Chapter 3
Cramer-Rao Lower Bound

What is the Cramer-Rao Lower Bound

Abbreviated: CRLB or sometimes just CRB

CRLB is a lower bound on the variance of any ***unbiased*** estimator:

If $\hat{\theta}$ is an unbiased estimator of θ, then

$$\sigma^2_{\hat{\theta}}(\theta) \geq CRLB_{\hat{\theta}}(\theta) \quad \Rightarrow \quad \sigma_{\hat{\theta}}(\theta) \geq \sqrt{CRLB_{\hat{\theta}}(\theta)}$$

The CRLB tells us the best we can ever expect to be able to do (w/ an unbiased estimator)

Some Uses of the CRLB

1. Feasibility studies (e.g. Sensor usefulness, etc.)
    - Can we meet our specifications?

2. Judgment of proposed estimators
    - Estimators that don't achieve CRLB are looked down upon in the technical literature

3. Can sometimes provide form for MVU est.

4. Demonstrates importance of physical and/or signal parameters to the estimation problem

   e.g. We'll see that a signal's BW determines delay est. accuracy
   $\Rightarrow$ Radars should use wide BW signals

3.3 Est. Accuracy Consideration

Q: What determines how well you can estimate θ?

Recall: Data vector is $\mathbf{x}$

samples from a random process that <u>depends</u> on an θ

$\Rightarrow$ the PDF describes that dependence: $p(\mathbf{x};\theta)$

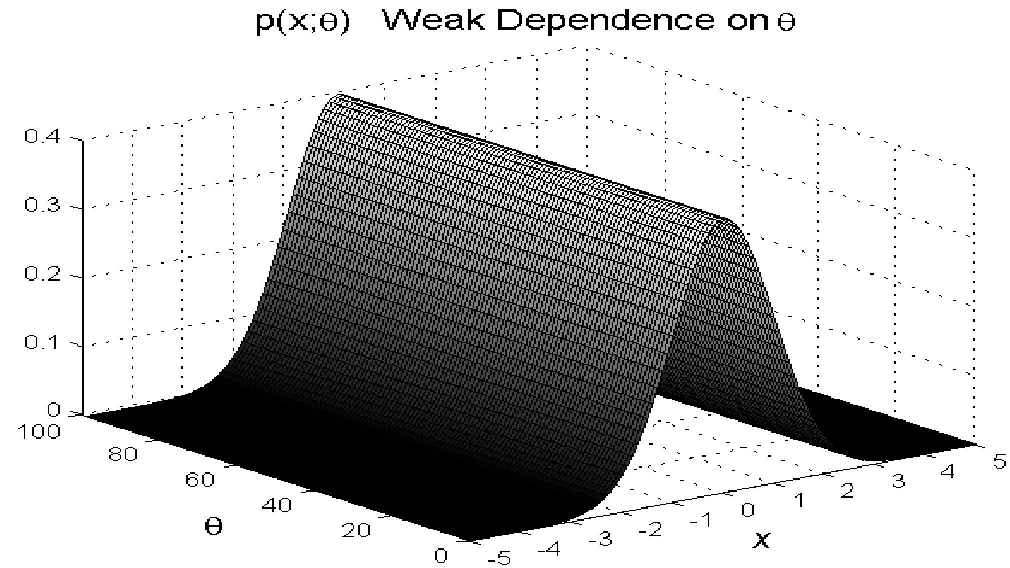
Clearly if $p(\mathbf{x};\theta)$ depends strongly/weakly on θ
…we should be able to estimate θ well/poorly.

