Solutions to Homework Set #5

(Prepared by TA Halyun Jeong)

1. Work on the midterm problems to make sure you understand everything clearly.
2. Read Sections 6.1–6.5 in the text. Try to work on all examples.
3. 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.

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

Solution:

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

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

(c) This can be 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 nonnegative definite. 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

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

4. 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) What is $P(2X_1 + X_2 + X_3 < 0)$?

(b) Find the joint pdf on $\mathbf{Y} = A\mathbf{X}$, where

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

Solution:

(a) Let $Y = X_1 + X_2 + 2X_3$. Now $Y \sim \mathcal{N}(9, 21)$, so

$$\begin{aligned} \mathrm{P}(X_1 + X_2 + 2X_3 < 0) &= \mathrm{P}(Y < 0) \\ &= \Phi\left(\frac{0 - \mu_Y}{\sigma_y}\right) \\ &= \Phi\left(\frac{-9}{\sqrt{21}}\right) \\ &= \Phi(-1.96) \\ &= Q(1.96) \\ &= 2.48 \times 10^{-2}. \end{aligned}$$

(b)

$$\begin{aligned} \mu_{\mathbf{Y}} &= A\mu_{\mathbf{X}} \\ &= \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \\ \Sigma_{\mathbf{Y}} &= A\Sigma_{\mathbf{X}}A^T \\ &= \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 21 & 6 \\ 6 & 12 \end{bmatrix}. \end{aligned}$$

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

5. *Packet switching.*
Let $N \sim \mathrm{P}(\lambda)$, i.e., Poisson with parameter λ, be the number of packets arriving at a switch per unit time. Each packet is routed to Output Port 1 with probability p and to Output Port 2 with probability $(1-p)$ independent of N and of other packets. Let X be the number of packets routed to Output Port 1 per unit time. Thus $X = 0$ if $N = 0$ and $X = \sum_{i=1}^{N} Z_i$ for $N > 0$, where

$$Z_i = \begin{cases} 1, & \text{packet } i \text{ routed to Port 1} \\ 0, & \text{packet } i \text{ routed to Port 2,} \end{cases}$$

and $Z_1, Z_2, \ldots, Z_N$ are conditionally independent given N.

(a) Find the mean and variance of X.

(b) Find the pmf of X. What is the pmf of $N - X$?

Solution:

(a) To find the mean of X we use iterated expectation. Note that given $N = n > 0$, $X|\{N = n\} \sim \mathrm{B}(n, p)$. Now consider

$$\begin{aligned} E(X) &= E_N(E_X(X|N)) \\ &= E_N(Np) = \lambda p. \end{aligned}$$

To find the variance consider

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - (EX)^2 \\ &= E_N(E_X(X^2|N)) - (\lambda p)^2 \\ &= E_N(Np(1-p) + N^2p^2) - (\lambda p)^2 \\ &= \lambda p(1-p) + p^2(\lambda^2 + \lambda) - (\lambda p)^2 \\ &= \lambda p. \end{aligned}$$

(b) From part (a), the mean and variance are equal to λp so we suspect that X is Poisson. To verify this let's use the law of total probability

$$\begin{aligned} p_X(x) &= \sum_{k=x}^{\infty} \mathrm{P}\{X = x|N = k\}\mathrm{P}\{N = k\} \\ &= \sum_{k=x}^{\infty} \binom{k}{x} p^x (1-p)^{(k-x)} \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \frac{(\lambda p)^x}{x!} e^{-\lambda p}, \quad \text{for } x \geq 0. \end{aligned}$$

Thus $X \sim \mathrm{Poisson}(\lambda p)$. Also $N - X \sim \mathrm{Poisson}(\lambda(1-p))$. Thus the Poisson input traffic is divided into two Poisson output traffics. This is the complement of the infinite divisibility property of the Poisson (sum of two independent Poissons is Poisson).

6. *Estimation (20 points).* Let X_1 and X_2 be independent identically distributed random variables. Let $Y = X_1 + X_2$.

(a) Find $E[X_1 - X_2|Y]$.

