430.714 – Fall 2017

Estimation Theory

Lecture 4

Prof. Songhwai Oh

ECE, SNU

Linear Models

$\underline{\text{Linear model}}$: $\mathbf{x} = \mathbf{H}\theta + \mathbf{w}$, where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ (white noise)

Example

$$x[n] = A + Bn + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad w[n] \sim \mathcal{N}(0, \sigma^2)$$

$$\begin{aligned} \mathbf{x} &= [x[0]\, x[1] \ldots x[N-1]]^T \\ \mathbf{w} &= [w[0]\, w[1] \ldots w[N-1]]^T \\ \boldsymbol{\theta} &= [A\, B]^T \end{aligned} \qquad \mathbf{H} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & N-1 \end{bmatrix} \qquad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$$

$\hat{\theta} = \mathbf{g}(\mathbf{x})$ is the MVU estimator if $\dfrac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{I}(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x}) - \boldsymbol{\theta})$

$$\begin{aligned} \frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{\partial}{\partial \boldsymbol{\theta}}\left[-\ln(2\pi\sigma^2)^{\frac{N}{2}} - \frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})\right] \\ &= -\frac{1}{2\sigma^2}\frac{\partial}{\partial \boldsymbol{\theta}}\left[\mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{H}\boldsymbol{\theta} + \boldsymbol{\theta}^T\mathbf{H}^T\mathbf{H}\boldsymbol{\theta}\right]. \\ &= \frac{1}{\sigma^2}[\mathbf{H}^T\mathbf{x} - \mathbf{H}^T\mathbf{H}\boldsymbol{\theta}] \\ &= \frac{\mathbf{H}^T\mathbf{H}}{\sigma^2}[(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x} - \boldsymbol{\theta}] \qquad (\text{if } \mathbf{H}^T\mathbf{H} \text{ is invertible}) \end{aligned}$$

MVU estimator: $\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$ $\qquad \hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2(\mathbf{H}^T\mathbf{H})^{-1})$

$\mathbf{I}(\boldsymbol{\theta}) = \dfrac{\mathbf{H}^T\mathbf{H}}{\sigma^2}.$ $\qquad \mathbf{C}_{\hat{\theta}} = \mathbf{I}^{-1}(\boldsymbol{\theta}) = \sigma^2(\mathbf{H}^T\mathbf{H})^{-1}$

Example: Fourier Analysis for Periodic Signals

$$x[n] = \sum_{k=1}^{M} a_k \cos\left(\frac{2\pi kn}{N}\right) + \sum_{k=1}^{M} b_k \sin\left(\frac{2\pi kn}{N}\right) + w[n] \qquad n = 0, 1, \ldots, N-1$$

$$\mathbf{x} = \mathbf{H}\theta + \mathbf{w}$$

$$\boldsymbol{\theta} = [a_1\, a_2 \ldots a_M\, b_1\, b_2 \ldots b_M]^T$$

$$\mathbf{H} = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \cos\left(\frac{2\pi}{N}\right) & \cdots & \cos\left(\frac{2\pi M}{N}\right) & \sin\left(\frac{2\pi}{N}\right) & \cdots & \sin\left(\frac{2\pi M}{N}\right) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \cos\left[\frac{2\pi(N-1)}{N}\right] & \cdots & \cos\left[\frac{2\pi M(N-1)}{N}\right] & \sin\left[\frac{2\pi(N-1)}{N}\right] & \cdots & \sin\left[\frac{2\pi M(N-1)}{N}\right] \end{bmatrix}$$

$$\mathbf{H} = [\mathbf{h}_1\, \mathbf{h}_2 \ldots \mathbf{h}_{2M}]$$

Orthogonality (DFT):

$$\mathbf{h}_i^T \mathbf{h}_j = 0 \qquad \text{for } i \neq j$$

$$\begin{aligned} \sum_{n=0}^{N-1} \cos\left(\frac{2\pi in}{N}\right)\cos\left(\frac{2\pi jn}{N}\right) &= \frac{N}{2}\delta_{ij} \\ \sum_{n=0}^{N-1} \sin\left(\frac{2\pi in}{N}\right)\sin\left(\frac{2\pi jn}{N}\right) &= \frac{N}{2}\delta_{ij} \\ \sum_{n=0}^{N-1} \cos\left(\frac{2\pi in}{N}\right)\sin\left(\frac{2\pi jn}{N}\right) &= 0 \quad \text{for all } i, j. \end{aligned}$$

