Estimation: chapter 2+3

Minimum variance unbiased estimation + the CRLB

Lifang Feng
lffeng@ustb.edu.cn

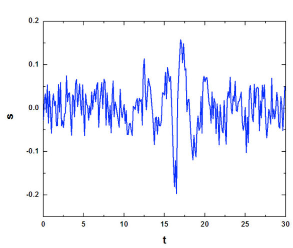
University of Science and Technology Beijing

Spring 2022

Estimation: a first example

- Estimate the DC level, A, of a signal given noisy measurements x[0], x[1], ... x[N-1] where

$$x[n] = A + w[n], \quad n = 0, 1, 2, \cdots N-1$$



x[n] are samples of this!

- Find a few estimators
- Compare their performance
  - mean?
  - variance?
  - pdf?

Estimation: a first example

- Estimators of the DC level, A

$$x[n] = A + w[n], \quad n = 0, 1, 2, \cdots N-1$$

- $\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$
- $\check{A} = x[0]$
- $\dot{A} = \frac{1}{N+2}\left(2x[0] + \sum_{n=1}^{N-2} x(n) + 2x[N-1]\right)$
- $\bar{A} = \frac{1}{2N}\sum_{n=0}^{N-1} x[n]$

Estimation: definitions

- Parameter: we wish to estimate the parameter θ from the observation(s) $\mathbf{x}$. These can be vectors $\theta = [\theta_1, \theta_2, \cdots, \theta_p]^T$ and $\mathbf{x} = [x[0], x[1], \cdots x[N-1]]^T$ or scalars.

- Parametrized PDF: the unknown parameter θ is to be estimated. θ parametrizes the probability density function of the received data $p(\mathbf{x};\theta)$. When dealing with Bayesian estimators, the notation $p(\mathbf{x}|\theta)$ will be used to highlight the fact that θ is a random variable.

- Estimator: a rule that assigns a value $\hat{\theta}$ to θ for each realization of $\mathbf{x}$.

- Estimate: the value of θ obtained for a given realization of $\mathbf{x}$. $\hat{\theta}$ will be used for the estimate, while θ will represent the true value of the unknown parameter.

Estimation: definitions

- Bias: an estimator $\hat{\theta}$ is called unbiased if $E(\hat{\theta}) = \theta$ for all possible θ. If this is not the case, we call $b(\theta) = E[\hat{\theta}] - \theta$ the bias. Expectation is taken with respect to $\mathbf{x}$ (or $p(\mathbf{x}; \theta)$).

- Variance: the variance of an estimator $\hat{\theta}$ is defined as $var(\hat{\theta}) = E[(\hat{\theta} - E[\hat{\theta}])^2]$. Expectations are taken over $\mathbf{x}$ (meaning $\hat{\theta}$ is random, not θ.)

How would you pick a "good" estimator?

Vector versions....

Minimum variance unbiased estimation

MVUE: the unbiased estimator $\hat{\theta}$ of the parameter θ that minimizes the estimation variance.

Why?

Suppose we want to minimize the mean squared error of our estimate, this can be shown to be:

$$\begin{aligned} \mathrm{MSE}(\hat{\theta}) &= E[(\hat{\theta} - \theta)^2] \\ &= var(\hat{\theta}) + b^2(\theta) \end{aligned}$$

So?

Minimum variance unbiased estimation

- Does a MVUE always exist?

Give a counter-example! (b1pg.20)

- If it does, can we always find it?

- Is there anything at all we can say about MVUE?

The Cramer-Rao Lower Bound

- the CRLB give a lower bound on the variance of ANY UNBIASED estimator
- does NOT guarantee bound can be obtained
- IF find an estimator whose variance = CRLB then it's MVUE
- otherwise can use Ch.5 tools (Rao-Blackwell-Lehmann-Scheffe Theorem and Neyman-Fisher Factorization Theorem) to construct a better estimator from any unbiased one - possibly the MVUE if conditions are met

The Cramer-Rao Lower Bound (CRLB)

- Use?
- Intuition?

Consider estimating the parameter A from $x[0] = A + w[0]$, where $w[0] \sim \mathcal{N}(0, \sigma^2)$.

