Appendix 3A

Derivation of Scalar Parameter CRLB

In this appendix we derive the CRLB for a scalar parameter $\alpha = g(\theta)$ where the PDF is parameterized by $\theta$. We consider all unbiased estimators $\hat{\alpha}$ or those for which

$$E(\hat{\alpha}) = \alpha = g(\theta)$$

or

$$\int \hat{\alpha} p(x; \theta) \, dx = g(\theta). \tag{3A.1}$$

Before beginning the derivation we first examine the regularity condition

$$E \left[ \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = 0, \tag{3A.2}$$

which is assumed to hold. Note that

$$\int \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = \int \frac{\partial p(x; \theta)}{\partial \theta} \, dx$$

$$= \frac{\partial}{\partial \theta} \int p(x; \theta) \, dx$$

$$= \frac{\partial 1}{\partial \theta}$$

$$= 0.$$

Hence, we conclude that the regularity condition will be satisfied if the order of differentiation and integration may be interchanged. This is generally true except when the domain of the PDF for which it is nonzero depends on the unknown parameter such as in Problem 3.1.

Now differentiating both sides of (3A.1) with respect to $\theta$ and interchanging the partial differentiation and integration produces

$$\int \hat{\alpha} \frac{\partial p(x; \theta)}{\partial \theta} \, dx = \frac{\partial g(\theta)}{\partial \theta}.$$
or
\[
\int \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = \frac{\partial g(\theta)}{\partial \theta}.
\tag{3A.3}
\]

We can modify this using the regularity condition to produce
\[
\int (\hat{\alpha} - \alpha) \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = \frac{\partial g(\theta)}{\partial \theta}
\tag{3A.4}
\]

since
\[
\int \alpha \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = \alpha E \left[ \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = 0.
\]

We now apply the Cauchy-Schwarz inequality
\[
\left[ \int w(x) g(x) h(x) \, dx \right]^2 \leq \int w(x) g^2(x) \, dx \int w(x) h^2(x) \, dx \tag{3A.5}
\]
which holds with equality if and only if \( g(x) = ch(x) \) for \( c \) some constant not dependent on \( x \). The functions \( g \) and \( h \) are arbitrary scalar functions, while \( w(x) \geq 0 \) for all \( x \). Now let
\[
w(x) = p(x; \theta) \\
g(x) = \hat{\alpha} - \alpha \\
h(x) = \frac{\partial \ln p(x; \theta)}{\partial \theta}
\]
and apply the Cauchy-Schwarz inequality to (3A.4) to produce
\[
\left( \frac{\partial g(\theta)}{\partial \theta} \right)^2 \leq \int (\hat{\alpha} - \alpha)^2 p(x; \theta) \, dx \int \left( \frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 p(x; \theta) \, dx
\]
or
\[
\text{var}(\hat{\alpha}) \geq \frac{\left( \frac{\partial g(\theta)}{\partial \theta} \right)^2}{E \left[ \left( \frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]}.\]

Now note that
\[
E \left[ \left( \frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right] = -E \left[ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right].
\]
This follows from (3A.2) as
\[
E \left[ \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = 0 \\
\int \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = 0.
\[ \frac{\partial}{\partial \theta} \int \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx = 0 \]

or

\[ -E \left[ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right] = \int \frac{\partial \ln p(x; \theta)}{\partial \theta} \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) \, dx \]

which is (3.16). If \( \alpha = g(\theta) = \theta \), we have (3.6).

Note that the condition for equality is

\[ \frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c} (\hat{\alpha} - \alpha) \]

where \( c \) can depend on \( \theta \) but not on \( x \). If \( \alpha = g(\theta) = \theta \), we have

\[ \frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c(\theta)} (\hat{\theta} - \theta). \]

