1. *Shot noise channel.* Consider an additive noise channel with input signal $X \sim U(0, 1)$ and output signal $Y = X + Z$, where the noise $Z|X = x \sim \mathcal{N}(0, ax)$, for some constant $a > 0$, i.e., the noise variance is proportional to the signal. Observing $Y$, find the MMSE linear estimate of $X$. Your answer should be in terms of only $a$ and $Y$.

**Solution** (10 points)

To find the MMSE linear estimate we need to find the means, variances and covariance of the signal and observation:

- $E(X) = \frac{1}{2}$
- $E(Y) = E(X) + E(Z)$
  
  $= \frac{1}{2} + E_X(E(Z|X)) = \frac{1}{2} + 0$
- $Var(X) = \frac{1}{12}$
- $E(Y^2) = E(X^2) + 2E(XZ) + E(Z^2)$
  
  $= \frac{1}{3} + 2E_X(XE(Z|X)) + E_X(E(Z^2|X))$
  
  $= \frac{1}{3} + 0 + E_X(aX) = \frac{1}{3} + \frac{a}{2}$
- $Var(Y) = \frac{1}{3} + \frac{a}{2} - \frac{1}{4} = \frac{1}{12} + \frac{a}{2}$
- $E(XY) = E(X^2) + E_X(XE(Z|X)) = \frac{1}{3} + 0$
- $Cov(X, Y) = \frac{1}{12}$

Thus

$$\hat{X} = \frac{Cov(X, Y)}{Var(Y)}(Y - E(Y)) + E(X)$$

$$= \frac{1}{1 + 6a} \left( Y - \frac{1}{2} \right) + \frac{1}{2}$$

$$= \frac{Y + 3a}{1 + 6a}.$$ 

2. *Additive-noise channel with path gain.* Consider the output $Y$ of an additive-noise channel with path gain, where $X$ and $Z$ are zero mean and uncorrelated, and $a$ and $b$ are constants. Find the MMSE linear estimate of $X$ given $Y$ and its MSE in terms only of $\sigma_X$, $\sigma_Z$, $a$ and $b$. 

**Solution**

To find the MMSE linear estimate we need to find the means, variances and covariance of the signal and observation:

- $E(X) = \frac{1}{2}$
- $E(Y) = E(X) + E(Z)$
  
  $= \frac{1}{2} + E_X(E(Z|X)) = \frac{1}{2} + 0$
- $Var(X) = \frac{1}{12}$
- $E(Y^2) = E(X^2) + 2E(XZ) + E(Z^2)$
  
  $= \frac{1}{3} + 2E_X(XE(Z|X)) + E_X(E(Z^2|X))$
  
  $= \frac{1}{3} + 0 + E_X(aX) = \frac{1}{3} + \frac{a}{2}$
- $Var(Y) = \frac{1}{3} + \frac{a}{2} - \frac{1}{4} = \frac{1}{12} + \frac{a}{2}$
- $E(XY) = E(X^2) + E_X(XE(Z|X)) = \frac{1}{3} + 0$
- $Cov(X, Y) = \frac{1}{12}$

Thus

$$\hat{X} = \frac{Cov(X, Y)}{Var(Y)}(Y - E(Y)) + E(X)$$

$$= \frac{1}{1 + 6a} \left( Y - \frac{1}{2} \right) + \frac{1}{2}$$

$$= \frac{Y + 3a}{1 + 6a}.$$
Solution (10 points)

First we find the mean and variance of $Y$ and its covariance with $X$. In the following we use the notation $\sigma^2_X = P$ and $\sigma^2_Z = N$.

\[
E(Y) = E(abX + bZ) = abE(X) + bE(Z) = 0
\]
\[
\text{Var} = E(abX + bZ)^2 - (E(abX + bZ))^2
\]
\[
= E(a^2b^2X^2 + 2ab^2XZ + b^2Z^2) - E(Y)^2
\]
\[
= a^2b^2 E(X^2) + 2ab^2 E(X) E(Z) + b^2 E(Z^2) - E(Y)^2
\]
\[
= a^2b^2 P + 0 + b^2 N - 0
\]
\[
= a^2b^2 P + b^2 N
\]
\[
\text{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))]
\]
\[
= E(XY)
\]
\[
= E[X(abX + bZ)]
\]
\[
= abE(X^2) + bE(XZ)
\]
\[
= abP + bE(X) E(Z)
\]
\[
= abP
\]

