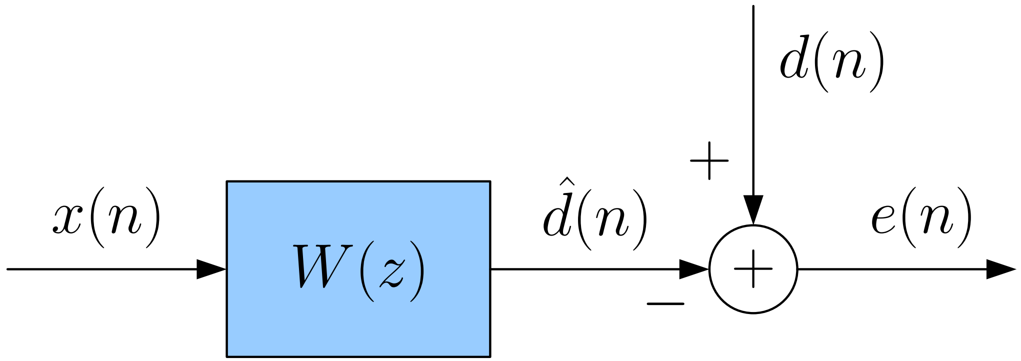
7. OPTIMUM FILTERS

We discuss how one signal can be optimally estimated from another signal (e.g., recover a signal in noise, prediction of future values, noise cancellation, deconvolution, ...).

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



TU Delft

The FIR Wiener filter

Suppose $x(n)$ is a WSS process that is correlated with the WSS process $d(n)$ and we want to estimate $d(n)$ from $x(n)$.

Suppose the estimate $\hat{d}(n)$ is obtained by applying an FIR filter $w(n)$ on $x(n)$:

$$\hat{d}(n) = \sum_{l=0}^{p-1} w(l)x(n-l) = [w(0), \ldots, w(p-1)] \begin{bmatrix} x(n) \\ \vdots \\ x(n-p+1) \end{bmatrix} = \mathbf{w}^T \mathbf{x}$$

The filter coefficients can now be found by minimizing the mean-square error:

$$\begin{aligned} \xi &= E\{|e(n)|^2\} = E\{|\hat{d}(n) - d(n)|^2\} = E\{|\mathbf{w}^T\mathbf{x} - d(n)|^2\} \\ &= \mathbf{w}^T \underbrace{E\{\mathbf{x}\mathbf{x}^H\}}_{\mathbf{R}_x^*} \mathbf{w}^* - \underbrace{E\{d(n)\mathbf{x}^H\}}_{\mathbf{r}_{dx}^T} \mathbf{w}^* - \mathbf{w}^T \underbrace{E\{\mathbf{x}d^*(n)\}}_{\mathbf{r}_{dx}^*} + \underbrace{E\{|d(n)|^2\}}_{r_d(0)} \\ &= \mathbf{w}^H \mathbf{R}_x \mathbf{w} - \mathbf{w}^H \mathbf{r}_{dx} - \mathbf{r}_{dx}^H \mathbf{w} + r_d(0) \end{aligned}$$

Note that $\mathbf{R}_x$ is the autocorrelation matrix containing $r_x(k) = E\{x(n)x^*(n-k)\}$ as entries and $\mathbf{r}_{dx}$ is the cross-correlation vector containing $r_{dx}(k) = E\{d(n)x^*(n-k)\}$ as entries.

TU Delft

The FIR Wiener filter

We can find the optimal filter coefficients by setting the gradient towards $\mathbf{w}^*$ to zero:

$$\Delta_{\mathbf{w}^*}\xi(\mathbf{w},\mathbf{w}^*) = \mathbf{R}_x\mathbf{w} - \mathbf{r}_{dx} = 0 \;\Rightarrow\; \mathbf{R}_x\mathbf{w} = \mathbf{r}_{dx}$$

These are the so-called *Wiener-Hopf equations* which can easily be solved.

It can be shown that the resulting error is uncorrelated with the known data:

$$E\{e(n)\mathbf{x}^H\} = E\{(\mathbf{w}^T\mathbf{x} - d(n))\mathbf{x}^H\} = \mathbf{w}^T\mathbf{R}_x^* - \mathbf{r}_{dx}^T = 0$$

This is known as the *orthogonality principle* or *projection theorem*.

