Best Linear Unbiased estimator

BLUE: Min variance unbiased estimator among the class of estimators of the form:

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] \qquad \text{Data set}: \{x[0], \cdots, x[N-1]\}$$

$p(x; \theta)$

$a_n \rightarrow$ constants to be optimized.

Sometimes MVUE is linear: $x[n] = A + w[n]$ → estimate; $w[n] \hookrightarrow N(0, \sigma^2)$

MVUE: $\hat{\theta} = \bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \rightarrow$ linear.

↳ So BLUE $\equiv$ MVUE

Other cases: MVUE is non-linear: uniform $[0, \beta]$

MVUE : $\hat{\theta} = \frac{N+1}{2N} \max x[n]$

So $\mathrm{Var}(\text{MVUE}) < \mathrm{Var}(\text{BLUE})$

In some cases, a non-trivial linear unbiased estimator might not even exist.:

Unknown variance of WGN: MVUE $\hat{\sigma}^2 = \frac{1}{N} \sum x^2[n]$

If in this case we force $\hat{\sigma}^2 = \sum a_n x[n]$ then

$E\hat{\sigma}^2 = 0 \rightarrow$ So no unbiased linear estimator.

(Q. Can this be fixed using transformation of data)

Computing BLUE: $\hat{\theta} = \sum a_n x[n] \Rightarrow E\hat{\theta} = \sum a_n E(x[n]) = \theta$

$$\mathrm{Var}(\hat{\theta}) = E\left[\left(\sum_{n=0}^{N-1} a_n x[n] - E\left(\sum a_n x[n]\right)\right)^2\right]$$

Let $a = [a_0 \; a_1 \; \cdots \; a_{N-1}]^T$

$$\mathrm{Var}(\hat\theta) = E\left[\left(a^T x - a^T E(x)\right)^2\right]$$
$$= E\left[a^T (x - E(x))(x - E(x))^T a\right] = a^T \underbrace{C}_{\mathrm{Cov}(x)} a \quad \text{— (1)}$$

Let $E(x[n]) = s[n]\,\theta$ where $s[n]$ are known (actually this is enough + C full pdf not required)

$\Rightarrow \sum a_n s[n]\, \theta = \theta$

$\Leftrightarrow \sum a_n s[n] = 1$

$\Leftrightarrow \boxed{a^T s = 1}$ — (2) $\quad s = [s[0] \; \cdots \; s[N-1]]^T$

P: Min (1) s.t. (2)

$$H = a^T C a + \lambda (a^T s - 1)$$

$$\frac{\partial H}{\partial a} = 2Ca + \lambda s = 0 \Rightarrow a = -\frac{\lambda}{2} C^{-1} s$$

$$\frac{\partial H}{\partial \lambda} = a^T s - 1 = 0 \Leftrightarrow -\frac{\lambda}{2} s^T C^{-1} s = 1 \Leftrightarrow -\frac{\lambda}{2} = \frac{1}{s^T C^{-1} s}$$

$$\boxed{a_{opt} = \frac{C^{-1} s}{s^T C^{-1} s}} \quad = \quad \mathrm{Var}(\hat\theta) = \frac{1}{s^T C^{-1} s} \longrightarrow \text{Exercise}$$

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

Moreover $$E\hat\theta = \frac{s^T C^{-1} E x}{s^T C^{-1} s} = \frac{s^T C^{-1} s\, \theta}{s^T C^{-1} s} = \theta$$

Ex: $x[n] = A + w[n] \quad n = 0, 1, \ldots, N-1$ $\quad$ $w[n]$ uncorrelated

$\hookrightarrow E(w[n]) = 0, \; \mathrm{Var}(w[n]) = \sigma^2$

[pdf unknown]

$E x[n] = A$ so $\quad$ $s[n] = 1 \;\; \forall n$

$$s = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} =: \mathbb{1}$$

BLUE: $$\hat A = \frac{\mathbb{1}^T \frac{1}{\sigma^2} I x}{\mathbb{1}^T \frac{1}{\sigma^2} I \mathbb{1}}$$

