EE 602 Term Paper on

Best Linear Unbiased Estimator (BLUE)

Submitted by:

Akhilesh Pathak (Y5044)

Ankur Saxena (Y5088)

Introduction

It frequently occurs that the MVU estimator, even if it exists, cannot be found. For example, if PDF is not known, CRLB and theory of sufficient statistics cannot be applied. Also, if PDF is known, it doesn't make ensure minimum variance.

In such cases, we have to resort to a suboptimal estimator approach. We can restrict the estimator to a linear form that is unbiased. It should also have minimum variance.

An example of this approach is the Best Linear Unbiased Estimator (BLUE) approach. In this case:

- Only first and second moments of PDF are required.
- It is useful in practical implementation as the complete PDF is not required.

BLUE Approach- 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
- Generalize it to vector parameter

Note that 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 $y[n] = x^2[n]$.

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

The expected value is:

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

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:

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

The variance is:

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

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

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

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

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

$$= \boldsymbol{a}^T\boldsymbol{C}\boldsymbol{a} \quad \textbf{(3)}$$

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:

$$\boldsymbol{E(x[n]) = s[n]\theta} \quad \textbf{(4)}$$

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:

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

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

$$=> \boldsymbol{a}^T \boldsymbol{s} = \mathbf{1}$$

The solution to this optimization problem is:

$$\boldsymbol{a}_{opt} = \frac{\boldsymbol{C}^{-1}\boldsymbol{s}}{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}}$$

So, the BLUE is:

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

Its minimum variance is:

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

As w 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*.

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:

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

In matrix form this is:

$$\theta = Ax$$

where A is a p×N matrix.

In order for the θ to be unbiased we have:

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

In matrix form it is:

$$E(\theta) = AE(x) = \theta \qquad \textbf{(9)}$$

From (9) we must have that

$$E(x) = H.\theta \qquad \textbf{(10)}$$

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

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

Hence, we have the unbiasedness constraint as:

$$A.H = I \qquad \textbf{(11)}$$

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

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

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

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

With these definitions the unbiasedness constraint is:

$$\boldsymbol{a}_i^T . \boldsymbol{h}_j = \delta_{ij} \quad i = 1,2, \dots, p; j = 1, 2, \dots, p$$

And the variance is:

$$Var(\boldsymbol{\theta}_i) = \boldsymbol{a}_i^T \boldsymbol{C} \boldsymbol{a}_i$$

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

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

And the covariance matrix is:

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

This form of BLUE is similar to the MVU estimator for the general linear model.

To summarize this work, 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:

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

References

1. "Fundamentals Of Statistical Signal Processing: Estimation Theory", Steven M Kay, Prentice Hall.
2. Gauss Markov Theorem article on Wikipedia (http://en.wikipedia.org/wiki/Gauss–Markov_theorem)