The MMSE linear estimate of $X$ given $Y$ is given by

\[
\hat{X} = \frac{\text{Cov}(X,Y)}{\sigma_Y^2}(Y - E(Y)) + E(X)
\]
\[
= \frac{aP}{a^2b^2 P + b^2 N}(Y - E(Y)) + E(X)
\]
\[
= \frac{aP}{b(a^2P + N)}Y.
\]

The MSE of the linear estimate is the minimum MMSE:

\[
\text{MMSE} = \sigma_X^2 - \frac{\text{Cov}^2(X,Y)}{\sigma_Y^2}
\]
\[
= P - \frac{a^2b^2 P^2}{a^2b^2 P + b^2 N}
\]
\[
= \frac{a^2P^2 + PN - a^2P^2}{a^2P + N}
\]
\[
= \frac{PN}{a^2P + N}
\]
3. **Camera measurement.** The measurement from a camera can be expressed as \( Y = AX + Z \), where \( X \) is the object position with mean \( \mu \) and variance \( \sigma_X^2 \), \( A \) is the occlusion indicator function and is equal to 1 (if the camera can see the object) with probability \( p \), and 0 (if the camera cannot see the object) with probability \( (1 - p) \), and \( Z \) is the measurement error with mean 0 and variance \( \sigma_Z^2 \). Assume that \( X, A, \) and \( Z \) are independent. Find the best linear MSE estimate of \( X \) given the camera measurement \( Y \). Your answer should be in terms of only \( \mu, \sigma_X^2, \sigma_Z^2, \) and \( p \).

**Solution** (10 points)

The MMSE linear estimate of \( X \) given \( Y \) is given by

\[
\hat{X} = \frac{\text{Cov}(X,Y)}{\sigma_Y^2}(Y - E(Y)) + E(X).
\]

Now,

\[
E(X) = \mu
\]
\[
E(Y) = E(AX + Z)
\]
\[
= E(A)E(X) + E(Z)
\]
\[
= p\mu
\]
\[
\text{Var}(Y) = E[(AX + Z - p\mu)^2]
\]
\[
= E[(AX - p\mu)^2]
\]
\[
= E[(AX - p\mu)^2] + \sigma_Z^2
\]
\[
= E[(AX)^2] - p^2\mu^2 + \sigma_Z^2
\]
\[
= E(A^2) E(X^2) - p^2\mu^2 + \sigma_Z^2
\]
\[
= p(\sigma_X^2 + \mu^2) - p^2\mu^2 + \sigma_Z^2
\]
\[
= p\sigma_X^2 + p(1 - p)\mu^2 + \sigma_Z^2
\]
\[
\text{Cov}(X,Y) = E[(X - \mu)(AX + Z - p\mu)]
\]
\[
= E[(X - \mu)(AX - p\mu)]
\]
\[
= E(A)E[(X - \mu)X]
\]
\[
= p\sigma_X^2.
\]

Substituting, we obtain

\[
\hat{X} = \frac{p\sigma_X^2}{p\sigma_X^2 + p(1 - p)\mu^2 + \sigma_Z^2}(Y - p\mu) + \mu.
\]

4. **Jointly Gaussian random variables.** Let \( X \) and \( Y \) be jointly Gaussian random variables with mean 0 and covariance matrix

\[
\begin{bmatrix}
\sigma_X^2 & \sigma_{XY}\rho_{X,Y} \\
\sigma_{XY}\rho_{X,Y} & \sigma_Y^2
\end{bmatrix}
\]

a. What is the pdf of \( E(X \mid Y) \)?

b. What is the minimum MSE estimate of \( Y^2 \) given \( X \)?

