UNIVERSITY OF CALIFORNIA, SAN DIEGO
Electrical & Computer Engineering Department
ECE 250 - Winter Quarter 2020
Random Processes

Solutions to P.S. #6

1. *Covariance matrices.* Which of the following matrices can be a covariance matrix? Justify your answer either by constructing a random vector $\mathbf{X}$, as a function of the i.i.d zero mean unit variance random variables Z_1, Z_2, and Z_3, with the given covariance matrix, or by establishing a contradiction.

$$\text{(a)} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \qquad \text{(b)} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{(c)} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \qquad \text{(d)} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 3 \end{bmatrix}$$

Solution:

(a) This cannot be a covariance matrix because it is not symmetric.

(b) This is a covariance matrix for $X_1 = Z_1 + Z_2$ and $X_2 = Z_1 + Z_3$.

(c) This is a covariance matrix for $X_1 = Z_1$, $X_2 = Z_1 + Z_2$, and $X_3 = Z_1 + Z_2 + Z_3$.

(d) This cannot be a covariance matrix. Suppose it is, then $\sigma_{23}^2 = 9 > \sigma_{22}\sigma_{33} = 6$, which contradicts the Schwartz inequality. You can also verify this by showing that the matrix is not positive semidefinite. For example, the determinant is -2. Also one of the eigenvalues is negative ($\lambda_1 = -0.8056$).
Alternatively, we can directly show that this matrix does not satisfy the definition of positive semidefiniteness by

$$\begin{bmatrix} 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} = -1 < 0.$$

since $(a^T\mathbf{X})^2$ is a nonnegative random variable.

2. *Spagetti.* We have a bowl with n spaghetti strands. You randomly pick two strand ends and join them. The process is continued until there are no ends left. Let L be the number of spaghetti loops formed. Find $\mathsf{E}[L]$.

Solution: First, we note that the process ends in exaclty n steps because, at each step, the number of strands decreases by 1. So, at the start of step i, there are $n - (i-1) = n - i + 1$ strands.

We use the method of indicators. Define the indicator random variables:

$$X_i = \mathbf{1}_{[\text{loop is formed at step i}]} = \begin{cases} 1, & \text{loop is formed at step i} \\ 0, & \text{otherwise} \end{cases}$$

Then the number of loops L at the end of the process is given by

$$L = \sum_{i=1}^{n} X_i.$$

Given j strands, the probability of forming a loop is the probability of picking 2 ends on one of the j strands. There are j such pairs of ends. There are $\binom{2j}{2}$ possible choices of pairs of ends altogether, each equally likely. Therefore,

$$\mathsf{P}(X_{n-j+1} = 1) = j\left(\frac{1}{\binom{2j}{2}}\right) = \frac{2j}{2j(2j-1)} = \frac{1}{2j-1},\ j = 1, \ldots, n.$$

and, so,

$$\mathsf{E}[X_{n-j+1}] = \frac{1}{2j-1},\ j = 1, \ldots, n.$$

It follows that

$$\mathsf{E}[L] = \mathsf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathsf{E}[X_i] = \sum_{i=1}^{n} \frac{1}{2j-1}.$$

3. *Gaussian random vector.* Given a Gaussian random vector $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu} = (1\,5\,2)^T$ and

$$\Sigma = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

   (a) Find the pdfs of
      i. X_1,
      ii. $X_2 + X_3$,
      iii. $2X_1 + X_2 + X_3$,
      iv. X_3 given (X_1, X_2), and
      v. (X_2, X_3) given X_1.
   (b) What is $\mathsf{P}\{2X_1 + X_2 - X_3 < 0\}$? Express your answer using the Q function.
   (c) Find the joint pdf on $\mathbf{Y} = A\mathbf{X}$, where

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

Solution:

(a) i. The marginal pdfs of a jointly Gaussian pdf are Gaussian. Therefore $X_1 \sim \mathcal{N}(1, 1)$.
   ii. Since X_2 and X_3 are independent ($\sigma_{23} = 0$), the variance of the sum is the sum of the variances. Also the sum of two jointly Gaussian random variables is also Gaussian. Therefore $X_2 + X_3 \sim \mathcal{N}(7, 13)$.

iii. Since $2X_1 + X_2 + X_3$ is a linear transformation of a Gaussian random vector,

$$2X_1 + X_2 + X_3 = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix},$$

it is a Gaussian random vector with mean and variance

$$\mu = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} = 9 \;\text{ and }\; \sigma^2 = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 21\,.$$

Thus $2X_1 + X_2 + X_3 \sim \mathcal{N}(9, 21)$.

iv. Since $\sigma_{13} = 0$, X_3 and X_1 are uncorrelated and hence independent since they are jointly Gaussian; similarly, since $\sigma_{23} = 0$, X_3 and X_2 are independent. Therefore the conditional pdf of X_3 given (X_1, X_2) is the same as the pdf of X_3, which is $\mathcal{N}(2, 9)$.

v. We use the general formula for the conditional Gaussian pdf:

$$\mathbf{X}_2 \,|\, \{\mathbf{X}_1 = \mathbf{x}_1\} \sim \mathcal{N}\left(\Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + \boldsymbol{\mu}_2\,, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$$

In the case of $(X_2, X_3)\,|\,X_1$,

$$\Sigma_{11} = \begin{bmatrix} 1 \end{bmatrix}, \quad \Sigma_{21} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Sigma_{22} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}.$$

Therefore the mean and variance of (X_2, X_3) given $X_1 = x_1$ are

$$\mu_{(X_2,X_3)|X_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 - 1 \end{bmatrix} + \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 + 4 \\ 2 \end{bmatrix},$$

$$\Sigma_{(X_2,X_3)|X_1} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 9 \end{bmatrix}.$$

Thus X_2 and X_3 are conditionally independent given X_1. The conditional densities are $X_2 \,|\, \{X_1 = x_1\} \sim \mathcal{N}(x_1 + 4, 3)$ and $X_3 \,|\, \{X_1 = x\} \sim \mathcal{N}(2, 9)$.

