Vector parameters, and CRLB contd.

Tuesday, 16 March 2021 13:53

Transformation of parameters:

$x[n] = A + w[n]$ $\qquad \alpha = g(A) = A^2$

$$\mathrm{var}(\hat{A}^2) \geq \frac{(2A)^2}{(N/\sigma^2)} = \frac{4A^2\sigma^2}{N}$$

$\bar{x}$ (sample mean); which is the MVU for A

$$= \frac{1}{N}\sum_{n=0}^{N-1} x[n]$$

Now, will performing $\bar{x}^2$ lead to an efficient estimator for A^2

$$\bar{x} \sim \mathcal{N}\left(A, \frac{\sigma^2}{N}\right)$$

$$E[\bar{x}^2] = E^2[\bar{x}] + \mathrm{var}[\bar{x}]$$

$$= A^2 + \frac{\sigma^2}{N} \neq A^2 \quad : \text{biased estimator}$$

"Efficiency is not preserved under non-linear transformation"

Suppose θ is the unknown parameter,

$\hat{\theta}$ is an efficient estimator

$$\alpha = g(\theta) = a\theta + b$$

$$\rightarrow \hat{\alpha} = g(\hat{\theta}) = a\hat{\theta} + b \quad \text{is this efficient?}$$

$$E[\hat{\alpha}] = a\,E[\hat{\theta}] + b$$

$$= a\theta + b = \alpha \quad \} \text{ unbiased}$$

$$\mathrm{var}(\hat{\alpha}) \geq \frac{a^2}{I(\theta)} = a^2\,\mathrm{var}(\hat{\theta})$$

$$\mathrm{var}(\hat{\alpha}) = \mathrm{var}(a\hat{\theta} + b)$$

$$= a^2\,\mathrm{var}(\hat{\theta})$$

CRLB is acheived.

"Efficiency is maintained under linear transformation"

$$\mathrm{var}(\bar{x}^2) = E[\bar{x}^4] - E^2[\bar{x}^2]$$

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

$$E[x^2] = \mu^2 + \sigma^2$$

$$E[x^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

$$\mathrm{Var}(x^2) = E[x^4] - E^2[x^2]$$

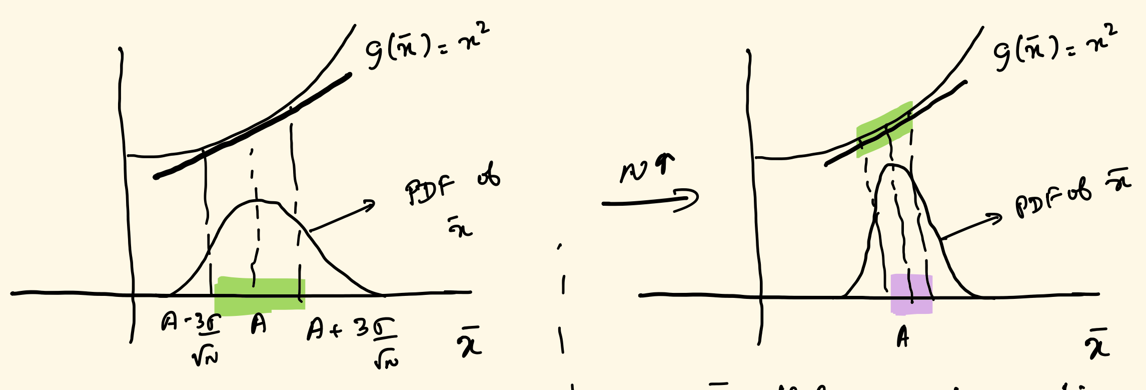
$$= 4\mu^2\sigma^2 + 2\sigma^4$$

$$\bar{x} \sim \mathcal{N}\left(A, \frac{\sigma^2}{N}\right)$$

$$\mathrm{Var}(\bar{x}^2) = \frac{4A^2\sigma^2}{N} + \frac{2\sigma^4}{N^2}$$

Asymptotically (as $N \to \infty$) efficient.

