Estimation Theory 2024-09-11

CRLB for Vector Parameters

Vector parameter

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$

if $\hat{\theta}$, an unbiased estimate of θ.

CRLB:

$$\mathrm{var}(\hat{\theta}_i) \geq \left[\mathbf{I}^{-1}(\boldsymbol{\theta})\right]_{ii}$$

$I(\theta)$, Fisher information matrix

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

ex1)

$x[n] = A + w[n]$, where $w[n] \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$

Unknowns: A and σ^2

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

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

$$= \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}$$

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

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

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

ex2)

$x[n] = A + Bn + w[n] \qquad n = 0, 1, \ldots, N-1$

$$\boldsymbol{\theta} = [A\ B]^T$$

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

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

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

$$= \frac{1}{\sigma^2}\begin{bmatrix} N & \frac{N(N-1)}{2} \\ \frac{N(N-1)}{2} & \frac{N(N-1)(2N-1)}{6} \end{bmatrix}$$

$$\mathbf{I}^{-1}(\boldsymbol{\theta}) = \sigma^2\begin{bmatrix} \frac{2(2N-1)}{N(N+1)} & -\frac{6}{N(N+1)} \\ -\frac{6}{N(N+1)} & \frac{12}{N(N^2-1)} \end{bmatrix} \rightarrow \begin{aligned} \mathrm{var}(\hat{A}) &\geq \frac{2(2N-1)\sigma^2}{N(N+1)} \\ \mathrm{var}(\hat{B}) &\geq \frac{12\sigma^2}{N(N^2-1)}. \end{aligned} \rightarrow$$

1. CRLB always increases as we estimate more parameters
   - When only A is unknown, $\mathrm{var}(\hat{A}) \geq \frac{\sigma^2}{N}$.
   - When both A and B are unknown, $\mathrm{var}(\hat{A}) \geq \frac{2(2N-1)\sigma^2}{N(N+1)}$.
   - For $N \geq 2$, $\frac{2(2N-1)\sigma^2}{N(N+1)} > \frac{\sigma^2}{N}$.
2. Some parameters are more sensitive than others

$$\frac{\mathrm{CRLB}(\hat{A})}{\mathrm{CRLB}(\hat{B})} = \frac{(2N-1)(N-1)}{6} > 1$$ for $N \geq 3$. i.e., B is easier to estimate

Vector Parameter CRLB for Transformation
함수화된 변수에 대한 추정

let $\boldsymbol{\alpha} = \mathbf{g}(\boldsymbol{\theta})$ r-dimensional function

$$\rightarrow \mathbf{C}_{\hat{\alpha}} - \frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{I}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{g}(\boldsymbol{\theta})^T}{\partial\boldsymbol{\theta}} \geq \mathbf{0}$$

ex) $x[n] = A + w[n]$, where $w[n] \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$.

- Unknowns: A and σ^2
- Estimate $\alpha = \frac{A^2}{\sigma^2}$ (SNR)

$$\begin{aligned}
\frac{\partial g(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial\theta_1} & \frac{\partial g(\boldsymbol{\theta})}{\partial\theta_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial A} & \frac{\partial g(\boldsymbol{\theta})}{\partial\sigma^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{I}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{g}(\boldsymbol{\theta})^T}{\partial\boldsymbol{\theta}} &= \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix}\begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix}\begin{bmatrix} \frac{2A}{\sigma^2} \\ -\frac{A^2}{\sigma^4} \end{bmatrix} \\
&= \frac{4A^2}{N\sigma^2} + \frac{2A^4}{N\sigma^4} \\
&= \frac{4\alpha + 2\alpha^2}{N}.
\end{aligned} \qquad \mathrm{var}(\hat{\alpha}) \geq \frac{4\alpha + 2\alpha^2}{N}$$

CRLB for the General Gaussian Case
일반 가우시안에서의 CRLB

$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$

$$\begin{aligned}
[\mathbf{I}(\boldsymbol{\theta})]_{ij} &= \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right]^T \mathbf{C}^{-1}(\boldsymbol{\theta})\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right] \\
&+ \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_i}\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_j}\right]
\end{aligned}$$

$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$ (scalar parameter)

$$\begin{aligned}
I(\theta) &= \left[\frac{\partial\boldsymbol{\mu}(\theta)}{\partial\theta}\right]^T \mathbf{C}^{-1}(\theta)\left[\frac{\partial\boldsymbol{\mu}(\theta)}{\partial\theta}\right] \\
&+ \frac{1}{2}\mathrm{tr}\left[\left(\mathbf{C}^{-1}(\theta)\frac{\partial\mathbf{C}(\theta)}{\partial\theta}\right)^2\right]
\end{aligned}$$

ex) $x[n] = A + w[n]$, where

- $w[n] \sim \mathcal{N}(0, \sigma^2)$
- $A \sim \mathcal{N}(0, \sigma_A^2)$

$$\begin{aligned}
[\mathbf{C}(\sigma_A^2)]_{ij} &= E\left[x[i-1]x[j-1]\right] \\
&= E\left[(A + w[i-1])(A + w[j-1])\right] \\
&= \sigma_A^2 + \sigma^2\delta_{ij}.
\end{aligned}$$

$$\mathbf{C}(\sigma_A^2) = \sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I}$$

$$\mathbf{C}^{-1}(\sigma_A^2) = \frac{1}{\sigma^2}\left(\mathbf{I} - \frac{\sigma_A^2}{\sigma^2 + N\sigma_A^2}\mathbf{1}\mathbf{1}^T\right)$$ (Woodbury identity) $\rightarrow$

$$\frac{\partial\mathbf{C}(\sigma_A^2)}{\partial\sigma_A^2} = \mathbf{1}\mathbf{1}^T, \qquad \mathbf{C}^{-1}(\sigma_A^2)\frac{\partial\mathbf{C}(\sigma_A^2)}{\partial\sigma_A^2} = \frac{1}{\sigma^2 + N\sigma_A^2}\mathbf{1}\mathbf{1}^T$$

$$\begin{aligned}
I(\theta) &= \left[\frac{\partial\boldsymbol{\mu}(\theta)}{\partial\theta}\right]^T \mathbf{C}^{-1}(\theta)\left[\frac{\partial\boldsymbol{\mu}(\theta)}{\partial\theta}\right] \\
&+ \frac{1}{2}\mathrm{tr}\left[\left(\mathbf{C}^{-1}(\theta)\frac{\partial\mathbf{C}(\theta)}{\partial\theta}\right)^2\right]
\end{aligned}$$

$$\begin{aligned}
I(\sigma_A^2) &= \frac{1}{2}\mathrm{tr}\left[\left(\frac{1}{\sigma^2 + N\sigma_A^2}\right)^2\mathbf{1}\mathbf{1}^T\mathbf{1}\mathbf{1}^T\right] \\
&= \frac{N}{2}\left(\frac{1}{\sigma^2 + N\sigma_A^2}\right)^2\mathrm{tr}(\mathbf{1}\mathbf{1}^T) \\
&= \frac{1}{2}\left(\frac{N}{\sigma^2 + N\sigma_A^2}\right)^2
\end{aligned} \qquad \mathrm{var}(\sigma_A^2) \geq 2\left(\sigma_A^2 + \frac{\sigma^2}{N}\right)^2$$