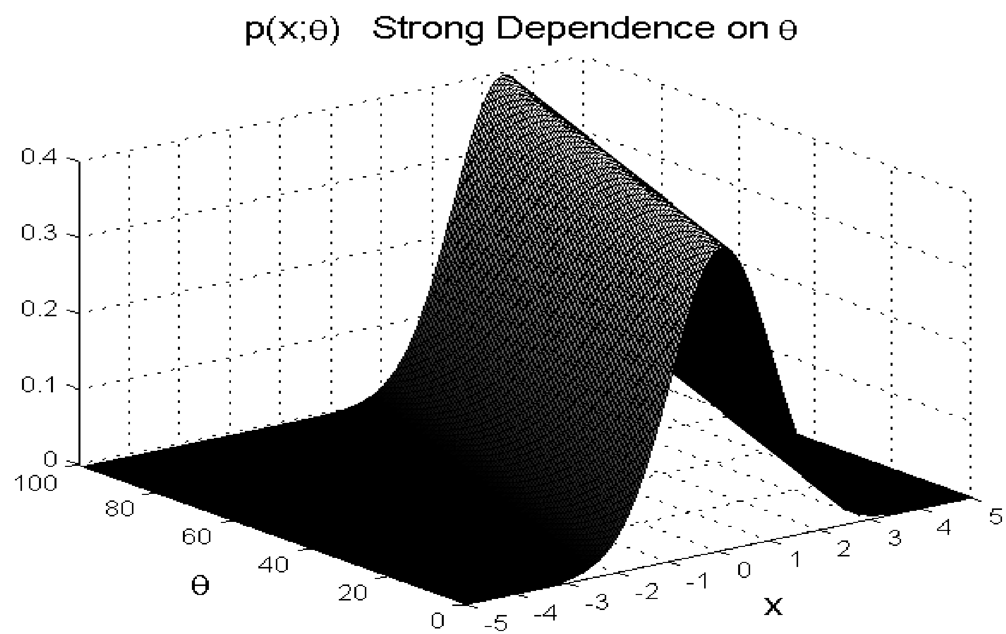
See surface plots vs. $\mathbf{x}$ & θ for 2 cases:

1. Strong dependence on θ
2. Weak dependence on θ

$\Rightarrow$ Should look at $p(\mathbf{x};\theta)$ as a <u>function of θ</u> for <u>fixed</u> value of <u>observed data $\mathbf{x}$</u>

