

Another approach to finding the mean square error is to recognize that the MMSE estimator is a linear estimator and thus must be the optimal linear estimator. Hence, the mean squared error of the optimal linear estimator given by Theorem 9.11 must equal $e_{X,Y}^*$. That is, $e_{X,Y}^* = \text{Var}[X](1 - \rho_{X,Y}^2)$. However, calculation of the correlation coefficient $\rho_{X,Y}$ is at least as much work as direct calculation of $e_{X,Y}^*$.

Problem 9.4.3

(a) The marginal PMFs of X and Y are listed below

$$P_X(x) = \begin{cases} 1/3 & x = -1, 0, 1 \\ 0 & \text{otherwise} \end{cases} \quad P_Y(y) = \begin{cases} 1/4 & y = -3, -1, 0, 1, 3 \\ 0 & \text{otherwise} \end{cases}$$

(b) No, the random variables X and Y are not independent since

$$P_{X,Y}(1, -3) = 0 \neq P_X(1)P_Y(-3)$$

(c) Direct evaluation leads to

$$\begin{aligned} E[X] &= 0 & \text{Var}[X] &= 2/3 \\ E[Y] &= 0 & \text{Var}[Y] &= 5 \end{aligned}$$

This implies

$$\sigma_{X,Y} = \text{Cov}[X, Y] = E[XY] - E[X]E[Y] = E[XY] = 7/6$$

(d) From Theorem 9.11, the optimal linear estimate of X given Y is

$$\hat{X}_L(Y) = \rho_{X,Y} \frac{\sigma_X}{\sigma_Y} (Y - \mu_Y) + \mu_X = \frac{7}{30}Y + 0$$

Therefore, $a^* = 7/30$ and $b^* = 0$.

(e) The conditional probability mass function is

$$P_{X|Y}(x|-3) = \frac{P_{X,Y}(x, -3)}{P_Y(-3)} = \begin{cases} \frac{1/6}{1/4} = 2/3 & x = -1 \\ \frac{1/12}{1/4} = 1/3 & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

(f) The minimum mean square estimator of X given that $Y = -3$ is

$$\hat{x}_M(-3) = E[X|Y = -3] = \sum_x x P_{X|Y}(x|-3) = -2/3$$

(g) The mean squared error of this estimator is

$$\begin{aligned} \hat{e}_M(-3) &= E[(X - \hat{x}_M(-3))^2 | Y = -3] = \sum_x (x + 2/3)^2 P_{X|Y}(x|-3) \\ &= (-1/3)^2(2/3) + (2/3)^2(1/3) = 2/9 \end{aligned}$$

Problem 9.4.4

These four joint PMFs are actually related to each other. In particular, completing the row sums and column sums shows that each random variable has the same marginal PMF. That is,

$$\begin{aligned} P_X(x) &= P_Y(x) = P_U(x) = P_V(x) = P_S(x) = P_T(x) = P_Q(x) = P_R(x) \\ &= \begin{cases} 1/3 & x = -1, 0, 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

This implies

$$E[X] = E[Y] = E[U] = E[V] = E[S] = E[T] = E[Q] = E[R] = 0$$

and that

$$E[X^2] = E[Y^2] = E[U^2] = E[V^2] = E[S^2] = E[T^2] = E[Q^2] = E[R^2] = 2/3$$

Since each random variable has zero mean, the second moment equals the variance. Also, the standard deviation of each random variable is $\sqrt{2/3}$. These common properties will make it much easier to answer the questions.

