(z)\Big) \supset \Gamma$$

with $[H(z)]_-$ containing the coefficients of the series $h(\ell)$ corresponding to positive power of z and $[H(z)]_+$ those to negative power.

Application 3 Lecture 3: Updating cross-correlation function $P_{y\boxed{\text{x}}}(z)$ when

$V(z) = G(z)X(z) \Rightarrow P_{y\boxed{\text{v}}}(z) = G^*(1/z^*)P_{y\boxed{\text{x}}}(z)$.

Highlights from Lecture 3 and 4

Example Lecture 3: Consider:

$$H(z) = \underbrace{\frac{z}{1-0.9z}}_{\text{anti-causal} \quad \left([\mathrm{H(z)}]_-\right)} + \underbrace{\frac{1}{1-0.9z^{-1}}}_{\text{causal} \quad \left([\mathrm{H(z)}]_+\right)} \qquad \text{ROC}\Big(H(z)\Big) \supset \Gamma$$

with $[H(z)]_-$ containing the coefficients of the series $h(\ell)$ corresponding to positive power of z and $[H(z)]_+$ those to negative power.

Application 3 Lecture 3: Updating cross-correlation function $P_{y\boxed{\text{x}}}(z)$ when

$V(z) = G(z)X(z) \Rightarrow P_{y\boxed{\text{v}}}(z) = G^*(1/z^*)P_{y\boxed{\text{x}}}(z)$.

SPECTRAL FACTORIZATION [KOLMOGOROV, WIENER] : A scalar postive real and rational function $P_x(e^{j\omega})$ satisfying $P_x(z) = P_x^*(1/z^*)$ can be factored as:

$$P_x(z) = \sigma Q(z) Q^*(1/z^*)$$

with $Q(z)$ stable, minimum-phase.

Highlights from Lecture 1

Theorem:[Cauchy's Integral Formula (Simplified)] Let $X(z)$ be the z-transform as given by $\sum_{n=-\infty}^{\infty} x(n)z^{-n}$ with ROC containing the unit circle, denoted by the curve Γ in the complex plane, then,

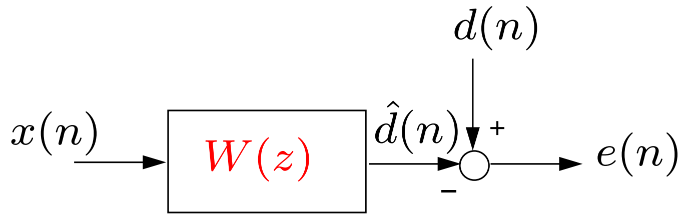
$$x(0) = \frac{1}{2\pi j} \oint_{\Gamma} \frac{X(z)}{z} dz$$

Optimal Causal IIR Filtering: Part II

1. **Recap the mixed causal, anti-causal IIR Wiener filter**
2. The causal IIR Wiener filter Problem
3. Why the causal solution is not the causal part of the mixed causal, anti causal IIR Wiener solution
4. Solution to the causal IIR Wiener filter
5. The filtering Problem

The IIR mixed causal, anti-causal Wiener filter problem

Let the filter $W(z)$ be LSI with a double sided impulse response.
Let $x(n), d(n)$ be WSS with mean zero.



Then the estimate $\hat{d}(n)$ is given by:

$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell)x(n-\ell) = w(n) \star x(n)$$

The optimality to find the coefficients $w(\ell)$ is expressed by minimizing the mean-square error:

$$\xi(\mathbf{w}) = E[|e(n)|^2] = E\left[\left|d(n) - \hat{d}(n)\right|^2\right] = E\left[\left|d(n) - w(n) \star x(n)\right|^2\right]$$

The solution to the IIR mixed causal, anti-causal Wiener filter

THEOREM: Let the conditions stipulated in the previous slide hold, let in addition $\{\mathbf{x}(\mathbf{n}), d(n)\}$ be jointly WSS and the following power and cross-spectra be given:

$$P_{\mathbf{x}}(e^{j\omega}) > 0 \quad P_{d\mathbf{x}}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{d\mathbf{x}}(k) e^{-j\omega k}$$

then the estimate of the signal $d(n)$ from $x(n)$ is derived from

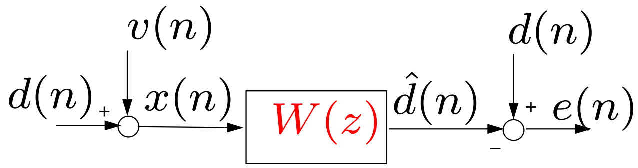
$$\hat{W}(z) = \arg \min_{W(z)} \xi\Big(W(z)\Big) \quad \text{and given as} \quad \hat{D}(z) = \hat{W}(z) X(z)$$

with the optimal filter $\hat{W}(z)$ given by,

$$\hat{W}(z) = P_{d\mathbf{x}}(z) P_{\mathbf{x}}(z)^{-1} \quad \text{("The Wiener-Hopf equations (WH)")}$$

$$\xi\Big(\hat{W}(z)\Big) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_d(e^{j\omega}) - P_{d\mathbf{x}}(e^{j\omega}) P_{\mathbf{x}}(e^{j\omega})^{-1} P_{d\mathbf{x}}^*(e^{j\omega}) d\omega$$