The possible dependence of \( c \) on \( \theta \) is noted. To determine \( c(\theta) \)

\[ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} = -\frac{1}{c(\theta)} + \frac{\partial}{\partial \theta} \left( \frac{1}{c(\theta)} \right) (\hat{\theta} - \theta) \]

\[ -E \left[ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right] = \frac{1}{c(\theta)} \]

or finally

\[ c(\theta) = \frac{1}{-E \left[ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right]} = \frac{1}{I(\theta)} \]

which agrees with (3.7).
Appendix 3B

Derivation of Vector Parameter CRLB

In this appendix the CRLB for a vector parameter \( \alpha = g(\theta) \) is derived. The PDF is characterized by \( \theta \). We consider unbiased estimators such that

\[
E(\hat{\alpha}_i) = \alpha_i = [g(\theta)]_i, \quad i = 1, 2, \ldots, r.
\]

The regularity conditions are

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

so that

\[
\int (\hat{\alpha}_i - \alpha_i) \frac{\partial \ln p(x; \theta)}{\partial \theta_i} p(x; \theta) \, dx = \frac{\partial [g(\theta)]_i}{\partial \theta_i}.
\] (3B.1)

Now consider for \( j \neq i \)

\[
\int (\hat{\alpha}_i - \alpha_i) \frac{\partial \ln p(x; \theta)}{\partial \theta_j} p(x; \theta) \, dx = \int (\hat{\alpha}_i - \alpha_i) \frac{\partial p(x; \theta)}{\partial \theta_j} \, dx
\]

\[
= \frac{\partial}{\partial \theta_j} \int \hat{\alpha}_i p(x; \theta) \, dx
\]

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

\[
= \frac{\partial \alpha_i}{\partial \theta_j}
\]

\[
= \frac{\partial [g(\theta)]_i}{\partial \theta_j}.
\] (3B.2)

Combining (3B.1) and (3B.2) into matrix form, we have

\[
\int (\hat{\alpha} - \alpha) \frac{\partial \ln p(x; \theta)}{\partial \theta}^T p(x; \theta) \, dx = \frac{\partial g(\theta)}{\partial \theta}.
\]
Now premultiply by $\mathbf{a}^T$ and postmultiply by $\mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ are arbitrary $r \times 1$ and $p \times 1$ vectors, respectively, to yield

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

Now let

$$
w(\mathbf{x}) = p(\mathbf{x}; \theta) \quad \quad g(\mathbf{x}) = \mathbf{a}^T (\hat{\mathbf{x}} - \mathbf{x}) \quad \quad h(\mathbf{x}) = \frac{\partial \ln p(\mathbf{x}; \theta)^T}{\partial \theta} \mathbf{b}
$$

and apply the Cauchy-Schwarz inequality of (3A.5)

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

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

$$
= \mathbf{a}^T \mathbf{C}_\alpha \mathbf{a} \mathbf{b}^T \mathbf{I}(\theta) \mathbf{b}
$$

since as in the scalar case

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

Since $\mathbf{b}$ was arbitrary, let

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

to yield

$$
\left( \mathbf{a}^T \frac{\partial \mathbf{g}(\theta)}{\partial \theta} \mathbf{I}^{-1}(\theta) \frac{\partial \mathbf{g}(\theta)^T}{\partial \theta} \mathbf{a} \right)^2 \leq \mathbf{a}^T \mathbf{C}_\alpha \mathbf{a} \left( \mathbf{a}^T \frac{\partial \mathbf{g}(\theta)}{\partial \theta} \mathbf{I}^{-1}(\theta) \frac{\partial \mathbf{g}(\theta)^T}{\partial \theta} \mathbf{a} \right).
$$

Since $\mathbf{I}(\theta)$ is positive definite, so is $\mathbf{I}^{-1}(\theta)$, and $\frac{\partial \mathbf{g}(\theta)}{\partial \theta} \mathbf{I}^{-1}(\theta) \frac{\partial \mathbf{g}(\theta)^T}{\partial \theta}$ is at least positive semidefinite. The term inside the parentheses is therefore nonnegative, and we have

$$
\mathbf{a}^T \left( \mathbf{C}_\alpha - \frac{\partial \mathbf{g}(\theta)}{\partial \theta} \mathbf{I}^{-1}(\theta) \frac{\partial \mathbf{g}(\theta)^T}{\partial \theta} \right) \mathbf{a} \geq 0.
$$