- (a) Random variables X and Y are independent since for all x and y ,

$$P_{X,Y}(x,y) = P_X(x)P_Y(y)$$

Since each other pair of random variables has the same marginal PMFs as X and Y but a different joint PMF, all of the other pairs of random variables must be dependent. Since X and Y are independent, $\rho_{X,Y} = 0$. For the other pairs, we must compute the covariances.

$$\begin{aligned} \text{Cov}[U, V] &= E[UV] = (1/3)(-1) + (1/3)(-1) = -2/3 \\ \text{Cov}[S, T] &= E[ST] = 1/6 - 1/6 + 0 + -1/6 + 1/6 = 0 \\ \text{Cov}[Q, R] &= E[QR] = 1/12 - 1/6 - 1/6 + 1/12 = -1/6 \end{aligned}$$

The correlation coefficient of U and V is

$$\rho_{U,V} = \frac{\text{Cov}[U, V]}{\sqrt{\text{Var}[U]}\sqrt{\text{Var}[V]}} = \frac{-2/3}{\sqrt{2/3}\sqrt{2/3}} = -1$$

In fact, since the marginal PMF's are the same, the denominator of the correlation coefficient will be $2/3$ in each case. The other correlation coefficients are

$$\rho_{S,T} = \frac{\text{Cov}[S, T]}{2/3} = 0 \quad \rho_{Q,R} = \frac{\text{Cov}[Q, R]}{2/3} = -1/4$$

(b) From Theorem 9.11, the least mean square linear estimator of U given V is

$$\hat{U}_L(V) = \rho_{U,V} \frac{\sigma_U}{\sigma_V} (V - E[V]) + E[U] = \rho_{U,V} V = -V$$

Similarly for the other pairs, all expected values are zero and the ratio of the standard deviations is always 1. Hence,

$$\begin{aligned}\hat{X}_L(Y) &= \rho_{X,Y} Y = 0 \\ \hat{S}_L(T) &= \rho_{S,T} T = 0 \\ \hat{Q}_L(R) &= \rho_{Q,R} R = -R/4\end{aligned}$$

From Theorem 9.11, the mean square errors are

$$\begin{aligned}e_L^*(X, Y) &= \text{Var}[X](1 - \rho_{X,Y}^2) = 2/3 \\ e_L^*(U, V) &= \text{Var}[U](1 - \rho_{U,V}^2) = 0 \\ e_L^*(S, T) &= \text{Var}[S](1 - \rho_{S,T}^2) = 2/3 \\ e_L^*(Q, R) &= \text{Var}[Q](1 - \rho_{Q,R}^2) = 5/8\end{aligned}$$

Problem 9.5.1

The problem statement tells us that

$$f_V(v) = \begin{cases} 1/12 & -6 \leq v \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, we are also told that $R = V + X$ where X is a zero mean Gaussian random variable with a variance of 3.

(a) The expected value of R is the expected value V plus the expected value of X . We already know that X is zero mean, and that V is uniformly distributed between -6 and 6 volts and therefore is also zero mean. So

$$E[R] = E[V + X] = E[V] + E[X] = 0$$

(b) Because X and V are independent random variables, the variance of R is the sum of the variance of V and the variance of X .

$$\text{Var}[R] = \text{Var}[V] + \text{Var}[X] = 12 + 3 = 15$$

(c) Since $E[R] = E[V] = 0$,

$$\text{Cov}[V, R] = E[VR] = E[V(V + X)] = E[V^2] = \text{Var}[V]$$

(d) the correlation coefficient of V and R is

$$\rho_{V,R} = \frac{\text{Cov}[V,R]}{\sqrt{\text{Var}[V]\text{Var}[R]}} = \frac{\text{Var}[V]}{\sqrt{\text{Var}[V]\text{Var}[R]}} = \frac{\sigma_V}{\sigma_R}$$

The LMSE estimate of V given R is

$$\hat{V}(R) = \rho_{V,R} \frac{\sigma_V}{\sigma_R} (R - E[R]) + E[V] = \frac{\sigma_V^2}{\sigma_R^2} R = \frac{12}{15} R$$

Therefore $a^* = 12/15 = 4/5$ and $b^* = 0$.

