Week 2 : Session 2

3. Cramer - Rao Bound

So far, we've looked at the Maximum Likelihood Estimation procedure for Noisy measurements of a Parameter and for Pilot based channel estimation. Now, let us look at another fundamental important concept in this context of Estimation, which is the Cramer - Rao Bound (CRB).

What is CRB ?

This is a fundamental lower bound on the Mean Square Error (MSE) of the parameter estimate.

CRB tells us the best / lowest MSE (i.e.) the best accuracy that is possible by any Unbiased estimator. (i.e.) The accuracy of any other estimator cannot be better than this.

Recall, MSE is a metric to characterize the accuracy of the Estimate. (i.e.) it tells us the spread of the estimate of the unknown parameter around the True value of the parameter. (in case of Unbiased estimator)

The Pathbreaking principle/ result is named in honor of Harald Cramér (Swedish) and C.R. Rao (Indian)

- Harald Cramér was a Swedish mathematician, specializing in Mathematical Statistics
- Calyampudi Radhakrishna Rao, known as C.R. Rao is an Indian - American mathematician and statistician. He has been described as "a living legend". His work has greatly influenced statistics and various other fields such as Economics, Genetics, Anthropology, Geology, Medicine, etc., He has also been described as one of the top 10 Indian Scientists of all time.

Mechanism of CRB

Let us consider the observation vector $\bar{y}$. The likelihood function is $p(\bar{y}; h)$. The log likelihood is the natural log of the likelihood function (ie) $\ln p(\bar{y}; h)$

Let $\hat{h}$ be any unbiased estimator of h. (ie) The Mean / Average of the Estimate yield the True value of the parameter. (ie) $E\{\hat{h}\} = h$.

Now, for this Unbiased estimator, the lower bound on the MSE of $\hat{h}$ is greater than or equal to the inverse of the Fisher information of the parameter $(I(h))$.

$$\text{(ie)} \quad E\{(\hat{h}-h)^2\} \geq \frac{1}{I(h)}$$

It is a measure of the information embodied by the parameter.

$$\text{where,} \quad I(h) = E\left\{\left(\underbrace{\frac{\partial}{\partial h} \ln p(\bar{y}; h)}_{\text{Partial derivative of Log Likelihood.}}\right)^2\right\}$$

As $I(h) \to \infty$, then the bound $\to$ zero.

So, for any Unbiased estimator, MSE $\geq$ Reciprocal of FI.

---

① Let us explore CRB for our first model (Noisy measurements)

$$\begin{aligned} y(1) &= h + v(1) \\ y(2) &= h + v(2) \\ &\vdots \\ y(N) &= h + v(n) \end{aligned}$$

Noisy observations

iid Gaussian noise samples with mean = 0, var = σ^2

Unknown parameter

Recall, Likelihood of h ← Unknown parameter

$$p(\bar{y}; h) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}.$$

Recall, Greater the Likelihood, better is that particular value of h explains the observation vector $\bar{y}$. And, Maximizing this gives the MLE of the parameter h, which we derived as Sample Mean / variance. We thoroughly analyzed it.

Now, the point here is, we are not concerned with any particular estimate (MLE / MMO / ...). What we are interested here in the CRB is "Given the Likelihood, what is the best possible estimation performance that can be achieved". It tells us that, "for any Unbiased estimator this is the Lowest possible variance that can be achieved".

If someone proposes a new estimator, we can always compare it with the CRB. Depending on how close it is with CRB, we can determine how good or bad is that proposed estimator. Of course, no Unbiased estimator can have MSE lower than the CRB, it can only be higher than CRB. The farther it is from CRB, the poorer is the estimation accuracy. So, the closeness to the CRB gives us a measure of how accurate that particular estimation procedure is.

So, this is the use of the CRB principle in practice.

Now, we take the natural log to obtain Log-likelihood.

$$\underbrace{\ln p(\bar{y}; h)}_{\text{Log-likelihood for noisy measurement problem}} = \frac{N}{2} \ln\left(\frac{1}{2\pi\sigma^2}\right) - \frac{1}{2\sigma^2} \sum_{k=1}^{N} \left(y(k) - h\right)^2$$

Now, we take the Partial derivative of Log-likelihood w.r.t unknown parameter 'h'.

