Alternate Form for CRLB

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

See Appendix 3A for Derivation

Sometimes it is easier to find the CRLB this way.

This also gives a new viewpoint of the CRLB:

Posted on BB

From Gardner's Paper (*IEEE Trans. on Info Theory*, July 1979)

Consider the Normalized version of this form of CRLB

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

We'll "derive" this in a way that will re-interpret the CRLB

Consider the "Incremental Sensitivity" of $p(\mathbf{x};\theta)$ to changes in θ:

If $\theta \rightarrow \theta + \Delta\theta$, then it causes $p(\mathbf{x};\theta) \rightarrow p(\mathbf{x};\theta + \Delta\theta)$

How sensitive is $p(\mathbf{x};\theta)$ to that change??

$$\tilde{S}_\theta^p(\mathbf{x}) \stackrel{\Delta}{=} \frac{\left[\dfrac{\Delta p(\mathbf{x};\theta)}{p(\mathbf{x};\theta)}\right]}{\left[\dfrac{\Delta\theta}{\theta}\right]} = \frac{\%\text{ change in } p(\mathbf{x};\theta)}{\%\text{ change in } \theta} = \left[\frac{\Delta p(\mathbf{x};\theta)}{\Delta\theta}\right]\left[\frac{\theta}{p(\mathbf{x};\theta)}\right]$$

Now let $\Delta\theta \rightarrow 0$: $S_\theta^p(\mathbf{x}) = \lim_{\Delta\theta \rightarrow 0} \tilde{S}_\theta^p(\mathbf{x}) = \left[\dfrac{\partial p(\mathbf{x};\theta)}{\partial\theta}\right]\left[\dfrac{\theta}{p(\mathbf{x};\theta)}\right] = \theta \dfrac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta}$

Recall from Calculus: $\dfrac{\partial \ln f(x)}{\partial x} = \dfrac{1}{f(x)} \dfrac{\partial f(x)}{\partial x}$

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

Interpretation
Norm. CRLB = Inverse Mean Square Sensitivity

Definition of Fisher Information

The denominator in CRLB is called the Fisher Information $I(\theta)$

It is a measure of the "expected goodness" of the data for the purpose of making an estimate

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

Has the needed properties for "info" (as does "Shannon Info"):

1. $I(\theta) \geq 0$ (easy to see using the alternate form of CRLB)
2. $I(\theta)$ is additive for independent observations

   follows from: $\ln p(\mathbf{x};\theta) = \ln\left[\prod_n p(x[n];\theta)\right] = \sum_n \ln\left[p(x[n];\theta)\right]$

If each $I_n(\theta)$ is the same: $I(\theta) = N \times I(\theta)$

3.5 CRLB for Signals in AWGN

When we have the case that our data is "signal + AWGN" then we get a simple form for the CRLB:

Signal Model: $x[n] = s[n;\theta] + w[n], \quad n = 0, 1, 2, \ldots, N\text{-}1$

White, Gaussian, Zero Mean

Q: What is the CRLB?

First write the likelihood function:

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

Differentiate Log LF twice to get:

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

Depends on random $x[n]$ so must take $E\{\}$

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

$$= \frac{-\sum_{n=0}^{N-1}\left[\frac{\partial s[n;\theta]}{\partial\theta}\right]^2}{\sigma^2}$$

Then using this we get the **CRLB for Signal in AWGN**:

$$\operatorname{var}(\hat{\theta}) \geq \frac{\sigma^2}{\sum_{n=0}^{N-1}\left[\frac{\partial s[n;\theta]}{\partial\theta}\right]^2}$$

Note: $\left[\frac{\partial s[n;\theta]}{\partial\theta}\right]^2$ tells how sensitive signal is to parameter

If signal is very sensitive to parameter change… then CRLB is small … can get very accurate estimate!

Ex. 3.5: CRLB of Frequency of Sinusoid

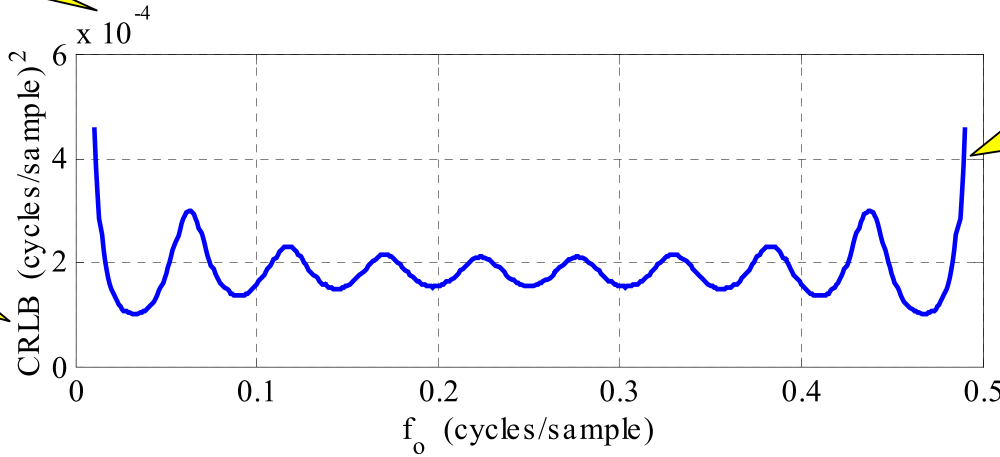
Signal Model: $x[n] = A\cos(2\pi f_o n + \phi) + w[n] \quad 0 < f_o < \frac{1}{2} \quad n = 0, 1, 2, \ldots, N-1$

