ECE531 Screencast 4.1: Introduction to Maximum Likelihood Estimation

D. Richard Brown III

Worcester Polytechnic Institute

Introduction

So far, our focus has been on finding MVU estimators.

- Constrain the class of candidate estimators to **unbiased** estimators.
- Performance quantified by the **variance** of the estimator.

Toolbox:

- Cramer-Rao lower bound and attainability condition.
- Rao-Blackwell-Lehmann-Sheffe theorem

These tools allow us to solve many estimation problems, but they still sometimes fail to work, e.g.

- CRLB attainability condition not satisfied.
- Can't find a complete sufficient statistic
- Can't compute the conditional expectation in the RBLS theorem

What can we do in these cases?

Maximum-Likelihood Estimation

Given a parameterized distribution $p_Y(y\,;\,\theta)$ and a particular observation $Y = y$, the idea of MLE is to find the parameter θ that makes that particular observation most likely.

You should think of y as fixed. The idea is to vary θ to maximize $p_Y(y\,;\,\theta)$ given y. The resulting value of θ, which will be a function of y, is the MLE.

Specifically, we want to solve

$$\hat{\theta}_{\mathsf{ML}}(y) = \arg\max_{\theta \in \Lambda} p_Y(y\,;\,\theta)$$

Example: Estimating the Mean of a Gaussian Observation

Suppose you receive one observation $Y \sim \mathcal{N}(\theta, 1)$ and wish to estimate θ.

We can write

$$p_Y(y\,;\,\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y-\theta)^2\right)$$

What value of θ maximizes $p_Y(y\,;\,\theta)$?

The value of θ that maximizes this pdf is clearly $\theta = y$. Hence $\hat{\theta}_{\mathsf{ML}}(y) = y$.

Note that $\hat{\theta}_{\mathsf{ML}}(y)$ is unbiased and equal to the MVU estimator here.

Example: Estimating the Mean of n Gaussian Observations

Suppose you receive n observations $Y_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, 1)$ for $k = 0, \dots, n-1$ and wish to estimate θ. We can write the joint pdf as

$$p_Y(y\,;\,\theta) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{k=0}^{n-1}(y_k - \theta)^2\right)$$

To maximize this pdf, we should find θ to minimize $\sum_{k=0}^{n-1}(y_k - \theta)^2$.

Let's take a derivative with respect to θ set the result to zero, and solve:

$$\frac{\partial}{\partial\theta}\sum_{k=0}^{n-1}(y_k - \theta)^2 = -\sum_{k=0}^{n-1} 2(y_k - \theta) = 0$$

which has a unique solution

$$\theta = \frac{1}{n}\sum_{k=0}^{n-1} y_k = \hat{\theta}_{\mathsf{ML}}(y)$$

Again, $\hat{\theta}_{\mathsf{ML}}(y)$ is unbiased and equal to the MVU estimator.