

430.714

Estimation Theory

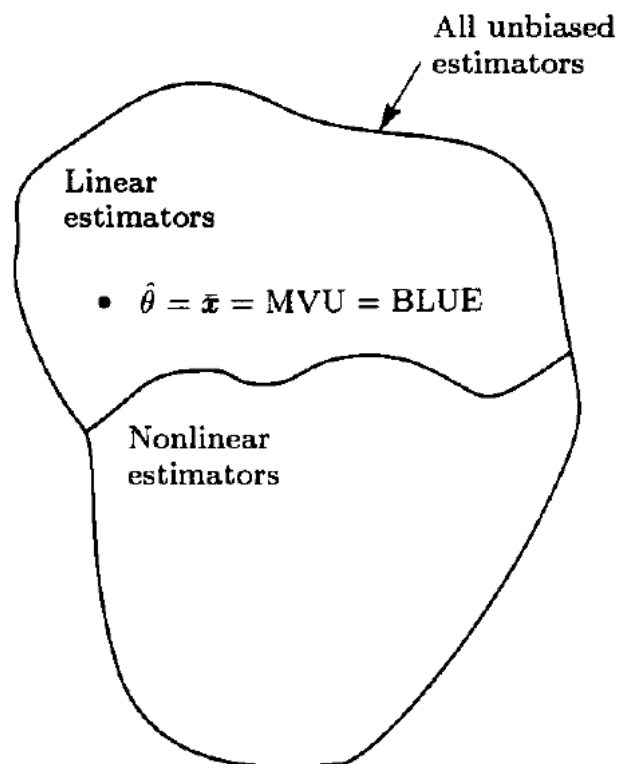
Prof. Songhwai Oh
ECE, SNU

Best Linear Unbiased Estimators (BLUE)

Best Linear Unbiased Estimator (BLUE)

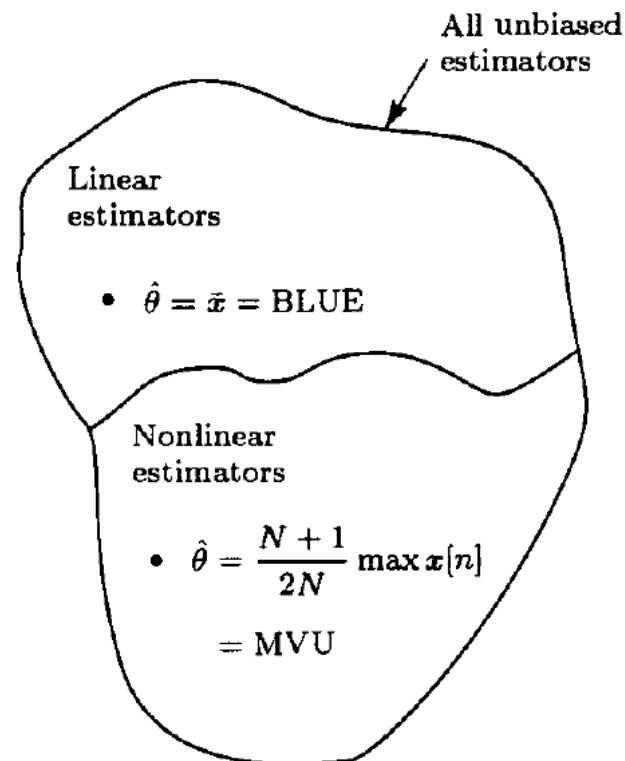
- The BLUE is an unbiased *linear* estimator with the minimum variance.

Linear estimator (linear in the data): $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n]$.



