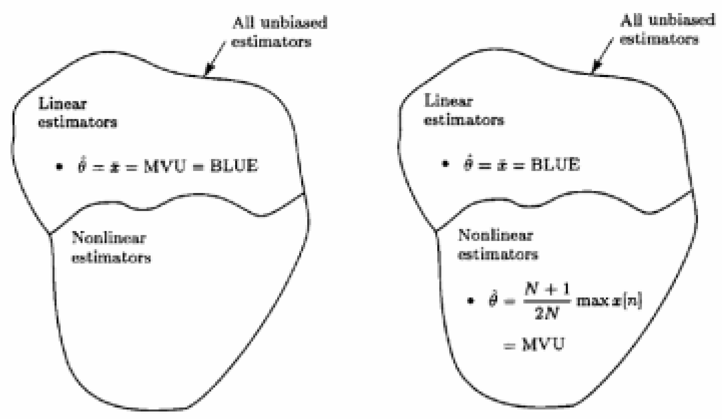
Best Linear Unbiased Estimator (BLUE)

1. What is BLUE?

BLUE : minimum variance를 가지는 unbiased, linear estimator

linear이기 때문에 데이터들의 선형 결합으로 표현 가능하다.

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



(a) DC level in WGN; BLUE is optimal $\hat{\theta} = \bar{x} = \sum_{n=0}^{N-1} \frac{1}{N} x[n]$

(b) Mean of uniform noise; BLUE is suboptimal

위 그림과 같이 선형 MVU의 경우에는 BLUE는 MVU와 동일하게 되지만 그렇지 않은 경우에는 BLUE는 다른 값을 가지게 된다.

2. How to find BLUE

우선 unbiased되어 있기 때문에,

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

를 만족한다.

그리고 분산은 $a = [a_0, a_1, \ldots, a_{N-1}]^T$라고 한다면, 정의에 따라

$$\begin{aligned} var(\hat{\theta}) &= E[(a^T x - a^T E(x))^2] \\ &= E[(a^T (x - E(x))^2] \\ &= E[a^T (x - E(x))(x - E(x))^T a] \\ &= a^T C a \end{aligned}$$

로 전개된다.

BLUE를 위한 최적의 벡터 a는 위 분산을 최소화해야 할 것이다.

이 때, 위 기댓값 식을 만족시키기 위해 $E(x[n])$을 θ에 대한 선형 결합으로 가정한다.

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

따라서 $x[n]$은

$$\begin{aligned} x[n] &= E(x[n]) + (x[n] - E(x[n])) \\ &= \theta s[n] + w[n] \end{aligned}$$

와 같이 표현가능하다.

다시 돌아와서, 맨 위의 기댓값 식을 전개해보면,

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

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

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

$$a^T s = 1$$

이 되고, BLUE는 위 조건을 만족하며 분산을 최소화시키는 최적화문제가 된다.

따라서 BLUE를 찾는 문제는 다음과 같이 정리된다.

$$BLUE: \arg\min(var(\hat{\theta}) = a^T Ca) \ \ subject\ to\ a^T s = 1$$

위 최적화 문제는 다음과 같이 풀 수 있다.
우선

$$J = a^T Ca + \lambda(a^T s - 1)$$

을 정의한다. $a^T s = 1$이므로 J가 최소가 되면, $a^T Ca$도 최소이다.

J의 최소를 알기 위해 a에 대해 미분하면

$$\frac{\partial J}{\partial a} = 2Ca + \lambda s = 0$$

$$a = -\frac{\lambda}{2} C^{-1} s$$

다음과 같이 정리된다. 그런데 $a^T s = 1$이므로

$$a^T s = -\frac{\lambda}{2} s^T C^{-1} s = 1$$

$$-\frac{\lambda}{2} = \frac{1}{s^T C^{-1} s}$$

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

이 된다.

따라서 구하고자 하는 최소 분산은

$$\begin{aligned} var(\hat{\theta}) &= a_{opt}^T C a_{opt} \\ &= \frac{s^T C^{-1} C C^{-1} s}{(s^T C^{-1} s)^2} \\ &= \frac{1}{s^T C^{-1} s} \end{aligned}$$

이고, BLUE는

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

이 된다.

Ex) $x[n] = A + w[n]$일 때, $(w[n] \sim N(0, \sigma^2))$

$$E(x[n]) = s[n]A = A$$

이므로 $s[n] = 1, s = 1$이다.

따라서 BLUE는

$$\hat{A} = \frac{1^T \frac{1}{\sigma^2} I x}{1^T \frac{1}{\sigma^2} I 1} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x}$$

이고 분산은

$$var(\hat{A}) = \frac{1}{1^T \frac{1}{\sigma^2} 1} = \frac{\sigma^2}{N}$$

이다.

3. Vector Parameter

추정하고자 하는 파라미터가 스칼라가 아닌 벡터일 경우, 각각의 추정값은 다음과 같다.

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] \quad i = 1, 2, ..., p$$

$$\hat{\theta} = Ax$$

$\hat{\theta}$는 unbiased이기 때문에 다음과 같다.

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

$$E(\hat{\theta}) = AE(x) = \theta$$

여기서 앞서 스칼라에 대해 진행했던 것과 동일하게 $E(x)$는 θ에 대해 선형이다.

$$E(x) = H\theta$$

$$H = \begin{bmatrix} s[0] \\ s[1] \\ \dots \\ s[N-1] \end{bmatrix} = [h_1\, h_2 \dots h_p]$$

이를 위 기댓값 식에 대입하면,

$$AH = I$$

가 된다. 따라서

$$a_i^T h_j = \delta_{ij}$$

를 만족한다.

분산은

$$var(\hat{\theta}_i) = a_i^T C a_i$$

로 쓸 수 있으므로 BLUE는 다음 최적화 문제와 같다.

$$\text{BLUE : } \arg\min(var(\hat{\theta}_i) = a_i^T C a_i) \quad subject\ to\ a_i^T h_j = \delta_{ij}$$

그리고 위 최적화 문제를 만족하는 BLUE와 최소 분산은 다음과 같다.

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

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

*Thoerem 6.1 (Gauss-Markov Theorem)

Theorem 6.1 (Gauss-Markov Theorem) *If the data are of the general linear model form*

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \tag{6.18}$$

where $\mathbf{H}$ is a known $N \times p$ matrix, $\boldsymbol{\theta}$ is a $p \times 1$ vector of parameters to be estimated, and $\mathbf{w}$ is a $N \times 1$ noise vector with zero mean and covariance $\mathbf{C}$ (the PDF of $\mathbf{w}$ is otherwise arbitrary), then the BLUE of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x} \tag{6.19}$$

and the minimum variance of $\hat{\theta}_i$ is

$$\mathrm{var}(\hat{\theta}_i) = \left[(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\right]_{ii}. \tag{6.20}$$

In addition, the covariance matrix of $\hat{\boldsymbol{\theta}}$ is

$$\mathbf{C}_{\hat{\theta}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}. \tag{6.21}$$