Example: Fourier Analysis for Periodic Signals

$$x[n] = \sum_{k=1}^{M} a_k \cos\left(\frac{2\pi kn}{N}\right) + \sum_{k=1}^{M} b_k \sin\left(\frac{2\pi kn}{N}\right) + w[n] \qquad n = 0, 1, \ldots, N-1$$

$$\mathbf{H}^T\mathbf{H} = \begin{bmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_{2M}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 & \ldots & \mathbf{h}_{2M} \end{bmatrix} \qquad \mathbf{h}_i^T\mathbf{h}_j = 0 \quad \text{for } i \neq j$$

$$= \begin{bmatrix} \mathbf{h}_1^T\mathbf{h}_1 & \mathbf{h}_1^T\mathbf{h}_2 & \ldots & \mathbf{h}_1^T\mathbf{h}_{2M} \\ \mathbf{h}_2^T\mathbf{h}_1 & \mathbf{h}_2^T\mathbf{h}_2 & \ldots & \mathbf{h}_2^T\mathbf{h}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{2M}^T\mathbf{h}_1 & \mathbf{h}_{2M}^T\mathbf{h}_2 & \ldots & \mathbf{h}_{2M}^T\mathbf{h}_{2M} \end{bmatrix} = \begin{bmatrix} \frac{N}{2} & 0 & \ldots & 0 \\ 0 & \frac{N}{2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{N}{2} \end{bmatrix} = \frac{N}{2}\mathbf{I}$$

MVU estimator: $\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$

$$= \frac{2}{N}\mathbf{H}^T\mathbf{x} = \frac{2}{N} \begin{bmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_{2M}^T \end{bmatrix} \mathbf{x}$$

$$= \begin{bmatrix} \frac{2}{N}\mathbf{h}_1^T\mathbf{x} \\ \vdots \\ \frac{2}{N}\mathbf{h}_{2M}^T\mathbf{x} \end{bmatrix}$$

$$\hat{a}_k = \frac{2}{N} \sum_{n=0}^{N-1} x[n] \cos\left(\frac{2\pi kn}{N}\right)$$

$$\hat{b}_k = \frac{2}{N} \sum_{n=0}^{N-1} x[n] \sin\left(\frac{2\pi kn}{N}\right)$$

Linear model: $\mathbf{x} = \mathbf{H}\theta + \mathbf{w}$, where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ (colored noise)

$\mathbf{C}^{-1} = \mathbf{D}^T\mathbf{D}$; covariance matrix is positive definite

D: whitening transformation

$$
\begin{aligned}
E\left[(\mathbf{D}\mathbf{w})(\mathbf{D}\mathbf{w})^T\right] &= \mathbf{D}\mathbf{C}\mathbf{D}^T \\
&= \mathbf{D}\mathbf{D}^{-1}\mathbf{D}^{T^{-1}}\mathbf{D}^T = \mathbf{I}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{x}' &= \mathbf{D}\mathbf{x} \\
&= \mathbf{D}\mathbf{H}\boldsymbol{\theta} + \mathbf{D}\mathbf{w} \\
&= \mathbf{H}'\boldsymbol{\theta} + \mathbf{w}' \qquad \mathbf{w}' = \mathbf{D}\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})
\end{aligned}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\theta}} &= (\mathbf{H}'^T\mathbf{H}')^{-1}\mathbf{H}'^T\mathbf{x}' \qquad \mathbf{C}_{\hat{\theta}} = (\mathbf{H}'^T\mathbf{H}')^{-1} \\
&= (\mathbf{H}^T\mathbf{D}^T\mathbf{D}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{D}^T\mathbf{D}\mathbf{x}
\end{aligned}
$$