Denoising signals



$$x(n) = d(n) + v(n) \quad E[d(n)v^*(\ell)] = 0 \ \ \forall n, \ell$$

$$\hat{d}(n) = \sum_{\ell=-\infty}^{\infty} w(\ell) x(n-\ell)$$

$$\hat{W}(z) = P_{d\mathbf{x}}(z) P_{\mathbf{x}}(z)^{-1}$$

Example 1 (Part I Ct'd): Denoising (real case)

Consider the AR(1) process $d(n)$ given by ($a = 0.8$):
$d(n+1) = ad(n) + r(n)$ for $r(n)$ ZMWN($\sigma_r^2 = 1 - a^2$) and let the noise $v(n)$ in $x(n) = d(n) + v(n)$ to be also ZMWN($\sigma_v^2 = 1$), then the optimal IIR Wiener filter is:

$$\begin{aligned}\hat{W}(z) &= \frac{P_d(z)}{P_d(z) + P_v(z)} = \frac{(1-a^2)}{(1-a^2) + (1-az^{-1})(1-az)} \\ &= \frac{0.225}{(1-0.5z^{-1})(1-0.5z)}\end{aligned}$$

Resulting in the value of the cost function

$$\xi_{\min} = \sigma_v^2 \hat{w}(0) = 0.3$$

Ex1Wf.m

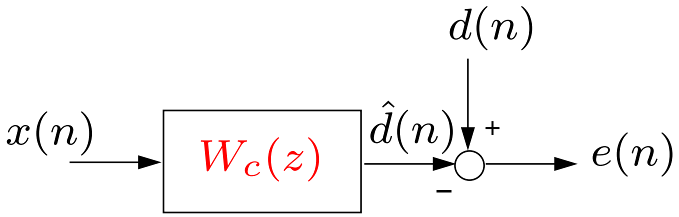
Optimal Causal IIR Filtering: Part II

1. Recap the mixed causal, anti-causal IIR Wiener filter
2. **The causal IIR Wiener filter Problem**
3. Why the causal solution is not the causal part of the mixed causal, anti causal IIR Wiener solution
4. Solution to the causal IIR Wiener filter
5. The filtering Problem

The causal IIR Wiener filter problem

Let the filter $W_c(z)$ be LSI with a **single** sided impulse response.

Let $x(n), d(n)$ be WSS with mean zero.



Then the estimate $\hat{d}(n)$ is given by:

$$\hat{d}(n) = \sum_{\ell=0}^{\infty} w_c(\ell) x(n-\ell) = w_c(n) \star x(n)$$

The optimality to find the coefficients $w_c(\ell)$ is expressed by minimizing the mean-square error:

$$\xi(\mathbf{w_c}) = E[|e(n)|^2] = E\left[\left|d(n) - \hat{d}(n)\right|^2\right] = E\left[\left|d(n) - w_c(n) \star x(n)\right|^2\right]$$

Optimal Causal IIR Filtering: Part II

1. Recap the mixed causal, anti-causal IIR Wiener filter
2. The causal IIR Wiener filter Problem
3. **Why the causal solution is not the causal part of the mixed causal, anti causal IIR Wiener solution**
4. Solution to the causal IIR Wiener filter
5. The filtering Problem

Why?

Minimizing the mean-square error:

$$\xi(\mathbf{w_c}) = E[|e(n)|^2] = E\Big[\Big|d(n) - \hat{d}(n)\Big|^2\Big] = E\Big[\Big|d(n) - \sum_{\ell=0}^{\infty} w_c(\ell)x(n-\ell)\Big|^2\Big]$$

The orthogonality condition for this problem reads,

$$\begin{aligned}
E[e(n)x^*(n-k)] &= 0 \quad \forall k : 0 \leq k < \\
E[d(n)x^*(n-k)] - \sum_{\ell=0}^{\infty} w_c(\ell)E[x(n-\ell)x^*(n-k)] &= 0 \\
r_{dx}(k) - \sum_{\ell=0}^{\infty} w_c(\ell)r_x(k-\ell) &= 0
\end{aligned}$$

Why (Ct'd)?