$$\frac{\partial}{\partial h} \ln p(\bar{y}; h) = \frac{\partial}{\partial h} \left\{ \cancel{\frac{N}{2} \ln\left(\frac{1}{2\pi\sigma^2}\right)}^{\text{constant}} - \frac{1}{2\sigma^2} \sum_{k=1}^{N} \left(y(k) - h\right)^2 \right\}$$

$$= 0 - \frac{1}{\cancel{2}\sigma^2} \sum_{k=1}^{N} \cancel{2}\left(y(k) - h\right)(-1)$$

$$= \frac{1}{\sigma^2} \sum_{k=1}^{N} \left(y(k) - h\right)$$

$y(k) = h + v(k)$
$\Rightarrow v(k) = y(k) - h$

$$= \frac{1}{\sigma^2} \sum_{k=1}^{N} v(k)$$

Now, the fisher Information (FI), which is the Expected value of square of the Partial derivative of the log-likelihood w.r.t the unknown parameter 'h' is

$$I(h) = E\left\{\left(\frac{\partial}{\partial h} \ln p(\bar{y}; h)\right)^2\right\}$$

$$= E\left\{\left(\frac{1}{\sigma^2} \sum_{k=1}^{N} v(k)\right)^2\right\}$$

$$= \frac{1}{\sigma^4} E\left\{\left(\sum_{k=1}^{N} v(k)\right)^2\right\}$$

$$= \frac{1}{\sigma 4} E\left\{\left(\sum_{k=1}^{N} v(k)\right)\left(\sum_{l=1}^{N} v(l)\right)\right\}$$

$$= \frac{1}{\sigma 4} E\left\{\sum_{k=1}^{N} \sum_{l=1}^{N} v(k) \cdot v(l)\right\}$$

$$= \frac{1}{\sigma 4} \sum_{k=1}^{N} \sum_{l=1}^{N} E\{v(k) \cdot v(l)\}$$

$\because v(k), v(l)$ are iid Gaussian noise samples with mean $= 0$, var $= \sigma^2$,

$$= \frac{1}{\sigma 4} \sum_{k=1}^{N} \sum_{l=1}^{N} \sigma^2 \delta(k-l)$$

$$E\{v(k) \cdot v(l)\} = \sigma^2 \delta(k-l) = \sigma^2, k = l \quad = 0, k \neq l$$

$$= \frac{1}{\sigma 4} \sum_{k=1}^{N} \sigma^2$$

$$= \frac{1}{\sigma 4} N \sigma^2$$

$$= \frac{N}{\sigma^2}.$$

Now, the Cramer-Rao Bound (CRB) for this problem is,

For any unbiased Estimator, $E\{(\hat{h}-h)^2\} \geq \frac{1}{I(h)}$

$$= \frac{1}{N/\sigma^2}$$

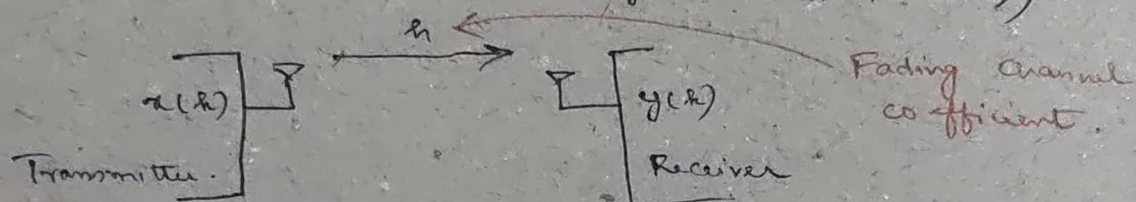
$$\boxed{E\{(\hat{h}-h)^2\} \geq \frac{\sigma^2}{N}}$$

(ie) MSE of any Unbiased Estimator has to be greater than or equal to $\frac{\sigma^2}{N}$. It cannot happen that any Unbiased estimator has MSE that is lower than $\frac{\sigma^2}{N}$. In that sense, this is the fundamental lower bound on the Estimation error of any Unbiased Estimate.

Recall that, MSE of the Maximum Likelihood Estimator is exactly $\frac{\sigma^2}{N}$, which means that MLE achieves the CRB, and any other estimator cannot have lower MSE!! Hence, MLE is the best possible unbiased estimator for this problem, coz it achieves the CRB. And it is termed as an "Efficient Estimator".

Week 2 : Session 3

(2) Let us now explore this CRB in the context of Channel Estimation Problem. Recall, in channel Estimation problem, we have the SISO communication system. (Later we are also going to look at the MIMO communication systems and so on.)



$$
\begin{aligned}
y(1) &= h\,x(1) + v(1) \\
y(2) &= h\,x(2) + v(2) \\
&\vdots \\
y(N) &= h\,x(N) + v(N)
\end{aligned}
$$

iid Gaussian noise samples with mean = 0, variance = σ^2.

Outputs

Pilots

Since $v(k) \sim \mathcal{N}(0, \sigma^2)$, $y(k)$ will also be Gaussian with mean shifted by h times.

