E9 211: Adaptive Signal Processing

Lecture 7: Linear estimation

Outline

1. Optimal estimator in the vector case (Ch. 2.1)
2. Normal equations (Ch. 3.1 and 3.2)

Vector-valued data

- Suppose $\mathbf{x} : p \times 1$ and $\mathbf{y} : q \times 1$ are vector valued. Then, we denote the estimator for $\mathbf{x}$ as $\hat{\mathbf{x}} = \mathbf{h}(\mathbf{y})$. Explicitly,

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_{p-1} \end{bmatrix} = \begin{bmatrix} h_0(\mathbf{y}) \\ h_1(\mathbf{y}) \\ \vdots \\ h_{p-1}(\mathbf{y}) \end{bmatrix}$$

- We can then seek optimal functions $\{h_k(\cdot)\}$ that minimizes the error in each component of $\mathbf{x}$:

$$\underset{h_k(\cdot)}{\text{minimize}} \quad E(|\tilde{x}_k|^2) = E(|x_k - h_k(\mathbf{y})|^2)$$

- Therefore, the optimal estimator for x_k given $\mathbf{y}$ in the least-mean-square error sense is $\hat{x}_k = E(x_k|\mathbf{y})$

Vector-valued data

- Suppose $\tilde{\mathbf{x}} = [x_0 - \hat{x}_0, x_1 - \hat{x}_1, \ldots, x_{p-1} - \hat{x}_{p-1}]^{\mathrm{T}}$. Then,

$$E(\tilde{\mathbf{x}}^{\mathrm{H}}\tilde{\mathbf{x}}) = E(|\tilde{x}_0|^2) + E(|\tilde{x}_1|^2) + \cdots + E(|\tilde{x}_{p-1}|^2) = \mathrm{Tr}(\mathbf{R}_{\tilde{\mathbf{x}}})$$

- Since each term depends only on the corresponding function $h_k(\cdot)$, minimizing the sum over each $h_k(\cdot)$ is equivalent to minimizing the sum over all $\{h_k(\cdot)\}$, i.e.,

$$\underset{h_k(\cdot)}{\text{minimize}} \quad E(|\tilde{x}_k|^2) = E(|x_k - h_k(\mathbf{y})|^2)$$

is equivalent to minimizing the trace of the error covariance matrix

$$\underset{\{h_k(\cdot)\}}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{R}_{\tilde{\mathbf{x}}})$$

Linear estimator

- Suppose
$$\bar{\mathbf{x}} = E(\mathbf{x}) = \mathbf{0};\ \bar{\mathbf{y}} = E(\mathbf{y}) = \mathbf{0};$$
and
$$\mathbf{R}_x = E(\mathbf{x}\mathbf{x}^{\mathrm{H}});\ \mathbf{R}_y = E(\mathbf{y}\mathbf{y}^{\mathrm{H}});\ \mathbf{R}_{xy} = E(\mathbf{x}\mathbf{y}^{\mathrm{H}})$$
- We restrict to a subclass of estimators of the form
$$\mathbf{h}(\mathbf{y}) = \mathbf{K}\mathbf{y} + \mathbf{b}$$
$\mathbf{K} : p \times q$ and vector $\mathbf{b} : p \times 1$
- Such linear estimators will depend on the first- and second-order moments of $\mathbf{x}$ and $\mathbf{y}$ and the full knowledge of the conditional pdf is not required.

Linear estimator

We find $\mathbf{K}$ and $\mathbf{b}$ such that the

- estimator is unbiased
- the trace of the error covariance matrix is minimized
- For unbiasedness, the following equation must be satisfied

$$E(\hat{\mathbf{x}}) = E(\mathbf{K}\mathbf{y} + \mathbf{b}) = \mathbf{K}E(\mathbf{y}) + \mathbf{b} = \mathbf{b}$$

  This means, we must have $\mathbf{b} = \mathbf{0}$.

- Explicitly,

$$\hat{\mathbf{x}} = \mathbf{K}\mathbf{y} = \begin{bmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_{p-1} \end{bmatrix} = \begin{bmatrix} \mathbf{k}_0^H \mathbf{y} \\ \mathbf{k}_1^H \mathbf{y} \\ \vdots \\ \mathbf{k}_{p-1}^H \mathbf{y} \end{bmatrix}$$

Linear estimator

To find $\mathbf{K}$ we solve

$$\underset{\mathbf{K}}{\text{minimize}} \quad \text{Tr}(\mathbf{R}_{\tilde{\mathbf{x}}})$$

or equivalently

$$\underset{\mathbf{k}_i}{\text{minimize}} \quad E(|\tilde{x}_i|^2) = E(|x_i - \mathbf{k}_i^H \mathbf{y}|^2)$$

- We denote the cost function

$$\begin{aligned} J(\mathbf{k}_i) &= E(|x_i|^2) - E(x_i \mathbf{y}^H)\mathbf{k}_i - \mathbf{k}_i^H E(\mathbf{y} x_i^H) + \mathbf{k}_i^H E(\mathbf{y}\mathbf{y}^{\mathrm{H}})\mathbf{k}_i \\ &= \sigma_{x,i}^2 - \mathbf{R}_{xy,i}\mathbf{k}_i - \mathbf{k}_i^H \mathbf{R}_{yx,i} + \mathbf{k}_i^H \mathbf{R}_y \mathbf{k}_i \end{aligned}$$

- Setting the gradient vector $J(\mathbf{k}_i)$ with respect to $\mathbf{k}_i$ to zero, we get

$$\mathbf{k}_i^H \mathbf{R}_y = \mathbf{R}_{xy,i}, \quad i = 0, 1, \ldots, p-1.$$

or the solution matrix should satisfy

$$\mathbf{K}\mathbf{R}_y = \mathbf{R}_{xy}$$

Normal equations

$$\mathbf{K}\mathbf{R}_y = \mathbf{R}_{xy}$$

- For a unique solution, $\mathbf{R}_y > \mathbf{0}$, so that

$$\mathbf{K} = \mathbf{R}_{xy}\mathbf{R}_y^{-1}$$

- Satisfies orthogonality criterion

$$\mathbf{k}_i^H \mathbf{R}_y = \mathbf{R}_{xy,i} \Rightarrow \mathbf{k}_i^H E(\mathbf{y}\mathbf{y}^H) = E(x_i \mathbf{y}^H) \Rightarrow E[(x_i - \mathbf{k}_i^H \mathbf{y})\mathbf{y}^H] = 0$$

- For the non-zero mean case, the solution is obtained by replacing $\mathbf{x}$ and $\mathbf{y}$ with centered variables $\mathbf{x} - \bar{\mathbf{x}}$ and $\mathbf{y} - \bar{\mathbf{y}}$

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{K}(\mathbf{y} - \bar{\mathbf{y}})$$