What will the estimator variance depend on and how?

The Cramer-Rao Lower Bound (CRLB)

Theorem: Cramer-Rao Lower Bound: Let $p(x;\theta)$ satisfy the regularity condition

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

Then the variance of any unbiased estimator $\hat{\theta}$ must satisfy

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

where the derivative is evaluated at the true value of θ and expectation is wrt $p(x;\theta)$. Furthermore an unbiased estimator may be found that attains the bound for all θ if and only if

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

for some $g(\cdot)$ and I. That estimator, which is the MVUE, is $\hat{\theta} = g(x)$ and the minimum variance is $1/I(\theta)$.

CRLB examples

Consider estimating the parameter A from $x[0] = A + w[0]$, where $w[0] \sim \mathcal{N}(0, \sigma^2)$. What is the CRLB in estimating A?

Consider estimating the parameter A from the observations $x[n] = A + w[n]$, where $n = 0, 1, 2, \cdots N-1$ is White Gaussian Noise of variance σ^2, i.e. $w[n] \sim \mathcal{N}(0, \sigma^2)$ and are independent. What is the CRLB in estimating A now?

Consider estimating the parameter ϕ from the observations

$$x[n] = A\cos(2\pi f_0 n + \phi) + w[n], \quad n = 0, 1, 2, \cdots N-1$$

in WGN of variance σ^2. The amplitude A and frequency f_0 are assumed to be known. What is the CRLB in estimating ϕ?

CRLB proof

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

$$= \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2\right] \quad (3-8)$$

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

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

$$= \frac{N}{\sigma^2}(\bar{x} - A)$$

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

$$\mathrm{var}(\hat{A}) \geq \frac{\sigma^2}{N}$$

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

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

$$= -\frac{A}{\sigma^2}\sum_{n=0}^{N-1}[x[n]\sin(2\pi f_0 n + \phi)\ -\frac{A}{2}\sin(4\pi f_0 n + 2\phi)]$$

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

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

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

$$= \frac{NA}{2\sigma^2}\frac{1}{N}\sum_{n=0}^{N-1}[1 - \cos(4\pi f_0 n + 2\phi)\ ]$$

$$\approx \frac{NA^2}{2\sigma^2}$$

since

$$\frac{1}{N}\sum_{n=0}^{N-1}\cos(4\pi f_0 n + 2\phi) \approx 0$$

for f_0 not near 0 or 1/2 (see Problem 3.7). Therefore,

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

CRLB T or F

- The CRLB always exists regardless of $p(\mathbf{x};\theta)$
- the CRLB applies to unbiased estimators only.
- Determining the CRLB requires statistics of all possible estimators $\hat{\theta}$.
- The CRLB depends on the observations $\mathbf{x}$.
- The CRLB depends on the parameter to be estimated, θ.
- The CRLB tells you whether or not a MVUE exists.

What is I(θ)?

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

- Why “information”?
  - non-negative
  - additive for independent observations

Vector form of the CRLB

We also have a vector form of the CRLB: when the parameter $\theta = [\theta_1\, \theta_2\, \cdots \theta_p]^T$, we now have a bound on the entire covariance matrix $\mathbf{C}_{\hat{\theta}}$ of any estimator $\hat{\theta}$ of θ. The *Fisher Information Matrix* is the quantity of importance here, and is the generalization of $I(\theta)$ to the vector case. It is a $p \times p$ matrix of second partials of the log likelihood function.

$$\text{Fisher Information Matrix's } i,j \text{ entry: } [\mathbf{I}(\theta)]_{i,j} = -E\left[\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial\theta_i \partial\theta_j}\right],$$

where expectations are again taken over $p(\mathbf{x};\theta)$.

Vector form of the CRLB

<u>Vector form of the CRLB:</u> Assuming $p(\mathbf{x};\theta)$ satisfies the regularity condition

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

where expectation is taken over $p(\mathbf{x};\theta)$, then the covariance matrix of any unbiased estimator $\hat{\theta}$ satisfies

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

where $\succeq \mathbf{0}$ means the matrix is positive semi-definite. Furthermore, an unbiased estimator may be found that attains the bound ($\mathbf{C}_{\hat{\theta}} = \mathbf{I}^{-1}(\theta)$ if and only if

$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta} = \mathbf{I}(\theta)\,(g(\mathbf{x}) - \theta)\,.$$

In that case, $\hat{\theta} = g(\mathbf{x})$ is the MVU estimator with variance $\mathbf{I}^{-1}(\theta)$.