The resulting error is given by

$$\xi_{\min} = r_d(0) - \mathbf{r}_{dx}^H\mathbf{w} = r_d(0) - \mathbf{r}_{dx}^H\mathbf{R}_x^{-1}\mathbf{r}_{dx}$$

The FIR Wiener filter

Filtering

We consider the following data model

$$x(n) = d(n) + v(n) \quad \text{or} \quad \mathbf{x} = \mathbf{d} + \mathbf{v}$$

where $\mathbf{d} = [d(n), \ldots, d(n-p+1)]^T$ and $\mathbf{v} = [v(n), \ldots, v(n-p+1)]^T$.

This results in the following expressions for $\mathbf{R}_x$ and $\mathbf{r}_{dx}$:

$$\mathbf{R}_x = E\{\mathbf{x}^*\mathbf{x}^T\} = E\{(\mathbf{d}+\mathbf{v})^*(\mathbf{d}+\mathbf{v})^T\} = E\{\mathbf{d}^*\mathbf{d}^T\} + E\{\mathbf{v}^*\mathbf{v}^T\} = \mathbf{R}_d + \mathbf{R}_v$$

$$\mathbf{r}_{dx} = E\{d(n)\mathbf{x}^*\} = E\{d(n)\mathbf{d}^*\} = \mathbf{r}_d$$

The Wiener-Hopf equations then become

$$(\mathbf{R}_d + \mathbf{R}_v)\mathbf{w} = \mathbf{r}_d$$

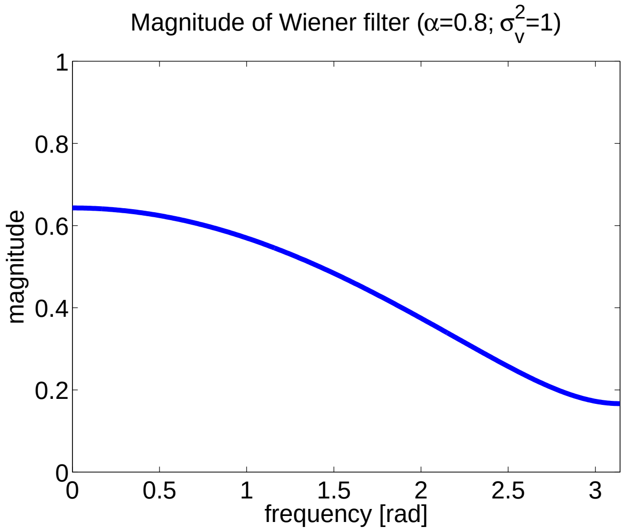
Filtering: example

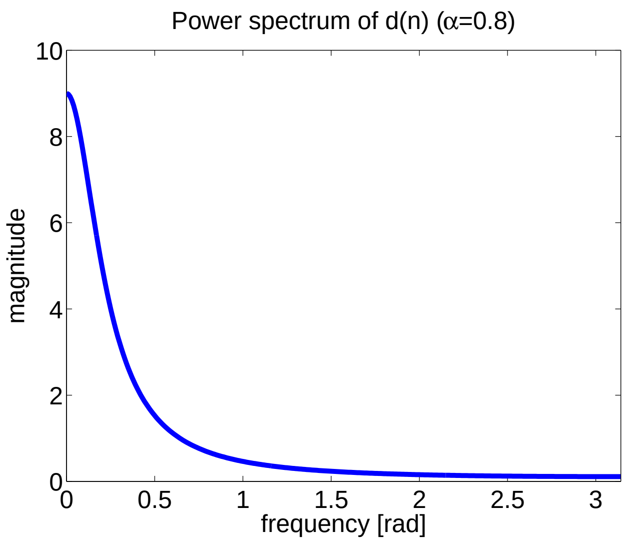
Let $d(n)$ be an AR(1) process with $r_d(k) = \alpha^{|k|}$ and $v(n)$ white noise with variance σ_v^2.

