Estimation Notes 3.7 pp.39 - 45

Key Points : Extension to a Vector Parameter

Extension to a Vector Parameter

1. Consider $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_p]^{\mathrm{T}}$, if $\hat{\boldsymbol{\theta}}$ is unbiased estimator (i.e., $E\{\hat{\boldsymbol{\theta}}\} = \hat{\boldsymbol{\theta}}$), the CRLB can be found as the $[i, i]$ element of the **inverse of a matrix** or

$$var(\hat{\theta}_i) \geq [\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{ii} \tag{3.20}$$

where $\boldsymbol{I}(\boldsymbol{\theta})$ is the $p \times p$ **Fisher Information Matrix**, which is defined as

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{ij} = -E\left\{\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\} \quad i = 1, \ldots, p; j = 1, \ldots, p \tag{3.21}$$

2. Example 3.6 (Revisited Example 3.3) DC level in White Gaussian Noise

Consider $x[n] = A + w[n]$, $n = 0, 1, \ldots, N-1$ and we don't know A or σ^2 of $w[n]$. The parameter vector is $\boldsymbol{\theta} = [A\sigma^2]^{\mathrm{T}}$, hence $p = 2$, then the 2×2 Fisher information matrix is

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{bmatrix} -E\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial A^2}\right\} & -E\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial A \partial\sigma^2}\right\} \\ -E\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\sigma^2 \partial A}\right\} & -E\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial(\sigma^2)^2}\right\} \end{bmatrix}$$

since

$$-E\left\{\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A \partial\sigma^2}\right\} = -E\left\{\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial\sigma^2 \partial A}\right\}$$

the matrix is **symmetric** and the log-likelihood function is, from Example 3.3,

$$\ln p(\mathbf{x}; [A, \sigma^2]) = \ln p(\mathbf{x}; \boldsymbol{\theta})) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n] - A)^2$$

then we can obtain that

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix} \Rightarrow \boldsymbol{I}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{N}{2\sigma^4} & -0 \\ -0 & \frac{N}{\sigma^2} \end{bmatrix} \det(\boldsymbol{I}(\boldsymbol{\theta})) = \boldsymbol{I}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\sigma^2}{N} & -0 \\ -0 & \frac{2\sigma^4}{N} \end{bmatrix}$$

finally, we have

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

recall (3.20), this is the case that $i = 1, 2$

$$var(\hat{\theta}_i) \geq [\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{ii} \quad i = 1, 2$$

2. Cramer-Rao Lower Bound - Vector Parameter

3.It is assumed that the PDF $p(\mathbf{x}; \boldsymbol{\theta})$ satisfies the **regularity** conditions

$$E\left\{\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\} = \mathbf{0} \quad \text{for all } \theta$$

then the **covariance matrix $\boldsymbol{C}_{\hat{\boldsymbol{\theta}}}$** of any unbiased estimator $\hat{\boldsymbol{\theta}}$ satisfies

$$\underbrace{\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{I}^{-1}(\boldsymbol{\theta}) \geq \mathbf{0}}_{(3.24)} \Rightarrow \boldsymbol{C}_{\hat{\boldsymbol{\theta}}} = \underbrace{var(\hat{\theta}_i) \geq \boldsymbol{I}^{-1}(\boldsymbol{\theta})}_{(3.20)}$$

and recall (3.21), the $\boldsymbol{I}(\boldsymbol{\theta})$ is given as

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{ij} = -E\left\{\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right\} \quad i = 1, \ldots, p; j = 1, \ldots, p$$

4. An unbiased estimator may be found that attains the bound in that $\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} = \boldsymbol{I}^{-1}(\boldsymbol{\theta})$ **if and only if**

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \boldsymbol{I}(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x}) - \boldsymbol{\theta})$$

for some p-dimensional function g and some $p \times p$ matrix $\boldsymbol{I}$. That estimator, which is the MVU estimator, is $\hat{\boldsymbol{\theta}} = \mathbf{g}(\mathbf{x})$, and its covariance matrix is $\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} = \boldsymbol{I}^{-1}(\boldsymbol{\theta})$.