Suppose that we would like to have an estimator for the random vector $\mathbf{X}$ in the form of

$$\hat{\mathbf{X}}_L = \mathbf{A}\mathbf{Y} + \mathbf{b},$$

where $\mathbf{A}$ and $\mathbf{b}$ are fixed matrices to be determined. Remember that for two random variables X and Y, the linear MMSE estimator of X given Y is

$$\begin{aligned}\hat{X}_L &= \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}(Y - EY) + EX \\ &= \frac{\text{Cov}(X,Y)}{\text{Cov}(Y,Y)}(Y - EY) + EX.\end{aligned}$$

We can extend this result to the case of random vectors. More specifically, we can show that the linear MMSE estimator of the random vector $\mathbf{X}$ given the random vector $\mathbf{Y}$ is given by

$$\hat{\mathbf{X}}_L = \mathbf{C_{XY}}\mathbf{C_Y}^{-1}(\mathbf{Y} - E[\mathbf{Y}]) + E[\mathbf{X}].$$

In the above equation, $\mathbf{C_Y}$ is the covariance matrix of $\mathbf{Y}$, defined as

$$\mathbf{C_Y} = E[(\mathbf{Y} - E\mathbf{Y})(\mathbf{Y} - E\mathbf{Y})^T],$$

and $\mathbf{C_{XY}}$ is the cross covariance matrix of $\mathbf{X}$ and $\mathbf{Y}$, defined as

$$\mathbf{C_{XY}} = E[(\mathbf{X} - E\mathbf{X})(\mathbf{Y} - E\mathbf{Y})^T].$$

The above calculations can easily be done using MATLAB or other packages. However, it is sometimes easier to use the orthogonality principle to find $\hat{\mathbf{X}}_L$. We now explain how to use the orthogonality principle to find linear MMSE estimators.

Using the Orthogonality Principle to Find Linear MMSE Estimators for Random Vectors:

Suppose that we are estimating a vector $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_m \end{bmatrix}$$

given that we have observed the random vector $\mathbf{Y}$. Let

$$\hat{\mathbf{X}}_L = \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{X}_m \end{bmatrix}$$

be the vector estimate. We define the MSE as

$$MSE = \sum_{k=1}^{m} E[(X_k - \hat{X}_k)^2].$$

Therefore, to minimize the MSE, it suffices to minimize each $E[(X_k - \hat{X}_k)^2]$ individually. This means that we only need to discuss estimating a random variable X given that we have observed the random vector $\mathbf{Y}$. Since we would like our estimator to be linear, we can write

$$\hat{X}_L = \sum_{k=1}^{n} a_k Y_k + b.$$

The error in our estimate $\tilde{X}$ is then given by

$$\begin{aligned} \tilde{X} &= X - \hat{X}_L \\ &= X - \sum_{k=1}^{n} a_k Y_k - b. \end{aligned}$$

Similar to the proof of Theorem 9.1, we can show that the linear MMSE should satisfy

$$\begin{aligned} &E[\tilde{X}] = 0, \\ &\textrm{Cov}(\tilde{X}, Y_j) = E[\tilde{X} Y_j] = 0, \quad \textrm{for all } j = 1, 2, \cdots, n. \end{aligned}$$

The above equations are called the **orthogonality principle**. The orthogonality principle is often stated as follows: The error ($\tilde{X}$) must be orthogonal to the observations (Y_1, Y_2, $\cdots$, Y_n). Note that there are $n+1$ unknowns (a_1, a_2, $\cdots$, a_n and b) and $n+1$ equations. Let us look at an example to see how we can apply the orthogonality principle.

Example 9.9

Let X be an unobserved random variable with $EX = 0$, $\textrm{Var}(X) = 4$. Assume that we have observed Y_1 and Y_2 given by

$$\begin{aligned} Y_1 &= X + W_1, \\ Y_2 &= X + W_2, \end{aligned}$$

where $EW_1 = EW_2 = 0$, $\mathrm{Var}(W_1) = 1$, and $\mathrm{Var}(W_2) = 4$. Assume that W_1, W_2 , and X are independent random variables. Find the linear MMSE estimator of X, given Y_1 and Y_2.

Solution

The linear MMSE of X given Y has the form

$$\hat{X}_L = aY_1 + bY_2 + c.$$

We use the orthogonality principle. We have

$$\begin{aligned} E[\tilde{X}] &= aEY_1 + bEY_2 + c \\ &= a \cdot 0 + b \cdot 0 + c = c. \end{aligned}$$

Using $E[\tilde{X}] = 0$, we conclude $c = 0$. Next, we note

$$\begin{aligned} \mathrm{Cov}(\hat{X}_L, Y_1) &= \mathrm{Cov}(aY_1 + bY_2, Y_1) \\ &= a\mathrm{Cov}(Y_1, Y_1) + b\mathrm{Cov}(Y_1, Y_2) \\ &= a\mathrm{Cov}(X + W_1, X + W_1) + b\mathrm{Cov}(X + W_1, X + W_2) \\ &= a(\mathrm{Var}(X) + \mathrm{Var}(W_1)) + b\mathrm{Var}(X) \\ &= 5a + 4b. \end{aligned}$$

Similarly, we find

$$\begin{aligned} \mathrm{Cov}(\hat{X}_L, Y_2) &= \mathrm{Cov}(aY_1 + bY_2, Y_2) \\ &= a\mathrm{Var}(X) + b(\mathrm{Var}(X) + \mathrm{Var}(W_2)) \\ &= 4a + 8b. \end{aligned}$$

We need to have

$$\mathrm{Cov}(\tilde{X}, Y_j) = 0, \quad \text{ for } j = 1, 2,$$

which is equivalent to

$$\mathrm{Cov}(\hat{X}_L, Y_j) = \mathrm{Cov}(X, Y_j), \quad \text{ for } j = 1, 2.$$

Since $\mathrm{Cov}(X, Y_1) = \mathrm{Cov}(X, Y_2) = \mathrm{Var}(X) = 4$, we conclude

$$\begin{aligned} 5a + 4b = 4, \\ 4a + 8b = 4. \end{aligned}$$

Solving for a and b, we obtain $a = \frac{2}{3}$, and $b = \frac{1}{6}$. Therefore, the linear MMSE estimator of X, given Y_1 and Y_2, is

$$\hat{X}_L = \frac{2}{3}Y_1 + \frac{1}{6}Y_2.$$