Recall that $\mathbf{a}$ was arbitrary, so that (3.30) follows. If $\mathbf{x} = \mathbf{g}(\theta) = \theta$, then $\frac{\partial \mathbf{g}(\theta)}{\partial \theta} = \mathbf{I}$ and (3.24) follows. The conditions for equality are $g(\mathbf{x}) = ch(\mathbf{x})$, where $c$ is a constant not dependent on $\mathbf{x}$. This condition becomes

$$
\mathbf{a}^T (\hat{\mathbf{x}} - \mathbf{x}) = c \frac{\partial \ln p(\mathbf{x}; \theta)^T}{\partial \theta} \mathbf{b}
$$

$$
= c \frac{\partial \ln p(\mathbf{x}; \theta)^T}{\partial \theta} \mathbf{I}^{-1}(\theta) \frac{\partial \mathbf{g}(\theta)^T}{\partial \theta} \mathbf{a}.
$$
Since \( a \) was arbitrary

\[
\frac{\partial g(\theta)}{\partial \theta} I^{-1}(\theta) \frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c} (\hat{\alpha} - \alpha).
\]

Consider the case when \( \alpha = g(\theta) = \theta \), so that \( \frac{\partial g(\theta)}{\partial \theta} = 1 \). Then,

\[
\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c} I(\theta)(\hat{\theta} - \theta).
\]

Noting that \( c \) may depend on \( \theta \), we have

\[
\frac{\partial \ln p(x; \theta)}{\partial \theta_i} = \sum_{k=1}^{p} \frac{[I(\theta)]_{ik}}{c(\theta)} (\hat{\theta}_k - \theta_k)
\]

and differentiating once more

\[
\frac{\partial^2 \ln p(x; \theta)}{\partial \theta_i \partial \theta_j} = \sum_{k=1}^{p} \left( \frac{[I(\theta)]_{ik}}{c(\theta)} (-\delta_{kj}) + \frac{\partial}{\partial \theta_j} \left( \frac{[I(\theta)]_{ik}}{c(\theta)} \right) (\hat{\theta}_k - \theta_k) \right).
\]

Finally, we have

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

since \( E(\hat{\theta}_k) = \theta_k \). Clearly, \( c(\theta) = 1 \) and the condition for equality follows.
Appendix 3C

Derivation of General Gaussian CRLB

Assume that $x \sim N(\mu(\theta), C(\theta))$, where $\mu(\theta)$ is the $N \times 1$ mean vector and $C(\theta)$ is the $N \times N$ covariance matrix, both of which depend on $\theta$. Then the PDF is

$$p(x; \theta) = \frac{1}{(2\pi)^{\frac{N}{2}} \det^{\frac{1}{2}}[C(\theta)]} \exp \left[ -\frac{1}{2} (x - \mu(\theta))^T C^{-1}(\theta) (x - \mu(\theta)) \right].$$

We will make use of the following identities

$$\frac{\partial \ln \det[C(\theta)]}{\partial \theta_k} = \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) \tag{3C.1}$$

where $\partial C(\theta)/\partial \theta_k$ is the $N \times N$ matrix with $[i,j]$ element $\partial[C(\theta)]_{ij}/\partial \theta_k$ and

$$\frac{\partial C^{-1}(\theta)}{\partial \theta_k} = -C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} C^{-1}(\theta). \tag{3C.2}$$

To establish (3C.1) we first note

$$\frac{\partial \ln \det[C(\theta)]}{\partial \theta_k} = \frac{1}{\det[C(\theta)]} \frac{\partial \det[C(\theta)]}{\partial \theta_k}. \tag{3C.3}$$

Since $\det[C(\theta)]$ depends on all the elements of $C(\theta)$

$$\frac{\partial \det[C(\theta)]}{\partial \theta_k} = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial \det[C(\theta)]}{\partial [C(\theta)]_{ij}} \frac{\partial [C(\theta)]_{ij}}{\partial \theta_k}$$

$$= \text{tr} \left( \frac{\partial \det[C(\theta)]}{\partial C(\theta)} \frac{\partial C^T(\theta)}{\partial \theta_k} \right) \tag{3C.4}$$

where $\partial \det[C(\theta)]/\partial C(\theta)$ is an $N \times N$ matrix with $[i,j]$ element $\partial \det[C(\theta)]/\partial [C(\theta)]_{ij}$ and the identity