Recall the result from the orthogonality condition,

$$r_{dx}(k) - \sum_{\ell=0}^{\infty} w_c(\ell) r_x(k-\ell) \quad = \quad 0 \quad \forall k : 0 \leq k < \infty$$

If we now **would** take the z-transform (double sided) we get

$$\begin{aligned}
P_{dx}(z) - \sum_{k=-\infty}^{\infty} \sum_{\ell=0}^{\infty} \hat{w}(\ell) r_x(k-\ell) z^{-k} &= 0 \\
P_{dx}(z) - \sum_{\ell=0}^{\infty} \hat{w}(\ell) \sum_{k=-\infty}^{\infty} r_x(k-\ell) z^{-k} &= 0 \\
P_{dx}(z) - \left(\sum_{\ell=0}^{\infty} \hat{w}(\ell) z^{-\ell} \right) P_x(z) &= 0
\end{aligned}$$

This equation is in general not satisfied since $P_{dx}(z) P_x(z)^{-1}$ is mixed causal, anti-causal.

Optimal Causal IIR Filtering: Part II

1. Recap the mixed causal, anti-causal IIR Wiener filter
2. The causal IIR Wiener filter Problem
3. Why the causal solution is not the causal part of the mixed causal, anti causal IIR Wiener solution
4. **Solution to the causal IIR Wiener filter**
5. The filtering Problem

The causal Wiener-Hopf equations: special case

Recall the causal WH equations,

$$\sum_{\ell=0}^{\infty} \hat{w}_c(\ell) r_x(k-\ell) = r_{dx}(k) \quad \text{for} \;\; 0 \leq k < \infty$$

For which $x(n)$ can we solve this equation exactly?

The causal Wiener-Hopf equations: special case

Recall the causal WH equations,

$$\sum_{\ell=0}^{\infty} \hat{w}_c(\ell) r_x(k-\ell) = r_{dx}(k) \quad \text{for } 0 \leq k < \infty$$

For which $x(n)$ can we solve this equation exactly?

For $x(n)$ ZMWN(1), we have that $r_x(k-\ell) = \Delta(k-\ell)$. Therefore,

$$\hat{w}_c(k) = r_{dx}(k) \quad \text{for } 0 \leq k < \infty$$

Formulated as a z-transform,

$$\hat{W}_c(z) = \Big[\; P_{dx}(z) \; \Big]_+$$

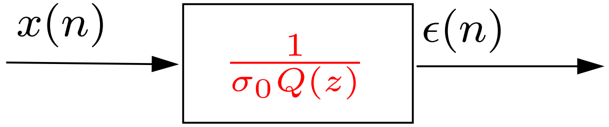
The causal Wiener-Hopf equations: the general case

Whitening $x(n)$

Recall from Lecture 4 - slide 6: Given the spectral factorization of the Power Spectrum $P_x(z)$ as:

$$P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$$

with $Q(z)$ minimum phase. Then filtering $x(n)$ by **causal** $\frac{1}{\sigma_0 Q(z)}$ makes the filtered signal $\epsilon(n)$ ZMWN(1).



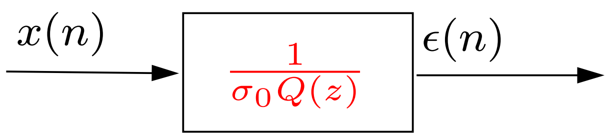
The causal Wiener-Hopf equations: the general case

Whitening $x(n)$

Recall from Lecture 4 - slide 6: Given the spectral factorization of the Power Spectrum $P_x(z)$ as:

$$P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$$

with $Q(z)$ minimum phase. Then filtering $x(n)$ by **causal** $\frac{1}{\sigma_0 Q(z)}$ makes the filtered signal $\epsilon(n)$ ZMWN(1).



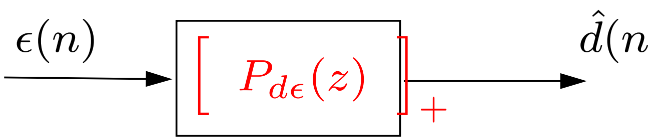
Causal IIR Wiener (ZMWN)

The causal IIR Wiener solution to the problem:

$$\min E[|d(n) - \sum_{\ell=0}^{\infty} w_c(\ell)\epsilon(n)|^2]$$

is given by

$$\left[\; P_{d\epsilon}(z) \; \right]_+$$



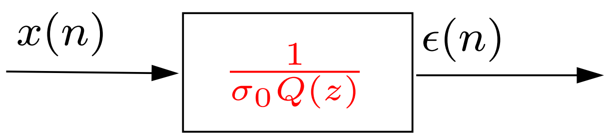
The causal Wiener-Hopf equations: the general case

Whitening $x(n)$

Recall from Lecture 4 - slide 6: Given the spectral factorization of the Power Spectrum $P_x(z)$ as:

$$P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$$

with $Q(z)$ minimum phase. Then filtering $x(n)$ by **causal** $\frac{1}{\sigma_0 Q(z)}$ makes the filtered signal $\epsilon(n)$ ZMWN(1).



