Estimation, Detection, and Identification

Graduate Course on the CMU/Portugal ECE PhD Program Spring 2008/2009

Chapter 3
Cramer-Rao Lower Bounds

Instructor:
Prof. Paulo Jorge Oliveira
pjcro @ isr.ist.utl.pt
Phone: +351 21 8418053 ou 2053 (inside IST)

INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa
Information and Communication Technologies Institute
Carnegie Mellon | PORTUGAL
AN INTERNATIONAL PARTNERSHIP

Syllabus:

Classical Estimation Theory

…

Chap. 2 - ***Minimum Variance Unbiased Estimation*** [1 week]

Unbiased estimators; Minimum Variance Criterion; Extension to vector parameters; Efficiency of estimators;

Chap. 3 - ***Cramer-Rao Lower Bound*** [1 week]
Estimator accuracy; Cramer-Rao lower bound (CRLB); CRLB for signals in white Gaussian noise; Examples;

Chap. 4 - ***Linear Models in the Presence of Stochastic Signals*** [1 week]

Stationary and transient analysis; White Gaussian noise and linear systems; Examples; Sufficient Statistics; Relation with MVU Estimators;

continues…

Estimator accuracy:

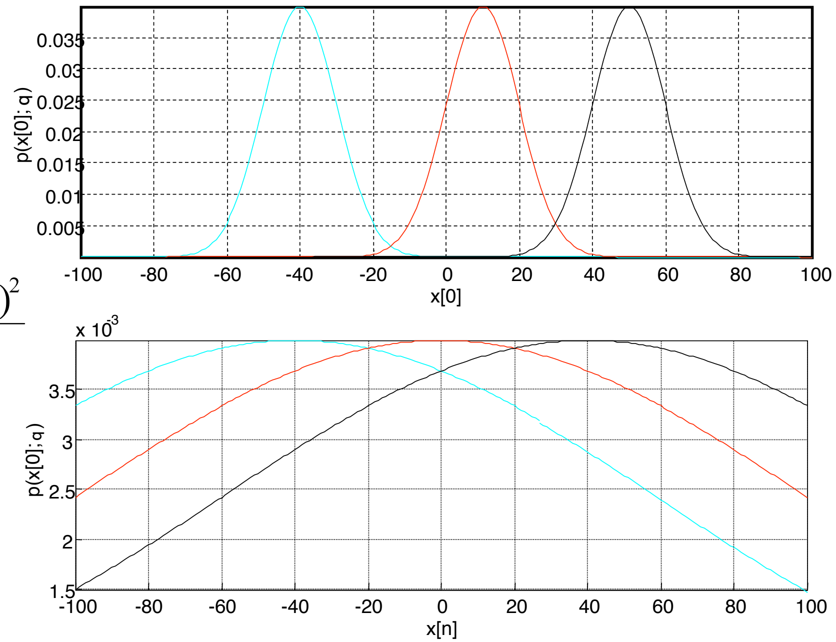
The accuracy on the estimates dependents very much on the PDFs

Example (revisited):

Model of signal $x[0] = A + w[0],$

Observation PDF $p\left(x[0];A\right) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x[0]-A)^2}{2\sigma^2}}$

for a disturbance $N(0,\sigma^2)$



Remarks:

If σ^2 is **Large** then the performance of the estimator is **Poor**;

If σ^2 is **Small** then the performance of the estimator is **Good**; or

If PDF concentration is **High** then the parameter accuracy is **High.**

How to measure sharpness of PDF (or concentration)?