(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

BLUE: a scalar case $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] \quad \mathbf{a} = [a_0, a_1, \dots, a_{N-1}]^T$

Unbiased: $E(\hat{\theta}) = \sum_{n=0}^{N-1} a_n E(x[n]) = \theta \quad \text{var}(\hat{\theta}) = E[(\mathbf{a}^T \mathbf{x} - \mathbf{a}^T E(\mathbf{x}))^2]$

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

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

$$= \mathbf{a}^T \mathbf{C} \mathbf{a}.$$

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

$$x[n] = \theta s[n] + w[n]$$

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

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

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

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

$$w[n] = x[n] - \mathbb{E}(x[n])$$

BLUE: $\arg \min (\text{var}(\hat{\theta}) = \mathbf{a}^T \mathbf{C} \mathbf{a}) \quad \text{subject to } \mathbf{a}^T \mathbf{s} = 1.$

BLUE: a scalar case

$$\text{BLUE: } \arg \min \left(\text{var}(\hat{\theta}) = \mathbf{a}^T \mathbf{C} \mathbf{a} \right) \quad \text{subject to } \mathbf{a}^T \mathbf{s} = 1.$$

$$J = \mathbf{a}^T \mathbf{C} \mathbf{a} + \lambda(\mathbf{a}^T \mathbf{s} - 1)$$

$$\frac{\partial J}{\partial \mathbf{a}} = 2\mathbf{C}\mathbf{a} + \lambda\mathbf{s} \quad (\text{set to } 0) \Rightarrow \mathbf{a} = -\frac{\lambda}{2}\mathbf{C}^{-1}\mathbf{s} \quad \mathbf{a}^T \mathbf{s} = -\frac{\lambda}{2}\mathbf{s}^T \mathbf{C}^{-1}\mathbf{s} = 1$$
$$\mathbf{a}_{\text{opt}} = \frac{\mathbf{C}^{-1}\mathbf{s}}{\mathbf{s}^T \mathbf{C}^{-1}\mathbf{s}} \quad -\frac{\lambda}{2} = \frac{1}{\mathbf{s}^T \mathbf{C}^{-1}\mathbf{s}}$$

$$\begin{aligned} \text{var}(\hat{\theta}) &= \mathbf{a}_{\text{opt}}^T \mathbf{C} \mathbf{a}_{\text{opt}} \\ &= \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{C} \mathbf{C}^{-1} \mathbf{s}}{(\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s})^2} \\ &= \frac{1}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}. \end{aligned}$$

$$\text{BLUE: } \hat{\theta} = \mathbf{a}_{\text{opt}}^T \mathbf{x} = \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}$$

BLUE: a scalar case example

$$x[n] = A + w[n], \quad \text{where } w[n] \sim \mathcal{N}(0, \sigma_n^2)$$

$$E(x[n]) = s[n]\theta \quad \Rightarrow s = 1$$

$$\begin{aligned} \text{BLUE: } \hat{A} &= \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \\ \text{var}(\hat{A}) &= \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}. \end{aligned}$$

$$\mathbf{C} = \begin{bmatrix} \sigma_0^2 & 0 & \dots & 0 \\ 0 & \sigma_1^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{N-1}^2 \end{bmatrix}$$

If $\sigma^2 = \sigma_n^2$ for all n ,

$$\begin{aligned} \hat{A} &= \frac{\sum_{n=0}^{N-1} \frac{x[n]}{\sigma_n^2}}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}} \\ \text{var}(\hat{A}) &= \frac{1}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}. \end{aligned}$$

$$\begin{aligned} \hat{A} &= \frac{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{x}}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{1}} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x} \\ \text{var}(\hat{A}) &= \frac{1}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{1}} \\ &= \frac{\sigma^2}{N}. \end{aligned}$$

BLUE: a vector case

$$\hat{\boldsymbol{\theta}} = \mathbf{A}\mathbf{x} \quad \hat{\boldsymbol{\theta}} \in \mathbb{R}^p, \mathbf{A} \in \mathbb{R}^{p \times N}.$$

Unbiased: $E(\hat{\boldsymbol{\theta}}) = \mathbf{A}E(\mathbf{x}) = \boldsymbol{\theta}$

$$E(\mathbf{x}) = \mathbf{H}\boldsymbol{\theta} \quad \mathbf{H} \in \mathbb{R}^{N \times p}$$

