STA 437/1005, Fall 2009 — Solutions to Assignment #1

Q1(a):

$E(X_1) = E(U_1) + E(U_2) + E(\epsilon_1) = 0$, and the same for X_2 and X_3. The the mean vector is $\mu = [0\ 0\ 0]$.

Since U_1, U_2, and ϵ_1 are independent, $\text{Var}(X_1) = \text{Var}(U_1)+\text{Var}(U_2)+\text{Var}(\epsilon_1) = 1+1+2^2 = 6$, and the same for X_2 and X_3.

To find the covariances:

$$\begin{aligned}
\text{Cov}(X_1, X_2) &= E[(X_1 - E(X_1))(X_2 - E(X_2)] \\
&= E[(U_1 + U_2 + \epsilon_1)(U_2 + U_3 + \epsilon_2)] \\
&= E(U_1U_2) + E(U_1U_3) + E(U_1\epsilon_2) \\
&\quad +E(U_2U_2) + E(U_2U_3) + E(U_2\epsilon_2) \\
&\quad +E(\epsilon_1U_2) + E(\epsilon_1U_3) + E(\epsilon_1\epsilon_2) \\
&= E(U_2^2) \quad = \quad 1
\end{aligned}$$

This uses the fact that if A and B are independent, then $E(AB) = E(A)E(B)$, which is zero if either $E(A) = 0$ or $E(B) = 0$ (or both).

The covariance matrix is therefore

$$\Sigma = \begin{bmatrix} 6 & 1 & 1 \\ 1 & 6 & 1 \\ 1 & 1 & 6 \end{bmatrix}$$

Q1(b):

From symmetry Σ^{-1} must have the following form, for some a and b:

$$\Sigma^{-1} = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$$

The requirement that $\Sigma\Sigma^{-1} = I$ gives the two equations

$$6a + 2b = 1, \qquad 7b + a = 0$$

Solving these give $a = 7/40$ and $b = -1/40$, so

$$\Sigma^{-1} = \begin{bmatrix} 7/40 & -1/40 & -1/40 \\ -1/40 & 7/40 & -1/40 \\ -1/40 & -1/40 & 7/40 \end{bmatrix}$$

Q1(c):

Result 4.6 on page 160 of the text says that the conditional distribution a group (1) of variables with a MVN distribution given values for another group (2) of variables is normal with

$$\begin{aligned} \textbf{mean} &= \mu_1 \;+\; \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2) \\ \textbf{covariance} &= \Sigma_{11} \;-\; \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \end{aligned}$$

Applying this to finding the conditional distribution of X_3 given $X_1 = x_1$ and $X_2 = x_2$ produces a normal distribution with

$$\begin{aligned} \textbf{mean} &= [1\ 1]\begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \textbf{variance} &= 6 \;-\; [1\ 1]\begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

One can easily find (as in part (b)) that

$$\begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1} = \begin{bmatrix} 6/35 & -1/35 \\ -1/35 & 6/35 \end{bmatrix}$$

From that, one can calculate that the conditional distribution for X_3 has

$$\begin{aligned} \textbf{mean} &= (x_1 + x_2)/7 \\ \textbf{variance} &= 40/7 \end{aligned}$$

Q2(a):

Let A be a $k \times k$ symmetric real matrix. By the spectral decomposition theorem, we can write it as

$$A \;=\; \lambda_1 e_1 e_1' \;+\; \cdots \;+\; \lambda_k e_k e_k'$$

where the e_i are orthogonal eigenvectors of length one, and the λ_i are the corresponding eigenvalues.

It is easy to see that the trace of a sum of matrices is equal to the sum of their traces, and that the trace of a scalar times a matrix is equal to the scalar times the trace of the matrix. So

$$\text{trace}(A) \;=\; \lambda_1 \text{trace}(e_1 e_1') \;+\; \cdots \;+\; \lambda_k trace(e_k e_k')$$

The diagonal elements of $e_i e_i'$ are $e_{i1}^2, \ldots, e_{ik}^2$. The sum of these is one, since the e_i have length one. So

$$\text{trace}(A) \;=\; \lambda_1 \;+\; \cdots \;+\; \lambda_k$$

Q2(b):

The covariance matrix of $Y = QX$ is $\Sigma_Y = Q\Sigma_X Q'$ (see p. 76 of the text). If e is an eigenvector of Σ_X, with eigenvalue λ, then Qe is an eigenvector of Σ_Y, with eigenvalue λ, since

$$\Sigma_Y(Qe) \;=\; Q\Sigma_X Q'Qe \;=\; Q\Sigma_X e \;=\; Q\lambda e \;=\; \lambda(Qe)$$

The set of eigenvalues for Σ_Y is therefore the same as for Σ_X. So by part (a), the trace of Σ_Y is the same as the trace of Σ_X.

