E9 211: Adaptive Signal Processing

Lecture 8: Linear models

Outline

1. Linear models (Ch. 5)
2. Constrained estimation (Ch. 6)

Linear model

- Suppose $\mathbf{x} : p \times 1$ and $\mathbf{y} : q \times 1$ are vector valued and we $\mathbf{x}$ and $\mathbf{y}$ are linearly related as
$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$$
with $\mathbf{H} : q \times p$ and $\mathbf{v} : q \times 1$ is the noise vector.

- Let $\mathbf{R}_x = E(\mathbf{x}\mathbf{x}^{\mathrm{H}})$ and $\mathbf{R}_v = E(\mathbf{v}\mathbf{v}^{\mathrm{H}})$. Assume $E(\mathbf{x}\mathbf{v}^{\mathrm{H}}) = \mathbf{0}$.

- We have
$$\mathbf{R}_y = \mathbf{H}\mathbf{R}_x\mathbf{H}^{\mathrm{H}} + \mathbf{R}_v \quad \text{and} \quad \mathbf{R}_{xy} = \mathbf{R}_x\mathbf{H}^{\mathrm{H}}$$

- Then, the linear least-mean-squares error estimator (or simply, linear minimum-mean-squared error estimator)
$$\text{lmmse:} \quad \hat{\mathbf{x}} = \mathbf{R}_{xy}\mathbf{R}_y^{-1}\mathbf{y} = \mathbf{R}_x\mathbf{H}^{\mathrm{H}}[\mathbf{H}\mathbf{R}_x\mathbf{H}^{\mathrm{H}} + \mathbf{R}_v]^{-1}\mathbf{y}$$

LMMSE - matrix inversion lemma

- Using the matrix inversion lemma

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$$

- The linear least-mean-squares error estimator can be simplified to

$$\text{lmmse:} \quad \hat{\mathbf{x}} = [\mathbf{R}_x^{-1} + \mathbf{H}^{\mathrm{H}}\mathbf{R}_v^{-1}\mathbf{H}]^{-1}\mathbf{R}_v^{-1}\mathbf{Hy}$$

  For $p=1$, $[\mathbf{R}_x^{-1} + \mathbf{H}^{\mathrm{H}}\mathbf{R}_v^{-1}\mathbf{H}]^{-1}$ will be a scalar, whereas $\mathbf{R}_x\mathbf{H}^{\mathrm{H}}[\mathbf{HR}_x\mathbf{H}^{\mathrm{H}} + \mathbf{R}_v]^{-1}$ will be a matrix.

- The minimum error is given by $[\mathbf{R}_x^{-1} + \mathbf{H}^{\mathrm{H}}\mathbf{R}_v^{-1}\mathbf{H}]^{-1}$

Constrained estimation

- Suppose $\mathbf{x}$ is deterministic/constant, and let us consider $P = 1$, i.e., x is a scalar. Then the linear model will simply to

$$\mathbf{y} = \mathbf{h}x + \mathbf{v}$$

- Assume that $E(\mathbf{v}) = \mathbf{0}$ and $E(\mathbf{v}\mathbf{v}^{\mathrm{H}}) = \mathbf{R}_v$.

- In this case, we constrain the estimator to have the form $\hat{x} = \mathbf{w}^H \mathbf{y}$ and find $\mathbf{w}$ such that the estimator is unbiased and has least-mean-squares error.

- This estimator is referred to as the *Best Linear Unbiased Estimator* (BLUE).

- For the estimator to beunbiased, we require

$$E(\hat{x}) = x \Leftrightarrow E(\mathbf{w}^H \mathbf{y}) = \mathbf{w}^{\mathrm{H}} \mathbf{h} E(x) \Rightarrow \mathbf{w}^{\mathrm{H}} \mathbf{h} = 1.$$

BLUE

- The least-mean-squares error is

$$E(\tilde{x}^2) = E(x - \hat{x})^2 = E(x - \mathbf{w}^{\mathrm{H}}\mathbf{y})^2 = \mathbf{w}^{\mathrm{H}}\mathbf{R}_v\mathbf{w} = E(\hat{x}^2)$$

  where the last equality is obtained by using $\mathbf{w}^{\mathrm{H}}\mathbf{h} = 1$.

- To find $\mathbf{w}$, we solve the constrained optimization problem

$$\underset{\mathbf{w}}{\text{minimize}} \quad \mathbf{w}^{\mathrm{H}}\mathbf{R}_v\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^{\mathrm{H}}\mathbf{h} = 1$$

  whose solution is

$$\mathbf{w}_{\mathrm{opt}} = \mathbf{R}_v^{-1}\mathbf{h}(\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1} \quad \Rightarrow \; \hat{x} = (\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1}\mathbf{h}^{\mathrm{H}}\mathbf{R}_v^{-1}\mathbf{y}$$

BLUE

- Using the method of the Lagrange multipliers, we should optimize the function

$$J(\mathbf{w}, \lambda) = \mathbf{w}^{\mathrm{H}}\mathbf{R}_v\mathbf{w} + \lambda(\mathbf{w}^{\mathrm{H}}\mathbf{h} - 1)$$

- Setting the gradient with respect to $\mathbf{w}$ to zero we get

$$\frac{\partial J}{\partial \mathbf{w}} = 2\mathbf{R}_v\mathbf{w} + \lambda\mathbf{h} = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{w} = -\frac{1}{2}\mathbf{R}_v^{-1}\mathbf{h}\lambda$$

- The Lagrange multiplier λ is obtained by the constraint

$$\mathbf{w}^T\mathbf{h} = -\frac{1}{2}\mathbf{h}^T R_v^{-1}\mathbf{h}\lambda = 1 \qquad \Rightarrow \qquad \lambda = -2(\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1}$$

- Substituting $\lambda = -2(\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1}$ in $\mathbf{w} = -\frac{1}{2}\mathbf{R}_v^{-1}\mathbf{h}\lambda$ we get

$$\mathbf{w}_{\mathrm{opt}} = \mathbf{R}_v^{-1}\mathbf{h}(\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1} \quad \Rightarrow \; \hat{x} = (\mathbf{h}^T\mathbf{R}_v^{-1}\mathbf{h})^{-1}\mathbf{h}^{\mathrm{H}}\mathbf{R}_v^{-1}\mathbf{y}$$