Surface Plot Examples of $p(\mathrm{x};\theta)$





Ex. 3.1: PDF Dependence for DC Level in Noise

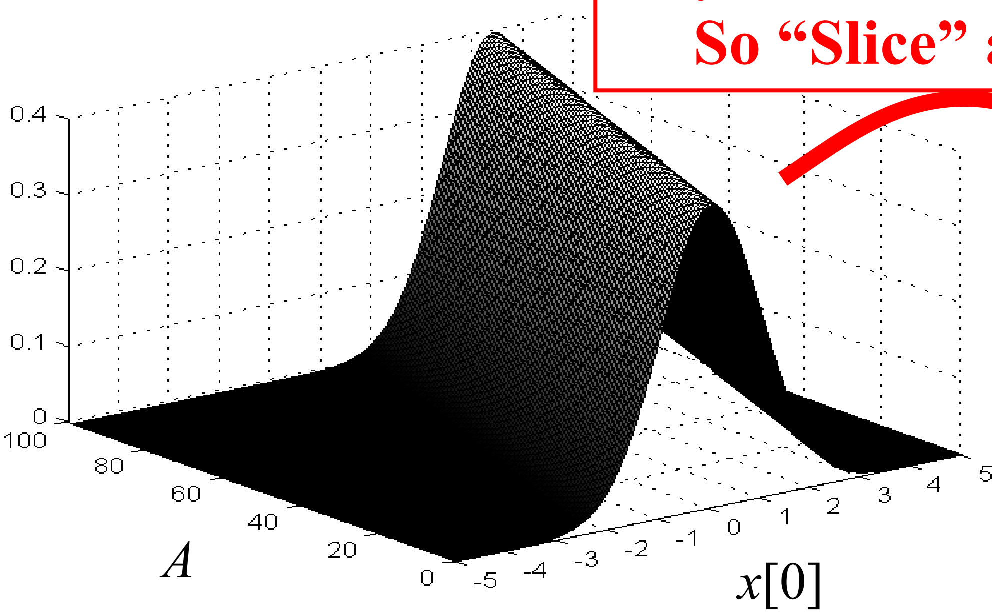
$$x[0] = A + w[0]$$

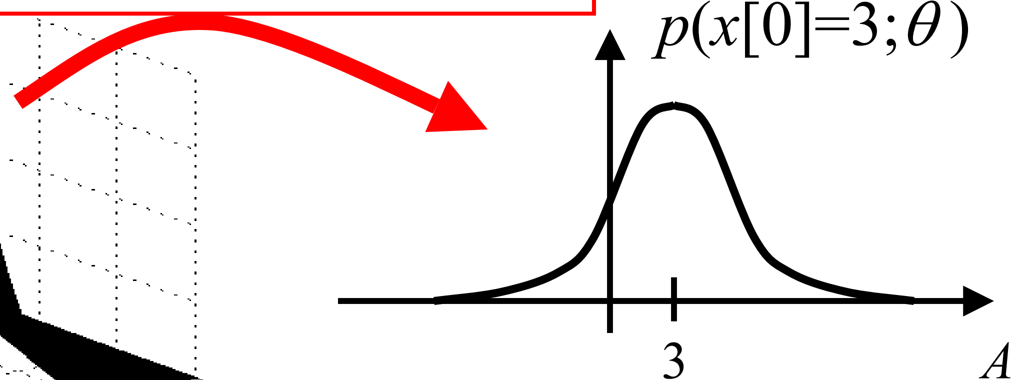
$w[0] \sim N(0,\sigma^2)$

Then the parameter-dependent PDF of the data point $x[0]$ is:

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

Say we observe $x[0] = 3$…
So "Slice" at $x[0] = 3$





Define: Likelihood Function (LF)

The LF = the PDF $p(\mathbf{x};\theta)$

…but as a function of parameter θ w/ the data vector **x** fixed

We will also often need the **Log Likelihood Function (LLF)**:

$$\text{LLF} = \ln\{\text{LF}\} = \ln\{\, p(\mathbf{x};\theta)\}$$

LF Characteristics that Affect Accuracy

Intuitively: "sharpness" of the LF sets accuracy… But How???

Sharpness is measured using curvature:

$$\left. \frac{-\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta^2} \right|_{\substack{\mathbf{x} = \text{given data} \\ \theta = \text{true value}}}$$

Curvature ↑ ⇒ PDF concentration ↑ ⇒ Accuracy ↑

But this is for a particular set of data… we want "in general":
So…Average over random vector to give the average curvature:

$$\left. -E\left\{ \frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta^2} \right\} \right|_{\theta = \text{true value}}$$

"Expected sharpness of LF"

$E\{\bullet\}$ is w.r.t $p(\mathbf{x};\theta)$

3.4 Cramer-Rao Lower Bound

Theorem 3.1 CRLB for Scalar Parameter

Assume "regularity" condition is met: $E\left\{\frac{\partial \ln p(x;\theta)}{\partial \theta}\right\} = 0 \quad \forall \theta$

Then

$$\sigma^2_{\hat{\theta}} \geq \frac{1}{-E\left\{\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta^2}\right\}\Bigg|_{\theta = \text{true value}}}$$

Right-Hand Side is CRLB

$E\{\bullet\}$ is w.r.t $p(\mathbf{x};\theta)$

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

Steps to Find the CRLB

1. Write log 1ikelihood function as a function of θ:
   - $\ln p(\mathbf{x};\theta)$

2. Fix **x** and take 2nd partial of LLF:
   - $\partial^2 \ln p(\mathbf{x};\theta)/\partial\theta^2$

3. If result still depends on **x**:
   - Fix θ and take expected value w.r.t. **x**
   - Otherwise skip this step

4. Result may still depend on θ:
   - Evaluate at each specific value of θ desired.

5. Negate and form reciprocal

Example 3.3 CRLB for DC in AWGN

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

$w[n] \sim N(0,\sigma^2)$ & white

Need likelihood function:

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

Due to whiteness

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

Property of exp

Now take ln to get LLF:

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

Now take first partial w.r.t. A:

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

sample mean

Now take partial again:

