现代信号处理

Lecture 15

唐晓颖
电子与电气工程系
南方科技大学

November 20, 2025

Chapter 2 – Minimum Variance Unbiased Estimation

2.1 Introduction

Search for good estimators of unknown deterministic parameters

- On the average yield the true parameter value — unbiased
- The one that exhibits the least variability — minimum variance

2.2 Summary

Minimum variance unbiased estimators do not, in general, exist.

When they do, several methods can be used to find them. The methods rely on the Cramer-Rao lower bound and the concept of a sufficient statistic.

If a minimum variance unbiased estimator does not exist or if approaches fail, a further constraint on the estimator, being linear in the data, leads to an easily implemented, but suboptimal estimator.

2.3 Unbiased Estimation

Bias: the difference between the expected value of an estimator and the true value of the parameter it is estimating.

Formula for the bias of an estimator $\hat{\theta}$ for the parameter θ:

$$Bias(\hat{\theta}) = E[\hat{\theta}] - \theta, \text{ where } E[\hat{\theta}] \text{ is the expected value of the estimator}$$

Biased estimator: An estimator is said to be biased if the expected value of the estimator does not equal the true parameter value it is estimating.

Unbiased estimator: An estimator is said to be unbiased if its expected value equals the true parameter value it is estimating.

2.3 Unbiased Estimation

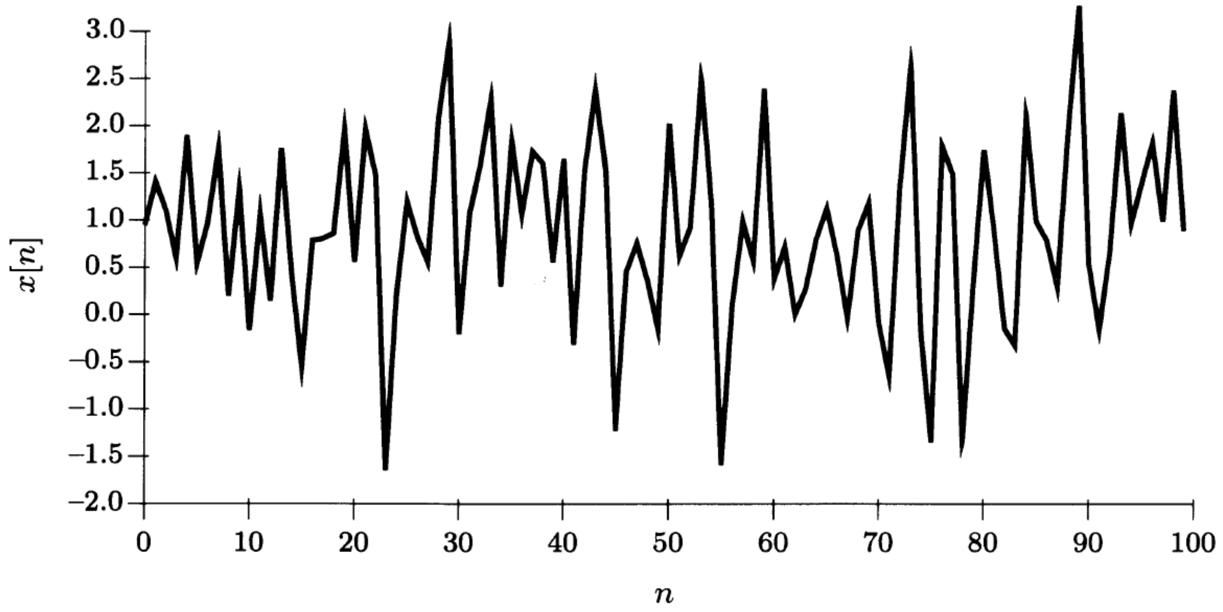
For an estimator to be unbiased we mean that ***on the average*** the estimator will yield the true value of the unknown parameter.

$$E(\hat{\theta}) = \theta \qquad a < \theta < b$$

(Note: $Bias(\hat{\theta}) = E[\hat{\theta}] - \theta$)

(a,b) denotes the range of possible values of θ.