(b) Let $Y = 2X_1 + X_2 - X_3$. Similarly as part (a)iii., $2X_1 + X_2 - X_3$ is a linear transformation of a Gaussian random vector,

$$2X_1 + X_2 - X_3 = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix},$$

it is a Gaussian random vector with mean and variance

$$\mu = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} = 5 \;\text{ and }\; \sigma^2 = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = 21\,.$$

Thus $2X_1 + X_2 - X_3 \sim \mathcal{N}(5, 21)$, i.e., $Y \sim \mathcal{N}(5, 21)$. Thus

$$\mathsf{P}\{Y < 0\} = \mathsf{P}\left\{\frac{(Y-5)}{\sqrt{21}} < \frac{(0-5)}{\sqrt{21}}\right\} = Q\left(\frac{5}{\sqrt{21}}\right).$$

(c) In general, $A\mathbf{X} \sim \mathcal{N}(A\mu_{\mathbf{X}},\, A\Sigma_{\mathbf{X}}A^T)$. For this problem,

$$\mu_{\mathbf{Y}} = A\mu_{\mathbf{X}} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ -2 \end{bmatrix},$$

$$\Sigma_{\mathbf{Y}} = A\Sigma_{\mathbf{X}}A^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 21 & 6 \\ 6 & 12 \end{bmatrix}.$$

Thus $\mathbf{Y} \sim \mathcal{N}\left(\begin{bmatrix} 9 \\ -2 \end{bmatrix}, \begin{bmatrix} 21 & 6 \\ 6 & 12 \end{bmatrix}\right)$.

4. *Gaussian Markov chain.* Let X, Y, and Z be jointly Gaussian random variables with zero mean and unit variance, i.e., $\mathsf{E}(X) = \mathsf{E}(Y) = \mathsf{E}(Z) = 0$ and $\mathsf{E}(X^2) = \mathsf{E}(Y^2) = \mathsf{E}(Z^2) = 1$. Let $\rho_{X,Y}$ denote the correlation coefficient between X and Y, and let $\rho_{Y,Z}$ denote the correlation coefficient between Y and Z. Suppose that X and Z are conditionally independent given Y.

(a) Find $\rho_{X,Z}$ in terms of $\rho_{X,Y}$ and $\rho_{Y,Z}$.

(b) Find the MMSE estimate of Z given (X, Y) and the corresponding MSE.

Solution:

(a) From the definition of $\rho_{X,Z}$, we have

$$\rho_{X,Z} = \frac{\mathrm{Cov}(X, Z)}{\sigma_X \sigma_Z},$$

where,

$$\begin{aligned} \mathrm{Cov}(X, Z) &= \mathsf{E}(XZ) - \mathsf{E}(X)\mathsf{E}(Z) = \mathsf{E}(XZ) - 0 = \mathsf{E}(XZ), \\ \sigma_X &= \mathsf{E}(X^2) - \mathsf{E}(X)^2 = 1 - 0 = 1, \\ \sigma_Y &= \mathsf{E}(Y^2) - \mathsf{E}(Y)^2 = 1 - 0 = 1. \end{aligned}$$

Thus, $\rho_{X,Z} = \mathsf{E}(XZ)$.
Moreover, since X and Z are conditionally independent given Y,

$$\mathsf{E}(XZ) = \mathsf{E}(\mathsf{E}(XZ|Y)) = \mathsf{E}[\mathsf{E}(X|Y)\mathsf{E}(Z|Y)].$$

Now $\mathsf{E}(X|Y)$ can be easily calculated from the bivariate Gaussian conditional density

$$\mathsf{E}(X|Y) = \mathsf{E}(X) + \frac{\rho_{X,Y}\sigma_X}{\sigma_Y}(Y - \mathsf{E}(Y)) = \rho_{X,Y}Y.$$

Similarly, we have

$$\mathsf{E}(Z|Y) = \rho_{Y,Z}Y.$$

Therefore, combining the above,

$$\begin{aligned}
\rho_{X,Z} &= \mathsf{E}(XZ) \\
&= \mathsf{E}[\mathsf{E}(X|Y)\mathsf{E}(Z|Y)] \\
&= \mathsf{E}(\rho_{X,Y}\rho_{Y,Z}Y^2) \\
&= \rho_{X,Y}\rho_{Y,Z}\mathsf{E}(Y^2) \\
&= \rho_{X,Y}\rho_{Y,Z}.
\end{aligned}$$

(b) X, Y and Z are jointly Gaussian random variables. Thus, the minimum MSE estimate of Z given (X, Y) is linear.