Vector form of the CRLB examples

Consider estimating the vector of parameters $\theta = [A\,\sigma^2]^T$ from the observations $x[n] = A + w[n]$, where $n = 0, 1, 2, \cdots N-1$ is White Gaussian Noise of variance σ^2, i.e. $w[n] \sim \mathcal{N}(0, \sigma^2)$ and are independent. What is the vector CRLB in estimating θ?

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

$$\begin{aligned} \operatorname{var}(\hat{A}) &\geq \frac{2(2N-1)\sigma^2}{N(N+1)} \\ \operatorname{var}(\hat{B}) &\geq \frac{12\sigma^2}{N(N^2-1)}. \end{aligned}$$

Consider the problem of determining A, B given the observations

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

where $w[n]$ is White Gaussian Noise. Determine the CRLB for $\theta = [A\,B]^T$.

What can we conclude?

CRLB for transformations

Sometimes we want to estimate a function of the parameter rather than the parameter itself, i.e. we want to estimate $\alpha = g(\theta)$, and we are given the dependence of the data on θ rather than on $g(\theta)$. One can prove that the CRLB is modified to include the function $g(\theta)$ as follows. If $\hat{\alpha}$ is any unbiased estimator of $g(\theta)$ (and all regularity conditions still hold) then

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

We can show that if $\hat{\theta}$ is an *efficient* estimator of θ that if $g(\theta)$ is an affine transformation, then $\hat{\alpha} = g(\hat{\theta})$ is an efficient estimator of $g(\theta)$. In general though, when $g(\theta)$ is a nonlinear function of θ efficiency is lost, and only asymptotic efficiency can be claimed. That is, as the number of samples $N \to \infty$, the estimator $\hat{\theta}$ converges on its mean, the true value of the unknown parameter, and at that point a function of also becomes efficient.

Vector CRLB for transformations

The co-variance matrix $\mathbf{C}_{\hat{\alpha}}$ of any unbiased estimator $\hat{\alpha}$ of $g(\theta)$, an $r-$dimensional function, satisfies:

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

Here $\frac{\partial \mathbf{g}(\theta)}{\partial \theta}$ is the $r \times p$ Jacobian matrix defined by

$$\left[\frac{\partial \mathbf{g}(\theta)}{\partial \theta}\right]_{ij} = \frac{\partial g_i(\theta)}{\partial \theta_i}$$

Example of vector CRLB with transformation

Consider estimating the SNR for a single example $\alpha = \frac{A^2}{\sigma^2}$ where both A and σ^2 are unknown, from the observations $x[n] = A + w[n]$, where $n = 0, 1, 2, \cdots N-1$ is White Gaussian Noise of variance σ^2. What is the CRLB in estimating α?

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}. \qquad \begin{aligned} \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial A} & \frac{\partial g(\boldsymbol{\theta})}{\partial \sigma^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix} \end{aligned}$$

$$\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{I}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{g}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix} \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix} \begin{bmatrix} \frac{2A}{\sigma^2} \\ -\frac{A^2}{\sigma^4} \end{bmatrix}$$

$$\operatorname{var}(\hat{\alpha}) \geq \frac{4\alpha + 2\alpha^2}{N}.$$

CRLB for General Gaussian Case

- When observations are Gaussian and one knows the dependence of the mean and covariance matrix on the unknown parameters, we know the closed form of the CRLB (or Fisher information matrix):

Given observations

$$\mathbf{x} \sim \mathcal{N}(\mu(\theta),\ \mathbf{C}(\theta)),$$

then the Fisher information matrix $\mathbf{I}(\theta)$ is given by:

$$[\mathbf{I}(\theta)]_{ij} = \left[\frac{\partial \mu(\theta)}{\partial \theta_i}\right]^T \mathbf{C}^{-1}(\theta) \left[\frac{\partial \mu(\theta)}{\partial \theta_j}\right] + \frac{1}{2} tr \left[\mathbf{C}^{-1}(\theta) \frac{\partial \mathbf{C}(\theta)}{\partial \theta_i} \mathbf{C}^{-1}(\theta) \frac{\partial \mathbf{C}(\theta)}{\partial \theta_j}\right]$$