Example - Unbiased Estimator for DC Level in White Gaussian Noise



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

$w[n]$ denotes some zero mean noise process

$-\infty < A < \infty$

Example - Unbiased Estimator for DC Level in White Gaussian Noise

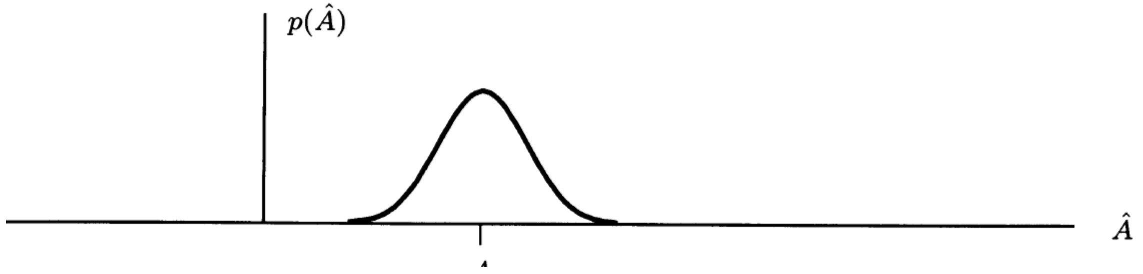
The sample mean of the data

$$\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$$

$$E(\hat{A}) = E\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right)$$
$$= \frac{1}{N}\sum_{n=0}^{N-1} E(x[n])$$
$$= A$$

The sample mean estimator is therefore unbiased.

Unbiased estimators tend to have symmetric PDFs centered about the true value of θ , although this is not necessary



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

The restriction that $E(\hat{\theta}) = \theta$ for all θ is an important one. Letting $\hat{\theta} = g(x)$, where $x = [x[0]x[1]...x[n-1]]^T$, it asserts that

$$E(\hat{\theta}) = \int g(x)p(x;\theta)dx = \theta \qquad \text{for all } \theta$$

Example - Biased Estimator for DC Level in White Gaussian Noise

The modified sample mean of the data

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

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

$$= \frac{1}{2}A$$

$$= A \text{ if } A = 0$$

$$\neq A \text{ if } A \neq 0$$

The modified sample mean estimator is therefore biased.

That an estimator is unbiased does not necessarily mean that it is a good estimator. It only guarantees that on the average it will attain the true value.

It sometimes occurs that multiple estimates of the same parameter are available, i.e., $\{\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_n\}$. A reasonable procedure is to combine these estimates into, hopefully, a better one by averaging them to form

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \hat{\theta}_i$$

Assuming the estimators are unbiased, with the same variance, and uncorrelated with each other,

$$E(\hat{\theta}) = \theta$$

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i$$

$$\mathrm{var}(\hat{\theta}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}(\hat{\theta}_i) = \frac{\mathrm{var}(\hat{\theta}_1)}{n}$$

$$n \to \infty,\ \ \mathrm{var}(\hat{\theta}) \to 0,\ \ \hat{\theta} \to \theta$$

As more estimates are averaged, the variance will decrease.

If the estimators are biased or $E(\hat{\theta}_i) = \theta + b(\theta)$

$$E(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}E(\hat{\theta}_i) = \theta + b(\theta)$$

No matter how many estimators are averaged, $\hat{\theta}$ will not converge to the true value.

$b(\theta) = E(\hat{\theta}) - \theta$ is defined as the *bias* of the estimator.