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

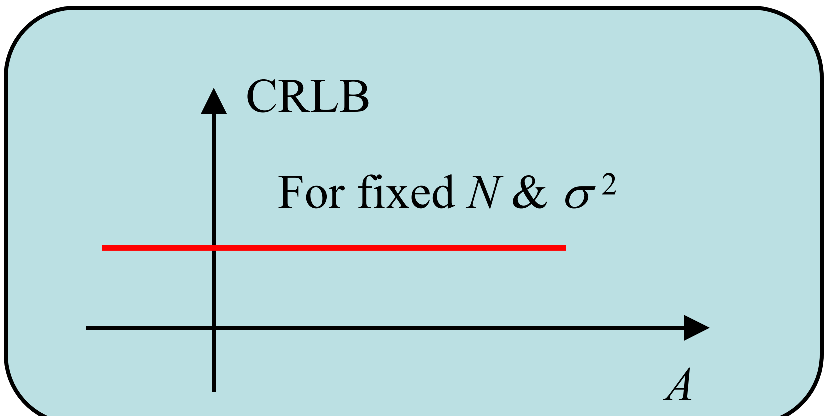
Doesn't depend on **x** so we don't need to do E{•}

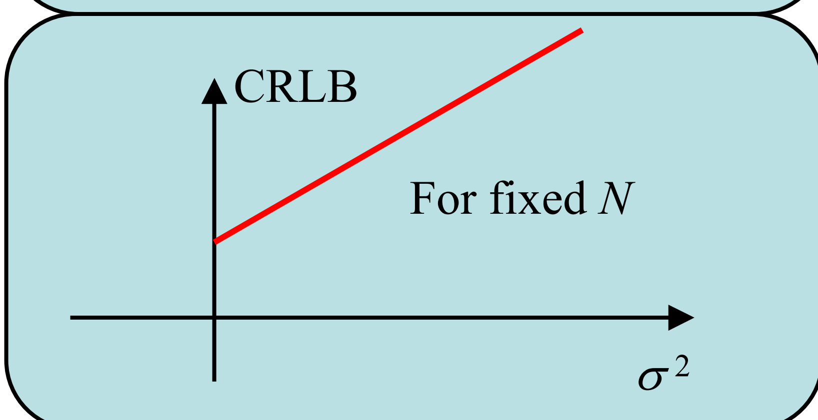
Since the result doesn't depend on **x** or A all we do is negate and form reciprocal to get CRLB:

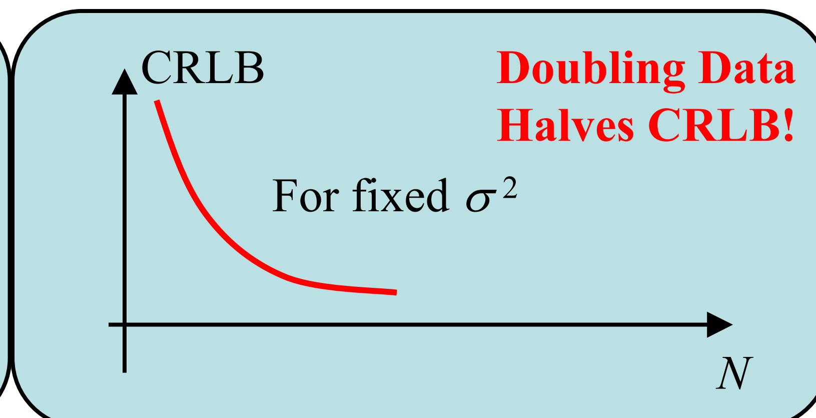
$$CRLB = \frac{1}{-E\left\{\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta^2}\right\}\Bigg|_{\theta = \text{true value}}} = \frac{\sigma^2}{N}$$

$$\Rightarrow \quad \operatorname{var}\{\hat{A}\} \geq \frac{\sigma^2}{N}$$

- **Doesn't depend on A**
- **Increases linearly with σ^2**
- **Decreases inversely with N**







Continuation of Theorem 3.1 on CRLB

There exists an unbiased estimator that **attains** the CRLB **iff**:

$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} = I(\theta)\left[g(\mathbf{x}) - \theta\right] \qquad (\blacktriangle)$$

for some functions $I(\theta)$ and $g(\mathbf{x})$

Furthermore, the estimator that achieves the CRLB is then given by:

$$\hat{\theta} = g(\mathbf{x})$$

$$\text{var}\{\hat{\theta}\} = \frac{1}{I(\theta)} = CRLB$$

Since no unbiased estimator can do better… this is the MVU estimate!!