MVU estimator:

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x} \qquad \mathbf{C}_{\hat{\theta}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}$$

General Linear Model:

$h_i[t]$: nonlinear function of t

$$x(t) = \sum_{i=1}^{M} \theta_i h_i(t) + w(t), \qquad w(t) \sim \mathcal{N}(0, \sigma^2)$$

$$h_i(t) = t^{i-1}$$

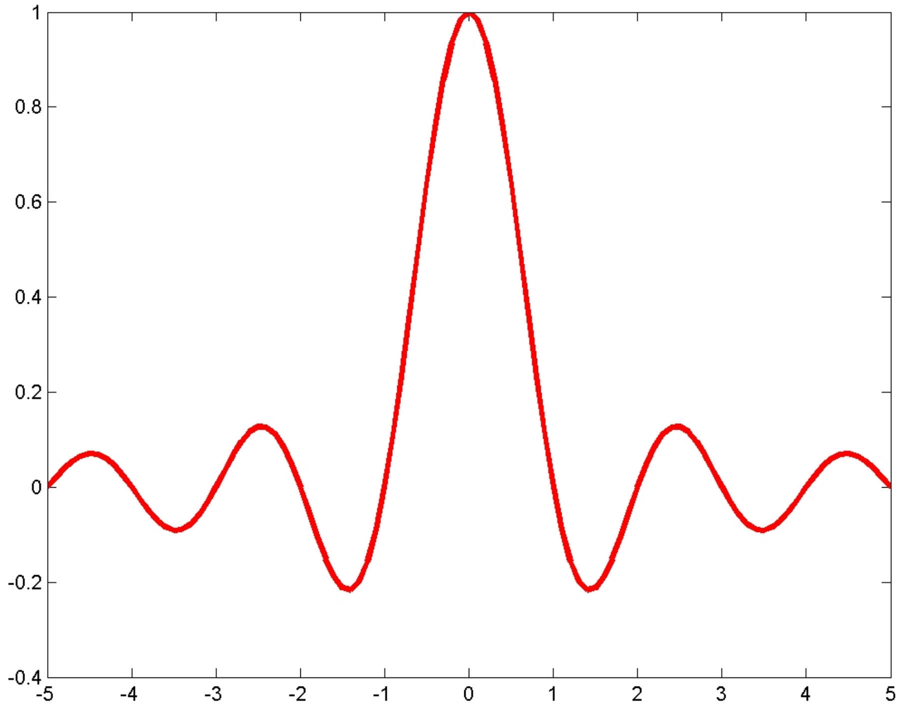
$$h_i(t) = \exp\left(-\frac{1}{2s^2}(t - m_i)^2\right)$$

$$\mathbf{x} = \mathbf{H}\theta + \mathbf{w}$$

$$\mathbf{H} = \begin{bmatrix} h_1(t_0) & h_2(t_0) & \cdots & h_M(t_0) \\ h_1(t_1) & h_2(t_1) & \cdots & h_M(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ h_1(t_{N-1}) & h_2(t_{N-1}) & \cdots & h_M(t_{N-1}) \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_M \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w(t_0) \\ w(t_1) \\ \vdots \\ w(t_{N-1}) \end{bmatrix}$$

$$\text{MVU Estimator: } \hat{\theta} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$$

Example: Linear modeling of the SINC function



Model 1: $x(t) = A + Bt + w(t), \quad w(t) \sim \mathcal{N}(0, \sigma^2)$

Model 2: $x(t) = \sum_{i=1}^{M} \theta_i h_i(t) + w(t), \quad w(t) \sim \mathcal{N}(0, \sigma^2)$ $\qquad h_i(t) = \exp\left(-\frac{1}{2s^2}(t - m_i)^2\right)$

Example: Linear modeling of the SINC function

N=50 | N=100 | N=1000

Data

Linear Model 1

Linear Model 2