Q3(a):

Let Y_j be the fuel consumption (in litres per 100 kilometres) of car j. We know that the Y_j are independent with $N(\mu, 0.5^2)$ distribution.

The observations X_{j1} and X_{j2} can be written as follows:

$$\begin{aligned} X_{j1} &= 0.1Y_j + \epsilon_{j1} \\ X_{j2} &= 0.9Y_j + \epsilon_{j2} \end{aligned}$$

The 0.1 and 0.9 factors come from the measurements being made after 10 kilometres and after 90 kilometres. The measurement errors, ϵ_{j1} and ϵ_{j2}, are independent (of each other, of Y_j, and of Y_k and $\epsilon_{k\ell}$ for $k \neq j$). The distribution of ϵ_{j1} and ϵ_{j2} is $N(0, 0.1^2)$.

The mean vector for $[X_{j1}\ X_{j2}]'$ is

$$\begin{bmatrix} E(0.1Y_j + \epsilon_{j1}) \\ E(0.9Y_j + \epsilon_{j1}) \end{bmatrix} \;=\; \begin{bmatrix} 0.1\mu \\ 0.9\mu \end{bmatrix}$$

The variances of X_{j1} and X_{j2} are

$$\begin{aligned} \text{Var}(X_{j1}) &= 0.1^2\text{Var}(Y_j) + \text{Var}(\epsilon_{j1}) &= 0.1^2 \times 0.5^2 + 0.1^2 &= 0.0125 \\ \text{Var}(X_{j1}) &= 0.9^2\text{Var}(Y_j) + \text{Var}(\epsilon_{j1}) &= 0.9^2 \times 0.5^2 + 0.1^2 &= 0.2125 \end{aligned}$$

The covariance of X_{j1} and X_{j2} is

$$\begin{aligned} \text{Cov}(X_{j1}, X_{j2}) &= E[(X_{j1} - E(X_{j1})(X_{j2} - E(X_{j2})] \;=\; E[(0.1Y_j + \epsilon_{j1})(0.9Y_j + \epsilon_{j2})] \\ &= E[(0.1Y_j)(0.9Y_j)] \;=\; 0.09E(Y_j^2) \;=\; 0.09 \times 0.5^2 \;=\; 0.0225 \end{aligned}$$

The covariance matrix of $[X_{j1}\ X_{j2}]'$ is therefore

$$\Sigma \;=\; \begin{bmatrix} 0.0125 & 0.0225 \\ 0.0225 & 0.2125 \end{bmatrix}$$

Q3(b):

(See p. 121 of the text.) The mean of $[\overline{X}_1 \;\; \overline{X}_2]'$ is the same as that of each $[X_{j1} \;\; X_{j2}]'$, which is $[0.1\mu \; 0.9\mu]'$. The covariance matrix of $[\overline{X}_1 \;\; \overline{X}_2]'$ is $\Sigma/10$, which is

$$\begin{bmatrix} 0.00125 & 0.00225 \\ 0.00225 & 0.02125 \end{bmatrix}$$

Q3(c):

The mean of $a\overline{X}_1 \; + \; b\overline{X}_2$ is $aE(\overline{X}_1) + bE(\overline{X}_2) \;\; = \;\; (0.1a + 0.9b)\mu$. For an unbiased estimator of μ, this must be equal to μ, so we must have $0.1a + 0.9b = 1$. It follows that $a = 10 - 9b$. We can therefore look for the value of b that gives the smallest variance, when setting $a = 10 - 9b$.

The variance of $(10 - 9a)\overline{X}_1 \; + \; b\overline{X}_2$ is (see p. 76 of the text)

$$[(10 - 9b) \;\; b] \begin{bmatrix} 0.00125 & 0.00225 \\ 0.00225 & 0.02125 \end{bmatrix} \begin{bmatrix} (10 - 9b) \\ b \end{bmatrix} \;\; = \;\; 0.082b^2 \; - \; 0.18b \; + \; 0.125$$

We want to find the value of b that minimizes this. We can do this by setting the derivative with respect to b to zero:

$$2 \times 0.082b \; - \; 0.18 \;\; = \;\; 0$$

from which we get $b \;=\; 0.18/(2 \times 0.082) \;=\; 1.09756\ldots$ and $a \;=\; 10 - 9b \;=\; 0.12195\ldots$