EE 602 Term Paper on Best Linear Unbiased Estimator (BLUE)

Akhilesh Pathak (Y5044)
Ankur Saxena (Y5088)

Introduction

- The MVU estimator cannot be easily found
  - If PDF is not known, CRLB and theory of sufficient statistics cannot be applied
  - If PDF is known, it doesn't make sure the minimum variance
- A suboptimal estimator approach
  - Restrict the estimator to linear that is unbiased
  - It should have minimum variance
- Best Linear Unbiased Estimator (BLUE)
  - Only first and second moments of PDF are required
  - Useful in practical implementation as the complete PDF is not required

BLUE – An overview

- We construct the BLUE in the following way
  - Define a linear estimator
  - Impose the condition to be unbiased
  - Get the minimum variance of the estimator
  - Find the conditions for minimum variance
  - We can generalize it to vector parameter

- In both the cases, scalar parameter or vector parameter, if the data is Gaussian, then BLUE is also MVU estimator

Definition of BLUE

- Given the data set { x[0], x[1], ….., x[N-1] } with PDF p(x; θ) where θ is an unknown parameter

$$\hat{\theta} = \sum_{n=0}^{n=N-1} a_n x[n]$$

where a_n's are constants to be determined.

- For example consider white noise x[n]=w. Let's define the estimator for variance as

$$\hat{\sigma}^2 = \sum_{n=0}^{n=N-1} a_n x[n]$$

The expected value for the estimator is

$$E(\hat{\sigma}^2) = \sum_{n=0}^{n=N-1} a_n E(x[n]) = 0$$

Hence we cannot find any linear estimator which is unbiased.

The solution to the problem could be that we can use transformed data and can produce a linear estimator with that. Let's suppose $y[n] = x^2[n]$.

$$E(\hat{\sigma}^2) = \sum_{n=0}^{n=N-1} a_n E(y[n]) = \sum_{n=0}^{n=N-1} a_n E(x^2[n])$$

Expected value

$$E(\hat{\sigma}^2) = \sum_{n=0}^{n=N-1} a_n \sigma^2 = \sigma^2$$

Finding the BLUE

- To determine BLUE we constrain **θ** to be linear and unbiased and then find the a_n's to minimize the variance. The unbiasedness constraint is:

$$E(\hat{\theta}) = \sum_{n=0}^{n=N-1} a_n E(x[n]) \qquad (1)$$

- The variance is:

$$Var(\boldsymbol{\theta}) = E[(\sum_{n=0}^{N-1} a_n x[n] - E(\sum_{n=0}^{N-1} a_n x[n]))^2] \qquad (2)$$

- By using (1) and letting $\mathbf{a} = [a_0\ a_1 \ldots a_{N-1}]^T$ we have:

$$Var(\boldsymbol{\theta}) = E[(\boldsymbol{a}^T\boldsymbol{x} - \boldsymbol{a}^T E(\boldsymbol{x}))^2]$$

$$= E[(\boldsymbol{a}^T(\boldsymbol{x} - E(\boldsymbol{x}))^2]$$

$$= E[a^T(x - E(x))(x - E(x))^T a]$$
$$= a^T C a \quad (3)$$

- But we need to assume some form for E(x[n]). In order to satisfy unbiasedness constraint E(x[n]) must be linear in θ or:

$$E(x[n]) = s[n]\theta \quad (4)$$

where the s[n]'s are known.

- This assumption means that ***BLUE is applicable to amplitude estimation of known signals in noise.***
- Combining (2) and (4) we get:

$$\sum_{n=0}^{N-1} a_n E(x|n) = \theta$$

$$\sum_{n=0}^{N-1} a_n s[n]\theta = \theta$$

$$=> a^T s = 1$$

- The solution to this optimization problem is:

$$a_{opt} = \frac{C^{-1}s}{s^T C^{-1} s}$$

- So, the BLUE is:

$$\theta = \frac{s^T C^{-1} x}{s^T C^{-1} s} \quad (5)$$

- Its minimum variance is:

$$Var(\theta) = \frac{1}{s^T C^{-1} s} \quad (6)$$

- As we asserted earlier, to determine the BLUE we only need information about

  1. **s** , the scaled mean

  2. **C**, the covariance

Or the first two moments but ***not the entire PDF***.

- Examples

Extension to Vector Parameter

- If the parameter to be estimated is a p×1 parameter, then for the estimator to be linear in the data we require

$$\theta_i = \sum_{n=0}^{N-1} a_{in} x[n] \qquad i=1, 2,\ldots,p \qquad (7)$$

- In matrix form this is:

$$\boldsymbol{\theta} = \boldsymbol{A}\boldsymbol{x}$$

where **A** is a p×N matrix.

- In order for the **θ** to be unbiased we have:

$$E(\theta_i) = \sum_{n=0}^{N-1} a_{in} E(x[n]) = \theta_i \qquad (8)$$

- In matrix form it is:

$$E(\boldsymbol{\theta}) = \boldsymbol{A}E(\boldsymbol{x}) = \boldsymbol{\theta} \qquad (9)$$

- From (9) we must have that

$$E(x) = \boldsymbol{H}.\boldsymbol{\theta} \tag{10}$$

for **H** a known N×p matrix, of the form:

$$E(x) = \begin{bmatrix} s[0] \\ s[1] \\ \vdots \\ s[N-1] \end{bmatrix}.\boldsymbol{\theta}$$

- Hence, we have the unbiasedness constraint as:

$$\boldsymbol{A}.\boldsymbol{H} = \boldsymbol{I} \tag{11}$$

- If we define $\mathbf{a_i} = [a_{i0}\ a_{i1} \ldots\ a_{i,(N-1)}]^T$, so that $\theta_i = \mathbf{a_i}^T \mathbf{x}$, the unbiasedness constraint may be rewritten for each $\mathbf{a_i}$ by noting that

$$\boldsymbol{A} = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_{p-1}^T \\ a_p^T \end{pmatrix}$$

- And letting denote h_i the i^{th} column of **H** , so that

$$\mathbf{H} = [h_1\ h_2\ \dots\ h_p]$$

- With these definitions the unbiasedness constraint is:

$$a_i^T . h_j = \delta_{ij} \quad i = 1,2, \dots, p; j = 1, 2, \dots, p$$

- And the variance is:

$$Var(\theta_i) = a_i^T C a_i$$

- The minimization of this variance subject to the unbiasedness constraints yields the BLUE as:

$$\ddot{\theta} = (H^T C^{-1} H)^{-1} H^T C^{-1} x$$

- And the covariance matrix is

$$C_\theta = (H^T C^{-1} H)^{-1}$$

- This form of BLUE is similar to the MVU estimator for the general linear model
- To summarize our presentation we now state the general BLUE for a vector parameter of a general linear model.

Gauss-Markov Theorem

If the data are of the general linear model form

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

where **H** is a known N×p matrix, **θ** is p×1 vector of parameters to be estimated and **w** is N×1 noise vector with zero mean and covariance **C**, then the BLUE of **θ** is

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

and the minimum variance of θ_i is

$$var(\hat{\theta}_i) = [(H^T C^{-1} H)^{-1}]_{ii}$$

The covariance matrix is

$$C_{\hat{\theta}} = (H^T C^{-1} H)^{-1}$$