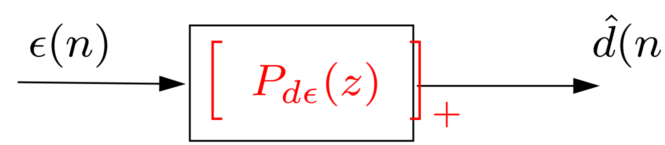
Causal IIR Wiener (ZMWN)

The causal IIR Wiener solution to the problem:

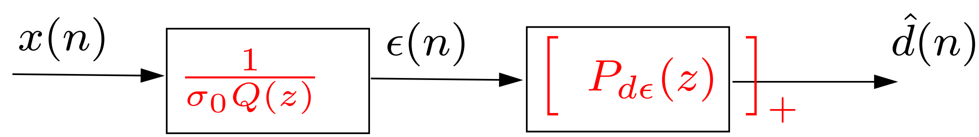
$$\min E[|d(n) - \sum_{\ell=0}^{\infty} w_c(\ell)\epsilon(n)|^2]$$

is given by

$$\left[\; P_{d\epsilon}(z) \; \right]_+$$



The General Solution:



The General Solution from “original” info $P_{dx}(z)$ and $P_x(z)$

1. Calculating $P_{d\epsilon}(z)$ from $P_{dx}(z)$ and $P_x(z)$: This is simply applying application 3 from Lecture 4:

   Since $\epsilon(z) = \frac{1}{\sigma_0 Q(z)} X(z)$,

$$P_{d\boxed{\epsilon}}(z) = \frac{1}{\sigma_0 Q^*(1/z^*)} P_{d\boxed{\text{x}}}(z)$$

The General Solution from "original" info $P_{dx}(z)$ and $P_x(z)$

1. Calculating $P_{d\epsilon}(z)$ from $P_{dx}(z)$ and $P_x(z)$: This is simply applying application 3 from Lecture 4:

   Since $\epsilon(z) = \frac{1}{\sigma_0 Q(z)} X(z)$,

$$P_{d\boxed{\epsilon}}(z) = \frac{1}{\sigma_0 Q^*(1/z^*)} P_{d\boxed{x}}(z)$$

2. *Summary general solution:*

$$x(n) \longrightarrow \boxed{\frac{1}{\sigma_0 Q(z)}} \xrightarrow{\epsilon(n)} \boxed{\left[\, P_{d\epsilon}(z) \,\right]_+} \longrightarrow \hat{d}(n)$$

   With $P_{d\epsilon}(z) = \frac{P_{dx}(z)}{\sigma_0 Q^*(1/z^*)}$, the optimal Causal IIR filter is:

$$\boxed{\hat{W}_c(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q^*(1/z^*)} \right]_+}$$

Optimal Causal IIR Filtering: Part II

1. Recap the mixed causal, anti-causal IIR Wiener filter
2. The causal IIR Wiener filter Problem
3. Why the causal solution is not the causal part of the mixed causal, anti causal IIR Wiener solution
4. Solution to the causal IIR Wiener filter
5. **Applications**

Example 1 (Ct'd): Denoising (real case)

Consider the AR(1) process $s(k)$ given by ($a = 0.8$):
$d(n+1) = ad(n) + r(n)$ for $r(k)$ zmwn with $\sigma_r^2 = 1 - a^2$ and let the noise $v(n)$ in $x(n) = d(n) + v(n)$ to be also zmwn with $\sigma_v^2 = 1$, then for computing the optimal causal IIR Wiener filter we need:

1. to perform a spectral factorization of
   $P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$
2. Compute the causal part of a transfer function

Example 1 (Ct'd): Denoising (real case)

Consider the AR(1) process $s(k)$ given by ($a = 0.8$):
$d(n+1) = ad(n) + r(n)$ for $r(k)$ zmwn with $\sigma_r^2 = 1 - a^2$ and let the noise $v(n)$ in $x(n) = d(n) + v(n)$ to be also zmwn with $\sigma_v^2 = 1$, then for computing the optimal causal IIR Wiener filter we need:

