MINIMUM VARIANCE UNBIASED ESTIMATION

Bias - Variance Tradeoff

The MSE of an estimator $\hat{\underline{\theta}}$ can be broken down into two components, the <u>bias</u> of $\hat{\underline{\theta}}$ and the <u>variance</u> of $\hat{\underline{\theta}}$.

In particular, the following bias-variance decomposition of the MSE holds:

$$\boxed{MSE_{\underline{\theta}}(\hat{\underline{\theta}}) = \left\|Bias_{\underline{\theta}}(\hat{\underline{\theta}})\right\|^2 + Var_{\underline{\theta}}(\hat{\underline{\theta}}).}$$

Let's prove this for the case of a scalar parameter (the vector case is left as an exercise).

$$MSE_\theta(\hat\theta) = E\{(\hat\theta-\theta)^2\}$$

$$= E\{(\hat\theta - E\hat\theta + E\hat\theta - \theta)^2\}$$

$$= E\{(\hat\theta - E\hat\theta)^2\} + 2E\{(\hat\theta - E\hat\theta)\cdot(E\hat\theta-\theta)\}$$

$$+ E\{(E\hat\theta-\theta)^2\}$$

The middle term is

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

(a) $=$

This leaves

$$MSE_\theta(\hat\theta) = E\{(\hat\theta - E\hat\theta)^2\} + E\{(E\hat\theta-\theta)^2\}$$

$$= E\{(\hat\theta - E\hat\theta)^2\} + (E\hat\theta-\theta)^2$$

$$= Var_\theta(\hat\theta) + Bias_\theta^2(\hat\theta) \quad ▨$$

When designing an estimator, you can typically trade off bias and variance. Decreasing the bias of your estimator will increase the variance, while increasing the bias will decrease the variance.

<u>Example</u> Suppose $\underline{x} = [x_1, \dots, x_N]^T$ where

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2), \qquad i = 1, \dots, N.$$

Consider the class of estimators for μ

$$\hat{\mu}_\alpha = \frac{\alpha}{N} \sum_{i=1}^{N} X_i , \qquad \alpha \in \mathbb{R}.$$

Let's see how α affects the bias-variance tradeoff.

Note that

$$\hat{\mu}_\alpha = \alpha \cdot \overline{X}$$

Recall

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{N}\right)$$

Thus

$$\mathrm{Var}_\mu(\hat{\mu}_\alpha) = \mathrm{Var}(\alpha \bar{X})$$

(b)

$$=$$

The bias of $\hat{\mu}_\alpha$ is

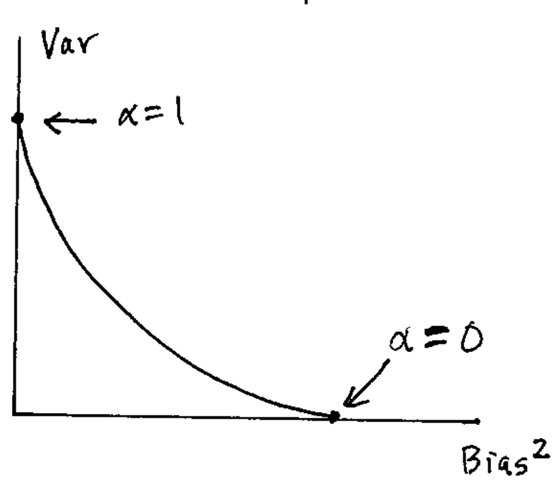
$$\mathrm{Bias}_\mu(\hat{\mu}_\alpha) = E\hat{\mu}_\alpha - \mu$$

$$= E\{\alpha \bar{X}\} - \mu$$

$$=$$

Therefore the MSE is

$$\mathrm{MSE}_\mu(\hat{\mu}_\alpha) = (\alpha-1)^2\mu^2 + \frac{\alpha^2\sigma^2}{N}$$



Minimum MSE?

