Solution of Problem 1

The multivariate normal (or Gaussian) distribution of a random vector \( \mathbf{Y} \in \mathbb{R}^p \) has the following pdf:

\[
f_Y(\mathbf{y}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{\mu})^T \Sigma^{-1} (\mathbf{y} - \mathbf{\mu}) \right\},
\]

where \( \mathbf{y} = (y_1, \ldots, y_p)^T \in \mathbb{R}^p \), and the parameters: \( \mathbf{\mu} \in \mathbb{R}^p \), \( \Sigma \in \mathbb{R}^{p \times p} \), where \( \Sigma > 0 \).

\( a) \) In our case we have that \( p = 2 \), yielding

\[
f_Y(\mathbf{y}) = \frac{1}{(2\pi)^{|\Sigma|^{1/2}}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{\mu})^T \Sigma^{-1} (\mathbf{y} - \mathbf{\mu}) \right\}.
\]

We start by calculating the determinant of \( \Sigma \in \mathbb{R}^{2 \times 2} \) as \( |\Sigma| = \sigma_1^2 \sigma_2^2 - \rho \sigma_1 \sigma_2 \sigma_1^2 \sigma_2^2(1 - \rho^2) \). This leads to \( |\Sigma|^{1/2} = \sigma_1 \sigma_2 \sqrt{1 - \rho^2} \) and

\[
\Sigma^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho \sigma_1 \sigma_2 \\ -\rho \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix}.
\]

Finally, we calculate

\[
-\frac{1}{2} (\mathbf{y} - \mathbf{\mu})^T \Sigma^{-1} (\mathbf{y} - \mathbf{\mu}) = -\frac{1}{2 \sigma_1^2 \sigma_2^2 (1 - \rho^2)} (y_1 - \mu_1)(y_2 - \mu_2) - \rho \sigma_1 \sigma_2 (y_1 - \mu_1)(y_2 - \mu_2)
\]

\[
= -\frac{1}{2 \sigma_1^2 \sigma_2^2 (1 - \rho^2)} \left[ (y_1 - \mu_1)^2 + (y_2 - \mu_2)^2 - 2 \rho \sigma_1 \sigma_2 (y_1 - \mu_1)(y_2 - \mu_2) \right]
\]

\[
= -\frac{1}{2(1 - \rho^2)} \left[ \frac{(y_1 - \mu_1)^2}{\sigma_1^2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2} - 2 \rho \frac{(y_1 - \mu_1)(y_2 - \mu_2)}{\sigma_1 \sigma_2} \right],
\]

this gives us the final expression for \( f_Y(\mathbf{y}) \) as

\[
f_Y(\mathbf{y}) = \frac{1}{(2\pi)^{1/2} \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \frac{(y_1 - \mu_1)^2}{\sigma_1^2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2} - 2 \rho \frac{(y_1 - \mu_1)(y_2 - \mu_2)}{\sigma_1 \sigma_2} \right] \right\}.
\]

\( b) \) From the definition of \( \mathbf{\mu} \) and \( \Sigma \) we directly get \( Y_1 \sim N(\mu_1, \sigma_1) \) and \( Y_2 \sim N(\mu_2, \sigma_2) \).
c) As stated in theorem 3.5 of the lecture’s script, the conditional density \( f_{Y_1|Y_2}(y_1|y_2) \) is given by the normal distribution \( f_{Y_1|Y_2}(y_1|y_2) \sim N_1(\mu_{1|2}, \Sigma_{1|2}) \), where \( \mu_{1|2} \) is
\[
\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - \mu_2) \\
= \mu_1 + (\rho\sigma_1\sigma_2)(1/\sigma_2^2)(y_2 - \mu_2) \\
= \mu_1 + \rho\sigma_1/\sigma_2(y_2 - \mu_2)
\]
and \( \Sigma_{1|2} \) is
\[
\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \\
= \sigma_1^2 - (\rho\sigma_1\sigma_2)(1/\sigma_2^2)(\rho\sigma_1\sigma_2) \\
= \sigma_1^2 - \rho^2\sigma_1^2 = \sigma_1^2(1 - \rho^2).
\]

**Solution of Problem 2**

Note that an estimator \( \hat{X} \) of a parameter \( X \) is unbiased if its expected value equals \( X \). Therefore it is enough to show:
\[
E(\bar{X}) = \mu = E(X), \quad E(S_n) = \Sigma = \text{Cov}(X).
\]

First see that:
\[
E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} E(X) = E(X).
\]

