EE 562

Homework 6

Due Wednesday, March 1, 2017 at 6:40 p.m.

Work all 5 problems.

Problem 1. Let $X_i,\ i = 1, \ldots, n$ be n random variables each with mean μ and variance σ^2. Consider the sample mean defined as

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Show that $\hat{\mu}$ is not the MMSE of μ by finding a constant a such that for any finite n, $a\hat{\mu}$ generates a lower MMSE of μ.

Solution:
Let's first compute the mean square error(MSE) of the estimator $a\hat{\mu}$. We know that $\mathbb{E}[\hat{\mu}_n] = \mu$ and $\text{Var}[\hat{\mu}_n] = \frac{\sigma^2}{n}$. So:

$$\begin{aligned} \text{MSE} = \mathbb{E}\left[|\mu - a\hat{\mu}|^2\right] &= a^2\mathbb{E}\left[\hat{\mu}_n^2\right] - 2a\mu\mathbb{E}[\hat{\mu}_n] + \mu^2 \\ &= a^2\left(\frac{\sigma^2}{n} + \mu^2\right) - 2a\mu^2 + \mu^2 \\ &= \frac{a^2\sigma^2}{n} + (a-1)^2\mu^2 \end{aligned}$$

If we differentiate the MSE with respect to a, we find that the MSE is minimized when $a = a^*$ where:

$$a^* = \frac{\mu^2}{\mu^2 + \frac{\sigma^2}{n}}$$

Since $a^* \neq 1$ for all n, we always have $a^*\hat{\mu}$ which gives us a lower MSE than $\hat{\mu}$. So $\hat{\mu}$ is not MMSE.

Problem 2. Stark and Woods Problems 6.16 and 6.17. Note: The problem statements should refer to equation 6.4-2. These problems are repeated here. Let $X_1, \ldots X_n$ be n independent random variables each with mean μ and variance $\sigma^2 < \infty$. Let us estimate the variance as

$$\hat{\sigma^2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

a. Show that $\hat{\sigma^2}$ is an unbiased estimator for σ^2.

b. Show that this sample variance $\hat{\sigma^2}$ is consistent for σ^2.

Solution:

a.

$$\begin{aligned}
\mathbb{E}\left[\hat{\sigma^2}\right] &= \frac{1}{n-1}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)\right] \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(1 - \frac{1}{n}\right)\sigma^2 \\
&= \sigma^2
\end{aligned}$$

Since $\mathbb{E}\left[\hat{\sigma^2}\right] = \sigma^2$ for all n, the estimator is unbiased.

b. To show $\hat{\sigma^2} \to \sigma^2$ in probability, we first find $\mathrm{Var}\left[\hat{\sigma^2}\right]$ and then use Chebychev's inequality. And since:

$$\begin{aligned}
\mathrm{Var}\left[\hat{\sigma^2}\right] &= \mathrm{E}\left[\left(\hat{\sigma^2} - \sigma^2\right)^2\right] \\
&= \mathbb{E}\left[\hat{\sigma^2}^2\right] - \sigma^4
\end{aligned}$$

But:

$$\begin{aligned}
\mathbb{E}\left[\hat{\sigma^2}\right] &= \frac{1}{(n-1)^2}\mathbb{E}\left[\left(\sum_{i=1}^{n}\left(X_i = \bar{X}\right)^2\right)^2\right] \\
&= \frac{1}{(n-1)^2}\left\{\sum_{i=1}^{n}\mathbb{E}\left[\left(X_i - \bar{X}\right)^4\right] + \sum_{i=1}^{n}\sum_{j\neq i}\mathbb{E}\left[\left(X_i - \bar{X}\right)^2\left(X_j - \bar{X}\right)^2\right]\right\}
\end{aligned}$$

By strong law of large number, $\bar{X} \to \mu$ with probability 1. So:

$$\mathbb{E}\left[\hat{\sigma^2}^2\right] \to \frac{1}{(n-1)^2} n m_4 + n(n-1)\sigma^4$$

where m_4 is the fourth central moment of X_i(Strictly speaking, we need the condition that m_4 is finite). So as n grows large:

$$\begin{aligned}\mathrm{Var}\left[\hat{\sigma^2}\right] &\to \frac{1}{(n-1)^2}\left\{nm_4 + n(n-1)\sigma^4\right\} - \sigma^4 \\ &\to \frac{m_4}{n} \\ &\to 0\end{aligned}$$

So by Chebychev's inequality,

$$\lim_{n\to\infty} \Pr\left\{\left|\hat{\sigma^2} - \sigma^2\right| > \epsilon\right\} < \frac{\mathrm{Var}\left[\hat{\sigma^2}\right]}{\epsilon^2} = 0$$

for any $\epsilon > 0$. So $\hat{\sigma^2} \to \sigma^2$ in probability.