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

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_p^T \end{bmatrix}$$

$$\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \dots \quad \mathbf{h}_p]$$

$$\mathbf{A}\mathbf{H} = \mathbf{I} \quad \Rightarrow \quad \mathbf{a}_i^T \mathbf{h}_j = \delta_{ij} \quad i = 1, 2, \dots, p; j = 1, 2, \dots, p$$

Appendix 6B

$$\text{BLUE: } \hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \mathbf{x}$$
$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}$$

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

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \quad (6.18)$$

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

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

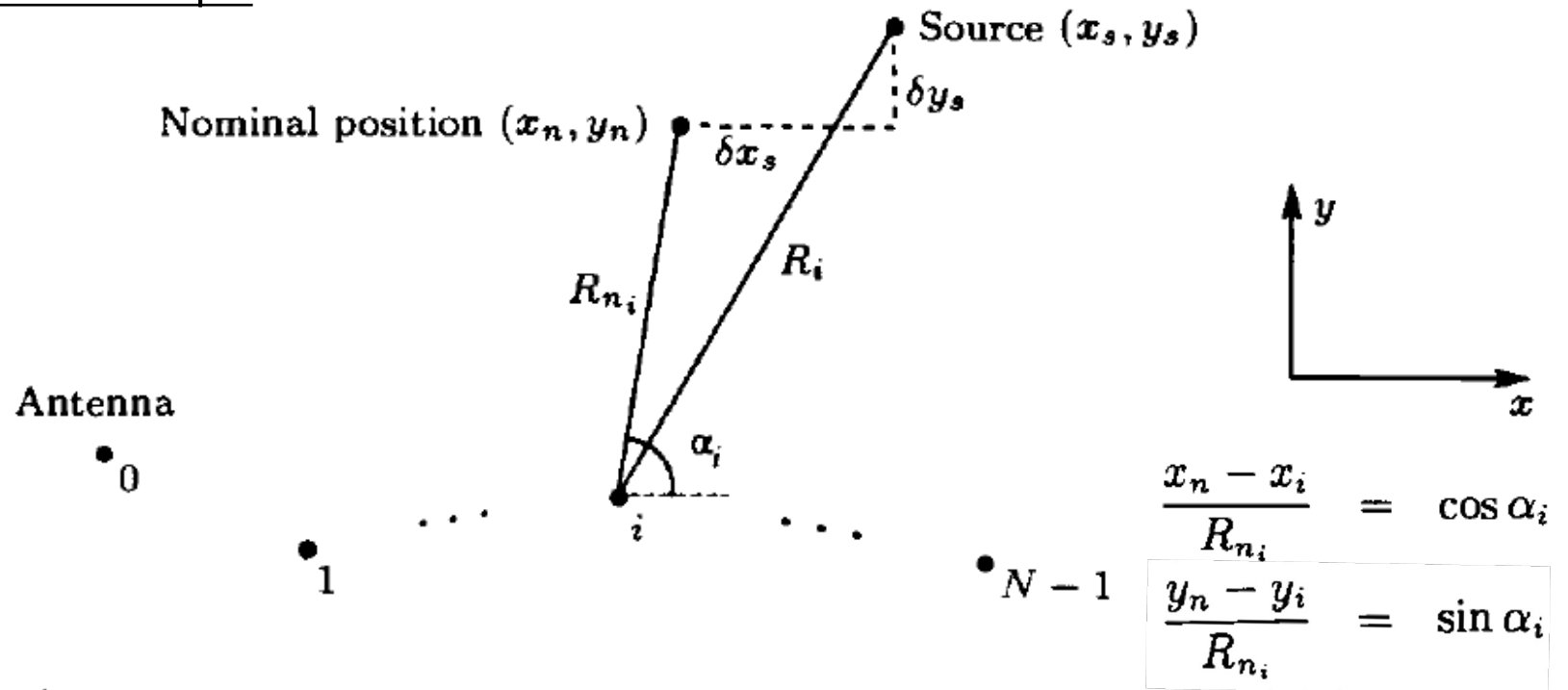
and the minimum variance of $\hat{\theta}_i$ is

