Minimum Mean-Square Error: Discrete-time

M. Sami Fadali
Professor of Electrical Engineering
University of Nevada

Outline

- Minimum mean square error.
- Orthogonality principle.
- Recursive versus batch computation.

Discrete-time Problem

- Discrete random signal \( x_i, i = 1, 2, \ldots, n, E\{x_i\} = 0 \)
  \[
x = [x_1 \ x_2 \ \ldots \ x_n]^T
\]
- Noisy zero-mean measurements
  \[
z = [z_1 \ z_2 \ \ldots \ z_n]^T = x + v
\]
- Estimate of \( x \): output \( \hat{x} \) of an FIR filter at time \( n \)
- \( \hat{x} \) = linear combination of \( \hat{x} = Kz \)
- Estimation error \( e = x - \hat{x} = x - Kz \)

Estimation Error

- Error: use \( tr[AB] = tr[BA] \)
  \[
  \|e\|^2 = \|x - \hat{x}\|^2 = (x - Kz)^T (x - Kz)
  = x^T x - 2x^T Kz + z^T K^T Kz
  = tr[xx^T] - 2 tr[z x^T K] + tr[Kzz^T K^T]
\]
- Mean-square error: use \( E\{tr[A]\} = tr[E\{A]\} \)
  \[
  E\{\|e\|^2\} = E\{tr[xx^T] - 2 z x^T K + tr[Kzz^T K^T]\}
  = tr[C_{xx}] - 2 tr[C_{xz} K] + tr[K C_{zz} K^T]
\]
Minimum Mean-square

- Mean-square error
  \[ E\{\|e\|^2\} = tr[C_{xx}] - 2tr[C_{xz}K] + tr[KC_{zz}K^T] \]
- Minimize mean-square error to find \( K \)
- Trace formulas
  \[ \frac{\partial tr[AB]}{\partial A} = B^T, A, B \text{ square} \]
  \[ \frac{\partial tr[ACA^T]}{\partial A} = 2AC, C \text{ symmetric} \]
  \[ \frac{\partial E\{\|e\|^2\}}{\partial K} = -2C_{xx} + 2KC_{zz} = [0] \]

Optimal Filter

- Necessary condition
  \[ -2C_{xx} + 2KC_{zz} = [0] \]
  \[ K = C_{xx}^{-1}C_{zz}^{-1} \]
- Sufficient condition: positive definite autocorrelation
  \[ \frac{\partial \{-2c_{xx,i} + 2k_i^T C_{zz}\}}{\partial k_i} = 2C_{zz} > 0, \quad K = \begin{bmatrix} k_1^T \\ \vdots \\ k_n^T \end{bmatrix} \]

Orthogonality

- \( e = x - Kz \)
- Optimality Condition: Error orthogonal to all components of the measurement vector.
  \[ E\{(x - Kz)z^T\} = C_{xz} - KC_{zz} = [0] \]
  \[ K = C_{xz}C_{zz}^{-1} \]

Non-Gaussian Case

- Minimizing mean square error can give bad results.
- Only uses covariance functions (auto/cross): optimum among linear filters.
- MMSE suitable for Gaussian signals.
- Optimum filter for Gaussian is a linear filter.
- Optimum filter for other cases may be nonlinear.
Batch Versus Recursive Computation

- So far, assumed all data available at time of computation.
- On-line computation: data measured at each sampling point \( \{z_1, z_2, z_3, \ldots \} \)
- Batch Computation: Must recalculate if a new data point becomes available.
- Recursive Computation: Update the calculated quantity without recalculation.
- Recursive computation is more efficient.

Mean Computation: Batch

- Initialization \( i \leftarrow 1, z \leftarrow z_1, \hat{m}_i \leftarrow z \)
- While \( i < i_{\text{final}} \) do
  - Increment and read \( i \leftarrow i + 1, z \leftarrow z_i \)
  - Calculate estimate of mean at time \( i \)
    \[ \hat{m}_i \leftarrow \frac{1}{i} \sum_{j=1}^{i} z_j \]
- End
  Recalculate when new data arrives

Mean Computation: Recursive

- Initialization \( i \leftarrow 1, z \leftarrow z_1, \hat{m}_i \leftarrow z \)
- While \( i < i_{\text{final}} \) do
  - Increment and read \( i \leftarrow i + 1, z \leftarrow z_i \)
  - Calculate estimate of mean at time \( i \)
    \[ \hat{m}_i \leftarrow \frac{(i - 1)\hat{m}_{i-1} + z_i}{i} \]
- End

Advantages of Recursive Estimation

- Advantages
  - Only last measurement \( z_i \) stored.
  - More efficient computation.