

EE6617: Tutorial 3

- 3.1** Assume that a deterministic signal $s[n; \theta]$, a function of an unknown parameter θ , is observed in White Gaussian noise,

$$x[n] = s[n; \theta] + w[n], \quad n = 0, \dots, N-1, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}).$$

Note that the dependence of $s[n; \theta]$ on θ is explicitly noted.

- (a) Find the expression for the likelihood function $p(\mathbf{x}; \theta)$.
- (b) Show that the regularity condition holds, i.e., show

$$E \left[\frac{\partial \ln(p(\mathbf{x}; \theta))}{\partial \theta} \right] = 0.$$

- (c) Evaluate CRLB of the estimator $\hat{\theta}$ and express it in terms of $s[n; \theta], \theta, \sigma^2$.
- (d) You are about to design the signal $s[n; \theta]$ so as to improve the estimation performance. One reasonable approach is to lower the CRLB found in (c). The lower the CRLB, the more accurate the estimator is.

What are the desired properties of $s[n; \theta]$ to minimize the CRLB?

- 3.2** From **3.1**, we set $s[n; \theta] = A$ with A being the DC value. Provided the observation $x[n] = A + w[n]$, $n = 0, 1, \dots, N-1$, $\mathbf{w} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$, we want to estimate $\alpha = g(A) = 3A + 5$.

- (a) We have learnt during the class that an estimator is called **efficient**, if it is unbiased and achieves the CRLB with equality.

Is an **efficient** estimator is MVU estimator?

- (b) Find the **efficient** estimator of A .
- (c) Find the **efficient** estimator of $\alpha = g(A) = 3A + 5$.
- (d) Suppose we want to estimate $\alpha = A^2$. One intuitive way is to design $\hat{\alpha} = \hat{A}^2$. Is $\hat{\alpha} = \hat{A}^2$ **efficient**?

- 3.3** Consider the same observation as **3.2**. In this exercise, by using CRLB theorem, we want to estimate two parameters A and σ^2 , i.e., the estimation parameters are represented as the vector $\boldsymbol{\theta} = [A \ \sigma^2]^T$.

- (a) Find the likelihood function $p(\mathbf{x}; \boldsymbol{\theta})$.
- (b) Compute the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$.
- (c) Find CRLB. Given the CRLBs for $\hat{A}$ and $\hat{\sigma}^2$, state which estimator is easier to be estimated than the other?
- (d) Design MVU estimators using CRLB Theorem. Are those estimators realizable?

Sol. 3.1 (a) The likelihood function is given by

$$p(\mathbf{x}; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta])^2 \right)$$

(b) To show the regularity condition, we compute

$$\begin{aligned} -\ln(p(\mathbf{x}; \theta)) &= \frac{N}{2} \ln(2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta])^2 \\ -\frac{\partial \ln(p(\mathbf{x}; \theta))}{\partial \theta} &= -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta]) \frac{\partial s[n; \theta]}{\partial \theta} \\ E \left[-\frac{\partial \ln(p(\mathbf{x}; \theta))}{\partial \theta} \right] &= -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} (E[x[n]] - s[n; \theta]) \frac{\partial s[n; \theta]}{\partial \theta} = 0 \end{aligned}$$

where the last equality follows from the fact that $E[x[n]] = s[n; \theta]$.

(c) Now, we need to compute the Fisher information $I(\theta) = E \left[-\frac{\partial^2 \ln(p(\mathbf{x}; \theta))}{\partial^2 \theta} \right]$.

$$\begin{aligned} -\frac{\partial^2 \ln(p(\mathbf{x}; \theta))}{\partial^2 \theta} &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left(\left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2 - (x[n] - s[n; \theta]) \frac{\partial^2 s[n; \theta]}{\partial^2 \theta} \right) \\ I(\theta) = -E \left[\frac{\partial^2 \ln(p(\mathbf{x}; \theta))}{\partial^2 \theta} \right] &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2 \\ \text{var}(\hat{\theta}) &\geq \frac{1}{I(\theta)} = \frac{\sigma^2}{\sum_{n=0}^{N-1} \left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2} = CRLB(\hat{\theta}) \end{aligned}$$

(d) In order to minimize the $CRLB(\hat{\theta})$, we must maximize $\left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2$. Thus, the desired $s[n; \theta]$ must change rapidly as θ changes, resulting in large $\left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2$ value.

Sol. 3.2 (a) Yes, it is an MVU estimator because it is unbiased and achieves the minimum variance.