Your answers should be in terms of \( \sigma_X, \sigma_Y, \rho_{X,Y}, \) and the random variables \( X \) and \( Y \).
**Solution** (10 points)

a. Since $X$ and $Y$ are zero-mean jointly Gaussian,
\[
E(X \mid Y) = \frac{\sigma_X \rho_{X,Y}}{\sigma_Y} Y.
\]

But $Y \sim \mathcal{N}(0, \sigma_Y^2)$. Therefore
\[
E(X \mid Y) \sim \mathcal{N}(0, \sigma_X^2 \rho_{X,Y}^2).
\]

b. The best MSE estimate of $Y^2$ given $X$ is $E(Y^2 \mid X)$. We have seen that
\[
Y \mid \{X = x\} \sim \mathcal{N}\left(\frac{\sigma_Y \rho_{X,Y}}{\sigma_X} x, \sigma_Y^2 (1 - \rho_{X,Y}^2)\right).
\]

Therefore
\[
E(Y^2 \mid X) = \sigma_Y^2 (1 - \rho_{X,Y}^2) + \left(\frac{\sigma_Y \rho_{X,Y}}{\sigma_X} X\right)^2
= \sigma_Y^2 \left(1 - \rho_{X,Y}^2 \left(1 - \frac{X^2}{\sigma_X^2}\right)\right).
\]

5. *Estimation vs. detection.* Signal $X$ and noise $Z$ are independent random variables, where
\[
X = \begin{cases} 
+1 & \text{with probability } \frac{1}{2} \\
-1 & \text{with probability } \frac{1}{2},
\end{cases}
\]
and $Z \sim \mathcal{U}[-2, +2]$. Their sum $Y = X + Z$ is observed.

a. Find the minimum MSE estimate of $X$ given $Y$ and the corresponding mean square error. What is the probability of error of this estimate?

b. Suppose that we decide whether $X = +1$ or $X = -1$ using a decoder that minimizes the probability of error. Find this optimal decoder and its probability of error. Compare the optimal decoder’s MSE to the minimum MSE.

**Solution** (15 points)

a. We can easily find the piecewise constant density of $Y$
\[
f_Y(y) = \begin{cases} 
\frac{1}{4} & |y| \leq 1 \\
\frac{1}{8} & 1 < |y| \leq 3 \\
0 & \text{otherwise}
\end{cases}
\]
The conditional probabilities of \( X \) given \( Y \) are
\[
P\{X = +1|Y = y\} = \begin{cases} 
0 & -3 \leq y < -1 \\
\frac{1}{2} & -1 \leq y \leq +1 \\
1 & +1 < y \leq +3 
\end{cases}
\]
\[
P\{X = -1|Y = y\} = \begin{cases} 
1 & -3 \leq y < -1 \\
\frac{1}{2} & -1 \leq y \leq +1 \\
0 & +1 < y \leq +3 
\end{cases}
\]

Thus the best MSE estimate is
\[
g(Y) = E(X|Y) = \begin{cases} 
-1 & -3 \leq Y < -1 \\
0 & -1 \leq Y \leq +1 \\
+1 & +1 < Y \leq +3 
\end{cases}
\]

The minimum mean square error is
\[
E_Y(\text{Var}(X|Y)) = E_Y(E(X^2|Y) - (E(X|Y))^2) = E(1 - g(Y)^2)
\]
\[
= 1 - E(g(Y)^2) = 1 - \int_{-\infty}^{\infty} g(y)^2 f_Y(y) \, dy \\
= 1 - \left( \int_{-3}^{-1} \frac{1}{8} \, dy + \int_{-1}^{1} \frac{1}{4} \, dy + \int_{1}^{+3} \frac{1}{8} \, dy \right) \\
= 1 - \int_{-3}^{-1} \frac{1}{8} \, dy - \int_{1}^{3} \frac{1}{8} \, dy = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}.
\]

b. The optimal decoder is given by the MAP rule. The \textit{a posteriori} pmf of \( X \) was found in part (a). Thus the MAP rule reduces to
\[
\hat{\Theta}(y) = \begin{cases} 
-1 & -3 \leq y \leq -1 \\
\pm 1 & -1 < y \leq +1 \\
+1 & +1 < y \leq +3 
\end{cases}
\]

Since either value can be chosen for \( d(y) \) in the center range of \( Y \), a \textit{symmetrical} decoder is sufficient, i.e.,
\[
\hat{\Theta}(y) = \begin{cases} 
-1 & y < 0 \\
+1 & y \geq 0 
\end{cases}
\]

The probability of decoding error is
\[
P\{\hat{\Theta}(Y) \neq X\} = P\{X = -1, Y \geq 0\} + P\{X = 1, Y < 0\} \\
= P\{X = -1|Y \geq 0\} P\{Y \geq 0\} + P\{X = 1|Y < 0\} P\{Y < 0\} \\
= \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{4}.
\]

