Lecture 9

November 3, 2010

Bounds on Estimates:

Cramer - Rao Lower Bound (CRLB).

Assume that $\hat{\theta}(y)$ is an unbiased estimate of θ.

Then:

$$E[\hat{\theta}(y) - \theta] = \int_{\Gamma} [\hat{\theta}(y) - \theta]\, p(y|\theta)\, dy = 0$$

Taking derivate with respect to θ, we get:

$$\frac{d}{d\theta}\int_{\Gamma} [\hat{\theta}(y) - \theta] p(y|\theta)\, dy = \int_{\Gamma} \frac{\partial}{\partial\theta}\{[\hat{\theta}(y) - \theta]\, p(y|\theta)\}\, dy$$

$$= -\int_{\Gamma} p(y|\theta)\, dy + \int_{\Gamma} \frac{\partial p(y|\theta)}{\partial\theta} [\hat{\theta}(y) - \theta]\, dy = 0$$

or

$$\int_{\Gamma} \frac{\partial}{\partial\theta} p(y|\theta)[\hat{\theta}(y) - \theta]\, dy = 1$$

We can write this as:

$$\int_{\Gamma} \frac{\partial \log p(y|\theta)}{\partial\theta}\, p(y|\theta)\, [\hat{\theta}(y) - \theta]\, dy = 1$$

where we have used the fact that:

$$\frac{d \log f(x)}{dx} = \frac{\frac{d f(x)}{dx}}{f(x)} \Rightarrow f(x) = f(x)\, \frac{d \log f(x)}{dx}$$

This further can be written as:

$$\int_\Gamma \left[\frac{\partial \log p(y|\theta)}{\partial\theta}\sqrt{p(y|\theta)}\right]\left[\sqrt{p(y|\theta)}\,(\hat{\theta}(y)-\theta)\right]dy = 1$$

Using Schwarz's inequality, i.e.,

$$\left[\int g(x)\,h(x)\,dx\right]^2 \leq \int g^2(x)\,dx \int h^2(x)\,dx$$

with equality if $g(x) = k\,h(x)$,

we get:

$$\int_\Gamma \left[\frac{\partial \log p(y|\theta)}{\partial\theta}\right]^2 p(y|\theta)\,dy \underbrace{\int_\Gamma (\hat{\theta}(y)-\theta)^2 p(y|\theta)\,dy}_{\mathrm{Var}[\hat{\theta}(y)-\theta]} \geq 1$$

with equality if and only if:

$$\frac{\partial \log p(y|\theta)}{\partial\theta} = [\hat{\theta}-\theta]\,k(\theta)$$

where $k(\theta)$ is not a function of y.

Finally, we have

$$\mathrm{var}[\hat{\theta}-\theta] = E\left\{(\hat{\theta}(y)-\theta)^2|\theta\right\} \geq \frac{1}{E\left\{\left[\frac{\partial \log p(y|\theta)}{\partial\theta}\right]^2\Big|\theta\right\}}$$

This is called the Cramer-Rao Lower Bound.

[Note: We have denoted natural logarithm by log and not ln.]

Alternative form of CRLB:

We have

$$\int_{\Gamma} p(y|\theta)\,dy = 1$$

Taking derivative:

$$\int_{\Gamma} \frac{\partial}{\partial\theta} p(y|\theta)\,dy = \int_{\Gamma} \frac{\partial \log p(y|\theta)}{\partial \theta}\, p(y|\theta)\,d\theta = 0$$

Taking derivative again:

$$0 = \int_{\Gamma} \frac{\partial^2 \log p(y|\theta)}{\partial\theta^2}\, p(y|\theta)\,d\theta + \int_{\Gamma}\left[\frac{\partial \log p(y|\theta)}{\partial\theta}\right]^2 p(y|\theta)\,dy$$

or

$$E\left[\frac{\partial^2 \log p(y|\theta)}{\partial\theta^2}\Big|\theta\right] = -E\left[\left(\frac{\partial \log p(y|\theta)}{\partial\theta}\right)^2\Big|\theta\right]$$

So:

$$E[(\hat{\theta}(y)-\theta)^2|\theta] \geq \frac{-1}{E\left[\frac{\partial^2 \log p(y|\theta)}{\partial\theta^2}\Big|\theta\right]}$$

An estimate that achieves CRLB is called an <u>efficient</u> <u>estimate</u>.

For an efficient estimate:

$$\frac{\partial \log p(y|\theta)}{\partial\theta} = [\hat{\theta}(y)-\theta]k(\theta)$$

Assume that an efficient estimate exists, i.e., we have:

$$\frac{\partial}{\partial \theta} \log p(y|\theta) = [\hat{\theta}(y) - \theta] k(\theta)$$

Substitute θ by $\hat{\theta}_{ML}$, i.e.,

$$\frac{\partial}{\partial \theta} \log p(y|\theta)\Big|_{\theta=\hat{\theta}_{ML}} = [\hat{\theta}(y) - \hat{\theta}_{ML}] k(\theta)$$

where $\hat{\theta}_{ML}$ is the Maximum-Likelihood estimate.