Estimator accuracy:

When PDFs are seen as function of the unknown parameters, for x fixed, they are called as ***Likelihood function.*** To measure the sharpness note that (and ln is monotone…)

$$\ln p\left(x[0];A\right) = -\ln\sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2}\left(x[0]-A\right)^2$$

Its first and second derivatives are respectively:

$$\frac{\partial}{\partial A}\ln p\left(x[0];A\right) = \frac{1}{\sigma^2}\left(x[0]-A\right) \qquad \text{and} \qquad -\frac{\partial^2}{\partial A^2}\ln p\left(x[0];A\right) = \frac{1}{\sigma^2}.$$

As we know that the estimator $\hat{A}$ has variance σ^2 (at least for this example)

$$\operatorname{var}\left(\hat{A}\right) = \frac{1}{-\dfrac{\partial^2}{\partial A^2}\ln p\left(x;A\right)} = \frac{1}{curvature}$$

We are now ready to present an important theorem…

Cramer-Rao lower bound:

Theorem 3.1 (Cramer-Rao lower bound, scalar parameter) *– It is assumed that the PDF p(**x;**θ) satisfies the "regularity" condition*

$$E\left[\frac{\partial}{\partial\theta}\ln p(\mathbf{x};\theta)\right]=0 \qquad \text{for all} \quad \theta \tag{1}$$

*where the expectation is taken with respect to p(**x;**θ). Then, the variance of any unbiased estimator* $\hat{\theta}$ *must satisfy*

$$\operatorname{var}(\hat{\theta}) \geq \frac{1}{-E\left[\frac{\partial^2}{\partial\theta^2}\ln p(\mathbf{x};\theta)\right]} \tag{2}$$

*where the derivative is evaluated at the true value of θ and the expectation is taken with respect to p(**x**, θ). Furthermore, an unbiased estimator can be found that attains the bound for all θ if and only if*

$$\frac{\partial}{\partial\theta}\ln p(\mathbf{x};\theta)=I(\theta)\left(g(\mathbf{x})-\theta\right) \tag{3}$$

for some functions g(.) and I *(.). The estimator, which is the MVU estimator, is* $\hat{\theta}=g(\mathbf{x})$, *and the minimum variance 1/*$I(\theta)$.

Cramer-Rao lower bound:

Proof outline:

Lets derive the CRLB for a scalar parameter $\alpha = g(\theta)$. We consider all unbiased estimators

$$E\left[\hat{\alpha}\right] = \alpha = g(\theta) \qquad \text{or} \qquad \int \hat{\alpha}\, p(\boldsymbol{x};\theta)\; d\boldsymbol{x} = g(\theta). \qquad \text{(p.1)}$$

Lets examine the regularity condition (1)

$$E\left[\frac{\partial}{\partial\theta}\ln p(\boldsymbol{x};\theta)\right] = \int \frac{\partial \ln p(\boldsymbol{x};\theta)}{\partial\theta} p(\boldsymbol{x};\theta) d\boldsymbol{x} = \int \frac{\partial p(\boldsymbol{x};\theta)}{\partial\theta} d\boldsymbol{x}$$

$$= \frac{\partial}{\partial\theta}\int p(\boldsymbol{x};\theta) d\boldsymbol{x} = \frac{\partial 1}{\partial\theta} = 0.$$

Remark: differentiation and integration are required to be interchangeable (Leibniz Rule)!

Lets differentiate (p.1) with respect to θ and use the previous results

$$\int \hat{\alpha}\frac{\partial p(\boldsymbol{x};\theta)}{\partial\theta} d\boldsymbol{x} = \frac{\partial g(\theta)}{\partial\theta} \qquad or \qquad \int \hat{\alpha}\frac{\partial \ln p(\boldsymbol{x};\theta)}{\partial\theta} p(\boldsymbol{x};\theta) d\boldsymbol{x} = \frac{\partial g(\theta)}{\partial\theta} \quad .$$

Cramer-Rao lower bound:

Proof outline (cont.):