(e) The minimum mean square error in the estimate is

$$e^* = \text{Var}[V](1 - \rho_{V,R}^2) = 12(1 - 12/15) = 12/5$$

Problem 9.5.2

The solution to this problem is to simply calculate the various quantities required for the optimal linear estimator given by Theorem 9.11. First we calculate the necessary moments of X and Y .

$$\begin{aligned} E[X] &= -1(1/4) + 0(1/2) + 1(1/4) = 0 \\ E[X^2] &= (-1)^2(1/4) + 0^2(1/2) + 1^2(1/4) = 1/2 \\ E[Y] &= -1(17/48) + 0(17/48) + 1(14/48) = -1/16 \\ E[Y^2] &= (-1)^2(17/48) + 0^2(17/48) + 1^2(14/48) = 31/48 \\ E[XY] &= 3/16 - 0 - 0 + 1/8 = 5/16 \end{aligned}$$

The variances and covariance are

$$\begin{aligned} \text{Var}[X] &= E[X^2] - (E[X])^2 = 1/2 \\ \text{Var}[Y] &= E[Y^2] - (E[Y])^2 = 493/768 \\ \text{Cov}[X,Y] &= E[XY] - E[X]E[Y] = 5/16 \\ \rho_{X,Y} &= \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{5\sqrt{6}}{\sqrt{493}} \end{aligned}$$

By reversing the labels of X and Y in Theorem 9.11, we find that the optimal linear estimator of Y given X is

$$\hat{Y}_L(X) = \rho_{X,Y} \frac{\sigma_Y}{\sigma_X} (X - E[X]) + E[Y] = \frac{5}{8}X - \frac{1}{16}$$

The mean square estimation error is

$$e_L^* = \text{Var}[Y](1 - \rho_{X,Y}^2) = 343/768$$

Problem 9.5.3

To solve this problem, we use Theorem 9.11. The only difficulty is in computing $E[X]$, $E[Y]$, $\text{Var}[X]$, $\text{Var}[Y]$, and $\rho_{X,Y}$. First we calculate the marginal PDFs

$$f_X(x) = \int_x^1 2(y+x) dy = y^2 + 2xy \Big|_{y=x}^{y=1} = 1 + 2x - 3x^2$$

$$f_Y(y) = \int_0^y 2(y+x) dx = 2xy + x^2 \Big|_{x=0}^{x=y} = 3y^2$$

The first and second moments of X are

$$E[X] = \int_0^1 (x + 2x^2 - 3x^3) dx = x^2/2 + 2x^3/3 - 3x^4/4 \Big|_0^1 = 5/12$$

$$E[X^2] = \int_0^1 (x^2 + 2x^3 - 3x^4) dx = x^3/3 + x^4/2 - 3x^5/5 \Big|_0^1 = 7/30$$

The first and second moments of Y are

$$E[Y] = \int_0^1 3y^3 dy = 3/4$$

$$E[Y^2] = \int_0^1 3y^4 dy = 3/5$$

Thus, X and Y each have variance

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{129}{2160} \quad \text{Var}[Y] = E[Y^2] - (E[Y])^2 = \frac{3}{80}$$

To calculate the correlation coefficient, we first must calculate the the correlation

$$E[XY] = \int_0^1 \int_0^y 2xy(x+y) dx dy$$

$$= \int_0^1 [2x^3y/3 + x^2y^2] \Big|_{x=0}^{x=y} dy$$

$$= \int_0^1 \frac{5y^4}{3} dy = 1/3$$

Hence, the correlation coefficient is

$$\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{E[XY] - E[X]E[Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{5}{\sqrt{129}}$$

Finally, we use Theorem 9.11 to combine these quantities in the optimal linear estimator.

$$\hat{X}_L(Y) = \rho_{X,Y} \frac{\sigma_X}{\sigma_Y} (Y - E[Y]) + E[X]$$

$$= \frac{5}{\sqrt{129}} \frac{\sqrt{129}}{9} (Y - 3/4) + 5/12$$

$$= 5Y/9$$

Problem 9.5.4

The linear mean square estimator of X given Y is

$$\hat{X}_L(Y) = \left(\frac{E[XY] - \mu_X \mu_Y}{\text{Var}[Y]} \right) (Y - \mu_Y) + \mu_X$$