$$\begin{aligned}
\Sigma_{(X,Y)^T} &= \begin{bmatrix} 1 & \rho_{X,Y} \\ \rho_{X,Y} & 1 \end{bmatrix}, \\
\Sigma_{(X,Y)^T Z} &= \begin{bmatrix} \mathsf{E}(XZ) \\ \mathsf{E}(YZ) \end{bmatrix} = \begin{bmatrix} \rho_{X,Z} \\ \rho_{Y,Z} \end{bmatrix}, \\
\Sigma_{Z(X,Y)^T} &= \begin{bmatrix} \rho_{X,Z} & \rho_{Y,Z} \end{bmatrix}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\hat{Z} &= \Sigma_{Z(X,Y)^T}\Sigma^{-1}_{(X,Y)^T}\begin{bmatrix} X \\ Y \end{bmatrix} \\
&= \begin{bmatrix} \rho_{X,Z} & \rho_{Y,Z} \end{bmatrix}\begin{bmatrix} 1 & \rho_{X,Y} \\ \rho_{X,Y} & 1 \end{bmatrix}^{-1}\begin{bmatrix} X \\ Y \end{bmatrix} \\
&= \begin{bmatrix} \rho_{X,Z} & \rho_{Y,Z} \end{bmatrix}\frac{1}{1-\rho^2_{X,Y}}\begin{bmatrix} 1 & -\rho_{X,Y} \\ -\rho_{X,Y} & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} \\
&= \frac{1}{1-\rho^2_{X,Y}}\begin{bmatrix} 0 & -\rho^2_{X,Y}\rho_{Y,Z} + \rho_{Y,Z} \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix},
\end{aligned}$$

where the last equality follows from the result of (a). Thus,

$$\hat{Z} = \begin{bmatrix} 0 & \rho_{Y,Z} \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} = \rho_{Y,Z}Y.$$

The corresponding MSE is

$$\begin{aligned}
\text{MSE} &= \Sigma^2_Z - \Sigma_{Z(X,Y)^T}\Sigma^{-1}_{(X,Y)^T}\Sigma_{(X,Y)^T Z} \\
&= 1 - \begin{bmatrix} \rho_{X,Z} & \rho_{Y,Z} \end{bmatrix}\begin{bmatrix} 1 & \rho_{X,Y} \\ \rho_{X,Y} & 1 \end{bmatrix}^{-1}\begin{bmatrix} \rho_{X,Z} \\ \rho_{Y,Z} \end{bmatrix} \\
&= 1 - \begin{bmatrix} \rho_{X,Z} & \rho_{Y,Z} \end{bmatrix}\frac{1}{1-\rho^2_{X,Y}}\begin{bmatrix} 1 & -\rho_{X,Y} \\ -\rho_{X,Y} & 1 \end{bmatrix}\begin{bmatrix} \rho_{X,Z} \\ \rho_{Y,Z} \end{bmatrix} \\
&= 1 - \begin{bmatrix} 0 & \rho_{Y,Z} \end{bmatrix}\begin{bmatrix} \rho_{X,Z} \\ \rho_{Y,Z} \end{bmatrix} \\
&= 1 - \rho^2_{Y,Z}.
\end{aligned}$$

5. *Sufficient statistic.* The bias of a coin is a random variable $P \sim \mathrm{U}[0,1]$. Let $Z_1, Z_2, \ldots, Z_{10}$ be the outcomes of 10 coin flips. Thus $Z_i \sim \mathrm{Bern}(P)$ and $Z_1, Z_2, \ldots, Z_{10}$ are conditionally independent given P. If X is the total number of heads, then $X|\{P = p\} \sim \mathrm{Binom}(10, p)$. Assuming that the total number of heads is 9, show that

$$f_{P|Z_1,Z_2,\ldots,Z_{10}}(p|z_1, z_2, \ldots, z_{10}) = f_{P|X}(p|9)$$

is independent of the order of the outcomes.

Solution: We have

$$\begin{aligned}
&f_{P|Z_1,Z_2,\ldots,Z_{10}}(p|z_1, z_2, \ldots, z_{10}) \\
&= \frac{p_{Z_1,\ldots Z_{10}|P}(z_1, \ldots z_{10}|p) f_P(p)}{\displaystyle\int_0^1 p_{Z_1,\ldots Z_{10}|P}(z_1, \ldots z_{10}|u) f_P(u) du} \\
&= \frac{p^9(1-p)}{\displaystyle\int_0^1 u^9(1-u) du} \\
&= 110 p^9(1-p).
\end{aligned}$$

(Note that, given P, the probability of any particular realization $z_1, \ldots z_{10}$with 9 heads is $P^9(1-P)$)

Given P, we see that $X \sim \mathrm{Bin}(10, P)$. Thus,

$$\begin{aligned}
&f_{P|X}(p|9) \\
&= \frac{p_{X|P}(9|p) f_P(p)}{\displaystyle\int_0^1 p_{X|P}(9|u) f_P(u) du} \\
&= \frac{\binom{10}{9} p^9(1-p)}{\displaystyle\int_0^1 \binom{10}{9} u^9(1-u) du} \\
&= 110 p^9(1-p).
\end{aligned}$$

Thus, assuming there are 9 heads, $f_{P|Z_1,Z_2,\ldots,Z_{10}}(p|z_1, z_2, \ldots, z_{10}) = f_{P|X}(p|9)$.

Thus, knowing the total number of heads conveys exactly the same amount of information about P, as knowing all the individual outcomes.

6. *Noise cancellation.* A classical problem in statistical signal processing involves estimating a weak signal (e.g., the heart beat of a fetus) in the presence of a strong interference (the heart beat of its mother) by making two observations; one with the weak signal present and one without (by placing one microphone on the mother's belly and another close to her heart). The observations can then be combined to estimate the weak signal by "canceling out" the interference. The following is a simple version of this application.

   Let the weak signal X be a random variable with mean μ and variance P, and the observations be $Y_1 = X + Z_1$ (Z_1 being the strong interference), and $Y_2 = Z_1 + Z_2$ (Z_2 is a measurement noise), where Z_1 and Z_2 are zero mean with variances N_1 and N_2, respectively. Assume that X, Z_1 and Z_2 are uncorrelated. Find the MMSE linear estimate of X given Y_1 and Y_2 and its MSE. Interpret the results.