If we use the decoder (detector) as an estimator, its MSE is
\[
E\left( (\hat{\Theta}(Y) - X)^2 \right) = \frac{3}{4} \cdot 0 + \frac{1}{4} \cdot 2^2 = 1.
\]

This MSE is twice that of the minimum mean square error estimator.
6. Jointly Gaussian random variables, redux. Consider the following joint pdf for $X$ and $Y$:

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{3/4}} e^{-\frac{1}{2} \left( \frac{1}{3} x^2 + \frac{16}{3} y^2 + \frac{8}{3} xy - 8x - 16y + 16 \right)}$$

a. Find $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, and $\text{Cov}(X,Y)$.

b. Find the minimum MSE estimate of $X$ given $Y$ and the corresponding MSE.

Solution (10 points)

a. We can write the joint pdf of any jointly Gaussian $X$ and $Y$ as

$$f_{X,Y}(x,y) = \frac{\exp\left(-\left( a(x-\mu_X)^2 + b(y-\mu_Y)^2 + c(x-\mu_X)(y-\mu_Y) \right) \right)}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho_{X,Y}^2}}$$

where

$$a = \frac{1}{2(1-\rho_{X,Y}^2)\sigma_X^2}, \quad b = \frac{1}{2(1-\rho_{X,Y}^2)\sigma_Y^2}, \quad c = \frac{-2\rho_{X,Y}}{2(1-\rho_{X,Y}^2)\sigma_X \sigma_Y}.$$ 

By inspection of the given $f_{X,Y}(x,y)$ we find that

$$a = \frac{2}{3}, \quad b = \frac{8}{3}, \quad c = \frac{4}{3}.$$ 

We can easily solve the above three equations for the unknowns:

$$\rho_{X,Y} = \frac{c}{2\sqrt{ab}} = -\frac{1}{2}$$

$$\sigma_X^2 = \frac{1}{2(1-\rho_{X,Y}^2)a} = 1$$

$$\sigma_Y^2 = \frac{1}{2(1-\rho_{X,Y}^2)b} = \frac{1}{4}$$

To find $\mu_X$ and $\mu_Y$, we solve a system of two linear equations:

$$2a\mu_X x + c\mu_Y x = 4x$$

$$2b\mu_X x + c\mu_Y y = 8y$$

obtaining $\mu_X = 2$, $\mu_Y = 1$ and $\text{Cov}(X,Y) = \rho_{X,Y} \sigma_X \sigma_Y = -\frac{1}{4}$.

b. Since $X$ and $Y$ are jointly Gaussian random variables, the minimum MSE estimate of $X$ given $Y$ is linear:

$$E(X|Y) = \frac{\rho_{X,Y} \sigma_X}{\sigma_Y} (Y - \mu_Y) + \mu_X = -(Y - 1) + 2 = 3 - Y$$

$$\text{MMSE} = \text{Var}(X|Y) = (1 - \rho_{X,Y}^2)\sigma_X^2 = \frac{3}{4}$$

7. Conditional Independence does not imply Independence. In class, we saw an example in which two independent, identically distributed random variables conditioned on a third random variable were no longer independent. Here, we examine an example of the opposite case: is it possible for conditionally independent random variables to be not independent?
Suppose $X_3 \sim U[0, 1]$, given $X_3, X_1, X_2 \overset{i.i.d.}{\sim} \text{Bern}(X_3)$. Show that $X_1, X_2$ are not independent. Work out the joint distribution $P_{X_1, X_2}$.

Hint: the Beta function is $B(x, y) = \int_0^1 t^{x-1} (1 - t)^{y-1} dt$.

**Solution** (10 points)

To check independence, we compare the marginal joint probability $P_{X_1, X_2}$ with marginal probabilities $P_{X_1}$ and $P_{X_2}$.