This gives a possible way to find the MVU:

- Compute $\partial \ln p(\mathbf{x};\theta)/\partial\theta$ (need to anyway)
- Check to see if it can be put in form like ($\blacktriangle$)
- If so… then $g(\mathbf{x})$ is the MVU esimator

Revisit Example 3.3 to Find MVU Estimate

For DC Level in AWGN we found in (✵) that:

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

Has form of $I(A)[\mathrm{g}(\mathbf{x}) - A]$

$$I(A) = \frac{N}{\sigma^2} \Rightarrow \operatorname{var}\{\hat{A}\} = \frac{\sigma^2}{N} = CRLB$$

$$\hat{\theta} = g(\mathbf{x}) = \bar{x} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$$

So… for the DC Level in AWGN:
the sample mean is the MVUE!!

Definition: Efficient Estimator

An estimator that is:

- unbiased and
- attains the CRLB

is said to be an "Efficient Estimator"

Notes:

- Not all estimators are efficient (see next example: Phase Est.)
- Not even all MVU estimators are efficient

So… there are times when our "1st partial test" won't work!!!!

Example 3.4: CRLB for Phase Estimation

This is related to the DSB carrier estimation problem we used for motivation in the notes for Ch. 1

Except here… we have a pure sinusoid and we only wish to estimate only its phase

Signal Model:

$$x[n] = \underbrace{A\cos(2\pi f_o n + \phi_o)}_{s[n;\phi_o]} + w[n]$$

AWGN w/ zero mean & σ^2

Signal-to-Noise Ratio:

Signal Power = $A^2/2$

Noise Power = σ^2

$$SNR = \frac{A^2}{2\sigma^2}$$

Assumptions:

1. $0 < f_o < ½$ (f_o is in cycles/sample)
2. A and f_o are known (we'll remove this assumption later)

Problem: Find the CRLB for estimating the phase.

We need the PDF:

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

Exploit Whiteness and Exp. Form

Now taking the log gets rid of the exponential, then taking partial derivative gives (see book for details):

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

Taking partial derivative again:

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

Still depends on random vector **x**… so need E{}

Taking the expected value:

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

$$= \frac{A}{\sigma^2}\sum_{n=0}^{N-1}\left(E\{x[n]\}\cos(2\pi f_o n + \phi) - A\cos(4\pi f_o n + 2\phi)\right)$$