Problem 3. Let $\mathbf{X}(u)$ and $\mathbf{Y}(u)$ be random vectors related to each other by the equation

$$\mathbf{Y}(u) = \mathbf{G}\mathbf{X}(u)$$

where

$$\mathbf{G} = \begin{bmatrix} 6 & 2 & 3 & 0 \\ 1 & 6 & 2 & 3 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 6 \end{bmatrix}, \quad \mu_{\mathbf{X}} = 0, \quad \mathbf{K}_{\mathbf{X}} = \mathbf{I}.$$

a. Compute the LMMSE estimator of $\mathbf{S}(u) = \begin{bmatrix} X(u,3) \\ X(u,4) \end{bmatrix}$ given $\mathbf{R}(u) = \begin{bmatrix} Y(u,1) \\ Y(u,2) \end{bmatrix}$.

b. Compute the LMMSE estimator of $\mathbf{S}(u) = \begin{bmatrix} X(u,1) \\ X(u,2) \end{bmatrix}$ given $\mathbf{R}(u) = \begin{bmatrix} Y(u,3) \\ Y(u,4) \end{bmatrix}$.

c. Compute the LMMSE estimator of $\mathbf{Y}(u)$ given $\mathbf{R}(u) = \begin{bmatrix} X(u,1) \\ X(u,2) \end{bmatrix}$.

Solution:

a. Recall that the LMMSE solution for estimating $\mathbf{S}(u)$ from $\mathbf{R}(u)$ is:

$$\hat{\mathbf{S}}(u) = \mathbf{R_{SR}}\mathbf{R_R^{-1}}\mathbf{R}(u)$$

And here we also have $\mathbf{R_X} = \mathbf{I}$ and $\mathbf{R_Y} = \mathbf{G}\mathbf{G}^T$. And $\mathbf{S}(u) = [X(u,3), X(u,4)]^T$ and $\mathbf{R}(u) = [Y(u,1), Y(u,2)]^T$.

$$\begin{aligned}\mathbf{R_R} &= \begin{bmatrix} \mathbb{E}[Y(u,1)Y(u,1)] & \mathbb{E}[Y(u,1)Y(u,2)] \\ \mathbb{E}[Y(u,2)Y(u,1)] & \mathbb{E}[Y(u,2)Y(u,2)] \end{bmatrix} \\ &= \begin{bmatrix} 49 & 24 \\ 24 & 50 \end{bmatrix}\end{aligned}$$

$$\begin{aligned}\mathbf{R_{SR}} &= \begin{bmatrix} \mathbb{E}[X(u,3)Y(u,1)] & \mathbb{E}[X(u,3)Y(u,2)] \\ \mathbb{E}[X(u,4)Y(u,1)] & \mathbb{E}[X(u,4)Y(u,2)] \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}\end{aligned}$$

The LMMSE estimator is thus:

$$\hat{\mathbf{S}}(u) = \frac{1}{1874}\begin{bmatrix} 102 & 26 \\ -76 & 147 \end{bmatrix}\mathbf{Y}(u)$$

b. Here $\hat{\mathbf{S}}(u) = \mathbf{Y}(u)$ and $\mathbf{R}(u) = [X(u,1), X(u,2)]^T$. Clearly $\mathbf{R_R} = \mathbf{I}$ and

$$\mathbf{R_{SR}} = \begin{bmatrix} 6 & 2 \\ 1 & 6 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

so the LMMSE estimator is:

$$\hat{\mathbf{Y}}(u) = \begin{bmatrix} 6 & 2 \\ 1 & 6 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{R}(u)$$

Problem 4. Let $X_1,\ X_2,\ X_3$ be real random variables with known means $E[X_i] = \mu_i$ and variances $\mathrm{Var}(X_i) = \sigma_i^2,\ i = 1,2,3$ and covariances $\mathrm{Cov}(X_i, X_j) = \sigma_{ij},\ i,j = 1,2,3$ for $i \neq j$ where

$$\sigma_2^2 = 1, \quad \sigma_3^2 = 2, \quad \sigma_{12} = 1/2, \quad \sigma_{13} = 4/3, \quad \sigma_{23} = 1.$$

a. Consider $\hat{X}_1$, the best linear predictor of X_1 given X_2, X_3, in the form

$$\hat{X}_1 = \alpha_2(X_2 - \mu_2) + \alpha_3(X_3 - \mu_3) + \mu_1.$$

Find the numerical values of the coefficients α_2 and α_3.

b. If the real vector (X_1, X_2, X_3) is multivariate normal with moments as given above, find $E[X_1|X_2, X_3]$, the conditional expectation of X_1 given $X_2,\ X_3$.

