Reading

Chapters 4 and 6 of Kay, and the Lecture Supplement on Concentration Ellipsoids.

Chapter 6 describes the (uniformly) best linear unbiased estimator (BLUE) (which is generally a suboptimal estimator, i.e., it is not usually the MVUE).

Chapter 4 shows the true optimality of BLUE under the Gaussian assumption (also see homework problem 5 below).

Section 3.4.2 of Moon gives a simple derivation of the BLUE for a noise vector having iid components.

Homework

1–4. Answer the four homework problems given in the lecture supplement on concentration ellipsoids. Note that the Cramér–Rao lower bound is developed in the class lecture as a matrix bound. The lecture supplement and homework problems are intended to developed an understanding of what it means to have one estimation error covariance matrix underbound another in the positive–definite partial ordering.

5. Let $y = Ax + n$ where $A$ has full column rank and $n \sim N(0, C)$. Show that the MLE of $x$ is unbiased. Also show that it attains the CRLB using two different ways.\(^1\) Why can we claim that we have found the BLUE?

6. Let $p_{\theta}(y)$ belong to a regular statistical family. Define the bias vector $b(\theta) = E_{\theta}(\tilde{\theta})$ for $\tilde{\theta} = \hat{\theta}(y) - \theta$.

   (a) Show that\(^2\)
   $$\text{MSE}_{\theta}(\hat{\theta}) = E_{\theta}(\tilde{\theta}\tilde{\theta}^T) = \text{Cov}_{\theta}(\tilde{\theta}) + b(\theta)b^T(\theta)$$
   and prove the bound\(^3\)
   $$\text{Cov}_{\theta}(\tilde{\theta}) \geq (I + b'(\theta)) J^{-1}(\theta) (I + b'(\theta))^T.$$ 

\(^1\)Again, that the MLE attains the CRLB, and is hence efficient, is not true in generally. The linear, gaussian problem considered here is therefore very special.

\(^2\)Note this shows that for $b(\theta) \neq 0$ we have,
   $$\text{MSE}_{\theta}(\hat{\theta}) = E_{\theta}(\tilde{\theta}\tilde{\theta}^T) \neq \text{Cov}_{\theta}(\hat{\theta}).$$

\(^3\)Hint: Note that the function $f(y) \triangleq \hat{\theta}(y)$ is an unbiased estimator of $g(\theta) = \theta + b(\theta)$ and apply the result of homework problem 5.6.
Note that this bound can be written as

\[
\text{MSE}_\theta(\hat{\theta}) \geq (I + b'(\theta)) J^{-1}(\theta) (I + b'(\theta))^T + b(\theta)b(\theta)^T.
\]

The right-hand-side of the latter bound shows that generally both the variance (the first term corresponding to a lower bound on the variance/covariance) and the bias contribute to the mean-square error. Often, decreasing one of these two terms comes at the expense of increasing the other, an effect known as bias/variance trade-off.

(b) We define efficiency of a biased estimator \( \hat{\theta} \) to hold when \( \text{MSE}_\theta(\hat{\theta}) \) attains the lower bound derived in part (a) above uniformly in \( \theta \). For the case where \( \theta \) is a scalar, show that when

\[-1 < b'(\theta) < -\epsilon < 0 \quad \text{for} \quad 0 < \epsilon < 1\]

a biased estimator may be more efficient than an unbiased estimator, for appropriate values of the bias function \( b(\theta) \). Do this by determine the range of such values for \( b(\theta) \) as a function of \( \epsilon \) and \( J(\theta) \). In particular show the conditions for an efficient biased estimator to be more efficient than an unbiased estimator when \( \epsilon = 0.2 \). [4]

7. Find the CRLB for an unbiased estimator of \( \lambda \) in the probability function \( P(y = k) = \frac{\lambda^k e^{-\lambda}}{k!} \) given \( m \) independent measurements of \( y \).

8. Kay 4.1 and 6.1. Why are these two problems different?


11. Let the real scalar random variable \( y \) be uniformly distributed over the closed interval \([0, \theta]\), for unknown \( \theta \), \( 0 < \theta < \infty \).

   (a) Given iid measurement samples, \( y_1, \ldots, y_m \) drawn from this distribution, determine the BLUE of \( \theta \) and the mse (estimation error covariance) of the BLUE.

---

[4]This example shows that UMVUE are not generally admissible. Recall that without imposing constraints such as the condition of uniform unbiasedness, a uniformly mean-squared error estimator generally does not exist. We do not claim here to be able to obtain an estimator which is globally optimal in the space of all estimators (and which generally could not anyway exist), but rather that there can exist a nonoptimal, biased estimator which is uniformly superior to the optimal unbiased estimator. This can occur because the optimal unbiased estimator is generally inadmissible in the space of all estimators, so it is possible that a uniformly better estimator could exist. Shrinkage estimators, a particular example of which is the Stein estimator, are systematic procedures to determine superior biased estimators starting from optimal unbiased estimators. An introductory discussion of this issue can be found in section 6.5 of Statistics and Econometric Models, Vol. 1, C. Gourieroux and A. Monfort, Cambridge University Press, 1995.
(b) In a future (ECE275B) homework assignment you will determine the actual optimal MVUE for \( \theta \) and will show that the mse of the MVUE is given by \( \frac{\sigma^2}{m(m+2)} \). Plot the ratio of the MVUE mse to the BLUE mse determined in part (a) for \( m = 1, 2, \cdots, 100 \). For greater dramatic effect, plot the ratio of the BLUE mse to the MVUE mse for the same values of \( m \) and state what the value of this ratio is as \( m \to \infty \). This shows that the BLUE (despite its optimality within the class of linear unbiased estimators) can be highly suboptimal in the class of all unbiased estimators.