Solution:

The LMMSE estimator is given by

$$\hat{X} = \mathbf{h}^\top(\mathbf{Y} - \mathsf{E}[\mathbf{Y}]) + \mathsf{E}[X]$$

where $\Sigma_{\mathbf{Y}X} = \Sigma_{\mathbf{Y}}\mathbf{h}$.

Now, $\mathsf{E}[Y_1] = \mu$ and $\mathsf{E}[Y_2] = 0$ from the definitions.

Also,

$$\begin{aligned} \mathrm{Cov}(X, Y_1) &= \mathrm{Cov}(X, X) + \mathrm{Cov}(X, Z_1) = P. \\ \mathrm{Cov}(X, Y_2) &= \mathrm{Cov}(X, Z_1) + \mathrm{Cov}(X, Z_2) = 0. \end{aligned}$$

And,

$$\begin{aligned} \mathrm{Var}(Y_1) &= \mathrm{Var}(X) + \mathrm{Var}(Z_1) = P + N_1 \quad (X, Z_1 \text{ are uncorrelated}) \\ \mathrm{Var}(Y_2) &= \mathrm{Var}(Z_1) + \mathrm{Var}(Z_2) = N_1 + N_2 \quad (Z_1, Z_1 \text{ are uncorrelated}) \\ \mathrm{Cov}(Y_1, Y_2) &= \mathrm{Cov}(X, Z_1) + \mathrm{Cov}(X, Z_2) + \mathrm{Cov}(Z_1, Z_1) + \mathrm{Cov}(Z_1, Z_2) = N_1. \end{aligned}$$

So,

$$\Sigma_{\mathbf{Y}X} = \begin{bmatrix} P \\ 0 \end{bmatrix}$$

$$\Sigma_{\mathbf{Y}} = \begin{bmatrix} P + N_1 & N_1 \\ N_1 & N_1 + N_2 \end{bmatrix} \qquad \Sigma_{\mathbf{Y}}^{-1} = \frac{1}{P(N_1 + N_2) + N_1 N_2} \begin{bmatrix} N + N_2 & -N_1 \\ -N_1 & P + N_1 \end{bmatrix}$$

So $\mathbf{h}$ is he product of P and the first column of $\Sigma_{\mathbf{Y}}^{-1}$:

$$\mathbf{h} = \Sigma_{\mathbf{Y}}^{-1}\Sigma_{\mathbf{Y}X} = \Sigma_{\mathbf{Y}}^{-1} \begin{bmatrix} P \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{P(N_1+N_2)}{P(N_1+N_2)+N_1N_2} \\ \frac{-N_1 P}{P(N_1+N_2)+N_1N_2} \end{bmatrix}$$

Therefore,

$$\hat{X} = \frac{1}{P(N_1 + N_2) + N_1 N_2}[P(N_1 + N_2)(Y_{-}\mu) - PN_1 Y_2) + \mu.$$

The MSE is given by

$$\begin{aligned} MSE &= \sigma_X^2 - \Sigma_{\mathbf{Y}X}^\top \Sigma_{\mathbf{Y}}^{-1} \Sigma_{\mathbf{Y}X} \\ &= P - [P\ 0]\mathbf{h} \\ &= P - \frac{P^2(N_1 + N_2)}{P(N_1 + N_2) + N_1 N_2} = \frac{PN_1N_2}{P(N_1 + N_2) + N_1 N_2} \end{aligned}$$

Interpretation: Assume $N_2 \ll N1$, so $N_1 + N_2 \approx N_1$ and $P(N + 1 + N_2) \approx PN_1$.

Then,

$$\hat{X} \approx= \frac{1}{(P+N_2)N_1}[PN_1(Y_{-}\mu) - PN_1Y_2) + \mu$$
$$\approx \frac{P}{P+N_2}[Y_1 - Y_2 - \mu] + \mu$$
$$\approx Y_1 - Y_2.$$

and

$$MSE \approx \frac{PN_1N_2}{PN_1 + N_1N_2} = \frac{PN_2}{P+N_2}.$$

7. *Nonlinear estimator.* Consider a channel with the observation $Y = XZ$, where the signal X and the noise Z are uncorrelated jointly Gaussian random variables. Let $E[X] = 1$, $E[Z] = 2$, $\sigma_X^2 = 5$, and $\sigma_Z^2 = 8$.

   (a) Using the fact that $\mathsf{E}(W^3) = \mu^3 + 3\mu\sigma^2$ and $\mathsf{E}(W^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$ for $W \sim \mathcal{N}(\mu, \sigma^2)$, find the mean and covariance matrix of $[X \;\; Y \;\; Y^2]^T$.

   (b) Find the MMSE linear estimate of X given Y and the corresponding MSE.

   (c) Find the MMSE linear estimate of X given Y^2 and the corresponding MSE.

   (d) Find the MMSE linear estimate of X given Y and Y^2 and the corresponding MSE.

   (e) Compare your answers in parts (b) through (d). Is the MMSE estimate of X given Y (namely, $\mathsf{E}(X|Y)$) linear?