(b) Find the minimum mean squared error estimate of X_1 given an observed value of $Y = X_1 + X_2$. (Hint: Consider $E[X_1 + X_2|X_1 + X_2]$.)

Solution:

(a) By symmetry, $E[X_1|Y] = E[X_2|Y]$. Hence, $E[X_1 - X_2|Y] = 0$.

(b) On one hand, we have $E[X_1+X_2|X_1+X_2] = X_1+X_2 = Y$. On the other hand, we have $E[X_1 + X_2|X_1 + X_2] = E[X_1|Y] + E[X_2|Y] = 2E[X_1|Y]$. Hence, the MMSE estimate of X_1 given Y is

$$E[X_1|Y] = \frac{Y}{2}.$$

7. *Gaussian Markov chain (from Spring 2007 Final).* Let X, Y, and Z be jointly Gaussian random variables with zero mean and unit variance, i.e., $EX = EY = EZ = 0$ and $EX^2 = EY^2 = EZ^2 = 1$. 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

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

where

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

Thus,

$$\rho_{X,Z} = E(XZ)$$

Moreover, since X and Z are conditionally independent given Y,

$$\begin{aligned}E(XZ) &= E(E(XZ|Y))\\ &= E[E(X|Y)E(Z|Y)]\end{aligned}$$

$E(X|Y)$ can be easily calculated from the bivariate Gaussian conditional density (Example 6.29, page 338 text)

$$
\begin{aligned}
E(X|Y) &= E(X) + \frac{\rho_{X,Y}\sigma_X}{\sigma_Y}(Y - E(Y)) \\
&= \rho_{X,Y} Y
\end{aligned}
$$

Similarly, we have

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

Therefore, combining above,

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

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

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

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,

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

The corresponding MSE is given by

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

8. *Prediction.*
Let $\mathbf{X}$ be a random process with zero mean and covariance matrix

$$\Sigma_{\mathbf{X}} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ \alpha & 1 & \alpha & & \\ \alpha^2 & \alpha & 1 & & \\ \vdots & & & \ddots & \\ \alpha^{n-1} & & & \cdots & 1 \end{bmatrix}$$

for $|\alpha| < 1$. $X_1, X_2, \ldots, X_{n-1}$ are observed, find the best linear MSE estimate (predictor) of X_n. Compute its MSE.

Solution: Since

$$\Sigma_{\mathbf{X}} = \begin{bmatrix} 1 & \cdots & \alpha^{n-2} & \alpha^{n-1} \\ \vdots & \ddots & \vdots & \vdots \\ \alpha^{n-2} & \cdots & 1 & \alpha \\ \alpha^{n-1} & \cdots & \alpha & 1 \end{bmatrix},$$

if we define

$$\mathbf{Y} = \begin{bmatrix} X_1 \\ \vdots \\ X_{n-1} \end{bmatrix},$$

then

$$\Sigma_{\mathbf{Y}} = \begin{bmatrix} 1 & \cdots & \alpha^{n-2} \\ \vdots & \ddots & \vdots \\ \alpha^{n-2} & \cdots & 1 \end{bmatrix},$$

$$\Sigma_{\mathbf{Y}X} = \begin{bmatrix} \alpha^{n-1} \\ \vdots \\ \alpha \end{bmatrix},$$

$$\Sigma_{X\mathbf{Y}} = \begin{bmatrix} \alpha^{n-1} & \cdots & \alpha \end{bmatrix},$$

$$\sigma_x{}^2 = 1.$$

Therefore,

$$\begin{aligned} \hat{X}_n &= \Sigma_{X\mathbf{Y}} \Sigma_Y^{-1} \mathbf{Y} \\ &= \begin{bmatrix} \alpha^{n-1} & \cdots & \alpha \end{bmatrix} \begin{bmatrix} 1 & \cdots & \alpha^{n-2} \\ \vdots & \ddots & \vdots \\ \alpha^{n-2} & \cdots & 1 \end{bmatrix}^{-1} \mathbf{Y} \\ &= \mathbf{h}^T \mathbf{Y} \quad \text{where } \mathbf{h}^T = \Sigma_{X\mathbf{Y}} \Sigma_{\mathbf{Y}}^{-1} \\ &= \begin{bmatrix} 0 & \cdots & 0 & \alpha \end{bmatrix} \mathbf{Y} \quad \text{by observing } \mathbf{h}^T \Sigma_{\mathbf{Y}} = \Sigma_{X\mathbf{Y}} \\ &= \alpha X_{n-1}; \end{aligned}$$

