20. Weiner filters and deconvolution

ECE 830, Spring 2014

Linear Minimum Mean Square Error Estimator

Suppose our data is $x \in \mathbb{R}^n$, a random vector governed by a distribution $p(x|\theta)$, which depends on the parameter θ. Moreover, the parameter $\theta \in \mathbb{R}^k$ is treated as a random variable with $\mathbb{E}[\theta] = 0$ and $\mathbb{E}[\theta\theta^\top] = \Sigma_{\theta\theta}$. Also, assume that $\mathbb{E}[x] = 0$ and let $\Sigma_{xx} := \mathbb{E}[xx^\top]$ and $\Sigma_{\theta x} := \mathbb{E}[\theta x^\top]$. Then, as we saw in the previous lecture, the linear filter that provides the minimum MSE is given by:

$$\widehat{A} = \underset{A \in \mathbb{R}^{n \times k}}{\arg\min}\, \mathbb{E}\big[\|\theta - A^\top x\|_2^2\big] = \Sigma_{xx}^{-1}\Sigma_{x\theta}$$

Definition: LMMSE

The *linear minimum MSE estimator* (LMMSE) estimator is:

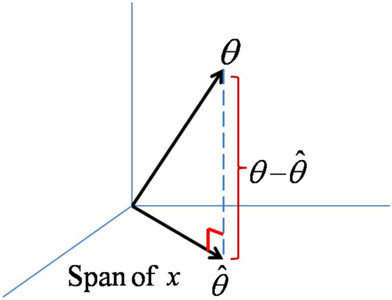
$$\widehat{\theta} = \widehat{A}^\top x = \Sigma_{\theta x}\Sigma_{xx}^{-1}x.$$

Orthogonality Principle

Let $\widehat{\theta} = \Sigma_{\theta x}\Sigma_{xx}^{-1}x$ be the LMMSE estimator, defined above. Then

$$\begin{aligned} \mathbb{E}\big[(\theta - \widehat{\theta})x^\top\big] &= \mathbb{E}\big[\big(\theta - \Sigma_{\theta x}\Sigma_{xx}^{-1}x\big)x^\top\big] \\ &= \Sigma_{\theta x} - \Sigma_{\theta x}\Sigma_{xx}^{-1}\Sigma_{xx} \\ &= 0\ . \end{aligned}$$

In other words, the error $(\theta - \widehat{\theta})$ is orthogonal to the data X. This is shown graphically in Fig. 1.



Orthogonality between the estimator $\widehat{\theta}$ and its error $\theta - \widehat{\theta}$.

LMMSE and Orthogonality

The orthogonality principle also provides a method for deriving the LMMSE filter. Consider any linear estimator of the form $\widehat{\theta} = B^\top x$. If we impose the orthogonality condition

$$0 = \mathbb{E}\big[(\theta - \widehat{\theta})x^\top\big] \quad = \quad \Sigma_{\theta x} - B^\top \Sigma_{xx}$$

then we see that $B^\top$ must be equal to $\Sigma_{\theta x}\Sigma_{xx}^{-1}$.

The Wiener Filter

When the expected values of the parameter $\theta \in \mathbb{R}^k$ and the data $x \in \mathbb{R}^n$ are zero, then the Wiener filter A_{opt} is obtained by minimizing the mean square error between the parameter and estimator:

$$A_{opt} = \underset{A:\widehat{\theta}=Ax}{\text{arg min}}\ \mathbb{E}\big[\|\theta - Ax\|_2^2\big]$$

which results in

$$A_{opt} = \Sigma_{\theta x}\Sigma_{xx}^{-1},$$

which is the LMMSE estimator when both the data and the parameter are jointly Gaussian distributed.

Classical derivation of the Wiener Filter

We start with the model $x = s + w$ where x, s, and w are wide-sense stationary processes. We express them as time series

$$x[n] = s[n] + w[n] \ .$$

We aim at defining a filter $h[n]$ that will be convolved with $x[n]$ to estimate $s[n]$

$$\widehat{s}[n] \ = \ \sum_k h[k]x[n-k] \ .$$

Our filter should minimize the MSE:

$$\begin{aligned}\mathsf{MSE}(\widehat{s}[n]) &= \mathbb{E}\big[(s[n] - \widehat{s}[n])^2\big] \\ &= \mathbb{E}\Big[s[n]^2 - 2s[n]\sum_k h[k]x[n-k] \\ &\quad + \left(\sum_k h[k]x[n-k]\right)^2\Big] \ .\end{aligned}$$

Differentiating with respect to $h[m]$ and setting the derivative equal to zero yields

$$\begin{aligned}\frac{\partial \mathsf{MSE}(\widehat{s}[n])}{\partial h[m]} &= \\ \\ &= \\ \\ &= \\ \\ &= 0 \,.\end{aligned}$$

Wiener-Hopf equations

The optimal filter satisfies

$$R_{ss}[m] = \sum_k h[k](R_{ss}[m-k] + R_{ww}[m-k])$$

which is the Wiener-Hopf equation.