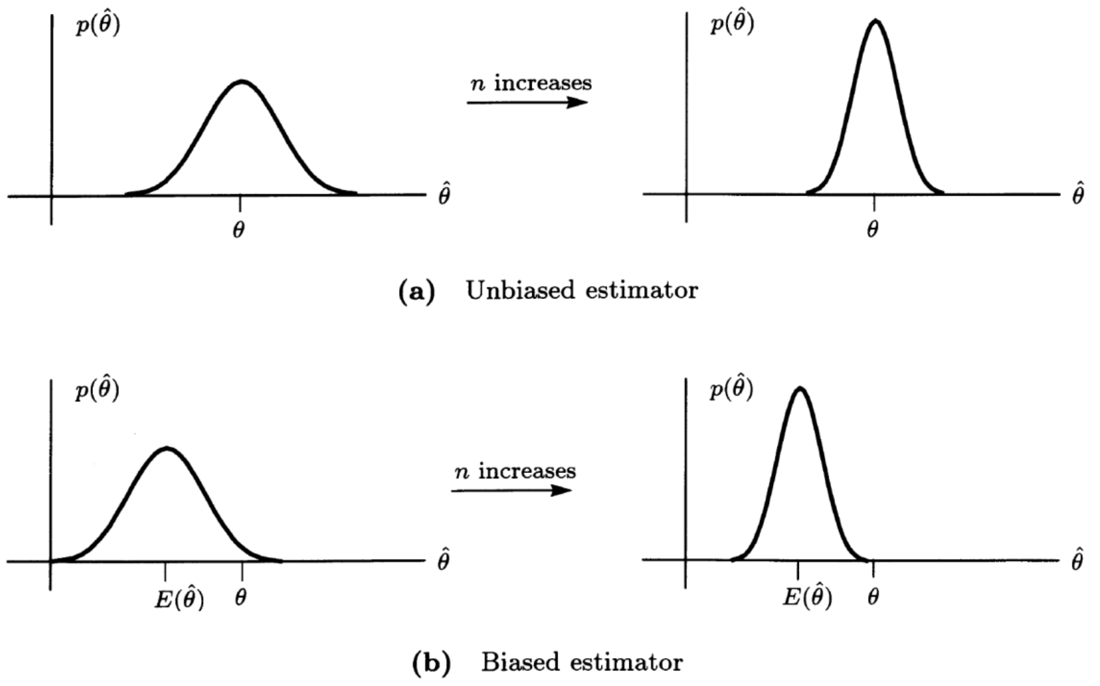


Figure 2.2 Effect of combining estimators

2.4 Minimum Variance Criterion

In searching for optimal estimators we need to adopt some optimality criterion. A natural one is the mean square error (MSE),

$$mse(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right]$$

This measures the average squared deviation of the estimator from the true value.

Adoption of MSE leads to unrealizable estimators, ones that cannot be written solely as a function of the data.

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

The MSE is composed of errors due to the variance of the estimator as well as the bias.

Example - Biased Estimator for DC Level in White Gaussian Noise

The modified sample mean of the data

$$\breve{A} = a\frac{1}{N}\sum_{n=0}^{N-1} x[n] \quad \text{for some constant } a$$

We will attempt to find the a which results in the minimum MSE.

$$E(\breve{A}) = E\left(a\frac{1}{N}\sum_{n=0}^{N-1} x[n] \right)$$
$$= aA$$

$$\mathrm{var}(\breve{A}) = \mathrm{var}\left(a\frac{1}{N}\sum_{n=0}^{N-1} x[n] \right)$$
$$= \frac{a^2\sigma^2}{N}$$

$$mse(\breve{A}) = \mathrm{var}(\breve{A}) + b^2(\breve{A})$$
$$= \mathrm{var}(\breve{A}) + \left[E(\breve{A}) - A \right]^2$$
$$= \frac{a^2\sigma^2}{N} + (a-1)^2 A^2$$

$$\frac{dmse(\breve{A})}{da} = \frac{2a\sigma^2}{N} + 2(a-1)A^2$$

$$a_{\mathrm{opt}} = \frac{A^2}{A^2 + \sigma^2 / N}$$

The optimal value depends upon the unknown parameter A. The estimator is therefore not realizable.

From a practical viewpoint the minimum MSE estimator needs to be abandoned.

An alternative approach is to constrain the bias to be zero and find the estimator which minimizes the variance. Such an estimator is termed ***the minimum variance unbiased (MVU) estimator***.

Note that the MSE of an unbiased estimator is just the variance.

Minimizing the variance of an unbiased estimator also has the effect of concentrating the PDF of the estimation error, $\hat{\theta}-\theta$, about zero. The estimation error will therefore be less likely to be large.

2.5 Existence of the Minimum Variance Unbiased Estimator

The question arises as to whether a MVU estimator exists (for all θ).