The Wiener-Hopf equations are given by

$$\begin{bmatrix} 1+\sigma_v^2 & \alpha \\ \alpha & 1+\sigma_v^2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix}$$

$$\Leftrightarrow \quad W(z) = \frac{1}{(1+\sigma_v^2)^2 - \alpha^2} [(1+\sigma_v^2-\alpha^2) + \alpha\sigma_v^2 z^{-1}]$$





The FIR Wiener filter

Prediction

We consider the following data model (α-step prediction)

$$d(n) = x(n + \alpha)$$

This results in the following expression for $\mathbf{r}_{dx}$:

$$\mathbf{r}_{dx} = E\{d(n)\mathbf{x}^*\} = E\{x(n+\alpha)\mathbf{x}^*\} = \mathbf{r}_\alpha$$

The Wiener-Hopf equations then become

$$\underbrace{\begin{bmatrix} r_x(0) & r_x^*(1) & \cdots & r_x^*(p-1) \\ r_x(1) & r_x(0) & \cdots & r_x^*(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & \cdots & r_x(0) \end{bmatrix}}_{\mathbf{R}_x} \underbrace{\begin{bmatrix} w(0) \\ w(1) \\ \vdots \\ w(p-1) \end{bmatrix}}_{\mathbf{w}} = \underbrace{\begin{bmatrix} r_x(\alpha) \\ r_x(\alpha+1) \\ \vdots \\ r_x(\alpha+p-1) \end{bmatrix}}_{\mathbf{r}_\alpha}$$

For $\alpha = 1$, this corresponds to the all-pole modeling using Prony's method, the autocorrelation method, or the Yule-Walker method.

The FIR Wiener filter

Prediction in noise

We consider the following data model (α-step prediction)

$$y(n) = x(n) + v(n) \;\text{ or }\; \mathbf{y} = \mathbf{x} + \mathbf{v} \;\text{ and }\; d(n) = x(n+\alpha)$$

where $\mathbf{y} = [y(n), \dots, y(n-p+1)]^T$ and $\mathbf{v} = [v(n), \dots, v(n-p+1)]^T$.

This results in the following expressions for $\mathbf{R}_y$ and $\mathbf{r}_{dy}$:

$$\mathbf{R}_y = E\{\mathbf{y}^*\mathbf{y}^T\} = E\{(\mathbf{x}+\mathbf{v})^*(\mathbf{x}+\mathbf{v})^T\} = E\{\mathbf{x}^*\mathbf{x}^T\} + E\{\mathbf{v}^*\mathbf{v}^T\} = \mathbf{R}_x + \mathbf{R}_v$$

$$\mathbf{r}_{dy} = E\{d(n)\mathbf{y}^*\} = E\{x(n+\alpha)(\mathbf{x}^* + \mathbf{v}^*)\} = E\{x(n+\alpha)\mathbf{x}^*) = \mathbf{r}_\alpha$$

The Wiener-Hopf equations then become

$$(\mathbf{R}_x + \mathbf{R}_v)\mathbf{w} = \mathbf{r}_\alpha$$

The FIR Wiener filter

Prediction: example

Let us focus on one-step prediction

Let $x(n)$ be an AR(1) process with $r_x(k) = \alpha^{|k|}$ and $v(n)$ white noise with variance σ_v^2.

The Wiener-Hopf equations are given by

$$\begin{bmatrix} 1+\sigma_v^2 & \alpha \\ \alpha & 1+\sigma_v^2 \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha^2 \end{bmatrix}$$

$$\Leftrightarrow \quad W(z) = \frac{\alpha}{(1+\sigma_v^2)^2 - \alpha^2}[(1+\sigma_v^2 - \alpha^2) + \alpha\sigma_v^2 z^{-1}]$$

In the noise-free case (take $\sigma_v^2 \to 0$), we obtain

$$W(z) = \alpha \qquad \text{or} \qquad \hat{x}(n+1) = \alpha x(n)$$

This makes sense since $x(n+1) = \alpha x(n) + i(n)$, where $i(n)$ is white noise that can not be predicted from past values of $x(n)$ or $i(n)$.

The FIR Wiener filter

Deconvolution

We consider a noisy convolutive model, with an FIR filter $g(n)$ of order L:

$$x(n) = g(n) * d(n) + v(n) \quad \text{or} \quad \mathbf{x} = \mathbf{G}\mathbf{d} + \mathbf{v}$$

where $\mathbf{d} = [d(n), \ldots, d(n-p+1), \ldots, d(n-L-p+1)]^T$, $\mathbf{v} = [v(n), \ldots, v(n-p+1)]^T$,

and $\mathbf{G} = \begin{bmatrix} g(0) & \cdots & g(L) & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & g(0) & \cdots & g(L) \end{bmatrix}$.