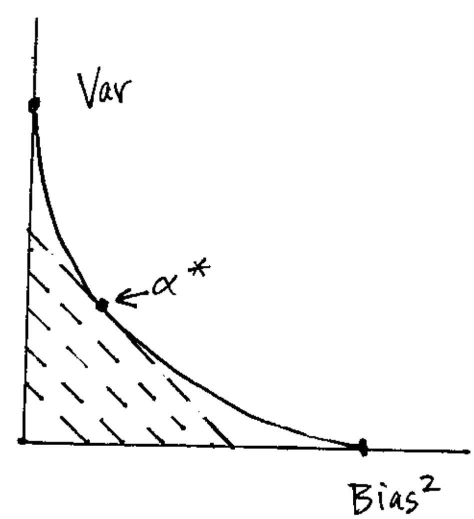
How practical is the MSE as a <u>design</u> criterion?

In the previous example, the MSE is minimized when

$$\frac{\partial \, MSE_\mu(\hat{\mu}_\alpha)}{\partial \alpha} = 2(\alpha-1)\mu^2 + 2\alpha\frac{\sigma^2}{N} = 0$$

$$\Rightarrow \quad \alpha^* = \frac{\mu^2}{\mu^2 + \frac{\sigma^2}{N}}.$$

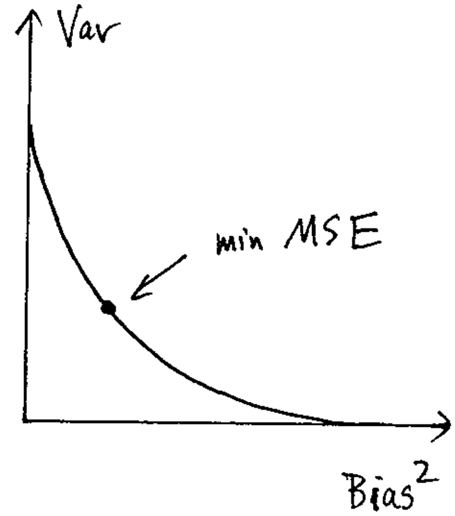
Unfortunately the solution depends on the unknown parameter μ. Therefore the estimator is not <u>realizable</u>.



This phenomenon occurs for many classes of problems, and therefore we need an alternative to direct MSE minimization.

Minimum Variance Unbiased Estimation

In general the minimum MSE estimator has <u>nonzero</u> bias <u>and</u> variance.



However, in many situations <u>only the bias</u> depends on the unknown parameter.

In our example, recall we had

$$\text{Bias}_\mu(\hat{\mu}_\alpha) = (\alpha-1)\mu$$

$$\text{Var}_\mu(\hat{\mu}_\alpha) = \frac{\alpha^2\sigma^2}{N}.$$

This suggests the following alternative:

- constrain the estimator to be unbiased and minimize the variance
- equivalently, minimize the MSE among all unbiased estimators.

Definition $\hat{\underline{\theta}}$ is said to be a (uniform) minimum variance unbiased estimator (MVUE) for $\underline{\theta}$ if

1. $E\,\hat{\underline{\theta}} = \underline{\theta} \qquad \forall\, \underline{\theta} \in \Theta$

2. If $E\,\hat{\underline{\theta}}' = \underline{\theta} \quad \forall\, \underline{\theta} \in \Theta$, then

$$Var_{\underline{\theta}}(\hat{\underline{\theta}}) \le Var_{\underline{\theta}}(\hat{\underline{\theta}}') \quad \forall\, \underline{\theta} \in \Theta$$

Remark Notice that the MVUE criterion requires an estimator to be optimal for all values of $\underline{\theta}$.

This highlights another drawback of unconstrained MSE minimization: it is absurd to require an estimator to have minimal MSE for all $\underline{\theta}$.

For example, the estimator $\hat{\theta} = 28$ can't be beat when $\theta = 28$, but it is terrible elsewhere.

By restricting the class of possible estimators, estimators with uniformly best MSE become possible.