(Efficiency is approximately preserved under non-linear transformations)



$\bar{x}$ that we observe lies in an even more smaller interval

$$g(\bar{x}) \approx g(A) + \frac{\partial g(A)}{\partial A}(\bar{x} - A)$$

$$E[g(\bar{x})] \approx g(A) = A^2$$

$$\mathrm{Var}(g(\bar{x})) \approx \left(\frac{\partial g(A)}{\partial A}\right)^2 \mathrm{var}(\bar{x}) = \frac{4A^2\sigma^2}{N}$$

Vector parameter:

$$\underline{\theta} = [\theta_1, \ldots \theta_p]^T : p \times 1$$

Suppose $p(\underline{x}; \underline{\theta})$ satisfies the regularity condition

$$E\left[\frac{\partial \ln p(\underline{x}; \underline{\theta})}{\partial \underline{\theta}}\right] = \underline{0} \quad \forall\, \underline{\theta}$$

Then the covariance matrix of any unbiased estimator $\hat{\underline{\theta}}$, $C_{\hat{\theta}}$, satisfies

$$\rightarrow C_{\hat{\theta}} - I^{-1}(\theta) \succeq 0$$

PSD $\rightarrow$ diagonal entries are non-negative

$C_{\hat{\theta}}$: $p \times p$ $\qquad$ $I^{-1}(\theta)$: $p \times p$ $\rightarrow$ "Fisher information matrix"

$$[C_{\hat{\theta}} - I^{-1}(\theta)]_{ii} \geq 0$$

$$\Rightarrow \mathrm{var}(\hat{\theta}_i) = [C_{\hat{\theta}}]_{ii} \geq [I^{-1}(\theta)]_{ii}$$

$$[I(\theta)]_{ij} = -E\left[\frac{\partial^2 \ln p(\underline{x}; \underline{\theta})}{\partial \theta_i\, \partial \theta_j}\right]$$

$$= E\left[\frac{\partial \ln p(\underline{x}; \underline{\theta})}{\partial \theta_i} \frac{\partial \ln p(\underline{x}; \underline{\theta})}{\partial \theta_j}\right]$$

$$\underbrace{\frac{\partial \ln p(\underline{x}; \underline{\theta})}{\partial \underline{\theta}}}_{p \times 1} = I(\theta)\,[g(\underline{x}) - \theta]$$

$\hat{\theta}$ will be MVU with $C_{\hat{\theta}} = I^{-1}(\underline{\theta})$

Transformed parameters:

$$\underline{\alpha} = g(\underline{\theta}) : r \times 1$$

$$\Rightarrow C_{\hat{\alpha}} - \underbrace{\frac{\partial g(\underline{\theta})}{\partial \underline{\theta}}}_{r \times p} I^{-1}(\underline{\theta}) \frac{\partial g(\underline{\theta})}{\partial \underline{\theta}^T} \succeq 0$$

unbiasedness:

$$\rightarrow \int (\underbrace{\hat{\underline{\alpha}}}_{r \times 1} - \underline{\alpha}) \underbrace{\frac{\partial \ln p(\underline{x}; \theta)}{\partial \underline{\theta}}^T}_{1 \times p} p(\underline{x}; \theta)\, d\underline{x} = \underbrace{\frac{\partial g(\underline{\theta})}{\partial \underline{\theta}}}_{r \times p}$$

$\underline{a} : r \times 1$
$\underline{b} : p \times 1$ $\quad \}$ arbitrary

$$\int \underline{a}^T (\hat{\underline{\alpha}} - \underline{\alpha}) \frac{\partial \ln p(\underline{x}; \theta)}{\partial \underline{\theta}}^T \underline{b}\; p(\underline{x}; \theta)\, d\underline{x} = \underline{a}^T \frac{\partial g(\underline{\theta})}{\partial \underline{\theta}} \underline{b}$$

Cauchy - Schwarz inequality:

$$\left(\underline{a}^T \frac{\partial g(\underline{\theta})}{\partial \theta} \underline{b}\right)^2 \leq \int \underline{a}^T (\hat{\underline{\alpha}} - \underline{\alpha})(\hat{\underline{\alpha}} - \underline{\alpha})^T \underline{a}\; p(\underline{x}; \underline{\theta})\, d\underline{x}$$

$$\times \int \underline{b}^T \frac{\partial \ln p(\underline{x}; \theta)}{\partial \theta} \frac{\partial \ln p(\underline{x}; \theta)}{\partial \theta}^T \underline{b}\; p(\underline{x}; \theta)\, d\underline{x}$$

$$= \underline{a}^T C_{\hat{\alpha}}\, \underline{a} \quad \underline{b}^T I(\underline{\theta})\, \underline{b}$$