Solution:

(a) Since X and Z are uncorrelated jointly Gaussian random variables, they are independent. We have

$$\mathsf{E}(X^2) = \sigma_X^2 + \mathsf{E}^2(X) = 5 + 1 = 6,$$
$$\mathsf{E}(X^3) = 1 + 3 \times 1 \times 5 = 16,$$
$$\mathsf{E}(X^4) = 1 + 6 \times 1 \times 5 + 3 \times 25 = 106.$$

$$\mathsf{E}(Z^2) = \sigma_Z^2 + \mathsf{E}^2(Z) = 8 + 4 = 12,$$
$$\mathsf{E}(Z^3) = 2 + 3 \times 2 \times 8 = 50,$$
$$\mathsf{E}(Z^4) = 2^4 + 6 \times 4 \times 8 + 3 \times 64 = 400.$$

Since X and Z are independent, we have

$$\mathsf{E}(Y) = \mathsf{E}(XZ) = \mathsf{E}(X)\mathsf{E}(Z) = 2,$$
$$\mathsf{E}(Y^2) = \mathsf{E}(X^2Z^2) = \mathsf{E}(X^2)\mathsf{E}(Z^2) = 6 \times 12 = 72,$$
$$\mathsf{E}(Y^3) = \mathsf{E}(X^3Z^3) = \mathsf{E}(X^3)\mathsf{E}(Z^3) = 16 \times 50 = 800,$$
$$\mathsf{E}(Y^4) = \mathsf{E}(X^4)\mathsf{E}(Z^4) = 106 \times 400 = 42400.$$

Therefore, the mean of $[X \;\; Y \;\; Y^2]^T$ is $[1 \;\; 2 \;\; 72]^T$.

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathsf{E}(Y^2) - \mathsf{E}^2(Y) = 72 - 4 = 68, \\
\mathrm{Var}(Y^2) &= \mathsf{E}(Y^4) - \mathsf{E}^2(Y^2) = 37216, \\
\mathrm{Cov}(X,Y) &= \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) = \mathsf{E}(X^2)\mathsf{E}(Z) - \mathsf{E}(X)\mathsf{E}(Y) = 10, \\
\mathrm{Cov}(X,Y^2) &= \mathsf{E}(XY^2) - \mathsf{E}(X)\mathsf{E}(Y^2) = \mathsf{E}(X^3)\mathsf{E}(Z^2) - \mathsf{E}(X)\mathsf{E}(Y^2) = 120, \\
\mathrm{Cov}(Y,Y^2) &= \mathsf{E}(YY^2) - \mathsf{E}(Y)\mathsf{E}(Y^2) = \mathsf{E}(X^3)\mathsf{E}(Z^3) - \mathsf{E}(Y)\mathsf{E}(Y^2) = 656.
\end{aligned}$$

Therefore, the covariance matrix of $[X \;\; Y \;\; Y^2]^T$ is

$$\begin{bmatrix} 5 & 10 & 120 \\ 10 & 68 & 656 \\ 120 & 656 & 37216 \end{bmatrix}.$$

(b) The MMSE linear estimate of X given Y is

$$\hat{X} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y - \mathsf{E}(Y)) + \mathsf{E}(X) = \frac{10}{68}(Y-2) + 1 = \frac{5}{34}Y + \frac{24}{34},$$

and its MSE is given by

$$\mathrm{MSE} = \mathrm{Var}(X) - \frac{\mathrm{Cov}^2(X,Y)}{\mathrm{Var}(Y)} = 5 - \frac{100}{68} = 3.5294.$$

(c) The MMSE linear estimate of X given Y^2 is

$$\hat{X} = \frac{\mathrm{Cov}(X,Y^2)}{\mathrm{Var}(Y^2)}(Y^2 - \mathsf{E}(Y^2)) + \mathsf{E}(X) = \frac{120}{37216}(Y^2 - 72) + 1 = \frac{15}{4652}Y^2 + \frac{893}{1163},$$

and its MSE is given by

$$\mathrm{MSE} = \mathrm{Var}(X) - \frac{\mathrm{Cov}^2(X,Y^2)}{\mathrm{Var}(Y^2)} = 5 - \frac{14400}{37216} = 4.6131.$$

(d) We first normalize the random variables by subtracting off their means to get

$$\begin{aligned}
X' &= X - \mathsf{E}(X) = X - 2, \\
Y' &= Y - \mathsf{E}(Y) = Y - 2, \\
Y'^2 &= Y^2 - \mathsf{E}(Y^2) = Y^2 - 72.
\end{aligned}$$

Using the covariance matrix in part a, we have

$$\begin{aligned}
\Sigma_{[Y \;\; Y^2]^T X} &= [10 \;\; 120]^T, \\
\Sigma_{[Y \;\; Y^2]^T} &= \begin{bmatrix} 68 & 656 \\ 656 & 37216 \end{bmatrix}.
\end{aligned}$$

Therefore,

$$\hat{X}' = \Sigma^T_{[Y\ Y^2]^T X} \Sigma^{-1}_{[Y\ Y^2]^T} \begin{bmatrix} Y' \\ Y'^2 \end{bmatrix} = 0.1397Y' + 0.0008Y'^2,$$

and hence

$\hat{X} = \hat{X}' + \mathsf{E}(X) = 0.1397(Y-2) + 0.0008(Y^2 - 72) + 1 = 0.1397Y + 0.0008Y^2 + 0.663.$

The corresponding MSE is given by

$$\text{MSE} = \text{Var}(X) - \Sigma^T_{[Y\ Y^2]^T X} \Sigma^{-1}_{[Y\ Y^2]^T} \Sigma_{[Y\ Y^2]^T X} = 3.5115.$$

(e) MSE linear estimate of X given Y and Y^2 results in the minimum MSE among the three. Therefore, MSE linear estimate of X given Y does not have the minimum MSE and MMSE estimate of X given Y is not linear.

8. *Sample mean convergence.* Consider the sequence of i.i.d. random variables $X_1, X_2, \ldots$ with

$$X_i = \begin{cases} 0 & \text{w.p. } \frac{1}{2}, \\ 2 & \text{w.p. } \frac{1}{2}, \end{cases}$$

for all $i \geq 1$.