CRLB for Gaussians examples

Find the CRLB for estimating θ from a deterministic signal in WGN

$$x[n] = s[n;\theta] + w[n], \quad n = 0, 1, 2, \cdots, N-1$$

Find the CRLB in estimating the variance σ^2 of WGN when

$$x[n] = w[n], \quad n = 0, 1, 2, \cdots, N-1$$

Find the CRLB in estimating the signal variance σ_A^2 in WGN when

$$x[n] = A + w[n], \quad n = 0, 1, 2, \cdots, N-1$$

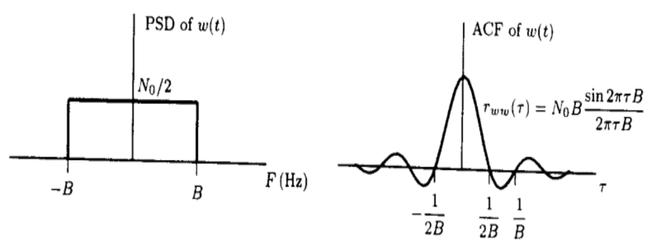
and $A \sim \mathcal{N}(0, \sigma_A^2)$. A is also independent of $w[n]$.

Example: range estimation

In radar or active sonar a signal pulse is transmitted. The round trip delay τ_0 from the transmitter to the target and back is related to the range R as $\tau_0 = 2R/c$, where c is the speed of propagation. Estimation of range is therefore equivalent to estimation of the time delay, assuming that c is known. If $s(t)$ is the transmitted signal, a simple model for the received continuous waveform is

$$x(t) = s(t - \tau_0) + w(t) \qquad 0 \le t \le T.$$

The transmitted signal pulse is assumed to be nonzero over the interval $[0, T_s]$. Additionally, the signal is assumed to be essentially bandlimited to B Hz. If the maximum time delay is $\tau_{0_{max}}$, then the observation interval is chosen to include the entire signal by letting $T = T_s + \tau_{0_{max}}$. The noise is modeled as Gaussian with PSD and ACF as

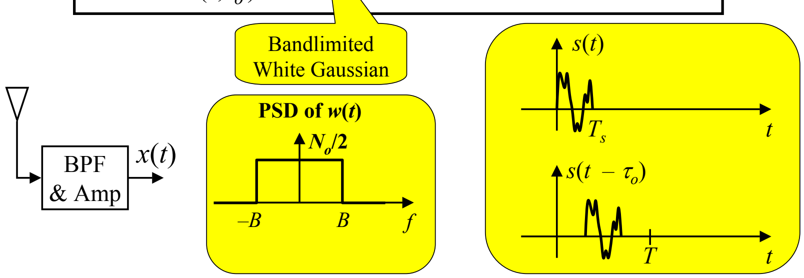


Example: range estimation

Transmit Pulse:	$s(t)$	nonzero over $t \in [0, T_s]$
Receive Reflection:	$s(t - \tau_o)$	
Measure Time Delay:	τ_o	

Continuous time signal model

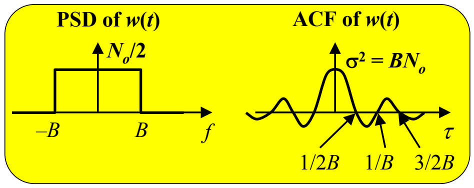
$$\boxed{x(t) = \underbrace{s(t - \tau_o)}_{s(t;\tau_o)} + w(t) \qquad 0 \le t \le T = T_s + \tau_{o,\max}}$$