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

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

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

Source Localization Example



$$t_i = T_0 + R_i/c + \epsilon_i \quad i = 0, 1, \dots, N-1$$

$$R_i = \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2} \quad R_i \approx R_{n_i} + \frac{x_n - x_i}{R_{n_i}} \delta x_s + \frac{y_n - y_i}{R_{n_i}} \delta y_s$$

$$t_i = T_0 + \frac{R_{n_i}}{c} + \frac{x_n - x_i}{R_{n_i} c} \delta x_s + \frac{y_n - y_i}{R_{n_i} c} \delta y_s + \epsilon_i \quad ; \text{1st order Taylor expansion}$$

$$t_i = T_0 + \frac{R_{n_i}}{c} + \frac{\cos \alpha_i}{c} \delta x_s + \frac{\sin \alpha_i}{c} \delta y_s + \epsilon_i$$

$$\tau_i = T_0 + \frac{\cos \alpha_i}{c} \delta x_s + \frac{\sin \alpha_i}{c} \delta y_s + \epsilon_i$$

ation Theory

$$\tau_i = t_i - \frac{R_{n_i}}{c}$$

Source Localization Example

$$\tau_i = T_0 + \frac{\cos \alpha_i}{c} \delta x_s + \frac{\sin \alpha_i}{c} \delta y_s + \epsilon_i$$

Unknowns: $(T_0, \delta x_s, \delta y_s)$

Time difference of arrival (TDOA):

$$\begin{aligned} \xi_1 &= \tau_1 - \tau_0 \\ \xi_2 &= \tau_2 - \tau_1 \\ &\vdots \\ \xi_{N-1} &= \tau_{N-1} - \tau_{N-2} \end{aligned}$$

$$\xi_i = \frac{1}{c} (\cos \alpha_i - \cos \alpha_{i-1}) \delta x_s + \frac{1}{c} (\sin \alpha_i - \sin \alpha_{i-1}) \delta y_s + \epsilon_i - \epsilon_{i-1}$$

$$\begin{aligned} \boldsymbol{\theta} &= [\delta x_s \quad \delta y_s]^T \\ \mathbf{H} &= \frac{1}{c} \begin{bmatrix} \cos \alpha_1 - \cos \alpha_0 & \sin \alpha_1 - \sin \alpha_0 \\ \cos \alpha_2 - \cos \alpha_1 & \sin \alpha_2 - \sin \alpha_1 \\ \vdots & \vdots \\ \cos \alpha_{N-1} - \cos \alpha_{N-2} & \sin \alpha_{N-1} - \sin \alpha_{N-2} \end{bmatrix} \\ \mathbf{w} &= \begin{bmatrix} \epsilon_1 - \epsilon_0 \\ \epsilon_2 - \epsilon_1 \\ \vdots \\ \epsilon_{N-1} - \epsilon_{N-2} \end{bmatrix} \end{aligned}$$

$$\mathbf{w} = \underbrace{\begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_{N-1} \end{bmatrix}}_{\boldsymbol{\epsilon}}$$

$$\mathbf{C} = E[\mathbf{A} \boldsymbol{\epsilon} \boldsymbol{\epsilon}^T \mathbf{A}^T] = \sigma^2 \mathbf{A} \mathbf{A}^T$$

$$\begin{aligned} \text{BLUE: } \hat{\boldsymbol{\theta}} &= (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \boldsymbol{\xi} \\ &= [\mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{H}]^{-1} \mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \boldsymbol{\xi} \end{aligned}$$

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = \sigma^2 [\mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{H}]^{-1}$$