$$\text{tr}(AB^T) = \sum_{i=1}^{N} \sum_{j=1}^{N} [A]_{ij} [B]_{ij}$$

73
has been used. Now by the definition of the determinant
\[
\det[C(\theta)] = \sum_{i=1}^{N} [C(\theta)]_{i,j} |M|_{i,j}
\]
where \(M\) is the \(N \times N\) cofactor matrix and \(j\) can take on any value from 1 to \(N\). Thus,
\[
\frac{\partial \det[C(\theta)]}{\partial [C(\theta)]_{i,j}} = |M|_{i,j}
\]
or
\[
\frac{\partial \det[C(\theta)]}{\partial C(\theta)} = M.
\]

It is well known, however, that
\[
C^{-1}(\theta) = \frac{M^T}{\det[C(\theta)]}
\]
so that
\[
\frac{\partial \det[C(\theta)]}{\partial C(\theta)} = C^{-1}(\theta) \det[C(\theta)].
\]

Using this in (3C.3) and (3C.4), we have the desired result.
\[
\frac{\partial \ln \det[C(\theta)]}{\partial \theta_k} = \frac{1}{\det[C(\theta)]} \text{tr} \left( C^{-1}(\theta) \det[C(\theta)] \frac{\partial C(\theta)}{\partial \theta_k} \right)
\]
\[
= \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right).
\]

The second identity (3C.2) is easily established as follows. Consider
\[
C^{-1}(\theta)C(\theta) = I.
\]
Differentiating each element of the matrices and expressing in matrix form, we have
\[
C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} + \frac{\partial C^{-1}(\theta)}{\partial \theta_k} C(\theta) = 0
\]
which leads to the desired result.

We are now ready to evaluate the CRLB. Taking the first derivative
\[
\frac{\partial \ln p(x; \theta)}{\partial \theta_k} = -\frac{1}{2} \frac{\partial \ln \det[C(\theta)]}{\partial \theta_k} - \frac{1}{2} \frac{\partial}{\partial \theta_k} \left[ (x - \mu(\theta))^T C^{-1}(\theta) (x - \mu(\theta)) \right].
\]

The first term has already been evaluated using (3C.1), so that we consider the second term:
\[
\frac{\partial}{\partial \theta_k} \left[ (x - \mu(\theta))^T C^{-1}(\theta) (x - \mu(\theta)) \right]
\]
\[ \frac{\partial}{\partial \theta_k} \sum_{i=1}^{N} \sum_{j=1}^{N} (x[i] - [\mu(\theta)],_j)(C^{-1}(\theta),_j)(x[j] - [\mu(\theta)],_j) \]
\[ = \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ (x[i] - [\mu(\theta)],_j) \left[ C^{-1}(\theta),_{ij} \left( -\frac{\partial[\mu(\theta)],_j}{\partial \theta_k} \right) \right. \right. \]
\[ + \left. \left. \frac{\partial[C^{-1}(\theta)],_{ij}}{\partial \theta_k} (x[j] - [\mu(\theta)],_j) \right\} \right. \]
\[ + \left. \left( -\frac{\partial[\mu(\theta)],_j}{\partial \theta_k} \right)[C^{-1}(\theta),_{ij}(x[j] - [\mu(\theta)],_j) \right\} \]
\[ = -(x - \mu(\theta))^T C^{-1}(\theta) \frac{\partial \mu(\theta)}{\partial \theta_k} + (x - \mu(\theta))^T \frac{\partial C^{-1}(\theta)}{\partial \theta_k} (x - \mu(\theta)) \]
\[ - \frac{\partial \mu(\theta)^T}{\partial \theta_k} C^{-1}(\theta)(x - \mu(\theta)) \]
\[ = -2 \frac{\partial \mu(\theta)^T}{\partial \theta_k} C^{-1}(\theta)(x - \mu(\theta)) + (x - \mu(\theta))^T \frac{\partial C^{-1}(\theta)}{\partial \theta_k} (x - \mu(\theta)). \]