Define the sequence

$$Y_n = \begin{cases} X_n, & \text{for all } n \quad \text{w.p. } \frac{1}{3}, \\ \frac{1}{2}X_n, & \text{for all } n \quad \text{w.p. } \frac{1}{3}, \\ 0, & \text{for all } n \quad \text{w.p. } \frac{1}{3}. \end{cases}$$

Let

$$M_n = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

(a) Determine the probability mass function (pmf) of Y_n.

(b) Determine the random variable (or constant) that M_n converges to (in probability) as n approaches infinity. Justify your answer.

(c) Use the central limit theorem to estimate the probability that the random variable M_{84} exceeds $\frac{2}{3}$.

Solution:

(a) Determine the probability mass function (pmf) of Y_n.

$Y_n \sim Y$, where Y has the pmf

$$\begin{aligned} p_Y(y) &= \mathsf{P}(Y{=}y|Y{=}X)\mathsf{P}(Y{=}X) + \mathsf{P}(Y{=}y|Y = \frac{X}{2})\mathsf{P}(Y{=}\frac{X}{2}) \\ &\quad + \mathsf{P}(Y{=}y|Y{=}0)\mathsf{P}(Y{=}0) \\ &= p_X(y)\frac{1}{3} + p_X(2y)\frac{1}{3} + p_0(y)\frac{1}{3} \end{aligned}$$

$$= \begin{cases} 0 & \text{w.p. } \frac{2}{3} \\ 1 & \text{w.p. } \frac{1}{6} \\ 2 & \text{w.p. } \frac{1}{6} \end{cases}$$

(b) Determine the random variable (or constant) that M_n converges to (in probability) as n approaches infinity. Justify your answer.

M_n is the sample mean of $Y_1, Y_2, \ldots, Y_n$.

By the weak law of large numbers (WLLN), it converges in probability to $\mathsf{E}[Y]$, where

$$\begin{aligned} \mathsf{E}[Y] &= \sum_{y \in \mathcal{Y}} y p_Y(y) \\ &= 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} \\ &= \frac{1}{2}. \end{aligned}$$

(c) Use the central limit theorem to estimate the probability that the random variable M_{84} exceeds $\frac{2}{3}$.

By the Central Limit Theorem (CLT)

$$\frac{M_n - \mathsf{E}[Y]}{\sigma_Y/\sqrt{n}} \to Z \sim \mathcal{N}(0,1).$$

$$\begin{aligned} \mathsf{E}[Y^2] &= \sum_{y^2 \in \mathcal{Y}} y^2 p_Y(y) \\ &= 0^2 \cdot \frac{2}{3} + 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} \\ &= \frac{5}{6}. \end{aligned}$$

So,

$$\mathrm{Var}(Y) = \mathsf{E}[Y^2] - (\mathsf{E}[Y])^2 = \frac{5}{6} - (\frac{1}{2})^2 = \frac{7}{12}$$

and

$$\sigma_Y = \sqrt{\left(\frac{7}{12}\right)}.$$

$$\mathsf{P}(M_n > x) = \mathsf{P}\left(\frac{M_n - \mathsf{E}[Y]}{\sigma_Y/\sqrt{n}} > \frac{x - \frac{1}{2}}{\sqrt{7/12n}}\right).$$

Setting $n = 84$ and $x = \frac{2}{3}$, we get

$$\mathsf{P}(M_n > \frac{2}{3}) \approx \mathsf{P}(Z > 2) = \mathsf{Q}(2).$$

9. *Minimum waiting time.* Let $X_1, X_2, \ldots$ be i.i.d. exponentially distributed random variables with parameter λ, i.e., $f_{X_i}(x) = \lambda e^{-\lambda x}$, for $x \geq 0$.

(a) Show that $Y_n = \min\{X_1, X_2, \ldots, X_n\}$ converges in probability as n approaches infinity. What is the limit?

(b) Does $Z_n = nY_n$ converge to the same limit in probability?

Solution:

(a) For any set of values X_i's, the sequence Y_n is monotonically decreasing in n. Since the random variables are nonnegative, it is reasonable to guess that Y_n converges to 0. Now Y_n will converge in probability to 0 if and only if for any $\epsilon > 0$, $\lim_{n\to\infty} P\{|Y_n| > \epsilon\} = 0$.

$$\begin{aligned}
P\{|Y_n| > \epsilon\} &= P\{Y_n > \epsilon\} \\
&= P\{X_1 > \epsilon, X_2 > \epsilon, .., X_n > \epsilon\} \\
&= P\{X_1 > \epsilon\}P\{X_2 > \epsilon\}...P\{X_n > \epsilon\} \\
&= (1 - F_X(\epsilon))(1 - F_X(\epsilon))...(1 - F_X(\epsilon)) \\
&= (1 - F_X(\epsilon))^n \\
&= (1 - (1 - e^{-\lambda\epsilon}))^n \\
&= e^{-\lambda\epsilon n}.
\end{aligned}$$

As n goes to infinity (for any finite $\epsilon > 0$) this converges to zero. Therefore Y_n converges to 0 in probability.