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{SNR \times \sum_{n=0}^{N-1} \left[2\pi n \sin(2\pi f_o n + \phi)\right]^2}$$

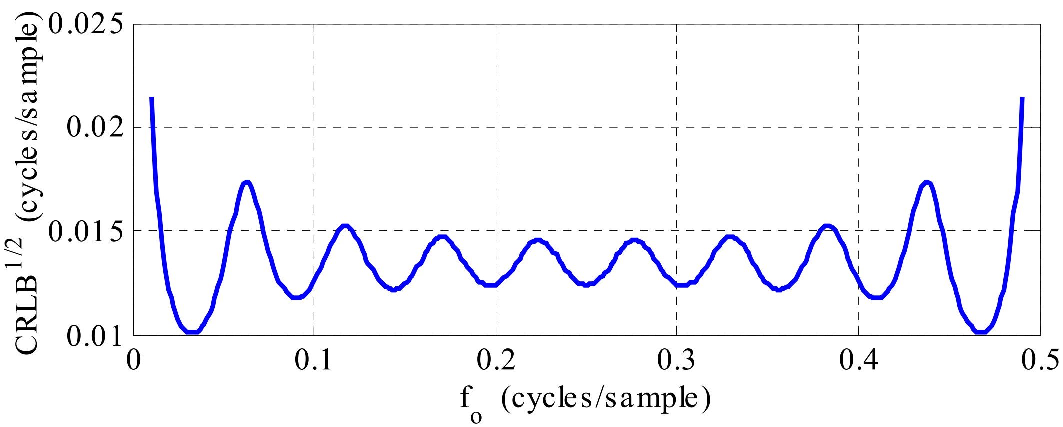
Error in Book

Signal is less sensitive if f_o near 0 or ½

Bound on Variance



Bound on Std. Dev.



3.6 Transformation of Parameters

Say there is a parameter θ with known CRLB_θ

But imagine that we instead are interested in estimating some other parameter α that is a function of θ:

$$\alpha = g(\theta)$$

Q: What is CRLB_α?

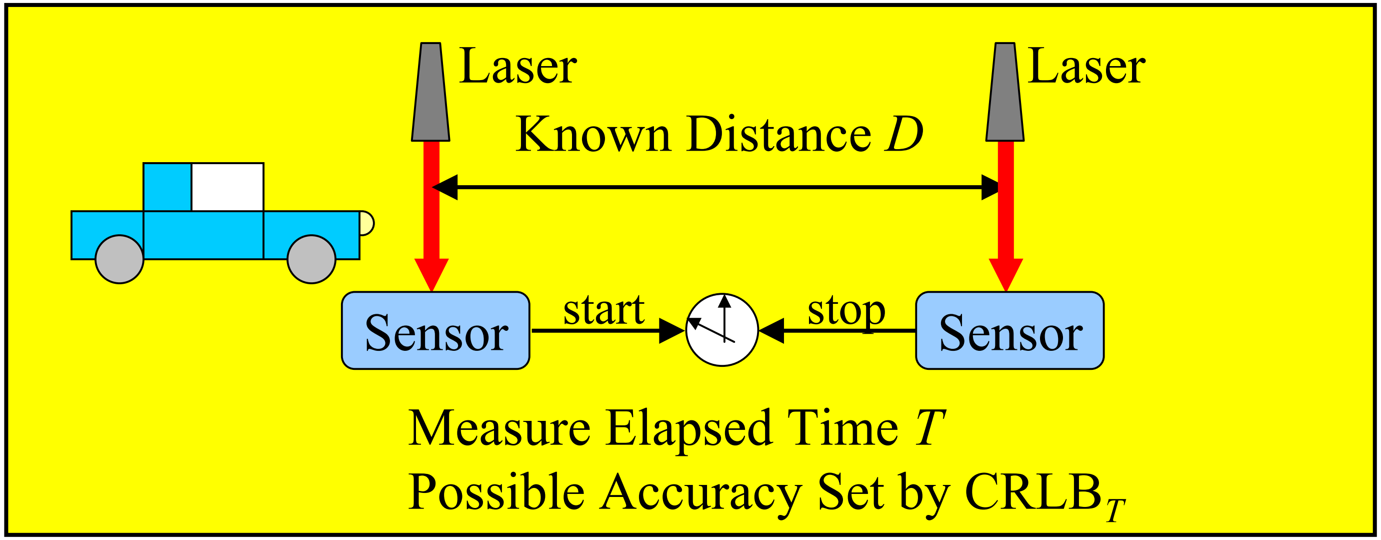
$$\text{var}(\alpha) \geq CRLB_\alpha = \left(\frac{\partial g(\theta)}{\partial \theta}\right)^2 CRLB_\theta$$