Using (3C.1) and the last result, we have

\[ \frac{\partial \ln p(x; \theta)}{\partial \theta_k} = -\frac{1}{2} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) + \frac{\partial \mu(\theta)^T}{\partial \theta_k} C^{-1}(\theta)(x - \mu(\theta)) \]
\[ - \frac{1}{2} (x - \mu(\theta))^T \frac{\partial C^{-1}(\theta)}{\partial \theta_k} (x - \mu(\theta)). \]

(3C.5)

Let \( y = x - \mu(\theta) \). Evaluating

\[ [I(\theta)]_{kl} = E \left[ \frac{\partial \ln p(x; \theta)}{\partial \theta_k} \frac{\partial \ln p(x; \theta)}{\partial \theta_l} \right] \]

which is equivalent to (3.23), yields

\[ [I(\theta)]_{kl} = \frac{1}{4} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_l} \right) \]
\[ + \frac{1}{2} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) E \left( y^T \frac{\partial C^{-1}(\theta)}{\partial \theta_l} y \right) \]
\[ + \frac{\partial \mu(\theta)^T}{\partial \theta_k} C^{-1}(\theta) E[y y^T] C^{-1}(\theta) \frac{\partial \mu(\theta)}{\partial \theta_l} \]
\[ + \frac{1}{4} E \left[ y^T \frac{\partial C^{-1}(\theta)}{\partial \theta_k} y y^T \frac{\partial C^{-1}(\theta)}{\partial \theta_l} y \right] \]
where we note that all odd order moments are zero. Continuing, we have

\[
[I(\theta)]_{kl} = \\
\frac{1}{4} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_l} \right) \\
- \frac{1}{2} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_l} \right) \\
+ \frac{\partial \mu(\theta)^T}{\partial \theta_k} C^{-1}(\theta) \frac{\partial \mu(\theta)}{\partial \theta_l} + \frac{1}{4} E \left[ y^T \frac{\partial C^{-1}(\theta)}{\partial \theta_k} y y^T \frac{\partial C^{-1}(\theta)}{\partial \theta_l} y \right] \quad (3C.6)
\]

where \( E(y^T z) = \text{tr}[E(zy^T)] \) for \( y, z \ N \times 1 \) vectors and (3C.2) have been used. To evaluate the last term we use [Porat and Friedlander 1986]

\[
E[y^T Ayy^T By] = \text{tr}(AC)\text{tr}(BC) + 2\text{tr}(ACBC)
\]

where \( C = E(yy^T) \) and \( A \) and \( B \) are symmetric matrices. Thus, this term becomes

\[
\frac{1}{4} \text{tr} \left( \frac{\partial C^{-1}(\theta)}{\partial \theta_k} C(\theta) \right) \text{tr} \left( \frac{\partial C^{-1}(\theta)}{\partial \theta_l} C(\theta) \right) \\
+ \frac{1}{2} \text{tr} \left( \frac{\partial C^{-1}(\theta)}{\partial \theta_k} C(\theta) \frac{\partial C^{-1}(\theta)}{\partial \theta_l} C(\theta) \right).
\]

Next, using the relationship (3C.2), this term becomes

\[
\frac{1}{4} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} \right) \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_l} \right) \\
+ \frac{1}{2} \text{tr} \left( C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_k} C^{-1}(\theta) \frac{\partial C(\theta)}{\partial \theta_l} \right) \quad (3C.7)
\]

and finally, using (3C.7) in (3C.6), produces the desired result.
Appendix 3D

Derivation of Asymptotic CRLB

It can be proven that almost any WSS Gaussian random process \( x[n] \) may be represented as the output of a causal linear shift invariant filter driven at the input by white Gaussian noise \( u[n] \) [Brockwell and Davis 1987] or

\[
x[n] = \sum_{k=0}^{\infty} h[k]u[n-k]
\]

(3D.1)

where \( h[0] = 1 \). The only condition is that the PSD must satisfy

\[
\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) \, df > -\infty.
\]