This can be modified to

$$\int (\alpha - \hat{\alpha}) \frac{\partial \ln p(\boldsymbol{x};\theta)}{\partial \theta} p(\boldsymbol{x};\theta) d\boldsymbol{x} = \frac{\partial g(\theta)}{\partial \theta},$$

as

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

Now applying the Cauchy-Schwarz inequality

$$\left[\int w(\boldsymbol{x}) g(\boldsymbol{x}) h(\boldsymbol{x}) d\boldsymbol{x}\right]^2 \leq \int w(\boldsymbol{x}) g^2(\boldsymbol{x}) d\boldsymbol{x} \int w(\boldsymbol{x}) h^2(\boldsymbol{x}) d\boldsymbol{x}$$

considering $w(\boldsymbol{x}) = p(\boldsymbol{x};\theta)$, $g(\boldsymbol{x}) = \hat{\alpha} - \alpha$, and $h(\boldsymbol{x}) = \frac{\partial \ln p(\boldsymbol{x};\theta)}{\partial \theta}$

results

$$\left(\frac{\partial g(\theta)}{\partial \theta}\right)^2 \leq \int (\alpha - \hat{\alpha})^2 p(\boldsymbol{x};\theta) d\mathbf{x} \int \left(\frac{\partial \ln p(\boldsymbol{x};\theta)}{\partial \theta}\right)^2 p(\boldsymbol{x};\theta) d\mathbf{x}$$

Cramer-Rao lower bound:

Proof outline (cont.):

It remains to relate this expression with the one in the Theorem $\int \left(\frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta} \right)^2 p(\mathbf{x};\theta) d\mathbf{x} = ?$

Starting with the previous result

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

thus, this function identically null verifies

$$\frac{\partial}{\partial\theta} \int \frac{\partial}{\partial\theta} \ln p(\mathbf{x};\theta)\, p(\mathbf{x};\theta)\, d\mathbf{x} = \int \left[\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial\theta^2} p(\mathbf{x};\theta) + \frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta} \frac{\partial p(\mathbf{x};\theta)}{\partial\theta} \right] d\mathbf{x} =$$

$$\int \left[\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial\theta^2} p(\mathbf{x};\theta) + \frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta} \frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta} p(\mathbf{x};\theta) \right] d\mathbf{x} = 0$$

And finally

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

Cramer-Rao lower bound:

Proof outline (cont.):

Taking this into consideration, i.e.

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

expression (2) results, in the case where $g(\theta)=\theta$.

The result (3) will be obtained next…

$\square$

See also appendix 3.B for the derivation in the vector case.

Cramer-Rao lower bound:

Summary:

- ***Being able to place a lower bound on the variance of any unbiased estimator is very useful.***
- ***It allow us to assert that an estimator is the MVU estimator (if it attains the bound for all values of the unknown parameter).***
- ***It provides in all cases a benchmark for the unbiased estimators that we can design.***
- ***It alerts to impossibility of finding unbiased estimators with variance lower than the bound.***
- ***Provides a systematic way of finding the MVU estimator, if it exists and if an extra condition is verified.***

Example:

Example (DC level in white Gaussian noise):

Problem: Find MVU estimator. **Approach:** Compute CRLB, if right form we have it.

Signal model: $x[n] = A + w[n], \qquad n = 0,\ldots,N-1, \qquad w[n] \sim N\left(0,\sigma^2\right)$

Likelihood function: $$p(\mathbf{x};A) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2}$$

$$\frac{\partial}{\partial A}\ln p(\mathbf{x};A) = \frac{\partial}{\partial A}\left(-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left(x[n]-A\right)^2\right) = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(x[n]-A\right) = \frac{N}{\sigma^2}\left(\bar{x}-A\right)$$

$$\frac{\partial^2}{\partial A^2}\ln p(\mathbf{x};A) = -\frac{N}{\sigma^2}$$

CRLB: $$\operatorname{var}\left(\hat{A}\right) \geq \frac{1}{N/\sigma^2} = \frac{\sigma^2}{N}$$