The marginal joint distribution $P_{X_1, X_2}$ is computed as follows by marginalizing over $X_3$:

\[
P_{X_1, X_2} = \int_{x_3} P_{X_3}(x_3) x_3^{x_1}(1 - x_3)^{1-x_1} x_3^{x_2}(1 - x_3)^{1-x_2} dx_3
\]

\[
= \int_0^1 x_3^{x_1+x_2}(1 - x_3)^{2-x_1-x_2} dx_3
\]

\[
= B(x_1 + x_2 + 1, 3 - x_1 - x_2).
\]

The marginal distribution $P_{X_1}(x_1)$ is computed as follows:

\[
P_{X_1}(x_1) = \int_0^1 P_{X_3}(x_3) x_3^{x_1}(1 - x_3)^{1-x_1} dx_3
\]

\[
= B(x_1 + 1, 2 - x_1).
\]

Similarly, $P_{X_2}(x_2) = B(X_2 + 1, 2 - x_1)$. Since in general $B(x_1 + x_2 + 1, 3 - x_1 - x_2) \neq B(x_1 + 1, 2 - x_1)B(X_2 + 1, 2 - x_1)$, which can be shown using the Gamma function, we know that $X_1, X_2$ are not independent.

The following problems are optional and need not be turned in for grading.

1. **Independence vs. Conditional Independence** Give an example of random variables $X, Y, Z$ where $f_{X,Z}(x, z) = f_X(x) f_Z(z)$ but $f_{X,Z|Y}(x, z|y) \neq f_{X|Y}(x|y) f_{Z|Y}(z|y)$ i.e. independence does not imply conditional independence.

   **Solution**

   One solution is let $X, Z \overset{iid}{\sim} \text{Bern}(1/2)$ and $Y = X + Z$.

2. **Sum and difference.** Let $X$ and $Y$ be two random variables, and define $U = X - Y$ and $V = X + Y$. Find the minimum MSE linear estimate of $V$ given $U$ as a function of the random variables and $E(X)$, $E(Y)$, $\sigma_X$, $\sigma_Y$, $\rho_{X,Y}$, where $\sigma_X = \sqrt{\text{Var}(X)}$, $\rho_{X,Y} = \text{corr}(X,Y)$.
Solution

First we calculate the first and second moments of $U$ and $V$.

\[
E(V) = E(X) + E(Y)
\]
\[
E(U) = E(X) - E(Y)
\]
\[
\sigma_V^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho_{X,Y}\sigma_X\sigma_Y
\]
\[
\sigma_U^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho_{X,Y}\sigma_X\sigma_Y
\]
\[
\text{Cov}(V,U) = \sigma_X^2 - \sigma_Y^2.
\]

The minimum MSE linear estimate of $V$ given $U$ is given by

\[
\hat{V} = \frac{\text{Cov}(V,U)}{\sigma_U^2}(U - E(U)) + E(V).
\]

Plugging in the moments of $U$ and $V$ gives the answer.

\[
\hat{V} = \frac{\sigma_X^2 - \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2 - 2\rho_{X,Y}\sigma_X\sigma_Y}(U - (E(X) - E(Y)) + (E(X) + E(Y))
\]

Note that $U$ and $V$ are positively correlated if $\sigma_X^2 > \sigma_Y^2$, negatively correlated if $\sigma_X^2 < \sigma_Y^2$, and uncorrelated if $\sigma_X^2 = \sigma_Y^2$.

3. Covariance matrices. Which of the following matrices can be a covariance matrix? Justify your answer. Either construct a random vector $X$ with the given covariance matrix as a function of the i.i.d. zero mean unit variance random variables $Z_1, Z_2, Z_3$, or establish a contradiction as was done in lecture.

\[
\begin{pmatrix}
1 & 2 \\
0 & 2
\end{pmatrix} \quad \begin{pmatrix}
2 & 1 \\
1 & 2
\end{pmatrix} \quad \begin{pmatrix}
1 & 1 & 1 \\
1 & 2 & 2 \\
1 & 2 & 3
\end{pmatrix} \quad \begin{pmatrix}
1 & 1 & 2 \\
1 & 2 & 3 \\
2 & 3 & 3
\end{pmatrix}
\]

Solution

a. No: not symmetric.

b. Yes: covariance matrix of $X_1 = Z_1 + Z_2$ and $X_2 = Z_1 + Z_3$.

c. Yes: covariance matrix of $X_1 = Z_1$, $X_2 = Z_1 + Z_2$, and $X_3 = Z_1 + Z_2 + Z_3$.

d. No: several justifications.
   - $\sigma_{23}^2 = 9 > \sigma_{22}\sigma_{33} = 6$, which contradicts the Schwarz inequality.
   - The matrix is not nonnegative definite since the determinant is $-2$.
   - One of the eigenvalues is negative ($\lambda_1 = -0.8056$).