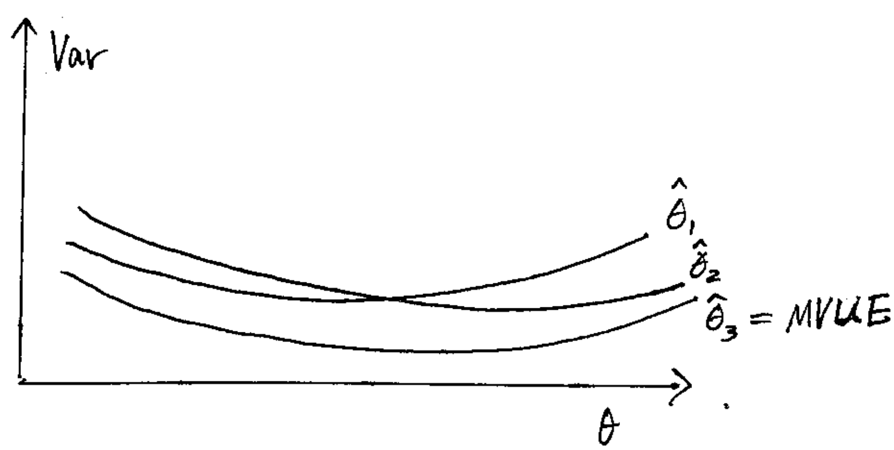
Existence of MVUE

Despite all of this talk, the MVUE does not always exist.

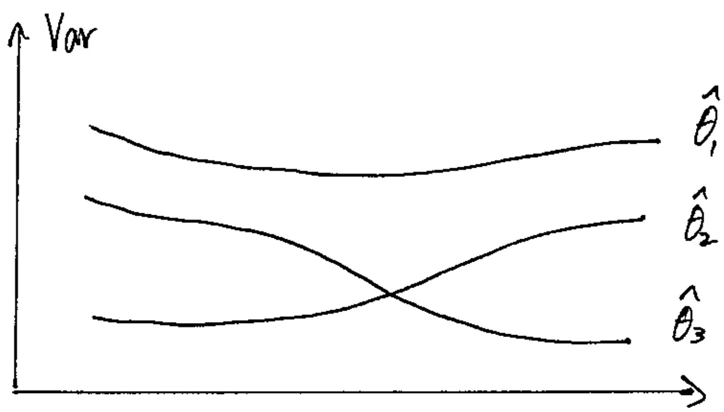
Suppose there are three unbiased estimators, $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$.

There are two possibilities

① One estimator has uniformly smaller variance



② No estimator has uniformly smaller variance



$\boxed{\text{No MVUE exists!}}$

In fact, sometimes there may not exist even a single unbiased estimator!

<u>Exercise</u> Suppose we observe a single scalar realization x of

$$X \sim \text{unif}\left(0, \frac{1}{\theta}\right), \quad \theta > 0.$$

Show that an unbiased estimator of θ does not exist.

Solution / The density of X is

$$f_\theta(x) = \theta \cdot I_{[0,\frac{1}{\theta}]}(x)$$

If $\hat{\theta}$ is unbiased then

$$\forall \theta > 0, \quad \theta = E\{\hat{\theta}\}$$

$$= \int_0^{\frac{1}{\theta}} \hat{\theta}(x) \cdot \theta \, dx$$

$$\Longrightarrow \int_0^{\frac{1}{\theta}} \hat{\theta}(x)\, dx = 1 \qquad \forall \theta > 0.$$

$$\Longrightarrow \hat{\theta}\left(\tfrac{1}{\theta}\right) = 0 \quad \forall \theta > 0 \quad \text{(FTC), a contradiction}$$

Finding the MVUE

Unfortunately there is no systematic procedure. We will discuss three potential ways of finding an MVUE

- Calculate the CRLB and see if some estimator achieves the bound
- Apply the Rao-Blackwell theorem with a complete sufficient statistic
- Further restrict the class of possible estimators to be linear.

$\boxed{\text{Summary}}$

- $MSE = Bias^2 + Variance$
- Minimizing MSE often requires knowledge of $\underline{\theta}$
- MVUE := uniformly best MSE / variance among all unbiased est.
  - may not always exist

Key

a. $= (E\hat{\theta} - \theta)\, E\{(\hat{\theta} - E\hat{\theta})\}$

$= (E\hat{\theta} - \theta)\cdot(E\hat{\theta} - E\hat{\theta}) = 0$

b. $= \alpha^2\, Var(\bar{X})$

$= \dfrac{\alpha^2 \sigma^2}{N}$

c. $= \alpha\mu - \mu$

$= (\alpha - 1)\mu$