(b) Does $Z_n = nY_n$ converges to 0 in probability ? No, it does not. In fact,

$$\begin{aligned}
P\{|Z_n| > \epsilon\} &= P\{nY_n > \epsilon\} \\
&= P\{Y_n > \frac{\epsilon}{n}\} \\
&= e^{-\lambda\frac{\epsilon}{n}n} \\
&= e^{-\lambda\epsilon}
\end{aligned}$$

which does not depend on n. So Z_n does not converge to 0 in probability. Note that the distribution of Z_n is exponential with parameter λ, the same as the distribution of X_i.

$$\begin{aligned}
F_{Z_n}(z) &= P\{Z_n < z\} \\
&= 1 - e^{-\lambda z}.
\end{aligned}$$

In conclusion, if X_i's are i.i.d. $\sim \exp(\lambda)$, then

$$Y_n = \min_{1\le i\le n}\{X_i\} \sim \exp(n\lambda)$$

and

$$Z_n = nY_n \sim \exp(\lambda).$$

Thus

$$P\{Y_n > \epsilon\} = e^{-\lambda\epsilon n} \to 0, \text{ so } Y_n \to 0 \text{ in probability,}$$

but

$$P\{Z_n > \epsilon\} = e^{-\lambda\epsilon} \not\to 0, \text{ so } Z_n \not\to 0 \text{ in probability.}$$

10. *Roundoff errors.* The sum of a list of 200 real numbers is to be computed. Suppose that these numbers are rounded off to the nearest integer so that each number has an error that is uniformly distributed in the interval $(-0.5, 0.5)$. Use the central limit theorem to estimate the probability that the total error in the sum of the 200 numbers exceeds 10.

Solution: Errors are independent and uniformly distributed in $(-0.5, 0.5)$, i.e. $e_i \sim \text{Unif}(-0.5, 0.5)$. Using the Central Limit Theorem, the total error $e = \sum_{i=1}^{200} e_i$ can be approximated as a Gaussian random Variable with mean

$$\begin{aligned} E(e) &= \sum_{i=1}^{200} E(e_i) \\ &= 0 \end{aligned}$$

and variance

$$\begin{aligned} \sigma_e^2 &= \sum_{i=1}^{200} \text{Var}(e_i) \\ &= 200 \cdot \frac{1}{12} \\ &\approx 16.66. \end{aligned}$$

Therefore the probability of the given event is

$$\begin{aligned} P\{e > 10\} &= P\left\{ \frac{e - E(e)}{\sigma_e} > \frac{10 - 0}{\sqrt{16.66}} \right\} \\ &\approx Q\left(\frac{10}{\sqrt{16.66}} \right) \\ &= 0.0072. \end{aligned}$$

If the error is interpreted as an absolute difference, then

$$\begin{aligned} P\{|e| > 10\} &= 2 \cdot P\{e > 10\} \\ &\approx 0.0144. \end{aligned}$$

11. The signal received over a wireless communication channel can be represented by two sums

$$\begin{aligned} X_{1n} &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Z_j \cos \Theta_j, \text{ and} \\ X_{2n} &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Z_j \sin \Theta_j, \end{aligned}$$

where $Z_1, Z_2, \ldots$ are i.i.d. with mean μ and variance σ^2 and $\Theta_1, \Theta_2, \ldots$ are i.i.d. $\text{U}[0, 2\pi]$ independent of $Z_1, Z_2, \ldots$. Find the distribution of $\begin{bmatrix} X_{1n} \\ X_{2n} \end{bmatrix}$ as n approaches ∞.

Solution: Since $Z_1, Z_2, \ldots$ are i.i.d. and $\Theta_1, \Theta_2, \ldots$ are also i.i.d. and independent of $Z_1, Z_2, \ldots$,

$\begin{bmatrix} Z_1 \cos\Theta_1 \\ Z_1 \sin\Theta_1 \end{bmatrix}$, $\begin{bmatrix} Z_2 \cos\Theta_2 \\ Z_2 \sin\Theta_2 \end{bmatrix}$, ... are i.i.d. random vectors.

Let $\mathbf{Y}_j = \begin{bmatrix} Z_j \cos\Theta_j \\ Z_j \sin\Theta_j \end{bmatrix}$. We have

$$\begin{aligned} \mathsf{E}[Z_j \cos\Theta_j] &= \mathsf{E}[Z_j]\mathsf{E}[\cos\Theta_j] \\ &= \mu \int_0^{2\pi} \frac{1}{2\pi} \cos\theta d\theta \\ &= 0. \end{aligned}$$

Similarly, $\mathsf{E}[Z_j \sin\Theta_j] = 0$. Thus, $\mathsf{E}[\mathbf{Y}_j] = \mathbf{0}$.

Moreover, we have

$$\begin{aligned} \mathsf{E}[(Z_j \cos\Theta_j)^2] &= \mathsf{E}[Z_j^2]\mathsf{E}[\cos^2\Theta_j] \\ &= (\mu^2+\sigma^2) \int_0^{2\pi} \frac{1}{2\pi} \cos^2\theta d\theta \\ &= \frac{\mu^2+\sigma^2}{2}. \end{aligned}$$

Similarly, $\mathsf{E}[(Z_j \sin\Theta_j)^2] = \dfrac{\mu^2+\sigma^2}{2}$. Also,

$$\begin{aligned} \mathsf{E}[(Z_j \cos\Theta_j)\cdot(Z_j \sin\Theta_j)] &= \mathsf{E}[Z_j^2]\mathsf{E}[\cos\Theta_j \sin\Theta_j] \\ &= (\mu^2+\sigma^2) \int_0^{2\pi} \frac{1}{2\pi} \cos\theta \sin\theta d\theta \\ &= 0. \end{aligned}$$

Thus,

$$\begin{aligned} \mathrm{Cov}(\mathbf{Y}_j) &= \mathsf{E}[\mathbf{Y}_j\mathbf{Y}_j^T] - \mathsf{E}[\mathbf{Y}_j]\mathsf{E}[\mathbf{Y}_j]^T \\ &= \mathsf{E}[\mathbf{Y}_j\mathbf{Y}_j^T] \\ &= \begin{bmatrix} \mathsf{E}[(Z_j \cos\Theta_j)^2] & \mathsf{E}[(Z_j \cos\Theta_j)\cdot(Z_j \sin\Theta_j)] \\ \mathsf{E}[(Z_j \cos\Theta_j)\cdot(Z_j \sin\Theta_j)] & \mathsf{E}[(Z_j \sin\Theta_j)^2] \end{bmatrix} \\ &= \frac{\mu^2+\sigma^2}{2}\mathbf{I}. \end{aligned}$$