Taken from http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522.htm

Example: range estimation

Discrete time signal model



Sample Every $\Delta = 1/2B$ sec

$w[n] = w(n\Delta)$

DT White Gaussian Noise Var $\sigma^2 = BN_o$

$$x[n] = \underbrace{s(n\Delta - \tau_o)}_{s[n;\tau_o]\ldots \text{ has } M \text{ non-zero samples starting at } n_o} + w[n] \qquad n = 0, 1, \ldots, N-1$$

$$\boxed{x[n] = \begin{cases} w[n] & 0 \le n \le n_o - 1 \\ s(n\Delta - \tau_o) + w[n] & n_o \le n \le n_o + M - 1 \\ w[n] & n_o + M \le n \le N - 1 \end{cases}}$$

Taken from http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522.htm

Example: range estimation

Now apply standard CRLB result for signal + WGN:

Plug in… and keep non-zero terms

$$\mathrm{var}(\hat{\tau}_o) \ge \frac{\sigma^2}{\sum_{n=0}^{N-1} \left(\frac{\partial s[n;\tau_o]}{\partial \tau_o} \right)^2} = \frac{\sigma^2}{\sum_{n=n_o}^{n_o+M-1} \left(\frac{\partial s(n\Delta - \tau_o)}{\partial \tau_o} \right)^2}$$

$$= \frac{\sigma^2}{\sum_{n=n_o}^{n_o+M-1} \left(\left. \frac{\partial s(t)}{\partial t} \right|_{t=n\Delta-\tau_o} \right)^2} = \frac{\sigma^2}{\sum_{n=0}^{M-1} \left(\left. \frac{\partial s(t)}{\partial t} \right|_{t=n\Delta} \right)^2}$$

Exploit Calculus!!!

Use approximation: $\tau_o = \Delta\, n_o$
Then do change of variables!!

Taken from http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522.htm

Example: range estimation

Assume sample spacing is small… approx. sum by integral…

$$\mathrm{var}(\hat{\tau}_o) \geq \frac{\sigma^2}{\frac{1}{\Delta}\int_0^{T_s}\left(\frac{\partial s(t)}{\partial t}\right)^2 dt} = \frac{N_o/2}{\int_0^{T_s}\left(\frac{\partial s(t)}{\partial t}\right)^2 dt} = \frac{1}{\frac{E_s}{N_o/2}\,\frac{\int_0^{T_s}\left(\frac{\partial s(t)}{\partial t}\right)^2 dt}{E_s}}$$

$$E_s = \int_0^{T_s} s^2(t)\,dt$$

FT Theorem & Parseval

$$\mathrm{var}(\hat{\tau}_o) \geq \frac{1}{\frac{E_s}{N_o/2}\,\frac{\int_0^{T_s}(2\pi f)^2|S(f)|^2 df}{E_s}}$$

$$= \frac{1}{\frac{E_s}{N_o/2}\,\frac{\int_{-\infty}^{\infty}(2\pi f)^2|S(f)|^2 df}{\int_{-\infty}^{\infty}|S(f)|^2 dt}}$$

Parseval

A type of "SNR"

Define a BW measure:

$$B_{rms} = \sqrt{\frac{\int_{-\infty}^{\infty}(2\pi f)^2|S(f)|^2 df}{\int_{-\infty}^{\infty}|S(f)|^2 dt}}$$

B_{rms} is "RMS BW" (Hz)

Taken from http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522.htm

Example: range estimation

Using these ideas we arrive at the CRLB on the delay:

$$\boxed{\mathrm{var}(\hat{\tau}_o) \geq \frac{1}{SNR \times B_{rms}^2} \quad (\sec^2)}$$

To get the CRLB on the **range**… use "transf. of parms" result:

$$CRLB_{\hat{R}} = \left(\frac{\partial R}{\partial \tau_o}\right)^2 CRLB_{\hat{\tau}_o} \quad \text{with} \quad R = c\tau_o/2 \quad \Rightarrow \quad \boxed{\mathrm{var}(\hat{R}) \geq \frac{c^2/4}{SNR \times B_{rms}^2} \quad (\mathrm{m}^2)}$$

CRLB is inversely proportional to:
- SNR Measure
- RMS BW Measure

So the CRLB tells us…
- Choose signal with large B_{rms}
- Ensure that SNR is large
- Better on Nearby/large targets
- Which is better?
  - Double transmitted energy?
  - Double RMS bandwidth?

Taken from http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522.htm