Where we can calculate the following

$$f_Y(y) = \int_0^y 6(y-x) dx = 6xy - 3x^2 \Big|_0^y = 3y^2 \quad (0 \leq y \leq 1)$$

$$f_X(x) = \int_x^1 6(y-x) dy = 3(1 + -2x + x^2) \quad (0 \leq x \leq 1) \quad 0 \quad \text{otherwise}$$

The moments of X and Y are

$$E[Y] = \int_0^1 3y^3 dy = 3/4 \quad E[X] = \int_0^1 3x(1 - 2x + x^2) dx = 1/4$$

$$E[Y^2] = \int_0^1 3y^4 dy = 3/5 \quad E[X^2] = \int_0^1 3x^2(1 + -2x + x^2) dx = 1/10$$

The correlation between X and Y is

$$E[XY] = 6 \int_0^1 \int_x^1 xy(y-x) dy dx = 1/5$$

Putting these pieces together, the optimal linear estimate of X given Y is

$$\hat{X}_L(Y) = \left(\frac{1/5 - 3/16}{3/5 - (3/4)^2} \right) \left(Y - \frac{3}{4} \right) + \frac{1}{4} = \frac{Y}{3}$$

Problem 9.5.5

We are told that random variable X has a second order Erlang distribution

$$f_X(x) = \begin{cases} \lambda x e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We also know that given $X = x$, random variable Y is uniform on $[0, x]$ so that

$$f_{Y|X}(y|x) = \begin{cases} 1/x & 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

- (a) Given $X = x$, Y is uniform on $[0, x]$. Hence $E[Y|X = x] = x/2$. Thus the minimum mean square estimate of Y given X is

$$\hat{Y}_M(X) = E[Y|X] = X/2$$

- (b) The minimum mean square estimate of X given Y can be found by finding the conditional probability density function of X given Y . First we find the joint density function.

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = \begin{cases} \lambda e^{-\lambda x} & 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Now we can find the marginal of Y

$$f_Y(y) = \int_y^\infty \lambda e^{-\lambda x} dx = \begin{cases} e^{-\lambda y} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

By dividing the joint density by the marginal density of Y we arrive at the conditional density of X given Y .

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \lambda e^{-\lambda(x-y)} & x \geq y \\ 0 & \text{otherwise} \end{cases}$$

Now we are in a position to find the minimum mean square estimate of X given Y . Given $Y = y$, the conditional expected value of X is

$$E[X|Y = y] = \int_y^\infty \lambda x e^{-\lambda(x-y)} dx$$

Making the substitution $u = x - y$ yields

$$E[X|Y = y] = \int_0^\infty \lambda(u+y)e^{-\lambda u} du$$

We observe that if U is an exponential random variable with parameter λ , then

$$E[X|Y = y] = E[U + y] = \frac{1}{\lambda} + y$$

The minimum mean square error estimate of X given Y is

$$\hat{X}_M(Y) = E[X|Y] = \frac{1}{\lambda} + Y$$

- (c) Since the MMSE estimate of Y given X is the linear estimate $\hat{Y}_M(X) = X/2$, the optimal linear estimate of Y given X must also be the MMSE estimate. That is, $\hat{Y}_L(X) = X/2$.
- (d) Since the MMSE estimate of X given Y is the linear estimate $\hat{X}_M(Y) = Y + 1/\lambda$, the optimal linear estimate of X given Y must also be the MMSE estimate. That is, $\hat{X}_L(Y) = Y + 1/\lambda$.

Problem 9.5.6

From the problem statement, we learn the following facts:

$$f_R(r) = \begin{cases} e^{-r} & r \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad f_{X|R}(x|r) = \begin{cases} re^{-rx} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Note that $f_{X,R}(x,r) > 0$ for all non-negative X and R . Hence, for the remainder of the problem, we assume both X and R are non-negative and we omit the usual “zero otherwise” considerations.