The estimator is unbiased and has the same variance, **thus it is a MVU** estimator! And it has the form:

$$\frac{\partial}{\partial A}\ln p(x;A) = I(\theta)\left(g(\boldsymbol{x})-\theta\right), \qquad \text{for} \quad I(\theta) = \frac{N}{\sigma^2} \qquad g(\boldsymbol{x}) = \bar{\mathbf{x}}.$$

Cramer-Rao lower bound:

Proof outline (second part of the theorem):

Still remains to prove that the CRLB is attained for the estimator $\hat{\theta} = g(\mathbf{x})$

$$\mathrm{var}\left(\hat{\theta}\right) = \frac{1}{I(\theta)}, \qquad \text{for} \quad I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \ln p(x;\theta)\right]$$

If

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

differentiation relative to the parameter gives

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

and then

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

i.e. the bound is attained. ■

INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Information and Communication Technologies Institute
Carnegie Mellon | PORTUGAL
AN INTERNATIONAL PARTNERSHIP

Example:

Example (phase estimation):

Signal model: $x[n] = A\cos(2\pi f_0 n + \phi) + w[n], \qquad n = 0,\dots,N-1$

A, f_0 known

Likelihood function: $p(\mathbf{x};\phi) = \dfrac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left(x[n]-A\cos(2\pi f_0 n+\phi)\right)^2}$

$$\frac{\partial}{\partial\phi}\ln p(\mathbf{x};\phi) = -\frac{A}{\sigma^2}\sum_{n=0}^{N-1}\left(x[n]\sin(2\pi f_0 n+\phi) - \frac{A}{2}\sin(4\pi f_0 n + 2\phi)\right)A$$

$$\frac{\partial^2}{\partial\phi^2}\ln p(\mathbf{x};\phi) = -\frac{A}{\sigma^2}\sum_{n=0}^{N-1}\left(x[n]\cos(2\pi f_0 n+\phi) - A\cos(4\pi f_0 n + 2\phi)\right)A$$

$$-E\left[\frac{\partial^2}{\partial\phi^2}\ln p(\mathbf{x};\phi)\right] = -\frac{A^2}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{1}{2} - \frac{1}{2}\cos(4\pi f_0 n + 2\phi)\right) \approx \frac{NA^2}{2\sigma^2}$$

Example:

Example (phase estimation cont.):

$$-E\left[\frac{\partial^2}{\partial\phi^2}\ln p(\mathbf{x}\,|\,\phi)\right] = -\frac{A^2}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{1}{2}-\frac{1}{2}\cos\left(4\pi f_0 n + 2\phi\right)\right) \approx \frac{NA^2}{2\sigma^2}$$

as $\sum_{n=0}^{N-1}\cos\left(4\pi f_0 n + 2\phi\right) \approx 0 \text{ for } f_0 \text{ not near 0 or 1/2.}$

$$\frac{1}{N}\sum_{n=0}^{N-1}\cos(\alpha n + \beta) = \frac{1}{N}\mathrm{Re}\left\{\sum_{n=0}^{N-1} e^{j(\alpha n+\beta)}\right\} = \ldots = \frac{\sin(N\alpha/2)}{N\sin(\alpha/2)}\cos\left(\alpha\frac{N-1}{2}+\beta\right)$$

for large *N*.

$$\mathrm{var}\left(\hat{\phi}\right) \geq \frac{2\sigma^2}{NA^2}$$

- Bound decreases as $SNR=A^2/2\sigma^2$ increases
- Bound decreases as *N* increases

Does an efficient estimator exists? Does a MVUE estimator exists?