With this representation the PSD of \( x[n] \) is

\[
P_{xx}(f) = |H(f)|^2 \sigma_u^2
\]

where \( \sigma_u^2 \) is the variance of \( u[n] \) and \( H(f) = \sum_{k=0}^{\infty} h[k]\exp(-j2\pi fk) \) is the filter frequency response. If the observations are \( \{x[0], x[1], \ldots, x[N-1]\} \) and \( N \) is large, then the representation is approximated by

\[
x[n] = \sum_{k=0}^{n} h[k]u[n-k] + \sum_{k=n+1}^{\infty} h[k]u[n-k]
\]

\[
\approx \sum_{k=0}^{n} h[k]u[n-k]. \quad (3D.2)
\]

This is equivalent to setting \( u[n] = 0 \) for \( n < 0 \). As \( n \to \infty \), the approximate representation becomes better for \( x[n] \). It is clear, however, that the beginning samples will be poorly represented unless the impulse response \( h[k] \) is small for \( k > n \). For large \( N \) most of the samples will be accurately represented if \( N \) is much greater than the impulse response length. Since

\[
r_{xx}[k] = \sigma_u^2 \sum_{n=0}^{\infty} h[n]h[n+k]
\]
the correlation time or effective duration of \( r_{xx}[k] \) is the same as the impulse response length. Hence, because the CRLB to be derived is based on (3D.2), the asymptotic CRLB will be a good approximation if the data record length is much greater than the correlation time.

To find the PDF of \( \mathbf{x} \) we use (3D.2), which is a transformation from \( \mathbf{u} = [u[0] \ldots u[N-1]]^T \) to \( \mathbf{x} = [x[0] x[1] \ldots x[N-1]]^T \) or

\[
\mathbf{x} = \begin{bmatrix}
h[0] & 0 & 0 & \cdots & 0 \\
h[1] & h[0] & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
h[N-1] & h[N-2] & h[N-3] & \cdots & h[0] \\
\end{bmatrix} \mathbf{u}.
\]

Note that \( \mathbf{H} \) has a determinant of \( (h[0])^N = 1 \) and hence is invertible. Since \( \mathbf{u} \sim \mathcal{N}(0, \sigma_u^2 \mathbf{I}) \), the PDF of \( \mathbf{x} \) is \( \mathcal{N}(0, \sigma_u^2 \mathbf{H} \mathbf{H}^T) \) or

\[
p(\mathbf{x}; \mathbf{\theta}) = \frac{1}{(2\pi)^{N/2} \det^{1/2}(\sigma_u^2 \mathbf{H} \mathbf{H}^T)} \exp \left[ -\frac{1}{2} \mathbf{x}^T(\sigma_u^2 \mathbf{H} \mathbf{H}^T)^{-1}\mathbf{x} \right].
\]

But

\[
\det(\sigma_u^2 \mathbf{H} \mathbf{H}^T) = \sigma_u^{2N} \det(\mathbf{H}) = \sigma_u^{2N}.
\]

Also,

\[
\mathbf{x}^T(\sigma_u^2 \mathbf{H} \mathbf{H}^T)^{-1}\mathbf{x} = \frac{1}{\sigma_u^2} (\mathbf{H}^{-1}\mathbf{x})^T(\mathbf{H}^{-1}\mathbf{x}) = \frac{1}{\sigma_u^2} \mathbf{u}^T \mathbf{u}
\]

so that

\[
p(\mathbf{x}; \mathbf{\theta}) = \frac{1}{(2\pi\sigma_u^2)^{N/2}} \exp \left( -\frac{1}{2\sigma_u^2} \mathbf{u}^T \mathbf{u} \right). \tag{3D.3}
\]

From (3D.2) we have approximately

\[
X(f) = H(f)U(f)
\]

where

\[
X(f) = \sum_{n=0}^{N-1} x[n] \exp(-j2\pi fn)
\]

\[
U(f) = \sum_{n=0}^{N-1} u[n] \exp(-j2\pi fn)
\]

are the Fourier transforms of the truncated sequences. By Parseval's theorem

\[
\frac{1}{\sigma_u^2} \mathbf{u}^T \mathbf{u} = \frac{1}{\sigma_u^2} \sum_{n=0}^{N-1} u^2[n]
\]
$$\ln \sigma_u^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \sigma_u^2 \, df$$
$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \left( \frac{P_{xx}(f)}{|H(f)|^2} \right) \, df$$
$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) \, df - \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln |H(f)|^2 \, df.$$