Proved in Appendix 3B

Captures the sensitivity of α to θ

Large $\partial g/\partial \theta$ → small error in θ gives larger error in α
→ increases CRLB (i.e., worsens accuracy)

Example: Speed of Vehicle From Elapsed Time



But… really want to measure speed $V = d/T$

Find the CRLB_V:

$$CRLB_V = \left[\frac{\partial}{\partial T}\left(\frac{D}{T}\right)\right]^2 \times CRLB_T$$

$$= \left(-\frac{D}{T}\right)^2 \times CRLB_T$$

$$= \frac{V^4}{D^2} \times CRLB_T$$

Accuracy Bound

$$\sigma_V \geq \frac{V^2}{D}\sqrt{CRLB_T} \quad (m/s)$$

- Less accurate at High Speeds (quadratic)
- More accurate over large distances

Effect of Transformation on Efficiency

Suppose you have an efficient estimator of θ: $\hat{\theta}$

But… you are really interested in estimating $\alpha = g(\theta)$

Suppose you plan to use $\hat{\alpha} = g(\hat{\theta})$

Q: Is this an efficient estimator of α ???

A: **Theorem**: If $g(\theta)$ has form $g(\theta) = \underbrace{a\theta + b}_{\text{"affine" transform}}$, then $\hat{\alpha} = g(\hat{\theta})$ is efficient.

Proof:

First: $\operatorname{var}(\hat{\alpha}) = \operatorname{var}(a\hat{\theta} + b) = a^2 \operatorname{var}(\hat{\theta}) = a^2 CRLB_\theta$

"=" because "efficient"

Now, what is CRB_α ? Using transformation result:

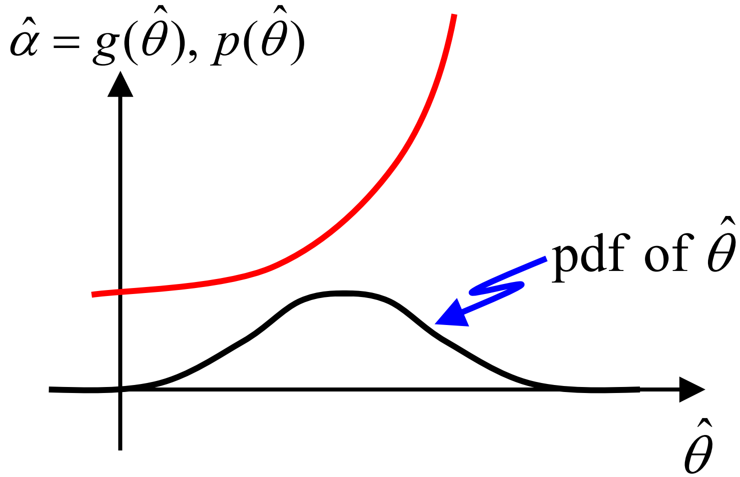
$$CRLB_\alpha = \underbrace{\left[\frac{\partial(a\theta + b)}{\partial\theta}\right]^2}_{=a^2} CRLB_\theta = a^2 CRLB_\theta \quad \Rightarrow \quad \boxed{\operatorname{var}(\hat{\alpha}) = CRLB_\alpha}$$

Efficient!

Asymptotic Efficiency Under Transformation

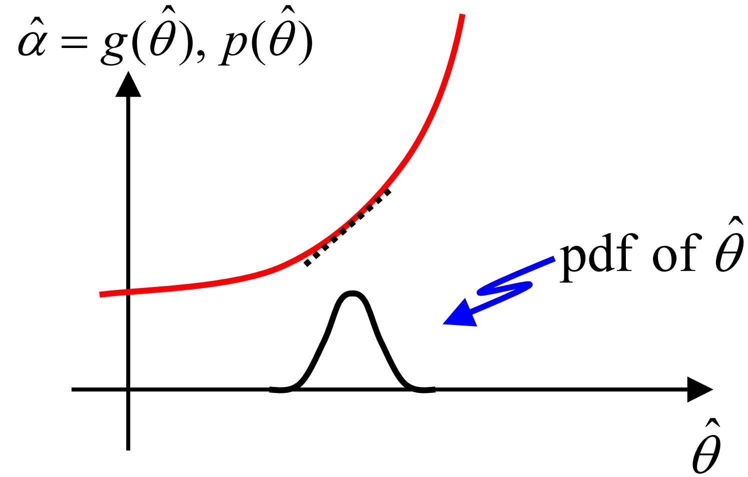
If the mapping $\alpha = g(\theta)$ is <u>not affine</u>… this result <u>does NOT hold</u>

<u>But</u>… if the number of data samples used is large, then the estimator is approximately efficient ("**<u>Asymptotically Efficient</u>**")



<u>Small *N* Case</u>

PDF is widely spread over nonlinear mapping



<u>Large *N* Case</u>

PDF is concentrated onto linearized section