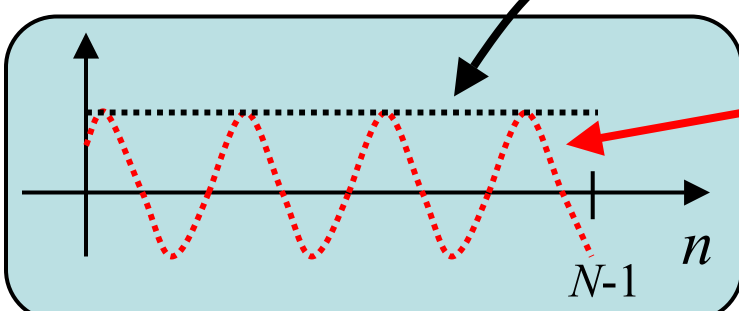
$E\{x[n]\} = A\cos(2\pi f_o n + \phi)$

So… plug that in, get a $\cos^2$ term, use trig identity, and get

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

$\sum_{n=0}^{N-1} 1 = N$

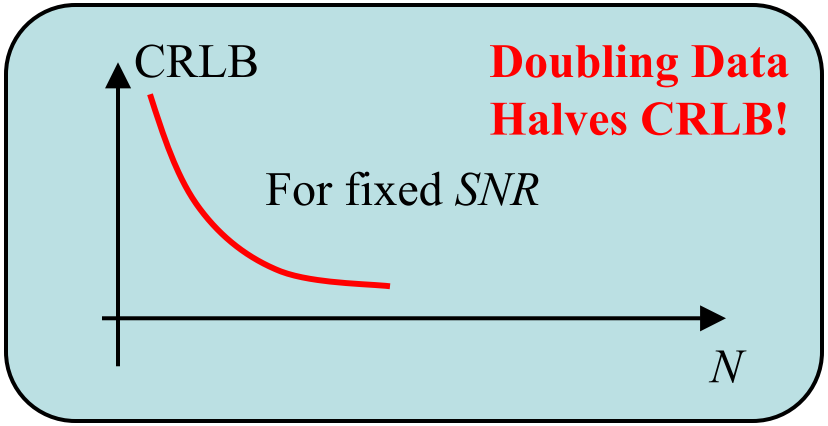
$\sum_{n-0}^{N-1}\cos(4\pi f_o n + 2\phi) \ll N$ if f_o not near 0 or ½

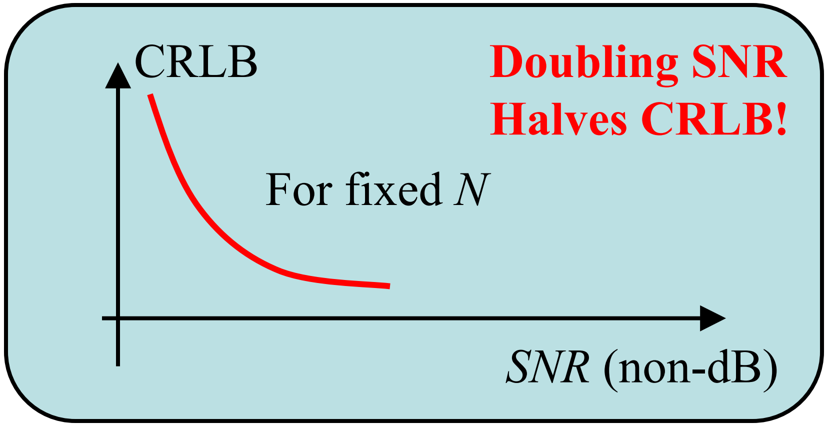


Now… invert to get **CRLB**:

$$\operatorname{var}\{\hat{\phi}\} \geq \frac{1}{N \times SNR}$$

Non-dB





Halve CRLB for every 3B in SNR

Does an efficient estimator exist for this problem? The CRLB theorem says there is only if

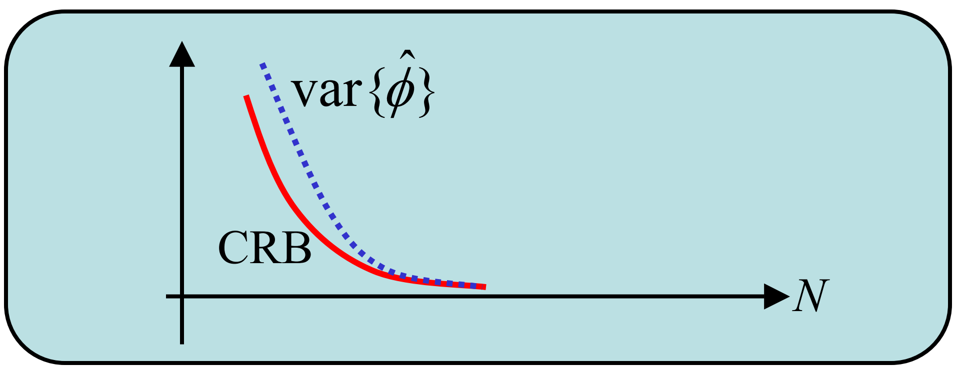
$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} = I(\theta)\left[g(\mathbf{x}) - \theta\right]$$

Our earlier result was:

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

Efficient Estimator does NOT exist!!!

We'll see later though, an estimator for which $\text{var}\{\hat{\phi}\} \to CRLB$ as $N \to \infty$ or as $SNR \to \infty$



Such an estimator is called an "asymptotically efficient" estimator
(We'll see such a phase estimator in Ch. 7 on MLE)