But

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln |H(f)|^2 \, df = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln H(f) + \ln H^*(f) \, df$$
$$= 2 \Re \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln H(f) \, df$$
$$= 2 \Re \int_C \ln \mathcal{H}(z) \frac{dz}{2\pi jz}$$
$$= 2 \Re [Z^{-1} \{\ln \mathcal{H}(z)\}]_{n=0}$$

where $C$ is the unit circle in the $z$ plane. Since $\mathcal{H}(z)$ corresponds to the system function of a causal filter, it converges outside a circle of radius $r < 1$ (since $\mathcal{H}(z)$ is assumed to exist on the unit circle for the frequency response to exist). Hence, $\ln \mathcal{H}(z)$ also converges outside a circle of radius $r < 1$, so that the corresponding sequence is causal. By the initial value theorem which is valid for a causal sequence

$$Z^{-1} \{\ln \mathcal{H}(z)\} = \lim_{z \to \infty} \ln \mathcal{H}(z)$$
$$= \ln \lim_{z \to \infty} \mathcal{H}(z)$$
$$= \ln h[0] = 0.$$

Therefore,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln |H(f)|^2 \, df = 0$$
and finally
\[ \ln \sigma_u^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) \, df. \] (3D.5)

Substituting (3D.4) and (3D.5) into (3D.3) produces for the log PDF
\[ \ln p(x; \theta) = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) \, df - \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{|X(f)|^2}{P_{xx}(f)} \, df. \]

Hence, the asymptotic log PDF is
\[ \ln p(x; \theta) = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \ln P_{xx}(f) + \frac{1}{N} \frac{|X(f)|^2}{P_{xx}(f)} \right] \, df. \] (3D.6)

To determine the CRLB
\[ \frac{\partial \ln p(x; \theta)}{\partial \theta_i} = -\frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left( \frac{1}{P_{xx}(f)} - \frac{1}{N} \frac{|X(f)|^2}{P_{xx}^2(f)} \right) \frac{\partial P_{xx}(f)}{\partial \theta_i} \, df \]
\[ + \left( -\frac{1}{P_{xx}^2(f)} + \frac{2}{N} \frac{|X(f)|^2}{P_{xx}^3(f)} \right) \frac{\partial P_{xx}(f)}{\partial \theta_i} \frac{\partial P_{xx}(f)}{\partial \theta_j} \, df. \] (3D.7)

In taking the expected value we encounter the term \( E(|X(f)|^2/N) \). For large \( N \) this is now shown to be \( P_{xx}(f) \). Note that \( |X(f)|^2/N \) is termed the periodogram spectral estimator.
\[ E \left( \frac{1}{N} |X(f)|^2 \right) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x[m]x[n] \exp[-j2\pi f(m-n)] \]
\[ = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} r_{xx}[m-n] \exp[-j2\pi f(m-n)] \]
\[ = \sum_{k=-(N-1)}^{N-1} \left( 1 - \frac{|k|}{N} \right) r_{xx}[k] \exp(-j2\pi fk) \] (3D.8)

where we have used the identity
\[ \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} g[m-n] = \sum_{k=-(N-1)}^{N-1} (N - |k|) g[k]. \]
**APPENDIX 3D. DERIVATION OF ASYMPTOTIC CRLB**

As \( N \to \infty \),

\[
\left( 1 - \frac{|k|}{N} \right) r_{xx}[k] \to r_{xx}[k]
\]

assuming that the ACF dies out sufficiently rapidly. Hence,

\[
E \left[ \frac{1}{N} |X(f)|^2 \right] \approx P_{xx}(f).
\]

Upon taking expectations in (3D.7), the first term is zero, and finally,

\[
[I(\theta)]_{ij} = \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{P_{xx}^2(f)} \frac{\partial P_{xx}(f)}{\partial \theta_i} \frac{\partial P_{xx}(f)}{\partial \theta_j} df
\]

\[
= \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial \ln P_{xx}(f)}{\partial \theta_i} \frac{\partial \ln P_{xx}(f)}{\partial \theta_j} df
\]

which is (3.34) without the explicit dependence of the PSD on \( \theta \) shown.