For the sample covariance matrix, we have:
\[
E(S_n) = E\left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})^T\right) \\
= \frac{1}{n-1} \sum_{i=1}^{n} E((X_i - \bar{X})(X_i - \bar{X})^T)
\]

Next see that:
\[
E \left( (X_i - \bar{X})(X_i - \bar{X})^T \right) = E \left( (X_i - \frac{1}{n} \sum_{j=1}^{n} X_j)(X_i - \frac{1}{n} \sum_{j=1}^{n} X_j)^T \right) \\
= E \left( (X_i - \mu - \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu))(X_i - \mu - \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu))^T \right) \\
= E \left( \left( \frac{n-1}{n}(X_i - \mu) - \frac{1}{n} \sum_{j=1, j \neq i}^{n} (X_j - \mu) \right) \left( \frac{n-1}{n}(X_i - \mu) - \frac{1}{n} \sum_{j=1, j \neq i}^{n} (X_j - \mu) \right)^T \right)
\]

It is easy to see that:
\[
E \left( (X_i - \mu)(X_j - \mu)^T \right) = \delta_{ij} \Sigma.
\]
Using this fact, it is easy to see that:

\[
E \left( \left( \frac{n-1}{n} (X_i - \mu) - \frac{1}{n} \sum_{j=1, j \neq i}^n (X_j - \mu) \right) \left( \frac{n-1}{n} (X_i - \mu) - \frac{1}{n} \sum_{j=1, j \neq i}^n (X_j - \mu) \right)^T \right) \\
= \frac{(n-1)^2}{n^2} E ((X_i - \mu)(X_i - \mu)^T) + \frac{1}{n^2} \sum_{j=1, j \neq i}^n E ((X_j - \mu)(X_j - \mu)^T) \\
= \frac{(n-1)^2}{n^2} \Sigma + \frac{n-1}{n^2} \Sigma = \frac{n-1}{n} \Sigma.
\]

Therefore \( E ((X_i - \bar{X})(X_i - \bar{X})^T) = \frac{n-1}{n} \Sigma \). We can finally find the expected value of sample covariance as follows:

\[
E(S_n) = \frac{1}{n-1} \sum_{i=1}^n E ((X_i - \bar{X})(X_i - \bar{X})^T) = \frac{1}{n-1} \sum_{i=1}^n \frac{n-1}{n} \Sigma = \Sigma.
\]

**Solution of Problem 3**

Consider four samples in \( \mathbb{R}^3 \) given as follows:

\[
\begin{align*}
\mathbf{x}_1 &= \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \\
\mathbf{x}_2 &= \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \\
\mathbf{x}_3 &= \begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix}, \\
\mathbf{x}_4 &= \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}.
\end{align*}
\]

**a)** The sample mean can be easily found as:

\[
\bar{\mathbf{x}} = \begin{bmatrix} -0.75 \\ 0.5 \\ 0.25 \end{bmatrix}
\]

To find the sample covariance, we have:

\[
S_n = \frac{1}{3} \sum_{i=1}^4 (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T = \frac{1}{3} \begin{bmatrix} 32.75 & -4.5 & -28.25 \\ -4.5 & 9 & -4.5 \\ -28.25 & -4.5 & 32.75 \end{bmatrix}.
\]

**b)** Step 1: find the sample covariance matrix \( S_n \) (previous part)

Step 2: find the eigenvalues and eigenvectors of the matrix. Sort them out and pick 2 orthonormal eigenvectors corresponding to 2 highest eigenvalues

\[
\lambda_1 = 20.333333, \lambda_2 = 4.5, \lambda_3 = 0.
\]

\[
\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} \frac{\sqrt{3}}{6} \\ \frac{\sqrt{3}}{6} \\ \frac{1}{\sqrt{6}} \end{bmatrix}.
\]

Step 3: Construct \( Q = \mathbf{v}_1 \mathbf{v}_1^T + \mathbf{v}_2 \mathbf{v}_2^T \).

Following this procedure, we have:

\[
Q = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.
\]
c) Note that all the points are already on the same plane \( x + y + z = 0 \), so intuitively, the projection should be the projection on the same plane. This projection leaves those points untouched (Check!). Each \( y \in \text{Im}(Q) \) is also on this plane. To see that assume that \( y = Qx \). Then \( y_1 + y_2 + y_3 = 0 \). Another way, is to observe that the kernel of \( Q \) is spanned by the vector \( (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) \), the last eigenvector. Note, how the eigenvalue is zero for this eigenvector. Therefore its image is the orthogonal complement of this vector which is the plane \( x + y + z = 0 \).