Fisher information:

We define the Fisher Information (Matrix) as

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

Note:

- I*(q)* ≥*0*
- *It is additive for independent observations*

$$\ln p\left(\mathbf{x};\theta\right) = \ln \pi_{n=0}^{N-1} p\left(x[n];\theta\right) = \sum_{n=0}^{N-1} \ln p\left(x[n];\theta\right)$$

$$I\left(\theta\right) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p\left(\mathbf{x};\theta\right)\right] = -\sum_{n=0}^{N-1} E\left[\frac{\partial^2}{\partial\theta^2}\ln p\left(x\left[n\right];\theta\right)\right]$$

- *If identically distributed (same PDF for each x[n])*

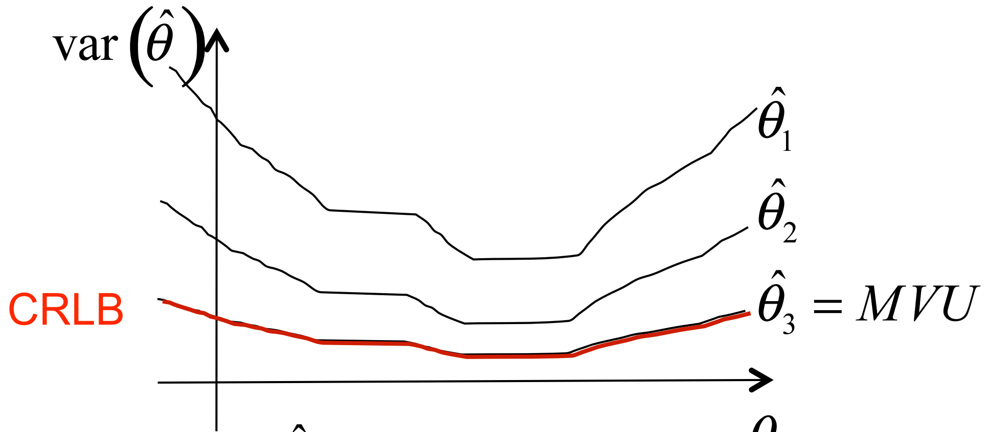
$$I\left(\theta\right) = N i\left(\theta\right) = -N\left[\frac{\partial^2}{\partial\theta^2}\ln\ p\left(x[.];\theta\right)\right]$$

As N->∞, for iid => CRLB-> 0

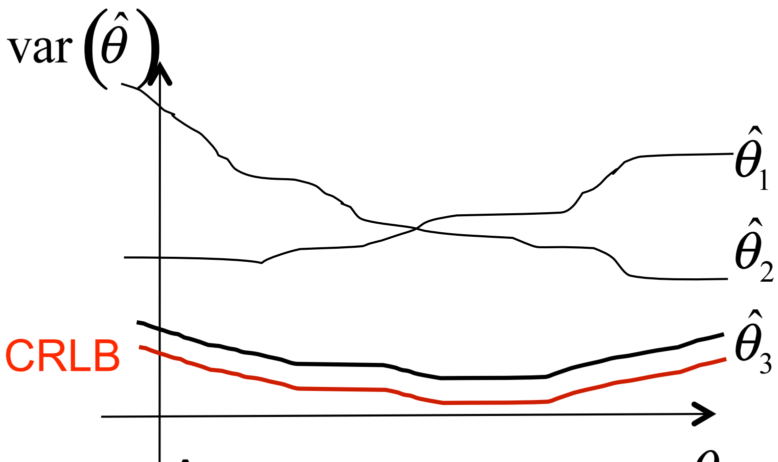
Other estimator characteristic:

Efficiency:

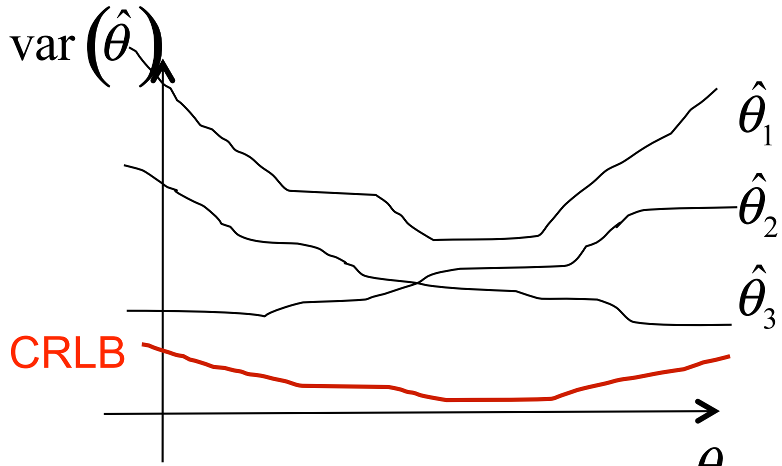
An estimator that is unbiased and attains the CRLB is said to be **efficient.**