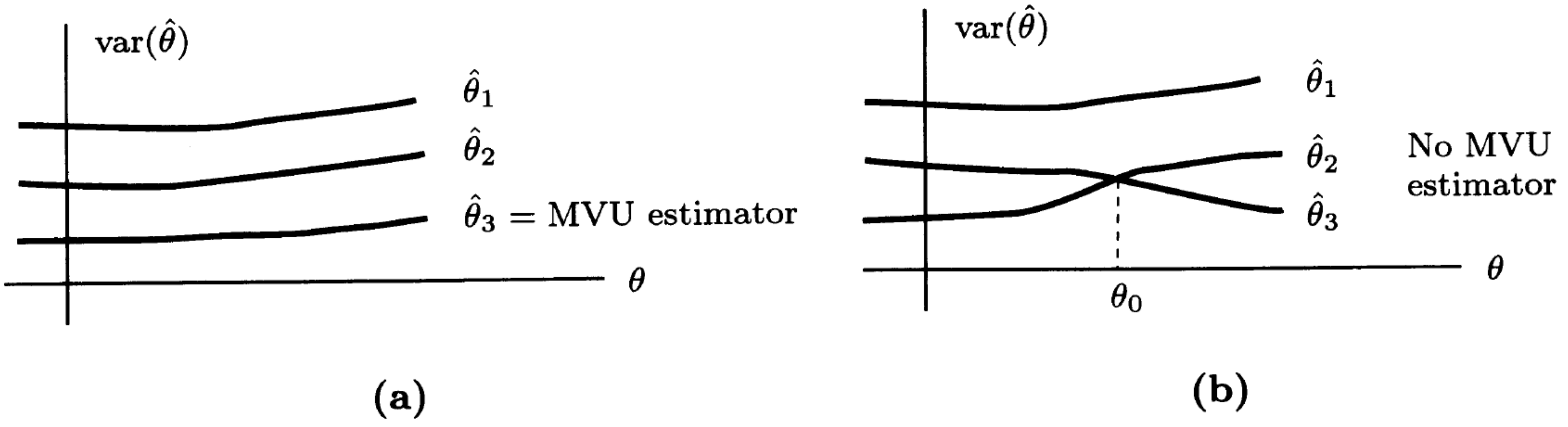


Figure 2.3 Possible dependence of estimator variance with θ

In the former case, $\hat{\theta}_3$ is sometimes referred to as the ***uniformly minimum variance unbiased estimator*** to emphasize that the variance is smallest for all θ.

In general, the MVU estimator does not always exist.

Example - Counterexample to Existence of MVU Estimator

If the form of the PDF changes with θ, then it would be expected that the best estimator would also change with θ.

Assume that we have two independent observations $x[0]$ and $x[1]$ with PDF

$$\begin{aligned} x[0] &\sim \mathcal{N}(\theta, 1) \\ x[1] &\sim \begin{cases} \mathcal{N}(\theta, 1) & \text{if } \theta \geq 0 \\ \mathcal{N}(\theta, 2) & \text{if } \theta < 0. \end{cases} \end{aligned}$$

The following two estimators are unbiased

$$\begin{aligned} \hat{\theta}_1 &= \frac{1}{2}(x[0] + x[1]) \\ \hat{\theta}_2 &= \frac{2}{3}x[0] + \frac{1}{3}x[1] \end{aligned}$$

We have

$$\begin{aligned} \operatorname{var}(\hat{\theta}_1) &= \frac{1}{4}(\operatorname{var}(x[0]) + \operatorname{var}(x[1])) \\ \operatorname{var}(\hat{\theta}_2) &= \frac{4}{9}\operatorname{var}(x[0]) + \frac{1}{9}\operatorname{var}(x[1]) \end{aligned}$$

And thus

$$\operatorname{var}(\hat{\theta}_1) = \begin{cases} \frac{18}{36} & \text{if } \theta \geq 0 \\ \frac{27}{36} & \text{if } \theta < 0 \end{cases}$$

$$\operatorname{var}(\hat{\theta}_2) = \begin{cases} \frac{20}{36} & \text{if } \theta \geq 0 \\ \frac{24}{36} & \text{if } \theta < 0. \end{cases}$$

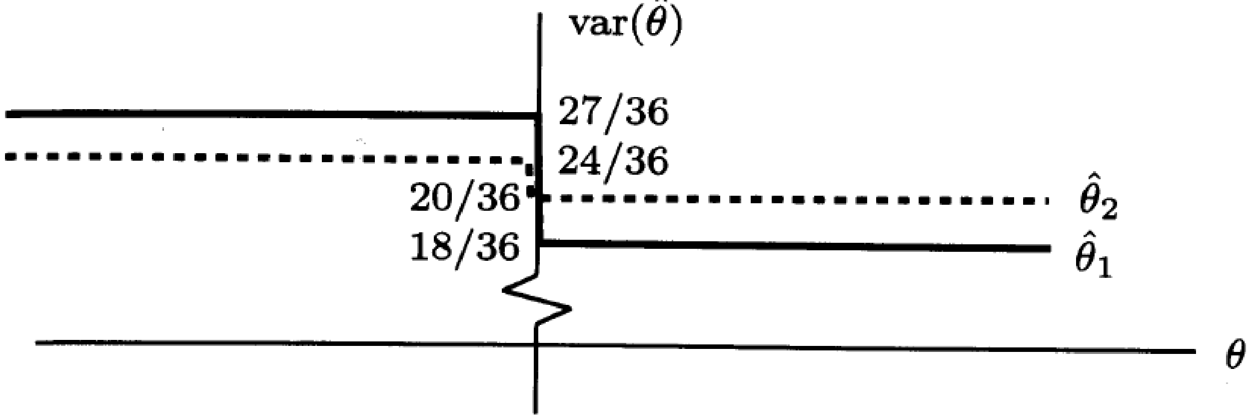


Figure 2.4 Illustration of nonexistence of minimum variance unbiased estimator

Clearly, between these two estimators no MVU estimator exists.