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

(We know here BLUE = MVUE)

Ex: $x[n] = A + w[n]$ $\quad$ $\left.\begin{array}{l}\operatorname{Var}[w[n]] = \sigma_n^2 \\ E[w[n]] = 0\end{array}\right]$ $w[n]$ is uncorrelated for diff. n.

$$\hat{A} = \frac{\mathbb{1}^T C^{-1} x}{\mathbb{1}^T C^{-1}\mathbb{1}}$$

$$\operatorname{Var}\hat{A} = \frac{1}{\mathbb{1} C^{-1}\mathbb{1}} \qquad C = \begin{bmatrix} \sigma_0^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{N-1}^2 \end{bmatrix}$$

$$C^{-1} = \begin{bmatrix} \frac{1}{\sigma_0^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_{N-1}^2} \end{bmatrix} \qquad \hat{A} = \frac{\sum_{n=0}^{N-1} \frac{x[n]}{\sigma_n^2}}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}$$

$$\operatorname{Var}\hat{A} = \frac{1}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}$$

For vector case: $\theta = [\theta_1 \cdots \theta_p]^T$

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

i.e $\hat{\theta} = A x \qquad A \in \mathbb{R}^{p\times N}$

$$E\hat{\theta}_i = \sum a_{in}\, E(x[n]) = \theta_i \qquad i = 1, \ldots, p$$

i.e. $E\hat{\theta} = A E(x) = \theta \quad \forall \theta$ (1) $\qquad A = [\qquad]_{p\times N}$

A) Assume an linear model $x = H\theta + w$ i.e.

$Ex = H\theta$ —— (2) $\quad \forall\theta \quad H \in \mathbb{R}^{N\times p}$

[S. Kay seems to suggest that (1) $\Rightarrow$ (2). That is not strictly correct]

If (A1) is an explicit assumption, then
(1) & (2) implies $\boxed{AH = I}$

Gauss-Markov Thm: If the data are of the general linear model $x = H\theta + w$ where H is known $N \times p$ matrix, θ is $p \times 1$ vector of parameter to be estimated & w is $N \times 1$ noise vector with zero mean & cov. C. (any pdf is OK), then the BLUE of θ is

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

with $\text{Cov}(\hat{\theta}) = (H^T C^{-1} H)^{-1}$

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

In addition if $w \sim N(0, C)$ then $\hat{\theta}$ given above is also MVUE (also attains CRLB)

Proof: Will be done later from a geometric viewp

[Recall defn. of min variance is elementwise i.e. $\text{Var}(\hat{\theta}_i) \leq \text{Var}(\bar{\theta}_i)$ $\forall \bar{\theta}_i$

However the above thm guarantees more:

$$\underbrace{\text{Cov}(\bar{\theta}) - \text{Cov}(\hat{\theta}) \geq 0}_{\text{psd}} \quad \forall \bar{\theta}_{p \times 1}$$
]

Ex: $x[n] = A + w[n]$ $\qquad w[n] \sim N(0, C)$

$H = [1 \cdots 1]^T$

$$\hat{A} = (H^T C^{-1} H)^{-1} H^T C^{-1} x = \frac{\mathbb{1}^T C^{-1} x}{\mathbb{1}^T C^{-1} \mathbb{1}}$$

$$\text{Var } \hat{A} = (H^T C^{-1} H)^{-1} = \frac{1}{\mathbb{1}^T C^{-1} \mathbb{1}}$$

We can stop here, but it is interesting to factorize $C^{-1} = D^T D$, (Cholesky factorization)

Then,
$$\hat{A} = \frac{\mathbb{1}^T D^T \overbrace{D\, x}^{x'}}{\mathbb{1}^T D^T D\, \mathbb{1}} = \frac{(D\mathbb{1})^T x'}{\mathbb{1}^T D^T D\, \mathbb{1}} = \sum d_n\, \underbrace{x'[n]}_{\text{pre-whitening}}$$

where $d_n = [D\mathbb{1}]_n / \mathbb{1}^T D^T D\, \mathbb{1}$