Solution:

a. Using the orthogonality principle,

$$\mathbb{E}\left[\left(X_1 - \hat{X}_1\right)X_2\right] = 0 \Rightarrow \mathbb{E}\left[\left(X_1 - \hat{X}_1\right)(X_2 - \mu_2)\right] = 0 \qquad (1)$$

$$\Rightarrow \mathbb{E}\left[\left(X_1 - \mu_1 - \left(\hat{X}_1 - \mu_1\right)\right)(X_2 - \mu_2)\right] = 0 \qquad (2)$$

$$\Rightarrow \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] = \mathbb{E}\left[\left(\hat{X}_1 - \mu_1\right)(X_2 - \mu_2)\right] \qquad (3)$$

$$\Rightarrow \sigma_{12} = \alpha_2\sigma_2^2 + \alpha_3\sigma_{23} \qquad (4)$$

or

$$\sigma_{12} = \alpha_2\sigma_2^2 + \alpha_3\sigma_{23}$$

In the same way,

$$\sigma_{12} = \alpha_2\sigma_{23} + \alpha_3\sigma_3^2$$

So,

$$\begin{pmatrix}\sigma_{12}\\ \sigma_{13}\end{pmatrix} = \begin{pmatrix}\sigma_2^2 & \sigma_{23}\\ \sigma_{23} & \sigma_3^2\end{pmatrix}\begin{pmatrix}\alpha_2\\ \alpha_3\end{pmatrix}$$

This last expression represents the desired system of equations. In our case, $\begin{pmatrix}\alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}-\frac{1}{3}\\ \frac{5}{6}\end{pmatrix}$.

b. Recall that for a nultivariate normal, the best linear predictor of X_1 given X_2,X_3 coincides with $\mathbb{E}[X_1|X_2, X_3]$. So,

$$\mathbb{E}[X_1|X_2, X_3] = \alpha_2(X_2 - \mu_2) + \alpha_3(X_3 - \mu_3) + \mu_1$$

where, α_2, α_3 satisfy the 2×2 system above.

Problem 5. Consider a real vector observation of the form

$$\mathbf{X}(u) = A(u)\mathbf{S} + \mathbf{n}(u)$$

where $A(u)$ and $\mathbf{n}(u)$ are independent and the noise $\mathbf{n}(u)$ is mean zero with known covariance matrix $\mathbf{K_n}$. Also

$$P(A(u) = 1) = p, \quad P(A(u) = -1) = 1 - p.$$

You may assume that $\mathbf{K_X}$ and $\mathbf{K_n}$ are nonsingular.

a. Compute the LMMSE estimator of $A(u)$ given the observation $\mathbf{X}(u)$.

b. Compute the mean-square estimation error for your answer in part (a).

Solutiuon:

a. The LMMSE estimator of $A(u)$ given $\mathbf{X}(u)$ is:

$$\hat{A}(u) = \mathbf{R}_{A\mathbf{X}}\mathbf{R_X}^{-1}\mathbf{X}(u)$$

So we need expressions for R_X and $\mathbf{R}_{A\mathbf{X}}$. Taking advantage of the independence of $A(u)$ and $\mathbf{n}(u)$:

$$\begin{aligned}
\mathbf{R_X} &= \mathbb{E}\Big[\mathbf{X}(u)\mathbf{X}(u)^T\Big] \\
&= \mathbb{E}\Big[(A(u)\mathbf{S} + \mathbf{n}(u))(A(u)\mathbf{S} + \mathbf{n}(u))^T\Big] \\
&= \mathbb{E}\big[A(u)^2\big]\mathbf{S}\mathbf{S}^T + \mathbb{E}[A(u)]\mathbf{S}\mathbb{E}\Big[\mathbf{n}(u)^T\Big] + \mathbb{E}[\mathbf{n}(u)]\mathbf{S}^T\mathbf{E}[A(u)] + \mathbb{E}\Big[\mathbf{n}(u)\mathbf{n}(u)^T\Big] \\
&= \mathbf{S}\mathbf{S}^T + \mathbf{K_n}
\end{aligned}$$

$$\begin{aligned}
\mathbf{R}_{A\mathbf{X}} &= \mathbb{E}\Big[A(u)\mathbf{X}(u)^T\Big] \\
&= \mathbb{E}\Big[A(u)(A(u)\mathbf{S}_{\mathbf{n}}(u))^T\Big] \\
&= \mathbb{E}\big[A(u)^2\big]\mathbf{S}^T + \mathbb{E}[A(u)]\mathbb{E}\Big[\mathbf{n}(u)^T\Big] \\
&= \mathbf{S}^T
\end{aligned}$$

The LMSSE estimator is thus:

$$\hat{A}(u) = \mathbf{S}^T\big[\mathbf{S}\mathbf{S}^T + \mathbf{K_n}\big]^{-1}\mathbf{X}(u)$$

b. The mean-square error is:

$$\begin{aligned} MSE &= \mathbb{E}\left[\left(A(u) - \hat{A}(u)\right)^2\right] \\ &= \mathbb{E}\left[A(u)^2\right] - \mathbf{R}_{A\mathbf{X}}\mathbf{R_X}^{-1}\mathbf{R}_{\mathbf{X}A} \\ &= 1 - \mathbf{S}^T\left[\mathbf{S}\mathbf{S}^T + \mathbf{K_n}\right]^{-1}\mathbf{S} \end{aligned}$$