No single estimator can have a variance uniformly less than or equal to the minima.

2.5 Existence of the Minimum Variance Unbiased Estimator

In general, the MVU estimator does not always exist.

Even if it does, we may not be able to find it.

There is not a single method that will always produce the MVU estimator.

Possible reasons:

Complex Parameter Spaces: For the cases with high-dimensional parameter spaces, finding an estimator that meets both the unbiasedness and minimum variance criteria simultaneously could be challenging or impossible.

Trade-offs Between Bias and Variance: There are instances where pursuing unbiasedness can lead to a significant increase in variance. In such scenarios, finding a MVU estimator faces challenges.

…

2.6 Finding the Minimum Variance Unbiased Estimator

In the next few chapters we shall discuss several possible approaches:

1. Determine the **Cramer-Rao lower bound (CRLB)** and check to see if some estimator satisfies it (Chapters 3 and 4).

2. Apply the **Rao-Blackwell-Lehmann-Scheffe (RBLS)** theorem (Chapter 5).

3. Further restrict the class of estimators to be not only unbiased but also *linear.* Then, find the minimum variance estimator within this restricted class (Chapter 6).

Approaches 1 and 2 may produce the MVU estimator, while 3 will yield it only if the MVU estimator is linear in the data.

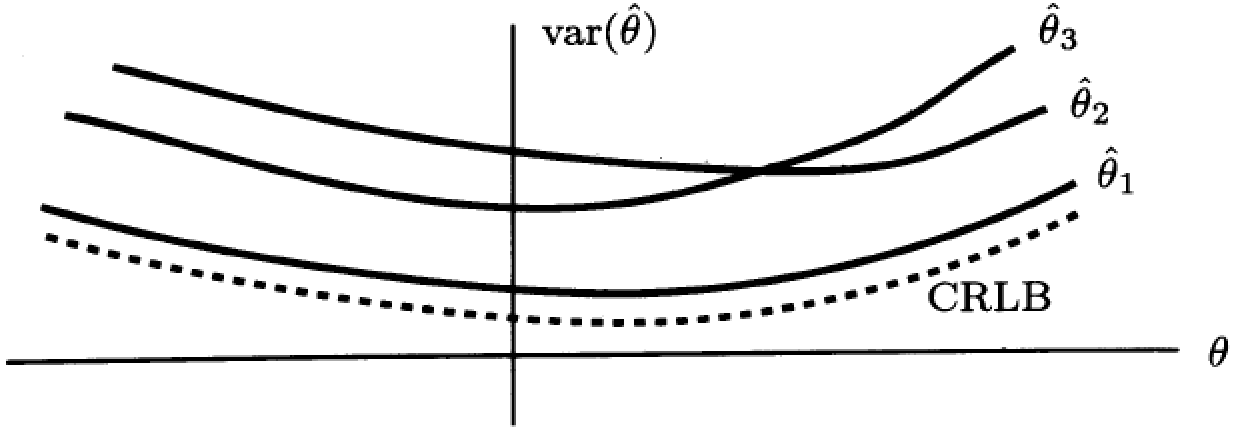


Figure 2.5 Cramer-Rao lower bound on variance of unbiased estimator

The CRLB allows us to determine that for any unbiased estimator the variance must be greater than or equal to a given value.

- If an estimator exists whose variance equals the CRLB for each value of θ, then it must be the MVU estimator.
- It may happen that no estimator exists whose variance equals the bound. Yet, a MVU estimator may still exist, as for instance in the case of $\hat{\theta}_1$. Then, we must resort to the Rao-Blackwell-Lehmann-Scheffe theorem. This procedure first finds a sufficient statistic, one which uses all the data efficiently and then finds a function of the sufficient statistic which is an unbiased estimator of θ.

2.7 Extension to a Vector Parameter

If $\theta = [\theta_1 \theta_2 \ldots \theta_p]^T$ is a vector of unknown parameters, then an estimator $\hat{\theta} = [\hat{\theta}_1 \hat{\theta}_2 \ldots \hat{\theta}_p]^T$ is unbiased if

$$E(\hat{\theta}_i) = \theta_i \qquad a_i < \theta_i < b_i$$

for $i = 1, 2, \ldots, p$.

A MVU estimator has the additional property that $\mathrm{var}(\hat{\theta}_i)$ for $i = 1, 2, \ldots, p$ is minimum among all unbiased estimators.