$\hat{\theta}_3$ MVU and efficient



$\hat{\theta}_3$ MVU but not efficient



No MVU available.

Transformation of parameters:

Imagine that the CRLB is known for the parameter θ. Can we compute easily the CRLB for a linear transformation of the form $\alpha = g(\theta) = a\theta + b$?

$$\hat{\alpha} = a\hat{\theta} + b, \qquad E\left[a\hat{\theta} + b\right] = aE\left[\hat{\theta}\right] + b = \alpha$$

$$\operatorname{var}\left[a\hat{\theta} + b\right] = a^2 \operatorname{var}\left[\hat{\theta}\right] = \frac{\left(\dfrac{\partial g(\theta)}{\partial \theta}\right)^2}{-E\left[\dfrac{\partial^2}{\partial \theta^2} \ln p(\boldsymbol{x};\theta)\right]}$$

Linear transformations preserve biasness and efficiency.

And for a nonlinear transformation of the form $\alpha = g(\theta)$?

Transformation of parameters:

Remark: after a nonlinear transformation, the good properties can be lost.

$$\operatorname{var}\left(\hat{\theta}\right) \geq \frac{\left(\dfrac{\partial g(\theta)}{\partial \theta}\right)^2}{-E\left[\dfrac{\partial^2}{\partial \theta^2} \ln p(\boldsymbol{x};\theta)\right]}$$

Example: Suppose that given a stochastic variable $\overline{x} \sim N\left(A, \dfrac{\sigma^2}{N}\right)$ we desire to have an estimator for $\alpha=g(A)=A^2$ (power estimator). Note that

$$\operatorname{var}(\overline{\mathbf{x}}) = E\left[\left(\overline{\mathbf{x}} - E\left[(\overline{\mathbf{x}})\right]\right)^2\right] = E\left[\overline{\mathbf{x}}^2 - 2\overline{\mathbf{x}}E[\overline{\mathbf{x}}] + E^2[\overline{\mathbf{x}}]\right] = E\left[\overline{\mathbf{x}}^2\right] - 2E^2[\overline{\mathbf{x}}] + E^2[\overline{\mathbf{x}}] = E\left[\overline{\mathbf{x}}^2\right] - E^2[\overline{\mathbf{x}}]$$

$$E\left[\overline{\mathbf{x}}^2\right] = \operatorname{var}(\overline{\mathbf{x}}) + E^2[\overline{\mathbf{x}}]$$

A bias estimate results. Efficiency is lost.

Cramer-Rao lower bound:

Theorem 3.1 (Cramer-Rao lower bound, Vector parameter) *– It is assumed that the PDF p(**x;θ**) satisfies the "regularity" condition*

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

*where the expectation is taken with respect to p(**x, θ**). Then, the variance of any unbiased estimator* $\hat{\theta}$ *must satisfy*

$$\mathbf{C}_{\hat{\theta}} - I^{-1}(\boldsymbol{\theta}) \geq 0,$$

where ≥ is interpreted as meaning the matrix is positive semi-definite. The Fisher information matrix I*(**θ**) is given as*

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

where the derivatives are evaluated at the true value of θ and the expectation is taken with respect to p*(**x**;θ). Furthermore, an unbiased estimator may be found that attains the bound for all **θ** if and only if*

$$\frac{\partial}{\partial\theta}\ln p(\mathbf{x};\boldsymbol{\theta}) = I(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x})-\boldsymbol{\theta}) \qquad (3)$$