(a) To find $\hat{r}_M(X)$, we need the conditional PDF

$$f_{R|X}(r|x) = \frac{f_{X|R}(x|r)f_R(r)}{f_X(x)}$$

The marginal PDF of X is

$$f_X(x) = \int_0^\infty f_{X|R}(x|r)f_R(r) dr = \int_0^\infty re^{-(x+1)r} dr$$

To use the integration by parts formula $\int u dv = uv - \int v du$ by choosing $u = r$ and $dv = e^{-(x+1)r} dr$. Thus $v = -e^{-(x+1)r}/(x+1)$ and

$$f_X(x) = \frac{-r}{x+1}e^{-(x+1)r} \Big|_0^\infty + \frac{1}{x+1} \int_0^\infty e^{-(x+1)r} dr = \frac{-1}{(x+1)^2}e^{-(x+1)r} \Big|_0^\infty = \frac{1}{(x+1)^2}$$

Now we can find the conditional PDF of R given X .

$$f_{R|X}(r|x) = \frac{f_{X|R}(x|r)f_R(r)}{f_X(x)} = (x+1)^2 re^{-(x+1)r}$$

By comparing, $f_{R|X}(r|x)$ to the Erlang PDF shown in Appendix A, we see that given $X = x$, the conditional PDF of R is an Erlang PDF with parameters $n = 1$ and $\lambda = x + 1$. This implies

$$E[R|X = x] = \frac{1}{x+1} \quad \text{Var}[R|X = x] = \frac{1}{(x+1)^2}$$

Hence, the MMSE estimator of R given X is

$$\hat{r}_M(X) = E[R|X] = \frac{1}{X+1}$$

(b) The MMSE estimate of X given $R = r$ is $E[X|R = r]$. From the initial problem statement, we know that given $R = r$, X is exponential with mean $1/r$. That is, $E[X|R = r] = 1/r$. Another way of writing this statement is

$$\hat{x}_M(R) = E[X|R] = 1/R$$

(c) Note that the mean of X is

$$E[X] = \int_0^{\infty} x f_X(x) dx = \int_0^{\infty} \frac{x}{(x+1)^2} dx = \infty$$

Because $E[X]$ doesn't exist, the LMSE estimate of X given R doesn't exist.

(d) Just as in part (c), because $E[X]$ doesn't exist, the LMSE estimate of R given X doesn't exist.

Problem 9.5.7

(a) As a function of a , the mean squared error is

$$e = E[(aY - X)^2] = a^2 E[Y^2] - 2aE[XY] + E[X^2]$$

Setting $de/da|_{a=a^*} = 0$ yields

$$a^* = \frac{E[XY]}{E[Y^2]}$$

(b) Using $a = a^*$, the mean squared error is

$$e^* = E[X^2] - \frac{(E[XY])^2}{E[Y^2]}$$

(c) We can write the LMSE estimator given in Theorem 9.11 in the form

$$\hat{x}_L(Y) = \rho_{X,Y} \frac{\sigma_X}{\sigma_Y} Y - b$$

where

$$b = \rho_{X,Y} \frac{\sigma_X}{\sigma_Y} E[Y] - E[X]$$

When $b = 0$, $\hat{X}(Y)$ is the LMSE estimate. Note that the typical way that $b = 0$ occurs when $E[X] = E[Y] = 0$. However, it is possible that the right combination of means, variances, and correlation coefficient can also yield $b = 0$.