Taking the Discrete-Time Fourier Transform (DTFT) of both sides, we get

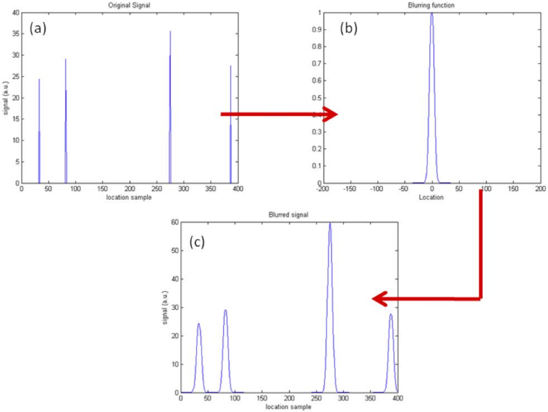
$$S_{ss}(\omega) = H(\omega)\big(S_{ss}(\omega) + S_{ww}(\omega)\big)$$

where $S_{ss}(\omega)$ and $S_{ww}(\omega)$ are the power spectra of the signal and the noise process, respectively. Therefore, the frequency response of the Wiener filter is

$$H(\omega) = \frac{S_{ss}(\omega)}{S_{ss}(\omega) + S_{ww}(\omega)} \ .$$

Deconvolution

The final topic of this lecture is deconvolution. We model the detected signal as $x = Gs + w$ where G is a circular convolution operator



Blurring process. (a) Original impulse signal. (b) Blurring function. (c). Blurred signal

As in the previous lecture,

$$s \sim \mathcal{N}(0, U\Lambda_s U^{\mathsf{H}}),$$
$$w \sim \mathcal{N}(0, U\Lambda_w U^{\mathsf{H}}).$$

Furthermore, since G is circulant,

$$G = UDU^{\mathsf{H}}$$

where D is a diagonal matrix, which is the frequency response of G. In this case, the Wiener filter solution is computed as follows:

$$\begin{aligned}
\widehat{s} &= \Sigma_{ss}G^{\mathsf{H}}(G\Sigma_{ss}G^{\mathsf{H}} + \Sigma_{ww})^{-1}x \\
&= U\Lambda_s U^{\mathsf{H}}G^{\mathsf{H}}(GU\Lambda_s U^{\mathsf{H}}G^{\mathsf{H}} + U\Lambda_w U^{\mathsf{H}})^{-1}x \\
&= U\Lambda_s U^{\mathsf{H}}UD^{\mathsf{H}}U^{\mathsf{H}}(UDU^{\mathsf{H}}U\Lambda_s U^{\mathsf{H}}UD^{\mathsf{H}}U^{\mathsf{H}} + U\Lambda_w U^{\mathsf{H}})^{-1}x \\
&= U\Lambda_s D^{\mathsf{H}}(D\Lambda_s D^{\mathsf{H}} + \Lambda_w)^{-1}U^{\mathsf{H}}x \\
&= U\widetilde{D}U^{\mathsf{H}}x
\end{aligned}$$

where $\widetilde{D}(k,k) = \frac{D^{\mathsf{H}}(k,k)}{|D(k,k)|^2 + P^{-1}(k,k)}$ and $P(k,k) = \frac{\Lambda_s(k,k)}{\Lambda_w(k,k)}$.

Classical Wiener Filter for Deconvolution

In the case of a blurred, noise time series modeled as

$$x[n] = g[n] * s[n] + w[n]$$

we aim at obtaining a filter $h[n]$ such that the estimator of the deblurred, noiseless signal is computed from

$$\widehat{s}[n] = \sum_k h[k]x[n-k].$$

Using a derivation similar to the classical derivation a few slides ago, we arrive at the Fourier domain filter

$$H(\omega) = \frac{G^*(\omega)S_{ss}(\omega)}{|G(\omega)|^2 S_{ss}(\omega) + S_{ww}(\omega)}$$

where $G(\omega)$ is the transfer function of the blurring filter $g[n]$ and G^* is its complex conjugate.