But,

$$\frac{\partial}{\partial \theta} \log p(y|\theta)\Big|_{\theta=\hat{\theta}_{ML}} = 0$$

So,

$$[\hat{\theta}(y) - \hat{\theta}_{ML}] k(\theta) = 0$$

or

$$\hat{\theta}(y) = \hat{\theta}_{ML}.$$

This means, if there is an efficient estimate then it is ML estimate and can be found by solving

$$\frac{\partial}{\partial \theta} \log p(y|\theta)\Big|_{\theta=\hat{\theta}_{ML}} = 0.$$

But if an efficient estimate does not exist, we cannot ~~know~~ how good the ML estimate is.

Example: Take the example that we considered in the previous lecture, i.e.,

$$p(\underline{y}|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i-\theta)^2}{2\sigma^2}\right]$$

Then

$$\frac{\partial \log p(\underline{y}|\theta)}{\partial \theta} = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} y_i - N\theta\right)$$

So,

$$\hat{\theta}_{ML} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

and

$$E[\hat{\theta}_{ML}] = \frac{1}{n}\sum_{i=1}^{n} E\{y_i\} = \frac{1}{n}\sum_{i=1}^{n}\theta = \theta$$

So $\hat{\theta}_{ML}$ is <u>unbiased</u>.

Since the condition of Cramer-Rao bound equality is satisfied, then the variance of the estimate is given by CRLB.

$$\frac{\partial^2 \log p(\underline{y}|\theta)}{\partial \theta^2} = -\frac{N}{\sigma^2}$$

So:

$$\mathrm{Var}[\hat{\theta}_{ML} - \theta] = \frac{\sigma^2}{N}.$$

Example: Now consider the estimation of a non random parameter of a signal of known form, e.g., the phase or frequency of a sinusoid.

$$y(t) = g(t, \theta) + N(t) \qquad t_0 \leq t \leq t_f$$

Let's assume that we have samples of the received wavefor $y(t)$ at points $t_k = kT$, $k = 1, 2, \ldots$

Then

$$y_k = g_k + N_k$$

where $g_k = g(t_k, \theta)$

It is clear that

$$E\{y_k\} = E\{g_k\}$$

and

$$Var\{y_k | \theta\} = Var(N_k) = \sigma^2$$

Let vector $\underline{y}_n$, be equal to $[y_1, \ldots, y_n]^T$

Then

$$p(\underline{y}_n | \theta) = \prod_{i=1}^{n} p(y_i | \theta)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i - g_i)^2}{2\sigma^2}\right]$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - g_i)^2\right]$$

We have

$$\frac{\partial \log p(\underline{y}_n|\theta)}{\partial\theta} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - g_i)\frac{\partial g_i}{\partial\theta}$$

if we let $n \to \infty$, then the summation tends to an integral and the ML estimate will be the solution to:

$$\int_{t_0}^{t_f}[y(t) - g(t,\theta)]\frac{\partial g(t,\theta)}{\partial\theta}dt\bigg|_{\theta=\hat{\theta}_{ML}} = 0$$

<u>Example</u>: Amplitude estimation:

$$y(t,\theta) = \theta\, g(t)$$

For this case

$$\int_{t_0}^{t_f}[y(t) - \theta g(t)]g(t)\,dt\bigg|_{\theta=\hat{\theta}_{ML}} = 0$$

So

$$\hat{\theta}_{ML} = \frac{\int_{t_0}^{t_f} y(t)\,g(t)\,dt}{\int_{t_0}^{t_f} g^2(t)\,dt}$$

Normalizing $g(t)$, i.e., letting:

$$\int_{t_0}^{t_f} g^2(t)\,dt = 1$$

we get:

$$\hat{\theta}_{ML} = \int_{t_0}^{t_f} y(t) g(t) dt$$

which is the correlation receiver.

We have

$$E\{\hat{\theta}_{ML}\} = E\left\{\int_{t_0}^{t_f} [\theta g(t) + N(t)] g(t) dt\right\} = \theta$$

So, the estimate is unbiased.

$$\frac{\partial \log p(y|\theta)}{\partial \theta} = \frac{2}{N_0} \int_{t_0}^{t_f} [y(t) - \theta g(t)] g(t) dt$$

$$= \frac{2}{N_0} [\hat{\theta}_{ML} - \theta]$$

So $\hat{\theta}_{ML}$ is efficient, i.e., it achieves CRLB.

Here $N_0 = \frac{\sigma^2}{W}$ where W is the bandwidth of the signal.

<u>Example:</u> Phase estimation

$$g(t, \theta) = A \sin(\omega_0 t + \theta)$$

$$\int_{t_0}^{t_f} y(t) \cos(\omega_0 t + \hat{\theta}_{ML}) dt = 0$$

or

$$\hat{\theta}_{ML} = \tan^{-1} \frac{\int_{t_0}^{t_f} y(t) \cos \omega_0 t \, dt}{\int_{t_0}^{t} y(t) \sin \omega_0 t \, dt}$$

CRLB for phase estimation.

By finding $E\left\{\frac{\partial^2 p(y|\theta)}{\partial \theta^2}\right\}$ and substituting in the Cramer-Rao Bound formula, we get

$$Var(\hat{\theta}-\theta) \geq \left[\frac{2E_s}{N_0}\right]^{-1} \frac{1}{n}$$

where E_s is the signal energy $\frac{A^2}{2}T$ and N_0 is the one sided noise density. So, the estimation improves as the length of observation increases and/or SNR increases.

For frequency estimation:

$$Var(\hat{\Omega}-\Omega) \geq \left[\frac{2E_s}{N_0}\right]^{-1} \frac{12}{(n-1)^2 n}$$