*for some functions p dimensional function **g(.)** and some p x p matrix **I (.)**. The estimator, which is the MVU estimator, is* $\hat{\boldsymbol{\theta}}=\mathbf{g}(x)$*, and its covariance matrix is **I⁻¹(θ)**.*

Vector Transformation of parameters:

The vector transformation of parameters $\boldsymbol{\alpha} = \mathbf{g}(\boldsymbol{\theta})$ impacts on the CRLB computation as

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

where the Jacobian is

$$\frac{\partial \mathbf{g}(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial g_1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial g_1(\theta)}{\partial \theta_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_r(\theta)}{\partial \theta_1} & & \frac{\partial g_r(\theta)}{\partial \theta_p} \end{bmatrix}$$

In the Gaussian general case for $x[n]=s[n]+w[n]$, where $\mathbf{w} \sim N\left(\mu(\theta), \mathbf{C}_\theta\right)$

the Fisher information matrix is

$$\left[I(\boldsymbol{\theta})\right]_{ij} = \left[\frac{\partial \mu(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i}\right]^T \mathbf{C}^{-1}(\boldsymbol{\theta}) \left[\frac{\partial \mu(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i}\right] \mathbf{C}_{\hat{\alpha}} + \frac{1}{2} tr\left[\mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial C(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i} \mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial C(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}\right].$$

Example:

Example (line fitting):

Signal model: $x[n] = A + Bn + w[n], \qquad n = 0,\dots,N-1$

A, B deterministic unknown quantities

Likelihood function: $p(\mathbf{x};\theta) = \dfrac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A-Bn)^2}, \qquad where \quad \theta = \begin{bmatrix} A & B \end{bmatrix}^T$

The Fisher Information Matrix is

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

where

$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial A} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}(x[n]-A-Bn), \qquad \text{and} \qquad \frac{\partial \ln p(\mathbf{x};\theta)}{\partial B} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}(x[n]-A-B)n.$$

Example:

Example (cont.):

Moreover

$$\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial A^2} = -\frac{N}{\sigma^2}, \frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial A \partial B} = -\frac{1}{\sigma^2}\sum_{n=0}^{N-1} n, \qquad \text{and} \qquad \frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial B^2} = -\frac{1}{\sigma^2}\sum_{n=0}^{N-1} n^2.$$

Since the second order derivatives do not depend on x, we have immediately that

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

And also,

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

Example:

Example (cont.):

Remarks:

For only one parameter to be determined $\mathrm{var}\left(\hat{A}\right) \geq \frac{\sigma^2}{N}$. Thus a general results was obtained: ***when more parameters are to be estimated the CRLB always degrades.***

Moreover

$$\frac{CRLB(\hat{A})}{CRLB(\hat{B})} = \frac{(2N-1)(N-1)}{6} > 1, \quad \text{for} \quad N \geq 3.$$

The parameter *B* is easier to be determined, as its CRLB decreases with $1/N^3$. This means that *x[n]* is more sensitive to changes in *B* than changes in *A*.

$$\Delta x[n] \approx \frac{\partial x[n]}{\partial A} \Delta A = \Delta A$$

$$\Delta x[n] \approx \frac{\partial x[n]}{\partial B} \Delta B = n\Delta B.$$

Bibliography:

Further reading

• Harry L. Van Trees, ***Detection, Estimation, and Modulation Theory, Parts I to IV,*** John Wiley, 2001.

• J. Bibby, H. Toutenburg, ***Prediction and Improved Estimation in Linear Models,*** John Wiley, 1977.

• C.Rao, ***Linear Statistical Inference and Its Applications,*** John Wiley, 1973.

• P. Stoica, R. Moses, *"On Biased Estimators and the Unbiased Cramer-Rao Lower Bound," Signal Process,* vol.21, pp. 349-350, 1990.