Let us choose:

$$\underline{b} = I^{-1}(\theta) \frac{\partial g^T(\theta)}{\partial \underline{\theta}} \underline{a}$$

$$\left(\underline{a}^T \frac{\partial g(\theta)}{\partial \theta} I^{-1}(\theta) \frac{\partial g^T(\theta)}{\partial \underline{\theta}} \underline{a}\right)^2 \leq \underline{a}^T C_{\hat{\alpha}}\, \underline{a}$$

$$\times \left(\underline{a}^T \frac{\partial g(\theta)}{\partial \theta} \underbrace{I^{-1}(\theta)}_{P.D} \frac{\partial g^T(\theta)}{\partial \theta} \underline{a}\right)$$

$$\underline{a}^T \left(C_{\hat{\alpha}} - \frac{\partial g(\theta)}{\partial \theta} I^{-1}(\theta) \frac{\partial g^T(\theta)}{\partial \theta}\right) \underline{a} \geq 0$$

PSD: $\underline{z}^T A\, \underline{z} \geq 0$

$$\Rightarrow C_{\hat{\alpha}} - \frac{\partial g(\underline{\theta})}{\partial \underline{\theta}} I^{-1}(\theta) \frac{\partial g^T(\theta)}{\partial \underline{\theta}} \succeq 0$$

$$\alpha = g(\theta) = \theta$$

$$\frac{\partial g(\theta)}{\partial \theta} = I$$

Example:

$$\underline{x} = H\underline{\theta} + \underline{w}$$

$\underline{\theta}: p \times 1$ $\quad$ $\underline{x}: N \times 1$ $\quad$ $H: N \times p$

$x[n] = A + Bn + w[n]$

$x[n] = A + Bn + Cn^2 + w[n]$

$$\underline{x} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \\ 1 & N-1 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} A \\ B \end{bmatrix}}_{\underline{\theta}}$$

$$\underline{w} \sim \mathcal{N}(\underline{0}, \sigma^2 I)$$

$$\ln p(\underline{x}; \underline{\theta}) = k - \frac{1}{2\sigma^2}[\underline{x} - H\underline{\theta}]^T[\underline{x} - H\underline{\theta}]$$

$$p \times 1 \quad \frac{\partial \ln p(\underline{x}; \underline{\theta})}{\partial \underline{\theta}} = \frac{1}{\sigma^2} H^T [\underline{x} - H\underline{\theta}]$$

$$\frac{\partial^2 \ln p(\underline{x}; \theta)}{\partial \underline{\theta}\, \partial \underline{\theta}^T} = -\frac{1}{\sigma^2} H^T H$$

$$-E\left[\frac{\partial^2 \ln p(\underline{x}; \theta)}{\partial \underline{\theta}\, \partial \underline{\theta}^T}\right] = \frac{H^T H}{\sigma^2} = I(\theta)$$

$$\frac{\partial}{\partial \theta} \ln p(\underline{x}; \theta) = \underbrace{\frac{1}{\sigma^2}(H^T H)}_{I(\theta)} \left[\underbrace{(H^T H)^{-1} H^T \underline{x}}_{g(\underline{x})} - \underline{\theta}\right]$$

$$\rightarrow I(\theta) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \qquad \underline{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

$$\rightarrow H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad x = A + B$$

$$\theta' = \begin{bmatrix} A - \alpha \\ B + \alpha \end{bmatrix} \qquad \theta = \begin{bmatrix} A \\ B \end{bmatrix}$$

$H^T H$ is ill-conditioned

Constant rank of FIM implies local identifiability

Converse is not true

$$x = \theta^2 + w$$

$$I(\theta) = 4\theta^2 \neq 0 \qquad \theta \neq 0$$

but $\theta = 0$ is identifiable.