1. to perform a spectral factorization of
   $P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$
2. Compute the causal part of a transfer function

(ad. 1) $P_x(z) = P_d(z) + P_v(z) = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} + 1 = 1.6\frac{(1-0.5z^{-1})(1-0.5z)}{(1-0.8z^{-1})(1-0.8z)}$. This means that,

$$\sigma_0^2 = 1.6 \quad Q(z) = \frac{(1 - 0.5z^{-1})}{(1 - 0.8z^{-1})}$$

Example 1 (Ct'd): Denoising

(ad. 2) We need to compute $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+$. For that purpose we first compute

$$\frac{P_d(z)}{Q^*(1/z^*)} = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} \frac{(1-0.8z)}{(1-0.5z)}$$

Example 1 (Ct'd): Denoising

(ad. 2) We need to compute $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+$. For that purpose we first compute

$$\frac{P_d(z)}{Q^*(1/z^*)} = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)}\frac{(1-0.8z)}{(1-0.5z)}$$

Then we factorize the right hand side into:

$$\frac{0.6}{1-0.8z^{-1}} + \frac{0.3z}{1-0.5z}$$

And therefore $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+ = \frac{0.6}{1-0.8z^{-1}}$. The end result is:

Example 1 (Ct'd): Denoising

(ad. 2) We need to compute $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+$. For that purpose we first compute

$$\frac{P_d(z)}{Q^*(1/z^*)} = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} \frac{(1-0.8z)}{(1-0.5z)}$$

Then we factorize the right hand side into:

$$\frac{0.6}{1-0.8z^{-1}} + \frac{0.3z}{1-0.5z}$$

And therefore $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+ = \frac{0.6}{1-0.8z^{-1}}$. The end result is:

$$\hat{W}_c(z) = \frac{(1-0.8z^{-1})}{1.6(1-0.5z^{-1})} \frac{0.6}{1-0.8z^{-1}} = \frac{0.375}{1-0.5z^{-1}}$$

Example 1 (Ct'd): Denoising

(ad. 2) We need to compute $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+$. For that purpose we first compute

$$\frac{P_d(z)}{Q^*(1/z^*)} = \frac{0.36}{(1-0.8z^{-1})(1-0.8z)} \frac{(1-0.8z)}{(1-0.5z)}$$

Then we factorize the right hand side into:

$$\frac{0.6}{1-0.8z^{-1}} + \frac{0.3z}{1-0.5z}$$

And therefore $\left[\frac{P_d(z)}{Q(1/z^*)}\right]_+ = \frac{0.6}{1-0.8z^{-1}}$. The end result is:

$$\hat{W}_c(z) = \frac{(1-0.8z^{-1})}{1.6(1-0.5z^{-1})} \frac{0.6}{1-0.8z^{-1}} = \frac{0.375}{1-0.5z^{-1}}$$

The value of the cost function is

$$J\left(\hat{W}_c(z)\right) = r_d(0) - \sum_{\ell=0}^{\infty} w_c(\ell) r_d(\ell) = 1 - 0.375 \sum_{\ell=0}^{\infty} 0.5^{\ell} 0.8^{\ell} = 0.375$$

Example 1 (Ct'd): Summary

A summary of the criterium value $\xi_{\min}$ for all the optimal Wiener filter methods and an ad-hoc one is:

Filter	Expression $W(z)$	Value Cost Function $\xi_{\min}$
FIR(2)	$w(0) + w(1)z^{-1}$	0.4048
IIR	$\frac{P_{dx}(z)}{P_x(z)}$	0.3
$\text{IR}_{\text{causal}}$	$\frac{1}{\sigma_0^2 Q(z)} \left[\frac{P_{dx}(z)}{Q^*(1/z^*)} \right]_+$	0.375
Ad Hoc	$\left[\frac{P_{dx}(z)}{P_x(z)}\right]_+$	0.4

Ex1Wf.m

Optimal One-Step Prediction

Approximating the signal $x(n+1)$ by,

$$\hat{x}(n+1) = W_c(z)x(n) = \sum_{\ell=0}^{\infty} w_c(\ell) z^{-\ell} x(n)$$

by minimizing $E\left[\left|x(n+1) - \hat{x}(n+1)\right|^2\right]$. Let the spectral factorization of $P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$ be given then the optimal causal Wiener filter is given as:

$$W_c(z) = \frac{1}{\sigma_0^2 Q(z)} \left[\frac{z P_x(z)}{Q^*(1/z^*)} \right]_+$$

Show that this is equal to $z\left[1 - \frac{1}{Q(z)}\right]$?