(i.e.) $y(k) \sim \mathcal{N}(h\,x(k), \sigma^2)$.

Recall, Likelihood of h ← Unknown channel coefficient

$$p(\bar{y}\,;\,h) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}\left(y(k) - h\,x(k)\right)^2}$$

Now, we take the natural log to obtain log-likelihood.

$$\ln p(\bar{y}\,;\,h) = \frac{N}{2}\ln\left(\frac{1}{2\pi\sigma^2}\right) - \frac{1}{2\sigma^2}\sum_{k=1}^{N}\left(y(k) - h\,x(k)\right)^2$$

Now, we take the partial derivative of log-likelihood w.r.t the unknown channel coefficient h.

$$\frac{\partial}{\partial h} \ln p(\bar{y}; h) = \frac{\partial}{\partial h}\left\{ \overbrace{\cancel{\frac{N}{2}\ln\left(\frac{1}{2\pi\sigma^2}\right)}}^{\text{constant}} - \frac{1}{2\sigma^2}\sum_{k=1}^{N}\left(y(k) - h\,x(k)\right)^2 \right\}$$

$$= 0 - \frac{1}{2\sigma^2}\sum_{k=1}^{N} 2\left(y(k) - h\,x(k)\right)\left(-x(k)\right)$$

$$= \frac{1}{\sigma^2}\sum_{k=1}^{N}\left(y(k) - h\,x(k)\right)\cdot x(k)$$

$y(k) = h\,x(k) + v(k)$
$\Rightarrow v(k) = y(k) - h\,x(k)$

$$= \frac{1}{\sigma^2}\sum_{k=1}^{N} v(k)\cdot x(k)$$

Now, the Fisher Information (FI), which is the Expected value of square of the partial derivative of the log-likelihood w.r.t the unknown channel coefficient 'h' is

$$I(h) = E\left\{\left(\frac{\partial}{\partial h}\ln p(\bar{y}; \bar{h})\right)^2\right\}$$

$$= E\left\{\left(\frac{1}{\sigma^2}\sum_{k=1}^{N} v(k)\cdot x(k)\right)^2\right\}$$

$$= \frac{1}{\sigma^4} E\left\{\left(\sum_{k=1}^{N} x(k)\cdot v(k)\right)^2\right\}$$

$$= \frac{1}{\sigma^4} E\left\{\left(\sum_{k=1}^{N} x(k)\cdot v(k)\right)\left(\sum_{l=1}^{N} x(l)\,v(l)\right)\right\}$$

$\because v(k), v(l)$ are zero mean iid noise samples,
$E\{v(k)\cdot v(l)\} = \sigma^2\delta(k-l)$

$$= \frac{1}{\sigma^4} E\left\{\sum_{k=1}^{N}\sum_{l=1}^{N} x(k)\,x(l)\,v(k)\,v(l)\right\}$$

$\delta(u) \rightarrow$ Discrete Delta function

$$= \frac{1}{\sigma^4}\sum_{k=1}^{N}\sum_{l=1}^{N} x(k)\,x(l)\,E\{v(k)\cdot v(l)\}$$

$$= \frac{1}{\sigma^4}\sum_{k=1}^{N}\sum_{l=1}^{N} x(k)\,x(l)\cdot\sigma^2\,\delta(k-l)$$

$$= \frac{1}{\sigma^{\cancel{4}\,2}}\sum_{k=1}^{N} x^2(k)\cdot\cancel{\sigma^2}$$

For simplicity, we are considering REAL Quantities.

$$= \frac{1}{\sigma^2}\|\bar{x}\|^2$$

Now, the Cramer - Rao Bound (CRB) for this problem is,

For any Unbiased Estimator, $\underbrace{E\{(\hat{h}-h)^2\}}_{\text{MSE of any Unbiased Estimator.}} \geq \frac{1}{I(h)}$

$$= \frac{1}{\|\bar{x}\|^2/\sigma^2} = \frac{\sigma^2}{\|\bar{x}\|^2}$$

$$\boxed{E\left[(\hat{h}-h)^2\right] = \frac{\sigma^2}{\|\bar{x}\|^2}}$$

This is the fundamental lower bound on MSE of any Unbiased estimator.

Recall that, MSE of the Maximum Likelihood Estimator is also exactly $\frac{\sigma^2}{\|\bar{x}\|^2}$. So, once again, this tells us that, there cannot be any better Unbiased estimator than the Maximum Likelihood estimator, coz the MSE coincides with the CRB.

Therefore, MLE achieves the CRB. And it is termed as an "Efficient Estimator", as there cannot be any better Unbiased Estimator !!