9.1.6 Linear MMSE Estimation of Random Variables

Suppose that we would like to estimate the value of an unobserved random variable X, given that we have observed $Y=y$. In general, our estimate $\hat{x}$ is a function of y

$$\hat{x}=g(y).$$

For example, the MMSE estimate of X given $Y=y$ is

$$g(y)=E[X|Y=y].$$

We might face some difficulties if we want to use the MMSE in practice. First, the function $g(y)=E[X|Y=y]$ might have a complicated form. Specifically, if X and Y are random vectors, computing $E[X|Y=y]$ might not be easy. Moreover, to find $E[X|Y=y]$ we need to know $f_{X|Y}(y)$, which might not be easy to find in some problems. To address these issues, we might want to use a simpler function $g(y)$ to estimate X. In particular, we might want $g(y)$ to be a linear function of y.
Suppose that we would like to have an estimator for X of the form

$$\hat{X}_L=g(Y)=aY+b,$$

where a and b are some real numbers to be determined. More specifically, our goal is to choose a and b such that the MSE of the above estimator

$$MSE=E[(X-\hat{X}_L)^2]$$

is minimized. We call the resulting estimator the **linear MMSE** estimator. The following theorem gives us the optimal values for a and b.

Theorem 9.1
Let X and Y be two random variables with finite means and variances. Also, let ρ be the correlation coefficient of X and Y. Consider the function

$$h(a,b)=E[(X-aY-b)^2].$$

Then,

1. The function $h(a,b)$ is minimized if

$$a = a^* = \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}, \quad b = b^* = EX - aEY.$$

2. We have $h(a^*, b^*) = (1-\rho^2)\text{Var}(X)$.
3. $E[(X - a^*Y - b^*)Y] = 0$ (orthogonality principle).

Proof: We have

$$\begin{aligned} h(a,b) &= E[(X - aY - b)^2] \\ &= E[X^2 + a^2Y^2 + b^2 - 2aXY - 2bX + 2abY] \\ &= EX^2 + a^2EY^2 + b^2 - 2aEXY - 2bEX + 2abEY. \end{aligned}$$

Thus, $h(a,b)$ is a quadratic function of a and b. We take the derivatives with respect to a and b and set them to zero, so we obtain

$$EY^2 \cdot a + EY \cdot b = EXY \qquad (9.4)$$

$$EY \cdot a + b = EX \qquad (9.5)$$

Solving for a and b, we obtain

$$a^* = \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}, \quad b^* = EX - aEY.$$

It can be verified that the above values do in fact minimize $h(a,b)$. Note that Equation 9.5 implies that $E[X - a^*Y - b^*] = 0$. Therefore,

$$\begin{aligned} h(a^*, b^*) &= E[(X - a^*Y - b^*)^2] \\ &= \text{Var}(X - a^*Y - b^*) \\ &= \text{Var}(X - a^*Y) \\ &= \text{Var}(X) + a^{*2}\text{Var}(Y) - 2a^*\text{Cov}(X,Y) \\ &= \text{Var}(X) + \frac{\text{Cov}(X,Y)^2}{\text{Var}(Y)^2}\text{Var}(Y) - 2\frac{\text{Cov}(X,Y)}{\text{Var}(Y)}\text{Cov}(X,Y) \\ &= \text{Var}(X) - \frac{\text{Cov}(X,Y)^2}{\text{Var}(Y)} \\ &= (1-\rho^2)\text{Var}(X). \end{aligned}$$

Finally, note that

$$\begin{aligned} E[(X - a^*Y - b^*)Y] &= EXY - a^*EY^2 - b^*EY \\ &= 0 \quad \text{(by Equation 9.4)}. \end{aligned}$$

Note that $\tilde{X} = X - a^*Y - b^*$ is the error in the linear MMSE estimation of X given Y. From the above theorem, we conclude that

$$E[\tilde{X}] = 0,$$
$$E[\tilde{X}Y] = 0.$$

In sum, we can write the linear MMSE estimator of X given Y as

$$\hat{X}_L = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y - EY) + EX.$$

If $\rho = \rho(X,Y)$ is the correlation coefficient of X and Y, then $\mathrm{Cov}(X,Y) = \rho\sigma_X\sigma_Y$, so the above formula can be written as

$$\hat{X}_L = \frac{\rho\sigma_X}{\sigma_Y}(Y - EY) + EX.$$

Linear MMSE Estimator

The **linear MMSE** estimator of the random variable X, given that we have observed Y, is given by