(b) We already know the efficient estimator for DC level in White Gaussian noise is the sample mean $\hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$.

(c) Recall estimating linear function of parameters A is equivalent to estimating the parameter A and doing a linear mapping $3\hat{A} + 5$. Thus, the efficient estimator for $\alpha = 3A + 5$ is $\hat{\alpha} = 3\hat{A} + 5 = \frac{3}{N} \sum_{n=0}^{N-1} x[n] + 5 = \frac{1}{N} \sum_{n=0}^{N-1} (3x[n] + 5)$.

(d) Recall the efficient estimator is unbiased and achieves the CRLB with the equality. First, try $E[\hat{\alpha}]$ and

$$E[\hat{A}^2] = \text{var}(\hat{A}) + E[\hat{A}]^2 = \frac{\sigma^2}{N} + A^2.$$

Since $E[\hat{A}^2] \neq A^2$, $\hat{\alpha} = \hat{A}^2$ is biased and thus it is not efficient.

Sol. 3.3 (a) The likelihood function is given by

$$p(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2\right).$$

(b) The Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ is given by

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} -E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 A}\right] & -E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A \partial \sigma^2}\right] \\ -E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2 \partial A}\right] & -E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 \sigma^2}\right] \end{bmatrix} \quad (1)$$

To compute this, we first need to check the regularity condition, $E_{\mathbf{x}}\left[\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \boldsymbol{\theta}}\right] = E_{\mathbf{x}}\left[\frac{\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A}}{\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2}}\right] = \mathbf{0}$.

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

$$\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{n=0}^{N-1} (x[n] - A)^2. \quad (3)$$

Since $E[x[n]] = A$, $E\left[\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A}\right] = 0$ and $E\left[\frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2}\right] = 0$, thus satisfying the regularity condition. We further compute

$$\begin{aligned} -\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 A} &= \frac{N}{\sigma^2} \\ -\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 \sigma^2} &= -\frac{N}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{n=0}^{N-1} (x[n] - A)^2 \\ -\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A \partial \sigma^2} &= \frac{1}{\sigma^4} \sum_{n=0}^{N-1} (x[n] - A) = -\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2 \partial A}. \end{aligned}$$

Then, $-E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 A}\right] = \frac{N}{\sigma^2}$, $-E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial^2 \sigma^2}\right] = -\frac{N}{2\sigma^4} + \frac{N}{\sigma^4} = \frac{N}{2\sigma^4}$, and $-E\left[\frac{\partial^2 \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A \partial \sigma^2}\right] = 0$.

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

- (c) Compute the inverse of $\mathbf{I}(\boldsymbol{\theta})$,

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

Thus, the bounds of variances of estimators $\hat{A}$ and $\hat{\sigma}^2$ are given by

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

Since $\frac{\sigma^2}{N} \leq \frac{2\sigma^2}{N}$, we can say that estimating the mean A is easier than estimating the variance σ^2 .

- (d) An unbiased estimator that achieves the bound CRLB in (4) is found if and only if

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

First, we factorize (2) and (3) in terms of the diagonal elements of $\mathbf{I}(\boldsymbol{\theta})$,

$$\begin{aligned} \frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A} &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A) = \frac{N}{\sigma^2} \left(\frac{1}{N} \sum_{n=0}^{N-1} x[n] - A \right) \\ \frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2} &= -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{n=0}^{N-1} (x[n] - A)^2 = \frac{N}{2\sigma^4} \left(\frac{1}{N} \sum_{n=0}^{N-1} (x[n] - A)^2 - \sigma^2 \right) \end{aligned}$$

Then, it follows

$$\begin{aligned} \begin{bmatrix} \frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial A} \\ \frac{\partial \ln(p(\mathbf{x}; \boldsymbol{\theta}))}{\partial \sigma^2} \end{bmatrix} &= \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix} \left(\begin{bmatrix} \frac{1}{N} \sum_{n=0}^{N-1} x[n] \\ \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - A)^2 \end{bmatrix} - \begin{bmatrix} A \\ \sigma^2 \end{bmatrix} \right) \\ &= \mathbf{I}(\boldsymbol{\theta}) (\boldsymbol{\alpha}(\mathbf{x}) - \boldsymbol{\theta}) \end{aligned}$$

The CRLB theorem tells us that we could design two estimators $\hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$ and

$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - A)^2$. In case of $\hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$, this is MVU with $\text{var}(\hat{A}) = \frac{\sigma^2}{N}$.

However, $\hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - A)^2$ is not realizable because of the dependency to A .