This results in the following expressions for $\mathbf{R}_x$ and $\mathbf{r}_{dx}$:

$$\mathbf{R}_x = E\{\mathbf{x}^*\mathbf{x}^T\} = \mathbf{G}^* E\{\mathbf{d}^*\mathbf{d}^T\}\mathbf{G}^T + E\{\mathbf{v}^*\mathbf{v}^T\} = \mathbf{G}^*\mathbf{R}_d\mathbf{G}^T + \mathbf{R}_v$$

$$\mathbf{r}_{dx} = E\{d(n)\mathbf{x}^*\} = \mathbf{G}^* E\{d(n)\mathbf{d}^*\} = \mathbf{G}^*\mathbf{r}_d$$

The Wiener-Hopf equations then become

$$(\mathbf{G}^*\mathbf{R}_d\mathbf{G}^T + \mathbf{R}_v)\mathbf{w} = \mathbf{G}^*\mathbf{r}_d$$

Noise cancellation

We consider the same data model as for filtering:

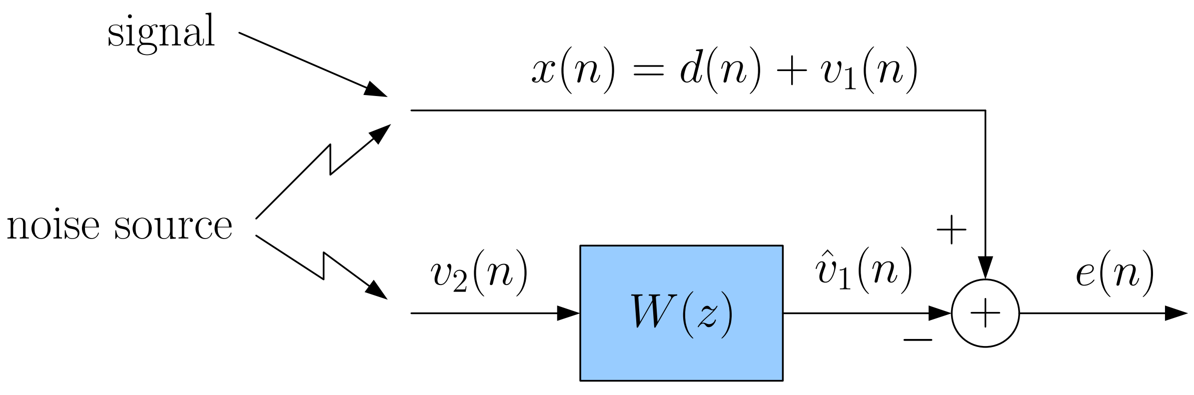
$$x(n) = d(n) + v_1(n) \;\text{ or }\; \mathbf{x} = \mathbf{d} + \mathbf{v}_1$$

where $\mathbf{d} = [d(n), \ldots, d(n-p+1)]^T$ and $\mathbf{v}_1 = [v_1(n), \ldots, v_1(n-p+1)]^T$.

This time we estimate $v_1(n)$ from a correlated noise source $v_2(n)$, and estimate $d(n)$ as

$$\hat{d}(n) = x(n) - \hat{v}_1(n) \;\text{ with }\; \hat{v}_1(n) = \mathbf{w}^T \mathbf{v}_2$$

where $\mathbf{v}_2 = [v_2(n), \ldots, v_2(n-p+1)]^T$.



The FIR Wiener filter

Noise cancellation

To estimate $v_1(n)$ from $v_2(n)$, we start from the Wiener-Hopf equations

$$\mathbf{R}_{v_2}\mathbf{w} = \mathbf{r}_{v_1 v_2}$$

Since $\mathbf{r}_{v1v2}$ is generally not known, we can rewrite this as

$$\mathbf{r}_{v_1 v_2} = E\{v_1(n)\mathbf{v}_2^*\} = E\{(d(n) + v_1(n))\mathbf{v}_2^*\} = E\{x(n)\mathbf{v}_2^*\} = \mathbf{r}_{x v_2}$$

and thus the Wiener-Hopf equations can be written as

$$\mathbf{R}_{v_2}\mathbf{w} = \mathbf{r}_{x v_2}$$

As already mentioned, $d(n)$ is then estimated as

$$\hat{d}(n) = x(n) - \hat{v}_1(n) \ \text{ with } \ \hat{v}_1(n) = \mathbf{w}^T\mathbf{v}_2$$

TUDelft

Echo cancellation: special case of noise cancellation