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

The estimation error, defined as $\tilde{X} = X - \hat{X}_L$, satisfies the **orthogonality principle**:

$$\begin{aligned}&E[\tilde{X}] = 0, \\ &\mathrm{Cov}(\tilde{X}, Y) = E[\tilde{X}Y] = 0.\end{aligned}$$

The MSE of the linear MMSE is given by

$$E\left[(X - X_L)^2\right] = E[\tilde{X}^2] = (1 - \rho^2)\mathrm{Var}(X).$$

Note that to compute the linear MMSE estimates, we only need to know expected values, variances, and the covariance. Let us look at an example.

Example 9.8

Suppose $X \sim Uniform(1,2)$, and given $X = x$, Y is exponential with parameter $\lambda = \frac{1}{x}$.

a. Find the linear MMSE estimate of X given Y.
b. Find the MSE of this estimator.
c. Check that $E[\tilde{X}Y]=0$.

Solution

We have

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

Therefore, we need to find EX, EY, $\text{Var}(Y)$, and $\text{Cov}(X,Y)$. First, note that we have $EX=\frac{3}{2}$, and

$$\begin{aligned}
EY &= E[E[Y|X]] && \big(\text{law of iterated expectations}\big)\\
&= E\left[X\right] && \big(\text{since } Y|X\sim Exponential(\frac{1}{X})\big)\\
&= \frac{3}{2}.
\end{aligned}$$

$$\begin{aligned}
EY^2 &= E[E[Y^2|X]] && (\text{law of iterated expectations})\\
&= E\left[2X^2\right] && \big(\text{since } Y|X\sim Exponential(\frac{1}{X})\big)\\
&= \int_1^2 2x^2dx\\
&= \frac{14}{3}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{Var}(Y) &= EY^2-(EY)^2\\
&= \frac{14}{3}-\frac{9}{4}\\
&= \frac{29}{12}.
\end{aligned}$$

We also have

$$
\begin{aligned}
EXY &= E[E[XY|X]] && \text{(law of iterated expectations)} \\
EXY &= E[XE[Y|X]] && \text{(given } X\text{, } X \text{ is a constant)} \\
&= E\left[X \cdot X\right] && \left(\text{since } Y|X \sim Exponential(\frac{1}{X})\right) \\
&= \int_1^2 x^2 dx \\
&= \frac{7}{3}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\text{Cov}(X, Y) &= E[XY] - (EX)(EY) \\
&= \frac{7}{3} - \frac{3}{2} \cdot \frac{3}{2} \\
&= \frac{1}{12}.
\end{aligned}
$$

a. The linear MMSE estimate of X given Y is

$$
\begin{aligned}
\hat{X}_L &= \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}(Y - EY) + EX \\
&= \frac{1}{29}\left(Y - \frac{3}{2}\right) + \frac{3}{2} \\
&= \frac{Y}{29} + \frac{42}{29}.
\end{aligned}
$$

b. The MSE of $\hat{X}_L$ is

$$
MSE = (1 - \rho^2)\text{Var}(X).
$$

Since $X \sim Uniform(1, 2)$, $\text{Var}(X) = \frac{1}{12}$. Also,

$$
\begin{aligned}
\rho^2 &= \frac{\text{Cov}^2(X, Y)}{\text{Var}(X)\text{Var}(Y)} \\
&= \frac{1}{29}.
\end{aligned}
$$

Thus,

$$
MSE = \left(1 - \frac{1}{29}\right)\frac{1}{12} = \frac{7}{87}.
$$

c. We have

$$
\begin{aligned}
\tilde{X} &= X - \hat{X}_L \\
&= X - \frac{Y}{29} - \frac{42}{29}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E[\tilde{X}Y] &= E\left[\left(X - \frac{Y}{29} - \frac{42}{29}\right)Y\right] \\
&= E[XY] - \frac{EY^2}{29} - \frac{42}{29}EY \\
&= \frac{7}{3} - \frac{14}{3 \cdot 29} - \frac{42}{29} \cdot \frac{3}{2} \\
&= 0.
\end{aligned}
$$