8.7 Multiparameter CRLB

In practice, we are more likely to have many parameters, some of which may be a nuisance. Is there any effect of the extra parameters on the CRLB of the parameter of interest? This question is meaningful since the CRLB is usually achieved by the MLE in large samples. This section is closely related to Section 3.3 on multiparameter (observed) Fisher information, but for completeness we repeat some of the notations.

Let $\theta = (\theta_1, \ldots, \theta_p)$. The score function is now a gradient vector

$$S(\theta) = \frac{\partial}{\partial \theta} \log L(\theta) = \begin{pmatrix} \frac{\partial}{\partial \theta_1} \log L(\theta) \\ \vdots \\ \frac{\partial}{\partial \theta_p} \log L(\theta) \end{pmatrix}.$$

The observed Fisher information is minus the Hessian of the log-likelihood function

$$I(\theta) = -\frac{\partial^2}{\partial \theta \partial \theta'} \log L(\theta).$$

The *expected* Fisher information is

$$\mathcal{I}(\theta) = E_\theta I(\theta).$$

Using similar methods as in the scalar case, assuming regularity conditions, we can show that

$$\begin{aligned} E_\theta S(\theta) &= 0 \\ \text{var}_\theta\{S(\theta)\} &= \mathcal{I}(\theta). \end{aligned}$$

The information matrix $\mathcal{I}(\theta)$ is now a $p \times p$ variance matrix, which means it is a nonnegative definite matrix.

Example 8.10: Let $x_1, \ldots, x_n$ be an iid sample from $N(\mu, \sigma^2)$ and let $\theta = (\mu, \sigma^2)$. Then we have the following:

$$\begin{aligned}
\log L(\theta) &= -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \\
S(\theta) &= \begin{pmatrix} \frac{\partial}{\partial\mu}\log L(\theta) \\ \frac{\partial}{\partial\sigma^2}\log L(\theta) \end{pmatrix} = \begin{pmatrix} \frac{n}{\sigma^2}(\overline{x} - \mu) \\ -\frac{n}{2\sigma^2} + \frac{\sum(x_i-\mu)^2}{2\sigma^4} \end{pmatrix} \\
I(\theta) &= \begin{pmatrix} \frac{n}{\sigma^2} & \frac{n}{\sigma^4}(\overline{x} - \mu) \\ \frac{n}{\sigma^4}(\overline{x} - \mu) & -\frac{n}{2\sigma^4} + \frac{\sum(x_i-\mu)^2}{\sigma^6} \end{pmatrix} \\
\mathcal{I}(\theta) &= \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}. \square
\end{aligned}$$

We now state the multiparameter version of the CRLB theorem:

Theorem 8.5 *Let $T(X)$ be a scalar function, $E_\theta T = g(\theta)$ and $\mathcal{I}(\theta)$ the expected Fisher information for θ based on data X. Then*

$$var_\theta(T) \geq \alpha' \mathcal{I}(\theta)^{-1}\alpha,$$

where $\alpha = \frac{\partial}{\partial\theta}g(\theta)$.

Proof: The proof relies on an extension of the covariance inequality involving a scalar random variable T and a vector of random variables S:

$$\mathrm{var}(T) \geq \mathrm{cov}(S,T)'\{\mathrm{var}(S)\}^{-1}\mathrm{cov}(S,T).$$

and showing that $\mathrm{cov}\{S(\theta), T\} = \frac{\partial}{\partial\theta}g(\theta)$. The proof of these statements is left as an exercise. $\square$