Now, $\begin{bmatrix} X_{1n} \\ X_{2n} \end{bmatrix} = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{j=1}^{n} \mathbf{Y}_j$, and thus, by the central limit theorem, as $n \to \infty$,

$$\begin{bmatrix} X_{1n} \\ X_{2n} \end{bmatrix} \longrightarrow \mathcal{N}\Big(\mathbf{0}, \frac{\mu^2+\sigma^2}{2}\mathbf{I}\Big) \text{ in distribution.}$$

12. *Convergence.* Consider the following sequences of random variables defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$, where $\Omega = \{0, 1, 2, \ldots, m-1\}$, $\mathcal{F}$ is the collection of all subsets of Ω, and P is the uniform distribution over Ω.

$$X_n(w) = \begin{cases} \frac{1}{n}, & \omega = n \bmod m \\ 0, & \text{otherwise} \end{cases}$$

$$Y_n(w) = \begin{cases} 2^n, & \omega = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$Z_n(w) = \begin{cases} 1, & \omega = 1 \\ 0, & \text{otherwise} \end{cases}$$

Which of these sequences converges to zero

(a) with probability one?

(b) in mean square?

(c) in probability?

Solution:

The random variables are defined as functions from the set of outcomes Ω of a probability space $[\Omega, \mathcal{F}, P]$ to the reals, $X : \Omega \to \mathbb{R}$. From these definitions, we find the pmfs to be:

$$X_n = \begin{cases} \frac{1}{n} & \text{w.p. } \frac{1}{m}, \\ 0 & \text{w.p. } \frac{m-1}{m}, \end{cases}$$

for all $n \geq 1$.

$$Y_n = \begin{cases} 2^n & \text{w.p. } \frac{1}{m}, \\ 0 & \text{w.p. } \frac{m-1}{m}, \end{cases}$$

for all $n \geq 1$.

$$Z_n = \begin{cases} 1 & \text{w.p. } \frac{1}{m}, \\ 0 & \text{w.p. } \frac{m-1}{m}, \end{cases}$$

for all $n \geq 1$.

1. with probability one?

   Let $\epsilon < 1$. For $M > \lceil \frac{1}{\epsilon} \rceil$,

$$\mathsf{P}(|X_n - 0| < \epsilon, \text{ for all } n \geq M) = \lim_{n \to \infty} \prod_{i=M}^{n} 1 = 1$$

   so X_n converges almost surely (i.e., with probability one) to zero.

For $\epsilon < 1$,

$$\mathsf{P}(|Y_n - 0| < \epsilon, \text{ for all} \mathrm{n} \geq \mathrm{M}) = \lim_{\mathrm{n=M}}^{\mathrm{n}} \prod_{\mathrm{i=M}}^{\mathrm{n}} \frac{\mathrm{m}-1}{\mathrm{m}} = 0$$

so Y_n **does not** converge almost surely (i.e., with probability one) to zero.

For $\epsilon < 1$,

$$\mathsf{P}(|Z_n - 0| < \epsilon, \text{ for all} \mathrm{n} \geq \mathrm{M}) = \lim_{\mathrm{n=M}}^{\mathrm{n}} \prod_{\mathrm{i=M}}^{\mathrm{n}} \frac{\mathrm{m}-1}{\mathrm{m}} = 0$$

so Z_n **does not** converge almost surely (i.e., with probability one) to zero.

2. in mean square?

For $\epsilon < 1$,

$$\mathsf{E}[(X_n - 0)^2)] = \mathsf{E}[X_n^2] = \frac{1}{n^2} \cdot \frac{1}{m}$$

so $\lim_{n\to\infty} \mathsf{E}[(X_n - 0)^2)] = 0$ and X_n converges in mean square to zero.

For $\epsilon < 1$,

$$\mathsf{E}[(Y_n - 0)^2)] = \mathsf{E}[Y_n^2] = 2^{2n}\frac{1}{m}$$

so $\lim_{n\to\infty} \mathsf{E}[(Y_n - 0)^2)] = \infty$ and Y_n **does not** converge in mean square to zero.

For $\epsilon < 1$,

$$\mathsf{E}[(Z_n - 0)^2)] = \mathsf{E}[Z_n^2] = \frac{1}{m}$$

so $\lim_{n\to\infty} \mathsf{E}[(Z_n - 0)^2)] = \frac{1}{m}$ and Z_n **does not** converge in mean square to zero.

3. in probability?

For $\epsilon < 1$,

$$\mathsf{P}(|X_n - 0| < \epsilon) = 1 \text{ for n} > \lceil \frac{1}{\epsilon} \rceil$$

so $\lim_{n\to\infty} \mathsf{P}(|X_n - 0| < \epsilon) = 1$ and X_n converges in probability to zero.

For $\epsilon < 1$,

$$\mathsf{P}(|Y_n - 0| < \epsilon) = \frac{m-1}{m}$$

so $\lim_{n\to\infty} \mathsf{P}(|Y_n - 0| < \epsilon) = \frac{m-1}{m} < 1$ and Y_n **does not** converge in probability to zero.

For $\epsilon < 1$,

$$\mathsf{P}(|Z_n - 0| < \epsilon) = \frac{m-1}{m}$$

so $\lim_{n\to\infty} \mathsf{P}(|Z_n - 0| < \epsilon) = \frac{m-1}{m} < 1$ and Z_n **does not** converge in probability to zero.

Note: The failure of Y_n and Z_n to converge almost surely and in mean square also follows from the failure to converge in probability.