Problem 9.5.8

The minimum mean square error linear estimator is given by Theorem 9.11 in which X_n and Y_{n-1} play the roles of X and Y in the theorem. That is, our estimate \hat{X}_n of X_n is

$$\hat{X}_n = \hat{X}_L(Y_{n-1}) = \rho_{X_n, Y_{n-1}} \left(\frac{\text{Var}[X_n]}{\text{Var}[Y_{n-1}]} \right)^{1/2} (Y_{n-1} - E[Y_{n-1}]) + E[X_n]$$

By recursive application of $X_n = cX_{n-1} + Z_{n-1}$, we obtain

$$X_n = a^n X_0 + \sum_{j=1}^n a^{j-1} Z_{n-j}$$

The expected value of X_n is $E[X_n] = a^n E[X_0] + \sum_{j=1}^n a^{j-1} E[Z_{n-j}] = 0$. The variance of X_n is

$$\text{Var}[X_n] = a^{2n} \text{Var}[X_0] + \sum_{j=1}^n [a^{j-1}]^2 \text{Var}[Z_{n-j}] = a^{2n} \text{Var}[X_0] + \sigma^2 \sum_{j=1}^n [a^2]^{j-1}$$

Since $\text{Var}[X_0] = \sigma^2 / (1 - c^2)$, we obtain

$$\text{Var}[X_n] = \frac{c^{2n} \sigma^2}{1 - c^2} + \frac{\sigma^2 (1 - c^{2n})}{1 - c^2} = \frac{\sigma^2}{1 - c^2}$$

Note that $E[Y_{n-1}] = dE[X_{n-1}] + E[W_n] = 0$. The variance of Y_{n-1} is

$$\text{Var}[Y_{n-1}] = d^2 \text{Var}[X_{n-1}] + \text{Var}[W_n] = \frac{d^2 \sigma^2}{1 - c^2} + \eta^2$$

Since X_n and Y_{n-1} have zero mean, the covariance of X_n and Y_{n-1} is

$$\text{Cov}[X_n, Y_{n-1}] = E[X_n Y_{n-1}] = E[(cX_{n-1} + Z_{n-1})(dX_{n-1} + W_{n-1})]$$

From the problem statement, we learn that

$$\begin{aligned} E[X_{n-1} W_{n-1}] &= 0 & E[X_{n-1}] E[W_{n-1}] &= 0 \\ E[Z_{n-1} X_{n-1}] &= 0 & E[Z_{n-1} W_{n-1}] &= 0 \end{aligned}$$

Hence, the covariance of X_n and Y_{n-1} is

$$\text{Cov}[X_n, Y_{n-1}] = cd \text{Var}[X_{n-1}]$$

The correlation coefficient of X_n and Y_{n-1} is

$$\rho_{X_n, Y_{n-1}} = \frac{\text{Cov}[X_n, Y_{n-1}]}{\sqrt{\text{Var}[X_n] \text{Var}[Y_{n-1}]}}$$

Since $E[Y_{n-1}]$ and $E[X_n]$ are zero, the linear predictor for X_n becomes

$$\hat{X}_n = \rho_{X_n, Y_{n-1}} \left(\frac{\text{Var}[X_n]}{\text{Var}[Y_{n-1}]} \right)^{1/2} Y_{n-1} = \frac{\text{Cov}[X_n, Y_{n-1}]}{\text{Var}[Y_{n-1}]} Y_{n-1} = \frac{cd \text{Var}[X_{n-1}]}{\text{Var}[Y_{n-1}]} Y_{n-1}$$

Substituting the above result for $\text{Var}[X_n]$, we obtain the optimal linear predictor of X_n given Y_{n-1} .

$$\hat{X}_n = \frac{c}{d} \frac{1}{1 + \beta^2 (1 - c^2)} Y_{n-1}$$

where $\beta^2 = \eta^2 / (d^2 \sigma^2)$. From Theorem 9.11, the mean square estimation error at step n

$$e_L^*(n) = E[(X_n - \hat{X}_n)^2] = \text{Var}[X_n] \left(1 - \rho_{X_n, Y_{n-1}}^2\right) = \sigma^2 \frac{1 + \beta^2}{1 + \beta^2(1 - c^2)} \quad (3)$$

We see that mean square estimation error $e_L^*(n) = e_L^*$, a constant for all n . In addition, e_L^* is an increasing function β .