and

$$\begin{aligned} \text{MSE} &= \sigma_x{}^2 - \Sigma_{X\mathbf{Y}} \Sigma_Y^{-1} \Sigma_{\mathbf{Y}X} \\ &= 1 - \mathbf{h}^T \Sigma_{\mathbf{Y}X} \\ &= 1 - \begin{bmatrix} 0 & \cdots & 0 & \alpha \end{bmatrix} \begin{bmatrix} \alpha^{n-1} \\ \vdots \\ \alpha \end{bmatrix} \\ &= 1 - \alpha^2. \end{aligned}$$

9. *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 "cancelling 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 best linear MSE estimate of X given Y_1 and Y_2 and its MSE. Interprete the results.

Solution: This is a vector linear MSE problem. Since Z_1 and Z_2 are zero mean, $\mu_X = \mu_{Y_1} = \mu$ and $\mu_{Y_2} = 0$. We first normalize the random variables by subtracting off their means to get $X' = X - \mu$, and

$$\mathbf{Y'} = \begin{bmatrix} Y_1 - \mu \\ Y_2 \end{bmatrix}.$$

Now using the orthogonality principle we can find the best linear MSE estimate $\hat{X}'$ of X'. To do so we first find

$$\Sigma_{\mathbf{Y}} = \begin{bmatrix} P + N_1 & N_1 \\ N_1 & N_1 + N_2 \end{bmatrix}$$

and

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

Thus,

$$\begin{aligned} \hat{X}' &= \Sigma_{\mathbf{Y}X}^T \Sigma_{\mathbf{Y}}^{-1} \mathbf{Y'} \\ &= \begin{bmatrix} P & 0 \end{bmatrix} \frac{1}{P(N_1+N_2)+N_1N_2} \begin{bmatrix} N_1 + N_2 & -N_1 \\ -N_1 & P + N_1 \end{bmatrix} \mathbf{Y'} \\ &= \frac{P}{P(N_1+N_2)+N_1N_2} \begin{bmatrix} (N_1+N_2) & -N_1 \end{bmatrix} \mathbf{Y'}. \end{aligned}$$

The best linear MSE estimate $\hat{X} = \hat{X}' + \mu$. Thus,

$$\begin{aligned} \hat{X} &= \frac{P}{P(N_1+N_2)+N_1N_2} \left((N_1+N_2)(Y_1-\mu) - N_1Y_2\right) + \mu \\ &= \frac{1}{P(N_1+N_2)+N_1N_2} \left(P((N_1+N_2)Y_1 - N_1Y_2)) + N_1N_2\mu\right). \end{aligned}$$

The MSE can be calculated by

$$\begin{aligned} \text{MSE} &= \sigma_X^2 - \Sigma_{\mathbf{Y}X}^T \Sigma_Y^{-1} \Sigma_{\mathbf{Y}X} \\ &= P - \frac{P}{P(N_1+N_2)+N_1N_2} \begin{bmatrix} (N_1+N_2) & -N_1 \end{bmatrix} \begin{bmatrix} P \\ 0 \end{bmatrix} \\ &= P - \frac{P^2(N_1+N_2)}{P(N_1+N_2)+N_1N_2} \\ &= \frac{PN_1N_2}{P(N_1+N_2)+N_1N_2} \end{aligned}$$

The equation for the MSE should make perfect sense. First, note that if N_1 and N_2 are held constant but P goes to infinity, the MSE goes to 0, *i.e.* the estimate becomes perfect. Next, note that if both N_1 and N_2 go to infinity, the MSE goes to σ_X^2, *i.e.* the estimate becomes worthless. Finally, note that if either N_1 or N_2 goes to 0, the MSE also goes to 0. This is because the estimator will then use the measurement with zero noise variance (that is, the one with no noise) and ignore the other measurement.